



โครงการ ตรวจสอบความเป็นนักศึกษา

ชื่อกลุ่ม BK

นางสาวสุประภา	บุญนาค	5901108608 (หัวหน้า)
นายจิราวุธ	พรมโลก	1710111109016

เสนอ

อาจารย์ ดร.ศิวฤทธิ์ สุนทรเสถียร

โครงการนี้เป็นส่วนหนึ่งของวิชา BC410 สัมมนาทางคอมพิวเตอร์ธุรกิจ

ประจำภาคเรียนที่ 1 ปีการศึกษา 2562

คณะบริหารธุรกิจ สาขาวิชาคอมพิวเตอร์ธุรกิจ

มหาวิทยาลัยหอการค้าไทย

สมาชิกในกลุ่ม



นางสาวสุประภา บุณนาค 5901108608 (หัวหน้า)

E-mail : Id5901108608@gmail.com

หน้าที่ : Web Design



นายจิราวุธ พรมโลก 1710111109016

E-mail : jpkee094@gmail.com

หน้าที่ : Python Program

คำนำ

โครงการเล่มนี้เป็นส่วนหนึ่งของวิชา สัมมนาทางคอมพิวเตอร์ธุรกิจ (BC410) เป็นโครงการพัฒนาเสนอการตรวจสอบความเป็นนักศึกษา โดยมีวัตถุประสงค์ คือ การใช้ระบบนี้ตรวจสอบว่าใครแต่งกายเป็นนักศึกษา เมื่อเวลานักศึกษาเข้ามาในมหาวิทยาลัย ห้องเรียน การแต่งกายนั้นถูกต้องตามความเป็นนักศึกษาของมหาวิทยาลัยกำหนดไว้หรือไม่ ได้นำ AI หรือปัญญาประดิษฐ์มาใช้ และเขียนด้วยภาษา Python โดยใช้ Model SSD เข้าสู่กระบวนการ Machine Learning เพื่อค้นหาทำการวิเคราะห์ ที่ได้ผลลัพธ์ถูกต้อง และแม่นยำที่สุดออกมาโดยอัตโนมัติ

คณะผู้จัดทำหวังว่าโครงการเล่มนี้จะให้ความรู้และเป็นประโยชน์กับผู้อ่านทุก ๆ ท่าน หากโครงการนี้มีข้อบกพร่องประการใด คณะผู้จัดทำต้องขออภัยและจะนำไปปรับปรุงแก้ไขในโอกาสต่อไป

คณะผู้จัดทำ

กลุ่ม BK

สารบัญ

สมาชิกในกลุ่ม	ข
คำนำ	ค
สารบัญ	ง
สารบัญภาพ	ฉ
บทที่ 1 บทนำ	1
1.1 ที่มาและความสำคัญ	1
1.2 วัตถุประสงค์	2
1.3 ขอบเขตการทำงาน	2
1.4 ขั้นตอนการดำเนินงาน	2
1.5 ประโยชน์ที่ได้รับ	3
1.6 ระยะเวลาการดำเนินงาน	4
1.7 ผู้รับผิดชอบโครงการ	4
บทที่ 2 หลักการและงานวิจัยที่เกี่ยวข้อง	5
2.1 Object Recognition	5
2.2 (SSD) Object Detection: Single Shot MultiBox Detector for real-time processing	6
2.3 OpenCV	25
2.4 งานวิจัยที่เกี่ยวข้อง	34
2.5 Flow Chart	41
บทที่ 3 วิธีการดำเนินงาน	44
3.1 วงจรการพัฒนาาระบบ (System Development : SDLC)	44
3.2 ตัวอย่าง วิธีทำการตรวจสอบความเป็นนักศึกษา	48
3.3 วิธีการใช้ Flask กับ Unicorn และ Nginx บน Ubuntu 18.04	60

บทที่ 4 ผลการดำเนินงาน	74
4.1 การใช้งานระบบตรวจสอบความเป็นนักศึกษา	74
บทที่ 5 สรุปผลการดำเนินงาน	80
5.1 สรุปผลการทำโครงการ	80
5.2 ปัญหาที่พบในการดำเนินการ	80
5.3 ข้อเสนอแนะ	80
บรรณานุกรม	81

สารบัญภาพ

รูปที่ 3.1	ภาพแสดงวงจรการพัฒนา ระบบ (System Development : SDLC)	47
รูปที่ 3.2	ทำการเก็บรูปภาพนักศึกษาที่แต่งกายถูกระเบียบและไม่ถูกระเบียบ	48
รูปที่ 3.3	รูปภาพที่ Crop แล้ว	51
รูปที่ 4.1	แสดงหน้าโปรแกรม Visual Studio Code	74
รูปที่ 4.2	แสดงถึงหน้ากำลังจะทำการ Run	75
รูปที่ 4.3	ระบบกำลังแสดงการตรวจสอบความเป็นนักศึกษา	76
รูปที่ 4.4	ภาพที่แสดงว่าเป็นนักศึกษา	77
รูปที่ 4.5	ภาพที่แสดงว่าไม่ใช่นักศึกษา	78
รูปที่ 4.6	หน้าหลักของเว็บไซต์	79
รูปที่ 4.7	ระบบตรวจสอบความเป็นนักศึกษานบนเว็บไซต์	79

บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญ

ในปัจจุบันมีการนำกล้องวิดีโอมาเป็นอุปกรณ์ตรวจจับ วัตถุประสงค์เพื่ออำนวยความสะดวกให้มนุษย์สามารถพบได้ใน ระบบตรวจรักษาความภัยบริเวณลานจอดรถ (Robot Navigation), การเชื่อมต่อระหว่างภาพเสมือนจริง (Virtual Reality Interface) , การลำเลียงสิ่งของในโรงงาน, ตรวจสอบหาความผิดปกติ, ระบบตรวจวัตถุต้องสงสัยโดยเฉพาะการตรวจเช็ควัตถุต้องสงสัย ของสาเหตุการเกิดโจรกรรมเหตุร้ายต่าง ๆ บนท้องถนน เป็นต้น และเพื่อให้บรรลุสู่ความสำเร็จ งานเหล่านี้ล้วนแล้วแต่ต้องการความสามารถของการประมวลผลภาพ การติดตามวัตถุรวมถึง การจำแนกประเภทวัตถุทั้งสิ้น

ระบบการจำแนกประเภทวัตถุ (Objects Classification System) ได้ถูกประยุกต์ใช้งานในหลายด้านมาเป็นเวลานานแล้ว สิ่งที่ตรวจจับได้ประกอบ เป็นวัตถุที่สนใจซึ่งถูกระบุด้วยตำแหน่งและความเร็ว, แหล่งกำเนิดสัญญาณรบกวนภาคพื้นหลัง (Background Noise Source) สภาพแปรปรวนของอากาศ, ทัศนวิสัย, อุณหภูมิความร้อน (Thermal Noise) ทำให้ระบบการจำแนกประเภทวัตถุจึงเข้ามามีบทบาท เกี่ยวกับระบบรักษาความปลอดภัยโดยเฉพาะภาคพื้นดิน เช่น ในเมืองที่มีประชากรอยู่หนาแน่น มี การจราจรคับคั่ง และมีพาหนะเคลื่อนที่ชนิดต่าง ๆ สันจรผ่านไปมาบนท้องถนน

ปัจจุบัน ปัญหาการแต่งกายของนักศึกษา ยังเป็นปัญหาที่ไม่สามารถควบคุมได้ ชุตนักศึกษาเป็นชุดที่แสดงถึงการมีระเบียบ วินัย เรียบร้อย แต่ละมหาวิทยาลัยต่างก็มีชุดนักศึกษาที่แสดงถึงสถาบันของตน ซึ่งอีกนัยยะหนึ่งถือเป็นการแบ่งแยกกลุ่มของนักศึกษา เครื่องแบบก็อาจจะเป็นตัวช่วยเตือนสติว่ากำลังศึกษาเล่าเรียนอยู่ การตรวจสอบความเป็นนักศึกษาเป็นส่วนหนึ่งของเรื่องการจำแนกประเภทวัตถุ ผู้จัดทำโครงการจึงได้เล็งเห็นถึงความสำคัญความถูกต้องของการแต่งกายนักศึกษาที่ถูกต้องตามระเบียบแบบแผนหรือกฎระเบียบที่ตั้งไว้ และยังสามารถนำไปพัฒนาใช้ได้กับทุก ๆ ด้าน ที่เกี่ยวกับเครื่องแต่งกาย เช่น โรงงานอุตสาหกรรมที่ต้องการตรวจสอบความถูกต้องของเครื่องแบบการแต่งกายของพนักงาน ว่าพนักงานแต่งกายถูกต้องตามระเบียบหรือไม่ และเพื่อความปลอดภัยในการทำงานของตัวพนักงานเอง หรือในด้านอื่น ๆ ก็ได้

1.2 วัตถุประสงค์

- 1.2.1 เพื่อศึกษาการทำงานของระบบตรวจสอบความเป็นนักศึกษา
- 1.2.2 เพื่อพัฒนาโปรแกรมตรวจสอบความเป็นนักศึกษาให้มีความแม่นยำว่าถูกหรือผิดในการแต่งกายชุดนักศึกษา
- 1.2.3 เพื่อวิเคราะห์ความถูกต้องการแต่งกายของนักศึกษา ด้วยวิธีการประมวลผลภาพ

1.3 ขอบเขตการทำงาน

- 1.3.1 ระบบที่พัฒนา หรือจัดทำขึ้นนั้นเพื่อใช้ในการตรวจสอบความเป็นนักศึกษา
- 1.3.2 ด้านโปรแกรมคอมพิวเตอร์หรือซอฟต์แวร์ (Software)
 - 1.3.2.1 Python ใช้เป็นภาษาหลักในการพัฒนาระบบ
 - 1.3.2.2 Visual Studio Code
- 1.3.3 ด้านอุปกรณ์คอมพิวเตอร์หรือฮาร์ดแวร์ (Hardware)
 - 1.3.3.1 หน่วยประมวลผล (CPU) Intel Core (TM) i5-8300H 2.30 GHz
 - 1.3.3.2 หน่วยความจำหลัก (RAM) 8.00 GB
 - 1.3.3.3 ระบบปฏิบัติการ (OS), Microsoft Windows 10 (64-bit)
- 1.3.4 กลุ่มเป้าหมาย
 - 1.3.4.1 บริษัทภาครัฐ และเอกชน หรือหน่วยงานราชการ รวมไปถึงโรงงานอุตสาหกรรม ที่มีความต้องการตรวจสอบความถูกต้องของเครื่องแบบการแต่งกายของพนักงาน ว่าพนักงานแต่งกายถูกต้องตามระเบียบหรือไม่

1.4 ขั้นตอนการดำเนินงาน

- 1.4.1 ปรึกษาอาจารย์ผู้สอนโครงการเกี่ยวกับขอบเขตของโครงการที่จะทำ
- 1.4.2 ศึกษาค้นคว้างานวิจัยที่ใกล้เคียงกับโครงการ
- 1.4.3 ศึกษาข้อมูลเกี่ยวกับอุปกรณ์ที่ต้องใช้ในโครงการ ได้แก่ กล้อง Camera Webcam C310 HD และทำการซื้ออุปกรณ์

1.4.4 ออกแบบและเขียนโปรแกรมตามวัตถุประสงค์ที่ตั้งไว้

1.4.5 ทดลองใช้งานและแก้ไขสิ่งที่มีผิดพลาด ทั้งสรุปผลการทดลอง

1.4.6 จัดทำรูปเล่มรายงานของโครงการเพื่อเสนออาจารย์ประจำสาขาวิชา

1.5 ประโยชน์ที่ได้รับ

1.5.1 ได้เรียนรู้การทำงานของระบบตรวจสอบความเป็นนักศึกษา

1.5.2 สามารถนำระบบไปช่วย หรือทำประโยชน์ให้แก่บริษัท องค์กร และคนที่มีความสนใจ ในด้านที่
อยากจะทำ การตรวจสอบความถูกต้องของเครื่องแต่งกาย

1.5.3 ระบบสามารถบอกลักษณะได้ว่า บุคคลที่ยืนอยู่หน้ากล้องนั้น แต่งกายใช่นักศึกษาหรือไม่

1.5.4 ได้เรียนรู้การทำงานเป็นทีมและได้นำความรู้ที่ได้จากการศึกษาภาคทฤษฎีมาประยุกต์ใช้งานจริง

1.6 ระยะเวลาการดำเนินงาน

ขั้นตอนในการดำเนินงาน	ระยะเวลา (ต่ออาทิตย์ พ.ศ. 2562)															
	สิงหาคม				กันยายน				ตุลาคม				พฤศจิกายน			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1 ปรึกษาอาจารย์ผู้สอน โครงการเกี่ยวกับขอบเขตของ โครงการที่จะทำ																
2 ศึกษาค้นคว้างานวิจัยที่ ใกล้เคียงกับโครงการ																
3 ศึกษาข้อมูลเกี่ยวกับอุปกรณ์ ที่ต้องใช้ในโครงการ ได้แก่ กล้อง Camera Webcam C310 HD และทำการซื้อ อุปกรณ์																
4 ออกแบบและเขียนโปรแกรม ตามวัตถุประสงค์ที่ตั้งไว้																
5 ทดลองใช้งานและแก้ไขสิ่งที่ ผิดพลาด ทั้งสรุปผลการ ทดลอง																
6 จัดทำรูปเล่มรายงานของ โครงการเพื่อเสนออาจารย์ ประจำสาขาวิชา																

1.7 ผู้รับรับผิดชอบโครงการ

1.7.1 นางสาวสุประภา บุนนาค (Web Design)

1.7.2 นายจิราวุธ พรมโลก (Python Program)

บทที่ 2

หลักการและงานวิจัยที่เกี่ยวข้อง

ในบทนี้เป็นการนำเสนอเนื้อหาที่ผู้จัดทำโครงการจะกล่าวถึงหลักการและงานวิจัยที่เกี่ยวข้องกับโครงการ
ชิ้นนี้ ซึ่งการจะทำโครงการชิ้นนี้จำเป็นต้องทำความเข้าใจเกี่ยวกับกระบวนการทำงานของ Object
Detection เพื่อนำมาใช้ประกอบขั้นตอนการทำงานในหัวข้อเรื่อง ตรวจสอบความเป็นนักศึกษา ตามลำดับ
ความสำคัญดังนี้

2.1 Object Recognition

เป็นเทคนิคการมองเห็นคอมพิวเตอร์เพื่อระบุวัตถุในภาพหรือวิดีโอ การรับรู้วัตถุเป็นผลผลิตที่สำคัญของ
deep learning และ machine learning algorithms เมื่อมนุษย์ดูรูปถ่ายหรือดูวิดีโอ สามารถมองเห็นคน วัตถุ
ฉากและรายละเอียดทางสายตาได้อย่างง่ายดาย วัตถุประสงค์คือการสอนคอมพิวเตอร์ให้ทำสิ่งที่เกิดขึ้นตาม
ธรรมชาติกับมนุษย์เพื่อให้ได้ระดับความเข้าใจในสิ่งที่รูปภาพที่กำลังแสดงอยู่

เทคโนโลยีตรวจจับ (Detection Technology)

ได้นำมาประยุกต์ใช้กันอย่างกว้างขวางใน อดีต ปัจจุบัน และยังมีแนวโน้มที่จะพัฒนาขึ้นใช้กับโครงข่าย
สื่อสารเพิ่มขึ้น ตัวตรวจจับ สัญญาณและเครื่องมือวัดชนิดต่าง ๆ จึงได้รับการพัฒนาขึ้นมาโดยลำดับ ได้แก่ ตัว
ตรวจจับอุณหภูมิ หรือเทอร์โมมิเตอร์ ตัวตรวจจับความชื้น (Hygrometer) ตัวตรวจจับความเร็ว ตัวแนวคิดในการ
ตรวจจับการเคลื่อนที่ที่มีการนำไปใช้เพื่อสร้างหุ่นยนต์สำหรับใช้งานตาม วัตถุประสงค์ที่ต้องการ บางกรณีใช้หุ่นยนต์
เพื่อเลียนแบบการเคลื่อนที่ตามมนุษย์ โดยทำการประมวลผลรวมจากข้อมูลที่ตรวจจับแล้วนำข้อมูลทั้งหมดไป
วิเคราะห์ ผลซึ่งเป็นส่วนของกรณีเวลาจริง (Real Time) ซึ่งยังมีปัญหาในด้านความถูกต้องของ การเคลื่อนที่

ปกติการตรวจจับวัตถุให้ละเอียดขึ้นจะใช้กล้องหลายตัว ใช้พิจารณาตึกรี ความเป็นอิสระของส่วนเคลื่อนที่
ด้วยการลดจำนวนจุดเคลื่อนที่งานของ ใช้ภาพขา เข้าเพื่อลดพื้นที่การค้นหา ทำให้ค้นหาจุดเคลื่อนที่ได้เร็วขึ้น
ประกอบกับการใช้เทคนิคขอบภาพ และสี ระบบสามารถจับภาพสามมิติโดยไม่ต้องใช้การตัวระบุตำแหน่ง

ส่วนการใช้กล้องเกี่ยวเน้นที่การได้ปริมาณข่าวสารจำนวนมากจากลำดับภาพ การแบ่ง บริเวณของการ
เคลื่อนที่ การใช้จุดใกล้สุดวนซ้ำ (Iterative Closest Point) การตรวจจับขอบ (Edge Detection) และการค้นหา
แบบลำดับขั้น (Hierarchical Search)

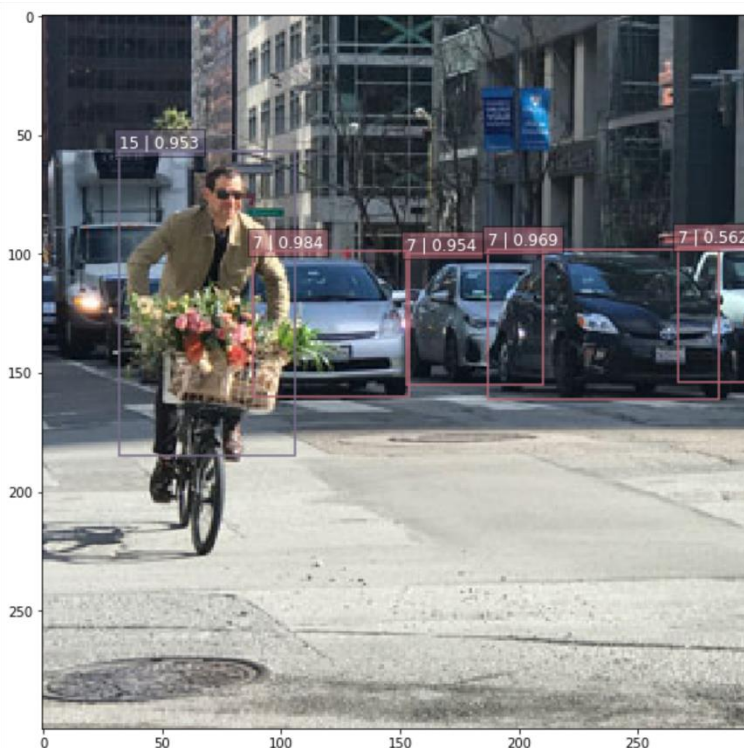
การตรวจจับขอบสามารถช่วยในการติดตามการเคลื่อนที่ของมนุษย์ ถ้ามีปริมาณจุดข้อมูล เพียงพอ ซึ่งโดยทั่วไปแล้วควรจะมีย่านอย่างน้อยหนึ่งล้านจุด ดังนั้นการพิจารณาการเคลื่อนที่ของมนุษย์ อาจต้องมีการกำหนดพื้นที่และการใช้เทคนิคขอบภาพร่วมกับการกำหนดฟังก์ชัน ซึ่งการกำหนด พื้นที่จะใช้การคำนวณที่ซับซ้อน

งานวิจัยที่นำเสนอเป็นการสร้างหน่วยตรวจจับการเคลื่อนที่ของวัตถุ โดยใช้พารามิเตอร์ ตรวจวัดคือ ระยะทาง เวลา ความเร็ว และความไว (Sensitivity) โดยใช้ตัวตรวจจับภาพที่มีอยู่ใน หน่วยสื่อสารในการตรวจจับ วัตถุเคลื่อนที่ระยะทางไกล (Remote) สำหรับโครงข่ายใช้สายและ โครงข่ายไร้สาย (Wireless Network) ของสื่อประสม (Multimedia)

ขั้นตอนวิธีที่ใช้ประมวลผลในคอมพิวเตอร์สำหรับการตรวจจับวัตถุแบบเวลาจริงใช้ เทคโนโลยีการตรวจจับขอบ (Edge Detection) เช่น SA-ED เนื่องจากมีความสามารถในการใช้เวลา ในการประมวลผลต่ำ ประกอบกับการใช้จุดศูนย์กลางภาพวัตถุที่ ผ่านกระบวนการ ED พร้อมกับการกำหนดความไวของการตรวจจับ ด้วยระยะทางยูคลิดีเนียน (Euclidian Distance) (หอมานน, 2551)

2.2 (SSD) Object Detection: Single Shot MultiBox Detector for real-time processing

การตรวจจับวัตถุ SSD: Single Shot MultiBox Detector สำหรับการประมวลผลแบบเรียลไทม์



SSD Object detection

SSD ถูกออกแบบมาสำหรับการตรวจจับวัตถุในแบบ real-time ,Faster R-CNN ใช้ region proposal network เพื่อสร้าง boundary boxesและใช้เพื่อจำแนกวัตถุ ซึ่งเป็นจุดเริ่มต้นความแม่นยำของกระบวนการทั้งหมด ซึ่งทำงานที่ 7 เฟรมต่อวินาที ซึ่งต่ำกว่าความต้องการของการประมวลผลแบบเรียลไทม์ SSD ทำให้กระบวนการให้เร็วขึ้น โดยไม่จำเป็นต้อง region proposal network ในการแก้ไขความแม่นยำที่ลดลง

SSD ใช้การปรับปรุงบางอย่าง รวมถึงคุณสมบัติหลายขนาดและกล่องเริ่มต้น การปรับปรุงเหล่านี้อนุญาตให้ SSD จับคู่กับ Faster R-CNN ที่รวดเร็วขึ้นโดยใช้ภาพความละเอียดที่ต่ำกว่า ซึ่งจะเพิ่มความเร็วให้สูงขึ้น จากการเปรียบเทียบดังต่อไปนี้ทำให้ได้ในการประมวลผลแบบเรียลไทม์เร็วขึ้นและยังทำให้ความแม่นยำ Faster R-CNN (ความแม่นยำวัดจากค่า mean average precision mAP: การทำนายที่แม่นยำ)

System	VOC2007 test mAP	FPS (Titan X)	Number of Boxes	Input resolution
Faster R-CNN (VGG16)	73.2	7	~6000	~1000 x 600
YOLO (customized)	63.4	45	98	448 x 448
SSD300* (VGG16)	77.2	46	8732	300 x 300
SSD512* (VGG16)	79.8	19	24564	512 x 512

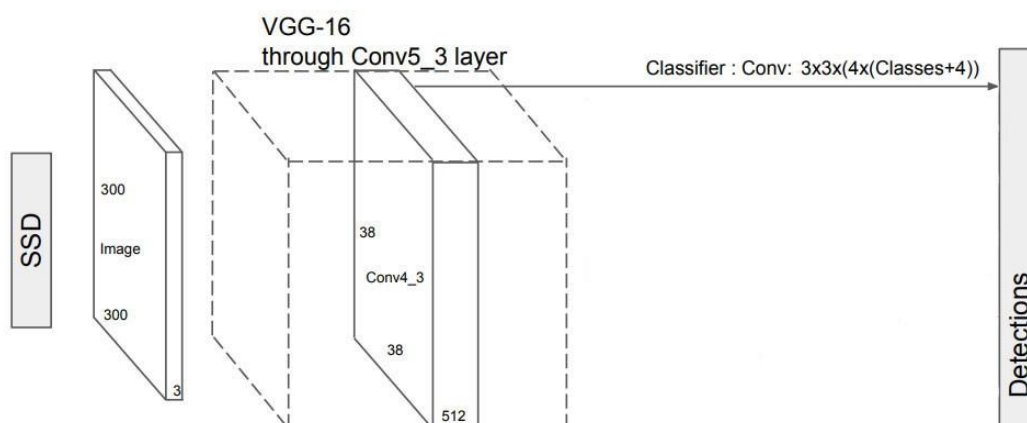
Performance comparison among object detection networks

การเปรียบเทียบประสิทธิภาพระหว่างเครือข่ายการตรวจจับวัตถุ

SSD

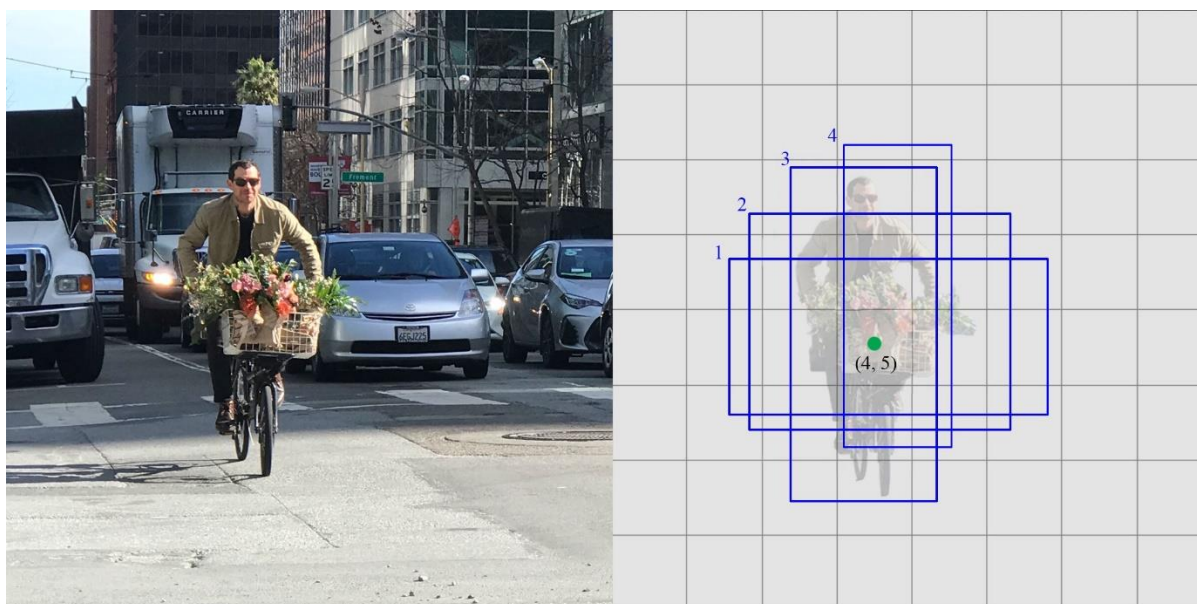
การตรวจจับวัตถุ SSD ประกอบด้วย 2 ส่วนคือ

1. แยก feature maps
2. ใช้ตัวกรอง Convolution เพื่อตรวจจับวัตถุ



Modified from SSD: Single Shot MultiBox Detector.

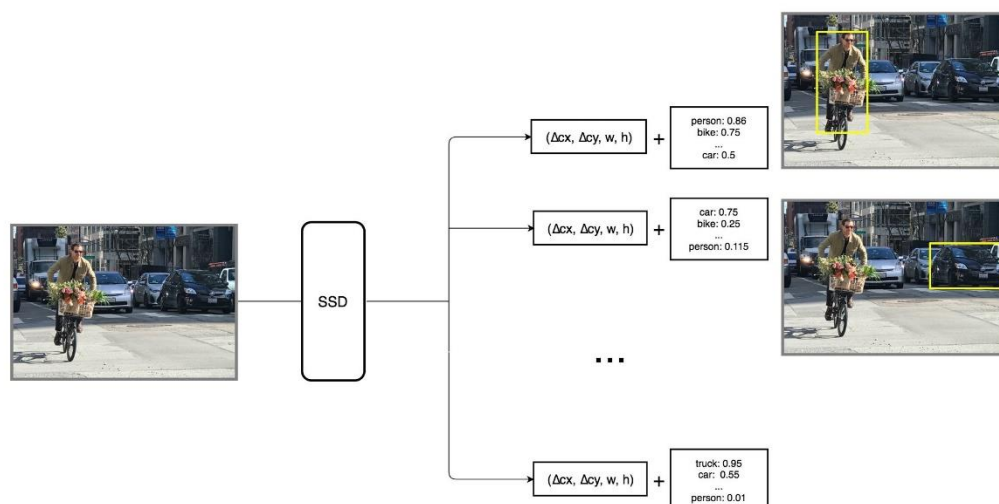
SSD ใช้ VGG16 เพื่อแยก feature maps จากนั้นจะตรวจจับวัตถุโดยใช้ Conv4_3 layer สำหรับภาพประกอบเรรวาด Conv4_3 ให้เป็น 8×8 เซลล์พื้นที่ (ควรเป็น 38×38) สำหรับแต่ละ cell (ตำแหน่ง) จะทำการคาดการณ์วัตถุ 4 อย่าง



ซ้าย: ภาพต้นฉบับ ขวา: 4 การคาดการณ์ในแต่ละเซลล์

Left: the original image. Right: 4 predictions at each cell.

