

การตรวจจับพฤติกรรมมนุษย์
HUMAN BEHAVIOR DETECTION

พิเชษฐ ปานเขียว
Pichet Pankhiaw
ภณิดา เอี่ยมวิจารณ์
Panida Aimvichan

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต
สาขาวิชาวิศวกรรมชีวการแพทย์
คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2560

การตรวจจับพฤติกรรมมนุษย์
HUMAN BEHAVIOR DETECTION

นายพิเชษฐ ปานเขียว รหัสนักศึกษา 57010899
นางสาวภณิดา เอี่ยมวิจารณ์ รหัสนักศึกษา 57010949

อาจารย์ที่ปรึกษา
รศ.ดร.สุรพันธุ์ เอื้อไพบูลย์

ปฏิญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต
สาขาวิชาวิศวกรรมชีวการแพทย์
คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2560

ปริญญาโทปีการศึกษา 2560

สาขาวิชาวิศวกรรมชีวการแพทย์ คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง การตรวจจับพฤติกรรมมนุษย์
HUMAN BEHAVIOR DETECTION

ผู้จัดทำ

- 1.นายพิเชษฐ ปานเขียว รหัสประจำตัว 57010899
- 2.นางสาวภณิดา เอี่ยมวิจารณ์ รหัสประจำตัว 57010949



..... อาจารย์ที่ปรึกษา
(รศ.ดร.สุรพันธุ์ เอื้อไพบูลย์)

หัวข้อปริญญานิพนธ์	การตรวจจับพฤติกรรมมนุษย์	
นักศึกษา	นายพิเชษฐ ปานเขียว	57010899
	นางสาวภณิดา เอี่ยมวิจารณ์	57010949
ปริญญา	วิศวกรรมศาสตรบัณฑิต	
สาขาวิชา	วิศวกรรมชีวการแพทย์	
ปีการศึกษา	2560	
อาจารย์ที่ปรึกษา	รศ.ดร.สุรพันธุ์ เอื้อไพบูลย์	

บทคัดย่อ

ปริญญานิพนธ์นี้ได้ถูกจัดทำขึ้นโดยมีจุดประสงค์เพื่อวิเคราะห์และตรวจจับพฤติกรรมของมนุษย์ เพื่อใช้ประโยชน์ในการเฝ้าระวังผู้สูงอายุหรือผู้ป่วย เมื่อมีการลุกออกจากเตียง โดยการใช้งานเว็บแคม (Web Camera) ในการจับภาพและนำภาพที่ได้เข้าสู่กระบวนการประมวลผลภาพ (Image processing) เพื่อหาวัตถุที่มีการเคลื่อนไหวภายในวิดีโอด้วยวิธีการออปติคัล โฟลว์ (Optical flow) ในกระบวนการนี้จะทำให้ได้มุม (Angle) ของการเคลื่อนที่ของวัตถุที่เปลี่ยนไปในทุกเฟรมของวิดีโอ ในการทำกิจกรรมต่าง ๆ ในชีวิตประจำวันของมนุษย์จากลักษณะของการขยับร่างกายที่ต่างกัน จึงทำให้ได้รูปแบบของค่ามุมเฉลี่ยที่เปลี่ยนไป ซึ่งเป็นลักษณะเฉพาะของแต่ละท่าทาง จากนั้นบันทึกข้อมูลค่ามุมของแต่ละบุคคลใน 5 กิจกรรมที่แตกต่างกัน คือ การเดิน การนั่ง การยืน การลุก และการลงจากเตียง แล้วนำข้อมูลที่ได้เข้าสู่กระบวนการจำแนกข้อมูล (Classification) โดยใช้โครงข่ายประสาทเทียม (Artificial neural network) เพอร์เซปตรอนหลายชั้น (Multi-layer perceptron) ในการจำแนกข้อมูลกิจกรรมต่าง ๆ ซึ่งเป็นการเรียนรู้แบบมีผู้สอน (Supervised learning) โดยแบ่งข้อมูลออกเป็น 2 ส่วน คือ ชุดฝึก (Training set) เป็นร้อยละ 75 ของข้อมูลทั้งหมด และชุดทดสอบ (Test set) เป็นร้อยละ 25 ของข้อมูลทั้งหมด และผลการทดสอบในการจำแนกกิจกรรมต่าง ๆ มีความถูกต้องถึงร้อยละ 96 ของชุดทดสอบ มีผลทำให้การตรวจจับในการเฝ้าระวังการลุกออกจากเตียงของผู้ป่วยหรือผู้สูงอายุมีความแม่นยำสูง จึงสามารถบอกได้หากมีการลุกออกจากเตียงเกิดขึ้น

Thesis	HUMAN BEHAVIOR DETECTION
Student	Pichet Pankhiaw 57010899 Panida Aimvichan 57010949
Degree	Bachelor of Engineering
Program	Biomedical Engineering
Year	2017
Advisor Assistant	Assoc. Prof. Dr. Surapan Airphaiboon

Abstract

This thesis is created to analyze and detect the behavior of humans. It is used to monitor patients and elder when they rise from bed. In capturing it use web camera and bring the video that is captured into image processing for finding the movement of an object by optical flow method. The optical flow will cause the angles of object movement in all frames of the video. In optical flow part, there are feature of different body movement in activities that is pattern of averaging angular changing and then record the pattern that from feature of 5 body movement in activity (Activities: walking, sitting down, standing up, getting up and getting off from the bed). After optical flow part, the pattern is taken into classification method. The classification method uses artificial neural network by multi - layer perceptron to classify all of activity data. That is supervised learning which divided data into 2 parts that is 75 percent of training set and 25 percent of testing set. The accuracy of testing reaches to 96 percent that cause of high precision for patient and elder monitoring that able to notify when they rise from bed.

กิตติกรรมประกาศ

งานวิจัยการตรวจจับพฤติกรรมมนุษย์สำเร็จลงได้ เนื่องจากได้รับคำแนะนำในการแก้ปัญหาต่างๆ ทั้งในกระบวนการเก็บข้อมูลวิดีโอ และวิธีการหารูปแบบในการเคลื่อนที่ของพฤติกรรมของมนุษย์ รวมทั้งแนวคิดในการสร้างประโยชน์จากงานวิจัยนี้ จาก รศ.ดร.สุรพันธุ์ เอื้อไพบูลย์ หัวหน้าภาควิชาอิเล็กทรอนิกส์ ผู้เป็นอาจารย์ที่ปรึกษาของงานวิจัยนี้และ รศ.ดร.ชูชาติ ปิณฑวิรุจน์ ที่ช่วยติดตามความคืบหน้าของงานวิจัยและให้คำปรึกษาด้านการประมวลผลภาพซึ่งเป็นส่วนสำคัญในงานวิจัย ขอขอบพระคุณอาจารย์ทั้งสองท่านเป็นอย่างสูงที่ทำให้งานวิจัยนี้สำเร็จลงและสามารถนำไปใช้ประโยชน์ในอนาคตได้

นายพิเชษฐ ปานเขียว
นางสาวภณิดา เอี่ยมวิจารณ์

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VI
สารบัญรูป.....	VII
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของการวิจัย.....	1
1.3 สมมุติฐานของงานวิจัย.....	2
1.4 ทฤษฎีหรือแนวความคิดที่ใช้ในการวิจัย.....	2
1.5 ขอบเขตของการศึกษา.....	2
1.6 ผลที่คาดว่าจะได้รับ.....	2
1.7 เครื่องมือที่ใช้การดำเนินงาน.....	2
1.8 โครงสร้างของงานวิจัย.....	3
บทที่ 2 ทฤษฎีที่เกี่ยวข้อง.....	4
2.1 การประมวลผลภาพ (Image Processing).....	4
2.1.1 Color System.....	4
2.1.2 Motion tracking.....	7
2.2 ปัญญาประดิษฐ์ (Artificial Intelligence : AI).....	10
2.3 แมชชีน เลิร์นนิง (Machine Learning).....	11
2.3.1 Machine Learning Algorithm.....	11
2.4.1 เซลล์ประสาทเทียม (Artificial Neuron).....	12
2.4.2 เซลล์ประสาท (Neuron).....	14
2.5 ภาษาไพธอน (Python).....	20
2.5.1 เวอร์ชันของไพธอน.....	20
2.5.2 การใช้งานภาษาไพธอน.....	21

สารบัญ(ต่อ)

	หน้า
2.7 กล้องเว็บแคม (Webcam).....	22
2.7.1 ชนิดของกล้องเว็บแคม.....	22
2.7.2 ส่วนประกอบของกล้องเว็บแคม (Web camera).....	20
2.7.3 คุณสมบัติทางเทคนิคของกล้องที่ใช้ในปริณูฎานิพนธ์.....	23
บทที่ 3 วิธีการดำเนินงานวิจัย.....	24
3.1 เตรียมข้อมูล (Video recording).....	24
3.2 กระบวนการวิเคราะห์ภาพวิดีโอ (Video analysis).....	27
3.2.1 ติดตั้งโปรแกรม (Program installing).....	27
3.2.2 การประมวลผลภาพ (Image processing).....	28
3.3 การฝึกเพื่อจดจำและแยกแยะข้อมูล (Machine learning).....	30
บทที่ 4 ผลการวิจัย.....	32
บทที่ 5 สรุปผลงานวิจัย อภิปรายผล และข้อเสนอแนะ.....	34
5.1 สรุปผลงานวิจัย.....	34
5.2 ข้อจำกัดของงานวิจัย.....	34
5.3 ข้อเสนอแนะ.....	35
เอกสารอ้างอิง.....	36
ภาคผนวก.....	37
ภาคผนวก ก Optical Flow.....	38
ภาคผนวก ข Neural Network Model.....	39
ภาคผนวก ค Real-time capture.....	40

สารบัญตาราง

ตารางที่	หน้า
ตาราง 4.1 ผลการทดสอบแสดงจำนวนของข้อมูลที่ถูกทดสอบแต่ละพฤติกรรม.....	32

สารบัญรูป

รูปที่	หน้า
2.1 RGB Cube.....	5
2.2 ระบบสี HSV.....	6
2.3 ระดับสีของ Grayscale.....	7
2.4 แสดงการหมุนรอบตัวเองของวัตถุ.....	7
2.5 แสดงออปติคอลลโฟล์ว (Optical flow).....	8
2.6 วิวัฒนาการของระบบ (Artificial Intelligence).....	10
2.7 Activate function.....	15
2.8 ส่วนประกอบของ Neuron Network.....	16
2.8 Perceptron.....	18
2.9 กล้องเว็บแคมแบบมีสายพร้อมขาตั้งกล้อง.....	23
3.1 สถานที่ในการถ่ายวิดีโอ.....	24
3.2 ตำแหน่งของบุคคลในการถ่ายวิดีโอ.....	25
3.3 ความสูงของกล้องจากพื้นดิน 0.87 เมตร.....	25
3.4 การเดิน.....	26
3.5 การนั่ง.....	26
3.6 การยืน.....	26
3.7 การลุกจากเตียง.....	26
3.8 การลงจากเตียง.....	27
3.9 ภายในโปรแกรม Anaconda3.....	27
3.10 โปรแกรม Spyder.....	28
3.11 ภาพ Gray scale ขนาด 384x216 พิกเซล (Pixels)	29
3.12 Optical flow ที่ได้จากภาพ Grayscale.....	30
3.13 Multi-layer perceptron.....	30
3.14 โครงสร้างการทำงานของวิเคราะห์พฤติกรรมของมนุษย์.....	31

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

สังคมประเทศไทยกำลังก้าวเข้าสู่สังคมผู้สูงอายุ เนื่องจากสถิติจำนวนผู้สูงอายุในปัจจุบันมีประมาณ 12 ล้านคน เพิ่มขึ้นจากปีพุทธศักราช 2553 ประมาณ 4 ล้านคน และมีแนวโน้มที่จะเพิ่มขึ้นในอนาคต ธรรมชาติของผู้สูงอายุมักจะมีโรคร้ายที่หลีกเลี่ยงไม่ได้ พร้อมกับร่างกายที่ชราลงจึงต้องมีคนดูแลอย่างใกล้ชิด ในกรณีของผู้สูงอายุบางรายอาจจำเป็นต้องได้รับการดูแลตลอดเวลาและนั่นหมายถึงการที่ลูกหลานหรือญาติอาจต้องสละเวลางานมาเพื่อดูแล หรืออาจเสียค่าจ้างพยาบาลพิเศษมาคอยดูแลอย่างใกล้ชิด เพื่อป้องกันการเกิดอุบัติเหตุ เช่น การพลัดตกบันได การหกล้ม หรืออาจหน้ามีดหมดสติ ผู้ดูแลจะได้เข้ามาช่วยได้อย่างทันเวลา

คนอีกกลุ่มหนึ่งที่ต้องได้รับการดูแลมากเทียบเท่ากับผู้สูงอายุ คือผู้ป่วยที่จำเป็นต้องพักรักษาตัวที่โรงพยาบาล รวมถึงผู้ป่วยที่กลับไปรักษาตัวที่บ้าน ทำให้เกิดปัญหาในการใช้แรงงานคนในการดูแลภายในโรงพยาบาลไม่สามารถมีพยาบาลดูแลผู้ป่วยได้ครบทุกคน หากผู้ป่วยต้องการความช่วยเหลือจะใช้การกดกริ่งเรียกพยาบาล สิ่งที่ตามคือผู้ป่วยคิดว่าสามารถช่วยเหลือตนเองได้จึงไม่ได้เรียกพยาบาล เช่น การเดินไปเข้าห้องน้ำด้วยตัวเอง จึงเป็นเหตุให้เกิดการลื่นล้มหรือหกล้มระหว่างเดินอยู่บ่อยครั้ง หรือในบางกรณีที่ผู้ป่วยมีอาการเดินละเมอ อาจเดินไปในที่เสี่ยงเป็นอันตรายต่อชีวิตในกรณีที่ผู้ป่วยที่กลับไปรักษาตัวที่บ้าน ก็ต้องได้รับการดูแลอย่างใกล้ชิดเช่นกัน

จะเห็นได้ว่าในการดูแลผู้ป่วยหรือผู้สูงอายุ จำเป็นต้องใช้แรงงานคนในการดูแล จึงทำให้เกิดการเสียค่าใช้จ่ายในการจ้างคนดูแล หรืออาจต้องสละเวลางานมาเพื่อดูแล งานวิจัยนี้จึงได้จัดทำขึ้นมาเพื่อสร้างนวัตกรรมที่สามารถใช้เฝ้าระวังดูแลผู้สูงและผู้ป่วย โดยใช้เว็บแคม (Web camera) แทนการใช้แรงงานคนได้ หากมีพฤติกรรมที่ผิดปกติ ให้มีการแจ้งเตือนเพื่อสามารถเข้าไปช่วยเหลือได้ทันเวลา โดยไม่จำเป็นต้องใช้เซ็นเซอร์ (Sensor) หรือมาร์คเกอร์ (Marker) ติดที่ผู้ป่วยหรือผู้สูงอายุ ไม่ทำให้เกิดอุปสรรคในการใช้ชีวิตประจำวัน และสามารถดูแลได้อย่างทั่วถึง ลดค่าใช้จ่ายในการจ้างคนดูแลและไม่กระทบกับการทำงานในชีวิตประจำวันของผู้ดูแล

1.2 วัตถุประสงค์ของการวิจัย

1.2.1 เพื่อวิเคราะห์และเรียนรู้พฤติกรรมเคลื่อนไหวของมนุษย์จากภาพวิดีโอ โดยใช้วิธีการทางแมชชีน เลิร์นนิง (Machine learning)

1.2.2 เพิ่มประสิทธิภาพในการเฝ้าระวังผู้ป่วยและผู้สูงอายุโดยใช้เว็บแคม (Web camera)

1.3 สมมติฐานของงานวิจัย

เว็บแคม (Web Camera) สามารถนำมาใช้ในการถ่ายวิดีโอเพื่อวิเคราะห์พฤติกรรมของมนุษย์ โดยใช้วิธีการออปติคัล โฟลว์ (Optical flow) และโครงข่ายประสาทเทียม (Artificial neural network) ในการวิเคราะห์ข้อมูลท่าทางหรือพฤติกรรมของมนุษย์ได้ และมีความถูกต้องในการเรียนรู้มากกว่าร้อยละ 80

1.4 ทฤษฎีหรือแนวความคิดที่ใช้ในการวิจัย

ใช้ทฤษฎีออปติคัล โฟลว์ (Optical flow) ในการหารูปแบบการเคลื่อนที่ของวัตถุในภาพ และนำข้อมูลไปวิเคราะห์แยกแยะพฤติกรรมกับแบบจำลองของโครงข่ายประสาทเทียม (Artificial neural-network)

1.5 ขอบเขตของการศึกษา

ในการศึกษาพฤติกรรมของมนุษย์ สามารถวิเคราะห์และแยกแยะพฤติกรรมของมนุษย์ได้ใน 5 พฤติกรรม ได้แก่ การนั่ง การยืน การเดิน การลุกและลงจากเตียง ในสภาพแวดล้อมที่มีเพียง 1 คนที่เคลื่อนไหวภายในภาพ ไม่มีสิ่งแวดล้อมที่เคลื่อนไหวรอบกวน และความสูงของคนในภาพต้องไม่เลยจากขอบภาพทั้งบนและล่าง และมีความสูงมากกว่าครึ่งหนึ่งของภาพ สามารถวิเคราะห์พฤติกรรมได้จาก 1 มุมกล้องเท่านั้น

1.6 ผลที่คาดว่าจะได้รับ

1.6.1 สามารถวิเคราะห์และระบุพฤติกรรมต่าง ๆ ภายใต้ขอบเขตของท่าทางที่กำหนดได้อย่างมีประสิทธิภาพ

1.6.2 สามารถนำไปใช้ประโยชน์ในการดูแลผู้ป่วยและผู้สูงอายุได้

1.7 เครื่องมือที่ใช้การดำเนินงาน

1.7.1 เว็บแคม (Web camera)

1.7.2 โปรแกรมพัฒนาภาษาไพธอน (Python): Anaconda3 (Spyder 3.5), Opencv3

1.7.3 Pytorch

1.8 โครงสร้างของงานวิจัย

งานวิจัยเรื่องการวิเคราะห์พฤติกรรมของมนุษย์นี้ ได้อธิบายถึงขั้นตอนและวิธีการในการวิเคราะห์ภาพและวิดีโอ โดยแบ่งเนื้อหาของบทต่างๆดังต่อไปนี้

บทที่ 1 บทนำ ในบทนี้ได้กล่าวถึงความเป็นมาและความสำคัญของปัญหา วัตถุประสงค์ แนวคิด หรือทฤษฎีที่ใช้ในการทำ ขอบเขตของงานวิจัย สมมุติฐานของงานวิจัย ผลที่คาดว่าจะได้รับ และเครื่องมือที่ใช้ในการดำเนินงาน

บทที่ 2 ทฤษฎีที่เกี่ยวข้องในบทนี้ได้กล่าวถึงความรู้ทั่วไปเกี่ยวกับ Web Camera และทฤษฎีทางการประมวลผลภาพ (Image Processing) ในการทำออปติคัล โฟลว์ (Optical flow) รวมถึงทฤษฎีโครงข่ายประสาทเทียม (Artificial neural network)

บทที่ 3 วิธีการดำเนินงาน ในบทนี้ได้กล่าวถึงวิธีการและขั้นตอนในการทำ Video analysis และการฝึกของข้อมูลด้วยมัลติเลเยอร์ เพอร์เซ็ปตรอน (Multi-layer perceptron) และวิธีการออกแบบการทดลอง

บทที่ 4 ผลการทดลอง

บทที่ 5 สรุปผลการทดลองและข้อเสนอแนะ ในบทนี้ได้กล่าวถึงผลสรุปของงานวิจัย ปัญหาและอุปสรรคที่พบ ข้อเสนอแนะและแนวทางในการพัฒนาต่อไป

บทที่ 2

ทฤษฎีที่เกี่ยวข้อง

2.1 การประมวลผลภาพ (Image Processing)

การประมวลผลภาพ (Image Processing) เป็นการประยุกต์ใช้งานการประมวลผลสัญญาณบนสัญญาณ 2 มิติ เช่น ภาพนิ่ง (ภาพถ่าย) หรือภาพวีดิทัศน์ (วิดีโอ) และยังรวมถึงสัญญาณ 2 มิติอื่นๆ ที่ไม่ใช่ภาพด้วย

การประมวลผลภาพดิจิทัล (Digital Image Processing : DIP) เป็นสาขาที่กล่าวถึงวิธีการต่าง ๆ ที่ใช้การประมวลผลภาพที่อยู่ในรูปแบบดิจิทัล และยังรวมถึงสัญญาณดิจิทัลใน 2 มิติอื่นๆ ครอบคลุมถึงสัญญาณวิดีโอ (Video) หรือภาพเคลื่อนไหวซึ่งจะเป็นชุดของภาพนิ่งเรียกว่าเฟรม (Frame) เมื่อหลายๆภาพต่อกันไปตามเวลาซึ่งก็คือสัญญาณ 3 มิติ เมื่อนับเวลาเป็นมิติที่ 3 หรืออาจจะครอบคลุมถึงสัญญาณ 3 มิติอื่น ๆ เช่น ภาพ 3 มิติทางการแพทย์[10]

