

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 Flutter

Flutter [1] คือ Framework ที่ใช้สร้าง UI สำหรับ mobile application ที่สามารถทำงานได้ทั้ง iOS และ Android ในเวลาเดียวกันโดยที่ใช้ source code ตัวเดียวกัน โดยภาษาที่ใช้ใน Flutter นั้นจะเป็นภาษา dart ซึ่งถูกพัฒนาโดย Google และ ยังเป็น open source ที่สามารถใช้งานได้ฟรี ตัวอย่าง syntax ของภาษา dart ที่ใช้ใน Flutter ซึ่งจะมีความคล้ายกับภาษา Java เนื่องจาก dart เป็นภาษาที่รองรับ OOP และมีแนวคิด เช่นเดียวกับภาษา Java

```
1 import 'package:flutter/material.dart';
2
3 void main() {
4   runApp(new MaterialApp(
5     home: new MyApp(),
6   ));
7 }
8
9 class MyApp extends StatelessWidget {
10  @override
11  Widget build(BuildContext context) {
12    return new Scaffold(
13      appBar: new AppBar(
14        title: new Text("Example App"),
15        backgroundColor: Colors.blue,
16      ),
17      backgroundColor: Colors.blue,
18      body: new Center(
19        child: new Column(
20          mainAxisAlignment: MainAxisAlignment.center,
21          children: <Widget>[
22            new Icon(Icons.favorite, color: Colors.redAccent, size: 200.0,
23          ),
24        ],
25      ),
26    );
27  };
28 }
29 }
```

ภาพประกอบที่ 2.1 ตัวอย่างโครงสร้าง Flutter

ซึ่งจะเห็นว่า Flutter นั้นจะมี Widget พื้นฐานมาให้ เพื่อให้การออกแบบ UI มีความง่าย และสะดวกยิ่งขึ้น โดย Widget พื้นฐานของ Flutter หลัก ๆ จะมีอยู่ 2 ชนิดคือ StatelessWidget และ StatefulWidget โดยที่ StatelessWidget จะใช้สร้าง Widget ที่ไม่มีการจัดการสถานะการทำงานใดหรือหน้านั้นๆจะไม่มีเปลี่ยนแปลง เช่น การแสดงข้อความ, Icon หรือรูปภาพที่ไม่มี animation เข้ามาเกี่ยวข้อง เป็นต้น ส่วน StatefulWidget จะใช้สร้าง Widget ที่มีการจัดการสถานะการทำงานต่างๆหรือมีปุ่มที่มี action เช่น การสร้าง Icon ที่มีการใส่ animation ให้สามารถขยับไปมาได้, ปุ่มกดต่างๆ บนหน้า UI เป็นต้น

จุดเด่นหลัก ๆ ของ Flutter คือ ระบบ Hot Reload โดยเมื่อมีการทดสอบ, การสร้าง, การ add features หรือการกระทำต่าง ๆ กับ UI จะต้องมีการ reload เพื่อให้หน้า UI update ซึ่งระบบ Hot Reload จะเข้ามาช่วยในส่วนของการ reload โดยจุดเด่นของระบบนี้คือการย่นระยะเวลาที่ใช้ในการ reload ให้เหลือเพียงเสี้ยววินาทีเท่านั้น ทำให้การพัฒนา UI ของ application มีความรวดเร็วขึ้นอย่างมาก และยังมีจุดเด่นอื่น ๆ ที่ช่วยให้การพัฒนาเป็นไปได้ง่ายขึ้นไม่ว่าจะเป็น Build-In ที่ช่วยในการออกแบบ UI ให้มีความสวยงามยิ่งขึ้นอย่าง Material Design และ Cupertino (iOS-flavor), มี Framework ที่ช่วยให้การทำ animation ต่าง ๆ หรือ gesture ของ UI เป็นเรื่องง่ายยิ่งขึ้น และยังสามารถใช้งานร่วมกับ IDE ที่กำลังเป็นที่นิยมอยู่ในปัจจุบันอย่าง VS Code และ Android Studio ได้

