



บทคัดย่อ

ทีมวิจัยกลุ่มปัญญาประดิษฐ์ ได้สร้างสรรค์ผลงาน และบริการ API ต่าง ๆ ให้ผู้ที่สนใจสามารถเรียกใช้งานได้ผ่าน AI for thai เช่น บริการตัดคำจากประโยคภาษาไทย เป็นต้น เพื่อเป็นประโยชน์ต่อสาธารณะ เนื่องจากไม่บริการตรวจจับวัตถุที่สามารถให้บริการ จึงจัดทำโครงการสร้างโมเดลสำหรับการตรวจจับวัตถุจากรูปภาพ เพื่อเป็นบริการใน AI for thai โดยใช้โมเดล YOLO โดยเริ่มต้นศึกษาโมเดลและทำการเทรนโมเดลด้วย YOLOv3 ซึ่งพบว่าใช้เวลาในการเทรนที่มากเกินไป จนเกิดปัญหา Timming out บน google colab ตลอดต่อมาเปลี่ยนมาใช้ YOLOv4 และ YOLOv4-tiny ทำการรวบรวมรูปภาพจากเว็บไซต์ต่าง ๆ แล้วทำการติดฉลากคำตอบและแบ่งชุดข้อมูลออกเป็น 3 ส่วน คือ ชุดข้อมูลฝึกฝน ชุดข้อมูลตรวจสอบและ ชุดข้อมูลทดสอบ ผลการทดสอบโมเดล ความแม่นยำโดยรวม ของ YOLOv4 สูงกว่าตัว YOLOv4-tiny อยู่ที่ 0.95 ในอนาคตสามารถนำไปพัฒนาต่อเป็นการติดตามวัตถุ การแยกประเภทวิดีโอได้จากการติดแท็กให้กับวิดีโอ และการสร้างคำอธิบายให้กับรูปภาพได้