วิธีการทางการประมวลผลภาพดิจิทัลสามารถแบ่งได้ดังนี้

1. Image Acquisition
2. Image Enhancement
3. Image Restoration
4. Color Image Processing
5. Wavelets and multi resolution processing
6. Compression
7. Morphological processing
8. Segmentation
9. Representation and Description
10. Object recognition

2.1.1 Color System

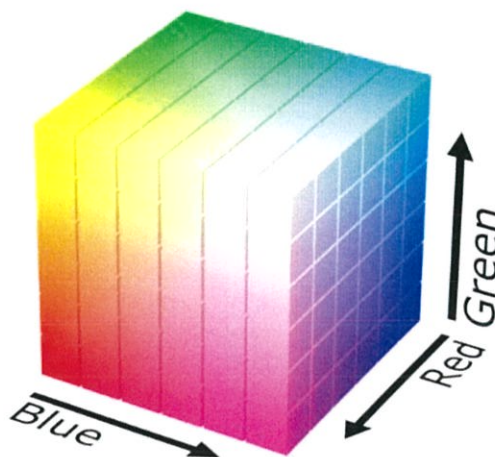
2.1.1.1 ระบบสี YCbCr

YCbCr เป็นหนึ่งในระบบสัญญาณสีที่นำมาใช้ในระบบภาพดิจิทัลวิดีโอ ความแตกต่างระหว่างระบบสี YCbCr กับ RGB คือ YCbCr แสดงภาพโดยใช้สัญญาณความสว่างและสัญญาณความต่างสี

(Color Difference) อีกสองสัญญาณ คำว่า YCbCr ตัวอักษร Y มาจากความสว่าง (Luminance) Cb คือสีน้ำเงินที่ตัดความสว่างออกไปและ Cr คือสีแดงที่ตัดความสว่างออกไป

สัญญาณ YCbCr เป็นสัญญาณชนิดดิจิทัล ซึ่งดิจิทัลทีวี (Digital Television) ใช้สัญญาณชนิด YCbCr รวมทั้งกล้องถ่ายภาพดิจิทัล (MiniDV , DV , Digital Betacam) ยังส่งสัญญาณ YCbCr ผ่านทางพอร์ตดิจิทัล[1]

2.1.1.2 ระบบสี RGB



รูปที่ 2.1 RGB Cube

(ที่มา : https://commons.wikimedia.org/wiki/File:RGB_color_solid_cube.png)

ระบบสี RGB เป็นระบบสีที่เกิดจากการรวมกันของแสงสีแดง สีเขียวและสีน้ำเงิน โดยมีการรวมกันแบบแอดดิทีฟ (Additive) ซึ่งโดยปกติถูกนำไปใช้ในจอภาพแบบซีอาร์ที (Cathode ray tube : CRT) ในการใช้งานระบบสี RGB และมีการสร้างมาตรฐานที่แตกต่างกันออกไปที่นิยมใช้งาน ได้แก่ RGBCIE และ RGBNTSC

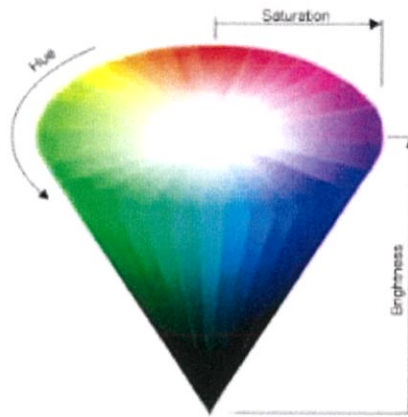
ระบบสีแบบ RGB ของซีไออี (Commission International Eclairage : CIE) เป็นระบบสีที่พัฒนาขึ้นโดยอ้างอิงสีด้วยสีแดงที่ 700 nm สีเขียวเท่ากับ 546.1 nm และสีน้ำเงิน 435.8 nm

ระบบสีแบบ RGB ของเอ็นทีเอสซี (National Television System Committee : NTSC) เป็นระบบที่พัฒนาโดย เพื่อใช้สำหรับการแสดงภาพของจอภาพแบบ CRT เป็นมาตรฐานสำหรับผู้ผลิตแบบ CRT ให้มีลักษณะเดียวกัน[1]

2.1.1.3 ระบบสี HSV

ระบบสี HSV เป็นการพิจารณาสีโดยใช้ Hue Saturation และ Value ซึ่ง Hue คือค่าสีของสีหลักได้แก่สีแดง สีเขียวและสีน้ำเงินในทางปฏิบัติจะมีค่าอยู่ระหว่าง 0 ถึง 255 ซึ่งถ้า Hue มีค่าเท่ากับ 0 จะแทนสีแดง เมื่อ Hue มีค่าเพิ่มขึ้นเรื่อย ๆ สีก็จะเปลี่ยนแปลงไปตามสเปกตรัมของสี

จนถึง 256 จึงจะกลับมาเป็นสีแดงอีกครั้ง ซึ่งสามารถแทนให้อยู่ในรูปขององศาได้ ดังนั้นคือ สีแดงเท่ากับ 0 องศา สีเขียวเท่ากับ 120 องศา สีน้ำเงินเท่ากับ 240 องศา Hue สามารถคำนวณได้จากระบบสี RGB ได้ดังนี้[1]



รูปที่ 2.2 ระบบสี HSV

(ที่มา : <http://support.robotis.com/en/product/darwinop/operating>)

จากลักษณะแบบจำลองของระบบ Hue พบว่าจะมีค่าน้อยหนึ่งค่าที่จะเท่ากับ 0 แต่ถ้ามีสองค่าเท่ากับ 0 แล้ว Hue จะเป็นมุมของสีจะมีค่าเป็นไปตามสีที่สามและถ้าทั้งสามสีมีค่าเท่ากับ 0 แล้วจะทำให้ไม่มีค่าของ Hue หรือสีที่ได้จะมีค่าเท่ากับสีขาว เช่น จอภาพสีขาวดำ ถ้าเกิดมีสีใดสีหนึ่งมีค่าเท่ากับ 0 จะทำให้ค่าสีที่ได้เป็นไปตามสีที่เหลือ การให้นำหนักในการพิจารณาเมื่อสีแดงมีค่าเท่ากับ 0[1]

$$\text{Hue} = \frac{(240 \times \text{blue}_k) + (120 \times \text{green}_k)}{\text{blue}_k + \text{green}_k} \quad (2.1)$$

Saturation คือความบริสุทธิ์ของสีซึ่ง Saturation มีค่าเท่ากับ 0 แล้วสีที่ได้จะไม่มี Hue ซึ่งจะเป็นสีขาวล้วนหาก Saturation มีค่าเท่ากับ 255 แสดงว่าไม่มีแสงสีขาวผสมอยู่ Saturation สามารถคำนวณได้ดังนี้[1]

$$\text{Saturation} = \frac{\max(\text{red}, \text{green}, \text{blue}) - \min(\text{red}, \text{green}, \text{blue})}{\max(\text{red}, \text{green}, \text{blue})} \quad (2.2)$$

Value คือความสว่างของสี ซึ่งสามารถวัดได้โดยค่าความเข้มของความสว่างของแต่ละสีที่ประกอบกันสามารถคำนวณได้จาก

$$value = \max(\text{red}, \text{green}, \text{blue}) \quad (2.3)$$

2.1.1.4 สี Grayscale

เป็นระบบสีที่มีเพียง 2 สี คือ สีขาวและสีดำแต่จะมีระดับความเข้มของสีดำ 255 ระดับ รวมกับสีขาวอีกสีหนึ่ง จึงมีเพียง 256 สี



รูปที่ 2.3 ระดับสีของ Grayscale

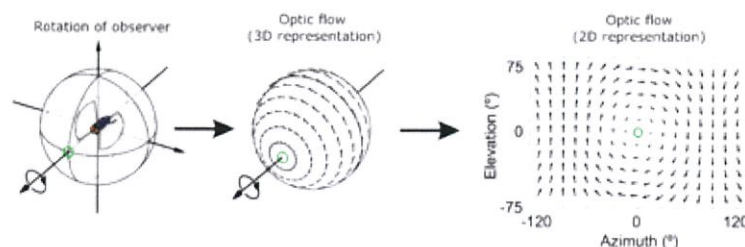
(ที่มา : <http://eliosh.info/library/grayscale-image.html>)

2.1.2 Motion tracking

การตรวจหาและติดตามการเคลื่อนไหวของวัตถุ (Motion tracking) เป็นการหาและแยกแยะวัตถุที่มีการเคลื่อนไหวในภาพวิดีโอ ซึ่งสามารถทำได้หลายวิธี เช่น มินชิฟ (Meanshift) แบคกราวด์ ซับแทรกชัน (Background Subtraction) หรือออปติคอลลโฟว์ (Optical flow) เป็นต้น โดยในส่วนของงานวิจัยนี้เลือกใช้วิธีการของออปติคอลลโฟว์ (Optical flow method)

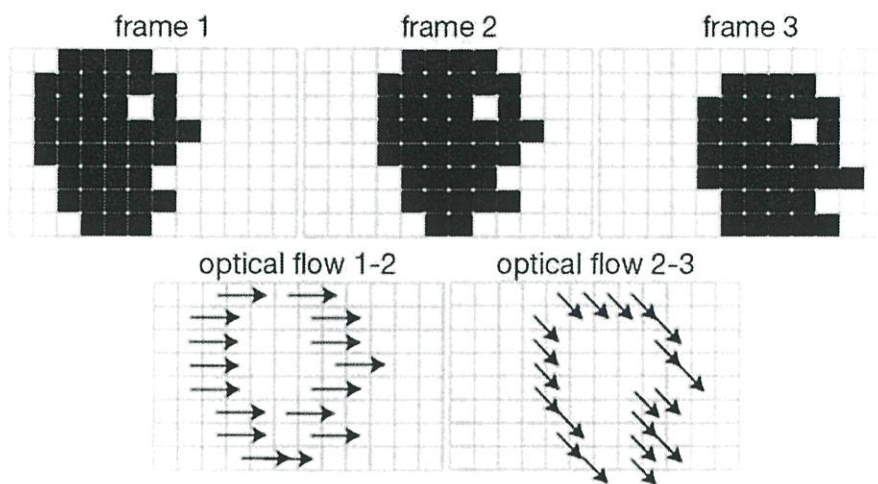
2.1.2.1 ออปติคอลลโฟว์ (Optical Flow)

ออปติคอลลโฟว์ (Optical Flow) เป็นแบบจำลองการคำนวณหาเวกเตอร์ลัพธ์ของภาพเคลื่อนไหว โดยจะแสดงการเคลื่อนไหวของวัตถุ พื้นผิว และเส้นขอบของภาพ



รูปที่ 2.4 แสดงการหมุนรอบตัวเองของวัตถุ จะเห็นเวกเตอร์แสดงการเคลื่อนที่ใน 3 มิติ และเมื่อนำเวกเตอร์ใน 3 มิติ มาเขียนใน 2 มิติจะเห็นเวกเตอร์ที่แสดงทั้งขนาดและมุมในการเคลื่อนที่

(ที่มา : https://en.wikipedia.org/wiki/Optical_flow)



รูปที่ 2.5 แสดงออปติคอลล โฟล์ว (Optical flow) ในการเคลื่อนที่ของวัตถุ
(ที่มา : <http://tcr.amegroups.com/article/view/3200/html>)

การคำนวณหาออปติคอลล โฟล์ว (Optical flow) ของวัตถุหากตำแหน่งแรกอยู่ที่พิกัด (x,y) หากเกิดการย้ายไปที่ตำแหน่งอื่นไปเป็น $x + \Delta x, y + \Delta y$ และ $t + \Delta t$ (เวลาระหว่างเฟรม) ดังสมการ [9]

$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t) \quad (2.4)$$

นำสมการที่ 1 ไปใช้กับ Taylor Series

$$I(x + \Delta x, y + \Delta y, t + \Delta t) = I(x, y, t) + \frac{\partial I}{\partial x} \Delta x + \frac{\partial I}{\partial y} \Delta y + \frac{\partial I}{\partial t} \Delta t + \dots$$

จะได้

$$\frac{\partial I}{\partial x} \Delta x + \frac{\partial I}{\partial y} \Delta y + \frac{\partial I}{\partial t} \Delta t = 0 \quad (2.5)$$

จะได้

$$\frac{\partial I}{\partial x} V_x + \frac{\partial I}{\partial y} V_y + \frac{\partial I}{\partial t} = 0$$

หรือ

$$\frac{\partial I}{\partial x} \frac{\Delta x}{\Delta t} + \frac{\partial I}{\partial y} \frac{\Delta y}{\Delta t} + \frac{\partial I}{\partial t} \frac{\Delta t}{\Delta t} = 0 \quad (2.6)$$

โดยที่ V_x, V_y คือ ออปติคอลล โฟล์ว (Optical flow) หรือความเร็ว (Velocity) ของภาพ $I(x,y,t)$

$\frac{\partial I}{\partial x}$, $\frac{\partial I}{\partial y}$, $\frac{\partial I}{\partial t}$ คืออนุพันธ์ (Derivatives) ของภาพ $I(x,y,t)$ เขียนแทนด้วย I_x , I_y และ I_t ใช้แสดงทิศทางในการเคลื่อนที่ของภาพ

จากสมการ (2.6) จะได้

$$I_x V_x + I_y V_y = -I_t \quad (2.7)$$

หรือ

$$\nabla I^T \cdot \vec{V} = -I_t \quad (2.8)$$

จากสมการนี้ เป็นสมการสองตัวแปรซึ่งไม่สามารถหาคำตอบได้ต้องใช้วิธีการทางคณิตศาสตร์อื่นในการแก้ปัญหา เช่น การใช้สมการพาราเมทริก (Parametric equation) ในการแก้ปัญหา

2.1.2.2 Parametric Equation

1. Lucas-kanade Method ใช้ในการเคลื่อนที่ที่มีขนาดเล็กเวกเตอร์ที่เกิดจะเล็กกว่าภายในพิกเซล (Pixel) การทำงานจะมีหน้าต่าง (Weighted Window) เคลื่อนที่ไปคำนวณทั้งภาพหาเวกเตอร์เฉพาะภายในหน้าต่างที่คำนวณออกมาได้เป็นดังสมการดังนี้[9]

$$\begin{aligned} I_x(q_1)V_x + I_y(q_1)V_y &= -I_t(q_1) \\ I_x(q_2)V_x + I_y(q_2)V_y &= -I_t(q_2) \\ &\vdots \\ I_x(q_n)V_x + I_y(q_n)V_y &= -I_t(q_n) \end{aligned} \quad (2.9)$$

q_1, q_2, \dots, q_n คือพิกเซล (Pixel) ภายในหน้าต่างภาพ $I_x(q_i), I_y(q_i), I_t(q_i)$ อนุพันธ์ย่อยในภาพ I ในตำแหน่ง x,y และเวลาขณะ t จากสมการเขียนเป็นรูปเมทริกได้ดังนี้

$$A = \begin{bmatrix} I_x(q_1) & I_y(q_1) \\ I_x(q_2) & I_y(q_2) \\ \vdots & \vdots \\ I_x(q_n) & I_y(q_n) \end{bmatrix} \quad v = \begin{bmatrix} V_x \\ V_y \end{bmatrix} \quad b = \begin{bmatrix} -I_t(q_1) \\ -I_t(q_2) \\ \vdots \\ -I_t(q_n) \end{bmatrix} \quad (2.10)$$

ใช้วิธี least squares ในระบบ 2×2 แก้ปัญหาหลายตัวแปร

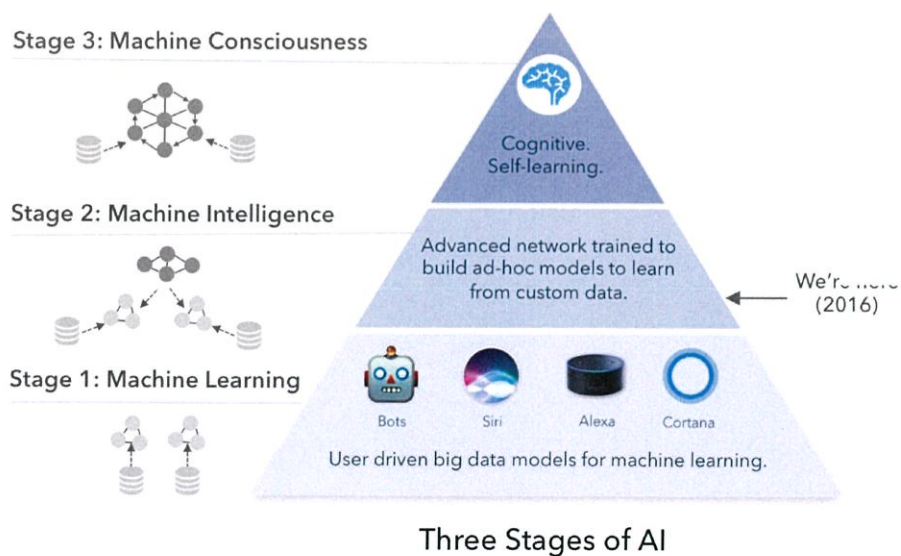
$$\begin{aligned} A^T A v &= A^T b \text{ or} \\ v &= (A^T A)^{-1} A^T b \end{aligned} \quad (2.11)$$

คำนวณ ทรานสโพส (Transpose) ของ A ได้คำตอบของสมการเป็น V_x, V_y

$$\begin{bmatrix} V_x \\ V_y \end{bmatrix} = \begin{bmatrix} \sum_i I_x(q_i)^2 & \sum_i I_x(q_i)I_y(q_i) \\ \sum_i I_y(q_i)I_x(q_i) & \sum_i I_y(q_i)^2 \end{bmatrix}^{-1} \begin{bmatrix} -\sum_i I_x(q_i)I_t(q_i) \\ -\sum_i I_y(q_i)I_t(q_i) \end{bmatrix} \quad (2.12)$$

2.2 ปัญญาประดิษฐ์ (Artificial Intelligence : AI)

เป็นกระบวนการพัฒนาให้เครื่องจักรมีทักษะและความสามารถเหมือนมนุษย์ เช่น การมองเห็น การพูด การตัดสินใจ และมีความรู้สึกนึกคิด คำว่าปัญญาประดิษฐ์ (AI) หมายถึงการสร้างเครื่องจักรที่มีความซับซ้อน โดยมีคุณลักษณะทางด้านสติปัญญาและความฉลาดคล้ายคลึงกับมนุษย์ หรือมีความสามารถมากกว่ามนุษย์ ด้วยแนวคิดนี้จึงถูกเรียกว่า General AI ดังนั้นในมุมมองของปัญญาประดิษฐ์ (AI) จะหมายถึงเครื่องจักรกลหรือระบบที่มีคุณลักษณะและความสามารถที่เหมือนหรือเหนือมนุษย์ ทั้งในแง่สติปัญญา ความสามารถที่มนุษย์สามารถทำได้ ความคิดเชิงเหตุและผล รวมทั้งความคิดเชิงสร้างสรรค์และทักษะทางสังคม เป็นต้น



@AmitPaka

รูปที่ 2.6 วิวัฒนาการของระบบ Artificial Intelligence

(ที่มา : <https://amitpaka.com/three-stages-of-ai-9d2df56dbd08>)

2.3 แมชชีน เลิร์นนิง (Machine Learning)

สมองของมนุษย์มีความสามารถที่มากมาย เช่น การตระหนักรู้ อารมณ์ ความรู้สึก ความทรงจำ ความสามารถในการควบคุมร่างกาย รวมถึงประสาทสัมผัสทั้งห้าที่ทำให้มนุษย์มีความสามารถในการรับรู้ แต่ปัญหาบางมีความซับซ้อนและไม่เหมาะที่จะแก้ปัญหาโดยใช้สมองของมนุษย์เพียงอย่างเดียว เมื่อต้องการจัดการกับข้อมูลจำนวนมากและมีรูปแบบที่แตกต่างกันออกไป เป็นเรื่องยากที่จะทำความเข้าใจข้อมูลและพัฒนาโปรแกรมที่จะตอบสนองต่อข้อมูลนั้น และเมื่อมีข้อมูลเข้ามาใหม่เพิ่มเข้ามาจึงเหมือนกับมีข้อมูลที่เปลี่ยนแปลงตลอดเวลา จึงต้องวิเคราะห์ข้อมูลใหม่และแก้ไขโปรแกรมตลอดเวลาเช่นกันซึ่งทำให้เกิดความยากลำบาก