ข้อเสียหลัก ๆ คือ การใช้ภาษา dart ในการเขียน ซึ่งคนส่วนใหญ่อาจจะยังไม่คุ้นเคยกับ syntax ของภาษา dart ประกอบกับ community ยังเล็กเนื่องจาก Flutter ยังเปิดตัวมาได้ไม่นานนัก เมื่อเทียบกับ Framework ตัวอื่น ๆ อย่าง React Native ที่มี community ค่อนข้างใหญ่จึงทำให้ document ต่าง ๆ ยังไม่เยอะเท่าที่ควร ทำให้เวลามีปัญหาเกี่ยวกับการใช้งานอาจจะต้องใช้เวลาในการหาวิธีแก้

2.1.2 Firebase

เป็นหนึ่งในผลิตภัณฑ์ของ Google โดย Firebase [2] คือ Platform ที่รวบรวมเครื่องมือต่าง ๆ สำหรับการจัดการในส่วนของการเก็บข้อมูล ซึ่งทำให้สามารถ สร้าง Mobile Application ได้ อย่างมีประสิทธิภาพ และยังลดเวลาและค่าใช้จ่ายของการทำ Server side หรือการวิเคราะห์ข้อมูลให้อีกด้วย โดยมีทั้งเครื่องมือที่ฟรี และเครื่องมือที่มีค่าใช้จ่าย Firebase มีบริการให้ใช้หลายอย่าง สามารถแบ่งเป็นหมวดหมู่ดังนี้

2.1.2.1 Better Apps

- Cloud Firestore จัดเก็บและซิงค์ข้อมูล ระหว่างผู้ใช้และอุปกรณ์ในระดับโลกโดยใช้ฐานข้อมูล NoSQL ที่โฮสต์บนคลาวด์ Cloud Firestore ให้การซิงโครไนซ์แบบสดและการสนับสนุนออฟไลน์พร้อมกับการสืบค้นข้อมูลที่มีประสิทธิภาพ การผสมผสานร่วมกับผลิตภัณฑ์ Firebase อื่นๆ ช่วยให้คุณสามารถสร้าง แอปแบบไร้เซิร์ฟเวอร์ได้อย่างแท้จริง

- Authentication จัดการผู้ใช้ของคุณด้วยวิธีที่ง่ายและปลอดภัย Firebase Auth มีหลายวิธีในการตรวจสอบสิทธิ์รวมถึง Email และ password ผู้ให้บริการบุคคลที่สามเช่น Google หรือ Facebook และใช้ระบบบัญชีที่คุณมีอยู่โดยตรง สร้างอินเทอร์เฟซของคุณเองหรือใช้ประโยชน์จากโอเพ่นซอร์ส UI ที่ปรับแต่งได้อย่างเต็มที่

- Hosting ลดความซับซ้อนของเว็บโฮสต์ของคุณด้วยเครื่องมือที่สร้างขึ้นเฉพาะสำหรับเว็บแอปสมัยใหม่ เมื่อคุณอัปโหลดเนื้อหาเว็บของคุณเราจะส่งเนื้อหาเหล่านั้นไปยัง CDN ทั่วโลกของเราโดยอัตโนมัติและมอบใบรับรอง SSL ฟรีเพื่อให้ผู้ใช้ของคุณได้รับประสบการณ์ที่ปลอดภัย เชื่อถือได้และมีเวลาแฝงต่ำไม่ว่าจะอยู่ที่ใดก็ตาม

- Realtime Database Realtime Database คือฐานข้อมูลดั้งเดิมของ Firebase เป็นโซลูชันที่มีประสิทธิภาพและมีเวลาแฝงต่ำสำหรับแอปบนอุปกรณ์เคลื่อนที่ที่ต้องการ

สถานะการชิงคราะห์หว่างไคลเอนต์แบบเรียลไทม์ เราขอแนะนำ Cloud Firestore แทน Realtime Database สำหรับนักพัฒนาส่วนใหญ่ที่เริ่มโปรเจกต์ใหม่