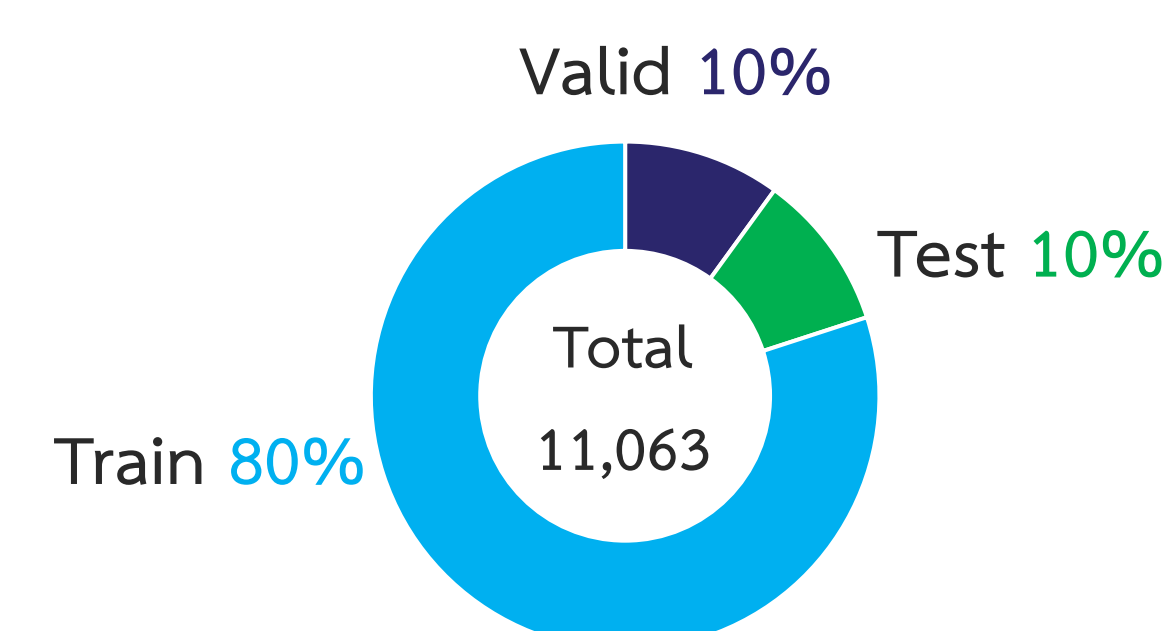
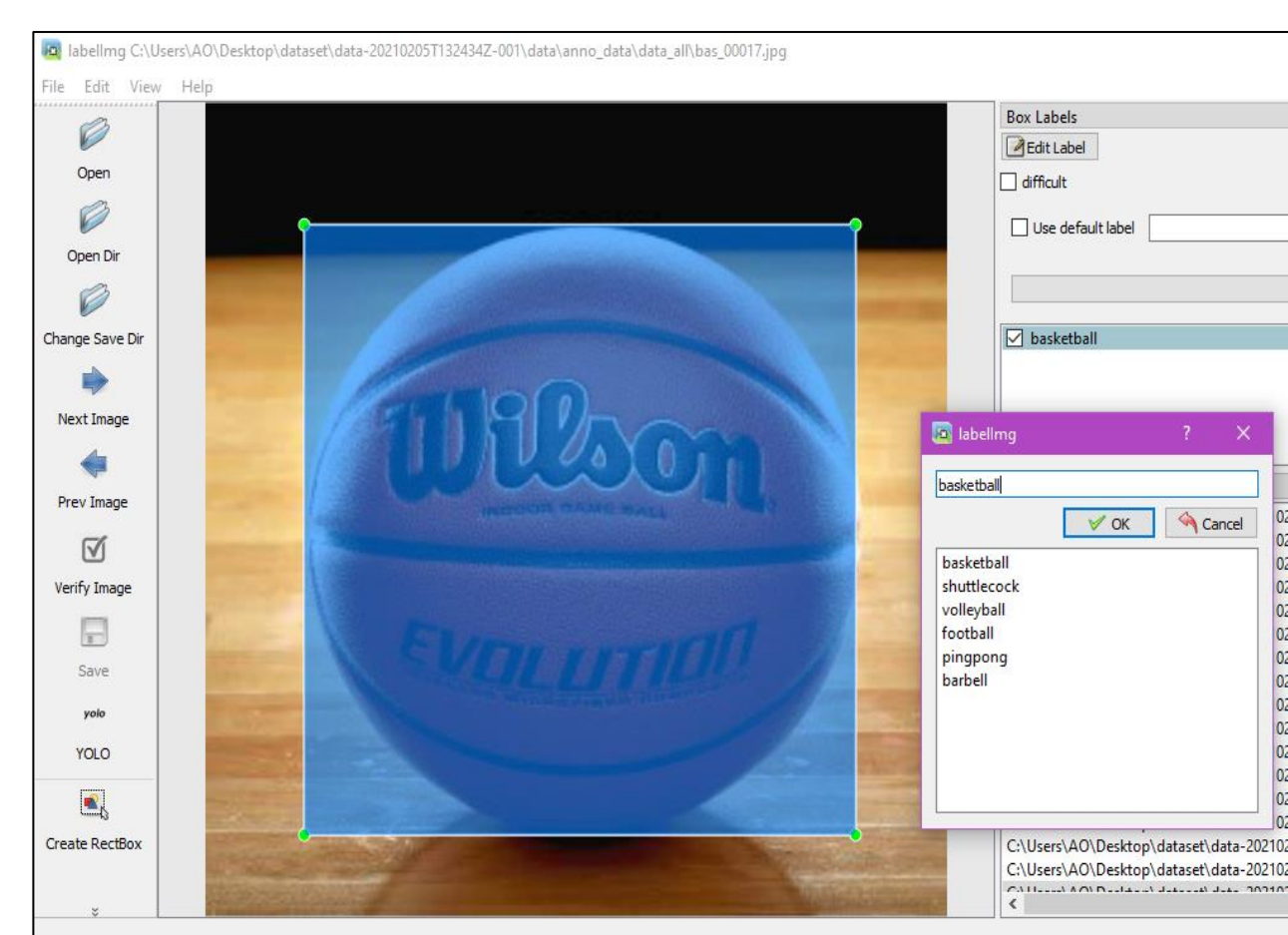
