

การศึกษาการจดจำวัตถุจากการประมวลผลภาพที่ได้จากการสัมผัส

STUDY OF OBJECT RECOGNITION FROM TACTILE IMAGE
PROCESSING

สมชาย เปาะทองคำ

ดุขฎฐินิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาวิศวกรรมศาสตรดุขฎฐิบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้า


คณะวิศวกรรมศาสตร์

มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

ปีการศึกษา 2564

ลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

การศึกษาการจดจำวัตถุจากการประมวลผลภาพที่ได้จากการสัมผัส

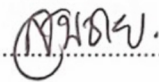


สมชาย เปาะทองคำ

ดุขฎฐินิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาวิศวกรรมศาสตรดุขฎฐินิพนธ์ สาขาวิชาวิศวกรรมไฟฟ้า
คณะวิศวกรรมศาสตร์
มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ปีการศึกษา 2564
ลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

วิทยานิพนธ์ฉบับนี้เป็นงานวิจัยที่เกิดจากการค้นคว้าและวิจัย ขณะที่ข้าพเจ้าศึกษาอยู่ในคณะ
วิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี ดังนั้นงานวิจัยในวิทยานิพนธ์ฉบับนี้ถือเป็น
ลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี และข้อความต่างๆ ในวิทยานิพนธ์ฉบับนี้ ข้าพเจ้า
ขอรับรองว่าไม่มีการคัดลอกหรือนำงานวิจัยของผู้อื่นมานำเสนอในชื่อของข้าพเจ้า

This thesis consists of research materials conducted at the Faculty of Engineering,
Rajamangala University of Technology Thanyaburi and hence the copyright owner. I
hereby certify that the thesis does not contain any forms of plagiarism.



(นายสมชาย เปาะทองคำ)

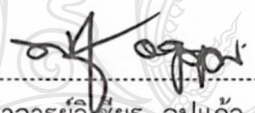


หัวข้อวิทยานิพนธ์ การศึกษาการจดจำวัตถุจากการประมวลผลภาพที่ได้จากการสัมผัส
Study of Object Recognition from Tactile Image Processing
ชื่อ - นามสกุล นายสมชาย เปาะทองคำ
สาขาวิชา วิศวกรรมไฟฟ้า
อาจารย์ที่ปรึกษา ผู้ช่วยศาสตราจารย์จักรี ศรีนนท์ฉัตร, Ph.D.
ปีการศึกษา 2564


คณะกรรมการสอบวิทยานิพนธ์


..... ประธานกรรมการ
(รองศาสตราจารย์สมเกียรติ อุดมธรรษากุล, Ph.D.)


..... กรรมการ
(ผู้ช่วยศาสตราจารย์อำนาจ เรืองวารี, Dr.-Ing.)


..... กรรมการ
(อาจารย์วิเชียร อุบแก้ว, Ph.D.)


..... กรรมการ
(ผู้ช่วยศาสตราจารย์นรเสฏฐ์ วิชัยพานิชย์, วศ.ด.)


..... กรรมการ
(ผู้ช่วยศาสตราจารย์จักรี ศรีนนท์ฉัตร, Ph.D.)

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี อนุมัติวิทยานิพนธ์ฉบับนี้เป็น
ส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาตรีบัณฑิต


..... คณบดีคณะวิศวกรรมศาสตร์
(รองศาสตราจารย์สรพงษ์ ภาวสุปรีดิ์, Ph.D.)
วันที่ 23 เดือน มีนาคม พ.ศ. 2565

หัวข้อคุณิพนธ์	การศึกษาการจดจำวัตถุจากการประมวลผลภาพที่ได้จากการสัมผัส
ชื่อ - นามสกุล	นายสมชาย เปาะทองคำ
สาขาวิชา	วิศวกรรมไฟฟ้า
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์จักรี ศรีนนท์ฉัตร, Ph.D.
ปีการศึกษา	2564

บทคัดย่อ

การพัฒนาเทคโนโลยีทางหุ่นยนต์เสมือนมนุษย์มุ่งเน้นให้หุ่นยนต์มีความสามารถในการสัมผัสรับรู้และจดจำวัตถุเลียนแบบการทำงานของมนุษย์ ซึ่งเป็นแขนงหนึ่งของปัญญาประดิษฐ์ (Artificial Intelligence) ทั้งนี้มีการนำเอาเทคโนโลยีคอมพิวเตอร์ที่มีประสิทธิภาพสูงมาใช้งาน แต่ประสิทธิภาพในการจดจำวัตถุจากการสัมผัสยังต่ำเนื่องจากขาดเซ็นเซอร์ที่เหมาะสม ดังนั้นในงานวิจัยนี้จึงได้ทำการศึกษาระบบการเรียนรู้และจดจำวัตถุจากการสัมผัสของหุ่นยนต์ โดยได้พัฒนาระบบประสาทสัมผัสทางกายเทียมขึ้น โดยการประยุกต์ใช้ Tactile Sensor ทำหน้าที่เป็นผิวหนังอิเล็กทรอนิกส์ (Electronic skin)

ในงานวิจัยนี้ได้พัฒนา Tactile Sensor แบบ Piezoresistive มีขนาด 16 x 16 Pixel มีขนาดกายภาพ 56.0 x 56.0 ตารางมิลลิเมตร และพัฒนาระบบอ่านค่าเซ็นเซอร์เพื่อแปลงค่าแรงกดที่ได้เป็นข้อมูลภาพ (Tactile Image) เพื่อนำมาใช้ในการวิเคราะห์การจดจำด้วยวิธีการประมวลผลภาพ การวิเคราะห์การจดจำพัฒนาบน Algorithm 2 วิธีคือ BoF และ DCNN โดยวิธี BoF พบว่าวิธี Polar Fourier Descriptor ร่วมกับ SVM ให้ผลการจดจำสูงที่สุด ส่วนการใช้วิธี DCNN ได้ทำการทดสอบกับ 19 Model พบว่า InceptionResNetV2 ให้ผลการจดจำสูงที่สุด

นอกจากนี้งานวิจัยนี้ได้พัฒนา Glove Tactile Sensor ซึ่งประกอบด้วย Tactile Sensor Array 15 จุดตามจุดสัมผัสของฝ่ามือ และทดสอบร่วมกับวิธี BoF และ DCNN ในส่วนของวิธี BoF พบว่าการใช้ SVM เป็น Classifier ร่วมกับ Moment Analysis Descriptor ให้ความถูกต้องสูงที่สุด โดยให้ความถูกต้องสูงกว่าร้อยละ 80.15 ในการจับวัตถุ 5 ครั้ง และวิธี DCNN พบว่า InceptionNetV3 ให้ความถูกต้องสูงสุดอยู่ที่ร้อยละ 98.28 ในการจับวัตถุเพียงครั้งเดียว

คำสำคัญ: การจดจำวัตถุจากการสัมผัส เซ็นเซอร์สัมผัส มือหุ่นยนต์ วิธีถ่วงค่า วิธีถ่วงพีตเจอร์ โครงข่ายประสาทเทียมแบบคอนโวลูชัน

Dissertation Title Study of Object Recognition from Tactile Image Processing
Name – Surname Mr. Somchai Pohtongkam
Program Electrical Engineering
Dissertation Advisor Assistant Professor Jakkree Srinonchat, Ph.D.
Academic Year 2021

ABSTRACT

With the development of human-like robotic technologies, robots have been advanced for the ability of sensing and recognizing objects, which is one branch of artificial intelligence, in order to mimic human reactions. Although high-performance computer technologies are used, the ability of object recognition by touch is still low due to the lack of proper sensors. Therefore, in this research, the system of object learning and recognition through robot touches was studied by developing an artificial sensory system acting as an electronic skin with tactile sensors.

This study built 16 × 16-pixel piezoresistive tactile sensors with the physical size of 56.0 × 56.0 mm² and developed a sensor reading system to convert obtained pressure values into tactile images for recognition analysis through image processing. Recognition analysis was based on two algorithm methods, namely BoF and DCNN. With BoF, Polar Fourier Descriptor with SVM provided the best recognition, and from the use of DCNN for testing with 19 models, InceptionResNetV2 provided the best recognition.

Glove tactile sensors comprising 15 points of tactile sensor array according to touchpoints of the palm were also developed and then tested together with BoF and DCNN methods. With BoF technique, using SVM as a classifier with Moment Analysis Descriptor provided the highest accuracy as it showed the accuracy of more than 80.15% from five grasping of an object. With DCNN approach, InceptionNetV3 provided the highest accuracy of 98.28% from only one capture of an object.

Keywords: tactile object recognition, tactile sensor, robot hand, Bag of Words, Bag of Features, CNN

กิตติกรรมประกาศ

ดุษฎีนิพนธ์นี้ ได้รับการสนับสนุนการวิจัยทุนพัฒนาบัณฑิตศึกษา จากสำนักงานการวิจัยแห่งชาติประจำปีงบประมาณ 2563 ผู้จัดทำขอขอบคุณมา ณ โอกาสนี้ และขอขอบคุณอาจารย์ที่ปรึกษา ผู้ช่วยศาสตราจารย์ ดร.จักรี ศรีนนท์ฉัตร ที่ได้ช่วยให้คำแนะนำและติดตามการทำดุษฎีนิพนธ์ครั้งนี้อย่างใกล้ชิดตลอดมา

ขอขอบคุณพ่อแม่ญาติพี่น้องและคณาจารย์ ที่ให้การสนับสนุนและประสิทธิ์ประสาทวิชาความรู้ให้ และขอขอบคุณสมาชิกร่วมห้องปฏิบัติการ Signal Processing Research (SPRL) ที่ช่วยให้คำแนะนำและเป็นกำลังใจที่ดีตลอดมา

สุดท้ายนี้ ผู้วิจัยหวังเป็นอย่างยิ่งว่างานค้นคว้าฉบับนี้จะเป็นประโยชน์สำหรับผู้ที่เกี่ยวข้อง หากดุษฎีนิพนธ์ในครั้งนี้ขาดตกบกพร่องหรือไม่สมบูรณ์ประการใด ผู้วิจัยขอกราบขออภัยมา ณ โอกาสนี้ด้วย

สมชาย เปาะทองคำ



สารบัญ

	หน้า
บทคัดย่อภาษาไทย	(3)
บทคัดย่อภาษาอังกฤษ	(4)
กิตติกรรมประกาศ	(5)
สารบัญ	(6)
สารบัญตาราง	(8)
สารบัญรูป	(9)
คำอธิบายสัญลักษณ์และคำย่อ	(13)
บทที่ 1 บทนำ	16
1.1 ความเป็นมาและความสำคัญของปัญหา	16
1.2 วัตถุประสงค์การวิจัย	17
1.3 ขอบเขตของการวิจัย	17
1.4 ขั้นตอนการวิจัย	17
1.5 ประโยชน์ที่คาดว่าจะได้รับ	18
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง	19
2.1 กายวิภาคของมือมนุษย์	19
2.2 มือหุ่นยนต์	22
2.3 Tactile sensor	25
2.4 การจดจำวัตถุจากภาพสัมผัส	31
2.5 การสกัดคุณลักษณะจากภาพ	39
2.6 วิธีการจัดคลัสเตอร์และคลาส	47
2.7 งานวิจัยที่เกี่ยวข้อง	52
2.8 การพัฒนาในงานวิจัยนี้	57
บทที่ 3 วิธีดำเนินการวิจัย	59
3.1 การทดสอบสมมติฐานเบื้องต้น	60
3.2 Tactile Sensor และ การจดจำวัตถุจาก Tactile Image	64
3.3 การออกแบบเซ็นเซอร์แบบใหม่	69

สารบัญ (ต่อ)

	หน้า
3.4 การจดจำวัตถุจากการสัมผัสสำหรับมือหุ่นยนต์	76
3.5 Glove Tactile Sensor	105
บทที่ 4 การทดลองและผลการทดลอง	116
4.1 ผลการทดสอบสมมติฐานเบื้องต้น	116
4.2 ผลการทดสอบ Tactile sensor และ การจดจำวัตถุจาก Tactile image	120
4.3 เซ็นเซอร์แบบใหม่	125
4.4 การจดจำวัตถุจากการสัมผัสสำหรับมือหุ่นยนต์	129
4.5 การจดจำวัตถุจาก Glove Tactile Sensor	137
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ	147
5.1 สรุปผลการวิจัย	147
5.2 ข้อเสนอแนะ	148
บรรณานุกรม	149
ภาคผนวก	164
ภาคผนวก ก วารสารทางวิชาการและบทความการประชุมวิชาการ	165
ประวัติผู้เขียน	170

สารบัญตาราง

	หน้า
ตารางที่ 2.1 มือหุ่นยนต์ในวรรณกรรม	23
ตารางที่ 2.2 การตรวจสอบวัตถุของหุ่นยนต์	25
ตารางที่ 2.3 Tactile Sensor แบบต่างๆ	27
ตารางที่ 2.4 Tactile Sensor แบบ Piezoresistive ในงานวรรณกรรม	30
ตารางที่ 2.5 Tactile object recognition ในงานวรรณกรรม	32
ตารางที่ 3.1 ขนาดของวัตถุ	77
ตารางที่ 3.2 Training Parameter	103
ตารางที่ 3.3 ค่า Initial Learning Rate ของ DCNN แต่ละ Model	104
ตารางที่ 3.4 Glove Tactile Sensor	106
ตารางที่ 3.5 ขนาดของวัตถุ	108
ตารางที่ 3.6 Training Parameter for Glove Tactile Sensor	113
ตารางที่ 3.7 ค่า Initial Learning Rate for Glove Tactile Sensor	113
ตารางที่ 4.1 ผลการจดจำจากการแบ่งส่วนมือเป็น 15 ส่วน.....	118
ตารางที่ 4.2 ผลการจดจำจากการแบ่งส่วนมือเป็น 20 ส่วน.....	118
ตารางที่ 4.3 ผลการจดจำจากการแบ่งส่วนมือเป็น 26 ส่วน.....	119
ตารางที่ 4.4 ผลการจดจำวัตถุจากการใช้ Tactile Sensor ร่วมกับการทำ Statistic Descriptor	123
ตารางที่ 4.5 ผลการทดสอบการจดจำวัตถุของมือหุ่นยนต์ด้วยวิธี DCNN	134
ตารางที่ 4.6 ผลการจดจำของวัตถุแต่ละ Class	135
ตารางที่ 4.7 ผลการจดจำจากการทดสอบ Glove Tactile Sensor ด้วย DCNN	142
ตารางที่ 4.8 ผลการจดจำของวัตถุแต่ละ Class สำหรับ Glove Tactile Sensor	143

สารบัญรูป

	หน้า
รูปที่ 2.1 กายวิภาคของมือมนุษย์	19
รูปที่ 2.2 จุดหมุนในจุดต่างๆ ของมือมนุษย์แบบ 16 DoF	20
รูปที่ 2.3 โครงสร้างของระบบประสาทสำหรับความรู้สึกทางกายของมนุษย์	21
รูปที่ 2.4 ลักษณะการสัมผัสวัตถุเพื่อประเมินลักษณะของวัตถุทั้ง 6 แบบ	22
รูปที่ 2.5 มือหุ่นยนต์แบบต่างๆ ที่ปรากฏในวรรณกรรม	24
รูปที่ 2.6 การตรวจสอบวัตถุของหุ่นยนต์ทั้ง 4 แบบ	26
รูปที่ 2.7 Tactile Sensor แบบต่างๆ	29
รูปที่ 2.8 มือหุ่นยนต์ และ Tactile Image	32
รูปที่ 2.9 การทำงานของการจัดกลุ่มภาพแบบ Machine Learning	34
รูปที่ 2.10 การทำงานของ BoF	35
รูปที่ 2.11 Histogram ของวัตถุ 3 ชนิดที่ได้จากการจัดคลัสเตอร์	36
รูปที่ 2.12 การทำงานของ CNN	37
รูปที่ 2.13 ผลของการทำคอนโวลูชันแบบ 2 มิติ	38
รูปที่ 2.14 การทำ Max Pooling ในชั้น Pooling Layer	38
รูปที่ 2.15 การทำชั้น Fully-Connected Layer และ Soft Max Function	39
รูปที่ 2.16 Tactile Image	40
รูปที่ 2.17 ภาพวัตถุที่มีขนาดตำแหน่งและการหมุนแตกต่างกัน	42
รูปที่ 2.18 การหาตำแหน่งของ Keypoint	44
รูปที่ 2.19 Histogram ของเวกเตอร์ของ SIFT	45
รูปที่ 2.20 ลำดับขั้นตอนในการทำการสกัดเอกลักษณ์โดยวิธี Polar Fourier	46
รูปที่ 2.21 K-Nearest Neighbors ที่ $K=3$	48
รูปที่ 2.22 โครงข่ายเพอร์เซปตรอน	50
รูปที่ 2.23 โครงข่ายเพอร์เซปตรอนแบบหลายชั้น	51
รูปที่ 2.24 Local Shape Tactile Image และ Global Shape Tactile Image	54
รูปที่ 3.1 ระบบการทำงานของเครื่องจดจำวัตถุจากการสัมผัส	59
รูปที่ 3.2 การจับวัตถุและจุดสัมผัสของฝ่ามือมนุษย์	61
รูปที่ 3.3 การดำเนินงานการจดจำวัตถุจากการสัมผัสของฝ่ามือมนุษย์ร่วมกับ ANN	61

สารบัญรูป (ต่อ)

	หน้า
รูปที่ 3.4 วัตถุและ Tactile Image	62
รูปที่ 3.5 การแบ่งส่วนของฝ่ามือเป็น 15 ส่วน 20 ส่วนและ 26 ส่วน	63
รูปที่ 3.6 การสร้าง Data Vector ตามการแบ่งตำแหน่งฝ่ามือ	64
รูปที่ 3.7 โครงข่ายประสาทเทียมแบบหลายชั้นที่ใช้ในงานวิจัยนี้	64
รูปที่ 3.8 การทำงานในส่วนเก็บภาพจากการสัมผัส	65
รูปที่ 3.9 การดำเนินงานการจดจำวัตถุจากการสัมผัสโดยใช้ Tactile Sensor ร่วมกับ ANN	65
รูปที่ 3.10 Tactile Sensor และกราฟความต้านทาน	66
รูปที่ 3.11 บอร์ดไมโครคอนโทรลเลอร์และวงจรรอ่านค่า	67
รูปที่ 3.12 วัตถุและภาพ Tactile Image	68
รูปที่ 3.13 โครงข่ายประสาทเทียมแบบหลายชั้นที่ใช้ในงานวิจัยนี้	69
รูปที่ 3.14 ขั้ว Electrode ของ Tactile Sensor แบบ Circle Electrode และ Square Electrode	70
รูปที่ 3.15 พื้นที่สัมผัสบางส่วนของ Circle Electrode Sensor	71
รูปที่ 3.16 พื้นที่สัมผัสของเซ็นเซอร์แบบ Circle Electrode และแบบ Square	72
รูปที่ 3.17 เปรียบเทียบสัดส่วน Active Area ต่อ % การสัมผัสเซ็นเซอร์	72
รูปที่ 3.18 เซ็นเซอร์ที่ใช้ในการวิจัยนี้	73
รูปที่ 3.19 การวัด Sensitivity ของเซ็นเซอร์	74
รูปที่ 3.20 ไดอะแกรมชุดเครื่องมือที่ใช้ในการทดสอบ	75
รูปที่ 3.21 GUI ของโปรแกรมสำหรับอ่านค่าและ Tactile Sensor Array	75
รูปที่ 3.22 ชั้นของเซ็นเซอร์ และ Tactile Sensor ที่ประกอบเสร็จ	76
รูปที่ 3.23 มือหุ่นยนต์เมคคาทรอนิกส์ที่ติดตั้ง Tactile Sensor และมือหุ่นยนต์ขณะจับวัตถุ	76
รูปที่ 3.24 วัตถุที่นำมาใช้ในการทดลอง	78
รูปที่ 3.25 ตัวอย่างภาพ Tactile Images	78
รูปที่ 3.26 การทำ Bicubic Interpolation	79
รูปที่ 3.27 ภาพ Tactile Image ที่ได้จากการทำ Bicubic Interpolation	80
รูปที่ 3.28 Work Flow สำหรับการจดจำวัตถุจากการสัมผัสด้วยวิธี BoF	81
รูปที่ 3.29 การแบ่ง Sub Image สำหรับทำ Descriptor	82

สารบัญรูป (ต่อ)

	หน้า
รูปที่ 3.30 ภาพ Sub image และ ค่า Feature ที่ได้จากการหาค่า Moment ทั้ง 7 ค่า	82
รูปที่ 3.31 การสร้าง Feature โดยใช้วิธี SIFT	84
รูปที่ 3.32 การสร้าง Feature โดยใช้วิธี PF	85
รูปที่ 3.33 ผลของการจัดกลุ่มด้วยวิธี K-Mean Clustering เมื่อกำหนดขนาด Cluster=10	86
รูปที่ 3.34 การสร้าง Histogram เมื่อทำการแบ่ง Cluster เป็น 10 Cluster, 20 Cluster และ 30 Cluster	87
รูปที่ 3.35 การใช้วิธี KNN ในการทำ Classification ตัวอย่างในปริภูมิ 3 มิติ	88
รูปที่ 3.36 การใช้วิธี ANN ในการทำ Classification ตัวอย่างสำหรับ BoF	88
รูปที่ 3.37 การแปลงข้อมูลที่ไม่เป็นเชิงเส้นให้เป็นข้อมูลเชิงเส้นด้วยฟังก์ชันเคอร์เนล	89
รูปที่ 3.38 Work Flow สำหรับการจดจำวัตถุจากการสัมผัสด้วยวิธี DCNN	90
รูปที่ 3.39 DCNN Transfer Learning ที่ใช้ในงานวิจัยนี้	91
รูปที่ 3.40 AlexNet	91
รูปที่ 3.41 VGG16	92
รูปที่ 3.42 VGG19	92
รูปที่ 3.43 GoogLeNet	93
รูปที่ 3.44 ResNet18	94
รูปที่ 3.45 ResNet50	94
รูปที่ 3.46 ResNet101	95
รูปที่ 3.47 InceptionV3	95
รูปที่ 3.48 EfficienNetB0	96
รูปที่ 3.49 SqueezeNet	97
รูปที่ 3.50 InceptionResNetV2	98
รูปที่ 3.51 DenseNet201	98
รูปที่ 3.52 DarkNet19	99
รูปที่ 3.53 DarkNet53	99
รูปที่ 3.54 XceptionNet	100
รูปที่ 3.55 NASNetMobile	101

สารบัญรูป (ต่อ)

	หน้า
รูปที่ 3.56 ShuffleNet	102
รูปที่ 3.57 MobileNetV2	102
รูปที่ 3.58 การจดจำวัตถุจากการสำรวจวัตถุ	105
รูปที่ 3.59 การประกอบชุดถุงมือ Sensor Array	106
รูปที่ 3.60 วงจรอ่านค่าและการสร้างภาพ Full Hand Tactile Image	107
รูปที่ 3.61 วัตถุที่ใช้ในการทดลอง	109
รูปที่ 3.62 การจับวัตถุแบบกำเต็มมือเพื่อสำรวจวัตถุของมนุษย์และการจับวัตถุแบบกำเต็มมือ โดย Glove Tactile Sensor	110
รูปที่ 3.63 ตัวอย่างของภาพ Full Hand Tactile Image ที่ได้จากการจับวัตถุ	111
รูปที่ 3.64 Work Flow สำหรับการจดจำวัตถุจากการสัมผัสด้วยวิธี BoF สำหรับ Glove Tactile Sensor	112
รูปที่ 3.65 การแบ่ง Sub Image ของ Full Hand Tactile Image สำหรับทำ Descriptor	112
รูปที่ 3.66 Work Flow สำหรับการจดจำวัตถุจากการสัมผัสด้วยวิธี DCNN สำหรับ Glove Tactile Sensor	114
รูปที่ 3.67 Transfer Learning Method สำหรับการจดจำวัตถุจาก Glove Tactile Sensor	115
รูปที่ 4.1 วัตถุประแจ, ภาพ Tactile Image ของวัตถุประแจ และภาพ Tactile Image ที่นำมา พล็อตเป็น 3 มิติ	116
รูปที่ 4.2 Data Vector ของภาพวัตถุประแจที่แบ่งส่วนฝ่ามือเป็น 26 ส่วน	117
รูปที่ 4.3 ผลเปรียบเทียบความถูกต้องในการจดจำของวัตถุในแต่ละ Class เมื่อใช้การแบ่งส่วนฝ่า มือเป็น 15, 20 และ 26 ส่วน	119
รูปที่ 4.4 เปรียบเทียบการจดจำวัตถุของการแบ่งส่วนทั้ง 3 แบบ	120
รูปที่ 4.5 ภาพ Tactile Image ของวัตถุประแจ แก้วกาแฟ และประแจหกเหลี่ยม	121
รูปที่ 4.6 ค่าเฉลี่ยของภาพที่ได้จากวัตถุแต่ละชนิด	122
รูปที่ 4.7 ค่าเบี่ยงเบนมาตรฐานของภาพที่ได้จากวัตถุแต่ละชนิด	122
รูปที่ 4.8 ผลการทดสอบเมื่อ Training ระบบด้วยค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐาน	124
รูปที่ 4.9 ตำแหน่งวัดความต้านทานของเซ็นเซอร์	125

สารบัญรูป (ต่อ)

	หน้า
รูปที่ 4.10 ค่าความต้านทานของเซ็นเซอร์ต่อแรงที่กระทำ	126
รูปที่ 4.11 คุณสมบัติของเซ็นเซอร์	127
รูปที่ 4.12 ตัวอย่างภาพ Tactile Images	128
รูปที่ 4.13 ภาพที่ได้จากการทำ Bicubic Interpolation	129
รูปที่ 4.14 ผลการจดจำเมื่อใช้ KNN	130
รูปที่ 4.15 ผลการจดจำเมื่อใช้ ANN	130
รูปที่ 4.16 ผลการจดจำเมื่อใช้ SVM	131
รูปที่ 4.17 ผลการจดจำเมื่อทำการปรับปรุงความละเอียดของภาพ	132
รูปที่ 4.18 ผลการจดจำเมื่อทำการปรับปรุงความละเอียดของภาพ 16 × 16 Pixels และ 512 × 512 Pixels	132
รูปที่ 4.19 Confusion Matrix แสดงผลที่ได้จากการทดสอบด้วย InceptionResNetV2, XceptionNet, และ InceptionNetV3	133
รูปที่ 4.20 ผลการจดจำที่ทำการทดสอบด้วย DCNN ของภาพปรับปรุงขนาด 512 × 512 Pixels	136
รูปที่ 4.21 ผลการจดจำเมื่อทำการสำรวจวัตถุก่อนทำนายผล	137
รูปที่ 4.22 Confusion Matrix แสดงผลการจดจำเมื่อทำการสำรวจวัตถุก่อนทำนายผล	137
รูปที่ 4.23 ภาพ Tactile Image ที่ได้จากการจับวัตถุด้วยเซ็นเซอร์ถ้งมือ	138
รูปที่ 4.24 ผลการทดสอบด้วยวิธี BoF โดยใช้วิธี KNN เป็น Classifier	139
รูปที่ 4.25 ผลการทดสอบด้วยวิธี BoF โดยใช้วิธี ANN เป็น Classifier	140
รูปที่ 4.26 ผลการทดสอบด้วยวิธี BoF โดยใช้วิธี SVM เป็น Classifier	141
รูปที่ 4.27 ผลการจดจำวัตถุด้วยวิธี BoF เมื่อจับวัตถุ 1 ครั้งด้วยวิธี SVM ร่วมกับ SIFT Descriptor	142
รูปที่ 4.28 ผลการจดจำวัตถุด้วย Glove Tactile Sensor ที่ได้รับการทดสอบด้วย DCNN	144
รูปที่ 4.29 Confusion Matrix แสดงผลการจดจำวัตถุด้วยวิธี DCNN AlexNet และ InceptionNetV3	145
รูปที่ 4.30 ผลการจดจำเมื่อทำการสำรวจวัตถุก่อนทำนายผล	145
รูปที่ 4.31 Confusion Matrix แสดงผลการจดจำเมื่อทำการสำรวจวัตถุ 10 ครั้งก่อนทำนายผล..	146

คำอธิบายสัญลักษณ์และคำย่อ

BoF	Bag of Feature
BoW	Bag of Word
CKCS	Concatenation Kernel Sparse Coding
CNN	Convolution Neural Network
DCNN	Deep Convolution Neural Network
DoF	Degree of Freedom
JKCS	Joint Kernel Sparse Coding
KNN	K-Nearest Neighbors
LDA	Liner Discriminant Analysis
MA	Moment Analysis
MPL	Multi-Layer Perceptron
MKL-SVM	Multiple Kernel Learning Support Vector Machine
PCA	Principal Component Analysis
PF	Polar Furrier
SIFT	Scale-Invariant Feature Transform
SKCS	Separate Kernel Sparse Coding
SOM	Self-Organizing Map
SVM	Support Vector Machine

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

จากการพัฒนาเทคโนโลยีหุ่นยนต์ที่ได้ประยุกต์เอาเทคโนโลยีคอมพิวเตอร์มาใช้งานร่วมกัน นั้น เนื่องจากในปัจจุบันคอมพิวเตอร์มีประสิทธิภาพที่สูงขึ้นมาก จึงตอบสนองต่อแนวคิดที่จะให้ระบบ หุ่นยนต์และคอมพิวเตอร์มาทำงานร่วมกัน และเพื่อให้หุ่นยนต์มีความสามารถในการเรียนรู้และจดจำ เลียนแบบมนุษย์ได้ อันเป็นส่วนหนึ่งของศาสตร์ทางปัญญาประดิษฐ์ (Artificial Intelligent : AI) การ พัฒนาระบบประสาทสัมผัสเทียมสำหรับหุ่นยนต์จึงเป็นงานวิจัยที่ถูกให้ความสนใจเป็นอย่างยิ่ง อย่างไรก็ตามงานวิจัยส่วนใหญ่จะเป็นการพัฒนาการมองเห็นของเครื่องจักร (Machine Vision) หรือ การมองเห็นของคอมพิวเตอร์ (Computer Vision) และการประมวลผลภาษาธรรมชาติ (Natural language processing : NLP) แต่การประมวลผลจากประสาทสัมผัสอื่นยังมีค่อนข้างน้อย เช่นการ จดจำวัตถุจากการสัมผัสของหุ่นยนต์ เป็นต้น จึงได้มีการพัฒนางานวิจัยที่ทำการศึกษาระบบการเรียนรู้ และจดจำวัตถุโดยใช้ระบบประสาทสัมผัสทางกายเพิ่มขึ้น โดยการทำงานจะใช้เซ็นเซอร์วัดแรงกดเป็น เสมือนปลายประสาทของตัวรับ (Receptors) ชนิด Ruffini Ending และ Pacinian ซึ่งเป็นปลาย ประสาทรับแรงสัมผัสและแรงกดของมนุษย์ จากนั้นก็ส่งข้อมูลเข้าคอมพิวเตอร์ทำการวิเคราะห์ ซึ่ง เปรียบเสมือนสมองของมนุษย์ที่ตีความจากข้อมูลที่ส่งมาจากปลายประสาท อย่างไรก็ตามประสิทธิภาพ ในการจดจำวัตถุจากการสัมผัสในปัจจุบันยังไม่ดีนักเนื่องจากขาดเซ็นเซอร์ (Sensor) และวิธีการ วิเคราะห์ (Algorithm) ที่เหมาะสม ดังนั้นในงานวิจัยนี้จึงได้พัฒนาระบบเซ็นเซอร์ที่ติดตั้งไว้ที่มือหุ่นยนต์ โดยเป็นเซ็นเซอร์สัมผัสแบบอิเล็กทรอนิกส์ (Tactile Sensor) ที่มีลักษณะเป็นเซ็นเซอร์วัดแรงกดแบบ อาเรย์ (Tactile Sensor Array) ซึ่งจะถูกฝังไว้ที่มือของหุ่นยนต์ เมื่อมือของหุ่นยนต์สัมผัสวัตถุจะเกิดแรง ที่วัตถุกระทำต่อมือหุ่นยนต์เกิดเป็นรูปแบบแรงกด (Pressure Profile) ซึ่งมีลักษณะเป็นเมตริกซ์ (Matrix) ข้อมูลของแรงกดของวัตถุต่างๆ ซึ่งถูกนำเสนอในรูปแบบภาพเรียกว่าภาพสัมผัส (Tactile Image)

Tactile Sensor สำหรับวัดการสัมผัสประกอบด้วยอาร์เรย์ของเซ็นเซอร์วัดแรงกด (Pressure Sensor) อันเป็นชิ้นส่วนที่พบบ่อยในระบบหุ่นยนต์สมัยใหม่ ซึ่งคาดการณ์ว่าการออกแบบมือหุ่นยนต์แบบ ใหม่ในปัจจุบันจะมี Tactile Sensor ฝังเอาไว้ด้วยที่ปลายนิ้วและในส่วนพื้นผิวต่างๆ ของมือ ค่าข้อมูลที่ อ่านได้จากแต่ละพิกเซล (Pixel) เป็นค่าแรงกดที่มากกระทำต่อเซ็นเซอร์ ทำให้เซ็นเซอร์มีการยุบตัวลง ค่า ความลึกที่ได้จะขึ้นอยู่กับคุณสมบัติของวัสดุที่นำมาทำตัวปิดด้านบน เมื่อเรามองรูปแบบข้อมูลที่ได้มา จาก Tactile Sensor Array ตัวข้อมูลถูกวางเป็นเมตริกซ์ซึ่งมีลักษณะเดียวกับรูปภาพ เราจึงสามารถทำ

การวิเคราะห์โดยใช้วิธีการ Computer Vision ได้ การตีความข้อมูลจากภาพที่ได้จากการสัมผัสยากกว่า การตีความข้อมูลจากภาพถ่ายเนื่องจากข้อมูลที่ได้จะเป็นข้อมูลที่เกิดจากการสัมผัสเพียงบางส่วนของ วัตถุจึงขาดรายละเอียดของวัตถุ บางครั้งจึงต้องทำการจับวัตถุหลายครั้งก่อนการทำนายผลจึงจะเกิด ความถูกต้อง การพัฒนาในงานวิจัยนี้ต้องการพัฒนาเพื่อเพิ่มประสิทธิภาพในการจดจำวัตถุของมือ หุ่นยนต์ โดยการทดสอบการประมวลผลการจดจำด้วยวิธีต่างๆ และค้นหาวิธีการที่มีประสิทธิภาพสูงสุด เพื่อการนำไปใช้กับระบบหุ่นยนต์จริงต่อไป

1.2 วัตถุประสงค์การวิจัย

ทำการพัฒนาระบบการเรียนรู้และจดจำวัตถุโดยใช้ระบบประสาทสัมผัสเทียมขึ้น เมื่อหุ่นยนต์ สัมผัสวัตถุที่ทราบชนิดแล้วทำให้เกิดการเรียนรู้ว่าวัตถุนั้นคือสิ่งใด จากนั้นก็ให้สัมผัสวัตถุที่ต้องการจะ ทดสอบ โดยการพัฒนาดังกล่าวจะทำให้หุ่นยนต์สามารถจดจำวัตถุจากการสัมผัสวัตถุนั้นได้

1.3 ขอบเขตของการวิจัย

การวิจัยครั้งนี้เป็นการศึกษาและพัฒนาเทคนิคเพื่อเพิ่มประสิทธิภาพของระบบหุ่นยนต์ในการ จดจำวัตถุจากการสัมผัสเพื่อนำไปใช้ในหุ่นยนต์เสมือนมนุษย์

1.3.1 สร้างชุดทดลองประกอบด้วยถุงมือที่ติดตั้งระบบ Tactile Sensor Array ในหลาย ตำแหน่ง

1.3.2 ใช้เทคนิคการประมวลผลภาพควบคู่กับเทคนิคการเรียนรู้แบบถุงฟีดเจอร์หรือแบ็กออฟ ฟีดเจอร์ (Bag of Feature) เพื่อใช้ในการตรวจจำวัตถุที่ได้จากการสัมผัสของมือหุ่นยนต์เสมือนมนุษย์

1.3.3 สามารถจดจำวัตถุจากการสัมผัสวัตถุได้อย่างน้อย 20 ชนิด โดยแบ่งเป็นวัตถุที่มีรูปทรง มาตรฐาน 10 วัตถุ และวัตถุทั่วไป 10 วัตถุ

1.3.4 มีความถูกต้องของการจดจำได้อย่างน้อยร้อยละ 80

1.4 ขั้นตอนการวิจัย

1.4.1 ศึกษาทฤษฎีการจับวัตถุ Tactile sensor การประมวลผลภาพ เครื่องมือทางการ ประมวลผลภาพ การ Classify ภาพ และรวบรวมวรรณกรรมที่เกี่ยวข้อง

1.4.2 ทดสอบสมมติฐานเบื้องต้น

1.4.3 ทดสอบการทำงานของ Tactile sensor ออกแบบวงจรอ่านค่า ระบบการเก็บรวบรวม ข้อมูลภาพ และทดสอบการจดจำวัตถุจากภาพ

1.4.4 ออกแบบ Tactile sensor สำหรับมือหุ่นยนต์และทดสอบการจดจำวัตถุจากภาพ

1.4.5 ออกแบบ Tactile sensor ชนิดถุงมือแบบอาร์เรย์ในหลายตำแหน่งและทดสอบการจดจำวัตถุจากภาพ

1.4.6 วิเคราะห์ผลการทดสอบ

1.4.7 จัดทำรูปเล่มคู่มือฉบับพิมพ์

1.5 ประโยชน์ที่คาดว่าจะได้รับ

การพัฒนาในงานวิจัยนี้ต้องการพัฒนาการเพิ่มประสิทธิภาพในการจดจำวัตถุของมือหุ่นยนต์ให้มีความสามารถในการจดจำวัตถุจากการสัมผัสได้เช่นเดียวกับระบบประสาทสัมผัสของมนุษย์ เพื่อให้หุ่นยนต์เสมือนมนุษย์มีความสามารถในการเรียนรู้จดจำวัตถุได้โดยกระบวนการเรียนรู้เป็นระบบการเรียนรู้แบบแบบปัญญาประดิษฐ์เมื่อหุ่นยนต์มีความสามารถในการเรียนรู้ในด้านการจดจำวัตถุอย่างมีประสิทธิภาพก็สามารถทำหน้าที่ได้ใกล้เคียงมนุษย์มากขึ้นทำให้เราสามารถนำหุ่นยนต์มาทำงานต่างๆ แทนมนุษย์ได้หลากหลายขึ้น



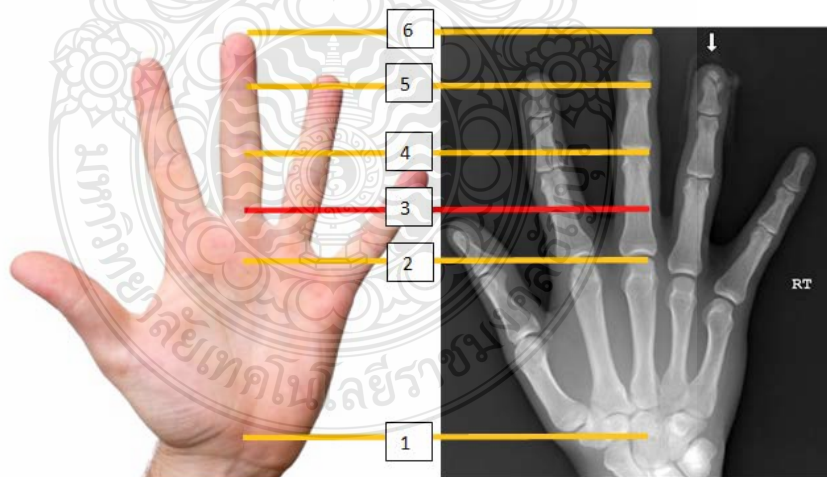
บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงหลักการและทฤษฎีที่ใช้ในการทำดัชนีนิพนธ์ โดยเนื้อหาจะประกอบด้วยกายวิภาคของมือมนุษย์ (Human Hand Anatomy) และงานวิจัยที่เกี่ยวข้อง

2.1 กายวิภาคของมือมนุษย์

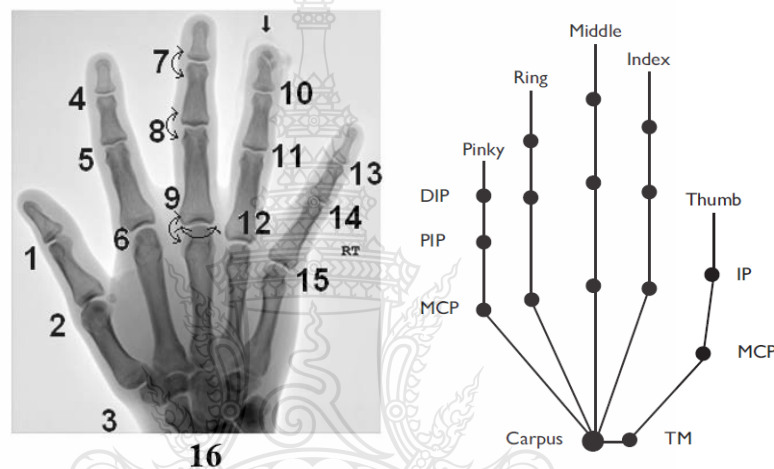
ในงานวิจัยนี้ศึกษาการจดจำวัตถุของหุ่นยนต์จากการสัมผัสซึ่งเป็นการพัฒนาระบบเครื่องจักรให้ทำงานเสมือนมนุษย์ดังนั้นจึงต้องศึกษารูปแบบการทำงานของมือมนุษย์เป็นลำดับแรก เพื่อเป็นต้นแบบในการพัฒนาของหุ่นยนต์ เนื่องจากปัจจุบันเทคโนโลยีหุ่นยนต์เสมือนมนุษย์หรือหุ่นยนต์ฮิวแมนนอยด์ (Humanoid Robot) ได้ถูกพัฒนาอย่างต่อเนื่องซึ่งการพัฒนาสามารถแยกได้เป็นสองส่วนคือทางด้านกายวิภาคของหุ่นยนต์ (Robot Anatomy) และทางด้านระบบประสาทของหุ่นยนต์ (Robot Nervous System) ทั้งสองอย่างล้วนแล้วแต่กระทำการเลียนแบบระบบของมนุษย์จริงๆ ในบทความนี้กล่าวถึงระบบประสาทสัมผัสทางกายของหุ่นยนต์เสมือนมนุษย์ เพื่อให้หุ่นยนต์เสมือนมนุษย์มีความสามารถในการจดจำวัตถุจากการสัมผัส พื้นฐานการทำงานพัฒนามาจากระบบประสาทสัมผัสทางกายของมนุษย์ [1]



รูปที่ 2.1 กายวิภาคของมือมนุษย์ [1]

จากรูปที่ 2.1 เป็นการแบ่งส่วนของมือ ซึ่งแบ่งเป็นสองส่วนคือฝ่ามือกับนิ้วมือ ส่วนของฝ่ามือจะมีกระดูกฝ่ามือยึดโยงเนื้อเยื่ออยู่โดยที่ส่วนของฝ่ามือจะยึดถึงกระดูกนิ้วมือส่วนต้นด้วยประมาณ

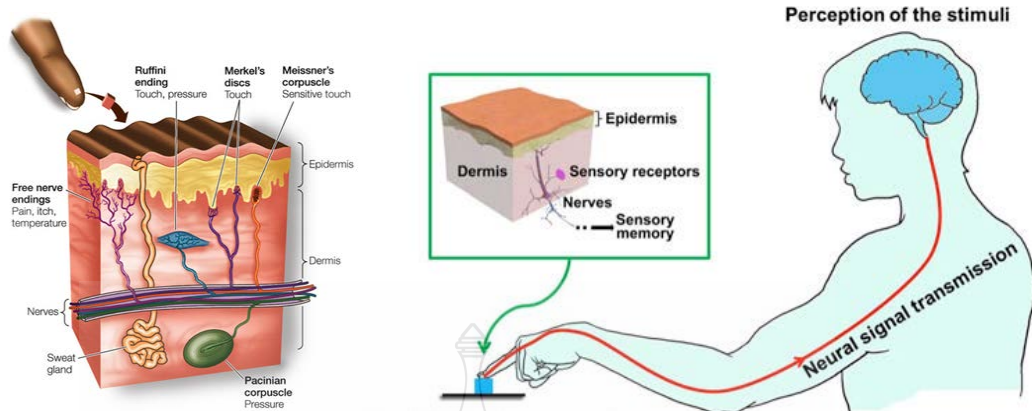
ครึ่งหนึ่งของความยาวกระดูก (จากจุดที่ 1 ถึงจุดที่ 3) ส่วนกระดูกนิ้วมือนั้น ในส่วนของนิ้วหัวแม่มือจะมีกระดูก 3 ชิ้น โดยจะฝังอยู่ในฝ่ามือ 1 ชิ้น และอีก 2 ชิ้นอยู่ที่นิ้ว ส่วนนิ้วชี้ นิ้วกลาง นิ้ว无名 และนิ้วก้อย จะมีกระดูก 4 ชิ้น โดยชิ้นแรกจะฝังอยู่ในฝ่ามือและชิ้นที่ 2 จะฝังอยู่ในฝ่ามือประมาณครึ่งหนึ่งและเป็น ส่วนของนิ้วอีกครึ่งหนึ่งและกระดูกนิ้วอีก 2 ชิ้นที่เหลือจะเป็นส่วนของปลายนิ้ว การแบ่งตำแหน่งจุดหมุน ของกระดูกนิ้วมือ (Degree of Freedom: DoF) และแนวการหมุนดังรูปที่ 2.2 แสดงความสามารถในการหมุนและท่ามุมของนิ้วมือ ตำแหน่ง 1, 2, 4, 5, 7, 8, 10, 11, 13, 14 จะเป็นจุดหมุนที่หมุนได้ในแนวก้มเงย และตำแหน่ง 3, 6, 9, 12, 15, 16 จะเป็นจุดหมุนที่หมุนได้ทั้งในแนวก้มเงยและแนวกวาด ซึ่งเราสามารถแบ่งออกได้เป็น 16 DoF



รูปที่ 2.2 จุดหมุนในจุดต่างๆ ของมือมนุษย์แบบ 16 DoF [2]

