บทที่ 2

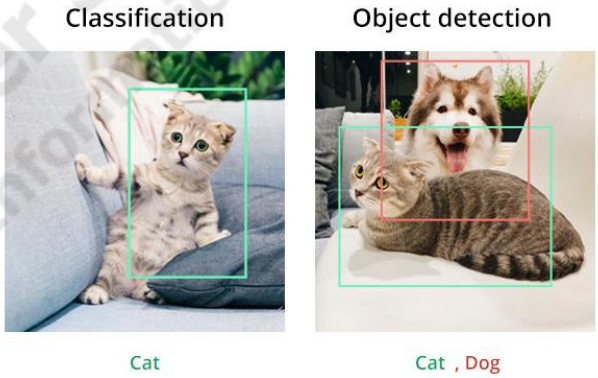
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 Objective Detection

"เทคโนโลยีตรวจจับวัตถุ" (Object detection) คือ หนึ่งในฟีเจอร์หลักของ AI (Artificial Intelligence) ที่ใช้กับกล้องวงจรปิด สามารถค้นหาสิ่งของโดยใช้ AI มาวิเคราะห์ข้อมูล จากการมองเห็นของคอมพิวเตอร์ (Computer Vision) และการประมวลผลภาพ (Image Processing) เพื่อตรวจจับวัตถุที่อยู่ในรูปหรือวิดีโอ เช่น มนุษย์ สัตว์ สิ่งของ รถยนต์ อาคาร และวัตถุอื่น ๆ ที่อยู่ในรูปภาพ หรือวิดีโอ

โดยตามหลักแล้วก่อนที่จะพัฒนามาเป็นเทคโนโลยีตรวจจับวัตถุ (Object detection) จะต้องผ่านการจัดหมวดของวัตถุ (Object Classification) มาก่อน โดยที่การจัดหมวดของวัตถุจะเป็นการจัดหมวดหมู่ของรูปภาพว่า รูปภาพนั้นคือภาพอะไร แต่เทคโนโลยีตรวจจับวัตถุจะเป็นการระบุเลยว่า ในรูปภาพนั้นมีวัตถุอะไรบ้าง ซึ่งจุดนี้จะต้องอาศัยการทำงานของ AI เข้ามาช่วยในการวิเคราะห์ข้อมูลด้วยเช่นกัน เช่น การจัดหมวดของวัตถุจะสามารถระบุได้ว่าวัตถุที่อยู่ในภาพ คือ แมว ในขณะที่เทคโนโลยีตรวจจับวัตถุจะระบุได้ว่าวัตถุที่อยู่ในภาพ มีแมว และสุนัข



ภาพประกอบที่ 2.1 หลักการการทำงาน Objective Detection

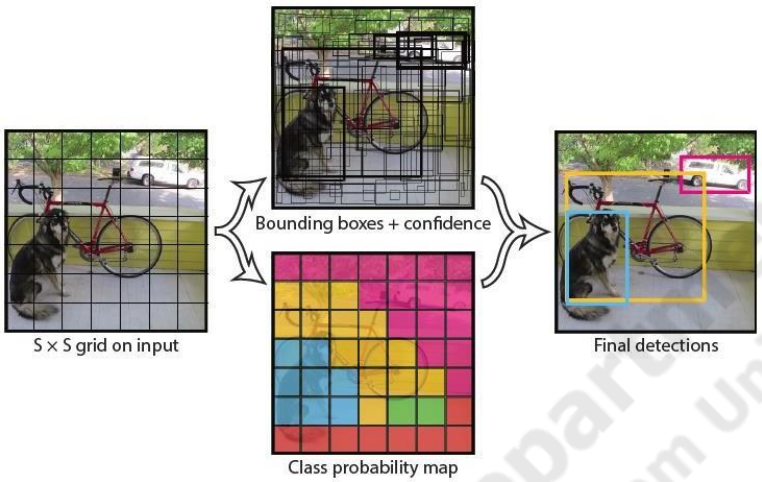
2.1.2 YOLO (You Only Look Once)

