รายงานสหกิจศึกษาฉบับสมบูรณ์

การพัฒนาการประมวลผลและวิเคราะห์ภาพสำหรับงานสื่อโฆษณาดิจิตอล
Development of Image Processing and Analytics for Digital Advertising

นายสัมฤทธิ์ เสริมศิริมั่นคง

ภาควิชาวิศวกรรมโทรคมนาคม
คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2561

รายงานสหกิจศึกษาฉบับสมบูรณ์

การพัฒนาการประมวลผลและวิเคราะห์ภาพสำหรับงานสื่อโฆษณาดิจิตอล

Development of Image Processing and Analytics for Digital Advertising

นายสัมฤทธิ์ เสริมศิริมั่นคง

ภาควิชาวิศวกรรมโทรคมนาคม

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2561

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชื่อโครงงานสหกิจศึกษา การพัฒนาการประมวลผลและวิเคราะห์ภาพสำหรับงานสื่อโฆษณาดิจิตอล

ชื่อ-สกุล นักศึกษา นายสัมฤทธิ์ เสริมศิริมั่นคง

คณะ วิศวกรรมศาสตร์ **ภาควิชา** วิศวกรรมโทรคมนาคม

ชื่อ-สกุล อาจารย์นิเทศ ผศ.ดร. ศรวัฒน์ ชิวปรีชา

ชื่อ-สกุล ผู้นิเทศงาน นายธีระเดช กุศลธรรมรัตน์

ชื่อสถานประกอบการ บริษัท เวริลี วิชั่น จำกัด

บทคัดย่อ

รายงานสหกิจศึกษาฉบับนี้มีจุดประสงค์เพื่อทำการออกแบบระบบประมวลผลและวิเคราะห์ภาพสำหรับตรวจจับและนับจำนวนคนที่เดินผ่านจุดป้ายโฆษณาดิจิตอล เพื่อที่จะตอบสนองความต้องการของบริษัท เวริลี วิชั่น จำกัดที่ต้องการระบบที่ทำการเก็บข้อมูลรูปภาพใบหน้าคน นับจำนวนใบหน้าคน ข้อมูลระยะเวลาในการติดตามใบหน้าคนและคนที่สนใจสื่อโฆษณาดิจิตอล เพื่อที่จะนำข้อมูลที่ได้ไปช่วยให้บริษัทที่ทำการติดตั้งป้ายสื่อโฆษณาดิจิตอล (Digital signage) ในการคิดค่าโฆษณากับเจ้าของโฆษณาได้อย่างถูกต้องและเพื่อที่จะลดต้นทุนในการนำเข้าซอฟแวร์ในการติดตั้งจากต่างประเทศ ซึ่งจะสามารถแบ่งการออกแบบระบบประมวลผลภาพได้ 2 ส่วน ส่วนแรกจะเป็นส่วนของฮาร์ดแวร์ ซึ่งในส่วนนี้จะมีการนำอุปกรณ์ ASUS Tinker Board มาทำประมวลผลภาพและทำการเก็บข้อมูลรูปภาพใบหน้าคน ส่วนที่สองจะเป็นส่วนของซอฟต์แวร์ที่มีการนำหลักการ Local Binary Pattern (LBP) มาทำการตรวจจับและนับจำนวนคนที่เดินผ่านจุดป้ายโฆษณาดิจิตอลและจะใช้หลักการ Convolution Neural Networks (CNNs) ในส่วนของการจำแนกเพศ และอารมณ์ที่แสดงออก เพื่อที่จะนำผลลัพธ์ที่ได้ไปพัฒนาและทำการต่อยอดป้ายโฆษณาประเภท interactive ต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Cooperative Title: Development of Image Processing and Analytics for Digital Advertising

Student intern name: Mr. Samrit Seamsirimonkong

Faculty: Engineering **Department:** Telecommunications Engineering

Advisor name: Asst. Prof. Dr. SorawatChivapreecha

Mentor name: Mr. Theeradet Kusonthammarat

Company: Verily Vision Co., Ltd.

ABSTRACT

This cooperative study aims to design an image processing and analysis system for detecting and counting people walking through digital signage. In order to meet the needs of the company, Verily Vision requires a system that stores the image of people faces, counts the faces, and tracks faces to find information about the eyeball rate and dwell time of passengers toward digital signage. The recorded information is sent to the end user company, the digital signage service provider, for them to charge to their client at an effective rate and to reduce the cost of importing the software from Overseas.The system can be divided into two parts, the first part is the hardware part. In this part, the ASUS Tinker Board will be used to process images and to store images of people face. The second part of the program is a software that adopts the Local Binary Pattern (LBP) method to detect and count the number of people walking through digital signage and Convolution Neural Networks (CNNs) are used in the classification of gender and emotional expression. In order to bring the results to the development of interactive advertisement and media platform.

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

รายงานสหกิจศึกษาฉบับนี้สำเร็จลุล่วงไปด้วยดีเนื่องจากข้าพเจ้าได้รับความกรุณา ความเมตตาจาก ผศ.ดร.ศรวัฒน์ ชิวปรีชา และคุณธีระเดช กุศลธรรมรัตน์ ที่คอยให้ความช่วยเหลือ ดูแลเอาใจใส่ ชี้แนะข้อบกพร่องและร่วมแก้ไขปัญหาติดตามความก้าวหน้าของงานวิจัย รวมทั้งฝึกฝนให้ข้าพเจ้ามีทักษะทางด้านความคิด การอ่าน การเขียนและการนำเสนอผลงานวิชาการ ซึ่งเป็นประโยชน์อย่างมากในการพัฒนา

ขอขอบคุณพี่ที่ บริษัท เวริลี วิชั่น จำกัด ทุกท่านที่ได้ให้คำปรึกษา แนะนำ และถ่ายทอดวิชาความรู้ในด้านเทคโนโลยีแก่ข้าพเจ้า

ขอขอบคุณ สถานีรถไฟฟ้าแอร์พอร์ตลิงค์ ที่ได้อนุเคราะห์ในการติดตั้งเพื่อเก็บผลในการทดลองครั้งนี้

สุดท้ายนี้ข้าพเจ้าขอกราบขอบพระคุณบิดามารดาและครอบครัวของข้าพเจ้าที่ได้ให้ชีวิตและโอกาสทางการศึกษา คอยเป็นกำลังใจและให้ความห่วงใยเสมอมา ตลอดจนคุณครูและอาจารย์ทุกท่านที่กรุณาประสิทธิ์ประสาทวิชาความรู้อันเป็นประโยชน์แก่ข้าพเจ้า

นายสัมฤทธิ์ เสริมศิริมั่นคง
ผู้จัดทำรายงานสหกิจศึกษา

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	I
บทคัดย่อภาษาอังกฤษ	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญตาราง	VI
สารบัญภาพ	VII
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญ	1
1.2 วัตถุประสงค์ของการวิจัย	2
1.3 ขอบเขตของการวิจัย	2
1.4 วิธีการดำเนินการวิจัย	2
1.5 ประโยชน์ที่คาดว่าจะได้รับ	2
บทที่ 2 แนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้อง	3
2.1 แบบจำลองการประมวลผลและวิเคราะห์สำหรับงานสื่อโฆษณาดิจิตอล	3
2.2 ความรู้พื้นฐานภาษาไพทอน	12
2.3 โครงข่ายประสาทเทียมเบื้องต้น (Artificial Neural Networks : ANN)	15
2.4 ความรู้พื้นฐานเกี่ยวกับ Convolutional Neural Networks	25
2.5 ความรู้พื้นฐานเกี่ยวกับ Local Binary Patten	28
2.6 ความรู้พื้นฐานของรูปภาพ	39
2.7 ความรู้พื้นฐานเกี่ยวกับการพัฒนาระบบประมวล	35
2.8 ความรู้พื้นฐานของการตรวจจับวัตถุ	39
2.9 โปรแกรมที่ใช้สำหรับในการพัฒนา	43
บทที่ 3 วิธีดำเนินการวิจัย	46
3.1 การสร้างฐานข้อมูลรูปภาพใบหน้าคนและไม่ใช่ใบหน้าคน	46
3.2 การออกแบบการ Training	51
3.3 การออกแบบระบบการตรวจจับและนับจำนวนคน	58
3.4 การออกระบบการตรวจจับเพศและสีหน้า	67

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
บทที่ 4 ผลการวิจัย	77
4.1 ระบบการตรวจจับและนับจำนวนคน	77
4.2 ระบบการจำแนกเพศและสีหน้า	82
4.3 ผลการทดสอบประสิทธิภาพของระบบ	88
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ	93
5.1 สรุปผลการวิจัย	93
5.2 ข้อเสนอแนะ	93
บรรณานุกรรม	94
ภาคผนวก ก	95
ภาคผนวก ข	105
ภาคผนวก ค	117

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่	หน้า
2.1 คุณลักษณะของ Adapter 3.0 A	4
2.2 คุณลักษณะของโมดูลกล้อง	5
2.3 คุณลักษณะของเราเตอร์ TL-MR6400	6
2.4 คุณลักษณะของอุปกรณ์ ASUS Tinker Board	9
4.1 สรุปผลการทดสอบประสิทธิภาพของระบบตรวจจับและนับจำนวนคน	89
4.2 สรุปผลการทดสอบประสิทธิภาพของระบบจำแนกเพศ	90
4.3 สรุปผลการทดสอบประสิทธิภาพของระบบจำแนกสีหน้า	91

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ

รูปที่		หน้า
2.1	ภาพรวมของแบบจำลองการประมวลผลและวิเคราะห์สำหรับงานสื่อโฆษณาดิจิตอล	3
2.2	(ก) (ข) และ (ค) ส่วนประกอบของระบบประมวลผลภาพ ได้แก่ (ก) ASUS Tinker Board (ข) โมดูลกล้อง และ(ค) TP-LINK Router	4
2.3	อุปกรณ์ Adapter USB AC to DC 3.0 A	5
2.4	โมดูลกล้อง	6
2.5	อุปกรณ์เราเตอร์ TL-MR6400	7
2.6	ชิป Rockchip RK3288 บนอุปกรณ์ ASUS Tinker Board	8
2.7	HD Audio บนอุปกรณ์ ASUS Tinker Board	9
2.8	พอร์ตในการเชื่อมต่อบนอุปกรณ์ ASUS Tinker Board	10
2.9	ส่วนประกอบของเซลล์ประสาทของมนุษย์	15
2.10	แบบจำลองหลักๆ ของ ANN	16
2.11	ตัวอย่างของโครงข่ายประสาทเทียม	17
2.12	บล็อกไดอะแกรมของ Neuron	18
2.13	ตัวอย่าง Activation Function ของโครงข่ายประสาทเทียม	19
2.14	รูปภาพ Linear Separability in the Perceptron	19
2.15	โครงข่ายประสาทเทียมแบบ Multilayer perceptron ที่มีเลเยอร์ hidden 2 เลเยอร์	21
2.16	โครงข่ายประสาทเทียมแบบ Back-Propagation 3 เลเยอร์	22
2.17	การทำงานของเลเยอร์ Convolutional (Conv)	26
2.18	การทำงานของเลเยอร์ Pooling	27
2.19	การทำงานของ Local Binary Patterns (LBP)	28
2.20	รูปภาพความละเอียดขนาด 1,000 x 750 พิกเซล	30
2.21	การไล่โทนระดับสีแบบ Grayscale จากค่า สีดำ (0) ไปถึง สีขาว (255)	31
2.22	(ก) พื้นที่สี Red Green Blue Color (RGB) และ (ข) ลูกบาศก์ Red Green Blue Color (RGB)	31
2.23	รูปภาพในพื้นที่สี RGB ที่แต่ละสีมารวมกันเป็นรูปภาพ 3 มิติ	32
2.24	รูปภาพตัวอักษร I ที่วางอยู่บนกระดาษกราฟ	33

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ (ต่อ)

รูปที่	หน้า
2.25 รูปด้านซ้ายเป็นรูปภาพต้นฉบับ และรูปด้านบนและด้านล่างเป็นรูปภาพที่ปรับขนาดความกว้าง (W) และความสูง (H)	34
2.26 รูปด้านซ้าย รูปภาพต้นฉบับจากกล้อง รูปด้านขวา รูปภาพที่ปรับขนาดความกว้าง (W) และความสูง (H)	35
2.27 รูปด้านซ้าย คือ ฐานข้อมูลรูปภาพใบหน้าคน และ รูปด้านขาว คือ ฐานข้อมูลที่ไม่ใช่ใบหน้าคน	36
2.28 ใบหน้าคนจากมุมมองจริงของฐานข้อมูลใบหน้าคน	37
2.29 รูปภาพใบหน้าแบบหลายมมุมมองของฐานข้อมูลใบหน้าคน	37
2.30 ข้อมูลรูปภาพที่ไม่ใช่ใบหน้าคนจากมุมมองจริง	38
2.31 ข้อมูลรูปภาพที่ไม่ใช่ใบหน้าคนจากความผิดพลาดจริง	39
2.32 รูปด้านซ้าย คือ ขั้นตอนพื้นฐานของการตรวจจับหาวัตถุ และรูปด้านขวา คือ กระบวนการ Image pyramid โดยรูปภาพต้นฉบับจะอยู่ที่ด้านล่าง	40
2.33 รูปภาพของการทำกระบวนการ Sliding Windows	41
2.34 รูปด้านซ้าย คือ รูปภาพที่ทำกระบวนการ Sliding window และรูปด้านขวา คือ รูปภาพที่ทำกระบวนการ Sliding window และ Image Pyramids	42
2.35 รูปด้านซ้าย แสดงการทำงานแบบไม่มี NMS และ รูปด้านขวา แสดงการทำงานแบบมี NMS	43
2.36 โปรแกรม Teamviewer	44
2.37 โปรแกรม Cascade-Trainer-GUI	44
2.38 โปรแกรม Visual Studio Code	45
3.1 รูปภาพของสถานีลาดกระบัง 45,513 รูป	47
3.2 การตัดรูปภาพใบหน้าคนจากมุมมองจริงของสถานีลาดกระบัง	47
3.3 รูปภาพใบหน้าแบบหลายมมุมมองของฐานข้อมูลใบหน้าคน	48
3.4 การตัดรูปภาพที่ไม่ใช่ใบหน้าคนจากมุมมองจริงของสถานีลาดกระบัง	49
3.5 ฐานข้อมูลรูปภาพที่ไม่ใช่ใบหน้าคนจากมุมมองจริงของทุกสถานี	49
3.6 ฐานข้อมูลรูปภาพที่ไม่ใช่ใบหน้าคนจากมุมมองจริงของสถานีทุกสถานี	50
3.7 ฐานข้อมูลที่ใช้สำหรับในการ Training	51

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ (ต่อ)

รูปที่	หน้า
3.8 โปรแกรม Cascade-Trainer-GUI	51
3.9 หน้าต่างโปรแกรม Casecade-Trainer-GUI	52
3.10 แสดงการเลือก Browse	52
3.11 แสดงการเลือกที่เก็บข้อมูลรูปภาพ	53
3.12 แสดงขั้นตอนในการเลือกชนิด และจำนวนรูปภาพ	54
3.13 แสดงการตั้งค่าของหน้าต่าง Common	55
3.14 แสดงการตั้งค่าของหน้าต่าง Casecade	56
3.15 แสดงการตั้งค่าของหน้าต่าง Boost	57
3.16 โฟว์ชาร์ตของระบบตรวจจับและนับจำนวนคน	58
3.17 โมเดลของ Real-time Convolutional Neural Networks for Emotion and Gender Classification	67
3.18 โฟว์ชาร์ตของระบบการตรวจจับสีหน้าและเพศ	68
4.1 การลดขนาดของรูปภาพ	77
4.2 ปริมาณข้อมูลของรูปภาพระหว่างรูปภาพต้นฉบับกับรูปภาพที่ลดขนาด	77
4.3 การแปลงสีจาก RGB เป็น Grayscale	78
4.4 ปริมาณข้อมูลของรูปภาพระหว่างรูปภาพสี RGB กับรูปภาพสี Grayscale	78
4.5 ค่าการตรวจจับใบหน้าคน และค่าจุดศูนย์การของใบหน้า	78
4.6 ผลจากการตรวจจับใบหน้าคน	79
4.7 ผลจากการตรวจสอบใบหน้าคน	79
4.8 การเก็บค่านับจำนวนคนและค่าที่บันทึกรูปภาพ	80
4.9 ผลจากการบันทึกรูปภาพใบหน้าคน	80
4.10 แสดงค่าการอัพเดทการตรวจจับใบหน้าคน	81
4.11 การบันทึกระยะเวลาในการตรวจจับใบหน้าคน	82
4.12 การรับวิดีโอใบหน้าคน	83
4.13 ปริมาณข้อมูลของรูปภาพ	83
4.14 การแปลงสีจาก RGB เป็น Grayscale ของภาพใบหน้าคน	84

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ (ต่อ)

รูปที่	หน้า
4.15 ปริมาณข้อมูลของรูปภาพระหว่างรูปภาพสี RGB กับรูปภาพสี Grayscale ของภาพใบหน้าคน	84
4.16 ค่าการตรวจจับใบหน้าคน และค่าจุดศูนย์การของใบหน้า	84
4.17 ผลจากการจำแนกเพศและสีหน้า	85
4.18 ผลจากการตรวจสอบใบหน้าคนของการจำแนกเพศและสีหน้า	85
4.19 การเก็บค่านับจำนวนคนและค่าที่บันทึกรูปภาพของการจำแนกเพศและสีหน้า	86
4.20 ผลจากการบันทึกรูปภาพใบหน้าคนของการจำแนกเพศและสีหน้า	86
4.21 แสดงค่าการอัพเดทการตรวจจับใบหน้าคนของการจำแนกเพศและสีหน้า	87
4.22 การบันทึกระยะเวลาในการตรวจจับใบหน้าคนของการจำแนกเพศและสีหน้า	87
4.23 กราฟความสัมพันธ์ระหว่างระยะเวลาในการตรวจจับใบหน้ากับลำดับใบหน้าคน	88
4.24 กราฟความสัมพันธ์ระหว่างระยะเวลาในการตรวจจับใบหน้ากับจำนวนใบหน้าคน	89
4.25 กราฟแสดงประสิทธิภาพความถูกต้องในการตรวจจับและนับจำนวนคน	90
4.26 กราฟแสดงประสิทธิภาพความถูกต้องในการจำแนกเพศ	91
4.27 กราฟแสดงประสิทธิภาพความถูกต้องในการจำแนกสีหน้า	92

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญ

ปัจจุบันป้ายสื่อโฆษณาดิจิตอลได้เข้ามามีบทบาทสำคัญในการเผยแพร่งานโฆษณาต่าง ๆ ซึ่งเราสามารถพบเห็นได้บ่อยตามสถานีรถไฟฟ้า หรือห้างสรรพสินค้า โดยบริษัทที่เผยแพร่งานสื่อเหล่านี้ตามสถานีรถไฟฟ้าต้องการที่จะทราบจำนวนผู้ที่ให้ความสนใจป้ายสื่อโฆษณาเพื่อนำไปคิดค่าบริการกับเจ้าของโฆษณา ดังนั้นบริษัทจึงต้องนำข้อมูลผู้ใช้บริการสถานีรถไฟฟ้าหรือในสถานที่ติดตั้งป้ายต่าง ๆ ในแต่ละวันเป็นตัวกำหนดมาตรฐานค่าบริการเผยแพร่โฆษณา ซึ่งพบว่าวิธีในการวัดผลจำนวนผู้เข้าใช้บริการสถานีรถไฟฟ้าในปัจจุบันจากจำนวนการใช้งานบัตรโดยสารนั้นไม่ได้บ่งบอกถึงปริมาณคนที่ให้ความสนใจป้ายสื่อโฆษณาดิจิตอล และสำหรับสถานที่อื่น ๆ ที่ไม่มีระบบวัดผลจากจำนวนบัตรโดยสารนั้นจะไม่สามารถคาดการณ์จำนวนการชมป้ายโฆษณาที่แท้จริงได้ นอกจากนี้ระบบควบคุมการแสดงผลของป้ายโฆษณาดิจิตอลในปัจจุบันนั้นต้องมีการนำเข้าจากต่างประเทศ ทำให้มีค่าใช้จ่ายที่สูง รวมถึงความยากในการปรับเปลี่ยนและพัฒนาต่อยอด ด้วยเหตุนี้จึงทำให้เกิดเป็นแนวคิดที่จะทำระบบประมวลผลภาพไปติดตั้งบริเวณป้ายสื่อโฆษณาดิจิตอล เพื่อตรวจจับและนับจำนวนคนที่ให้ความสนใจป้ายสื่อโฆษณาดิจิตอล รวมทั้งจำแนกเพศ และสีหน้าที่แสดงออกของคนที่รับชมป้ายสื่อโฆษณาดิจิตอล รวมถึงระบบปฏิบัติการสำหรับการบริหารจัดการการแสดงผลของสื่อโฆษณาบนจอป้ายโฆษณาดิจิตอลที่ถ้าสามารถตั้งค่าการแสดงผลได้ทั้งแบบออนไลน์และออฟไลน์ ซึ่งผลลัพธ์จากการพัฒนานี้ยังสามารถนำไปต่อยอดในการพัฒนาระบบป้ายโฆษณาที่มีลักษณะเป็นป้ายโฆษณาประเภท interactive ที่สามารถโต้ตอบกับผู้ชมป้ายได้ในหลากหลายรูปแบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2 วัตถุประสงค์ของการวิจัย

1. เพื่อทำการออกแบบระบบประมวลผลภาพตรวจจับและนับจำนวนคนที่เดินผ่านจุดป้ายโฆษณาดิจิตอล
2. เพื่อศึกษาการทำงานเสริมในการตรวจจับใบหน้าบุคคลที่ให้ความสนใจจุดป้ายโฆษณาดิจิตอลสำหรับในการจำแนกเพศ และสีหน้าที่แสดงออก

1.3 ขอบเขตของการวิจัย

1. การออกแบบระบบประมวลผลภาพตรวจจับและนับจำนวนคนที่เดินผ่านจุดป้ายโฆษณาดิจิตอล
2. ศึกษาการทำงานเสริมในการตรวจจับใบหน้าบุคคลที่ให้ความสนใจจุดป้ายโฆษณาดิจิตอลสำหรับในการจำแนกเพศ และสีหน้าที่แสดงออก

1.4 วิธีการดำเนินการวิจัย

1.4.1 ทำการศึกษาพื้นฐานเกี่ยวกับระบบประมวลภาพ
1.4.2 ทำการศึกษาวิธีการใช้งานของโปรแกรม Visual Studio Code
1.4.3 ทำการศึกษาภาษาไพทอนและพื้นฐานการเขียนโปรแกรม
1.4.4 ศึกษาโครงข่ายประสาทเทียมเบื้องต้น
1.4.5 ศึกษาความรู้พื้นฐานเกี่ยวกับ Convolutional Neural Networks
1.4.6 ศึกษาความรู้พื้นฐานของรูปภาพ
1.4.7 ศึกษาความรู้พื้นฐานเกี่ยวกับการพัฒนาระบบประมวลผล
1.4.8 ศึกษาความรู้พื้นฐานของการตรวจจับวัตถุ
1.4.9 ทำการเขียนโปรแกรมของระบบตรวจจับและนับจำนวนคน และทำการเขียนโปรแกรมในส่วนของระบบการจำแนกเพศ และสีหน้า
1.4.10 ทำการวิเคราะห์เปรียบเทียบผลที่ได้จากการตรวจจับและนับจำนวนคน และระบบในการจำแนกเพศ และสีหน้า ว่าเป็นไปตามผลลัพธ์ทางบริษัทที่ต้องการมากน้อยแค่ไหน

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1.5.1 สามารถที่จะสร้างสมรรถภาพสำหรับในงานด้านการเขียนโปรแกรมให้ชำนาญมากยิ่งขึ้น
1.5.2 สามารถที่จะออกแบบระบบประมวลผลภาพตรวจจับและนับจำนวนคนที่ผ่านจุดป้ายโฆษณาดิจิตอลได้
1.5.3 สามารถที่จะเข้าใจความรู้พื้นฐานเกี่ยวกับระบบประมวลภาพได้มากยิ่งขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

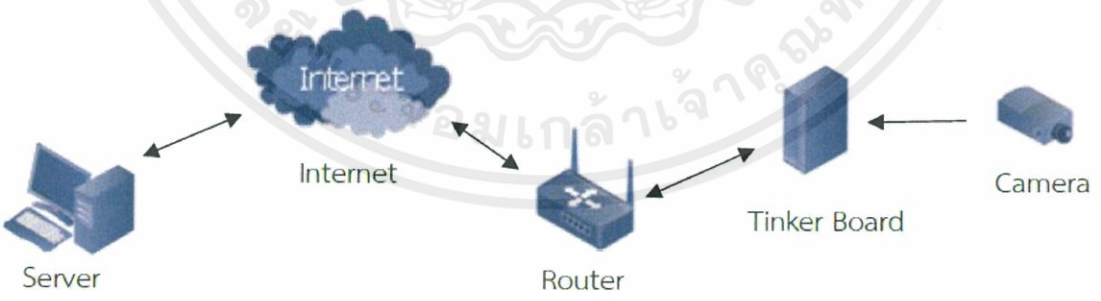
บทที่ 2

แนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทที่ 2 นี้จะกล่าวถึงทฤษฎีและหลักการที่เกี่ยวข้องในการนำมาใช้ในการพัฒนาการประมวลผลและวิเคราะห์ภาพสำหรับงานสื่อโฆษณาดิจิตอล เช่น แบบจำลองการประมวลผลและวิเคราะห์สำหรับงานสื่อโฆษณาดิจิตอล ความรู้พื้นฐานของรูปภาพ ความรู้พื้นฐานของการตรวจจับรูปภาพ และการ Tracking เป็นต้น โดยในทฤษฎีและหลักการที่อธิบายจะมีที่มาจากบริษัท เวริสี วิชั่น จำกัด

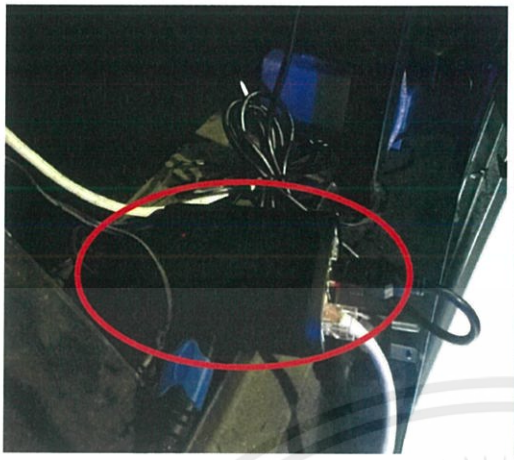
2.1 แบบจำลองการประมวลผลและวิเคราะห์สำหรับงานสื่อโฆษณาดิจิตอล

แบบจำลองการประมวลผลและวิเคราะห์สำหรับงานสื่อโฆษณาดิจิตอลในรายงานสหกิจศึกษาเป็นระบบที่ถูกติดตั้งจริงที่สถานีแอร์พอร์ตเรลลิงค์ทั้ง 3 สถานีได้แก่ สถานีพญาไท สถานีลาดกระบัง และสถานีท่าอากาศยานสุวรรณภูมิ โดยในทั้ง 3 สถานีนี้จะมีภารกิจหลักทั้ง 3 ภารกิจคือ 1) การตรวจจับ โดยการตรวจจับจะต้องทำการตรวจจับทั้งคนที่เดินผ่านจุดสื่อโฆษณาดิจิตอลและคนที่ยืนรอรถไฟฟ้าแอร์พอร์ตเรลลิงค์ 2) นับจำนวนคน โดยในการทำงานของในส่วนนี้คล้ายกับการตรวจจับโดยจะทำการนับคนที่เดินผ่านจุดโฆษณาดิจิตอลและคนที่ยืนรอรถไฟฟ้าแอร์พอร์ตเรลลิงค์ 3) สามารถนำมาใช้งานจริงได้ในการติดตั้งที่สถานีทั้ง 3 สถานี โดยภาพรวมของแบบจำลองการประมวลผลและวิเคราะห์สำหรับงานสื่อโฆษณาดิจิตอลสามารถแสดงได้ดังรูปที่ 2.1 และรูปภาพอุปกรณ์ที่ทำการติดตั้งที่สถานีแสดงดังรูปที่ 2.2



รูปที่ 2.1 ภาพรวมของแบบจำลองการประมวลผลและวิเคราะห์สำหรับงานสื่อโฆษณาดิจิตอล

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ก)

(ข)

(ค)

รูปที่ 2.2 (ก) (ข) และ (ค) ส่วนประกอบของระบบประมวลผลภาพ ได้แก่
(ก) ASUS Tinker Board (ข) โมดูลกล้อง และ(ค) TP-LINK Router

จากรูปที่ 2.1 และรูปที่ 2.2 แบบจำลองการประมวลผลและวิเคราะห์สำหรับงานสื่อโฆษณาดิจิตอลจะประกอบไปด้วยอุปกรณ์การทำงานหลักถึง 4 อย่าง คือ อุปกณ์แปลงไฟฟ้า ASUS Tinker Board โมดูลกล้อง และ TP-LINK Router

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.1 อุปกรณ์แปลงไฟฟ้า

อุปกรณ์แปลงไฟฟ้าเป็นหนึ่งในส่วนประกอบที่สำคัญที่ทำให้แบบจำลองการประมวลผลและวิเคราะห์สำหรับงานสื่อโฆษณาดิจิตอลสามารถปฏิบัติงานได้ ในรายงานสหกิจศึกษานี้จะใช้อุปกรณ์ Adapter USB AC to DC 3.0 A ที่สามารถแปลงแรงดันไฟฟ้าแบบสลับจาก 220 โวลต์ เป็นแรงดันไฟฟ้าคงที่ 5 โวลต์ และจ่ายกระแสไฟฟ้าได้ถึง 3 แอมป์สามารถแสดงดังรูปที่ 2.3 และคุณลักษณะของ Adapter 3.0 A ดังตารางที่ 2.1

รูปที่ 2.3 อุปกรณ์ Adapter USB AC to DC 3.0 A

ตารางที่ 2.1 คุณลักษณะของ Adapter 3.0 A

รายละเอียด	คำอธิบาย
แรงดันไฟฟ้าเข้า	110 ถึง 240 โวลต์
แรงดันไฟฟ้าออก	5 โวลต์
กระแสขาออก	3 แอมป์
หัวสาย	พอร์ต Micro USB

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.2 อุปกรณ์โมดูลกล้อง

โมดูลกล้องที่ใช้งานในแบบจำลองการประมวลผลและวิเคราะห์สำหรับงานสื่อโฆษณาดิจิตอล จะต้องเป็นโมดูลกล้องที่มีขนาดกระทัดรัด มุมมองภาพขนาดกว้าง และมีความละเอียดสูง ซึ่งในการทำสหกิจครั้งนี้ได้ใช้โมดูลกล้อง ELP-USB8MP02G-BL21 ดังรูปที่ 2.4 และมีคุณลักษณะของโมดูลกล้องดังตารางที่ 2.2

รูปที่ 2.4 โมดูลกล้อง

ตารางที่ 2.2 คุณลักษณะของโมดูลกล้อง

รายละเอียด	**คำอธิบาย**
ความคมชัด	3,264 x 2,448 พิกเซล
ขนาดเลนส์	2.1 มิลลิเมตร
โมลเดล	ELP-USB8MP02G-BL21
ความละเอียด	8 เมกะพิกเซล
รูปแบบสี	Color
หัวการเชื่อมต่อ	พอร์ต USB

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.3 อุปกรณ์เราเตอร์

อุปกรณ์เราเตอร์ที่ใช้งานในการแบบจำลองการประมวลผลและวิเคราะห์สำหรับงานสื่อโฆษณาดิจิตอลจะต้องเป็นอุปกณ์ที่มีขนาดปานกลาง สามารถทำงานโดยผ่านซิมการ์ด และรองรับความเร็ว 3G/4G ทุกเครือข่ายในไทย ดังนั้นในการทำสหกิจครั้งนี้จึงได้ใช้อุปกรณ์เราเตอร์ TL-MR6400 ที่สามารถปล่อยความเร็วได้ถึง 300 เมกะไบร์ตต่อวินาที ซึ่งสามารถแสดงได้ในรูปที่ 2.5 และคุณลักษณะของเราเตอร์ TL-MR6400 ดังตารางที่ 2.3

รูปที่ 2.5 อุปกรณ์เราเตอร์ TL-MR6400

ตารางที่ 2.3 คุณลักษณะของเราเตอร์ TL-MR6400

รายละเอียด	**คำอธิบาย**
รูปแบบ	ซิมการ์ด
สายอากาศ	สายอากาศ 2 ด้านในแบบ Wi-FI Antennas และ สายอากาศ 2 ด้านนอกแบบ 4G LTE Antennas
มาตรฐานไวไฟ	IEEE 802.11b/g/n 2.4GHz
เครือข่ายรองรับ 4G	AIS DTAC TRUE H CAT
เครือข่ายรองรับ 3G	AIS TOT TRUE H DTAC
เครือข่ายรองรับ 2G	AIS DTAC TRUE

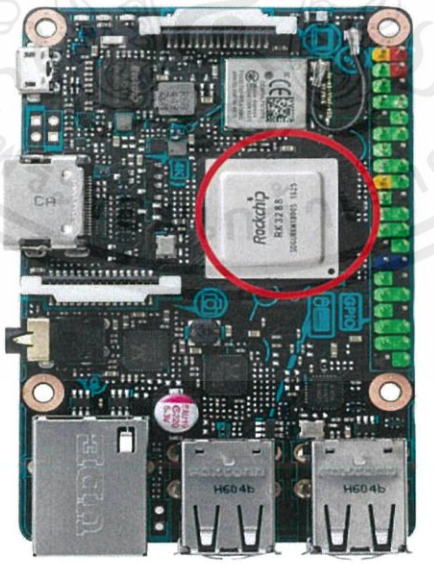
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.3 (ต่อ) คุณลักษณะของเราเตอร์ TL-MR6400

รายละเอียด	คำอธิบาย
ความเร็วดาวโหลด	150 เมกะไบต์ (ต่ำสุด) ถึง 300 เมกะไบต์ (สูงสุด) ที่ความถี่ 2.4 กิกะเฮิร์ต
ความเร็วอัพโหลด	50 เมกะไบต์
อุณหภูมิทำงาน	-40 ถึง 70 องศาเซลเซียส
ขนาด	20 x 15 x 4 เซนติเมตร

2.1.4 อุปกรณ์ ASUS Tinker Board

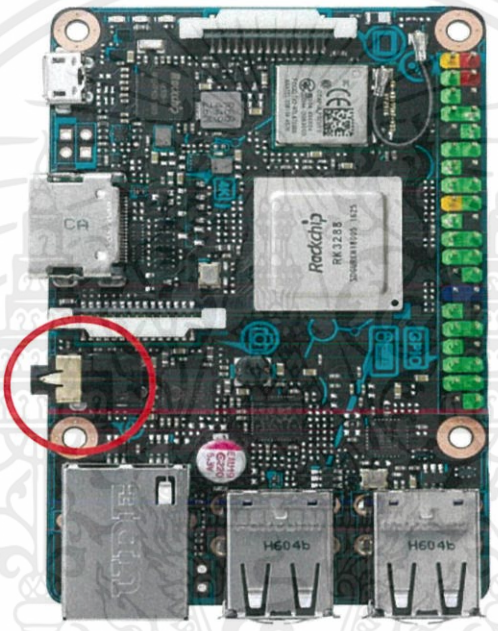
ASUS Tinker Board เป็น Single Board Computer (SBC) ที่เป็นบอร์ดประมวลผลขนาดเล็กที่มีประสิทธิภาพสูงสำหรับการทำงาน ซึ่งเหมาะสมที่จะนำใน ASUS Tinker Board มาทำการประมวลผลภาพเพื่อตรวจจับและนับจำนวนคนที่ผ่านบริเวณจุดป้ายโฆษณาดิจิตอล เพราะด้วยชิปประมวลผล Rockchip RK3288แบบ Quad-core จึงทำให้ ASUS Tinker Board มีประสิทธิภาพสูงในการทำงาน พร้อมกับหน่วยความจำ LPDDR3 2GB แบบ Dual-channel และยังมาพร้อมกับอินเตอร์เฟส SD 3.0 เพื่อให้ประสิทธิภาพความเร็วการอ่านและเขียนลงในหน่วยความจำ (microSD) สำหรับระบบปฏิบัติการ แอพพลิเคชั่นและการเก็บข้อมูล



รูปที่ 2.6 ชิป Rockchip RK3288 บนอุปกรณ์ ASUS Tinker Board

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 2.6 อุปกรณ์ ASUS Tinker Board ยังถูกออกแบบมาเพื่อประหยัดพลังงาน พร้อมทั้งรองรับการประมวลผลกราฟฟิกระดับสูง โดย GPU ARM-based Mali™-T764 GPU ของ ASUS Tinker Board เป็นหน่วยประมวลผลแบบ Fixed-point ทำให้สามารถใช้งานได้อย่างหลากหลาย รวมถึงการเล่นไฟล์มัลติมีเดียคุณภาพสูง เล่นเกม แสดงผล ตรวจสอบการเคลื่อนไหว ลดอาการภาพสั่น รวมถึงการประมวลผลสำหรับรูปภาพและอื่นๆ อีกมากมาย โดยมัลติมีเดียมีการรองรับการเล่น H.264 และ H.265 รวมถึงการเล่นไฟล์วิดีโอ HD และ UHD ที่ 30 fps



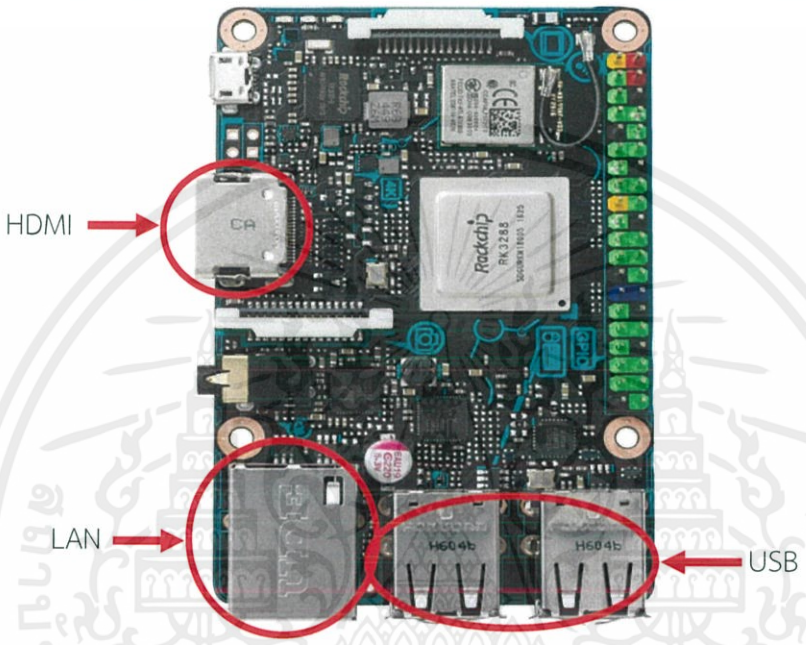
รูปที่ 2.7 HD Audio บนอุปกรณ์ ASUS Tinker Board

จากรูปที่ 2.7 ASUS Tinker Board ได้รับการติดตั้งภาคถอดสัญญาณเสียงแบบ HD ที่รองรับเสียง 192kHz/24-bit ช่องเชื่อมต่อเสียงยังสามารถรองรับเอาท์พุตเสียงและไมโครโฟนโดยไม่ต้องใช้สายต่อเพิ่มเติม

ASUS Tinker Board ยังมีการเชื่อมต่อ HD MIPI 2 พอร์ตเพื่อการแสดงผล HD และกล้อง HD Tinker board ยังมีพอร์ต LAN อัตราเร็วขนาด Gigabit/sec เพื่อการโอนถ่ายข้อมูลขนาดใหญ่ ซึ่งเหมาะแก่การเป็นศูนย์กลางระบบเครือข่ายและเหมาะที่จะเป็นอุปกรณ์จัดเก็บข้อมูลบนระบบเครือข่าย อีกทั้งยังมีตัวอุปกรณ์ WiFi และ Bluetooth ที่มีการติดตั้งแผ่นโลหะเพื่อใช้สำหรับการป้องกันสัญญาณรบกวนและทำให้ประสิทธิภาพในการรับ WiFi และ Bluetooth ดีขึ้น และยังมีสายอากาศแบบ IPEX ที่มีการทำงานที่

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ความถี่ 1.8 GHz ซึ่งทำให้อุปกรณ์ ASUS Tinker Board มีการรับสัญญาณ WiFi และ Bluetooth ที่ระยะได้ไกลมากขึ้นและมีประสิทธิภาพในการรับที่ดีขึ้น อุปกรณ์ ASUS Tinker Board มีพอร์ต HDMI แบบมาตรฐานที่ใช้สำหรับการเชื่อมต่อกับจอ TV และยังมีพอร์ต USB 2.0 จำนวน 4 พอร์ตเพื่อการเชื่อมต่ออุปกรณ์เสริม



รูปที่ 2.8 พอร์ตในการเชื่อมต่อบนอุปกรณ์ ASUS Tinker Board

ตารางที่ 2.4 คุณลักษณะของอุปกรณ์ ASUS Tinker Board

รายละเอียด	คำอธิบาย
CPU	Rockchip Quad-Core RK3288 processor
GPU	Integrated Graphics Processor ARM® Mali™-T764 GPU*1
ขนาดความจำ	ขนาดความจำ 2 กิกะเฮิร์ตแบบ Dual Channel DDR3
ตัวเก็บข้อมูล	Micro SD (TF) card slot
สายแลน	RTL GB LAN
มาตรฐาน WiFi	802.11 b/g/n และ Bluetooth V4.0 + EDR
พอร์ต USB	4 พอร์ต USB 2.0
มาตราเสียง	RTL ALC4040 CODEC