2.1.1 โครงสร้างของผิวหนังและระบบประสาทสัมผัสทางกาย

โครงสร้างของผิวหนังและระบบประสาทสัมผัสทางกายของมนุษย์มีโครงสร้างดังรูปที่ 2.3 [3] โครงสร้างระบบประกอบด้วยส่วนของปลายประสาทที่ทำหน้าที่แตกต่างกัน 5 แบบดังนี้ 1) Free Nerve Ending เป็นส่วนของการรับรู้ความเจ็บปวดและอุณหภูมิ, 2) Merkel's Discs เป็นส่วนของการรับรู้การสัมผัส, 3) Meissner's Corpuscle เป็นส่วนของการรับรู้การสัมผัสแบบแผ่วเบา, 4) Ruffini Ending และ 5) Pacinian Corpuscle ที่เป็นส่วนของการรับรู้แรงกด ซึ่งการทำงานของระบบร่างกายมนุษย์นี้จะเริ่มจากระบบประสาทสัมผัสได้สัมผัสกับตัวกระตุ้นเช่นอุณหภูมิภายนอก, หรือสัมผัสโดนวัตถุใดๆ แล้วทำให้เกิดแรงกดขึ้นเป็นต้น ระบบประสาทจะเปลี่ยนข้อมูลที่ได้เป็นสัญญาณไฟฟ้าเรียกว่าสื่อกระแสประสาท ส่งผ่านเส้นประสาท (Nerves) เข้าสู่สมองเพื่อตีความว่าวัตถุที่สัมผัสคืออะไร ในทำนองเดียวกันนี้ การจะพัฒนาหุ่นยนต์เสมือนมนุษย์ให้มีความสามารถในการรับรู้ทางกายก็อาศัยกลไกเดียวกันนี้เป็นพื้นฐานในการออกแบบ



รูปที่ 2.3 โครงสร้างของระบบประสาทรับความรู้สึกทางกายของมนุษย์ [3]

จากรูปที่ 2.3 เมื่อมีการจับวัตถุข้อมูลจากระบบประสาทสัมผัส (Pressure) และข้อมูลจากมุมของนิ้วมือ (Geometry) จะถูกส่งไปยังสมองด้วยสื่อกระแสประสาท ทำให้สมองมนุษย์เกิดภาพในสมองเพื่อวิเคราะห์ว่าวัตถุที่จับอยู่คืออะไร ในกระบวนการเรียนรู้เดียวกันเป็นแนวทางที่ผู้วิจัยทำการวิจัยเพื่อการเรียนรู้ของระบบหุ่นยนต์คล้ายมนุษย์

2.1.2 ลักษณะการสัมผัสวัตถุและชนิดข้อมูล

รูปแบบการสัมผัสเพื่อเก็บข้อมูลของมนุษย์ที่ใช้เพื่อการตีความว่าวัตถุที่สัมผัสอยู่นั้นคืออะไรมี 6 ลักษณะ [3] ดังรูปที่ 2.4 ประกอบด้วย การลูบเพื่อประเมินพื้นผิว (Lateral motion: texture) การกดเพื่อประเมินความแข็ง (Pressure: Hardness) การแตะเพื่อประเมินพื้นผิวและอุณหภูมิ (Static Contact: Texture, Temperature) การถือเพื่อประเมินรูปร่างและน้ำหนัก (Unsupported Holding: Shape, Weight) การกำเพื่อประเมินรูปร่างและพื้นผิว (Enclosure: Global Shape, Volume) และการหมุนเพื่อสำรวจรูปร่างและพื้นผิว (Contour Following: Global Shape, Texture) รูปแบบของการสัมผัสแบ่งออกเป็น 6 ลักษณะซึ่งได้ข้อมูลอันเป็นคุณสมบัติของวัตถุ (Object properties) ดังนี้

1. การลูบวัตถุ รูปแบบข้อมูลที่ได้คือความขรุขระของพื้นผิวของวัตถุ
2. การกด รูปแบบของข้อมูลที่ได้คือความแข็งของวัตถุ
3. การแตะ รูปแบบของข้อมูลที่ได้คือลักษณะพื้นผิวและอุณหภูมิของวัตถุ
4. การถือ รูปแบบข้อมูลที่ได้จะเป็นรูปทรงและน้ำหนักของวัตถุ
5. การจับ รูปแบบข้อมูลที่ได้คือรูปทรงและขนาดของวัตถุ
6. การหมุนวัตถุ รูปแบบข้อมูลที่ได้คือรูปร่างโดยรอบตัวและพื้นผิวของวัตถุ



รูปที่ 2.4 ลักษณะการสัมผัสวัตถุเพื่อประเมินลักษณะของวัตถุทั้ง 6 แบบ [3]

จากข้อมูลแต่ละชนิดที่ได้จากการสัมผัสจะถูกส่งผ่านสื่อกระแสประสาทไปยังสมองเพื่อแปลข้อมูลเป็นคำตอบว่าวัตถุคืออะไร การจดจำจากการสัมผัสพื้นผิว คือการจดจำวัตถุจากข้อมูลจากระบบประสาทสัมผัสซึ่งจะทำให้สมองตีความลักษณะพื้นผิวของวัตถุ และการเปลี่ยนแปลงของแรงกดที่กระทำต่อผิวหนังเมื่อทำการเปลี่ยนแรงกด จากหลักการทำงานของมือคนถูกนำไปพัฒนามือหุ่นยนต์เสมือนมนุษย์ โดยการสร้างลักษณะมือเลียนแบบทั้งโครงสร้างของมือ และผิวหนัง โดยการใช้ Tactile Sensor มาเป็นผิวหนังอิเล็กทรอนิกส์ (Electronic Skin)

2.2 มือหุ่นยนต์

ในงานวิจัยนี้ศึกษาการจดจำวัตถุของหุ่นยนต์จากการสัมผัสซึ่งเป็นการพัฒนาระบบเครื่องจักรให้ทำงานเสมือนมนุษย์ดังนั้นจึงต้องศึกษารูปแบบการทำงานของมือมนุษย์เป็นลำดับแรกเพื่อเป็นต้นแบบในการพัฒนามือหุ่นยนต์เนื่องจากมือมนุษย์จะเป็นพื้นฐานในการสร้างมือหุ่นยนต์ ในการสร้างมือหุ่นยนต์ ปัจจุบันมีอยู่ 2 แนวคิด [4], [5] แนวคิดแรกคือ สร้างเพื่อได้ตามเป้าหมาย แบบนี้จะเป็นมือที่อาจจะมี 2 หรือ 3 นิ้ว มีจุดหมุนไม่มาก แต่ทำงานได้ตรงตามจุดประสงค์ กับ แนวคิดที่สอง คือ มือที่มีความซับซ้อนในการออกแบบ มีนิ้วครบ 5 นิ้วเหมือนมือมนุษย์ ดังรูปที่ 2.5 โดยทฤษฎีแล้ว มือมนุษย์ถือเป็นผลงานการออกแบบอย่างชาญฉลาดของธรรมชาติผ่านกระบวนการวิวัฒนาการมาเป็นล้านๆ ปี ดังนั้นการสร้างมือหุ่นยนต์ให้มีความสามารถยอดเยี่ยมดังเช่นมือมนุษย์ก็ควรจะใช้วิธีสร้างเลียนแบบมือมนุษย์ให้ได้มากที่สุด การพัฒนามือหุ่นยนต์สำหรับงานจับวัตถุอย่างมีประสิทธิภาพแสดงดังตารางที่ 2.1

ตารางที่ 2.1 มือหุ่นยนต์ในวรรณกรรม

มือหุ่นยนต์	ข้อดี	ข้อเสีย	การติดตั้งเซ็นเซอร์
Allegro hand [6]	มีโครงสร้างคล้ายมือคน, เซ็นเซอร์เป็นอาเรย์	มีจำนวนนิ้วสั้น และติดเซ็นเซอร์ที่ปลายนิ้วอย่างเดียว	PPS Robot touch
iCub humanoid robot [7]	มีโครงสร้างเหมือนมือคน จับในรูปทรงกลมได้ เซ็นเซอร์เป็นอาเรย์	ความละเอียดเซ็นเซอร์ต่ำ	12 Sensor ในแต่ละนิ้ว และ 48 Sensor ที่ฝ่ามือ
Shadow hand & BioTac [8]	มีโครงสร้างเหมือนมือคน จับในรูปทรงกลมได้ เซ็นเซอร์เป็นอาเรย์	ติดเซ็นเซอร์เพียงจุดเดียวที่ปลายนิ้ว	BioTac Sensor
Anthropomorphic robotic [9]	มีโครงสร้างเหมือนมือคน จับในรูปทรงกลมได้	นิ้วมือมี 2 ข้อ ติดเซ็นเซอร์แบบอาเรย์นิ้วละ 2 จุด, ไม่มีเซ็นเซอร์ที่ฝ่ามือ	PVDF sensor 4 x 7 พิกเซล
Shadow hand & Tekscan [10]	มีโครงสร้างเหมือนมือคน จับในรูปทรงกลมได้, เซ็นเซอร์กระจายทั่วทั้งมือ	ไม่มีเซ็นเซอร์ตรงกลางฝ่ามือ, ไม่มีชั้นยืดหยุ่น	Tekscan tactile sensing
ARMAR-IIIb Hand [11]	มีโครงสร้างห้านิ้ว, จับในรูปทรงกลมได้, มีการติดตั้งเซ็นเซอร์ที่ฝ่ามือและนิ้วมือ	ติดตั้งเซ็นเซอร์ 3 นิ้ว	Weiss Robotics
PESA Hand [12]	มีโครงสร้างห้านิ้ว, มีการติดตั้งเซ็นเซอร์ที่นิ้วละ 3 จุด	ไม่มีการติดตั้งเซ็นเซอร์ที่ฝ่ามือ	pressure sensor
PR2 robot gripper [13]	โครงสร้างง่าย	โครงสร้างไม่เหมือนมือคน	5x3 pixels 2 ชุด
Gripper [14]	โครงสร้างง่าย	โครงสร้างไม่เหมือนมือคน	5x3 pixels 2 ชุด
Barrett Hand [15]	โครงสร้างง่าย	โครงสร้างไม่เหมือนมือคน	Weiss matrix sensor
AUBO OUR-i5 [16]	โครงสร้างง่าย, เซ็นเซอร์มีความละเอียดสูง	โครงสร้างไม่เหมือนมือคน, จับวัตถุไม่ได้	Tekscan 28 x 50 pixels
ROBOTIS & Gripper [17]	โครงสร้างง่าย, เซ็นเซอร์มีความละเอียดสูง	โครงสร้างไม่เหมือนมือคน	Tekscan 28 x 50 pixels



Allegro hand [6]



iCub humanoid robot [7]



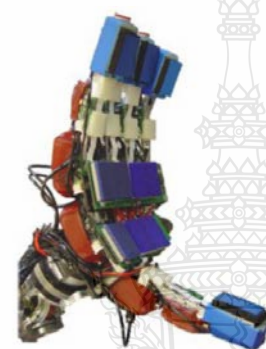
Shadow hand & BioTac [8]



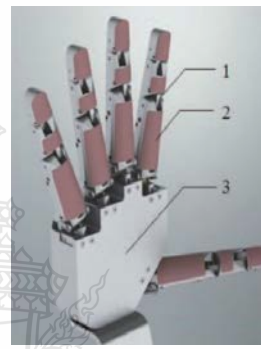
Anthropomorphic robotic [9]



Shadow hand & Tekscan [10]



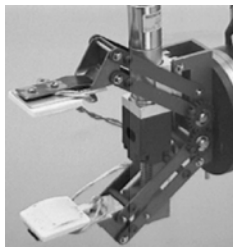
ARMAR-IIIb Hand [11]



PESA Hand [12]



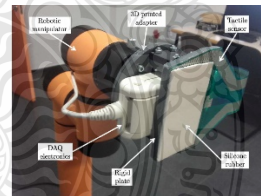
PR2 robot gripper [13]



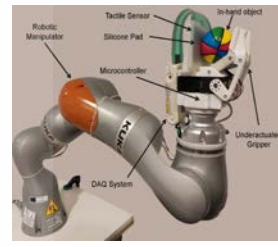
Gripper [14]



Barrett Hand [15]



AUBO OUR-i5 [16]



ROBOTIS & Gripper [17]

รูปที่ 2.5 มือหุ่นยนต์แบบต่างๆ ที่ปรากฏในวรรณกรรม

เนื่องจากมือมนุษย์มีความซับซ้อนมาก การสร้างมือหุ่นยนต์เลียนแบบมือมนุษย์จึงต้องลดรายละเอียดบางส่วนลง ในบางงานวิจัยการออกแบบชุดทดลองด้วยหุ่นยนต์ไม่ได้ออกแบบมือหุ่นยนต์ให้มีโครงสร้างแบบเดียวกับมือมนุษย์ [13, 14, 15, 16, 17] ซึ่งบางโครงสร้างไม่สามารถจับวัตถุได้ [16] การออกแบบให้มือจับวัตถุได้นั้นอย่างน้อยต้องมี 2 นิ้ว เป็นโครงสร้างมือแบบ Gripper ซึ่งมีทั้งแบบ 2

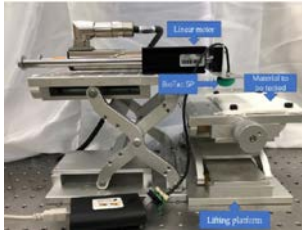
นิ้ว [13, 14, 17] และ 3 นิ้ว [14] ลักษณะมือของหุ่นยนต์แบบนี้ถูกออกแบบมาสำหรับการทำงานเฉพาะทาง ความหลากหลายในการจับและประเมินวัตถุยังไม่มีประสิทธิภาพมากนัก อย่างไรก็ตามก็ตีการพัฒนาที่หุ่นยนต์เสมือนมนุษย์ให้มีลักษณะเหมือนมือมนุษย์จริงก็ได้รับการพัฒนาขึ้นอย่างหลากหลาย [6, 7, 8, 9, 10, 11, 12] ผลจากการลดความซับซ้อนทำให้บางงานวิจัยมีนิ้วมือแค่ 4 นิ้ว [6] หรือมีข้อมือแค่ 2 ข้อ [9, 11] และการติดตั้ง Tactile Sensor ก็ไม่เต็มฝ่ามือ [6, 7, 8, 9, 11, 12]

2.3 Tactile sensor

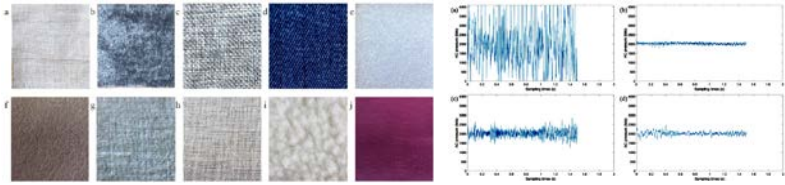
หุ่นยนต์เสมือนมนุษย์ไม่มีความสามารถในการระบุวัตถุโดยการสัมผัสได้ ซึ่งแตกต่างจากมนุษย์ที่สามารถทำได้เป็นอย่างดี อันเนื่องจากหุ่นยนต์ยังขาดเซ็นเซอร์สัมผัสที่เหมาะสมและขาดการประมวลผลการจดจำที่มีประสิทธิภาพ ดังนั้นจึงมีงานวิจัยหลากหลายที่พัฒนาให้หุ่นยนต์มีความสามารถนี้ การพัฒนาที่สำคัญของเทคโนโลยีหุ่นยนต์เสมือนมนุษย์สามารถแบ่งออกเป็นสองส่วน: 1) กายวิภาคของหุ่นยนต์, 2) ระบบประสาทหุ่นยนต์ การพัฒนาการด้านโครงสร้างและความสามารถในการเรียนรู้เหมือนมนุษย์เป็นสิ่งจำเป็นเพื่อให้หุ่นยนต์ทำงานในบ้านได้ หรือทำงานในสภาพแวดล้อมในสำนักงาน นอกจากนี้การพัฒนาหุ่นยนต์ที่มีโครงสร้างมือเหมือนมนุษย์เป็นสิ่งที่ต้องการ ในระบบการเรียนรู้และจดจำวัตถุของมนุษย์ใช้ประสาทสัมผัสทางกายแบบต่างๆ ซึ่งอยู่ในชั้นผิวหนังและใช้สมองในการแปลผล ในส่วนของหุ่นยนต์ใช้การรับรู้จากประสาทสัมผัสทางกายเทียมโดยประยุกต์ใช้ Pressure Sensor หรือ Tactile Sensor [4, 5, 20] ในการรับรู้ข้อมูลและข้อมูลที่ได้รับจะถูกส่งไปยังคอมพิวเตอร์เพื่อการวิเคราะห์วัตถุ [5] ในงานวิจัยที่ปรากฏมีการนำ Tactile Sensor มาใช้โดยการติดตั้งเข้ากับมือและส่วนต่างๆ ของหุ่นยนต์ แต่อย่างไรก็ตามในความหลากหลายของการทดสอบและลักษณะข้อมูลที่ได้ ตัวเซ็นเซอร์ที่ใช้จึงแตกต่างกันไป วิธีการตรวจสอบวัตถุของระบบหุ่นยนต์มีลักษณะเช่นเดียวกับการตรวจสอบวัตถุของมนุษย์ คือการลูบ การกด การจับ และการหมุนสำรวจวัตถุ ดังตารางที่ 2.2

ตารางที่ 2.2 การตรวจสอบวัตถุของหุ่นยนต์

วิธีการ	เซ็นเซอร์ที่ใช้	ข้อมูลที่ได้	การประยุกต์ใช้
การลูบวัตถุ [18]	BioTac Sensor	ข้อมูล 1 มิติ	การจดจำ texture
การกดวัตถุ [19]	Weiss Robotics	ข้อมูล 2 มิติ	การจดจำวัตถุ
การจับวัตถุ [11]	Weiss Robotics	ข้อมูล 2 มิติ	การจดจำวัตถุ
การหมุนสำรวจวัตถุ [20]	Force Sensing Resistor (FSR)	ข้อมูล 1 มิติ	การจดจำวัตถุ



BioTac Sensor [18]

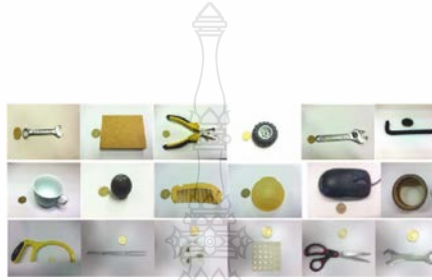


Texture

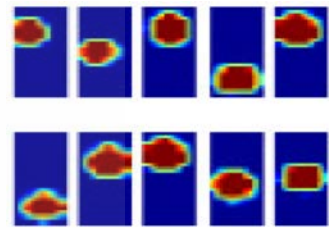
1-D Data



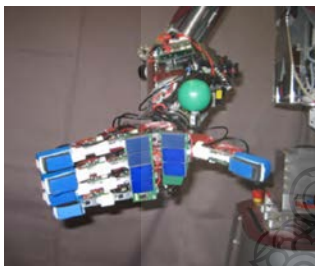
Weiss Robotics [19]



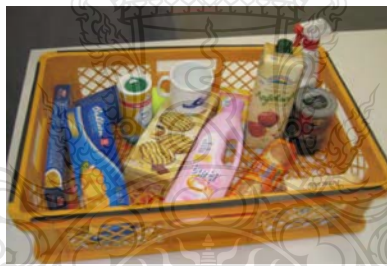
Object



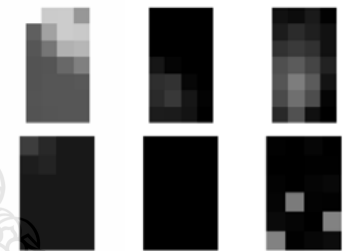
2-D Data



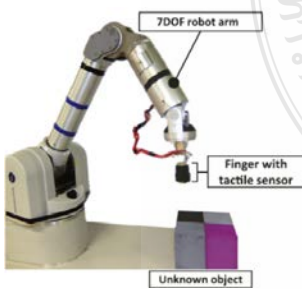
Weiss Robotics [11]



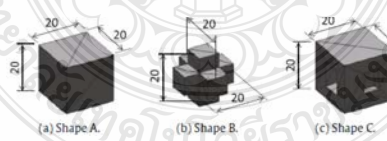
Object



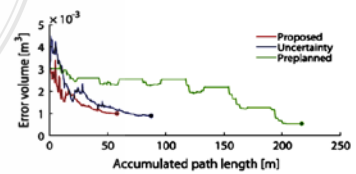
2-D Data



FSR [20]



Object



1-D Data

รูปที่ 2.6 การตรวจสอบวัตถุของหุ่นยนต์ทั้ง 4 แบบ

จากรูปที่ 2.6 การรับรู้วัตถุจากการสัมผัสของหุ่นยนต์แบ่งออกได้เป็น 3 โหมด โหมดแรกคือจากการเคลื่อนไหว เช่น การถูหรือลูบวัตถุ สำหรับหุ่นยนต์โหมดการรับรู้ข้อมูลนี้ได้จาก Tactile Sensor ที่ปลายนิ้วหรือฝ่ามือ ข้อมูลได้จากการลากเซ็นเซอร์ไปบนวัตถุและนำเสนอข้อมูลในรูปแบบ 1 มิติ (1-D) [18] โหมดที่สองคือการรับรู้ข้อมูลของหุ่นยนต์จากการกดหรือจับวัตถุโดยการใช้ Tactile Sensor Array สัมผัสวัตถุและนำเสนอข้อมูลในรูปแบบ 2 มิติ (2-D) หรือเรียกว่า Tactile Image [11, 19] โหมดการรับรู้ที่สามคือการเคลื่อนไหวร่างกาย การรับรู้โหมดนี้คือการรับรู้คือรูปร่างของวัตถุซึ่งข้อมูลจะได้จากข้อต่อของแขนเพื่อระบุการเปลี่ยนแปลงทิศทางและข้อมูลที่ได้รับใช้เพื่อระบุรูปร่างหรือเรขาคณิตของวัตถุ [20]

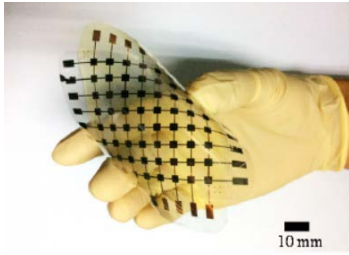
Tactile Sensor ถูกนำมาใช้เป็นประสาทสัมผัสทางกายในรูปแบบผิวหนังของหุ่นยนต์โดยการติดตั้งเข้ากับส่วนต่างๆ ของมือหุ่นยนต์ Tactile Sensor มีการพัฒนามาในหลายรูปแบบและหลายขนาดอีกทั้งมีหลักการทำงานของเซ็นเซอร์ที่หลากหลายเช่น การเปลี่ยนแปลงค่าความต้านทาน (Resistance) การเปลี่ยนแปลงค่าความจุไฟฟ้า (Capacitance) การกระจายแสง (Optical Distribution) และค่าแรงดันไฟฟ้า (Electrical) เป็นต้น [5, 21] ในงานวิจัยที่พัฒนา Tactile Sensor โดยมีการนำเสนอในประเด็นหลักการทำงานของเซ็นเซอร์ในแบบต่างๆ แสดงดังตาราง 2.3 และรูปที่ 2.7

ตารางที่ 2.3 Tactile Sensor แบบต่างๆ

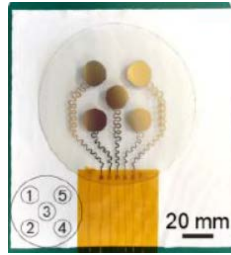
ชนิด	จุดเด่น	จุดด้อย	ตัวอย่าง
คาปาซิทีฟ (Capacitive)	สามารถออกแบบให้มีขนาดเล็กได้, spatial resolution สูง, มีความไวสูง, ทนต่อแรงกด, ตอบสนองได้รวดเร็ว	มีค่า Hysteresis สูง, มีค่าความจุค้ำง, ไวต่อคลื่นแม่เหล็กไฟฟ้ารบกวน, ไวต่ออุณหภูมิ, ไมลีนีเยอร์, Cross talk, วงจรอ่านค่าวงจรซับซ้อน	PPS Digi-Tacts [22], tactile skin for iCub [23], carbon micro coils [24], Polydimethylsiloxane [25]
เพียโซอิเล็กทริก (Piezoelectric)	ใช้งานได้ดี, มีความไวสูง, ลิเนียร์, ทนต่อสารเคมี, ตอบสนองได้ดีต่อการเปลี่ยนแปลงแรงกด	ไวต่ออุณหภูมิ, ขั้วต่อไฟฟ้าเปราะบาง, วัสดุเฉพาะช่วงที่มีการเปลี่ยนแปลงแรงกด	aPS fiber [26], Flexible P(VDF-TrFE) [27], Triboelectrification-enabled thin-film [28]
แสง (Optical)	ทนต่อสนามแม่เหล็กไฟฟ้า, ความไวสูง, ตอบสนองได้รวดเร็ว	มีขนาดใหญ่, สิ้นเปลืองพลังงาน, วงจรอ่านค่าซับซ้อน	Fiber optics tactile array [29], Optoelectronic [30], Fiber Optical [31]

ตารางที่ 2.3 Tactile Sensor แบบต่างๆ (ต่อ)

ชนิด	จุดเด่น	จุดด้อย	ตัวอย่าง
เพียโซรีซิสทีฟ (Piezoresistive)	ต้นทุนต่ำ, ความไวสูง, สัญญาณรบกวนต่ำ, วงจรรอ่านค่าไม่ซับซ้อน	ไม่ลีนีเยร์, มี Hysteresis สูง, ไวต่ออุณหภูมิ, มี drift ของสัญญาณสูง	Conductive Silicone [32], Low Density Polyethylene [33], Pressure sensitive rubber [34]
Tunnel effect	spatial resolution สูง, มีช่วงไดนามิกกว้าง, มี ความไวสูง	มีขนาดใหญ่, ไม่ลีนีเยร์	Quantum tunnel effect [35]
อัลตราโซนิก (Ultrasonic based)	ตอบสนองได้เร็ว, มีความ ละเอียดสูง	วงจรรอ่านค่าซับซ้อน, ไวต่ออุณหภูมิ	Ultrasonic based [36], Ultrasonic based [37]
แม่เหล็ก (Magnetism based)	ความไวสูง, มีช่วงไดนามิก กว้าง, ค่า Hysteresis ต่ำ, มีความทนทาน	ไวต่อสนามแม่เหล็ก รบกวน, มีการคำนวณซับซ้อน, มีขนาดใหญ่	Magnetic Field-Based [38], Magnetic Tactile Sensor [39]
Camera base	spatial resolution สูง มาก	มีขนาดใหญ่, ไวต่อแสง ภายนอก	vision-based tactile sensor [40], Gel Sight [41], Shadow [42]
Resonator base	ความไวสูง, ตอบสนองได้ เร็ว, ความละเอียดสูง	วงจรรอ่านค่าซับซ้อน, ชิ้นส่วนเปราะบาง	Perylene base [43], Stress-Sensitive Resonator [44]



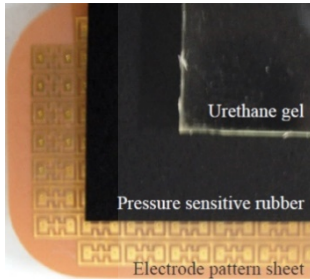
Capacitive [25]



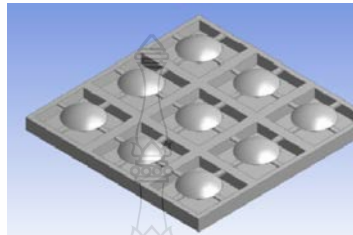
Piezoelectric [27]



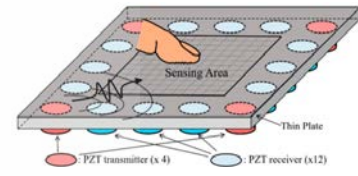
Optical [30]



Piezoresistive [34]



Tunnel effect [35]



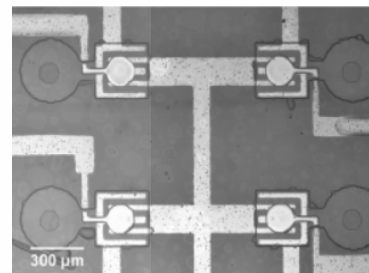
Ultrasonic based [37]



Magnetism based [38]



Camera base [40]



Resonator [43]

รูปที่ 2.7 Tactile Sensor แบบต่างๆ

Tactile Sensor Array บนพื้นฐานหลักการของการทำงานแบบ Piezoresistive มีอย่างหลากหลายเนื่องจากมีโครงสร้างที่ไม่ซับซ้อน วงจรการทำงานไม่ยุ่งยาก มีความไวสูง มีต้นทุนต่ำ และมีความทนทานสูง ดังนั้นจึงได้รับการพัฒนาสำหรับหลากหลายการประยุกต์ใช้งานเช่น การประยุกต์ใช้งานทางด้านการแพทย์ [45, 46], การประยุกต์ใช้ในงานอุตสาหกรรม [47], การประยุกต์ใช้ในงานวิศวกรรม [48], และการติดตามการทำกิจกรรมของมนุษย์ [49] เป็นต้น การประยุกต์ใช้ที่มีการวิจัยพัฒนาโดยใช้หลักการทำงานของเซ็นเซอร์ชนิดนี้มากที่สุดคือ Humanoid Robot [5-32, 33, 34] หลักการทำงานของ Piezoresistive Tactile Sensor สามารถอธิบายด้วย Piezoresistive Effect [32, 50] คือค่าความต้านทานจะเปลี่ยนแปลงไปเมื่อรูปร่างทางกลของวัสดุเปลี่ยนไป ซึ่งสามารถอธิบายได้ด้วยสมการที่ 2.1 และสมการที่ 2.2 [50]

$$\frac{\Delta R}{R} = (1 + 2\sigma + \pi E) \quad (2.1)$$

$$R = \rho \frac{L}{A} \quad (2.2)$$

ในสมการที่ 2.1, R คือความต้านทานของวัสดุตัวนำไฟฟ้าตามความยาว L , ΔR คือค่าความต้านทานที่เปลี่ยนแปลงไปตาม strain ที่เกิดขึ้นกับวัสดุตัวนำ, σ คือค่า Poisson's ratio of the material, π คือค่า Piezoresistive Coefficient, E คือค่า Young's modulus, χ คือค่าความเครียดของวัสดุที่เกิดจากแรงที่กระทำ, สมการที่ 2.2 เป็นสมการแสดงค่าความต้านทานของวัสดุตัวนำไฟฟ้า

งานวิจัยที่เกี่ยวข้องกับการพัฒนา Tactile Sensor Array สำหรับ Humanoid Robot สรุปไว้ในตารางที่ 2.4

ตารางที่ 2.4 Tactile Sensor แบบ Piezoresistive ในงานวรรณกรรม

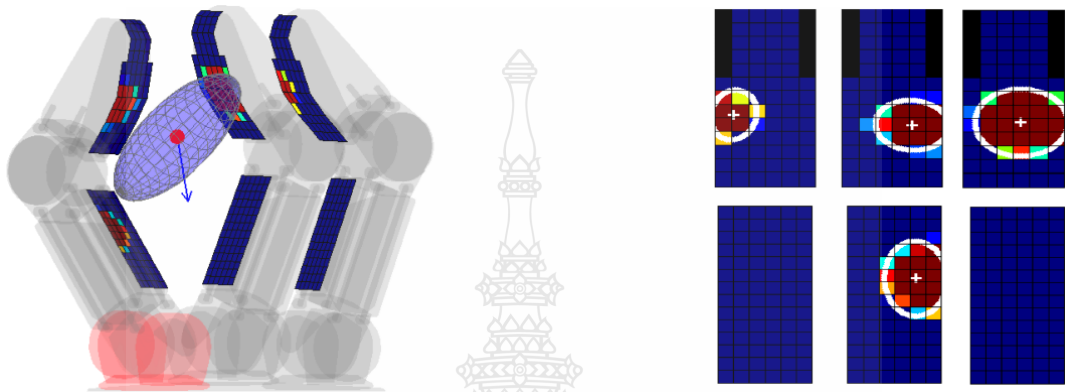
Year	Sensor Resolution (pixel)	Physical Size (mm ²)	Sensor size (mm ²)	Transducer	Application
2007 [34]	9 × 8	20 × 30	1.8 × 3.4	Conductive rubber	Finger
2009 [9]	4 × 7	20 × 35	20 × 35	Polyvinylidene Fluoride	Finger
2009 [51]	25 × 25	150 × 150	5.0 × 5.0	Organic resistance	Flex e-Skin
2009 [52]	5 × 5	32 × 32	50	FSR	Touch sensing
2010 [53]	16 × 16	50 × 50	3.44	Conductive polymer	Flex e-Skin
2011 [54]	8 × 8	20 × 20	2.5 × 2.5	Conductive rubber	Gripper Finger
2011 [55]	8 × 16	70 × 120	3.0 × 3.0	Nickel powder +PDMS	Flex e-Skin
2014 [56]	8 × 8	-	2.0 × 2.0, 3.0 × 3.0	P(VDF-TrFE), MWCNT/PDMS	Flex e-Skin
2014 [57]	10 × 10	20 × 20	1.3 × 1.3	Nanoparticles of carbon and silica	Finger
2015 [58]	3 × 3	50 × 50	10 × 10	Conductive pillars	Flex e-Skin
2017 [59]	8 × 8	300 × 220	37.5 × 2.5	Conductive polymer	e-Skin
2017 [60]	8 × 8	100 × 100	480	Polymer composite	e-Skin
2018 [33]	1 × 1	20 × 25	20 × 25	Conductive polymer	Finger tip
2019 [61]	4 × 4	10 × 10	1900	Polymer composite	e-Skin

การออกแบบ Tactile Sensor แบบ Piezoresistive ในงานวรรณกรรมนั้นถูกพัฒนาอย่างต่อเนื่องและมีความหลากหลายของงาน โดยงานแรกที่เป็นต้นแบบของการออกแบบ Tactile Sensor แบบ Piezoresistive มาประยุกต์ใช้กับมือของหุ่นยนต์คือในปี ค.ศ. 2007 Nakamoto และคณะ [34] ได้การออกแบบ Tactile Sensor Array สำหรับติดตั้งที่นิ้วของ Humanoid Robot มีขนาด $20 \times 30 \text{ mm}^2$ ตัวเซ็นเซอร์มีจุดเด่นที่มีโครงสร้างที่ง่ายเป็นแบบแผ่นวงจรพิมพ์ (Printed Circuit Board : PCB) แต่มีข้อเสียเรื่องการออกแบบ Pixel ที่ไม่สมมาตร และมีขนาดเล็กจึงไม่เหมาะกับการติดตั้งที่ฝ่ามือของ Humanoid Robot [34], การออกแบบ Tactile Sensor Array สำหรับติดตั้งที่นิ้วของ Humanoid Robot ซึ่งมีขนาดอยู่ระหว่าง $20 \times 20 \text{ mm}^2$ [54, 57] และ $20 \times 35 \text{ mm}^2$ [9] ด้วยขนาดเล็กไปจึงไม่เหมาะกับการนำมาติดตั้งที่ฝ่ามือของ Humanoid Robot เช่นกัน, Drimus และคณะ [54] ได้พัฒนาเซ็นเซอร์ที่มีโครงสร้างที่ง่ายและในงานนี้ได้นำเสนอการจดจำวัตถุจากการสัมผัสวัตถุบนพื้นฐานของ Tactile Image Recognition โดยใช้โอเตอร์ลำดับหนึ่ง (First Order Feature) ทำงานร่วมกับ K-Nearest Neighbors (KNN) แต่ในการทดสอบ ใช้ Gripper ทำหน้าที่เป็นนิ้วมือ ดังนั้นตัวเซ็นเซอร์จึงมีขนาดเล็กเกินไปสำหรับการนำมาติดตั้งที่ฝ่ามือของ Humanoid Robot, ส่วนงานที่พัฒนา Tactile Sensor Array แบบความละเอียด (Resolution) สูง แต่กลับมีขนาดใหญ่เกินไปจึงไม่เหมาะกับการนำมาติดตั้งที่ฝ่ามือของ Humanoid Robot [51, 55, 59, 60], มีงานวิจัยอีกหลายงานที่ทำการออกแบบ Tactile Sensor Array ที่มีความละเอียดต่ำเกินไป ไม่เหมาะสำหรับการนำมาประมวลผลในเชิง Tactile Image Recognition [52, 58, 33, 61], ในปี 2014 Khan และคณะ [56] ได้พัฒนาเซ็นเซอร์ขนาด 8×8 pixels ซึ่งความละเอียดนี้สามารถนำมาติดตั้งกับฝ่ามือหุ่นยนต์และสามารถใช้เทคนิคการประมวลผลภาพในการคำนวณได้แต่งานนี้ไม่ได้ระบุขนาดทางกายภาพไว้, Ramos และคณะ [53] ได้นำเสนอเซ็นเซอร์ที่เหมาะสมกับฝ่ามือของ Humanoid Robot ทั้งขนาดและความละเอียดแต่ในงานไม่ได้นำเสนอเรื่องอัลกอริทึม (Algorithm) ของ Object Recognition ที่เหมาะสมกับเซ็นเซอร์นี้ไว้ จากที่กล่าวมาทั้งหมดพบว่างานวิจัยส่วนใหญ่ได้ออกแบบเซ็นเซอร์ที่มีขนาดเล็กจึงไม่เหมาะกับการนำมาติดตั้งที่ฝ่ามือของ Humanoid Robot ดังนั้นในงานวิจัยนี้จึงได้นำเสนอการออกแบบเซ็นเซอร์แบบใหม่ที่มีความละเอียดและมีขนาดที่เหมาะสมกับการนำไปติดตั้งที่ฝ่ามือของ Humanoid Robot เพื่อนำมาใช้ในการประมวลผลการจดจำด้วยวิธี Tactile Image Recognition ต่อไป

2.4 การจดจำวัตถุจากภาพสัมผัส

การพัฒนา Humanoid Robot Hand ถูกพัฒนาด้วยการติดตั้งเซ็นเซอร์แบบต่างๆ เพื่อการประยุกต์ใช้งานที่หลากหลาย การพัฒนาหนึ่งที่สำคัญคือ Object Recognition in Hand Method เป็น การจดจำวัตถุจากการประมวลผลข้อมูลที่ได้จากเซ็นเซอร์ ข้อมูลที่สำคัญที่ถูกนำมาใช้ในการประมวลผล

คือข้อมูลที่ได้จากการติดตั้ง Tactile Sensor Array ซึ่งเราเรียกวิธีการนี้ว่า Tactile Image Recognition ลักษณะของมือที่ติดตั้ง Tactile Sensor Array แสดงดังรูปที่ 2.8 เมื่อมีการจับวัตถุ Tactile Sensor Array ที่ติดตั้งไว้ที่มือของหุ่นยนต์จะถูกสัมผัสด้วยวัตถุที่อยู่ในมือ และเมื่อมีการจับที่แรงพอจะทำให้สามารถอ่านค่าที่ได้จากเซ็นเซอร์เมื่อนำข้อมูลที่อ่านได้จากเซ็นเซอร์มาจัดเรียงในรูปแบบเมทริกซ์และนำเสนอในรูปแบบ 2 มิติ จะได้ภาพ Tactile Image ดังรูปที่ 2.8



รูปที่ 2.8 มือหุ่นยนต์ และ Tactile Image [62]

การวิเคราะห์ Tactile object recognition ด้วยเทคนิคการประมวลผลภาพจึงถูกนำมาใช้ในการตีความข้อมูลเพื่อทำนายวัตถุที่อยู่ในมือ งานวรรณกรรมที่เกี่ยวข้องกับการประมวลผลข้อมูล Tactile image จากการจับวัตถุด้วยมือหุ่นยนต์ได้สรุปไว้ในตารางที่ 2.5 ดังนี้

ตารางที่ 2.5 Tactile object recognition ในงานวรรณกรรม

Year	Sensor Resolution (pixel)	Number of Class	Descriptor	Classification Method	Number of grasps	Recognition rate (%)
2009 [63]	6 x 14 (2ea)	21	All Data Vector	BoF	10	84.6
2010 [64]	4 x 7 (9ea)	7	Moment Analysis	SOM	1	≈80.0
2011 [65]	8 x 8(2ea)	10	Mean, SD	KNN	1	92.0
2011 [66]	6 x 6	5	Polar Furrier	PCA, BoF	>50	90.0

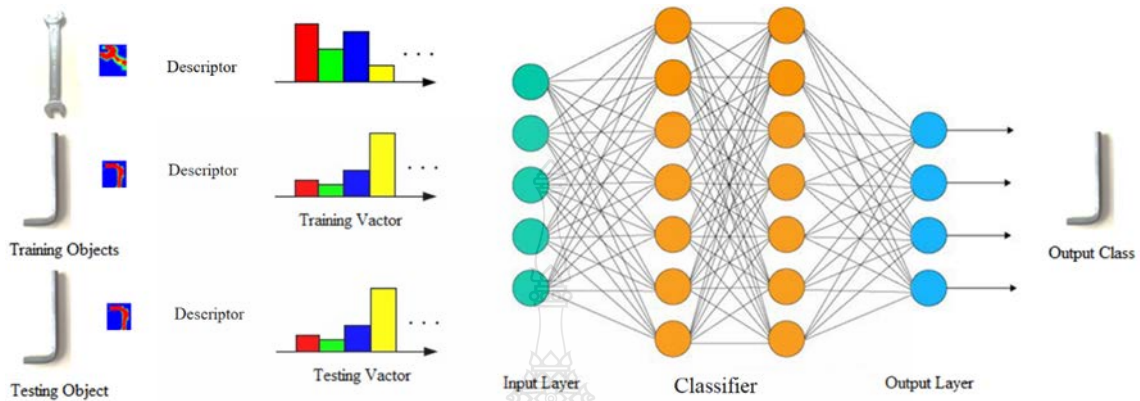
ตารางที่ 2.5 Tactile object recognition ในงานวรรณกรรม (ต่อ)

Year	Sensor Resolution (pixel)	Number of Class	Descriptor	Classification Method	Number of grasps	Recognition rate (%)
2012 [67]	24 x 16	4	Maximum Vector	PCA, KNN	1	81.4
2012 [10]	5 x 9, 12 x 10	4	3x3 Segmentation	ANN	1	91.0
2013 [68]	32 x 32	10	Haar Wavelet	KNN, LDA, Naïve Bayes	1	86.0
2014 [57]	10 x 10	12	Kernel PCA, FD,GE	MKL-SVM	1	85.54
2015 [19]	6 x 14	18	Segmentation SIFT	BoF	15	89.9
2015 [69]	16 x 16	25	Reduced Vector	ANN	1	96.0
2015 [70]	108 + 133	20	All Data Vector	DNN	1	91.1
2016 [71]	8 x 3 (3ea), 7 x 4	7	All Data Vector	JKSD	1	94.0
2016 [72]	6 x 14	20	Zernike moment	iCLAP	20	85.36
2017 [73]	6 x 6	15	Ensemble of Shape Functions	ANN	30	94.67
2017 [74]	768	2	Convolutional	AlexNet DCNN	1	98.3
2017 [75]	28 x 50	8	Convolutional	AlexNet DCNN-SVM	1	91.6
2019 [76]	28 x 50	22	Convolutional	RestNet DCNN	1	95.36

จากตารางที่ 2.5 สามารถแบ่งวิธีในการวิเคราะห์ผลได้เป็น 3 วิธี คือ 1) วิธีการจัดกลุ่มภาพแบบทั่วไป (General Image Classification) หรือวิธีการแมชชีนเลิร์นนิง (Machine learning : ML), 2) การจัดกลุ่มภาพด้วยวิธีแบ็กออฟเวด หรือ ถุงคำ (Bag of Word : BoW) ซึ่งเรียกอีกอย่างว่า แบ็ก ออฟ

พีเจอร์ หรือ ถุงพีเจอร์ (Bag of Feature : BoF) และ 3) วิธีโครงข่ายประสาทแบบคอลโวลูชัน (Convolution Neural Network :CNN)

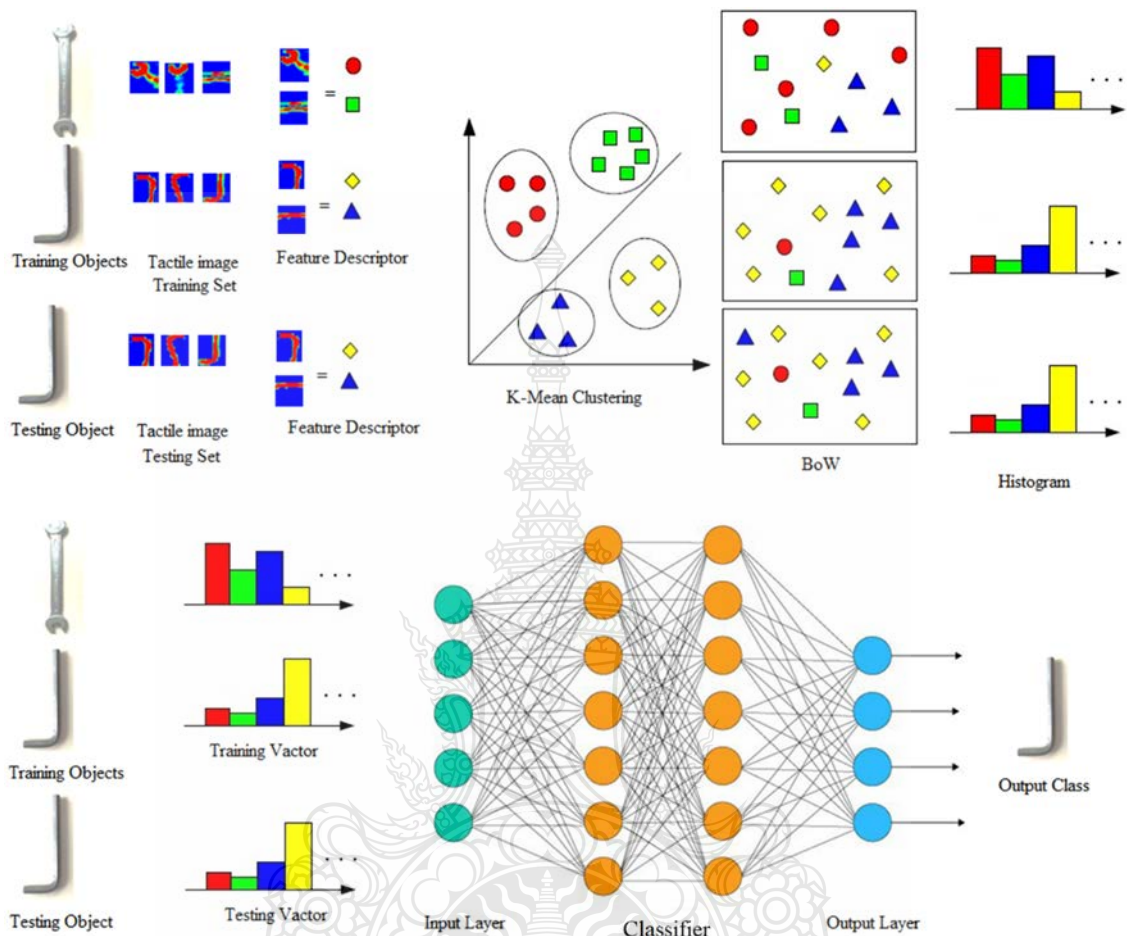
2.4.1 การจัดกลุ่มภาพแบบ Machine Learning