คือสถาปัตยกรรมที่ทาง ultralytics ได้ออกแบบไว้เพื่อทำ Image Detection ได้อย่าง รวดเร็ว และมีประสิทธิภาพ

ขั้นตอนการทำงานของ YOLO จะเป็นในลักษณะการแบ่งรูปภาพเป็นส่วนๆ หรือกริด (grid) จากนั้นทำการเลื่อนการคำนวนไปทีละจุดตามที่แบ่งกริดไว้ (Sliding Windows) พร้อมกับคำนวณ

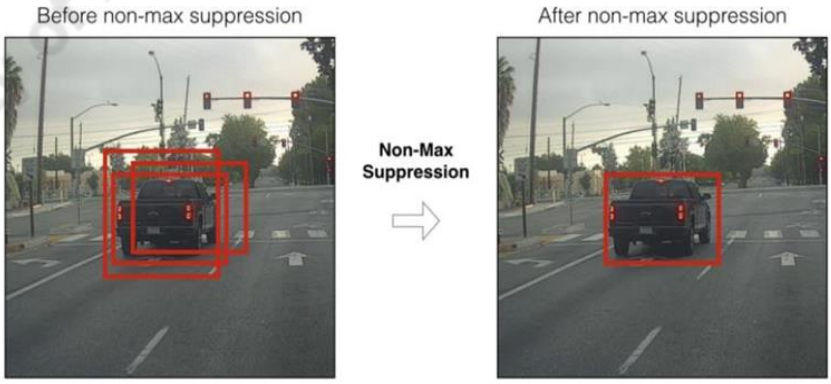
Computer Science Department
Faculty of Informatics, Mahasarakham University

จำแนกหาว่าวัตถุจะมีอยู่จริงหรือไม่จากความน่าจะเป็นของวัตถุที่ปรากฏ ในพื้นที่นั้นๆด้วยกระบวนการ Intersection over Union (IoU) เป็นการวัดประสิทธิภาพ ของโมเดล



ภาพประกอบที่ 2.2 การทำงานของอัลกอริทึม YOLO

ภาพประกอบที่ 2.3 การทำงาน Anchor box



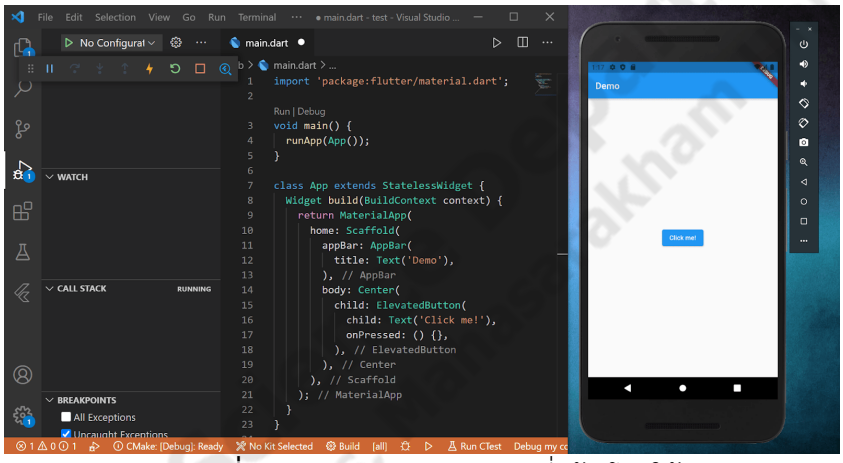
ภาพประกอบที่ 2.4 การทำงาน Non-max Suppression

ในการทำนายกรอบล้อมวัตถุจะได้ข้อมูลเป็นชุดข้อมูล ประเภทอาร์เรย์ ซึ่ง ประกอบด้วยข้อมูลการมีอยู่จริงของวัตถุตำแหน่งและขนาดของกรอบล้อมวัตถุ และชนิด ของวัตถุ กรณีถ้าในแต่ละกริดมี