การทำนายแต่ละครั้งประกอบด้วย boundary box และ 21 scores สำหรับแต่ละ class (หนึ่ง extra class สำหรับไม่มีวัตถุ) และเลือกคะแนนสูงสุดเป็นชั้นเรียนสำหรับวัตถุที่ถูกผูกไว้ Conv4_3 ทั้งหมดของ $38 \times 38 \times 4$: การคาดการณ์สี่ครั้งต่อ cell โดยไม่คำนึงถึงความลึกของ feature maps อย่างที่คาดไว้ การคาดการณ์หลายครั้งเมื่อไม่มีวัตถุ SSD จะส่งคลาส “0” เพื่อระบุว่าไม่มีวัตถุใด ๆ



การทำนายแต่ละครั้งจะมี boundary box และ 21 scores สำหรับ 21 classes (หนึ่งคลาสสำหรับไม่มีวัตถุ)

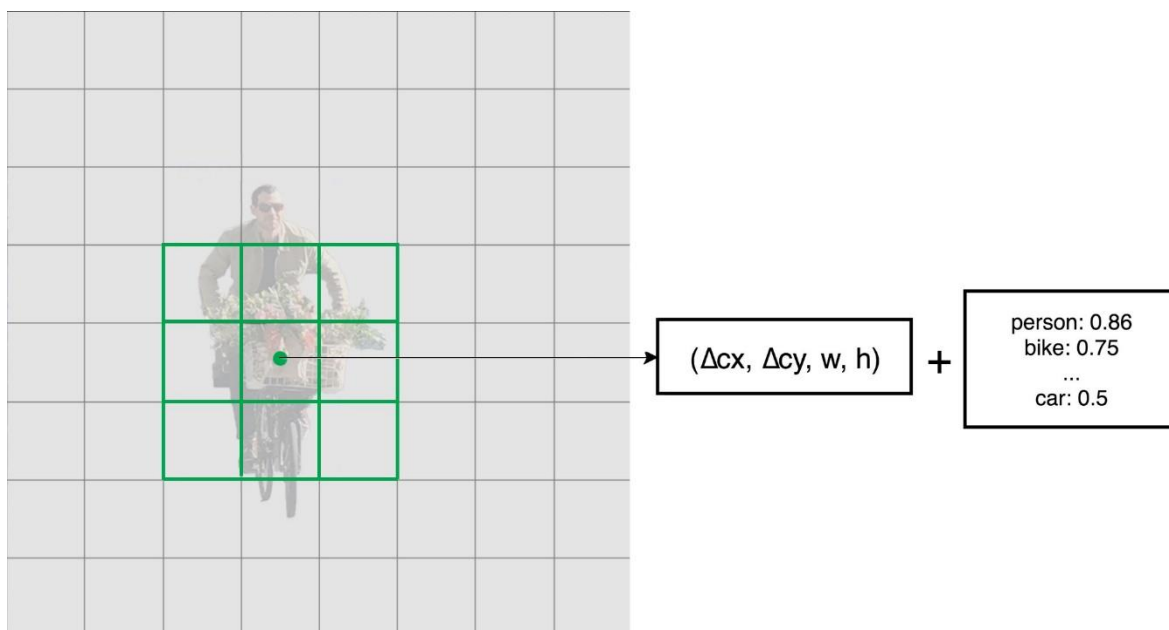
Each prediction includes a boundary box and 21 scores for 21 classes (one class for no object).

การคาดคะเนหลายครั้งที่มี boundary box และ confidence scores เรียกว่า multibox

ตัวทำนาย Convolutional สำหรับการตรวจจับวัตถุ

Convolutional predictors for object detection

SSD ไม่ใช่ region proposal network ที่ได้รับ เป็นวิธีที่ง่ายมาก สามารถคำนวณทั้งสถานที่และ class scores การใช้ small convolution filters หลังจากแยก feature maps แล้ว SSD จะใช้ตัวกรอง Convolution 3×3 สำหรับแต่ละ cell เพื่อทำการคาดการณ์ (filters เหล่านี้คำนวณผลลัพธ์เช่นเดียวกับตัวกรอง CNN ทั่วไป) filters แต่ละ outputs 25 channels: 21 scores สำหรับแต่ละ class พร้อม boundary box



ใช้ตัวกรองการสนทนา 3x3 เพื่อคาดการณ์ตำแหน่งและชั้นเรียน

Apply a 3x3 convolution filter to make a prediction for the location and the class.

ตัวอย่างเช่นใน Conv4_3 ใช้ตัวกรอง 3×3 สีตัวเพื่อ filters to map 512 input channels เข้ากับ output channels 25 ช่อง

$$(38 \times 38 \times 512) \xrightarrow{(4 \times 3 \times 3 \times 512 \times (21+4))} (38 \times 38 \times 4 \times (21 + 4))$$

แผนที่คุณสมบัติหลายขนาดสำหรับการตรวจจับ

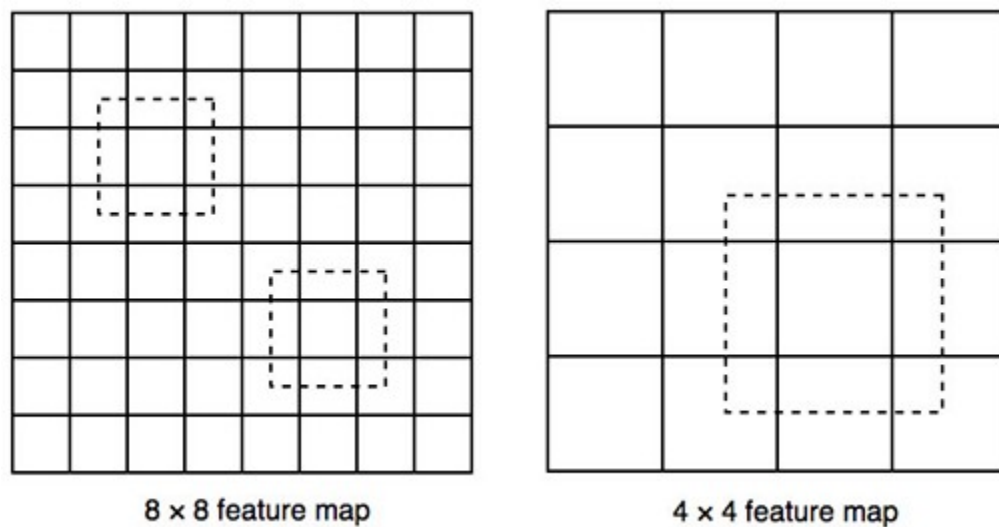
Multi-scale feature maps for detection



Scale matter มุมมองการเปลี่ยนแปลงขนาดของวัตถุ

Scale matter. Perspective changes the scale of the objects.

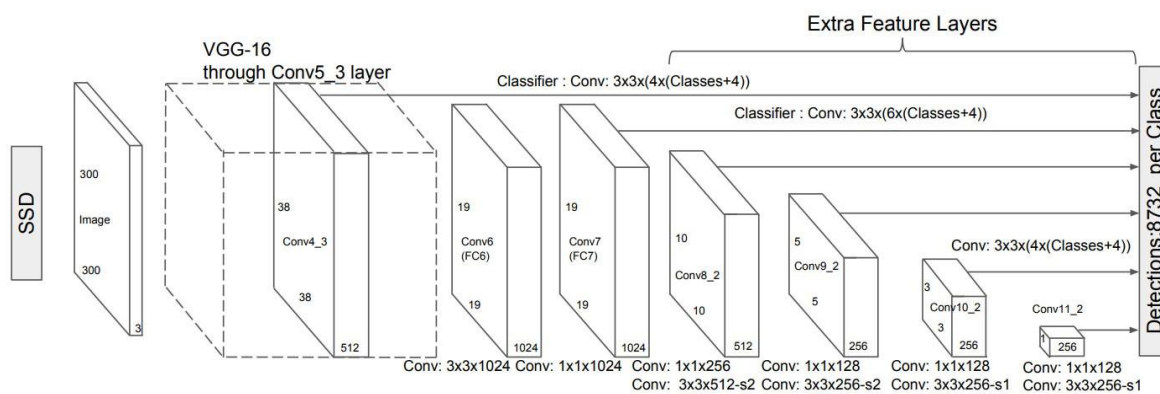
อย่างแรกจะอธิบายวิธีที่ SSD ตรวจจับวัตถุจากเลเยอร์เดียว จริงแล้วใช้หลายเลเยอร์ (multi-scale feature maps) เพื่อตรวจจับวัตถุ detect objects อย่างอิสระ เมื่อ CNN ลดมิติเชิงพื้นที่ลงเรื่อย ๆ ความละเอียดของ feature maps ก็ลดลงเช่นกัน SSD ใช้เลเยอร์ความละเอียดต่ำกว่าเพื่อตรวจจับวัตถุขนาดใหญ่ ตัวอย่างเช่น feature maps 4×4 ใช้สำหรับวัตถุขนาดใหญ่



แมปฟีเจอร์ความละเอียดที่ต่ำกว่า (ขวา) ตรวจจับวัตถุขนาดใหญ่

Lower resolution feature maps (right) detects larger scale objects.

SSD เพิ่มเลเยอร์ Convolution เพิ่มเติมอีก 6 เลเยอร์ 6 หลังจาก VGG16 เพิ่มสำหรับการตรวจจับวัตถุ ในสามเลเยอร์เหล่านั้นเราทำการคาดการณ์ 6 ครั้งแทน 4 โดยรวม SSD ทำการคาดการณ์ 8732 การใช้ 6 เลเยอร์



ที่มา: SSD: Single Shot MultiBox Detector

Multi-scale feature maps ช่วยปรับปรุง accuracy อย่างมาก นี่คือ accuracy ที่มีเลเยอร์ feature map ต่าง ๆ ที่ใช้สำหรับการตรวจจับวัตถุ

Prediction source layers from:						mAP		# Boxes
38 × 38	19 × 19	10 × 10	5 × 5	3 × 3	1 × 1	use boundary boxes?		
✓	✓	✓	✓	✓	✓	Yes	No	8732
✓	✓	✓				74.3	63.4	9864
	✓					70.7	69.2	8664
						62.4	64.0	

กล่องขอบเขตเริ่มต้น Default boundary box

The default boundary boxes เทียบเท่ากับที่ติดตั้งใน Faster R-CNN

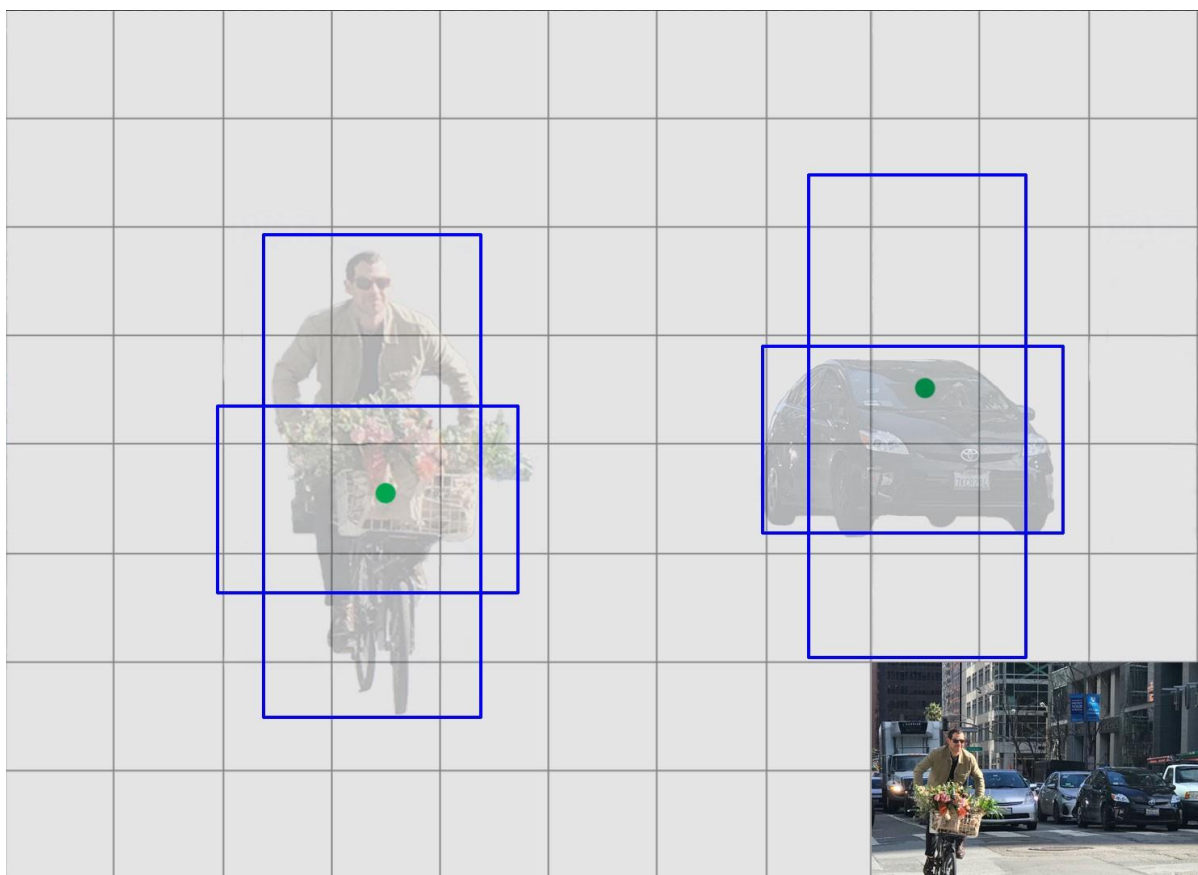
เราจะทำนาย boundary boxes ได้อย่างไร เช่นเดียวกับ Deep Learning สามารถเริ่มต้นด้วยการทำนายแบบสุ่มและใช้การไล่ระดับสี แบบไล่ระดับเพื่อเพิ่มประสิทธิภาพโมเดล อย่างไรก็ตามในระหว่างการ training เบื้องต้น แบบจำลองจะกำหนดว่ารูปร่างใด (คนเดินเท้าหรือรถยนต์) ที่จะปรับให้เหมาะสำหรับการพยากรณ์ ผลลัพธ์คือ training ในช่วงแรกอาจไม่เสถียร การคาดคะเน boundary box ด้านล่างทำงานได้ดีสำหรับ 1 หมวดหมู่ แต่ไม่เหมาะสำหรับอื่น ๆ เราต้องการให้การคาดการณ์เริ่มต้นของเรามีความหลากหลายและไม่เหมือนกัน



หากการคาดการณ์หลากหลายโมเดลจะไม่ทำงาน

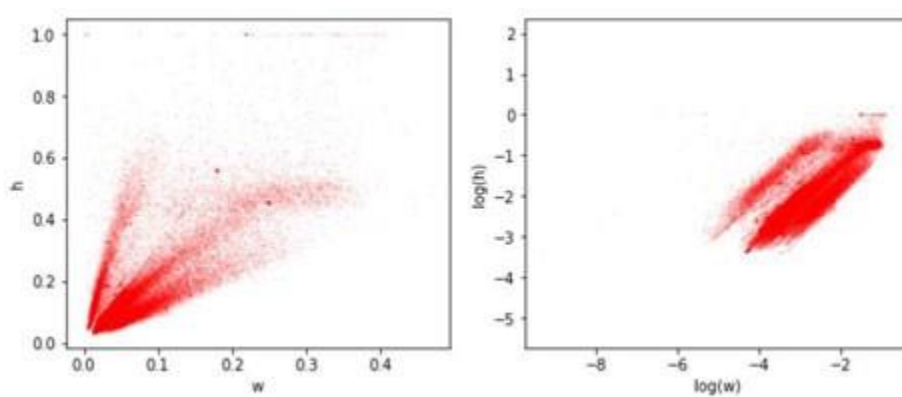
If predictions are not diverse, the model will not perform.

หากการคาดการณ์ของครอบคลุมรูปร่างมากกว่า เช่นด้านล่างโมเดล ของเราสามารถตรวจจับประเภทวัตถุได้มากขึ้น การเริ่มต้นประเภทนี้ทำให้ training ง่ายขึ้นและมีเสถียรภาพมากขึ้น

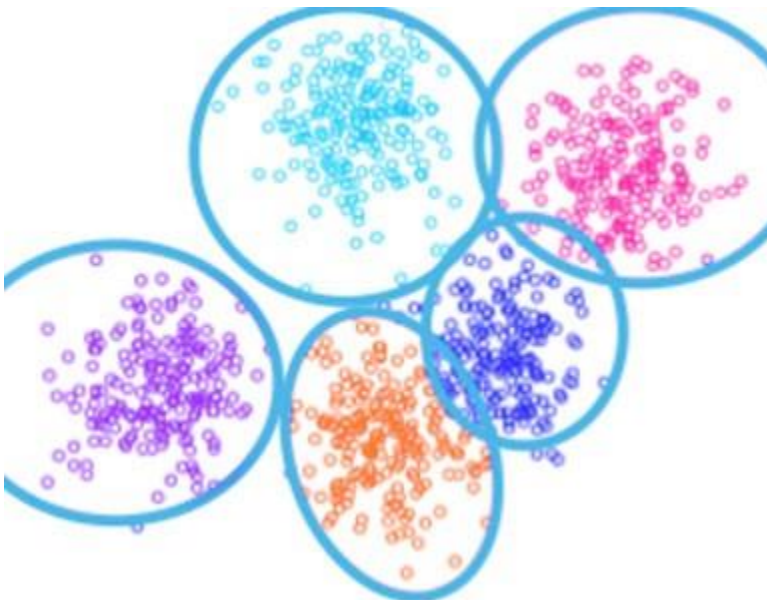


Diverse predictions cover more object types.

ในชีวิตจริง boundary boxes ไม่มีรูปร่างและ arbitrary shape รถยนต์มีรูปร่างและคนเดินถนนที่คล้ายกันมีอัตราส่วนโดยประมาณเท่ากับ 0.41 ในชุดข้อมูล KITTI ที่ใช้ในการขับเคลื่อนแบบอิสระการกระจายความกว้างและความสูงสำหรับ arbitrary shape มีการรวมกันเป็นกลุ่ม



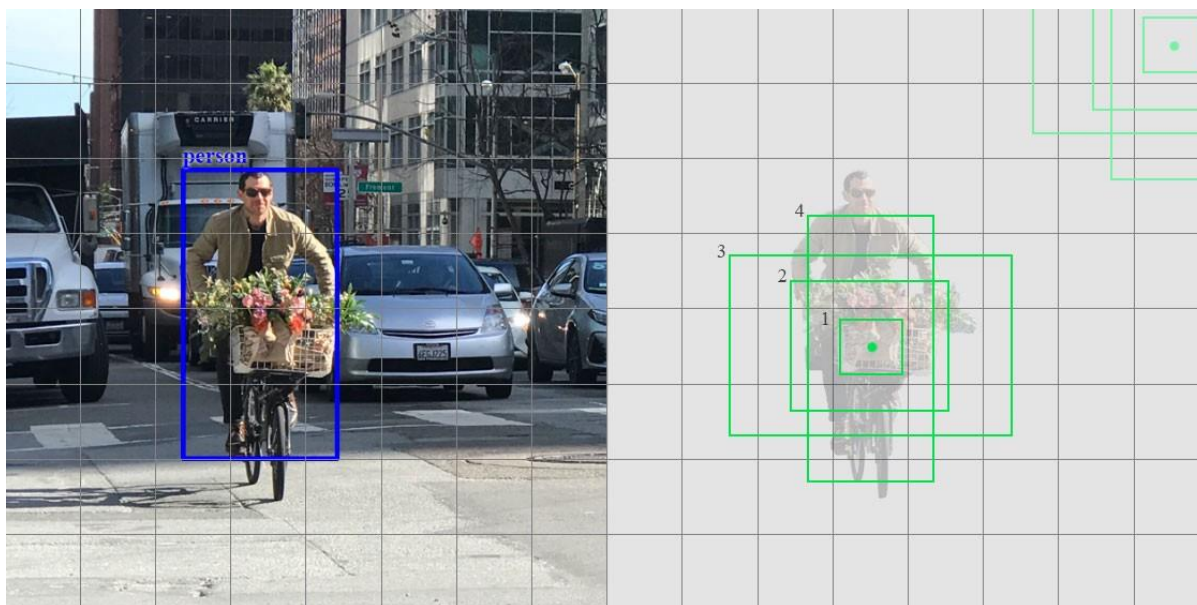
ตามแนวคิดแล้ว boundary boxes ความจริงของภาคพื้นดินสามารถแบ่งพาร์ติชันเป็นกลุ่มโดยแต่ละ cluster แสดงโดย boundary boxes เริ่มต้น (centroid ของ cluster) ดังนั้นแทนที่จะทำการสุ่มเลือกสามารถเริ่มการเดาโดยใช้ default boxes เหล่านั้น



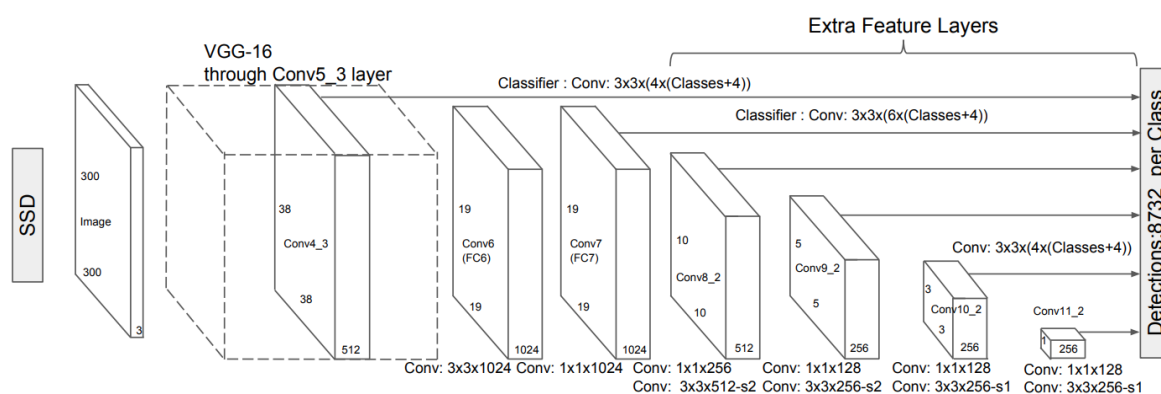
(รูปภาพที่แก้ไขจาก k-means cluster)

เพื่อให้ความซับซ้อนที่ต่ำลง default boxes จะถูกเลือกล่วงหน้าด้วยตนเองและรอบคอบ เพื่อครอบคลุม spectrum ของ real-life objects SSD ยังช่วยให้ default boxes อยู่ในระดับต่ำสุด (4 หรือ 6) โดยมีการทำนาย 1 รายการต่อ default box ตอนนี้แทนที่จะใช้การประสานงานร่วมกันทั่วโลกสำหรับ box location การคาดคะเน boundary box จะสัมพันธ์กับ default boundary box ที่แต่ละเซลล์ (Δcx , Δcy , w , h) นั่นคือค่า offsets ไปยัง default box ที่แต่ละเซลล์

สำหรับแต่ละ feature map layers จะแชร์ชุด default boxes เดียวกัน ที่กึ่งกลางที่เซลล์ที่เกี่ยวข้องแต่ละเยอร์ต่าง ๆ ใช้ชุด default boxes ที่แตกต่างกันในการปรับแต่งการตรวจจับวัตถุ ด้วยความละเอียดที่แตกต่างกัน กล้องสีเขียว 4 กล้องด้านล่างแสดง default boundary boxes 4 กล้อง



การเลือก default boundary boxes



เลือก Default boundary boxes เอง SSD จะกำหนดค่ามาตราส่วนสำหรับแต่ละชั้นของ feature map layer เริ่มจากด้านซ้าย Conv4_3 ตรวจจับวัตถุในระดับที่เล็กที่สุด 0.2 (หรือ 0.1 บางครั้ง) จากนั้นเพิ่มเลเยอร์เชิงเส้นเป็นชั้นขวาสุดที่ระดับ 0.9 การรวมค่าสัดส่วนกับอัตราส่วนภาพเป้าหมาย และคำนวณความกว้างและความสูงของกล่อง-เริ่มต้น สำหรับเลเยอร์ที่ทำการคาดการณ์ 6 ครั้ง SSD เริ่มต้นด้วยอัตราส่วนภาพ 5 เป้าหมาย: 1, 2, 3, 1/2 และ 1/3 จากนั้นคำนวณความกว้างและความสูงของ default boxes ดังนี้

$$w = scale \cdot \sqrt{\text{aspect ratio}}$$

$$h = \frac{scale}{\sqrt{\text{aspect ratio}}}$$

Then SSD adds an extra default box with scale:

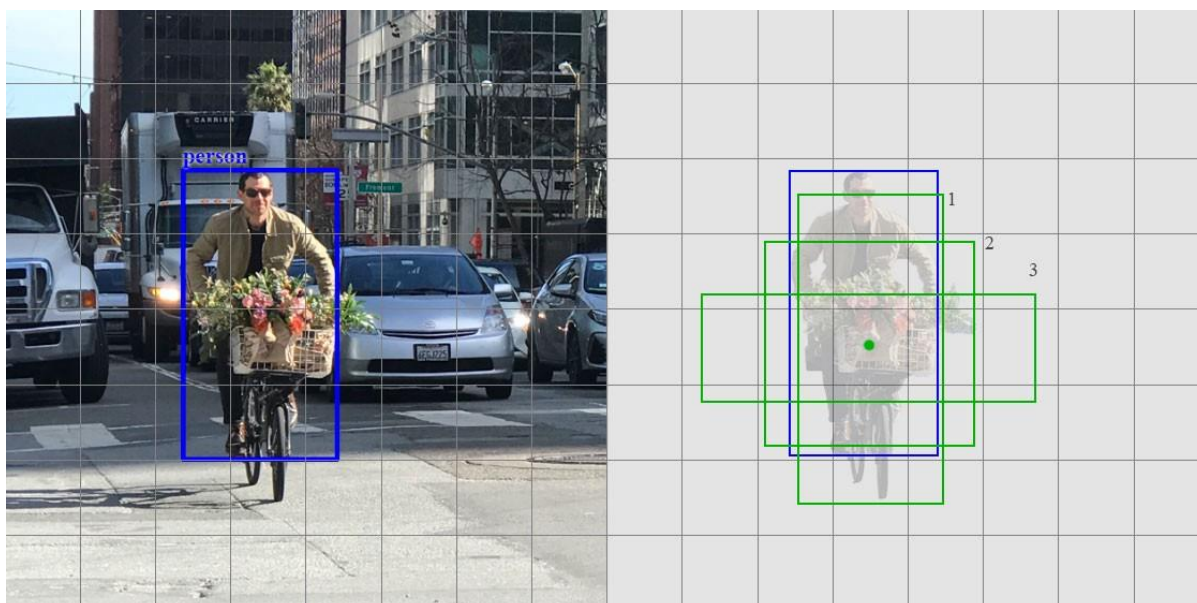
$$scale = \sqrt{\text{scale} \cdot \text{scale at next level}}$$

และอัตราส่วน = 1

YOLO ใช้การจัดกลุ่ม k-mean บนชุดข้อมูล training เพื่อกำหนด default boundary เหล่านั้น