2.1.2.2 Improve app quality

- Crashlytics ลดเวลาในการแก้ไขปัญหาของคุณด้วยการเปลี่ยนข้อขัดข้องจากหิมะถล่มให้เป็นรายการปัญหาที่จัดการได้ รับข้อมูลเชิงลึกที่ชัดเจนและนำไปปฏิบัติได้ว่าปัญหาใดที่ต้องจัดการก่อนโดยเห็นผลกระทบของผู้ใช้ในแดชบอร์ด Crashlytics การแจ้งเตือนแบบเรียลไทม์จะช่วยให้คุณมีความเสถียรแม้ในขณะเดินทาง Crashlytics เป็นตัวรายงานข้อขัดข้องหลักของ Firebase
- Performance Monitoring วินิจฉัยปัญหาประสิทธิภาพของแอปที่เกิดขึ้นบนอุปกรณ์ของผู้ใช้ ใช้การติดตามเพื่อตรวจสอบประสิทธิภาพของบางส่วนของแอปและดูมุมมองสรุปในคอนโซล Firebase อยู่เหนือเวลาเริ่มต้นของแอปและตรวจสอบคำขอ HTTP โดยไม่ต้องเขียนโค้ดใด ๆ
- Test Lab เรียกใช้การทดสอบอัตโนมัติและกำหนดเองสำหรับแอปของคุณบนอุปกรณ์เสมือนและจริงที่โฮสต์โดย Google ใช้ Firebase Test Lab ตลอดวงจรการพัฒนาของคุณเพื่อค้นหาจุดบกพร่องและความไม่สอดคล้องกันเพื่อให้คุณสามารถนำเสนอประสบการณ์ที่ยอดเยี่ยมบนอุปกรณ์หลากหลายประเภท

2.1.2.3 Grow your business

- Google Analytics วิเคราะห์คุณลักษณะและพฤติกรรมของผู้ใช้ในแดชบอร์ดเดียวเพื่อทำการตัดสินใจอย่างชาญฉลาดเกี่ยวกับแผนงานผลิตภัณฑ์ของคุณ รับข้อมูลเชิงลึกแบบเรียลไทม์จากรายงานหรือส่งออกข้อมูลเหตุการณ์ดิบไปยัง Google BigQuery สำหรับการวิเคราะห์ที่กำหนดเอง
- Remote Config กำหนดวิธีการแสดงผลแอปของคุณสำหรับผู้ใช้แต่ละคน เปลี่ยนรูปลักษณะเปิดตัวพีเจอร์ที่ละน้อยเรียกใช้การทดสอบ A/B ส่งมอบเนื้อหาที่กำหนดเองให้กับผู้ใช้บางรายหรือทำการอัปเดตอื่น ๆ โดยไม่ต้องปรับใช้เวอร์ชันใหม่ทั้งหมดนี้ทำได้จากคอนโซล Firebase ตรวจสอบผลกระทบของการเปลี่ยนแปลงของคุณและทำการปรับเปลี่ยนในเวลาไม่กี่นาที
- Cloud Messaging ส่งข้อความและการแจ้งเตือนไปยังผู้ใช้ข้ามแพลตฟอร์มทั้ง Android, iOS และเว็บได้ฟรี สามารถส่งข้อความไปยังอุปกรณ์เดี่ยวกลุ่มอุปกรณ์หรือหัวข้อเฉพาะหรือกลุ่มผู้ใช้ Firebase Cloud Messaging (FCM) ปรับขนาดเป็นแอปที่ใหญ่ที่สุดโดยส่งข้อความหลายแสนล้านข้อความต่อวัน

2.1.3 การประมวลผลภาพ (Image Processing)

การประมวลผลภาพ (Image Processing) [3]เป็นการนำภาพมาประมวลผลหรือคิดคำนวณด้วย คอมพิวเตอร์แล้วใช้กรรมวิธีใดๆมากระทำกับข้อมูลภาพเพื่อให้ได้ภาพที่มีคุณสมบัติตามต้องการทั้งในเชิง คุณภาพและปริมาณ โดยมีขั้นตอนต่างๆที่สำคัญ คือ การทำให้ภาพมีความคมชัดมากขึ้น การกำจัดสัญญาณรบกวนออกจากภาพ การแบ่งส่วนของวัตถุที่สนใจออกมาจากภาพ เพื่อนำภาพวัตถุที่ได้ ไปวิเคราะห์หาข้อมูลเชิงปริมาณ เช่น ขนาด รูปร่าง และทิศทางการเคลื่อนของวัตถุในภาพ คอมพิวเตอร์มีความสามารถในการคำนวณและประมวลผลข้อมูลจำนวนมากได้ในเวลาอันสั้นจึงมี

