

# บทที่ 1

## บทนำ

### 1.1 หลักการและเหตุผล

ในปัจจุบันมีผู้พิการทางสายตาตั้งแต่กำเนิดและจากอุบัติเหตุ การสูญเสียการมองเห็นนั้นเป็นอุปสรรคในการดำเนินชีวิตอย่างมาก เช่น การต้องการหาสิ่งของภายในบ้าน หรือสิ่งของที่อยู่บนชั้นวาง แต่ไม่สามารถระบุได้ว่าอยู่ส่วนไหน และถ้าหากต้องการหาสิ่งของที่อยู่รวมกับของชิ้นอื่นภายในกล่อง อาจจะต้องใช้เวลานาน ดังนั้นการระบุสิ่งของภายในบ้านจึงเป็นประโยชน์ต่อผู้พิการทางสายตาที่อาจจะมีคนคอยช่วยเหลือในการหาสิ่งของภายในบ้าน

เนื่องจากของแต่ละชั้นมีลักษณะเฉพาะ แต่มีบางชั้นที่ลักษณะเหมือนหรือคล้ายกันแต่คุณสมบัติเฉพาะ เช่น ด้ายถักที่มีสีต่างกัน จึงเป็นเรื่องยากที่ผู้พิการทางสายตาจะแยกด้วยการสัมผัส เราจึงเล็งเห็นปัญหานี้จึงอยากนำปัญญาประดิษฐ์(Artificial Intelligence) และการเรียนรู้แบบอัตโนมัติ (Deep Learning) ด้วยอัลกอริทึมYOLO( You only look once) มาจำแนกประเภทสิ่งของใช้ภายในบ้านและคุณสมบัติเฉพาะสิ่งของ พร้อมยังนำ Speech Recognition มาช่วยสั่งการผ่านทางเสียงเพื่อความสะดวกในการใช้งาน

### 1.2 วัตถุประสงค์ของโครงการ

เพื่อพัฒนาแอปพลิเคชันที่ช่วยเหลือผู้พิการทางสายตา ตัวแอปพลิเคชันเป็นดังเหมือนตาให้แก่ผู้พิการที่สามารถหาและระบุสิ่งของเครื่องใช้ภายในบ้าน เช่น อาจเคยวางสิ่งของไว้ตรงนี้แต่กลับมาอีกครั้งของสิ่งนั้นกลับไม่อยู่ที่เดิมอีกแล้ว หรือไม่สามารถสัมผัสหาได้ มีโหมดระบุสิ่งของที่ต้องการจะหา โดยส่งข้อมูลเข้าไปในการพูดชื่อสิ่งของนั้นๆ หรือสิ่งของที่จะอยู่ในกล่องที่มีสิ่งของอื่นๆอยู่รวมปะปน ตัวแอปพลิเคชันจะสามารถบอกระบุตำแหน่งได้ โดยจะส่งเสียงเตือนเมื่ออยู่ใกล้วัตถุสิ่งของนั้นๆ

### 1.3 ขอบเขตของโครงการ

1.3.1 ชุดข้อมูล (Data set) ประกอบด้วยของสิ่งของภายในบ้าน 24 ชิ้น คือ

1. โต้ะ
2. แก้ว
3. จาน
4. ช้อน
5. ช้อม
6. มีด
7. แก้ว

8. แจกกัน
9. กุญแจ
10. รีโมททีวี
11. สายชาร์จโทรศัพท์
12. กล่อง
13. หนังสือ
14. หูฟัง
15. ไม้เท้า
16. พัดลม
17. แมสปิดปาก
18. กระเป๋า
19. ธนบัตร 1000 บาท
20. ธนบัตร 100 บาท
21. ธนบัตร 20 บาท
22. ธนบัตร 500 บาท
23. ธนบัตร 50 บาท
24. ปลั๊กไฟ

โดยแบ่งเป็น

- ข้อมูล Train อย่างน้อย 100 ภาพต่อสิ่งของภายในบ้าน
- ข้อมูล Test อย่างน้อย 30 ภาพต่อสิ่งของภายในบ้าน

1.3.2 คุณสมบัติเบื้องต้นที่ใช้ในการทำ Object Detection คือ

- ภาพสิ่งของ

1.3.3 อัลกอริทึมที่ใช้ในโปรแกรมเพื่อการสร้างโมเดลสำหรับการทำนาย คือ

- YOLO ( You only look once)