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.4 (ต่อ) คุณลักษณะของอุปกรณ์ ASUS Tinker Board

รายละเอียด	คำอธิบาย
แรงดันไฟฟ้าเข้า	5 โวลต์
กระแสไฟฟ้าเข้า	2 ถึง 2.5 แอมป์
น้ำหนัก	55 กรัม
ระบบปฏิบัติการ	Debian และKODI
พอร์ต อินพุต	**1 x 40-pin header** - up to 28 x GPIO pins - up to 2 x SPI bus - up to 2 x I2C bus - up to 4 x UART - up to 2 x PWM - up to 1 x PCM/I2S - 2 x 5V power pins - 2 x 3.3V power pins - 8 x ground pins **1 x 2-pin contact pin** - 1 x PWM - 1 x S/PDIF **1 x 15-pin MIPI DSI** **1 x 15-pin MIPI CSIt**

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2 ความรู้พื้นฐานภาษาไพทอน

ภาษาไพทอน เป็นอีกส่วนที่สำคัญที่ได้ถูกนำมาใช้งานในการทำแบบจำลองการประมวลผลและวิเคราะห์สำหรับงานสื่อโฆษณาดิจิตอล และเป็นภาษาเขียนที่ใช้งานกันอย่างกว้างขวาง โดยภาษาไพทอนเป็นภาษาที่ได้รับความนิยมเป็นอย่างมาก เพราะเป็นภาษาที่เข้าใจง่าย เข้าถึงง่าย และเป็นภาษาที่ผู้พัฒนาโปรแกรมส่วนใหญ่ให้การสนับสนุนการทำงานอย่างแพร่หลาย เช่น การเขียนโปรแกรมเชิงวัตถุ การเขียนโปรแกรมเชิงฟังก์ชั่น และการเขียนโปรแกรมแบบขั้นตอน แล้วยังมีไลบรารี่ที่ครอบคลุมการทำในส่วนของการเขียนโปรแกรมเชิงฟังก์ชั่นมากกว่าภาษา C C++ เช่น ไลบรารี่-Numpy ไลบรารี่-cv2 และไลบรารี่-pyimagesearch เป็นต้น ซึ่งทำให้เหมาะกับการนำภาษา Python มาใช้งานในด้าน image processing เพื่อการตรวจจับและนับจำนวนคนที่ผ่านบริเวณจุดป้ายโฆษณาดิจิตอล โดยการใช้งานภาษาไพทอน จำเป็นต้องศึกษา 2 ส่วนสำคัญ ส่วนของไลบรารี่ และส่วนของคำสั่งที่ใช้งาน

2.2.1 ส่วนของไลบรารี่

ไลบรารี่ เป็นอีกส่วนที่สำคัญสำหรับใช้งานในการเขียนโปรแกรมตรวจจับและนับจำนวนคนที่ผ่านบริเวณจุดป้ายโฆษณาดิจิตอล โดยในแต่ละไลบรารี่จะมีการทำงานที่แตกต่างกันออกไป ดังนี้

2.2.1.1 ไลบรารี่ numpy

NumPy (Numeric Python) เป็นหนึ่งในไลบรารี่ส่วนเสริมของภาษาไพทอน ที่มีฟังก์ชันเกี่ยวกับคณิตศาสตร์และการคำนวณต่างๆ มาให้ใช้งาน โดยทั่วไปจะเกี่ยวกับการจัดการข้อมูลชุด (Array) ขนาดใหญ่และเมทริกซ์ ซึ่งเป็นโมดูลเสริมที่สำคัญที่จะนำมาใช้งานในการสร้างระบบประมวลผลภาพเพื่อตรวจจับและนับจำนวนคนที่ผ่านบริเวณจุดป้ายโฆษณาดิจิตอล

2.2.1.2 ไลบรารี่ OpenCV

OpenCV เป็นหนึ่งใน open source library ที่ได้นำมาใช้สำหรับทำงานจำพวก image processing ต่างๆ เช่น การเรียกไฟล์ อ่านรูปภาพและวีดีโอ Correlation Region of interest convert color Image Filtering Mean shift และ Convolution เป็นต้น พร้อมทั้งมีการ interface ที่หลากหลายภาษาให้เลือกใช้งาน เช่น C++ จาวา และไพทอน จึงทำให้นำ OpenCV มาทำการประยุกต์ใช้งานในด้าน computer vision ในการประมวลผลภาพเพื่อตรวจจับและนับจำนวนคนที่ผ่านบริเวณจุดป้ายโฆษณาดิจิตอลซึ่งส่วนที่นำมาใช้ในการทำสหกิจครั้งนี้ เช่น cv2.VideoCapture() ใช้สำหรับในการเปิดกล้อง cv.Tracker() ใช้สำหรับในการติดตามวัตถุที่ตรวจจับ cv2.namedWindow() ใช้สำหรับในการตั้งชื่อไฟล์วีดีโอที่แสดงผล cv2.resize() เป็นส่วนสำคัญในปรับขนาดของรูปภาพหรือวีดีโอที่แสดงผลออกมา cv2.cvtColor()

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เป็นส่วนสำคัญการแปลงสีของรูปภาพหรือวีดีโอ cv2.rectangle() เป็นส่วนที่ทำการตีกรอบสี่เหลี่ยมกับข้อมูลที่ตรวจจับ cv2.imshow() ใช้สำหรับแสดงผลลัพธ์จากการประมวลผล และcv2.destroyAllWindows() ใช้สำหรับลบหน้าจอทั้งหมดที่แสดงผลลัพธ์ เป็นต้น

2.2.1.3 ไลบรารี่ time

ไลบรารี่ time เป็นไลบรารี่หนึ่งในภาษาไพทอน ที่ใช้สำหรับแสดงค่าเวลาในปัจจุบันออกมา ซึ่งเป็นไลบรารี่สำคัญที่นำมาใช้ในการตรวจจับคนที่สนใจจุดป้ายสื่อโฆษณาดิจิตอล โดยทำการใช้ time.time() ในการเก็บช่วงระยะเวลาที่คนให้ความสนใจจุดป้ายสื่อโฆษณาดิจิตอล

2.2.1.4 ไลบรารี่ dlib

ไลบรารี่ dlib เป็น open source library ที่ใช้โมเดลของภาษา C++ ที่นำมาใช้งานสำหรับงานพวก image processing เพื่อมาทำการตรวจจับและติดตามคนที่ผ่านบริเวณจุดป้ายโฆษณาดิจิตอล โดย dlib ในการใช้อัลกอริทึมที่จำเป็นที่ประกอบด้วย Machine learning algorithms Numerical algorithms Face recognition และDetecting objects เป็นต้น

2.2.1.5 ไลบรารี่ threading

ไลบรารี่ threading เป็นไลบรารี่ที่สามารถทำงานแบบขนาน โดยแต่ละ threading จะสามารถทำงานไปพร้อมกัน โดยปกติการทำงานของโปรแกรมจะทำงานทีละบรรทัด และต้องรอให้บรรทัดก่อนหน้าเสร็จก่อนจึงจะทำการประมวลผลต่อไป โดยการใช้ threading นั้นสามารถทำให้โปรแกรมทำงานพร้อมกันได้ในแต่ละ threading ที่ถูกสร้างขึ้นภายในโปรแกรม เช่น การใช้งานในการตรวจจับและนับจำนวนคนไปพร้อมกับการเพิ่มจำนวนคนที่เพิ่มเข้ามาในระบบ

2.2.1.6 ไลบรารี่ csv

ไลบรารี่ csv เป็นอีกหนึ่งส่วนที่สำคัญที่ใช้สำหรับในการบันทึกข้อมูลในการตรวจจับและนับจำนวนคนออกมาอยู่ในรูปของตารางแล้วบันทึกว่าจำนวนคนที่เดินผ่านมีจำนวนกี่คน และแต่ละคนใช้เวลาเท่าไรสำหรับในการติดตาม

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.2 ส่วนของคำสั่งที่ใช้งาน

นอกจากที่จะทำความเข้าใจเกี่ยวกับการทำงานส่วนของไลบรารี่แล้วจะต้องทำความเข้าใจเกี่ยวกับส่วนของคำสั่งที่นำมาใช้งาน เนื่องจากในการเขียนโปรแกรมจะต้องมีความเข้าใจเกี่ยวกับคำสั่ง ดังนี้

2.2.2.1 คำสั่ง if

คำสั่ง if เป็นคำสั่งเลือกเงื่อนไขในภาษาไพทอน โดยจะพูดถึงการควบคุมการทำงานโปรแกรมด้วยคำสั่ง if เพื่อให้โปรแกรมสามารถทำงานซับซ้อนและมีประสิทธิภาพมากขึ้น เช่น การตรวจจับและนับจำนวนคนจะมีการทำงานทุกๆ 5 เฟรม และจะทำการเช็คว่าถ้าความกว้างของรูปภาพมีค่ามากเกินขนาดที่กำหนดจะทำการลบทิ้ง เป็นต้น โดยสามารถดูการทำงานได้ดังนี้

```
tracked_position =  faceTrackers[fid].get_position()
t_w = int(tracked_position.width())
if t_w > 52:
    fidsToDelete.append( fid )
```

ในส่วนของคำสั่ง if ข้างต้นจะมีการเช็คขนาดความกว้างของรูปภาพในแต่ละคนโดยถ้าค่าขนาดความกว้างมีขนาดที่มากเกินกว่าค่า 52 พิกเซล จะทำการลบรูปภาพที่กำหนดทิ้งไป

2.2.2.2 คำสั่ง exception

คำสั่ง exception เป็นคำสั่งที่ใช้สำหรับในการจัดการข้อผิดพลาดในภาษาไพทอน โดยในการเขียนโปรแกรมแล้วทำการรันแบบปกติ ถ้าพบข้อผิดพลาดโปรแกรมจะทำการหยุดโปรแกรมทันที แต่ในคำสั่ง exception เมื่อเกิดข้อผิดพลาดก็ยังสามารถรันโปรแกรมต่อไปได้ สามารถดูการทำงานได้ดังนี้

```
try:
    rc,fullSizeBaseImage = cap.read()
    ...
except KeyboardInterrupt as e:
    pass
```

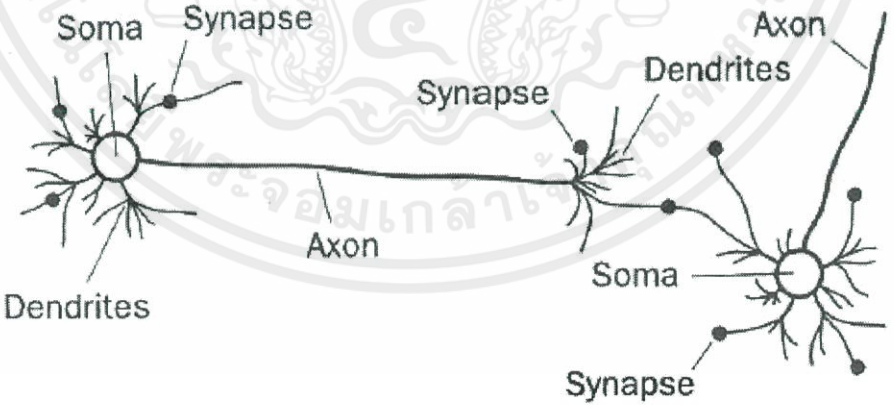

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในส่วนของคำสั่ง exception นี้จะมีการทำงานในส่วนของคำสั่ง try ก่อนแต่เมื่อเกิดข้อผิดพลาดผิดพลาดขึ้น จะทำงานในส่วนของคำสั่ง excep แต่ในส่วนของคำสั่ง excep ที่เขียนคือการเกิดข้อผิดพลาดจากคีบอร์ดตัวอักษร e โปรแกรมจะไม่สนใจแล้วผ่านไป และกลับไปทำงานที่ try อีกครั้ง

2.3 โครงข่ายประสาทเทียมเบื้องต้น (Artificial Neural Networks : ANN)

โครงข่ายประสาทเทียมหรือที่เรียกกันว่า Artificial Neural Networks (ANN) เป็นแบบจำลองแบบหนึ่งในแขนงของงานวิจัยทางด้านปัญญาประดิษฐ์หรือ Artificial Intelligence (AI) โดยมีการอ้างอิงมาจากการทำงานของเซลล์ประสาทในสมองของมนุษย์แต่ในโครงข่ายของเซลล์ประสาทในสมองของมนุษย์จริงมีขั้นตอนการทำงานที่ซับซ้อนกว่ามาก แต่กระนั้น ANN ก็ยังสามารถที่จะนำมาใช้ประโยชน์ได้จริงในงานด้านวิทยาการคอมพิวเตอร์และวิศวกรรมโดยทั่วไป

ในสมองของมนุษย์นั้นจะประกอบไปด้วยเซลล์ประสาทหรือ Neuron ประมาณ 10^{11} ตัว โดยที่เซลล์ประสาทหนึ่งๆ จะมีการเชื่อมโยงต่อไปยังเซลล์ประสาทอื่นๆ อีกประมาณ 10^{4} ตัว และเวลาที่ใช้ในการแลกเปลี่ยนข้อมูลระหว่าง Neurons นั้นจะอยู่ในหลัก 10^{-3} วินาที โดยการทางานของ neurons ในสมองมนุษย์นั้นจะมีลักษณะการประมวลผลแบบขนานขั้นสูงอีกด้วย (Highly Parallel Processing) จึงทำให้มนุษย์สามารถทำการเรียนรู้และจดจำหน้าแม่ของตัวเองได้ภายในเวลาเพียงไม่กี่วินาที



รูปที่ 2.9 ส่วนประกอบของเซลล์ประสาทของมนุษย์

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

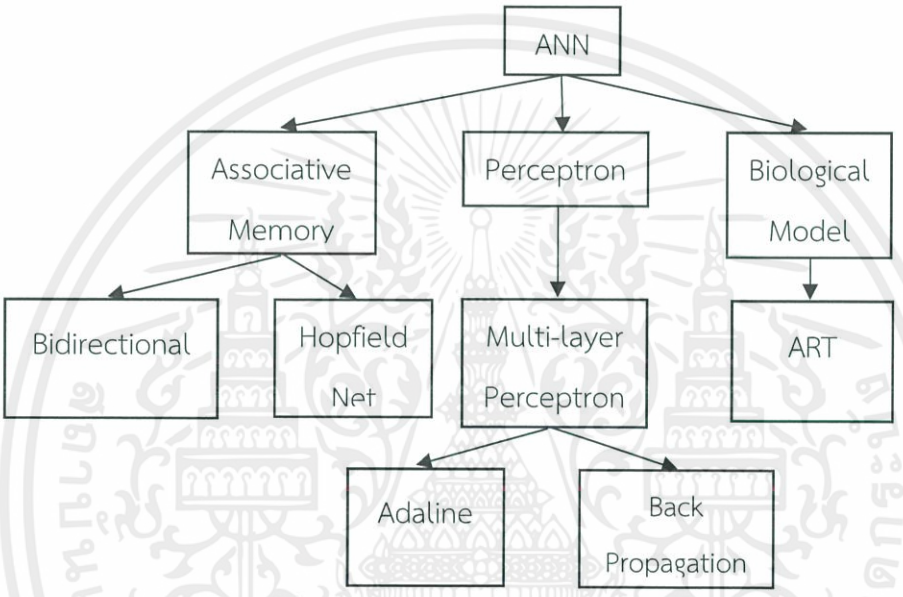
แบบจำลองของ ANN ที่ได้รับความนิยมสามารถจำแนกออกได้เป็น 3 กลุ่มใหญ่ ดังนี้คือ

1) Perceptron : Multi-layer Perceptron

2) Associative Memory : Hopfield Net and Bidirectional

3) Biological Model : ART

โดยสามารถแบ่งย่อยลงไปได้อีกดังในรูปที่ 2.10



รูปที่ 2.10 แบบจำลองหลักๆ ของ ANN

คุณสมบัติและความสามารถที่เป็นประโยชน์ของ ANN

1) มีความยืดหยุ่นสูงจนสามารถจาลองขบวนการของปัญหาใดๆ ได้

2) มีความสามารถในการจาลองชุดของคู่ input-output ที่มีความซับซ้อนมากจนไม่สามารถจาลองแบบในเชิงของความน่าจะเป็นได้

3) มีความสามารถในการปรับตัวเข้ากับการเปลี่ยนแปลงของสิ่งแวดล้อม

4) มีความสามารถในการตอบสนองต่อข้อมูลที่ไม่เคยเห็น

5) ความรู้จะกระจายอยู่ทั่วทั้งโครงสร้างของโครงข่ายประสาทเทียม

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลักษณะของปัญหาที่เหมาะสมกับ ANN

1) ตัวอย่างข้อมูลสาหรับการฝึกหัดประกอบด้วย attribute หลายๆ ค่า

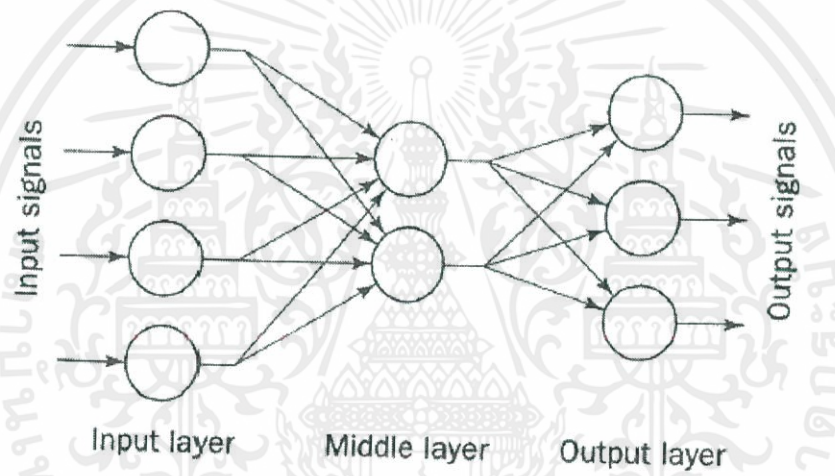
2) Output ที่ต้องการสามารถเป็นค่าจานวนเต็ม หรือจานวนจริง หรือเป็น vector ของจานวนเต็ม หรือจานวนจริงก็ได้

3) ตัวอย่างข้อมูลสามารถมีความผิดพลาดปะปนอยู่ได้

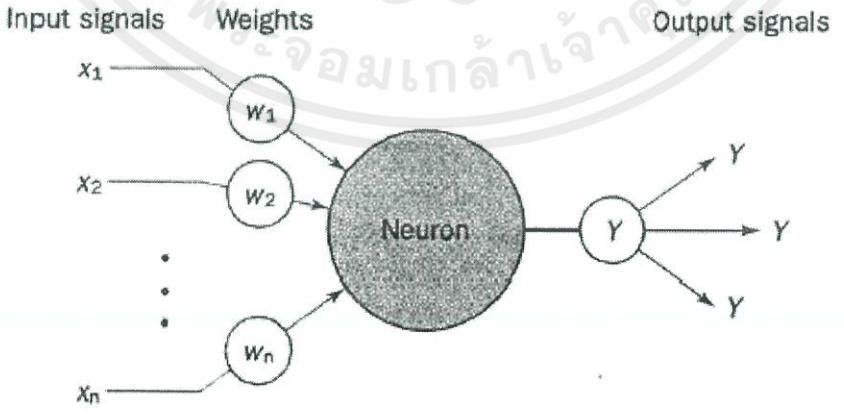
4) ไม่มีปัญหากับขั้นตอนการฝึกหัดซึ่งอาจใช้เวลานานมาก

5) ต้องการประเมินผล output อย่างรวดเร็ว

6) ไม่ต้องการเข้าใจวิธีการแก้ปัญหา ต้องการเพียงแค่ให้ปัญหาได้รับการแก้ไขเท่านั้น



รูปที่ 2.11 ตัวอย่างของโครงข่ายประสาทเทียม



รูปที่ 2.12 บล็อกไดอะแกรมของ Neuron

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ซึ่งตามแบบจำลองที่นำมาเสนอโดย McCullorh and Pitts โดยถือเป็นแบบจำลองของ Neuron ในยุคแรกแสดงไว้ดังในสมการที่ 2.1

$$X = \sum_{i=1}^{n} x_i w_i \tag{2.1}$$

โดยที่ X คือ net weighted input to the neuron

จากนั้นพิจารณาต่อดังนี้ (ซึ่งถือเป็น Transfer Characteristic หรือที่นิยมเรียกว่า Activation Function)ซึ่งแสดงได้ดังสมการที่ 2.2

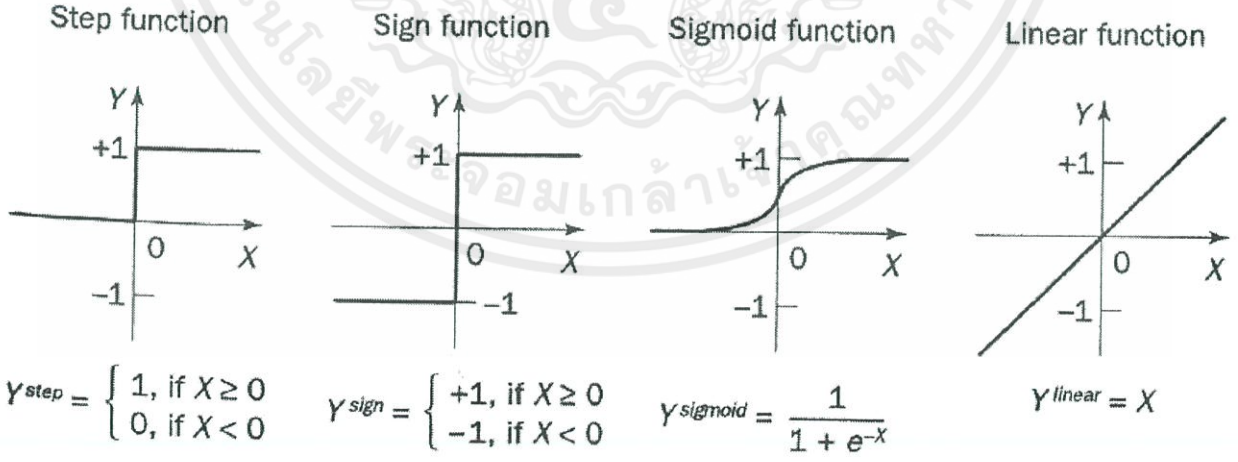
$$Y = \begin{cases} +1, X \geq \theta \\ -1, X < \theta \end{cases} \tag{2.2}$$

โดยที่ θ คือ ค่าเกณฑ์

ซึ่ง Activation Function ดังกล่าวนี้จะเรียกว่า Sign function หรือ Signum function ดังนั้นจึงสามารถแสดง Actual output ของ Neuron ที่ใช้ Sign activation function ได้ดังสมการที่ 2.3

$$Y = sign[\sum_{i=1}^{n} x_i w_i - \theta] \tag{2.3}$$

ซึ่ง Activation Function ในแบบต่างๆ ที่นิยมใช้ในโครงข่ายประสาทเทียมได้ดังรูปที่ 2.13

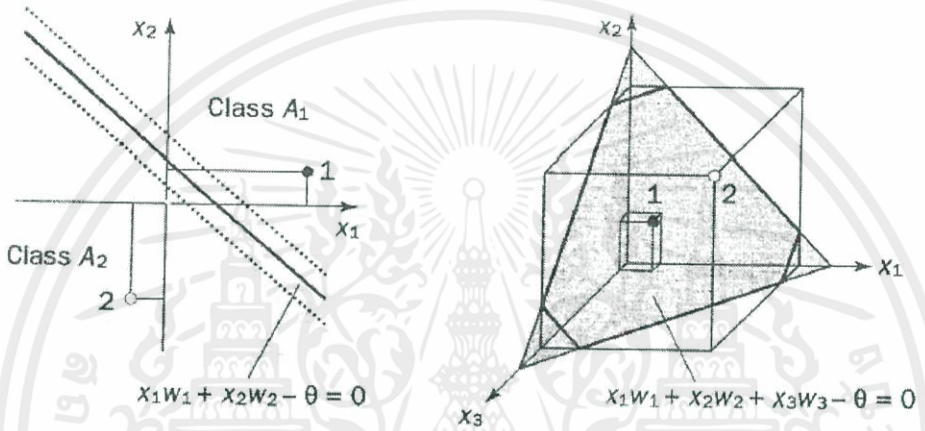


รูปที่ 2.13 ตัวอย่าง Activation Function ของโครงข่ายประสาทเทียม

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยพื้นฐานต่างๆ ที่จำเป็นต้องเรียนรู้คือวิธีการใช้งานที่เป็นเพียง Sigle-Layer ก่อนโดยใช้วิธีการแบบ Linear Peceptron ซึ่งมีทั้งหมด 4 ขั้นตอนซึ่งใน Linear Peceptron นั้นยังมี Hyperplane ที่ถูกใช้ในการแบ่ง Decision Region แสดงได้ดังสมการที่ 2.4 ซึ่งในสมการยังถูกได้ถูกนำมาใช้ในการทำงานแบบ Linear Separability Function ดังรูปที่ 2.14

$$\sum_{i=1}^{n} x_i w_i - \theta \tag{2.4}$$



รูปที่ 2.14 รูปภาพ Linear Separability in the Perceptron (ก) อินพุตสองค่าของ Perceptron (ข) อินพุตสามค่าของ Perceptron

ซึ่งในขั้นตอนสำหรับการเรียนรู้แบบ Linear Perceptron มีทั้งหมด 4 ขั้นตอน ดังนี้

1) Initialization

กำหนดค่า Initial weights $w_1, w_2, w_3, \ldots, w_n$ และค่า Threshold θ ให้มีค่าแบบสุ่มในช่วง [-0.5,0.5]

2) Activation

ทำการ Activate the Perceptron โดยการ Apply inputs $x_1(p), x_2(p), x_3(p), \ldots, x_n(p)$ และ Desired output $Y_d(p)$ จากนั้นทำการคำนวณค่า Actual output ที่ Iteration $p = 1$ โดยใช้สมการที่ 2.5

$$Y = step[\sum_{i=1}^{n} x(p)w(p) - \theta] \tag{2.5}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดย n คือจำนวนของ Perceptron input และ Step หมายถึงการใช้ Activation Function

3) Weight Training

กำหนดค่า Error signal ซึ่งค่าที่ได้จะทำให้รู้ว่าข้อมูลในจุดไหนเกิดการ Error ได้โดยแทนด้วย $e(p)$ ซึ่งแสดงดังสมการที่ 2.6

$$e(p) = Y_d(p) - Y(p) \tag{2.6}$$

โดยที่ $Y_d(p)$ คือค่า Desired output และ $Y(p)$ คือค่า Actual output ที่ได้จากการทดลอง

กำหนดค่า Learning Rate (α) ซึ่งเป็นค่าคงที่ที่มีเป็นค่าบวกและมีขนาดน้อยกว่า 1 โดยค่าที่ทำให้ระบบสามารถทำการรู้จำได้ดีจะต้องมีค่า α มียิ่งเข้าใกล้ศูนย์แล้วทำให้ความสามารถในการรู้จำก็จะยิ่งแม่นยำมากขึ้น แต่ก็ต้องแลกมาด้วยระบบจะมีค่าดีเลย์ที่มากขึ้น โดยสามารถแสดงได้ดังสมการที่ 2.7

$$\Delta w_i(p) = \alpha \times x_i(p) \times e(p) \tag{2.7}$$

โดยค่า $\Delta w_i(p)$ คือค่า Weight correction factor ที่ Iteration p และเสร็จแล้วจึงทำการ Update weights ของ Perceptron ดังสมการที่ 2.8

$$w_i(p+1) = w_i(p) + \Delta w_i(p) \tag{2.8}$$

4) Iteration

ทำการเพิ่ม iteration *p* ไปอีก *1* จากนั้นกลับไปทำในตอนที่ *2* และ repeat process ทั้งหมดจนกระทั่งระบบเข้าสู่สภาวะเสถียร

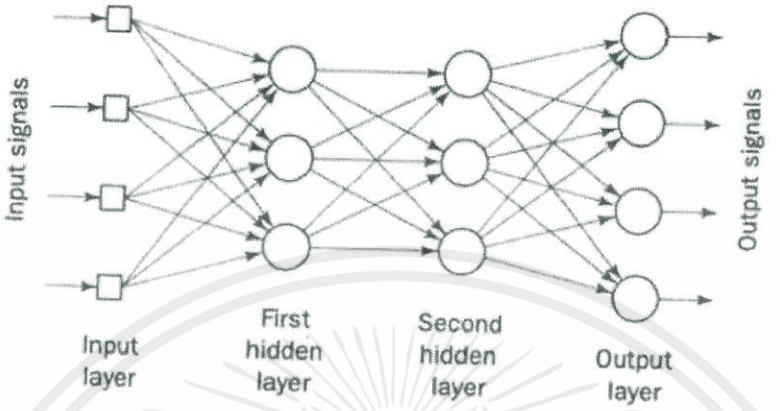
ซึ่งในการเรียนรู้โครงข่ายประสาทเทียมเบื้องต้น (Artificial Neural Networks *: ANN)* ยังมีในส่วนของโครงข่ายประสาทเทียมอื่นๆ ซึ่งจะสามารถดูการทำงานได้ดังนี้

2.3.1 Multilayer Neural Networks

Multilayer Neural Networks หรือ Multilayer perceptron เป็นโครงข่ายประสาทเทียมที่มีการทำงานแบบ feedforward ในการประมวลผลจากอินพุตไปยังเอาท์พุตเพียงอย่างเดียวและในส่วนของ

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เลเยอร์ hidden สามารถมีการทำงานได้ตั้งแต่ 1 หรือ มากกว่าขึ้นไปขึ้นอยู่กับการออกแบบ ซึ่งสามารถดูโครงข่ายประสาทเทียมแบบ Multilayer perceptron ได้ในรูปที่ 2.15



รูปที่ 2.15 โครงข่ายประสาทเทียมแบบ Multilayer perceptron ที่มีเลเยอร์ hidden 2 เลเยอร์

โดยในการทำงานของเลเยอร์ hidden 1 เป็นส่วนในการทำงานระหว่างอินพุตกับระบบที่งานแบบต่อเนื่อง ในส่วนของเลเยอร์ hidden 2 จะเป็นส่วนของการทำงานที่ประมวลผลของฟังก์ชั่นที่มีการทำงานแบบไม่ต่อเนื่องกันขึ้นอยู่ข้อมูลที่ป้อน และในส่วนการทำงานของเลเยอร์ hidden 3 และ 4 จะมีจำนวนของโหนดที่ทำงานตั้งแต่ 10 ถึง 1,000 โหนด โดยในส่วนของการทำงานจริงส่วนมากที่สุดจะใช้เพียง 3 เลเยอร์ เพราะเนื่องมากจากในแต่ละเลเยอร์ที่ทำงานเพิ่มไปจะส่งผลให้ระบบมีการคำนวณที่มากเกินไป ซึ่งในการเรียนรู้ของโครงข่ายมีอยู่มากมายที่ใช้งานแต่ที่นิยมที่สุดคือ Back-Propagation โดยในการใช้งานโครงข่ายแบบ Back-Propagation จะมีการใช้ในส่วนเลเยอร์ hidden 3 หรือ 4 เลเยอร์ในการออกแบบ

ในการออกแบบเอาต์พุตจะมีการออกแบบที่คล้ายคลึงกับโครงข่ายแบบ Perceptron

$$X = \sum_{i=1}^{n} x_i w_i - \theta \tag{2.9}$$

โดยที่ n คือจำนวนของอินพุต และ θ คือเกณฑ์ที่ใช้กับเซลล์ประสาท

ถัดไปเป็นส่วน weighted ของอินพุตที่เปิดใช้งาน โดยส่วนของการเปิดใช้งานของ back propagation network จะใช้ฟังก์ชั่นของ Sigmoid Activation ที่ถูกนำมาใช้งาน

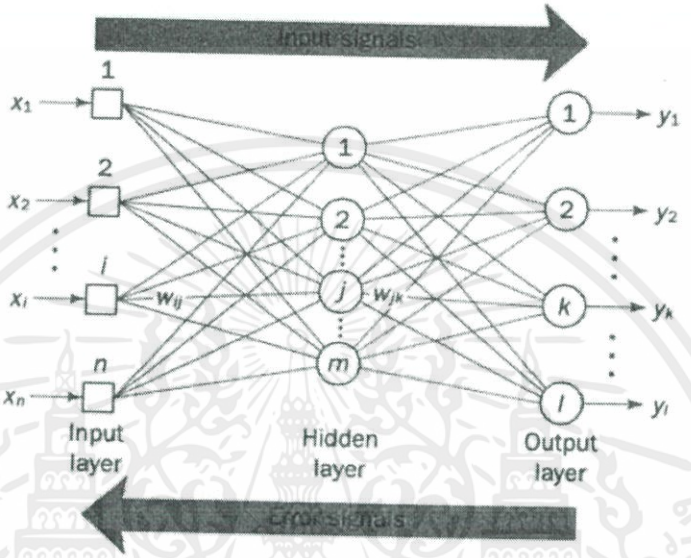
$$Y^{sigmoid} = \frac{1}{1+e^{-X}} \tag{2.10}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

นอกจากนี้ยังรับประกันว่าผลลัพธ์ที่ได้จะมีค่าอยู่ระหว่าง 0 ถึง 1

2.3.2 Back-Propagation Networks

ซึ่งสามารถพิจารณา Back-Propagation Networks ได้ดังรูปที่ 2.16



รูปที่ 2.16 โครงข่ายประสาทเทียมแบบ Back-Propagation 3 เลเยอร์

ซึ่งในรูปที่ 2.6 ค่า I j และ *k* จะอ้างอิงถึงเซลล์ประสาทในอินพุต เลเยอร์ hidden และเอาต์พุตในส่วนจองค่า $x_1, x_2, ..., x_n$ จะเป็นส่วนของการเคลื่อนที่ของข้อมูลจากซ้ายไปขวา และในส่วนของสัญญาณที่ผิดพลาดจะเป็นส่วนของการเคลื่อนที่ของข้อมูลจากขวามาซ้ายเพื่อแก้ไขข้อผิดพลาด

ในส่วนของข้อมูลที่ผิดพลาดจะเริ่มต้นที่เลเยอร์เอาต์พุตและทำการย้อนกลับมายังเลเยอร์แรก ข้อมูลผิดพลาดที่เอาต์พุตของเซลล์ประสาท k ที่ทำซ้ำกับ p ซึ่งสามารถดูสมการได้ดังข้างล่าง

$$e_k(p) = y_{d,k}(p) - y_k(p) \tag{2.11}$$

โดย $y_{d,k}(p)$ คือผลลัพธ์ที่ได้จากเอาต์พุตของเซลล์ประสาท k ที่ทำซ้ำ p

ดังนั้นจะใช้สำหรับการอัปเดต Weights w_{jk} ที่เลเยอร์เอาต์พุต ซึ่งสามารถดูการทำงานได้ดังสมการข้างล่าง

$$w_{jk}(p+1) = w_{jk}(p) + \Delta w_{jk}(p) \tag{2.12}$$

โดย $\Delta w_{jk}(p)$ คือค่าการแก้ไข Weights

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การคำนวณของ Weights ในโครงข่ายแบบหลายชั้น สามารถคำนวณได้ดังสมการข้างล่าง

$$\Delta w_{jk}(p) = \alpha \times y_j(p) \times \delta_k(p) \tag{2.13}$$

โดย $\delta_k(p)$ คือ การไล่ความผิดพลาดของเซลล์ประสาท k ในเลเยอร์เอาต์พุตที่ซ้ำ p

การไล่ระดับสีข้อผิดพลาดจะถูกกำหนดเป็นอนุพันธ์ของฟังก์ชั่นการเปิดใช้งานคูณด้วยข้อผิดพลาดที่เอาต์พุตของเซลล์ประสาท ดังนั้นสำหรับเซลล์ประสาท k ได้ดังสมการข้างล่าง

$$\delta_k(p) = \frac{\partial y_k(p)}{\partial X_k(p)} \times e_k(p) \tag{2.14}$$

โดย $y_k(p)$ คือ เอาต์พุตของซลล์ประสาท k ที่ซ้ำ p และ $X_k(p)$ คือ ค่า Weighted ของเซลล์ประสาทอินพุต k

สำหรับฟังก์ชั่น Sigmoid Activation สามารถได้ดังสมการดังนี้

$$\delta_k(p) = \frac{\partial\left\{\frac{1}{1+\exp[-X_k(p)]}\right\}}{\partial X_k(p)} \times e_k(p) = \frac{\exp[-X_k(p)]}{\{1+\exp[-X_k(p)]\}^2} \times e_k(p) \tag{2.15}$$

หรือ

$$\delta_k(p) = y_k(p) \times [1 - y_k(p)] \times e_k(p) \tag{2.16}$$

เมื่อต้องการคำนวณการแก้ไข Weight สำหรับเลเยอร์ที่ hidden ที่อยู่สามารถใช้สมการได้ดังนี้

$$\Delta w_{ij}(p) = \alpha \times x_i(p) \times \delta_j(p) \tag{2.17}$$

โดย $\delta_j(p)$ คือ การไล่ระดับสีข้อผิดพลาดเซลล์ประสาท j ในชั้นของเลเยอร์ hidden

$$\Delta w_{ij}(p) = \alpha \times x_i(p) \times \delta_j(p) \tag{2.18}$$

โดย j คือ จำนวนของเซลล์ประสาทในเลเยอร์เอาต์พุต

$$y_j(p) = \frac{1}{1+e^{-X_j(p)}}; \qquad X_j(p) = \sum_{i=1}^{n} x_i(p) \times w_{ij}(p) - \theta_j \tag{2.19}$$

โดย n คือจำนวนของเซลล์ประสาทในชั้นอินพุต

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ซึ่งในขั้นตอนสำหรับการเรียนรู้แบบ Back propagationมีทั้งหมด 4 ขั้นตอน ดังนี้

1) Initialization

ตั้งค่า Weight และระดับเกณฑ์ทั้งหมดของเครือข่ายให้เป็นตัวเลขสุ่มกระจายอย่างสม่ำเสมอภายในช่วงเล็กๆ ซึ่งสามารถดูสมการได้ดังสมการด้านล่าง

$$(-\frac{2.4}{F_i},+\frac{2.4}{F_i}) \tag{2.20}$$

โดย F_i คือ จำนวนอินพุตของเซลล์ประสาททั้งหมดในเครือข่าย

2) Activation

เปิดใช้งานเครือข่ายประสาทเทียมหลังการถ่ายทอดข้อมูลโดยใช้อินพุท $x_1, x_2, ..., x_n$ และส่วนของเอาต์พุต $y_{d,1}(p), y_{d,2}(p), ..., y_{d,l}(p)$

คำนวณเอาต์พุตที่เกิดขึ้นจริงของเซลล์ประสาทในเลเยอร์ hidden ได้ดังสมการดังนี้

$$y_j(p) = sigmoid[\sum_{i=1}^{n} x_i(p) \times w_{ij}(p) - \theta_j] \tag{2.21}$$

โดย n คือ จำนวนอินพุตของเซลล์ประสาท j ในเลเยอร์ hidden

คำนวณเอาต์พุตที่เกิดขึ้นจริงของเซลล์ประสาทในเลเยอร์เอาต์พุต

$$y_j(p) = sigmoid[\sum_{i=1}^{n} x_i(p) \times w_{ij}(p) - \theta_j] \tag{2.22}$$

โดย m คือ จำนวนอินพุตของเซลล์ประสาท k ในเลเยอร์เอาต์พุต

3) Weight Training

อัปเดตค่า Weight ในเครือข่ายแบบ *Back propagation* การขยายส่วนของข้อผิดพลาดที่เกี่ยวข้องกับเซลล์ประสาทเอาต์พุต

คำนวณการไล่ระดับสีข้อผิดพลาดของเซลล์ประสาทในเลเยอร์เอาต์พุต

$$\delta_k(p) = y_k \times [1 - y_k(p)] \times e_k(p) \tag{2.23}$$

คำนวณการแก้ไข Weight

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\Delta w_{ij}(p) = \alpha \times y_j(p) \times \delta_k(p) \quad (2.24)$$

อัปเดตน้ำหนักที่เซลล์ประสาทเอาต์พุต

$$w_{jk}(p+1) = w_{jk}(p) + \Delta w_{jk}(p) \quad (2.25)$$

คำนวณการไล่ระดับสีข้อผิดพลาดของเซลล์ประสาทในเลเยอร์ hidden

$$\delta_j(p) = y_j(p) \times [1 - y_j(p)] \times \sum_{k=1}^{l} \delta_k(p) \times w_{jk}(p) \quad (2.26)$$