แมชชีน เลิร์นนิง (Machine Learning) เป็นโปรแกรมที่ทำความเข้าใจความสัมพันธ์ของข้อมูล แล้วสร้างวิธีการตอบสนองต่อข้อมูลด้วยตนเอง ในเมื่อโปรแกรมสามารถเปลี่ยนแปลงวิธีการตอบสนองต่อข้อมูลได้ด้วยตัวเอง เราจึงไม่จำเป็นต้องคอยวิเคราะห์ข้อมูลและแก้โปรแกรมทุกครั้งที่ข้อมูลใหม่เข้ามาอีกต่อไป[6]

2.3.1 Machine Learning Algorithm

สามารถจำแนกตามหน้าที่และขอบเขตปัญหาได้หลักๆ คือการเรียนรู้แบบมีผู้สอน (Supervised learning) การเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning) และการเรียนรู้แบบเสริมกำลัง (Reinforcement learning)

2.3.1.1 การเรียนรู้แบบมีผู้สอน (Supervised Learning)

รูปแบบการเรียนรู้แบบมีผู้สอน เริ่มด้วยการส่งสิ่งเร้าที่ใช้ในการสอนเป็นอินพุต (Input) ในโครงข่ายประสาทเทียม (Neural network) เพื่อให้โครงข่ายประสาทเทียมสร้างผลลัพธ์ตอบออกมาเป็นเอาต์พุต (Output) ซึ่งผลลัพธ์จะเป็นอย่างไรจะขึ้นอยู่กับสถานะในขณะเริ่มต้นเรียนรู้ของโครงข่ายประสาทเทียม ผลลัพธ์ดังกล่าวจะถูกนำมาเปรียบเทียบกับผลลัพธ์เป้าหมาย (Target response) ซึ่งผู้สอน (Teacher) จะเป็นผู้สร้างขึ้นหากผลตอบทั้งสองมีความแตกต่างกันนั้นคือมีความคลาดเคลื่อน (Error) เกิดขึ้น และความคลาดเคลื่อนดังกล่าวจะถูกนำไปคำนวณการปรับแต่งค่าน้ำหนัก (Weight) ต่าง ๆ ในโครงข่ายประสาทเทียมเพื่อลดความคลาดเคลื่อนลงให้เหลือน้อยที่สุด

การปรับแต่งค่าน้ำหนักโดยพิจารณาจากความคลาดเคลื่อนนี้จะขึ้นอยู่กับกฎการเรียนรู้หรือขั้นตอนการคำนวณซึ่งเรียกว่า อัลกอริทึม (Algorithm) ที่แตกต่างกัน โดยแต่ละอัลกอริทึมจะมีคุณลักษณะและความสามารถแตกต่างกัน อย่างไรก็ตามวิธีการส่วนใหญ่ของการเรียนรู้แบบมีผู้สอนนี้ จะถูกดัดแปลงมาจากวิธีการทางคณิตศาสตร์ ในเรื่องของวิธีการหาค่าเหมาะสม (Optimization technique)[6]

เนื่องจากเจตนาของมนุษย์ในการพัฒนาเครื่องมือขึ้นมาใช้งานนั้นจะขึ้นอยู่กับการทำงานที่มนุษย์ต้องการสั่งการและควบคุมเครื่องมืออื่น ๆ ให้ทำงานได้ตามต้องการ จึงทำให้โครงข่ายประสาทเทียมประเภทที่ใช้การเรียนรู้แบบมีผู้สอนได้รับความนิยมในการนำไปประยุกต์ใช้มากที่สุด เนื่องจากสามารถควบคุมได้ การสั่งการโครงข่ายประสาทเทียมจะเป็นไปโดยทางอ้อม ลักษณะของการฝึกสอนโครงข่ายประสาทเทียมเป็นการสร้างข้อมูลตัวอย่าง รวมทั้งค่าเป้าหมายที่จะให้โครงข่ายเรียนรู้ เมื่อโครงข่ายประสาทเทียมเรียนรู้ข้อมูลตัวอย่างได้ถูกต้องหมดแล้ว ความรู้ที่โครงข่ายประสาทเทียมได้เก็บไว้ในลักษณะของค่าน้ำหนักต่าง ๆ จะเป็นสิ่งที่ถูกนำไปใช้งานจริงเพื่อสร้างผลตอบต่อข้อมูลใหม่ที่โครงข่ายไม่เคยเห็นมาก่อน ดังนั้นสำหรับความรู้ของโครงข่ายประสาทเทียมหนึ่งๆ และวิธีการเรียนรู้แบบมีผู้สอนวิธีการหนึ่งๆ นั้นจะสามารถนำไปใช้งานจริงได้เพียงใด จะขึ้นอยู่กับคุณภาพของข้อมูลตัวอย่างที่นำไปสอนนั้น หากข้อมูลตัวอย่างมีจำนวนมากพอโครงข่ายประสาทเทียมจะสามารถสร้างความรู้ได้อย่างถูกต้อง อย่างไรก็ตามด้วยระดับความเจริญก้าวหน้าด้านโครงข่ายประสาทเทียมในปัจจุบันวิธีการวิเคราะห์ปัญหาเพื่อสร้างข้อมูลตัวอย่าง สำหรับการเรียนรู้ที่ประสิทธิภาพนั้นยังคงเป็นงานวิจัยที่ต้องมีการค้นคว้าต่อไป[6]

2.3.1.2 การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning)

การเรียนรู้แบบไม่มีผู้สอนนั้น ไม่จำเป็นต้องมีค่าเป้าหมายของแต่ละข้อมูลตัวอย่าง ในระหว่างการเรียนรู้โครงข่ายประสาทเทียมจะได้รับข้อมูลกระตุ้นในรูปแบบต่าง ๆ และจะทำการจัดกลุ่มรูปแบบต่าง ๆ เหล่านั้นตามต้องการ ผลลัพธ์ของโครงข่ายประสาทเทียมที่ใช้การเรียนรู้แบบไม่มีผู้สอนนี้ จะเป็นการระบุกลุ่มของข้อมูลที่ใส่เข้าไป โดยจะขึ้นอยู่กับวิธีการจัดกลุ่มซึ่งได้เรียนรู้จากข้อมูลที่โครงข่ายประสาทเทียมเคยพบ

ตัวอย่างการเรียนรู้แบบไม่มีผู้สอน เช่น การให้เด็กเล็ก ๆ จัดเก็บสิ่งของไว้บนชั้นวางของให้เป็นระเบียบเรียบร้อย หากเด็กคนหนึ่งเลือกเก็บหนังสือต่าง ๆ ไว้ที่ชั้นบน เก็บตุ๊กตาไว้ที่ชั้นล่างและเก็บของเล่นอื่น ๆ ไว้ที่ชั้นกลาง หลังจากนั้นหากเด็กคนนั้นมีตุ๊กตาตัวใหม่ ก็จะไปเก็บไว้ที่ชั้นล่าง เป็นต้น

แม้ว่าการเรียนรู้แบบไม่มีผู้สอนจะไม่ต้องการผู้สอนแต่ต้องการแนวทางในการจัดกลุ่ม เช่น การจัดกลุ่มอาจจะจัดตามรูปทรง สี หรือวิธีการใช้งานของวัตถุต่าง ๆ ที่ถูกนำมาจัด เป็นต้น ดังนั้นหากไม่มีการให้แนวทางที่ชัดเจนว่าการจัดกลุ่มควรเป็นไปตามคุณลักษณะใด การจัดกลุ่มอาจไม่ประสบความสำเร็จ ในแง่การนำมาใช้งานจริงได้ เช่น การให้เด็กจัดของไว้บนชั้นวางของนั้น เด็กอาจจะจัดตามใจชอบและไม่เป็นหมวดหมู่ ทำให้ไม่สะดวกต่อการนำสิ่งของใหม่ๆ เข้าไปเก็บรวมด้วย ดังนั้นการใช้งานโครงข่ายประสาทเทียมที่ใช้การเรียนรู้แบบไม่มีผู้สอนจึงมักต้องมีการดำเนินการปรับแต่งข้อมูลเพื่อให้เกิดการเน้นสภาพของคุณลักษณะสำคัญที่ต้องการนำมาเป็นแนวทางในการจัดกลุ่มให้เด่นชัด

หรืออาจเป็นการปรับกฎการเรียนรู้ เพื่อให้เน้นไปที่คุณลักษณะที่ต้องการ เนื่องจากขั้นตอนการเรียนรู้แบบไม่มีผู้สอนนี้ จะมีการระบุกลุ่มของข้อมูลตัวอย่างก่อน เมื่อตัดสินใจได้แล้วว่า ข้อมูลใหม่มีลักษณะที่ควรจัดรวมเข้ากลุ่มใด (หรืออาจถือเป็นกลุ่มใหม่ ในกรณีที่เห็นว่าไม่ควรจัดเข้ากลุ่มใดเลย) หลังจากนั้นจึงมีการปรับคุณลักษณะของกลุ่ม โดยการนำลักษณะของข้อมูลใหม่นี้มาช่วยกำหนดแนวทางการจัดกลุ่มด้วย ในการตัดสินใจว่าข้อมูลใหม่นี้ควรจัดรวมเข้ากลุ่มใด โครงข่ายประสาทเทียมบางชนิดอาจจะใช้วิธีการแข่งขันกันของกลุ่มต่าง ๆ ว่ากลุ่มใดควรได้ข้อมูลดังกล่าวไป การเรียนรู้ในลักษณะนี้จะถูกเรียกว่า การเรียนรู้แบบแข่งขันกัน (Competitive Learning) ในอีกแง่มุมหนึ่ง นับจากจุดเริ่มต้นของการเรียนรู้ ซึ่งไม่มีการจัดกลุ่มข้อมูลในแบบใดเลยจนถึงเวลาที่การจัดเสร็จสิ้นแล้ว จะพบว่าการจัดกลุ่มข้อมูลเกิดขึ้นตามคุณลักษณะบางอย่างของข้อมูลตัวอย่าง ซึ่งการจัดกลุ่มนี้เกิดจากการที่โครงข่ายประสาทเทียมประเมินข้อมูลต่าง ๆ ที่ถูกป้อนเข้าไป ในระหว่างการเรียนรู้จนสร้างเป็นวิธีการจัดกลุ่มขึ้นมาได้ ดังนั้นการเรียนรู้ในลักษณะดังกล่าวจึงถูกเรียกว่าการเรียนรู้แบบจัดตัวเอง (Self-organizing)[6]

2.3.1.2 การเรียนรู้แบบเสริมกำลัง (Reinforcement learning)

แนวความคิด คือ จะเรียนรู้จากสิ่งแวดล้อมรอบตัว เรียนรู้จากธรรมชาติรอบตัวที่มีอยู่ในชีวิตจริง นำมาดัดแปลงให้กับคอมพิวเตอร์ เช่น การเล่นเกมหมากรุก จะต้องมีการทำนายล่วงหน้าว่าจะสามารถเกิดอะไรขึ้นได้ ซึ่งการเดินแต่ละครั้งอาจจะไม่เป็นผลดีต่อครั้งนั้นแต่อาจมีผลดีในครั้งต่อ ๆ ไป[6]

2.3.2 เซลล์ประสาทเทียม (Artificial Neuron)

เซลล์ประสาทเทียมเป็นหน่วยรากฐานของโครงข่ายประสาทเทียม และไม่สามารถใช้เป็นแบบในการอธิบายการทำงานของเซลล์ประสาทของสิ่งมีชีวิตได้ถูกต้อง แต่เป็นการนำเอาแนวคิดที่ได้จากความเข้าใจการทำงานของเซลล์ประสาทของสิ่งมีชีวิตมาประยุกต์ใช้ แบบจำลองพื้นฐานของเซลล์ประสาทเทียมมีการทำงานแบบเซลล์ประสาทของสิ่งมีชีวิตคือทำหน้าที่รวมสัญญาณที่เข้ามายังเซลล์ประสาทเทียมเปรียบเสมือนว่าเป็นสัญญาณที่เข้าทางเดนไดรต์ (Dendrite) ของเซลล์ประสาทของสิ่งมีชีวิตแล้วปล่อยสัญญาณกระตุ้นออกไป หากผลรวมของสัญญาณอินพุตนั้นมีค่าสูงเกินค่าระดับที่ตั้งไว้ (Threshold) เปรียบเหมือนการปล่อยสัญญาณไฟฟ้าออกทางแอกซอน (Axon) จากเซลล์ประสาทของสิ่งมีชีวิต[4]

อย่างไรก็ตามสิ่งที่สำคัญในการจำลองเซลล์ประสาทคือการจำลองไซแนปส์ (Synapses) ในโครงข่ายประสาทเทียมซึ่งเปรียบเสมือนแหล่งสะสมความรู้ของสมอง การจำลองไซแนปส์นั้นใช้ไซแนปส์แต่ละตัวทำหน้าที่เป็นตัวปรับเปลี่ยนสภาพสัญญาณไฟฟ้าที่ส่งมาจากเซลล์ประสาทตัวอื่นๆ ก่อนส่งสัญญาณนั้นผ่านเดนไดรต์เข้าสู่ตัวเซลล์ประสาทและการปรับเปลี่ยนสัญญาณดังกล่าว

และจะขึ้นอยู่กับความเหนียวแน่นของการเชื่อมต่อบริเวณรอยต่อไซแนปส์ โดยความแข็งแรงของการเชื่อมต่อจะเปลี่ยนไปตามความรู้ที่สมองได้เรียนเข้าไป

โดยตัวแปรตัวหนึ่งที่เรียกว่าค่าน้ำหนัก (Weight) ซึ่งมีทั้งค่าบวกและลบ ในการจำลองไซแนปส์ หากค่าน้ำหนักนี้มีขนาดใหญ่หมายความว่าความเหนียวแน่นของรอยต่อไซแนปส์มีค่าสูง ทำให้การส่งผ่านสัญญาณทำได้ดี หากค่าน้ำหนักนี้มีขนาดเล็กจะหมายความว่าสัญญาณจะส่งผ่านรอยต่อไซแนปส์ได้น้อย นอกจากนี้หากค่าน้ำหนักมีค่าเป็นบวกจะหมายความว่า สัญญาณที่วิ่งผ่านรอยต่อไซแนปส์เข้าสู่เซลล์ประสาทเทียมจะเป็นสัญญาณกระตุ้น แต่หากค่าน้ำหนักมีค่าเป็นลบจะหมายความว่า สัญญาณที่ผ่านรอยต่อไซแนปส์เข้าสู่เซลล์ประสาทเทียมจะมีผลเป็นสัญญาณกด การทำให้มีค่าน้ำหนักและแต่ละค่ามีผลต่อสัญญาณเข้าดังที่กล่าวข้างต้น สามารถทำได้โดยกำหนดให้สัญญาณที่วิ่งเข้าเซลล์ประสาทเป็นผลคูณของค่าน้ำหนักกับสัญญาณอินพุตที่ถูกส่งมาจากเซลล์ประสาทเทียมตัวอื่น

ในทางคณิตศาสตร์สัญญาณที่ถูกส่งมาจากเซลล์ประสาทเทียมตัวอื่นเป็นตัวแปร x และหากให้ค่าน้ำหนักเป็น w และเมื่อสัญญาณ x วิ่งผ่านไซแนปส์จะกลายเป็นสัญญาณที่มีค่า wx วิ่งเข้าสู่เซลล์ประสาทเทียม[4]

2.3.3 เซลล์ประสาท (Neuron)

ส่วนที่เล็กที่สุดของโครงข่ายประสาทเทียม (Neural Network) คือเซลล์ประสาท (Neuron) ซึ่งทำหน้าที่คำนวณอินพุต (Input) ที่เข้ามาเพื่อให้ได้ผลลัพธ์ โดยมีส่วนประกอบสำคัญดังนี้

2.3.3.1 สัญญาณเข้า (Input)

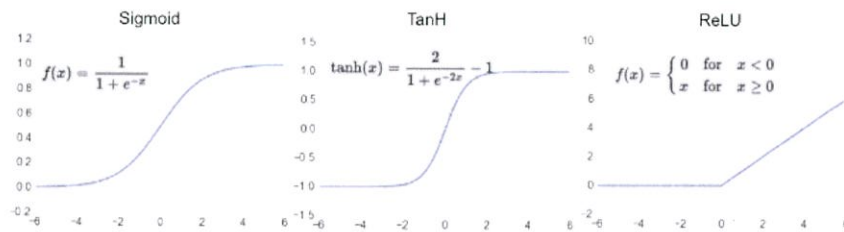
แบ่งข้อมูลวิธีการแบ่งส่วน (Split Test) เป็นการแบ่งส่วนของข้อมูลด้วยการสุ่มออกเป็น 2 ส่วน เช่น ร้อยละ 70 ต่อร้อยละ 30 หรือร้อยละ 80 ต่อร้อยละ 20 โดยข้อมูลส่วนที่หนึ่งคือร้อยละ 70 และร้อยละ 80 ใช้ในการสร้างแบบจำลองของโครงข่ายประสาทเทียมเรียกว่าชุดฝึกสอน (Training set) และข้อมูลส่วนที่สองคือร้อยละ 30 หรือร้อยละ 20 ใช้ในการทดสอบประสิทธิภาพของแบบจำลองเรียกว่าชุดทดสอบ (Test set) และอาจมีข้อมูลอีกชุดหนึ่งที่เรียกว่า Validation set คือข้อมูลที่มีคำตอบเหมือนกันเพื่อส่งให้กับอัลกอริทึมเพื่อปรับแต่งแบบจำลองใช้ในการแก้ปัญหาเรื่องโอเวอร์ฟิต (Overfitting)[4]

2.3.3.2 ค่าน้ำหนัก (Weight)

การให้น้ำหนักของไซแนปส์แต่ละไซแนปส์ที่ส่งเข้ามา โดยมีค่าระหว่าง 0 ถึง 1 เมื่อเริ่มต้นจะเป็นการสุ่ม (Random) จากนั้นเซลล์ประสาท (Neuron) จะทำการเรียนรู้ เป็นการปรับค่าน้ำหนัก (Weight) ในตัวเองให้ได้คำตอบที่ใกล้เคียงที่สุด[4]

2.3.3.3 ไบแอส (Bias)

ค่าที่จะช่วยทำให้ค่าที่เข้ามาอยู่ในระหว่าง 0 ถึง 1 ได้ โดยจะเป็นเลขสุ่ม (Random) และปรับไปทุกครั้งที่เราเรียนรู้[4]



รูปที่ 2.7 Activate function

(ที่มา : <https://www.kdnuggets.com/2017/09/neural-network-foundations-explained-activation-function.html>)

2.3.3.4 ฟังก์ชันกระตุ้น (Activate Function)

1. ฟังก์ชันเส้นตรง (Linear function)

มีความง่ายและสะดวก แต่ไม่สามารถทำให้ผลลัพธ์ (Output) มีค่าอยู่ในช่วงใดช่วงหนึ่ง ซึ่งทำให้ผลออกมาไม่น่าสนใจ

2. ซิกมอยด์ (Sigmoid)

เป็นที่นิยมใช้มายาวนานในอดีตเนื่องจากซิกมอยด์ (Sigmoid) สามารถทำให้ผลลัพธ์ (Output) ให้อยู่ระหว่าง 0 ถึง 1 ซึ่งจะคล้ายกับพฤติกรรมของเซลล์ประสาทจริงคือมีการปล่อยสัญญาณออกมา แต่ปัญหาของซิกมอยด์ (Sigmoid) คือที่ปลายทั้งสองข้างที่แบนราบทำให้เกิดปัญหาการเหวี่ยง (Vanishing gradient) หากสร้างแบบจำลองโดยใช้วิธีการถ่ายทอดย้อนกลับ (Back propagation)[2]

3. แทนห์ (Tanh)

แทนห์ (Tanh) ดีกว่าซิกมอยด์ (Sigmoid) เพราะสามารถทำให้ผลลัพธ์ (Output) อยู่ระหว่าง -1 ถึง 1 ซึ่งจะให้ค่าอยู่ระหว่างศูนย์ทำให้ไม่เกิดปัญหาการเหวี่ยง (Vanishing gradient)[2]

4. ลีดู (ReLU)

ลีดู (Rectified Linear Unit : ReLU) จะให้ค่ากระโดดเฉพาะจากข้างล่างที่ศูนย์ แต่ข้างบนไม่มีการกระโดดของค่าจึงเป็นการแก้ปัญหาการเหวี่ยง (Vanishing gradient) แต่เนื่องจากค่าที่กระโดดจากข้างล่างที่ศูนย์ทำให้มีเซลล์ประสาท (Neuron) ที่มีค่าศูนย์ออกมามาก[2]

2.3.3.5 ผลลัพธ์ (Output)