1.3.4 คุณสมบัติเบื้องต้นที่ใช้ในการทำ Speech Recognition คือ

- ชื่อสิ่งของ
- ชื่อแอปพลิเคชัน

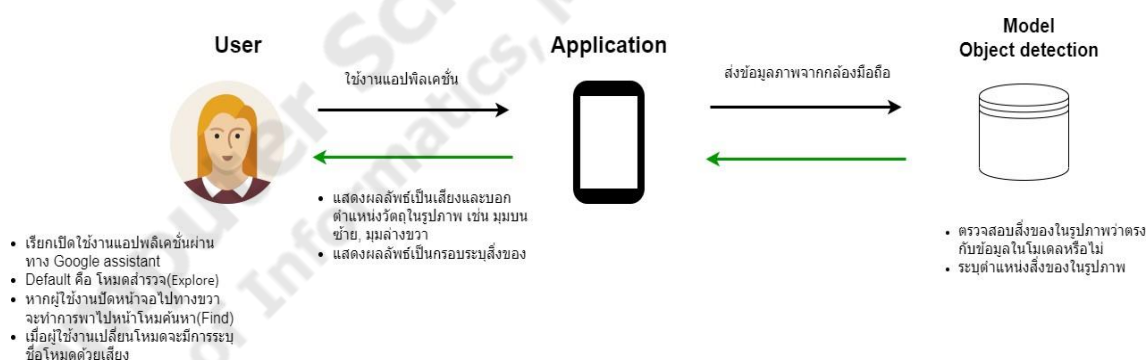
1.3.5 โมเดลที่ใช้ในการทำ Speech recognition คือ

- Wav2Vec2

### 1.3.6 ทำ Mobile application เพื่อทำการใช้งาน

1. โหมดสำรวจ(Explore) ตรวจสอบจับสิ่งของที่อยู่ด้านหน้าพร้อมแจ้งเตือนผ่านเสียง โดยระบุชื่อสิ่งของที่อยู่ด้านหน้า แต่จะไม่ระบุตำแหน่ง
2. โหมดการค้นหา(Find) ตรวจสอบเฉพาะสิ่งของที่ระบุพร้อมบอกตำแหน่ง โดยกดปุ่มแล้วจะมีเสียงแจ้งเตือนจากนั้นให้พูดสิ่งของที่ต้องการจะหา
  - หากตรวจพบจะแจ้งเตือนผ่านทางเสียง ชื่อสิ่งของและตำแหน่งในรูปภาพ เช่น “Cup Upper Right”
  - หากไม่ตรวจพบจะแจ้งเตือนผ่านทางเสียงว่า “Not detected.Please try again.”
3. แจ้งเตือนโหมดปัจจุบันบนแอปพลิเคชันผ่านทางเสียง
  - โหมดสำรวจ(Explore) แจ้งเตือนว่า “Explore mode”
  - โหมดการค้นหา(Find) แจ้งเตือนว่า “Find mode”
4. สั่งการเสียงได้เฉพาะ ภาษาอังกฤษ
5. เปิดแอปพลิเคชันด้วย Google assistant

### 1.4 ภาพรวมของระบบ



### ภาพประกอบที่ 1.1 การทำงานแอปพลิเคชัน

#### 1.4.1 ลักษณะการทำงานของระบบ

1. เปิดแอปพลิเคชันด้วย Google assistant
2. หน้าเริ่มต้นของแอปพลิเคชันคือหน้าโหมดสำรวจ(Explore)
  - มีเสียงแจ้งเตือน “Explore Mode” เพื่อแจ้งให้ผู้ใช้ทราบ
  - ปิดไปด้านขวาจะเป็น โหมดค้นหา(Find)

- ฟังก์ชันของโหมดสำรวจ(Explore) ระบุสิ่งของที่อยู่ด้านหน้าทุกชิ้น และแจ้งเตือนชื่อสิ่งของที่ตรวจจับ แต่จะไม่ระบุตำแหน่งสิ่งของ