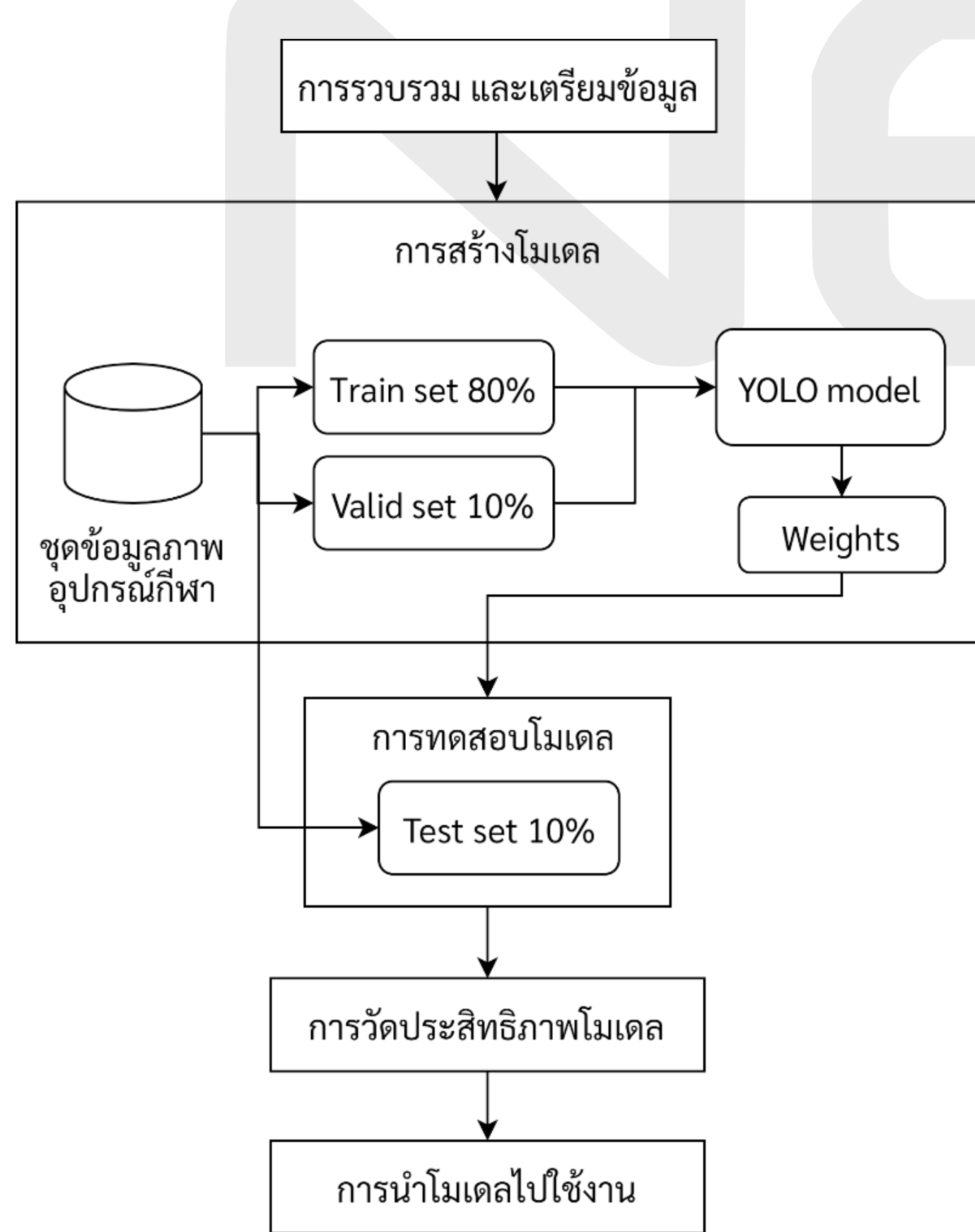
วัตถุประสงค์และขอบเขต