ก่อนผลลัพธ์ (Output) จะมีฟังก์ชันซอฟต์แม็กซ์ (Softmax) ซึ่งเป็นขั้นตอนสุดท้ายที่จะใส่ค่าความน่าจะเป็นของเป้าหมายทั้งหมดออกมาหรือเป็นผลการจำแนก โดยค่าความน่าจะเป็นจะคำนวณได้ดังนี้[2]

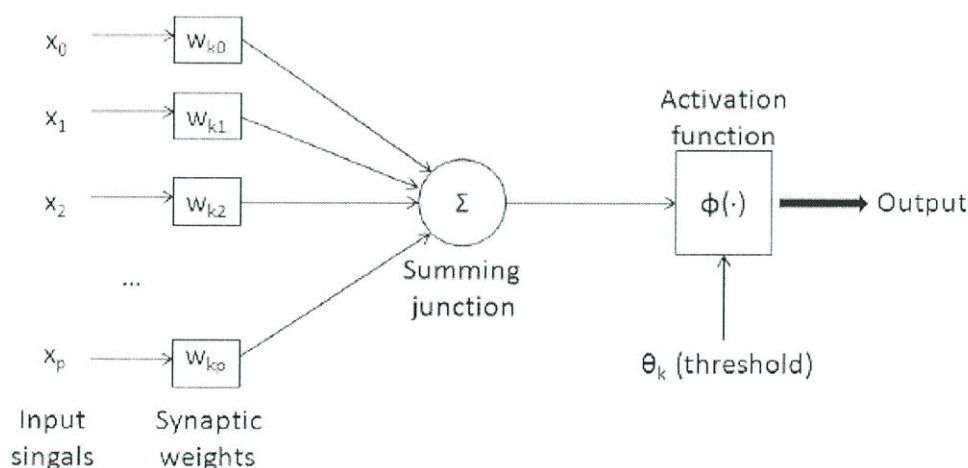
$$P(c_i | x) = \frac{P(x | c_i)P(c_i)}{\sum_{j=1}^K P(x | c_j)P(c_j)} = \frac{e^{o_i}}{\sum_{j=1}^K e^{o_j}} \quad (2.13)$$

โดยที่ $P(c_i)$ คือ ค่าความน่าจะเป็นที่จะเกิดคลาสที่ i

$P(x | c_i)$ คือ ความน่าจะเป็นที่จะเป็นตัวอย่าง x เมื่อเป็นคลาสที่ i

K คือ จำนวนคลาสทั้งหมด

o_i คือ ค่าผลลัพธ์ของเซลล์ประสาทตัวที่ i



รูปที่ 2.8 ส่วนประกอบของ Neuron Network

(ที่มา : https://www.researchgate.net/figure/Mathematical-model-of-an-artificial-neural-network-ANN_fig2_284516167)

การเรียนรู้ของโครงข่ายประสาทเทียม ผลการวิจัยทางพฤติกรรมศาสตร์พบว่าการเรียนรู้ของสิ่งมีชีวิตชนิดต่าง ๆ นั้น มีกระบวนการแตกต่างกันในหลากหลายแบบ และแต่ละแบบอาจเหมาะสมกับแต่ละเผ่าพันธุ์ของสิ่งมีชีวิตนั้น ๆ ในสาขาโครงข่ายประสาทเทียมนั้น แนวคิดของกระบวนการเรียนรู้จะประยุกต์มาจากผลการศึกษาทางพฤติกรรมศาสตร์ อาจกล่าวโดยทั่วไปได้ว่าการเรียนรู้คือ

กระบวนการซึ่งระบบประสาทปรับตัวเองไปตามสิ่งเร้า จนกระทั่งสามารถให้ผลลัพธ์ได้ตามต้องการ โดยใช้การปรับตัวแปรที่ควบคุมสภาพของระบบเอง

การเรียนรู้สามารถถูกมองได้ว่าเป็นกระบวนการจัดกลุ่มของสิ่งเร้าที่เข้ามาอย่างต่อเนื่องเมื่อได้รับสิ่งเร้า หากระบบประสาทรู้จักสิ่งเร้านั้นจะให้ผลตอบได้ตามที่เคยเข้าใจไว้ แต่หากไม่รู้จักจะพยายามปรับความเข้าใจในการจัดกลุ่มขึ้นใหม่ ในทางปฏิบัตินั้นระบบประสาทของสิ่งมีชีวิตจะปรับความเหนียวแน่นของการเชื่อมต่อที่ไซแนปส์จนสร้างผลตอบต่อสิ่งเร้าได้ตามที่ต้องการ สถานะที่กระบวนการของการเรียนรู้จึงจะสิ้นสุดลงเป็นสถานะที่ถือว่าระบบประสาทได้รับความรู้ไปแล้ว คำจำกัดความของกระบวนการเรียนรู้หมายถึงขั้นตอนต่อไปนี้

ขั้นที่ 1 โครงข่ายประสาทถูกกระตุ้นด้วยสิ่งแวดลอม

ขั้นที่ 2 โครงข่ายประสาทเกิดการเปลี่ยนแปลง อันเป็นผลมาจากการกระตุ้นดังกล่าว

ขั้นที่ 3 โครงข่ายประสาทตอบสนองต่อสิ่งแวดลอมในแนวทางใหม่ อันเป็นผลมาจากการเปลี่ยนแปลงที่เกิดขึ้นในโครงสร้างภายในโครงข่าย

เมื่อพิจารณาไปที่เฉพาะบริเวณหนึ่งๆ ของโครงข่ายประสาทจะพบว่าการเชื่อมต่อของเซลล์ประสาทที่บริเวณต่าง ๆ นั้นสามารถมีรูปแบบที่แตกต่างกันได้หลายแบบ และกระบวนการเรียนรู้ของแต่ละบริเวณไม่เหมือนกัน ในทำนองเดียวกันวิธีการเรียนรู้ของโครงข่ายประสาทเทียมจึงแตกต่างกันไปสำหรับแต่ละชนิดของโครงข่าย

2.3.3.6 สถาปัตยกรรมโครงข่ายประสาทเทียม

การเชื่อมโยงเซลล์ประสาทเทียมจำนวนหนึ่งเข้าด้วยกันเป็นโครงข่ายประสาทเทียมนั้นสามารถเชื่อมโยงแบบใดก็ได้โดยไม่ขอบเขตจำกัด อย่างไรก็ตามในทางปฏิบัติวิธีการเรียนรู้ของโครงข่ายประสาทเทียมมักจะถูกออกแบบให้ใช้งานได้กับสถาปัตยกรรมโครงข่ายประสาทเทียมที่มีลักษณะเฉพาะเท่านั้น สถาปัตยกรรมโครงข่ายประสาทเทียมที่พบทั่วไปจะมีลักษณะการจัดเซลล์ประสาทเทียมเป็นชั้น ๆ (Layer) ชั้นที่รับข้อมูลเข้าเรียกว่าชั้นอินพุต (Input layer) ชั้นที่ผลิตผลตอบของโครงข่ายเรียกว่าชั้นเอาต์พุต (Output layer) ส่วนชั้นอื่นที่มีส่วนในการช่วยทำการประมวลผลอยู่ภายในเรียกว่าชั้นซ่อน (Hidden Layer : HL) ในโครงข่ายประสาทเทียมอาจมีชั้นซ่อนได้หลายชั้น โครงสร้างพื้นฐานจะมีลักษณะเป็นการประกอบกันของรูปแบบดังต่อไปนี้

1. เพอร์เซ็ปตรอน (Perceptron)

เพอร์เซ็ปตรอน (Perceptron) ประกอบด้วยเซลล์ประสาทเทียมอย่างน้อยหนึ่งชั้น อินพุตจะถูกส่งตรงไปยังเอาต์พุตโดยผ่านชุดค่าน้ำหนัก ด้วยเหตุนี้จึงสามารถพิจารณาว่าเป็นโครงข่ายการใส่ข้อมูลไปหน้า (Feedforward network) แบบพื้นฐานที่สุดเซลล์ประสาทแต่ละตัวจะคำนวณผลรวมค่าถ่วงน้ำหนักจากอินพุต ซึ่งคือผลรวมของผลคูณระหว่างค่าอินพุตกับค่าน้ำหนักของแต่ละตัว หากค่าที่ได้

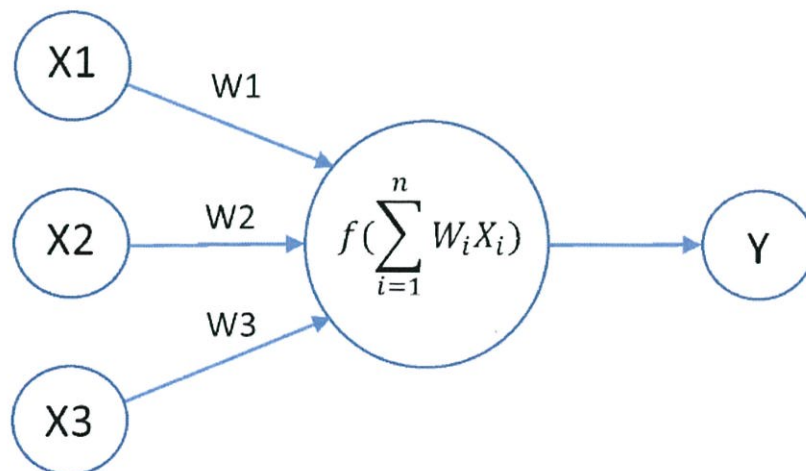
จากการคำนวณมีค่ามากกว่าระดับ (Threshold) เซลล์ประสาทจึงจะให้ค่าเอาต์พุตเท่ากับ 1 ถ้าน้อยกว่าก็จะให้ค่า -1 เพื่อให้ง่ายจึงมักจะกำหนดให้ระดับเป็นค่าน้ำหนักตัวหนึ่งของอินพุตที่เป็นค่าคงที่ โดยฟังก์ชันค่าระดับจะมีศูนย์กลางอยู่ที่ 0

เซลล์ประสาทที่ใช้ฟังก์ชันกระตุ้น (Activation function) อาจจะเรียกว่าเซลล์ประสาทแบบแมคคัลลอค-พิตส์ (McCulloch-Pitts neurons) ในเอกสารทางวิชาการส่วนใหญ่ มักจะใช้คำว่าเพอร์เซ็ปตรอน เมื่อพูดถึงข่ายงานที่ประกอบด้วยเซลล์ประสาทแบบหน่วยเดียว

ขั้นตอนวิธีการเรียนรู้ที่ง่ายที่สามารถนำมาใช้กับเพอร์เซ็ปตรอนจะเรียกว่า กฎเดลต้า (Delta-rule) โดยจะคำนวณค่าความผิดพลาดระหว่างเอาต์พุตที่ได้กับเอาต์พุตของตัวอย่าง แล้วนำค่าที่ได้ไปปรับค่าน้ำหนัก โดยนำไปใช้ในรูปแบบของการเคลื่อนลงตามความชัน (Gradient descent)

แม้ในช่วงแรกเพอร์เซ็ปตรอนจะทำงานได้ตามที่ต้องการแต่ต่อมาก็สามารถพิสูจน์ได้ว่าไม่สามารถสอนให้เพอร์เซ็ปตรอนอย่างง่ายจำรูปแบบต่าง ๆ หลายแบบจนกระทั่งมีการค้นพบว่าโครงข่ายประสาทแบบสามชั้นหรือมากกว่าให้ผลที่ดีกว่าเพอร์เซ็ปตรอนอย่างมาก

เพอร์เซ็ปตรอนอย่างง่าย ที่ประกอบด้วยหนึ่งหรือสองชั้นนั้น สามารถเรียนรู้เฉพาะรูปแบบที่สามารถแบ่งแยกเชิงเส้น (Linearly separable) ซึ่งแสดงให้เห็นว่าไม่สามารถใช้โครงข่ายประสาทเทียมนี้เรียนรู้ฟังก์ชัน XOR และคาดเดาว่าจะได้ผลเหมือนกันแม้ว่าจะใช้โครงข่ายประสาทเทียมที่มีสามชั้นหรือมากกว่า



รูปที่ 2.9 Perceptron

(ที่มา : <https://stackoverflow.com/questions/40537503/deep-neural-networks-precision-for-image-recognition-float-or-double>)

2. เพอร์เซ็ปตรอนหลายชั้น (Multi-layer perceptron : MLP)

โครงข่ายประสาทเทียมแบบ MLP เป็นรูปแบบหนึ่งของโครงข่ายประสาทเทียมที่มีโครงสร้างเป็นแบบหลายชั้น ใช้สำหรับงานที่มีความซับซ้อนได้ผลเป็นอย่างดีโดยมีกระบวนการฝึกฝนเป็นแบบมีผู้สอน (Supervise learning) และใช้ขั้นตอนการส่งค่าย้อนกลับ (Back propagation) สำหรับการฝึกฝนกระบวนการส่งค่าย้อนกลับ ประกอบด้วย 2 ส่วนย่อยคือการส่งผ่านไปข้างหน้า (Forward Pass) การส่งผ่านย้อนกลับ (Backward Pass) สำหรับการส่งผ่านไปข้างหน้าข้อมูลจะผ่านเข้าโครงข่ายประสาทเทียมที่ชั้นข้อมูลเข้า (Input layer) และจะส่งผ่านจากอีกชั้นหนึ่งไปสู่อีกชั้นหนึ่งจนกระทั่งถึงชั้นข้อมูลออก (Output) ส่วนการส่งผ่านย้อนกลับค่าน้ำหนักการเชื่อมต่อจะถูกปรับเปลี่ยนให้สอดคล้องกับกฎการแก้ข้อผิดพลาด (Error-Correction) คือผลต่างของผลลัพธ์ที่แท้จริง (Actual Response) กับผลตอบเป้าหมาย (Target Response) เกิดเป็นสัญญาณผิดพลาด (Error Signal) ซึ่งสัญญาณผิดพลาดนี้จะถูกส่งย้อนกลับเข้าสู่โครงข่ายประสาทเทียมในทิศทางตรงกันข้ามกับการเชื่อมต่อ และค่าน้ำหนักของการเชื่อมต่อจะถูกปรับจนกระทั่งผลตอบที่แท้จริงเข้าใกล้ผลตอบเป้าหมาย

สัญญาณที่มีโครงข่ายประสาทเทียมแบบ MLP มี 2 ประเภทคือฟังก์ชันสัญญาณ (Function Signal) และสัญญาณผิดพลาด (Error Signal) โดยฟังก์ชันสัญญาณเป็นสัญญาณเข้าที่มาจากเซลล์ประสาทในชั้นก่อนหน้าและจะส่งผ่านไปข้างหน้าจากเซลล์ประสาทหนึ่งไปสู่อีกเซลล์ประสาทหนึ่ง สัญญาณผิดพลาด (Error Signal) เป็นสัญญาณย้อนกลับที่เกิดขึ้นที่เซลล์ประสาทในชั้นข้อมูลออกของโครงข่ายประสาทเทียมและถูกส่งผ่านย้อนกลับจากชั้นหนึ่งไปสู่อีกชั้นหนึ่ง

หลักการทำงานของ MLP คือในแต่ละชั้นของชั้นซ่อนตัว (Hidden Layer) จะมีฟังก์ชันสำหรับคำนวณเมื่อได้รับสัญญาณ (Output) จากเซลล์ประสาทในชั้นก่อนหน้านี้ เรียกว่าฟังก์ชันกระตุ้น (Activation Function) โดยในแต่ละชั้นไม่จำเป็นต้องเป็นฟังก์ชันเดียวกัน ชั้นซ่อนตัวนั้นมีหน้าที่สำคัญในแปลงข้อมูลที่เข้ามาในชั้น (Layer) นั้น ๆ ให้สามารถแยกแยะความแตกต่างโดยใช้เส้นตรงเส้นเดียว (Linearly Separable) และก่อนที่ข้อมูลจะถูกส่งไปถึงชั้นข้อมูลออก (Output Layer) ในบางครั้งอาจจำเป็นต้องใช้ชั้นซ่อนตัวมากกว่า 1 ชั้น

ในการคำนวณหาผลลัพธ์ (Output) ในปัญหาการจำแนกทำได้โดยการใส่ข้อมูลเข้า (Input) ไปยังโครงข่ายประสาทเทียมที่ได้สร้างไว้ จากนั้นให้ทำการเปรียบเทียบค่าของผลลัพธ์ (Output) ในชั้นข้อมูลออกและให้ทำการเลือกค่าของผลลัพธ์ (Output) ที่มีค่าสูงกว่า (Neuron ที่มีค่าสูงกว่า) และทำการรับค่าของพยากรณ์ที่ตรงกับเซลล์ประสาทที่เลือก และให้นำค่าของผลลัพธ์มาเปรียบเทียบกับค่าที่ยอมรับได้ หากค่าอยู่ในช่วงที่ยอมรับได้หมายถึงค่าความผิดพลาด (Error) น้อยกว่าที่กำหนด จึงให้ทำการรับข้อมูลชุดถัดไป แต่หากค่าที่ได้มากกว่าค่าที่ยอมรับได้ ให้ทำการปรับค่าน้ำหนักและไบแอส

(Biased) ตามขั้นตอนที่ได้กล่าวไว้ข้างต้น เมื่อทำการปรับน้ำหนักแล้วให้ทำการรับข้อมูลชุดถัดไปและทำตามขั้นตอนซ้ำอีกรอบจนกระทั่งถึงข้อมูลชุดสุดท้าย และเมื่อทำข้อมูลชุดสุดท้ายเสร็จจะนับเป็น 1 รอบของการคำนวณ (1 Epoch) จากนั้นจะทำการหาค่าผิดพลาดรวมและเก็บค่าไว้เพื่อใช้ในการตรวจสอบว่าค่า โดยเฉลี่ยในการจำแนกนั้นมีค่าน้อยกว่าค่าผิดพลาดที่ยอมรับได้หรือไม่ หากใช้แสดงว่าโครงข่ายประสาทเทียมที่สร้างขึ้นนั้นสามารถให้ผลลัพธ์ที่ถูกต้องของทุกข้อมูล จึงจะทำการจบการเรียนรู้ได้ หากไม่ใช่ให้กลับไปทำตามขั้นตอนแรกโดยเริ่มรับข้อมูลชุดที่ 1 ใหม่

2.4 ภาษาไพธอน (Python)

ภาษาโปรแกรมส่วนใหญ่เป็นภาษาระดับสูง ซึ่งในจำนวนนั้นก็ยิ่งแบ่งย่อยออกได้เป็นหลายระดับ ถ้าระดับสูงมากจะง่ายต่อการเขียนแต่จะทำงานช้าลง แต่หากอยู่ในระดับกลางจะเขียนยากกว่าแต่จะทำงานได้เร็ว ซึ่งภาษาซีและฟอร์แทรนถือเป็นภาษาระดับที่สูงไม่มาก ในขณะที่ภาษาไพธอน (Python) เป็นภาษาระดับสูงมากจึงเขียนง่ายและทำงานช้าเมื่อเทียบกับภาษาซีหรือฟอร์แทรนแล้ว แม้จะรู้ว่าทำงานช้าแต่เหตุผลที่ยังคงควรจะใช้ ก็คงเป็นเพราะว่าการใช้งานง่ายและเหมาะกับคนทั่วไปที่ไม่ได้เชี่ยวชาญด้านคอมพิวเตอร์

แม้ภาษาไพธอนจะทำงานช้าแต่ด้วยความเรียบง่ายจะทำให้สามารถประหยัดทั้งเวลาในการเรียนรู้และเวลาในการคิดและพัฒนาโปรแกรม การเริ่มฝึกจากภาษาไพธอนก่อนแล้วจึงต่อยอดไปยังภาษาที่ยากขึ้นทีหลังจะทำได้ไม่ยาก

จุดเด่นของภาษาไพธอน เป็นภาษาที่มีรูปแบบการเขียนที่เข้าใจง่าย เป็นระเบียบ และมีคำสั่งที่อำนวยความสะดวกอยู่ในตัวช่วยให้การเขียนสั้นลง สร้างตัวแปรได้ง่ายไม่ต้องประกาศชนิดของตัวแปรสามารถใช้งานได้หลายแพลตฟอร์ม (Platform) มีชุดคำสั่งเสริม (Library) ที่มีผู้พัฒนาไว้ค่อนข้างมาก สามารถนำมาใช้เพื่อเพิ่มประสิทธิภาพการทำงานได้อย่างกว้างไกล และยังเป็นภาษาที่ถูกนำไปใช้กับซอฟต์แวร์ต่าง ๆ หลากหลาย เช่น Maya, Houdini, Metasequoia[5]

2.4.1 เวอร์ชันของไพธอน

ภาษาไพธอนมีการพัฒนาขึ้นมาตลอดเวลา และมีการปรับปรุงเวอร์ชันใหม่อยู่ตลอดเวลา ในปัจจุบันมีการพัฒนาถึงเวอร์ชัน 3.xx และภาษาไพธอนเวอร์ชัน 2.7 ได้รับความนิยมอย่างมากเพราะในปัจจุบันโปรแกรมต่าง ๆ ที่ใช้ภาษาไพธอนยังสนับสนุนเวอร์ชัน 2.xx เป็นหลัก อย่างไรก็ตามแนวโน้มในอนาคตเวอร์ชัน 3.xx จะได้รับการสนับสนุนมากขึ้น ผู้ฝึกใช้ใหม่ควรใช้ 3.xx มากกว่า ดังนั้นปริญญาบัตรนี้จะเน้นเวอร์ชัน 3.xx เป็นหลัก