ภาพเคลื่อนไหวนั้น โปรแกรมจะมองเห็นเป็นภาพนิ่งหลายภาพที่เรียงต่อกัน จึงมีประโยชน์อย่าง มากในการเพิ่ม ประสิทธิภาพการประมวลผลภาพและวิเคราะห์ข้อมูลได้จากภาพในระบบต่าง ๆ ดังกล่าวข้างต้น การประมวลผลภาพ สามารถจำแนกตาม วัตถุประสงค์ในการใช้งานดังนี้

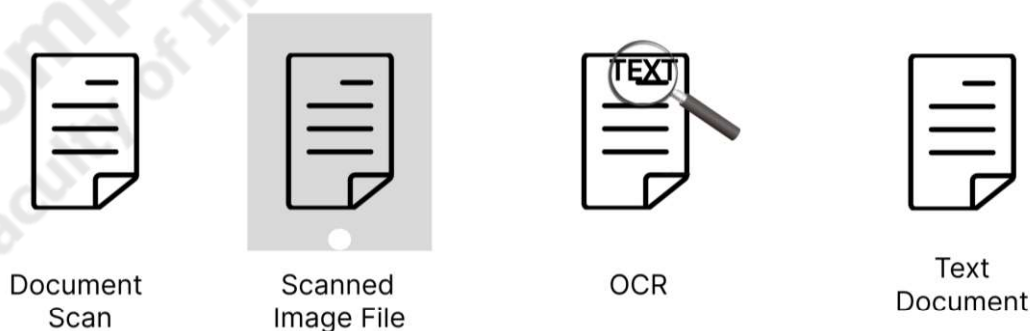
- การปรับปรุงคุณภาพของภาพ (Image Enhancement and Restoration) เช่น ปรับภาพ ให้คมชัด มากขึ้นการปรับ Contrast หรือการปรับเน้นเส้นขอบภาพ (Edge Enhancement) หรือการ กรองสัญญาณรบกวน (Image Filtering) เพื่อกำจัดสัญญาณรบกวนได้

- การบีบอัดภาพ (Image Compression) เนื่องจากข้อมูลภาพนั้นมีขนาดใหญ่ มากโดยเฉพาะ ภาพสี ซึ่งจะทำให้การจัดเก็บหรือรับส่งไฟล์ข้อมูลขนาดใหญ่เสียเวลามาก เพื่อให้ ขนาดภาพเล็กลงจึง ทำให้การบีบอัด ข้อมูลภาพ เช่น JPEG หรือ GIF และตัวอย่างการบีบอัดข้อมูลวีดีโอ เช่น MPEG หรือ AVI เป็นต้น

- การวิเคราะห์ ข้อมูลภาพ (Image Analysis) เป็นการสังเคราะห์ ข้อมูลที่มีความหมายจากภาพ แล้ว นำไปใช้งาน ซึ่งอาจต้องมีรู้จำวัตถุ (object recognition) การแยกส่วนวัตถุ (segmentation) ตรวจจับการ เคลื่อนไหวที่ของวัตถุ (motion detection)