คำนวณการแก้ไข Weight

$$\Delta w_{ij}(p) = \alpha \times_i (p) \times \delta_j(p) \quad (2.27)$$

อัปเดตน้ำหนักที่เซลล์ประสาทที่เลเยอร์ hidden

$$w_{ij}(p+1) = w_{ij}(p) + \Delta w_{ij}(p) \quad (2.28)$$

4) Iteration

เพิ่มการที่ซ้ำทีละรายการกลับไปที่ขั้นตอนที่ 2 และทำซ้ำขั้นตอนจนกว่าจะถึงเกณฑ์

2.4 ความรู้พื้นฐานเกี่ยวกับ Convolutional Neural Networks

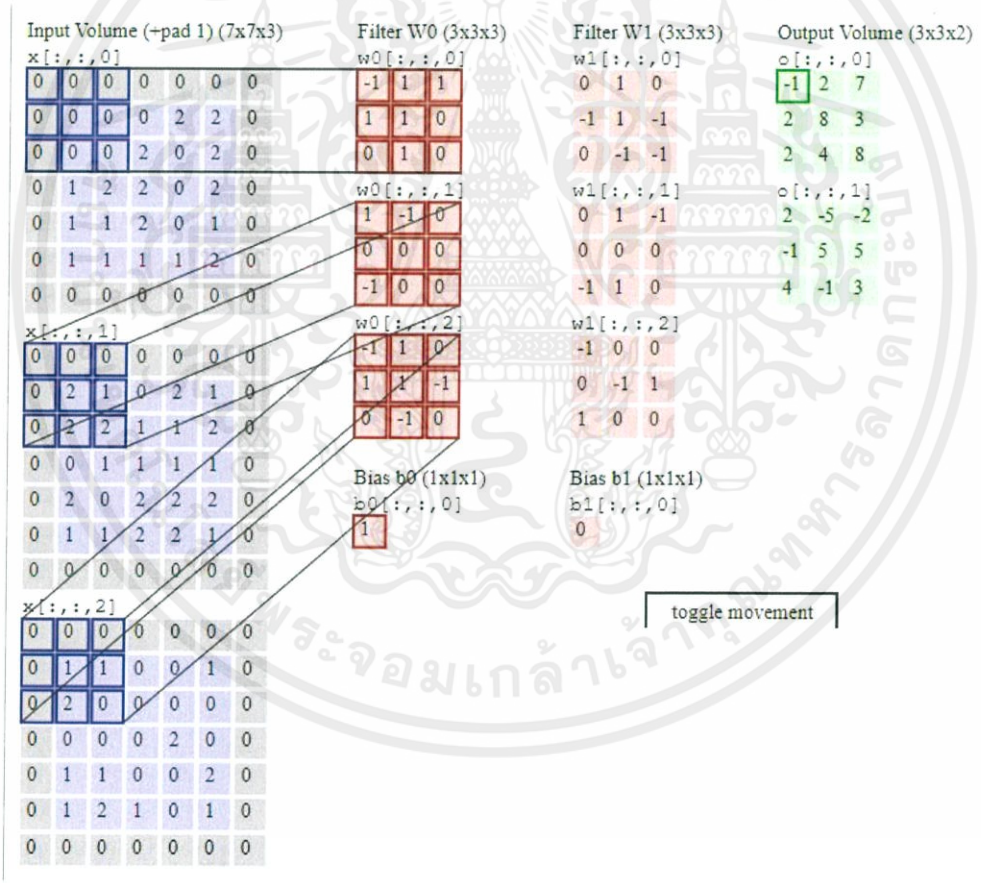
Convolutional Neural Networks (CNNs/ConvNets) เป็นโครงข่ายประสาทเทียมที่ได้นำมาใช้สำหรับในการจำแนกเพศ และสีหน้าของใบหน้าคนออกมา โดยโครงข่ายประสาทเทียมแบบ CNNs มีลักษณะคล้ายคลึงโครงข่ายประสาทเทียมทั่วไป แต่การทำงานจะมีการเพิ่มในส่วนของ Layer ให้มีความซับซ้อนที่มากขึ้น เพื่อที่จะทำให้โครงสร้างในการเรียนรู้มีการรู้จำที่มากขึ้น และมีประสิทธิภาพในการจำแนกได้มากขึ้น

CNNs ที่นำมาใช้งานจะมีการทำงานในการแปลงแต่ละโหนด โดยจะมีการทำงานผ่านฟังก์ชั่น Differentiable ที่จะมีการทำงานในเลเยอร์หลักอยู่สามประเภทเพื่อสร้างโครงข่ายประสาทเทียมแบบ CNNs ได้แก่ เลเยอร์ Convolutional (Conv) เลเยอร์ Pooling และเลเยอร์ Fully-Connected เพื่อใช้สำหรับในการประมวลผล

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4.1 เลเยอร์ Convolutional

Convolutional (Conv) จะประกอบด้วยการทำงานในส่วนของชุดตัวกรองที่สามารถเรียนรู้ได้ด้วยตัวเอง โดยในแต่ละตัวกรองทุกตัวจะมีขนาดเล็กกว่าขนาดของรูปภาพอินพุตที่ป้อนเข้ามา ยกตัวอย่างเช่น ในภาที่ 2.15 รูปภาพอินพุตมีขนาด 7 x 7 x 3 พิกเซล (ความสูงและความกว้าง 7 พิกเซลและความลึกของสี 3 สี) มีตัวกรองขนาด *3* x *3* พิกเซลในเลเยอร์ Conv จะมีค่าน้ำหนักที่เกิดกับขนาดของ *3* x *3* x 3 พิกเซล จะได้ปริมาณน้ำหนักทั้งหมด 27 น้ำหนักในการประมวลผล และในรูปที่ 2.17 จะมีการสร้างของตัวกรองอยู่ 2 ตัวในทำ Conv โดยในการทำงานต่อมาจะมีการเลื่อนของตัวกรองแต่ละตัวไปทั่วความกว้างและความสูงของปริมาณการป้อนข้อมูลและคำนวณค่าแต่ละการเลื่อนของตัวกรองในอินพุตที่ตำแหน่งต่างๆ เพื่อที่จะทำการเก็บค่าคุณลักษณะของรูปภาพเก็บไว้



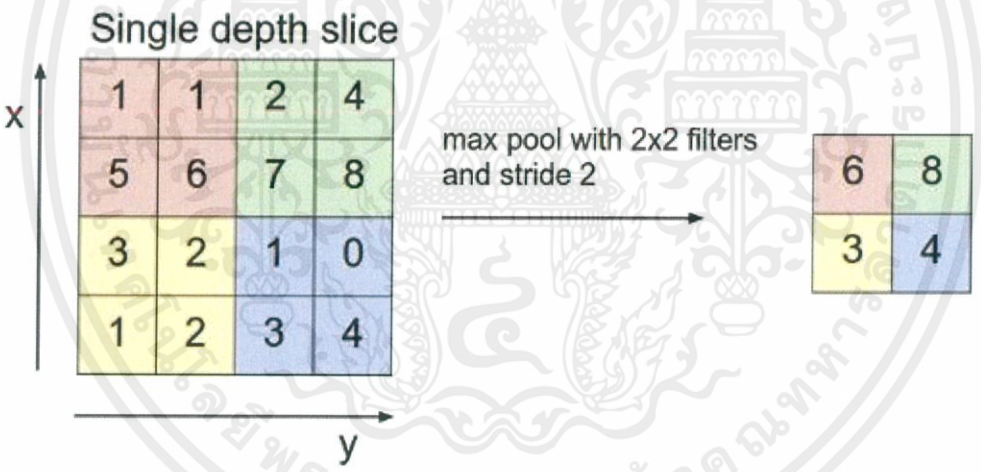
รูปที่ 2.17 การทำงานของเลเยอร์ Convolutional (Conv)

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4.2 เลเยอร์ Pooling

Pooling เป็นเลเยอร์ที่จะต่อจากการทำงานของเลเยอร์ Conv เนื่องจากจะทำหน้าที่ในการลดขนาดเชิงพื้นที่ของข้อมูล เพื่อที่จะลดจำนวนของพารามิเตอร์และการคำนวณในเครือข่ายให้น้อยลง โดยในการทำงานของ Pooling จะทำการปรับขนาดของพื้นที่โดยใช้ค่าสูงสุดของการทำงาน หรือ MaxPooling โดยส่วนใหญ่ที่จะปรับขนาดจะมีการใช้ตัวกรองของ MaxPooling ที่มีการหาค่าสูงสุดของพื้นที่ โดยขนาดที่นำมาใช้ส่วนมากจะเป็นขนาด *2* x 2 พิกเซลในการเลื่อน และในการเลื่อนของการทำ MaxPooling สามารถที่จะกำหนดการเลื่อนของการทำงานได้ว่าจะให้มีการเลื่อนมากน้อยเท่าไร ซึ่งจะสามารถดูการทำงานได้ในรูปที่ 2.18

ในรูปที่ 2.16 จะมีรูปภาพอินพุตขนาด 4 x 4 พิกเซลที่มีการเลื่อนของการทำ MaxPooling ขนาด 2 x 2 พิกเซลและระยะการเลื่อนอยู่ที่ 2 ช่อง ในด้านขวาคือขนาดของพิกเซลที่มีขนาดลดลงจากเดิม 4 เท่า



รูปที่ 2.18 การทำงานของเลเยอร์ Pooling

2.4.3 เลเยอร์ Fully-Connected

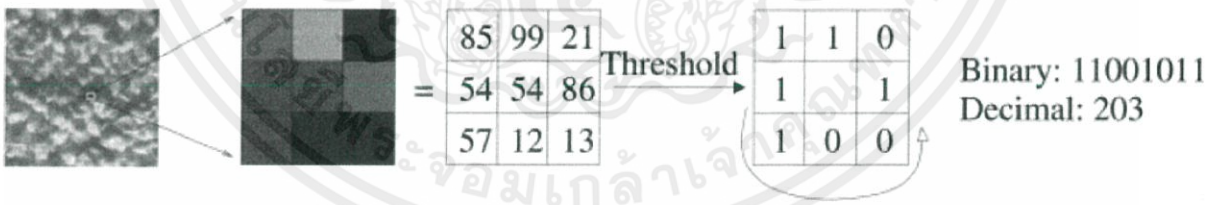
Fully-Connected (FC) เป็นส่วนที่สำคัญส่วนสุดท้ายที่ใช้สำหรับในการเชื่อมต่อแต่ละพิกเซลเข้าด้วยกันทั้งหมดของพิกเซลก่อนหน้า เพื่อที่จะคำนวณหาค่าการคูนของเมทริกซ์และทำการจำแนกชนิดของวัถุออกมา

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5 ความรู้พื้นฐานเกี่ยวกับ Local Binary Pattern

Local Binary Pattern (LBP) เป็นเทคนิคที่ได้นำมาใช้ในการดึงลักษณะของรูปภาพใบหน้าคน เนื่องจากความสามารถในการจำแนกและความเรียบง่ายในการคำนวณ ดั้งนั้นจึงทำการนำ LBP มาใช้งานในการพัฒนาการตรวจจับและนับจำนวนคน

LBP เป็นเทคนิคที่สำคัญที่ได้ถูกนำมาใช้ในการตรวจจับและนับจำนวนคนที่ผ่านบริเวณจุดป้ายโฆษณาดิจิตอลเนื่องจาก Local Binary Pattern เป็นเทคนิคที่มีประสิทธิภาพที่เหมาะสมมากสำหรับในการตรวจจับใบหน้าคน โดยหลักการทำงานพื้นฐานของ Local Binary Pattern จะมีการทำงานดังรูปที่ 2.17 ซึ่งในการทำงานของรูปภาพจะมีการนำพื้นที่ขนาด 3x3 พิกเซลจากรูปภาพมาทำการแปลงสีจาก Red Green Blue (RGB) เป็น Grayscale จากนั้นจะทำการตั้งค่า threshold ของรูปภาพจากพิกเซลตรงกลางของพื้นที่ขนาด 3x3 พิกเซลโดยในพื้นที่ตรงกลางของรูปภาพนี้คือค่า 54 จากนั้นจะทำการใช้ค่า threshold ที่ได้มาทำการเปรียบเทียบกับแต่ละพิกเซลในพื้นที่รอบตรงกลางให้อยู่ในค่าตัวเลขฐาน 2 โดยถ้าพิกเซลรอบข้างตรงกลางมีค่าน้อยกว่าค่า threshold ที่กำหนดไว้จะทำการแปลงค่าที่ได้เป็นค่าไบนารี 0 แต่ถ้าพิกเซลรอบข้างตรงกลางมีค่ามากกว่าค่า threshold ที่กดไว้จะทำการแปลงค่าที่ได้เป็นเป็นค่าไบนารี 1 จากนั้นจะทำการรวมค่าไบนารีที่ได้ทั้งหมดจากการที่ทำการเปรียบเทียบจากค่า threshold โดยจะทำการรวมโดยใช้พิกเซลข้างซ้ายของค่า threshold เป็นพิกเซลตัวแรกจากนั้นจะเก็บค่าตัวต่อไปโดยการมุมทวนนาฬิกา พอได้ค่าไบนารีออกมาต่อไปจะทำการแปลงค่าไบนารีที่อยู่ในตัวเลขฐาน 2 มาอยู่ในตัวเลขฐาน 10 ซึ่งเหตุผลนี้จึงทำให้เรียกว่า Local Binary Pattern หรือ LBP-codes



รูปที่ 2.19 การทำงานของ Local Binary Patterns (LBP)

จากกระบวนการทำงานข้างต้นที่ได้กล่าวมาในตอนต้นสามารถที่จะนำมาเขียนออกมาเป็นสมการโดยในขั้นตอนที่ต้องนำค่า threshold มาทำการเปรียบเทียบกับแต่ละพิกเซลในพื้นที่รอบตรงกลางแล้วอยู่ในรูปไบนารี ได้ดังสมการ 2.29

$$LBP(x_c, y_c) = \sum_{n=0}^{N-1} f(i_n - i_c)2^n \qquad (2.29)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ i_c คือค่าของความเข้มสีของพิกเซลตำแหน่งศูนย์กลาง i_n คือค่าความเข้มสีของของพิกเซลรอบตำแหน่งตรงกลางของพื้นที่ (x_c, y_c) n คือจำนวนของพิกเซลรอบตำแหน่งตรงกลางของพื้นที่ โดยจะมีฟังก์ชั่น $f(x)$ เป็นตัวกำหนดค่าไบนารี่ที่ได้ดังแสดงในสมการที่ 2.30

$$f(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} \quad (2.30)$$

หลังจากที่ได้ค่าไบนารีออกมาต่อไปจะทำการแปลงค่าไบนารีที่อยู่ในตัวเลขฐาน 2 มาอยู่ในตัวเลขฐาน 10 เพื่อทำการเก็บรูปแบบของรูปภาพไว้ในฐานข้อมูล

2.6 ความรู้พื้นฐานของรูปภาพ

ความรู้พื้นฐานของรูปภาพเป็นส่วนหนึ่งที่สำคัญในการทำการพัฒนาการประมวลผลและวิเคราะห์ภาพสำหรับงานสื่อโฆษณาดิจิตอล โดยในการพัฒนาการตรวจจับและนับจำนวนคนจำเป็นที่จะต้องเข้าใจในความรู้พื้นฐานของรูปภาพ เช่น พิกเซลในรูปภาพที่แสดงออกเป็นข้อมูลอาร์เรย์ สี RGB Grayscales และอัตราส่วนของรูปภาพ เป็นต้น

2.6.1 พิกเซลของรูปภาพ

พิกเซลของรูปภาพเป็นส่วนหนึ่งที่อยู่ข้างในรูปภาพของทุกรูปภาพที่ทำให้เกิดเป็นรูปภาพ โดยปกติแล้วพิกเซลจะเป็นสี่เหลี่ยมที่มีขนาด 1x1 และพิกเซลถือว่าเป็นข้อมูลชนิด สี หรือ ความเข้มของแสงที่ปรากฏในตำแหน่งที่กำหนดในรูปภาพ ซึ่งในรูปที่ 2.20 จะเป็นรูปภาพที่มีความละเอียดขนาด 1,000 x 750 พิกเซล หมายความว่ารูปภาพทด้านล่างจะเป็นขนาด 2 มิติที่เกิดขึ้นจะมีมิติของความกว้าง 1,000 พิกเซล และมิติของความสูง 750 พิกเซล โดยพิกเซลรวมของรูปภาพจะมีขนาดเท่ากับ 1,000 x 750 = 750,000 พิกเซลทั้งหมดในรูปภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 2.20 รูปภาพความละเอียดขนาด 1,000 x 750 พิกเซล

พิกเซลของรูปภาพส่วนใหญ่ที่นิยมใช้สำหรับในการพัฒนาการประมวลผลจะมีอยู่ 2 แบบ

1) พิกเซลสีแบบ Grayscale หรือ Single channel

2) พิกเซลสีแบบ Red Green Blue Color (RGB)

2.6.1.1 พิกเซลสีแบบ Grayscale หรือ Single channel

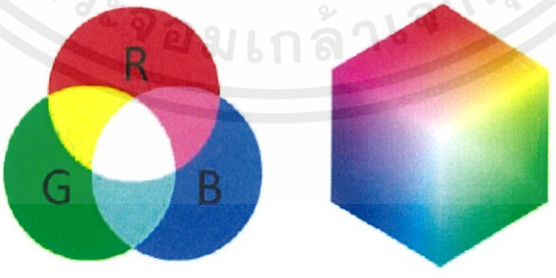
พิกเซลสีแบบ Grayscale หรือ Single channel จะเป็นพิกเซลสีที่ทำให้รูปภาพเป็นสีแบบขาวดำ โดยการกำหนดค่าในแต่ละพิกเซลของรูปภาพให้มีค่าอยู่ระหว่าง 0 ถึง 255 แบ่งออกเป็น 256 ค่า โดยถ้าพิกเซลของรูปภาพมีค่าเป็น 0 จะส่งผลให้สีที่ออกมาเป็นสีดำ แต่ถ้าพิกเซลของรูปภาพมีค่าเป็น 255 จะส่งผลให้สีที่ออกมาเป็นสีขาว ส่วนพิกเซลของรูปภาพที่อยู่ระหว่าง 0 ถึง 255 จะมีเฉดสีที่แตกต่างกันออกไป โดยที่พิกเซลของรูปภาพที่มีค่าใกล้เคียงกับค่า 0 ก็จะมืดลงจากสีขาวนิดหน่อย ส่วนค่าที่ใกล้เคียง 255 ก็จะสว่างขึ้นจากสีดำนิดหน่อย โดยในรูปภาพไล่โทนระดับสีแบบ Grayscale จะแสดงดังรูปที่ 2.21 ซึ่งจะเห็นพิกเซลของรูปภาพที่เป็นสีดำทางด้านซ้ายมือจากนั้นจะค่อยๆสว่างขึ้นไปด้านขวามือ โดยในการพัฒนาการตรวจจับและนับจำนวนคนจะทำการนำค่าการไล่ระดับสีที่ได้มาทำการแทนด้วยตัวเลขฐาน 2 จำนวน 8 บิตมาใช้งาน เพื่อนำพิกเซลสีแบบ Grayscale ไปช่วยให้ระบบประมวลผลเร็วขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 2.21 การไล่โทนระดับสีแบบ Grayscale จากค่า สีดำ (0) ไปถึง สีขาว (255)

2.6.1.2 พิกเซลสีแบบ Red Green Blue Color (RGB)

พิกเซลสีแบบ Red Green Blue Color (RGB) จะเป็นพิกเซลสีที่ทำให้รูปภาพสีแบบปกติ โดยในพิกเซลสีแบบ RGB จะไม่เหมือนกับการไล่โทนระดับสีแบบ Grayscale หรือ Single channel ที่มีแต่การไล่ระดับสีดำไปถึงสีขาว แต่พิกเซลสีแบบ RGB จะมีการแทนสีด้วยค่า 3 สี ได้แก่ สีแดง สีเขียว และสีน้ำเงิน เพื่อในการกำหนดสีในรูปแบบสี RGB ซึ่งในแต่ละสีของสีแดง สีเขียว และสีน้ำเงินจะมีค่าที่กำหนดไว้อยู่ระหว่าง 0 ถึง 255 โดยการไล่โทนสีของพิกเซลทั้ง 3 สีจะมีการไล่โทนสีทั้งหมด 256 ค่า โดยถ้าพิกเซลของรูปภาพมีค่าเป็น 0 จะส่งผลให้สีที่ออกมาทั้ง 3 สีเป็นสีดำ ส่วนถ้าค่าพิกเซลของรูปภาพมีค่าเป็น 255 จะส่งผลให้สีที่ออกมาทั้ง 3 สีเป็นสีเฉพาะของตัวเองคือ สีแดง สีเขียว และสีน้ำเงิน โดยในการพัฒนาการตรวจจับและนับจำนวนคนจะทำการนำค่าการไล่ระดับสีทั้ง 3 สีที่ได้มาทำการแทนด้วยตัวเลขฐาน 2 จำนวน 8 บิตรวมเป็นชุดข้อมูล RGB ในรูปแบบ (R,G,B) ซึ่งจะแสดงถึงสีที่กำหนดในแต่ละพื้น เช่น ถ้ากำหนดชุดข้อมูล RGB เป็น (255,0,0) จะส่งผลให้สีที่ออกมาเป็นสีแดง แต่ถ้ากำหนดชุดข้อมูล RGB เป็น (0,255,0) จะส่งผลให้สีที่ออกมาเป็นสีเขียว และถ้ากำหนดชุดข้อมูล RGB (255,255,255) จะส่งผลให้สีที่ออกมาเป็นสีขาว สามารถแสดงความสัมพันธ์ของสีแบบ RGB ได้ดังรูปที่ 2.20 (ก) และพื้นที่สีแบบ RGB สามารถที่จะมองอยู่ในรูปแบบลูกบาศก์ได้ดังรูปที่ 2.22 (ข) ซึ่งในชุดข้อมูล RGB ถูกกำหนดให้เป็นข้อมูล 3 ค่าซึ่งแต่ละค่าสามารถมีช่วงสีได้ตั้งแต่ 0 ถึง 255 จึงส่งผลทำให้สีที่เกิดขึ้นจากชุดข้อมูล RGB ใน 1x1 พิกเซลจะมีจำนวนสีทั้งหมด 256 x 256 x 256 = 16,777,216 สี



(ก) (ข)

รูปที่ 2.22 (ก) พื้นที่สี Red Green Blue Color (RGB) และ
(ข) ลูกบาศก์ Red Green Blue Color (RGB)

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ส่วนในการพัฒนาการตรวจจับและนับจำนวนคนสำหรับงานสื่อโฆษณาดิจิตอลส่วนที่ใช้งานคือพิกเซลสีแบบ Grayscale หรือ Single channel แทนที่การใช้พิกเซลสีแบบ Red Green Blue Color (RGB) เพราะเนื่องมาจากข้อบกพร่องหลักของพิกเซลสีแบบ RGB ที่มีสีที่หลากหลายถึง 16,777,216 สี ส่งผลให้การพัฒนาการทำงานหนักและทำงานได้ไม่ต่อเนื่อง จึงทำให้การใช้พิกเซลสีแบบ RGB มาแสดงในส่วนของการแสดงผลให้มนุษย์เห็นแล้วเข้าใจง่าย แต่ในส่วนของการประมวลผลที่เกิดขึ้นจะใช้พิกเซลส์แบบ Grayscale หรือ Single channel ในการประมวลผล เพราะปริมาณข้อมูลที่ใช้จะน้อยกว่าพิกเซลสีแบบ RGB จึงทำพิกเซลสีแบบ Grayscale มีการประมวลผลเร็วกว่า

2.6.2 การสร้างรูปภาพ

การสร้างรูปภาพจากในหัวข้อที่ 2.6.1.2 ที่ได้กล่าวถึงพิกเซลของรูปภาพในรูปภาพหนึ่งจะประกอบไปด้วยชุดข้อมูล RGB ที่จะแสดงค่าสีทั้ง 3 สีได้แก่ สีแดง สีเขียว และสีน้ำเงินตามลำดับ ซึ่งในการสร้างรูปภาพจำเป็นจะต้องใช้ชุดข้อมูล RGB ที่ประกอบด้วยความกว้าง (W) ความสูง (H) ในแต่ละองค์ประกอบของชุดข้อมูล RGB มาทำการสร้างรูปภาพ ซึ่งแสดงได้ดังรูปที่ 2.23 ที่จะมีการนำสีทั้ง 3 สีมารวมกันเป็นรูปภาพเดี่ยวกันที่มีความหลากหลายของอาร์เรย์ของชุดข้อมูล RGB ที่มีรูปร่างขนาด W x H x D โดย D คือความลึก หรือ จำนวนของชุดข้อมูล RGB ที่เกิดขึ้น

รูปที่ 2.23 รูปภาพในพื้นที่สี RGB ที่แต่ละสีมารวมกันเป็นรูปภาพ 3 มิติ

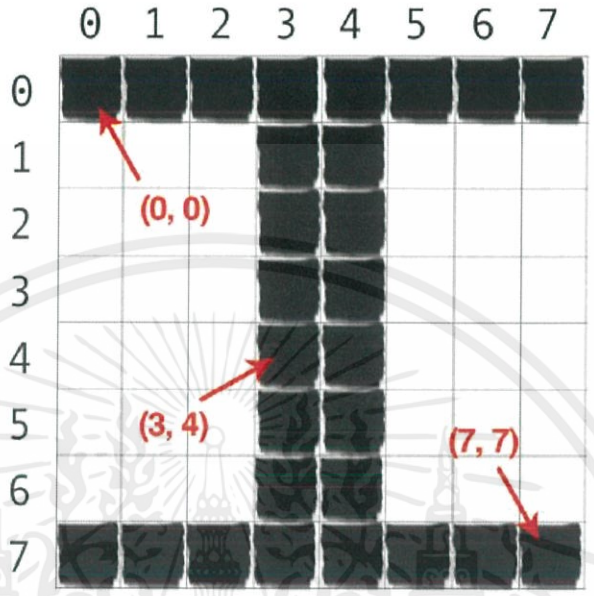
2.6.3 ระบบพิกัดของรูปภาพ

ระบบพิกัดของรูปภาพเป็นอีกส่วนหนึ่งใช้ในการติดตามแล้วหาพิกัดของใบหน้าคนและบันทึกข้อมูลใบหน้าคนของการตรวจจับ จากที่ได้กล่าวไว้ในหัวข้อที่ 2.6.1 ที่เกี่ยวข้องกับพิกเซลของรูปภาพ เพื่อให้ทำความเข้าใจที่ชัดเจนยิ่งขึ้นให้นึกถึงรูปภาพที่มีเส้นของตารางกราฟที่ทับอยู่ โดยจุดเริ่มต้นพิกัดของรูปภาพจะอยู่ที่ตำแหน่งที่มุมซ้ายบนของรูปภาพ ส่วนในการจะเลื่อนไปทางขวาและการเลื่อนลงจะเป็นการเพิ่มของค่า x และค่า y ตามลำดับในแนวแกน

ในรูปที่ 2.24 จะแสดงรูปภาพของกระดาษกราฟที่มีตัวอักษร I บนกระดาษกราฟที่มีขนาดตาราง 8 x 8 ที่จำนวนพิกเซลทั้งหมด 64 พิกเซลจะมีการชี้พิกัดของรูปภาพออกมาทั้ง 3 พิกัดได้แก่ พิกัด (0,0)

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

พิกัด (3,4) และพิกัด (7,7) โดยถ้าทำการพิจารณาที่พิกัดที่ (3,4) จะมีการเลื่อนไปทางด้านขวา 4 ช่อง และมีการเลื่อนไปทางด้านล่าง 5 ช่อง โดยจุดเริ่มต้นของพิกเซลจะที่อยู่จุด (0,0)



รูปที่ 2.24 รูปภาพตัวอักษร I ที่วางอยู่บนกระดาษกราฟ

2.6.4 การปรับขนาดและอัตราส่วนของรูปภาพ

การปรับขนาดและอัตราส่วนภาพจะถูกนำมาใช้งานในการพัฒนา เพื่อลดขนาดของรูปภาพให้เล็กลงและไม่ส่งผลต่อรูปลักษณะของรูป โดยในกระบวนการปรับขนาดจะมีทั้งการเพิ่มหรือการลดขนาดของรูปภาพโดยใช้ความกว้าง (W) และความสูง (H) ในการปรับขนาด โดยการปรับขนาดจะมีทั้งการปรับขนาดที่ให้อัตราส่วนที่เหมือนเดิมกับการปรับขนาดที่ให้อัตราส่วนที่เปลี่ยนแปลงไปจนทำให้รูปภาพถูกบีบอัดและเกิดการบิดเบี้ยวดังรูปที่ 2.25

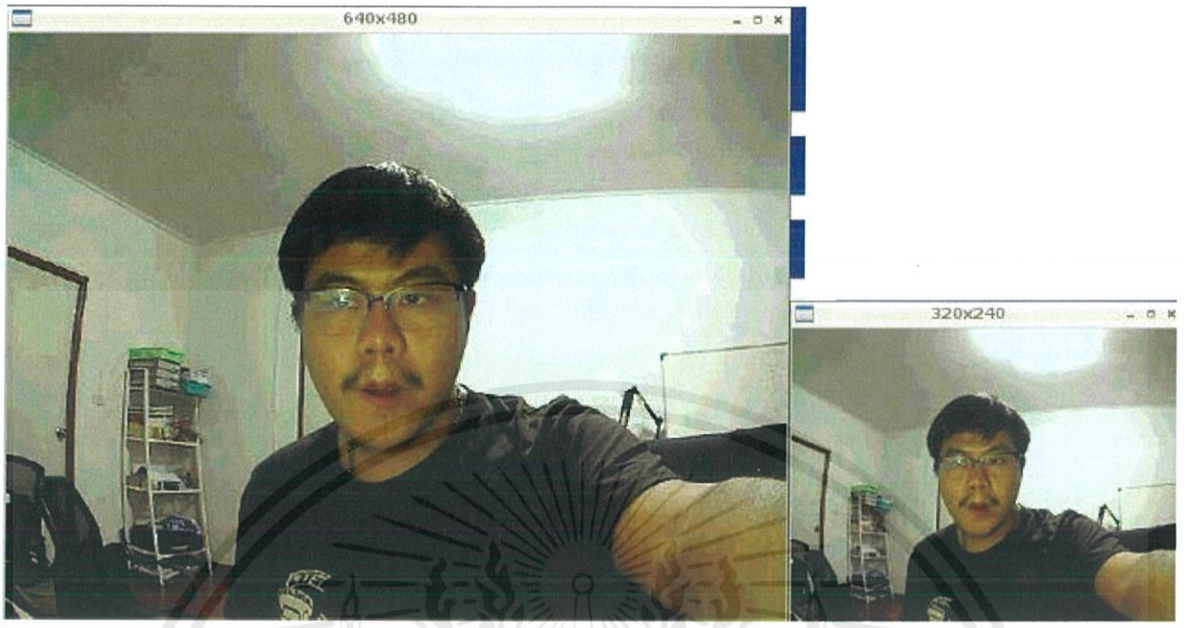
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 2.25 รูปด้านซ้ายเป็นรูปภาพต้นฉบับ และรูปด้านบนและด้านล่างเป็นรูปภาพที่ปรับขนาดความกว้าง (W) และความสูง (H)

โดยในรูปที่ 2.25 รูปภาพด้านซ้ายบนจะเป็นรูปภาพต้นฉบับที่ไม่มีการปรับขนาดความกว้าง (W) และความสูง (H) ส่วนรูปภาพด้านขวาบนและรูปด้านล่างขวาเป็นรูปภาพที่มีการปรับขนาดความกว้าง (W) และความสูง (H) ซึ่งผลที่ได้คือในรูปภาพด้านบนขวาและด้านล่างขวาจะมีการถูกบีบอัดและบิดเบี้ยว เพื่อจะทำการป้องกันไม่ให้เกิดการถูกบีบอัดและบิดเบี้ยวในรูปภาพที่ถูกนำมาใช้งานในการพัฒนาการประมวลผลและวิเคราะห์ภาพสำหรับงานสื่อโฆษณาดิจิตอล จึงจำเป็นที่จะต้องปรับขนาดความกว้าง (W) และความสูง (H) ให้มีอัตราส่วนของรูปภาพเท่าเดิมกับรูปภาพต้นฉบับ

โดยในการพัฒนาการประมวลผลและวิเคราะห์ภาพสำหรับงานสื่อโฆษณาดิจิตอลจะมีการลดรูปภาพให้อัตราส่วนของรูปภาพจากขนาดความกว้าง 640 พิกเซล เป็น 320 พิกเซล และขนาดความสูง 480 พิกเซล เป็น 240 พิกเซลมาใช้งาน เนื่องจากที่ต้องทำการปรับขนาดให้เล็กลงเพราะขนาดของรูปภาพที่ 640 x 480 พิกเซลจะมีปริมาณพิกเซลทั้งหมด 307,200 พิกเซล แต่ขนาดของรูปภาพที่ 320 x 240 พิกเซลจะมีปริมาณพิกเซลทั้งหมด 76,800 พิกเซลจะส่งผลให้ระบบการทำงานในการตรวจจับและนับจำนวนคนมีการประมวลผลลดลงถึง 4 เท่า โดยสามารถแสดงการลดขนาดของรูปภาพได้ดังรูปที่ 2.26

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.26 รูปด้านซ้าย รูปภาพต้นฉบับจากกล้อง รูปด้านขวา รูปภาพที่ปรับขนาด ความกว้าง (W) และความสูง (H)

2.7 ความรู้พื้นฐานเกี่ยวกับการพัฒนาระบบประมวลผล

ความรู้พื้นฐานเกี่ยวกับการพัฒนาจะมีส่วนประกอบด้วยกันทั้งหมด 2 ส่วนได้แก่ ชนิดของการเรียนรู้ในการสอนระบบให้มีการรู้จำ และการเตรียมฐานข้อมูลรูปภาพให้มีประสิทธิภาพแม่นยำ สามารถดูได้ดังนี้

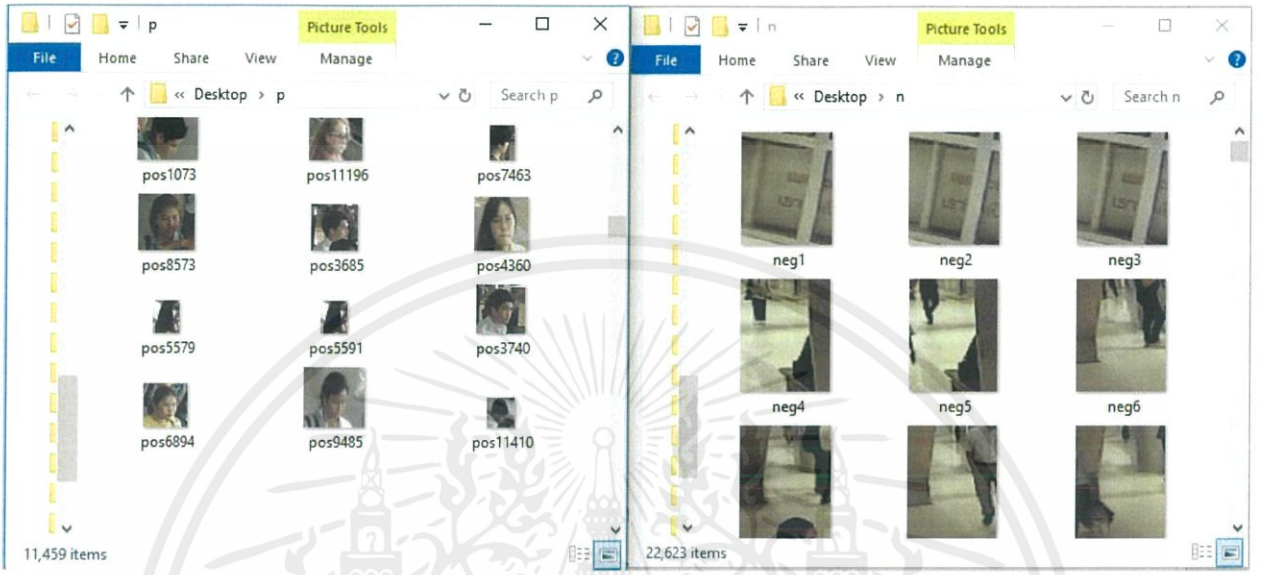
2.7.1 ชนิดของการเรียนรู้ (Learning)

ชนิดของการเรียนรู้ (Learning) จะเป็นส่วนในการสอนระบบให้มีการรู้จำ โดยจะประกอบด้วยชนิดของการเรียนรู้ 3 แบบ ได้แก่ การเรียนรู้แบบมีคนสอน (Supervised Learning) การเรียนรู้แบบไม่มีคนสอน (Unsupervised Learning) และการเรียนรู้แบบกึ่งคนสอน (Semi-supervised Learning) แต่สำหรับในการทำการพัฒนาตรวจจับและนับจำนวนคนที่ได้นำมาใช้งาน คือ การเรียนรู้แบบมีคนสอน (Supervised Learning) เนื่องจากการเรียนรู้แบบมีคนสอนเป็นรูปแบบที่ทำได้ง่ายและสะดวกต่อการใช้งาน

การเรียนรู้แบบมีคนสอน (Supervised Learning) เป็นการเรียนรู้ที่มีการสอนจากคนสอน ซึ่งในการเรียนรู้แบบมีคนสอน จำเป็นจะต้องมีการแยกระหว่างรูปภาพใบหน้าคนและรูปภาพที่ไม่ใช่ใบหน้าคนออกเป็น 2 ส่วนดังรูปที่ 2.27 ที่แยกข้อมูลรูปภาพออกเป็น 2 ส่วนได้แก่ รูปด้านซ้ายเป็นข้อมูลรูปภาพใบหน้า

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คน และรูปด้านขวาเป็นรูปภาพที่ไม่ใช่ใบหน้าคน เพื่อที่จะนำข้อมูลทั้งสองไปทำการสอนระบบให้ระบบมีการรู้จำ



รูปที่ 2.27 รูปด้านซ้าย คือ ฐานข้อมูลรูปภาพใบหน้าคน และ
รูปด้านขาว คือ ฐานข้อมูลที่ไม่ใช่ใบหน้าคน

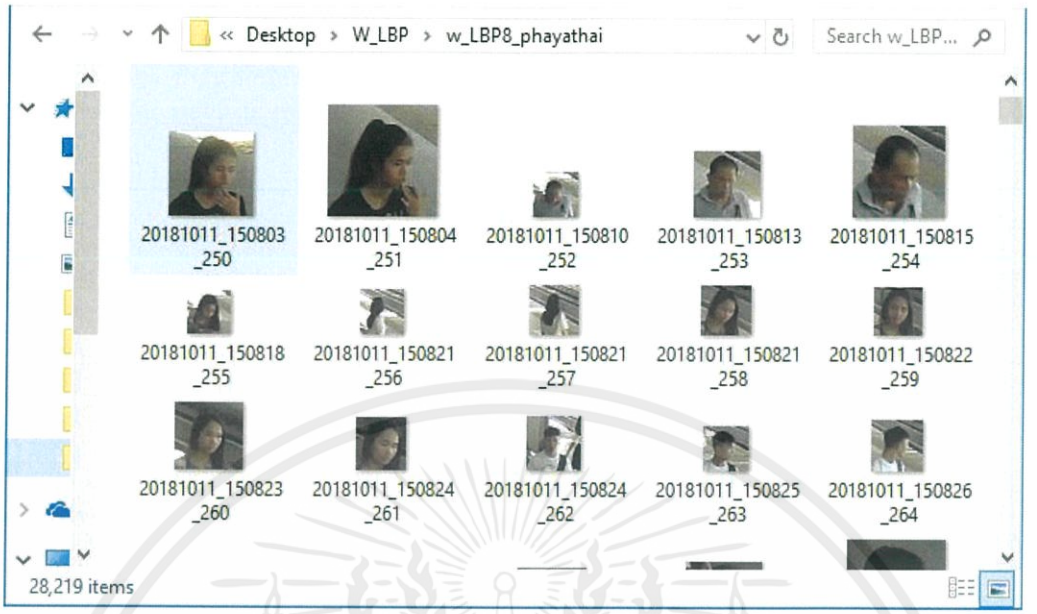
2.7.2 การเตรียมฐานข้อมูลรูปภาพ

การเตรียมฐานข้อมูลรูปภาพเป็นอีกส่วนที่สำคัญสำหรับในการพัฒนาการตรวจจับและนับจำนวนคน เพราะในการรู้จำใบหน้าคนจะต้องมีการเตรียมฐานข้อมูลรูปภาพใบหน้าคนและรูปภาพที่ไม่ใช่ใบหน้าคนเพื่อที่จะนำข้อมูลที่ได้ไปทำการ Training ให้ระบบมีการตรวจจับที่มีความแม่นยำและการตรวจจับจำนวนคนที่มากที่สุด โดยจะมีการเทคนิคในการเตรียมฐานข้อมูลอยู่ทั้งหมด 4 เทคนิคดังนี้

2.7.2.1 การจับคู่ภาพมุมมองจริง

ระบบรู้จำใบหน้าที่มีอยู่ส่วนใหญ่สามารถทำงานได้ดีกับลักษณะใบหน้าที่เป็นหน้าตรง แต่ในทางปฏิบัติภาพใบหน้าที่ต้องการให้ระบบตรวจจับและนับจำนวนคนที่ผ่านบริเวณจุดป้ายโฆษณาดิจิตอลที่สถานีส่วนใหญ่จะเป็นรูปภาพในลักษณะที่หลายมุมมอง ส่งผลให้การตรวจจับและนับจำนวนคนไม่มีประสิทธิภาพ จากนั้นจึงได้นำหลักการ Real View-based Matching หรือ การจับคู่ภาพมุมมองจริง ซึ่งเป็นวิธีที่นำเอาฐานข้อมูลใบหน้าจากสถานที่จริงของแต่ละบุคคลมาใช้แทนฐานข้อมูลรูปภาพแบบเดิมตามภาพหที่ 2.28 ซึ่งส่งผลให้มีประสิทธิภาพในการรู้จำเพิ่มมากขึ้น เนื่องจากมีปริมาณฐานข้อมูลที่นำมาสอนระบบมากขึ้นและเป็นฐานข้อมูลที่ตรงกับสภาพแวดล้อมจริงลองระบบ