2.4.2 การใช้งานภาษาไพธอน

ภาษาไพธอนเปรียบได้กับภาษาไทยหมายถึงเป็นภาษาง่ายที่เข้าใจดีอยู่แล้ว แต่ที่ไม่สามารถนำไปใช้งานได้โดยตรง ส่วนภาษาอังกฤษจึงเปรียบได้กับภาษาคอมพิวเตอร์คือภาษาที่สามารถนำไปใช้งานได้จริงแต่มนุษย์ไม่เข้าใจ ในคอมพิวเตอร์สิ่งที่ทำหน้าที่ตีความภาษาจะเรียกว่าคอมไพเลอร์ (Compiler) นอกจากคอมไพเลอร์แล้วส่วนประกอบที่สำคัญอีกคือส่วนที่มีไว้ใช้สำหรับเขียนข้อความซึ่งเรียกว่าอิดิเตอร์ (Editor) เหมือนกับการเขียนข้อความในกระดาษบอกว่าการให้ทำอะไร

อิดิเตอร์กับคอมไพเลอร์อาจไม่จำเป็นต้องอยู่ภายในโปรแกรมเดียวกัน อาจโค้ดผ่านโปรแกรมง่าย ๆ เช่น notepad จากนั้นค่อยนำไปรัน (Run) กับคอมไพเลอร์ดังนั้นอิดิเตอร์จึงไม่มีความสำคัญเท่าคอมไพเลอร์

อย่างไรก็ตามโปรแกรมที่ใช้ทำงานกับภาษาไพธอนนั้นมักจะประกอบไปด้วยอิดิเตอร์ที่ช่วยอำนวยความสะดวกให้การเขียนนั้นง่าย เช่น การใส่สีให้ข้อความสำคัญ และมีตัวตรวจไวยากรณ์ทำให้มีการฟ้องเวลาเจอข้อผิดพลาดทางวากยสัมพันธ์ (Syntax error) ในขณะที่หากเขียนในโปรแกรม notepad จะไม่สามารถรู้ได้ว่าผิดที่ใด นอกจากนี้ยังมีส่วนที่ทำหน้าที่ค้นหาข้อผิดพลาดในการทำงานของโปรแกรม (Debug)

นอกจากการเขียนโปรแกรมและให้คอมไพเลอร์อ่านแล้ว ในบางภาษาซึ่งรวมถึงภาษาไพธอนนั้นยังมีอีกวิธีหนึ่งในการใช้งานคือการสั่งให้ทำงานแบบคำต่อคำ ส่วนที่ใช้สั่งงานโปรแกรมแบบคำต่อคำนั้นเรียกว่าเชลล์โต้ตอบ (Interactive shell) และในกรณีนี้ตัวประมวลผลจะถูกเรียกว่าอินเทอร์พรีเตอร์ (Interpreter) ข้อดีคือเห็นผลทันทีอย่างรวดเร็ว แต่ข้อเสียคือใส่คำสั่งได้ทีละน้อยและต้องสั่งไปเรื่อย ๆ ไม่สามารถสั่งงานทิ้งไว้แล้วให้ทำงานยาวๆได้

สรุปโดยรวม สิ่งที่ต้องมีเพื่อจะทำงานกับภาษาไพธอน (Python) คือ คอมไพเลอร์ที่ใช้ในการแปลข้อความโค้ดที่เขียนเพื่อสั่งให้คอมพิวเตอร์ทำงานอิดิเตอร์ มีหน้าที่ใช้เขียนโค้ดเพื่อให้คอมไพเลอร์อ่านแล้วสั่งคอมพิวเตอร์ และเชลล์โต้ตอบมีหน้าที่เขียนโค้ดเพื่อสั่งการคอมพิวเตอร์ทันที

สุดท้ายจะกล่าวถึง IDE ซึ่งเป็นโปรแกรมสำเร็จรูปสำหรับพัฒนาการเขียนโปรแกรมซึ่งประกอบไปด้วยส่วนประกอบต่าง ๆ ที่จำเป็น เช่น คอมไพเลอร์ อิดิเตอร์ และอื่นๆ สำหรับภาษาไพธอนนั้นโดยพื้นฐานจะมี IDE ที่โหลดจากเว็บหลักของผู้พัฒนาไพธอน มีชื่อเรียกว่า IDLE และโปรแกรมสไปเดอร์ (Spyder) จะได้นามสกุลไฟล์ .py ซึ่งจะมาพร้อมกับโปรแกรมอนาคอนดา (Anaconda) สามารถติดตั้งและใช้งานได้ง่าย โดยที่มีเครื่องมือเสริมที่สำคัญที่ติดตั้งมาให้แล้ว เช่น นัมพาย (Numpy) ไชพาย (Scipy) และแพนดาส (Pandas) เป็นต้น[5]

2.5 เว็บแคม (Web camera)

เว็บแคม (Web camera) หรือที่เรียกว่า Video Conference เป็นอุปกรณ์ที่สามารถจับภาพเคลื่อนไหวไปปรากฏในจอมอนิเตอร์ (Display monitor) และสามารถส่งภาพเคลื่อนไหวผ่านระบบเครือข่ายเพื่อให้คนอื่นอีกฟากหนึ่งสามารถเห็นตัวเราเคลื่อนไหวได้เหมือนอยู่ต่อหน้า ถือว่าเป็นอุปกรณ์ที่มีประโยชน์อีกชนิดหนึ่งและเริ่มมีความจำเป็นมากขึ้นในปัจจุบัน

2.5.1 ชนิดของกล้องเว็บแคม

เว็บแคม (Web camera) สามารถแบ่งออกได้ 2 ชนิดได้แก่

2.5.1.1 กล้องเว็บแคมแบบมีสาย

จะมีความยุ่งยากในเรื่องการใช้สายต่อพ่วงเข้ากับเครื่องคอมพิวเตอร์ แต่จะมีราคาถูกกว่าแบบไร้สาย ทำให้คนส่วนใหญ่นิยมซื้อเว็บแคมแบบมีสายมาใช้งาน และข้อเสียของเว็บแคมแบบมีสายคือไม่สามารถวางตัวกล้องไกลจากเครื่องคอมพิวเตอร์ ทำให้ไม่สามารถจับภาพเคลื่อนไหวในระยะไกลได้[8]

2.5.1.2 กล้องเว็บแคมแบบไร้สาย

จะมีราคาค่อนข้างแพงมากเมื่อเทียบกับแบบมีสาย เนื่องจากตัวกล้อง ต้องใช้เทคโนโลยีแบบไร้สายที่เรียกว่า Wireless Wi-Fi หรือ IEEE 802.11 ที่ค่อนข้างมีต้นทุนสูงจึงส่งผลให้ตัวกล้องมีราคาสูงจึงไม่ค่อยได้รับความนิยมนัก จุดเด่นของกล้องเว็บแคมแบบไร้สายคือสามารถนำไปติดตั้งที่จุดใดก็ได้โดยไม่ต้องคำนึงระยะห่างระหว่างตัวกล้องกับคอมพิวเตอร์[8]

2.5.2 ส่วนประกอบของเว็บแคม (Web camera)

2.5.2.1 เลนส์กล้อง

ทำหน้าที่ในการจับภาพเคลื่อนไหวต่าง ๆ ที่เคลื่อนไหวผ่านไปมาอยู่หน้ากล้องหรืออยู่ในตำแหน่งที่เลนส์กล้องสามารถมองเห็นภาพได้

2.5.2.2 ตัวปรับระยะโฟกัส

จะทำหน้าที่ในการปรับโฟกัสของภาพเพื่อให้ภาพมีความชัดเจนมากขึ้น

2.5.2.3 ฐานรองกล้อง

มีไว้สำหรับเป็นที่ตั้งของตัวกล้องซึ่งช่วยให้เราสามารถวางกล้องบนหน้าจอคอมพิวเตอร์ได้สะดวก[8]



ภาพที่ 2.10 กล้องเว็บแคมแบบมีสายพร้อมขาตั้งกล้อง

(ที่มา : <https://www.logitech.com/th-th/product/hd-pro-webcam-c920>)

2.5.3 คุณสมบัติทางเทคนิคของกล้องที่ใช้ในปริิญญาณิพนธ์ ยี่ห้อ Logitech รุ่น c920

2.5.3.1 วิดีโอ Full HD 1080p (สูงสุด 1920 x 1080 พิกเซล)

2.5.3.2 การบีบอัดวิดีโอ H.264 (H.264 คือมาตรฐานการบีบอัดข้อมูลของสัญญาณภาพและเสียงที่ความคมชัดเท่ากัน แต่ขนาดไฟล์จะเล็กกว่าเมื่อเทียบกับ Mpeg4 และ Mpeg2 ตามลำดับ)

2.5.3.3 ไมค์ระบบสเตอริโอคู่แบบในตัวพร้อมการป้องกันเสียงรบกวนอัตโนมัติ

2.5.3.4 แก้วใสภาพแสงน้อยโดยอัตโนมัติ[7]

บทที่ 3

วิธีการดำเนินงานวิจัย

ในการตรวจจับและวิเคราะห์พฤติกรรมของมนุษย์ โดยใช้ Web camera แบ่งขั้นตอนในการทำเป็น 3 ขั้นตอนดังนี้

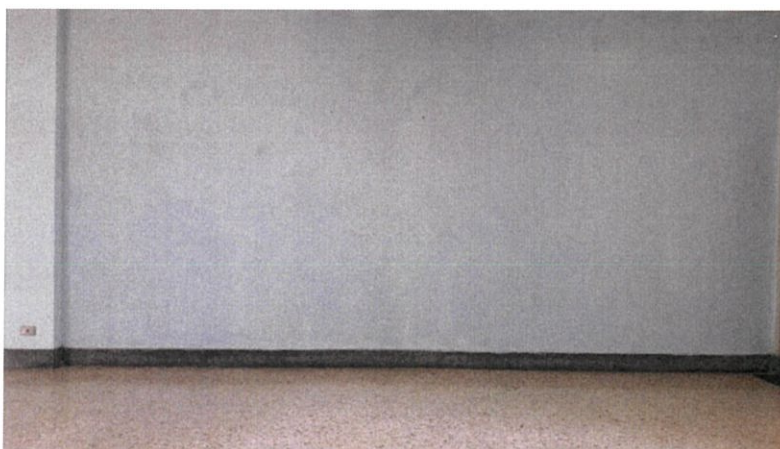
- 3.1 เตรียมข้อมูล (Video recording)
- 3.2 กระบวนการวิเคราะห์ภาพวิดีโอ (Video analysis)
- 3.3 การฝึกเพื่อจดจำและแยกแยะข้อมูล (Machine learning)

3.1 เตรียมข้อมูล (Video recording)

การเตรียมข้อมูลพฤติกรรมของมนุษย์ในการทำกิจกรรมต่าง ด้วยการบันทึกวิดีโอจากกลุ่มตัวอย่างจำนวน 5 พฤติกรรมได้แก่ การเดิน การนั่ง การยืน การลุกและลงจากเตียง

โดยกลุ่มตัวอย่างเลือกจากวัยรุ่นชาย/หญิงที่มีสภาพร่างกายปกติและสามารถดำเนินชีวิตประจำวันได้อย่างปกติ ที่ความสูงอยู่ระหว่าง 160 เซนติเมตรถึง 180 เซนติเมตรจำนวน 404 คน

- 3.1.1 เตรียมพื้นที่ในการถ่ายวิดีโอที่เป็นที่โล่ง ไม่มีปัจจัยภายนอกมารบกวน



รูปที่ 3.1 สถานที่ในการถ่ายวิดีโอ ที่โล่งและไม่มีปัจจัยภายนอกมารบกวน

3.1.2 ให้บุคคลยืนห่างจากพื้นหลังประมาณ 1.2 เมตร และห่างจาก Web camera ประมาณ 3.6 เมตร



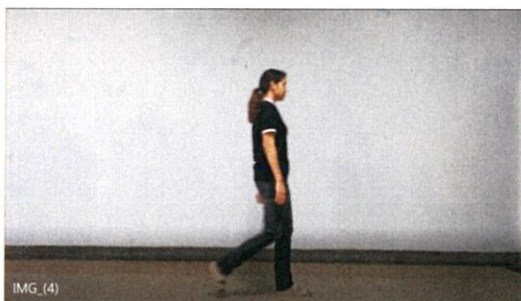
รูปที่ 3.2 ตำแหน่งของบุคคลในการถ่ายวิดีโอ

3.1.3 ใช้ Web camera ยี่ห้อ Logitech รุ่น C920 มีความละเอียด 1920x1080 พิกเซล (Pixels) ตั้งกล้องสูงจากพื้นประมาณ 0.87 เมตรและถ่ายในระนาบเดียวกับทิศทางการเดิน

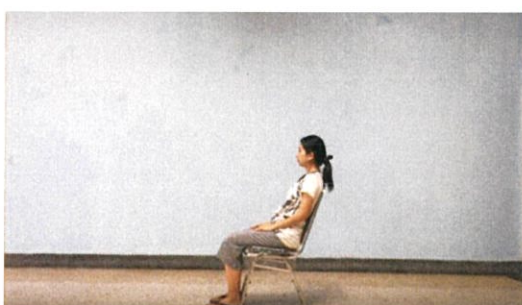


รูปที่ 3.3 ความสูงของกล้องจากพื้นดิน 0.87 เมตร

3.1.4 ถ่ายวิดีโอเก็บข้อมูลในการทำพฤติกรรมต่าง 5 พฤติกรรมได้แก่ การเดิน การนั่ง การยืน การลุกและลงจากเตียง



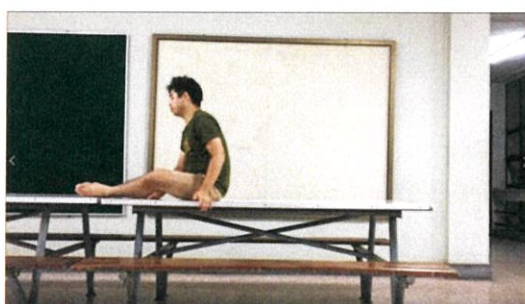
รูปที่ 3.4 การเดิน



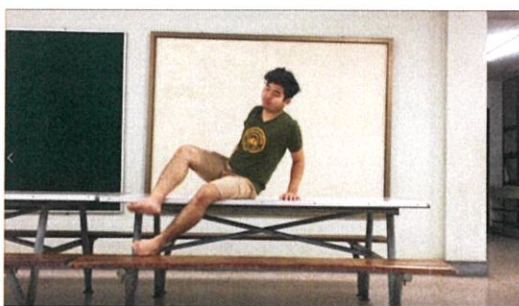
รูปที่ 3.5 การนั่ง



รูปที่ 3.6 การยืน



รูปที่ 3.7 การลุกจากเตียง



รูปที่ 3.8 การลงจากเตียง

3.2 กระบวนการวิเคราะห์ภาพวิดีโอ (Video analysis)

กระบวนการวิเคราะห์ภาพวิดีโอ เพื่อวิเคราะห์พฤติกรรมของมนุษย์มีขั้นตอนดังนี้

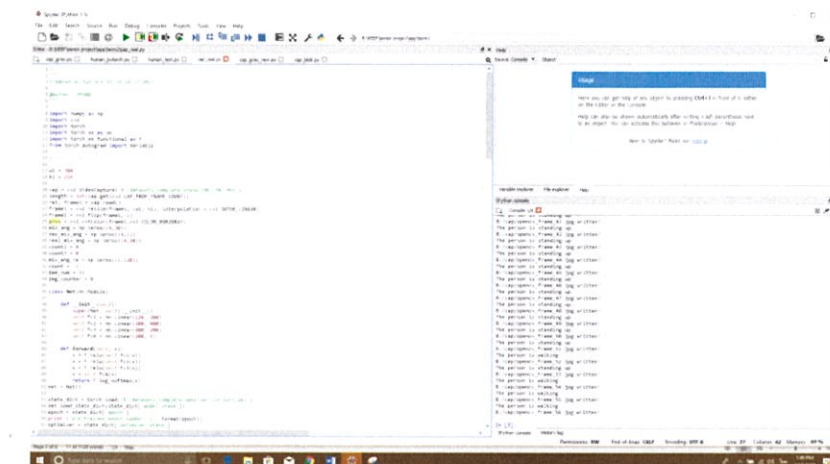
3.2.1 ติดตั้งโปรแกรม (Program installing)

ในงานวิจัยนี้โปรแกรมการวิเคราะห์พฤติกรรมของมนุษย์ในอิริยาบถต่าง ๆ จะถูกพัฒนาจากภาษาไพธอน (Python) โดยโปรแกรมที่ใช้พัฒนาได้แก่ Anaconda3 (Python version 3.5) ข้อดีของโปรแกรม Anaconda3 เป็นโปรแกรมที่ติดตั้งเครื่องมือพื้นฐาน (Library) ที่จำเป็นสำหรับการพัฒนาไว้แล้วขณะติดตั้งโปรแกรม และเลือกใช้โปรแกรม Spyder ในการพัฒนา

ติดตั้งเครื่องมือเพิ่มเติมโดยใช้ชุดคำสั่งเพื่อติดตั้งใน Windows Command หรือ Anaconda Command ในการประมวลผลภาพต้องติดตั้ง Opencv โดยใช้ชุดคำสั่ง `conda install -c conda-forge opencv` และในการสร้างโครงข่ายประสาทเทียม (Artificial- neural network) ติดตั้ง Pytorch โดยใช้ชุดคำสั่ง `conda install -c peterjc123 pytorch`



รูปที่ 3.9 ภายในโปรแกรม Anaconda3



รูปที่ 3.10 โปรแกรม Spyder

3.2.2 การประมวลผลภาพ (Image processing)

การประมวลผลภาพ (Image processing) เพื่อวิเคราะห์พฤติกรรมและแยกแยะท่าทางการเคลื่อนไหวของมนุษย์แบ่งเป็นขั้นตอนได้ดังนี้

3.2.2.1 Pre-Processing

3.2.2.2 คำนำหนาออปติคอล โฟล์ว (Optical flow)

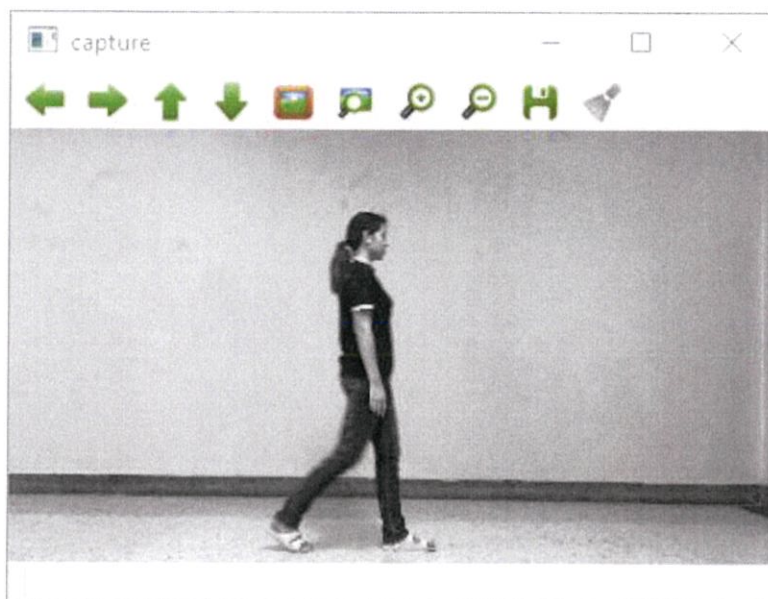
3.2.2.3 การเตรียมชุดข้อมูล (Dataset)

ในขั้นต้นก่อนเริ่มการพัฒนาต้องเรียกใช้เครื่องมือที่สำคัญในการทำการวิเคราะห์พฤติกรรมมนุษย์ด้วยคำสั่ง “Import (module name)” เครื่องมือที่สำคัญได้แก่ numpy เป็นเครื่องมือที่ใช้ในการคำนวณทางคณิตศาสตร์ และ cv2 เป็นการเรียกใช้งาน Opencv เพื่อทำการประมวลผลภาพ