กลยุทธ์การจับคู่

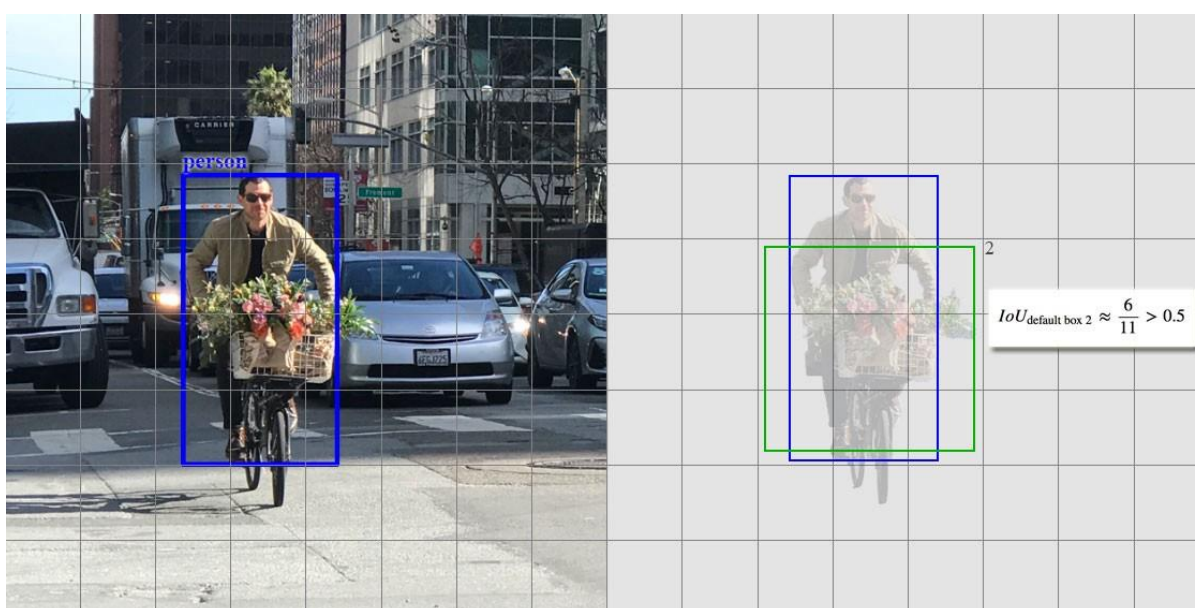
เป็นการจับคู่เชิงบวกหรือเชิงลบ SSD ใช้การจับคู่ที่เป็นบวกเท่านั้นในการคำนวณ ต้นทุนการแปล(กล่องไม่ตรงกัน) หากกล่องขอบเขตเริ่มต้นที่สอดคล้องกัน(ไม่ใช่กล่องขอบเขตที่คาดการณ์ไว้) มี IoU มากกว่า 0.5 ด้วยความจริงพื้นฐานการแข่งขันจะเป็นบวก มิฉะนั้นจะเป็นลบ (IoU, จุดตัดกันของสหภาพคืออัตราส่วนระหว่างพื้นที่ที่ถูกตัดผ่านพื้นที่เชื่อมต่อสำหรับสองภูมิภาค)



วัตถุความจริงภาคพื้นดิน (สีฟ้า) และ 3 กล่องขอบเขตเริ่มต้น (สีเขียว)

มาลดความซับซ้อนของการสนทนาของเราเป็น 3 กล้องเริ่มต้น เฉพาะกล้องเริ่มต้น 1 และ 2 (แต่ไม่ใช่ 3) มี IoU มากกว่า 0.5 โดยมีกล้องความจริงภาคพื้นดินด้านบน (กล้องสีน้ำเงิน) ดังนั้นเฉพาะกล้องที่ 1 และ 2 เท่านั้นที่เป็นค่าบวก เมื่อเราระบุการแข่งขันที่เป็นบวกแล้วเราจะใช้กล่องขอบเขตการทำนายที่สอดคล้องกันเพื่อคำนวณต้นทุน กลยุทธ์การจับคู่นี้แบ่งส่วนใดของความจริงพื้นฐานที่การทำนายรับผิดชอบ

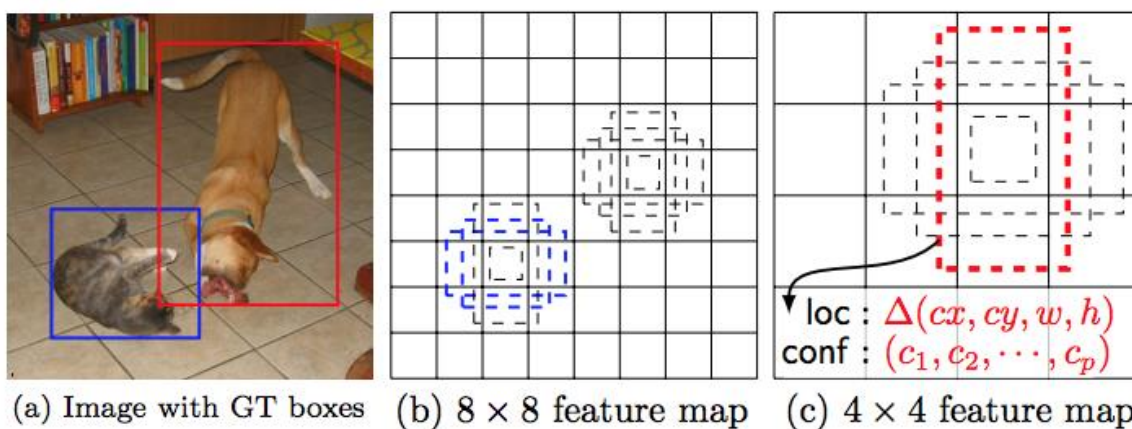
กลยุทธ์การจับคู่นี้ส่งเสริมให้การทำนายแต่ละครั้งเพื่อทำนายรูปร่างที่ใกล้เคียงกับกล้องเริ่มต้นที่สอดคล้องกัน ดังนั้นการทำนายของเราจึงมีความหลากหลายและมีเสถียรภาพมากขึ้นในการฝึกอบรม



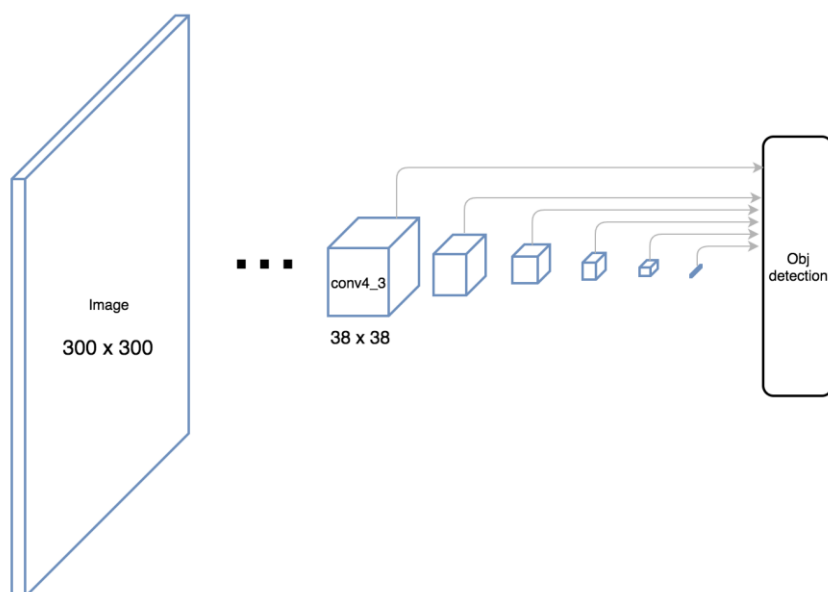
กล้องเริ่มต้น 2 มี IoU มากกว่า 0.5 ด้วยความจริงพื้นฐาน

แผนที่คุณลักษณะหลายขนาดและกล่องขอบเขตเริ่มต้น

นี่คือตัวอย่างของวิธีที่ SSD ผสานการแมปฟีเจอร์หลายสเกลและกล่องขอบเขตเริ่มต้นเพื่อตรวจจับวัตถุในระดับที่ต่างกัน สุนัขด้านล่างตรงกับกล่องเริ่มต้นหนึ่งกล่อง (สีแดง) ในเลเยอร์แผนที่คุณลักษณะ 4×4 แต่ไม่มีกล่องเริ่มต้นใด ๆ ในแผนที่คุณลักษณะความละเอียดสูงกว่า 8×8 แมวที่มีขนาดเล็กกว่านั้นจะถูกตรวจพบโดยชั้นแผนที่คุณลักษณะ 8×8 ในกล่องเริ่มต้น 2 กล่อง (เป็นสีน้ำเงิน)



แผนที่คุณลักษณะความละเอียดสูงมีหน้าที่ตรวจจับวัตถุขนาดเล็กเลเยอร์แรกสำหรับการตรวจจับวัตถุ conv4_3 มีมิติเชิงพื้นที่ที่ 38×38 ซึ่งลดลงมาจากภาพอินพุต ดังนั้น SSD มักทำงานได้ไม่ดีสำหรับวัตถุขนาดเล็กเมื่อเปรียบเทียบกับวิธีตรวจจับอื่น ๆ หากเป็นปัญหาเราสามารถบรรเทาได้โดยใช้ภาพที่มีความละเอียดสูงกว่า



ฟังก์ชันการสูญเสีย

การสูญเสียการแปลไม่ตรงกันระหว่างพื้นกล่องจริงและที่คาดการณ์กล่องเขตแดน SSD ลงโทษการคาดการณ์จากการแข่งขันที่เป็นบวกเท่านั้น เราต้องการคำนวณจากการจับคู่ที่เป็นบวกเพื่อเข้าใจความจริงพื้นฐาน การจับคู่เชิงลบสามารถถูกละเว้น

The localization loss between the predicted box l and the ground truth box g is defined as the smooth L1 loss with cx, cy as the offset to the default bounding box d of width w and height h .

$$L_{loc}(x, l, g) = \sum_{i \in Pos} \sum_{m \in \{cx, cy, w, h\}} x_{ij}^k \text{smooth}_{L1}(l_i^m - \hat{g}_j^m)$$

$$\hat{g}_j^{cx} = (g_j^{cx} - d_i^{cx})/d_i^w \quad \hat{g}_j^{cy} = (g_j^{cy} - d_i^{cy})/d_i^h$$

$$\hat{g}_j^w = \log\left(\frac{g_j^w}{d_i^w}\right) \quad \hat{g}_j^h = \log\left(\frac{g_j^h}{d_i^h}\right)$$

$$x_{ij}^p = \begin{cases} 1 & \text{if IoU} > 0.5 \text{ between default box } i \text{ and ground true box } j \text{ on class } p \\ 0 & \text{otherwise} \end{cases}$$

การสูญเสียความมั่นใจคือการสูญเสียในการทำนายผลการเรียน สำหรับการคาดคะเนการแข่งขันที่เป็นบวกทุกครั้งเราจะลงโทษการสูญเสียตามคะแนนความเชื่อมั่นของคลาสที่เกี่ยวข้อง สำหรับการทำนายการจับคู่เชิงลบเราลงโทษการสูญเสียตามคะแนนความเชื่อมั่นของคลาส“ 0”: คลาส“ 0” จัดประเภทไม่พบวัตถุใด ๆ

It is calculated as the softmax loss over multiple classes confidences c (class score).

$$L_{conf}(x, c) = - \sum_{i \in Pos} x_{ij}^p \log(\hat{c}_i^p) - \sum_{i \in Neg} \log(\hat{c}_i^0) \quad \text{where} \quad \hat{c}_i^p = \frac{\exp(c_i^p)}{\sum_p \exp(c_i^p)}$$

where N is the number of matched default boxes.

ฟังก์ชันการสูญเสียครั้งสุดท้ายคำนวณเป็น :

$$L(x, c, l, g) = \frac{1}{N} (L_{conf}(x, c) + \alpha L_{loc}(x, l, g))$$

โดยที่ N คือจำนวนของการจับคู่ที่เป็นบวกและ α คือน้ำหนักสำหรับการสูญเสียการแปล

การชุดเชิงลบอย่างหนัก

อย่างไรก็ตามเราคาดการณ์ได้มากกว่าจำนวนวัตถุที่มีอยู่ ดังนั้นจึงมีการจับคู่เชิงลบมากกว่าการแข่งขันที่เป็นบวก สิ่งนี้สร้างความไม่สมดุลในชั้นเรียนซึ่งทำร้ายการฝึกอบรม เรากำลังฝึกจำลองเพื่อเรียนรู้พื้นที่พื้นหลังมากกว่าการตรวจจับวัตถุ อย่างไรก็ตาม SSD ยังต้องการการสุ่มตัวอย่างเชิงลบเพื่อให้สามารถเรียนรู้สิ่งที่ถือเป็นการคาดการณ์ที่ไม่ดี ดังนั้นแทนที่จะใช้เชิงลบทั้งหมดเราจัดเรียงเชิงลบเหล่านั้นด้วยการสูญเสียความมั่นใจที่คำนวณได้ SSD เลือกเนกาทีฟที่มีการสูญเสียสูงสุดและตรวจสอบให้แน่ใจว่าอัตราส่วนระหว่างเนกาทีฟที่เลือกและบวกมีค่าสูงสุด 3: 1 สิ่งนี้นำไปสู่การฝึกอบรมที่รวดเร็วและมีเสถียรภาพมากขึ้น

การเพิ่มข้อมูล

การเพิ่มข้อมูลมีความสำคัญในการปรับปรุงความแม่นยำ ข้อมูลเพิ่มเติมพร้อมการพลิกการครอบตัดและการบิดเบือนสี เมื่อต้องการจัดการชุดรูปแบบในขนาดและรูปร่างของวัตถุที่แตกต่างกันแต่ละรูปแบบการฝึกอบรมจะถูกสุ่มตัวอย่างโดยหนึ่งในตัวเลือกต่อไปนี้ :

- ใช้ต้นฉบับ
- ตัวอย่างแพทช์ที่มี IoU 0.1, 0.3, 0.5, 0.7 หรือ 0.9
- สุ่มแพทช์ตัวอย่าง

แพทช์ตัวอย่างจะมีอัตราส่วนระหว่าง 1/2 และ 2 จากนั้นจะถูกปรับขนาดเป็นขนาดคงที่และเราจะพลิกครึ่งหนึ่งของข้อมูลการฝึกอบรม นอกจากนี้เราสามารถใช้การบิดเบือนภาพถ่าย



นี่คือการปรับปรุงประสิทธิภาพหลังจากการเพิ่มข้อมูล:

data augmentation	SSD300		
horizontal flip	✓	✓	✓
random crop & color distortion		✓	✓
random expansion			✓
VOC2007 test mAP	65.5	74.3	77.2

เวลาอนุมาน

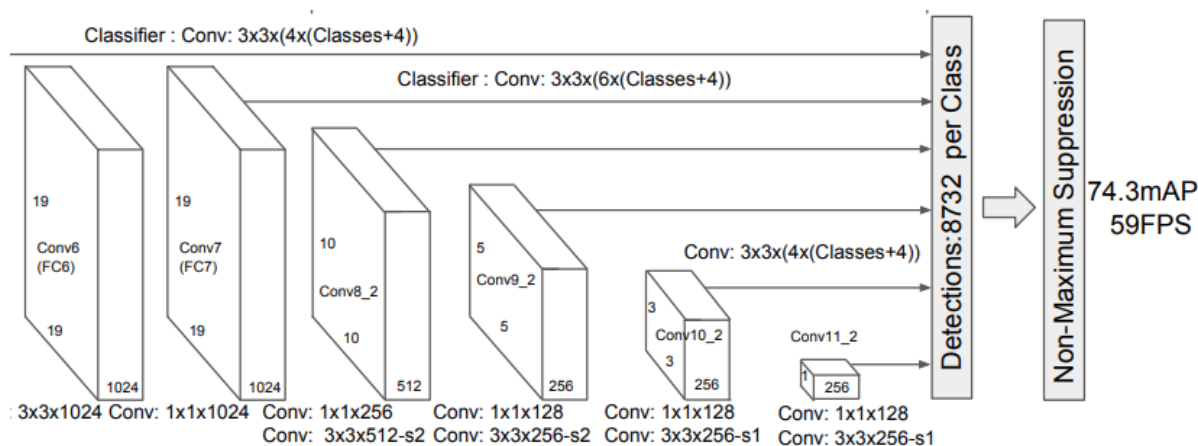
SSD ทำการคาดการณ์ได้หลายอย่าง (8732) เพื่อความครอบคลุมที่ดีขึ้นของตำแหน่งขนาดและอัตราส่วน ภาพมากกว่าวิธีตรวจจับอื่น ๆ อย่างไรก็ตามการคาดการณ์จำนวนมากไม่มีวัตถุ ดังนั้นการคาดการณ์ใด ๆ ที่มีคะแนนความมั่นใจในระดับต่ำกว่า 0.01 จะถูกกำจัด

Method	mAP	FPS	batch size	# Boxes	Input resolution
Faster R-CNN (VGG16)	73.2	7	1	~ 6000	~ 1000 × 600
Fast YOLO	52.7	155	1	98	448 × 448
YOLO (VGG16)	66.4	21	1	98	448 × 448
SSD300	74.3	46	1	8732	300 × 300
SSD512	76.8	19	1	24564	512 × 512
SSD300	74.3	59	8	8732	300 × 300
SSD512	76.8	22	8	24564	512 × 512

SSD ทำให้การคาดการณ์มากขึ้น การปรับปรุงอนุญาตให้ SSD ใช้ภาพความละเอียดต่ำลงเพื่อความแม่นยำที่ใกล้เคียงกัน

การปราบปรามที่ไม่สูงสุด (nms)

SSD ใช้การปราบปรามที่ไม่มากที่สุดเพื่อลดการทำนายที่ซ้ำกันซึ่งชี้ไปยังวัตถุเดียวกัน SSD จัดเรียงการคาดคะเนด้วยความมั่นใจ เริ่มจากการคาดการณ์ความเชื่อมั่นสูงสุด SSD ประเมินว่ากล่องขอบเขตที่คาดการณ์ไว้ก่อนหน้านี้มี IoU สูงกว่า 0.45 ด้วยการทำนายปัจจุบันสำหรับคลาสเดียวกันหรือไม่ หากพบการคาดการณ์ปัจจุบันจะถูกละเว้น อย่างน้อยที่สุดเราเก็บการคาดคะเนสูงสุด 200 อันดับต่อภาพ



ผล

แบบจำลองนี้ได้รับการฝึกฝนโดยใช้ SGD โดยมีอัตราการเรียนรู้เริ่มต้น 0.001, 0.9 โมเมนตัม, 0.0005 น้ำหนักลดลงและขนาดชุดที่ 32 การใช้ Nvidia Titan X กับการทดสอบ VOC2007 SSD บรรลุ 59 FPS กับ mAP 74.3% ในการทดสอบ VOC2007 เทียบกับ Faster R- FN ของ CNN 7 ที่มี mAP 73.2% หรือ YOLO 45 FPS ที่มี mAP 63.4%

นี่คือการเปรียบเทียบความแม่นยำสำหรับวิธีการที่ต่างกัน สำหรับ SSD ใช้ขนาดภาพ 300×300 หรือ 512×512

Method	data	mAP	aero	bike	bird	boat	bottle	bus	car	cat	chair	cow	table	dog	horse	mbike	person	plant	sheep	sofa	train	tv
Fast [6]	07	66.9	74.5	78.3	69.2	53.2	36.6	77.3	78.2	82.0	40.7	72.7	67.9	79.6	79.2	73.0	69.0	30.1	65.4	70.2	75.8	65.8
Fast [6]	07+12	70.0	77.0	78.1	69.3	59.4	38.3	81.6	78.6	86.7	42.8	78.8	68.9	84.7	82.0	76.6	69.9	31.8	70.1	74.8	80.4	70.4
Faster [2]	07	69.9	70.0	80.6	70.1	57.3	49.9	78.2	80.4	82.0	52.2	75.3	67.2	80.3	79.8	75.0	76.3	39.1	68.3	67.3	81.1	67.6
Faster [2]	07+12	73.2	76.5	79.0	70.9	65.5	52.1	83.1	84.7	86.4	52.0	81.9	65.7	84.8	84.6	77.5	76.7	38.8	73.6	73.9	83.0	72.6
Faster [2]	07+12+COCO	78.8	84.3	82.0	77.7	68.9	65.7	88.1	88.4	88.9	63.6	86.3	70.8	85.9	87.6	80.1	82.3	53.6	80.4	75.8	86.6	78.9
SSD300	07	68.0	73.4	77.5	64.1	59.0	38.9	75.2	80.8	78.5	46.0	67.8	69.2	76.6	82.1	77.0	72.5	41.2	64.2	69.1	78.0	68.5
SSD300	07+12	74.3	75.5	80.2	72.3	66.3	47.6	83.0	84.2	86.1	54.7	78.3	73.9	84.5	85.3	82.6	76.2	48.6	73.9	76.0	83.4	74.0
SSD300	07+12+COCO	79.6	80.9	86.3	79.0	76.2	57.6	87.3	88.2	88.6	60.5	85.4	76.7	87.5	89.2	84.5	81.4	55.0	81.9	81.5	85.9	78.9
SSD512	07	71.6	75.1	81.4	69.8	60.8	46.3	82.6	84.7	84.1	48.5	75.0	67.4	82.3	83.9	79.4	76.6	44.9	69.9	69.1	78.1	71.8
SSD512	07+12	76.8	82.4	84.7	78.4	73.8	53.2	86.2	87.5	86.0	57.8	83.1	70.2	84.9	85.2	83.9	79.7	50.3	77.9	73.9	82.5	75.3
SSD512	07+12+COCO	81.6	86.6	88.3	82.4	76.0	66.3	88.6	88.9	89.1	65.1	88.4	73.6	86.5	88.9	85.3	84.6	59.1	85.0	80.4	87.4	81.2

นี่คือบทสรุปของประสิทธิภาพความเร็วในเฟรมต่อวินาที

Method	mAP	FPS	batch size	# Boxes	Input resolution
Faster R-CNN (VGG16)	73.2	7	1	~ 6000	~ 1000 × 600
Fast YOLO	52.7	155	1	98	448 × 448
YOLO (VGG16)	66.4	21	1	98	448 × 448
SSD300	74.3	46	1	8732	300 × 300
SSD512	76.8	19	1	24564	512 × 512
SSD300	74.3	59	8	8732	300 × 300
SSD512	76.8	22	8	24564	512 × 512

ผลการวิจัย

นี่คือข้อสังเกตที่สำคัญบางประการ :

- SSD ทำงานได้เร็วกว่า Faster R-CNN ที่เร็วกว่าสำหรับวัตถุขนาดเล็ก ใน SSD วัตถุขนาดเล็กสามารถตรวจพบได้ในเลเยอร์ความละเอียดสูงกว่า (เลเยอร์ซ้ายสุด) แต่เลเยอร์เหล่านั้นมีคุณสมบัติระดับต่ำเช่นขอบหรือแพทช์สีซึ่งมีข้อมูลน้อยกว่าสำหรับการจำแนกประเภท
- ความแม่นยำจะเพิ่มขึ้นตามจำนวนกล่องขอบเขตเริ่มต้นที่ค่าความเร็ว
- แมปคุณสมบัติหลายสเกลปรับปรุงการตรวจจับวัตถุในระดับต่างๆ
- การออกแบบกล่องขอบเขตเริ่มต้นที่ดีขึ้นจะช่วยให้แม่นยำ
- ชุดข้อมูล COCO มีวัตถุขนาดเล็ก เพื่อปรับปรุงความแม่นยำให้ใช้กล่องเริ่มต้นที่เล็กลง (เริ่มต้นด้วยขนาดที่เล็กกว่าที่ 0.15)
- SSD มีข้อผิดพลาดในการแปลที่ต่ำกว่าเมื่อเปรียบเทียบกับ R-CNN แต่มีข้อผิดพลาดการจัดหมวดหมู่เพิ่มเติมที่เกี่ยวข้องกับหมวดหมู่ที่คล้ายกัน ข้อผิดพลาดในการจำแนกประเภทที่สูงกว่านั้นน่าจะเป็นเพราะเราใช้กล่องขอบเขตเดียวกันเพื่อทำการพยากรณ์หลายชั้น
- SSD512 มีความแม่นยำที่ดีกว่า (2.5%) กว่า SSD300 แต่ทำงานที่ 22 FPS แทน 59

ข้อสรุป

SSD เป็นเครื่องตรวจจับครั้งเดียว มันไม่มีเครือข่ายข้อเสนอที่ได้รับมอบหมายภูมิภาคและคาดการณ์กล่องขอบเขตและชั้นเรียนโดยตรงจากแผนที่คุณลักษณะในหนึ่งผ่าน

เพื่อปรับปรุงความแม่นยำ SSD แนะนำ :

- ตัวกรอง convolutional ขนาดเล็กเพื่อทำนายคลาสอ็อบเจกต์และออฟเซตไปยังกล่องขอบเขตดีฟอลต์
- ตัวกรองแยกต่างหากสำหรับกล่องเริ่มต้นเพื่อจัดการความแตกต่างในอัตราส่วนภาพ
- แมปพีเจอร์หลายระดับสำหรับการตรวจจับวัตถุ

SSD สามารถผ่านการฝึกอบรมตั้งแต่ต้นจนจบเพื่อความแม่นยำที่ดีขึ้น SSD ทำให้การคาดการณ์มากขึ้นและมีความครอบคลุมที่ดีกว่าเกี่ยวกับตำแหน่งขนาดและอัตราส่วนภาพ ด้วยการปรับปรุงข้างต้น SSD สามารถลดความละเอียดของภาพที่นำเข้าไปเหลือ 300×300 ด้วยประสิทธิภาพความแม่นยำในการเปรียบเทียบ ด้วยการลบ

ข้อเสนอภูมิภาคที่ได้รับมอบหมายและใช้ภาพความละเอียดต่ำกว่ารุ่นนั้นสามารถวิ่งด้วยความเร็วแบบเรียลไทม์ และยังคงความแม่นยำของ Faster R-CNN ที่ล้ำสมัย

2.3 OpenCV

Exploring OpenCV's Deep Learning Object Detection Library

Deep learning เกี่ยวกับการตรวจจับวัตถุบนภาพและวิดีโอทำให้ผู้ปฏิบัติงานและโปรแกรมเมอร์สามารถเข้าถึงได้ง่ายขึ้น เหตุผลหนึ่งที่ทำให้แนวโน้มนี้คือการแนะนำของ software libraries ใหม่ เช่น TensorFlow Object Detection API, OpenCV Deep Neural Network Module, and ImageAI. libraries เหล่านี้มีสิ่งหนึ่งที่เหมือนกัน: ทั้งหมดได้รวมโมเดลการตรวจจับวัตถุ deep-learning เข้าไปในระบบ เป็นผลให้ผู้ใช้งานไลบรารีเหล่านี้สามารถเข้าถึงโมเดลที่ผ่าน pre-trained มาแล้วจำนวนมากและตรวจสอบไลบรารีที่ดีที่สุด อย่างไรก็ตามการประเมินรูปแบบที่แตกต่างกัน (แม้ใน library เดียวกัน) อาจไม่ใช่เรื่องง่าย สิ่งนี้อธิบายถึงการศึกษาเบื้องต้นเกี่ยวกับรูปแบบการตรวจจับวัตถุที่เรียนรู้ลึกสองแบบภายใต้โมดูล Deep Neural Network ใน OpenCV 3.4.1 :

- SSD / MobileNet ดำเนินการโดย Tensorflow และ
- YOLOv2 ดำเนินการโดย Darknet

ซึ่งทั้งสอง models ก่อนผ่าน pre-trained โดยชุด COCO เรามุ่งเน้นที่นี่คือการแสดงเวิร์กโฟลว์ทั่วไปของการใช้แบบจำลองเหล่านี้ภายใต้ OpenCV และสำรวจส่วนที่ยุ่งยากบางอย่างในกระบวนการเช่นการดำเนินการที่แตกต่างกันสำหรับการประมวลผลภาพล่วงหน้าตัวแทนที่แตกต่างกันสำหรับการอธิบาย bounding boxes สูตรที่แตกต่างกัน / ระดับความน่าจะเป็น ฯลฯ

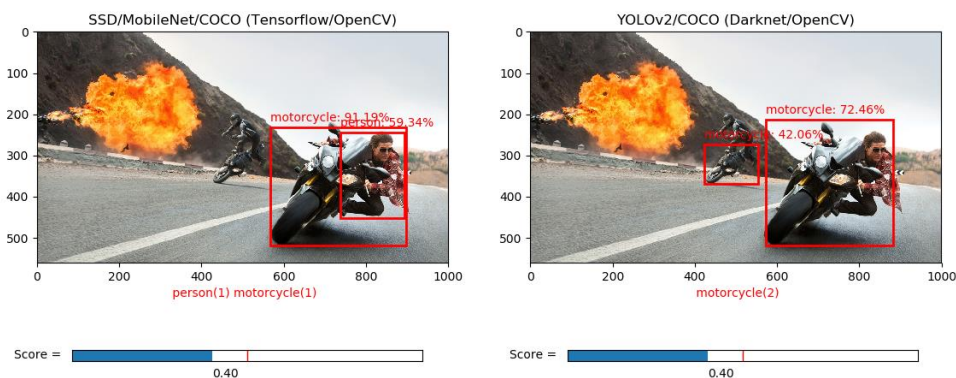
SSD (ย่อมาจาก Single Shot MultiBox Detector) และ YOLO (ย่อมาจาก You Only Look Once) เป็น models การตรวจจับวัตถุล้ำสมัยที่ล้ำสมัย อย่างไรก็ตามมีการปรับปรุงและการใช้งานมากมายสำหรับโมเดลเหล่านี้ ตัวอย่างเช่นนักวิจัยได้ใช้ deep neural networks ที่แตกต่างกัน (เช่น VGG, ResNet หรือ MobileNet) เป็นตัวแยกคุณลักษณะหรือตัวแยกประเภทวัตถุสำหรับ SSD ผู้คนยังใช้ SSD ภายใต้แพลตฟอร์มซอฟต์แวร์การเรียนรู้เชิงลึกที่แตกต่างกันเช่น Caffe, PyTorch หรือ Tensorflow ในทางกลับกัน YOLO ยังมีตัวแปรอีกมากมาย เช่น YOLOv2 และ YOLOv3 ตารางที่ 1 สรุปโมเดลภายใต้ OpenCV 3.4.1 สำหรับการศึกษา รายละเอียดเพิ่มเติม

Object Detection Model	SSD	YOLOv2
Object Classification Model	<u>MobileNet</u>	Darknet-19
Pre-trained Dataset	COCO	COCO
Deep Learning Software Platform	<u>Tensorflow</u>	Darknet
OpenCV <u>ReadNet</u> Function Call	<u>readNetFromTensorflow</u>	<u>readNetFromDarknet</u>
Config File	<u>frozen_inference_graph.pb</u>	yolov2.cfg
Weight File	<u>ssd_mobilenet_v1_coco_2017_11_17_pbt.txt</u>	yolov2.weights
OpenCV Forward Function Return	(1, 1, 100, 7)	(845, 85)
Representation of Bounding Box	(left, top, right, bottom)	(<u>x_center</u> , <u>y_center</u> , width, height)
Representation of Class Name	90	80

Table 1. SSD/MobileNet and YOLOv2 in OpenCV 3.4.1.

