Computer Vision Engineer Intern - Questions

Adsist Innovation | Candidate Version

คำชี้แจง: กรุณาตอบคำถามต่อไปนี้โดยอ้างอิงจากการทดลองและผลลัพธ์ที่คุณได้จาก assignment จริง คำตอบควรอธิบายเหตุผลและ trade-off

1) การแบ่ง Train/Validation

- คุณใช้**อัตราส่วนเท่าไหร่ในการแบ่ง train และ validation**? เพราะอะไร?
- ถ้าใน dataset มีรูปของรถยนต์คันเดียวกันหลายเฟรม เช่น เฟรมต่อเนื่องจากวิดีโอ คุณจะจัดการกับ data leakage ระหว่าง train/validation อย่างไร?

2) การเลือก Augmentation

อ้างอิงจาก Ultralytics training config ที่คุณใช้จริง ให้**ระบุ augmentation** ต่อไปนี้ว่าคุณเปิด, ปิด หรือปรับค่า พร้อมเหตุผลที่เฉพาะเจาะจงสำหรับงาน frontal car detection บน dataset 1,000 รูปนี้

- ตัวอย่างรายการ augmentation ในการพิจารณา: hsv_h, hsv_s, hsv_v, degrees, translate, scale, shear, perspective, flipud, fliplr, mosaic, mixup, copy_paste, erasing, auto_augment, close_mosaic

3) Failure Analysis

หลังเทรนเสร็จ ให้**รัน model บน validation set** แล้วเลือก 3-5 ตัวอย่างที่ทำนายแย่ที่สุด เช่น false positive, false negative หรือ IoU ต่ำ พร้อมชี้แจงและอธิบายเหตุผล

- แสดงรูปพร้อม prediction เทียบกับ ground truth
- ตั้งสมมุติฐานว่าทำไม model ถึงล้มเหลว เช่น occlusion, lighting, small object, wrong label, distribution gap
- ถ้ามีเวลามากกว่านี้ จะแก้ปัญหานี้อย่างไร

4) Confidence Threshold สำหรับ Inference

หลังเทรนเสร็จ ให้**รัน model บนวิดีโอ inference (Video.mp4)** ที่บริษัทจัดเตรียมให้

- คุณใช้ confidence threshold เท่าไหร่สำหรับวิดีโอ inference สุดท้าย และทำไม?
- หาก model ทำงานผิดพลาด เช่น ไม่ detect หน้ารถ, detect ผิดตำแหน่ง, หรือ confidence ต่ำผิดปกติในบางเฟรม Video ให้อธิบายสาเหตุที่เป็นไปได้และแนวทางปรับปรุง
- ทดสอบ 2 versions ที่ threshold ต่างกัน เช่น 0.25 และ 0.5 พร้อมอภิปราย trade-off ระหว่าง precision และ recall
- ถ้า model นี้ถูกใช้เพื่อ trigger การประมวลผล License Plate Recognition (LPR) downstream คุณจะเลือก threshold สูงหรือต่ำ? เพราะอะไร?

5) ถ้ามีเวลาเพิ่ม 1 เดือน และข้อมูลเพิ่ม 10 เท่า คุณจะเพิ่มประสิทธิภาพอย่างไร

- ให้ระบุอย่างชัดเจนว่าข้อมูลแบบไหนที่ต้องการเพิ่ม
- จะปรับ labeling หรือ dataset quality อย่างไร
- จะรัน experiment อะไรเพิ่มเติม
- จะวัดผลอย่างไรว่า model ดีขึ้นจริง