วัตถุมากกว่าหนึ่งอย่าง YOLO มีกระบวนการ Anchor box เพื่อแก้ปัญหาดังกล่าว แนวคิดคือการสร้าง Anchor box ใน รูปทรงต่างๆ และคำนวณใหม่ เพื่อให้ได้ผลลัพธ์ที่ครอบคลุมและ แม่นยำมากขึ้น เมื่อระบบสร้างกรอบล้อมวัตถุจนหมดแล้ว จะเข้าสู่กระบวนการ Non-max Suppression ซึ่งขั้นตอนนี้จะเป็นการลดจำนวนกรอบล้อม วัตถุที่เป็นวัตถุเดียวกันถ้ากล่องใดมีค่านี้สูง (เกิน Threshold ของ iou ที่กำหนด) แสดงว่า มันคือ Object เดียวกัน

2.1.3 Flutter

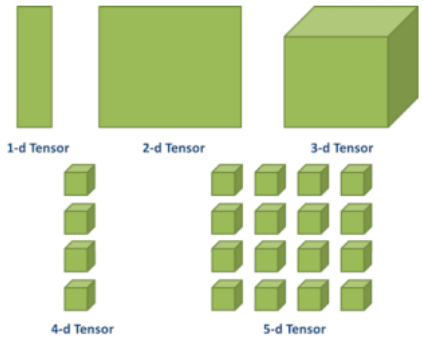
Flutter คือ Framework ที่ใช้สร้าง UI สำหรับ mobile application ที่สามารทำงานข้ามแพลตฟอร์มได้ทั้ง iOS และ Android ในเวลาเดียวกัน โดยภาษาที่ใช้ใน Flutter นั้นจะเป็นภาษา dart ซึ่งถูกพัฒนาโดย Google และที่สำคัญคือเป็น open source ที่สามารถใช้งานได้แบบ



ภาพประกอบที่ 2.5 ตัวอย่างแอปพลิเคชันที่สร้างโดยใช้ Flutter

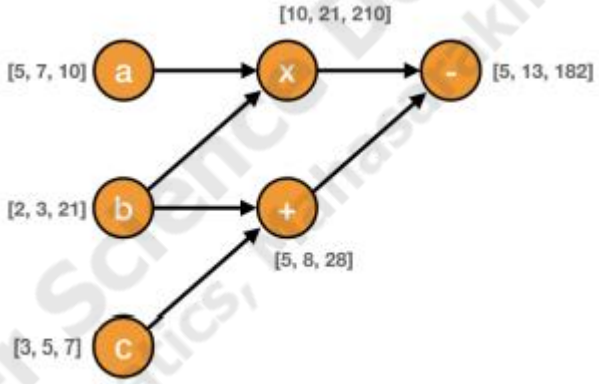
2.1.1 TensorFlow

TensorFlow เป็นไลบรารีโอเพนซอร์ซ (Library open source) ได้รับการ พัฒนาโดยทีม Google Brain ของบริษัท Google ได้ทำการเปิดตัวเมื่อวันที่ 11 กุมภาพันธ์ 2017 สามารถทำงานบน CPU และ GPUs รองรับระบบปฏิบัติการ Linux, macOS, Windows และ Android เป็นการทำงานสำหรับการคำนวณเชิงตัวเลข เป็นการสร้างกราฟกระแส ข้อมูลและโครงสร้าง เพื่อกำหนดการทำงานของข้อมูลผ่านกราฟ โดยรับข้อมูลโหนดเข้ามา เป็นอาร์เรย์หลายมิติที่เรียกว่าเทนเซอร์ (tensor) ไหลผ่านการเชื่อมต่อการทำงานและ แสดงผลข้อมูลแต่ละครั้ง 15 ส่วนประกอบเบื้องต้นของ TensorFlow คือ เทนเซอร์ เป็นเวกเตอร์ (vector) หรือเมตริกซ์ (Metrix) ของมิติ n ที่แสดง ข้อมูลประเภทเดียวกัน เป็นการคำนวณทั้งหมดที่เกี่ยวข้องกับเมตริกซ์ ที่มีการดำเนินงาน ในกราฟ