1. Tom Cruise ใน Mission Impossible 6

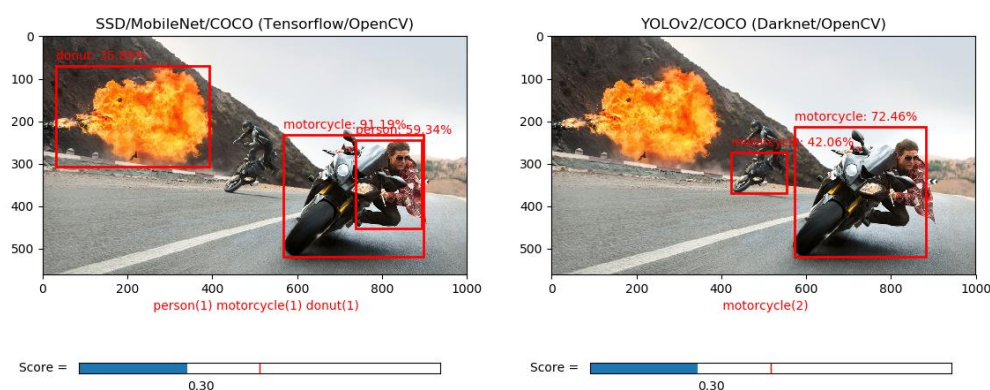
ตรวจสอบผลการตรวจจับจาก SSD / MobileNet และ YOLOv2 ในภาพของ Tom Cruise ใน Mission Impossible 6



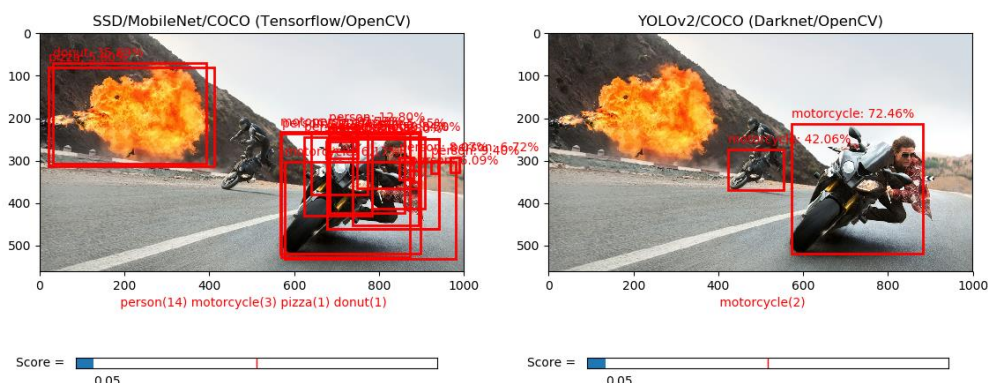
รูปที่ 1 ผลลัพธ์ของการตรวจจับวัตถุจาก SSD / MobileNet และ YOLOv2 (คะแนน = 0.40)

เป็นที่น่าสนใจที่จะทราบว่ารูปแบบที่ต่างกันให้ประโยชน์แก่วัตถุที่แตกต่างในกรณีนี้: SSD / MobileNet ตรวจพบคนคนหนึ่งและเป็นส่วนหนึ่งในรถจักรยานยนต์ในขณะที่ YOLOv2 ตรวจพบสองรถจักรยานยนต์

เนื่องจากทั้งสองโมเดลได้รับ trained โดยชุดข้อมูลเดียวกันเราอาจสันนิษฐานว่าความแตกต่างอาจเกิดจากค่าเกณฑ์ของคะแนน การพูดโดยประมาณค่าคะแนนบ่งบอกถึงความมั่นใจ / ความน่าจะเป็นของวัตถุที่ตรวจพบ ค่าเกณฑ์คะแนนของคะแนนในรูปที่ 1 คือ 0.40 โดยที่คะแนนที่สูงกว่าแสดงถึงความมั่นใจ / ความน่าจะเป็นที่สูงขึ้นของวัตถุที่ตรวจพบและคะแนนที่ต่ำกว่าอาจส่งผลให้วัตถุที่ตรวจพบได้มากขึ้น ตอนนี้ถ้ามาเปลี่ยนคะแนนเกณฑ์เป็น 0.30 และ 0.05 ว่าเกิดอะไรขึ้น



รูปที่ 2 ผลลัพธ์ของการตรวจจับวัตถุจาก SSD / MobileNet และ YOLOv2 (คะแนน = 0.30)



รูปที่ 3 ผลลัพธ์ของการตรวจจับวัตถุจาก SSD / MobileNet และ YOLOv2 (คะแนน = 0.05)

โดยการเปรียบเทียบและทำสัญญารูปที่ 2 และรูปที่ 3

- SSD / MobileNet เป็นเครื่องมือตรวจจับวัตถุชัดเจนซึ่งอาจตรวจจับวัตถุได้มากขึ้น แต่ทำผิดพลาดน้อย
- YOLOv2 เป็นเครื่องมือตรวจจับแบบเก่า ซึ่งอาจตรวจจับวัตถุบางอย่างและดี

2. ขั้นตอนการทำงานของการทำงานตรวจจับวัตถุใน OpenCV Deep Neural Network Module

บทสรุปเกี่ยวกับกรอบทั่วไปและความแตกต่างระหว่างทั้งสอง model

1. เลือกและกำหนดรูปแบบ : ขั้นตอนนี้ดำเนินการโดย `cv.dnn.readNetFromTensorflow()` ใน Line 4 ของรูปที่ 4 และ `cv.dnn.readNetFromDarknet()` ใน Line 4 ของรูปที่ 5; หมายเหตุ: ไฟล์การกำหนดค่าจะถูกดาวน์โหลดจาก official website
2. ประมวลผลภาพล่วงหน้า : ขั้นตอนนี้ดำเนินการโดย `cv.dnn.blobFromImage()` ใน Line 6 ของรูปที่ 4 และใน Line 6 ของรูปที่ 5 โดยที่ `blobFromImage()` ถูกใช้เพื่อดำเนินการชุดของการดำเนินการประมวลผลล่วงหน้าเช่น การปรับสเกลการทำให้เป็นมาตรฐานการลบหมายถึง ฯลฯ
3. ดำเนินการทำนาย : ขั้นตอนนี้ทำโดยส่งต่อ หมายเหตุ: ฟังก์ชันนี้ใช้เพื่อรับข้อมูลของวัตถุที่ถูกทำนาย เช่นดัชนีคลาสกล่องขอบเขตความเชื่อมั่น / ความน่าจะเป็น เป็นต้น

หลังจากรันโค้ด จะได้รับการทำนายรูปร่างดังต่อไปนี้

- รูปร่างการทำนาย SSD / MobileNet = (1, 1, 100, 7)
- รูปร่างทำนาย YOLOv2 = (845, 85)

3. การประมวลผลการทำนาย

(1) รูปร่างทำนายของ SSD / MobileNet = (1, 1, 100, 7)

SSD / MobileNet คาดการณ์วัตถุ 100 ขึ้นบนภาพอินพุต แต่ละวัตถุถูกระบุโดยคุณลักษณะสามอย่าง: a class index, a score, and a bounding box แสดงรหัสสำหรับการประมวลผลคุณลักษณะเหล่านี้และการวาด bounding boxes.

(2) YOLOv2 prediction shape = (845, 85)

YOLOv2 คาดการณ์ 845 bounding boxes บนภาพอินพุต หมายเลขมายากล 845 มาจากการออกแบบดังต่อไปนี้: YOLOv2 แบ่งภาพอินพุตให้เป็นตาราง 13 เซลล์จาก 13 เซลล์และแต่ละเซลล์มีหน้าที่ในการทำนาย 5 bounding boxes ดังนั้นจำนวนรวมของกล่องที่ถูกผูกไว้ในภาพที่ป้อนคือ $13 \times 13 \times 5 = 845$ ที่นี่แต่ละวัตถุจะถูกระบุโดยสามคุณลักษณะ: คะแนน, กล่องขอบเขต (x_center, y_center , ความกว้าง, ความสูง) และ ความน่าจะเป็น softmax มากกว่า 80 คลาสในชุดข้อมูล COCO

(3) Class Index and Class Name ในชุดข้อมูล COCO

จนถึงขณะนี้เรามีการตรวจสอบความแตกต่างระหว่างการทำนายผลการทำนายภายใต้ SSD MobileNet และ YOLOv2 มีข้อเท็จจริงที่ไม่ดีอีกอย่างหนึ่ง แม้ว่าทั้งสองโมเดลจะได้รับการ trained โดยชุดข้อมูล COCO เดียวกัน แต่ใช้ฟังก์ชันการทำแผนที่ที่แตกต่างกันใน class index และ class names

➤ แสดงชื่อคลาส 80 คลาสซึ่งถูกอ้างอิงโดย YOLOv2 ที่นำมาใช้โดย Darknet

```
classes_80 =
["person",
"bicycle",
"car",
"motorcycle",

"airplane", "bus", "train", "truck", "boat", "traffic light", "fire hydrant",

"stop sign", "parking meter", "bench", "bird", "cat", "dog", "horse",

"sheep", "cow", "elephant", "bear", "zebra", "giraffe", "backpack",

"umbrella", "handbag", "tie", "suitcase", "frisbee", "skis",

"snowboard", "sports ball", "kite", "baseball bat", "baseball glove",
"skateboard",

"surfboard", "tennis racket", "bottle", "wine glass", "cup", "fork", "knife",

"spoon", "bowl", "banana", "apple", "sandwich", "orange", "broccoli", "carrot",
"hot dog",

"pizza", "donut", "cake", "chair", "couch", "potted plant", "bed", "dining
table",

"toilet", "tv", "laptop", "mouse", "remote", "keyboard",

"cell phone", "microwave", "oven", "toaster", "sink", "refrigerator",

"book", "clock", "vase", "scissors", "teddy bear", "hair drier", "toothbrush" ]
```

80 class names

- แสดงชื่อคลาส 90 รายการ (ที่มีรายการ “ไม่รู้จัก”) ซึ่งอ้างอิงโดย SSD / MobileNet ที่ TensorFlow นำมาใช้

classes_90 =

```
[ "background", "person", "bicycle", "car", "motorcycle", "airplane", "bus", "train", "truck",
  "boat", "traffic light", "fire hydrant", "unknown", "stop sign", "parking meter", "bench", "bird",
  "cat", "dog", "horse", "sheep", "cow", "elephant", "bear", "zebra", "giraffe", "unknown",
  "backpack", "umbrella", "unknown", "unknown", "handbag", "tie", "suitcase", "frisbee", "skis",
  "snowboard", "sports ball", "kite", "baseball bat", "baseball glove", "skateboard",
  "surfboard", "tennis racket", "bottle", "unknown", "wine glass", "cup", "fork", "knife", "spoon",
  "bowl", "banana", "apple", "sandwich", "orange", "broccoli", "carrot", "hot dog", "pizza",
  "donut", "cake", "chair", "couch", "potted plant", "bed", "unknown", "dining table",
  "unknown", "unknown", "toilet", "unknown", "tv", "laptop", "mouse", "remote", "keyboard",
  "cell phone", "microwave", "oven", "toaster", "sink", "refrigerator", "unknown", "book",
  "clock", "vase", "scissors", "teddy bear", "hair drier", "toothbrush" ]
```

90 class names

4. OpenCV การประมวลผลภาพ

OpenCV เป็นไลบรารีแบบหลาย Platform เพื่อใช้ในการพัฒนา Application Vision System แบบ Realtime ซึ่งทำงานเกี่ยวกับการประมวลผลภาพ การจับภาพ วิดีโอ สามารถวิเคราะห์คุณลักษณะต่าง ๆ ตัวอย่างเช่น ตรวจสอบใบหน้า ตรวจสอบวัตถุ

คุณลักษณะของ OpenCV Library

- ดำเนินการตรวจหาคุณลักษณะ
- วิเคราะห์วิดีโอเช่นประมาณการเคลื่อนไหวในวิดีโอลบพื้นหลังและติดตามวัตถุในวิดีโอ
- จับภาพและบันทึกวิดีโอ
- ประมวลผลภาพ (กรอง, แปลง)

- ตรวจหาวัตถุเฉพาะเช่นใบหน้าดวงตารอยนตในวิดีโอหรือภาพ
- อ่านและเขียนภาพ

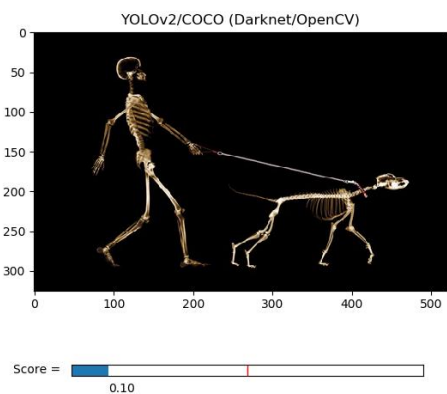
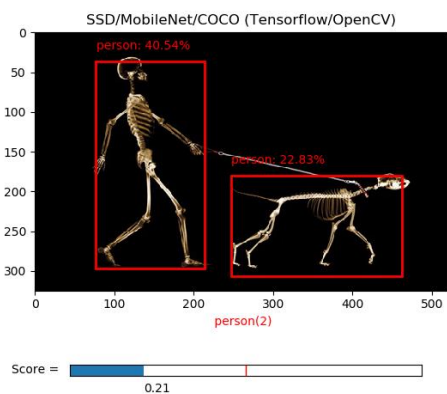
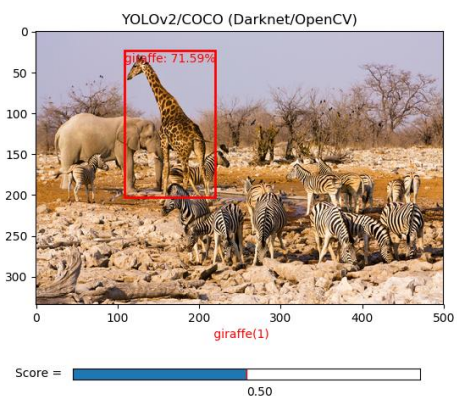
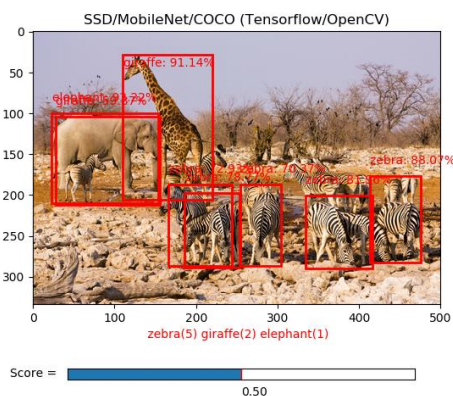
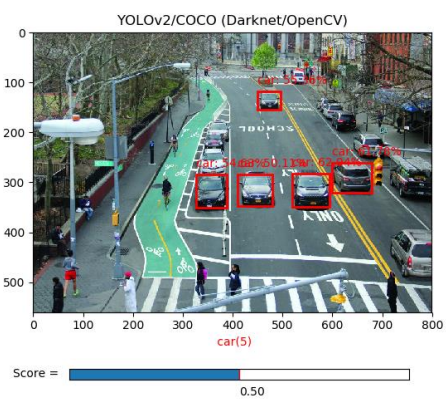
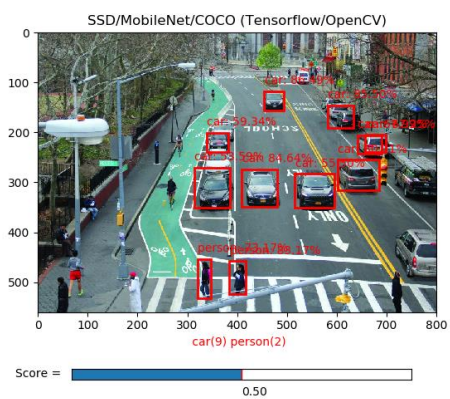
งานย่อยในการประมวลผลภาพสามารถแบ่งได้เป็นดังนี้

- การปรับปรุงและฟื้นฟูภาพ : สำหรับการปรับปรุงข้อมูลภาพ
- การได้มาของภาพ, การจัดเก็บ, การส่งผ่าน : การแปลงเป็นตัวเลข / ดิจิทัล, การบีบอัด, การเข้ารหัส / ถอดรหัส
- การสกัดข้อมูล : สำหรับการวิเคราะห์คอมพิวเตอร์เพิ่มเติม

5. สรุป

ได้อธิบายการศึกษาเบื้องต้นเกี่ยวกับกลไกของ the deep learning object detection library ใน OpenCV ไม่มีโมเดลที่ชัดเจนระหว่างSSD / MobileNet โดย Tensorflow และYOLOv2ดำเนินการโดย Darknet ขึ้นอยู่กับความต้องการของแอปพลิเคชันรวมถึงความเข้าใจที่ชัดเจนเกี่ยวกับการวัดประสิทธิภาพที่แตกต่างและการใช้หน่วยความจำเพื่อเป็นแนวทางในกระบวนการเลือกรูปแบบ

ตัวอย่างเพิ่มเติมสำหรับการอ้างอิง



2.4 งานวิจัยที่เกี่ยวข้อง

(หอมานน, 2551) งานวิจัยนี้เป็นการศึกษาและสร้างหน่วยตรวจจับขอบสำหรับการตรวจจับวัตถุเคลื่อนที่ภายใต้ โครงข่ายพื้นที่ท้องถิ่นไร้สายโดยใช้ปัญญาประดิษฐ์และกรรมวิธีสัญญาณดิจิทัล ในงานวิจัยมี พารามิเตอร์ การเคลื่อนที่ที่ถูกกำหนด ได้แก่ ความไว เวลา ระยะทาง และความเร็ว โดยกระบวนการ สัญญาณประกอบด้วย 2 ขั้นตอนคือ 1) ขั้นตอนการตรวจจับขอบ ได้แก่ ขั้นตอนวิธีดิจิทัล และขั้นตอน วิธีปัญญาประดิษฐ์และเชิงเลข ซึ่งงานวิจัยเล่มนี้นำเสนอขั้นตอนวิธีดิจิทัล โดยใช้ขั้นตอนวิธีไฮบริด (HA: Hybrid Algorithm) ซึ่งผสมผสานขั้นตอนวิธีโซเบล (SA: Sobel Algorithm) และขั้นตอนวิธี แคนนี่ (CA: Canny Algorithm) โดยผลของขอบภาพที่ดีขึ้นจาก HA ได้รับการเน้นด้วยดัชนีการเรียนรู้ และดัชนีกำลังจะให้ขอบเข้มจาก SA และขอบจางจาก CA นอกจากนี้ HA ยังมีความยืดหยุ่นกว่า ขั้นตอนวิธีดั้งเดิมเนื่องจากมีระดับความอิสระมากกว่า 2) ขั้นตอนการคำนวณความเร็วบนพื้นฐาน ระยะทางยูคลิเดียน เพื่อตรวจวัดคุณลักษณะของการเคลื่อนที่ด้วยการหาจุดศูนย์กลางขอบ ผลที่ได้ พบว่าความเร็วแบบเวลาจริงสอดคล้องกับการเคลื่อนที่ของวัตถุตรวจจับ

เราเสนอวิธีการตรวจจับวัตถุที่ปรับปรุงความแม่นยำของ SSD ทั่วไป (Single Shot Multibox Detector) ซึ่งเป็นหนึ่งในอัลกอริทึมการตรวจจับวัตถุชั้นนำทั้งในแง่ของความแม่นยำและความเร็ว ประสิทธิภาพของเครือข่ายเชิงลึกนั้นได้รับการปรับปรุงให้ดีขึ้นเมื่อจำนวนของคุณสมบัติแผนที่เพิ่มขึ้น อย่างไรก็ตามเป็นการยากที่จะปรับปรุงประสิทธิภาพด้วยการเพิ่มจำนวนของคุณสมบัติแผนที่ ในบทความนี้เราเสนอและวิเคราะห์วิธีการใช้คุณสมบัติแผนที่อย่างมีประสิทธิภาพเพื่อปรับปรุงประสิทธิภาพของ SSD ทั่วไป ประสิทธิภาพที่เพิ่มขึ้นนั้นได้มาจากการเปลี่ยนโครงสร้างใกล้กับเครือข่ายลักษณะนามแทนที่จะเพิ่มเลเยอร์ใกล้กับข้อมูลอินพุตเช่นโดยแทนที่ VGGNet ด้วย ResNet เครือข่ายที่เสนอนั้นเหมาะสำหรับการแบ่งปันน้ำหนักในเครือข่ายลักษณะนามโดยคุณสมบัติการฝึกรอบมจะเร็วขึ้นด้วยการวางนัยทั่วไปที่ดีขึ้น สำหรับชุดทดสอบ Pascal VOC 2007 ที่ผ่านการฝึกรอบมด้วยชุดการฝึกรอบม VOC 2007 และ VOC 2012 เครือข่ายที่นำเสนอด้วยขนาดอินพุต 300×300 ได้รับ 78.5% mAP (ความแม่นยำเฉลี่ยเฉลี่ย) ที่ความเร็ว 35.0 FPS (เฟรมต่อวินาที) ในขณะที่ เครือข่ายที่มีอินพุตขนาด 512×512 ได้รับ 80.8% mAP ที่ 16.6 FPS โดยใช้ Nvidia Titan X GPU เครือข่ายที่นำเสนอแสดง mAP ที่ทันสมัยซึ่งดีกว่าของ SSD ทั่วไป, YOLO, Faster-RCNN และ RFCN นอกจากนี้ยังเร็วกว่า Faster-RCNN และ RFCN

(Jeong, Park, & Kwak, 2560) ใช้วิธีการตรวจจับวัตถุที่ปรับปรุงความแม่นยำของ SSD ทั่วไป (Single Shot Multibox Detector) ซึ่งเป็นหนึ่งในอัลกอริทึมการตรวจจับวัตถุชั้นนำ ทั้งในแง่ของความแม่นยำและความเร็ว ประสิทธิภาพของเครือข่ายเชิงลึกนั้นได้รับการปรับปรุงให้ดีขึ้นเมื่อจำนวนของ feature maps เพิ่มขึ้น อย่างไรก็ตามยากที่จะปรับปรุงประสิทธิภาพด้วยการเพิ่มจำนวนของ feature maps ในบทความนี้เสนอและ

วิเคราะห์วิธีการใช้ feature maps อย่างมีประสิทธิภาพเพื่อปรับปรุงประสิทธิภาพของ SSD ทั่วไป ประสิทธิภาพที่เพิ่มขึ้นนั้นได้มาจากการเปลี่ยนโครงสร้างใกล้กับ classifier network แทนที่จะเพิ่มเลเยอร์ใกล้กับข้อมูลอินพุต เช่นโดยแทนที่ VGGNet ด้วย ResNet เครือข่ายที่เสนอนั้นเหมาะสำหรับการแบ่งปันน้ำหนักใน classifier networks โดยคุณสมบัติการ training จะเร็วขึ้นด้วยพลังการ generalization ทั่วไปที่ดีขึ้น สำหรับ test ทดสอบ Pascal VOC 2007 ที่ผ่าน test ด้วยชุด trained VOC 2007 และ VOC 2012 proposed network ด้วยขนาดอินพุต 300×300 ได้รับ 78.5% mAP ((mean average precision) ที่ความเร็ว 35.0 FPS (เฟรมต่อวินาที) ในขณะที่ network ที่มีอินพุตขนาด 512×512 ได้รับ 80.8% mAP ที่ 16.6 FPS โดยใช้ Nvidia Titan X GPU เครือข่ายที่นำเสนอแสดง mAP ที่ทันสมัยซึ่งดีกว่าของ SSD ทั่วไป, YOLO, Faster-RCNN และ RFCN นอกจากนี้ยังเร็วกว่า Faster-RCNN และ RFCN

(งามวงศ์ชน, 2549) วิทยานิพนธ์ฉบับนี้กล่าวถึงการตรวจจับการเคลื่อนไหวของวัตถุในภาพจากกล้องเดียว ชนิดคงที่ เป็นการนำสัญญาณภาพต่อเนื่องจากกล้องวิดีโอมาหาความแตกต่างโดยเทียบกับ สัญญาณภาพตั้งต้น สัญญาณภาพปัจจุบันมีความแตกต่างจากภาพตั้งต้นในส่วนที่เป็นวัตถุที่กำลัง เคลื่อนที่ จากนั้นทำการปรับปรุงคุณภาพของสัญญาณเพื่อให้วัตถุที่กำลังเคลื่อนที่มีความชัดเจน และเด่นชัดพอที่สามารถหาขอบเขตของแต่ละวัตถุที่กำลังเคลื่อนที่ เพื่อที่จะทำการตีกรอบ ล้อมรอบวัตถุแต่ละวัตถุได้อย่างถูกต้อง การวิจัยนี้ทดสอบกับทั้ง ภาพนิ่ง จำนวน 35 คู่ภาพ ภาพ วิดีทัศน์ (วิดีโอ) จำนวน 29 ชุด 13665 เฟรม และภาพจากกล้องวิดีโอ จำนวน 10 ชุด ระบบ ระบุขอบเขตได้อย่างถูกต้องคิดเป็นร้อยละ 96.50