### 3. หน้าโหมดค้นหา(Find)

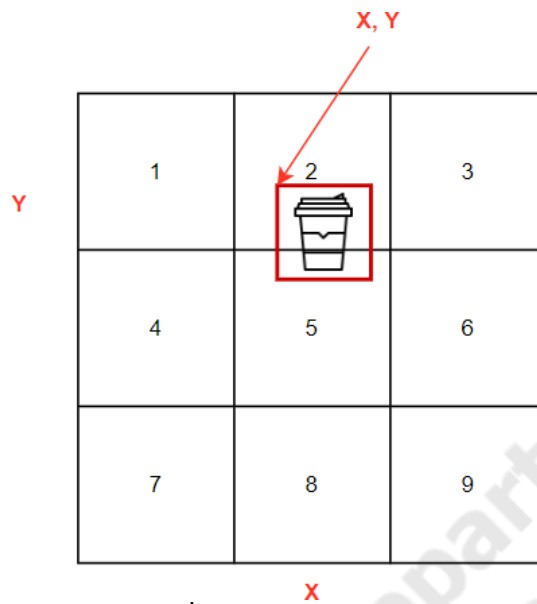
- จะมีเสียงแจ้งเตือนว่า “Find Mode” เพื่อแจ้งให้ผู้ใช้ทราบ
- ปิดไปด้านซ้ายจะเป็น โหมดสำรวจ(Explore)
- ฟังก์ชันของโหมดค้นหา(Find) ระบุเฉพาะสิ่งของที่ต้องการหา พร้อมระบุตำแหน่งสิ่งของ

- ผู้ใช้งานกดปุ่ม Speak เพื่อพูดสิ่งที่ต้องการหา ในการกดปุ่มจะมีการ แจ้งเตือนให้ผู้ใช้งานทราบเป็นเสียงเตือนพร้อมสั้น เช่น หากผู้ใช้งาน ต้องการหากระจกจะต้องทำการพูด “Mirror” แอปพลิเคชันจะระบุ เฉพาะกระจก หากตรวจสอบเจอ แอปพลิเคชันจะทำการแจ้งเตือน ชื่อสิ่งของและระบุตำแหน่ง เป็นเสียงว่า “Mirror, Upper Righth” ในกรณีที่ไม่พบจะแจ้งเตือนว่า “Not found. Please try again.”

### 4. การหาดำแหน่งสิ่งของในรูปภาพ

- เมื่อโมเดลตรวจเจอสิ่งของ จะส่งค่า (x, y) กลับคืนมา จากนั้นจะนำ ค่า x, y มาคำนวณเช็คหาตำแหน่งบริเวณในรูปภาพ จากนั้นจะแจ้ง เตือนตำแหน่งผ่านทางเสียงกลับไปยังผู้ใช้งาน

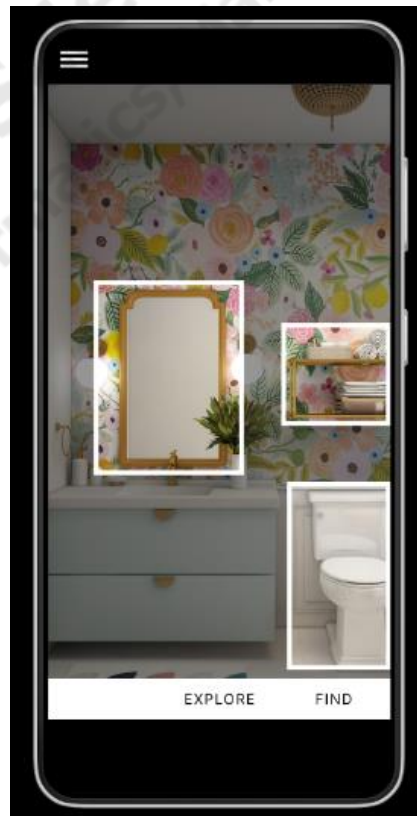
- ตำแหน่งในรูป
  1. Upper Left
  2. Upper Middle
  3. Upper Right
  4. Middle Left
  5. Middle
  6. Middle Right
  7. Lower Left
  8. Lower Middle
  9. Lower Right