เพื่อตรวจจับอุปกรณ์กีฬา และระบุตำแหน่ง(Detection)ของอุปกรณ์ในภาพ โดยต่อยอดมาจากงานจำแนกประเภท(Classification)กีฬาจากภาพถ่าย

ตรวจจับอุปกรณ์กีฬาภายในภาพ จำนวน 10 อุปกรณ์ เพื่อสร้างโมเดลการตรวจจับวัตถุ โดยใช้ YOLOv4 และ YOLOv4 -tiny ได้แก่อุปกรณ์ ดังนี้

- 1) แบ๊สบาสเก็ตบอล
- 2) บาร์เบล
- 3) ลูกบาสเก็ตบอล
- 4) ไม้ปิงปอง
- 5) ไม้แบดมินตัน
- 6) ลูกขนไก่
- 7) โต๊ะปิงปอง
- 8) ลูกฟุตบอล
- 9) ลูกวอลเลย์บอล
- 10) ประตูฟุตบอล

ขั้นตอนการดำเนินงาน



การรวบรวมและเตรียมข้อมูล

ผู้ศึกษาได้รวบรวมภาพจากเว็บไซต์ต่าง ๆ และทำการปรับขนาดรูปภาพที่ขนาด 416x416 px จากนั้นทำการระบุชื่อและตำแหน่งวัตถุในภาพ โดยใช้เครื่องมือ Labelimg เพื่อสร้างฉลากคำตอบในรูปแบบ YOLO จำนวน 4,269 ภาพ จากทั้งหมด 10 คลาส แล้วทำการแบ่งข้อมูลออกเป็น 3 ส่วนคือ ชุดข้อมูลฝึกฝน(Train set) ชุดข้อมูลตรวจสอบ(Valid set) และชุดข้อมูลทดสอบ(Test set)

เนื่องจากข้อมูลฝึกฝนในบางคลาสมีจำนวนที่น้อยและลักษณะที่คล้าย กัน จึงใช้เทคนิคการทำ Data augmentation เพื่อเพิ่มจำนวนรูปภาพและความหลากหลาย เทคนิคที่ใช้คือ Flip Rotate 90 และ Rotation -15,15 ได้รูปภาพเพิ่มเป็น 11,063 ภาพ

ผลการทดสอบโมเดล

- ความแม่นยำโดยรวม(mAP) ของ YOLOv4 สูงกว่าตัว YOLOv4-tiny เล็กน้อย
 - ค่าความแม่นยำ(precision) YOLOv4 สูงกว่าตัว YOLOv4-tiny อยู่ 0.95 และ 0.86
- โดยภาพรวมแล้ว YOLOv4 สามารถตรวจจับอุปกรณ์กีฬาที่มีประสิทธิภาพมากกว่า YOLOv4-tiny