(เวฬุวนาร, 2555) ปัจจุบันมีความพยายามสร้างระบบการเฝ้าระวังบุคคลผู้ต้องสงสัยแบบอัตโนมัติ ด้วยการรู้จำใบหน้า (Face Recognition) เพื่อลดปัญหาที่เกิดจากข้อจำกัดของการเฝ้าระวัง ด้วยคนที่มีการเมื่อยล้า และยากที่จะสามารถจดจำใบหน้าผู้ต้องสงสัยได้ทั้งหมด จากการศึกษา งานวิจัยที่ผ่านมาพบว่า งานวิจัยส่วนใหญ่ใช้ภาพใบหน้าทั้งใบหน้าเพียงอย่างเดียวในการรู้จำ ซึ่งทำให้ได้ประสิทธิภาพของการรู้จำยังไม่ดีเท่าที่ควร งานวิจัยนี้จึงนำเสนอวิธีการเพิ่ม ประสิทธิภาพการรู้จำใบหน้า โดยใช้คุณลักษณะของภาพใบหน้าทั้งใบหน้าร่วมกับคุณลักษณะ เฉพาะส่วนของใบหน้าที่อีก 4 ส่วน คือ ตาขวา ตาซ้าย จมูก และปาก นอกจากนี้ยังได้ศึกษา พบว่าเทคนิคที่สามารถแก้ปัญหการจำแนกข้อมูล โดยเฉพาะข้อมูลภาพได้อย่างมีประสิทธิภาพและน่าเชื่อถือในปัจจุบันคือ ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine : SVM) แบบเรเดียลเบสฟังก์ชัน (Radial Basis Function : RBF) ซึ่งมีพารามิเตอร์ที่สำคัญ 2 ตัว คือ C และ 1 งานวิจัยที่ผ่านมาโดยส่วนใหญ่ได้กำหนดค่าของพารามิเตอร์ด้วยมือ ซึ่งอาจจะไม่ใช่ค่าความเหมาะสมที่ดีที่สุด จึงส่งผลกระทบต่อการใช้จำได้ไม่ดีเท่าที่ควร จากปัญหาที่ได้อธิบายมา งานวิจัยนี้จึงนำเสนอขั้นตอนวิธีการเพิ่มประสิทธิภาพการรู้จำใบหน้า บนวิดีโอโดยใช้หลาย

คุณลักษณะ ด้วยซอฟต์แวร์แมชชีนแบบเรเดียลเบสฟังก์ชัน นอกจากนี้แล้ว ยังนำเสนอวิธีการหาค่าความเหมาะสมที่ดีที่สุดให้กับพารามิเตอร์ C และ " แบบอัตโนมัติด้วยพาร์ทิเคิลสวอมออฟทิโมเซชัน (Particle Swarm Optimization : PsO) ผลของการวิจัยแสดงให้เห็นว่า ขั้นตอนวิธีการที่ได้นำเสนอในงานวิจัยนี้ เมื่อเปรียบเทียบกับขั้นตอนวิธีการเดิมแล้ว สามารถเพิ่มประสิทธิภาพการรู้จำใบหน้าโดยเฉลี่ยได้ 9.94% โดยมี ประสิทธิภาพการรู้จำสูงสุดคือ 97.92%

(ขุนอาสา, 2552) สารนิพนธ์ฉบับนี้ได้ นำแบบแผนการตรวจจับวัตถุของ Viola และ Jones มาประยุกต์ใช้ในการ ตรวจจับคนเดินถนนในเวลากลางคืนจากภาพวิดีโอที่คนในโหมดถ่ายภาพในที่มืดที่บันทึกในรถยนต์ ขณะ เคลื่อนที่ ผู้วิจัยได้นำวิธีการปรับปรุงภาพวิดีโอที่คนไม่ซับซ้อนมาใช้เพื่อเพิ่มประสิทธิภาพในการตรวจจับ รวมถึง ได้นำวิธีการปรับลดอัตราความผิดพลาดในการตรวจจับด้วยค่าเฉลี่ยของความสว่าง โดยทำการทดสอบใน 4 กรณีคือ 1) ก่อนปรับความเข้ม 2) การบวกด้วยค่าคงที่ 3) การคูณด้วยค่าคงที่ และ 4) การปรับค่าแกมมา และนำผลในแต่ละกรณีมาปรับลดอัตราความผิดพลาด ซึ่งผลการทดลองวิธีการปรับปรุงภาพได้อัตราการ ตรวจจับที่ถูกต้อง เพิ่มขึ้นจาก 72.50% เป็น 96.75%, 97.0% และ 98.0% ตามลำดับ และอัตราการตรวจจับที่ ผิดพลาดหลังจาก ปรับลดแล้วเหลือ 4.33%, 3.81%, 3.95% และ 3.22% ตามลำดับ ดังนั้นวิธีการนี้สามารถ นำไปใช้งานในระบบ ช่วยเหลือการขับขี่ได้อย่างมีประสิทธิภาพ

(บำรุงเกียรติ, 2551) การติดตามวัตถุในเวลาจริงนั้นถูกนำมาประยุกต์ใช้ในงานร่วมกับวิดิทัศน์ในหลากหลายประเภท เช่น ใช้เพื่อประโยชน์ในการติดตามบุคคลที่สนใจของระบบรักษาความปลอดภัย ใช้ในการติดตามวัตถุต่าง ๆ ด้วยหุ่นยนต์ติดตาม หรือใช้ติดตามรถยนต์ในระบบขนส่งจราจรอัจฉริยะ เป็นต้น ซึ่งงานประยุกต์เกี่ยวกับการติดตามวัตถุในเวลาจริงดังกล่าว นั้น ต้องการขั้นตอนวิธีการติดตามที่มีประสิทธิภาพ แม่นยำ โดยเฉพาะอย่างยิ่งเมื่อถูกนำมาใช้งานในข้อจำกัดที่ยากต่อการติดตามหรือในสถานการณ์จำกัดต่าง ๆ เช่น เมื่อมุมมองของวิดิทัศน์มีการเปลี่ยนแปลง ความสว่างของบริเวณที่ต้องการติดตามมีการเปลี่ยนแปลง หรือเมื่อมีการใช้งานส่งผ่านข้อมูลระหว่างกล้องวิดิทัศน์หลายตัวให้ทำงานร่วมกัน เป็นต้น วิทยานิพนธ์ฉบับนี้นำเสนอขั้นตอนวิธีการติดตามบุคคลอย่างต่อเนื่องในเวลาจริง โดยใช้กล้องวิดิทัศน์แบบสาย ก้มเงย และซูม (Pan Tilt Zoom Camera; PTZ) 2 ตัว ทำงานร่วมกัน โดยอาศัยขั้นตอนวิธีการย้ายตามค่าเฉลี่ย ซึ่งส่วนใหญ่แล้วคุณลักษณะสำคัญที่ใช้ในการติดตามวัตถุหรือบุคคลนั้น จะอาศัยเพียงค่าสีเท่านั้น ซึ่งจะเกิดปัญหาเป็นอย่างมากเมื่อความสว่างหรือมุมมองของกล้องวิดิทัศน์มีการเปลี่ยนแปลง ซึ่งวิทยานิพนธ์เล่มนี้จะใช้คุณลักษณะสำคัญสองประการคือ สี เช่น สีผิวหรือสีเสื้อผ้าและขอบของร่างกายของบุคคล โดยที่กล้องวิดิทัศน์ตัวที่หนึ่งในระบบจะทำหน้าที่เก็บข้อมูลคุณลักษณะสำคัญของบุคคลที่สนใจและทำการติดตามด้วยขั้นตอนวิธีดังกล่าวข้างต้น และเมื่อบุคคลที่สนใจนั้นนอกจากมุมมอง

ของกล้องวิดีโอตัวที่หนึ่งมายังมุมมองของกล้องวิดีโอตัวที่สอง กล้องวิดีโอตัวที่สองในระบบจะรับค่าคุณลักษณะสำคัญนั้น ๆ จากกล้องวิดีโอตัวที่หนึ่ง และทำการติดตามบุคคลที่สนใจนั้นอย่างต่อเนื่องต่อไป นอกจากนี้ยังอาศัยการเทียบมาตรฐานทางสีระหว่างกล้องวิดีโอ เพื่อให้การส่งข้อมูลสีระหว่างกล้องวิดีโอทั้งสองตัวในระบบสามารถติดตามได้อย่างถูกต้อง

(ทูลธรรม, 2556) วิทยานิพนธ์ฉบับนี้มีวัตถุประสงค์เพื่อพัฒนาชุดการสอนเรื่องประยุกต์ใช้ขบวนการวิเคราะห์ภาพ สำหรับตรวจจับวัตถุ ด้วยวิธีจัดการเรียนรู้แบบร่วมมือ โดยใช้เทคนิคจิกซอร์มีขั้นตอนในการจัดการเรียน 6 ขั้นตอน คือ 1) ขั้นเตรียม 2) ขั้นจัดกลุ่มผู้เรียน 3) ขั้นแจกเอกสารและมอบหมายภารกิจ 4) จัดกลุ่มผู้เชี่ยวชาญ 5) สมาชิกกลับกลุ่มบ้านเพื่อนำเสนอความรู้ 6) ทดสอบความรู้ ประกอบด้วยเนื้อหาเกี่ยวกับการปรับปรุงภาพโดยการประมวลผลภาพแบบจุด การปรับปรุงภาพโดยการประมวลผลภาพแบบบริเวณ การประยุกต์ใช้การประมวลผลภาพด้วย Microsoft Visual C++ โดยใช้ MFC และเรื่องการประยุกต์ใช้การประมวลผลภาพสำหรับการคัดแยกวัตถุ รูปทรง สามเหลี่ยม สี่เหลี่ยม และวงกลม ด้วยวิธีฮัฟฟทรานฟอร์ม โดยใช้โปรแกรม Microsoft Visual C++ ร่วมกับ Open CV เครื่องมือที่ใช้ในการวิจัยได้แก่ แผนการจัดการเรียนรู้ แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน และ แบบสอบถามความความคิดเห็น กลุ่มตัวอย่างที่ใช้ในการทดลองจำนวนทั้งหมด 25 คน ได้แก่ผู้เรียนที่ลงทะเบียนวิชาการศึกษาประมวลผลสัญญาณภาพ และผ่านการเรียนในรายวิชาการโปรแกรมคอมพิวเตอร์มาแล้ว ในหลักสูตรครุศาสตรบัณฑิต สาขาวิศวกรรมคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี โดยเลือกแบบเจาะจง ผลของการวิจัยสรุปได้ดังนี้ 1) ความคิดเห็นของผู้เชี่ยวชาญที่มีต่อชุดการสอนเรื่องประยุกต์ ใช้ขบวนการวิเคราะห์ภาพ สำหรับตรวจจับวัตถุ ด้วยวิธีจัดการเรียนรู้แบบร่วมมือ โดยใช้เทคนิคจิกซอร์มีความเหมาะสมอยู่ในระดับมาก 2) ประสิทธิภาพของชุดการสอนเรื่องประยุกต์ใช้ขบวนการวิเคราะห์ภาพ สำหรับตรวจจับวัตถุ ด้วยวิธีจัดการเรียนรู้แบบร่วมมือ โดยใช้เทคนิคจิกซอร์ที่พัฒนาขึ้นมีประสิทธิภาพ (1.34) สูงกว่าเกณฑ์การหาคุณภาพของเมกยูแกนส์ (Meguigans) 3) ผลสัมฤทธิ์ของการเรียนการสอนด้วยชุดการสอนเรื่องประยุกต์ใช้ขบวนการวิเคราะห์ภาพ สำหรับตรวจจับวัตถุ ด้วยวิธีจัดการเรียนรู้แบบร่วมมือ โดยใช้เทคนิคจิกซอร์ที่พัฒนาขึ้นมีค่าร้อยละ 77.36 ซึ่งมีค่าสูงกว่าเกณฑ์ที่กำหนดร้อยละ 70 และ 4) ความพึงพอใจของผู้เรียนที่ผ่านการเรียนด้วยชุดการสอนเรื่องประยุกต์ใช้ขบวนการวิเคราะห์ภาพ สำหรับตรวจจับวัตถุ ด้วยวิธีจัดการเรียนรู้แบบร่วมมือ โดยใช้เทคนิคจิกซอร์อยู่ในระดับมากดังนั้นสรุปได้ว่าชุดการสอนเรื่องประยุกต์ใช้ขบวนการวิเคราะห์ภาพ สำหรับตรวจจับวัตถุ ด้วยวิธีจัดการเรียนรู้แบบร่วมมือ โดยใช้เทคนิค จิกซอร์สามารถนำไปประยุกต์ใช้กับการเรียนการสอนรายวิชาการประมวลผลสัญญาณภาพ ได้อย่างมีประสิทธิภาพ

(ออกเวหา, 2555) การระบุตำแหน่งของวัตถุโดยอยู่บนของภาพ 2 มิติ นั้นจำเป็นอย่างยั้งที่ต้องใช้ข้อมูลภาพถ่ายในหลายๆ มุมมองเพื่อนำมาประมวลผล จึงเป็นผลให้มีการประมวลผล จึงเป็นผลให้มีการประมวลผลที่ซ้ำและต้องการข้อมูลที่มาก ดังนั้นวิทยานิพนธ์นี้จึงทำการศึกษาคำตรวจหาและระบุวัตถุที่อยู่บนสุดของภาพ 2 มิติ ที่ทับซ้อนโดยใช้ข้อมูลภาพมูมบนเพียงภาพเดียวร่วมกับเทคนิคมาร์จินอลสเปซและมอร์โฟโลยี การทดลองได้ใช้ภาพถ่ายวัตถุขนาด 480x640 พิกเซล มาทำการหาขอบภาพโดยใช้การหาขอบภาพอนุพันธ์อันดับสอง เทคนิคที่นำมาใช้ คือ เทคนิคการลาปลาซของเกาส์เซียน (Laplacian of a Gaussian) กรองภาพโดยใช้ตัวกรองแบบโมชัน (Motion Filter) และทำการปรับปรุงภาพโดยใช้เทคนิคมอร์โฟโลยี ผลการทดลองแสดงให้เห็นว่าเทคนิคดังกล่าวมีประสิทธิภาพในการตรวจหาวัตถุที่ต้องการ จากนั้นทำการระบุวัตถุที่อยู่ด้านบนโดยใช้เทคนิคดิสแตนทรานฟอร์ม เทคนิคการแบ่งส่วนภาพสันปันน้ำ (Watershed Segmentation) และเทคนิคหาจุดศูนย์กลาง ผลการทดลองการระบุวัตถุที่อยู่ด้านบนสุดพบว่า เทคนิคดิสแตนทรานฟอร์มจะมีประสิทธิภาพในการระบุกลุ่มวัตถุมากกว่าระบุวัตถุเดียว ส่วนเทคนิคการแบ่งส่วนภาพสันปันน้ำมีประสิทธิภาพในการระบุวัตถุในภาพที่มีการทับซ้อน เช่น ภาพวัตถุทรงกลม ภาพวัตถุรูปทรงสี่เหลี่ยมผืนผ้า และภาพวัตถุรูปทรงสี่เหลี่ยมจัตุรัส โดยให้ประสิทธิภาพร้อยละ 93.33 29.03 และ 74.36 ตามลำดับ และเทคนิคจุดศูนย์กลางให้มีประสิทธิภาพในการระบุวัตถุบนสุดร้อยละ 68.88 ในภาพที่มีการทับซ้อนมาก

(David G. , 2542) ระบบการจำแนกวัตถุได้รับการพัฒนาที่ใช้ new class ของ f local image features feature ดังกล่าวไม่แปรเปลี่ยนไปตามการปรับขนาดภาพการแปลและการหมุนและคงที่บางส่วนสำหรับการเปลี่ยนแปลงความสว่างและการเลียนแบบหรือการฉายภาพ 3 มิติ features เหล่านี้แบ่งปัน similar properties ที่คล้ายกันกับ neurons ใน inferior temporal cortex ที่ใช้สำหรับการรับรู้วัตถุใน primate vision ตรวจจับคุณสมบัติได้อย่างมีประสิทธิภาพผ่านวิธีการกรองแบบฉากซึ่งระบุจุดคงที่ในพื้นที่ขนาด ปุ่มรูปภาพถูกสร้างขึ้นที่อนุญาตสำหรับการเปลี่ยนรูปทรงเรขาคณิตในพื้นที่โดยแสดงการไล่ระดับสีของภาพเบลอ ในระนาบการวางแผนหลายจุดและหลายระดับ คีย์ถูกใช้เป็นอินพุตไปยังวิธีการจัดทำดัชนี nearest-neighbor ซึ่งระบุการจับคู่วัตถุ การตรวจสอบขั้นสุดท้ายของการแข่งขันแต่ละครั้งนั้นทำได้โดยการหาวิธีแก้ปัญหากำลังสองน้อยที่สุดที่ต่ำสำหรับค่าพารามิเตอร์รุ่นที่ไม่รู้จัก ผลการทดลองแสดงให้เห็นว่าการรับรู้วัตถุที่มีประสิทธิภาพสามารถทำได้ในรูปแบบที่กระจัดกระจายบางส่วนที่มีเวลาในการคำนวณน้อยกว่า 2 วินาที

(สุรสารินทร์, 2557) ลักษณะของชนิดเสื้อผ้ามีความสำคัญในการใช้ชีวิตประจำวัน ซึ่งเสื้อผ้าจะเป็นส่วนแรกที่บรรยายตัวบุคคลว่าเป็นอย่างไรและแสดงให้เห็นถึงลักษณะการใช้ชีวิตประจำวัน วัฒนธรรมและ สถานะของสังคมของผู้ที่สวมใส่เสื้อผ้าชนิดนั้น ๆ ตัวอย่างเช่น เราสามารถจดจำผู้คนแต่ละคนได้ว่ามี ลักษณะเป็นอย่างไรจากเสื้อผ้าของพวกเขา เป็นต้น และอีกหลายๆ บริบทที่สามารถบรรยายได้ด้วย ชนิดของเสื้อผ้า เช่น ถ้ามีผู้สวมใส่ชุดสูทจะสามารถบรรยายได้ว่าพวกเขาอยู่ในสถานการณ์ทาง ธุรกิจ เป็นต้น ซึ่งการจดจำชนิดของเสื้อผ้ายังเป็นปัญหาที่ท้าทายต่อการคอมพิวเตอร์วิทัศน์ ดังนั้น ในวิทยานิพนธ์ฉบับนี้ได้นำเสนอวิธีการใหม่ในการจดจำและแบ่งแยกชนิดของเสื้อผ้าโดยใช้ SURF ร่วมกับ LBP ด้วยเทคนิคคุณภาพของคุณลักษณะสำคัญ ซึ่งรูปภาพที่นำเข้าทดสอบและสอนจะนำมา จากแคตตาล็อกแฟชั่นที่เป็นภาพเสื้อผ้าและใบหน้าของผู้ที่สวมใส่ชัดเจนสำหรับขั้นตอนการประมวลผลภาพและการสร้างคุณลักษณะชนิดของเสื้อผ้าจากเทคนิคคุณภาพของคุณลักษณะเพื่อรู้จำและแบ่งแยกชนิดของเสื้อผ้ามีขั้นตอนการทำงานทั้งหมด 4 ขั้นตอน คือ ขั้นแรกเสื้อผ้าในรูปภาพจะถูกค้นหาและระบุตำแหน่ง แล้วจึงแบ่งส่วนด้วยวิธีการของ Grab Cut ต่อมาพื้นที่ที่เป็นเสื้อผ้าจะถูกแบ่งออกเป็นสามส่วนย่อย ๆ กล่าวคือ ด้านขวา ตรงกลาง และด้านซ้าย โดยในแต่ละส่วนย่อย ๆ จะถูกดึงคุณลักษณะสำคัญด้วยวิธีของ SURF และ LBP เพื่อสร้างพจนานุกรม ของคุณลักษณะสำคัญต่อไป ผลสุดท้ายคุณลักษณะสำคัญต่าง ๆ จะถูกทำนายว่าเป็นเสื้อผ้าชนิดอะไร ด้วยซัพพอร์ตเวกเตอร์แมชชีน ฐานข้อมูลที่ใช้ในวิทยานิพนธ์ได้ใช้รูปภาพทั้งหมด 1131 รูปภาพแบ่งเป็นชุดสอนให้เครื่อง เรียนรู้ 991 รูปภาพและทดสอบ 140 รูปภาพ โดยเราได้แบ่งประเภทชนิดของเสื้อผ้าไว้ทั้งหมด 7 ชนิด ได้แก่ เสื้อแจ็คเก็ต เสื้อเชิ้ต เสื้อสูท เสื้อกันหนาว เสื้อยืด เสื้อโปโล และเสื้อกล้าม ผลที่ได้จากการทดลองแสดงให้เห็นว่าการรู้จำชนิดของเสื้อผ้ามีความถูกต้องเฉลี่ย 73.57%

(Anguelov, และคนอื่นๆ, 2560) เรานำเสนอวิธีการตรวจจับวัตถุในภาพโดยใช้โครงข่ายประสาทเทียมเดี่ยว วิธีการของเราที่ชื่อว่า SSD นั้นจะลดพื้นที่เอาต์พุตของกล่อง bounding ลงในชุดของกล่องเริ่มต้นผ่านอัตราส่วนและอัตราส่วนที่แตกต่างกันตามตำแหน่งของคุณสมบัติแผนที่ ในเวลาที่คาดการณ์เครือข่ายจะสร้างคะแนนสำหรับการปรากฏตัวของแต่ละประเภทของวัตถุในแต่ละกล่องเริ่มต้นและสร้างการปรับเปลี่ยนให้เหมาะสมกับรูปร่างของวัตถุ นอกจากนี้เครือข่ายยังรวมการทำนายจากแผนที่คุณลักษณะหลายอย่างที่มีความละเอียดต่างกันเพื่อจัดการกับวัตถุที่มีขนาดต่างๆ SSD นั้นค่อนข้างง่ายเมื่อเทียบกับวิธีการที่ต้องการข้อเสนอวัตถุเพราะมันจะช่วยลดการสร้างข้อเสนอและฟังก์ชันที่ตามมาหรือขั้นตอนการ resampling คุณสมบัติและ encapsulate การคำนวณทั้งหมดในเครือข่ายเดียว สิ่งนี้ทำให้ SSD ง่ายต่อการฝึกอบรมและตรงไปตรงมาเพื่อรวมเข้ากับระบบที่ต้องการส่วนประกอบตรวจจับ ผลการทดลองบนชุดข้อมูล PASCAL VOC, COCO และ ILSVRC ยืนยันว่า SSD มีความแม่นยำในการแข่งขันสำหรับวิธีการที่ใช้ขั้นตอนการเสนอวัตถุเพิ่มเติมและเร็วขึ้นมากในขณะเดียวกันก็มี

กรอบการทำงานแบบครบวงจรสำหรับการฝึกอบรมและการอนุมาน สำหรับ อินพุต (300×300) , SSD ได้รับ 74.3% mAP ในการทดสอบ VOC2007 ที่ 59 FPS บน Nvidia Titan X และสำหรับ (512×512) อินพุต SSD บรรลุ 76.9% mAP ดีกว่าสถานะที่เทียบเท่า ศิลปะรุ่น R-CNN ที่เร็วขึ้น เมื่อเทียบกับวิธีการขั้นตอนเดียวอื่น ๆ SSD มีความแม่นยำที่ดีกว่ามากแม้จะมีขนาดภาพเล็กลง รหัสที่มีอยู่

(Cao, และคนอื่น ๆ, 2561) การตรวจจับวัตถุขนาดเล็กเป็นงานที่ท้าทายในการมองเห็นคอมพิวเตอร์เนื่องจากความละเอียดและข้อมูลที่ จำกัด เพื่อที่จะแก้ปัญหาวิธีการส่วนใหญ่ที่มีอยู่จะต้องเสียสละความเร็วเพื่อปรับปรุงความแม่นยำ ในบทความนี้เรามุ่งมั่นที่จะตรวจจับวัตถุขนาดเล็กด้วยความเร็วที่รวดเร็วโดยใช้เครื่องตรวจจับวัตถุที่ดีที่สุด Single Shot Multibox Detector (SSD) ที่เกี่ยวกับความแม่นยำเทียบกับการแลกเปลี่ยนความเร็วเป็นสถาปัตยกรรมพื้นฐาน เราเสนอวิธีการพิกเซลคุณสมบัติหลายระดับสำหรับการแนะนำข้อมูลเชิงบริบทใน SSD เพื่อปรับปรุงความแม่นยำสำหรับวัตถุขนาดเล็ก ในการดำเนินการพิกเซลอย่างละเอียดเราออกแบบโมดูลพิกเซลสองโมดูลโมดูลเชื่อมต่อและโมดูลผลรวมองค์ประกอบแตกต่างกันในการเพิ่มข้อมูลตามบริบท ผลการทดลองแสดงให้เห็นว่าโมดูลพิกเซลทั้งสองนี้ได้รับ mAP ที่สูงขึ้นใน PASCAL VOC2007 มากกว่า SSD พื้นฐาน 1.6 และ 1.7 คะแนนตามลำดับโดยเฉพาอย่างยิ่งกับการปรับปรุงจุด 2-3 ในบางประเภทวัตถุขนาดเล็ก ความเร็วในการทดสอบของพวกเขา คือ 43 และ 40 FPS ตามลำดับซึ่งเหนือกว่าเครื่องตรวจจับกระสุนนัดเดียว Deconvolutional (DSSD) 29.4 และ 26.4 FPS

(Lowe, 2542) ระบบการจำแนกวัตถุได้รับการพัฒนาที่ใช้คลาสใหม่ของคุณสมบัติภาพท้องถิ่น คุณสมบัติดังกล่าวไม่แปรเปลี่ยนไปตามการปรับขนาดภาพการแปลและการหมุนและคงที่เพียงบางส่วนในการเปลี่ยนแปลงความสว่างและเลียนแบบหรือฉายภาพ 3 มิติ คุณสมบัติเหล่านี้แบ่งปันคุณสมบัติที่คล้ายกันกับเซลล์ประสาทในเยื่อหุ้มสมองส่วนล่างที่ใช้สำหรับการรับรู้วัตถุในการมองเห็นเข้าคณะ ตรวจจับคุณสมบัติได้อย่างมีประสิทธิภาพผ่านวิธีการกรองแบบฉากซึ่งระบุจุดคงที่ในพื้นที่ขนาด ปุ่มรูปภาพถูกสร้างขึ้นที่อนุญาตสำหรับการเปลี่ยนรูปทรงเรขาคณิตในพื้นที่โดยแสดงการไล่ระดับสีของภาพเบลอในระนาบการวางแนวหลายจุดและในระดับต่าง ๆ ศัพท์ถูกใช้เป็นอินพุตไปยังวิธีการจัดทำดัชนีเพื่อนบ้านที่ใกล้ที่สุดซึ่งระบุการจับคู่วัตถุ การตรวจสอบขั้นสุดท้ายของการแข่งขันแต่ละครั้งนั้นทำได้โดยการหาวิธีแก้ปัญหากำลังสองน้อยที่สุดที่ตกค้างต่ำสำหรับพารามิเตอร์รุ่นที่ไม่รู้จัก ผลการทดลองแสดงให้เห็นว่าการรับรู้วัตถุที่มีประสิทธิภาพสามารถทำได้ในรูปแบบที่กระจัดกระจายบางส่วนด้วยเวลาการคำนวณน้อยกว่า 2 วินาที

2.5 Flow Chart

Flow Chart ระบบการทำงาน Model

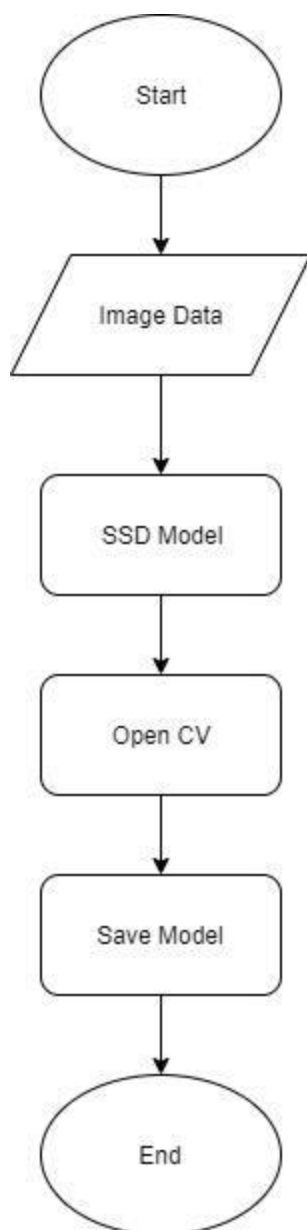


Image Data

เริ่มจากเก็บตัวอย่างรูปภาพนักศึกษาที่มีเครื่องแต่งกายนักศึกษาทั้งฤดูกระเป๊ยบและผิตระเป๊ยบ

SSD-Model

ต่อมาเราจะทำการ SSD ซึ่งถูกออกแบบมาสำหรับการตรวจจับวัตถุในแบบ real-time ,Faster R-CNN ใช้ region proposal network เพื่อสร้าง boundary boxes และใช้เพื่อจำแนกวัตถุ