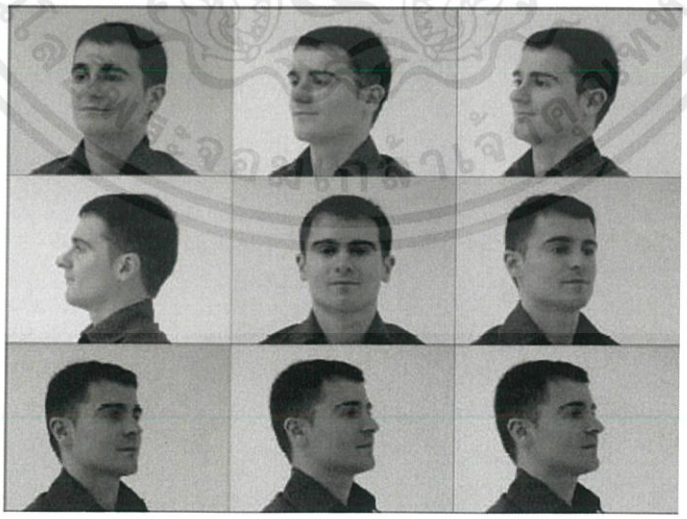
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.28 ใบหน้าคนจากมุมมองจริงของฐานข้อมูลใบหน้าคน

2.7.2.2 การรู้จำใบหน้าแบบหลายมุมมอง

การรู้จำใบหน้าแบบหลายมุมมอง คือ เทคนิคที่ใช้ในการแก้ปัญหาภาพใบหน้าแบบหลายมุมมองอย่างหนึ่งโดยจัดเป็นวิธีที่ง่ายและตรงไปตรงมาหลักการพื้นฐานคือ การเพิ่มฐานข้อมูลที่เป็นแบบมุมมองจริงให้มีความหลากหลายในแต่ละบุคคล เช่นฐานข้อมูลใบหน้าที่มีองศาการหมุนของใบหน้าตั้งแต่ 0,±30, ±45, ±60 และ ±90 องศา ตามรูปที่ 2.29 ทำให้ประสิทธิภาพการรู้จำเพิ่มมากขึ้น เนื่องจากมีปริมาณฐานข้อมูลที่นำมาสอนระบบมากขึ้น

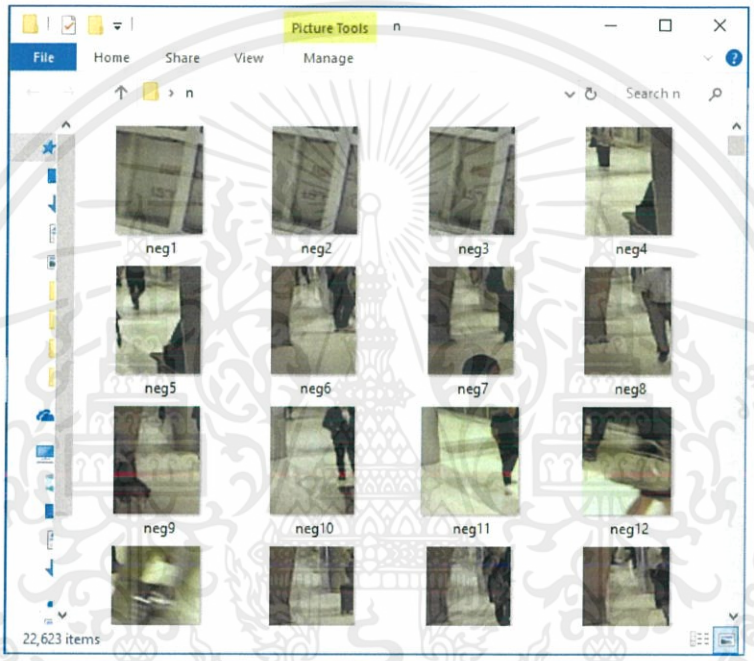


รูปที่ 2.29 รูปภาพใบหน้าแบบหลายมมุมมองของฐานข้อมูลใบหน้าคน

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.7.2.3 การเตรียมฐานข้อมูลที่ไม่ใช่ใบหน้าคนจากมุมมองจริง

จากการที่ได้ทำการเตรียมฐานข้อมูลใบหน้าคนจากการใช้เทคนิคการจับคู่ภาพจากมุมของจริง และการใช้เทคนิครู้จำใบหน้าหลายมุมมอง แต่ส่วนที่สำคัญอีกส่วนคือการเตรียมฐานข้อมูลทำการเตรียมฐานข้อมูลรูปภาพที่ไม่ใช่ใบหน้าคนจากมุมมองจริงของสถานี โดยจะทำการตัดรูปภาพที่ไม่ใช่ใบหน้าคนจากมุมมองจริงจากสถานีออกมา ดังรูปที่ 2.30 เพื่อที่จะลดความผิดพลาดในการรู้จำให้น้อยลง เนื่องจากมีปริมาณฐานข้อมูลที่ไม่ใช่ใบหน้าคนนำมาสอนระบบมากขึ้น

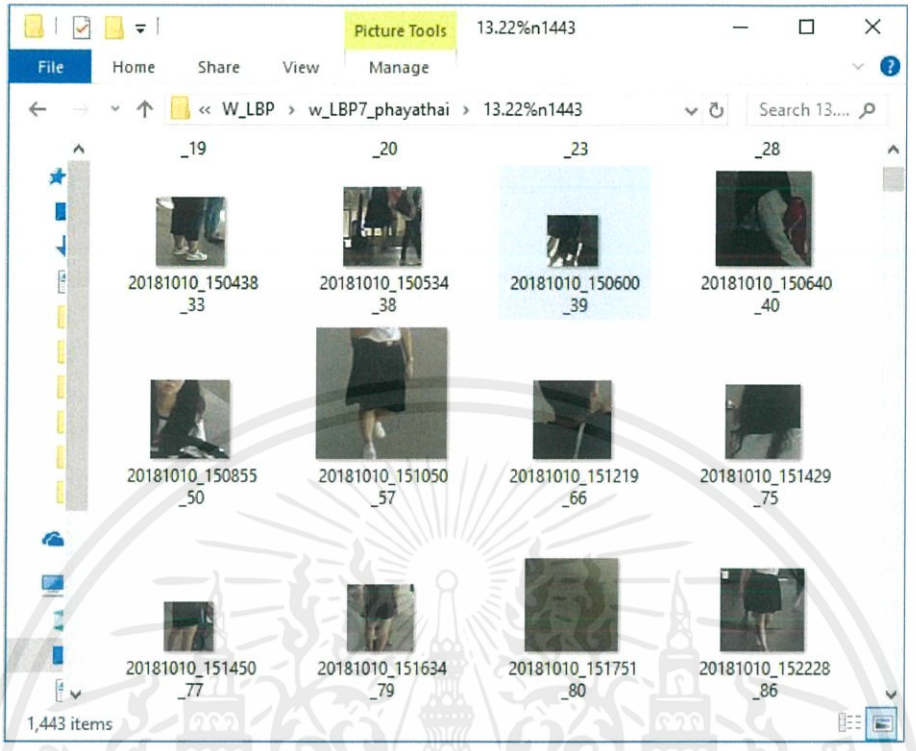


รูปที่ 2.30 ข้อมูลรูปภาพที่ไม่ใช่ใบหน้าคนจากมุมมองจริง

2.7.2.4 การเตรียมฐานข้อมูลที่ไม่ใช่ใบหน้าคนจากความผิดพลาดจริง

จากการที่ได้เตรียมฐานข้อมูลที่ไม่ใช่ใบหน้าคนจากมุมมองจริงจะช่วยให้ระบบประมวลผลภาพมีประสิทธิภาพที่เพิ่มขึ้นและแม่นยำขึ้น แต่ยังไม่เพียงพอสำหรับในการใช้งานจริง จึงทำให้ต้องมีการพัฒนาเพิ่มโดยการใช้เทคนิคการดึงความผิดพลาดจริงของฐานข้อมูลมาทำการคัดแยกและทำการเพิ่มข้อมูลรูปภาพที่ไม่ใช่ใบหน้าคนไปเก็บไว้ในฐานข้อมูลรูปภาพที่ไม่ใช่ใบหน้าคนเพิ่ม โดยจะมีการดึงข้อมูลรูปภาพที่ผิดพลาดจากทุกสถานีมาทำการเก็บไว้ในฐานข้อมูลรูปภาพที่ไม่ใช่ใบหน้าคน ตามรูปที่ 2.31 แล้วนำกลับไปสอนระบบใหม่อีกครั้ง เพื่อให้มีประสิทธิภาพในการรู้จำใบหน้าที่ผิดพลาดน้อยลง และประสิทธิภาพในการรู้จำใบหน้าแม่นยำมากขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.31 ข้อมูลรูปภาพที่ไม่ใช่ใบหน้าคนจากความผิดพลาดจริง

2.8 ความรู้พื้นฐานของการตรวจจับวัตถุ

ความรู้พื้นฐานของการตรวจจับวัตถุเป็นส่วนที่จำเป็นต่อการพัฒนาการประมวลผลและวิเคราะห์ภาพสำหรับงานสื่อโฆษณาดิจิตอล เพื่อสำหรับในการทำความเข้าใจในหลักการทำงานของระบบตรวจจับและนับจำนวนคน โดยพื้นฐานในการตรวจจับใบหน้าคนที่ถูกนำมาใช้งานในสถานีแอร์พอตลิ้งทั้ง 3 สถานีคือวิธีการ Sliding windows Image pyramid และNon-maxima suppression

เมื่อต้องการสร้างระบบตรวจจับและนับจำนวนคนที่ถูกนำมาใช้งานในการจำแนกประเภทจะมีการใช้ในกระบวนการในการตรวจจับวัตถุ โดยมีองค์ประกอบสำคัญหลักอยู่ 4 อย่าง

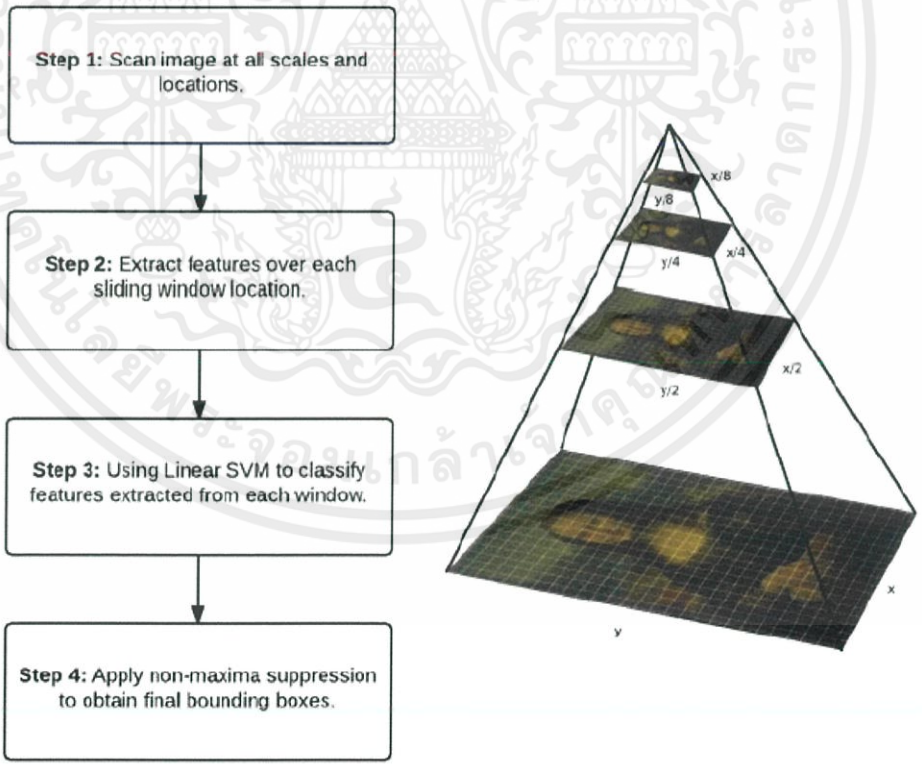
1) กระบวนการสำคัญแรกคือกระบวนการทำ Sliding window โดยในการทำจะมีการสร้างพื้นสี่เหลี่ยมที่หนึ่งขึ้นเพื่อใช้สำหรับในการเลื่อนหาวัตถุจากซ้ายไปขวาและจากบนลงล่าง ซึ่งในแต่ละการตรวจจับจะทำไปพร้อมกับการหาพื้นที่ Region of Interest (ROI) ที่ทำการเลื่อนจากซ้ายไปขวาและจากบนลงล่าง โดยในกระบวนการทำ Sliding window จะช่วยให้เราสามารถตรวจจับตำแหน่งของวัตถุได้แม่นยำมากขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) จากที่ผ่านกระบวนการทำ Sliding window จากนั้นจะทำการใช้กระบวนการ Image pyramid ที่ช่วยลดขนาดของรูปภาพอินพุต โดยการทำงานของกระบวนการ Image pyramid จะทำการจัดเรียงรูปภาพจากรูปภาพขนาดใหญ่ไปเป็นรูปภาพขนาดเล็กลงไปเรื่อยๆ ดังรูปที่ 2.32 โดยกระบวนการทำ Sliding window จะทำงานในแต่ละขนาดของรูปภาพที่ได้จนเสร็จและจากนั้นจะทำการลดขนาดโดยใช้กระบวนการ Image pyramid เพื่อที่จะตรวจจับวัตถุที่อยู่ใกล้และวัตถุที่อยู่ไกลออกมาจากรูปภาพ

3) หลังจากที่ทำกระบวนการทำ Sliding window และ Image pyramid ซึ่งในการตรวจจับวัตถุหลายวัตถุที่หลายๆขนาดของการทำ Sliding window จะทำให้เกิดปรากฏการณ์ที่เมื่อมีการตรวจจับวัตถุเดี่ยวกันหลายๆครั้ง ดังนั้นจึงเสนอกระบวนการทำอีกหนึ่งกระบวนการคือการทำ Non-maxima suppression ขึ้นมา เพื่อรวมกลุ่มวัตถุที่ตรวจจับให้เป็นวัตถุเดี่ยวกัน

4) สุดท้ายเพื่อที่จะทำให้แน่ใจว่าการทำงานของการตรวจจับวัตถุจะทำงานได้รวดเร็วที่สุด โดยจำเป็นต้องสนใจการทำงานของ Graphics Processing Unit (GPU) ในการทำงานสำหรับการประมวลผลในแต่ละ ROI ที่ทำการเลื่อนไปให้มีประสิทธิภาพที่สูงสุด



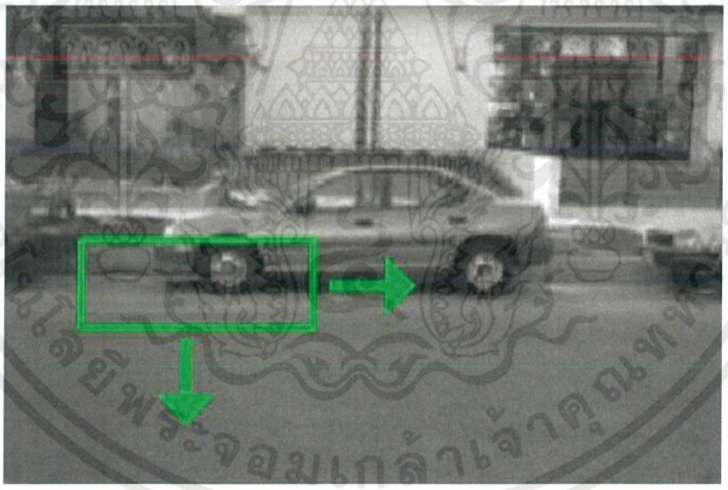
รูปที่ 2.32 รูปด้านซ้าย คือ ขั้นตอนพื้นฐานของการตรวจจับหาวัตถุ และรูปด้านขวา คือ กระบวนการ Image pyramid โดยรูปภาพต้นฉบับจะอยู่ที่ด้านล่าง

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.8.1 Sliding Windows

Sliding Windows เป็นเป็นการทำงานของการเลื่อนที่มีความสำคัญในการพัฒนาการประมวลผลและวิเคราะห์ภาพสำหรับงานสื่อโฆษณาดิจิตอลเพื่อใช้สำหรับในการตรวจจับวัตถุเนื่องจากการทำกระบวนการ Sliding Window สามารถที่จะระบุตำแหน่งของวัตถุออกมาเป็นค่า (x,y) ในรูปภาพที่วัตถุได้ ดังนั้นการทำกระบวนการ Sliding Window จะสามารถหาวัตถุและสถานที่ต่างๆ โดยในการทำกระบวนการ Sliding Windows จะมีการสร้างพื้นที่สีเหลี่ยมที่มีความกว้าง (W) และความสูง (H) คงที่ขึ้น เพื่อนำพื้นที่สี่เหลี่ยมไปทำการเลื่อนจากด้านซ้ายไปด้านขาวและจากด้านบนไปด้านล่างมาทำตรวจจับวัตถุ

ในรูปที่ 2.33 แสดงรูปภาพของการทำงานของกระบวนการ Sliding Windows โดยพื้นหลังของรูปภาพที่ทำการตรวจจับวัตถุในขณะนั้นจะมีพื้นที่สีเหลี่ยมสีเขียวคือการทำกระบวนการ Sliding Windows ซึ่งในพื้นที่สี่เหลี่ยมกำลังจะเลื่อนจากด้านซ้ายไปด้านขวาและจากด้านบนไปด้านล่าง สำหรับพื้นที่สี่เหลี่ยมแต่ละ Sliding Windows จะใช้พื้นที่สีเหลี่ยมหรือเรียกว่า ROI ส่งไปทำการประมวลผล เพื่อตรวจสอบวัตถุ (Adrian, 2018, p. 209)



รูปที่ 2.33 รูปภาพของการทำกระบวนการ Sliding Windows

2.8.2 Image Pyramids

Image Pyramids เป็นหนึ่งในกระบวนการที่สำคัญในการพัฒนาการประมวลผลและวิเคราะห์ภาพสำหรับงานสื่อโฆษณาดิจิตอล เพื่อใช้สำหรับในการค้นหาวัตถุในรูปภาพที่ระยะขนาดใกล้และไกลของรูปภาพ โดยการทำงานจะมีการลดขนาดของรูปภาพ ซึ่งในการลดขนาดของรูปภาพจะแสดงดังรูปที่ 2.32 ที่รูปภาพด้านล่างจะเป็นรูปภาพต้นฉบับเดิมขนาดความกว้าง (W) และความสูง (H) เท่าเดิม แต่ในชั้นถัดไปของ

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปภาพจะลดขนาดของรูปภาพให้ลดลงเรื่อยๆ จนกว่ารูปภาพจะมีขนาดเล็กกว่าพื้นที่สี่เหลี่ยม ROI แล้วหยุดการทำ Image Pyramids

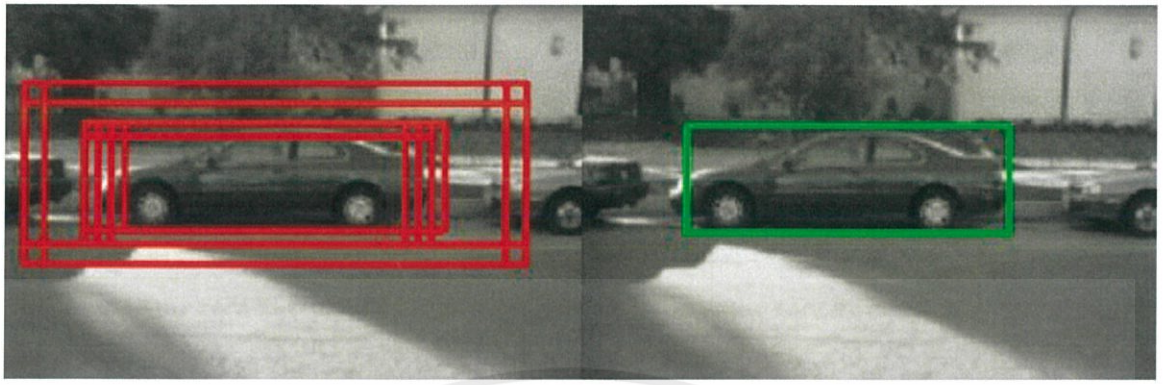
จากรูปที่ 2.34 ในรูปด้านซ้ายจะเห็นการทำกระบวนการ Sliding Windows เพียงอย่างเดียวจากรูปภาพต้นฉบับ จะสังเกตได้ว่าวัตถุที่ตรวจจับได้จะมีขนาดใหญ่เกินไปซึ่งเป็นความผิดพลาดที่เกิดขึ้นจากการตรวจจับวัตถุ แต่ในรูปด้านขวาจะแก้ปัญหาเรื่องวัตถุที่ตรวจจับมีขนาดใหญ่เกินไป โดยทำการลดขนาดของรูปภาพ ซึ่งจะทำให้การทำกระบวนการ Sliding Windows มีขนาดที่พอดีกับวัตถุ (Adrian, 2018, p. 211)

รูปที่ 2.34 รูปด้านซ้าย คือ รูปภาพที่ทำกระบวนการ Sliding window และรูปด้านขวา คือ รูปภาพที่ทำกระบวนการ Sliding window และ Image Pyramids

2.8.3 Non-maxima Suppression

Non-maxima Suppression (NMS) เป็นกระบวนการทำที่สำคัญในการพัฒนาการประมวลผลและวิเคราะห์ภาพสำหรับงานสื่อโฆษณาดิจิตอลเพื่อใช้ในการแก้ไขปัญหาที่เกิดจากการตรวจจับวัตถุหลายๆรอบในวัตถุอันเดิมซึ่งแสดงในรูปที่ 2.35 รูปด้านซ้ายจะไม่มีการทำกระบวนการ NMS ทำให้ในรูปภาพมีการตรวจจับวัตถุและมีพื้นที่ ROI หลายพื้นที่ออกมา ส่วนในรูปด้านขวามีการทำกระบวนการ NMS (Adrian, 2018, p. 213)

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.35 รูปด้านซ้าย แสดงการทำงานแบบไม่มี NMS และ
รูปด้านขวา แสดงการทำงานแบบมี NMS

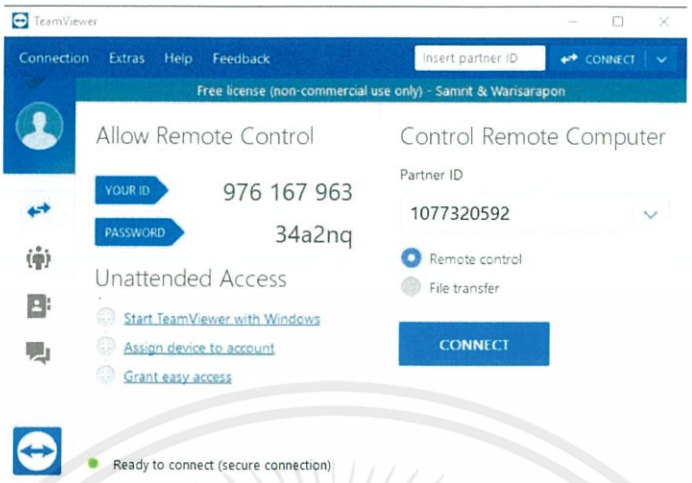
2.9 โปรแกรมที่ใช้สำหรับในการพัฒนา

โปรแกรมที่ใช้สำหรับในการพัฒนาเป็นอีกหนึ่งสิ่งสำคัญในการอำนวยความสะดวก สำหรับการพัฒนาการประมวลผลและวิเคราะห์ภาพสำหรับงานสื่อโฆษณาดิจิตอล โดยจะมีโปรแกรมหลัก 3 โปรแกรม ได้แก่ โปรแกรม Teamviewer เป็นโปรแกรมที่ใช้สำหรับในการส่งข้อมูล หรือรับข้อมูลที่สถานีแอร์พอร์ตลิ้งทั้ง 3 สถานี โปรแกรม Cascade-Trainer-GUI จะเป็นโปรแกรมในส่วนของการสอนระบบให้มีการรู้จำ และโปรแกรม Visual Studio Code เป็นโปรแกรแกรมในส่วนเสริมสำหรับในการอำนวยความสะดวกในการเขียนโปรแกรมภาษาไพทอน โดยจะมีรายละเอียดดังนี้

2.9.1 โปรแกรม Teamviewer

โปรแกรม Teamviewer เป็นส่วนหนึ่งในการทำงานในการพัฒนาการประมวลผลและวิเคราะห์ภาพสำหรับงานสื่อโฆษณาดิจิตอล เนื่องจากโปรแกรม Teamviewer จะช่วยให้สารารถเข้าถึงข้อมูลและควบคุม ASUS Tinker Board จากที่แต่ละสถานีจากทั้ง 3 สถานีได้โดยจะแสดงรูปภาพโปรแกรมในรูปที่ 2.36

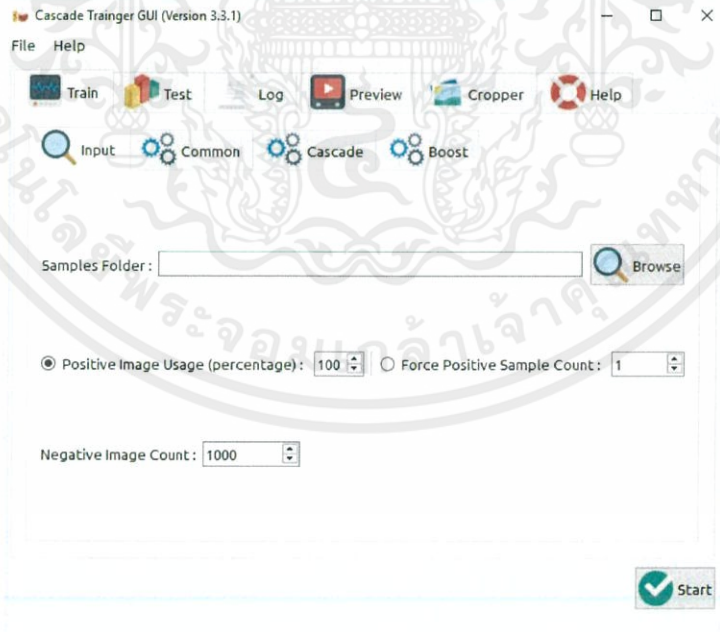
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.36 โปรแกรม Teamviewer

2.9.2 โปรแกรม Cascade-Trainer-GUI

โปรแกรม Cascade-Trainer-GUI เป็นหนึ่งในโปรแกรมที่ได้ทำการนำมาใช้งานในการสอนและการเรียนรู้ของการ Training ซึ่งในโปรแกรมจะมีการปรับทั้งชนิดของการ Training ได้แก่ Haar LBP และ HOG และโปรแกรมสามารถปรับได้ทั้งขนาดความกว้าง (W) ความสูง (H) และความลึกของการ Training เป็นต้น ซึ่งจะแสดงรูปภาพโปรแกรรูปที่ 2.37

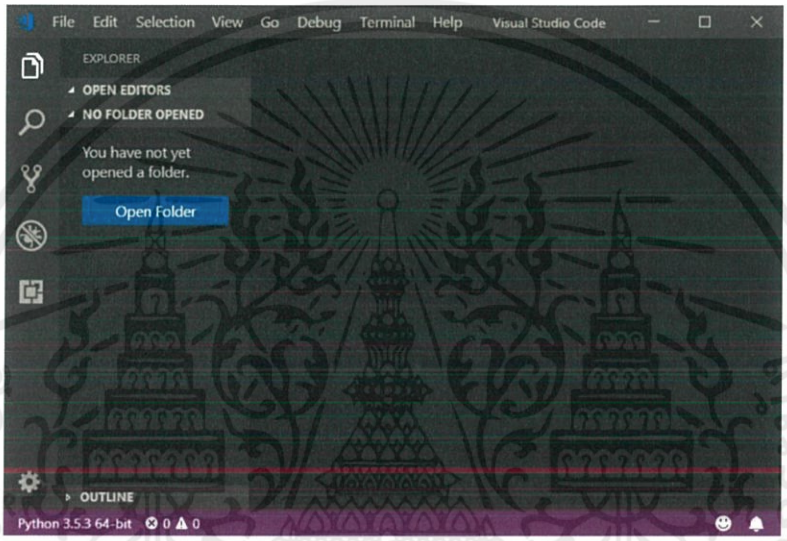


รูปที่ 2.37 โปรแกรม Cascade-Trainer-GUI

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.9.3 โปรแกรม Visual Studio Code

โปรแกรม Visual Studio Code เป็นโปรแกรมที่สำคัญสำหรับในการพัฒนาโปร โดยโปรแกรมนี้นี้จะได้สามารถใช้งานได้ทั้งระบบปฏิบัติการวินโดว์ แมค และวิลุกซ์ ซึ่งเป็นโปรแกรมที่ใช้งานง่ายและไม่ซับซ้อน พร้อมทั้งได้รับการยอมรับมากกว่า 30 โปรแกรมภาษา อธิเช่น C++ C# CSS Dockerfile HTML JavaScript JSON Less Markdown PHP และ Python เป็นต้น ซึ่งในส่วนของโปรแกรที่ได้ทำการนำมาใช้การพัฒนาจะทำการใช้ภาษาไพทอนในการเขียนโปรแกรม จะแสดงโปรแกรมดังรูปที่ 2.38



รูปที่ 2.38 โปรแกรม Visual Studio Code

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

วิธีดำเนินการวิจัย

ในบทนี้จะกล่าวถึงการออกแบบและการสร้างแบบจำลองการประมวลผลและวิเคราะห์สำหรับงานสื่อโฆษณาดิจิตอล เช่น การสร้างฐานข้อมูลรูปภาพใบหน้าคนและไม่ใช่ใบหน้าคน การออกแบบในการ Training การออกแบบระบบการตรวจจับใบหน้าและนับจำนวนคนและการออกระบบการตรวจจับสีหน้า และเพศ เป็นต้นซึ่งจะมีรายละเอียดดังนี้

3.1 การสร้างฐานข้อมูลรูปภาพใบหน้าคนและไม่ใช่ใบหน้าคน

ในการสร้างฐานข้อมูลรูปภาพสำหรับแบบจำลองการประมวลผลและวิเคราะห์สำหรับงานสื่อโฆษณาดิจิตอล จำเป็นที่จะต้องใช้ฐานข้อมูลรูปภาพทั้งหมด 2 แบบ คือ ข้อมูลรูปภาพใบหน้าคน และข้อมูลรูปภาพที่ไม่ใช่ใบหน้าคน เพื่อที่จะนำฐานข้อมูลทั้ง 2 แบบนี้ไปทำการ Training ระบบต่อไป โดยในการสร้างฐานข้อมูลรูปภาพจำเป็นที่จะต้องมีการใช้เทคนิคเหล่านี้

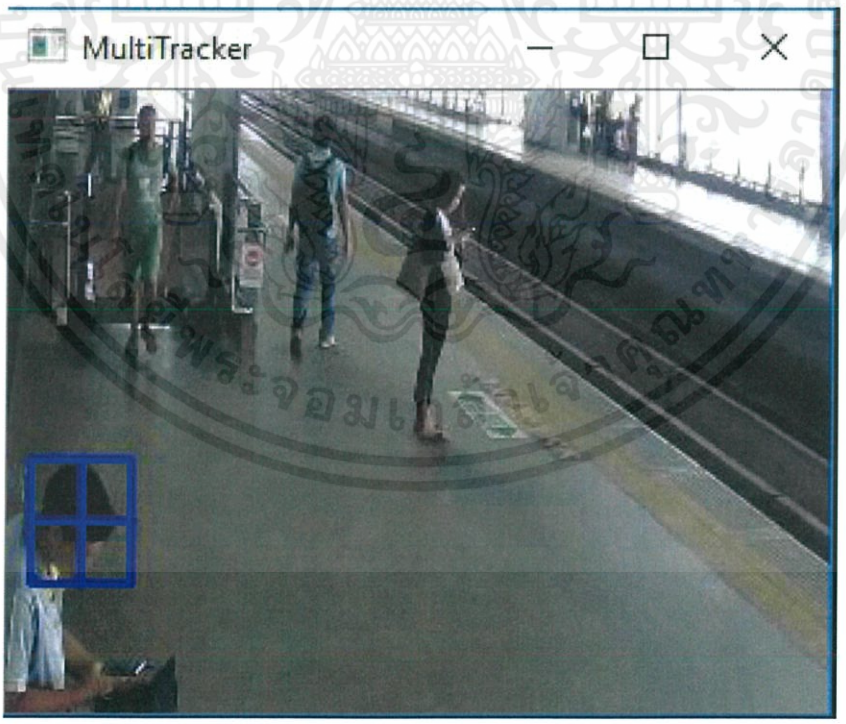
3.1.1 การจับคู่ภาพมุมมองจริง

ในการจับคู่ภาพมุมมองจริงที่ได้ใช้กันในการตรวจจับและนับจำนวนคนจำเป็นที่จะต้องมีข้อมูลรูปภาพใบหน้าคนจากสถานีจริงมาใช้งานสำหรับในการ Training โดยในการที่จะนำข้อมูลใบหน้าคนจากสถานีจริงออกมาจึงจำเป็นที่จะต้องมีการบันทึกวิดีโอจากสถานีลาดกระบังขนาด 800 x 600 พิกเซล ความร็ว 30 เฟรมต่อวินาทีความยาว 37.55 นาที ออกมาเพื่อจะนำวิดีโอที่บันทึกไว้จากสถานีลาดกระบังไปทำการแยกออกเป็นรูปภาพ โดยการใช้ภาษาไพทอนในการแยกวิดีโอออกมาเป็นรูปภาพซึ่งจำนวนรูปภาพที่ได้จากวิดีโอจะมีรูปภาพทั้งหมดจำนวน 45,516 รูป ตามรูปที่ 3.1

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 3.1 รูปภาพของสถานีลาดกระบัง 45,513 รูป

จากที่ได้ทำการแยกรูปภาพจากวิดีโอคลิปเรียบร้อยแล้วจากนั้นจะทำการจับคู่ภาพมุมมองจริงจากการตัดส่วนของรูปภาพใบหน้าคนจากมุมมองจริงของสถานีลาดกระบัง โดยใช้ภาษาไพทอนในการตัดรูปภาพใบหน้าคนจากมุมมองจริงออกมา ตามรูปที่ 3.2



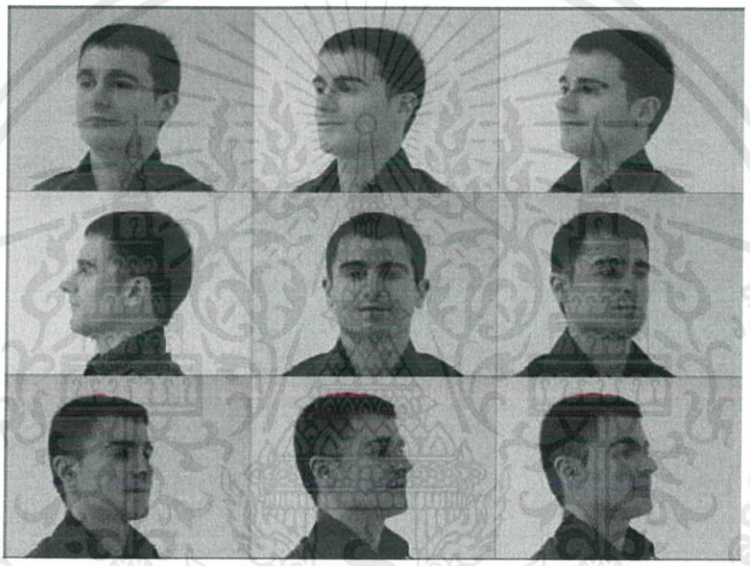
รูปที่ 3.2 การตัดรูปภาพใบหน้าคนจากมุมมองจริงของสถานีลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หลังจากที่ได้ทำการตัดส่วนของรูปภาพใบหน้าคนจากมุมจริงของสถานีลาดกระบังออกมาจะได้จำนวนของรูปภาพใบหน้าคนออกมาทั้งหมด 13,013 รูปที่จะนำไปใช้สำหรับการ Training ต่อไป

3.1.2 การรู้จำใบหน้าแบบหลายมุมมอง

การรู้จำใบหน้าแบบหลายมุมมองเป็นส่วนหนึ่งในการออกแบบที่จะเพิ่มข้อมูลรูปภาพใบหน้าคน โดยจะทำการนำข้อมูลรูปภาพใบหน้าคนที่หลายมุมตั้งแต่ 0,±30, ±45, ±60 และ ±90 องศา ดังรูปที่ 3.3 ที่มากจากเว็ปไซด์ และฐานข้อมูลรูปภาพใบหน้าคนจากบริษัท เวริลี วิชั่น จำกัด มาทำการเพิ่มประสิทธิภาพรู้จำให้มากขึ้น

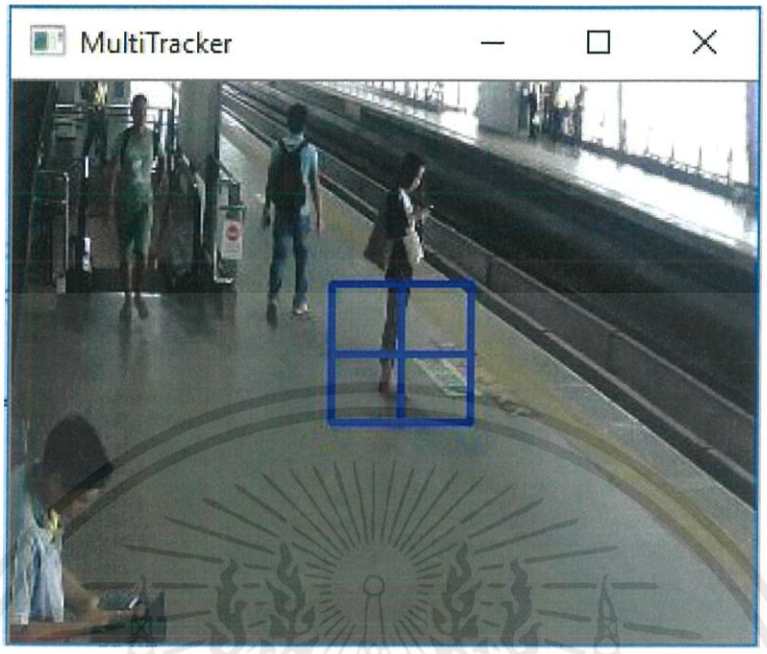


รูปที่ 3.3 รูปภาพใบหน้าแบบหลายมมุมมองของฐานข้อมูลใบหน้าคน

3.1.3 การเตรียมฐานข้อมูลที่ไม่ใช่ใบหน้าคนจากมุมมองจริง

จากการที่ได้ทำการสร้างฐานข้อมูลใบหน้าคนจากการใช้เทคนิคการจับคู่ภาพจากมุมของจริงและการใช้เทคนิครู้จำใบหน้าหลายมุมมอง แต่ก็ยังไม่เพียงพอสำหรับในการทำให้การตรวจจับมีความแม่นยำมากขึ้น เพราะในการตรวจจับใบหน้าคนที่สถานีจริงจะเกิดข้อผิดพลาดมากมาย เช่น ตรวจจับหู ขา แขน และองค์ประกอบต่างๆที่ไม่ได้สนใจ เป็นต้น ดังนั้นจึงต้องมีการออกแบบการสร้างฐานข้อมูลที่ไม่ใช่ใบหน้าคนเพื่อไปใช้สำหรับในการ Training ให้การตรวจจับใบหน้าคนมีความผิดพลาดให้น้อยลง โดยการใช้วิธีเหมือนกับการจับคู่ภาพมุมมองจริงที่ได้นำวิดีโอคลิปที่บันทึกไว้ที่สถานีลาดบังขนาด 800 x 600 ความเร็ว 30 เฟรมต่อวินาที ความยาว 37.55 นาที ไปทำการแยกออกเป็นรูปภาพ โดยการใช้ภาษาไพทอนในการแยกวิดีโอออกมาเป็นรูปภาพจากนั้นก็ใช้ภาษาไพทอนในการตัดข้อมูลรูปภาพที่ไม่ใช่ใบหน้าคนออกมาเก็บไว้ในฐานข้อมูล ดังรูปที่ 3.4

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.4 การตัดรูปภาพที่ไม่ใช่ใบหน้าคนจากมุมมองจริงของสถานีลาดกระบัง

หลังจากที่ได้ทำการตัดส่วนของรูปภาพที่ไม่ใช่ใบหน้าคนจากวิดีโอที่สถานีลาดกระบังออกมาจนได้ฐานข้อมูลรูปภาพที่ไม่ใช่ใบหน้าเป็นจำนวนทั้งหมด 22,623 รูป ดังรูปที่ 3.5 เพื่อที่จะนำไปใช้สำหรับในการลดความผิดพลาดในการตรวจจับใบหน้าที่เกิดขึ้นจากสถานีลาดกระบัง

รูปที่ 3.5 ฐานข้อมูลรูปภาพที่ไม่ใช่ใบหน้าคนจากมุมมองจริงของทุกสถานี

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.4 การเตรียมฐานข้อมูลที่ไม่ใช่ใบหน้าคนจากความผิดพลาดจริง

จากการที่ได้เตรียมฐานข้อมูลที่ไม่ใช่ใบหน้าคนจากมุมมองจริงจะช่วยให้ระบบประมวลผลภาพมีประสิทธิภาพที่เพิ่มขึ้นและความผิดพลาดในการตรวจจับที่น้อยลง แต่ยังไม่เพียงพอสำหรับในการใช้งานจริงเพราะในการใช้งานจริงยังพบกับปัญญหาที่เกี่ยวข้องกับการตรวจจับที่ไม่ใช่ใบหน้าคนออกมา ดังนั้นจึงทำให้ต้องมีการออกแบบฐานข้อมูลขึ้นมาใหม่เพื่อที่จะเพิ่มประสิทธิภาพในการตรวจตรวจจับให้ดีขึ้น โดยในการออกแบบฐานข้อมูลนี้จะมีการดึงความผิดพลาดจริงของฐานข้อมูลมาทำการคัดแยกและทำการเพิ่มข้อมูลของรูปภาพที่ไม่ใช่ใบหน้าคนมาเก็บไว้ ดังรูปที่ 3.6 โดยจะทำการดึงข้อมูลรูปภาพที่ผิดพลาดจากทุกสถานีมาทำการเก็บไว้ฐานข้อมูลรูปภาพที่ไม่ใช่ใบหน้าคนทั้งหมด เพื่อที่จะให้มีประสิทธิภาพในการผิดพลาดให้น้อยลง และประสิทธิภาพในการรู้จำใบหน้าคนแม่นยำมากขึ้น