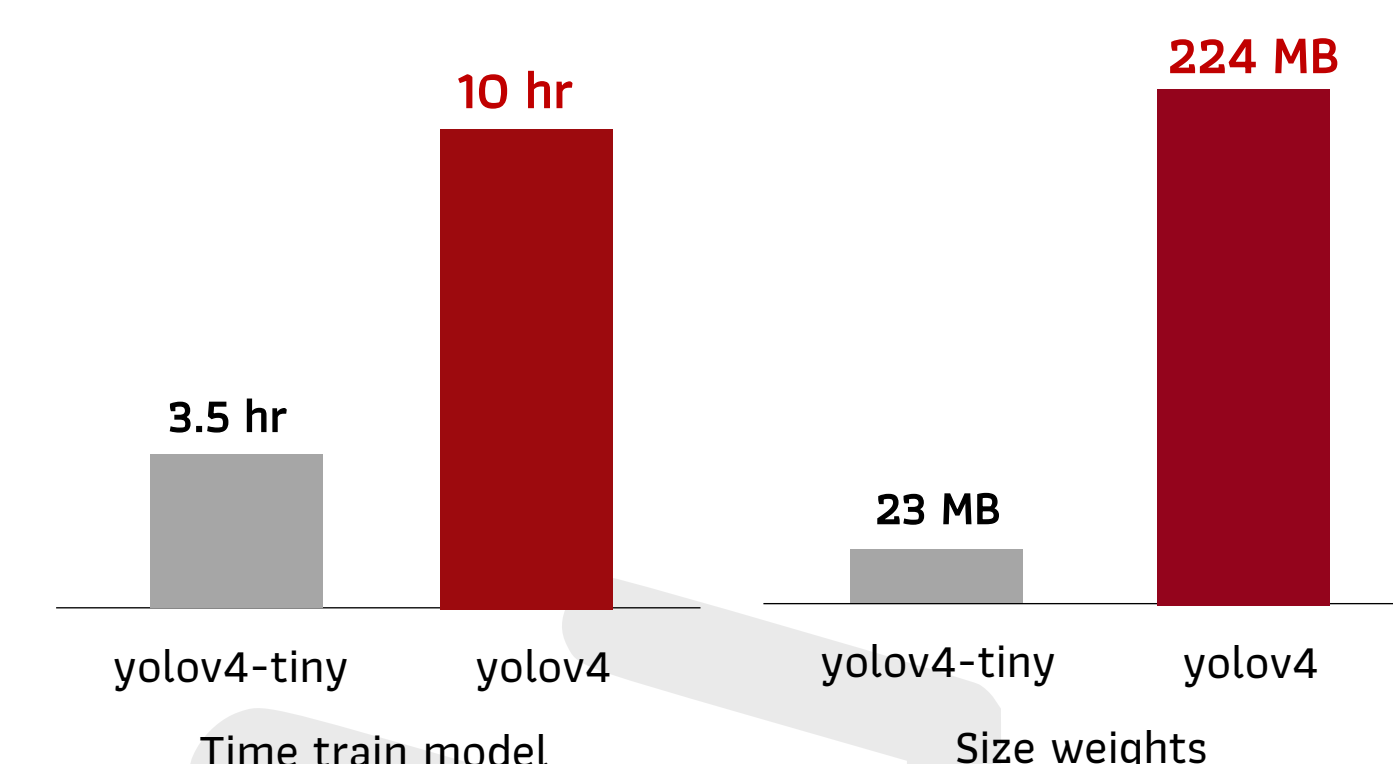
การสร้างโมเดล

การเขียนโปรแกรมทั้งหมดโดยใช้ภาษา Python ร่วม โมเดล YOLO Darknet บน Google Colab สามารถใช้ทรัพยากร GPU ได้ฟรี ช่วยแก้ปัญหาความเขากันโดของไลบรารี ต่าง ๆ ที่จำเป็นต้องใช้ในการเทรนโมเดล การที่จะทำการใช้โมเดล Darknet นั้นจะต้องเตรียมไฟล์และทำการโคลนไฟล์เตอร์จากเว็บไซต์ GitHub ได้จาก <https://github.com/AlexeyAB/darknet.git> สำหรับวิธีการฝึก (เพื่อตรวจจับวัตถุที่กำหนดเอง) ดังตาราง