รูปที่ 2.9 การทำงานของการจัดกลุ่มภาพแบบ Machine Learning

โดยส่วนใหญ่วิธีดำเนินการ Tactile Object Recognition โดยวิธี Machine learning นี้จะมีรูปแบบการทำงานดังรูปที่ 2.9 [7, 10, 67, 65, 67, 68, 69, 70, 73] การดำเนินงานเริ่มจากนำเข้าสู่ชุดข้อมูลภาพเข้าสู่ระบบ ซึ่งข้อมูลภาพจะถูกแบ่งออกเป็นชุดสอน (Training Set) และชุดทดสอบ (Testing Set) และเพื่อลดขนาดเวกเตอร์ของภาพชุดภาพดังกล่าวจะถูกนำมาทำ Normalization และ Thresholding ซึ่งในขั้นตอนนี้ในบางงานวิจัยอาจทำการปรับปรุงภาพด้วยวิธีอื่นเช่นการแปลงเป็นภาพสี [68], การแปลงร่วมกับฟังก์ชันอื่นๆ [64] การแปลงร่วมกับฟังก์ชันภาพจากกล้อง [73], และการขยายขนาดภาพ [74] เป็นต้น จากนั้นภาพที่ผ่านกระบวนการจะเข้าสู่การสกัดคุณลักษณะ (Feature Extraction) เพื่อดึงค่าที่สามารถเป็นตัวแทนของภาพดังกล่าวได้ เรียกค่าตัวแทนนี้ว่า Feature หรือ Descriptor วิธีการทำ Feature Extraction มีหลากหลายวิธีเช่นเดียวกับงานทางด้าน Visual Image เช่นวิธีทางสถิติ (Statistic) และ First Order Feature, วิธี Moment Analysis (MA), วิธี Polar Furrier (PF), Haar Wavelet, Principal Component Analysis (PCA), Scale-Invariant Feature Transform (SIFT), Zernike moment, Ensemble of Shape Functions และวิธี Convolutional เป็นต้น จากนั้นภาพ Training Set จะถูกนำไปสอนให้กับโมเดลสำหรับจัดกลุ่ม (Classification Model) ซึ่งมีหลายวิธีเช่น KNN, Liner Discriminant Analysis (LDA), Support Vector Machine (SVM), Naïve Bayes, Self-Organizing Map (SOM) และ Artificial Neural Network (ANN) เป็นต้น วิธีนี้จะเหมาะกับ Tactile Image ที่มีลักษณะ Global Shape [62] หรือมี Sensor ขนาดใหญ่ใกล้เคียงกับวัตถุจะให้ประสิทธิภาพในการจดจำได้สูง

2.4.2 Bag of Feature



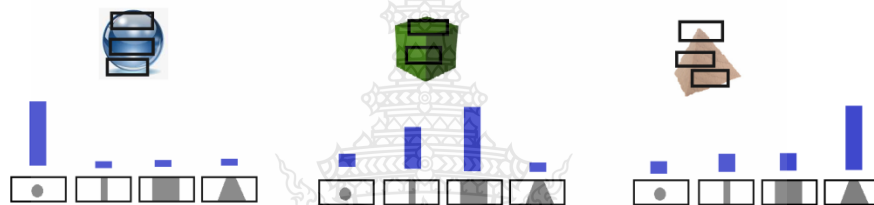
รูปที่ 2.10 การทำงานของ BoF

Bag of Feature หรือ BoF [19, 63, 66] ดังรูปที่ 2.10 เป็นโมเดลที่ถูกนำมาใช้ในการจดจำวัตถุจากการสัมผัสที่เหมาะสมกับ Tactile Image ที่มีลักษณะ Local Shape [62] หรือมีเซ็นเซอร์ขนาดเล็กกว่าวัตถุซึ่งต้องสัมผัสวัตถุหลายครั้งเพื่อหาเอกลักษณ์ของวัตถุ ลำดับการทำงานเริ่มจากการนำเข้าสู่ชุดข้อมูล ซึ่งประกอบไปด้วย 2 ชุดคือชุด Training Set หนึ่งชุดและชุด Testing Set อีกหนึ่งชุด ลำดับขั้นตอนการทำงานเป็นดังนี้

1. นำข้อมูลภาพชุด Training Set เข้ามาเพื่อทำ Feature Extraction หรือ Descriptor ซึ่งจะได้ Feature ของข้อมูลแต่ละตัวมา

2. Feature ซึ่งถือเป็นตัวแทนของภาพในแต่ละคลาสจะถูกนำมาจัดกลุ่มด้วยเทคนิค K-Mean Clustering เพื่อระบุว่าภาพนั้นถูกจัดอยู่ใน Cluster ไต การใช้เทคนิค K-Mean Clustering เป็น

เทคนิคเรียนรู้แบบ Unsupervised หรือการเรียนรู้แบบไม่ต้องสอน ซึ่งเราเรียกกระบวนการนี้ว่า Bag of Feature ความหมายคือการสร้างถุงที่ไว้สำหรับ Feature ถุงนี้เรียกอีกอย่างว่าคลัสเตอร์ (Cluster) และจำนวนของคลัสเตอร์เรียกว่าบิน (bin) เมื่อ Feature ที่เข้ามาถูกจัดอยู่ในกลุ่มใดก็จะถูกจัดเก็บใน bin นั้น เมื่อข้อมูลภาพในคลาสนั้นๆ ถูกนำเข้ามาจัดกลุ่มแล้วก็จะได้อิสโตแกรม (Histogram) ของคลาสนั้นๆ ซึ่งจะถูกนำไปใช้ในการ Training ในลำดับถัดไป ตัวอย่างของภาพที่แตกต่างกันตัวอย่างเช่นในรูปที่ 2.11 วัตถุ 3 ชนิดเมื่อนำภาพจัดกลุ่มคลัสเตอร์และสร้าง Histogram จะได้ว่าวัตถุทรงกลมจะมีรูปที่ซ้ำกันบ่อยที่สุดคือ Feature จุดกลมเล็กๆ วัตถุลูกบาศก์จะมีภาพซ้ำกันเยอะสุดคือ Feature รูปสี่เหลี่ยม และพีระมิดมีภาพที่มี Feature ซ้ำกันบ่อยที่สุดคือสี่เหลี่ยมคางหมู Histogram ที่ได้ของวัตถุแต่ละ class จะแสดงความถี่มากน้อยของ Feature ของวัตถุแต่ละตัว ดังนั้น Histogram ที่ได้จากการจัด Cluster คือตัวแทนของวัตถุในคลาสนั้นๆ นั่นเอง



รูปที่ 2.11 Histogram ของวัตถุ 3 ชนิดที่ได้จากการจัดคลัสเตอร์

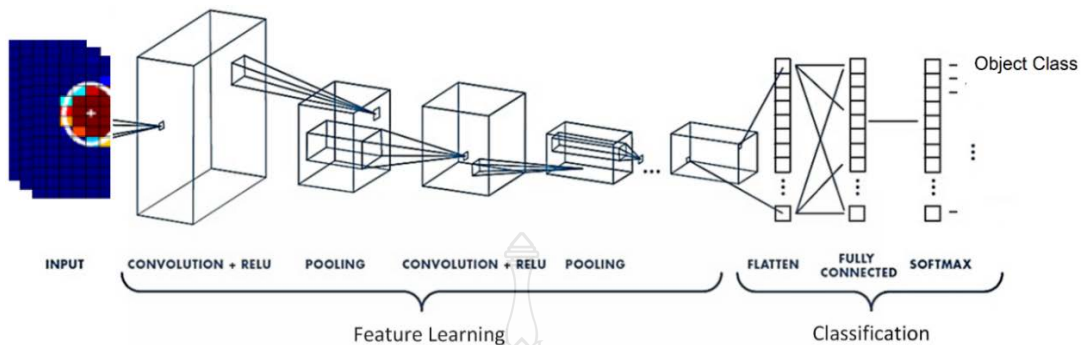
3. จากนั้นนำ Histogram เข้าสู่การ Training โดย Classification Model การ Training เพื่อการเรียนรู้ของระบบในขั้นตอนนี้ต้องเป็นการเรียนรู้แบบ Supervised หรือกระบวนการเรียนรู้แบบมีการสอน Model ที่ใช้ในการเรียนรู้มีหลาย Model เช่น KNN, SVM และ ANN เป็นต้น

4. เมื่อ Training ระบบเสร็จจากนั้นก็จะเป็นการทดสอบระบบหรือ Testing ซึ่งเริ่มจากการนำข้อมูลภาพชุด Testing Set เข้ามาสู่กระบวนการทำ Feature Extraction ซึ่งจะได้ Feature ของข้อมูลแต่ละตัวมา

5. จากนั้นนำ Feature ของภาพชุด Testing Set มาสร้าง Histogram ด้วยเทคนิค K-Mean Clustering เช่นเดียวกับชุด Training เมื่อได้ Histogram แล้วยก็นำเข้าไปทดสอบกับ Classifier

6. เมื่อทดสอบเสร็จก็จะได้คำตอบมาว่าภาพที่นำมาทดสอบอยู่ในคลาสนี้ใดนั่นเอง

2.4.3 โครงข่ายประสาทเทียมแบบ Convolution Neural Network (CNN)



รูปที่ 2.12 การทำงานของ CNN [77]

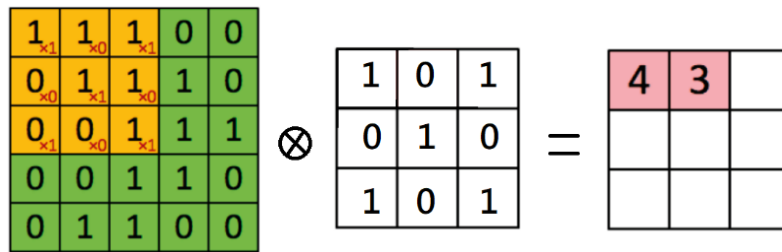
โครงข่ายประสาทเทียมแบบ Convolution Neural Network หรือ CNN มีโครงสร้างดังรูปที่ 2.12 เป็นสถาปัตยกรรม (Architecture) ในรูปแบบหนึ่งของ Feed-Forward Neural Networks ซึ่งถูกพัฒนาขึ้นเพื่อวัตถุประสงค์ในงานการเรียนรู้ด้วยภาพโดยในปี 1990 LeCun และคณะ [77] ได้นำวิธีการคำนวณแบบคอนโวลูชันมาผนวกเข้ากับโครงข่ายประสาทเทียมจากนั้นในปี 1998 LeCun และคณะ [78] ได้นำเสนอโครงข่ายประสาทเทียมแบบคอนโวลูชันที่ชื่อ LeNet-5 ประกอบไปด้วย 3 ชั้นที่สำคัญดังนี้

1. ชั้นคอนโวลูชัน (Convolutional Layer) การทำคอนโวลูชัน (Convolutional) เพื่อทำ Feature Extraction หรือดึงลักษณะเด่นที่สำคัญของภาพตัวอย่างออกมา โดยการคำนวณแบบ Dot Product ระหว่างพื้นที่ส่วนย่อยของรูปภาพ (Sub-region) กับ Filter หรือเคอร์เนล (Kernel) ขนาดเล็กซึ่งมีลักษณะเป็นเมทริกซ์จัตุรัส (Square Matrix) เช่นขนาด 3×3 เป็นต้น ซึ่งผลจะได้ชุดข้อมูลชุดใหม่ที่เกิดจาก $h * F$ เรียกข้อมูลชุดใหม่ที่ถูกสร้างขึ้นว่า Feature Map การคำนวณ Convolutional สามารถทำได้ดังสมการที่ 2.3 และ 2.4 และผลที่ได้จากการทำคอนโวลูชันแสดงดังรูปที่ 2.13

$$G = h * F \quad (2.3)$$

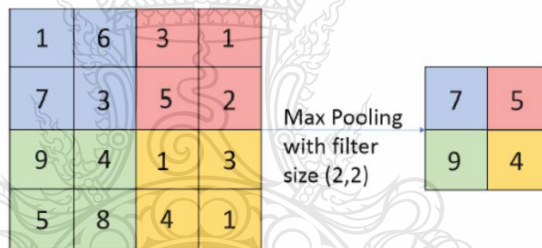
$$G[i, j] = \sum_{u=-k}^k \sum_{v=-k}^k h[u, v] F[i - u, j - v] \quad (2.4)$$

เมื่อ h คือเมทริกซ์ของภาพ และ F คือ เมทริกซ์ของเคอร์เนลขนาด $k * k$



รูปที่ 2.13 ผลของการทำคอนโวลูชันแบบ 2 มิติ [79]

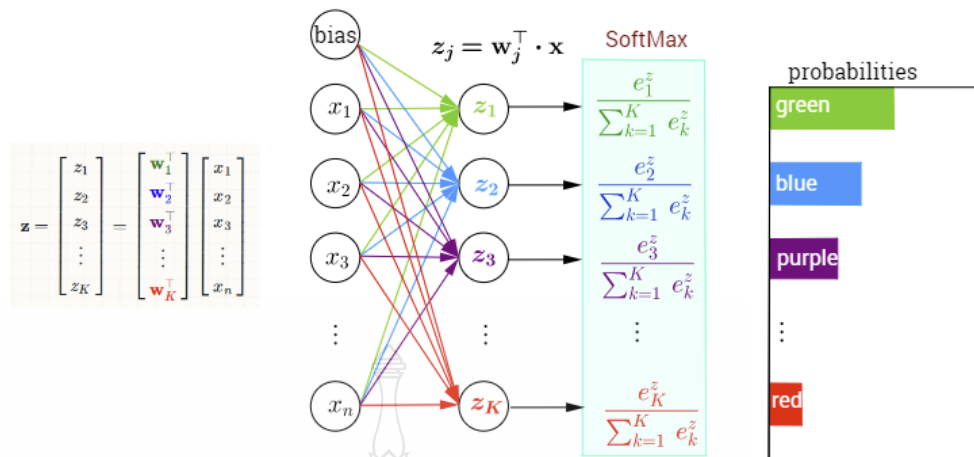
2. ชั้นพูลลิ่ง (Pooling Layer) เป็นชั้นที่ถัดจากชั้น Convolution Layer หน้าที่ของชั้นนี้เพื่อลดมิติของ Feature Map ให้มีขนาดเล็กลง (Downsampling) เพื่อความไวในการคำนวณ ด้วยการหาค่าที่สูงที่สุดด้วยการใช้ฟังก์ชันค่าสูงสุดเรียกว่า Max Pooling หรือหาค่าเฉลี่ยด้วยฟังก์ชันค่าเฉลี่ยเรียกว่า Average Pooling วิธีที่นิยมใช้มากที่สุดคือวิธี Max Pooling ทำได้โดยนำ Feature Map มาแบ่งออกเป็นพื้นที่ย่อย (Sub-region) ขนาดเป็น $P \times P$ เช่นขนาด 2×2 แล้วหาค่าที่มากที่สุดในแต่ละพื้นที่ย่อยเพื่อสร้าง Feature Map ใหม่ดังรูปที่ 2.14



รูปที่ 2.14 การทำ Max Pooling ในชั้น Pooling Layer [79]

3. ชั้นเชื่อมต่อแบบสมบูรณ์ (Fully-Connected Layer) เป็นชั้นสุดท้ายของ network ซึ่งเป็นส่วนของการ Classification โดยใช้สถาปัตยกรรมเพอร์เซปตรอนแบบหลายชั้น (Multi-Layer Perceptron : MLP) โดยที่ Feature Map ในชั้นสุดท้ายจะเป็นอินพุต (Input) ที่นำเข้าสู่โครงข่ายประสาทเทียม ผลลัพธ์ที่ได้จากชั้น Fully Connected Layer จะเป็น Output Prediction ซึ่งคำนวณโดย Softmax Function แปลงค่าการทำนายอยู่ในรูปแบบของความน่าจะเป็นสามารถหาได้ตามสมการที่ 2.5 และลักษณะการทำงานแสดงดังรูปที่ 2.15

$$\sigma(j) = \frac{\exp(z_j)}{\sum_{k=1}^K \exp(z_k)} \quad (2.5)$$



รูปที่ 2.15 การทำชั้น Fully-Connected Layer และ Soft Max Function [80]

ปัจจุบัน Model ของ CNN ถูกพัฒนาให้มีความลึกในการสกัด Feature โดยโครงสร้างหลายชั้นซึ่งถูกนำมาใช้กับการ Classify ข้อมูลภาพ เรียก CNN ที่มีโครงสร้างลึกหลายชั้นว่า Deep Convolution Neural Network หรือ DCNN เช่น AlexNet [88] เป็นต้น

2.5 การสกัดคุณลักษณะจากภาพ

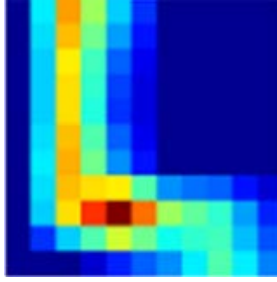
การจดจำรูปแบบเป็นส่วนสำคัญของระบบวิเคราะห์ภาพระดับสูง การวิเคราะห์ภาพแบ่งเป็นสี่ขั้นตอนได้แก่ 1) การจัดหารูปภาพ, 2) การประมวลผลเบื้องต้น (Pre-image Processing), 3) การสกัดคุณลักษณะ (Feature Extraction) หรือ Descriptors และ 4) การจำแนกกลุ่ม (Classification) การสกัดคุณลักษณะภาพคือการลดมิติของข้อมูลและดึงค่าเอกลักษณ์ของภาพออกมาเพื่อใช้ในการวิเคราะห์ต่อไป คุณลักษณะจากภาพสามารถดึงมาได้จากหลายหลายวิธีการดังนี้

2.5.1 Statistic Descriptor

Statistic Descriptor เป็นวิธีทางสถิติแบบ First Order กล่าวคือเป็นวิธีที่ใช้ข้อมูลดิบมาหาค่าทางสถิติ เช่นการจัดเรียงเวกเตอร์ของข้อมูล (Vectorize), ฮิสโตแกรมของข้อมูล (Histogram), ค่าเฉลี่ยของข้อมูล (Average) และค่าเบี่ยงเบนมาตรฐาน (Standard Deviation) เป็นต้น

2.5.1.1 เวกเตอร์ของข้อมูล

เวกเตอร์ของข้อมูลคือการจัดเรียงข้อมูลภาพใหม่จากรูปแบบเมทริกให้เป็นเวกเตอร์จากภาพในรูปที่ 2.16 ตัวข้อมูลภาพ ดังกล่าวเป็นรูปแบบเมทริกของข้อมูลซึ่งแทนด้วยเมทริก I ดังสมการที่ 2.6



รูปที่ 2.16 Tactile Image [19]

$$I = \begin{bmatrix} x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(m)} \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(m)} \\ \vdots & \vdots & \ddots & \vdots \\ x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(m)} \end{bmatrix} \quad (2.6)$$

เมื่อนำข้อมูลมาจัดเรียงเป็นข้อมูลเวกเตอร์ 1 มิติจะได้เวกเตอร์ v ดังสมการที่ 2.7

$$v = x_1^{(1)}, x_2^{(2)}, x_3^{(3)}, \dots, x_n^{(m)} \quad (2.7)$$

การทดสอบระบบสามารถนำเวกเตอร์ข้อมูลเข้าทำการทดสอบได้โดยตรงโดยไม่ต้องเปลี่ยนแปลงหรือทำอะไรกับข้อมูลเดิมเลย ลักษณะการทดสอบระบบโดยการไม่เปลี่ยนแปลงข้อมูลนี้เมื่อตำแหน่งของวัตถุเปลี่ยนหรือเกิดการหมุนของวัตถุรูปแบบข้อมูลจะเปลี่ยนแปลงไปตามการหมุน

2.5.1.2 ฮิสโตแกรม

ฮิสโตแกรมคือกราฟแห่งที่แสดงความสัมพันธ์ระหว่างข้อมูลเป็นหมวดหมู่ที่เรียกว่าชั้นข้อมูลกับความถี่ของข้อมูล เพื่อดูการกระจายของข้อมูล ลักษณะของข้อมูลที่เป็นหมวดหมู่จะเรียงลำดับจากน้อยไปหามากโดยแกน y จะเป็นตัวเลขแสดง “ความถี่” และแกนอนจะเป็นลำดับของข้อมูล เราสามารถนำฮิสโตแกรมมาใช้ในการทดสอบระบบได้เนื่องจากภาพจากการสัมผัสวัตถุที่ต่างกันจะมีฮิสโตแกรมภาพที่ต่างกันไปด้วย การหาค่าฮิสโตแกรมของข้อมูลเมทริกสามารถหาได้จากสมการที่ 2.8

$$v_i = \sum_{j=1}^k c_i \quad (2.8)$$

เมื่อ v_i คือลำดับของข้อมูล 0-255 และ c_i คือข้อมูลในลำดับ i นั้นๆ

ข้อดีประการหนึ่งของฮิสโตแกรมคือแม้ว่าตำแหน่งของวัตถุจะเปลี่ยนหรือเกิดการหมุนฮิสโตแกรมก็ยังคงเหมือนเดิมแต่ข้อด้อยคือแม้ว่าภาพจะมีความแตกต่างกันแต่สามารถมีฮิสโตแกรมที่เหมือนกันได้

2.5.1.3 ค่าเฉลี่ย

จากข้อมูลที่มีเวกเตอร์เป็นจำนวนมากซึ่งเท่ากับจำนวน pixel ของภาพการหาค่าเฉลี่ยของภาพเป็นวิธีการหนึ่งที่สำคัญในการลดเวกเตอร์ข้อมูลโดยยังคงเป็นตัวแทนของข้อมูลอยู่ การหาค่าเฉลี่ยของภาพสามารถหาได้จากสมการที่ 2.9

$$\mu = \frac{1}{NM} \sum_{i,j} x_i^j \quad (2.9)$$

2.5.1.4 ค่าเบี่ยงเบนมาตรฐาน

ค่าเบี่ยงเบนมาตรฐาน (Standard Deviation) เป็นการวัดการกระจายตัวแบบหนึ่งของข้อมูลภาพซึ่งจะบอกถึงคุณสมบัติภาพนั้นๆ ว่ามีการกระจายตัวของความเข้มอย่างไร จึงสามารถนำมาเป็นตัวแทนของภาพที่นำไปใช้ในการจดจำภาพได้ ข้อดีคือเป็นการลดจำนวนเวกเตอร์ข้อมูลเหมือนกับค่าเฉลี่ยของข้อมูลแต่ก็มีข้อด้อยเช่นเดียวกันคือภาพที่ต่างกันอาจจะมีค่าเบี่ยงเบนมาตรฐานที่เหมือนกันด้วยเช่นกัน ค่าเบี่ยงเบนมาตรฐานหาได้จากสมการที่ 2.10

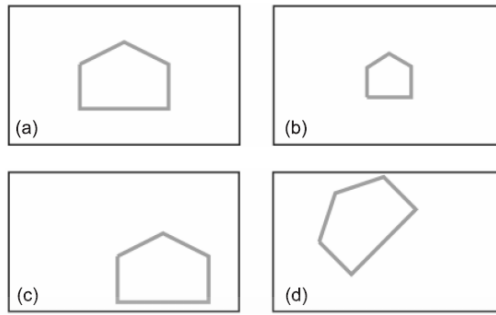
$$s = \left(\frac{1}{NM - 1} \sum_{i,j} (x_i^j - \mu)^2 \right)^{\frac{1}{2}} \quad (2.10)$$

2.5.2 วิธีการประมวลผลภาพเป็นวิธีแบบลำดับสอง (Second Order)

วิธีการประมวลผลภาพเป็นวิธีแบบลำดับสอง เป็นการสกัดคุณลักษณะที่อธิบายรูปร่างของวัตถุเนื่องจากรูปร่างเป็นคุณสมบัติพื้นฐานของวัตถุ ตัวอธิบายรูปร่างมีสองประเภท (1) ตัวอธิบายรูปร่างตามคอนทัวร์ (Contour-based Shape Descriptors) และ (2) ตัวอธิบายรูปร่างตามภูมิภาค (Region-based Shape Descriptors) วิธีการสกัดคุณลักษณะที่ไม่แปรปรวนไปตามการหมุนของภาพมีหลายวิธี เช่นวิธี Moment-Analysis, SIFT, Polar Fourier เป็นต้นซึ่งเป็นวิธีที่ถูกนำมาทดลองใช้ในการทำภารกิจคุณลักษณะในงานวิจัยนี้

2.5.2.1 Moment Analysis Descriptor

Moment Analysis หรือ MA ถูกนำเสนอโดย Hu ในปี 1962 [81] เป็นวิธีที่นิยมนำมาใช้สำหรับการสกัดคุณลักษณะของภาพในงานเกี่ยวกับ Machine Vision อย่างแพร่หลายเนื่องจากข้อดีเรื่องการไม่แปรปรวนไปตามระยะทางและทิศทางของภาพ ดังนั้นระบบการมองเห็นจึงมีประสิทธิภาพสูงเนื่องจากไม่เปลี่ยนแปลงไปตามระยะทางและทิศทางของการมองเห็นดังแสดงในรูปที่ 2.17



รูปที่ 2.17 ภาพวัตถุที่มีขนาดตำแหน่งและการหมุนแตกต่างกัน (A) วัตถุ 2 มิติ, (B) ขนาดเปลี่ยนแปลง, (C) ตำแหน่งเปลี่ยนแปลง, (D) ทิศทางเปลี่ยนแปลง

Hu ได้นำเสนอวิธีการคำนวณหาชุดของค่าคงที่พีชคณิตของภาพวัตถุ ค่า Moment ในปริภูมิ 2 มิติของชุดข้อมูลเมทริกขนาด $M \times M$ ซึ่งเป็นภาพในระดับความสว่างสีเทามีรูปแบบฟังก์ชันเป็น $f(x, y), (x, y = 0, \dots, M - 1)$ จะได้ว่า

$$m_{pq} = \sum_{x=0}^{x=M-1} \sum_{y=0}^{y=M-1} (x)^p (y)^q f(x, y) \quad p, q = 0, 1, 2, 3, \dots \quad (2.11)$$

โมเมนต์ $f(x, y)$ แทนโดยจำนวน (a, b) จะได้ว่า

$$\mu_{p,q} = \sum_x \sum_y (x + a)^p \cdot (y + b)^q f(x, y) \quad (2.12)$$

ดังนั้นศูนย์กลางของโมเมนต์ m'_{pq} หรือ μ_{pq} สามารถคำนวณได้จาก โดยการแทนค่า $a = -\bar{x}$ และ $b = -\bar{y}$ ดังนี้

$$\bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}} \quad \mu_{pq} = \sum_x \sum_y (x - \bar{x})^p \cdot (y - \bar{y})^q f(x, y) \quad (2.13)$$

เมื่อทำ Normalize ค่า Central Moment จะเป็น μ

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^r}, \quad r = \left[\frac{(p + q)}{2} \right] + 1 \quad (2.14)$$

Hu ได้กำหนดการคำนวณค่าโมเมนต์ไว้ 7 ค่า คำนวณโดยการทำให้ Normalize ค่า Central Moment ลำดับที่ 3 ซึ่งไม่แปรเปลี่ยนไปตามขนาด ตำแหน่ง การหมุน ค่าโมเมนต์ทั้ง 7 ค่าหาได้จาก

$$\begin{aligned}
M1 &= (\eta_{20} + \eta_{02}), \\
M2 &= (\eta_{20} + \eta_{02})^2 + 4\eta_{11}^2, \\
M3 &= (\eta_{30} + \eta_{12})^2 + (3\eta_{21} + \eta_{03})^2, \\
M4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2, \\
M5 &= (\eta_{30} + 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
&\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2], \\
M6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
&\quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}), \\
M7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
&\quad - (\eta_{30} + 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2].
\end{aligned} \tag{2.15}$$

จากนั้นก็นำค่าโมเมนต์ทั้ง 7 ค่าที่คำนวณได้มาเป็น Descriptor สำหรับอธิบายคุณลักษณะของภาพโดยการสร้างเป็นฮิสโตแกรม ขนาด 7 bin และให้ฮิสโตแกรมนี้เป็นตัวแทนของภาพสำหรับการวิเคราะห์ต่อไป

2.5.2.2 Scale Invariant Feature Transform Descriptor

Scale Invariant Feature Transform หรือวิธี SIFT [82] เป็นวิธีที่นำมาใช้ในการวิเคราะห์เพื่อหาจุดเด่นของรูปภาพด้วยการค้นหาและสร้าง Keypoint ของภาพ โดย Keypoint ที่สร้างขึ้นจะไม่ขึ้นกับมุมมองทั้งขนาดและทิศทางของวัตถุในภาพ Feature ที่ได้จากวัตถุลักษณะเดียวกันจะคล้ายคลึงกันแม้ขนาดและทิศทางจะแตกต่าง การทำ Feature Extraction ด้วยวิธี SIFT ประกอบไปด้วย 4 ขั้นตอนหลักคือ

1. การหาปริภูมิค่าในมิติขนาดและระยะทาง (Scale-space extrema detection)

การทำ Feature ด้วยการหาลักษณะเด่นของภาพ โดยการทำการเบลอ (Blur) ภาพด้วยฟังก์ชันเกาส์เซียน (Gaussian Function) ขั้นตอนแรกของการค้นหา Keypoint ของภาพที่มีคุณสมบัติไม่ขึ้นอยู่กับขนาดภาพ จะทำการค้นหาโดยการเปลี่ยนแปลงขนาดของภาพและทำการเบลอภาพ (Blur) ภาพด้วยฟังก์ชันเกาส์เซียน (Gaussian) ดังสมการที่ 2.16

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{2.16}$$

เมื่อ (x, y) คือพิกัดบนภาพ, L คือ ภาพที่กรองผ่านการเบลอด้วย Blur Filter, G คือตัวกรองแบบ Gaussian และ I คือภาพอินพุต ทั้งนี้ตัวกรองแบบ Gaussian แสดงดังสมการที่ 2.17

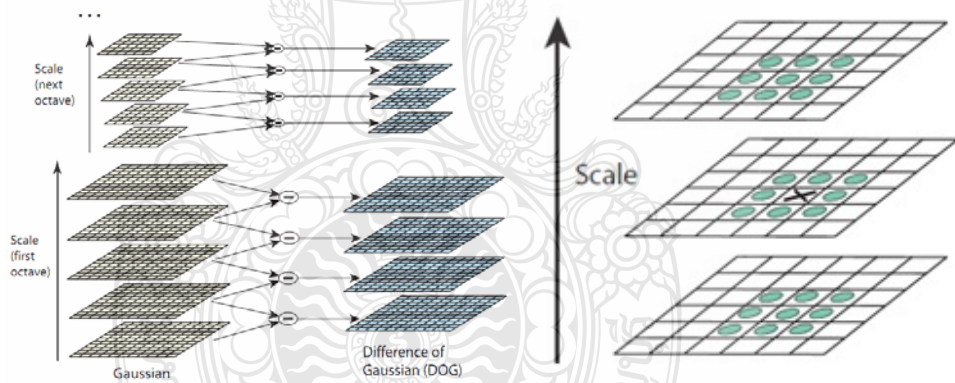
$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (2.17)$$

จากนั้นทำการหาผลต่างของแต่ละภาพ (Difference of Gaussian : DOG) โดยการนำภาพที่ได้จากการเบลอมาเปรียบเทียบกับภาพก่อนทำการเบลอตั้งสมการที่ 2.18

$$\begin{aligned} G(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad (2.18)$$

2. การกำหนดตำแหน่งจุดสนใจ (Keypoint Localization)

การหาตำแหน่งของ Keypoint เป็นการหาค่าจุดภาพ (Pixel) ที่มีค่าสูงสุดเมื่อทำการเปรียบเทียบกับจุดรอบข้าง ประกอบด้วย 8 จุดที่อยู่ติดกับจุดที่พิจารณา และ 9 จุดบนและล่างที่อยู่ติดกับจุดที่พิจารณา แสดงดังรูปที่ 2.18 ถ้าพบว่าจุดที่พิจารณามีค่าสูงสุดหรือต่ำสุดจุดนั้นจะถูกกำหนดให้เป็น Keypoint แต่ถ้าจุดที่ได้เป็นภาพที่ราบเรียบมีความสว่างและความมืดต่ำ (Low Contrast) หรือถ้าพิจารณาแล้วว่ามีลักษณะเป็นขอบ (Edge) Keypoint นั้นจะไม่ถูกนำมาใช้



รูปที่ 2.18 การหาตำแหน่งของ Keypoint [82]

3. การหาขนาดและทิศทางของเกรเดียน (Gradient) ให้กับ Keypoint (Orientation Assignment)

เมื่อได้ Keypoint แล้วขั้นตอนถัดมาจะเป็นการหาขนาด $m(x, y)$ และทิศทาง $\theta(x, y)$ ของ gradient ที่อยู่ในขอบเขตของ Keypoint เพื่อกำหนดขนาดและทิศทางของ Gradient ที่ Keypoint นั้นๆ โดยใช้สมการที่ 2.19 และ 2.20

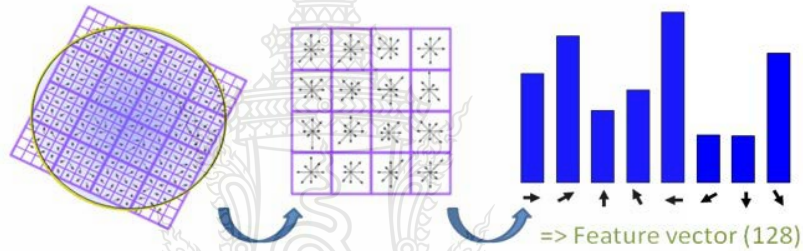
$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (2.19)$$

$$\theta(x, y) = \tan^{-1}((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))) \quad (2.20)$$

เมื่อ $m(x, y)$ คือ ขนาดของเกรเดียน และ คือ $\theta(x, y)$ ทิศทางของ Keypoint

4. การสร้างคำอธิบายลักษณะเด่น (Keypoints descriptor)

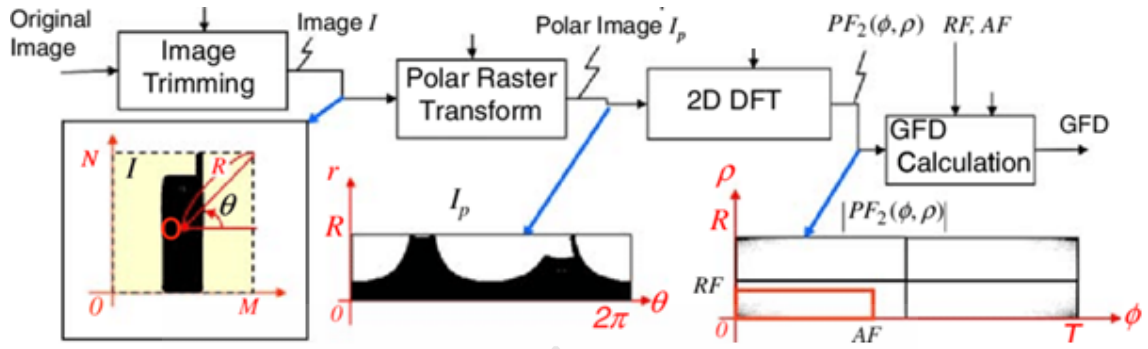
ในการสร้าง Keypoints descriptor ทำโดยการสร้าง Window ขนาด 16×16 รอบ Keypoint ให้ทำการแบ่งออกเป็น Windows ขนาด 4×4 ทั้งหมด 16 ชุด ผลการคำนวณหาค่าเกรเดียนและทิศทางบริเวณรอบๆ Keypoint แสดงดังรูป 2.17 และแบ่งออกเป็นพื้นที่ย่อยขนาด Window 4×4 แสดงดังรูป 2.19 แล้วสร้าง Histogram ของทิศทาง (Orientation Histograms) ขนาด 8 ข้อมูล ในแต่ละพื้นที่ย่อย ซึ่งจะมีแกน x ช่วงขององศา และแกน y เป็นขนาดของ Gradient ซึ่งจะได้เวกเตอร์ของแต่ละ Keypoint ที่มีขนาดเท่ากับ 128 ข้อมูล



รูปที่ 2.19 Histogram ของเวกเตอร์ของ SIFT [82]

2.5.2.3 Polar Fourier Descriptor

Polar Fourier หรือ PF [82] เป็นตัวแปลง 2 มิติอีกตัวที่ไม่ขึ้นกับการหมุนของภาพซึ่งมีขั้นตอนในการแปลง 3 ขั้นตอนหลักๆ ดังรูปที่ 2.20 ซึ่งประกอบด้วย การแปลงโพล่า 2 มิติ ซึ่งจะได้รูปเชิงขั้ว (Polar Image) จากนั้นขั้นตอนถัดไปก็ทำการแปลงฟูเรียร์ 2 มิติ (2-D Fourier Transform) จากนั้นก็ทำ Descriptor ซึ่งจะได้ฮิสโตแกรมที่เป็นตัวแทนของข้อมูลสำหรับทำวิเคราะห์ต่อไป ลำดับขั้นตอนในการแปลงทั้ง 3 ขั้นตอนเป็นดังนี้



รูปที่ 2.20 ลำดับขั้นตอนในการทำการสกัดเอกลักษณ์โดยวิธี Polar Fourier [83]

1. การแปลงภาพเป็นภาพเชิงขั้วโดยการนำภาพซึ่งมีคุณสมบัติเป็นข้อมูลแบบเมทริกซ์ภาพอินพุตที่นำเข้ามาแทนด้วย I ซึ่งแสดงดังสมการที่ 2.21

$$I = \{f(x, y) | 0 \leq x \leq W, 0 \leq y \leq H\} \quad (2.21)$$

จากนั้นภาพ I นี้จะถูกนำไปแปลงเชิงขั้ว ต่อไป ซึ่งการแปลงสามารถทำได้โดยสมการที่ 2.22

$$I_p = \{f(\theta, y) | 0 \leq r \leq R, 0 \leq \theta \leq 2\pi\} \quad (2.22)$$

$$x = r \cos \theta + W/2, y = r \sin \theta + H/2$$

เมื่อ R คือจำนวนที่แสดงถึงรัศมีของภาพ I_p จากเซ็นทรอยด์ $(W/2, H/2)$ ของภาพ I_p , ซึ่งจะได้สมการที่ 2.23 ว่า

$$R = \left\lceil \sqrt{(W/2)^2 + (H/2)^2} \right\rceil \quad (2.23)$$

2. การแปลงฟูเรียร์ เมื่อได้ภาพเชิงขั้วขั้นตอนถัดไปคือการแปลงฟูเรียร์แบบไม่ต่อเนื่องแบบ 2-D ถูกนำไปใช้กับ I_p โดย ผลการแปลงจะได้ดังสมการที่ 2.24

$$PF_2(\varphi, \rho) = \sum_r^{R-1} \sum_i^{T-1} f(r, \theta_i) \exp \left[-j2\pi \left(\frac{r}{R} \rho + \theta_i \varphi \right) \right] \quad (2.24)$$

$$0 \leq r \leq R, \theta_i = 2\pi i | T, 0 \leq \rho < R, 0 \leq \varphi < T$$

เมื่อ ρ และ φ คือ ρ_{th} ความถี่ในหน่วยเรเดียนและ φ_{th} คือความถี่ปกติ

3. การสร้างเวกเตอร์ของคุณลักษณะภาพ ซึ่งในการสร้างเวกเตอร์คุณลักษณะภาพจะใช้ Cartesian grids สามารถหาได้จากสมการ

$$PFD = \{PF_2(00,01), PF_2(00,01), \dots, PF_2(AF, RF)\} \quad (2.25)$$

2.6 วิธีการจัดคลัสเตอร์และคลาส

การเรียนรู้สำหรับ Machine Learning แบ่งเป็นการเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning) กับแบบมีผู้สอน (Supervised learning) การจัดกลุ่มคลัสเตอร์ (Clustering) เป็นการจัดการเรียนรู้แบบไม่มีผู้สอน หมายถึงการเรียนรู้จากตัวอย่างที่ไม่มีการกำหนดเป้าหมาย (Target) หรือฉลาก (Label) ของคลาส (Class) แต่ต้องการให้ Model จัดกลุ่มให้ โดยใช้ Attribute หรือ คุณลักษณะ เป็นตัวแจกแจง แนวทางในการจัดกลุ่มซึ่งวิธีการที่นิยมนำมาใช้งานคือวิธี K-Mean Clustering จึงถูกนำมาใช้สำหรับ Model BoF รูปแบบ Model ใช้ Algorithm แบบ Centroid models ส่วนการเรียนรู้แบบมีผู้สอนชุดข้อมูลที่นำเข้าสู่ระบบจะแบ่งเป็น 2 ชุดคือชุด Training และชุด Testing โดยข้อมูลชุดแรกจะเป็นข้อมูลชุดสอนให้ระบบเรียนรู้ ส่วนชุดที่สองเป็นชุดที่นำเข้าทดสอบการทำงาน วิธีที่ใช้ในการทำ Classification เช่น KNN SVM และ ANN เป็นต้น

2.6.1 K-mean Clustering

การจัดกลุ่มด้วยวิธี K-Mean Clustering [84] เป็นวิธีการจัดกลุ่มข้อมูลแบบ Unsupervised Learning หรือการเรียนรู้แบบไม่ต้องสอน K-means ทำหน้าที่แบ่งกลุ่ม แบบ Clustering ซึ่งการแบ่งกลุ่มในลักษณะนี้จะใช้พื้นฐานทางสถิติ จำนวนกลุ่มข้อมูล K ถูกกำหนดขึ้นมา และทำการจัดกลุ่มของเซตตัวอย่าง $\{x_1, x_2, x_3, \dots, x_n\}$ เมื่อ $x_i \in \mathcal{R}^d$ ข้อมูลการจัดกลุ่มข้อมูลวิธีการจัดกลุ่มจะทำการจัดข้อมูลที่คล้ายคลึงกันให้อยู่ในกลุ่มเดียวกัน ผลลัพธ์ของการจัดกลุ่มคือกลุ่มของตัวอย่างและเซนทรอยด์ (Centroid) ซึ่งมักเป็นตัวแทนของกลุ่มและมักอยู่ตรงกลาง สามารถสร้างฟังก์ชันวัตถุประสงค์สำหรับการจัดกลุ่มนี้ได้จากการหาค่าที่น้อยที่สุดของระยะห่างของตัวอย่างและ Centroid ของแต่ละกลุ่ม c_k ดังนี้

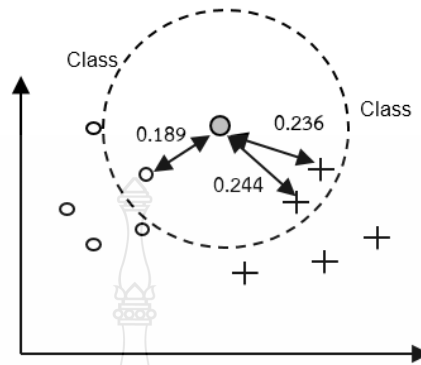
$$J(r, c) = \sum_{i=1}^n \sum_{k=1}^K r_{ik} \|x_i - c_k\|^2 \quad (2.26)$$

โดย $r_{ik} \in \{0,1\}$ เป็นตัวแปรที่บอกความเป็นสมาชิกของกลุ่มที่ k ของตัวอย่างที่ i เมื่อ x_i ถูกกำหนดให้อยู่ในกลุ่มที่ k จะมีค่าเป็น 1 และมีค่าเป็น 0 สำหรับกลุ่มอื่นๆ ที่ไม่ใช่กลุ่ม k

2.6.2 K-Nearest Neighbors

K-Nearest Neighbors หรือ KNN [84] เป็น Algorithm ในการจำแนกข้อมูล (Classification) ซึ่งเหมาะสมกับข้อมูลที่มีการกระจายตัวโดยการทำนายตัวอย่างจากปริภูมิตัวอย่างที่ใกล้เคียงที่สุด หรือจำแนกตัวอย่างโดยการวัดระยะทางเปรียบเทียบกับข้อมูลในชุด Training ระยะทางที่ได้จากการคำนวณเพื่อหาระยะห่างของตัวอย่างกับข้อมูลในชุด Training จะถูกจัดเรียงลำดับและหา Class ของชุดข้อมูลที่ใกล้กับตัวอย่างมากที่สุด โดยกำหนดจำนวนของข้อมูลที่ใกล้ที่สุดจำนวน K

ตัว ดังรูปตัวอย่างที่ 2.21 เปรียบเทียบระยะทางกับข้อมูลเพื่อนบ้านที่ $K=3$ ซึ่งพบว่าตัวอย่าง \bullet อยู่ใกล้กับกลุ่ม $+$ มากกว่า



รูปที่ 2.21 K-Nearest Neighbors ที่ $K=3$

วิธีการคำนวณระยะทางที่นิยมใช้มากที่สุดคือระยะทางแบบยูคลิด (Euclidean Distance) ซึ่งระยะทางแบบยูคลิดจากจุด $p(p_1, p_2, p_3, \dots, p_n)$ ไปจุด $q(q_1, q_2, q_3, \dots, q_n)$ ในปริภูมิ n มิติสามารถหาได้จาก

$$D_{Euclidean}(q,p) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (2.27)$$

เมื่อ p คือข้อมูลที่ต้องการจำแนกและ q คือข้อมูลในแต่ละกลุ่ม

2.6.3 Support Vector Machine

Support Vector Machine หรือ SVM [84] (Cortes et al., 1995) [85] เป็น Algorithm ในกลุ่มวิธีการเรียนรู้ของเครื่องแบบมีผู้สอนที่ได้รับความนิยมอย่างแพร่หลาย ในงานที่เกี่ยวข้องกับการจดจำรูปแบบตลอดจนการแก้ปัญหาการจัดกลุ่ม เหมาะกับการจัดกลุ่มข้อมูลที่มีขนาดของข้อมูลไม่ใหญ่มาก อาศัยหลักการหาเส้นแบ่งกลุ่มหรือไฮเปอร์เพลน (Hyperplane) เป็นสัมประสิทธิ์ของสมการเพื่อสร้างเส้นแบ่งกลุ่มชุดข้อมูลชุด Training เป็นชุดข้อมูลสอนให้ระบบเรียนรู้ และหาไฮเปอร์เพลนที่ใช้แยกคลาสของข้อมูลให้ได้จุดที่ดีที่สุด (Optimal hyperplane) ไฮเปอร์เพลนมีทั้งแบบเชิงเส้นและไม่เป็นเชิงเส้นโดยอาศัยการปรับค่าสมการด้วยเคอร์เนล (Kernel) โดยที่ Kernel Function มีอยู่หลายตัวที่นิยมใช้ Polynomial, RBF หรือ Sigmoid เป็นต้น หลักการของ SVM เป็นการสร้างเส้นแบ่งกลุ่มดังสมการที่ 2.28

$$D = \{(x_i, y_i); i = 1, 2, \dots, n\} \quad (2.28)$$

เมื่อ $x_i = (x_{i1}, x_{i2}, \dots, x_{in}) \in R^m$ และ $y_i \in \{1, -1\}$ โดย 1 คือข้อมูล Class ที่ 1 และ -1 คือข้อมูล Class ที่ 2 ซึ่งเป็นการกำหนดกลุ่มเป้าหมายให้ SVM โดยที่ SVM นั้นมุ่งเป้าเพื่อหาฟังก์ชันการตัดสินใจที่สามารถแบ่งแยกค่าที่ไม่ทราบได้ดังสมการที่ 2.29

$$f(X) = \text{sign} \left\{ \sum_{k=1}^{n_v} W_k \varphi_k(X) \varphi_k(X_k) + b \right\} \quad (2.29)$$

$$\varphi(X) = [\varphi_1(X_1), \varphi_2(X_2), \dots, \varphi_n(X_n)]^T \quad (2.30)$$