2.1.4 OCR หรือ Optical Character Recognition

OCR หรือ Optical Character Recognition [4] หรือ "การรู้จำอักขระด้วยแสง" เป็นเทคโนโลยีที่แปลงเอกสาร รูปภาพ หรือแม้แต่ ลายเซ็นอิเล็กทรอนิกส์ ให้อยู่ในรูปแบบตัวอักษร หรือข้อความ (Plain Text) ไฟล์ดิจิทัลที่สามารถสืบค้นได้ด้วยคำค้นหา (Keyword) ยกตัวอย่าง ไฟล์เอกสาร ที่มาในรูปแบบ PDF มีข้อดีก็คือไม่สามารถปลอมแปลง แก้ไขได้ แต่ถ้าต้องการแก้ไขข้อความใด ๆ ในนั้นก็ เป็นเรื่องยากลำบาก วิธีแก้ไขปัญหาก็คือ แปลงไฟล์ PDF ให้เป็น OCR ซึ่งวิธีการที่ทุกคนคุ้นเคยก็คือ การนำแผ่นเอกสาร มาสแกนข้อความ ตัวอักษร เส้นตารางและอื่น ๆ ให้เป็นไฟล์ Word หรือ PDF ด้วย เครื่องสแกนเนอร์ หรือกล้องถ่ายภาพ



ภาพประกอบที่ 2.3 ขบวนการการทำงานของ Optical Character Recognition

ส่วนการทำงานของ OCR นั้น หลัก ๆ คือ ใช้การจดจำรูปแบบ เพื่อกำหนดอักขระของแต่ละประเภทไฟล์ จากนั้น ซอฟต์แวร์จะทำการอ่านข้อความและอักขระ แล้วแปลงเป็นไฟล์ที่สืบค้นได้

นอกจากนี้ การทำงานของ OCR ยังขึ้นอยู่กับซอฟต์แวร์ที่ใช้ประมวลผล เพราะนอกเหนือจากการแปลงรูปภาพ อักษรใด ๆ เป็นข้อความแล้ว ซอฟต์แวร์บางตัวสามารถจัดวางรูปแบบข้อความในไฟล์ OCR ได้

2.1.5 Object Detection

Object Detection [5] การตรวจจับวัตถุ คือ เทคโนโลยีในทางคอมพิวเตอร์ หลักการที่เกี่ยวกับ Computer Vision และ Image Processing ที่ใช้ในงาน AI ตรวจจับวัตถุชนิดที่กำหนด เช่น มนุษย์ รถยนต์ อาคาร ที่อยู่ในรูปภาพ หรือวิดีโอ งาน Object Detection การตรวจจับวัตถุในรูปภาพ สามารถเจาะลึกลงไปได้อีกหลายแขนง เช่น การทำ ตรวจจับหน้าคน ตรวจจับคนเดินถนน สามารถประยุกต์ใช้ได้หลากหลาย เช่น ใช้ในงานรักษาความปลอดภัย และรถยนต์ไร้คนขับ เป็นต้น

2.1.6 YOLO (You Only Look Once)

เป็นอัลกอริทึม [6] ที่นำแนวความคิดของการทำนายตำแหน่งและขนาดของ กล่องจาก ความน่าจะเป็นที่กล่องนั้น จะเป็นกรอบล้อม วัตถุ แต่สิ่งที่ YOLO มีความสามารถและความเร็ว เหนือกว่าอัลกอริทึมอื่น เช่น Faster R-CNN ที่มีการ ทำงานในลักษณะที่จะทำนายตำแหน่งของกรอบ ล้อม วัตถุและค่อนำวัตถุในกล่องนั้น ไปผ่านแบบจำลองเพื่อ ทำนายวัตถุในกล่อง แต่ YOLO นั้นจะ ทำนายทั้งกรอบ ล้อมวัตถุ และความน่าจะเป็นของวัตถุบางส่วนที่อยู่ใน กรอบออกมาพร้อมกันทีเดียว YOLO จัดว่าเป็นเทคนิคการตรวจจับวัตถุในภาพซึ่ง เป็นซอฟต์แวร์เปิดสำหรับงานปัญญาประดิษฐ์แบบ โครงข่ายประสาทเทียม (Neural Network) ที่พัฒนาด้วย โปรแกรมภาษา C++ และสามารถทำงานบน หน่วยประมวลผล CUDA ของ GPU ได้เป็นอย่างดี เหมาะกับ การประมวลผลภาพแบบ Real Time ภาพจากกล้องหรือ วิดีโอ ปัจจุบัน YOLO มีการพัฒนามาแล้ว 5 เวอร์ชันปัจจุบันคือ YOLO v5 โครงสร้างโครงข่ายประสาทเทียมของ YOLO แต่ละ เวอร์ชันจะมี Convolution Box ที่ แตกต่างกัน แต่ หลักการโดยทั่วไปแล้ว YOLO จะแบ่งภาพออกเป็น Grid Cell เล็กๆ และแต่ละ Grid Cell จะถูก ทำนายผ่าน แบบจำลอง เพื่อหาตำแหน่งจุดกึ่งกลางของวัตถุ และ ความน่าจะเป็นที่จะมีวัตถุใด ใน Grid Cell