ภาพประกอบที่ 1.2 การหาตำแหน่งในรูปภาพ

#### 1.4.1 ภาพรวมของแอปพลิเคชัน

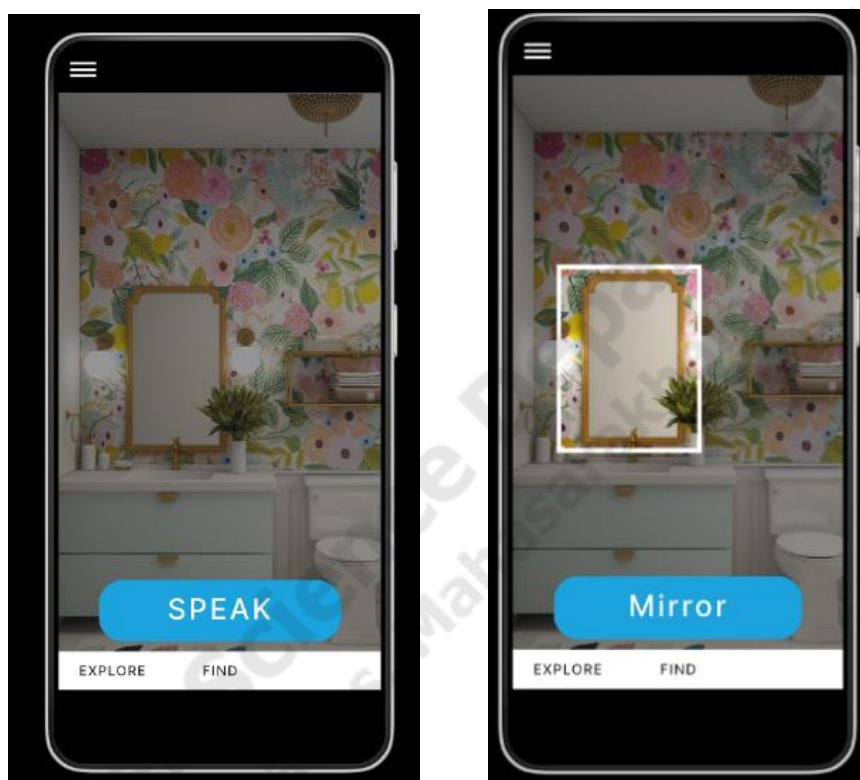
##### 1.4.1.1 หน้าโฮมสำรวจ(Explore)



ภาพประกอบที่ 1.3 หน้าโฮมสำรวจ

- หน้าแรกจะเป็นโหมดไม่ระบุสิ่งของ เป็นการตรวจจับทุกอย่างที่อยู่ด้านหน้า
- ปิดหน้าจอไปทางขวาจะเป็นโหมดระบุสิ่งของ

#### 1.4.1.2 หน้าโหมดการค้นหา(Find)



ภาพประกอบที่ 1.4 หน้าโหมดการค้นหา

- โหมดระบุสิ่งของ โดยรับอินพุตผ่านทางเสียง ในโหมดนี้จะช่วยหาสิ่งของได้ง่ายขึ้น จะทำการระบุสิ่งของนั้นๆ อย่างเดียว เพราะหากว่ามีของหลายๆ ชิ้นอยู่รวมกันจะทำให้การระบุยากขึ้น และในโหมดนี้สามารถบอกตำแหน่งผ่านเสียง หากของชิ้นนั้นอยู่ทางขวา ก็แจ้งว่า “ทางขวา”
- ปิดหน้าจอไปทางซ้ายจะกลับไปโหมดไม่ระบุสิ่งของ

#### 1.5 ประโยชน์ที่คาดว่าจะได้รับ

- 1.5.1 สามารถวิเคราะห์การจำแนกระบุสิ่งของภายในบ้าน
- 1.5.2 อำนวยความสะดวกให้ผู้พิการทางสายตาในการหาสิ่งของ

## 1.6 อุปกรณ์และเครื่องมือที่ใช้ในการดำเนินงาน

### 1.6.1 ฮาร์ดแวร์(Hardware)

#### 1.6.1.1 เครื่องคอมพิวเตอร์เครื่องที่ 1

1. Laptop MSI
2. Intel(R) Core(TM) i7-10750H CPU @ 2.60GHz 2.59 GHz
3. Ram 16.0 Gb
4. SSD 512 GB
5. NVIDIA GeForce GTX 1650 Ti with Max-Q Design

#### 1.6.1.2 เครื่องคอมพิวเตอร์เครื่องที่ 2

1. Personal Computer
2. Intel(R) Core(TM) i5-10400F CPU @ 2.90GHz 2.90 GHz
3. Ram 16.0 GB
4. SSD 240 GB
5. NVIDIA GeForce GTX 970

### 1.6.2 ซอฟต์แวร์(Software)

1. Python
2. Google Colab
3. Flutter
4. Android Studio
5. TensorFlow
6. YOLO
7. Wav2Vec2

## 1.7 แผนการดำเนินงาน

โครงการปริญญาโทฉบับนี้ ดำเนินงาน ณ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคามระหว่างเดือน ตุลาคม 2565 ถึง ตุลาคม 2566