กลุ่มข้อมูลที่ไม่สามารถใช้ฟังก์ชันเชิงเส้นแบ่งกลุ่มได้นั้นเมื่อข้อมูล X จากสมการที่ 3 แบ่งแยกได้ด้วยสมการเส้นตรงไม่ได้จะถูกแปลงให้อยู่ในรูปแบบที่สามารถใช้สมการเส้นตรงแบ่งแยกได้โดยใช้เคอร์เนลฟังก์ชัน (Kernel Function) ดังสมการที่ 2.31

$$K(X, X_k) = \varphi(X) \varphi(X_k) \quad (2.31)$$

เมื่อ $\varphi(X)$ คือฟังก์ชันสำหรับแปลงข้อมูลที่ไม่เป็นเชิงเส้นให้เป็นข้อมูลที่อยู่ในรูปเชิงเส้นให้สามารถแบ่งแยกได้, W_k คือค่าน้ำหนักที่เชื่อมโยงจาก Feature Space ไปสู่ Output Space, b คือค่าไบอัส (Bias), X_k ซัพพอร์ตเวกเตอร์ โดย $k = 1, 2, \dots, n_v$, n_v จำนวนซัพพอร์ตเวกเตอร์

วิธีการที่ใช้ในการหาเส้นแบ่งที่ดีที่สุดคือการเพิ่มเส้นขอบ (Margin) ให้กับเส้นแบ่งทั้งสองข้างและสร้างเส้นขอบที่สัมผัสกับค่าข้อมูลใน Feature Space ที่ไกลที่สุดดังนั้นเส้นแบ่งที่มีเส้นขอบกว้างที่สุดจึงเป็นเส้นแบ่งที่ดีที่สุดและเรียกตำแหน่งการสัมผัสข้อมูลที่ไกลที่สุดจากการเพิ่มขอบนี้ว่า “ซัพพอร์ตเวกเตอร์” (Support Vector) แต่หากกลุ่มข้อมูลมีการปนเปกันจนในบางกรณีการแบ่งแยกกลุ่มไม่สามารถทำได้ถูกต้องโดยสมบูรณ์ ดังนั้นจึงต้องมีการกำหนดตัวแปรสำหรับยอมรับค่าความผิดพลาดโดยการเพิ่มตัวแปร ξ (Slack Variable) ดังสมการที่ 2.32 และ 2.33 ดังนี้

$$W^T X + b \geq y - \xi_i \quad \text{เมื่อ } y = 1 \quad (2.32)$$

$$W^T X + b \leq y + \xi_i \quad \text{เมื่อ } y = -1 \quad (2.33)$$

จากการกำหนดค่า $\xi_i > 0$ ทำให้โครงสร้างของซัพพอร์ตเวกเตอร์เมทซ์ขึ้นบรรลู่วัตถุประสงค์ใน 2 ส่วนคือการเพิ่มระยะแบ่งแยกให้มากที่สุด และลดข้อผิดพลาดในการทำนายให้ต่ำที่สุดดังสมการที่ 2.34

$$\text{Minimize } \frac{1}{2} \|W\|^2 + c \sum_{i=1}^N \xi_i \quad (2.34)$$

$$\text{เมื่อ } y_i (W^T \varphi(X) + b) + \xi_i - 1 \geq 0, \xi_i \geq 0, i = 1, 2, \dots, N$$

และมีเคอร์เนลสำหรับ SVM ฟังก์ชัน ที่นิยมใช้อยู่ 3 ชนิดด้วยกันคือ ฟังก์ชันโพลิโนเมียล (Polynomial) :

$$K(X_i, X_j) = (X_i^T X_j + r)^\gamma; \gamma > 0 \quad (2.35)$$

ฟังก์ชันเรเดียลเบสิส (Radial Basis Function-RBF) :

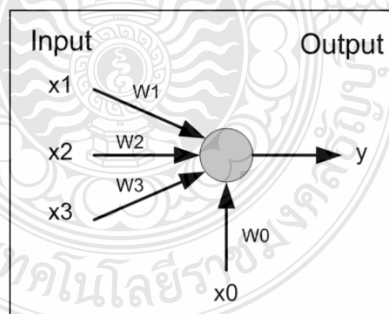
$$K(X_i, X_j) = \exp(-\gamma \|X_i - X_j\|^2); \gamma > 0 \quad (2.36)$$

ฟังก์ชันซิกมอยด์ (Sigmoid) :

$$K(X_i, X_j) = \tanh(\gamma X_i^T X_j - r) \quad (2.37)$$

2.6.4 ANN

วิธีโครงข่ายประสาทเทียม ANN [86, 87] เป็นกระบวนการ ทางคณิตศาสตร์ที่จำลองมาจากการทำงานของระบบสมองของมนุษย์เพื่อนำมาใช้ประโยชน์ในการตัดสินใจ การทำนาย การจำแนก และอื่นๆ รูปแบบของโครงข่ายประสาทเทียมพื้นฐานจะประกอบด้วยเซลล์ประสาท (Neuron) เพียง 1 เซลล์ ซึ่งเราเรียกว่าสถาปัตยกรรมเพอร์เซปตรอน (Perceptron Architecture) โดยที่อินพุตต่างๆ จะถูกส่งผ่านกลุ่มใยประสาท นำเข้า (Dendrites) เข้าสู่เซลล์ประสาท เพื่อทำการประมวลผลและส่งผลลัพธ์ออกมาทางแกนประสาท (Axon) โดยการประมวลผลจะใช้ผลรวมของผลคูณ (Sum of Product) ของข้อมูลอินพุตและค่าน้ำหนัก (Weight) ที่ใยประสาทนำเข้า ตัวอย่างของรูปแบบสถาปัตยกรรมเพอร์เซปตรอน แสดงดังภาพที่ 2.22 โดยข้อมูล x_1 ถึง x_3 คือข้อมูลอินพุต w_1 ถึง w_3 คือค่าน้ำหนัก และ y คือ เอาต์พุตที่ได้จากการประมวลผลของเซลล์ประสาท



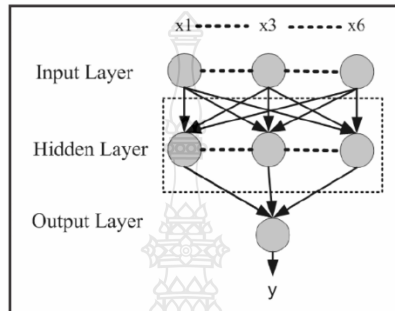
รูปที่ 2.22 โครงข่ายเพอร์เซปตรอน [86]

โครงข่ายเพอร์เซปตรอนเป็นพื้นฐานและเป็นต้นแบบของโครงข่ายประสาทเทียมชนิดอื่นๆ ในงานวิจัยนี้ได้ใช้ Multi-Layer Perceptron และใช้ขั้นตอนการเรียนรู้แบบแพร่ย้อนกลับ (Back Propagation Learning Algorithm) ซึ่งวิธีการนี้สามารถแก้ปัญหาที่มีความซับซ้อนได้อย่างมีประสิทธิภาพ สถาปัตยกรรมเพอร์เซปตรอนแบบหลายชั้นจะประกอบไปด้วยชั้นอินพุต (Input Layer)

ชั้นซ่อน (Hidden Layer) และ ชั้นเอาต์พุต (Output Layer) แสดงดังภาพที่ 2.23 ข้อมูลที่เข้ามาจะถูกส่งไปคำนวณในชั้นซ่อนเพื่อหาผลรวมของผลคูณของข้อมูลเข้าและค่าน้ำหนักแสดงดังสมการที่ 2.38

$$y = \sum_{i=0}^n x_i w_i \quad (2.38)$$

เมื่อ y คือค่าผลรวมของผลคูณข้อมูลเข้า x_i (x_i) และน้ำหนัก w_i , i คือจำนวนข้อมูลเข้าหรือจำนวนค่าน้ำหนัก



รูปที่ 2.23 โครงข่ายเพอร์เซปตรอนแบบหลายชั้น [86]

นำผลลัพธ์ที่ได้ไปคำนวณในฟังก์ชันการแปลงถ่ายทอดข้อมูล (Sigmoid Function) ดังสมการที่ 2.39

$$o = g(y) = \frac{1}{1 + e^{-y}} \quad (2.39)$$

จากผลลัพธ์ของชั้นซ่อนก็จะถูกส่งไปยังชั้นเอาต์พุตซึ่งในส่วนของชั้นเอาต์พุตจะมีการเปรียบเทียบกับค่าผลลัพธ์ที่ประมวลผลได้และผลลัพธ์เป้าหมายซึ่งถ้าได้ผลลัพธ์ที่ยอมรับได้ก็จะหยุดการปรับค่าน้ำหนัก แต่ถ้ายังไม่อยู่ในช่วงของผลลัพธ์ที่ยอมรับได้ ก็จะเข้าสู่กระบวนการเรียนรู้แบบแพร่ย้อนกลับ ซึ่งจะเป็กระบวนการปรับค่าน้ำหนักจนกว่าจะได้ค่าที่เหมาะสม โดยสามารถคำนวณได้จากสมการที่ 2.40 และ 2.41

$$\delta_k = O_k(1 - O_k)(T_k - O_k) \quad (2.40)$$

$$\delta_l = O_l(1 - O_l) \left(\sum_{k \in O} W_{k_l} \delta_k \right) \quad (2.41)$$

เมื่อ δ_k คือค่าความผิดพลาดที่คำนวณจากชั้นผลลัพธ์

δ_l คือค่าความผิดพลาดที่คำนวณจากชั้นซ่อน

T_k คือค่าผลลัพธ์เป้าหมาย

O_k คือค่าผลลัพธ์ที่ประมวลผลได้จากชั้นผลลัพธ์

O_l คือค่าผลลัพธ์ที่ประมวลผลได้จากชั้นซ่อน

W คือค่าน้ำหนัก

k และ l คือดัชนีของโหนดชั้นผลลัพธ์และชั้นซ่อน

ในกรณีค่าผลลัพธ์ที่ได้จากการประมวลผลและผลลัพธ์เป้าหมายยังมีความแตกต่างกันสูงจะกระทำการปรับค่าน้ำหนักเพื่อหาค่าน้ำหนักที่เหมาะสมกับงานการปรับค่าน้ำหนักแสดงดังสมการที่ 2.42 และ 2.43

$$w_i^{new} = w_i^{old} + \Delta w_i \quad (2.42)$$

$$\Delta w_i = \delta_i x_i \quad (2.43)$$

เมื่อ w_i^{new} คือค่าน้ำหนักใหม่ที่ได้จากการคำนวณ

w_i^{old} คือค่าน้ำหนักเก่า

Δw_i คืออัตราการเปลี่ยนแปลง

δ_i คือค่าความผิดพลาดของผลลัพธ์

x_i คือค่าข้อมูลชั้นนำเข้า

i คือจำนวนข้อมูลเข้าหรือจำนวนค่าน้ำหนัก

การทำงานของสถาปัตยกรรมเพอร์เซปตรอนแบบหลายชั้นและการเรียนรู้แบบแพร่ย้อนกลับ จะกระทำการฝึกสอน (Training) โดยการปรับค่าน้ำหนักไปจนกระทั่งได้ค่าความผิดพลาดที่น้อยที่สุดหรือได้ค่าความผิดพลาดที่ยอมรับได้เมื่อได้ค่าน้ำหนักที่เหมาะสมแล้วก็จะนำไปใช้ทดสอบ (Testing) และนำผลลัพธ์ที่ได้ทำการทดสอบไปคำนวณด้วยฟังก์ชันการแปลงถ่ายทอดข้อมูล (Threshold Function) เพื่อให้ได้คำตอบที่เป็นจริงหรือเท็จใช่หรือไม่ใช่ หรือ “0” หรือ “1” ดังสมการที่ 2.44

$$f(x) = \begin{cases} 1 & \text{if } x > T \\ 0 & \text{if } x < T \\ \text{Random} & \text{if } x = T \end{cases} \quad (2.44)$$

ค่าผลลัพธ์ที่ได้จากการประมวลผลจะเป็น 1 ในกรณีที่ค่าผลลัพธ์ x มีค่ามากกว่าค่า Threshold เปรียบเสมือนการตัดสินใจว่าใช่ ค่าผลลัพธ์จะเป็น 0 ในกรณีที่ค่าผลลัพธ์ x มีค่าน้อยกว่า Threshold เปรียบเสมือนการตัดสินใจว่าไม่ใช่และถ้าค่าผลลัพธ์ที่ได้เท่ากับค่า Threshold ให้ทำการสุ่มการตัดสินใจว่าจะเป็น “0” หรือ “1”

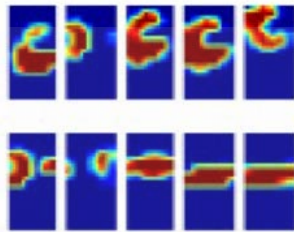
2.7 งานวิจัยที่เกี่ยวข้อง

งานวรรณกรรมทางด้านกราฟิกภาพมีอย่างแพร่หลาย แต่งานที่เกี่ยวข้องกับการจดจำภาพจากการสัมผัสเป็นงานที่นักวิจัยให้ความสนใจค่อนข้างน้อย เนื่องจากเซ็นเซอร์มีน้อยและการสร้างชุดข้อมูลในการทดสอบมีความยุ่งยากกว่างานวิจัยทางภาพแบบอื่นๆ ในช่วงต้นๆ งานวิจัยเน้นที่เรื่องการหา

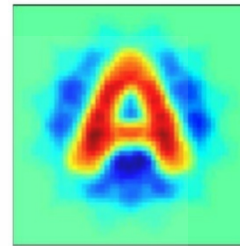
จุดสัมผัสเพื่อวัดแรงกดแบบจุดเดียวและการดึงข้อมูลจากจีโอเมทรีของวัตถุ โดยที่นักวิจัยทำออกแบบระบบให้เก็บข้อมูลจากจีโอเมทรีของวัตถุจากการจับวัตถุ และบางงานอาจใช้ Tactile Sensor แต่โดยทั่วไปแล้วเพียงเพื่อวัตถุประสงค์ในการบอกตำแหน่งจุดสัมผัสหรือประมาณการพื้นผิวที่เกี่ยวข้องกับจุดสัมผัส งานเกือบทั้งหมดต้องควบคุมตำแหน่งที่แน่นอนและควบคุมแรงสุทธิตั้งใช้ด้วย โดยทั่วไปตัวอย่างที่ได้จากแพทช์เล็กๆ ข้อมูลเหล่านี้จะถูกดึงเวกเตอร์คุณลักษณะออกมา และทำการกระจายข้อมูลบนเวกเตอร์คุณลักษณะเพื่อทำการวิเคราะห์ข้อมูล งานวิจัยด้านการจับ, สัมผัสและจดจำวัตถุนี้มีหลายกลุ่มซึ่งพอแบ่งได้เป็นทางด้านการพัฒนาตัวฮาร์ดแวร์ คือระบบกลไกแขนของหุ่นยนต์, ระบบเซ็นเซอร์ทั้งเซ็นเซอร์วัดตำแหน่ง, วัดมุมมองของนิ้ว, เซ็นเซอร์วัดแรงกด และเซ็นเซอร์อาร์เรย์สำหรับการสัมผัส มีทั้งพัฒนาตัวเซ็นเซอร์แบบต่างๆ อย่างหลากหลาย และงานที่ทำการพัฒนาระบบอ่านค่าจากเซ็นเซอร์ และงานที่พัฒนาควบคุมทั้งระบบกลไกของมือหุ่นยนต์และระบบเซ็นเซอร์ งานวิจัยอีกกลุ่มหนึ่งคือการพัฒนาโปรแกรมสำหรับจดจำ โดยมีทั้งงานที่อ่านค่าจากเซ็นเซอร์จากตัวฮาร์ดแวร์ชนิดต่างๆ ทั้งแบบที่อ่านค่าจีโอเมทรีวัตถุที่สัมผัส และการพัฒนาเพื่อวิเคราะห์ข้อมูลจากเซ็นเซอร์อาร์เรย์โดยตรง และการประมวลผลค่าที่ได้จากเซ็นเซอร์อาร์เรย์ในลักษณะของภาพจากการสัมผัส และงานที่พัฒนาทั้งฮาร์ดแวร์และซอฟต์แวร์ไปพร้อมกัน งานวิจัยกลุ่มแรกคือกลุ่มที่เน้นพัฒนาโปรแกรมที่ทำการวิเคราะห์ซึ่งข้อมูลที่น่าวิเคราะห์มีหลากหลายรูปแบบ ทั้งลักษณะที่มองข้อมูลเป็นภาพ และนำภาพขนาดเล็กมาทำการวิเคราะห์ จากตารางที่ 2.5 เป็นกลุ่มงานวิจัยที่พัฒนา Algorithm สำหรับ Tactile Sensor ซึ่งสามารถแบ่งออกได้เป็น 2 กลุ่มคือกลุ่มที่ใช้ Tactile Sensor ขนาดเล็กทดสอบกับวัตถุ ซึ่งหากเซ็นเซอร์มีขนาดเล็กกว่าวัตถุมากเราเรียก Tactile Image นี้ว่า Local Shape Tactile Image ส่วนงานที่ใช้ Tactile Sensor ขนาดใหญ่หรือขนาดใกล้เคียงกับวัตถุ เราเรียก Tactile Image นี้ว่า Global Tactile Image [62] ส่วนใหญ่ Local Tactile Image เรียกว่า Local shape Recognition [62] จะใช้วิธี BoF ในการวิเคราะห์ส่วน Global Shape Tactile Image เรียกว่า Global Shape Recognition จะใช้วิธี Machine Learning และ CNN ในการวิเคราะห์ [62]

$$\text{Tactile Sensor size} \ll \text{Object size} \quad (2.45)$$

$$\text{Tactile Sensor size} \approx \text{Object size} \quad (2.46)$$



(A) Local Shape Tactile Image



(B) Global Shape Tactile Image

รูปที่ 2.24 Local Shape Tactile Image และ Global Shape Tactile Image [19]

2.7.1 Local Shape Tactile Image Recognition

งานวรรณกรรมที่นำเสนอเกี่ยวกับการจำแนก Local Shape Tactile Image จากตารางที่ 2.5 ใช้แผ่นเซ็นเซอร์ขนาดเล็กสัมผัสส่วนหนึ่งของวัตถุรวบรวมข้อมูลมาในจำนวนหนึ่งแล้วใช้วิธี BoF ในการวิเคราะห์ รูปแบบการทดลองจะติด Tactile Sensor ที่ Griper แล้วจับวัตถุเมื่ออ่านค่าจาก Sensor จะได้ Local Shape Tactile Image มารายละเอียดของแต่ละงานเป็นดังนี้

ในปี 2009 Schneider และคณะ [63] ใช้ Tactile Sensor ขนาด 6×14 Pixels หรือขนาดภาพ $24.4 \times 51.4 \text{ mm}^2$ จำนวน 2 ชั้นติดเข้ากับ Gripper ทดสอบกับวัตถุ 21 ชนิด ใช้ Vectorize Feature และ Model BoF ซึ่งใช้ K-Means Clustering เป็นตัวจัด Cluster และใช้ Bayes เป็นตัวจัด Class ให้ความถูกต้อง 84.6% เมื่อจับวัตถุ 10 ครั้ง

ในปี 2011 Pezzementi และคณะ [66] ใช้การ Simulation Tactile Sensor ที่มี Resolution เท่ากับ 6×6 Pixels และ วัตถุจำนวน 10 ชนิดเพื่อทดสอบ ใช้วิธี BoF ทดสอบกับ Feature 6 เทคนิค ประกอบด้วย Vectorize, SIFT, MR-8, Moment-Normalized, Polar Fourier และ Moment-Normalized Translation-Invariant และใช้ PCA เป็นตัวลดมิติข้อมูล ใช้ K-Means เป็นตัวจัด Clusters และใช้ Descriptor Comparison เป็นตัวจัด Class เมื่อทำการทดลองเปรียบเทียบประสิทธิภาพโดยการปรับค่าตัวแปรหลากหลาย พบว่าเทคนิค Polar Fourier ให้ผลดีที่สุด จากนั้นได้ทดสอบกับ เซ็นเซอร์ PSS DigiTacts เป็น Tactile Sensor สำหรับทำการวัดตัวอย่าง มีขนาดเซ็นเซอร์ ประมาณ 1.2 cm^2 และ Resolution เท่ากับ 6×6 Pixels ทดสอบกับวัตถุที่เป็นตัวอักษร 5 ตัว พบว่า ในการจำลองการทำงาน Descriptor แบบ Moment-Normalized Translation-Invariant ให้ผลดี

สุด แต่ในการทดลองจริงพบว่า Polar Fourier ให้ผลที่ดีที่สุดมีความถูกต้อง 90% ที่การสัมผัสตัวอย่าง 60 ครั้ง

ในปี 2015 Luo และคณะ [19] ใช้ Tactile Sensor ขนาด 6×14 Pixels หรือขนาด ภายภาพ $24 \times 51 \text{ mm}^2$ ติดเข้ากับปลายแขนหุ่นยนต์ ทดสอบกับวัตถุ 18 ชนิด ใช้ SIFT Feature และ Model BoF ซึ่งใช้ K-Means Clustering เป็นตัวจัด Cluster และใช้ KNN เป็นตัวจัด Class ให้ความถูกต้อง 89.98% เมื่อจับวัตถุ 15 ครั้ง

ในปี 2016 Luo และคณะ [72] ใช้ Tactile Sensor ขนาด 6×14 Pixels หรือขนาด ภายภาพ $24 \times 51 \text{ mm}^2$ ติดเข้ากับปลายแขนหุ่นยนต์ ทดสอบกับวัตถุ 20 ชนิด ใช้วิธี Iterative Closest Labeled Point (iCLAP) ร่วมกับ Feature Zernike moment ในการวิเคราะห์ผลและจัด Class ให้ความถูกต้อง 85.36% เมื่อจับวัตถุ 20 ครั้ง

ในปี 2017 Falco และคณะ [73] ใช้ Tactile Sensor ขนาด 6×6 Pixels หรือขนาด ภายภาพ $50 \times 50 \text{ mm}^2$ ติดเข้ากับแขนปลายหุ่นยนต์อุตสาหกรรม ทดสอบกับวัตถุ 15 ชนิด ใช้วิธี KNN ร่วมกับ Feature Cross-Modal Ensemble of Shape Functions (CMESF) ในการวิเคราะห์ผล และจัด Class ให้ความถูกต้อง 94.67% เมื่อจับวัตถุ 30 ครั้ง

2.7.2 Global Shape Tactile Image Recognition

งานวรรณกรรมที่นำเสนอเกี่ยวกับการจำแนก Global Shape Tactile Image จากตารางที่ 2.5 ใช้แผ่นเซ็นเซอร์ขนาดใหญ่หรือใกล้เคียงกับวัตถุสัมผัสวัตถุซึ่งทำให้ได้ข้อมูลที่มีสัดส่วนค่อนข้างสูง เมื่อเทียบกับขนาดวัตถุ รูปแบบการทดลองจะติด Tactile Sensor ที่แขนหรือมือของหุ่นยนต์ แล้วจับวัตถุเมื่ออ่านค่าจาก Sensor จะได้ Global Shape Tactile Image จากนั้นใช้วิธีวิเคราะห์ด้วยวิธี Machine Learning หรือ CNN รายละเอียดของแต่ละงานเป็นดังนี้

ในปี 2010 Gorges และคณะ [64] ใช้ Tactile Sensor ขนาด 4×7 Pixels จำนวน 9 ชั้น ติดเข้ากับมือของหุ่นยนต์ จับวัตถุ 7 ชนิด ใช้ Moment Analysis Feature และ SOM เป็นตัวจัด Class ให้ความถูกต้องประมาณ 80%

ในปี 2011 Drimus และคณะ [65] สร้าง Tactile Sensor ขนาด 8×8 Pixels จำนวน 2 ชั้นติดเข้ากับ Gripper ของหุ่นยนต์ จับวัตถุ 10 ชนิด ใช้ Statistic Feature (Mean, SD) และ KNN เป็นตัวจัด Class ให้ความถูกต้องประมาณ 92.0%

ในปี 2012 Bhattacharjee และคณะ [67] ได้ติด Tactile Sensor ขนาด 24×16 Pixels ที่ แขนของหุ่นยนต์ ใช้การเคลื่อนที่ของแขนหุ่นยนต์เคลื่อนที่สัมผัสวัตถุจำนวน 4 Class และ 18 Class

ใช้ Statistic Feature (Maximum Vector) และ KNN เป็นตัวจัด Class ให้ความถูกต้องที่ 81.40% (4 Class) และ 72.14% (18 Class)

ในปี 2012 Liu และคณะ [10] ได้ติดตั้ง Tactile Sensor แบบเต็มฝ่ามือให้กับมือหุ่นยนต์ Shadow Hand เซ็นเซอร์มีหลายขนาดกระจายตัวตามนิ้วมือ และฝ่ามือ ทดสอบจับวัตถุ 4 Class และใช้วิธี 3 x 3 Segmentation มาผ่าน Filter แล้วนำมาจัดเรียงเป็น Vector Input ใช้ ANN ในการจัด Class ให้ความถูกต้องประมาณ 91.0%

ในปี 2013 Datta และคณะ [68] ได้ทำการทดลองโดยนำวัตถุจำนวน 10 Class มากดบน Tactile Sensor ขนาด 32 x 32 Pixels ใช้วิธี Haar Wavelet สำหรับทำ Feature Extraction และใช้ KNN LDA และ Naïve Bayes ในการจัด Class พบว่าวิธี KNN ให้ค่าความถูกต้องมากที่สุดอยู่ที่ 84.0%

ในปี 2014 [57] Liu และคณะ [57] ได้ทำการสร้าง Tactile Sensor ขนาด 10 x 10 Pixels หรือ 20 x 20 mm^2 และทดลองนำวัตถุมากดลงบนเซ็นเซอร์ จำนวน 12 Class และใช้ Kernel PCA, FD และ GE สำหรับทำ Feature Extraction และใช้ MKL-SVM ในการจัด Class พบว่าผลที่ได้ให้ค่าความถูกต้องมากที่สุดอยู่ที่ 85.54%

ในปี 2015 Cretu และคณะ [69] ได้ใช้ Tactile Sensor ขนาด 16 x 16 Pixels ติดตั้งบนมือของหุ่นยนต์และจับวัตถุที่เป็นรูปตัวเลขและตัวอักษร จำนวน 10 Class และ 25 Class นำ Data Vector เข้าทดสอบด้วย ANN พบว่า ให้ค่าความถูกต้องมากที่สุดอยู่ที่ 98.6% (10 Class) และ 96% (25 Class)

ในปี 2015 Schmitz และคณะ [70] ได้ติดตั้ง Tactile Sensor กระจายตัวอยู่บนมือของหุ่นยนต์ TWENDY-ONE ทั้งในส่วนนิ้วมือและฝ่ามือรวมทั้งหมด 241 จุด ทดลองจับวัตถุ 20 Class ใช้ Vector ของข้อมูลทั้ง 241 ข้อมูลเข้าทดสอบด้วย DNN ให้ความถูกต้องในการจดจำวัตถุ 91.1%

ในปี 2016 Liu และคณะ [71] ใช้ BarrettHand ซึ่งติดตั้ง Tactile sensor ไว้ที่นิ้วมือ 3 นิ้ว ขนาด 8 x 3 Pixels นิ้วละจุด และที่ฝ่ามือขนาด 7 x 4 Pixels อีกหนึ่งจุด รวมเป็นทั้งหมด 4 จุดได้ทำการทดลองจับวัตถุ 10 Class และใช้วิธีสร้าง Data Vector จากข้อมูลที่ได้จากเซ็นเซอร์ที่เกิดขึ้นตามเวลาที่เปลี่ยนแปลงไปขณะจับวัตถุ และใช้วิธีการจัด Class เปรียบเทียบ 4 วิธีคือ KNN, SKSC, CKSC, and JKSC พบว่า JKSC ให้ค่าความถูกต้องสูงที่สุดอยู่ที่ 94.0 %

ในปี 2017 Albini และคณะ [74] ใช้ Tactile Sensor ขนาด 768 pixels ติดที่แขนของหุ่นยนต์ Baxter ทดสอบกับวัตถุ 2 Class โดยทำการเปรียบเทียบผลระหว่าง BoF และ Alexnet-DCNN พบว่า AlexNet DCNN ให้ความถูกต้องมากกว่าอยู่ที่ 98.3%

ในปี 2017 Gandarias และคณะ [75] ใช้ Tactile Sensor ขนาด 28 x 50 Pixels ติดที่ปลายแขนของหุ่นยนต์ AUBO Our-i5 ทดสอบกับวัตถุ 8 Class โดยทำการเปรียบเทียบผลระหว่าง BoF และ Alexnet-DCNN พบว่า AlexNet DCNN ให้ความถูกต้องมากกว่าอยู่ที่ 91.6%

ในปี 2019 Gandarias และคณะ [76] ใช้ Tactile Sensor ขนาด 28 x 50 Pixels ติดที่ปลายแขนของหุ่นยนต์ AUBO Our-i5 ทดสอบกับวัตถุ 22 Class โดยทำการเปรียบเทียบ DCNN 11 Model พบว่า ResNet DCNN ให้ความถูกต้องมากกว่าอยู่ที่ 95.36%

2.8 การพัฒนาในงานวิจัยนี้

จากในหัวข้อที่ 2.7 พบว่างานวิจัยที่นำเสนอการพัฒนา Algorithm สำหรับการจดจำวัตถุจากการสัมผัส ที่นำเสนอการประมวลผลข้อมูลที่ได้จาก Tactile Sensor Array มีการพัฒนาอย่างต่อเนื่อง แต่งานที่พัฒนา Experiment บน Humanoid Robot Hand ร่วมกับการใช้วิเคราะห์ด้วยวิธี Tactile Image Recognition มีน้อยมาก [10, 64, 69, 70] ดังนั้นในวรรณกรรมที่เกี่ยวข้องส่วนใหญ่เป็นงานที่ทดลองลักษณะอื่นเช่นการทำ Simulation [66] หรือทดลองบน Experiment ที่เป็น Griper [63, 65, 71], Industrial Robot [19, 72, 73, 75, 76] ที่แขนของหุ่นยนต์ [14, 74] หรือกดวัตถุลงบนเซ็นเซอร์โดยตรง [57, 68] เพื่อหลีกเลี่ยงความยุ่งยาก และข้อจำกัดต่างๆ ของ Humanoid Robot ดังนั้นผลที่ได้จึงเป็นผลเบื้องต้นที่ยังไม่ได้นำมาประยุกต์ใช้กับ Humanoid Robot Hand จริง

จากความแตกต่างของขนาด Tactile Sensor จึงเกิดลักษณะการวิเคราะห์ 2 รูปแบบตามขนาดของ Tactile Sensor คือหากเซ็นเซอร์มีขนาดเล็กกว่าวัตถุมากเราเรียก Tactile Image นี้ว่า Local Shape Tactile Image ส่วนงานที่ใช้ Tactile Sensor ขนาดใหญ่หรือขนาดใกล้เคียงกับวัตถุ เราเรียก Tactile Image นี้ว่า Global Tactile Image [62] ส่วนใหญ่ Local Tactile Image เรียกว่า Local Shape Recognition [62] จะใช้วิธี BoF ในการวิเคราะห์ส่วน Global Shape Tactile Image เรียกว่า Global Shape Recognition จะใช้วิธี Machine Learning และ CNN ในการวิเคราะห์ ซึ่งจากงานก่อนหน้าสามารถสรุปได้ว่างานที่ใช้ BoF สามารถทำงานได้ดีกับ Sensor ที่มี Physical Size ขนาดเล็กและมีจำนวน Pixel น้อย หรือเป็นแบบ Low Resolution Tactile Sensor (36 Pixels [66], 84 Pixels [19, 63, 72]) และเนื่องจากมี Physical Size ขนาดเล็ก จึงเหมาะกับการติดตั้งที่นิ้วมือหรือ Griper ส่วนวิธี Global Shape Tactile Image ซึ่งใช้วิธี Machine Learning และ DCNN จะสามารถทำงานได้ดี กับ Sensor ที่มี Physical Size ขนาดใหญ่ (ขนาดใหญ่สามารถครอบคลุม Global Shape ของวัตถุได้มากกว่า) และด้วยขนาดที่ใหญ่จึงมักเป็น High Resolution Tactile Sensor (768 Pixels [74], 1400 Pixels [75, 76]), แต่ในงานวรรณกรรมที่ผ่านมา เซ็นเซอร์ที่มีขนาดใหญ่เกินไปจึงไม่เหมาะสม

กับการติดตั้งให้กับ Humanoid Robot Hand และใน Experiment ของงานก็ไม่ได้ติดตั้งบน Humanoid Robot Hand โดยเล็งไปที่ติดตั้งที่ Robot Arm [74] และ Industrial Robot [75, 76] ดังนั้นในการพัฒนาในงานวิจัยนี้

1) ได้นำเสนอการพัฒนา Tactile Sensor Array สำหรับ Humanoid Robot เป็นแบบ Full Hand แบบใหม่ที่มีโครงสร้างแบบ กระจายตัวตามจุดสัมผัสต่างๆ บนฝ่ามือ โดยทำการติดตั้งบนถ้งมือ สำหรับสวมใส่ให้กับ Humanoid Robot ซึ่งมีประสิทธิภาพสูงกว่า Sensor ที่ติดตั้งเพียงจุดเดียวที่ปลาย นิ้วหรือฝ่ามือ [6, 7, 8, 13] และมี Resolution สูงกว่า Sensor แบบจุดเดียว

2) ทดสอบการจดจำวัตถุจากการจับวัตถุในมือ โดยทดสอบกับวิธี Local Shape Recognition เปรียบเทียบกับวิธี Global Shape Recognition

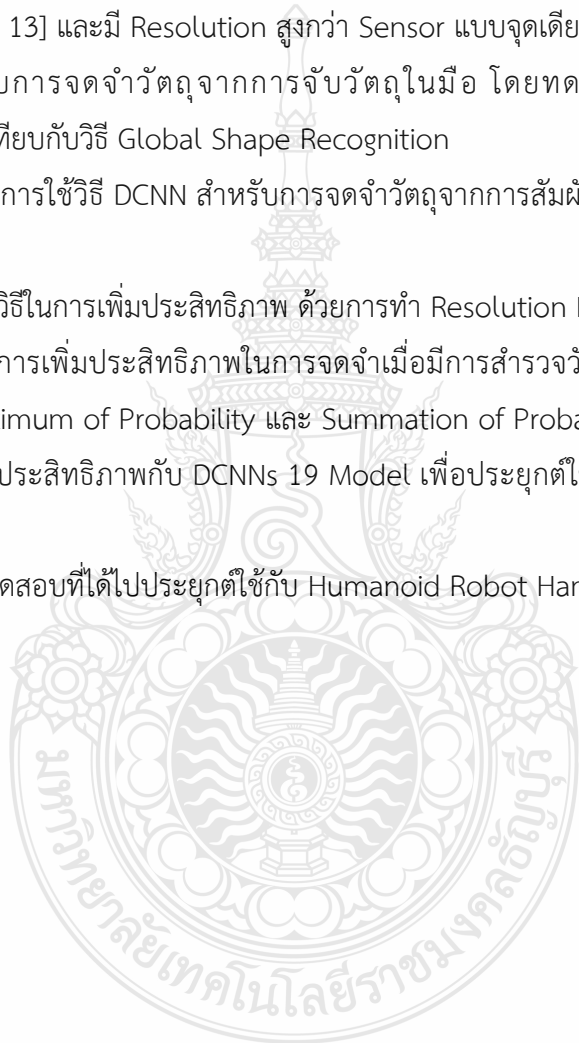
3) นำเสนอการใช้วิธี DCNN สำหรับการจดจำวัตถุจากการสัมผัสที่ได้จาก Tactile Sensor Array แบบใหม่นี้

4) นำเสนอวิธีในการเพิ่มประสิทธิภาพ ด้วยการทำ Resolution Enhance

5) ทดสอบการเพิ่มประสิทธิภาพในการจดจำเมื่อมีการสำรวจวัตถุมากขึ้นด้วยการจับวัตถุ หลายครั้ง โดยวิธี Maximum of Probability และ Summation of Probability

6) ทดสอบประสิทธิภาพกับ DCNNs 19 Model เพื่อประยุกต์ใช้ Model ที่มีประสิทธิภาพ สูงสุด และ

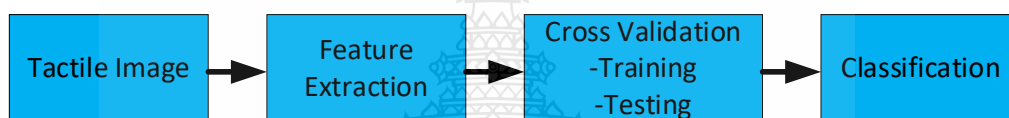
7) นำการทดสอบที่ได้ไปประยุกต์ใช้กับ Humanoid Robot Hand ได้จริง



บทที่ 3

วิธีดำเนินการวิจัย

ในงานวิจัยนี้ศึกษาการจดจำวัตถุจากการสัมผัส การดำเนินงานของระบบประกอบด้วย 4 ขั้นตอนหลัก ซึ่งเป็นระบบการทำงานของ การจดจำวัตถุจากการสัมผัสดังรูปที่ 3.1 ลำดับการดำเนินงาน ประกอบด้วย 1) ส่วนของ Tactile Image ที่นำเข้าสู่การทดสอบหรือเรียกว่าชุดข้อมูลหรือ Data Set, 2) ส่วนของการทำ Feature Extraction หรือการสกัดคุณลักษณะของภาพ, 3) ส่วน Cross Validation ซึ่งทำหน้าที่ Training และ Testing ชุดข้อมูล และ ส่วนสุดท้ายคือ 4) ผลที่ได้จากการทำ Classification เป็นส่วนของการทำนายผลการจดจำ แต่อย่างไรก็ตามกระบวนการทำงานจริงในแต่ละส่วนจะประกอบไปด้วยขั้นตอนย่อยหลายขั้นตอนดังนี้



รูปที่ 3.1 ระบบการทำงานของ การจดจำวัตถุจากการสัมผัส

1. Tactile Image

ในการทดลองจำเป็นต้องสร้าง Data Set ขึ้นมาใช้ ดังนั้นการสร้างชุดข้อมูลดังกล่าว ประกอบด้วยองค์ประกอบที่สำคัญดังนี้

- 1.1 วัตถุและวิธีการสัมผัสวัตถุหรือ Experiment ในการสัมผัสวัตถุ
- 1.2 เซ็นเซอร์ที่ใช้ในการสัมผัสวัตถุ
- 1.3 วงจรที่ใช้ในการอ่านค่าจากเซ็นเซอร์
- 1.4 โปรแกรมที่ใช้ในการบันทึกภาพ

2. Feature Extraction

- 2.1 การทำ Pre-Image Processing
- 2.2 การทำ Descriptor

3. การทำ Cross Validation

- 3.1 วิธีการที่ใช้ในการทำ Cross Validation
- 3.2 Training Set
- 3.3 Testing Set

4. การรายงานผล Classification

ในงานวิจัยนี้มีแนวคิดที่จะให้ระบบหุ่นยนต์และคอมพิวเตอร์มาทำงานร่วมกันเพื่อให้มีความสามารถในการเรียนรู้และจดจำเลียนแบบมนุษย์ เป็นระบบปัญญาประดิษฐ์ โดยจะทำการพัฒนาระบบการเรียนรู้และจดจำวัตถุโดยใช้ระบบประสาทสัมผัสเทียมขึ้น ซึ่งทำหน้าที่เป็นประสาทสัมผัสรับแรงสัมผัสและแรงกดของคล้ายกับมือมนุษย์ จากนั้นก็ส่งข้อมูลเข้าคอมพิวเตอร์ทำการวิเคราะห์ ซึ่งในส่วนนี้เปรียบเสมือนสมองของมนุษย์ที่ทำหน้าที่ตีความจากข้อมูลที่ส่งมาจากประสาทสัมผัส อย่างไรก็ตามได้มีงานวิจัยก่อนหน้าที่ได้ทำการทดสอบการทำงานนี้มาแล้ว ซึ่งเป็นการจดจำวัตถุจากการทดสอบกดวัตถุบน Tactile Sensor และใช้วิธีการวิเคราะห์การจดจำแบบต่างๆ ดังนั้นในงานนี้จึงได้ทำการทดลองเพื่อพิสูจน์สมมติฐานที่ว่า การออกแบบเซ็นเซอร์แบบ Glove Tactile Sensor มีประสิทธิภาพในการจดจำสูงกว่าเซ็นเซอร์แบบแผ่นเดียว ดังนั้นในงานนี้ได้ออกแบบการทดลองไว้ทั้งหมด 5 การทดลอง ดังนี้

1. การทดสอบสมมติฐานเบื้องต้น
2. Tactile Sensor การจดจำวัตถุจาก Tactile Image
3. การออกแบบเซ็นเซอร์แบบใหม่ และการทดสอบคุณสมบัติของเซ็นเซอร์
4. การจดจำวัตถุจากการสัมผัสสำหรับมือหุ่นยนต์
 - 4.1 การทดลองอัลกอริทึมแบบ BoF สำหรับมือหุ่นยนต์
 - 4.2 การทดลองอัลกอริทึมแบบ DCNN สำหรับมือหุ่นยนต์
5. การจดจำวัตถุจากเซ็นเซอร์แบบ Glove Tactile Sensor
 - 5.1 การทดลองอัลกอริทึมแบบ BoF สำหรับเซ็นเซอร์แบบ Glove Tactile Sensor
 - 5.2 การทดลองอัลกอริทึมแบบ DCNN สำหรับเซ็นเซอร์แบบ Glove Tactile Sensor

3.1 การทดสอบสมมติฐานเบื้องต้น

ในหัวข้อนี้มีวัตถุประสงค์เพื่อศึกษาหลักการการทำงานของระบบสัมผัสของมือมนุษย์เป็นพื้นฐานในการออกแบบระบบผิวหนังเทียมของมือหุ่นยนต์ เพื่อให้สามารถจดจำวัตถุจากการสัมผัส ทำให้สามารถทำงานได้ใกล้เคียงกับสรีรวิทยาของมนุษย์มากที่สุด การศึกษามือมนุษย์ในการจดจำวัตถุเป็นการทำงานร่วมกันของผิวหนังที่มีประสาทสัมผัสและส่งไปให้สมองตีความ ดังนั้นลักษณะของข้อมูลที่ได้จากการสัมผัสจากมือหากนำข้อมูลนี้ไปตีความด้วยคอมพิวเตอร์จะสามารถทำให้คอมพิวเตอร์รู้ชนิดของวัตถุได้

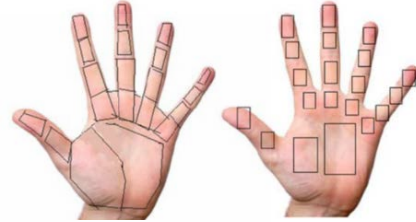
3.1.1 วิธีการทดสอบการจดจำวัตถุจาก Tactile Image

การจับวัตถุของมือมนุษย์มีลักษณะดังรูปที่ 3.2 (A) ส่วนของผิวหนังฝ่ามือที่สัมผัสวัตถุสามารถแบ่งได้เป็นส่วนต่างๆ ตามฝ่ามือดังรูปที่ 3.2 (B) เมื่อมีการจับวัตถุ ผิวหนังที่ฝ่ามือและนิ้วจะ

สัมผัสกับวัตถุ การสัมผัสนี้ทำให้ประสาทสัมผัสที่อยู่บนฝ่ามือได้รับการตอบสนอง ส่งผลให้มีการส่งกระแสประสาทไปยังสมอง ลักษณะของข้อมูลที่ถูกส่งไปจะถูกตีความด้วยสมอง ซึ่งกลไกการตีความของสมองมีความสลับซับซ้อนอันเกิดจากการเรียนรู้และสะสมประสบการณ์มายาวนาน แต่หากนำข้อมูลมาตีความด้วยคอมพิวเตอร์ มีความจำเป็นต้องลดความซับซ้อนดังกล่าวอย่างหลีกเลี่ยงไม่ได้ ดังนั้นการลดมิติข้อมูลเราสามารถลดลงให้เป็นไปตามส่วนต่างๆ ของฝ่ามือ ดังรูปที่ 3.2 (B) ได้



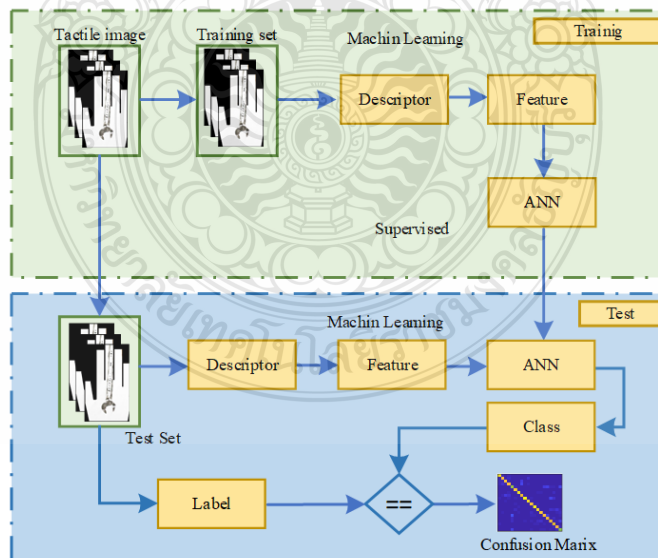
(A) การจับวัตถุของมือมนุษย์



(B) จุดสัมผัสและส่วนของฝ่ามือมนุษย์

รูปที่ 3.2 การจับวัตถุและจุดสัมผัสของฝ่ามือมนุษย์

เมื่อสัมผัสกับวัตถุใดๆ จะมีแรงมากระทำกับผิวหนังในตำแหน่งที่สัมผัส เมื่อเราทำการแบ่งส่วนของผิวหนังของมือจะสามารถแบ่งได้เป็นส่วนตามรูป 3.2 ดังนั้นหากทำการวิเคราะห์ผลเราสามารถแบ่งส่วนของฝ่ามือเป็นส่วนๆ เพื่อนำข้อมูลนี้ไปวิเคราะห์ได้ ในการทดลองนี้จึงใช้ลักษณะข้อมูลที่ได้จากประสาทสัมผัสที่ผิวหนังของมือ และนำข้อมูลที่ได้เข้าไปวิเคราะห์ด้วยคอมพิวเตอร์ ลำดับในการดำเนินงานแสดงดังรูปที่ 3.3 ซึ่งมีลำดับดังนี้



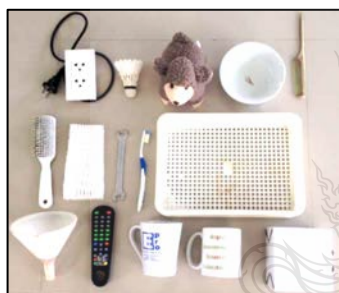
รูปที่ 3.3 การดำเนินงานการจดจำวัตถุจากการสัมผัสของฝ่ามือมนุษย์ร่วมกับ ANN