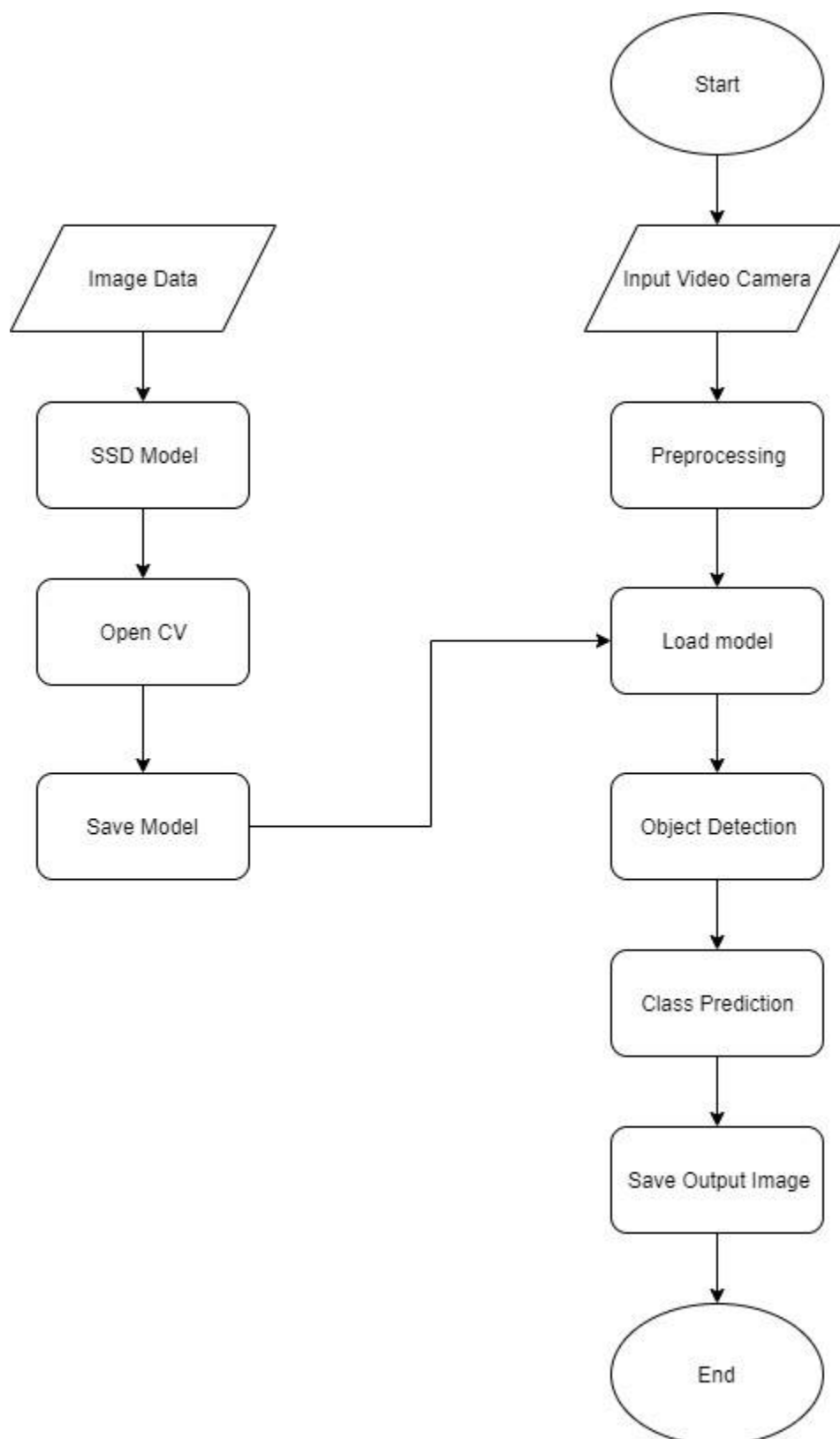
Open CV

ซึ่งทำงานเกี่ยวกับการประมวลผลภาพ การจับภาพ วีดีโอ สามารถวิเคราะห์คุณลักษณะต่าง ๆ ตัวอย่างเช่น ตรวจแสกนใบหน้า ตรวจจับวัตถุ

Save Model

หลังจากทำเสร็จ เราก็จะทำการ Save Model และเป็นอันที่สิ้นสุด

Flow Chart ระบบการทำงาน ตรวจสอบความเป็นนักศึกษา



บทที่ 3

วิธีการดำเนินงาน

วิธีการดำเนินงานครั้งนี้มีวัตถุประสงค์เพื่อนำเอาข้อมูลรูปภาพนักศึกษา มาทำการวิเคราะห์โดยใช้วิธีการ Machine Learning ด้วยโปรแกรม Python เพื่อทำการตรวจสอบว่านักศึกษามีการแต่งกายถูกระเบียบหรือไม่

3.1 วงจรการพัฒนากระบวนการ (System Development : SDLC)

วงจรการพัฒนากระบวนการ หมายถึง ขั้นตอนหรือกระบวนการที่ใช้ในการพัฒนากระบวนการ โดยมี จุดเริ่มต้นในการทำงานและจุดสิ้นสุดของการปฏิบัติงานและการพัฒนาซอฟต์แวร์ โดยจะประกอบไปด้วย กลุ่มกิจกรรม 3 ส่วนหลักๆ คือ การวิเคราะห์ (Analysis) การออกแบบ (Design) และการนำไปใช้ (Implementation) ทั้งสามนี้สามารถใช้งานได้ดีกับโครงการซอฟต์แวร์ที่มีขนาดเล็ก และในขณะที่โครงการ ซอฟต์แวร์ขนาดใหญ่ มักจะใช้เป็นแบบแผนการพัฒนาซอฟต์แวร์ตามแนวทางของ SDLC จนครบทุก กิจกรรม และภายในวงจรนี้จะแบ่งกระบวนการพัฒนาออกเป็นระยะ (Phases) ได้แก่ ระยะการวางแผน (Planning Phase) ระยะการวิเคราะห์ (Analysis Phase) ระยะการออกแบบ (Design Phase) และระยะการสร้างและพัฒนา (Implementation Phase) เราจึงจะต้องทำความเข้าใจในและกระบวนการทำงานของ แต่ละขั้นตอนจะต้องทำอะไรบ้าง โดยขั้นตอนการพัฒนากระบวนการมีอยู่ด้วยกัน 7 ขั้นตอน มีดังต่อไปนี้

3.1.1 เข้าใจปัญหา (Problem Recognition)

เป็นการทำความเข้าใจถึงปัญหาของการทำงานของระบบโดยตระหนักถึงปัญหาที่เกิดขึ้นว่าต้องการระบบสารสนเทศหรือการจัดการระบบเดิมและระบบต่าง ๆ มีความสามารถในการใช้งานมากแค่ไหน และมีประสิทธิภาพเพียงพอที่ตอบสนองความต้องการในปัจจุบันได้มากน้อยเพียงใด สำหรับระบบที่เราจะทำการพัฒนานั้นจะต้องมีความทันสมัย และสามารถเข้าถึงเทคโนโลยีที่ไม่ซับซ้อน ปัญหาที่สำคัญที่เราพบของระบบสารสนเทศในปัจจุบันคือ ระบบเขียนมานานแล้ว และส่วนใหญ่เขียนมาใช้งานซับซ้อน และนาเทคโนโลยีเข้ามาใช้ สิ่งที่เราควรที่จะทำการศึกษาระบบที่จะแก้ไขนั้นจากระบบเดิมที่มีอยู่แล้วมาพัฒนา หรือเริ่มการสร้างระบบใหม่ ดังนั้นเราควรมีการศึกษาความต้องการของผู้ใช้งานให้เพียงพอที่เป็นไปได้ หรือไม่ก็ได้แก่ การศึกษาความเป็นไปได้ (Feasibility Study)

3.1.2 ศึกษาความเป็นไปได้ (Feasibility Study)

เรามีจุดประสงค์ของการศึกษาถึงความเป็นไปได้ และกำหนดว่าปัญหาคืออะไรและทำการ ตัดสินใจว่าการพัฒนาหรือสร้างระบบสารสนเทศออกมาแบบไหน และทำการแก้ไขระบบสารสนเทศที่มีอยู่เดิมให้มีความเป็นไปได้หรือไม่มีการเสียค่าใช้จ่ายและใช้เวลาในการดำเนินงานให้น้อยที่สุดและได้ผลเป็นที่น่าพอใจเราทำการวิเคราะห์ระบบและกำหนดให้ได้ว่าเราจะทำการแก้ไขปัญหาดังกล่าว ให้อยู่ได้ 22 ในทางเทคนิคและบุคลากร โดยปัญหาทางเทคนิคมีส่วนเกี่ยวข้องกับเครื่องมือต่าง ๆ รวมทั้งเครื่องคอมพิวเตอร์ซอฟต์แวร์อีกด้วย สุดท้ายเราทำการวิเคราะห์ระบบให้เห็นถึงความเป็นไปได้อันจะนำมาใช้ในการพัฒนาระบบและที่สำคัญคือผลประโยชน์ที่จะได้รับ

3.1.3 วิเคราะห์ (Analysis)

เราได้เริ่มเข้าสู่การทำความเข้าใจและการวิเคราะห์ระบบโดยเริ่มจากการวิเคราะห์ระบบเริ่มต้นตั้งแต่ต้น การศึกษาระบบการทำงานของโครงการเราทำการออกแบบระบบใหม่โดยที่ทราบว่าจะระบบเดิมที่มีการทำงานอย่างไร หลังจากนั้นกำหนดความต้องการของระบบใหม่ ด้านการวิเคราะห์ระบบที่จะต้องใช้เทคนิคในการเก็บข้อมูลและศึกษาเอกสารที่มีอยู่ และตรวจสอบวิธีการทำงานของระบบในปัจจุบัน หา ข้อมูลจากผู้ใช้งานและผู้จัดทำที่มีส่วนเกี่ยวข้องกับระบบที่เราจะจัดทำให้ระบบใหม่มีการทำงานที่แตกต่างจาก ระบบเดิม

3.1.4 การออกแบบ (Design)

ในขั้นตอนของการเริ่มต้นในระยะแรกของการออกแบบ เราได้ทำการวิเคราะห์ระบบเราทำการตัดสินใจเลือกซื้อคอมพิวเตอร์ ฮาร์ดแวร์และซอฟต์แวร์และอุปกรณ์ต่าง ๆ และเราได้ทำการวิเคราะห์ระบบจะนำมาใช้ออกมาในรูปแบบแผนภาพต่าง ๆ ที่เขียนขึ้นในขั้นตอนการวิเคราะห์นำมาแปลงเป็นแผนภาพลำดับขั้น หรือแบบต้นไม้เพื่อให้มองเห็นภาพลักษณ์ที่แน่นอนของโปรแกรมว่ามีความสัมพันธ์กันอย่างไร และโปรแกรมอะไรบ้างที่จะต้องเขียนในระบบ จากนั้นก็เริ่มทำการตัดสินใจว่าควรจะทำโครงสร้าง ออกมาในรูปแบบที่เราได้กำหนด และทำการเชื่อมระหว่างโปรแกรมที่กำหนดในการออกแบบโปรแกรมเราคำนึงถึงความปลอดภัย (Security) ของระบบเพื่อป้องกันการผิดพลาดที่อาจจะเกิดขึ้น เช่น รหัสสำหรับผู้ใช้ที่มีสิทธิ์เข้าใช้งาน เป็นต้น

3.1.5 การพัฒนาระบบ (Construction)

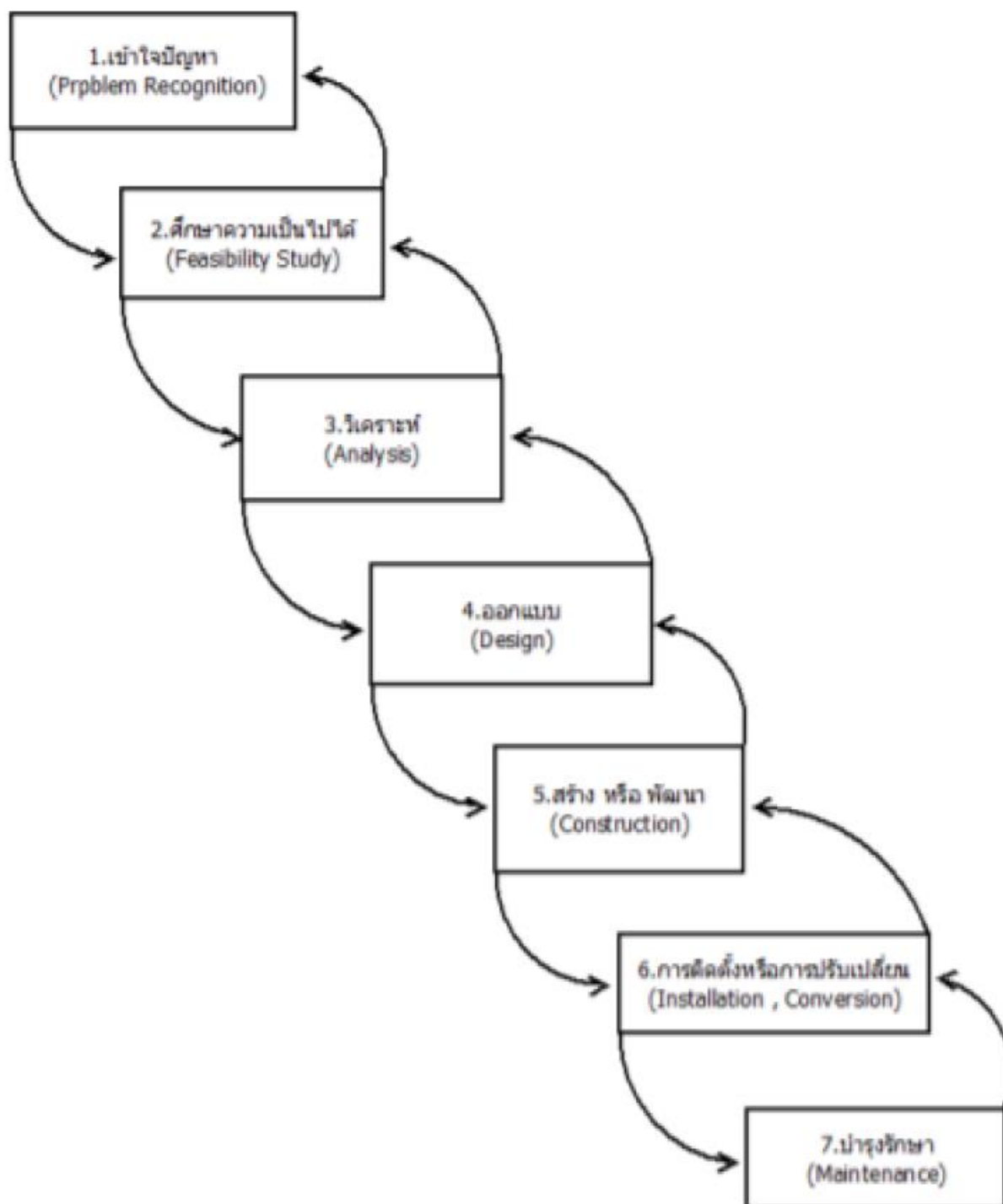
ในขั้นตอนนี้ นักพัฒนาระบบจะเริ่มเขียนและทำการทดสอบโปรแกรมว่ามีการทำงานที่ถูกต้องหรือไม่และมีการทดสอบกับข้อมูลจริงที่ที่นำมาใช้ และตรวจสอบความพร้อมที่จะนำไปใช้งานจริง จากนั้น ทำการจัดเตรียมคู่มือการใช้และการฝึกอบรมให้กับผู้ใช้งาน ในระยะแรกขั้นตอนนี้เราจะทำการวิเคราะห์ ระบบและจะต้องจัดเตรียมสถานที่และเครื่องคอมพิวเตอร์จะต้องตรวจสอบว่าคอมพิวเตอร์ทำงานเรียบร้อยหรือไม่ จากนั้นต้องควบคุมดูแลการเขียนคู่มือซึ่งจะประกอบด้วยข้อมูลการใช้งานสารบัญการอ้างอิงบน จอภาพ เป็นต้น นอกจากข้อมูลการใช้งานแล้วต้องมีการฝึกอบรมพนักงานที่จะเป็นผู้ใช้งานจริงของระบบ เพื่อให้เข้าใจและทำงานได้โดยไม่มีปัญหาอาจจะอบรมตัวต่อตัวหรือเป็นกลุ่มก็ได้ 23

3.1.6 การปรับเปลี่ยน (Construction)

ในขั้นตอนนี้ เรานำระบบใหม่มาใช้แทนของเก่าภายใต้การดูแลของนักวิเคราะห์ระบบและทำการป้อนข้อมูลที่ต้องทำให้เรียบร้อย และทำการเริ่มต้นใช้งานระบบใหม่ การนำระบบเข้ามาเราจะทำอย่างค่อย เป็นค่อยไป ที่ละน้อยในทางที่ดีที่สุดคือ ทำการใช้งานระบบใหม่ควบคู่ไปกับระบบเก่าไปสักระยะหนึ่ง โดย ใช้ข้อมูลชุดเดียวกัน แล้วเปรียบเทียบผลลัพธ์ว่าตรงกันหรือไม่ ถ้าเรียบร้อยแล้วเอาระบบเก่าออกได้ แล้วใช้ระบบใหม่ได้เลย

3.1.7 บำรุงรักษา (Maintenance)

การบำรุงรักษาได้แก่ การแก้ไขโปรแกรมหลังจากการใช้งานแล้ว สาเหตุที่จะต้องแก้ไขโปรแกรม หลังจากใช้งานแล้ว คือ มีปัญหาในโปรแกรม หรือการดำเนินงานการใช้งานระบบเปลี่ยนไป จากสถิติของระบบที่พัฒนาแล้วทั้งหมดประมาณ 40% ของค่าใช้จ่ายในการแก้ไขโปรแกรม เนื่องจากมี Bug จึงทำให้นักวิเคราะห์ระบบควรให้ความสำคัญกับการบำรุงรักษาถือว่าเป็นสิ่งสำคัญอย่างมาก การบำรุงรักษา ระบบควรจะอยู่ภายใต้การดูแลของนักวิเคราะห์ระบบ



รูปที่ 3.1 ภาพแสดงวงจรการพัฒนาระบบ (System Development : SDLC)

3.2 ตัวอย่าง วิธีทำการตรวจสอบความเป็นนักศึกษา



รูปที่ 3.2 ทำการเก็บรูปภาพนักศึกษาที่แต่งกายถูกระเบียบและไม่ถูกระเบียบ

หลังจากนั้นเราจะเข้าสู่กระบวนการเตรียมรูป (prepair_image.py)

Code

```
import cv2
import numpy as np
import os, glob
import uuid
from keras.models import load_model
import imutils

CLASSES = ['background', 'aeroplane', 'bicycle', 'bird', 'boat', 'bottle', 'bus',
'car', 'cat', 'chair', 'cow', 'diningtable', 'dog', 'horse', 'motorbike', 'perso
n', 'pottedplant', 'sheep', 'sofa', 'train', 'tvmonitor']
#CLASSES = ['motorbike', 'person']
COLORS = np.random.uniform(0, 255, size=(len(CLASSES), 3))
```

```

# load our serialized model from disk
print("[INFO] loading model...")
net = cv2.dnn.readNetFromCaffe('MobileNetSSD_deploy.prototxt.txt', 'MobileNetSSD_deploy.caffemodel')

loaded_model = load_model('new_helmet_model.h5')
loaded_model.compile(loss='binary_crossentropy', optimizer='rmsprop', metrics=['accuracy'])

def crop_image(filename):
    image = cv2.imread('./raw/' + filename)
    # resizing the images
    image = imutils.resize(image, width=600, height=600)

    # grab the frame dimensions and convert it to a blob
    (h, w) = image.shape[:2]

    # Resizing to a fixed 300x300 pixels and normalizing it.
    # Creating the blob from image to give input to the Caffe Model
    blob = cv2.dnn.blobFromImage(cv2.resize(image, (300, 300)), 0.007843, (300, 300), 127.5)

    # pass the blob through the network and obtain the detections and predictions
    net.setInput(blob)

    detections = net.forward() # getting the detections from the network

    persons = []
    person_roi = []

    # loop over the detections
    for i in np.arange(0, detections.shape[2]):
        # extract the confidence associated with the prediction
        confidence = detections[0, 0, i, 2]

        # filter out weak detections by ensuring the confidence
        # is greater than minimum confidence
        if confidence > 0.5:

            # extract index of class label from the detections
            idx = int(detections[0, 0, i, 1])

            if idx == 15:

```

```

        box = detections[0, 0, i, 3:7] * np.array([w, h, w, h])
        (startX, startY, endX, endY) = box.astype("int")
        persons.append((startX, startY, endX, endY))

xsdiff = 0
xediff = 0
ysdiff = 0
yediff = 0
p = ()

for i in persons:
    # cv2.rectangle(image, (i[0], i[1]), (i[2], i[3]), (0,0,255), 2)
    # print("{},{},{},{}".format(i[0],i[1],i[2],i[3]))
    # cv2.imshow('Frame', image) # Displaying the frame
    # cv2.waitKey(0)
    sub_image = image[i[1]:i[3], i[0]:i[2]]
    img = cv2.cvtColor(sub_image, cv2.COLOR_BGR2GRAY)
    cv2.imwrite('./data_crop/' + filename, img)

if __name__ == "__main__":
    try:
        os.mkdir('./data_crop')
    except:
        pass

    for file in os.listdir('./raw'):
        print(file)
        crop_image(file)

```

และเราก็จะได้ก็จะได้รูปภาพที่ Crop หรือตัดกรอบมาแล้ว



รูปที่ 3.3 รูปภาพที่ Crop แล้ว

และนำรูปภาพไปแยกใส่ Folder ที่มีชื่อว่า data โดยในโฟลเดอร์ดาด้า เราจะแยกออกเป็นอีก 2 โฟลเดอร์ย่อย ก็คือ Not_Student และ Student จะทำการแยกว่าใครแต่งตัวดูกระเปียบเป็นนักศึกษา หรือไม่ใช่นักศึกษาเสร็จแล้วก็เริ่มทำการ Train รออีก 5-10 นาที เราก็จะได้ Model ออกมา โดยใช้ควบคู่กับ (label.py) หน้าที่หลักของลาเบล คือ ใช้บ่งบอกลักษณะว่าเป็นนักศึกษา หรือไม่ใช่นักศึกษา

Code (label.py)

```
from __future__ import absolute_import
from __future__ import division
from __future__ import print_function

import argparse
import sys
import time

import numpy as np
import tensorflow as tf

def load_graph(model_file):
```

```

graph = tf.Graph()
graph_def = tf.GraphDef()

with open(model_file, "rb") as f:
    graph_def.ParseFromString(f.read())
with graph.as_default():
    tf.import_graph_def(graph_def)

return graph

def read_tensor_from_image_file(file_name, input_height=299, input_width=299,
                                input_mean=0, input_std=255):
    input_name = "file_reader"
    output_name = "normalized"
    file_reader = tf.read_file(file_name, input_name)
    if file_name.endswith(".png"):
        image_reader = tf.image.decode_png(file_reader, channels = 3,
                                            name='png_reader')
    elif file_name.endswith(".gif"):
        image_reader = tf.squeeze(tf.image.decode_gif(file_reader,
                                                       name='gif_reader'))
    elif file_name.endswith(".bmp"):
        image_reader = tf.image.decode_bmp(file_reader, name='bmp_reader')
    else:
        image_reader = tf.image.decode_jpeg(file_reader, channels = 3,
                                            name='jpeg_reader')
    float_caster = tf.cast(image_reader, tf.float32)
    dims_expander = tf.expand_dims(float_caster, 0)
    resized = tf.image.resize_bilinear(dims_expander, [input_height, input_width])
    normalized = tf.divide(tf.subtract(resized, [input_mean]), [input_std])
    sess = tf.Session()
    result = sess.run(normalized)

    return result

def load_labels(label_file):
    label = []
    proto_as_ascii_lines = tf.gfile.GFile(label_file).readlines()
    for l in proto_as_ascii_lines:
        label.append(l.rstrip())
    return label

def main(img):
    model_file = "retrained_graph.pb"
    label_file = "retrained_labels.txt"

```

```

input_height = 224
input_width = 224
input_mean = 128
input_std = 128
input_layer = "input"
output_layer = "final_result"
graph = load_graph(model_file)

file_name = img
t = read_tensor_from_image_file(file_name, input_height=input_height, input_width=input_width, input_mean=input_mean, input_std=input_std)

input_name = "import/" + input_layer
output_name = "import/" + output_layer
input_operation = graph.get_operation_by_name(input_name)
output_operation = graph.get_operation_by_name(output_name)

with tf.Session(graph=graph) as sess:
    start = time.time()
    results = sess.run(output_operation.outputs[0],
                       {input_operation.outputs[0]: t})
    end=time.time()
results = np.squeeze(results)

top_k = results.argsort()[-5:][::-1]
labels = load_labels(label_file)

data = {}
for i in range(len(labels)):
    data[labels[i]] = results[i]
return data

# for i in top_k:
#     return labels[i]

```

ต่อมาเริ่มทำการทดสอบกับกล้อง Camera Webcam C310 HD เพื่อหาค่าว่าอ่านได้หรืออ่านไม่ได้ ถ้าอ่านได้จะเป็น True ถ้าอ่านไม่ได้จะเป็น False แปลงรูปภาพให้เป็นขนาด 600 x 600 และจะทำการเลือกรูปภาพเฉพาะที่เป็นคนมาเท่านั้น (main.py)

Code

```
# import necessary packages
from imutils.video import VideoStream
import numpy as np
from imutils.video import FPS
import imutils
import time
import cv2
import os, glob
from keras.models import load_model
import operator
import label
import traceback

# initialize the list of class labels MobileNet SSD was trained to detect
# generate a set of bounding box colors for each class
CLASSES = ['background', 'aeroplane', 'bicycle', 'bird', 'boat', 'bottle', 'bus',
            'car', 'cat', 'chair', 'cow', 'diningtable', 'dog', 'horse', 'motorbike', 'person',
            'pottedplant', 'sheep', 'sofa', 'train', 'tvmonitor']
#CLASSES = ['motorbike', 'person']
COLORS = np.random.uniform(0, 255, size=(len(CLASSES), 3))

# load our serialized model from disk
print("[INFO] loading model...")
net = cv2.dnn.readNetFromCaffe('MobileNetSSD_deploy.prototxt.txt', 'MobileNetSSD_deploy.caffemodel')

loaded_model = load_model('new_helmet_model.h5')
loaded_model.compile(loss='binary_crossentropy', optimizer='rmsprop', metrics=['accuracy'])

# Loading the video file
cap = cv2.VideoCapture(0)

# time.sleep(2.0)

# Starting the FPS calculation
# fps = FPS().start()
```



```

        if idx == 15:
            box = detections[0, 0, i, 3:7] * np.array([w, h, w, h])
            (startX, startY, endX, endY) = box.astype("int")
            persons.append((startX, startY, endX, endY))

xsdiff = 0
xediff = 0
ysdiff = 0
yediff = 0
p = ()

for i in persons:
    sub_image = frame[i[1]:i[3], i[0]:i[2]]
    color = (0, 0, 0)
    if cv2.waitKey(33) == ord('a'):
        img = cv2.cvtColor(sub_image, cv2.COLOR_BGR2GRAY)
        cv2.imwrite('input.jpg', img)
        result = label.main('input.jpg')
        result = sorted(result.items(), key=operator.itemgetter(1))[-
1]

        if result[0] == 'student':
            color = (0, 0, 255)
            cv2.rectangle(frame, (i[0], i[1]), (i[2], i[3]), color, 2)
            cv2.putText(frame, result[0] + ' ' + str(result[1]), (i[0], i
[1]), cv2.FONT_HERSHEY_SIMPLEX, 0.5, color, 1, cv2.LINE_AA)
            cv2.imshow('Frame', frame)
            cv2.waitKey(0)
        else:
            cv2.rectangle(frame, (i[0], i[1]), (i[2], i[3]), color, 2)
    else:
        print('Cannot connect camera')
except Exception as e:
    traceback.print_exc()

if ret:
    cv2.imshow('Frame', frame) # Displaying the frame
    key = cv2.waitKey(1) & 0xFF
    if key == ord('q'): # if 'q' key is pressed, break from the loop
        break

cv2.destroyAllWindows()
cap.release() # Closing the video stream