ภาพประกอบที่ 2.6 รูปร่างเทนเซอร์

กราฟ เป็นการคำนวณตัวดำเนินการทางคณิตศาสตร์ (operator) เช่น บวก ลบ คูณ หาร เป็นต้น ทั้งหมดภายในกราฟที่เรียกว่าโหนด (node) เป็นการ เชื่อมต่อเทนเซอร์เข้าด้วยกัน



ภาพประกอบที่ 2.7 กราฟที่มีการเชื่อมเทนเซอร์เข้าด้วยกัน

2.1.2 Flutter Text to Speech (TTS)

Flutter Text to Speech (TTS) คือเทคโนโลยีที่ช่วยแปลงข้อความเป็นเสียงในแอปพลิเคชันที่พัฒนาด้วย Flutter framework ซึ่งเป็นเครื่องมือสำหรับสร้างแอปพลิเคชันมัลติแพลตฟอร์มที่สามารถทำงานได้ทั้งบน Android และ iOS ด้วยโค้ดเดียวกัน การใช้งาน Flutter TTS ช่วยให้คุณสามารถให้แอปพลิเคชันของคุณอ่านข้อความออกเสียงหรือพูดในภาษาอื่น ๆ ได้อย่างง่ายดาย ขั้นตอนการใช้งานเบื้องต้นของ Flutter TTS:

1. การนำเข้า Flutter TTS: เริ่มด้วยการนำเข้าแพ็กเกจ flutter_tts ในโปรเจค Flutter ของคุณ นี้เป็นการเชื่อมต่อและใช้งาน TTS ในแอปของคุณ.

2. สร้าง Instance ของ Flutter TTS: คุณจะต้องสร้างอ็อบเจ็กต์ของ Flutter TTS ซึ่งเป็นตัวแทนในการควบคุมการอ่านข้อความออกเสียง.

Computer Science Department
Faculty of Informatics, Mahasarakham University

3. การอ่านข้อความ: คุณสามารถใช้ Flutter TTS เพื่ออ่านข้อความออกเสียง โดยระบุข้อความที่คุณต้องการให้อ่าน. คุณสามารถปรับแต่งการตั้งค่าเสียงเช่น ภาษา ระดับเสียง และอื่น ๆ ตามความต้องการ.

4. ควบคุมการอ่าน: คุณสามารถควบคุมกระบวนการอ่านเสียง เช่น หยุดการอ่านเสียงหรือรับข้อมูลเกี่ยวกับสถานะการอ่านเสียงปัจจุบัน.

5. ตัวอย่างการใช้งาน: สามารถใช้ Flutter TTS ในแอปพลิเคชันของคุณเพื่อสร้างประสบการณ์การอ่านข้อความออกเสียงเพื่อผู้ใช้ ตัวอย่างการใช้งานรวมถึงการอ่านข้อความสำหรับข้อความบนหน้าจอหรือสร้างเสียงเชิงสร้างสรรค์สำหรับฟีเจอร์ในแอปพลิเคชันของคุณ.

6. การปรับแต่ง: คุณสามารถปรับแต่งการตั้งค่าเสียงและส่วนต่าง ๆ ของการใช้งาน Flutter TTS เพื่อตอบสนองความต้องการของแอปพลิเคชันและผู้ใช้ของคุณ.

2.1.3 Flutter Speech to Text (STT)

Flutter Speech to Text (STT) เป็นเทคโนโลยีที่อนุญาตให้แอปพลิเคชันใน Flutter รับข้อมูลเสียงจากผู้ใช้และแปลงเสียงนั้นเป็นข้อความ เพื่อให้แอปพลิเคชันสามารถเข้าใจและประมวลผลข้อมูลที่เสียงบอกได้ ขั้นตอนการใช้งานเบื้องต้นของ Flutter Speech to Text:

1. เริ่มต้นโปรเจค: เริ่มโปรเจค Flutter ใหม่หรือใช้โปรเจคที่มีอยู่แล้ว และตรวจสอบว่าคุณมี dependencies สำหรับ Flutter Speech to Text ในไฟล์ pubspec.yaml.