3.1.2 วัตถุและภาพจากการสัมผัส Object และ Tactile Image

การสร้างภาพ Tactile Image ทำการทดลองโดยใช้มือมนุษย์จริงสวมถุงมือกระดาษและสัมผัสวัตถุที่ระบายสีสำหรับประทับตราเอาไว้ ลำดับการสร้าง Tactile Image สำหรับการทดลองนี้เป็นดังนี้

1. สร้างถุงมือกระดาษ
2. ระบายสีที่วัตถุ
3. ติดถุงมือกระดาษและสัมผัสวัตถุ
4. สแกนภาพสัมผัสที่ได้

เพื่อสร้างชุดข้อมูลในการทดสอบนี้ใช้วัตถุจำนวน 15 ชิ้น ประกอบด้วย ปลั๊กไฟฟ้า, ลูกแบดมินตัน, ตู๊กตา, ขาม, ทวี, แปรง, โฟมกันกระแทก, ประแจ, แปรงสีฟัน, ตะกร้า, กรวย, รีโมทคอนโทรล, แก้วกาแฟ1, แก้วกาแฟ2 และกล่องกระดาษ เมื่อสัมผัสถุงมือกับวัตถุ สีที่วัตถุจะติดมือมามีลักษณะเป็นภาพตามลักษณะของวัตถุ ดังรูปที่ 3.4



(A) วัตถุที่ใช้ในการทดสอบ



(B) ประแจ



(C) Tactile Image

รูปที่ 3.4 วัตถุและ Tactile Image

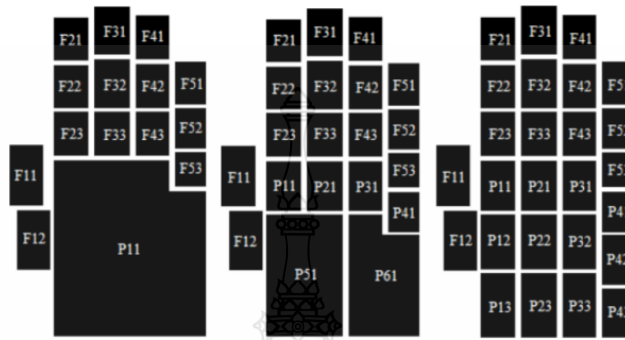
โดยการสัมผัสวัตถุในตำแหน่งต่างๆ เพื่อเก็บภาพโดยรอบวัตถุจากวัตถุ 15 ชิ้น ทำการทดสอบ 40 ภาพต่อวัตถุซึ่งจะได้ภาพทั้งหมด 600 ภาพ และทำการสแกนภาพที่ได้จากการจับวัตถุด้วยเครื่องสแกนภาพ จากนั้นบันทึกลงคอมพิวเตอร์เพื่อเป็นชุดข้อมูลสำหรับการทดสอบ

3.1.3 Descriptor

การทำ Descriptor เพื่อดึง Feature จากภาพ ในการทดลองนี้ได้ทำการ Descriptor ด้วยการแบ่งภาพ และสร้าง Data Vector เพื่อเป็นตัวแทนของข้อมูลด้วยการหาผลรวมของข้อมูลในตำแหน่งที่ถูกแบ่งไว้ วิธีการทำ Descriptor มีลำดับดังนี้

1. การแบ่งส่วนภาพ

นำภาพ Tactile Image ที่ได้มาแบ่งเป็นส่วนย่อยตามลักษณะของฝ่ามือที่สัมผัสวัตถุตั้งรูปที่ 3.5 ซึ่งฝ่ามือจริงหากแบ่งตามแผ่นของผิวหนังจะสามารถแบ่งออกได้เป็น 15 ส่วน แต่ในการทดลองนี้เราแบ่งออกเป็นสามแบบ คือ 15 ส่วน 20 ส่วนและ 26 ส่วน เมื่อทำการแบ่งจะได้ภาพย่อยๆ ที่ได้จากการแบ่งเป็น 15 ภาพ 20 ภาพและ 26 ภาพตามลำดับ



รูปที่ 3.5 การแบ่งส่วนของฝ่ามือเป็น 15 ส่วน 20 ส่วนและ 26 ส่วน

2. ทำ Feature Extraction

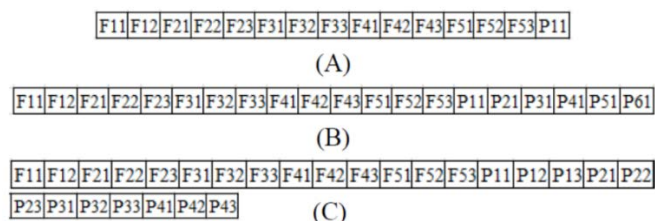
จากนั้นในภาพย่อยๆ ที่ถูกแบ่งมานั้นมีเวกเตอร์ข้อมูลจำนวนมาก ซึ่งเท่ากับจำนวน pixel ของภาพ ในงานนี้ใช้วิธีการหาผลรวมของภาพย่อยเพื่อลดจำนวนเวกเตอร์ข้อมูลลง โดยยังคงเป็นตัวแทนของข้อมูลอยู่ การหาค่าผลรวมของภาพย่อยสามารถหาได้จากสมการที่ 3.1

$$F = \sum_{i,j}^{N,M} x_i^j \quad (3.1)$$

เมื่อ F คือ Feature ได้จากผลรวมของเวกเตอร์ข้อมูล และ x_i^j คือเวกเตอร์ข้อมูลในภาพย่อยที่ถูกแบ่ง

3.สร้าง Data Vector

จากนั้นนำผลรวมของภาพที่ได้มาสร้าง Data Vector เพื่อเป็นตัวแทนของข้อมูลของแต่ละภาพ จากการแบ่งภาพเป็นภาพย่อย 15, 20 และ 26 ภาพย่อย จะได้ Vector ขนาด 15, 20 และ 26 เวกเตอร์ตามลำดับ ดังรูปที่ 3.6 แล้วเรียงลำดับข้อมูลเวกเตอร์เพื่อดำเนินการ Training ให้กับ ANN ซึ่งจากชุด Data Set ทั้งหมด 600 ภาพทำการแบ่งออกเป็นชุด Training Set 50% (300 ภาพ) และชุด Testing 50% (300 ภาพ) และทำการทดสอบ 3 ครั้งตามการแบ่งส่วนฝ่ามือออกเป็น 3 แบบ โดยทดสอบกับการทำ Descriptor 15 Feature, 20 Feature และ 26 Feature ตามลำดับ

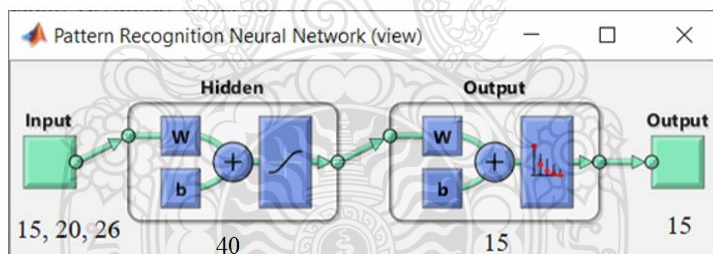


(A) 15 ภาพย่อย (B) 20 ภาพย่อย (C) 26 ภาพย่อย

รูปที่ 3.6 การสร้าง Data Vector ตามการแบ่งตำแหน่งฝ่ามือ

3.1.4 Training และ Testing ระบบด้วย ANN

โครงข่ายประสาทเทียมที่ใช้ในการทดสอบเป็นแบบ MLP เป็นรูปแบบหนึ่งของโครงข่ายประสาทเทียมที่มีโครงสร้างเป็นแบบหลายชั้น แบ่งกระบวนการทำงานเป็น 3 Layer ประกอบด้วยชั้นอินพุต (Input Layer), ชั้นซ่อน (Hidden Layer) และชั้นเอาต์พุต (Output Layer) โครงข่ายประสาทเทียมแบบหลายชั้นที่ใช้ในงานวิจัยนี้เป็นแบบฟีดฟอร์เวิร์ด (Feed Forward Back Propagation ANN) มีโครงสร้าง ดังรูปที่ 3.7 ประกอบไปด้วย 3 ชั้นคือ ชั้นอินพุต, ชั้นซ่อนจำนวน 40 โหนด และชั้นเอาต์พุต 15 Class ซึ่งออกแบบมาใช้ในการทดสอบการจดจำวัตถุจากการสัมผัสโดยใช้ Tactile Image ที่ได้จากมือมนุษย์



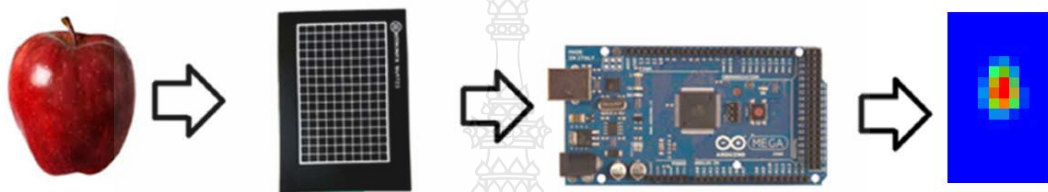
รูปที่ 3.7 โครงข่ายประสาทเทียมแบบหลายชั้นที่ใช้ในงานวิจัยนี้

3.2 Tactile Sensor และ การจดจำวัตถุจาก Tactile Image

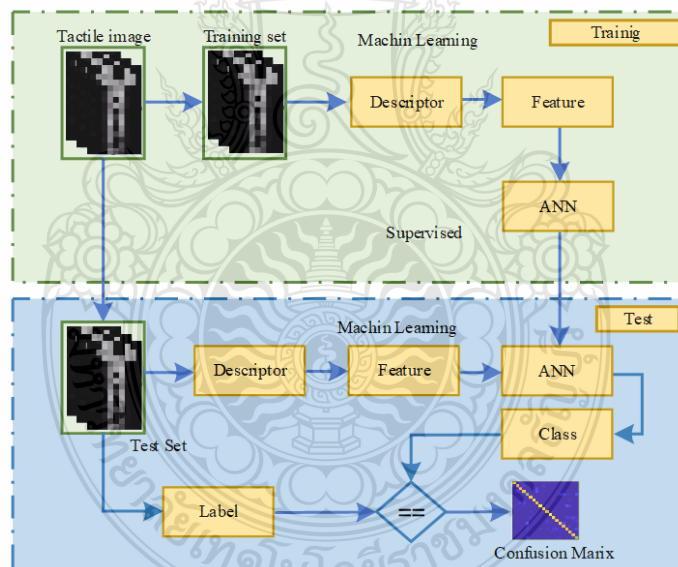
3.2.1 วิธีการทดสอบการจดจำวัตถุจาก Tactile Sensor

ในหัวข้อนี้มีวัตถุประสงค์เพื่อศึกษาการทำงานของ Tactile Sensor วงจรอ่านค่า และการจดจำวัตถุจากการสัมผัสที่ประยุกต์ใช้ Tactile Sensor จากงานวิจัยของ Drimus และคณะ [65] ในปี 2011 ได้ออกแบบเซ็นเซอร์อาร์เรย์ ขนาด 8 x 8 pixel จำนวน 2 ชั้นติดเข้ากับ Gripper ของหุ่นยนต์ตัวเซ็นเซอร์ทำมาจากวัสดุยางเพียโซเรซิสทีฟ (Piezoresistive Rubber) มีขนาด 25 x 25 mm² วงจรอ่านค่าเซ็นเซอร์ได้ออกแบบบนพื้นฐานไมโครคอนโทรลเลอร์เบอร์ dsPIC33FJ256 และทดลองจับวัตถุ

10 ชนิด จากนั้นใช้ Descriptor วิธี Statistic Feature (Mean, SD) และใช้วิธี KNN เป็นตัวจัด Class ซึ่งในงานนี้ให้ความถูกต้องอยู่ที่ประมาณ 92.0%, ในปี 2017 Gerlach และคณะ [60] ได้ออกแบบและสร้างเซ็นเซอร์อาร์เรย์ 8 x 8 pixel ขนาด 100 x 100 mm² และได้ใช้ LabVIEW and Arduino ในการอ่านค่าเซ็นเซอร์ และในปี 2019 Ramalingame และคณะ [61] ได้ออกแบบและสร้างเซ็นเซอร์อาร์เรย์ 4 x 4 pixel ขนาด 10 x 10 mm² และได้ใช้ LabVIEW and Arduino ในการอ่านค่าเซ็นเซอร์ ในงานนี้จึงได้ทำการทดลองการอ่านค่าของ Tactile Sensor ด้วย Arduino ดังรูปที่ 3.8 และทำการ Classify วัตถุจาก Tactile Image ที่ได้โดยการสร้าง Descriptor วิธี Statistic Feature (Mean, SD) และทำงานร่วมกับวิธี ANN ดังรูปที่ 3.9



รูปที่ 3.8 การทำงานในส่วนเก็บภาพจากการสัมผัส

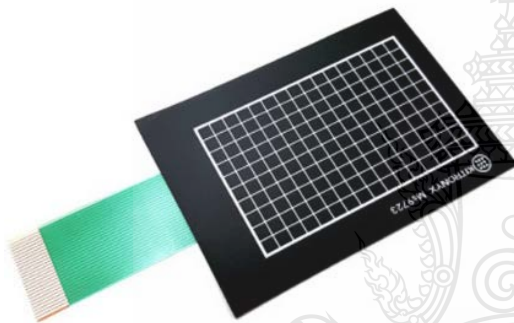


รูปที่ 3.9 การดำเนินการจดจำวัตถุจากการสัมผัสโดยใช้ Tactile Sensor ร่วมกับ ANN

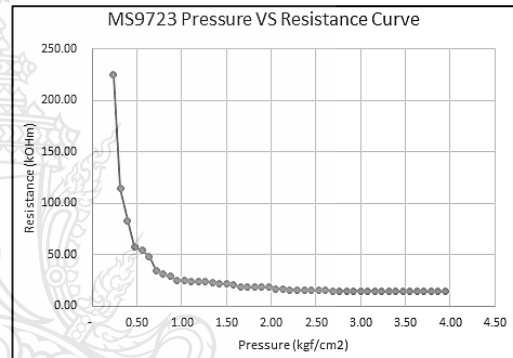
3.2.2 Tactile Sensor และวงจรรอ่านค่า

อุปกรณ์หลักที่ใช้ในการรับแรงกดคือ Tactile Sensor array ในงานวิจัยนี้เลือกใช้เซ็นเซอร์ของ SNOWFORCE ขนาด m คอลัมน์ x n แถว เท่ากับ 16 x 10 pixel ซึ่งมีลักษณะดังรูปที่ 3.10 การจัดวางเซ็นเซอร์เป็นแบบเมทริกซ์ ประกอบไปด้วยทั้งหมด 16 คอลัมน์ x 10 แถว ซึ่งจะได้จำนวน

pixel ทั้งหมดเป็น 160 Pixel ขนาดทางกายภาพของเซ็นเซอร์เท่ากับ $80 \times 50 \text{ mm}^2$ ซึ่งมีขนาดใกล้เคียงกับฝ่ามือ และความละเอียดเพียงพอในการจับภาพวัตถุขนาดใกล้เคียงกับฝ่ามือได้ ทั้งนี้หากใช้เซ็นเซอร์ที่มีความละเอียดต่ำไปจะทำให้ภาพที่ได้มีความละเอียดต่ำยากต่อการวิเคราะห์ และหากใช้เซ็นเซอร์ที่มีความละเอียดสูงขึ้นความละเอียดของภาพก็จะสูงกว่าซึ่งจะเหมาะกับการทดสอบกับวัตถุที่มีขนาดเล็กหรือวัตถุที่มีรายละเอียดสูง ในงานวิจัยนี้ได้ทำการทดสอบกับวัตถุขนาดใหญ่จึงทดลองโดยใช้เซ็นเซอร์ขนาด $16 \times 10 \text{ Pixel}$ ในแต่ละ pixel ขณะไม่มีแรงใดๆ มากระทำ ความต้านทานจะอยู่ที่ประมาณ $10 \text{ M}\Omega$ ความสัมพันธ์ของแรงที่กระทำกับค่าความต้านทานของเซ็นเซอร์แสดงดังรูปที่ 3.10 เมื่อมีแรงมากระทำต่อเซ็นเซอร์ขนาด 1 kg/cm^2 ความต้านทานของเซ็นเซอร์จะลดลงเหลือประมาณ $25 \text{ k}\Omega$



(A) Tactile Sensor



(B) กราฟแสดงความสัมพันธ์ระหว่างแรงกดและค่าความต้านทาน

รูปที่ 3.10 Tactile Sensor และกราฟความต้านทาน

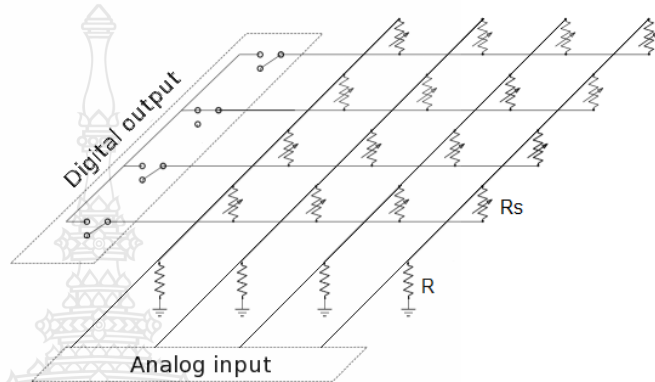
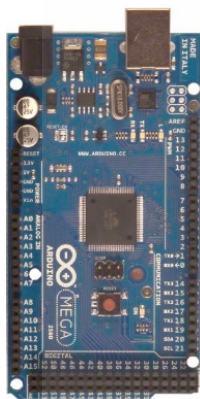
ค่าความต้านทานของเซ็นเซอร์สามารถหาได้จากสมการที่ 3.2 [89]

$$R_s = kF^{-n} \quad (3.2)$$

เมื่อ k เป็นค่าคงที่ของจีโอเมตริกคอนดิชันและค่าความต้านทานของวัสดุเพียโซรีซิสทีฟ และ F คือค่าแรงที่กระทำให้เกิดแรงกดบนพื้นผิวของเซ็นเซอร์ ส่วนค่า n เป็นค่าที่ขึ้นอยู่กับธรรมชาติของค่าการยืดหยุ่นของวัสดุ

ส่วนเก็บข้อมูลภาพในงานวิจัยนี้ใช้บอร์ดไมโครคอนโทรลเลอร์ Arduino รุ่น mega 2560 ซึ่งมีลักษณะดังรูปที่ 3.11(A) และได้พัฒนาโปรแกรมเพื่อทำการอ่านค่าจากเซ็นเซอร์แบบสแกน โดยต่อสัญญาณจาก 16 คอลัมน์ของเซ็นเซอร์เข้ากับขา Analogue Input $A_0 - A_{15}$ และทำการการสแกน

การอ่านทีละแถวโดยการต่อสัญญาณในแต่ละแถวเข้ากับขา Digital Output จนครบทั้ง 10 แถว วงจรปรับแต่งสัญญาณ (Signal Condition) สำหรับวัดค่าแรงกดที่กระทำบนตัวเซ็นเซอร์ เป็นวงจรที่ออกแบบบนพื้นฐานของวงจรแบ่งแรงดัน (Voltage Divider) ซึ่งการจัดวงจรแสดงดังรูปที่ 3.11(B) โดยที่ตัวเซ็นเซอร์ในแต่ละอีลีเมนต์จะมีความต้านทานที่เปลี่ยนแปลงไปตามแรงที่มากระทำโดยที่ค่าความต้านทานของเซ็นเซอร์จะอยู่ที่ $10\text{ M}\Omega - 10\text{ k}\Omega$ และต่ออนุกรมอยู่กับตัวต้านทาน R ค่าคงที่ขนาด $100\text{ k}\Omega$ ค่าความสว่างของภาพสามารถหาได้จากสมการ 3.3



(A) บอร์ดไมโครคอนโทรลเลอร์

(B) วงจรอ่านค่าเซ็นเซอร์

รูปที่ 3.11 บอร์ดไมโครคอนโทรลเลอร์และวงจรอ่านค่า

$$I = \left(\frac{255R}{kF - n + R} \right) \quad (3.3)$$

เมื่อ I คือค่าความเข้มของสัญญาณในระดับ 0-255 ขณะมีแรง F มากระทำและ R คือค่าความต้านทานที่ต่อไบอัส

การสร้างภาพจากการสัมผัสเมื่อทำการอ่านสัญญาณจากในแต่ละ pixel ของเซ็นเซอร์แล้วนำข้อมูลที่ได้อ่านมาจัดเรียงในรูปแบบสัญญาณสองมิติตามรูปแบบเดียวกับการจัดวางเซ็นเซอร์จะได้เมทริกซ์ข้อมูลขนาดเท่ากับเซ็นเซอร์ ซึ่งในที่นี้จะป็นเมทริกซ์ขนาด 16×10 ดังสมการที่ 3.4 จากนั้นนำข้อมูลที่ได้อ่านเป็นข้อมูล 2 มิติที่มีความสูงเป็นความเข้มของภาพจะได้เป็นภาพที่เกิดจากการสัมผัส

$$I = \begin{bmatrix} x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(m)} \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(m)} \\ \vdots & \vdots & \ddots & \vdots \\ x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(m)} \end{bmatrix} \quad (3.4)$$

เมื่อ I คือภาพจากการสัมผัส x คือค่าที่อ่านได้จากเซ็นเซอร์แต่ละ pixel

3.2.3 วัตถุและภาพจากการสัมผัสวัตถุและ Tactile Image

การเก็บข้อมูลภาพได้ทำการทดสอบระบบโดยการนำวัตถุ 10 ชนิดคือ แก้วน้ำ, ลูกกอล์ฟ, ขาม, รีโมทคอลโทรล, ประแจ, ประแจหกเหลี่ยม, แม่เหล็กบาศก์, แม่เหล็กทรงสามเหลี่ยม, แม่เหล็กกลมยาว และแม่เหล็กทรงรี ทำการสัมผัสวัตถุเพื่อเก็บภาพวัตถุในระนาบด้านเดียวแต่หมุนวัตถุให้เกิดการเปลี่ยนมุม ทำการสัมผัสวัตถุละ 100 ครั้ง ดังนั้นได้ภาพจากการสัมผัสทั้งหมดจำนวน 1000 ภาพ ขนาดของภาพในแต่ละภาพมีขนาด 160 Pixel จาก (16x10) เมื่อพล็อตเป็นภาพได้ Tactile Image ลักษณะดังรูปที่ 3.12



(A) วัตถุประแจ



(B) Tactile Image ของวัตถุประแจในมุมต่างๆ

รูปที่ 3.12 วัตถุและภาพ Tactile Image

3.2.4 สร้าง Data Vector

การดึง Feature ออกจากภาพเพื่อเป็นตัวแทนของภาพนั้นใช้ 2 วิธีประกอบด้วย ค่าเฉลี่ยของภาพ (μ) ที่ได้ซึ่งสามารถหาได้โดยสมการที่ 3.5

$$\mu = \frac{1}{NM} \sum_{i,j} x_i^j \quad (3.5)$$

และในส่วนของ Feature ที่สองเป็นค่าเบี่ยงเบนมาตรฐาน (s) ของ pixel ของภาพซึ่งได้จากสมการที่ 3.6

$$s = \left(\frac{1}{NM-1} \sum_{i,j} (x_i^j - \mu)^2 \right)^{\frac{1}{2}} \quad (3.6)$$

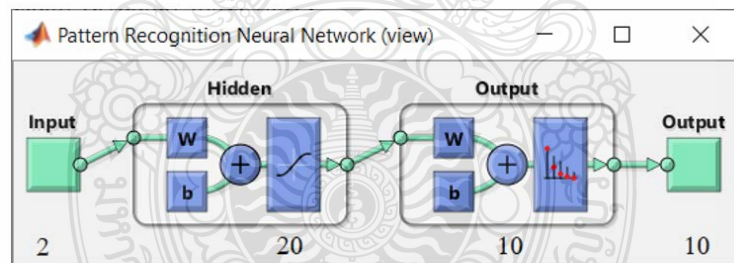
ค่าเฉลี่ยของภาพที่สัมผัสได้จะทำให้ค่าประมาณของแรงกดโดยรวมที่ใช้กับพื้นที่สัมผัส ซึ่งค่าที่ได้จะเปลี่ยนไปตามจำนวนของจุดสัมผัสที่เกิดขึ้นเนื่องจากบางวัตถุมีปริมาณจุดสัมผัสหลายจุด บางวัตถุจะมีปริมาณจุดสัมผัสที่น้อยกว่าอันเกิดจากลักษณะรูปร่างของวัตถุที่แตกต่างกัน และค่าส่วนเบี่ยงเบนมาตรฐานเป็นการอธิบายถึงประมาณจำนวน pixel ที่ได้จากการสัมผัส มีการกระจายของแรงมากหรือน้อย ทั้งสองคุณลักษณะนี้เป็นการลดมิติของข้อมูลจากข้อมูลทั้งหมด 160 ข้อมูลต่อภาพเหลือ 2 ข้อมูลต่อภาพ จากนั้นก็นำมาทำการจัดเรียงเป็นเวกเตอร์ข้อมูล Z ตามจำนวนภาพทั้งหมด N ภาพดังสมการที่ 3.7 และสมการที่ 3.8 และนำเข้าทดสอบการจดจำด้วยวิธี ANN

$$Z\mu = \{\mu_1, \mu_2, \dots, \mu_N\} \quad (3.7)$$

$$Zs = \{s_1, s_2, \dots, s_N\} \quad (3.8)$$

3.2.5 Training และ Testing ระบบด้วย ANN

โครงข่ายประสาทเทียมที่ใช้ในการทดลองนี้ ปรับปรุงจากงานในหัวข้อ 3.1 ดังรูปที่ 3.13 ประกอบไปด้วย 3 ชั้นคือ ชั้นอินพุต กำหนด 2 ข้อมูลต่อภาพ , ชั้นซ่อนจำนวน 20 โหนด และชั้นเอาต์พุต 10 Class ซึ่งออกแบบมาใช้สำหรับการทดสอบการจดจำวัตถุจากการสัมผัสโดยใช้ Tactile Image ที่ได้จาก Tactile sensor

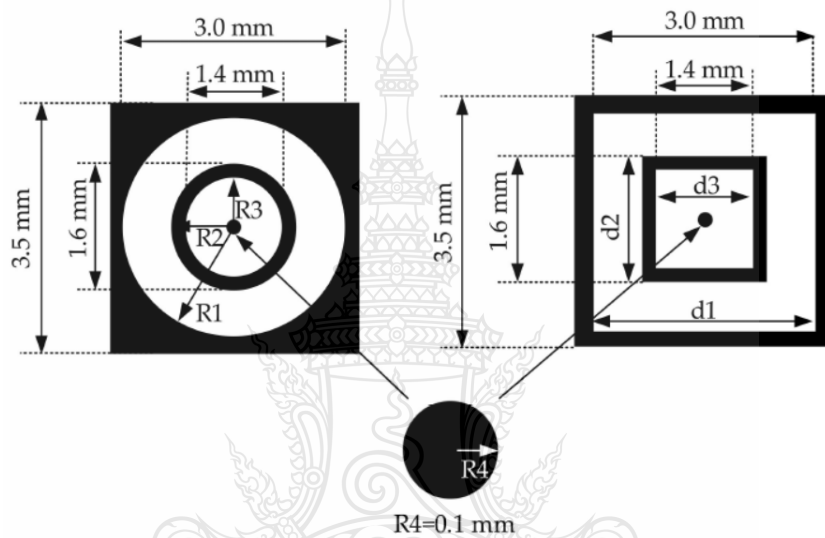


รูปที่ 3.13 โครงข่ายประสาทเทียมแบบหลายชั้นที่ใช้ในงานวิจัยนี้

3.3 การออกแบบเซ็นเซอร์แบบใหม่

มีงานวิจัยหลายงานที่ทำการออกแบบและสร้าง Tactile Sensor [9, 22-44] ซึ่งมีการออกแบบและสร้างจากหลักการทำงานที่หลากหลาย [22-44] ประเภทหนึ่งของ Tactile Sensor ที่ได้รับความนิยมเนื่องจากมีโครงสร้างที่ง่ายและวงจรการทำงานไม่ซับซ้อนคือ Tactile Sensor แบบ Piezoresistive [9, 32-34, 51-61] แต่พบว่าในงานวิจัยที่ได้นำเสนอก่อนหน้านี้มีปัญหาเรื่อง Resolution และ Physical Size ของ Tactile Sensor ไม่เหมาะสมกับการนำมาติดตั้งบนมือของ Humanoid Robot [9, 34, 51-61] ในงานนี้จึงได้นำเสนอการออกแบบเซ็นเซอร์แบบใหม่ที่มี

Resolution และ Physical Size เหมาะสมกับการนำไปติดตั้งที่ฝ่ามือของ Humanoid Robot เพื่อมาประมวลผลการ Recognition ด้วยวิธี Tactile Image Recognition ที่มีประสิทธิภาพ โดยพัฒนา Tactile Sensor แบบ Piezoresistive ปกติแล้วการออกแบบ Electrode ของ Tactile Sensor แบบ Piezoresistive ส่วนใหญ่เป็นแบบ Circle Electrode ดังรูปที่ 3.14 (A) [52] ในงานนี้แนะนำการออกแบบขั้ว Square Electrode ดังรูปที่ 3.14 (B) ซึ่งสามารถลด Dead Area ลงได้ มีพื้นที่หน้าสัมผัสที่เพิ่มขึ้น โดยพื้นที่หน้าสัมผัสของ Circle Electrode สามารถหาได้จากสมการ 3.10 เมื่อ R คือ Radius และ พื้นที่หน้าสัมผัสของ Square Electrode สามารถหาได้จากสมการ 3.11 เมื่อ d คือความกว้าง



(A) ขั้ว Electrode แบบ Circle Electrode) (B) ขั้ว Electrode แบบ Square Electrode
รูปที่ 3.14 ขั้ว Electrode ของ Tactile Sensor แบบ Circle Electrode และ Square Electrode

การคำนวณพื้นที่หน้าสัมผัสของ Electrode จะเป็นส่วนพื้นที่เฉพาะส่วนที่เป็นตัวนำไฟฟ้าซึ่งสามารถหาพื้นที่เฉพาะส่วนนี้ได้จากสมการที่ 3.9

$$A = A_1 - A_2 + A_3 - A_4 \quad (3.9)$$

ดังนั้นจะได้พื้นที่ Active Area ของ Circle Electrode เป็น

$$A = \pi R_1^2 - \pi R_2^2 + \pi R_3^2 - \pi R_4^2 \quad (3.10)$$

และได้พื้นที่ Active Area ของ Square Electrode เป็น

$$A = d_1^2 - d_2^2 + d_3^2 - \pi R_4^2 \quad (3.11)$$

จากการคำนวณพื้นที่หน้าสัมผัสของ Circle Electrode ของเซ็นเซอร์ในรูปที่ 3.14 (A) เท่ากับ 8.3692 mm^2 ส่วนของ Square Electrode ในรูปที่ 3.14 (B) เท่ากับ 6.5665 mm^2 ซึ่งมากกว่าแบบแรก 21.45 % จึงให้ความไวของเซ็นเซอร์สูงกว่า เมื่อมีการสัมผัสของวัตถุกระทำต่อเซ็นเซอร์

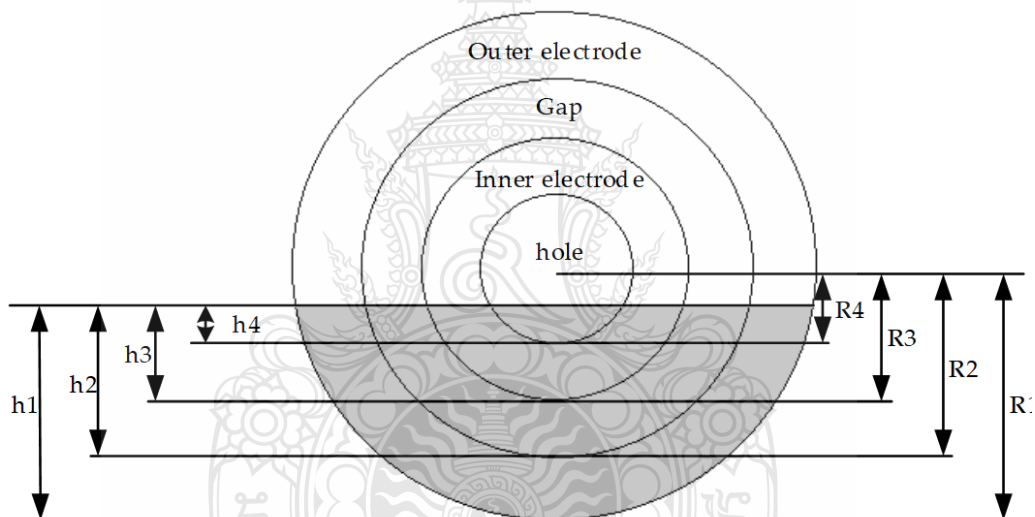
และเมื่อมีการสัมผัสกับเซ็นเซอร์แค่บางส่วนดังในรูปที่ 3.15 พื้นที่บางส่วนที่ถูกสัมผัสสามารถหาได้จากสมการที่ 3.12 และ 3.13 จากนั้นนำพื้นที่ในแต่ละส่วนไปแทนในสมการที่ 3.9 เพื่อหาพื้นที่เฉพาะส่วนที่เป็น Electrode

$$A = \frac{R^2}{2} \left(\frac{\alpha\pi}{180} - \sin \frac{\alpha\pi}{180} \right) \quad (3.12)$$

และ

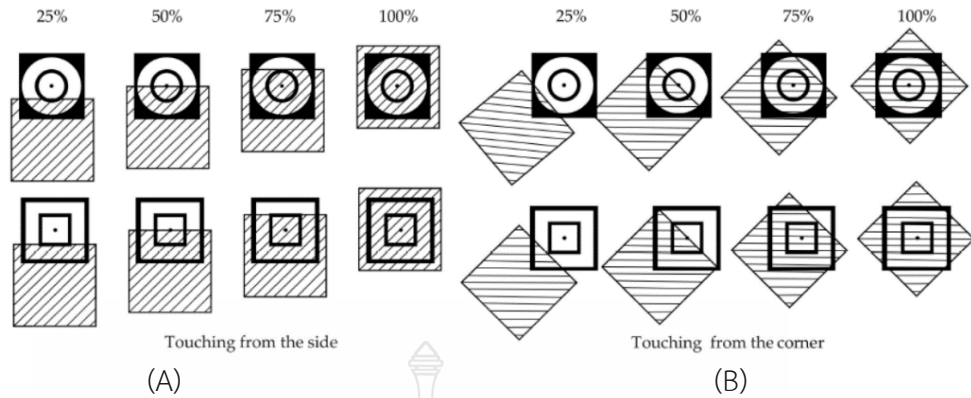
$$\alpha = 2\cos^{-1} \frac{(R - h)}{R} \quad (3.13)$$

เมื่อ R1 และ R3 เป็นรัศมีของอิเล็กโทรดภายนอกและภายในตามลำดับ, R2 และ R4 คือรัศมีของช่องว่างและรูตามลำดับ, นอกจากนี้ h1 และ h3 คือระยะสัมผัสของอิเล็กโทรดด้านนอกและด้านในตามลำดับ, h2 และ h4 คือระยะสัมผัสของช่องว่างและรูตามลำดับ, R คือ radius และ h คือระยะที่สัมผัส



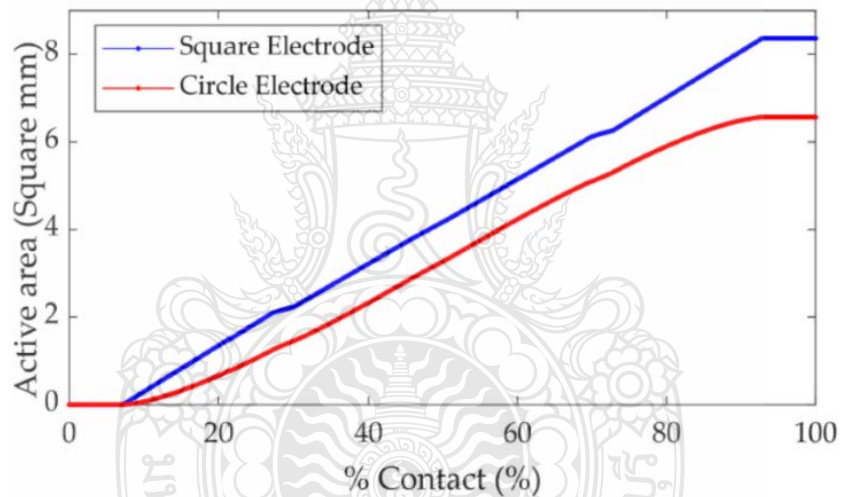
รูปที่ 3.15 พื้นที่สัมผัสบางส่วนของ Circle Electrode Sensor

รูปที่ 3.16 แสดงพื้นที่สัมผัสของเซ็นเซอร์แบบ Circle Electrode และแบบ Square Electrode, (A) เมื่อมีการสัมผัสที่ 25%, 50%, 75%, และ 100% จากด้านข้าง และ (B) เมื่อมีการสัมผัสที่ 25%, 50%, 75%, and 100% จากมุม เมื่อมีการสัมผัสที่เซ็นเซอร์แค่บางส่วนการคำนวณค่า % Contact ต่อ Active Area เพื่อเปรียบเทียบระหว่าง Electrode ทั้ง 2 แบบผลที่ได้พบว่า Square Electrode มีสัดส่วน Active Area ต่อ % Contact ดีกว่าแสดงดังรูปที่ 3.17



- (A) เมื่อมีการสัมผัสที่ 25%, 50%, 75%, และ 100% จากด้านข้าง
 (B) เมื่อมีการสัมผัสที่ 25%, 50%, 75%, and 100% จากมุม

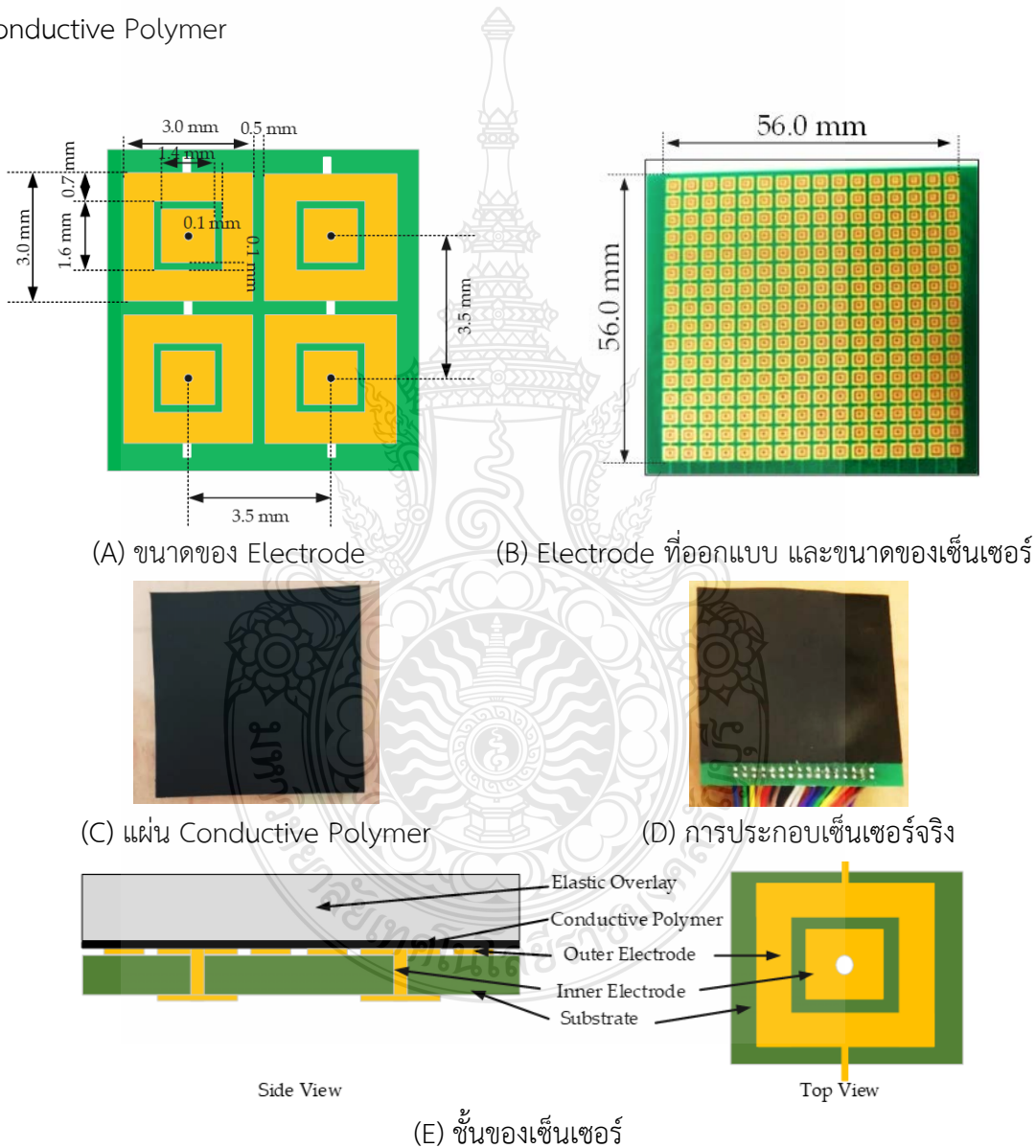
รูปที่ 3.16 พื้นที่สัมผัสของเซ็นเซอร์แบบ Circle Electrode และแบบ Square



รูปที่ 3.17 เปรียบเทียบสัดส่วน Active Area ต่อ % การสัมผัสเซ็นเซอร์

ในงานนี้เซ็นเซอร์ออกแบบเป็นชนิด Piezoresistive ใช้ Conductive Polymer ยี่ห้อ capLINQ (รหัส MVCF-40012BT50KS / 2A) เป็น Piezoresistive แบบแผ่นมีความหนา 0.1 mm และความต้านทานพื้นผิว 50,000 Ω/cm^2 โครงสร้างของเซ็นเซอร์ใช้เทคโนโลยี PCB โดยใช้แผ่น PCB ชนิด Epoxy เป็น Substrate ที่มีความหนา 0.5 mm และ Electrode เป็น Cu ความหนา 0.2 mm ชุบ Au หนา 18 μm ชั้น Electrode ได้รับการออกแบบเป็นขนาด 16 Row x 16 Column Electrode ออกแบบเป็น Square Electrode เพื่อลด Dead Area ของเซ็นเซอร์ซึ่งมีลักษณะดังรูปที่ 3.18 (A) โดยมีขนาด Pixel ของเซ็นเซอร์เท่ากับ 3.25 x 3.25 mm² มี External Electrode ขนาด 3.0 x 3.0 mm² และมี Internal Electrode ขนาด 1.4 x 1.4 mm², ช่องว่างระหว่าง Internal และ External

Electrode ห่างกัน 0.1 mm และระยะห่างของเซ็นเซอร์แต่ละ Pixel เท่ากับ 0.5 mm ขนาดของ Sensor Array เท่ากับ $56.0 \times 56.0\text{ mm}^2$ ดังแสดงในรูปที่ 3.18 (B) ในรูปที่ 3.18 (D) แสดงชั้น Layer ของเซ็นเซอร์ประกอบด้วยชั้นล่างสุดเป็น Substrate Layer ทำหน้าที่เป็น base และ Insulator, ชั้นถัดมาเป็นชั้นของ Electrode Layer ทำหน้าที่เป็น Conductor ให้กับเซ็นเซอร์แต่ละตัว, ชั้นที่ 3 เป็น Conductive Polymer ทำหน้าที่ Piezoresistive ซึ่งค่าความต้านทานจะเปลี่ยนแปลงไปตามแรงที่มากระทำ, และชั้นบนสุดเป็นวัสดุยืดหยุ่น (Elastic Overlay) ทำหน้าที่เป็นตัวรับและส่งผ่านแรงไปยัง Conductive Polymer

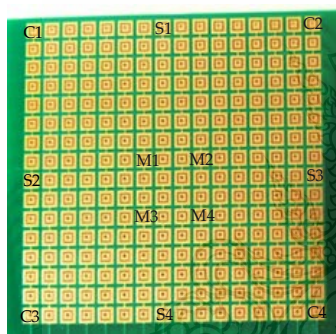


รูปที่ 3.18 เซ็นเซอร์ที่ใช้ในการวิจัยนี้

จากนั้นได้ทำการทดสอบคุณสมบัติของเซ็นเซอร์ที่ออกแบบในครั้งนี ซึ่งแบ่งเป็น 2 ส่วนคือ คุณสมบัติด้านความต้านทานที่เกิดขึ้น เมื่อมีแรงมากระทำ และภาพ Tactile ที่ได้เมื่อมีการสัมผัสวัตถุ โดยทั้งสองส่วนนี้ได้ทำการ Setup Experiment ที่ต่างกันซึ่งมีรายละเอียดดังนี้

3.3.1 Sensor Sensitivity

Sensitivity เป็นการวัดคุณสมบัติด้านความต้านทานที่เกิดขึ้นต่อแรงที่มากระทำ สามารถทำได้โดยการ Setup Experiment ดังรูปที่ 3.19 (B) ซึ่งอุปกรณ์ที่ใช้ประกอบด้วยเกจวัดแรงกด (Force Gauge), อุปกรณ์ส่งแรง (Transmission Rod) ที่รู้ขนาดพื้นที่หน้าตัด และ โอห์มมิเตอร์ (Ohmmeter) สำหรับวัดค่าความต้านทาน เนื่องจากจำนวน Element ของเซ็นเซอร์มีถึง 256 Pixel ในการทดสอบในครั้งนีจึงต้องลดตำแหน่งทดสอบให้เหมาะสม ซึ่งในการกำหนดตำแหน่งแสดงดังรูปที่ 3.19 (A) แบ่งกลุ่มในการทดสอบออกเป็น 3 กลุ่ม คือ 1) เซ็นเซอร์ตำแหน่งมุมประกอบด้วย C1, C2, C3 และ C4 2) ตำแหน่งด้านข้าง ประกอบด้วย S1, S2, S3 และ S4 และ 3) ตำแหน่งตรงกลางประกอบด้วย M1, M2, M3 และ M4