รูปที่ 3.6 ฐานข้อมูลรูปภาพที่ไม่ใช่ใบหน้าคนจากมุมมองจริงของสถานีทุกสถานี

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 การออกแบบการ Training

การออกแบบการ Training เป็นการออกแบบหลักในการที่จะทำให้การตรวจจับให้มีประสิทธิ โดยจะทำการใช้โปรแกรมสำหรับทำการ Training ซึ่งจะต้องนำฐานข้อมูลรูปภาพใบหน้าคนและรูปภาพที่ไม่ใช่ใบหน้าคนมาทำการแยกข้อมูลออกเป็น 2 ส่วนได้แก่ 1) ส่วนของรูปภาพใบหน้าคนเป็นรูปภาพ Positive เก็บไว้ที่โพลเดอร์ p 2) ส่วนของรูปภาพที่ไม่ใช่ใบหน้าคนเป็นรูปภาพ Negative เก็บไว้ที่โพลเดอร์ n ตามรูปที่ 3.7 ที่ได้ทำการแยกไว้

รูปที่ 3.7 ฐานข้อมูลที่ใช้สำหรับในการ Training

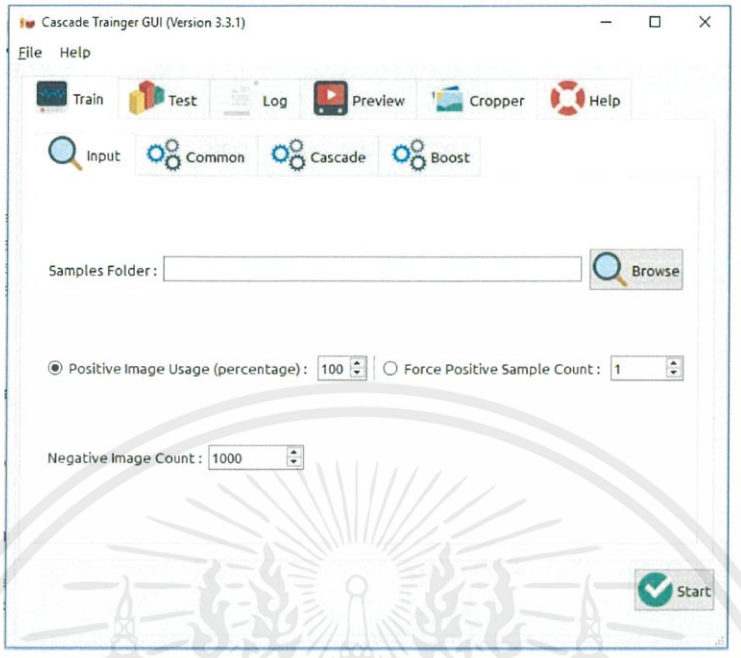
หลังจากที่ได้ทำการแยกฐานข้อมูลออกมาเป็นส่วนของรูปภาพ Positive ที่เก็บไว้ในโพลเดอร์ p และส่วนของรูปภาพ Negative ที่เก็บไว้ในโพลเดอร์ n ต่อไปจะทำการใช้โปรแกรม Cascade-Trainer-GUI ดังรูปที่ 3.8 ใช้สำหรับในการออกแบบการ Training

รูปที่ 3.8 โปรแกรม Cascade-Trainer-GUI

ต่อมาจะทำการใช้โปรแกรมและการตั้งค่าโปรแกรม Cascade-Trainer-GUI สำหรับในการ Training ดังนี้

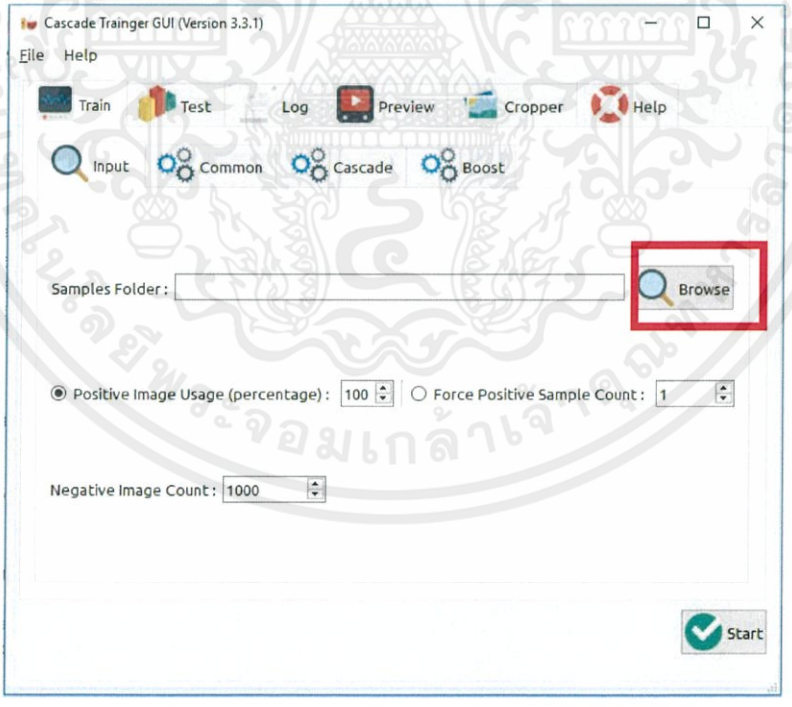
1) หลังจากที่ทำการเปิดโปรแกรม Cascade-Trainer-GUI จะพบหน้าต่างโปรแกรม ดังรูปที่ 3.9 ที่จะทำการตั้งค่าอินพุตที่จะใช้สำหรับในการ Training

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.9 หน้าต่างโปรแกรม Casecade-Trainer-GUI

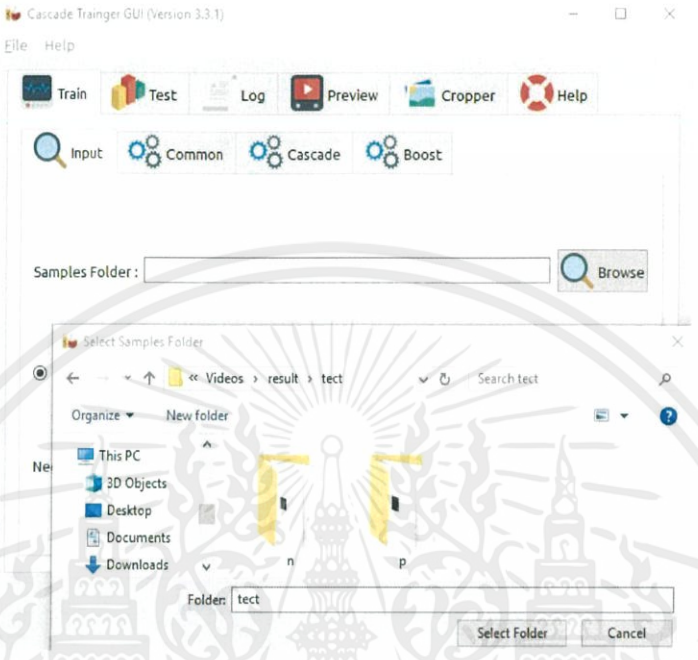
2) เลือกไปที่ Browse ที่จะทำการ Training ดังรูปที่ 3.10



รูปที่ 3.10 แสดงการเลือก Browse

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3) เลือกโฟลเดอร์ที่เก็บข้อมูลรูปภาพ Positive และรูปภาพ Negative ที่จะทำการ Training ดังรูปที่ 3.11



รูปที่ 3.11 แสดงการเลือกที่เก็บข้อมูลรูปภาพ

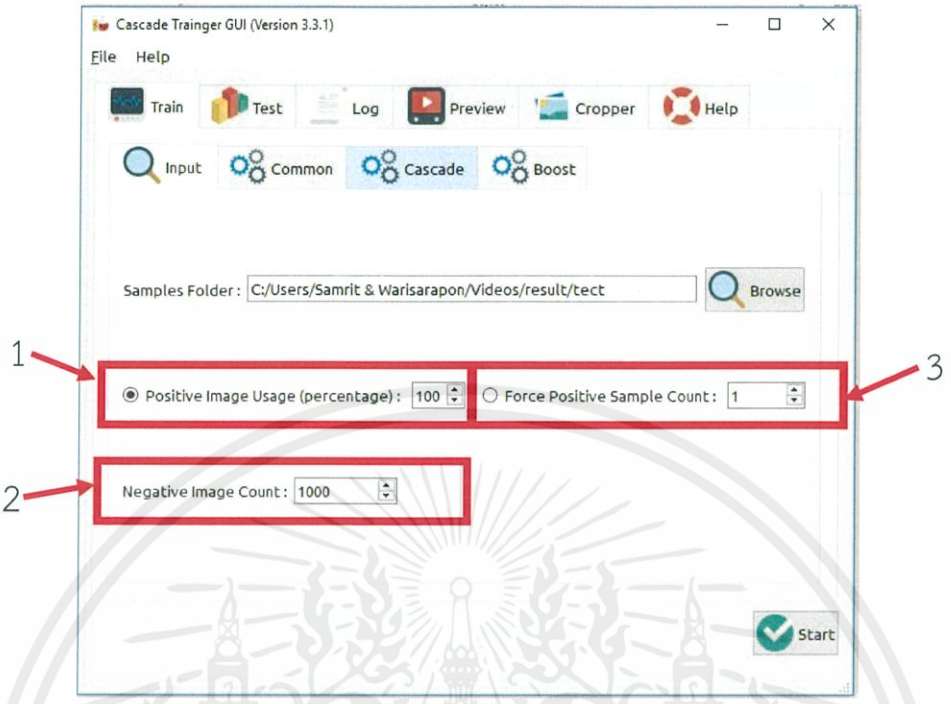
4) เลือกชนิด และจำนวนของรูปภาพ Positive และรูปภาพ Negative ดังรูปที่ 3.12 ที่จะทำการ Training

4.1) ส่วนที่ 1 Positive Image Usage (percentage) คือ การเลือกจำนวนรูปภาพ Positive จาก 0 - 100 เปอร์เซ็นของรูปภาพ Positive ของทั้งหมด

4.2) ส่วนที่ 2 Negative Image Count คือ การเลือกจำนวนรูปภาพ Negative ที่ต้องการโดยจำนวนรูปภาพไม่มากกว่ารูปภาพ Negative ที่มี

4.3) ส่วนที่ 3 Force Positive Sample Count คือ การเลือกจำนวนรูปภาพ Positive ที่ต้องการโดยจำนวนรูปภาพไม่มากกว่ารูปภาพ Positive ที่มี

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.12 แสดงขั้นตอนในการเลือกชนิด และจำนวนรูปภาพ

สำหรับในการออกแบบการเลือกชนิด และจำนวนของรูปภาพ Positive และรูปภาพ Negative ที่จะทำการใช้สำหรับในการ Training และการ Testing โดยในการออกแบบที่ได้ใช้คือการแบ่งข้อมูลรูปภาพ Positive ไปทำการ Training 80 เปอร์เซ็นต์ และรูปภาพ Positive อีก 20 เปอร์เซ็นไปทำการ Testing ต่อไปจะทำการเลือกการทำงานของหน้าต่าง Common ที่ใช้งาน

5) เลือกการทำงานของหน้าต่าง Common ดังรูปที่ 3.13

5.1) ส่วนที่ 1 Number of Stages คือ จำนวนของการ Training

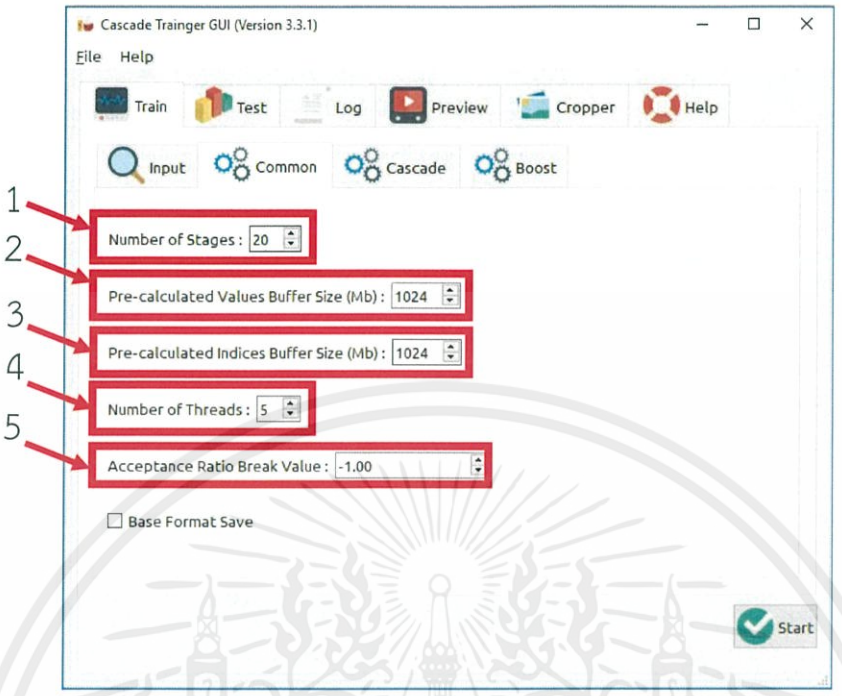
5.2) ส่วนที่ 2 Pre-calculated Values Buffer Size คือ การกำหนดค่าขนาดของบัฟเฟอร์ที่ใช้ในการ Training

5.3) ส่วนที่ 3 Pre-calculated Indices Buffer Size คือ การกำหนดค่าดัชนีของบัฟเฟอร์ที่ใช้ในการ Training

5.4) ส่วนที่ 4 Number of Threads คือ ค่าที่กำหนดกระบวนการทำงานของ CPU โดยค่า Thead จะสารถมารถทำงานได้สูงสุดตามค่า CPU ของเครื่องคอมพิวเตอร์

5.5) ส่วนที่ 5 Acceptance Ratio Break Value คือ ค่าที่กำหนดรูปแบบของการเรียนรู้ และการหยุดของการ Trainning

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.13 แสดงการตั้งค่าของหน้าต่าง Common

สำหรับในการ Training ในส่วนของหน้าต่าง Common ที่ได้ใช้สำหรับในการออกแบบระบบประมวลผลจะมีการใช้ Number of Stages จำนวน 20 stagesในการ Training เพื่อให้ระบบมีความแม่นยำมากขึ้น มีการใช้ Pre-calculated Values Buffer Size Pre-calculated Indices Buffer Size และNumber of Threads ตามค่าที่ได้กำหนดตามมาตรฐานในการทำงานของคอมพิวเตอร์ ส่วนของ Acceptance Ratio Break Value จะไม่มีการกำหนดการทำงานค่าในการหยุดของการ Training ต่อไปจะทำการเลือกการทำงานของหน้าต่าง Casecade ที่ใช้งาน

6) เลือกการทำงานของหน้าต่าง Casecade ดังรูปที่ 3.14

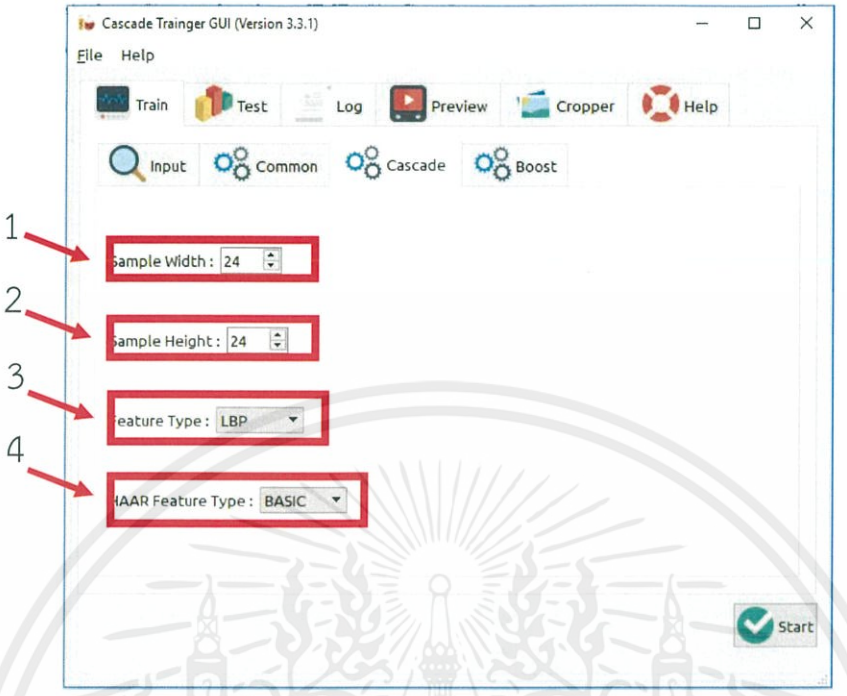
6.1) ส่วนที่ 1 Sample Width คือ การกำหนดค่าความกว้างของการ Training โดยขนาดความกว้างเท่ากับขนาดของการตรวจจับและนับจำนวนคน

6.2) ส่วนที่ 2 Sample Height คือ การกำหนดค่าความสูงของการ Training โดยขนาดความสูงเท่ากับขนาดของการตรวจจับและนับจำนวนคน

6.3) ส่วนที่ 3 Feature Type คือ การกำหนดชนิดของการ Training เช่น Haar LBP และHog เป็นต้น

6.4) ส่วนที่ 4 Haar Feature Type คือ การกำหนดลักษณะของการ Training แบบ Haar

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.14 แสดงการตั้งค่าของหน้าต่าง Casecade

สำหรับการทำงานในส่วนของหน้าต่าง Casecade จะทำการออกแบบให้ Sample Width และ Sample Height มีขนาด 24 พิกเซลเท่ากับขนาดของการตรวจจับและนับจำนวนคน Feature Type จะทำการเลือกชนิดของการ Training เป็น LBP ส่วนของ HAAR Feature Type จะไม่ส่งเกี่ยวข้องในการทำการ Training ของ LBP เพราะ HAAR Feature Type จะใช้ปรับการทำงานของ Haar ส่วนต่อไปจะทำการเลือกการทำงานของหน้าต่าง Boost ที่ใช้งาน

7) เลือกการทำงานของหน้าต่าง Boost ดังรูปที่ 3.15

7.1) ส่วนที่1 Boost Type คือ การกำหนดชนิดของ Boost ใช้สำหรับในการจำแนกประเภท

7.2) ส่วนที่ 2 Minimal Hit Rate คือ การกำหนดค่าความเร็วต่ำสุดของการทำงานในแต่ละ Stage

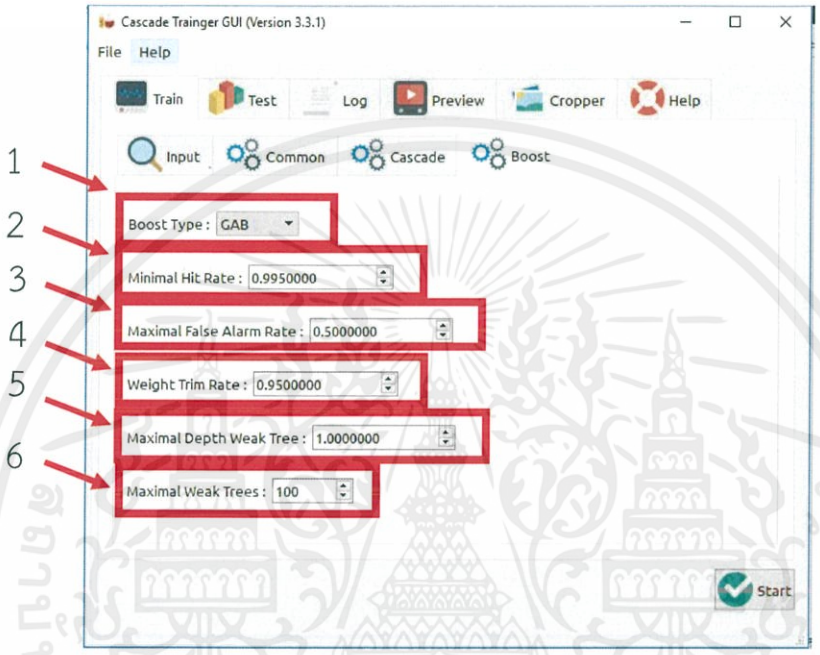
7.3) ส่วนที่ 3 Maximal False Alarm Rate คือ การกำหนดค่าการเตือนความผิดพลาดสูงสุดของการทำงานในแต่ละ Stage

7.4) ส่วนที่ 4 Weight Trim Rate คือ การกำหนดอัตราในการปรับค่าค่า Weight ในแต่ละ Stage

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

7.5) ส่วนที่ 5 Maximal Depth Weak Tree คือ การกำหนดค่าความลึกสูงสุดของการ Training

7.6) ส่วนที่ 6 Maximal Weak Trees คือ การกำหนดค่าในการนับเวลาสำหรับในการทำงานของแต่ละ Stage



รูปที่ 3.15 แสดงการตั้งค่าของหน้าต่าง Boost

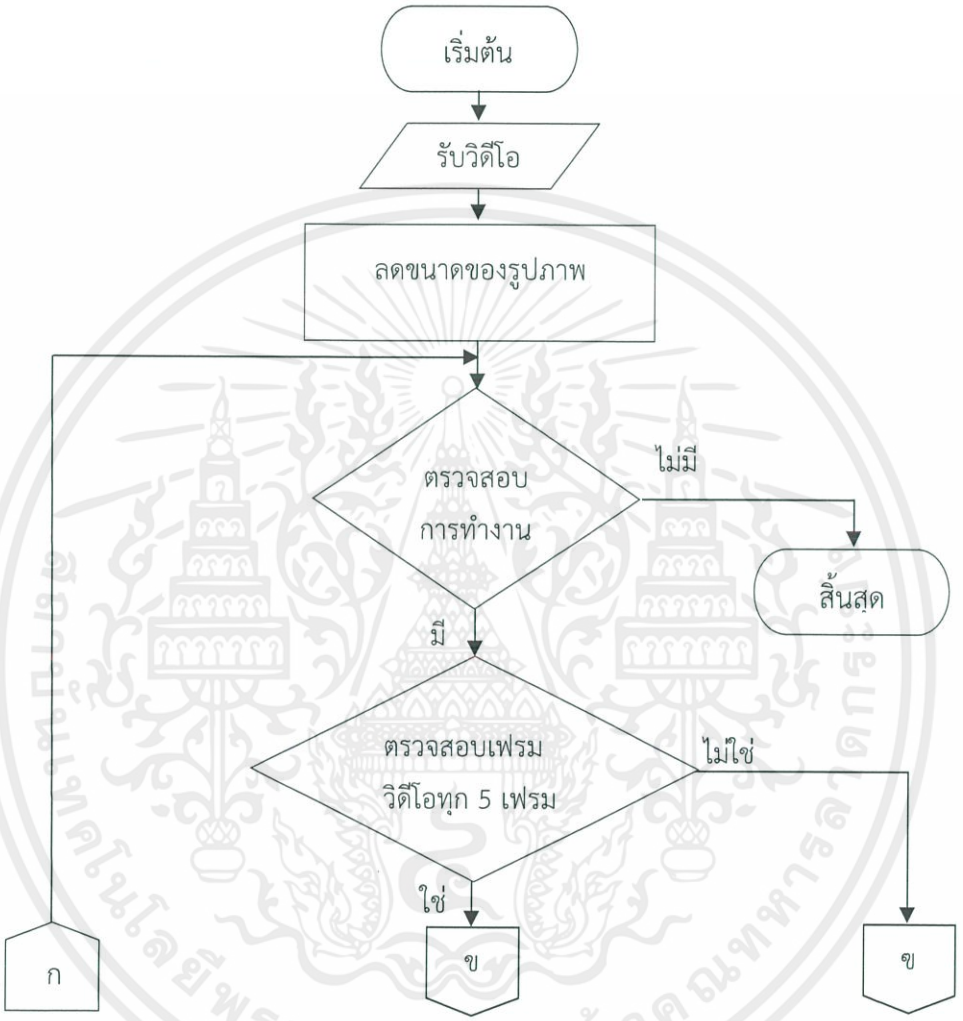
สำหรับในส่วนของหน้าต่าง Boost ได้แก่ Boost Type Minimal Hit Rate Maximal False Alarm Weight Trim Rate Maximal Depth Weak Tree และMaximal Weak Trees ได้กำหนดค่าต่างๆ เป็นค่ามาตราฐานเริ่มต้นในการทำงานของ LBP

สำหรับในส่วนของการเตรียมการ Training ที่ได้กำหนดค่าต่างๆในโปรแกรม Cascade-Trainer-GUI เป็นการออกแบบที่สำคัญในการที่ตรวจจับและนับจำนวนคนเนื่องจากถ้ามีการกำหนดค่าที่ไม่ได้มาตราฐานและกำหนดการตรวจจับผิดจะทำให้ในการตรวจจับมีความผิดพลาดที่เกิดขึ้นได้

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3 การออกแบบระบบการตรวจจับและนับจำนวนคน

ในการออกแบบระบบตรวจจับและนับจำนวนคนในการที่ใช้งานจริงสามารถแสดงภาพโฟวชาร์ตของระบบตรวจจับและนับจำนวนคนดังรูปที่ 3.16

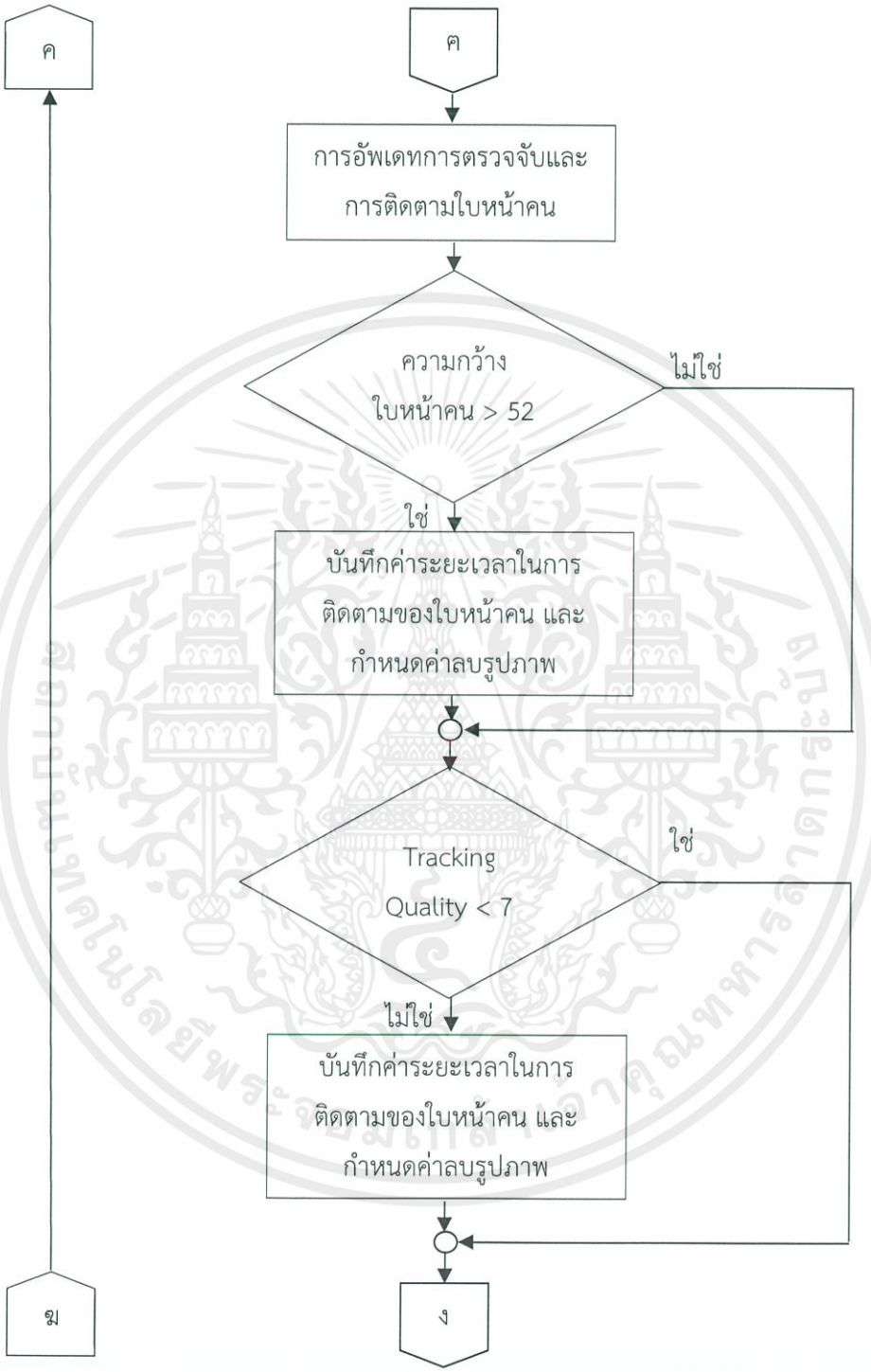


รูปที่ 3.16 โฟว์ชาร์ตของระบบตรวจจับและนับจำนวนคน

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

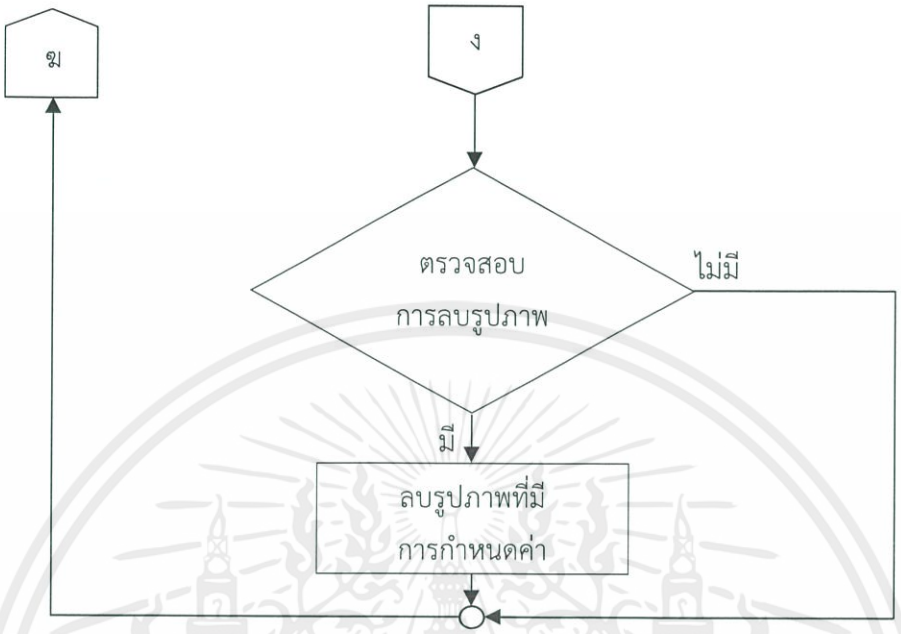
รูปที่ 3.16 (ต่อ) โฟว์ชาร์ตของระบบตรวจจับและนับจำนวนคน

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.16 (ต่อ) โฟว์ชาร์ตของระบบตรวจจับและนับจำนวนคน

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.16 (ต่อ) โฟว์ชาร์ตของระบบตรวจจับและนับจำนวนคน

ในรูปที่ 3.16 เป็นโฟว์ชาร์ตของระบบการตรวจจับและนับจำนวนคนในแต่ละขั้นตอนจะมีการออกแบบหน้าที่ในการทำงานดังนี้

3.3.1 การรับวิดีโอ

การรับวิดีโอจะเป็นการข้อมูลวิดีโอจากโมดูลกล้องมา โดยในการรับวิดีจากโมดูลกล้องแบบปกติจะมีการรับข้อมูลรูปภาพเข้ามาขนาด 640 x 480 พิกเซล ความเร็ว 120 เฟรมต่อวินาที ปริมาณพิกเซลใน 1 วินาทีจะมีจำนวนพิกเซลทั้งหมด 36,864,000 พิกเซลที่เข้าไปทำการประมวลผล แต่สำหรับการออกแบบในส่วนของการรับวิดีโอจะถูกปรับขนาดจากที่ตัวกล้องเป็น 320 x 240 พิกเซล ความเร็ว 120 เฟรมต่อวินาที ปริมาณพิกเซลใน 1 วินาทีจะมีจำนวนพิกเซลทั้งหมด 9,216,000 พิกเซลจะทำให้ปริมาณที่รับข้อมูลมีขนาดน้อยลง 4 เท่าทำให้ระบบประมวลผลได้เร็ว 4 เท่าโดยที่รูปภาพแบบปกติกับแบบปรับขนาดมีลักษณะรูปภาพที่เหมือนกัน โดยสามารถปรับขนาดได้ด้วยการเขียนโปรแกรมดังนี้

```
def detectAndTrackMultipleFaces():
    cap = cv2.VideoCapture(0)
    cap.set(cv2.CAP_PROP_FRAME_WIDTH,320)
    cap.set(cv2.CAP_PROP_FRAME_HEIGHT,240)
```


เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในส่วนของการเขียนโปรแกรมข้างต้นจะกำหนดวิดีโอที่จะทำโดยการปรับขนาดจากตัวกล้องให้เหลือขนาด 320 x 240 พิกเซลเพื่อที่จะให้คุณลักษณะของรูปใบหน้าคนไม่หายและทำให้ระบบมีการประมวลผลที่เร็วขึ้น

3.3.2 การแปลงสีจาก RGB เป็น Grayscale

ในส่วนของการทำงานในการแปลงสีจาก RGB เป็น Grayscale หลังจากที่ได้รับวิดีโอขนาด 320 x 240 พิกเซล ความเร็ว 120 เฟรมต่อวินาที ปริมาณพิกเซลใน 1 วินาทีจะมีจำนวนพิกเซลทั้งหมด 9,216,000 พิกเซลจะทำการออกแบบด้วยการแปลงสีจาก RGB เป็น Grayscale เนื่องจากขนาดของรูปภาพสีแบบ RGB จะมีความลึกของสีทั้งหมด 3 สีทำให้รูปภาพสีแบบ RGB ในแต่ละพิกเซลสีจะมีจำนวนบิตข้อมูลทั้งหมด 8 บิตต่อพิกเซลซึ่งทำให้ใน 1 พิกเซลรูปภาพสีแบบ RGB จะมีจำนวนบิตข้อมูลทั้งหมด 24 บิตต่อพิกเซลทำให้จำนวนบิตที่ประมวลผลรูปภาพทั้งหมดใน 1 วินาทีมีจำนวนบิตทั้งหมด 221,184,000 บิตต่อวินาทีที่ได้ไปทำการประมวลผลจึงทำให้ในการประมวลผลไม่ต่อเนื่องและเกิดการดีเลย์ ส่งผลให้ต้องนำการออกแบบด้วยการแปลงสี RGB เป็น Grayscale ด้วยแปลงเป็นรูปภาพสีแบบ Grayscale จะมีจำนวนบิตข้อมูลทั้งหมด 24 บิตต่อพิกเซลทำให้จำนวนบิตที่ประมวลผลรูปภาพทั้งหมใน 1 วินาทีจะมีจำนวนบิตทั้งหมด 73,728,00 บิตต่อวินาทีซึ่งทำให้ความเร็วในการประมวลผลมีความเร็วขึ้น 3 เท่าเพราะว่าการตัดสีในส่วนที่ไม่จำเป็นต่อการใช้งานออกไปจึงทำให้ระบบประมวลผลเร็วขึ้น ซึ่งจะสามารถแสดงการเขียนโปรแกรมได้ดังนี้

```
rc,fullSizeBaseImage = cap.read()
baseImage = cv2.resize( fullSizeBaseImage, ( 320, 240))
if (frameCounter % 5) == 0:
    gray = cv2.cvtColor(baseImage, cv2.COLOR_BGR2GRAY)
```

3.3.3 การตรวจจับใบหน้าคน และหาจุดศูนย์การของใบหน้า

ในส่วนของการตรวจจับใบหน้าคน และหาจุดศูนย์กลางของใบหน้า จะทำการออกแบบด้วยการกำหนดค่า Image Pyramids หรือ scaleFactor ให้ขนาดในการลดรูปภาพลดลงทุกๆ 10 % ในการตรวจจับใบหน้าคน ต่อมาทำการออกแบบในการหาใบหน้าคนด้วยการตรวจหาใบหน้าที่ซ้ำกันต่ำสุด 2 หน้าจึงจะตัดสินใจว่าใบหน้านั้นคือใบหน้าคน ส่วนของ minSize และmaxSize จะเป็นตัวกำหนดกรอบในการตรวจจับใบหน้าคนต่ำสุดที่ 17 x 17 พิกเซล และกรอบในการตรวจจับใบหน้าคนใหญ่สุดที่ 52 x 52 พิกเซลเพราะรูปภาพใบหน้าคนจะมีขนาดไม่เกิน 52 x 52 พิกเซลและขนาดเล็กไม่เกิน 17 x 17 พิกเซลทำให้ส่วนของ minSize และmaxSize ยังสามารถที่จะลดความผิดพลาดที่เกิดจากการตรวจจับใบหน้าคนได้ และยังมีการคำนวณหาจุดศูนย์กลางของใบหน้าคนเพื่อที่จะเก็บค่าไว้เช็คว่าคนเดิมหรือคนใหม่พร้อมกับเพิ่มค่า matchedFid สำในการเช็ครูปภาพว่าเป็นใบหน้าคนเดิมหรือไม่ ซึ่งจะสามารถแสดงการเขียนโปรแกรมได้ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
faces =LBPCascade.detectMultiScale(gray,scaleFactor=1.10,minNeighbors=2,
        minSize=(17, 17),maxSize=(52, 52))
for (_x,_y,_w,_h) in faces:
        x = int(_x)
        y = int(_y)
        w = int(_w)
        h = int(_h)
        x_bar = x + 0.5 * w
        y_bar = y + 0.5 * h
        matchedFid = None
```

3.3.4 การตรวจสอบใบหน้าคน

หลังจากที่ได้ทำการตรวจจับใบหน้าคน และหาจุดศูนย์การของใบหน้าต่อไปจะทำการเช็คใบหน้าคนว่าเป็นใบหน้าของคนเดิมหรือไม่ โดยในการออกแบบจะทำการอัพเดทข้อมูลตำแหน่งของรูปภาพใบหน้าและเช็คว่าคือตำแหน่งเดิมไม ถ้าเป็นรูปภาพใบหน้าคนเดิมจะมี keys สำหรับในการเก็บค่า fid เอาไว้เพื่อที่จะเข้าไปทำงานใน faceTrackers.keys() ดังที่แสดงการเขียนโปรแกรมไว้ด้านล่าง แต่ถ้ารูปภาพใบหน้าคนใหม่ก็จะไปสู่กระบวนการการเพิ่มการติดตาม นับจำนวนคน และกำหนดค่าที่บันทึกรูปที่จะกล่าวในส่วนถัดไป

```
for fid in faceTrackers.keys():
        tracked_position =  faceTrackers[fid].get_position()
        t_x = int(tracked_position.left())
        t_y = int(tracked_position.top())
        t_w = int(tracked_position.width())
        t_h = int(tracked_position.height())
        t_x_bar = t_x + 0.5 * t_w
        t_y_bar = t_y + 0.5 * t_h
        if ( ( t_x <= x_bar   <= (t_x + t_w)) and
                ( t_y <= y_bar   <= (t_y + t_h)) and
                ( x   <= t_x_bar <= (x   + w  )) and
```


เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
( y   <= t_y_bar <= (y   + h  ))):
matchedFid = fid
```

3.3.5 การเพิ่มการติดตาม นับจำนวนคน และกำหนดค่าที่บันทึกรูป

ในส่วนของการเพิ่มการติดตาม นับจำนวนคน และกำหนดค่าที่บันทึกรูป จะออกแบบด้วยการเพิ่มการติดตามโดยถ้าค่า matchedFid เป็น None หมายความว่ารูปภาพใบหน้าคนนี้เป็นอันใหม่ที่ตรวจจับพบ จากนั้นก็จะทำการเพิ่มการติดตามด้วย dlib และนับจำนวนด้วยการใช้ threading.Thread() และกำหนดค่าที่บันทึกรูป ซึ่งจะสามารถแสดงการเขียนโปแกรมได้ดังนี้

```
if matchedFid is None:
    tracker = dlib.correlation_tracker()
    tracker.start_track(baseImage,
    dlib.rectangle( x, y, x+w, y+h))
    faceTrackers[ currentFaceID ] = tracker
    t = threading.Thread( target = doRecognizePerson ,
        args=(faceNames, currentFaceID))
    t.start()
    currentFaceID += 1
```

ในส่วนการทำงาน threading.Thread() จะทำการเพิ่มค่าในการนับจำนวนคนโดยจะทำงานที่ฟังก์ชั่น doRecognizePerson ซึ่งจะมีการส่ง 2 ค่าไปยังฟังก์ชั่นได้แก่ จำนวนคนที่นับ และค่า fid ประจำรูปภาพใบหน้าไปทำการกำหนดต่อในฟังก์ชั่นเพื่อที่จะได้รับค่า saveimg[fid] ที่จะนำไปใช้ในส่วนของการบันทึกรูปภาพใบหน้าคน

```
def doRecognizePerson(faceNames, fid):
    faceNames[fid] = "Person " + str(fid)
    start_time[fid] = time.time()
    timestamp[fid] = time.strftime("%Y%m%d_%H%M%S",time.localtime())
    saveimg[fid] = False
```


เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.6 การเช็คค่ากำหนดที่บันทึกรูป และการบันทึกรูปภาพใบหน้าคน