2. นำเข้าแพ็กเกจ: ในโค้ด Dart ของแอปพลิเคชัน Flutter ของคุณ, คุณจะต้องนำเข้าแพ็กเกจ speech_to_text เพื่อให้แอปสามารถใช้งาน Speech to Text:

3. สร้าง Instance ของ Speech to Text: คุณจะต้องสร้างอ็อบเจ็กต์ของ Speech to Text ซึ่งเป็นตัวแทนในการควบคุมการรับข้อมูลเสียงและการแปลงเสียงเป็นข้อความ.

4. ตรวจสอบความพร้อมใช้งาน: ก่อนที่คุณจะเริ่มใช้งาน Speech to Text, คุณควรตรวจสอบว่าเทคโนโลยี STT พร้อมใช้งานบนอุปกรณ์ของผู้ใช้หรือไม่ ดังนั้นคุณจะไม่พบปัญหาการใช้งานขณะที่โปรแกรมพร้อมเข้ารับคำพูด.

5. เริ่มการรับข้อมูลเสียง: หลังจากตรวจสอบความพร้อมใช้งานแล้ว, คุณสามารถเริ่มการรับข้อมูลเสียงจากผู้ใช้โดยใช้เมธอด listen. แอปพลิเคชันจะรับข้อมูลเสียงและพยายามแปลงเป็นข้อความ.

6. หยุดการรับข้อมูลเสียง: คุณสามารถหยุดการรับข้อมูลเสียงโดยใช้เมธอด stop เมื่อคุณไม่ต้องการรับข้อมูลเสียงอีกต่อไป

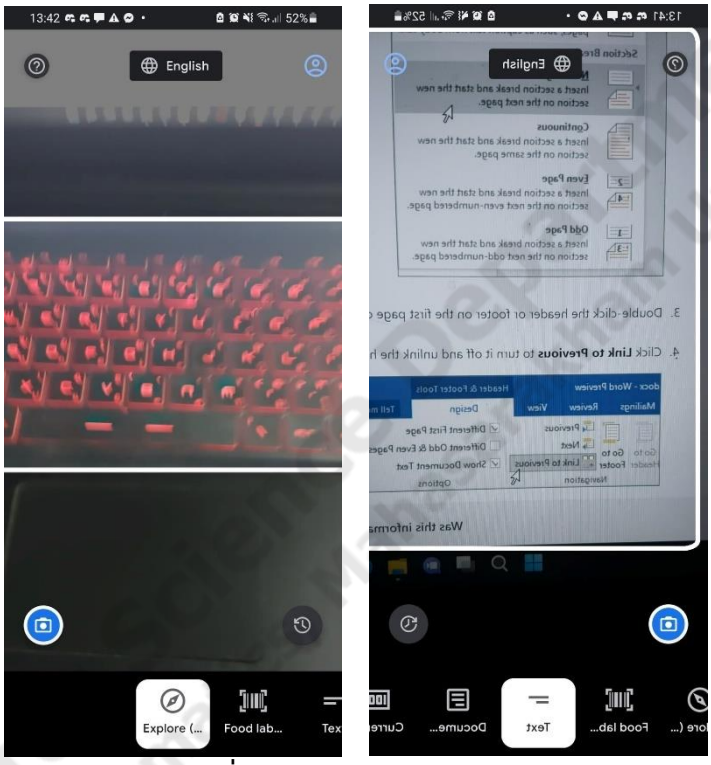
7. การใช้งาน Speech to Text ช่วยให้แอปพลิเคชันของคุณสามารถรับข้อมูลเสียงจากผู้ใช้และนำมาใช้งานเพื่อประโยชน์ต่าง ๆ ในแอปพลิเคชัน เช่น การพูดเพื่อค้นหา การแปลภาษา และฟีเจอร์อื่น ๆ ที่เกี่ยวข้องกับเสียงและข้อความในแอปของคุณ.