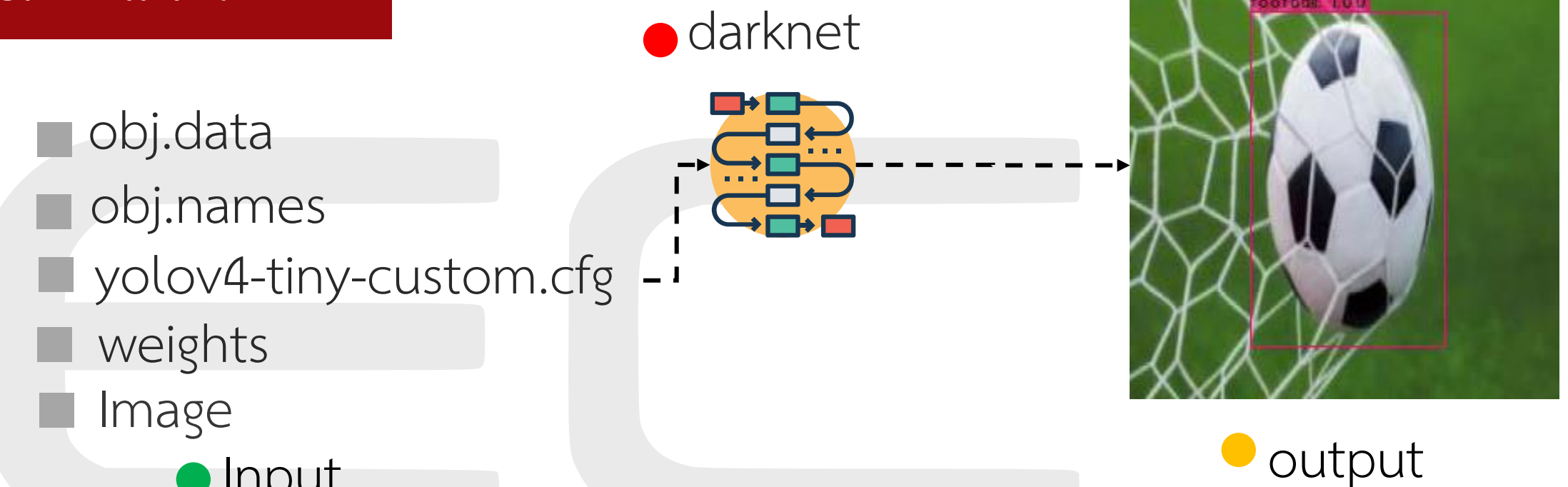
parameter	YOLOv4	YOLOv4-tiny
max_batches	20,000	20,000
Number of filters yolo layer	45	45
Number of output classes	10	10
ConvNet weights	Yolov4.conv.137	Yolov4.conv.29

การเทรนโมเดลจะใช้ชุดข้อมูลชุดฝึกฝน และตรวจสอบ และไฟล์เพื่อเชื่อมต่อกับ darknet โหลดไฟล์ที่เตรียมไว้และทำการดูค่าในระหว่างการเทรน โดยดูจากค่า training loss หากมีค่าที่เข้าใกล้ 0 มากที่สุดหรือค่าไม่ค่อยเปลี่ยนแปลง ให้ทำการหยุดเทรนได้ จากนั้นโมเดลทำการบันทึกค่าน้ำหนักที่ได้จากการเทรน

ผลจากการที่เทรนโมเดล YOLOv4 ค่า training loss เริ่มคั้งที่ประมาณรอบที่ 10,000 ค่า mAP อยู่ที่ 97.02% จึงทำการหยุดเทรนใช้เวลาทั้งสิ้น 10 ชั่วโมง ในขณะที่ YOLOv4-tiny ค่า training loss เริ่มคั้งที่ประมาณรอบที่ 20,000 ค่า mAP อยู่ที่ 94.80% จึงทำการหยุดเทรนใช้เวลาทั้งสิ้น 3.5 ชั่วโมง