(A) ตำแหน่งที่วัด



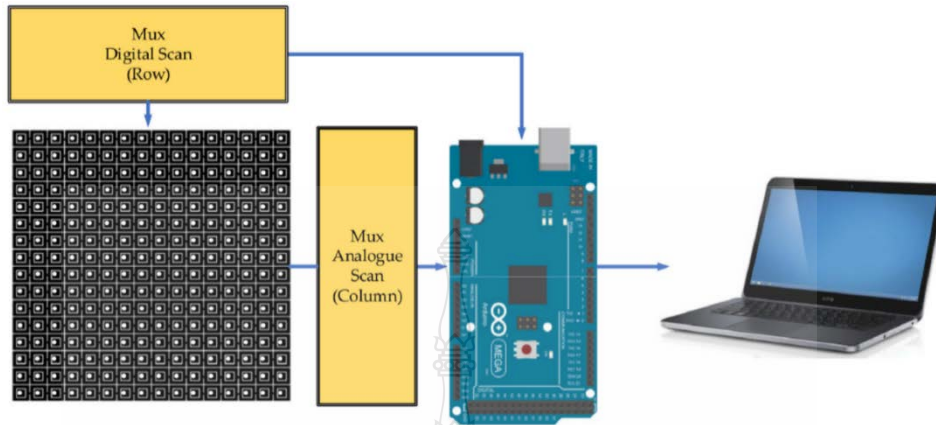
(B) เครื่องมือที่ใช้

รูปที่ 3.19 การวัด Sensitivity ของเซ็นเซอร์

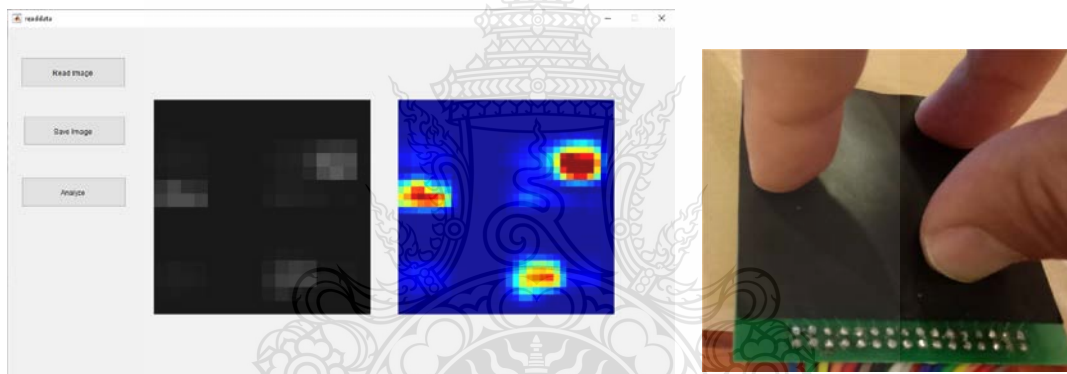
3.3.2 Tactile sensor และวงจรสำหรับอ่านค่าจากเซ็นเซอร์

การอ่านค่าจาก Tactile Sensor Array ใช้บอร์ดไมโครคอนโทรลเลอร์ Arduino รุ่น mega 2560 การทำงานจะเป็นการสแกนเซ็นเซอร์ในแต่ละ Row โดยการควบคุม Output ของ Digital Line ที่ Row ของเซ็นเซอร์และทำการสแกนและอ่านค่า Analogue จาก Column ของเซ็นเซอร์ ลำดับการทำงานให้เริ่มจากให้ Output เป็น Logic "1" ใน Row ที่ต้องการอ่านค่า และ Logic "0" ใน Row อื่นๆ ($O_i = 1$), ($O_1, \dots, O_{i-1}, O_{i+1}, \dots, O_n = 0$) จากนั้นใช้ Analogue Port ทำการสแกนเพื่ออ่านค่าในแต่ละ Column และใช้ A/D เพื่อแปลงเป็นข้อมูลดิจิทัลและส่งข้อมูลไปยังส่วนการประมวลผลด้วยคอมพิวเตอร์ จากนั้นทำการสแกน Digital Line ในตำแหน่งถัดไปจนกว่าจะสามารถอ่านค่าจากเซ็นเซอร์ได้จนครบทั้งหมด ในส่วนของคอมพิวเตอร์รับผิดชอบการประมวลผลข้อมูล ทั้งส่วนการ

แสดงผล Tactile Image และ Recognition ระบบการทำงานแสดงดังรูปที่ 3.20 และโปรแกรมที่พัฒนาสำหรับอ่านและวิเคราะห์ผลดังรูปที่ 3.21



รูปที่ 3.20 ไดอะแกรมชุดเครื่องมือที่ใช้ในการทดสอบ

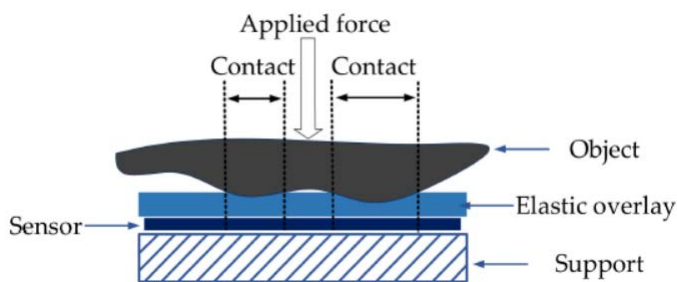


(A) GUI ของโปรแกรมสำหรับอ่านค่า

(B) Tactile Sensor Array

รูปที่ 3.21 GUI ของโปรแกรมสำหรับอ่านค่าและ Tactile Sensor Array

ส่วนสำคัญของ Tactile Sensor คือส่วนของผิวด้านบนที่ใช้ในการกระจายแรงกด หรือที่เรียกว่าชั้นผิวหนังดังรูปที่ 3.22 ในกรณีที่เซ็นเซอร์ไม่มีชั้นผิวหนัง เมื่อวัตถุสัมผัสลงบนเซ็นเซอร์ ตำแหน่งที่นูนสุดของวัตถุจะสัมผัส Pixel ใด Pixel หนึ่ง ซึ่ง Pixel นั้นจะเป็นตัวรับแรงกดทั้งหมดเซ็นเซอร์ ทำให้ไม่สามารถอ่านลักษณะพื้นผิวของวัตถุได้ เนื่องจากปรากฏแรงกดแค่จุดเดียว แต่เมื่อมีส่วนกระจายแรงกดเมื่อวัตถุสัมผัสลงบนชั้นผิวหนัง ซึ่งเป็นวัสดุยืดหยุ่น หรือมีชั้นผิวหนังเป็นตัวกระจายแรงกดลงบนเซ็นเซอร์หรือชั้น Elastic Overlay ก็จะได้รูปแบบแรงกดหรือ Pressure Profile ที่เหมาะสมของวัตถุ ทั้งนี้หากชั้นผิวหนังไม่เหมาะสม เช่นมีความหนามากเกินไป หรือมีความแข็งมากเกินไปจะไม่สามารถกระจายแรงกดได้ดีเช่นกัน ดังนั้นชั้นผิวหนังที่เหมาะสมจึงมีความสำคัญสำหรับ Tactile Sensor ที่นำมาประยุกต์ใช้งานเป็นผิวหนังของหุ่นยนต์



(A) ชั้นของเซ็นเซอร์



(B) Tactile Sensor ที่ประกอบเสร็จ

รูปที่ 3.22 ชั้นของเซ็นเซอร์ และ Tactile Sensor ที่ประกอบเสร็จ

3.4 การจดจำวัตถุจากการสัมผัสสำหรับมือหุ่นยนต์

ในการศึกษานี้ได้มีการพัฒนามือหุ่นยนต์เมคคาทรอนิกส์ที่มีโครงสร้างเหมือนมือมนุษย์ ประกอบด้วยฝ่ามือและนิ้วมือทำนิ้ว ซึ่งเป็นการออกแบบพื้นฐานของหุ่นยนต์ฮิวแมนนอยด์สมัยใหม่ ในการทดลองได้ใช้ Tactile Sensor ที่ประกอบขึ้นมาติดตั้งเป็นระบบประสาทสัมผัสทางกาย โดยได้ติดตั้งไว้กับฝ่ามือของหุ่นยนต์ดังรูปที่ 3.23 เมื่อมีการจับวัตถุแรงกดจากนิ้วมือจะส่งแรงกดลงบนวัตถุ ซึ่งทำให้วัตถุถูกกดและสัมผัสกับชั้นผิวหนัง แรงกดจะถูกส่งและกระจายไปยัง Tactile Sensor โดยวัสดุยืดหยุ่น การกระจายและส่งผ่านแรงทำให้ Tactile Sensor เกิดแรงกดในลักษณะตามรูปร่างของวัตถุ เมื่ออ่านค่าจาก Tactile Sensor และนำมานำเสนอเป็นภาพจะได้ภาพแสดง Pressure Profile ของวัตถุ



(A) มือหุ่นยนต์ที่ติดตั้ง Tactile Sensor



(B) มือหุ่นยนต์ขณะจับวัตถุ

รูปที่ 3.23 มือหุ่นยนต์เมคคาทรอนิกส์ที่ติดตั้ง Tactile Sensor และมือหุ่นยนต์ขณะจับวัตถุ

3.4.1 วัตถุ และ Tactile Image

การเก็บภาพตัวอย่างทำโดยการส่มตำแหน่งในการจับ ภาพที่ได้จากการสัมผัสถูกรวบรวมมาจากการจับในตำแหน่งต่างๆ ของวัตถุ รูปที่ 3.23 (B) แสดงให้เห็นถึงขั้นตอนการรวบรวมข้อมูลจากวัตถุ ตัวอย่าง การวางวัตถุในตำแหน่งเริ่มต้นจะถูกส่มโดยการหมุนวัตถุในพิภัก 3 มิติ หรือหมุนในแนว X-Y-Z เมื่อวัตถุอยู่ในมือแล้วก็ปล่อยให้มีการจับอย่างอิสระ ข้อมูลภาพที่ได้มีข้อมูลจำนวน 256 เวกเตอร์ ได้จากเซ็นเซอร์ขนาด 16 x 16 Pixels ชุดข้อมูลถูกสร้างขึ้นจากการทดสอบกับวัตถุจำนวน 20 วัตถุ (หรือ 20 Class) ประกอบด้วย แบตเตอรี่ รีโมตคอนโทรล คีมพลาสติก ไขควง ถ้วยกาแฟ กรรไกร ประแจ ประแจหกเหลี่ยม ลูกกอล์ฟ ตลับเมตร เมาส์คอมพิวเตอร์ แปรง แอมป์มิเตอร์ ขวดโคล่า ปากกา เครื่องชาร์จ ขวดโซดา ประแจเลื่อน ขวดน้ำ และขวดครีม ขนาดของวัตถุเหล่านี้แสดงอยู่ในตารางที่ 3.1 และวัตถุแสดงในรูปที่ 3.24 ชุดวัตถุที่ใช้ทดลองเป็นวัตถุที่ใช้ในชีวิตประจำวันซึ่งคล้ายกับงานของ [59]

ตารางที่ 3.1 ขนาดของวัตถุ

Class	Object	Size
1.	Battery	Ø=33 mm, H=60 mm
2.	Remote controller	L=172 mm, W=45 mm, H=15 mm
3.	Plastic tongs	L=165 mm, W=85 mm, H=15 mm
4.	Screw driver	Ø=38 mm, L=220 mm
5.	Cup	Ø=85 mm, W=120 mm, H=95 mm
6.	Scissors	L=18 mm, W=80 mm, H=12 mm
7.	Fixed wrench	L=168 mm, W=30 mm, H=6 mm
8.	Allen key	L=90 mm, W=35 mm, H=6 mm
9.	Golf ball	Ø=40 mm
10.	Tape measure	Ø=70 mm, W=92 mm, H=47 mm
11.	Computer mouse	L=100 mm, W=50 mm, H=35 mm
12.	Brush	Ø=80 mm, L=390 mm
13.	Amp meter	Ø=35 mm, L=35 mm
14.	Cola bottle	Ø=60 mm, H=22 mm
15.	Pen	Ø=12 mm, L=147 mm
16.	Charger	L=90 mm, W=70 mm, H=22 mm

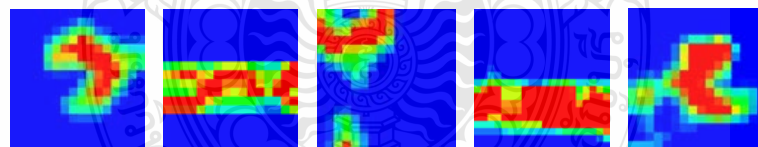
ตารางที่ 3.1 ขนาดของวัตถุ (ต่อ)

Class	Object	Size
17.	Soda Bottle	$\varnothing=65$ mm, H=155 mm
18.	Variable wrench	L=200 mm, W=60 mm, H=14 mm
19.	Water bottle	$\varnothing=60$ mm, H=175 mm
20.	Cream bottle	$\varnothing=75$ mm, H=160 mm

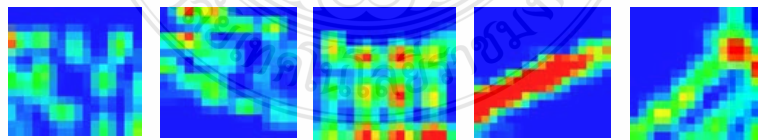


รูปที่ 3.24 วัตถุที่นำมาใช้ในการทดลอง

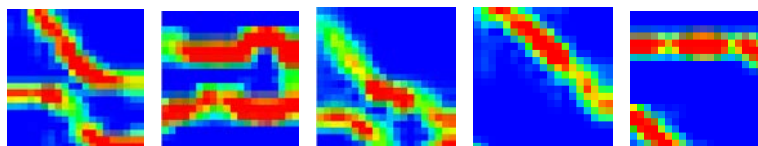
ในการทดลองนี้ได้สร้างชุดข้อมูลจากการจับวัตถุโดยจับวัตถุละ 200 ครั้ง จำนวน 20 วัตถุ ส่งผลให้ได้ภาพสำหรับชุดข้อมูล 4000 ภาพเป็นฐานข้อมูล เนื่องจากในการทดลองนี้ไม่ขึ้นกับทิศทางและตำแหน่ง ภาพที่ได้จากการจับจึงมีการหมุนวัตถุจึงได้ภาพที่แตกต่างกันของวัตถุดังรูปที่ 3.25 แสดงภาพ Tactile Image จากวัตถุ 3 ชนิดในตำแหน่งต่างๆ



(A) "Fixed wrench"



(B) "Remote controller"



(C) "Plastic tongs"

รูปที่ 3.25 ตัวอย่างภาพ Tactile Images

3.4.2. Resolution Enhancement

ในการศึกษานี้ใช้การปรับปรุงความละเอียดภาพเพื่อปรับความละเอียดของ Tactile Image ให้สูงขึ้นให้เกิดความเหมาะสมกับการนำไปทำการวิเคราะห์ด้วยเทคนิคต่างๆ เช่นเป็นอินพุตที่เหมาะสมสำหรับ DCNN การปรับปรุงความละเอียดของภาพใช้วิธีการ Bicubic Interpolation (Resize Tactile Image) วิธีการแสดงในสมการ 3.14 ถึงสมการที่ 3.17 เนื่องจากวิธีนี้เคยถูกนำมาใช้โดย [61, 62] และให้ผลการทดสอบที่ดีกว่า การทำงานคือเป็นฟังก์ชันเพื่อประมาณความต่อเนื่องของการกระจายแรงกดบนพื้นผิวโดยใช้ข้อมูลพิกัดอันดับสอง การทำ Bicubic Interpolation แสดงในรูปที่ 3.26 การแก้ไขแบบไบคิวบิกถูกคำนวณที่ 16 จุดรอบๆ ตำแหน่งที่กำลังพิจารณา ฟังก์ชันหลัก $W(x)$ ของการแก้ไขแบบไบคิวบิกคือ

$$W(x) = \begin{cases} (a+2)|x|^3 - (a+3)|x|^2 + 1 & \text{for } |x| \leq 1 \\ a|x|^3 - 5a|x|^2 + 8a|x| - 4a & \text{for } 1 < |x| < 2 \\ 0 & \text{otherwise} \end{cases} \quad (3.14)$$

เมื่อระยะตามแนวแกน y และแกน x สำหรับการประมาณค่า $B(x, y)$ (จำนวน 16 pixel) ถูกกำหนดเป็น K_{im} และ K_{jn} ตามลำดับ ดังแสดงในสมการ (3.15)

$$\begin{aligned} K_{i0} &= 1 + \mu; & K_{i1} &= \mu; & K_{i2} &= 1 - \mu; & K_{i3} &= 2 - \mu; \\ K_{j0} &= 1 + \nu; & K_{j1} &= \nu; & K_{j2} &= 1 - \nu; & K_{j3} &= 2 - \nu; \end{aligned} \quad (3.15)$$

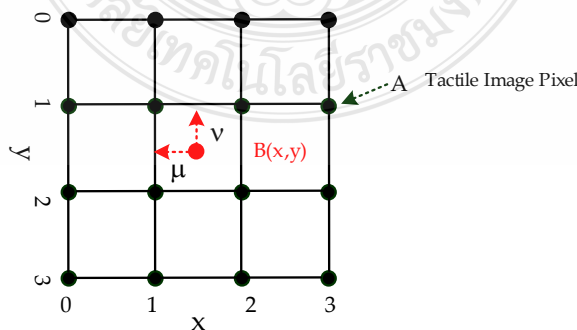
ค่าสัมประสิทธิ์หาได้จาก

$$a_{mn} = W(K_{im})W(K_{jn}) \quad (3.16)$$

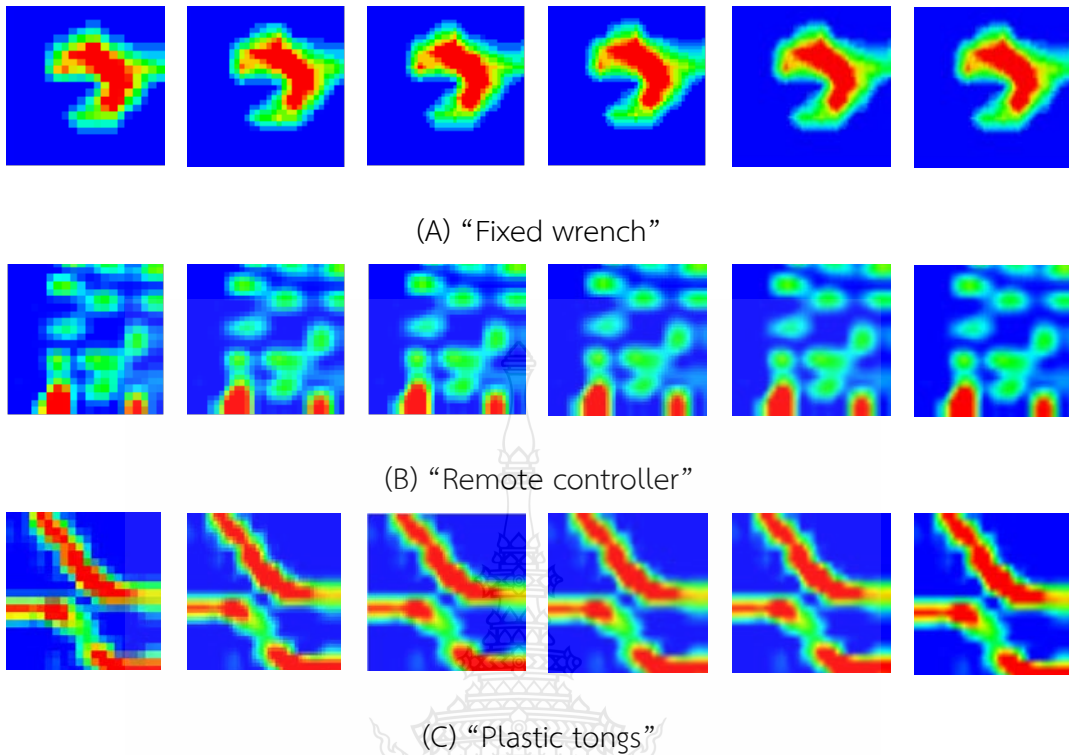
ดังนั้นภาพที่ทำ Interpolation จึงสามารถอธิบายได้ด้วยสมการ (3.17)

$$B(x, y) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} A(x_i, y_j) \quad (3.17)$$

เมื่อ A คือค่าของ Tactile image ของแต่ละ pixel ของ Tactile Sensor, B คือจุดภาพที่ทำ Interpolation, a_{ij} แทนพารามิเตอร์น้ำหนัก และ i, j แทนพิกัดแกน x และแกน y ตามลำดับ



รูปที่ 3.26 การทำ Bicubic Interpolation



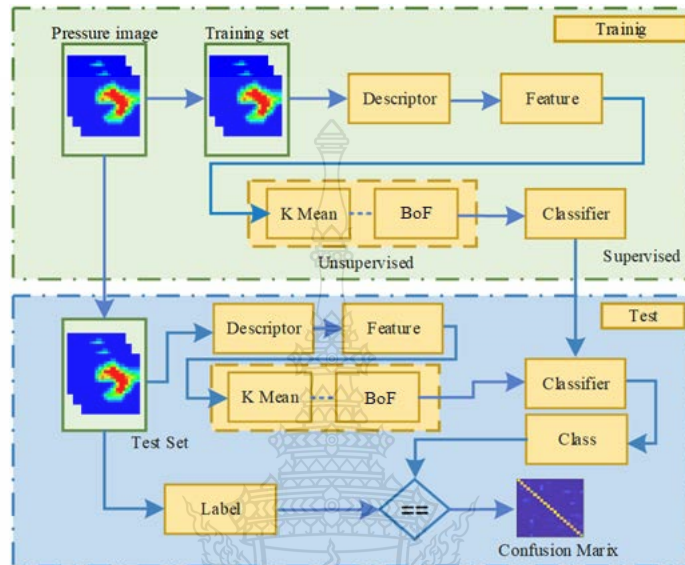
รูปที่ 3.27 ภาพ Tactile Image ที่ได้จากการทำ Bicubic Interpolation

ในรูปที่ 3.27 แสดงการทำ Bicubic เพื่อเพิ่มความละเอียดของภาพ จากภาพที่ได้จาก Tactile Sensor array ขนาด 16×16 pixels เป็นภาพขนาด 32×32 , 64×64 , 128×128 , 256×256 และ 512×512 Pixels พบว่าคุณภาพของภาพดีขึ้นอย่างมาก เนื่องจากภาพความละเอียดต่ำการระบุโพรไฟล์ทางเรขาคณิตของวัตถุนั้นเป็นเรื่องยาก เมื่อความละเอียดสูงขึ้นรายละเอียดทางเรขาคณิตของภาพที่สัมผัสได้ก็ชัดเจนขึ้น ภาพที่ปรับปรุงแล้วเหล่านี้ถูกใช้เป็นข้อมูลป้อนเข้าสำหรับการพัฒนาระบบการจดจำ

3.4.3 อัลกอริทึมแบบ BoF สำหรับมือหุ่นยนต์

จากหลักการทำงานของ BoF ในหัวข้อ 2.4.2 เมื่อนำมาประยุกต์ใช้กับ Tactile image การดำเนินงานหรือ Work Flow แสดงดังรูปที่ 3.28 การทดลองได้ทดสอบกับ Descriptor ทั้งหมด 3 วิธี ประกอบด้วย Moment Analysis (MA), Scale Invariant Feature Transform (SIFT) และ Polar Furrier (PF) (วิธีการหาค่าตามหัวข้อ 2.5.2.1, 2.5.2.2 และ 2.5.2.3 ตามลำดับ) สำหรับวิธี Classifier ทำการทดลองกับ K-Nearest Neighbors (KNN), Support Vector Machine (SVM) และ Artificial Neural Network (ANN) (มีวิธีการหาค่าตามหัวข้อ 2.6.2, 2.6.3 และ 2.6.4 ตามลำดับ) ขั้นตอนการทำงานเริ่มจากการแบ่ง Data Set ออกเป็น 2 ชุด ซึ่งจากชุด Data Set ทั้งหมด 4000 ภาพทำการแบ่ง

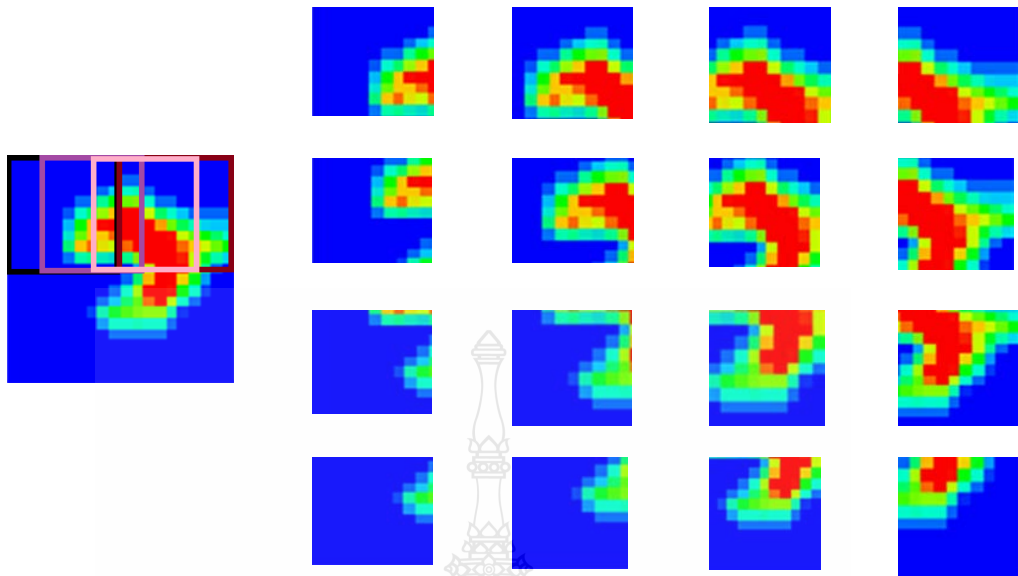
ออกเป็นชุด Training Set 50% (2000 ภาพ) และชุด Testing Set 50% (2000 ภาพ) ในส่วนของ Training Set นำมาใช้ในการสอน Model โดยการจัด Cluster คุณสมบัติของวัตถุด้วยวิธี K-Mean Clustering แล้วสร้าง Histogram เพื่อเป็นตัวแทนของวัตถุแต่ละ Class จากนั้นนำ Histogram ที่ได้ไป Training ให้ตัว Classifier เพื่อนำไปใช้ในการทดสอบกับ Test Set ต่อไป



รูปที่ 3.28 Work Flow สำหรับการจดจำวัตถุจากการสัมผัสด้วยวิธี BoF

1. การแบ่งส่วนภาพ

เนื่องจากการใช้วิธี BoF ถูกนำมาใช้สำหรับการทดสอบการจดจำวัตถุจากภาพ Tactile image ขนาดเล็กซึ่งเป็นภาพที่ได้จากการติดตั้ง Tactile sensor ขนาดเล็กในส่วนของนิ้วมือ แต่ในงานนี้ได้พัฒนาเซ็นเซอร์ขนาดใหญ่และนำเสนอการใช้วิธี BoF ในการทดสอบ ดังนั้นจึงได้นำเสนอวิธีการแบ่งส่วนภาพเป็นภาพย่อย (Sub image) ขนาด 4×4 ภาพ หรือ 16 ภาพย่อย แต่เนื่องจากเมื่อมีการแบ่งภาพเป็นภาพย่อยทำให้ขนาดของภาพเล็กลงไปมาก เพื่อคงขนาดของภาพให้เหมาะสมกับการคำนวณค่า จึงได้นำเสนอวิธีการแบ่งภาพแบบ Overlapping ซึ่งมีการซ้อนทับกันบางส่วนของภาพดังรูปที่ 3.29 ในการทดลองใช้การแบ่งภาพเป็นภาพย่อย ขนาด 4×4 แบบ Overlapping เป็นทั้งหมด 16 ภาพย่อย (ภาพย่อยขนาด 8×8 pixel) ดังรูปที่ 3.29 แล้วนำไปทำ Descriptor สำหรับสร้าง Data Vector เพื่อเป็นตัวแทนของภาพ

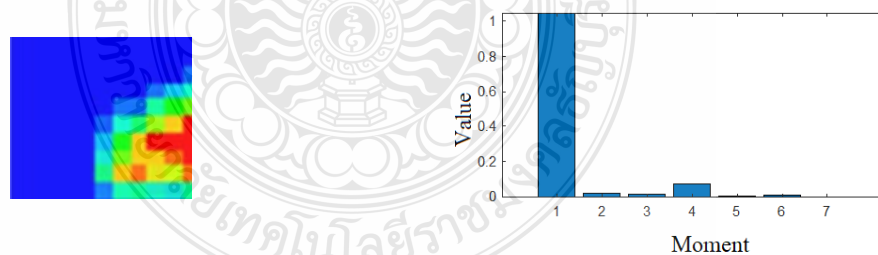


รูปที่ 3.29 การแบ่ง Sub Image สำหรับทำ Descriptor

2. การทำ Descriptor

เมื่อได้ Sub image มาแล้วก็เข้าสู่กระบวนการทำ Descriptor เพื่อสร้าง Feature สำหรับภาพ Sub Image แต่ละตัว ซึ่งในการทดลองนี้ได้ทำการทดสอบกับการทำ Descriptor 3 วิธี ดังนี้

1. Moment Analysis [81] หรือค่า Hu Moment สามารถหาได้จากสมการที่ 3.18 ซึ่งจะได้ Feature มาทั้งหมด 7 ค่า จากนั้นนำมาจัดเรียงเป็นเวกเตอร์ข้อมูลดังสมการที่ 3.19 สำหรับนำไปใช้วิธี K-mean Clustering ในการจัด Cluster ของแต่ละภาพย่อยต่อ



รูปที่ 3.30 ภาพ Sub image และ ค่า Feature ที่ได้จากการหาค่า Moment ทั้ง 7 ค่า

$$\begin{aligned}
M1 &= (\eta_{20} + \eta_{02}), \\
M2 &= (\eta_{20} + \eta_{02})^2 + 4\eta_{11}^2, \\
M3 &= (\eta_{30} + \eta_{12})^2 + (3\eta_{21} + \eta_{03})^2, \\
M4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2, \\
M5 &= (\eta_{30} + 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
&\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2], \\
M6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
&\quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}), \\
M7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
&\quad - (\eta_{30} + 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2].
\end{aligned} \tag{3.18}$$

เมื่อ M คือค่า Moment แต่ละตัว η คือค่า normalized central moments

$$v = M1, M2, M3, \dots, M7 \tag{3.19}$$

เมื่อ v คือเวกเตอร์ข้อมูลที่เป็นตัวแทนของภาพ เมื่อ M คือค่า Moment แต่ละตัว

เนื่องจากได้ทำการแบ่งภาพย่อย Sub image เป็น 16 ภาพ ดังนั้นจึงต้องทำ Descriptor เพื่อหาค่า Feature สำหรับ Sub image แต่ละภาพ ภาพ Sub image แต่ละภาพถือเป็น ข้อมูลภาพ 1 ข้อมูลที่จะนำเข้าสู่การทำ K-mean Clustering ดังนั้นจะมี Data Vector สำหรับจัด Cluster จำนวน 16 เวกเตอร์ต่อภาพ Tactile image ที่นำเข้ามาทดสอบ 1 ภาพ

2. SIFT [82] เริ่มจากนำ Sub image มาทำการเบลอภาพด้วย Gaussian Function ด้วยสมการที่ 3.20 ซึ่งจะได้ภาพที่ผ่านการเบลอมาดังรูปที่ 3.31

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{3.20}$$

เมื่อ (x, y) คือพิกัดบนภาพ, L คือ ภาพที่กรองผ่านการเบลอด้วย Blur Filter, G คือ ตัวกรองแบบ Gaussian และ I คือภาพอินพุต ทั้งนี้ตัวกรองแบบ Gaussian แสดงดังสมการที่ 3.21

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \tag{3.21}$$

จากนั้นทำการหาผลต่างของภาพ DOG โดยการนำภาพที่ได้จากการเบลอมาเปรียบ เทียบเพื่อหาผลต่างดังสมการที่ 3.22

$$\begin{aligned}
G(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\
&= L(x, y, k\sigma) - L(x, y, \sigma)
\end{aligned} \tag{3.22}$$

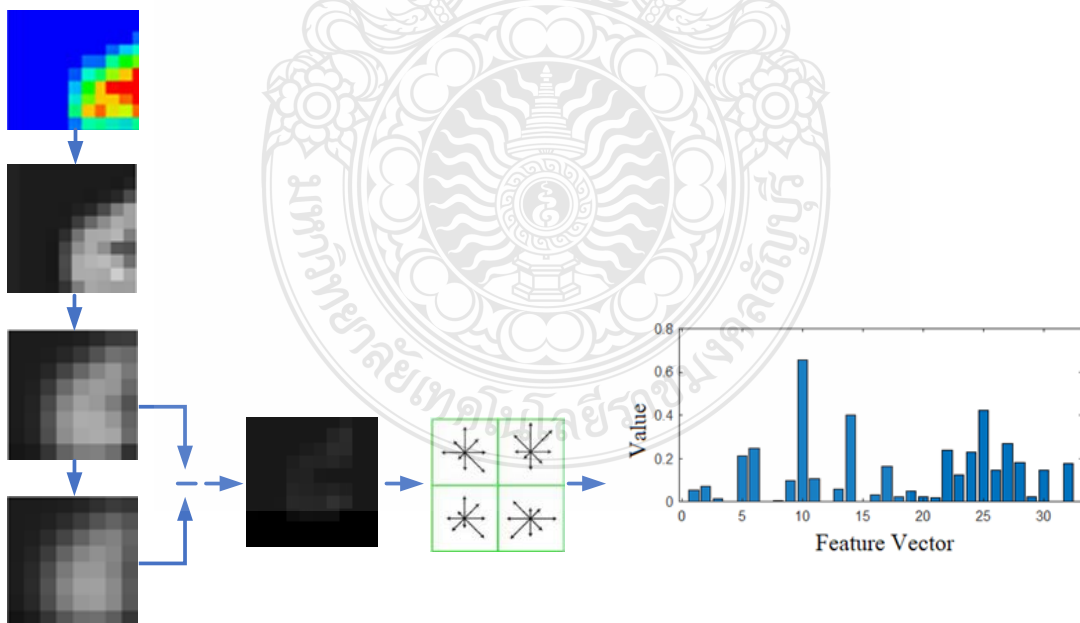
จากนั้นกำหนดจุดสนใจหรือ Keypoint เนื่องจากภาพมีขนาดเล็กจึงกำหนดจุดสนใจเป็น pixel กึ่งกลางภาพแล้วขั้นตอนถัดมาจะเป็นการหาค่าขนาด $m(x, y)$ และ ทิศทาง $\theta(x, y)$ ของ gradient ที่อยู่ในขอบเขตของ Keypoint เพื่อกำหนดขนาดและทิศทางของ Gradient ที่ Keypoint นั้นๆ โดยใช้สมการที่ 3.23 และ 3.24

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (3.23)$$

$$\theta(x, y) = \tan^{-1}((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))) \quad (3.24)$$

เมื่อ $m(x, y)$ คือ ขนาดของเกรเดียน และ คือ $\theta(x, y)$ ทิศทางของ Keypoint

แล้วสร้าง Keypoint Descriptor เนื่องจากภาพมีขนาดเล็กคือขนาด 8×8 pixel ในงานนี้จึงปรับการสร้าง Keypoints descriptor โดยการกำหนด Keypoint เป็นกึ่งกลางภาพแล้วทำการแบ่ง Windows เป็นขนาด 4×4 ทั้งหมด 4 จุด แล้วคำนวณหาค่าเกรเดียนและทิศทางบริเวณรอบๆ Keypoint ซึ่งจะทำได้ค่าขนาดและทิศทางของแต่ละ pixel แล้วแบ่งออกเป็นพื้นที่ย่อยขนาด Window 4×4 ดังรูปที่ 3.30 และสร้าง Histogram ของทิศทาง (Orientation Histograms) ขนาด 8 bin ($b=8$) โดยกำหนดจาก bin ละ 45° จากนั้นหาผลรวมของขนาดในแต่ละ bin ซึ่งจะได้เวกเตอร์ของแต่ละ Keypoint ที่มีขนาดเท่ากับ 32 ข้อมูล ($b=32$) ดังรูปที่ 3.31 ดังนั้น Histogram ที่ได้คือตัวแทนของภาพ Sub image ในแต่ละภาพ



รูปที่ 3.31 การสร้าง Feature โดยใช้วิธี SIFT

$$v = b1, b2, b3, \dots, b32 \quad (3.25)$$

เมื่อ v คือเวกเตอร์ข้อมูลที่เป็นตัวแทนของภาพ เมื่อ b คือค่าผลรวมของในแต่ละ bin

3. PF [82] หรือการแปลงโพล่า 2 มิติ ซึ่งจะได้ Polar Image ด้วยสมการที่ 3.26 จากนั้นขั้นตอนถัดไปก็ทำการแปลงฟูเรียร์ 2 มิติ (2-D Fourier Transform) จากนั้นก็ทำ Descriptor ซึ่งจะได้ Feature Vector ที่เป็นตัวแทนของข้อมูลสำหรับทำวิเคราะห์ดังรูปที่ 3.32

$$I_p = \{f(\theta, y) | 0 \leq r \leq R, 0 \leq \theta \leq 2\pi\} \quad (3.26)$$

$$x = r \cos \theta + W/2, y = r \sin \theta + H/2$$

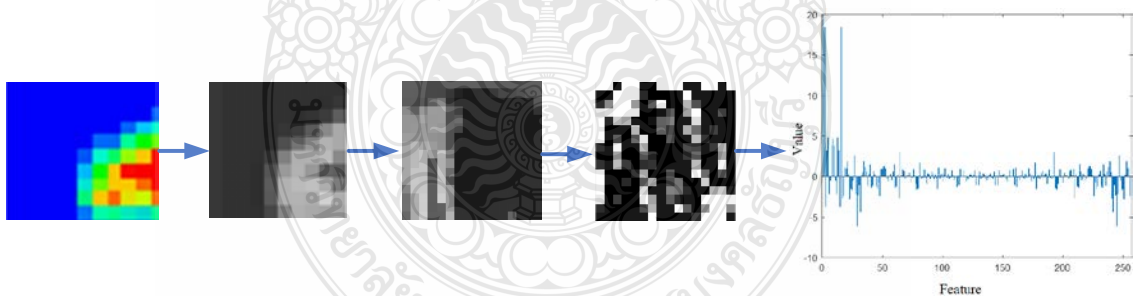
การแปลงฟูเรียร์เมื่อได้ภาพเชิงขั้วขั้นตอนถัดไปคือการแปลงฟูเรียร์แบบไม่ต่อเนื่องแบบ 2-D ถูกนำไปใช้กับ I_p โดย ผลการแปลงจะได้ดังสมการที่ 3.27

$$PF_2(\varphi, \rho) = \sum_r^{R-1} \sum_i^{T-1} f(r, \theta_i) \exp \left[-j2\pi \left(\frac{r}{R} \rho + \theta_i \varphi \right) \right] \quad (3.27)$$

$$0 \leq r \leq R, \theta_i = 2\pi i / T, 0 \leq \rho < R, 0 \leq \varphi < T$$

เมื่อ ρ และ φ คือ ρ_{th} ความถี่ในหน่วยเรเดียนและ φ_{th} คือความถี่ปกติ

การสร้างเวกเตอร์ของคุณลักษณะภาพ ซึ่งในการสร้างเวกเตอร์คุณลักษณะภาพ จะโดยใช้ Cartesian grids สามารถหาได้จากสมการ 3.28



รูปที่ 3.32 การสร้าง Feature โดยใช้วิธี PF

$$v = \{PF_2(00,01), PF_2(00,01), \dots, PF_2(AF, RF)\} \quad (3.28)$$

เมื่อ v คือเวกเตอร์ข้อมูลที่เป็นตัวแทนของภาพ เมื่อ PF คือค่าขนาดของภาพที่แปลง PF ในแต่ละตำแหน่ง

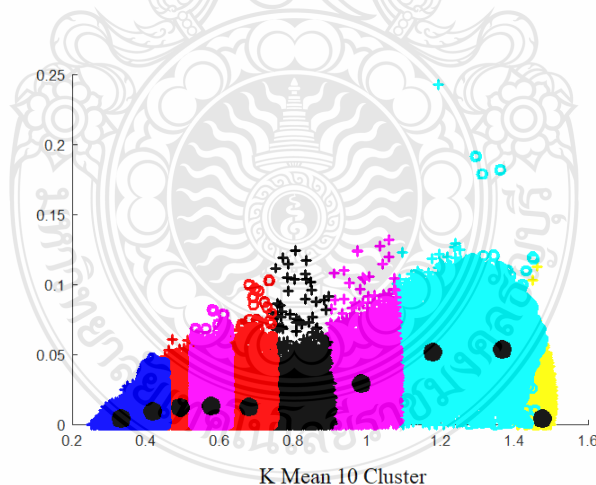
3. การทำ K-Mean Clustering

จากนั้นเมื่อได้ Feature ที่ได้จากการทำ Descriptor ทั้ง 3 วิธี ในขั้นตอนถัดไปเป็นการจัดกลุ่มของ Feature ด้วยวิธี K-Mean Clustering [84] โดยในงานวิจัยนี้ทำการทดสอบด้วยการจัดกลุ่มเป็นจำนวน Cluster 3 ขนาด คือ จำนวน 10 Cluster, 20 Cluster และ 30 Cluster การจัดกลุ่มของเซตตัวอย่างที่ได้จากการทำ Descriptor ของภาพที่นำมาใช้ในการ Training 2000 ภาพ ซึ่งแบ่งเป็น 16 ภาพย่อยต่อ Tactile image หนึ่งภาพ จะได้ Feature Vector ทั้งหมด 32,000 ชุด และนำมาจัดเรียงเป็น $\{v_1, v_2, v_3, \dots, v_n\}$ เมื่อทำการจัดกลุ่มข้อมูลวิธีการจัดกลุ่มจะทำการจัดข้อมูลที่คล้ายคลึงกันให้อยู่ในกลุ่มเดียวกัน ผลลัพธ์ของการจัดกลุ่มคือกลุ่มของตัวอย่างและ Centroid ของกลุ่ม ซึ่งมักเป็นตัวแทนของกลุ่มและมีที่อยู่ตรงกลาง สามารถสร้างฟังก์ชันวัตถุประสงค์สำหรับการจัดกลุ่มนี้ได้จากการหาค่าที่น้อยที่สุดของระยะห่างของตัวอย่างและ Centroid ของแต่ละกลุ่ม c_k ดังสมการที่ 3.29

$$J(r, c) = \sum_{i=1}^n \sum_{k=1}^K r_{ik} \|v_i - c_k\|^2 \quad (3.29)$$

โดย $r_{ik} \in \{0,1\}$ เป็นตัวแปรที่บอกความเป็นสมาชิกของกลุ่มที่ k ของตัวอย่างที่ i เมื่อ v_i ถูกกำหนดให้อยู่ในกลุ่มที่ k จะมีค่าเป็น 1 และมีค่าเป็น 0 สำหรับกลุ่มอื่นๆ ที่ไม่ใช่กลุ่ม k

เมื่อจัดกลุ่มผลการจัดกลุ่มและ Centroid ของกลุ่มแสดงดังรูปที่ 3.33 เมื่อกำหนด Cluster=10

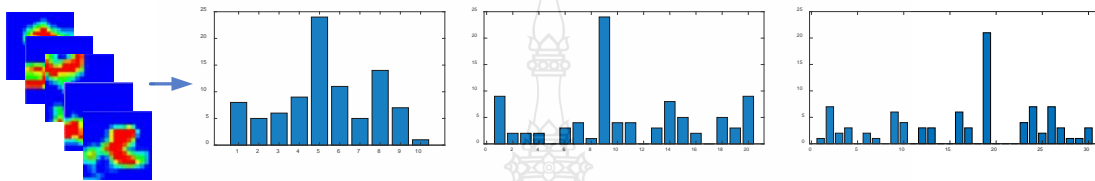


รูปที่ 3.33 ผลของการจัดกลุ่มด้วยวิธี K-Mean Clustering เมื่อกำหนดขนาด Cluster=10

4. การสร้าง Histogram

การสร้าง Histogram เพื่อเป็นตัวแทนของข้อมูล ในการทดสอบเพื่อทำนายผลของวัตถุที่ถูกจับจะเป็นการทำนายผลจากการเปรียบเทียบ Histogram ของภาพชุด Training set และภาพชุด

Testing set ซึ่งในขั้นตอนนี้จะเป็นการสร้าง Histogram ของข้อมูลภาพชุด Training Set เพื่อใช้ในการ Training ให้ระบบและการสร้าง Histogram ของข้อมูลภาพชุด Testing Set เพื่อใช้ในการ Testing เพื่อทำนายผล Histogram ที่ได้จะขึ้นอยู่กับภาพที่นำเข้ามาทดสอบถูกจัดอยู่ใน Cluster ใดบ้างและมีสมาชิกเท่าไร ตัวอย่างดังรูปที่ 3.34 เมื่อนำภาพวัตถุประแจ 4 ภาพมาสร้าง Histogram ด้วยการจัด Cluster ขนาด 10 Cluster พบว่าข้อมูลถูกจัดใน Cluster ที่ 5 มากที่สุด แต่เมื่อนำมาจัดด้วยขนาด 20 Cluster พบว่าข้อมูลถูกจัดใน Cluster 9 มากที่สุดเป็นต้น Histogram ที่ได้จะเป็นตัวแทนของข้อมูลในแต่ละ Class เพื่อนำไปทดสอบด้วยการทำ Classification ต่อไป



รูปที่ 3.34 การสร้าง Histogram เมื่อทำการแบ่ง Cluster เป็น 10 Cluster, 20 Cluster และ 30 Cluster

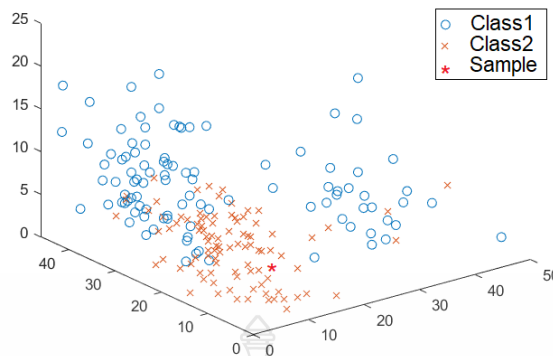
5. การทำ Classification

1. KNN [84] เป็นอัลกอริทึมในการจำแนกข้อมูลโดยการวัดระยะทางเปรียบเทียบกับข้อมูลในชุด Training ระยะทางที่ได้จากการคำนวณเพื่อหาระยะทางของตัวอย่างกับข้อมูลในชุด Training จะถูกจัดเรียงลำดับและหา Class ของชุดข้อมูลที่ใกล้กับตัวอย่างมากที่สุด โดยกำหนดจำนวนของข้อมูลที่ใกล้ที่สุดจำนวน K ตัว ซึ่งในงานวิจัยนี้กำหนดขนาด $K=10$ และวิธี Euclidean Distance ในการคำนวณระยะทางแบบยูคลิดจากจุด $p(p_1, p_2, p_3, \dots, p_n)$ ไปจุด $q(q_1, q_2, q_3, \dots, q_n)$ ในปริภูมิ n มิติสามารถหาได้จากสมการที่ 3.30

$$D_{Euclidean}(q,p) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (3.30)$$