3.2.2.1 Pre-Processing

ในการเปิดข้อมูลวิดีโอ ใช้คำสั่ง `cv2.VideoCapture("path file/file name")` จะได้ภาพวิดีโอที่ขนาดเท่ากับความละเอียดเริ่มต้นของวิดีโอ นั่นๆ ซึ่งจากข้อมูลวิดีโอที่ถ่ายจาก Web camera จะมีขนาด 1920x1080 พิกเซล (Pixels) ซึ่งเป็นภาพที่มีขนาดใหญ่มาก จะทำให้การประมวลผลภาพนั้นช้าลงจึงปรับขนาดของภาพวิดีโอลงให้มีขนาด 384x216 พิกเซล (Pixels) ด้วยคำสั่ง `cv2.resize`

จากนั้นเปลี่ยนระบบสีของภาพจาก RGB เป็น Gray scale เพื่อลดความผิดพลาดในการวิเคราะห์ที่อาจเกิดจากสีของภาพ ด้วยคำสั่ง `cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)`



รูปที่ 3.11 ภาพ Gray scale ขนาด 384x216 พิกเซล (Pixels)

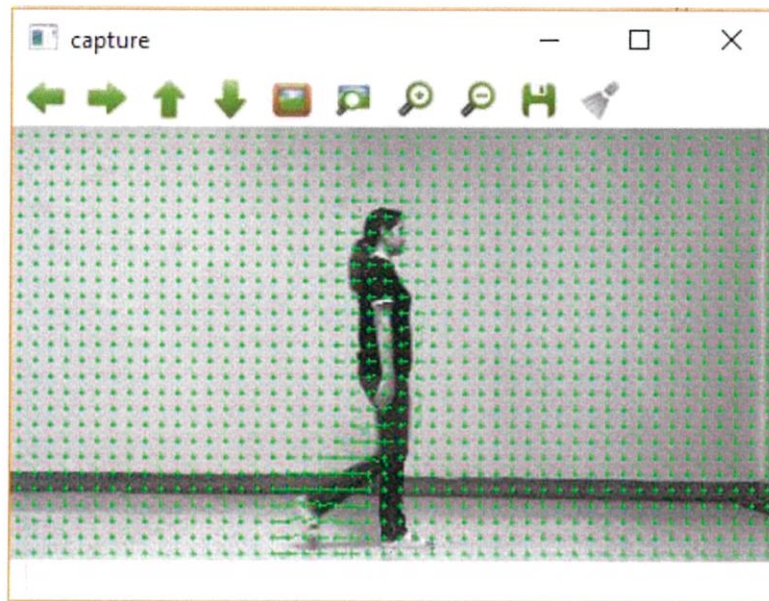
3.2.2.2 คำนวณหาออปติคอลล โฟล์ว (Optical flow)

ในการคำนวณหาออปติคอลล โฟล์ว (Optical flow) เพื่อหารูปแบบของท่าทางในพฤติกรรมต่างๆ โดยใช้คำสั่ง `cv2.calcOpticalFlowFarneback()` เป็นอัลกอริทึมของ Gunnar Farneback

ในการวิเคราะห์หารูปแบบของพฤติกรรมต่างเริ่มจากการแบ่งส่วนของภาพออกเป็น 4 ส่วนในแนวราบและหาออปติคอลล โฟล์ว (Optical flow) ในแต่ละส่วน จากนั้นนำเวกเตอร์ของออปติคอลล โฟล์ว มาแยกองค์ประกอบเป็นขนาด (Magnitude) และมุม (Angle) ซึ่งมุมและขนาดที่ได้ ต้องนำมาผ่านกระบวนการ Normalization เพื่อปรับให้ค่าของขนาดอยู่ระหว่าง 0 ถึง 255 และค่ามุมอยู่ระหว่าง 0 ถึง 180 องศา

จากนั้นเลือกเวกเตอร์ของออปติคอลล โฟล์ว ที่มีขนาดมากกว่า 80 และหาค่ามุมเฉลี่ยของเวกเตอร์ที่เลือกในแต่ละส่วนของภาพ จะทำให้ได้ค่ามุมเฉลี่ย 4 ค่าในภาพ จะเป็นตัวแทนของการเปลี่ยนแปลงท่าทางของมนุษย์ในภาพระหว่างเฟรมภาพ 2 เฟรม

บันทึกค่ามุมเฉลี่ยใน 30 เฟรมเก็บในเมทริกซ์ (Matrix) ขนาด 4x30 (4 Rows 30 columns) โดยแต่ละแถว (Row) คือค่ามุมเฉลี่ยในแต่ละส่วนของภาพ และแต่ละหลัก (Column) คือจำนวนเฟรมที่เพิ่มขึ้น จะทำให้ได้รูปแบบของมุมที่เปลี่ยนแปลงในแต่ละพฤติกรรมที่ต่างกัน



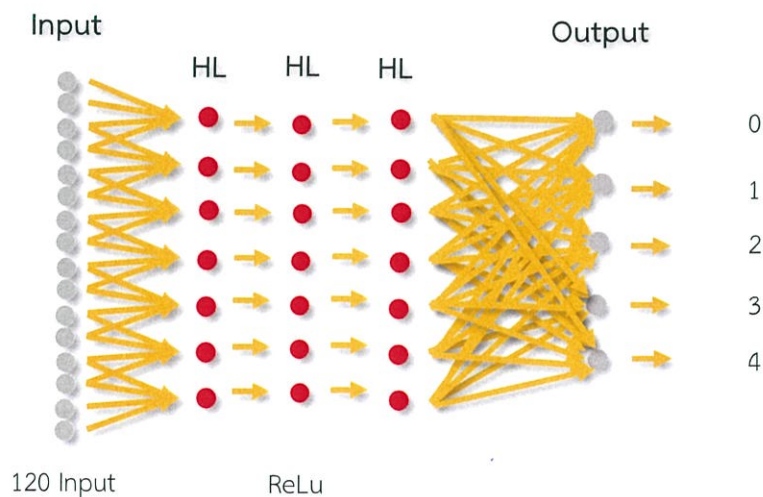
รูปที่ 3.12 Optical flow ที่ได้จากภาพ Grayscale

3.2.2.3 การเตรียมชุดข้อมูล (Dataset)

จากเมทริกซ์ (Matrix) ขนาด 4×30 ที่ได้ นำมาเปลี่ยนแปลงขนาดเป็น 1×120 เพื่อเตรียมเป็น อินพุต (I) สำหรับโครงข่ายประสาทเทียม โดยใช้คำสั่ง `cv2.reshape`

3.3 การฝึกเพื่อจดจำและแยกแยะข้อมูล (Machine learning)

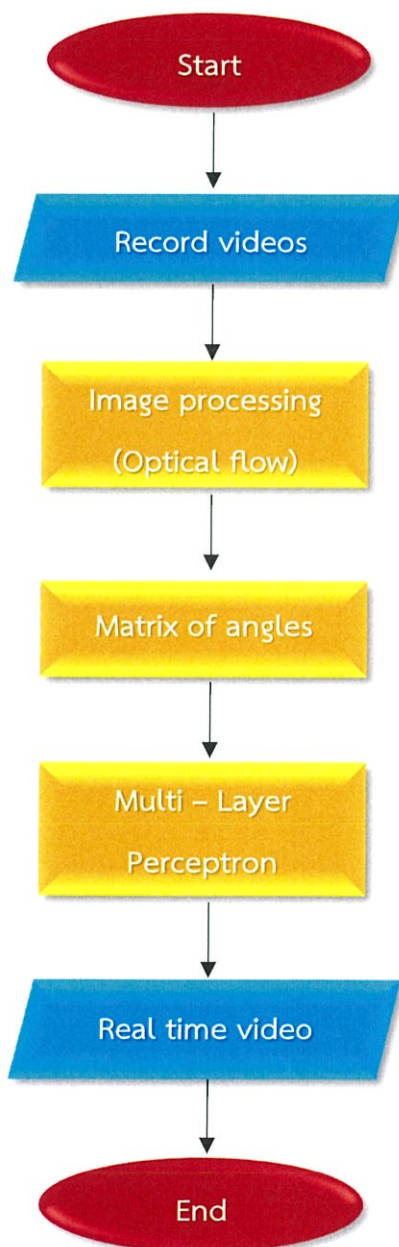
เรียกใช้เครื่องมือในการสร้างโครงข่ายประสาทเทียมของ Pytorch โดยการ import torch และ สร้างคลาสของโครงข่ายประสาทเทียม `class Net(nn.Module)` เป็น MLP ที่มี 120 อินพุต (I) 5 เอาต์พุต (O) และ 3 HL โดยมี Activate function เป็นสมการของ ReLu



รูปที่ 3.13 Multi-layer perceptron

จากข้อมูลวิดีโอทั้งหมด 404 วิดีโอ ทำให้ข้อมูลของรูปแบบของมุมที่เปลี่ยนแปลงในแต่ละพฤติกรรมจำนวน 404 ข้อมูล จากนั้นแบ่งส่วนของข้อมูลออกเป็น 2 กลุ่ม กลุ่มแรกก็กลุ่มที่ใช้ฝึก (Training set) ประมาณร้อยละ 75 จากข้อมูลทั้งหมด และอีกร้อยละ 25 เป็นชุดทดสอบ (Test set) โค้งข่ายประสาทเทียมชุดนี้เป็นการเรียนรู้แบบมีผู้สอน ดังนั้นในแต่ละข้อมูลจะมีการกำหนดชื่อให้แต่ละพฤติกรรมครบทุกข้อมูล ทั้งหมด 5 พฤติกรรมจาก 0 ถึง 4 เช่น 0 ชื่อข้อมูลการเดิน เป็นต้น

จากกระบวนการและขั้นตอนวิธีการในการวิเคราะห์พฤติกรรมของมนุษย์ทั้งหมด สามารถสรุปได้ดังภาพต่อไปนี้



รูปที่ 3.14 โครงสร้างการทำงานของการทำงานของการวิเคราะห์พฤติกรรมของมนุษย์

บทที่ 4

ผลการวิจัย

การทดลองศึกษาและวิเคราะห์พฤติกรรมของมนุษย์ใน 5 อิริยาบถ ได้แก่ การเดิน การนั่ง การยืน การลุกและลงจากเตียง โดยใช้ Web camera การจับภาพจากกลุ่มตัวอย่าง 404 คน มาวิเคราะห์ด้วยการประมวลผลภาพและโครงข่ายประสาทเทียม นำผลที่ได้มาหาค่าความถูกต้องเพื่อตอบวัตถุประสงค์ในการวิจัย โดยมีขั้นตอนดังนี้

1. บันทึกวิดีโอพฤติกรรมของมนุษย์ใน 5 อิริยาบถ รวม 404 คน
2. วิเคราะห์หารูปแบบเฉพาะของแต่ละพฤติกรรมโดยใช้การประมวลผลภาพ
3. แบ่งข้อมูลเป็น 2 กลุ่ม ได้แก่ ชุดฝึก (Training set) และชุดทดสอบ (Test set) ร้อยละ 75 และ 25 ตามลำดับ
4. วิเคราะห์พฤติกรรมโดยใช้โครงข่ายประสาทเทียม และหาค่าความถูกต้องจากการทดสอบ

จากข้อมูลของชุดทดสอบ (Test set) ประกอบด้วย การเดิน 21 ข้อมูล การนั่ง 20 ข้อมูล การยืน 20 ข้อมูล การตื่น 19 ข้อมูล และการลุกออกจากเตียง 19 ข้อมูล

ตาราง 4.1 ผลการทดสอบแสดงจำนวนของข้อมูลที่ถูกทดสอบแต่ละพฤติกรรม

Predicted Actual	การเดิน	การตื่น	การลุกออก จากเตียง	การนั่ง	การยืน
การเดิน	21	0	0	0	0
การตื่น	0	19	0	0	0
การลุกออก จากเตียง	2	0	16	1	0
การนั่ง	0	0	0	20	0
การยืน	0	0	0	0	20

จากผลการทดสอบชุดทดสอบ (Test set) กับแบบจำลองของโครงข่ายประสาทเทียมได้ผลทดสอบ ดังนี้

การเดินจำนวน	21	ข้อมูลมีความถูกต้องร้อยละ	100.00
การตื่นจำนวน	19	ข้อมูลมีความถูกต้องร้อยละ	100.00
การลุกออกจากเตียงจำนวน	19	ข้อมูลมีความถูกต้องร้อยละ	84.21
การนั่งจำนวน	20	ข้อมูลมีความถูกต้องร้อยละ	100.00
การยืนจำนวน	20	ข้อมูลมีความถูกต้องร้อยละ	100.00

บทที่ 5

สรุปผลงานวิจัย อภิปรายผล และข้อเสนอแนะ

5.1 สรุปผลงานวิจัย

งานวิจัยการตรวจจับพฤติกรรมของมนุษย์ กระทำโดยอาศัยหลักการของออปติคอลล โฟลว์ (Optical flow) เพื่อใช้ในการหาลักษณะของรูปแบบเฉพาะของมุมที่เปลี่ยนแปลงในการเคลื่อนไหวในแต่ละพฤติกรรม โดยใช้เว็บแคม (Web camera) ในการบันทึกภาพ จากนั้นนำรูปแบบของมุมที่เปลี่ยนแปลงไปฝึกสอนกับ Machine learning ซึ่งโครงข่ายประสาทเทียม (Artificial neural network) แบบ MLP

เพื่อให้เว็บแคม (Web camera) สามารถวิเคราะห์พฤติกรรมของมนุษย์ได้เอง โดยความถูกต้องจากการทดสอบของชุดทดสอบ (Test set) ที่ได้ทดสอบกับแบบจำลองของ MLP สรุปผลได้ดังนี้ ในการเดิน การยืน การนั่ง และการตื่นนอน มีความถูกต้องร้อยละ 100 และการลุกออกจากเตียง มีความถูกต้องร้อยละ 84.21 ทำให้มีความถูกต้องเฉลี่ยประมาณร้อยละ 96

จากผลการทดลองจะเห็นได้ว่ามีความผิดเกิดขึ้นในการจับการออกจากเตียง ซึ่งจากการทดลองระบบวิเคราะห์การลงจากเตียงเป็นการเดิน 2 ข้อมูล และเป็นการนั่ง 1 ข้อมูล ความผิดพลาดอาจเกิดจากลักษณะการเคลื่อนที่มีมุมในการขยับร่างกายใกล้เคียงกัน และอาจเกิดจากการเลือกพื้นที่ในการหาค่ามุมเฉลี่ยในการทำการประมวลผลภาพไม่เหมาะสมสำหรับในบางข้อมูล

อย่างไรก็ตามจากผลการทดลองที่ได้ทำให้ประสบความสำเร็จไปอีกขั้นในการทำให้คอมพิวเตอร์สามารถเรียนรู้ วิเคราะห์ และแยกแยะพฤติกรรมของมนุษย์ได้ด้วยตัวเอง เตรียมพร้อมเข้าสู่ยุคของ AI (Artificial Intelligence) โดยสามารถนำงานวิจัยนี้ไปใช้ประโยชน์ได้หลายทาง เช่น การเฝ้าระวังผู้ป่วยหรือผู้สูงอายุแทนพยาบาลหรือผู้เฝ้าไข้ เพื่อเป็นการลดภาระและใช้เวลาให้คุ้มค่าในการทำอย่างอื่นได้มากขึ้น รวมทั้งสามารถลดต้นทุนในการจ้างพยาบาลหรือผู้เฝ้าไข้แล้วผู้ป่วยหรือผู้สูงอายุได้ด้วย

5.2 ข้อจำกัดของงานวิจัย

เนื่องจากจุดประสงค์ของงานวิจัยนี้คือ ผู้สูงอายุและผู้ป่วย แต่กลุ่มตัวอย่างนั้นมาจากช่วงอายุ 20-25 ปี ซึ่งเป็นช่วงอายุที่ไม่ได้อยู่ในช่วงวัยสูงอายุ และกลุ่มตัวอย่างนั้นเป็นคนที่มีความสุขภาพปกติ ไม่ได้เป็นโรคใด ๆ แบบจำลองที่ได้นั้นเมื่อนำไปทดลองกับผู้สูงอายุและผู้ป่วยจริง ร้อยละความถูกต้องอาจคลาดเคลื่อนได้

5.3 ข้อเสนอแนะ

สำหรับการทำงานวิจัยต่อเนื้องนี้ เพื่อความสมบูรณ์ของงานวิจัย การวิเคราะห์พฤติกรรมครอบครัวครอบคลุมปัญหาของผู้สูงอายุและผู้ป่วย เช่น ปัญหาของผู้สูงอายุคือการล้ม เนื่องจากการทรงตัวของผู้สูงอายุค่อนข้างที่จะมีปัญหาตามวัยที่มากขึ้น และปัญหาของผู้ป่วยที่ค่อนข้างน่าเป็นห่วงคือ โรคที่เกิดฉับพลันเวลาผู้ป่วยอยู่คนเดียว ไม่สามารถขอความช่วยเหลือได้ทันท่วงที เช่น โรคลมชัก ซึ่งถ้าสามารถทำกล้องตรวจจับได้ เมื่อเจอพฤติกรรมที่เป็นโรคฉับพลัน จึงจะสามารถที่จะแจ้งเตือนไปยังพยาบาลได้ จึงส่งผลให้ไม่เกิดการสูญเสียกับผู้ป่วย และทำให้สามารถช่วยเหลือผู้ป่วยได้อย่างเร็วที่สุดเท่าที่เป็นไปได้

เอกสารอ้างอิง

- [1] VIKAS GUPTA. (2017). **Color spaces in OpenCV (C++ / Python)**. August 18, 2017, Web site: <https://www.learnopencv.com/color-spaces-in-opencv-cpp-python/>
- [2] ไม่ปรากฏนามผู้แต่ง. **Convolutional neural network for visual recognition**. December 10, 2017, Web site: <http://cs231n.github.io/neural-networks-1/>
- [3] จตุภัทร เมฆพ่ายัพ และ กิตติการ สายธน. (2011). “Performance of Neural Networks, Multi-Layer Perceptron and Radial Basis Function,for Multivariate Quality Control Charts”.
- [4] Anonymous. (2017). **สรุปแนวคิด Neural Network แบบไม่มี Math**. December 20, 2017, Web site: <https://coladev.com/machine-learning/neural-network/2017/02/22/neural-network-basic>
- [5] ไม่ปรากฏนามผู้แต่ง. (2017). **แนะนำภาษา Python**. December 22, 2017, Web site: <http://marcuscode.com/lang/python/introduction>
- [6] ไม่ปรากฏนามผู้แต่ง. **Machine learning คืออะไร**. December 20, 2017, Web site: <https://medium.com/@dumpdatasci.th/machine-learning>
- [7] ไม่ปรากฏนามผู้แต่ง. **SPECIFICATIONS C920 Web camera**. August 30, 2017, Web site : http://support.logitech.com/en_us/product/hd-pro-webcam-c920/specs
- [8] Dounjai Khuntasamon.(2556). **Web camera**. August 30, 2017, Web site: <http://dounjai.blogspot.com/2013/08/webcam.html>
- [9] doxygen. (2017). **Optical flow**. September 10, 2017 Web site: https://docs.opencv.org/3.3.1/d7/d8b/tutorial_py_lucas_kanade.html
- [10] ไม่ปรากฏนามผู้แต่ง. **Digital image processing**. February 10, 2018, Web site: http://www.researchsystem.siam.edu/images/thesistee/AUTOMATIC_VEHICLE_MODEL_DETECTION_SYSTEM/09_%E0%B8%9A%E0%B8%97%E0%B8%97_2.pdf

ภาคผนวก

ภาคผนวก ก

Optical flow

```

77 flow = cv2.calcOpticalFlowFarneback(pvs,gray, None, 0.5, 3, 15, 3, 5, 1.2, 0)
78 prvs=gray
79 mag, ang = cv2.cartToPolar(flow[...,0], flow[...,1])
80 angle = ang*180/np.pi/2
81 magnitude = cv2.normalize(mag, None, 0, 255, cv2.NORM_MINMAX)
--

```

รูปที่ 1 ฟังก์ชันการทำออปติคอลล โฟล์ว (Optical flow) ของ Opencv

```

59 def draw_flow(img, flow, step=8):
60     h, w = img.shape[:2]
61     y, x = np.mgrid[step/2:h:step, step/2:w:step].reshape(2,-1).astype(int)
62     fx, fy = flow[y,x].T
63     lines = np.vstack([x, y, x+fx, y+fy]).T.reshape(-1, 2, 2)
64     lines = np.int32(lines + 0.5)
65     vis = cv2.cvtColor(img, cv2.COLOR_GRAY2RGB)
66     cv2.polylines(vis, lines, 0, (0, 255, 0))
67     for (x1, y1), (x2, y2) in lines:
68         cv2.circle(vis, (x1, y1), 1, (0, 255, 0), -1)
69     return vis
70

```

รูปที่ 2 ฟังก์ชันในการวาดเวกเตอร์ของออปติคอลล โฟล์ว (Optical flow)