2.2 งานวิจัยที่เกี่ยวข้อง

2.2.1 Safety Helmet Detection Based on YOLOv5

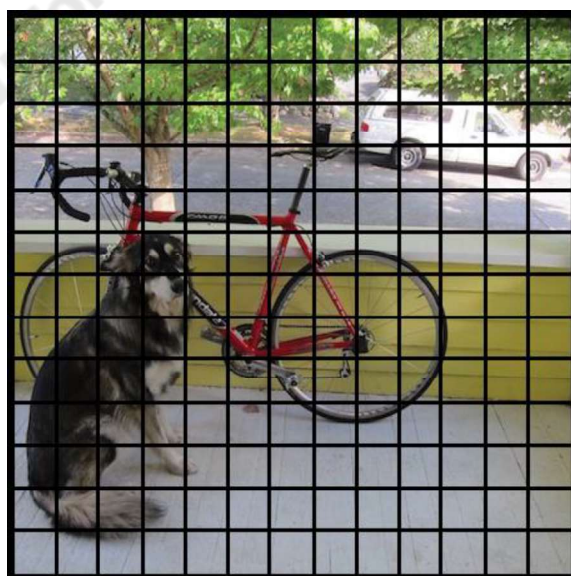
"Safety Helmet Detection Based on YOLOv5" [7] เป็นงานวิจัย ของ Fangbo Zhou, Huailin Zhao , Zhen Nie มหาวิทยาลัย Shanghai Institute of Technology [7] ที่พัฒนาระบบตรวจสอบความปลอดภัยของคนงาน ซึ่งระบบจะทำการตรวจจับภาพของคนงาน และ จะแสดงกรอบที่มีคำอธิบาย โดยได้นำรูปภาพจากอินเทอร์เน็ตจำนวน 6045 ภาพมาทำการทดสอบ โดยสีระ ของคนงานที่ไม่สวมหมวกนิรภัยจะมีคำอธิบายว่า "Alarm" และสำหรับสีระของคนงานที่สวมหมวกนิรภัยจะมีคำอธิบายว่า "Helmet" โดยใช้ algorithm ที่ทันสมัยที่สุดในการตรวจสอบ algorithm งาน ตรวจจับวัตถุถูกใช้อย่างแพร่หลายในความเป็นจริง เป้าหมายของการตรวจจับคือการค้นหาวัตถุทั้งหมดที่น่าสนใจในภาพ ซึ่งจะมี 2 งานย่อยคือการกำหนดหมวดหมู่และการระบุตำแหน่งของวัตถุ แม้ว่า

อัลกอริธึมการตรวจจับแบบดั้งเดิมสามารถทำงานได้ผลลัพธ์ที่ดี ในบางสถานการณ์ในสภาพแวดล้อมที่แปรปรวน เช่น การเปลี่ยนแปลงของสภาพอากาศ การกระจายตัวของคนงานที่ไม่สม่ำเสมอ และหมวกกันน็อคแบบต่างๆ ทำให้ความแม่นยำนั้นรับประกันได้ยากกว่ามันสามารถทำงานได้ถูกต้อง จึงได้นำตัวYOLOv5มาใช้ซึ่งมีความเร็วและความแม่นยำสูง



ภาพประกอบที่ 2.4 ภาพการตรวจจับหมวกกันน็อคของคนงาน [8]