```

และคราวนี้ทางคณะผู้จัดทำ ได้พัฒนาโปรแกรมมากขึ้นโดยที่เราได้เขียนออกแบบเว็บไซต์ไว้สำหรับการตรวจสอบแบบออนไลน์ ซึ่งสามารถทำการตรวจสอบบนเครื่องคอมพิวเตอร์ หรือ Notebook เครื่องใดก็ได้ ขอแค่เครื่องที่จะทำการตรวจสอบมีอินเทอร์เน็ต แค่นั้นผู้ใช้งานก็สามารถทำการตรวจสอบความเป็นนักศึกษาได้ (index.html),(web.py)

Code

```
<!DOCTYPE html>
<html>
<head>
<meta charset="utf-8">
<title>Bootstrap Example</title>
<meta charset="utf-8">
<meta name="viewport" content="width=device-width, initial-scale=1">
<link rel="stylesheet" href="https://maxcdn.bootstrapcdn.com/bootstrap/3.4.0/css/bootstrap.min.css">
<script src="https://ajax.googleapis.com/ajax/libs/jquery/3.4.1/jquery.min.js"></script>
<script src="https://maxcdn.bootstrapcdn.com/bootstrap/3.4.0/js/bootstrap.min.js"></script>

<style>
/* #container {
    margin: 0px auto;
    width: 500px;
    height: 375px;
    border: 10px #333 solid;
} */
#videoElement {
    width: 500px;
    height: 375px;
    background-color: #666;
}
</style>
</head>

<body>
<div class="row">
<div class="col-md-12 text-center"><h1>Detect Student</h1></div>
<div class="col-md-6 text-center">
<video autoplay="true" id="videoElement">
</video>
<canvas style="display:none;"></canvas>
```

```

    </div>
    <div class="col-md-6 text-center">
        
    </div>
    <div class="col-md-12 text-center">
        <button id="screenshot-button" type="button" class="btn btn-primary">
            Capture
        </button>
        <button id="clear-button" type="button" class="btn btn-primary">
            Clear
        </button>
    </div>
</div>
<script>
    var video = document.querySelector("#videoElement");
    var screenshotButton = document.querySelector('#screenshot-button');
    var clearButton = document.querySelector('#clear-button');
    const canvas = document.createElement('canvas');

    if (navigator.mediaDevices.getUserMedia) {
        navigator.mediaDevices.getUserMedia({ video: true })
            .then(function (stream) {
                video.srcObject = stream;
            })
            .catch(function (error) {
                console.log("Something went wrong!");
            });
    }

    clearButton.onclick = function(){
        document.getElementById('result_image').src='static/white.png?a=' + (new Date()).getTime();
    }

    screenshotButton.onclick = function() {
        canvas.width = video.videoWidth;
        canvas.height = video.videoHeight;
        canvas.getContext('2d').drawImage(video, 0, 0);
        var httpPost = new XMLHttpRequest(),

        path = "https://www.easycoding.co/authen",
        data = JSON.stringify({image: canvas.toDataURL('image/webp')});
        httpPost.onreadystatechange = function(err) {
            if (httpPost.readyState == 4 && httpPost.status == 200){

```



```

        document.getElementById('result_image').src='static/output.png?a=' + (new Date()).getTime();
        console.log(httpPost.responseText);
        if(httpPost.responseText == 'success'){
            window.location.href="/main"
        }
        else{
            // alert("Login Fail");
        }
    } else {
        console.log(err);
    }
};

// Set the content type of the request to json since that's what's being
sent
httpPost.open("POST", path, true);
httpPost.setRequestHeader('Content-Type', 'application/json');
httpPost.send(data);
// img.src = canvas.toDataURL('image/webp');
};

</script>
</body>
</html>

```

Code

```

from flask import Flask, render_template, request, session, redirect, url_for, es
cape, abort, jsonify
import base64
import os
import json
from main import Main
import cv2

app = Flask(__name__)

instance = Main()
net = cv2.dnn.readNetFromCaffe('MobileNetSSD_deploy.prototxt.txt', 'MobileNetSSD_
deploy.caffemodel')

@app.route('/')
def hello():

```

```

        return render_template("index.html")

@app.route('/authen', methods=['POST'])
def authen():
    global instance
    image = request.json.get('image', '').replace('data:image/webp;base64,', '')
    imgdata = base64.b64decode(image)
    filename = 'static/some_image.png'
    with open(filename, 'wb') as f:
        f.write(imgdata)
    res = instance.getData(net)
    return res

if __name__ == '__main__':
    instance.initial()
    app.run(debug=True, host='0.0.0.0:5000', port=80)

```

3.3 วิธีการใช้ Flask กับ Gunicorn และ Nginx บน Ubuntu 18.04

จะต้องสร้างแอปพลิเคชัน Python โดยใช้ Flask microframework บน Ubuntu 18.04 ส่วนใหญ่ของบทความนี้จะเกี่ยวกับวิธีการตั้งค่าเซิร์ฟเวอร์แอปพลิเคชัน Gunicorn และวิธีเปิดใช้แอปพลิเคชันและกำหนดค่า Nginx ให้ทำหน้าที่เป็น front-end reverse proxy

สิ่งที่ต้องมี

- เซิร์ฟเวอร์ที่ติดตั้ง Ubuntu 18.04 และ non-root user ด้วย sudo ทำตามคำแนะนำในการตั้งค่าเซิร์ฟเวอร์เริ่มต้น
- ติดตั้ง Nginx บน Ubuntu 18.04
- ชื่อโดเมนที่กำหนดค่าให้ไปที่เซิร์ฟเวอร์ของเรา คุณสามารถเรียนรู้วิธีการที่จะซื้อโดเมน DigitalOcean สร้าง DNS record ต่อไปนี้:
 - บันทึก A record your_domain ซึ่ไปยังที่อยู่ IP สาธารณะของเซิร์ฟเวอร์ของเรา
 - บันทึก A record ไปยังที่อยู่ IP สาธารณะของเซิร์ฟเวอร์ของเรา www.your_domain
- ทำความเข้าใจในข้อมูลจำเพาะของ WSGI ซึ่งเซิร์ฟเวอร์ Gunicorn จะใช้ในการสื่อสารกับแอปพลิเคชัน Flask ของเรา รายละเอียดWSGI ในเพิ่มเติม

ขั้นตอนที่ 1 - การติดตั้งส่วนประกอบจากที่เก็บ Ubuntu

ขั้นตอนแรกคือการติดตั้งชิ้นส่วนทั้งหมดที่เราต้องการจากที่เก็บ Ubuntu ซึ่งรวมถึง pip ตัวจัดการแพ็คเกจ Python ซึ่งจะจัดการส่วนประกอบ Python จะได้รับไฟล์ Python development ที่จำเป็นสำหรับการสร้างองค์ประกอบของ Gunicorn

อันดับแรกให้อัปเดต local package index และติดตั้งแพ็คเกจที่จะช่วยสร้าง Python environment รวมถึง python3-pip แพ็คเกจเพิ่มเติมและเครื่องมือการพัฒนาที่จำเป็นสำหรับการเขียนโปรแกรมที่มีประสิทธิภาพ

```
$ sudo apt update
$ sudo apt install python3-pip python3-dev build-essential libssl-dev libffi-dev python3-setuptools
```

ด้วยแพ็คเกจเหล่านี้จะมาสสร้าง virtual environment สำหรับ project

ขั้นตอนที่ 2 - การสร้าง Python Virtual Environment

ต่อไปจะตั้งค่า virtual environment เพื่อแยกแอปพลิเคชัน Flask ของเราออกจากไฟล์ Python อื่น ๆ ในระบบ เริ่มต้นด้วยการติดตั้ง python3-venv แพ็คเกจซึ่งจะติดตั้ง venv module:

```
$ sudo apt install python3-venv
```

ต่อไปมาสสร้าง parent directory สำหรับ Flask ย้ายไปยัง directory หลังจากสร้าง:

```
$ mkdir ~/myproject
$ cd ~/myproject
```

สร้าง virtual environment เพื่อเก็บข้อกำหนด Python ของ Flask โดยพิมพ์:

```
$ python3.6 -m venv myprojectenv
```

สิ่งนี้จะติดตั้ง local copy ของ Python และ pip ลงใน directory ที่เรียกว่า **myprojectenv** ภายใน directory ของ Project

ก่อนที่จะติดตั้งแอปพลิเคชันภายใน virtual environment ต้องเปิดใช้งาน โดยพิมพ์:

```
$ source myprojectenv/bin/activate
```

Prompt จะเปลี่ยนเพื่อระบุว่าขณะนี้ทำงานอยู่ใน virtual environment จริง ก็จะมีลักษณะบางอย่างเช่นนี้
 (myprojectenv)user@host:~/myproject\$

ขั้นตอนที่ 3 - การตั้งค่า Flask Application

ตอนนี้อยู่ใน virtual environment จริง สามารถติดตั้ง Flask และ Gunicorn และเริ่มต้นในการออกแบบแอปพลิเคชัน

อย่างแรกติดตั้ง wheel กับ local instance ของ pip เพื่อให้แน่ใจว่าแพ็คเกจจะติดตั้งแม้ว่าจะหายไปจาก wheel archives:

```
$ pip install wheel
```

ไม่ว่าจะใช้ Python เวอร์ชันใดเมื่อเปิดใช้งาน virtual environment ควรใช้ คำสั่ง pip (ไม่ใช่ pip3)

ต่อมาติดตั้ง Flask และ Gunicorn :

```
(myprojectenv) $ pip install gunicorn flask
```

การสร้างแอปตัวอย่าง

ตอนนี้มี Flask ให้ใช้งานแล้วสามารถสร้างแอปพลิเคชันที่เรียบง่าย Flask เป็น microframework. ไม่รวมเครื่องมือจำนวนมากที่อาจมีเฟรมเวิร์กที่มีคุณสมบัติครบถ้วนมากกว่าและส่วนใหญ่โมดูลที่สามารถนำเข้าสู่โปรเจ็คเพื่อช่วยในการเริ่มต้นโปรแกรมประยุกต์บนเว็บ

ในขณะที่แอปพลิเคชันอาจซับซ้อนกว่านี้ ดังนั้นจะสร้างแอป Flask ในไฟล์เดียวชื่อ **myproject.py**:

```
(myprojectenv) $ nano ~/myproject/myproject.py
```

application code จะอยู่ในไฟล์นี้ จะนำเข้า Flask และยกตัวอย่าง Flask object สามารถใช้ฟังก์ชันนี้เพื่อ

```
~ / myproject / myproject.py
```

```
from flask import Flask
```

```
app = Flask(__name__)
```

```
@app.route("/")
```

```
def hello():
```

```
    return "<h1 style='color:blue'>Hello There!</h1>"
```

```
if __name__ == "__main__":
    app.run(host='0.0.0.0')
```

สิ่งนี้จะกำหนดเนื้อหาที่จะนำเสนอเมื่อเข้าถึง root domain บันทึกและปิดไฟล์เมื่อดำเนินการเสร็จ

หากทำตามคำแนะนำในการตั้งค่าเซิร์ฟเวอร์เริ่มต้นควรเปิดใช้งานไฟร์วอลล์ UFW ในการทดสอบแอปพลิเคชันต้องอนุญาตการเข้าถึงพอร์ต 5000:

```
(myprojectenv) $ sudo ufw allow 5000
```

ตอนนี้คุณสามารถทดสอบแอป Flask โดยพิมพ์:

```
(myprojectenv) $ python myproject.py
```

จะเห็น output ดังต่อไปนี้รวมถึง helpful warning reminding เพื่อเตือนให้คุณไม่ใช้การตั้งค่าเซิร์ฟเวอร์นี้ :

Output

```
* Serving Flask app "myproject" (lazy loading)
* Environment: production
  WARNING: Do not use the development server in a production environment.
  Use a production WSGI server instead.
* Debug mode: off
* Running on http://0.0.0.0:5000/ (Press CTRL+C to quit)
```

ที่อยู่ IP ของเซิร์ฟเวอร์ ตามด้วย:5000 ของเว็บเบราว์เซอร์:

```
http://your_server_ip:5000
```

จะเห็นสิ่งนี้:

Hello There!

เมื่อเสร็จแล้วให้กด CTRL-C terminal window เพื่อหยุดเซิร์ฟเวอร์ Flask development

การสร้าง WSGI Entry Point

ต่อไปทำการสร้างไฟล์ที่จะทำหน้าที่เป็นจุดเริ่มต้นสำหรับแอปพลิเคชันของเรา จะอธิบายเวิร์กเกอร์ Gunicorn ว่า จะโต้ตอบกับแอปพลิเคชันอย่างไร

เรียกไฟล์ `wsgi.py`:

```
(myprojectenv) $ nano ~/myproject/wsgi.py
```

ไฟล์นี้ให้นำเข้า Flask instance จากแอปพลิเคชันและเรียกใช้:

```
~ / myproject / wsgi.py

from myproject import app

if __name__ == "__main__":
    app.run()
```

บันทึกและปิดไฟล์เมื่อดำเนินการเสร็จ

ขั้นตอนที่ 4 - การกำหนดค่า Gunicorn

แอปพลิเคชันของเราจะถูกเขียนขึ้นโดยมีจุดเริ่มต้นที่ที่ทำไว้ก่อนหน้านี้แล้ว สามารถกำหนดค่า Gunicorn ก่อนดำเนินการต่อควรตรวจสอบว่า Gunicorn สามารถให้บริการแอปพลิเคชันได้อย่างถูกต้องหรือไม่

สามารถทำได้โดยเพียงแค่ส่งชื่อ entry point ของเรา จะถูกสร้างขึ้นเป็นชื่อของโมดูล (minus.py extension) บวกกับชื่อของ callable ภายในแอปพลิเคชัน ในกรณีนี้เป็น `wsgi:app`

นอกจากนี้เราจะระบุอินเทอร์เฟซและพอร์ตที่จะผูกไว้เพื่อให้แอปพลิเคชันเริ่มต้นบนอินเทอร์เฟซที่มีอยู่ทั่วไป:

```
(myprojectenv) $ cd ~/myproject
(myprojectenv) $ gunicorn --bind 0.0.0.0:5000 wsgi:app
```

จะเห็นผลลัพธ์ดังนี้:

```
Output
[2018-07-13 19:35:13 +0000] [28217] [INFO] Starting gunicorn 19.9.0
[2018-07-13 19:35:13 +0000] [28217] [INFO] Listening at: http://0.0.0.0:5000 (28217)
[2018-07-13 19:35:13 +0000] [28217] [INFO] Using worker: sync
```

```
[2018-07-13 19:35:13 +0000] [28220] [INFO] Booting worker with pid: 28220
```

Visit ที่อยู่ IP ของเซิร์ฟเวอร์:5000 ต่อท้ายเว็บเบราว์เซอร์อีกครั้ง:

```
http://your_server_ip:5000
```

จะเห็นผลลัพธ์ของแอปพลิเคชัน:

Hello There!

กด CTRL-C เพื่อยืนยันว่าทำงานอย่างถูกต้อง ในหน้าต่าง terminal

ตอนนี้เสร็จแล้วด้วย virtual environment จริงดังนั้นเราจึงสามารถปิดการใช้งานได้:

```
(myprojectenv) $ deactivate
```

คำสั่ง Python ใด ๆ จะใช้ Python environment ของระบบอีกครั้ง

ต่อไปทำการสร้างไฟล์ systemd service unit การสร้างไฟล์ systemd unit จะทำให้ระบบเริ่มต้นของ Ubuntu ทำให้ Gunicorn เริ่มการทำงาน และให้บริการแอปพลิเคชัน Flask โดยอัตโนมัติเมื่อใดก็ตามที่ server boots

สร้างไฟล์หน่วยที่ลงท้ายด้วย .service ภายใน /etc/systemd/systemdirectory เพื่อเริ่ม:

```
$ sudo nano /etc/systemd/system/myproject.service
```

Inside จะเริ่มต้นด้วย [Unit] ส่วนซึ่งใช้เพื่อระบุ metadata และ dependencies ใส่คำอธิบาย sever และบอกให้ init system เริ่มต้นเมื่อเป้าหมายเครือข่ายถึงแล้วเท่านั้น:

```
/etc/systemd/system/myproject.service
```

```
[Unit]
```

```
Description=Gunicorn instance to serve myproject
```

```
After=network.target
```

ต่อไปเปิด [Service] สิ่งนี้จะระบุผู้ใช้และกลุ่มที่ต้องการให้กระบวนการทำงาน มาเป็นเจ้าของบัญชีผู้ใช้ตามปกติ เพราะเป็นเจ้าของไฟล์ที่เกี่ยวข้องทั้งหมด มอบความเป็นเจ้าของกลุ่มให้กับ www-data เพื่อให้ Nginx สามารถสื่อสารกับกระบวนการ Gunicorn ได้ง่าย เปลี่ยนชื่อผู้ใช้ที่เป็นของตัวเอง:

```
/etc/systemd/system/myproject.service
```

```
[Unit]
```

```
Description=Gunicorn instance to serve myproject
```

```
After=network.target
```

```
[Service]
```

```
User=sammy
```

```
Group=www-data
```

ขั้นต่อไปให้ทำแผนที่ directory การทำงานและตั้งค่า PATH environmental variable เพื่อให้ระบบ init ทราบว่าไฟล์ประมวลผลสำหรับกระบวนการนั้นอยู่ใน virtual environmentจริงของเรา ลองระบุคำสั่งเพื่อเริ่มบริการ คำสั่งนี้จะทำสิ่งต่อไปนี้:

- เริ่ม 3 กระบวนการผู้ปฏิบัติงาน (ควรปรับเปลี่ยนตามความจำเป็น)
- สร้างและผูกเข้ากับ Unix socket file **myproject.sock**, ใน directory งานของเรา จะตั้งค่า an umask value 007 เพื่อให้ socket file ถูกสร้างขึ้นให้เข้าถึงเจ้าของและกลุ่มในขณะที่ จำกัดการเข้าถึงอื่น ๆ
- ระบุชื่อ entry point file ใช้งาน WSGI พร้อมกับ Python callable ภายในไฟล์นั้น (wsgi:app)

Systemd ต้องการให้เส้นทางแบบเต็มชื่อของ Gunicorn ซึ่งติดตั้งใน virtual environmentจริง

แทนที่ชื่อผู้ใช้และProject paths ด้วยข้อมูลของเราเอง:

```
/etc/systemd/system/myproject.service
```

```
[Unit]
```

```
Description=Gunicorn instance to serve myproject
```

```
After=network.target
```

```
[Service]
```

```
User=sammy
```

```
Group=www-data
```

```
WorkingDirectory=/home/sammy/myproject
```

```
Environment="PATH=/home/sammy/myproject/myprojectenv/bin"
```

```
ExecStart=/home/sammy/myproject/myprojectenv/bin/gunicorn --workers 3 --bind
```

```
unix:myproject.sock -m 007 wsgi:app
```

สุดท้ายเรามาเพิ่ม [Install] section สิ่งนี้จะทำให้ทราบว่า systemd จะเชื่อมโยงบริการนี้กับอะไรหากเปิดใช้งานเพื่อเริ่มการทำงาน ต้องให้บริการนี้เริ่มต้นขึ้นเมื่อ regular multi-user system และ up and running

```
/etc/systemd/system/myproject.service
```

```
[Unit]
```

```
Description=Gunicorn instance to serve myproject
```

```
After=network.target
```

```
[Service]
```

```
User=sammy
```

```
Group=www-data
```

```
WorkingDirectory=/home/sammy/myproject
```

```
Environment="PATH=/home/sammy/myproject/myprojectenv/bin"
```

```
ExecStart=/home/sammy/myproject/myprojectenv/bin/gunicorn --workers 3 --bind
```

```
unix:myproject.sock -m 007 wsgi:app
```

```
[Install]
```

```
WantedBy=multi-user.target
```

ไฟล์บริการ systemd เสร็จสมบูรณ์ บันทึกและปิดทันที
สามารถเริ่มบริการ Gunicorn ที่สร้างขึ้นและเปิดใช้งานเพื่อให้ starts at boot

```
$ sudo systemctl start myproject
$ sudo systemctl enable myproject
```

ตรวจสอบสถานะ:

```
$ sudo systemctl status myproject
```

ผลลัพธ์เช่นนี้

Output

```
● myproject.service - Gunicorn instance to serve myproject
   Loaded: loaded (/etc/systemd/system/myproject.service; enabled; vendor preset:
   enabled)
   Active: active (running) since Fri 2018-07-13 14:28:39 UTC; 46s ago
 Main PID: 28232 (gunicorn)
    Tasks: 4 (limit: 1153)
   CGroup: /system.slice/myproject.service
           └─28232 /home/sammy/myproject/myprojectenv/bin/python3.6
             /home/sammy/myproject/myprojectenv/bin/gunicorn --workers 3 --bind
             unix:myproject.sock -m 007
           └─28250 /home/sammy/myproject/myprojectenv/bin/python3.6
             /home/sammy/myproject/myprojectenv/bin/gunicorn --workers 3 --bind
             unix:myproject.sock -m 007
           └─28251 /home/sammy/myproject/myprojectenv/bin/python3.6
             /home/sammy/myproject/myprojectenv/bin/gunicorn --workers 3 --bind
             unix:myproject.sock -m 007
```

```

└─28252 /home/sammy/myproject/myprojectenv/bin/python3.6
/home/sammy/myproject/myprojectenv/bin/gunicorn --workers 3 --bind
unix:myproject.sock -m 007

```

หากเห็นข้อผิดพลาดใด ๆ แก้ไขก่อนที่จะดำเนินงานต่อ

ขั้นตอนที่ 5 - การกำหนดค่า Nginx ให้เป็นคำขอพร็อกซี

เซิร์ฟเวอร์แอปพลิเคชัน Gunicorn ตอนนี้ ควรเปิดใช้งานแล้วรอคำขอใน socket file ใน project directory กำหนดค่า Nginx เพื่อส่งคำร้องขอทางเว็บไปยัง socket โดยทำการเพิ่มเติมเล็กน้อยไปยัง configuration file

เริ่มด้วยการสร้างไฟล์การกำหนดค่า server block ใหม่ใน sites-available directory ของ Nginx เรียกว่า myproject เพื่อให้สอดคล้องกัน:

```
$ sudo nano /etc/nginx/sites-available/myproject
```

เปิด server block และบอก Nginx เพื่อฟัง default port 80 ใช้บล็อกนี้เพื่อขอชื่อ server's domain :

```

/etc/nginx/sites-available/myproject

server {
    listen 80;
    server_name your_domain www.your_domain;
}

```

ต่อไปเราจะเพิ่มบล็อกที่ตั้งที่ตรงกับทุกคำขอ ภายในบล็อกนี้เราจะรวม proxy_params ไฟล์ที่ระบุ proxying parameters ทั่วไปบางอย่างที่จำเป็นต้องตั้งค่า จากนั้นจะส่งคำขอไปยัง socket ที่กำหนดโดยใช้ คำสั่ง proxy_pass

```
/etc/nginx/sites-available/myproject

server {
    listen 80;

    server_name your_domain www.your_domain;

    location / {
        include proxy_params;
        proxy_pass http://unix:/home/sammy/myproject/myproject.sock;
    }
}
```

บันทึกและปิดไฟล์เมื่อดำเนินการเสร็จ

หากต้องการเปิดใช้งานการกำหนดค่าบล็อกเซิร์ฟเวอร์ Nginx ที่เพิ่งสร้างขึ้นให้เชื่อมโยงไฟล์ไปยัง sites-enabled directory:

```
$ sudo ln -s /etc/nginx/sites-available/myproject /etc/nginx/sites-enabled
```

ด้วยไฟล์ใน directory นั้นสามารถทดสอบหาข้อผิดพลาดทาง syntax errors:

```
$ sudo nginx -t
```

หากไม่มีปัญหาใด ๆ ให้รีสตาร์ทกระบวนการ Nginx เพื่ออ่านการกำหนดค่าใหม่:

```
$ sudo systemctl restart nginx
```

สุดท้าย ปรับ firewall อีกครั้ง ไม่ต้องการเข้าถึงผ่านพอร์ต 5000 อีกต่อไป ดังนั้นจึงสามารถลบกฎนั้นได้ จากนั้นสามารถอนุญาตให้เข้าถึงเซิร์ฟเวอร์ Nginx ได้อย่างสมบูรณ์:

```
$ sudo ufw delete allow 5000
```

```
$ sudo ufw allow 'Nginx Full'
```

ตอนนี้ควรจะสามารภ navigate ไปยังชื่อโดเมนของเซิร์ฟเวอร์ในเว็บเบราว์เซอร์ของเรา:

```
http://your_domain
```

ผลลัพธ์ของแอปพลิเคชัน:

Hello There!

หากพบข้อผิดพลาดลองตรวจสอบสิ่งต่อไปนี้:

- `sudo less /var/log/nginx/error.log`: ตรวจสอบบันทึกข้อผิดพลาด Nginx
- `sudo less /var/log/nginx/access.log`: ตรวจสอบบันทึกการเข้าถึง Nginx
- `sudo journalctl -u nginx`: ตรวจสอบบันทึกกระบวนการ Nginx
- `sudo journalctl -u myproject`: ตรวจสอบบันทึก Unicorn ของแอป Flask

ขั้นตอนที่ 6 - Securing the Application

เพื่อให้แน่ใจว่าเราพิกไปยังเซิร์ฟเวอร์ของเรายังคงปลอดภัยขอใบรับรอง SSL สำหรับโดเมนของเรามีหลายวิธีที่จะทำ นี้นรวมทั้งได้รับใบรับรองการเป็นอิสระจากการเป็น Let's Encrypt, การสร้างใบรับรองลงนามด้วยตัวเองหรือซื้อจากผู้ให้บริการอื่นและการกำหนดค่า Nginx ที่จะใช้โดยขั้นตอนที่ 2 ต่อไปทั้งหมด 6 ของ วิธีการสร้างตัวเองลงนาม ใบรับรอง SSL สำหรับ Nginx ใน Ubuntu 18.04

เพิ่มที่เก็บ Certbot Ubuntu:

```
$ sudo add-apt-repository ppa:certbot/certbot
```

กด ENTER เพื่อยอมรับ

ติดตั้ง Certbot's Nginx package ด้วย apt:

```
$ sudo apt install python-certbot-nginx
```

Certbot มีวิธีการมากมายในการขอใบรับรอง SSL ผ่านทางปลั๊กอิน ปลั๊กอิน Nginx จะดูแลการกำหนดค่า Nginx อีกครั้งและโหลดการกำหนดค่า เมื่อใดก็ตามที่จำเป็น ในการใช้ปลั๊กอินนี้ให้พิมพ์ดังต่อไปนี้:

```
$ sudo certbot --nginx -d your_domain -d www.your_domain
```

สิ่งนี้จะทำงาน certbot กับ—nginx ปลั๊กอินโดยใช้-d เพื่อระบุชื่อที่ต้องการให้ใบรับรองถูกต้อง

ถ้าเป็นครั้งแรกที่ใช้งาน certbot จะได้รับแจ้งให้ป้อนที่อยู่อีเมลและยอมรับข้อกำหนดในการให้บริการ หลังจากนั้น certbot จะสื่อสารกับเซิร์ฟเวอร์ Let's Encrypt จากนั้นเรียกใช้ challenge เพื่อยืนยันว่าเป็นผู้ควบคุมโดเมนที่กำลังขอใบรับรอง

หากสำเร็จ certbot จะถามว่าต้องการกำหนดการตั้งค่า HTTPS อย่างไร:

Output

Please choose whether or not to redirect HTTP traffic to HTTPS, removing HTTP access.