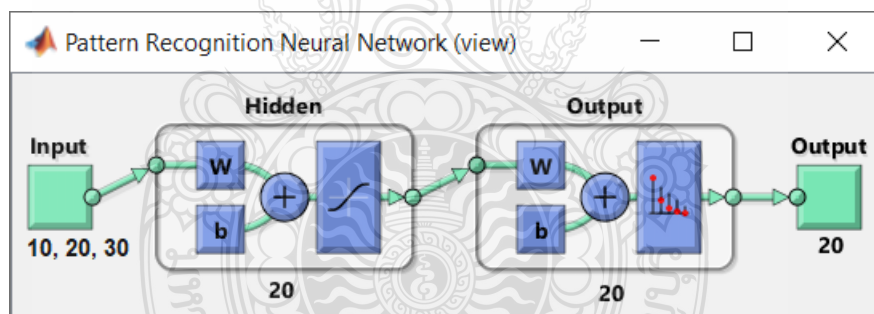
เมื่อ p คือข้อมูลที่ต้องการจำแนกและ q คือข้อมูลในแต่ละกลุ่ม

ตัวอย่างผลในการทำ Classification แสดงดังรูปที่ 3.35 ซึ่งเป็นการทำ Classification ตัวอย่างจากชุด Training จำนวน 2 Class ในปริภูมิ 3 มิติ ซึ่งผลที่ได้พบว่า Sample มีความใกล้เคียงกับ Class2 ทั้งนี้ในการทดสอบจริงได้ทำการทดสอบกับ ชุด Training จำนวน 20 Class ในปริภูมิ 10, 20 และ 30 มิติ



รูปที่ 3.35 การใช้วิธี KNN ในการทำ Classification ตัวอย่างในปริภูมิ 3 มิติ

2. ANN [86, 87] โครงข่ายประสาทเทียมที่ใช้ในการทดสอบเป็นแบบ MLP เป็นรูปแบบหนึ่งของโครงข่ายประสาทเทียมที่มีโครงสร้างเป็นแบบหลายชั้น แบ่งกระบวนการทำงานเป็น 3 Layer ประกอบด้วยชั้น Input Layer, ชั้น Hidden Layer และชั้น Output Layer โครงข่ายประสาทเทียมแบบหลายชั้นที่ใช้ในงานวิจัยนี้เป็นแบบฟีดฟอร์เวิร์ดมีโครงสร้างดังรูปที่ 3.36 ประกอบไปด้วย 3 ชั้นคือ ชั้นอินพุต, ชั้นซ่อนจำนวน 20 โหนด และชั้นเอาต์พุต 20 Class ซึ่งออกแบบมาใช้สำหรับการทดสอบการจดจำวัตถุจากการสัมผัสโดยการใช้วิธี BoF



รูปที่ 3.36 การใช้วิธี ANN ในการทำ Classification ตัวอย่างสำหรับ BoF

3. SVM [84, 85] อาศัยหลักการหาเส้นแบ่งกลุ่มหรือ Hyperplane เป็นสัมประสิทธิ์ของสมการเพื่อสร้างเส้นแบ่งกลุ่มชุดข้อมูลชุด Training ใช้แยกคลาสของข้อมูลให้ได้จุดที่ดี Hyperplane มีทั้งแบบเชิงเส้นและที่ไม่เป็นเชิงเส้นโดยอาศัยการปรับค่าสมการด้วย Kernel สำหรับงานวิจัยนี้ใช้ Kernel แบบเชิงเส้นซึ่งแสดงดังสมการที่ 3.31

$$f(x) = w^T x + b \quad (3.31)$$

เมื่อ w คือเวกเตอร์น้ำหนักและ b คือค่าไบอัส

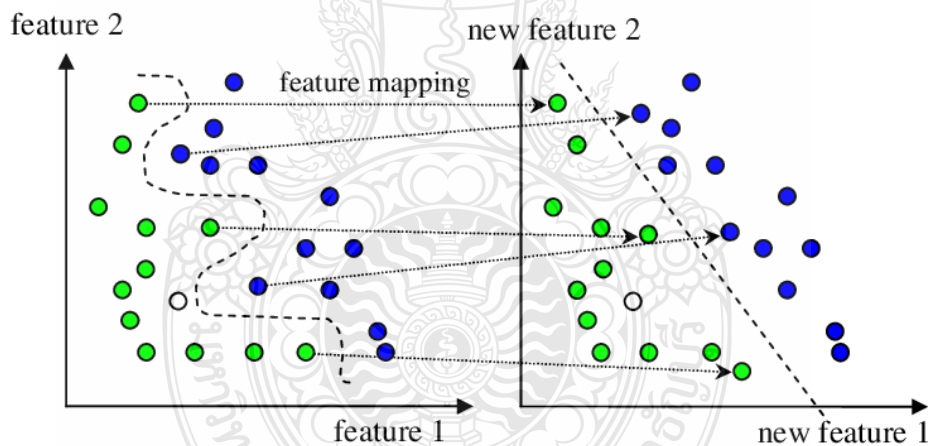
ก่อนการสร้าง Hyperplane สถาปัตยกรรมของ SVM มีการแมปข้อมูลจากปริภูมิอินพุตไปอยู่ในปริภูมิที่สูงขึ้นเรียกว่าปริภูมิลักษณะเด่นด้วยฟังก์ชันเคอร์เนล เพื่อให้สามารถหาฟังก์ชันเชิงเส้นที่เหมาะสมกับ Hyperplane การแมปข้อมูลด้วยฟังก์ชันเคอร์เนลแสดงดังสมการที่ 3.32 จะได้ว่าระนาบ Hyperplane เป็นสมการที่ 3.33

$$K(X_i, X_j) = (\varphi(X_i) \cdot \varphi(X_j)) \quad (3.32)$$

$$f(x) = (w \cdot K(x_j, x_j)) + b \quad (3.33)$$

เมื่อ K คือฟังก์ชันเคอร์เนล, φ คือฟังก์ชันการแมป, w คือเวกเตอร์น้ำหนักและ b คือค่าไบอัส

การแปลงดังกล่าวสามารถทำให้ข้อมูลที่ไม่เป็นเชิงเส้นให้เป็นข้อมูลที่อยู่ในรูปเชิงเส้นเพื่อให้สามารถแบ่งแยกได้ดังแสดงในรูปที่ 3.37 ซึ่งจากนั้นก็ยังสามารถจัด Class ตัวอย่างได้จาก Hyperplane ที่ได้ ในการทดลองได้ทำการทดสอบกับตัวอย่าง 20 Class ซึ่งมี Input ที่ได้จากการแบ่ง Cluster เป็น 10 Cluster, 20 Cluster และ 30 Cluster หรือมิติของ Input มี 10, 20 และ 30 มิติ และเป็นข้อมูลที่ไม่เป็นเชิงเส้น

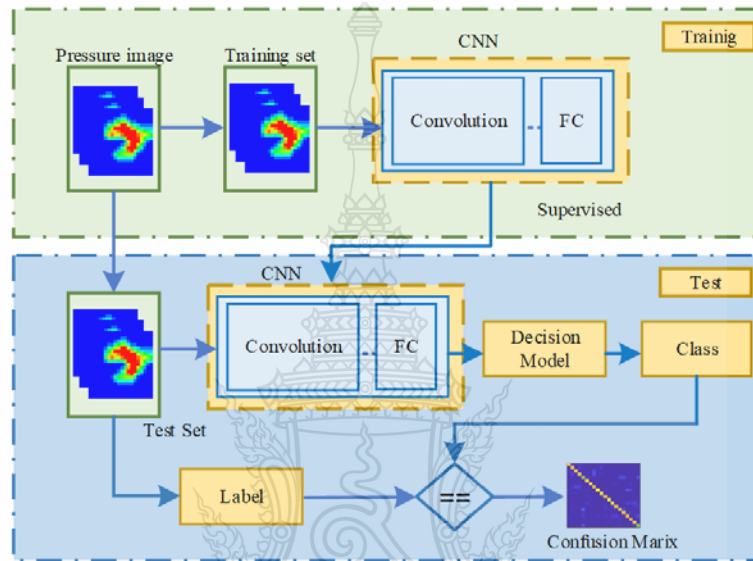


รูปที่ 3.37 การแปลงข้อมูลที่ไม่เป็นเชิงเส้นให้เป็นข้อมูลเชิงเส้นด้วยฟังก์ชันเคอร์เนล [84]

3.4.4 การทดลองอัลกอริทึมแบบ DCNN สำหรับมือหุ่นยนต์

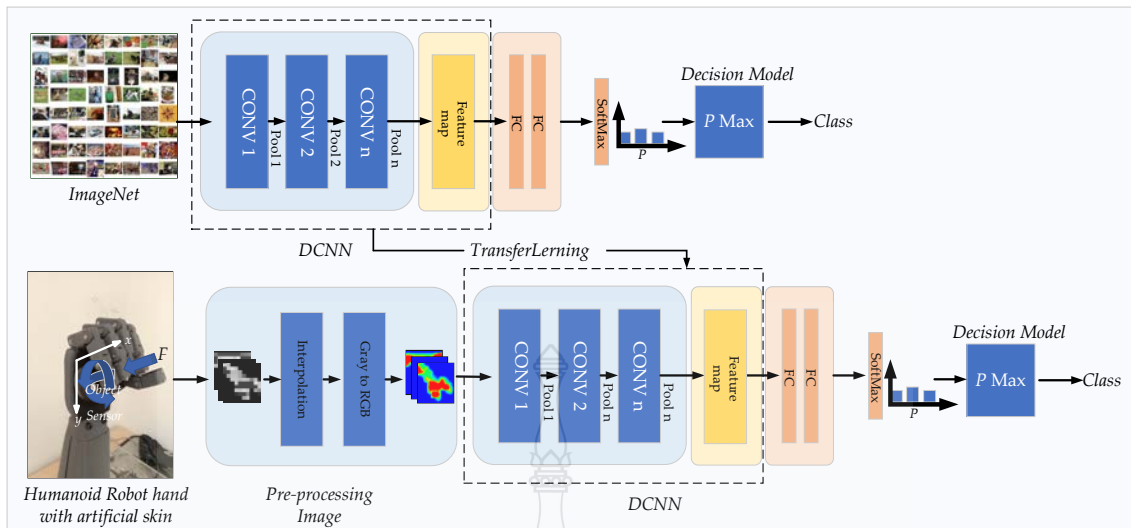
การทดลองอัลกอริทึมแบบ DCNN มี Work Flow ดังรูปที่ 3.38 ขั้นตอนการทำงานเริ่มจากการแบ่ง Data Set ออกเป็น 2 ชุด เช่นเดียวกัน ในส่วนของ Training Set ถูกนำมาใช้ training ให้กับ Network จากนั้นนำ Network ที่เรียนรู้แล้วไปใช้ทดสอบกับ Test Set ในงานนี้ได้ทดสอบกับ DCNN ทั้งหมด 19 Model ประกอบด้วย AlexNet [88], VGG16 [90], VGG19 [90], GoogLeNet [91], ResNet18 [92], ResNet50 [92], ResNet101 [92], Place365GoogLeNet [93], InceptionV3 [94], EfficientNetB0 [95], SqueezeNet [96], InceptionResNetV2 [97], DenseNet201 [98],

DarkNet19 [99], DarkNet53 [100], XceptionNet [101], NASNetMobile [102], ShuffleNet [103] และ MobileNetV2 [104] ทั้งหมดเป็น DCNN ที่ใช้กันอย่างแพร่หลายใน Computer Vision และได้ถูกนำมาใช้ในการจดจำวัตถุจาก Tactile Image และพบว่าให้ประสิทธิภาพสูง สำหรับในงานนี้ได้นำมาประยุกต์ใช้กับการจดจำวัตถุจากการสัมผัส โดย Network ทั้งหมดที่นำมาใช้ได้ใช้วิธี Transfer Learning Method [105] ซึ่งถูกนำมาใช้ในการจดจำวัตถุจากการสัมผัสโดยใช้เซ็นเซอร์แบบติดฝ่ามือจุดเดียวโดยใช้เซ็นเซอร์แบบใหม่ที่พัฒนา



รูปที่ 3.38 Work Flow สำหรับการจดจำวัตถุจากการสัมผัสด้วยวิธี DCNN

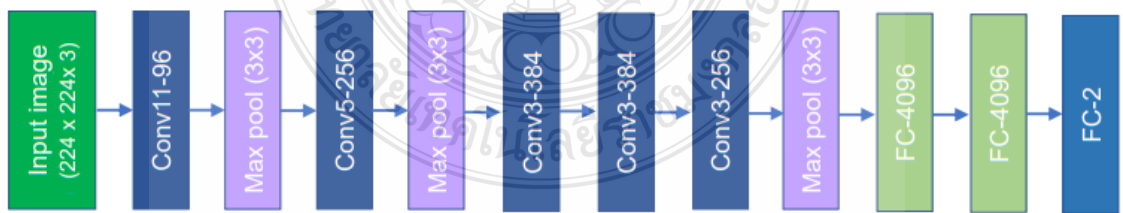
องค์ประกอบที่เกี่ยวข้องกับ DCNN แสดงไว้ในรูปที่ 3.39 ซึ่งใช้กับการจดจำวัตถุสัมผัสโดยใช้ Transfer Learning Method จาก ImageNet DCNN ประกอบด้วยฟังก์ชัน SoftMax 20 เวกเตอร์ เลเยอร์สุดท้าย (Fully Connected Layer) เพื่อคำนวณการกระจายความน่าจะเป็นสำหรับ Class ของวัตถุ 20 Class และทำนายผลจากค่าความน่าจะเป็นสูงสุดที่ได้จากฟังก์ชัน SoftMax



รูปที่ 3.39 DCNN Transfer Learning ที่ใช้ในงานวิจัยนี้

DCNN ที่นำมาใช้ในงานวิจัยนี้เป็นการประยุกต์ใช้ DCNN แบบ Transfer Learning DCNN ซึ่งสถาปัตยกรรมของ DCNN แต่ละ Model มีโครงสร้างดังนี้

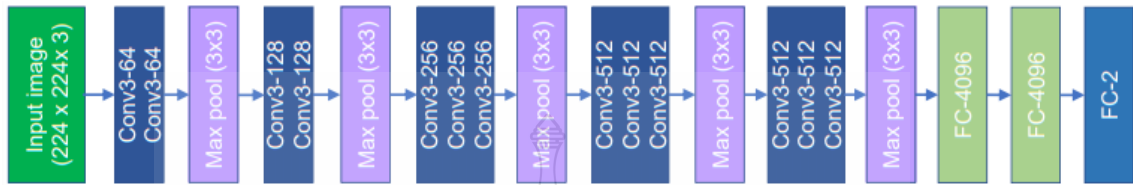
1. AlexNet [88] ประกอบไปด้วยชั้นทั้งหมด 8 ชั้น ขนาดของภาพ Input กำหนดขนาดที่ $224 \times 224 \times 3$ pixels (ภาพสี) โครงสร้างประกอบด้วยชั้นคอนโวลูชัน 5 ชั้นและชั้น Fully connected 3 ชั้น โดยชั้นคอนโวลูชัน 1 (Conv1) มีขนาด 96 เคอร์เนล ชั้นคอนโวลูชัน 2 (Conv2) มี 256 เคอร์เนล ชั้นคอนโวลูชัน 3 (Conv3) มี 384 เคอร์เนล ชั้นคอนโวลูชัน 4 (Conv4) มี 384 เคอร์เนล และ ชั้นคอนโวลูชัน 5 (Conv5) มี 256 เคอร์เนล โดยที่ Conv 1-5 ใช้ฟังก์ชัน Rectified Linear Units (ReLU) เป็นฟังก์ชันกระตุ้น (Activation Function) และในส่วนของ 3 ชั้นสุดท้ายเป็น Fully connected layer ซึ่งชั้นสุดท้ายเป็นชั้นคำนวณความน่าจะเป็นด้วย SoftMax Function โครงสร้างของ AlexNet แสดงดังรูปที่ 3.40 AlexNet มีพารามิเตอร์ 60 ล้านพารามิเตอร์



รูปที่ 3.40 AlexNet [88, 106]

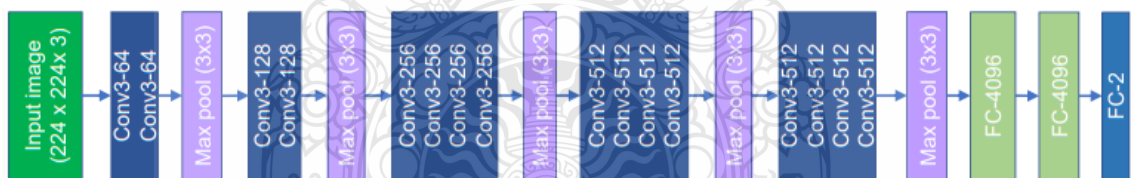
2. VGG16 [90] โครงสร้างของ VGGNet-16 ประกอบไปด้วยชั้นทั้งหมด 16 ชั้น ขนาดของภาพ Input กำหนดขนาดที่ $224 \times 224 \times 3$ pixels โครงสร้างเป็นชั้นคอนโวลูชัน (Conv) 13 ชั้น และชั้น Fully connected 3 ชั้น โดยชั้นคอนโวลูชันแบ่งออกเป็น 5 ชุด ชุดแรกมี 2 ชั้น (Conv1-2) มี

ขนาด 64 เคอร์เนล ชุดที่ 2 มี 2 ชั้น (Conv3-4) มีขนาด 128 เคอร์เนล ชุดที่ 3 มี 3 ชั้น (Conv5-7) มีขนาด 256 เคอร์เนล ชุดที่ 4 มี 3 ชั้น (Conv8-10) มีขนาด 512 เคอร์เนล และชุดที่ 5 มี 3 ชั้น (Conv11-13) มีขนาด 512 เคอร์เนล และในส่วนของ 3 ชั้นสุดท้ายเป็น Fully connected layer โครงสร้างของ VGG16 แสดงดังรูปที่ 3.41 VGG16 มีพารามิเตอร์ 138.3 ล้านพารามิเตอร์



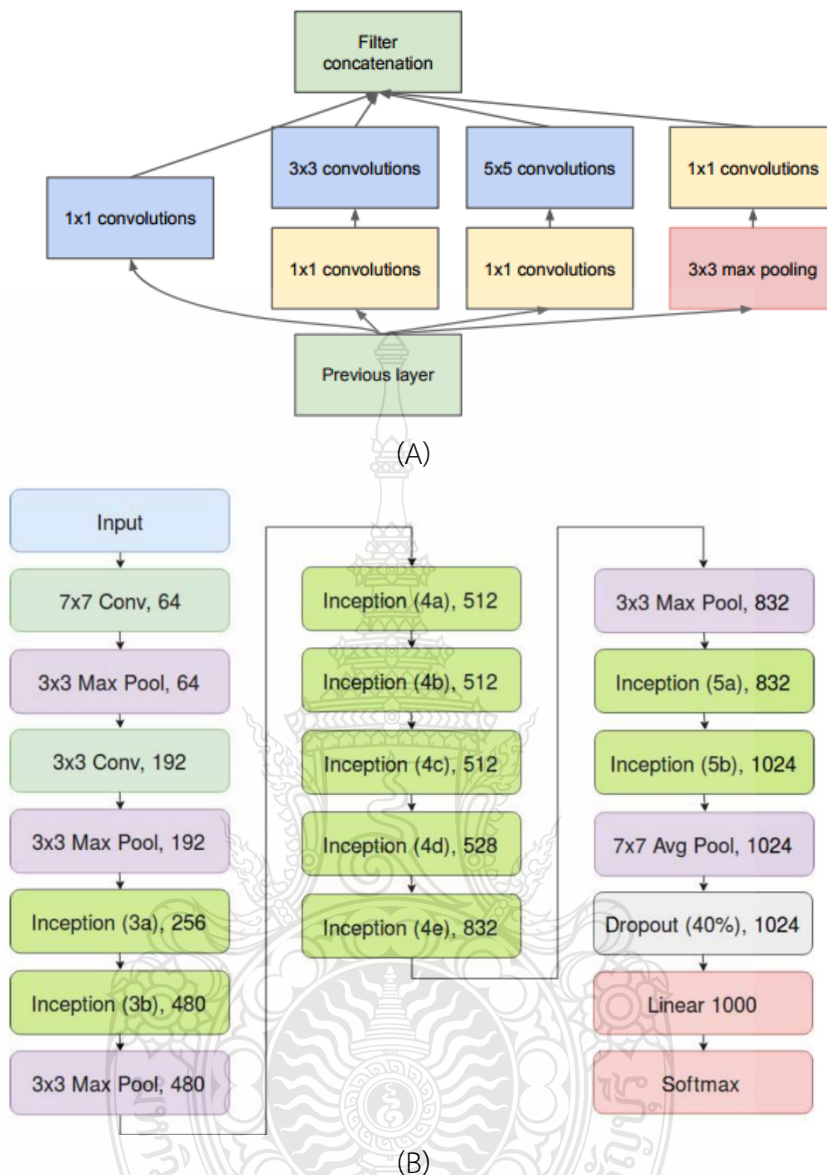
รูปที่ 3.41 VGG16 [90, 106]

3. VGG19 [90] โครงสร้างของ VGGNet-19 ประกอบไปด้วยชั้นทั้งหมด 19 ชั้น ขนาดของภาพ Input กำหนดขนาดที่ $224 \times 224 \times 3$ pixels โครงสร้างเป็นชั้นคอนโวลูชัน (Conv) 16 ชั้น และชั้น Fully connected 3 ชั้น โดยชั้นคอนโวลูชันแบ่งออกเป็น 5 ชุด ชุดแรกมี 2 ชั้น (Conv1-2) มีขนาด 64 เคอร์เนล ชุดที่ 2 มี 2 ชั้น (Conv3-4) มีขนาด 128 เคอร์เนล ชุดที่ 3 มี 4 ชั้น (Conv5-8) มีขนาด 256 เคอร์เนล ชุดที่ 4 มี 4 ชั้น (Conv9-12) มีขนาด 512 เคอร์เนล และชุดที่ 5 มี 4 ชั้น (Conv13-16) มีขนาด 512 เคอร์เนล และในส่วนของ 3 ชั้นสุดท้ายเป็น Fully connected layer โครงสร้างของ VGG19 แสดงดังรูปที่ 3.42 VGG19 มีพารามิเตอร์ 143.6 ล้านพารามิเตอร์



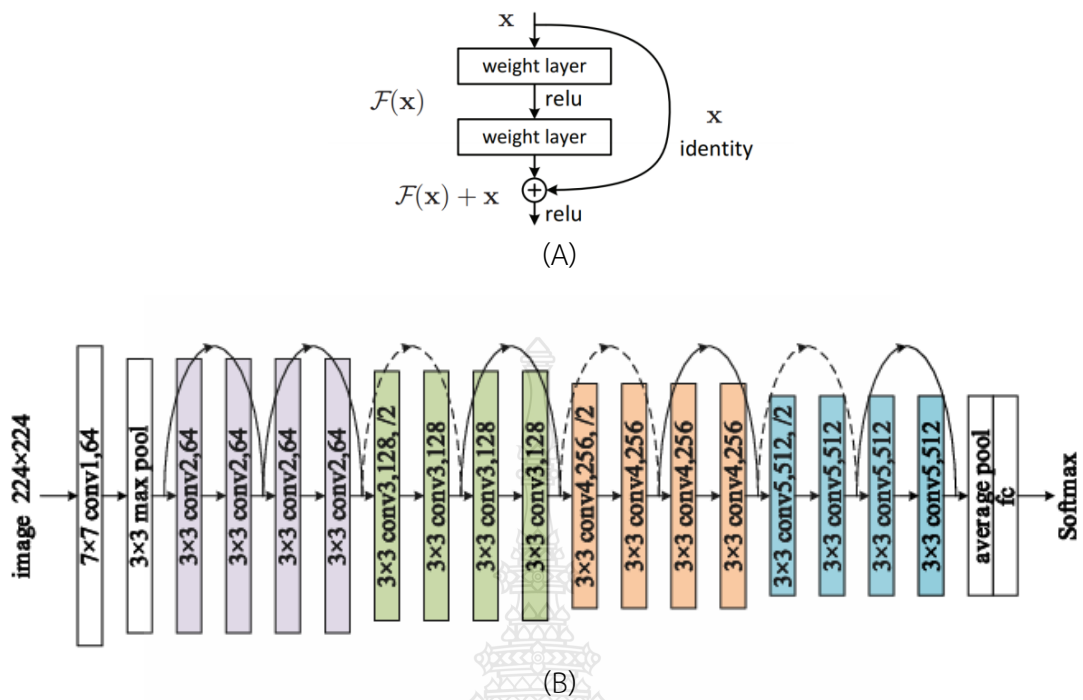
รูปที่ 3.42 VGG19 [90, 106]

4. GoogLeNet [91] หรือ InceptionV1 โครงสร้างแบบ GoogLeNet ออกแบบมาให้มีชั้นรวมทั้งสิ้น 22 ชั้น กำหนดขนาดของภาพ Input ที่ขนาด $224 \times 224 \times 3$ pixels สถาปัตยกรรมของโครงข่ายใช้โครงสร้างแบบ Inception Architecture ซึ่งมีโครงสร้างดังรูปที่ 3.43 (A) โครงสร้าง Inception Architecture มีการทำงานแบบขนาน ใช้เคอร์เนล 3 ขนาดในการคำนวณ ซึ่งมีขนาดเคอร์เนลเป็น 1×1 , 3×3 และ 5×5 โครงสร้างแสดงดังรูปที่ 3.43 (B) โดยมีชั้นแรกเป็นชั้นคอนโวลูชันขนาด 64 เคอร์เนล ชั้นถัดมาเป็นคอนโวลูชันขนาด 192 เคอร์เนล 2 ชุด จากนั้นเป็นชั้นคอนโวลูชันแบบ Inception จำนวน 9 ชั้นที่ประกอบด้วยชั้นคอนโวลูชันภายใน และชั้นสุดท้ายเป็นชั้นคำนวณความน่าจะเป็นด้วย SoftMax Function GoogLeNet มีพารามิเตอร์ 7 ล้านพารามิเตอร์



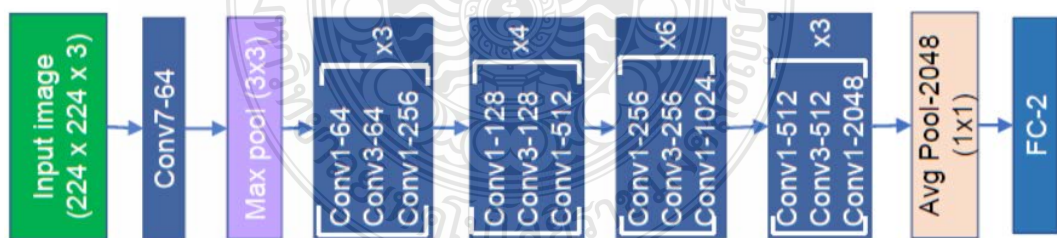
รูปที่ 3.43 GoogLeNet [91, 107]

5. ResNet18 [92] มาจาก Deep residual Network โครงสร้างของ Deep residual Network มีการรวมผลของข้อมูลก่อนทำ Convolution กับผลที่ได้จากการทำ Convolution แล้วดังแสดงในรูปที่ 3.44 (A) โครงสร้างของ ResNet18 มีจำนวนชั้นทั้งหมด 18 ชั้น ขนาด Input ของ ResNet18 กำหนดที่ $224 \times 224 \times 3$ pixels ประกอบด้วยชั้นแรกเป็นชั้นคอนโวลูชันขนาด 64 เคอร์เนล ถัดมาชั้น Deep residual Network 16 ชั้น และชั้นสุดท้ายเป็นชั้นคำนวณความน่าจะเป็นด้วย SoftMax Function โครงสร้างดังรูปที่ 3.44 (B) ResNet18 มีพารามิเตอร์ 25.6 ล้านพารามิเตอร์



รูปที่ 3.44 ResNet18 [92 108]

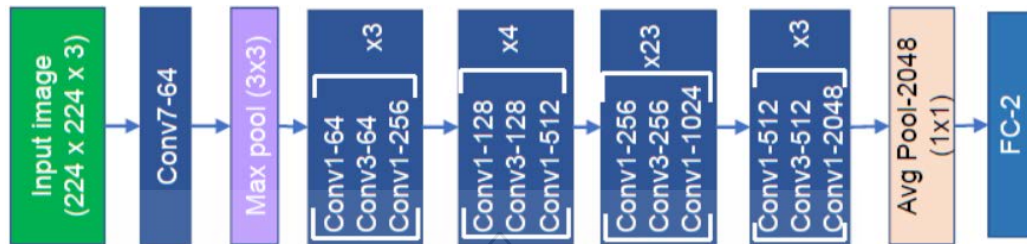
6. ResNet50 [92] มีโครงสร้างเป็น Deep residual Network เหมือน ResNet18 โครงสร้างมีชั้นรวม 50 ชั้น ขนาด Input ของ ResNet50 กำหนดที่ $224 \times 224 \times 3$ pixels มีชั้น Convolution 49 ชั้น ประกอบด้วยชั้นแรกเป็นชั้นคอนโวลูชันขนาด 64 เคอร์เนล และถัดมาเป็น Deep residual Network 48 ชั้น และชั้นสุดท้ายเป็นชั้นคำนวณความน่าจะเป็นด้วย SoftMax Function โครงสร้างดังรูปที่ 3.45 ResNet50 มีพารามิเตอร์ 25.6 ล้านพารามิเตอร์



รูปที่ 3.45 ResNet50 [92, 106]

7. ResNet101 [92] มีโครงสร้างเป็น Deep residual Network เหมือน ResNet18 และ ResNet50 ขนาดโครงสร้างมีชั้นรวม 101 ชั้น Input ของ ResNet101 กำหนดที่ $224 \times 224 \times 3$ pixels โครงสร้างมีชั้น Convolution 100 ชั้น ประกอบด้วยชั้นแรกเป็นชั้นคอนโวลูชันขนาด 64 เคอร์

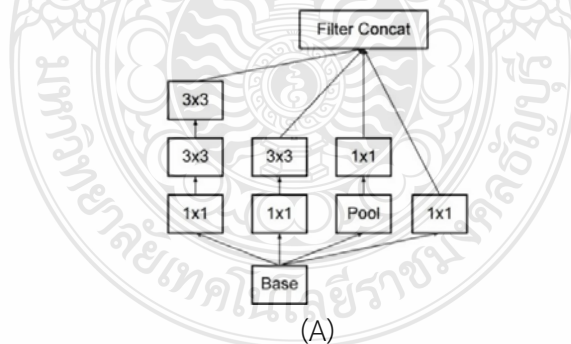
นอล และถัดมาเป็น Deep residual Network 99 ชั้น และชั้นสุดท้ายเป็นชั้นคำนวณความน่าจะเป็นด้วย SoftMax Function โครงสร้างดังรูปที่ 3.46 ResNet101 มีพารามิเตอร์ 40 ล้านพารามิเตอร์



รูปที่ 3.46 ResNet101 [92, 106]

8. Places365-GoogLeNet [93] มีโครงสร้างเช่นเดียวกับ GoogLeNet แต่เป็น Network ที่ปรับปรุงสำหรับการใช้งานกับการจดจำสถานที่ซึ่งถูก Pre-train มาด้วยข้อมูลภาพที่เป็นสถานที่ต่างๆ

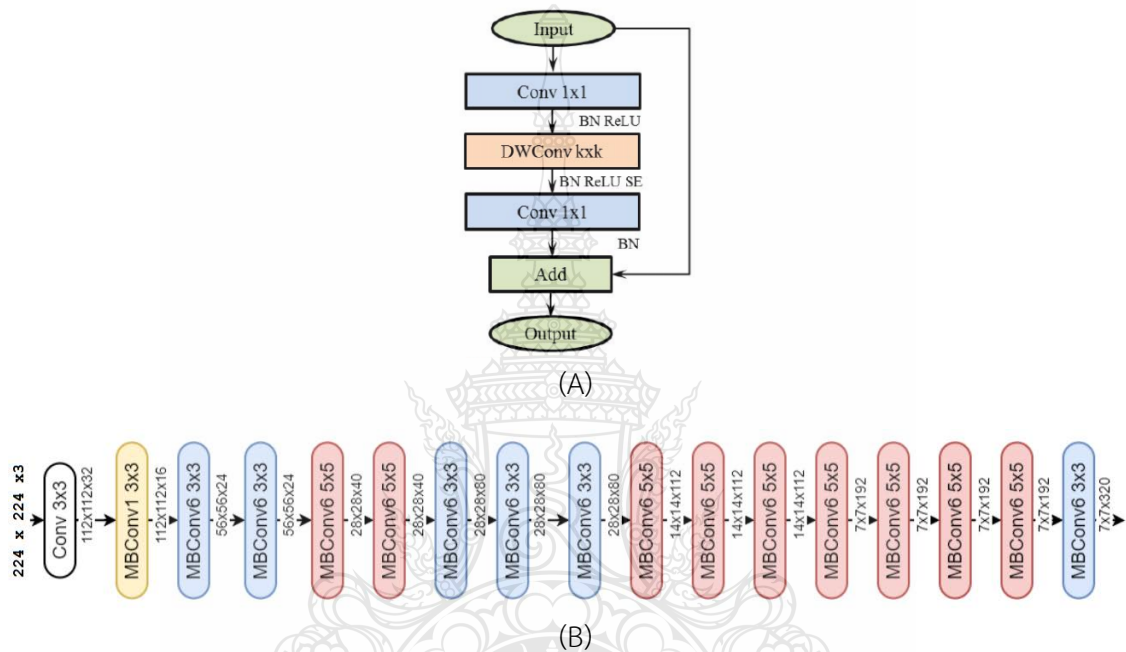
9. InceptionV3 [94] โครงสร้างของ InceptionV3 เป็นการพัฒนาต่อยอดจาก InceptionV1 และ InceptionV2 มีชั้นรวม 48 ชั้น กำหนดขนาด Input เป็น $299 \times 299 \times 3$ pixels โดยมีโครงสร้างของส่วน Inception และจำนวนชั้นที่เปลี่ยนแปลงไป โดยการปรับปรุง convolutions เดิมที่มีขนาด 5×5 ลงเหลือ 3×3 pixels ดังรูปที่ 3.47 (A) ในส่วนของชั้น Inception แบ่งเป็น Inception Module A, Inception Module B, Inception Module C, Inception Module D และ Inception Module E การปรับปรุงดังกล่าวทำให้พารามิเตอร์ลดลงจากเดิมแต่ยังคงมีประสิทธิภาพสูง โครงสร้างของ InceptionV3 แสดงดังรูปที่ 3.46 (B) มีพารามิเตอร์ 23.6 ล้านพารามิเตอร์



(B)

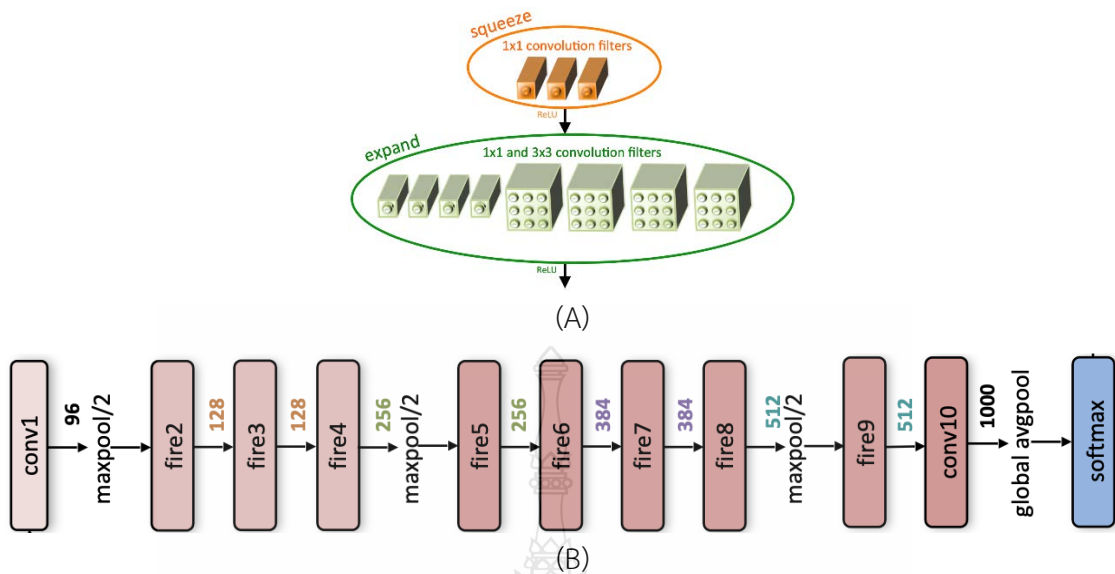
รูปที่ 3.47 InceptionV3 [94, 106]

10. EfficientNetB0 [95] คือกลุ่มของโครงข่ายประสาทเทียมที่มีโครงสร้างพื้นฐานจาก Mobile Inverted Bottleneck Conv Block (MBconv) [104] ดังรูปที่ 3.48 (A) สถาปัตยกรรมของ EfficientNetB0 แสดงในรูปที่ 3.48 (B) กำหนดขนาด Input เป็น $224 \times 224 \times 3$ pixels มีชั้นรวม 28 ชั้น มีพารามิเตอร์ 4.2 ล้านพารามิเตอร์ โดยแนวคิดหลักของการออกแบบสถาปัตยกรรม EfficientNet คือการลดขนาดของ Network ลง เป็นผลให้ประสิทธิภาพของ EfficientNet ในด้านความเร็วและการลดการใช้ทรัพยากรดีกว่า Network ทั้งหมดรุ่นก่อนหน้า



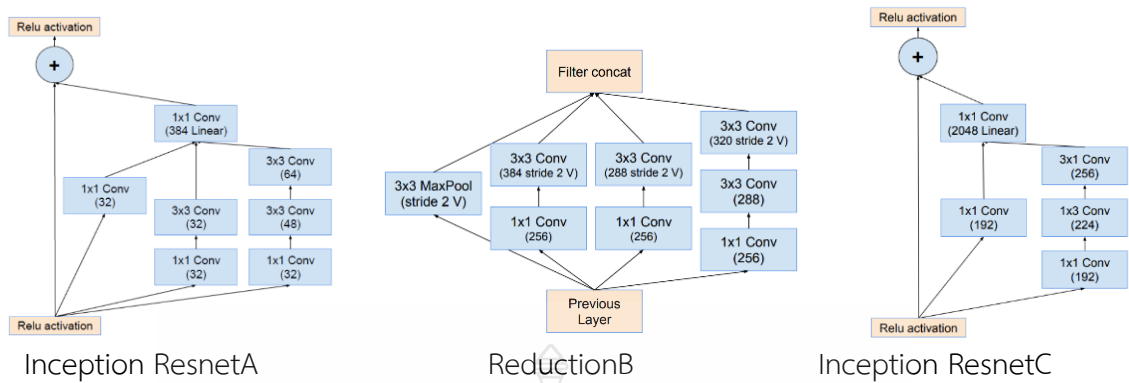
รูปที่ 3.48 EfficientNetB0 [95, 109]

11. SqueezeNet [96] เป็นโครงข่ายประสาทเทียมที่พัฒนาเพื่อความรวดเร็วในการทำงานเรียกว่าการ Squeeze ทำงานโดยการใช้ตัวกรองขนาด 1×1 จำนวนมากเพื่อลดจำนวน weight โครงสร้างการทำงานในหนึ่งชุดเรียกว่า Fire module ซึ่งมีลักษณะดังรูปที่ 3.49 (A) จะใช้ตัวกรองคอนโวลูชันขนาด 1×1 เพื่อทำ Squeeze ก่อนแล้วจึงขยายด้วยตัวกรองคอนโวลูชัน 1×1 และ 3×3 หลายตัว การดำเนินการลักษณะนี้สามารถลดจำนวน weight โดยรวมได้ 50 เท่าเมื่อเทียบกับ AlexNet โครงสร้างของ SqueezeNet กำหนดขนาด Input เป็น $224 \times 224 \times 3$ pixels ชั้นแรกเป็นชั้นคอนโวลูชันขนาด 96 เคอร์เนล ตามด้วยชั้น Fire module จำนวน 8 ชั้น และตามด้วยชั้นคอนโวลูชันขนาด 1×1 อีกหนึ่งชั้น และชั้นสุดท้ายเป็นชั้นคำนวณความน่าจะเป็นด้วย SoftMax Function โครงสร้างดังรูปที่ 3.49 (B) มีชั้นรวม 18 ชั้น มีพารามิเตอร์ 1.2 ล้านพารามิเตอร์

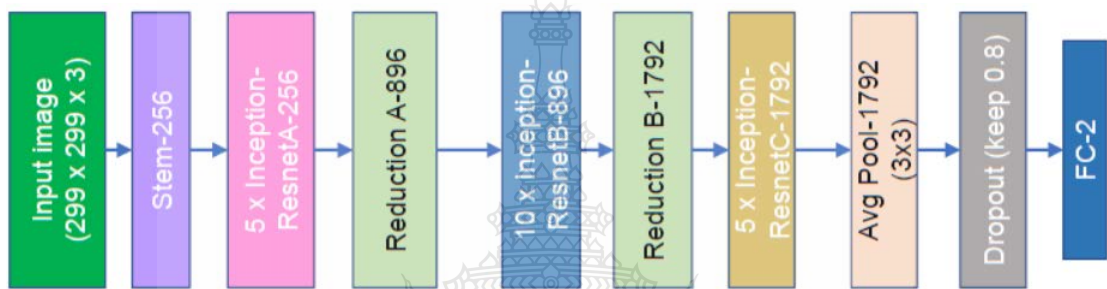


รูปที่ 3.49 SqueezeNet [96]

12. InceptionResNetV2 [97] เป็นโครงข่ายประสาทเทียมที่มีโครงสร้างผสมผสานระหว่างการทำงานของทั้งเครือข่ายแบบ Inception และแบบ Residual เครือข่ายมีความลึก 164 ชั้น กำหนดขนาด Input เป็น $299 \times 299 \times 3$ pixels ลักษณะของโครงข่าย ประกอบไปด้วยโครงสร้าง 3 แบบคือ Inception Resnet, Reduction และ Inception Resnet ซึ่งมีการกำหนดขนาดของเคอร์เนลที่แตกต่างกัน โดยสามารถแบ่งออกได้เป็นโมดูล 5 โมดูลคือ Inception ResnetA, ReductionA, Inception ResnetB, ReductionB, Inception ResnetC ดังรูปที่ 3.50 (A) โครงข่ายประสาทนี้ถูกพัฒนามาเพื่อให้มีประสิทธิภาพที่ดีแต่ใช้ทรัพยากรที่ต่ำกว่าโครงข่ายอื่น สถาปัตยกรรมโครงข่าย ประกอบด้วยส่วนแรก Stem265 เป็นชั้นคอนโวลูชันขนาด 9 ชั้น จากนั้นต่อด้วย Inception ResnetA จำนวน 5 ชั้น ต่อด้วยชั้น ReductionA และตามด้วยชั้น Inception ResnetB จำนวน 10 ชั้น ต่อด้วย ReductionB และชั้น Inception ResnetC ตามลำดับ และชั้นสุดท้ายเป็นชั้นคำนวณความน่าจะเป็นด้วย SoftMax Function โครงสร้างดังรูปที่ 3.50 (B) โครงสร้างสถาปัตยกรรมนี้มีความลึก 572 มีขนาดพารามิเตอร์ 55 ล้านพารามิเตอร์ และได้รับการพิสูจน์แล้วว่ามีความมีประสิทธิภาพที่ดีมากโดยมีต้นทุนในการคำนวณที่ค่อนข้างต่ำ



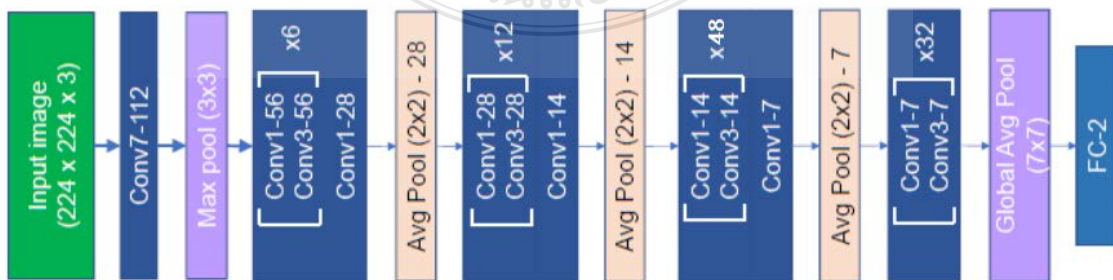
(A)



(B)

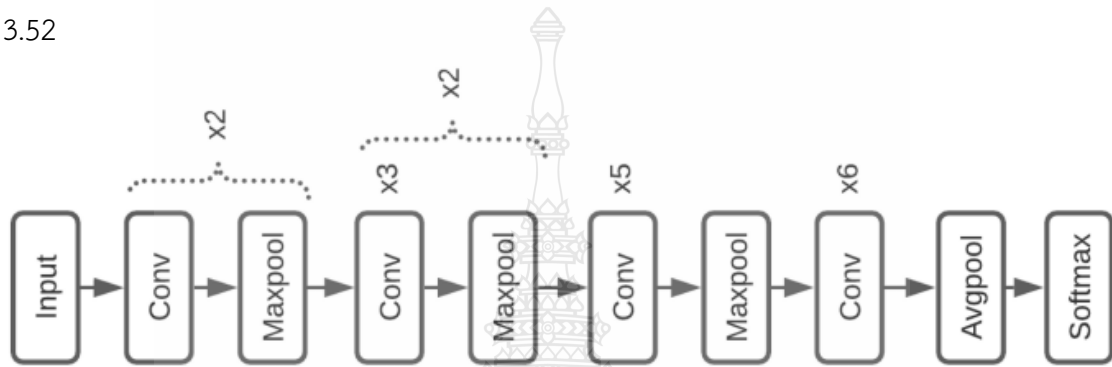
รูปที่ 3.50 InceptionResNetV2 [97, 106]

13. DenseNet201 [98] เป็นโครงข่ายประสาทเทียมที่มีความลึก 201 ชั้น เครือข่ายมีภาพอินพุตขนาด $224 \times 224 \times 3$ pixels โครงสร้างของ DenseNet จะเชื่อมต่อแต่ละเลเยอร์กับทุกๆ เลเยอร์ในรูปแบบพิตฟอร์เวิร์ด ในขณะที่โครงข่ายแบบอื่นจะมีการต่อไปแบบอนุกรม แต่โครงข่าย DenseNet จากเลเยอร์แรกจะเชื่อมต่อไปยังเลเยอร์ถัดไปทุกเลเยอร์ และเลเยอร์ถัดไปก็จะต่อในลักษณะเดียวกัน แต่จะไม่ต่อย้อนกลับมายังเลเยอร์ก่อนหน้า DenseNets มีข้อดีที่น่าสนใจหลายประการเช่นการลดปัญหา vanishing-gradient, เพิ่มการแพร่ Feature และการใช้ Feature ซ้ำ และลดจำนวนพารามิเตอร์ลงได้มาก โครงสร้างของ DenseNet201 แสดงดังรูปที่ 3.51 มี 20 ล้านพารามิเตอร์