ในส่วนของการออกแบบการเช็ครูปภาพใบหน้าคนที่มีการตรวจจับเป็นครั้งแรกจะได้รับค่า saveimg[fid] เป็นค่า False มาและจะได้รับ Key มาเพื่อที่จะนำมาใช้สำหรับในตรวจสอบแล้วทำการบันทึกรูปภาพใบหน้าคนเป็นครั้งแรกไปที่โฟลเดอร์ที่เก็บรูปภาพ ซึ่งจะแสดงการเขียนโปรแกรมไว้ที่ด้านล่าง

```
for fid in faceTrackers.keys():
        tracked_position =  faceTrackers[fid].get_position()
        t_x = int(tracked_position.left())
        t_y = int(tracked_position.top())
        t_w = int(tracked_position.width())
        t_h = int(tracked_position.height())
        cv2.rectangle(resultImage, (t_x, t_y),(t_x + t_w , t_y + t_h),
                rectangleColor ,2)
        if saveimg[fid] == False:
                f t_w > 0 and t_h > 0:
                if t_x < 0:
                        t_x = 0
                if t_y < 0:
                        t_y = 0
                if (t_x + t_w) > 319:
                        t_xx = 319
                if (t_y + t_h) > 239:
                        t_yy = 239
                t_xx = t_x+t_w
                t_yy = t_y+t_h
                img[fid] = baseImage[t_y:t_yy,t_x:t_xx]
                saveimg[fid] = True
                path = 'C:/Users/Samrit & Warisarapon/Desktop/Learn
                        OpenCV by pyimagesearch/Demo/OMAS/log/'
                pathimg[fid] = os.path.join(path, str(timestamp[fid]) + '_'
```


เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
                + str(fid) + '.jpg')
                cv2.imwrite(pathimg[fid],img[fid])
```

3.3.7 การอัพเดทการตรวจจับใบหน้าคน เช็คความกว้างรูปภาพ และเช็ค TrackingQuality

จากที่ได้บันทึกรูปภาพใบหน้าคน หลังจากนั้นจะทำการออกแบบในการอัพเดทรูปภาพใบหน้าคน เช็คความกว้างรูปภาพถ้าความกว้างของรูปภาพมีขนาดมากกว่า 52 พิกเซลจะทำการกำหนดค่าลบรูปภาพและการเช็ค TrackingQuality จะต้องไม่น้อยกว่า 7 โดยถ้าน้อยกว่า 7 จะทำการกำหนดค่าลบรูปภาพและทำการบันทึกระยะเวลาในการติดตามของใบหน้าคนลงไฟล์ Excel ที่ถูกจัดเก็บไว้ในโฟลเดอร์ซึ่งสามารถดูการเขียนโปรแกรมได้ด้านล่าง

```
for fid in faceTrackers.keys():
        trackingQuality = faceTrackers[ fid ].update( baseImage )
        tracked_position =  faceTrackers[fid].get_position()
        t_w = int(tracked_position.width())
        if t_w > 52:
            fidsToDelete.append( fid )
        if trackingQuality < 7:
            fidsToDelete.append( fid )
            end_time[fid] = time.time()
            duration = end_time[fid] - start_time[fid]
            with open("C:/Users/Samrit & Warisarapon/Desktop/Learn
                OpenCV by pyimagesearch/Demo/OMAS/log/log.csv",'a',
                newline='',encoding='utf-8') as csvfile:
            a = csv.writer(csvfile, delimiter= ',', quotechar='|', quoting=
                csv.QUOTE_MINIMAL)
            a.writerow([fid,timestamp[fid],duration,pathimg[fid]])
```

3.3.8 การเช็คค่าที่กำหนดลบรูปภาพ และทำการลบรูปภาพใบหน้าคน

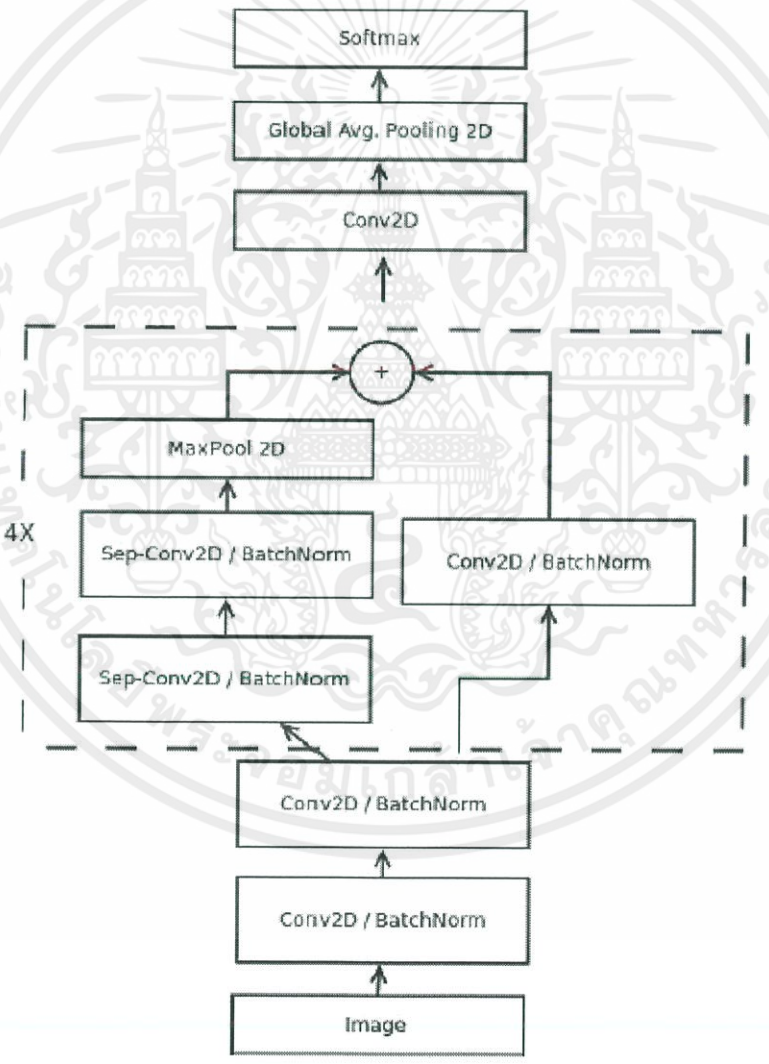
จากที่ได้ทำการอัพเดทการตรวจจับใบหน้าคน เช็คความกว้างรูปภาพ และเช็ค TrackingQuality ต่อจากนั้นจะทำการลบรูปภาพใบหน้าคนที่มีขนาดความกว้างเกิน 52 พิกเซล และค่า TrackingQuality น้อยกว่า 7 แสดงดังโปรแกรมด้านล่าง

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
for fid in fidsToDelete:
    faceTrackers.pop( fid , None )
```

3.4 การออกระบบการตรวจจับเพศและสีหน้า

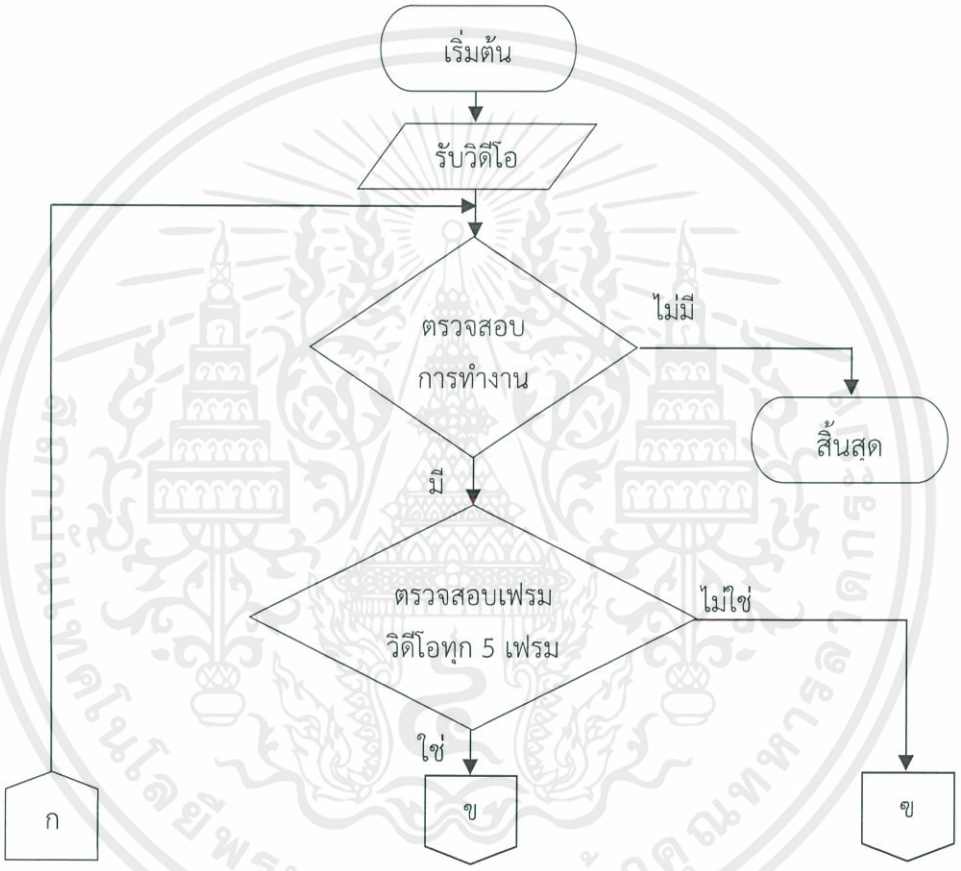
การออกแบบระบบตรวจจับเพศ และสีหน้าจะทำการออกแบบโดยใช้หลักการ Convolutional Neural Networks (CNNs) ในการออกแบบระบบโดยยึดหลักโครงสร้างตามโมเดลของ Real-time Convolutional Neural Networks for Emotion and Gender Classification (Octavio Arriaga, Paul G. Ploger, Matias Valdenegro) เป็นหลัก ซึ่งสามารถดูโมเดลได้ดังรูปที่ 3.17



รูปที่ 3.17 โมเดลของ Real-time Convolutional Neural Networks for Emotion and Gender Classification

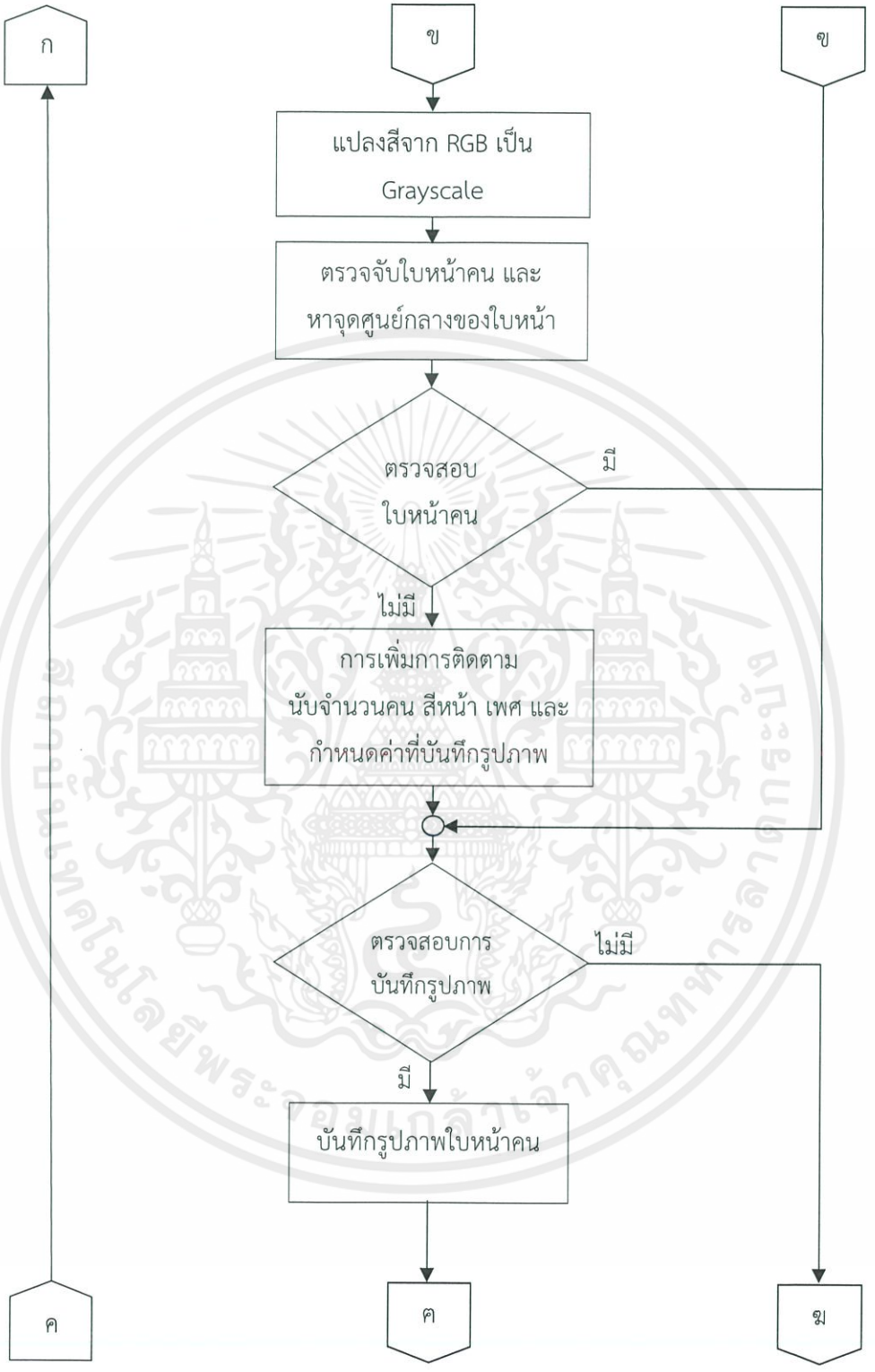
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การออกแบบระบบการตรวจจับเพศและสีหน้า ได้มีการออกมีการออกแบบการทำงานจริง แต่ไม่สามารถที่จะนำไปใช้งานในระบบที่สถานีทั้ง 3 สถานี เนื่องมาจากการประมวลผลมากทำให้ประสิทธิภาพการรอง ASUS Tinker Board ไม่สามารถรันโปรแกรมได้ แต่มีการออกแบบโฟวชาร์ตของการตรวจจับสีหน้า และเพศดังรที่ 3.17



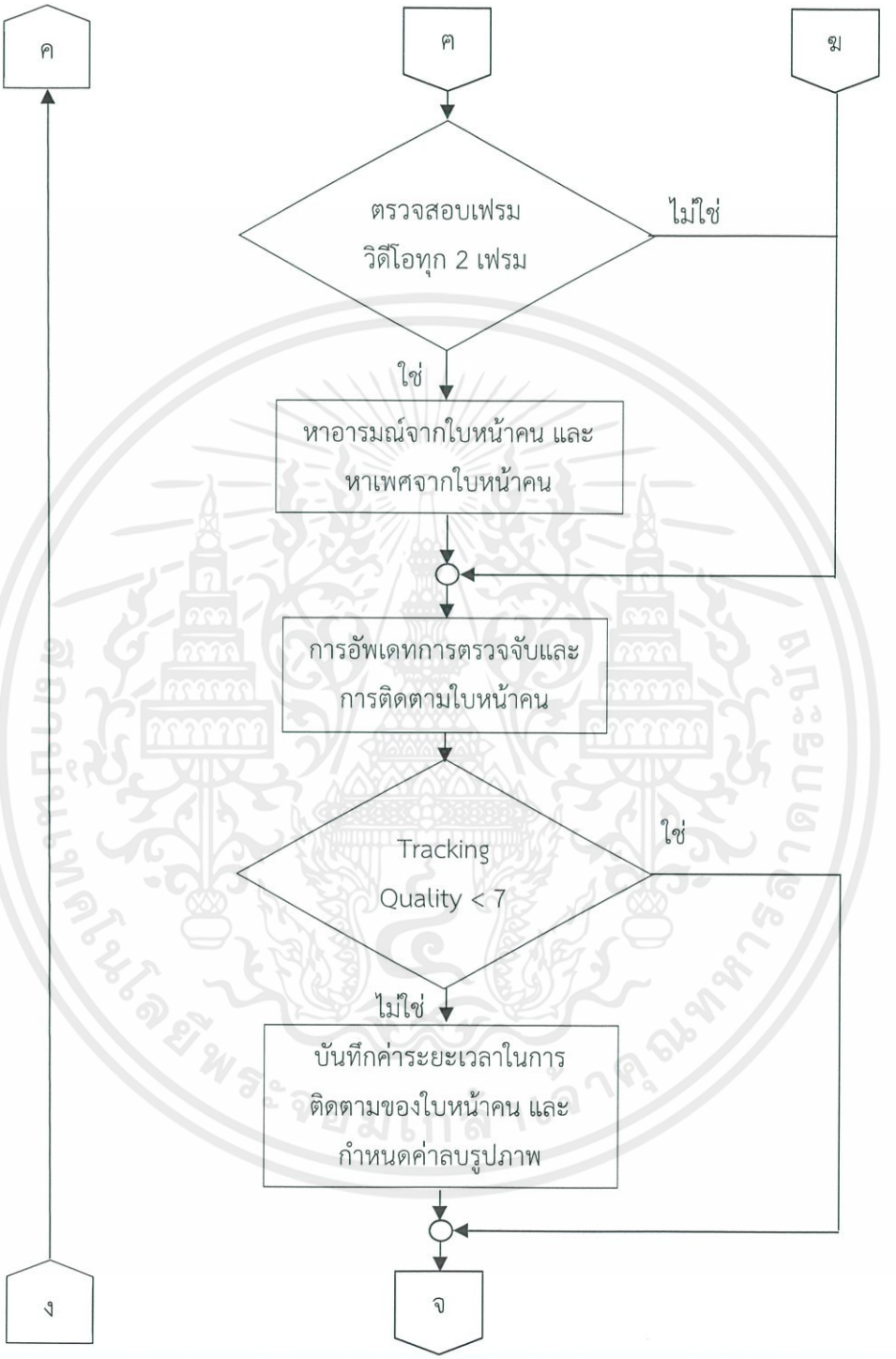
รูปที่ 3.18 โฟว์ชาร์ตของระบบการตรวจจับสีหน้าและเพศ

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



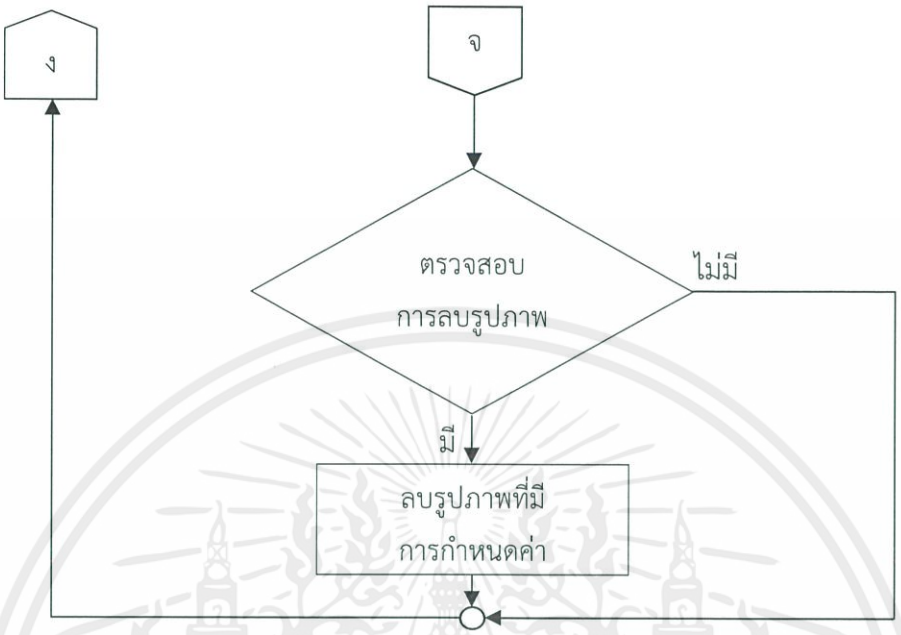
รูปที่ 3.18 (ต่อ) โฟว์ชาร์ตของระบบการตรวจจับสีหน้าและเพศ

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.18 (ต่อ) โฟว์ชาร์ตของระบบการตรวจจับสีหน้าและเพศ

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.18 (ต่อ) โฟว์ชาร์ตของระบบการตรวจจับสีหน้าและเพศ

ในรูปที่ 3.17 เป็นโฟว์ชาร์ตของระบบการตรวจจับสีหน้าและเพศในแต่ละขั้นตอนจะมีการออกแบบหน้าที่ในการทำงานดังนี้

3.4.1 การรับวิดีโอ

การรับวิดีโอจะเป็นการข้อมูลวิดีโอจากโมดูลกล้องมา โดยในการรับวิดีจากโมดูลกล้องแบบปกติจะมีการรับข้อมูลรูปภาพเข้ามาขนาด 640 x 480 พิกเซล ความเร็ว 120 เฟรมต่อวินาที ปริมาณพิกเซลใน 1 วินาทีจะมีจำนวนพิกเซลทั้งหมด 36,864,000 พิกเซลที่เข้าไปทำการประมวลผล โดยการตรวจจับสีหน้าและเพศจำเป็นที่ต้องใช้ขนาดรูปภาพ 640 x 480 พิกเซล เพื่อที่จะดูองค์ประกอบของใบหน้าเพื่อหาอารมณ์และเพศ ซึ่งที่จะแสดงการเขียนโปรแกรมได้ดังนี้

```
def detectAndTrackMultipleFaces():
    cap = cv2.VideoCapture(0)
```

3.4.2 การแปลงสีจาก RGB เป็น Grayscale

ในส่วนของการทำงานในการแปลงสีจาก RGB เป็น Grayscale หลังจากที่ได้รับวิดีโอขนาด 640 x 480 พิกเซล ความเร็ว 120 เฟรมต่อวินาที ปริมาณพิกเซลใน 1 วินาทีจะมีจำนวนพิกเซลทั้งหมด 36,864,000 พิกเซลจะทำการออกแบบด้วยการแปลงสีจาก RGB เป็น Grayscale เนื่องจากขนาดของรูปภาพสีแบบ RGB จะมีความลึกของสีทั้งหมด 3 สีทำให้รูปภาพสีแบบ RGB ในแต่ละพิกเซลสีจะมีจำนวนบิตข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ทั้งหมด 8 บิตต่อพิกเซลซึ่งทำให้ใน 1 พิกเซลรูปภาพสีแบบ RGB จะมีจำนวนบิตข้อมูลทั้งหมด 24 บิตต่อพิกเซล ทำให้จำนวนบิตที่ประมวลผลรูปภาพทั้งหมดใน 1 วินาทีมีจำนวนบิตทั้งหมด 884,736,000 บิตต่อวินาทีที่ได้ไปทำการประมวลผลจึงทำให้ในการประมวลผลไม่ต่อเนื่องและเกิดการดีเลย์ ส่งผลให้ต้องนำการออกแบบด้วยการแปลงสี RGB เป็น Grayscale ด้วยแปลงเป็นรูปภาพสีแบบ Grayscale จะมีจำนวนบิตข้อมูลทั้งหมด 24 บิตต่อพิกเซลทำให้จำนวนบิตที่ประมวลผลรูปภาพทั้งหมใน 1 วินาทีจะมีจำนวนบิตทั้งหมด 294,912,00 บิตต่อวินาที ซึ่งทำให้ความเร็วในการประมวลผลมีความเร็วขึ้น 3 เท่าเพราะว่าการตัดสีในส่วนที่ไม่จำเป็นต่อการใช้งานออกไปจึงทำให้ระบบประมวลผลเร็วขึ้น สามารถแสดงการเขียนโปรแกรมได้ดังนี้

```
rc,fullSizeBaseImage = cap.read()
baseImage = cv2.resize( fullSizeBaseImage, ( 640, 480))
if (frameCounter % 5) == 0:
    gray = cv2.cvtColor(baseImage, cv2.COLOR_BGR2GRAY)
```

3.4.3 การตรวจจับใบหน้าคน และหาจุดศูนย์การของใบหน้า

ส่วนของการตรวจจับใบหน้าคน และหาจุดศูนย์กลางของใบหน้า จะทำการออกแบบด้วยการกำหนดค่า Image Pyramids หรือ scaleFactor ให้ขนาดในการลดรูปภาพลดลงทุกๆ 10 % ในการตรวจจับใบหน้าคน ต่อมาทำการออกแบบในการหาใบหน้าคนด้วยการตรวจหาใบหน้าที่ซ้ำกันต่ำสุด 5 หน้าจึงจะตัดสินใจว่าใบหน้านั้นคือใบหน้าคน ส่วนของ minSize จะเป็นตัวกำหนดกรอบในการตรวจจับใบหน้าคนขนาดต่ำสุด 100 x 100 พิกเซล เพราะในการตรวจจับใบหน้าเพื่อที่จะไปทำการจำแนกสีหน้าและเพศจำเป็นที่จะต้องมีขนาดของใบหน้าไม่ต่ำกว่า 100 x 100 พิกเซลที่ไปทำการประมวลผล และการกำหนดขนาดต่ำสุดยังสามารถที่จะลดความผิดพลาดที่เกิดขึ้นจากตรวจจับใบหน้าคนได้ พร้อมทั้งโปรแกรมจะทำการคำนวณหาจุดศูนย์กลางของใบหน้าคนเพื่อที่จะเก็บค่าไว้เช็คว่าคนเดิมหรือคนใหม่พร้อมกับเพิ่มค่า matchedFid สำในการเช็ครูปภาพว่าเป็นใบหน้าคนเดิมหรือไม่ ซึ่งสามารถแสดงการเขียนโปรแกรมได้ดังนี้

```
faces =LBPCascade.detectMultiScale(gray,scaleFactor=1.10,minNeighbors=5,
        minSize=(100, 100))
for (_x,_y,_w,_h) in faces:
        x = int(_x)
        y = int(_y)
        w = int(_w)
        h = int(_h)
        x_bar = x + 0.5 * w
```


เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
y_bar = y + 0.5 * h
matchedFid = None
```

3.4.4 การตรวจสอบใบหน้าคน

หลังจากที่ได้ทำการตรวจจับใบหน้าคน และหาจุดศูนย์การของใบหน้าต่อไปจะทำการเช็คใบหน้าคนว่าเป็นใบหน้าของคนเดิมหรือไม่ โดยในการออกแบบจะทำการอัพเดทข้อมูลตำแหน่งของรูปภาพใบหน้าและเช็คว่าคือตำแหน่งเดิมไม ถ้าเป็นรูปภาพใบหน้าคนเดิมจะมี keys สำหรับในการเก็บค่า fid เอาไว้เพื่อที่จะเข้าไปทำงานใน faceTrackers.keys() ดังที่แสดงการเขียนโปรแกรมไว้ด้านล่าง แต่ถ้ารูปภาพใบหน้าคนใหม่ก็จะไปสู่กระบวนการการเพิ่มการติดตาม นับจำนวนคน และกำหนดค่าที่บันทึกรูปที่จะกล่าวในส่วนถัดไป

```
for fid in faceTrackers.keys():
    tracked_position =  faceTrackers[fid].get_position()
    t_x = int(tracked_position.left())
    t_y = int(tracked_position.top())
    t_w = int(tracked_position.width())
    t_h = int(tracked_position.height())
    t_x_bar = t_x + 0.5 * t_w
    t_y_bar = t_y + 0.5 * t_h
    if ( ( t_x <= x_bar   <= (t_x + t_w)) and
            ( t_y <= y_bar   <= (t_y + t_h)) and
            ( x   <= t_x_bar <= (x   + w  )) and
            ( y   <= t_y_bar <= (y   + h  ))):
            matchedFid = fid
```

3.4.5 การเพิ่มการติดตาม นับจำนวนคน สีหน้า เพศ และกำหนดค่าที่บันทึกรูป

ในส่วนของการเพิ่มการติดตาม นับจำนวนคน สีหน้า เพศ และกำหนดค่าที่บันทึกรูป จะออกแบบด้วยการเพิ่มการติดตาม โดยถ้าค่า matchedFid เป็น None หมายความว่ารูปภาพใบหน้าคนนี้เป็นอันใหม่ที่ตรวจจับพบ จากนั้นก็จะทำการเพิ่มการติดตามด้วย dlib และนับจำนวนด้วยการใช้ threading.Thread() และกำหนดค่าที่บันทึกรูป ซึ่งจะสามารถแสดงการเขียนโปแกรมได้ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
if matchedFid is None:
    tracker = dlib.correlation_tracker()
    tracker.start_track(baseImage,
    dlib.rectangle( x, y, x+w, y+h))
    faceTrackers[ currentFaceID ] = tracker
    t = threading.Thread( target = doRecognizePerson ,
        args=(faceNames, currentFaceID,
        faceEmotion , faceGenders))
    t.start()
    currentFaceID += 1
```

ในส่วนการทำงาน threading.Thread() จะทำการเพิ่มค่าในการนับจำนวนคนโดยจะทำงานที่ฟังก์ชั่น doRecognizePerson ซึ่งจะมีการส่ง 4 ค่าไปยังฟังก์ชั่นได้แก่ จำนวนคนที่นับ สีหน้า เพศ และค่า fid ประจำรูปภาพใบหน้าไปทำการกำหนดต่อในฟังก์ชั่นเพื่อที่จะได้รับค่า saveimg[fid] ที่จะนำไปใช้ในส่วนของการบันทึกรูปภาพใบหน้าคนและจะนำค่าที่ได้จากสีหน้าและเพศไปงานต่อในส่วนต่อไป

```
def doRecognizePerson(faceNames, fid, faceEmotion, faceGenders):
    faceNames[fid] = "Person " + str(fid)
    faceEmotion[fid] = “ ”
    faceGenders[fid] = “ ”
    start_time[fid] = time.time()
    timestamp[fid] = time.strftime("%Y%m%d_%H%M%S",time.localtime())
    saveimg[fid] = False
```

3.4.6 การเช็คค่ากำหนดที่บันทึกรูป และการบันทึกรูปภาพใบหน้าคน

ในส่วนของการออกแบบการเช็ครูปภาพใบหน้าคนที่มีการตรวจจับเป็นครั้งแรกจะได้รับค่า saveimg[fid] เป็นค่า False มาและจะได้รับ Key มาเพื่อที่จะนำมาใช้สำหรับในตรวจสอบแล้วทำการบันทึกรูปภาพใบหน้าคนเป็นครั้งแรกไปที่โฟลเดอร์ที่เก็บรูปภาพ พร้อมทั้งจะมีการทำงานในส่วนของการตรวจหาสีหน้าและเพศในส่วนของฟังก์ชั่น detectioneMotion() และฟังก์ชั่น detectionGenders () ซึ่งจะแสดงการเขียนโปรแกรมไว้ที่ด้านล่าง

```
for fid in faceTrackers.keys():
    tracked_position =  faceTrackers[fid].get_position()
```


เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
t_x = int(tracked_position.left())
t_y = int(tracked_position.top())
t_w = int(tracked_position.width())
t_h = int(tracked_position.height())
cv2.rectangle(resultImage, (t_x, t_y),(t_x + t_w , t_y + t_h),
        rectangleColor ,2)

if (frameCounter % 2) == 0:
        detectionMotion(baseImage[t_y-30:t_y+t_h+30,t_x-
        20:t_x+t_w+20], fid, faceEmotion)
        detectionGenders(baseImage[t_y-30:t_y+t_h+30,t_x-
        20:t_x+t_w+20], fid, faceGenders)
if saveimg[fid] == False:
        if t_w > 0 and t_h > 0:
                if t_x < 0:
                        t_x = 0
                if t_y < 0:
                        t_y = 0
                if (t_x + t_w) > 639:
                        t_xx = 639
                if (t_y + t_h) > 239:
                        t_yy = 239
                t_xx = t_x+t_w
                t_yy = t_y+t_h
                img[fid] = baseImage[t_y-30:t_yy+30,t_x-20:t_xx+20]
                saveimg[fid] = True
                path = 'C:/Users/Samrit & Warisarapon/Desktop/Learn
                        OpenCV by pyimagesearch/Demo/OMAS/log/'
                        pathimg[fid] = os.path.join(path,
                        str(timestamp[fid]) + '_' + str(fid) + '.jpg')
                cv2.imwrite(pathimg[fid],img[fid])
```


เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.4.7 การอัพเดทการตรวจจับใบหน้าคน และเช็ค TrackingQuality

จากที่ได้บันทึกรูปภาพใบหน้าคน หลังจากนั้นจะทำการออกแบบในการอัพเดทรูปภาพใบหน้าคน การเช็ค TrackingQuality จะต้องไม่น้อยกว่า 7 โดยถ้าน้อยกว่า 7 จะทำการกำหนดค่าลบรูปภาพและทำการบันทึกระยะเวลาในการตรวจจับของใบหน้าคนลงไฟล์ Excel ที่ถูกจัดเก็บไว้ในโฟลเดอร์ ซึ่งสามารถดูการเขียนโปรแกรมได้ด้านล่าง

```
for fid in faceTrackers.keys():
    trackingQuality = faceTrackers[ fid ].update( baseImage )
    tracked_position =  faceTrackers[fid].get_position()
    t_w = int(tracked_position.width())
    if trackingQuality < 7:
        fidsToDelete.append( fid )
        end_time[fid] = time.time()
        duration = end_time[fid] - start_time[fid]
        with open("C:/Users/Samrit & Warisarapon/Desktop/Learn
            OpenCV by pyimagesearch/Demo/OMAS/log/log.csv",'a',
            newline='',encoding='utf-8') as csvfile:
        a = csv.writer(csvfile, delimiter= ',', quotechar='|', quoting=
            csv.QUOTE_MINIMAL)
        a.writerow([fid,timestamp[fid],duration,pathimg[fid]])
```

3.4.8 การเช็คค่าที่กำหนดลบรูปภาพ และทำการลบรูปภาพใบหน้าคน

จากที่ได้ทำการอัพเดทการตรวจจับใบหน้าคน และเช็ค TrackingQuality ต่อจากนั้นจะทำการลบรูปภาพใบหน้าคนที่มค่า TrackingQuality น้อยกว่า 7 จะแสดงดังโปรแกรมด้านล่าง

```
for fid in fidsToDelete:
    faceTrackers.pop( fid , None )
```


เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

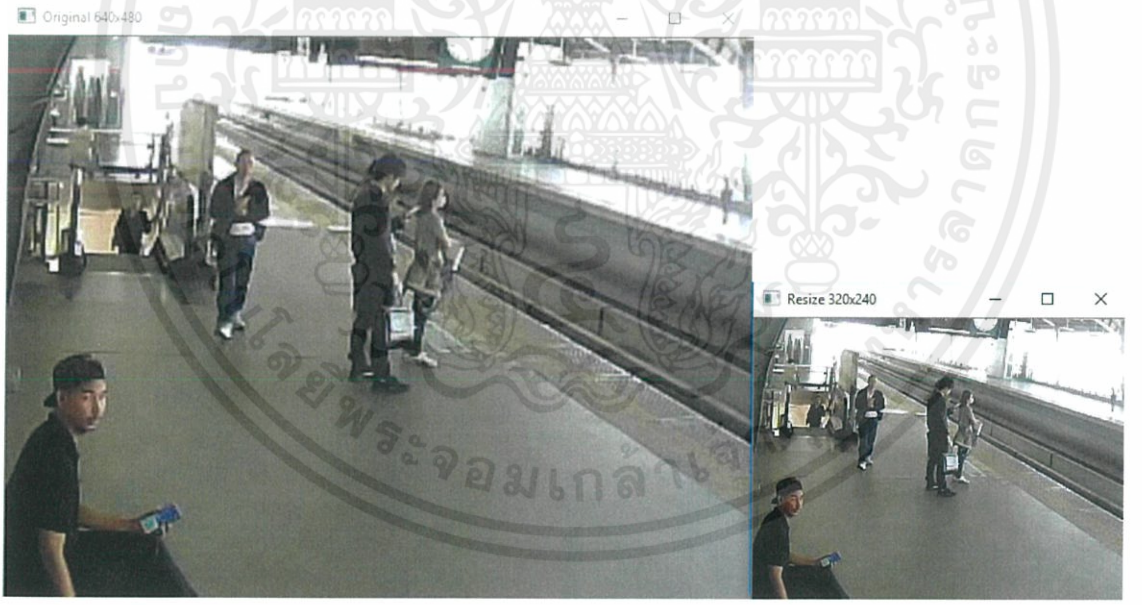
ผลการวิจัย

4.1 ระบบการตรวจจับและนับจำนวนคน

จากในบทที่ 3.3 การออกแบบระบบการตรวจจับและนับจำนวนคนต่อไปจะทำการแสดงผลจากการทำงานในแต่ละขั้นตอนที่ได้ทำการออกแบบการทดลอง สามารถแสดงได้ดังนี้

4.1.1 การลดขนาดของรูปภาพ

การลดขนาดของรูปภาพสามารถที่จะทำให้ระบบประมวลผลเร็วขึ้นถึง 4 เท่า เพราะว่าปริมาณของข้อมูลที่เข้าไปประมวลมีขนาดที่น้อยลง 4 เท่าโดยสามารถแสดงการลดขนาดได้ดังรูปที่ 4.1 และปริมาณข้อมูลรูปภาพที่น้อยลง 4 เท่าได้ในรูปที่ 4.2



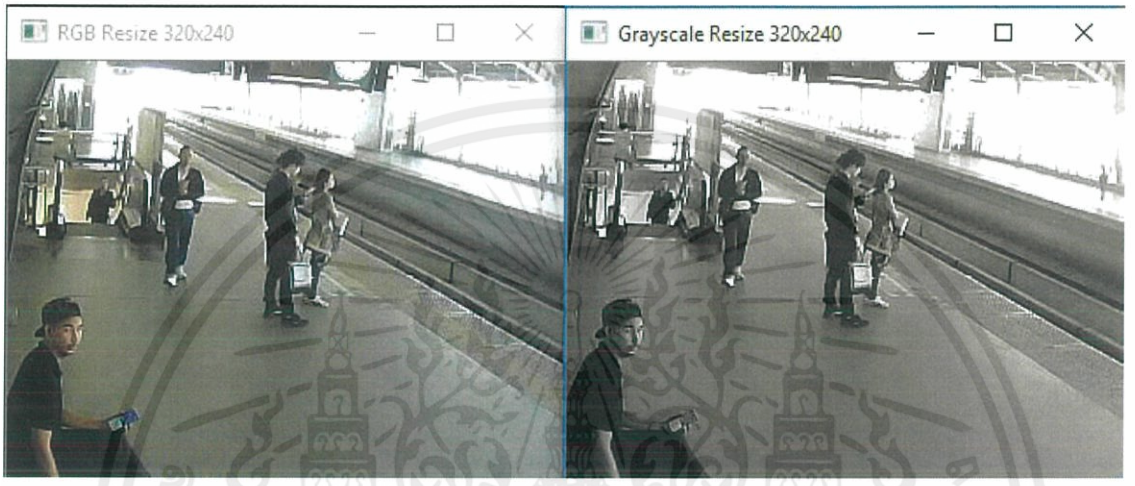
รูปที่ 4.1 การลดขนาดของรูปภาพ

รูปที่ 4.2 ปริมาณข้อมูลของรูปภาพระหว่างรูปภาพต้นฉบับกับรูปภาพที่ลดขนาด

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.2 การแปลงสีจาก RGB เป็น Grayscale

การแปลงสีจาก RGB เป็น Grayscale จะสามารถทำให้ระบบประมวลผลเร็วขึ้น 3 เพราะปริมาณข้อมูลที่เข้าไปทำการตรวจจับและนับจำนวนคนมีขนาดที่เล็กลงถึง 3 เท่าโดยสามารถดูการแปลงสีจาก RGB เป็น Grayscale ดังรูปที่ 4.3 และดูปริมาณข้อมูลที่ลดลงได้ในรูปที่ 4.4

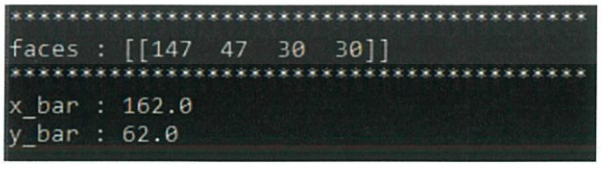


รูปที่ 4.3 การแปลงสีจาก RGB เป็น Grayscale

รูปที่ 4.4 ปริมาณข้อมูลของรูปภาพระหว่างรูปภาพสี RGB กับรูปภาพสี Grayscale

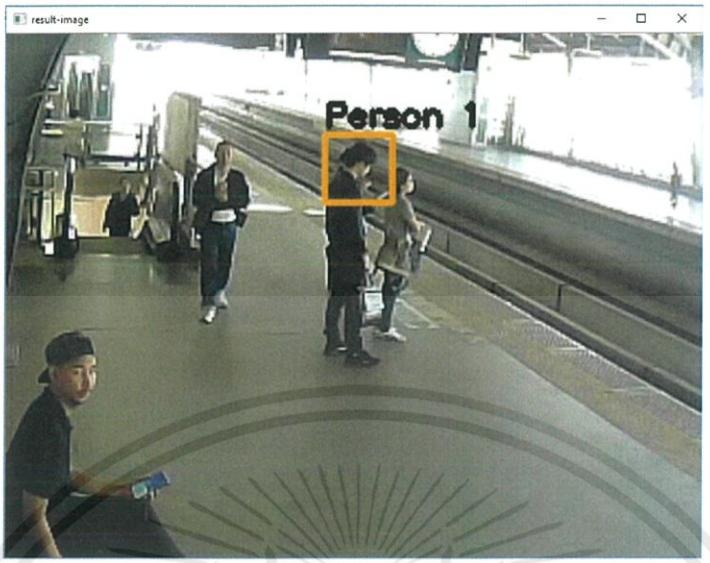
4.1.3 การตรวจจับใบหน้าคน และหาจุดศูนย์การของใบหน้า

ในรูปที่ 4.5 การตรวจจับใบหน้าคนจะพบกับตำแหน่งของคนแล้วไปทำการคำนวณจุดศูกย์กลางของรูปภาพเพื่อเก็บไว้ในฐานข้อมูลสำหรับนำไปเช็คอีกครั้งในตอนที่ตรวจจับว่าใช่คนเดิมหรือคนใหม่ พร้อมทั้งแสดงผลจะการตรวจจับใบหน้าคนได้ในรูปที่ 4.6



รูปที่ 4.5 ค่าการตรวจจับใบหน้าคน และค่าจุดศูนย์การของใบหน้า

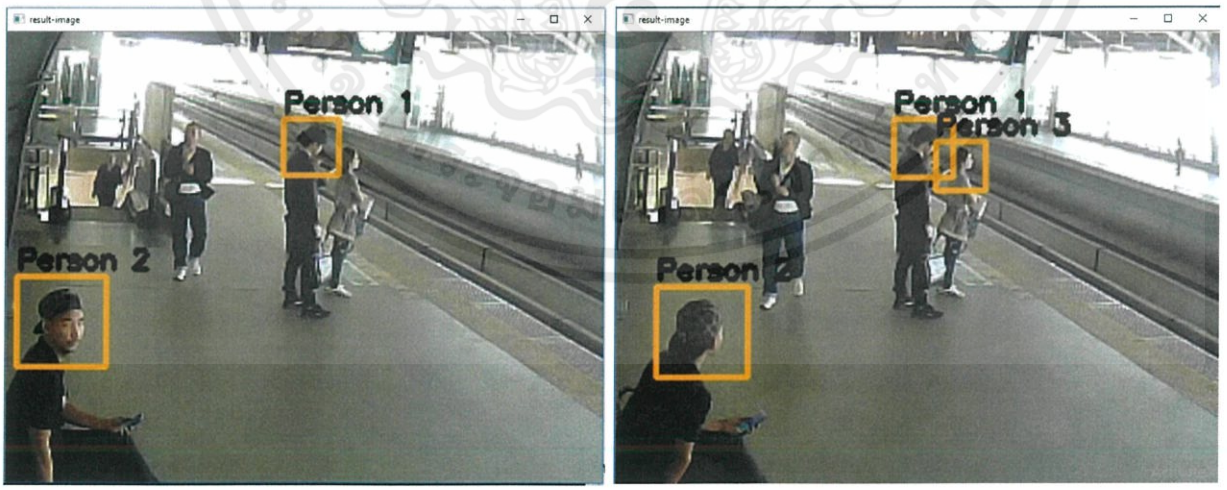
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.6 ผลจากการตรวจจับใบหน้าคน

4.1.4 การตรวจสอบใบหน้าคน

ในการตรวจสอบว่าใบหน้าคนนี้จะมีการเช็คจากจุดศูนย์กลางของรูปภาพที่ได้ทำการตรวจจับโดย ถ้าใบหน้าคนเป็นภาพเดิมจะไม่ทำการตรวจจับแต่ถ้าใบหน้าคนนั้นเป็นใบหน้าคนใหม่จะทำการตรวจจับใหม่ เพื่อลดการซ่อนทับกันของการตรวจจับใบหน้าคน โดยสามารถแสดงการทำงานได้ในรูปที่ 4.7 (ก) รูปภาพที่ได้ทำการเพิ่มรูปภาพใบหน้าคน ส่วนรูปภาพ (ข) รูปภาพที่ผ่านกระบวนการตรวจสอบใบหน้าคนและไม่ทำการตรวจจับใบหน้าเดิม



(ก) (ข)

รูปที่ 4.7 ผลจากการตรวจสอบใบหน้าคน

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.5 การเพิ่มการติดตาม นับจำนวนคน และกำหนดค่าที่บันทึกรูป

จากที่ได้ทำการตรวจจับใบหน้าคนเจอเป็นครั้งแรกระบบจะทำการประมวลผลโดยจะทำการเพิ่มการติดติดตามและทำการนับจำนวนคน พร้อมทั้งเก็บค่าวันเดือนปีเวลา และกำหนดค่าที่จะบันทึกรูปภาพเพื่อจะไปทำการบันทึกต่อในส่วนทัดไป โดยในการทำการเพิ่มการติดตาม นับจำนวนคน และกำหนดค่าที่บันทึกรูปสามารถแสดงได้ในรูปที 4.8 ทำการนับจำนวนคน เก็บค่าวันเดือนปีเวลา และเก็บค่าการบันทึกรูปภาพ