-
- 1: No redirect - Make no further changes to the webserver configuration.
 - 2: Redirect - Make all requests redirect to secure HTTPS access. Choose this for new sites, or if you're confident your site works on HTTPS. You can undo this change by editing your web server's configuration.
-

Select the appropriate number [1-2] then [enter] (press 'c' to cancel):

เลือกที่ตัวเลือกแล้วกด ENTER การกำหนดค่าจะได้รับการอัปเดตและ Nginx จะโหลดซ้ำเพื่อรับการตั้งค่าใหม่ certbot จะสรุปข้อความที่บอกว่ากระบวนการนี้สำเร็จและที่เก็บใบรับรอง:

Output

IMPORTANT NOTES:

- Congratulations! Your certificate and chain have been saved at:
/etc/letsencrypt/live/your_domain/fullchain.pem
Your key file has been saved at:
/etc/letsencrypt/live/your_domain/privkey.pem
Your cert will expire on 2018-07-23. To obtain a new or tweaked version of this certificate in the future, simply run certbot again with the "certonly" option. To non-interactively renew **all** of your certificates, run "certbot renew"
- Your account credentials have been saved in your Certbot configuration directory at /etc/letsencrypt. You should make a secure backup of this folder now. This configuration directory will also contain certificates and private keys obtained by Certbot so making regular backups of this folder is ideal.
- If you like Certbot, please consider supporting our work by:

Donating to ISRG / Let's Encrypt: <https://letsencrypt.org/donate>

Donating to EFF: <https://eff.org/donate-le>

การติดตั้ง Nginx ไม่ต้องใช้ HTTP profile allowance ที่ซ้ำซ้อนอีกต่อไป:

```
$ sudo ufw delete allow 'Nginx HTTP'
```

หากต้องการตรวจสอบการกำหนดค่าให้นำทางอีกครั้งไปยังโดเมนโดยใช้https://:

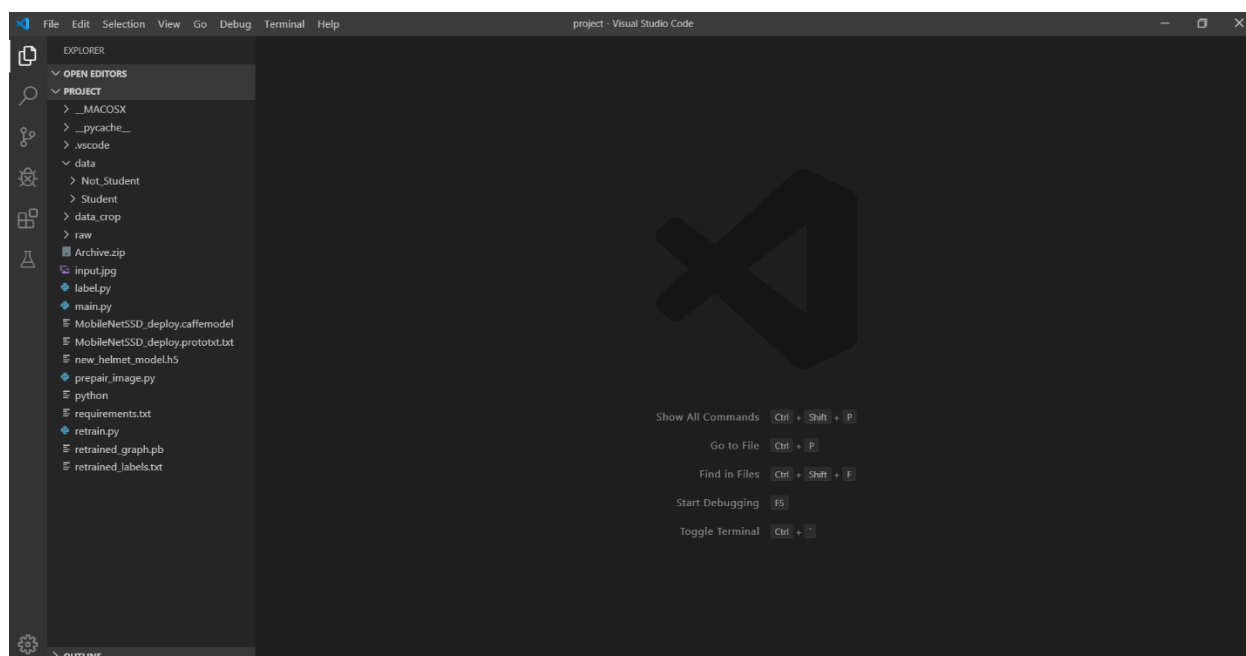
```
https://your_domain
```

บทที่ 4

ผลการดำเนินงาน

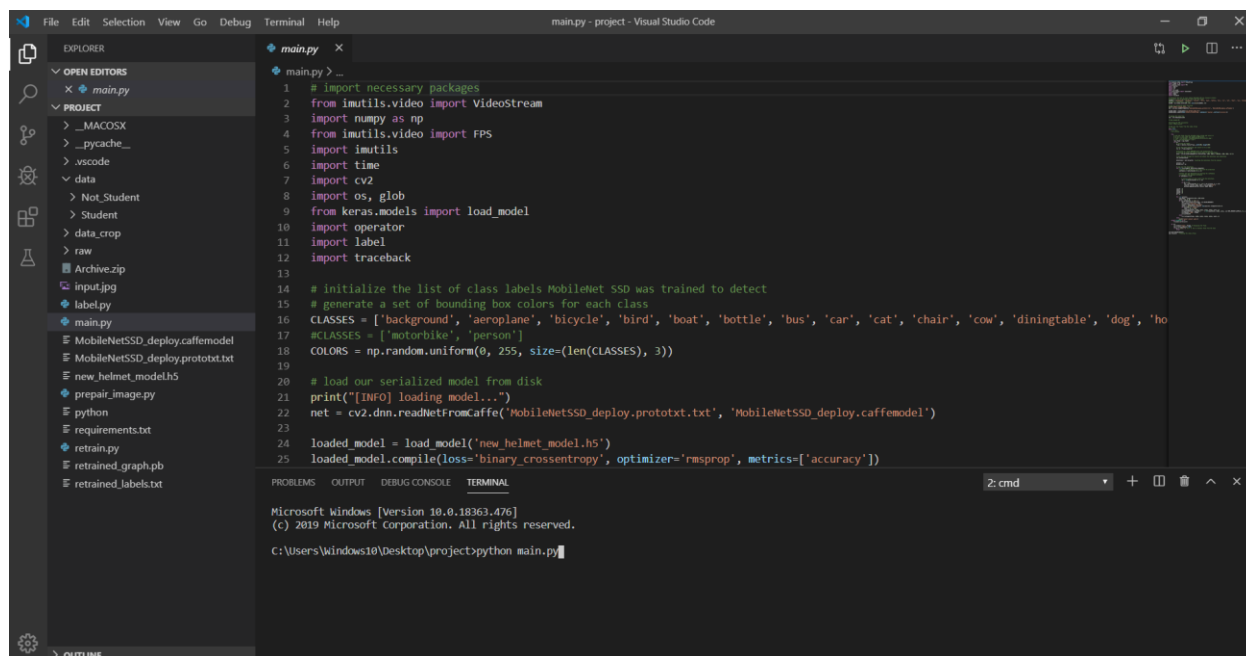
4.1 การใช้งานระบบตรวจสอบความเป็นนักศึกษา

4.1.1 เปิดโปรแกรม Visual Studio Code



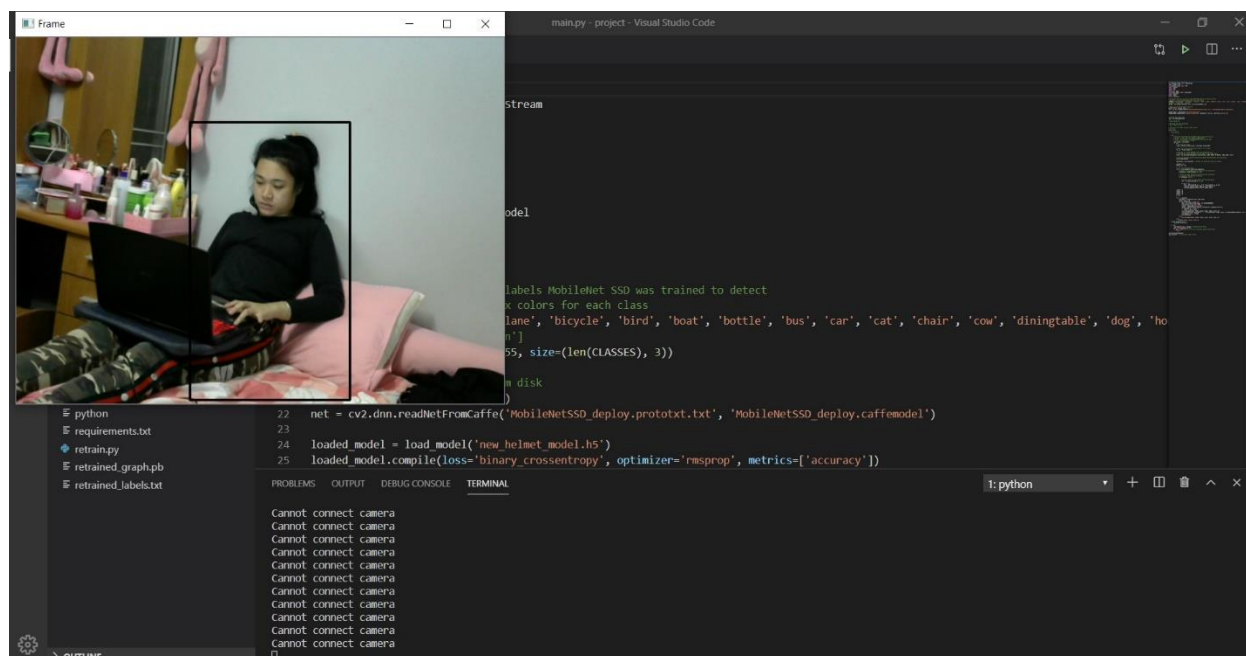
รูปที่ 4.1 แสดงหน้าโปรแกรม Visual Studio Code

4.1.2 ไปที่ main.py เพื่อจะทำการ Run และคำสั่งที่ใช้ในการรันนั้นคือ python main.py



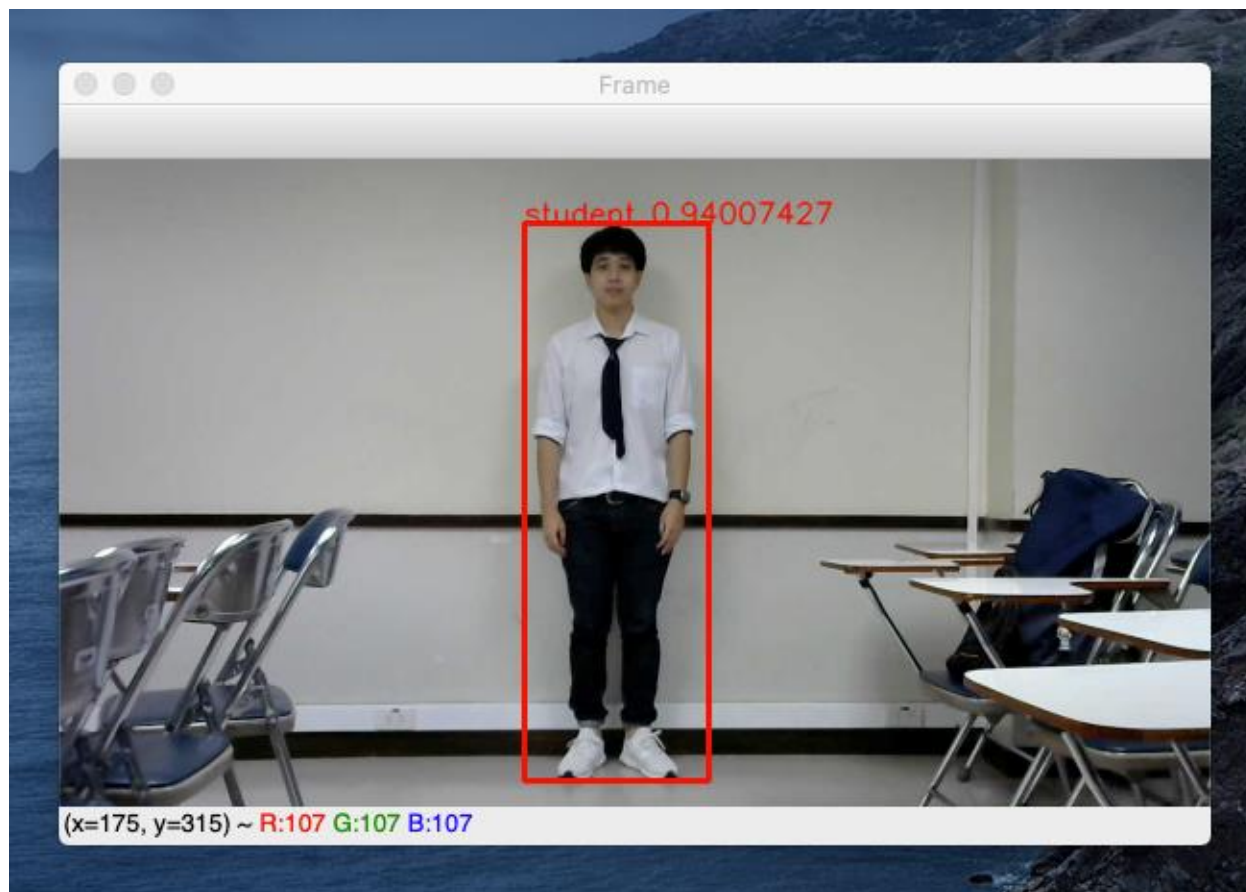
รูปที่ 4.2 แสดงถึงหน้ากำลังจะทำการ Run

4.1.3 เมื่อทำการป้อนคำสั่ง Run เสร็จเรียบร้อยแล้ว จะมีหน้าจอโชว์ขึ้นมาทำการจับบุคคลที่อยู่หน้ากล้อง โดยระบบจะมีการติกรอบรอบตัวของคนคนนั้นเป็นสีดำ



รูปที่ 4.3 ระบบกำลังแสดงการตรวจสอบความเป็นนักศึกษา

4.1.4 และเมื่อเรากดปุ่ม A ระบบจะทำการ Crop ภาพนั้นและประมวลผลว่า บุคคลที่อยู่หน้ากล้องนั้น ใช่นักศึกษาหรือไม่ (ถ้าใช่) กรอบจะเปลี่ยนเป็นสีแดงและขึ้นคำว่า Student แต่ถ้า (ไม่ใช่) กรอบจะเป็นสีดำเหมือนเดิม แต่จะขึ้นคำว่า Not Student ส่วนตัวเลขที่ขึ้นอยู่ด้านข้างประโยคคำที่บอกว่าใช่นักศึกษาหรือไม่นั้น เป็นตัวเลขที่บอกเกี่ยวกับเปอร์เซ็นต์ว่าจาก 100% นั้น รูปมีความเหมือนหรือไม่เหมือนนักศึกษาที่เปอร์เซ็นต์

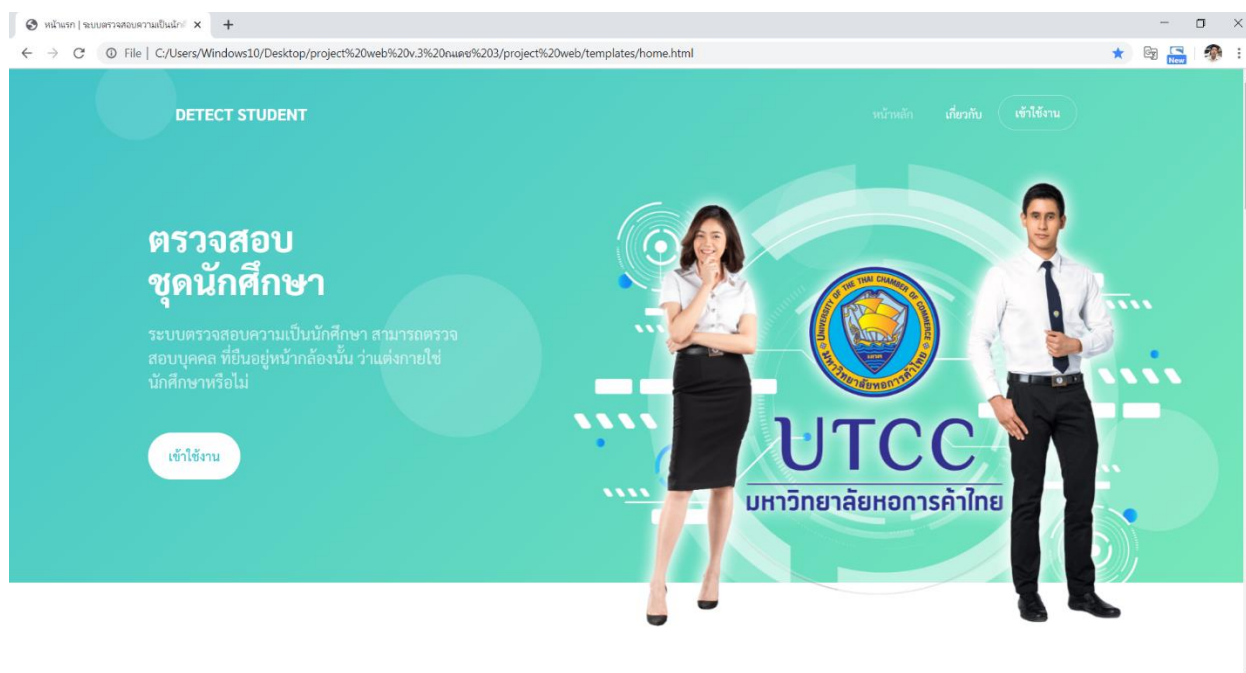


รูปที่ 4.4 ภาพที่แสดงว่าเป็นนักศึกษา

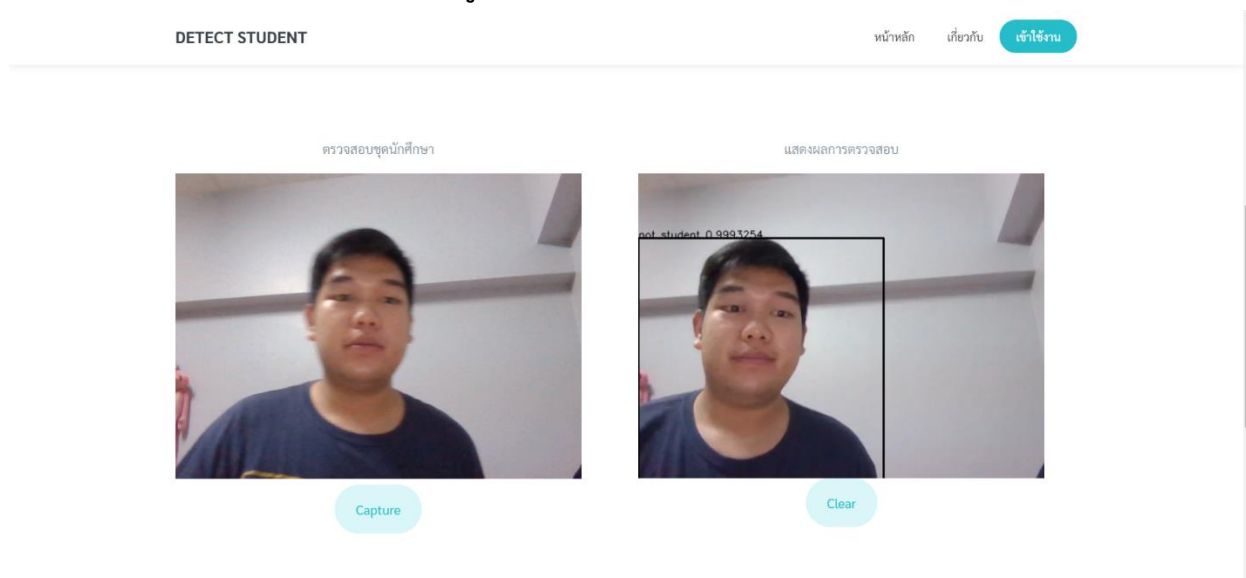


รูปที่ 4.5 ภาพที่แสดงว่าไม่ใช่นักศึกษา

4.1.5 ส่วนคราวนี้ทางเรา “คณะผู้จัดทำ” ได้ทำการออกแบบเว็บไซต์ในการตรวจสอบความเป็นนักศึกษาแบบออนไลน์ สามารถตรวจสอบบนเครื่องคอมพิวเตอร์ หรือ Notebook เครื่องใดก็ได้



รูปที่ 4.6 หน้าหลักของเว็บไซต์



รูปที่ 4.7 ระบบตรวจสอบความเป็นนักศึกษานเว็บไซต์

บทที่ 5

สรุปผลการดำเนินงาน

ในส่วนนี้จะกล่าวถึงการสรุปผลของระบบตรวจสอบความเป็นนักศึกษา โดยการประมวลผลภาพตลอดจน ปัญหาและอุปสรรคของการจัดทำโครงงานนี้รวมทั้งข้อเสนอแนะในการนำโครงงานไปพัฒนาเพื่อเป็นแนวทาง สำหรับผู้ที่สนใจ

5.1 สรุปผลการทำโครงการ

ระบบตรวจสอบความเป็นนักศึกษา ใช้ในการตรวจสอบเครื่องแต่งกายของนักศึกษาว่าถูกระเบียบหรือไม่ สามารถ แยกแยะระหว่างเป็นนักศึกษา หรือ ไม่ได้เป็นนักศึกษา โดยทำการเก็บรวบรวมข้อมูลรูปภาพนักศึกษาที่ถูกระเบียบ และไม่ถูกระเบียบเพื่อเป็นข้อมูลสำหรับกระบวนการ Machine Learning เพื่อทำให้เครื่องคอมพิวเตอร์เกิดการเรียนรู้ ซึ่งไฟล์ข้อมูลทั้งหมดได้จากการบันทึกจากกล้อง Camera Webcam C310 HD โดยถ่ายรูปลงนักศึกษาเต็มตัว ด้านหลังและด้านข้าง เพื่อความแม่นยำที่ดีขึ้น โดยใช้ ทฤษฎีอื่น ๆ หรืออัลกอริทึม มาช่วยในการประมวลผลภาพ ตัวอย่างเช่น Object Recognition Open CV และ Model SSD เป็นต้น เพื่อต้องการผลลัพธ์ว่านักศึกษาแต่งกาย ถูกระเบียบหรือไม่ ผลที่ได้จากการทดลองแสดงให้เห็นว่าการตรวจสอบความถูกต้องของเครื่องแต่งกายนักศึกษามี ความถูกต้องเฉลี่ย 91.98% โดยแบ่งออกเป็นคนถูกระเบียบ 22 คน ผิดระเบียบ 8 คน จากกลุ่มตัวอย่าง 30 คน

5.2 ปัญหาที่พบในการดำเนินการ

5.2.1 การตรวจสอบความถูกต้องอาจเกิดผิดพลาดด้วยสถานการณ์ต่าง ๆ เช่น แสงสว่าง ความคมชัด เครื่องแต่งกายที่คล้ายนักศึกษา เป็นต้น

5.2.2 ข้อมูลที่นำมาทำการฝึกฝนมีน้อยเกินไป

5.2.3 เวลาในการทำโครงการมีน้อย อาจทำให้ความรู้ในด้านทฤษฎียังมีไม่มากพอ

5.3 ข้อเสนอแนะ

การพัฒนา ระบบตรวจสอบความเป็นนักศึกษาจะต้องมีประสิทธิภาพและมีความเชื่อถือได้ จึงอยากเพิ่มใน ส่วนของโปรแกรมดังนี้

5.3.1 เพิ่มชุดข้อมูลรูปภาพของนักศึกษาทั้งถูกระเบียบและผิดระเบียบที่นำมาทดสอบควรมีมากกว่านี้

5.3.2 ทดสอบการใช้โปรแกรมหลายครั้ง เพื่อหาความแม่นยำที่ดีกว่านี้

5.3.3 เพิ่มการทำแบบ Real time เพื่อให้โปรแกรมมีประสิทธิภาพที่สูงขึ้น

5.3.4 ใช้กล้องที่คุณภาพสูงทำให้ภาพมีความละเอียดสูงและทำให้ผลลัพธ์ที่ถูกต้องและแม่นยำมากกว่าเดิม

5.3.5 สามารถนำระบบไปช่วย หรือทำประโยชน์ให้แก่บริษัท องค์กร และคนที่มีความสนใจ ในด้านที่ อยากจะทำการตรวจสอบความถูกต้องของเครื่องแต่งกาย

บรรณานุกรม

- วิศรุต สุรการินทร์. (2557). *การรู้จำชนิดของเสื้อผ้า โดยใช้วิธีการจำแนกประเภท*. จุฬาลงกรณ์มหาวิทยาลัย, วิทยาศาสตร์มหาบัณฑิต. **เข้าถึงได้จาก**
<https://cuir.car.chula.ac.th/bitstream/123456789/45671/1/5670385021.pdf>
- David G Lowe. (2542). *Object Recognition from Local Scale-Invariant Features*. British Columbia, Computer Science Department. **เข้าถึงได้จาก**
<http://new.csd.uwo.ca/Courses/CS9840a/PossibleStudentPapers/iccv99.pdf>
- Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng Yang Fu, และ Alexander C. Berg. (2560). *SSD: Single Shot MultiBox Detector*. **เข้าถึงได้จาก**
https://link.springer.com/chapter/10.1007/978-3-319-46448-0_2
- Guimei Cao, Xuemei Xie, Wenzhe Yang, Quan Liao, Guangming Shi, และ Jinjian Wu. (2561). *Feature-fused SSD: fast detection for small objects*. **เข้าถึงได้จาก**
<https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10615/106151E/Feature-fused-SSD-fast-detection-for-small-objects/10.1117/12.2304811.short?SSO=1>
- Jisoo Jeong, Hyojin Park, และ Nojun Kwak. (2560). *Enhancement of SSD by concatenating feature maps for object detection*. **เข้าถึงได้จาก** <https://arxiv.org/pdf/1705.09587.pdf>
- Lowe David G. . (2542). *Object Recognition from Local Scale-Invariant Features*. British Columbia, Computer Science Department. **เข้าถึงได้จาก**
<http://new.csd.uwo.ca/Courses/CS9840a/PossibleStudentPapers/iccv99.pdf>
- ขวัญจิต ออกเวหา. (2555). *การศึกษาการตรวจหาวัตถุที่อยู่บนสุดของภาพที่ทับซ้อนโดยใช้เทคนิคมาร์จิ้นอลสเปซและมอร์โฟโลยี*. มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี, วิศวกรรมศาสตรมหาบัณฑิต. **เข้าถึงได้จาก**
<file:///C:/Users/Acer/Downloads/1.Front.pdf>

จารุวรรณ พูลธรรม. (2556). การเรียนรู้การประยุกต์ใช้ขบวนการวิเคราะห์ภาพสำหรับตรวจจับวัตถุ ด้วยวิธีการเรียนรู้แบบร่วมมือ. บัณฑิตวิทยาลัยมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, สาขาวิชาไฟฟ้าศึกษา ภาควิชาวิศวกรรมไฟฟ้า. **เข้าถึงได้จาก** <file:///C:/Users/Acer/Downloads/B1577417x.pdf>

ตะวัน ขุนอาสา. (2552). ระบบตรวจจับคนเดินเท้าตามถนนในเวลากลางคืนด้วยโหมดถ่ายภาพในที่มืด. มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, วิทยาการคอมพิวเตอร์. **เข้าถึงได้จาก** <file:///C:/Users/Acer/Downloads/mcs0072.pdf>

ธรัช งามวงศ์ชน. (2549). การตรวจจับเคลื่อนไหวของวัตถุในภาพจากกล้องเดี่ยวคงที่. **เข้าถึงได้จาก** <file:///C:/Users/Acer/Downloads/B14779663.pdf>

นนทรฐ บำรุงเกียรติ. (2551). การตรวจหาและติดตามบุคคลอย่างต่อเนื่องด้วยกล้องวิดีโอหลายตัวโดยใช้สีและการเคลื่อนไหว. จุฬาลงกรณ์มหาวิทยาลัย, ภาควิชาวิศวกรรมไฟฟ้า. **เข้าถึงได้จาก** file:///C:/Users/Acer/Downloads/Nontarat_bu.pdf

บงการ หอมนาน. (2551). ระบบตรวจจับวัตถุชาญฉลาดผ่านโครงข่ายพื้นที่ท้องถิ่นไร้สายโดยใช้ปัญญาประดิษฐ์และกรรมวิธีสัญญาณ. ขั้นตอนวิธีดิจิทัล. ภาควิชาวิศวกรรมโทรคมนาคม . **เข้าถึงได้จาก** <http://libdoc.dpu.ac.th/research/117373.pdf>

สมปอง เวฬุวนาธร. (2555). การรู้จำใบหน้าบนวีดิทัศน์โดยใช้หลายคุณลักษณะ. **เข้าถึงได้จาก** <file:///C:/Users/Acer/Downloads/B15736155.pdf>

Justin Ellingwood และ Kathleen Juell. (2561). **เข้าถึงได้จาก** <https://www.digitalocean.com/community/tutorials/how-to-serve-flask-applications-with-unicorn-and-nginx-on-ubuntu-18-04?fbclid=IwAR1UI8T3kgS1V5skPk9t8lUvPT6qWEvJE2hu7gRTTrH3yiGkzJunWI0zUxtw>