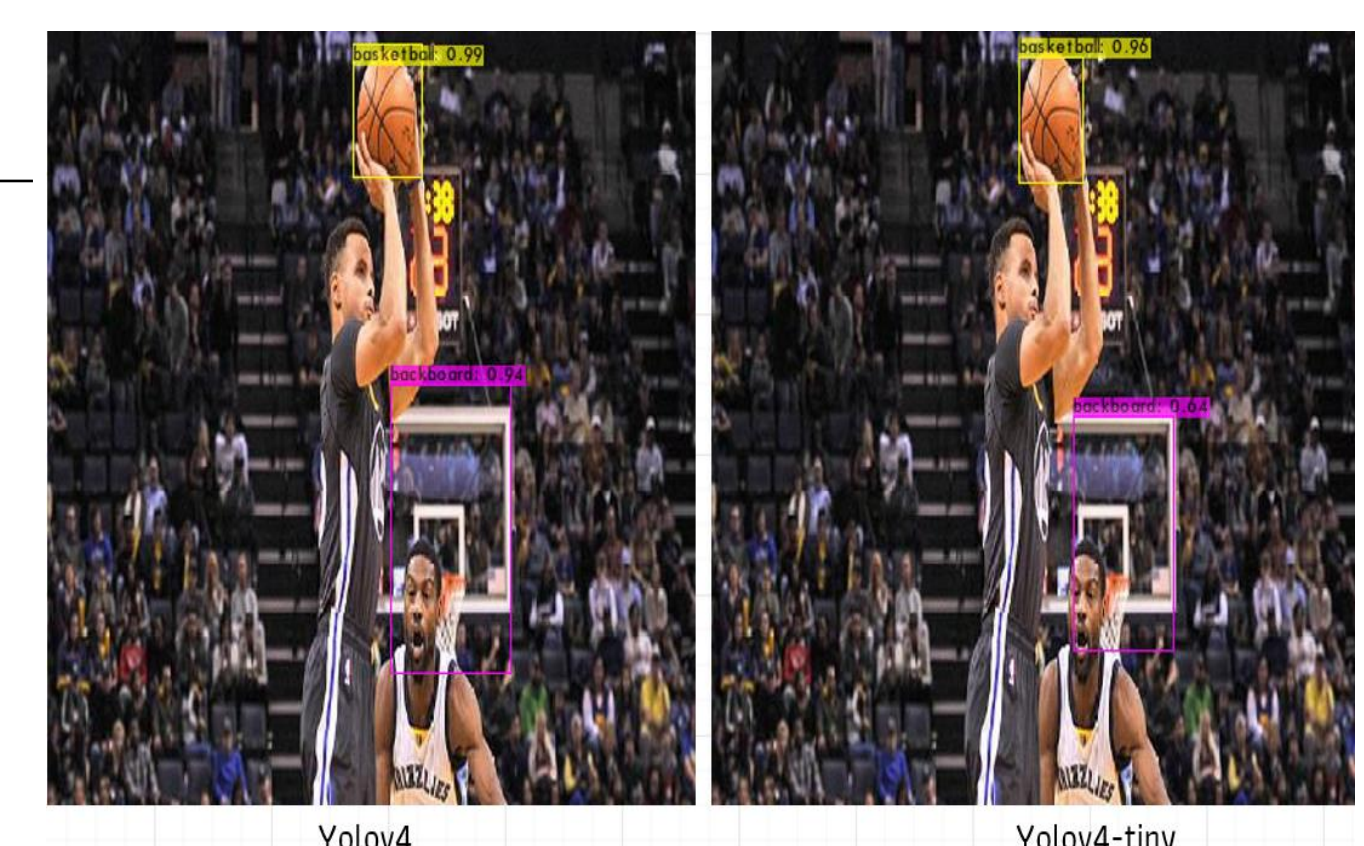
การวัดประสิทธิภาพโมเดล



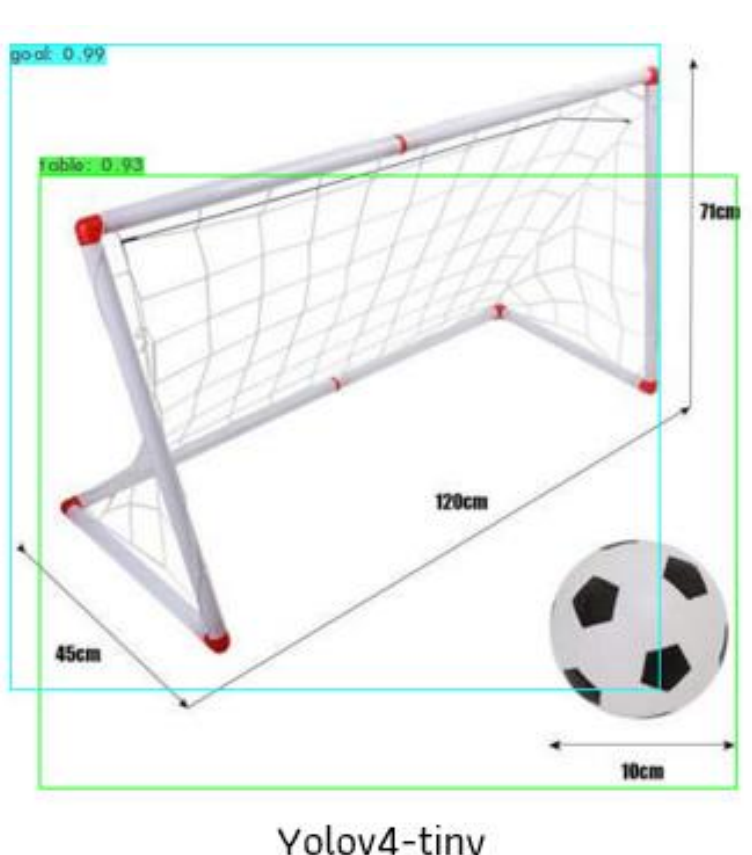
ผลการทดสอบเห็นได้ว่า YOLOv4 สามารถตรวจจับอุปกรณ์กีฬาที่มีความถูกต้องสูงกว่า อีกทั้งยังตีกรอบวัตถุตรงที่อุปกรณ์ได้ครอบคลุมกว่าด้านของ YOLOv4

YOLOv4 สามารถตรวจจับประตูฟุตบอลได้แม่นยำที่ระดับ 1.00 แต่ YOLOv4-tiny นั้นสามารถตรวจจับประตูฟุตบอลได้ที่ระดับ ซึ่งในความเป็นจริงแล้วไม่น่าจะเกิดขึ้นได้โมเดลอาจจะมีความผิดพลาดที่ในบางส่วนในรูปภาพจะคล้ายกับประตูฟุตบอล ทำให้โมเดลตรวจจับผิดพลาด

Performance	YOLOv4	YOLOv4-tiny
Precision	0.95	0.86
Recall	0.96	0.95
F1-score	0.95	0.90
Average IoU	80.24%	67.13%
mAP@0.50	94.83%	94.28%