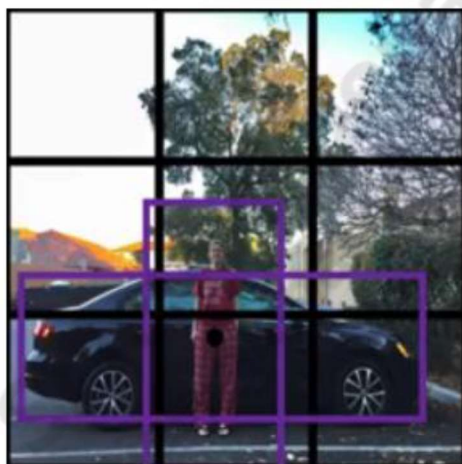
โดยการทำงานของYOLOv5 [8]คือจากรูป 1 รูปเต็มๆ จะทำการแบ่ง Grid cell ออกมาเป็น $n \times n$ grid ยิ่งแบ่งมากก็จะละเอียด trade off กับการคำนวณ



ภาพประกอบที่ 2.5 การแบ่งช่องของรูปภาพออกเป็น grid [9]

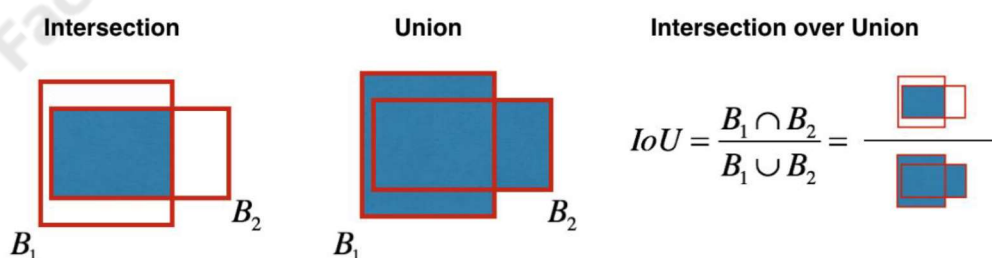
และในแต่ละ grid จะมี Label ทุกช่อง เช่น [Pc, bx, by, bh, bw, c1, c2, ..., cn] โดยที่ Pc คือ ความน่าจะเป็นที่มีวัตถุอยู่ใน Grid นั้นๆ ถ้าไม่มีคือจะได้ค่า 0 ถ้ามีจะได้ค่า 1 bx, by คือ ตำแหน่งตรงกลางของ Object ว่าอยู่ที่พิกัดไหน ส่วน bh, bw คือขนาดความสูงและกว้างของ Object ว่าสูง, กว้างขนาดไหน และ c1, c2, ..., cn คือ ผลลัพธ์ว่าเป็น class อะไร ถ้าโจทย์มีแค่ detect หมวกนิรภัย ก็จะมี class เดียว และ grid นั้นมีค่าเป็น 1 แต่ถ้ามีหลาย Object ก็จะมีเลขต่อไป

จากภาพประกอบที่ 2.6 จะรับได้เฉพาะ 1 grid คือ 1 Object แต่กรณีถ้ามีหลาย Object เราจะใช้หลักการที่เรียกว่า Anchor Box โดยเราสามารถกำหนดจำนวน Box นี้ได้ ก็จะมี Label แบบ ด้านบน 2 อันใน 1 Grid ได้ และตัว YOLO ก็จะสามารถให้ว่ารูปนั้นใกล้ Anchor อันไหนสุดจากค่า IOU ก็จะถูกกำหนดไปที่ Anchor นั้น



ภาพประกอบที่ 2.6 ภาพของเทคนิค Anchor Box [9]

IOU จะมีส่วนหลักๆ หลายส่วนใน YOLO เช่น เวลาโมเดลทำนายออกมา มันอาจจะทำนายรูปรถเดียวกันแต่หลายกล่องได้ จึงต้องใช้ metric IOU ในเลือกกล่องเดียวเป็นตัวแทนของ Object โดยหลักการคำนวณคือ หาส่วนที่ Intersect หาส่วนที่ Union กัน ถ้ากล่องใดมีค่านี้สูง เกินเกณฑ์ที่กำหนดแสดงว่ามันคือ Object เดียวกัน



ภาพประกอบที่ 2.7 การทำงานของขบวนการ IOU [9]