```
**********************************************
fid : 1
faceNames : {1: 'Person 1'}
timestamp : {1: '20181114_195428'}
saveimg : {1: False}
**********************************************
```

รูปที่ 4.8 การเก็บค่านับจำนวนคนและค่าที่บันทึกรูปภาพ

4.1.6 การบันทึกรูปภาพใบหน้าคน

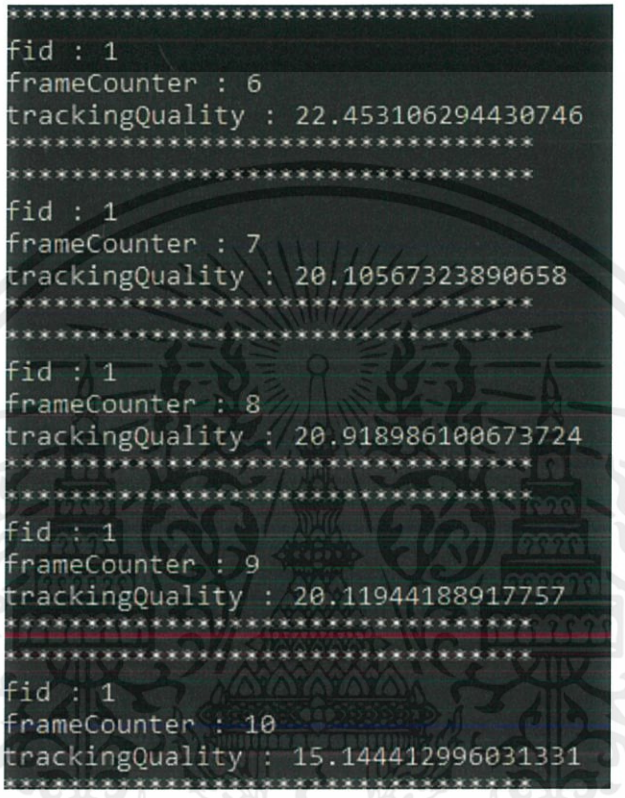
จากที่ได้ทำการกำหนดค่าที่บันทึกรูปภาพไว้เป็น False จากนั้นจะทำการนำค่าที่ได้ไปทำการเช็คค่าในการบันทึกรูปภาพเพื่อจะเก็บฐานข้อมูลรูปภาพคนที่ตรวจจับใบหน้าคนเป็นครั้งแรก สามารถแสดงผลลัพธ์ของการบันทึกได้จากรูปที่ 4.9

รูปที่ 4.9 ผลจากการบันทึกรูปภาพใบหน้าคน

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.7 การอัพเดทการตรวจจับใบหน้าคน

การอัพเดทการตรวจจับใบหน้าคนจะมีการอัพเดทค่าทุกการทำงานของระบบประมวลผล ซึ่งสามารถแสดงได้ดังรูปที่ 4.11

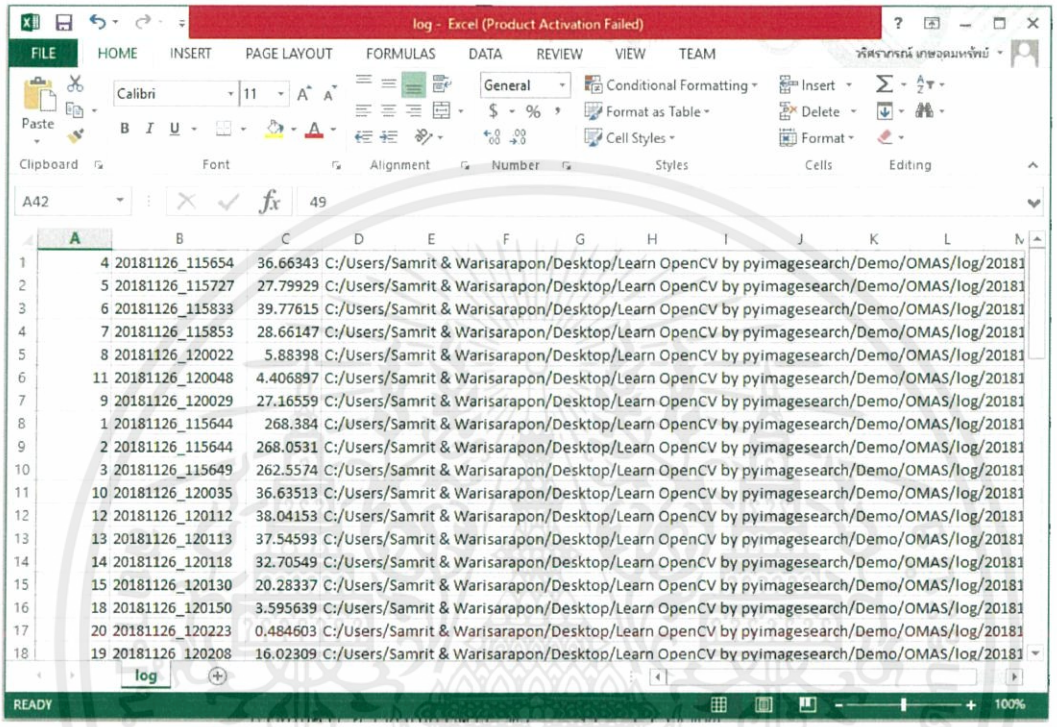


รูปที่ 4.10 แสดงค่าการอัพเดทการตรวจจับใบหน้าคน

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.8 การบันทึกระยะเวลาในการตรวจจับใบหน้าคน

ในการบันทึกระยะเวลาในการตรวจจับใบหน้าคนจะทำการเก็บข้อมูลอยู่ในรูปของไฟล์ Excel ซึ่งจะแสดงได้ดังรูปที่ 4.12



log - Excel (Product Activation Failed)

	A	B	C	D
1	4	20181126_115654	36.66343	C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by pyimagesearch/Demo/OMAS/log/20181
2	5	20181126_115727	27.79929	C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by pyimagesearch/Demo/OMAS/log/20181
3	6	20181126_115833	39.77615	C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by pyimagesearch/Demo/OMAS/log/20181
4	7	20181126_115853	28.66147	C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by pyimagesearch/Demo/OMAS/log/20181
5	8	20181126_120022	5.88398	C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by pyimagesearch/Demo/OMAS/log/20181
6	11	20181126_120048	4.406897	C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by pyimagesearch/Demo/OMAS/log/20181
7	9	20181126_120029	27.16559	C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by pyimagesearch/Demo/OMAS/log/20181
8	1	20181126_115644	268.384	C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by pyimagesearch/Demo/OMAS/log/20181
9	2	20181126_115644	268.0531	C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by pyimagesearch/Demo/OMAS/log/20181
10	3	20181126_115649	262.5574	C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by pyimagesearch/Demo/OMAS/log/20181
11	10	20181126_120035	36.63513	C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by pyimagesearch/Demo/OMAS/log/20181
12	12	20181126_120112	38.04153	C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by pyimagesearch/Demo/OMAS/log/20181
13	13	20181126_120113	37.54593	C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by pyimagesearch/Demo/OMAS/log/20181
14	14	20181126_120118	32.70549	C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by pyimagesearch/Demo/OMAS/log/20181
15	15	20181126_120130	20.28337	C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by pyimagesearch/Demo/OMAS/log/20181
16	18	20181126_120150	3.595639	C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by pyimagesearch/Demo/OMAS/log/20181
17	20	20181126_120223	0.484603	C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by pyimagesearch/Demo/OMAS/log/20181
18	19	20181126_120208	16.02309	C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by pyimagesearch/Demo/OMAS/log/20181

รูปที่ 4.11 การบันทึกระยะเวลาในการตรวจจับใบหน้าคน

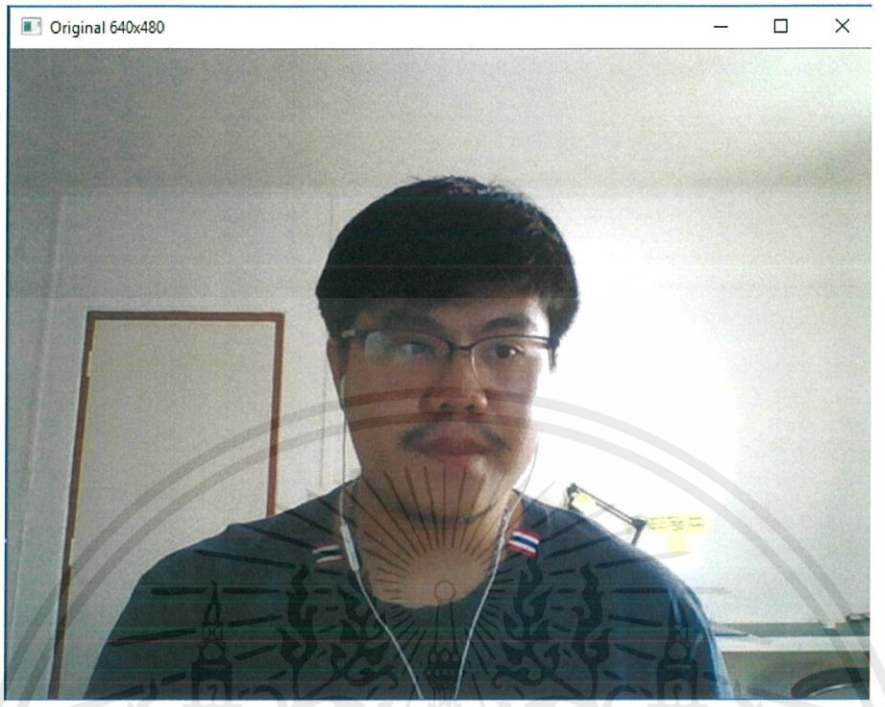
4.2 ระบบการจำแนกเพศและสีหน้า

จากในบทที่ 3.4 การออกแบบระบบการจำแนกเพศและสีหน้าต่อไปจะทำการแสดงผลจากการทำงานในแต่ละขั้นตอนที่ได้ทำการออกแบบการทดลอง สามารถแสดงได้ดังนี้

4.2.1 การรับวิดีโอ

การรับวิดีโอจะมีการรับวิดีโอขนาด 640 x 480 พิกเซลเข้ามา เพื่อใช้สำหรับในการจำแนกเพศและสีหน้าในการประมวลผล ซึ่งจะสามารถดูการรับวิดีโอได้ดังรูปที่ 4.12 และสามารถดูปริมาณข้อมูลรูปภาพได้ในรูปที่ 4.2

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.12 การรับวิดีโอใบหน้าคน

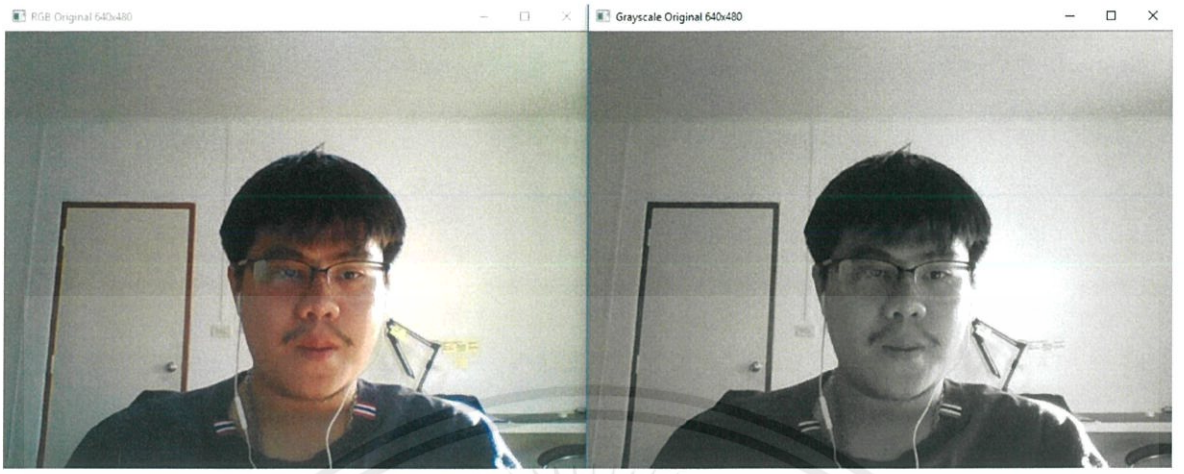
```
Size Original 640x480 : 921728
```

รูปที่ 4.13 ปริมาณข้อมูลของรูปภาพ

4.2.2 การแปลงสีจาก RGB เป็น Grayscale

การแปลงสีจาก RGB เป็น Grayscale จะสามารถทำให้ระบบประมวลผลเร็วขึ้น 3 เพราะปริมาณข้อมูลที่เข้าไปทำการจำแนกเพศและสีหน้ามีขนาดเล็กลงถึง 3 เท่าโดยสามารถดูการแปลงสีจาก RGB เป็น Grayscale ดังรูปที่ 4.14 และดูปริมาณข้อมูลที่ลดลงได้ในรูปที่ 4.15

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.14 การแปลงสีจาก RGB เป็น Grayscale ของภาพใบหน้าคน

```
**********************************************************
Size RGB Original 640x480 : 921728
Size Grayscale Original 640x480 : 307312
**********************************************************
```

รูปที่ 4.15 ปริมาณข้อมูลของรูปภาพระหว่างรูปภาพสี RGB กับรูปภาพสี Grayscale ของภาพใบหน้าคน

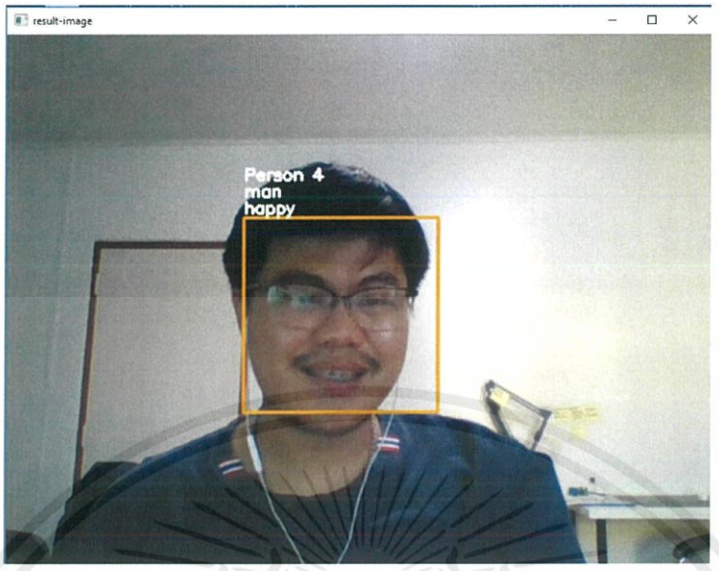
4.2.3 การตรวจจับใบหน้าคน และหาจุดศูนย์การของใบหน้า

ในรูปที่ 4.16 การจำแนกเพศและสีหน้าจะพบกับตำแหน่งของใบหน้าคนแล้วจะไปทำการคำนวณจุดศูกย์กลางของรูปภาพเพื่อเก็บไว้ในฐานข้อมูลสำหรับนำไปเช็คอีกครั้งในตอนที่ตรวจจับว่าใช่คนเดิมหรือคนใหม่ พร้อมทั้งแสดงผลจะการตรวจจับใบหน้าคนได้ในรูปที่ 4.17

```
**********************************************************
faces : [[215 175 173 173]]
**********************************************************
x_bar : 301.5
y_bar : 261.5
```

รูปที่ 4.16 ค่าการตรวจจับใบหน้าคน และค่าจุดศูนย์การของใบหน้า

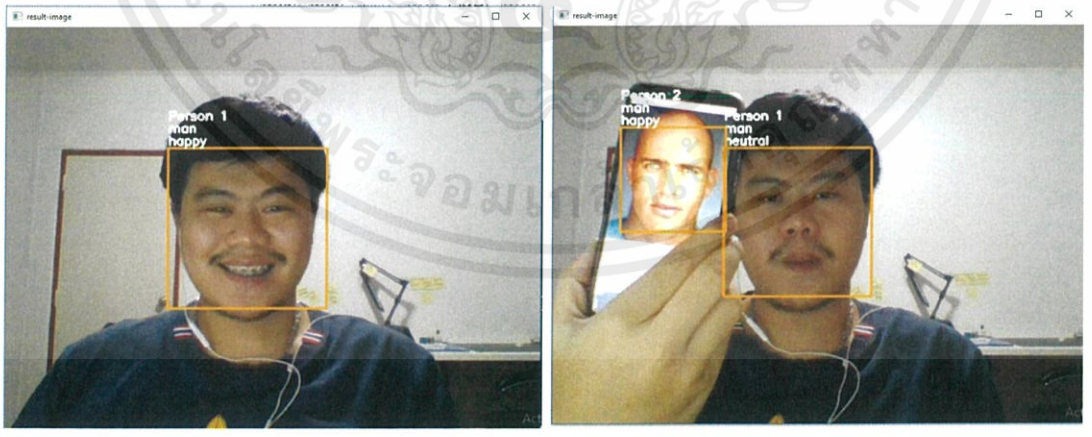
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพ 4.17 ผลจากการจำแนกเพศและสีหน้า

4.2.4 การตรวจสอบใบหน้าคน

ในการตรวจสอบว่าใบหน้าคนนี้จะมีการเช็คจากจุดศูกย์กลางของรูปภาพที่ได้ทำการตรวจจับโดย ถ้าใบหน้าคนเป็นภาพเดิมจะไม่ทำการตรวจจับแต่ถ้าใบหน้าคนนั้นเป็นใบหน้าคนใหม่จะทำการตรวจจับใหม่ เพื่อลดการซ่อนทับกันของการตรวจจับใบหน้าคน โดยสามารถแสดงการทำงานได้ในรูปที่ 4.18 (ก) รูปภาพที่ได้ทำการเพิ่มรูปภาพใบหน้าคน ส่วนรูปภาพ (ข) รูปภาพที่ผ่านกระบวนการตรวจสอบใบหน้าคนและไม่ทำการตรวจจับใบหน้าเดิม



(ก) (ข)

รูปที่ 4.18 ผลจากการตรวจสอบใบหน้าคนของการจำแนกเพศและสีหน้า

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.5 การเพิ่มการติดตาม นับจำนวนคน และกำหนดค่าที่บันทึกรูป

จากที่ได้ทำการตรวจจับใบหน้าคนเจอเป็นครั้งแรกระบบจะทำการประมวลผลโดยจะทำการเพิ่มการติดติดตามและทำการนับจำนวนคน พร้อมทั้งเก็บค่าวันเดือนปีเวลา สีหน้า เพศ และกำหนดค่าที่จะบันทึกรูปภาพเพื่อจะไปทำการบันทึกต่อในส่วนทัดไป โดยในการทำการเพิ่มการติดตาม นับจำนวนคน และกำหนดค่าที่บันทึกรูป สามารถแสดงได้ในภาพที่ 4.19 ทำการนับจำนวนคน เก็บค่าวันเดือนปีเวลา สีหน้า เพศ และเก็บค่าการบันทึกรูปภาพ

```
**************************************************
fid : 2
faceNames : {1: 'Person 1', 2: 'Person 2'}
faceEmotion : {1: 'sad', 2: ' '}
faceGenders : {1: 'man', 2: ' '}
timestamp : {1: '20181213_180116', 2: '20181213_180151'}
saveimg : {1: True, 2: False}
**************************************************
```

รูปที่ 4.19 การเก็บค่านับจำนวนคนและค่าที่บันทึกรูปภาพของการจำแนกเพศและสีหน้า

4.2.6 การบันทึกรูปภาพใบหน้าคน

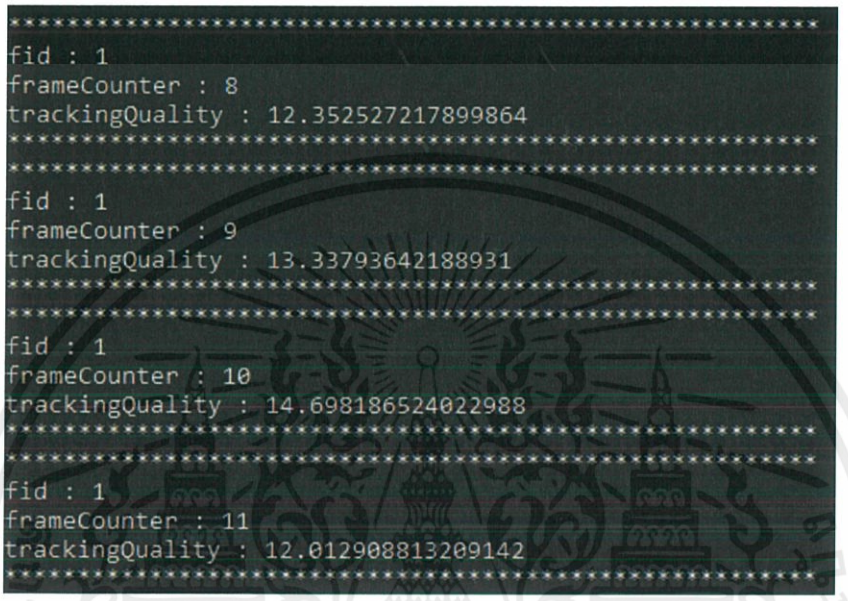
จากที่ได้ทำการกำหนดค่าที่บันทึกรูปภาพไว้เป็น False จากนั้นจะทำการนำค่าที่ได้ไปทำการเช็คค่าในการบันทึกรูปภาพเพื่อจะเก็บฐานข้อมูลรูปภาพคนที่ตรวจจับใบหน้าคนเป็นครั้งแรก สามารถแสดงผลลัพธ์ของการบันทึกได้จากรูปที่ 4.20

รูปที่ 4.20 ผลจากการบันทึกรูปภาพใบหน้าคนของการจำแนกเพศและสีหน้า

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.7 การอัพเดทการตรวจจับใบหน้าคน

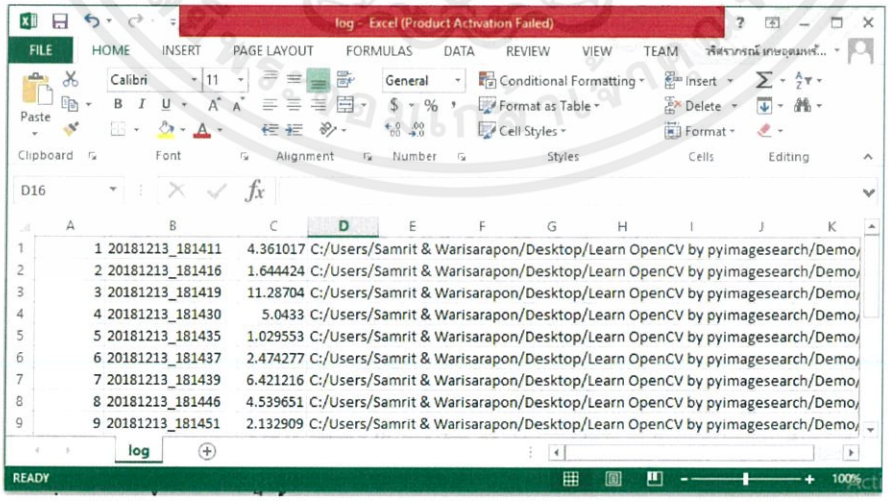
การอัพเดทการตรวจจับใบหน้าคนจะมีการอัพเดทค่าทุกการทำงานของระบบประมวลผล ซึ่งสามารถแสดงได้ดังรูปที่ 4.21



รูปที่ 4.21 แสดงค่าการอัพเดทการตรวจจับใบหน้าคนของการจำแนกเพศและสีหน้า

4.2.8 การบันทึกระยะเวลาในการตรวจจับใบหน้าคน

ในการบันทึกระยะเวลาในการตรวจจับใบหน้าคนจะทำการเก็บข้อมูลอยู่ในรูปของไฟล์ Excel ซึ่งจะแสดงได้ดังรูปที่ 4.22



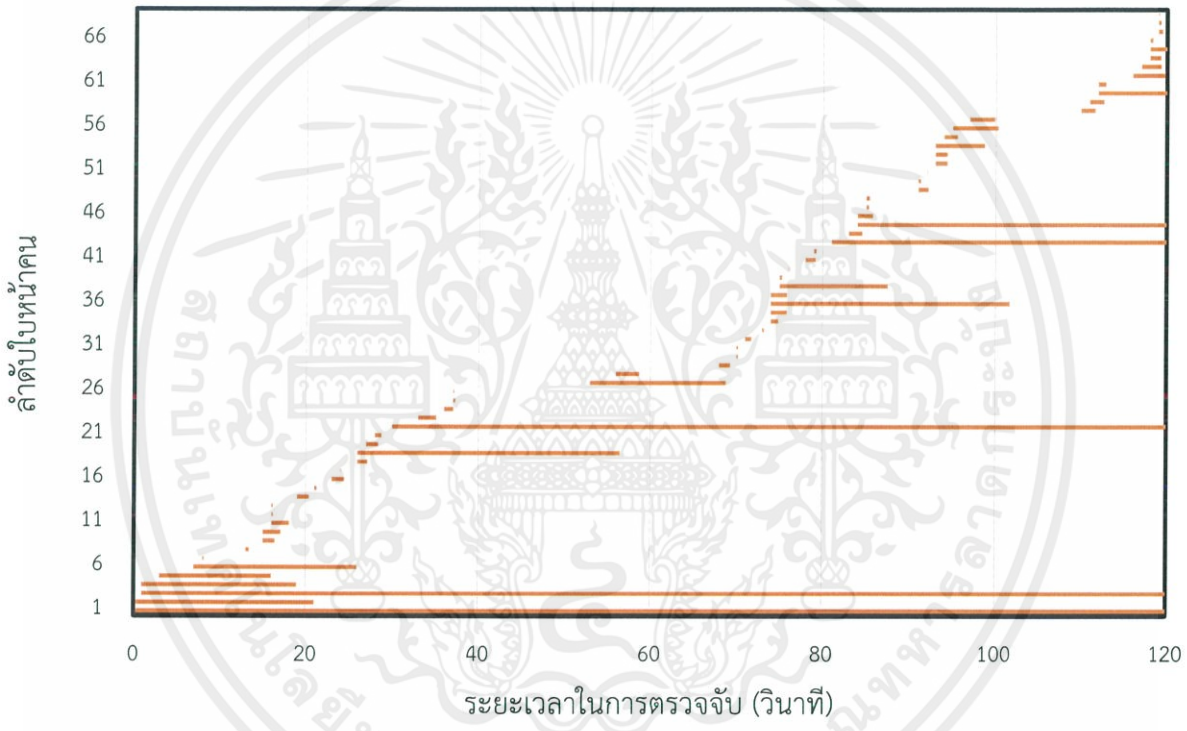
รูปที่ 4.22 การบันทึกระยะเวลาในการตรวจจับใบหน้าคนของการจำแนกเพศและสีหน้า

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3 ผลการทดสอบประสิทธิภาพของระบบ

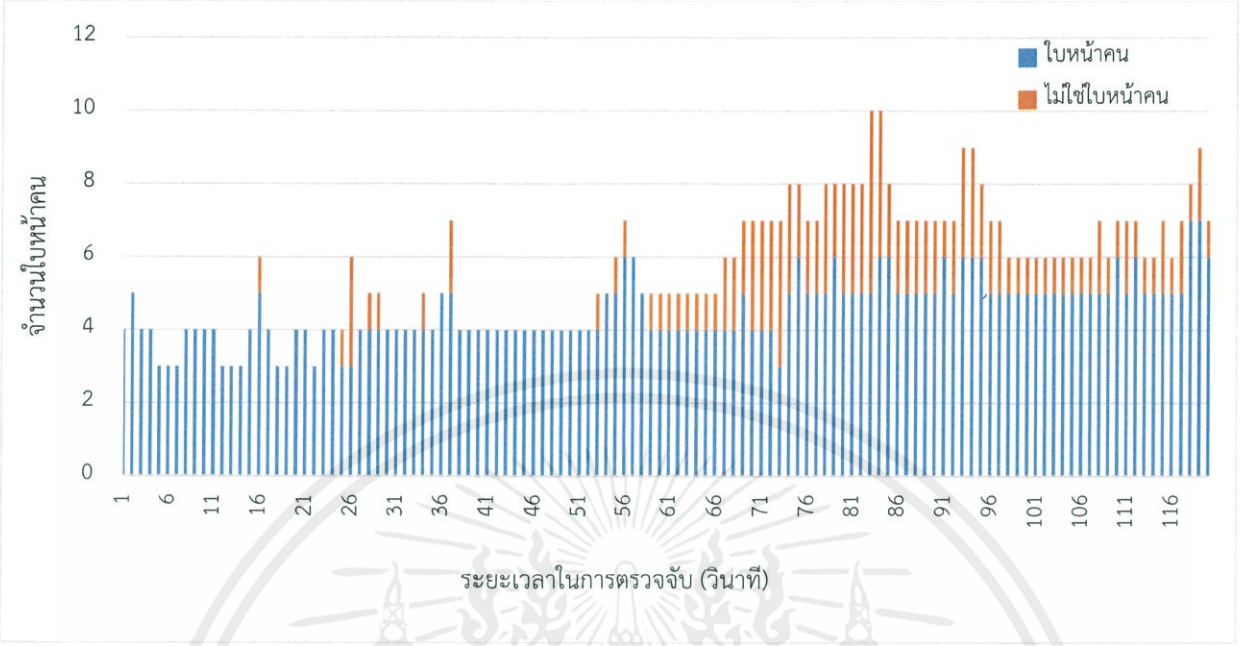
4.3.1 ผลการทดสอบประสิทธิภาพของระบบตรวจจับ

ในผลการทดสอบนี้จะทำการนับวิดีโอที่อัดจากสถานีลาดกระบังขนาด 320x240 พิกเซล ความเร็ว 120 เฟรมต่อวินาที ความยาว 2 นาทีมาทำการทดสอบประสิทธิภาพของระบบตรวจจับ โดยจะทำการแสดงกราฟความสัมพันธ์ระหว่างระยะเวลาในการตรวจจับใบหน้ากับลำดับใบหน้าคนได้ดังรูปที่ 4.23 และแสดงกราฟความสัมพันธ์ระหว่างระยะเวลาในการตรวจจับใบหน้ากับจำนวนคนได้ดังรูปที่ 4.24



รูปที่ 4.23 กราฟความสัมพันธ์ระหว่างระยะเวลาในการตรวจจับใบหน้ากับลำดับใบหน้าคน

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



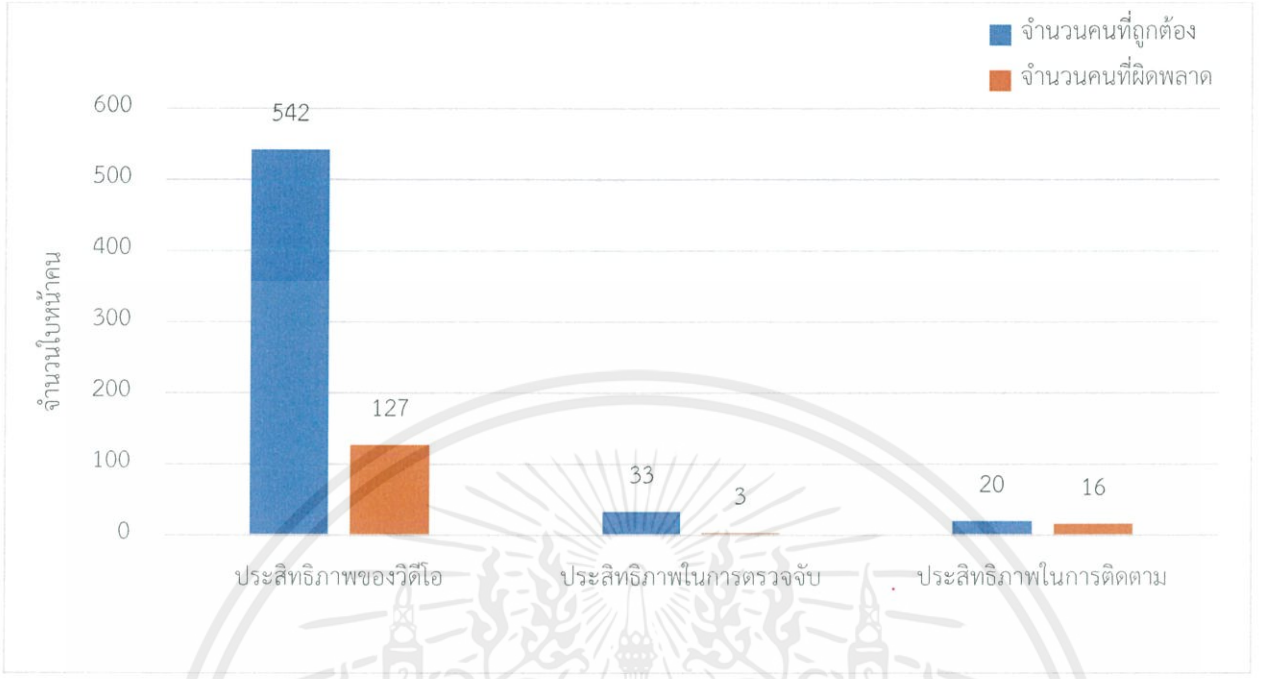
รูปที่ 4.24 กราฟความสัมพันธ์ระหว่างระยะเวลาในการตรวจจับใบหน้ากับจำนวนใบหน้าคน

จากการทดสอบวิดีโอที่อัดจากสถานีลาดกระบังจะสามารถสรุปผลการทดสอบสำหรับการตรวจจับใบหน้าคนออกมาได้ดังตารางที่ 4.1 และสามารถแสดงออกมาเป็นกราฟแสดงประสิทธิภาพความถูกต้องในการตรวจจับและนับจำนวนคนได้ในรูปที่ 4.25

ตารางที่ 4.1 สรุปผลการทดสอบประสิทธิภาพของระบบตรวจจับและนับจำนวนคน

	จำนวนคนที่ถูกต้อง	จำนวนคนที่ผิดพลาด	เปอร์เซ็นความถูกต้อง	เปอร์เซ็นความผิดพลาด
ประสิทธิภาพของวิดีโอ	542	127	81.02	18.98
ประสิทธิภาพในการตรวจจับ	33	3	91.67	8.33
ประสิทธิภาพในการติดตาม	20	16	55.56	44.44

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.25 กราฟแสดงประสิทธิภาพความถูกต้องในการตรวจจับและนับจำนวนคน

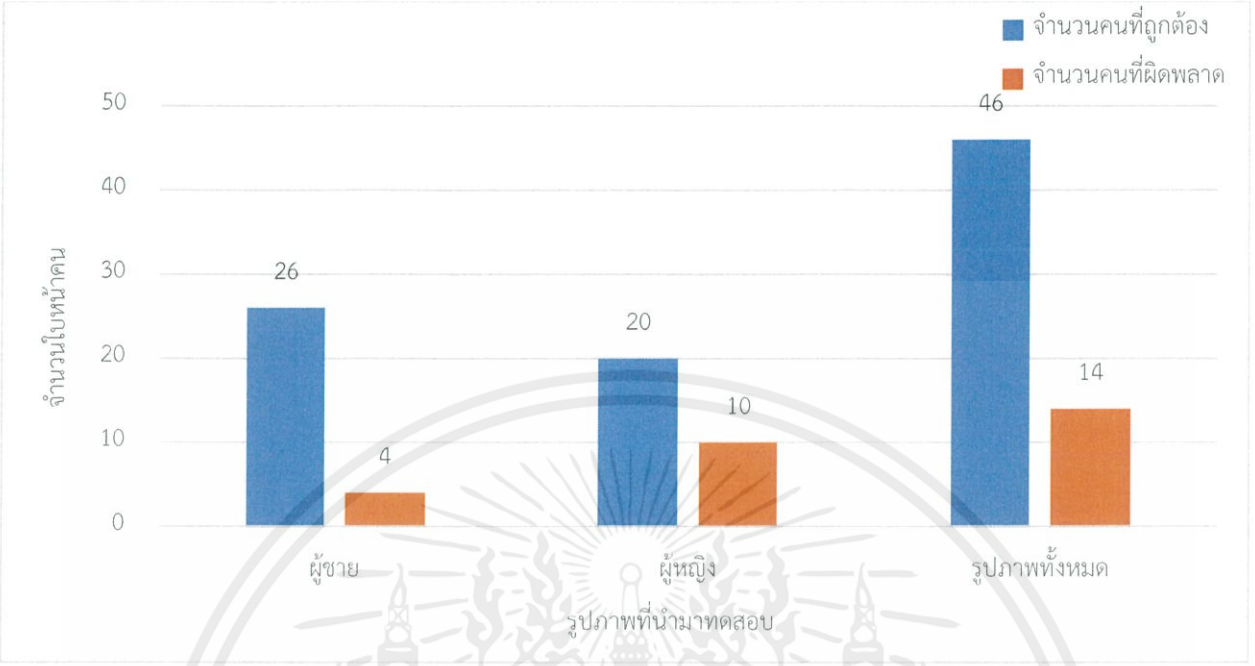
4.3.2 ผลทดสอบประสิทธิภาพของระบบจำแนกสีหน้า

จากการทดลองรูปภาพใบหน้าคน จากนั้นจะทำการสุ่มตัวอย่างใบหน้าผู้ชายและใบหน้าผู้หญิงมาอย่างละ 30 ภาพ มาทำการทดสอบประสิทธิภาพของระบบจำแนกเพศ ซึ่งจะสรุปผลการทดสอบได้ดังตารางที่ 4.2 และสามารถแสดงออกมาเป็นกราฟแสดงประสิทธิภาพความถูกต้องในการจำแนกเพศได้ในรูปที่ 4.26

ตารางที่ 4.2 สรุปผลการทดสอบประสิทธิภาพของระบบจำแนกเพศ

	จำนวนคนที่ถูกต้อง	จำนวนคนที่ผิดพลาด	เปอร์เซ็นความถูกต้อง	เปอร์เซ็นความผิดพลาด
ผู้ชาย 30 ภาพ	26	4	86.67	13.33
ผู้หญิง 30 ภาพ	20	10	66.67	33.33
รูปภาพทั้งหมด	46	14	76.67	23.33

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.26 กราฟแสดงประสิทธิภาพความถูกต้องในการจำแนกเพศ

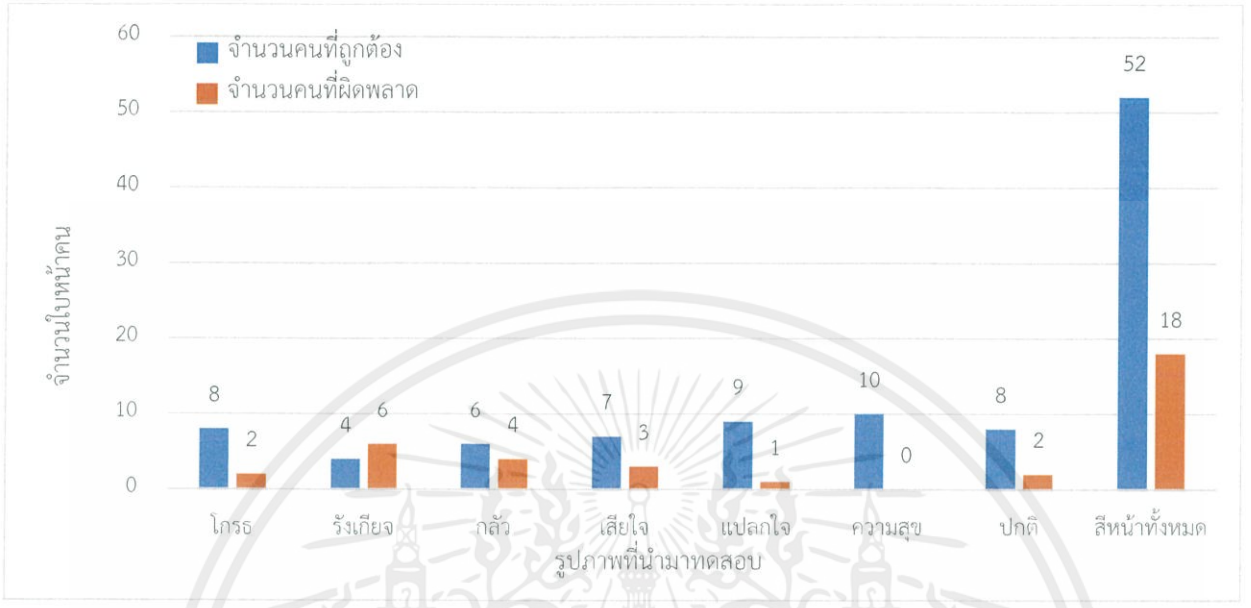
4.3.3 ผลทดสอบประสิทธิภาพของระบบจำแนกสีหน้า

จากที่ทำการทดลองระบบจำแนกเพศ จากนั้นจะทำการสุ่มรูปภาพสีหน้าของใบหน้าอย่างละ 10 ภาพ มาทำการทดสอบประสิทธิภาพซึ่งจะแสดงจากการทดสอบได้ดังตารางที่ 4.3 และสามารถแสดงออกมาเป็นกราฟแสดงประสิทธิภาพความถูกต้องในการจำแนกสีหน้าได้ในรูปที่ 4.27

ตารางที่ 4.3 สรุปผลการทดสอบประสิทธิภาพของระบบจำแนกสีหน้า

รูปภาพสีหน้าที่นำมาทดสอบ	จำนวนคนที่ถูกต้อง	จำนวนคนที่ผิดพลาด	เปอร์เซ็นความถูกต้อง	เปอร์เซ็นความผิดพลาด
โกรธ 10 ภาพ	8	2	80	20
รังเกียจ 10 ภาพ	4	6	40	60
กลัว 10 ภาพ	6	4	60	40
เสียใจ 10 ภาพ	7	3	70	30
แปลกใจ 10 ภาพ	9	1	90	10
ความสุข 10 ภาพ	10	0	100	0
ปกติ 10 ภาพ	8	2	80	0
สีหน้าทั้งหมด	52	18	74.29	25.71

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.27 กราฟแสดงประสิทธิภาพความถูกต้องในการจำแนกสีหน้า

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปผลและข้อเสนอแนะ

5.1 สรุปผลการทดลอง

ในรายงานสหกิจศึกษานี้จะนำเสนอเกี่ยวกับการพัฒนาการประมวลผลและวิเคราะห์ภาพสำหรับงานสื่อโฆษณาดิจิตอล เพื่อใช้สำหรับในการตรวจจับและนับจำนวนคนที่เดินผ่านสื่อโฆษณาดิจิตอลด้วยวิธีการ Local Binary Pattern (LBP) ซึ่งเมื่อระบบได้ทำการประมวลผลรูปภาพใบหน้าคนจากวิดีโอที่อัดจากสถานีลาดกระบัง จะมีการเก็บข้อมูลรูปภาพใบหน้าคน และระยะเวลาในการตรวจจับใบหน้าคน ซึ่งผลการทดลองจะมีประสิทธิภาพของระบบอยู่ที่ 81.02% ประสิทธิภาพในการตรวจจับใบหน้าอยู่ที่ 91.67% และประสิทธิภาพในการติดตามอยู่ที่ 55.56% รวมทั้งเพื่อใช้สำหรับในการจำแนกเพศ และสีหน้าที่แสดงออกมาด้วยวิธีการใช้ Convolutional Neural Networks (CNNs) ซึ่งเมื่อระบบได้ทำการประมวลผลรูปภาพใบหน้าคนที่ได้นำมาทดสอบจะสามารถบอกประสิทธิภาพในการจำแนกเพศได้อยู่ที่ 76.67% และประสิทธิภาพในการจำแนกสีหน้าได้อยู่ที่ 74.29%

5.2 ข้อเสนอแนะ

เนื่องจากการพัฒนาการประมวลผลและวิเคราะห์ภาพสำหรับงานสื่อโฆษณาดิจิตอล จะมีอัตราความถูกต้องที่ขึ้นอยู่กับปริมาณฐานข้อมูลรูปภาพที่นำมาใช้ จึงจำเป็นที่จะต้องมีการเตรียมฐานข้อมูลรูปใบหน้าหน้าคนและรูปภาพที่ไม่ใช่ใบหน้าคนจำนวนมากในการใช้งาน โดยจะต้องเลือกใช้ข้อมูลรูปภาพใบหน้าคนจากสถานีรถไฟฟ้าแอร์พอร์ตลิ้งและข้อมูลรูปภาพที่ไม่ใช่ใบหน้าคนจากความผิดพลาดจริงนำมาใช้งาน เนื่องจากระบบในการตรวจจับและนับจำนวนคนจะค่อนข้างมีความอ่อนไหวสูงสำหรับข้อมูลรูปภาพที่ไม่มีมาตราฐาน และในส่วนของการจำแนกเพศ และสีหน้าจำเป็นที่จะต้องเพิ่มฐานข้อมูลรูปภาพใบหน้าผู้ชาย ผู้หญิงและสีหน้าให้เพิ่มขึ้นเพื่อที่จะลดความผิดพลาดผลาดที่เกิดขึ้นของระบบในการตรวจจับและนับจำนวนคน

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

[1] นายศิรเสกข์ ตราชื่นต้อง, นายอธิษฐ์ แย้มสุข. "*การพัฒนาระบบรู้จำเสียงพูด/ภาพใบหน้าและทางการประยุกต์ใช้งาน.*" หลักสูตรวิศวกรรมศาสตรบัณฑิต, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2559.

[2] นายธีระเดช กุศลธรรมรัตน์, นายธุวานนท์ เทพสิทธา, นายภีมวิชญ์ สุทธิกาญจน์."*การออกแบบและสร้างแบบจาลองทางวิศวกรรมของดาวเทียมคิวบ์แซทและการเข้ารหัสช่องสัญญาณ.*" หลักสูตรวิศวกรรมศาสตรบัณฑิต, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2559.

[3] นายศิวกร เดชดำเกิงชัย. "*ระบบเรียลไทม์มอนิเตอร์ริ่งรุ่นที่ 4 : ส่วนที่ 2.*" หลักสูตรวิศวกรรมศาสตรบัณฑิต, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2555.

[4] Octavio, Arriaga. Real-time Convolutional Neural Networks for Emotion and Gender Classification. 2018.

[5] Mr. karpathy. CS231n Convolutional Neural Networks for Visual Recognition. http://cs231n.github.io/convolutional-networks/.

[6] Rosebrock, Adrian. Deep Learning for Computer Visition with Python vol 1. 1st ed. 2018.

[7] Rosebrock, Adrian. Deep Learning for Computer Visition with Python vol 2. 1st ed. 2018.

[8] Rosebrock, Adrian. Deep Learning for Computer Visition with Python vol 3. 1st ed. 2018.

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ซอร์สโค้ดระบบตรวจจับและนับจำนวนคน