ภาคผนวก ข

Neural Network Model

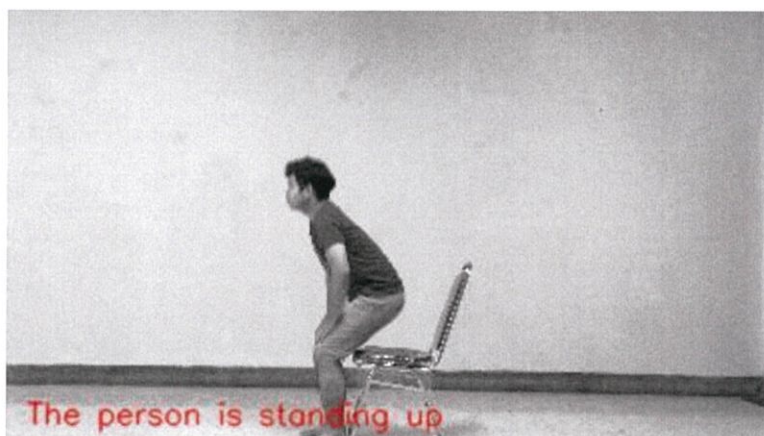
```
36 class Net(nn.Module):
37
38     def __init__(self):
39         super(Net, self).__init__()
40         self.fc1 = nn.Linear(120, 200)
41         self.fc2 = nn.Linear(200, 400)
42         self.fc3 = nn.Linear(400, 200)
43         self.fc4 = nn.Linear(200, 5)
44
45     def forward(self, x):
46         x = F.relu(self.fc1(x))
47         x = F.relu(self.fc2(x))
48         x = F.relu(self.fc3(x))
49         x = self.fc4(x)
50         return F.log_softmax(x)
51 net = Net()
--
```