ผู้ศึกษาได้เปรียบเทียบผลลัพธ์จาก YOLOv4 และ YOLOv4-tiny ด้วยข้อมูลและค่าพารามิเตอร์เดียวกัน ผลลัพธ์เป็นไปดังตาราง จะเห็นได้ว่า YOLOv4 มี AP สูงกว่า YOLOv4-tiny ค่า Precision ค่า Recall F1-score และมีค่า average IoU สูงกว่า อย่างไรก็ตาม YOLOv4 มี และ YOLOv4-tiny มีค่า mAP ใกล้เคียงกัน



สรุปผลและข้อเสนอแนะ

การสร้างโมเดลการตรวจจับอุปกรณ์กีฬาภายในภาพด้วยโมเดล YOLOv4 และ YOLOv4-tiny ทั้งหมด 10 อุปกรณ์ รวบรวมภาพจากเว็บไซต์ต่าง ๆ ทำการเทรนโมเดลบน google colab YOLOv4 ใช้เวลาในการเทรน 10 ชั่วโมง และ YOLOv4-tiny ใช้เวลาในการเทรน 3.5 ชั่วโมง

ข้อเสนอแนะ

- การนำโมเดลไปใช้ตรวจจับอุปกรณ์กีฬาด้วยวิดีโอ สามารถพัฒนาต่อเป็นการติดตามวัตถุ
- การนำโมเดลพัฒนาต่อเป็นการสร้างความอธิบายให้กับรูปภาพ
- การปรับจำนวนรูปภาพฝึกฝนให้มีความสมดุลเท่า ๆ กัน ทุกประเภทเพื่อป้องกันการเกิด Overfitting