Computer Science Department
Faculty of Informatics, Mahasarakham University

2.2 ระบบงานที่เกี่ยวข้อง

2.2.1 Lookout

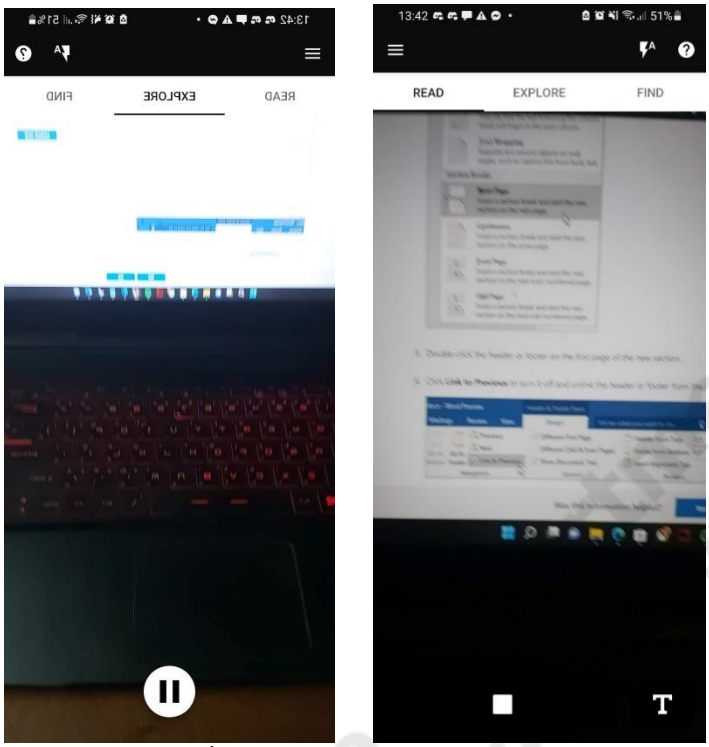
Lookout เป็นแอปพลิเคชันสำหรับช่วยเหลือคนพิการทางสายตาที่มีฟังก์ชันการใช้งานเยอะมากไม่ว่าจะเป็นการอ่านเอกสาร อ่านฉลากสินค้า อ่านบาร์โค้ดสินค้านับว่าเป็นแอปพลิเคชันที่เป็นประโยชน์อย่างมาก



ภาพประกอบที่ 2.8 ตัวอย่างแอปพลิเคชัน Lookout

2.2.2 Supersense

Supersense เป็นแอปพลิเคชันสำหรับช่วยเหลือคนพิการทางสายตาที่มีฟังก์ชันการใช้งานแค่สามอย่างคือ โหมดสำรวจ โหมดค้นหาที่ต้องกดเลือกหรือพิมพ์เอาและโหมดการอ่านข้อความ



ภาพประกอบที่ 2.9 ตัวอย่างแอปพลิเคชัน Supersense

2.3 ตารางเปรียบเทียบ

ตารางที่ 2.1 ตารางเปรียบเทียบการทำงานของระบบ

การทำงานของระบบ	Lookout	Supersense	ระบบที่กำลังพัฒนา
สมัครสมาชิก			
เข้าสู่ระบบ	✓		
ตรวจจับวัตถุ	✓	✓	✓
ระบุสีของสิ่งของที่ตรวจจับ			✓
อ่านข้อความ	✓	✓	
อ่านเอกสาร	✓		
ระบบสั่นเมื่อตรวจจับเจอสิ่งของ	✓	✓	✓
เสียงแจ้งเตือนเมื่อตรวจจับเจอสิ่งของ		✓	✓

ตารางที่ 2.1 ตารางเปรียบเทียบการทำงานของระบบ(ต่อ)

การทำงานของระบบ	Lookout	Supersense	ระบบที่กำลังพัฒนา
ตั้งค่าเสียง		✓	✓
ตั้งค่าการสั่น		✓	✓

Computer Science Department
Faculty of Informatics, Mahasarakham University