```
#Import the OpenCV and dlib libraries
import os
import cv2
import csv
import dlib
import time
import imutils
import threading
import numpy as np

LBPCascade = cv2.CascadeClassifier('C:/Users/Samrit & Warisarapon/Desktop/Learn
OpenCV by pyimagesearch/Demo/OMAS/LBP10.xml')

fourcc = cv2.VideoWriter_fourcc(*'XVID')
out = cv2.VideoWriter('C:/Users/Samrit & Warisarapon/Desktop/
videoLadkrabang320x240_.avi', fourcc,120, (800,600))

start_time = {}
end_time = {}
timestamp = {}
img = {}
saveimg = {}
pathimg = {}

def doRecognizePerson(faceNames, fid):
    faceNames[fid] = "Person " + str(fid)
    start_time[fid] = time.time()
    timestamp[fid] = time.strftime("%Y%m%d_%H%M%S",time.localtime())
    saveimg[fid] = False
def detectAndTrackMultipleFaces():
    #Open the first webcame device
```

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
cap = cv2.VideoCapture(0)

#The deisred output width and height
cap.set(cv2.CAP_PROP_FRAME_WIDTH,320)
cap.set(cv2.CAP_PROP_FRAME_HEIGHT,240)
width = int(cap.get(3))
height = int(cap.get(4))
framerate = int(cap.get(5))
print("frame rate : {}".format(framerate))
print("Width : {} || Height : {}".format(width,height))

#Create two opencv named windows
cv2.namedWindow("result-image", cv2.WINDOW_AUTOSIZE)

#Position the windows next to eachother
cv2.moveWindow("result-image",400,100)

#Start the window thread for the two windows we are using
cv2.startWindowThread()

#The color of the rectangle we draw around the face
rectangleColor = (0,165,255)

#variables holding the current frame number and the current faceid
frameCounter = 1
currentFaceID = 1

#Variables holding the correlation trackers and the name per faceid
faceTrackers = {}
faceNames = {
try:
    while True:
        #Retrieve the latest image from the webcam
```

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
rc,fullSizeBaseImage = cap.read()

#Resize the image to 320x240
baseImage = cv2.resize( fullSizeBaseImage, ( 320, 240))

 #Result image is the image we will show the user, which is a
#combination of the original image from the webcam and the
#overlayed rectangle for the largest face
resultImage = baseImage.copy()

#STEPS:
# * Update all trackers and remove the ones that are not
#   relevant anymore
# * Every 5 frames:
#       + Use face detection on the current frame and look
#         for faces.
#       + For each found face, check if centerpoint is within
#         existing tracked box. If so, nothing to do
#       + If centerpoint is NOT in existing tracked box, then
#         we add a new tracker with a new face-id

#Increase the framecounter
frameCounter += 1

#Update all the trackers and remove the ones for which the update
#indicated the quality was not good enough
fidsToDelete = []


for fid in faceTrackers.keys():
    trackingQuality = faceTrackers[ fid ].update( baseImage )
    t_w = int(tracked_position.width())
```

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            if t_w > 52:
                fidsToDelete.append( fid )
                end_time[fid] = time.time()
                print("Face Id : " + str(fid))
                print("Duration")
                duration = end_time[fid] - start_time[fid]
                print("%.2f"%duration + '  sec')
                with open("C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by
pyimagesearch/Demo/OMAS/log/log.csv",'a', newline='',encoding='utf-8') as csvfile:
                    a = csv.writer(csvfile, delimiter= ',', quotechar='|', quoting=
csv.QUOTE_MINIMAL)
                    a.writerow([fid,timestamp[fid],duration,pathimg[fid]])

            #If the tracking quality is good enough, we must delete
            #this tracker
            if trackingQuality < 7:
                fidsToDelete.append( fid )
                end_time[fid] = time.time()
                print("Face Id : " + str(fid))
                print("Duration")
                duration = end_time[fid] - start_time[fid]
                print("%.2f"%duration + '  sec')
                with open("C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by
pyimagesearch/Demo/OMAS/log/log.csv",'a', newline='',encoding='utf-8') as csvfile:
                    a = csv.writer(csvfile, delimiter= ',', quotechar='|', quoting=
csv.QUOTE_MINIMAL)
                    a.writerow([fid,timestamp[fid],duration,pathimg[fid]])

        for fid in fidsToDelete:
            faceTrackers.pop( fid , None )

        #Every 10 frames, we will have to determine which faces
        #are present in the frame
```

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
        if (frameCounter % 5) == 0:
            gray = cv2.cvtColor(baseImage, cv2.COLOR_BGR2GRAY)
            faces =
LBPCascade.detectMultiScale(gray,scaleFactor=1.10,minNeighbors=2,minSize=(17,
17),maxSize=(52, 52))

            #Loop over all faces and check if the area for this
            #face is the largest so far
            #We need to convert it to int here because of the
            #requirement of the dlib tracker. If we omit the cast to
            #int here, you will get cast errors since the detector
            #returns numpy.int32 and the tracker requires an int
            for (_x,_y,_w,_h) in faces:
                x = int(_x)
                y = int(_y)
                w = int(_w)
                h = int(_h)


                #calculate the centerpoint
                x_bar = x + 0.5 * w
                y_bar = y + 0.5 * h
                matchedFid = None

                #Now loop over all the trackers and check if the
                #centerpoint of the face is within the box of a
                #tracker
                for fid in faceTrackers.keys():
                    tracked_position =  faceTrackers[fid].get_position()


                    t_x = int(tracked_position.left())
                    t_y = int(tracked_position.top())
```

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
t_w = int(tracked_position.width())
t_h = int(tracked_position.height())

#calculate the centerpoint
t_x_bar = t_x + 0.5 * t_w
t_y_bar = t_y + 0.5 * t_h

#check if the centerpoint of the face is within the
#rectangleof a tracker region. Also, the centerpoint
#of the tracker region must be within the region
#detected as a face. If both of these conditions hold
#we have a match
if ( ( t_x <= x_bar   <= (t_x + t_w)) and
     ( t_y <= y_bar   <= (t_y + t_h)) and
     ( x   <= t_x_bar <= (x   + w  )) and
     ( y   <= t_y_bar <= (y   + h  ))):
     matchedFid = fid

#If no matched fid, then we have to create a new tracker
if matchedFid is None:

#Create and store the tracker
tracker = dlib.correlation_tracker()
tracker.start_track(baseImage,
                    dlib.rectangle( x,
                                    y,
                                    x+w,
                                    y+h))
faceTrackers[ currentFaceID ] = tracker

#Start a new thread that is used to simulate
#face recognition. This is not yet implemented in this
#version :)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            t = threading.Thread( target = doRecognizePerson ,
                                  args=(faceNames, currentFaceID))
            t.start()

            #Increase the currentFaceID counter
            currentFaceID += 1

    #Now loop over all the trackers we have and draw the rectangle
    #around the detected faces. If we 'know' the name for this person
    #(i.e. the recognition thread is finished), we print the name
    #of the person, otherwise the message indicating we are detecting
    #the name of the person
    for fid in faceTrackers.keys():
        tracked_position =  faceTrackers[fid].get_position()
        t_x = int(tracked_position.left())
        t_y = int(tracked_position.top())
        t_w = int(tracked_position.width())
        t_h = int(tracked_position.height())

        cv2.rectangle(resultImage, (t_x, t_y),
                                   (t_x + t_w , t_y + t_h),
                                   rectangleColor ,2)
        faceTrackers[ fid ].update( baseImage )
        if saveimg[fid] == False:
            if t_w > 0 and t_h > 0:
                if t_x < 0:
                    t_x = 0
                if t_y < 0:
                    t_y = 0
                if (t_x + t_w) > 319:
                    t_xx = 319
                if (t_y + t_h) > 239:
                    t_yy = 239
```

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
                t_xx = t_x+t_w
                t_yy = t_y+t_h
                img[fid] = baseImage[t_y:t_yy,t_x:t_xx]
                saveimg[fid] = True
                path = 'C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by
pyimagesearch/Demo/OMAS/log/'
                pathimg[fid] = os.path.join(path, str(timestamp[fid]) + '_' + str(fid) +
'.jpg')
                cv2.imwrite(pathimg[fid],img[fid])
            if fid in faceNames.keys():
                cv2.putText(resultImage, faceNames[fid] ,
                    (int(t_x), int(t_y)-4),
                    cv2.FONT_HERSHEY_SIMPLEX,
                    0.5, (0, 0, 0), 2)
            else:
                cv2.putText(resultImage, faceNames[fid] ,
                    (int(t_x), int(t_y)-4),
                    cv2.FONT_HERSHEY_SIMPLEX,
                    0.5, (255, 255, 255), 2)

        #Since we want to show something larger on the screen than the
        #original 320x240, we resize the image again
        #
        #Note that it would also be possible to keep the large version
        #of the baseimage and make the result image a copy of this large
        #base image and use the scaling factor to draw the rectangle
        #at the right coordinates.
        largeResult = cv2.resize(resultImage,
                                 (800,600))
        out.write(largeResult)

        #Finally, we want to show the images on the screen
        cv2.imshow("result-image", largeResult)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            cv2.waitKey(1)

    #To ensure we can also deal with the user pressing Ctrl-C in the console
    #we have to check for the KeyboardInterrupt exception and break out of
    #the main loop
    except KeyboardInterrupt as e:
        pass

    #Destroy any OpenCV windows and exit the application
    cv2.destroyAllWindows()
    exit(0)

if __name__ == '__main__':
    detectAndTrackMultipleFaces
```

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ข

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ซอร์สโค้ดระบบจำแนกเพศและสีหน้า

```
#Import the OpenCV and dlib libraries
import os
import cv2
import csv
import dlib
import time
import imutils
import threading
import numpy as np

from utils.datasets import get_labels
from utils.inference import detect_faces
from utils.preprocessor import preprocess_input
from utils.inference import load_detection_model

predictor = dlib.shape_predictor("shape_predictor_68_face_landmarks.dat")
face_detection = load_detection_model('trained_models/detection_
        models/haarcascade_frontalface_default.xml')
emotion_classifier = load_model('trained_models/emotion_models/fer2013_mini_
        XCEPTION.102-0.66.hdf5', compile=False)
gender_classifier = load_model('trained_models/gender_models/simple_CNN.81-
        0.96.hdf5', compile=False)
LBPCascade = cv2.CascadeClassifier('C:/Users/Samrit & Warisarapon/Desktop/Learn
        OpenCV by pyimagesearch/Demo/OMAS/haarcascade_frontalface_default.xml')

gender_labels = get_labels('imdb')
emotion_labels = get_labels('fer2013')


fourcc = cv2.VideoWriter_fourcc(*'XVID')
```

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
out = cv2.VideoWriter('C:/Users/Samrit & Warisarapon/Desktop/
videoLadkrabang320x240_.avi', fourcc,120, (800,600))

start_time = {}
end_time = {}
timestamp = {}
img = {}
saveimg = {}
pathimg = {}
check = ()
emotion_window = []
gender_window = []
time1 = []
time2 = []

def detectionemotion(image, fid, faceEmotion):
    imagg = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
    frontalfaces = detect_faces(face_detection, imagg)
    if frontalfaces == check:
        faceEmotion[fid] = "neutral"
    else:
        try:
            imgrgb = cv2.resize(imagg, (64,64))
        except:
            return 0
        gray_face = preprocess_input(imgrgb, False)
        gray_face = np.expand_dims(gray_face, 0)
        gray_face = np.expand_dims(gray_face, -1)
        emotion_label_arg = np.argmax(emotion_classifier.predict(gray_face))
        emotion_text = emotion_labels[emotion_label_arg]
        emotion_window.append(emotion_text)
        if len(emotion_window) > 1:
            emotion_window.pop(0)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
        try:
            emotion_mode = mode(emotion_window)
            faceEmotion[fid] = str(emotion_mode)
        except:
            faceEmotion[fid] = "neutral"

def detectionGenders(image, fid, faceGenders):
    imagg = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
    frontalfaces = detect_faces(face_detection, imagg)
    # print("frontalfaces : {}".format(frontalfaces))
    if frontalfaces == check:
        faceGenders[fid] = "Don't sure"
    else:
        try:
            imgrgb = cv2.resize(image, (48,48))
        except:
            return 0
        imgrgb = cv2.cvtColor(imgrgb, cv2.COLOR_BGR2RGB)
        rgb_face = np.expand_dims(imgrgb, 0)
        rgb_face = preprocess_input(rgb_face, False)
        gender_prediction = gender_classifier.predict(rgb_face)
        gender_label_arg = np.argmax(gender_prediction)
        gender_text = gender_labels[gender_label_arg]
        gender_window.append(gender_text)
        if len(gender_window) > 1:
                gender_window.pop(0)
        try:
            gender_mode = mode(gender_window)
            faceGenders[fid] = str(gender_mode)
        except:
            faceGenders[fid] = "Don't sure"

        if gender_text == gender_labels[0]:
```

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            path = 'C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by
pyimagesearch/Demo/OMAS/log/log women/'
            pathimg[fid] = os.path.join(path, str(timestamp[fid]) + '_' + str(fid) + '.jpg')
            cv2.imwrite(pathimg[fid],image)
        else:
            path = 'C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by
pyimagesearch/Demo/OMAS/log/log man/'
            pathimg[fid] = os.path.join(path, str(timestamp[fid]) + '_' + str(fid) + '.jpg')
            cv2.imwrite(pathimg[fid],image)
def doRecognizePerson(faceNames, fid, faceEmotion,faceGenders):
    faceNames[fid] = "Person " + str(fid)
    faceEmotion[fid] = " "
    faceGenders[fid] = " "
    start_time[fid] = time.time()
    timestamp[fid] = time.strftime("%Y%m%d_%H%M%S",time.localtime())
    saveimg[fid] = False

def detectAndTrackMultipleFaces():
    #Open the first webcame device
    cap = cv2.VideoCapture(0)

    #The deisred output width and height
    cap.set(cv2.CAP_PROP_FRAME_WIDTH,640)
    cap.set(cv2.CAP_PROP_FRAME_HEIGHT,480)
    width = int(cap.get(3))
    height = int(cap.get(4))
    framerate = int(cap.get(5))
    print("frame rate : {}".format(framerate))
    print("Width : {} || Height : {}".format(width,height))

    #Create two opencv named windows
    cv2.namedWindow("result-image", cv2.WINDOW_AUTOSIZE)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
#Position the windows next to eachother
cv2.moveWindow("result-image",400,100)

#Start the window thread for the two windows we are using
cv2.startWindowThread()

#The color of the rectangle we draw around the face
rectangleColor = (0,165,255)

#variables holding the current frame number and the current faceid
frameCounter = 1
currentFaceID = 1

#Variables holding the correlation trackers and the name per faceid
faceTrackers = {}
faceNames = {
try:
    while True:
        #Retrieve the latest image from the webcam
        rc,fullSizeBaseImage = cap.read()

        #Resize the image to 320x240
        baseImage = cv2.resize( fullSizeBaseImage, ( 640, 480))

        #Result image is the image we will show the user, which is a
        #combination of the original image from the webcam and the
        #overlayed rectangle for the largest face
        resultImage = baseImage.copy()

        #STEPS:
        # * Update all trackers and remove the ones that are not
        #   relevant anymore
        # * Every 5 frames:
```

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
        #       + Use face detection on the current frame and look
        #         for faces.
        #       + For each found face, check if centerpoint is within
        #         existing tracked box. If so, nothing to do
        #       + If centerpoint is NOT in existing tracked box, then
        #         we add a new tracker with a new face-id

        #Increase the framecounter
        frameCounter += 1

        #Update all the trackers and remove the ones for which the update
        #indicated the quality was not good enough
        fidsToDelete = []

          #If the tracking quality is good enough, we must delete
          #this tracker
          if trackingQuality < 7:
            fidsToDelete.append( fid )
            end_time[fid] = time.time()
            print("Face Id : " + str(fid))
            print("Duration")
            duration = end_time[fid] - start_time[fid]
            print("%.2f"%duration + ' sec')
            with open("C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by
pyimagesearch/Demo/OMAS/log/log.csv",'a', newline='',encoding='utf-8') as csvfile:
                a = csv.writer(csvfile, delimiter= ',', quotechar='|', quoting=
csv.QUOTE_MINIMAL)
                a.writerow([fid,timestamp[fid],duration,pathimg[fid]])

        for fid in fidsToDelete:
          faceTrackers.pop( fid , None )

        #Every 10 frames, we will have to determine which faces
```

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            #are present in the frame
            if (frameCounter % 5) == 0:
                gray = cv2.cvtColor(baseImage, cv2.COLOR_BGR2GRAY)
                faces =
LBPCascade.detectMultiScale(gray,scaleFactor=1.10,minNeighbors=2,minSize=(100, 100))

                #Loop over all faces and check if the area for this
                #face is the largest so far
                #We need to convert it to int here because of the
                #requirement of the dlib tracker. If we omit the cast to
                #int here, you will get cast errors since the detector
                #returns numpy.int32 and the tracker requires an int
                for (_x,_y,_w,_h) in faces:
                    x = int(_x)
                    y = int(_y)
                    w = int(_w)
                    h = int(_h)

                    #calculate the centerpoint
                    x_bar = x + 0.5 * w
                    y_bar = y + 0.5 * h
                    matchedFid = None

                    #Now loop over all the trackers and check if the
                    #centerpoint of the face is within the box of a
                    #tracker
                    for fid in faceTrackers.keys():
                        tracked_position =  faceTrackers[fid].get_position()
                        t_x = int(tracked_position.left())
                        t_y = int(tracked_position.top())
                        t_w = int(tracked_position.width())
                        t_h = int(tracked_position.height())
                        #calculate the centerpoint
```

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
t_x_bar = t_x + 0.5 * t_w
t_y_bar = t_y + 0.5 * t_h

#check if the centerpoint of the face is within the
#rectangleof a tracker region. Also, the centerpoint
#of the tracker region must be within the region
#detected as a face. If both of these conditions hold
#we have a match
if ( ( t_x <= x_bar   <= (t_x + t_w)) and
     ( t_y <= y_bar   <= (t_y + t_h)) and
     ( x   <= t_x_bar <= (x   + w  )) and
     ( y   <= t_y_bar <= (y   + h  ))):
     matchedFid = fid

#If no matched fid, then we have to create a new tracker
if matchedFid is None:

    #Create and store the tracker
    tracker = dlib.correlation_tracker()
    tracker.start_track(baseImage,
                        dlib.rectangle( x,
                                        y,
                                        x+w,
                                        y+h))
    faceTrackers[ currentFaceID ] = tracker

    #Start a new thread that is used to simulate
    #face recognition. This is not yet implemented in this
    #version :)
    t = threading.Thread( target = doRecognizePerson ,
                          args=(faceNames, currentFaceID
                           faceEmotion, faceGender))
    t.start()
```

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            #Increase the currentFaceID counter
            currentFaceID += 1

    #Now loop over all the trackers we have and draw the rectangle
    #around the detected faces. If we 'know' the name for this person
    #(i.e. the recognition thread is finished), we print the name
    #of the person, otherwise the message indicating we are detecting
    #the name of the person
    for fid in faceTrackers.keys():
        tracked_position =  faceTrackers[fid].get_position()
        t_x = int(tracked_position.left())
        t_y = int(tracked_position.top())
        t_w = int(tracked_position.width())
        t_h = int(tracked_position.height())

        cv2.rectangle(resultImage, (t_x, t_y),
                              (t_x + t_w , t_y + t_h),
                              rectangleColor ,2)
        if (frameCounter % 2) == 0:
            detectionemotion(baseImage[t_y-30:t_y+t_h+30,t_x-20:t_x+t_w+20], fid,
                faceEmotion)
            detectionGenders(baseImage[t_y-30:t_y+t_h+30,t_x-20:t_x+t_w+20], fid,
                faceGenders)
        faceTrackers[ fid ].update( baseImage )
        if saveimg[fid] == False:
            if t_w > 0 and t_h > 0:
                if t_x < 0:
                    t_x = 0
                if t_y < 0:
                    t_y = 0
                if (t_x + t_w) > 319:
                    t_xx = 319
```

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
                if (t_y + t_h) > 239:
                    t_yy = 239
                t_xx = t_x+t_w
                t_yy = t_y+t_h
                img[fid] = baseImage[t_y:t_yy,t_x:t_xx]
                saveimg[fid] = True
                path = 'C:/Users/Samrit & Warisarapon/Desktop/Learn OpenCV by
pyimagesearch/Demo/OMAS/log/'
                pathimg[fid] = os.path.join(path, str(timestamp[fid]) + '_' + str(fid) +
'.jpg')
                cv2.imwrite(pathimg[fid],img[fid])
            if fid in faceNames.keys():
              cv2.putText(resultImage, faceNames[fid] ,
                        (int(t_x), int(t_y)-34),
                        cv2.FONT_HERSHEY_SIMPLEX,
                        0.5, (255, 255, 255), 2)
              cv2.putText(resultImage, faceGenders[fid] ,
                        (int(t_x), int(t_y)-19),
                        cv2.FONT_HERSHEY_SIMPLEX,
                        0.5, (255, 255, 255), 2)
              cv2.putText(resultImage, faceEmotion[fid] ,
                        (int(t_x), int(t_y)-4),
                        cv2.FONT_HERSHEY_SIMPLEX,
                        0.5, (255, 255, 255), 2)
            else:
              cv2.putText(resultImage, faceNames[fid] ,
                        (int(t_x), int(t_y)-4),
                        cv2.FONT_HERSHEY_SIMPLEX,
                        0.5, (255, 255, 255), 2)
              cv2.putText(resultImage, faceGenders[fid] ,
                        (int(t_x), int(t_y)-4),
                        cv2.FONT_HERSHEY_SIMPLEX,
                        0.5, (255, 255, 255), 2)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
                cv2.putText(resultImage, faceEmotion[fid] ,
                        (int(t_x), int(t_y)-4),
                        cv2.FONT_HERSHEY_SIMPLEX,
                        0.5, (255, 255, 255), 2)

                #Since we want to show something larger on the screen than the
                #original 320x240, we resize the image again
                #
                #Note that it would also be possible to keep the large version
                #of the baseimage and make the result image a copy of this large
                #base image and use the scaling factor to draw the rectangle
                #at the right coordinates.
                largeResult = cv2.resize(resultImage,
                                (800,600))
                out.write(largeResult)

                #Finally, we want to show the images on the screen
                cv2.imshow("result-image", largeResult)
                cv2.waitKey(1)

        #To ensure we can also deal with the user pressing Ctrl-C in the console
        #we have to check for the KeyboardInterrupt exception and break out of
        #the main loop
        except KeyboardInterrupt as e:
            pass

        #Destroy any OpenCV windows and exit the application
        cv2.destroyAllWindows()
        exit(0)

if __name__ == '__main__':
    detectAndTrackMultipleFac()
```

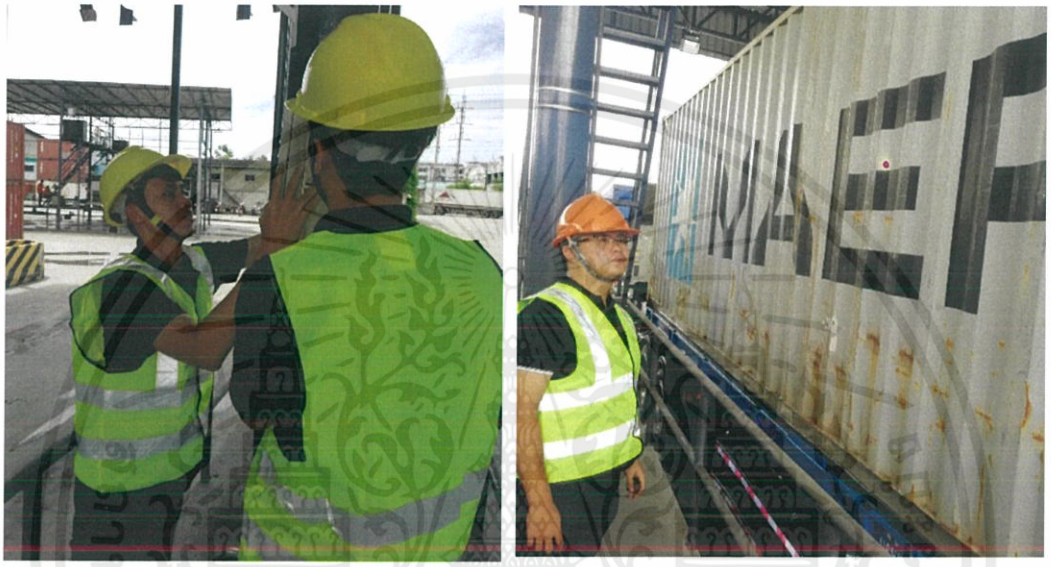
เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ค

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การปฏิบัติงานอื่นๆ ที่ได้รับมอบหมาจากทางบริษัท

จากที่ได้ไปทำงานสหกิจศึกษากับทางบริษัท เวริลี วิชั่น จำกัดทำให้ผมได้มีโอกาสไปสัมผัสงานที่หลากหลายมากยิ่งขึ้น เช่น ออกแบบและแก้ปัญหาการติดตั้งกล้องตรวจจับตัวเลขบนตู้คอนเทนเนอร์ที่ท่าเรือ Sahathai Terminal เมื่อวันที่ 15 สิงหาคม พ.ศ.2561 และได้ไปทำการออกบูธบริษัท เวริลี วิชั่น จำกัด ที่งาน TI Log Logistix 2018 ณ ไบเทคบางนา เมื่อวันที่ 29 สิงหาคม พ.ศ.2561



(ก) (ข)

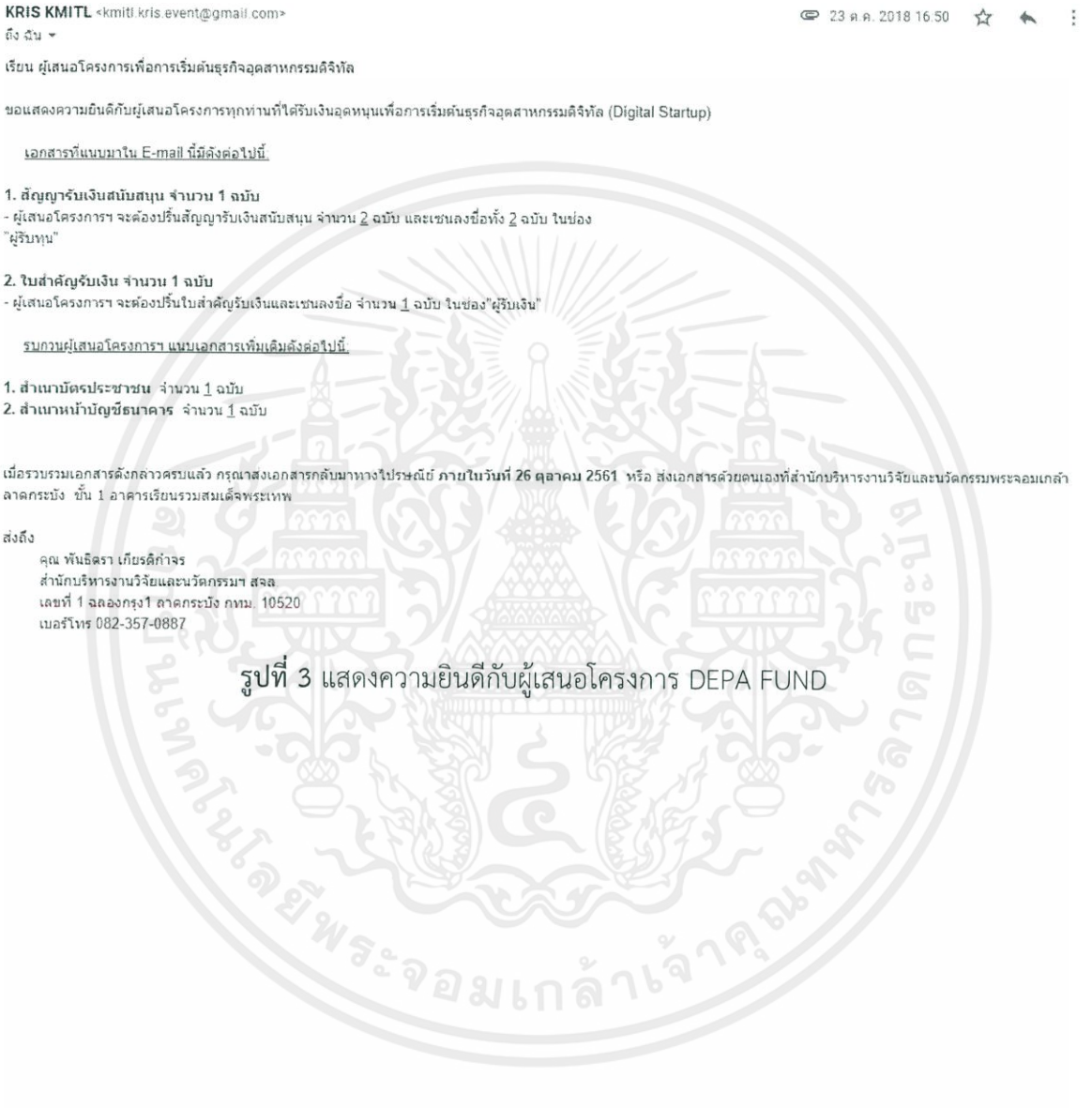
รูปที่ 1 (ก) และ (ข) ออกแบบและแก้ปัญหาการติดตั้งกล้องตรวจจับตัวเลขบนตู้คอนเทนเนอร์ที่ท่าเรือ Sahathai Terminal

(ก) (ข)

รูปที่ 2 (ก) และ (ข) ออกบูธบริษัท เวริลี วิชั่น จำกัดที่งาน TI Log Logistix 2018

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในอีกส่วนหนึ่งที่ได้มีการทำงานคือการได้มีสิทธิที่เข้าร่วมในส่วนของโครงการเพื่อการเริ่มต้นธุรกิจอุตสาหกรรมดิจิทัล (DEPA FUND) หัวข้อ ป้ายสื่อโฆษณาดิจิตอลอัจฉริยะ และเป็นตัวแทนในการรับเงินสนับสนุนเพื่อการเริ่มต้นธุรกิจอุตสาหกรรมดิจิตอล (Digital Startup) เมื่อวันที่ 23 ต.ค. 2561 ซึ่งมีรายละเอียดดังภาพที่ 3



KRIS KMITL <kmitl.kris.event@gmail.com> 23 ต.ค. 2018 16:50

ถึง ฉัน

เรียน ผู้เสนอโครงการเพื่อการเริ่มต้นธุรกิจอุตสาหกรรมดิจิทัล

ขอแสดงความยินดีกับผู้เสนอโครงการทุกท่านที่ได้รับเงินอุดหนุนเพื่อการเริ่มต้นธุรกิจอุตสาหกรรมดิจิทัล (Digital Startup)

เอกสารที่แนบมาใน E-mail นี้มีดังต่อไปนี้

1. สัญญารับเงินสนับสนุน จำนวน 1 ฉบับ
- ผู้เสนอโครงการฯ จะต้องปริ้นสัญญารับเงินสนับสนุน จำนวน 2 ฉบับ และเซ็นลงชื่อทั้ง 2 ฉบับ ในช่อง "ผู้รับทุน"

2. ใบสำคัญรับเงิน จำนวน 1 ฉบับ
- ผู้เสนอโครงการฯ จะต้องปริ้นใบสำคัญรับเงินและเซ็นลงชื่อ จำนวน 1 ฉบับ ในช่อง"ผู้รับเงิน"

รบกวนผู้เสนอโครงการฯ แนบเอกสารเพิ่มเติมดังต่อไปนี้

1. สำเนาบัตรประชาชน จำนวน 1 ฉบับ
2. สำเนาหน้าบัญชีธนาคาร จำนวน 1 ฉบับ

เมื่อรวบรวมเอกสารดังกล่าวครบแล้ว กรุณาส่งเอกสารกลับมาทางไปรษณีย์ **ภายในวันที่ 26 ตุลาคม 2561** หรือ ส่งเอกสารด้วยตนเองที่สำนักบริหารงานวิจัยและนวัตกรรมพระจอมเกล้าลาดกระบัง ชั้น 1 อาคารเรียนรวมสมเด็จพระเทพ

ส่งถึง
คุณ พันธิตรา เกียรติกำจร
สำนักบริหารงานวิจัยและนวัตกรรมฯ สจล
เลขที่ 1 ฉลองกรุง1 ลาดกระบัง กทม. 10520
เบอร์โทร 082-357-0887

รูปที่ 3 แสดงความยินดีกับผู้เสนอโครงการ DEPA FUND

เอกสารนี้เป็นเอกสารที่สงวนไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใด ๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้