รูปที่ 3 Multi-layer perceptron 3 Hidden layer, 120 input with ReLu function

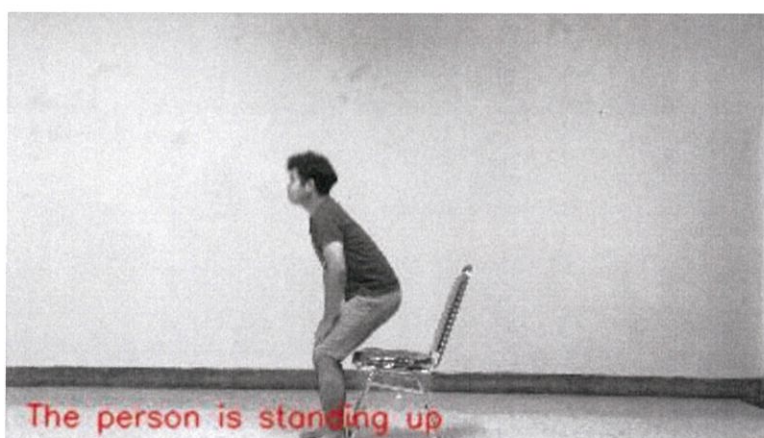
ภาคผนวก ค

Realtime Capture

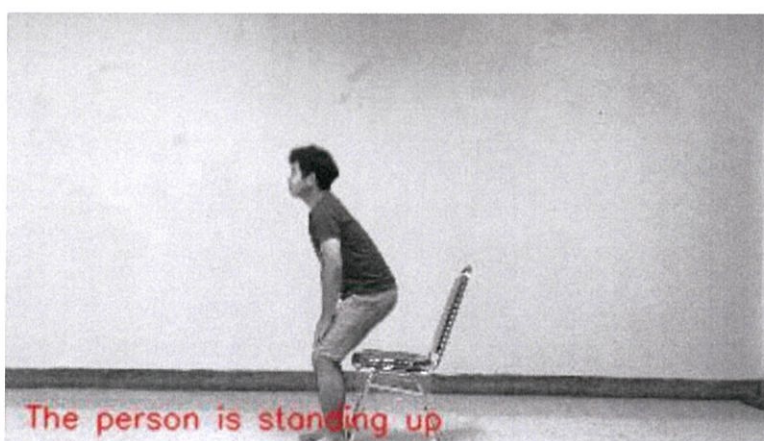
Real-time capture while stand up



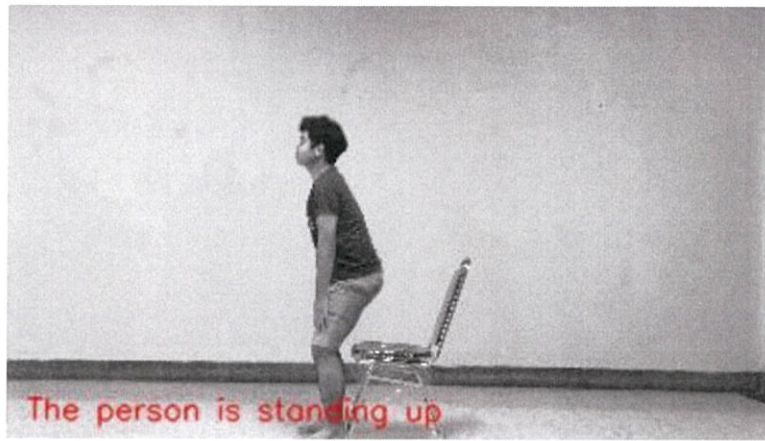
รูปที่ 4 The person is standing up



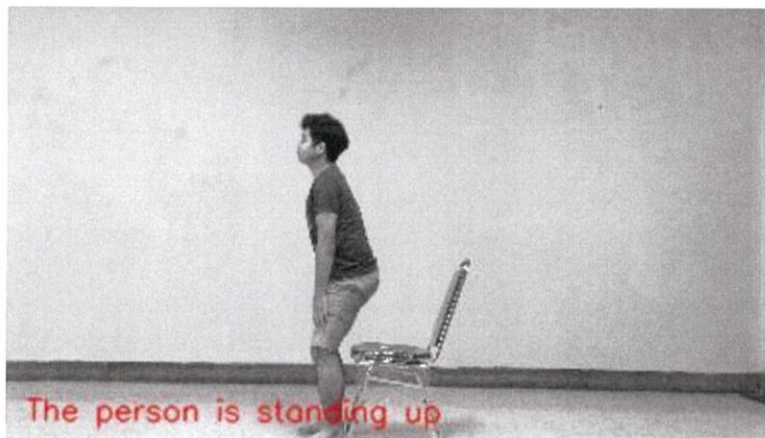
รูปที่ 5 The person is standing up



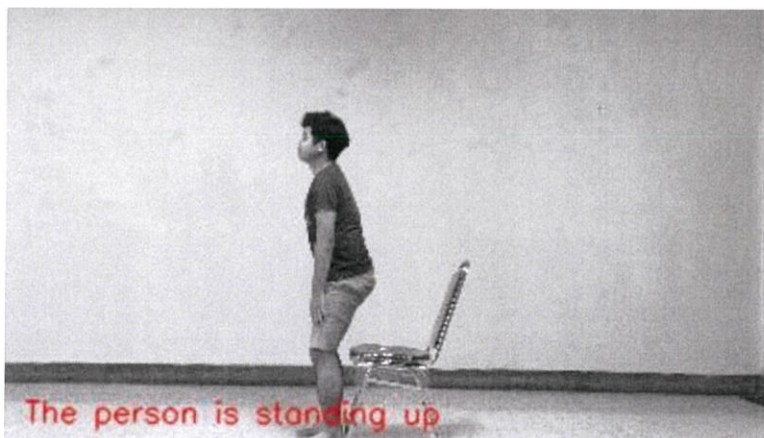
รูปที่ 6 The person is standing up



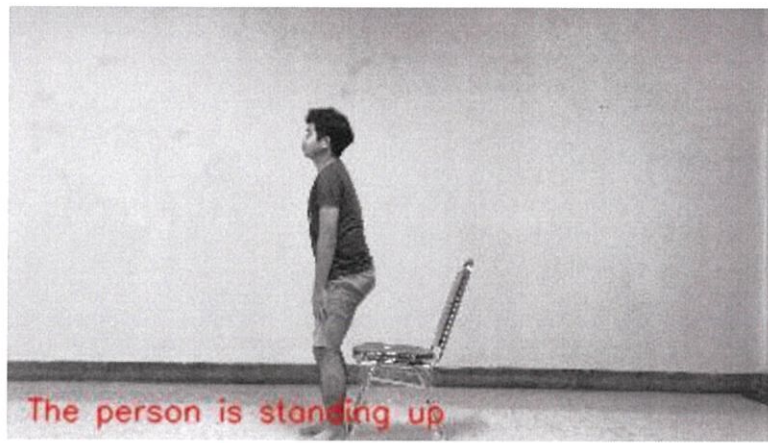
รูปที่ 7 The person is standing up



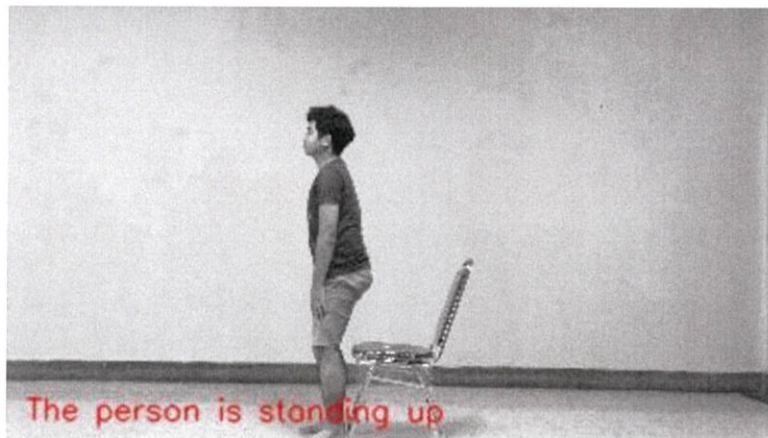
รูปที่ 8 he person is standing up



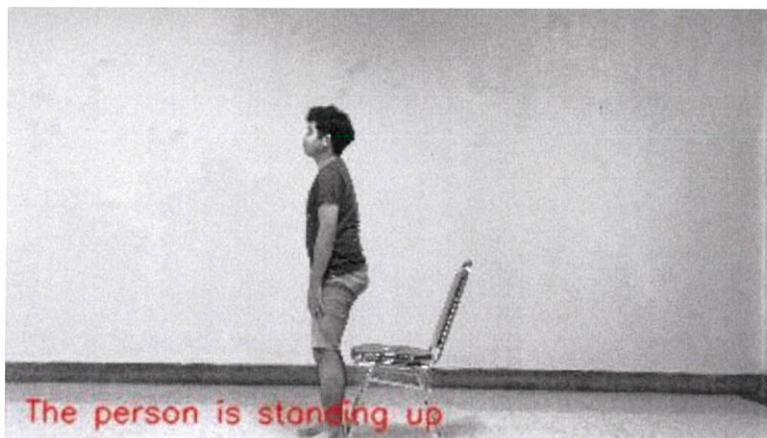
รูปที่ 9 The person is standing up



รูปที่ 10 The person is standing up

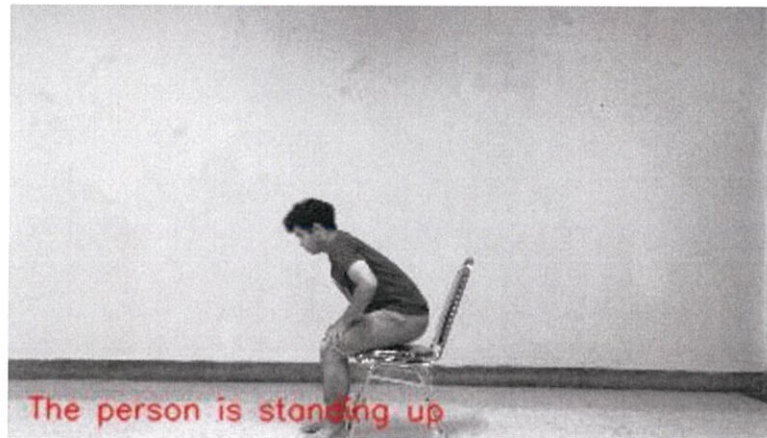


รูปที่ 11 The person is standing up

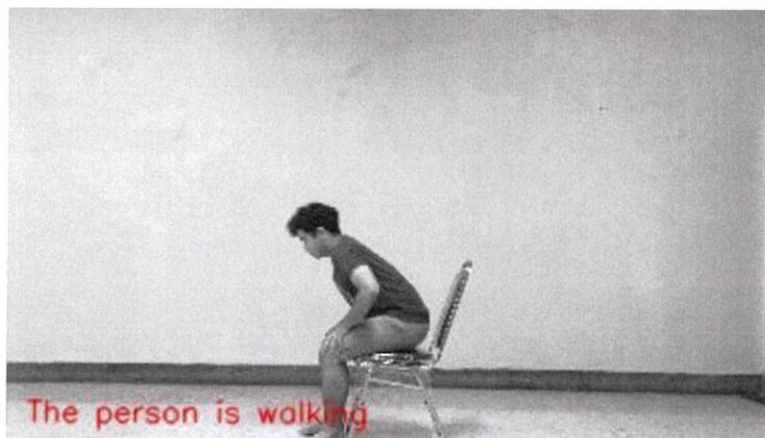


รูปที่ 12 The person is standing up

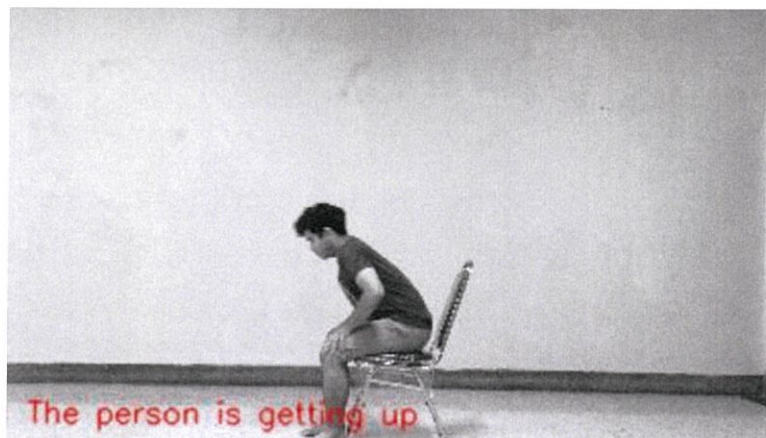
Realtime capture while sit up



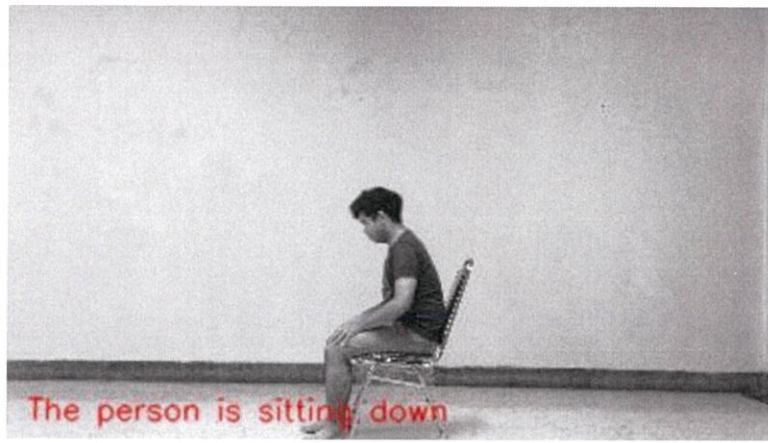
รูปที่ 13 The person is sitting up



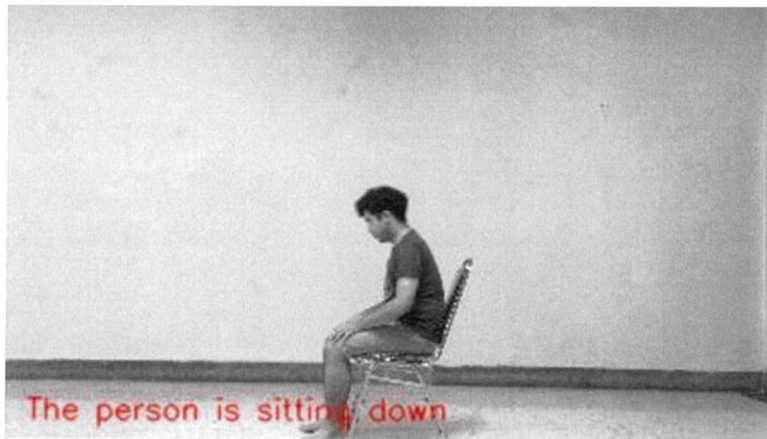
รูปที่ 14 The person is sitting up



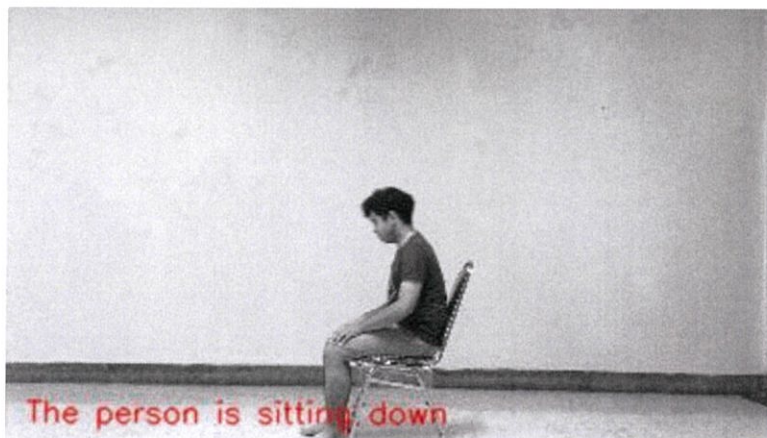
รูปที่ 15 The person is sitting up



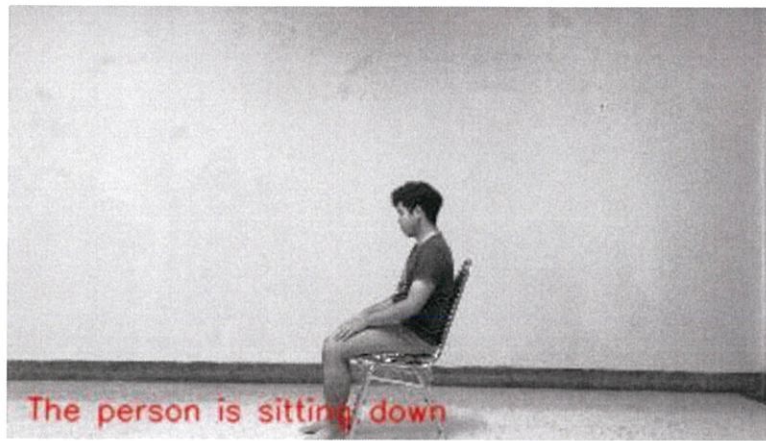
รูปที่ 16 The person is sitting up



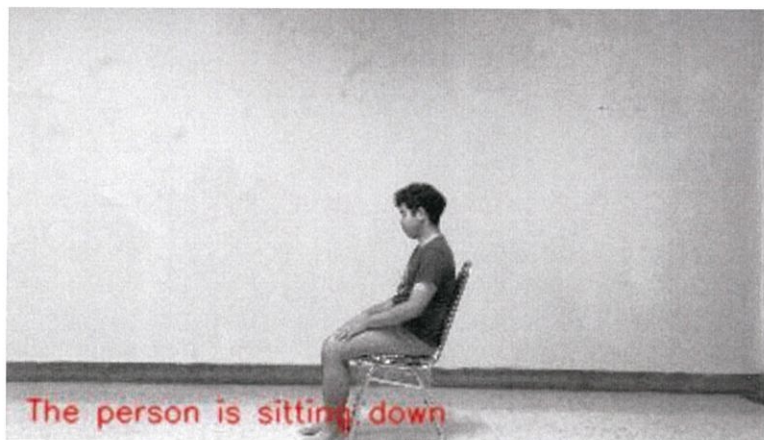
รูปที่ 17 The person is sitting up



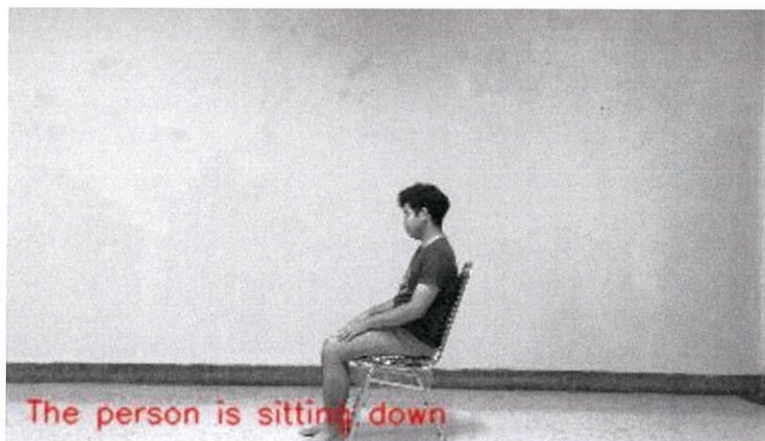
รูปที่ 18 The person is sitting up



รูปที่ 19 The person is sitting up

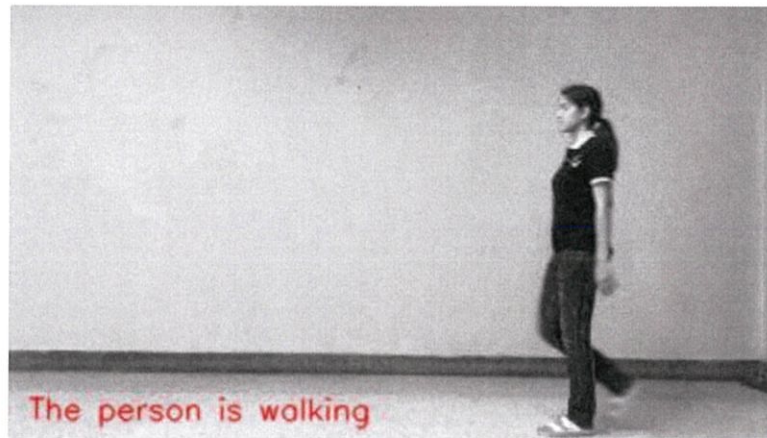


รูปที่ 20 The person is sitting up

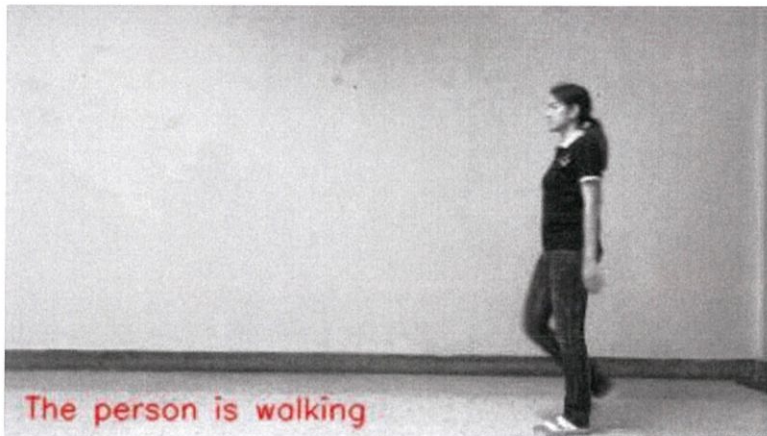


รูปที่ 21 The person is sitting up

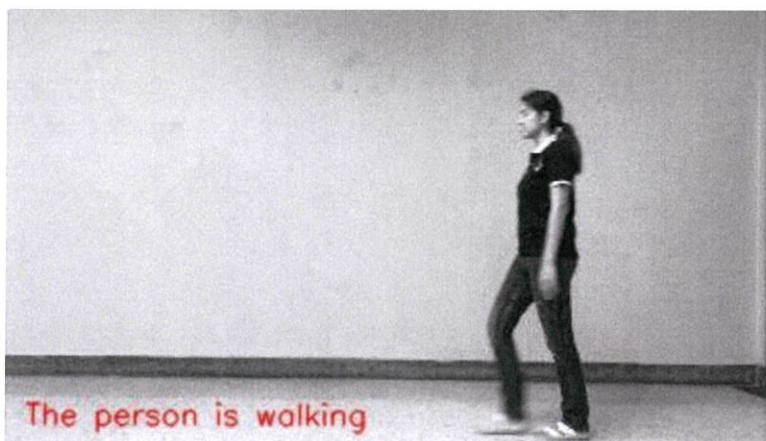
Realtime capture while walk



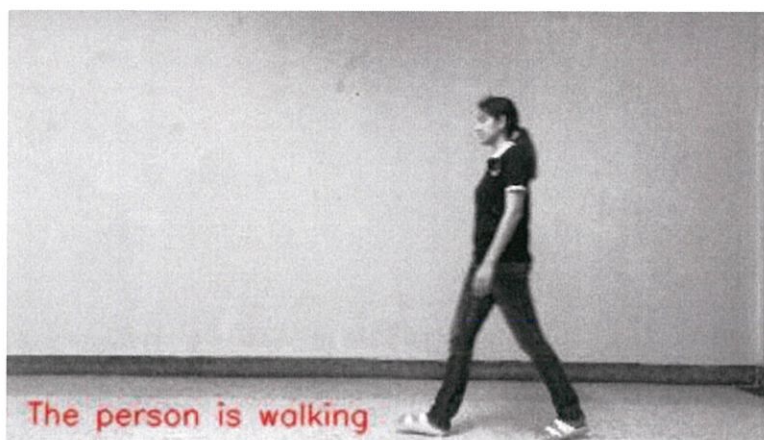
รูปที่ 22 The person is walking



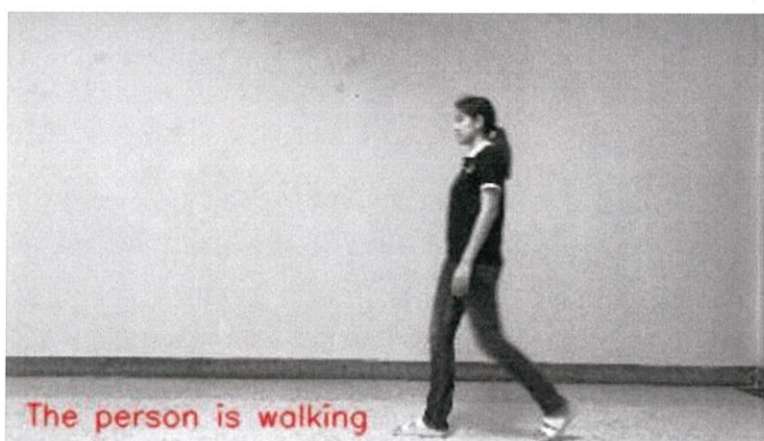
รูปที่ 23 The person is walking



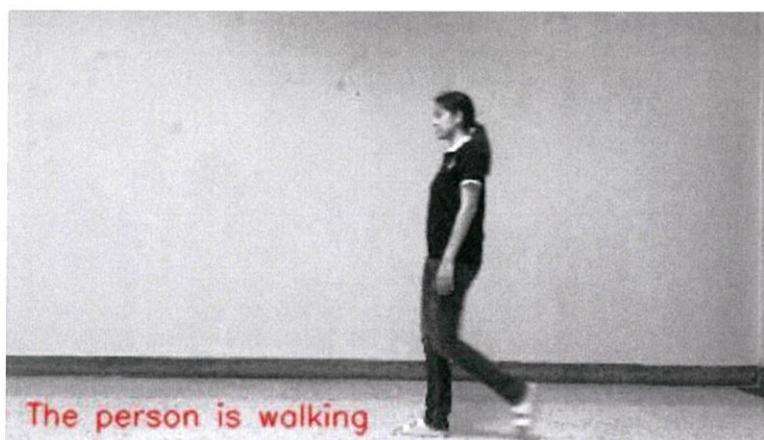
รูปที่ 24 The person is walking



รูปที่ 25 The person is walking

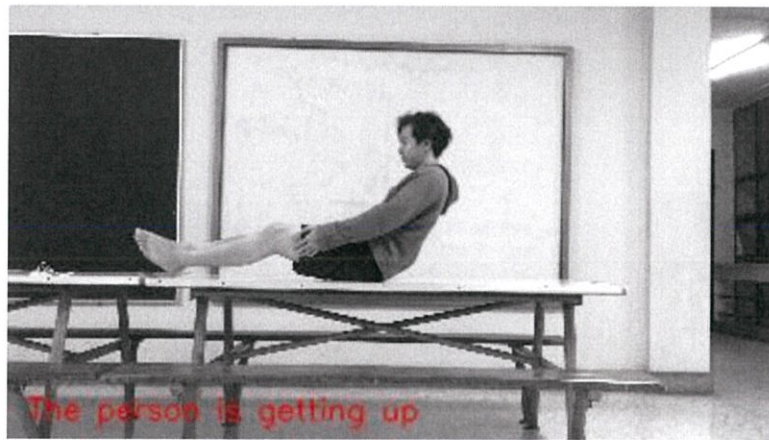


รูปที่ 26 The person is walking

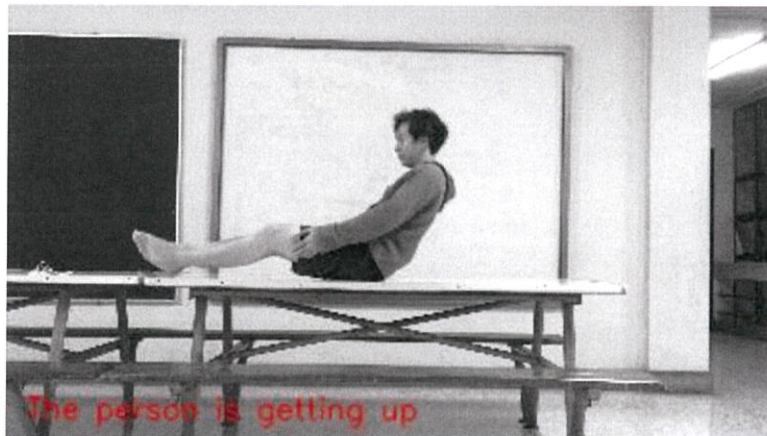


รูปที่ 27 The person is walking

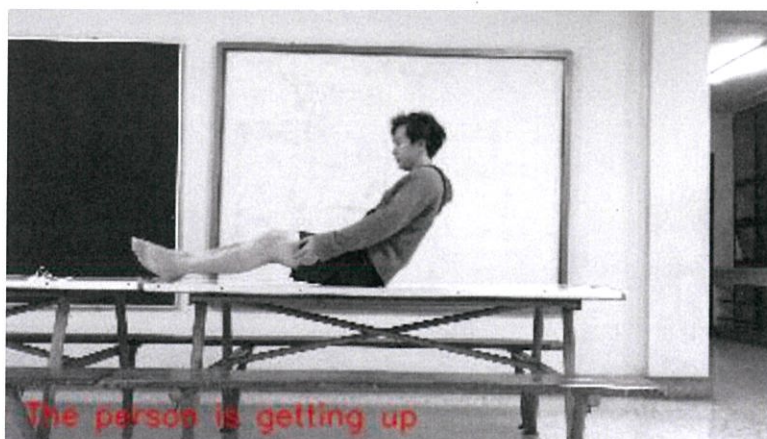
Realtime capture while get up



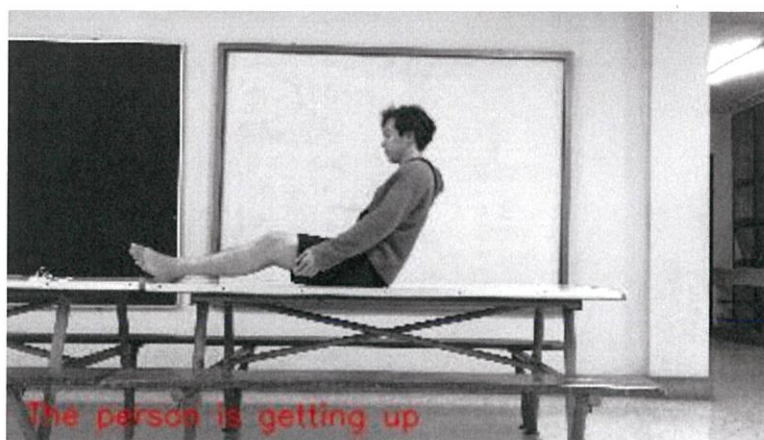
รูปที่ 28 The person is getting up



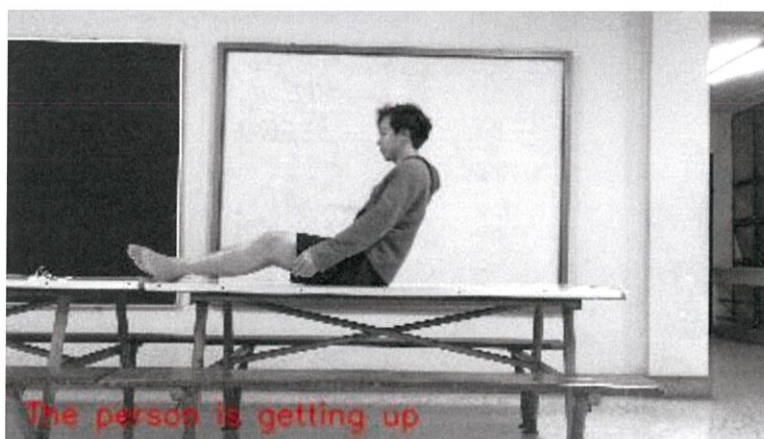
รูปที่ 29 The person is getting up



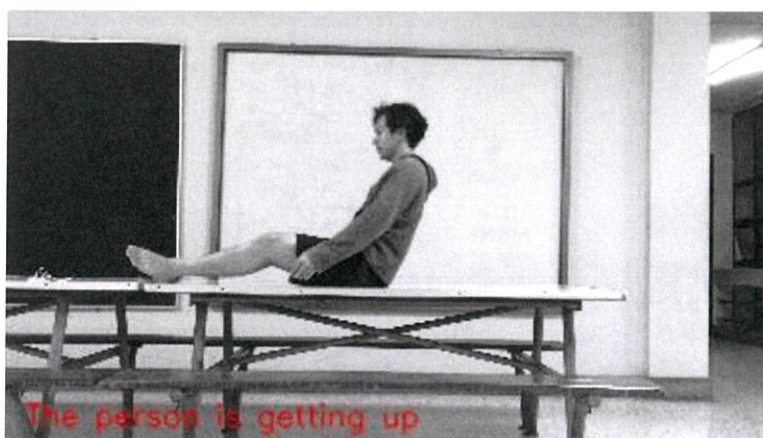
รูปที่ 30 The person is getting up



รูปที่ 31 The person is getting up

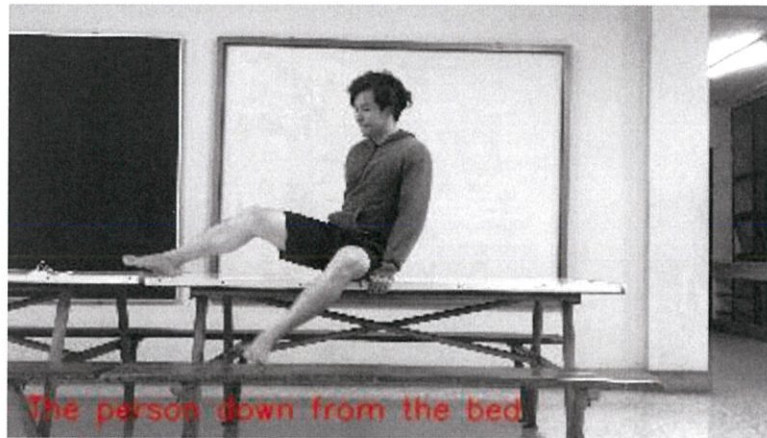


รูปที่ 32 The person is getting up

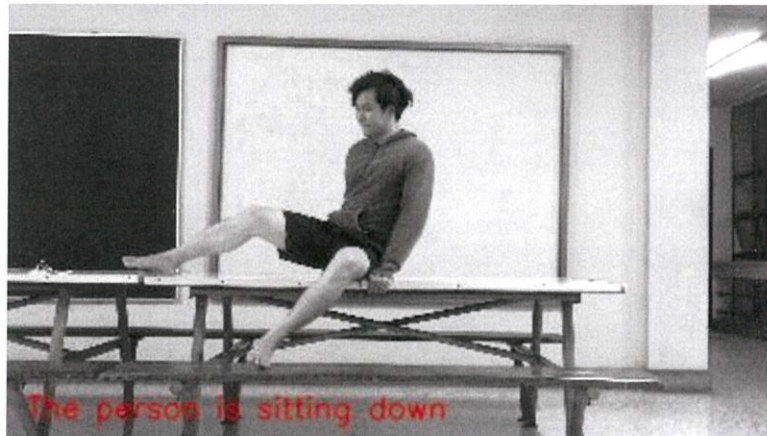


รูปที่ 33 The person is getting up

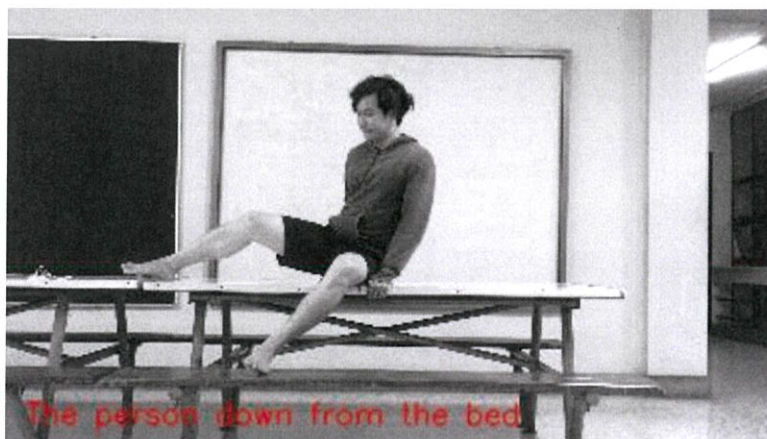
Realttime capture while get out from the bed



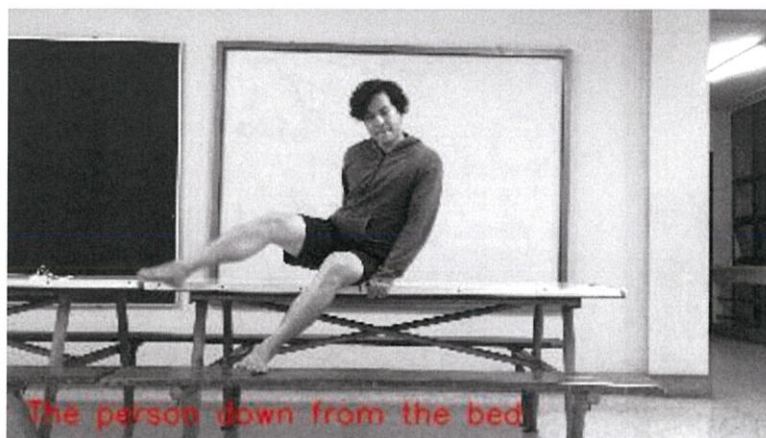
រូបទី 34 The person is getting off from the bed



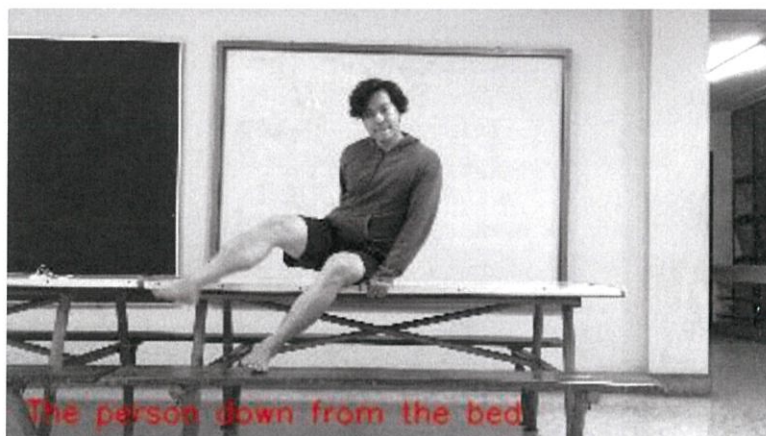
រូបទី 35 The person is getting off from the bed



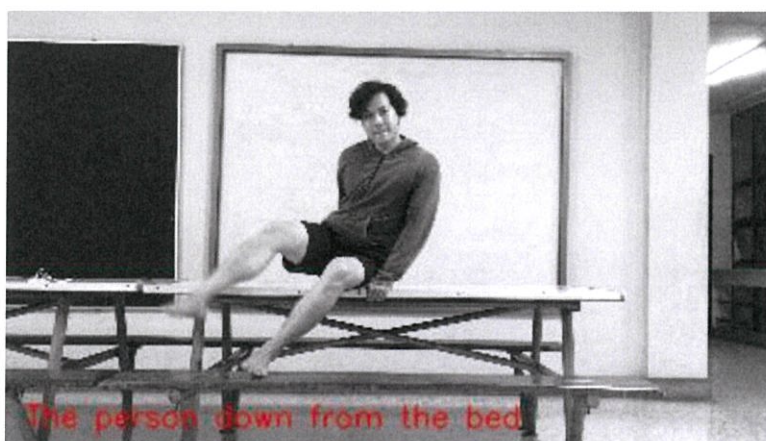
រូបទី 36 The person is getting off from the bed



รูปที่ 37 The person is getting off from the bed



รูปที่ 38 The person is getting off from the bed



รูปที่ 39 The person is getting off from the bed