

สรุปการทำงาน week 4

มีการพัฒนาระบบตรวจจับพฤติกรรม “ยกของ” โดยมุ่งเน้นการทำงานแบบ **Real-Time Action Recognition** ด้วยการผสานหลายเทคโนโลยีหลัก ได้แก่ YOLOv11 สำหรับตรวจจับวัตถุ, MediaPipe Pose สำหรับตรวจจับจุดบนร่างกาย และการวิเคราะห์พฤติกรรมจากการเปลี่ยนแปลงตำแหน่งท่าทางของบุคคล (Heuristic-based)

รายละเอียดการทำงานของระบบ

1. ตรวจจับมนุษย์และวัตถุ (กล่อง)

- ใช้โมเดล YOLOv11 จำนวน 2 โมเดล:
 - human.pt สำหรับตรวจจับมนุษย์
 - box.pt สำหรับตรวจจับวัตถุที่อาจถูกยก (กล่อง)
- ทำการ track วัตถุและบุคคลแยกกัน โดยใช้ track_id และ object_id เพื่อเก็บลิงก์ระหว่างคนกับของที่เขากำลังยก

2. เก็บลักษณะท่าทางจาก MediaPipe Pose

- ตรวจจับ keypoints ของท่าทางจาก MediaPipe (เฉพาะ 22 จุดที่เกี่ยวข้องกับการเคลื่อนไหว เช่น แขน ขา หลัง)
- เก็บข้อมูลท่าทางแต่ละบุคคลแยกกันด้วย deque เพื่อใช้วิเคราะห์ความต่อเนื่อง

3. วิเคราะห์พฤติกรรมจากเปลี่ยนแปลงของท่าทาง

- คำนวณค่า std deviation ของ keypoints เพื่อแยกแยะพฤติกรรม:
 - ถ้าเปลี่ยนแปลงน้อย → แสดงว่าอยู่นิ่ง (idle)
 - ถ้ามีการเคลื่อนไหวมาก → แสดงว่าเคลื่อนไหว (moving)
- เมื่อวัตถุอยู่ในบริเวณเดียวกับร่างกายของบุคคลนั้น และ track อย่างต่อเนื่อง → ระบบจะเปลี่ยน label เป็น carrying

4. การแมตช์ระหว่างบุคคลและวัตถุ

- ระบบใช้ระยะห่างของจุดศูนย์กลาง (center point) ของ bounding box และการตรวจว่า "จุดของกล่อง" อยู่ใน box ของมนุษย์หรือไม่
- ใช้การเก็บสถิติความถี่ของการจับคู่ในแต่ละ frame เพื่อเพิ่มความแม่นยำ เช่น ถ้าคนเดียวกันแมตช์กับ object เดียวกันมากกว่า 15 ครั้ง → ถือว่าเป็น carrying

5. บันทึกผลการตรวจจับ

- เมื่อพฤติกรรมของบุคคลเปลี่ยน (จาก idle → carrying หรือในทางกลับกัน) ระบบจะทำการบันทึกข้อมูลลงฐานข้อมูล:
 - person_id, action, object_type, object_id, start_time, end_time
- ใช้โมเดล Action จากระบบฐานข้อมูลเพื่อเก็บประวัติการกระทำของแต่ละบุคคล

6. การแสดงผลแบบ Real-Time

- แสดงผล bounding box ของมนุษย์และวัตถุ
- แสดง Action, Confidence, Avg STD และข้อมูล ID บนจอวิดีโอ
- แสดง FPS และ Frame number เพื่อประเมิน performance

ผลลัพธ์เบื้องต้น

- ระบบสามารถตรวจจับและระบุการยกของได้แม่นยำในกรณีที่กำลังอยู่ใน bounding box ของบุคคล และบุคคลมีการเคลื่อนไหวสม่ำเสมอ
- การแมตช์ object-person มีความเสถียรเมื่ออยู่ในเฟรมต่อเนื่องมากพอ (≥ 15)
- ระบบสามารถบันทึก log ของการกระทำที่เปลี่ยนแปลงลงในฐานข้อมูลได้ครบถ้วน