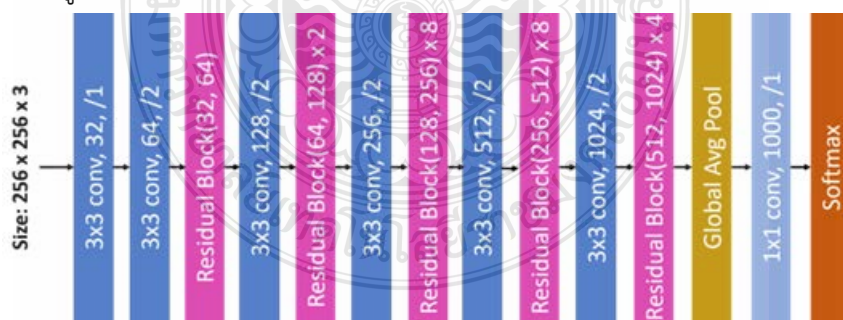
รูปที่ 3.51 DenseNet201 [98, 106]

14. DarkNet19 [99] เป็นโครงข่ายประสาทเทียมที่มีความลึก 19 ชั้น หรือเรียกอีกอย่างว่า YOLO เครือข่ายมีขนาดอินพุตภาพ $256 \times 256 \times 3$ pixels มักถูกใช้เป็นรากฐานสำหรับปัญหาการตรวจจับวัตถุและ YOLO workflows สถาปัตยกรรมโครงข่ายคล้ายคลึงกับ VGG ส่วนใหญ่จะใช้ตัวกรองและเพิ่มจำนวนช่องสัญญาณเป็นสองเท่าหลังจากทุกชั้นตอน Darknet ประกอบด้วย 19 ชั้น Convolution ส่วนใหญ่ใช้ตัวกรอง 3×3 พร้อม max-pooling layer 5 ชั้น โครงสร้างแสดงดังรูปที่ 3.52



รูปที่ 3.52 DarkNet19 [99, 110]

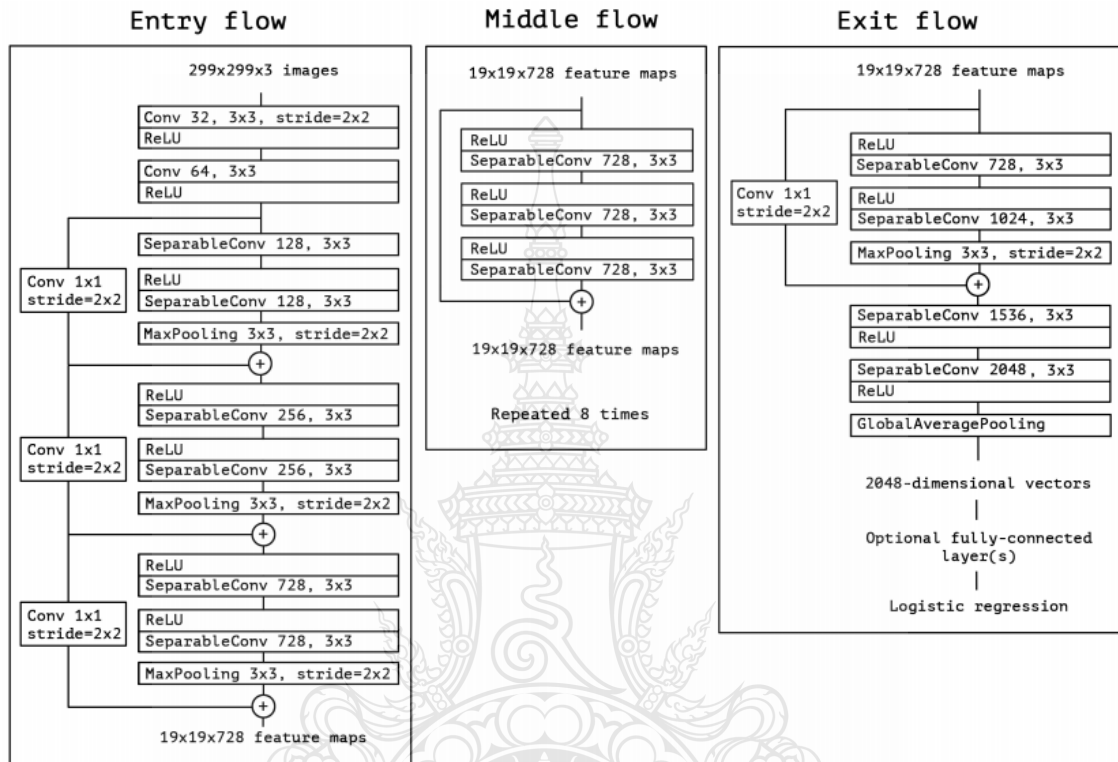
15. DarkNet53 [100] เป็นโครงข่ายประสาทเทียมที่มีความลึก 53 ชั้น หรือเรียกอีกอย่างว่า YOLOv3 เครือข่ายมีขนาดอินพุตภาพ $256 \times 256 \times 3$ pixels Darknet53 เป็นโครงข่ายประสาทเทียมที่ทำหน้าที่เป็นแกนหลักสำหรับแนวทางการตรวจจับวัตถุ สถาปัตยกรรมถูกปรับปรุงจาก Darknet19 รุ่นก่อนนั้นรวมถึงการใช้การเชื่อมต่อ residual connections และชั้นเครือข่ายเพิ่มเติม โครงสร้างแสดงดังรูปที่ 3.53



รูปที่ 3.53 DarkNet53 [100, 111]

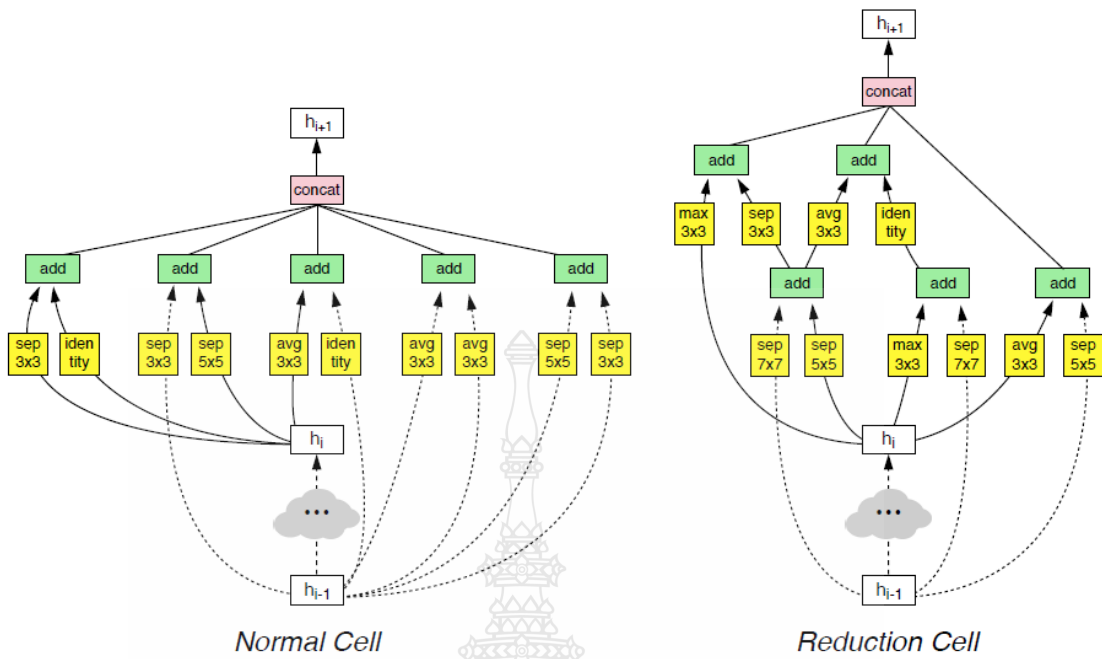
16. XceptionNet [101] มาจาก Extreme version of Inception ปรับปรุงโดย Google พัฒนามาจาก GoogLeNet ให้มีประสิทธิภาพดีกว่า InceptionNetv3 โครงสร้างเป็นแบบ Separable Convolution หรือ SeparableConv กล่าวคือมีการแยกส่วนของการทำคอนโวลูชันดังรูป

ที่ 3.54 ซึ่งแบ่งส่วนของการทำคอนโวลูชันออกเป็น 3 ส่วนคือ Entry flow, Middle flow และ Exit flow โดยกำหนดขนาดของภาพ Input ไว้ที่ 299 x 299 x 3 pixels XceptionNet มีความลึก 126 มีจำนวนพารามิเตอร์ 22.9 ล้านพารามิเตอร์ สำหรับการทดสอบด้วย ImageNet ให้ประสิทธิภาพสูงกว่า InceptionNetv3

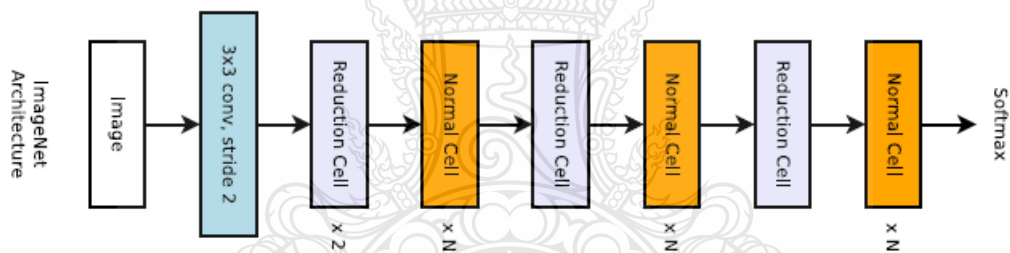


รูปที่ 3.54 XceptionNet [101]

17. NASNetMobile [102] หรือ Neural Architecture Search Network ได้รับการพัฒนาโดย Google Brain Team ซึ่งใช้สองฟังก์ชันคือ 1) Normal cell 2) Reduction cell ดังแสดงในรูปที่ 3.55 (A) เริ่มแรก NASNet ใช้การดำเนินการกับชุดข้อมูลขนาดเล็กแล้วโอนบล็อกไปยังชุดใหญ่ มีการพัฒนาในส่วน droppath เรียกว่า Scheduled droppath ส่งผลให้ประสิทธิภาพสูงขึ้น สถาปัตยกรรมของ NASNet กำหนดขนาดของภาพ Input ไว้ที่ 224 x 224 x 3 pixels NASNetMobile มีจำนวนพารามิเตอร์ 5.3 ล้านพารามิเตอร์



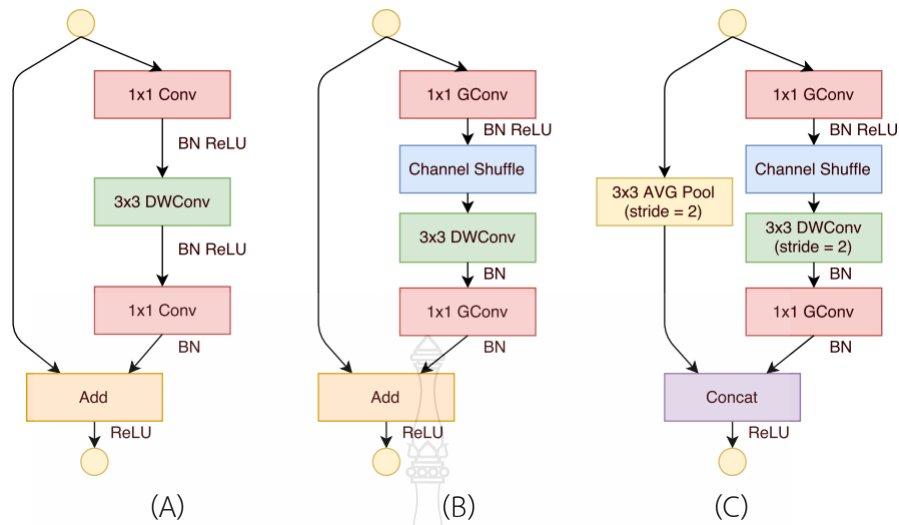
(A)



(B)

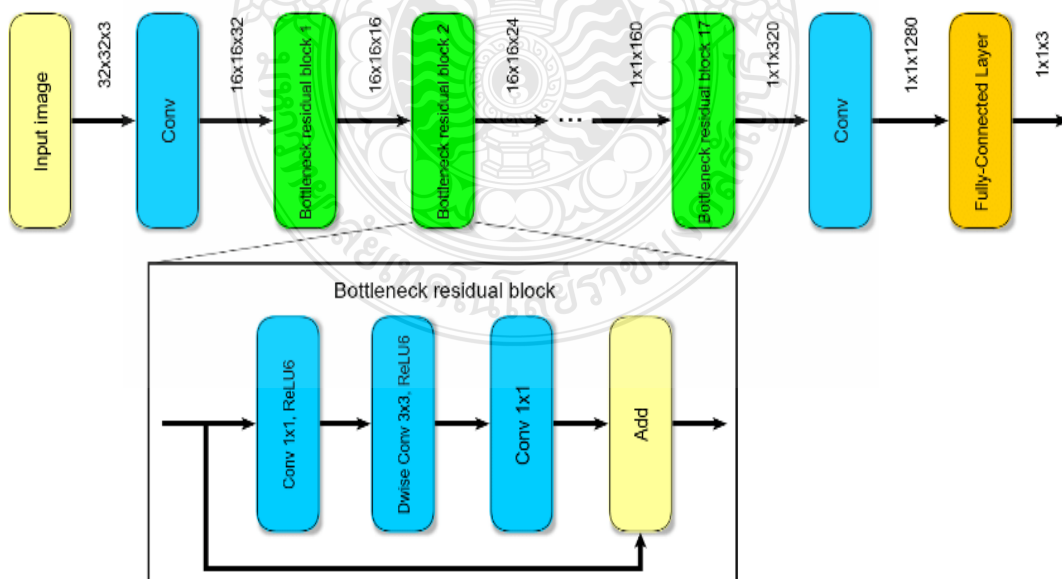
รูปที่ 3.55 NASNetMobile [102]

18. ShuffleNet [103] เป็นโครงข่ายที่ออกแบบมาเป็นพิเศษให้เป็นโครงข่ายขนาดเล็ก เริ่มจากการออกแบบ bottleneck unit ดังแสดงในรูปที่ 3.56 (A) มีโครงสร้างเป็น residual block ประกอบด้วยการทำคอนโวลูชันขนาด 3×3 แบบ depthwise convolution จากนั้นนำมาคั่นด้วยการทำคอนโวลูชันขนาด 1×1 แบบ group convolution และตามด้วย channel shuffle เพื่อสร้าง ShuffleNet ดังแสดงในรูปที่ 3.56 (B) จุดประสงค์เพื่อเพิ่มมิติในการทำคอนโวลูชัน ในกรณีที่มีการ stride ได้ทำการเพิ่มการทำคอนโวลูชันขนาด 3×3 ด้วยการทำ average pooling ดังรูปที่ 3.56 (C) ShuffleNet กำหนดขนาดของภาพ Input ไว้ที่ $224 \times 224 \times 3$ pixels มีจำนวน 11 เลเยอร์ มีจำนวนพารามิเตอร์ 5.3 ล้านพารามิเตอร์



รูปที่ 3.56 ShuffleNet [103]

19. MobileNetV2 [104] เป็นสถาปัตยกรรมโครงข่ายประสาทเทียมที่พยายามจะทำงานได้ดึบนอุปกรณ์มือถือ มีโครงสร้างแบบ inverted residual ที่ต่ออยู่กับ bottleneck residual block สถาปัตยกรรมของ MobileNetV2 ดังแสดงในรูปที่ 3.57 ประกอบด้วยชั้นทั้งหมด 28 ชั้น กำหนดขนาด Input เป็น $224 \times 224 \times 3$ pixels โครงสร้างประกอบด้วยชั้นคอนโวลูชัน 1 ชั้น ตามด้วย residual bottleneck layers 19 ชั้น และตามด้วยชั้นคอนโวลูชัน 1 ชั้น และชั้นสุดท้ายเป็น Fully connected layer ซึ่งเป็นชั้นคำนวณความน่าจะเป็นด้วย SoftMax Function มีจำนวนพารามิเตอร์ 3.5 ล้านพารามิเตอร์



รูปที่ 3.57 MobileNetV2 [104, 112]

เนื่องจาก DCNN ทั้งหมดที่นำมาใช้ทดสอบในงานวิจัยนี้เป็น Network ที่ถูกพัฒนามาสำหรับงาน Image Net ซึ่งถูก Training มาด้วยข้อมูลภาพจำนวน 1000 Class โดยมีจำนวนภาพในชุดข้อมูลมากกว่า 14 ล้านภาพ เมื่อนำ Network มาใช้กับชุดข้อมูลที่มีขนาดเล็ก ซึ่งในงานนี้กำหนด Class ของวัตถุไว้ 20 Class และมีจำนวนภาพที่เป็นชุดข้อมูลเพียง 4000 ภาพจึงได้ทำการดัดแปลง Network โดยการกำหนด Output Layer ซึ่งถูกคำนวณด้วย SoftMax Function ให้เหลือ Output Layer เพียง 20 Class ในการทดสอบการจดจำด้วยวิธี DCNN ได้ทำการทดสอบเพื่อหาพารามิเตอร์ที่เหมาะสมในการฝึกอบรมด้วย Data Set ที่สร้างขึ้น พารามิเตอร์ของ DCNN ที่เหมาะสมสรุปไว้ในตารางที่ 3.2 ส่วนค่า Initial Learning Rate ที่เหมาะสมของแต่ละ Network แสดงดังตารางที่ 3.3 ในการทดสอบกับ Network แต่ละ Model ได้ทำการทดสอบ Model ละ 10 ครั้ง และหาค่า Average Accuracy เพื่อนำมาเปรียบเทียบผลที่ได้

ตารางที่ 3.2 Training Parameter

Parameters	Values
NumEpochs	30
NumBatchSize	16
Momentum	0.9
LearnRateDropFactor	0.1
LearnRateDropPeriod	8

ในการทดสอบชุดข้อมูลทั้งหมดด้วย DCNN พารามิเตอร์การฝึกอบรมที่เกี่ยวข้องจะเหมือนกันในทุก Model ดังสรุปไว้ในตารางที่ 3.2 แต่อัตราการเรียนรู้เริ่มต้นหรือ initial learning อยู่ในช่วง 0.00001–0.1 โดยได้รับการปรับให้เหมาะสมในแต่ละ Model ซึ่งได้ทำการทดสอบและสรุปไว้ดังตารางที่ 3.3 ซึ่งเป็นอัตราการเรียนรู้ที่เหมาะสมที่สุดสำหรับ Network แต่ละ Model ในการทดสอบในแต่ละ Model ได้ทำการทดสอบซ้ำ จำนวน 10 ซ้ำและหาค่าเฉลี่ย การใช้ชุดข้อมูลทดสอบใช้ชุดข้อมูลชุดเดียว มีจำนวน Tactile Image จำนวน 4000 ภาพ ชุดข้อมูลถูกสุ่มและแบ่งออกเป็นชุดการฝึก (50%) และชุดทดสอบ (50%) ในการทดสอบแต่ละครั้ง อัตราการเรียนรู้จำของแต่ละรุ่นได้รับการประเมินจากนั้นจึงได้ทำการเปรียบเทียบผลที่ได้ และได้ทำการทดสอบการเพิ่มประสิทธิภาพโดยการเพิ่มความละเอียดของภาพ

ตารางที่ 3.3 ค่า Initial Learning Rate ของ DCNN แต่ละ Model

Initial Learning Rate	0.00001	0.0001	0.001	0.01
				ResNet18
				ResNet50
				ResNet101
		VGGNet16	AlexNet	InceptionV3
		VGGNet19	GoogLeNet	InceptionResNetV2
CNN Model	SqueezeNet	DarkNet19	Place365	GoogleNet
		DarkNet53	EfficienNetB0	DensNet201
				XceptionNet
				NASNetMobile
				ShuffleNet
				MobileNetV2

3.4.5 Recognition Rate from Object Exploration

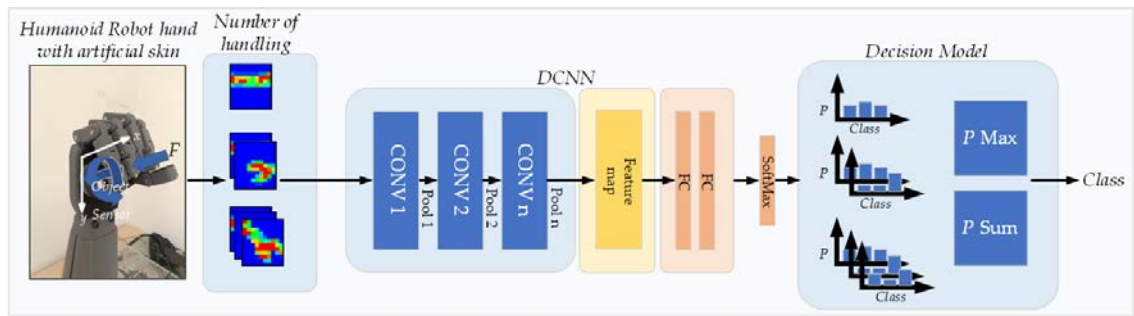
วิธีการ BoF จะสำรวจวัตถุอย่างละเอียดโดยการจัดการกับวัตถุหลายครั้งเพื่อให้สามารถทำนายผลลัพธ์ได้อย่างถูกต้อง เมื่อใช้วิธี DCNN สามารถเพิ่มประสิทธิภาพการจดจำวัตถุได้โดยการเพิ่มจำนวนครั้งในการจัดการกับวัตถุเพื่อสำรวจวัตถุอย่างละเอียด การจะเพิ่มประสิทธิภาพในการจดจำสามารถทำได้โดยการจับวัตถุหลายครั้งแล้วหาค่า Maximum of Probability (P_{max}) หรือ Summation of Probability (P_{sum}) โดยสามารถใช้สมการ 3.34 และ 3.35 เป็นแบบจำลองการตัดสินใจดังรูปที่ 3.58

$$P = \max(P(I_1), P(I_2), P(I_3), \dots, P(I_n)) \quad (3.34)$$

สมการที่ 3.34 ใช้ในการทำนายผลลัพธ์โดยใช้วิธีความน่าจะเป็นสูงสุด กล่าวคือความน่าจะเป็น P สำหรับการทำนายผลลัพธ์ที่ได้จากความน่าจะเป็นสูงสุดของ Tactile Image (I) ที่ทดสอบ

$$P = P(I_1) + P(I_2) + P(I_3) + \dots + P(I_n) \quad (3.35)$$

สมการที่ 3.35 ใช้สำหรับทำนายผลลัพธ์โดยใช้วิธีผลรวมของค่าความน่าจะเป็น นั่นคือความน่าจะเป็น P สำหรับการทำนายผลลัพธ์ที่ได้จากผลรวมของความน่าจะเป็นจาก Tactile Image (I) ที่ทดสอบ



รูปที่ 3.58 การจดจำวัตถุจากการสำรวจวัตถุ

2.4.6 การวัดประสิทธิภาพ

การวัดประสิทธิภาพในการทดสอบการจดจำวัตถุจากการสัมผัสในงานวิจัยนี้ใช้วิธีการการวัดค่าความถูกต้อง (Accuracy) ซึ่งรายงานผลเป็นค่าร้อยละ หรือเปอร์เซ็นต์ (%) สามารถหาค่าได้จาก

$$Accuracy = \left[\frac{(TP + TN)}{TP + TN + FP + FN} \right] \times 100 \quad (3.36)$$

เมื่อ TP คือ ค่าที่พยากรณ์ถูกต้องเชิงบวก, TN คือ ค่าที่พยากรณ์ถูกต้องเชิงลบ, FP คือ ค่าที่พยากรณ์ผิดพลาดเชิงบวก และ FN คือ ค่าที่พยากรณ์ผิดพลาดเชิงลบ

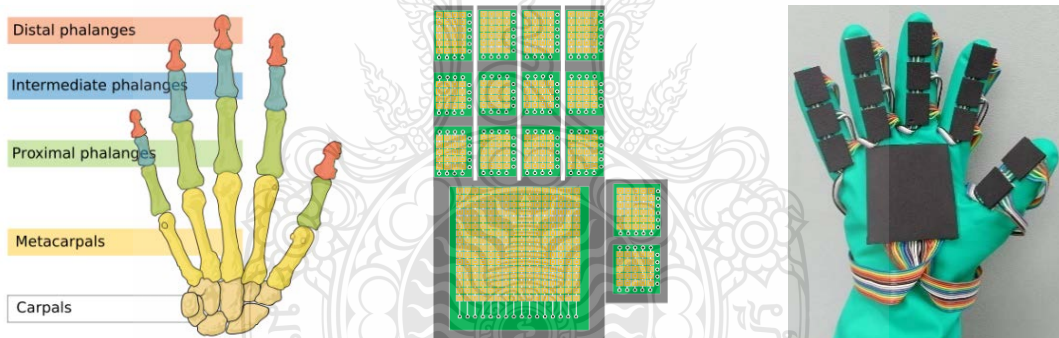
3.5 Glove Tactile Sensor

ในงานวิจัยนี้ได้ทำการพัฒนา Glove Tactile Sensor ซึ่งเป็นชุด Tactile Sensor ที่เป็นไปตามโครงสร้างของมือนมนุษย์ โดยการนำ Sensor Array ที่พัฒนาในหัวข้อ 3.3 ที่ถูกออกแบบมาประกอบเป็น Sensor แบบ Full Hand ซึ่งแสดงดังรูปที่ 3.59 โดยตำแหน่งที่ติดตั้งอ้างอิงจากรูปแบบโครงสร้างของนิ้วมือ แบ่งเป็น 3 รูปแบบ คือ 1) ติดตั้งที่ตำแหน่งฝ่ามือจำนวน 1 จุด โดยติดตั้ง Sensor ขนาด 16×16 pixels หรือ $56.0 \times 56.0 \text{ mm}^2$ ที่ตำแหน่งตรงกลางฝ่ามือ 2) ตำแหน่งนิ้วหัวแม่มือ 1 นิ้ว โดยติดตั้ง Sensor จำนวน 2 จุดตามข้อนิ้วมือ คือตำแหน่ง Distal Phalanges ขนาด 5×6 Pixels หรือ $17.5 \times 21.0 \text{ mm}^2$ และติดตั้งที่ตำแหน่ง Proximal Phalanges ขนาด 5×5 Pixels หรือ $17.5 \times 17.5 \text{ mm}^2$ และ 3) ตำแหน่งนิ้วมือ 4 นิ้ว โดยติดตั้ง Sensor จำนวน 3 จุดตามข้อนิ้วมือ คือตำแหน่ง Distal Phalanges 4×6 Pixels หรือ $14.0 \times 21.0 \text{ mm}^2$, ติดตั้งที่ตำแหน่ง Intermediate Phalanges ขนาด 4×2 Pixels หรือ $14.0 \times 7.0 \text{ mm}^2$ และตำแหน่ง Proximal Phalanges ขนาด 4×4 Pixels หรือ $14.0 \times 14.0 \text{ mm}^2$ ดังนั้นจึงมีจุดที่ติดตั้ง Sensor ทั้งหมด 15 จุด รวมเป็น 551 Pixels ดังตารางที่ 3.4 การประกอบชุด Sensor ทั้งหมดประกอบลงบนถุงมือที่ไว้สำหรับสวมให้กับ Humanoid Robot Hand

ชุด Sensor ที่ประกอบเสร็จแสดงดังรูปที่ 3.59 (C) และ Sensor ที่ประกอบเสร็จเพื่อพัฒนาไปใช้กับ Humanoid Robot Hand

ตารางที่ 3.4 Glove Tactile Sensor

Sensor	Number of set	resolution (Pixels)	Size (mm ²)	Number of pixels (Pixels)
Palm	1	16x16	56.0x56.0	256
Thump	1	5x6, 5x5	17.5x21.0, 17.5x17.5	55
Finger	4	4x6, 4x4, 4x5	14.0x21.0, 14.0x14.0, 14.0x17.5	240

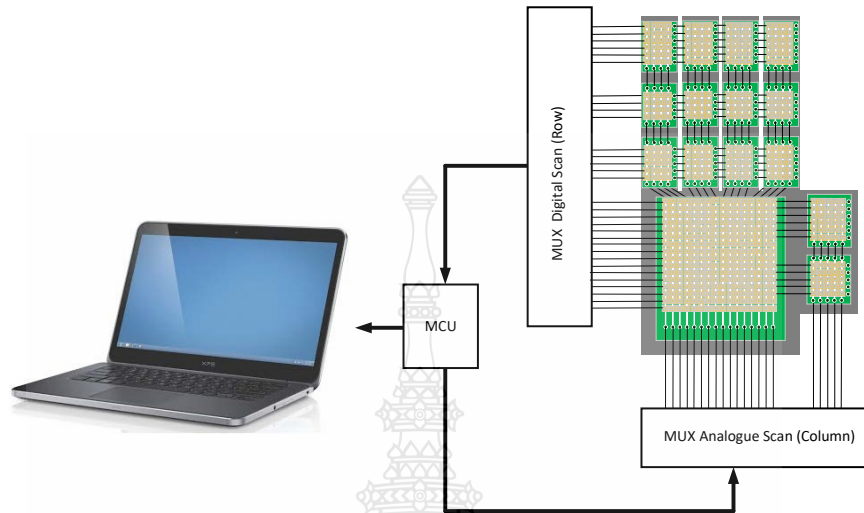


(A) โครงสร้างของฝ่ามือมนุษย์ (B) การจัดวาง Sensor Array (C) Glove Tactile Sensor
รูปที่ 3.59 การประกอบชุดถุงมือ Sensor Array

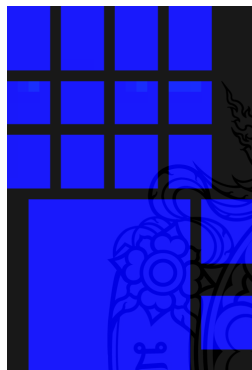
3.5.1 วงจรอ่านค่าและการสร้างภาพ Full Hand Tactile Image

การอ่านค่าจาก Sensor ทำงานแสดงดังรูปที่ 3.60 (A) ใช้บอร์ดไมโครคอนโทรลเลอร์ Arduino รุ่น mega 2560 ร่วมกับ Multiplexer เบอร์ CD74HC4067 การทำงานจะเป็นการ Multiplex เพื่อ Scan Sensor ในแต่ละตำแหน่ง ลำดับการทำงานให้เริ่มจากทำ Multiplex ในตำแหน่ง Row เพื่อให้ Output ในตำแหน่ง R1 เป็น +5V จากนั้น Multiplex ตำแหน่ง Column เพื่ออ่านค่าที่เซ็นเซอร์แต่ละตัว ตั้งแต่ Column แรกโดยเริ่มจาก C1 ไปจนถึง C21 (ทั้งหมด 21 Column) การอ่านค่าใช้ A / D ภายในไมโครคอนโทรลเลอร์แปลงเป็นข้อมูลดิจิทัลและส่งข้อมูลไปยังส่วนการประมวลผล

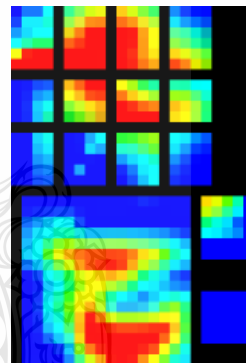
สัญญาณคอมพิวเตอร์ เมื่ออ่านค่าครบทุก Column แล้วก็สแกน Digital Line ในตำแหน่ง Row ถัดไป จนกว่าจะอ่านเซ็นเซอร์ได้ทั้งหมด (ทั้งหมด 31 Row) จึงนำข้อมูลที่อ่านได้มา Plot เป็นภาพ โดยการ จัดเรียงภาพตาม Pattern ในรูปที่ 3.60 (B)



(A) วงจรอ่านค่าเซ็นเซอร์



(B) Full Hand Tactile Image Pattern



(C) ตัวอย่าง Full Hand Tactile Image

รูปที่ 3.60 วงจรอ่านค่าและการสร้างภาพ Full Hand Tactile Image





















3.5.2 วัตถุ และ Full Hand Tactile Image

ภาพตัวอย่างวัตถุจะถูกรวบรวมโดยการสแกนตัวอย่างตำแหน่งของวัตถุ ภาพที่สัมผัสได้นั้นมาจากตำแหน่งต่างๆ ของวัตถุ ภาพที่สัมผัสเป็นผลมาจากตำแหน่งของวัตถุที่แตกต่างกัน ตำแหน่งการจับจะใช้เพื่อหมุนวัตถุในแนว 3 มิติ ก่อนที่จะจับด้วยมือของหุ่นยนต์ เมื่อหุ่นยนต์จัดการและปล่อยวัตถุ ลักษณะทางกายภาพของมือหุ่นยนต์จะถูกควบคุมวัตถุอย่างอิสระ สำหรับการทดลองนี้ในการทดสอบระบบจดจำวัตถุได้สร้างชุดข้อมูลจากการจับวัตถุจริงจำนวน 20 Object หรือ 20 Class เป็นวัตถุที่มีรูปทรงมาตรฐาน 10 วัตถุและวัตถุในชีวิตประจำวันอีก 10 วัตถุ ทั้งหมดประกอบด้วย Battery, Cup,

Plastic tongs, Golf Ball, Remote controller, Scissors, Allen key, Screwdriver, Tape measure, Cube, Rectangular, Triangular prism, Pyramid, Tetrahedron, Large cylinder, Small cylinder, Cone, Oval และ Sphere ลักษณะวัตถุแสดงดังรูปที่ 3.61 ขนาดแสดงดังตารางที่ 3.5 การสร้าง Data Set เกิดจากการจับวัตถุและบันทึกภาพวัตถุๆ ละ 200 ภาพ รวมภาพทั้งหมด 4,000 ภาพ

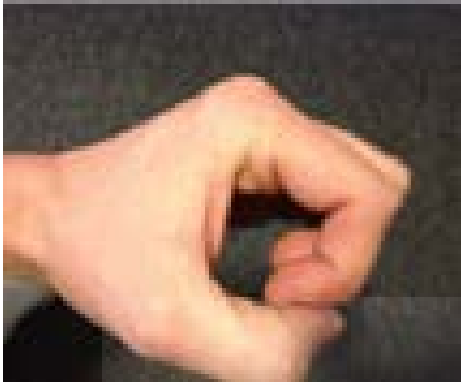
ตารางที่ 3.5 ขนาดของวัตถุ

Class	Object	Size
1.	Battery	$\varnothing=33$ mm, H=60 mm
2.	Cup	$\varnothing=85$ mm, W=120 mm, H=95 mm
3.	Plastic tongs	L=165 mm, W=85 mm, H=15 mm
4.	Remote controller	L=172 mm, W=45 mm, H=15 mm
5.	Golf Ball	$\varnothing=40$ mm
6.	Scissors	L=18 mm, W=80 mm, H=12 mm
7.	Variable wrench	L=200 mm, W=60 mm, H=14 mm
8.	Allen key	L=90 mm, W=35 mm, H=6 mm
9.	Screwdriver	$\varnothing=38$ mm, L=220 mm
10.	Tape measure	$\varnothing=70$ mm, W=92 mm, H=47 mm
11.	Cube	L=25 mm, W=25 mm, H=25 mm
12.	Rectangular	L=50 mm, W=25 mm, H=25 mm
13.	Triangular prism	L=50 mm, W=25 mm, H=20 mm
14.	Pyramid	L=25 mm, W=25 mm, H=50 mm
15.	Tetrahedron	L=25 mm, W=25 mm, H=50 mm
16.	Large cylinder	$\varnothing=120$ mm, L=110 mm
17.	Small cylinder	$\varnothing=25$ mm, L=50 mm
18.	Cone	$\varnothing=25$ mm, L=50 mm
19.	Oval	L=25 mm, W=50 mm
20.	Sphere	$\varnothing=25$ mm

			
(1)	(2)	(3)	(4)
			
(5)	(6)	(7)	(8)
			
(9)	(10)	(11)	(12)
			
(13)	(14)	(15)	(16)
			
(17)	(18)	(19)	(20)

รูปที่ 3.61 วัตถุที่ใช้ในการทดลอง

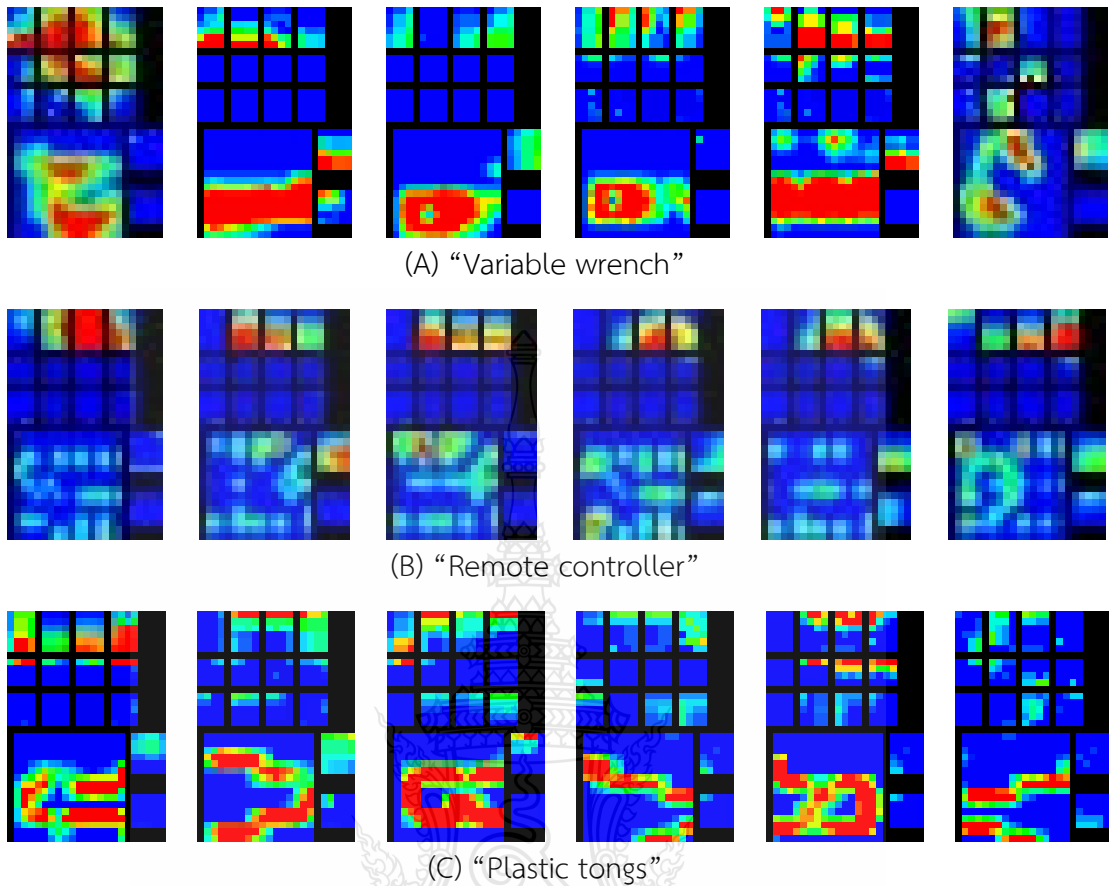
การจับวัตถุมีลักษณะแบบรูปที่ 3.62 คือการกำวัตถุโดยรอบ ซึ่งท่าทางของมือจะเปลี่ยนไปตามลักษณะของวัตถุ เช่น ทรงกระบอก ทรงกลม, การสำรวจวัตถุด้วยวิธีนี้จะเป็นการสำรวจวัตถุโดยรอบวัตถุ เนื่องจากการกำวัตถุโดยรอบ และในการสำรวจวัตถุของหุ่นยนต์มีลักษณะเดียวกับมนุษย์ ในรูป 3.62 แสดงการสำรวจวัตถุจากการสัมผัสวัตถุเต็มฝ่ามือ และการจับวัตถุของมือมนุษย์



(A) การจับวัตถุแบบกำเต็มมือเพื่อสำรวจวัตถุของมนุษย์ (B) การจับวัตถุแบบกำเต็มมือเพื่อสำรวจวัตถุโดย Glove Tactile Sensor

รูปที่ 3.62 การจับวัตถุแบบกำเต็มมือเพื่อสำรวจวัตถุของมนุษย์และการจับวัตถุแบบกำเต็มมือโดย Glove Tactile Sensor

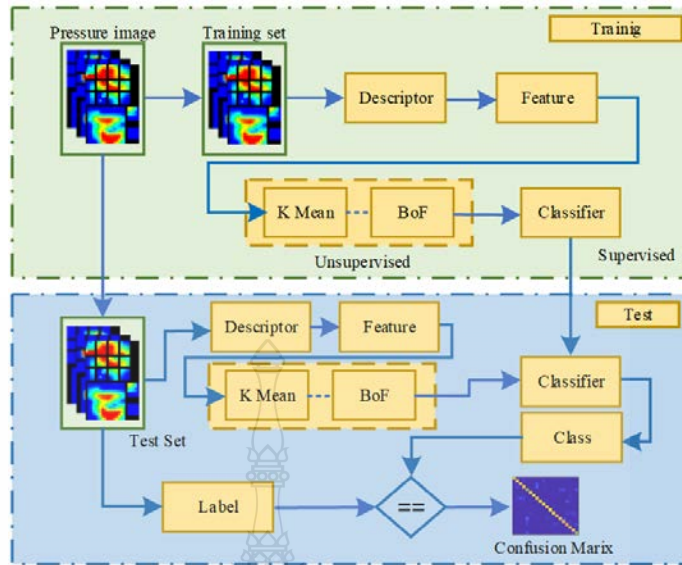
ภาพ Full Hand Tactile Image ที่ได้จากการจับวัตถุด้วย Glove Tactile Sensor เป็นภาพ Tactile Image ที่ได้จากการจับหรือกำวัตถุ ภาพตัวอย่างที่ได้แสดงดังรูปที่ 3.63 ภาพที่ได้จะเป็นคุณลักษณะของวัตถุโดยรอบของวัตถุหากวัตถุมีขนาดเล็กหรือมีรูปทรงที่สามารถกำได้เต็มมือ หรืออาจจะเป็น Pressure Profile จากวัตถุ 2 ด้านคือด้านที่สัมผัสด้วยฝ่ามือ และด้านที่สัมผัสด้วยนิ้ว ตัวอย่างเช่นในรูปที่ 3.63 (A) มือที่จับวัตถุ “Wrench” Tactile Image ที่ได้จะเป็น Pressure Profile จากวัตถุสองด้าน คือด้านที่สัมผัสด้วยฝ่ามือ และด้านที่สัมผัสด้วยนิ้วมือ ในรูปที่ 3.63 (B) มือที่สัมผัสวัตถุ “Remote controller” ภาพที่ได้จะเป็น Pressure Profile Tactile Image ที่ได้จะเป็นจากด้านปุ่มของ Remote Controller และด้านหลังของ Remote Controller และในรูปที่ 3.63 (C) เมื่อสัมผัสกับ Object “Plastic tongs” Tactile Image ที่ได้จะเป็น Pressure Profile ของ Plastic Tongs ที่เกิดขึ้นกับ Sensor บนฝ่ามือ และอีกด้านจะเป็น Pressure Profile ที่สัมผัสด้วยนิ้วมือ



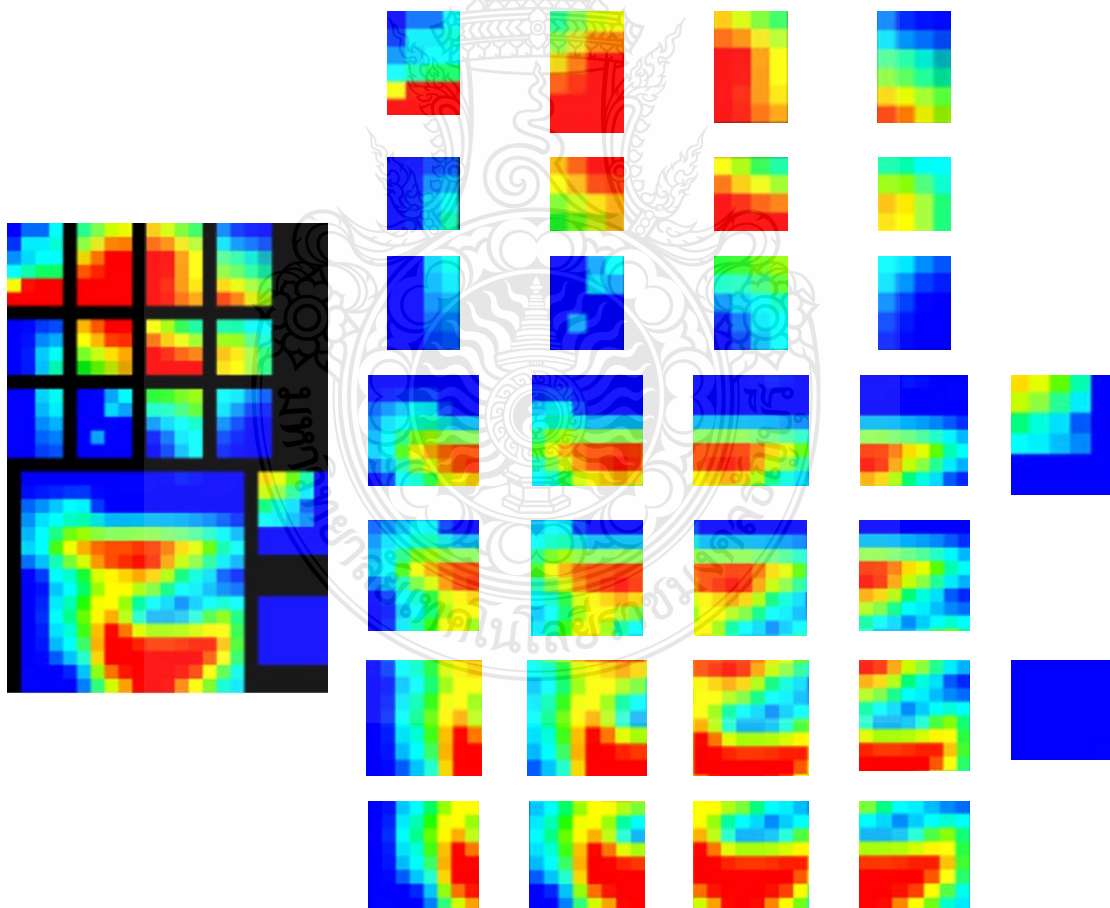
รูปที่ 3.63 ตัวอย่างของภาพ Full Hand Tactile Image ที่ได้จากการจับวัตถุ

3.5.3 การทดลองอัลกอริทึมแบบ BoF สำหรับ Glove Tactile Sensor

ในการทดลองที่ใช้ BoF นั้นได้ทดสอบเช่นเดียวกับวิธีในหัวข้อ 3.4.3 คือการทดสอบกับ Classifier 3 วิธีประกอบด้วย KNN, SVM และ ANN ในแต่ละ Classifier ทดสอบกับ Descriptor 3 วิธี คือ Moment Analysis, SIFT และ Pola Furrier และได้ทดสอบกับ Cluster จำนวน 10, 20 และ 30 Cluster ผลที่ได้จากการทดสอบแบ่งตามวิธี Classifier โดยทุกการทดสอบได้ทำการทดสอบ 10 ซ้ำและทำการหาค่า average accuracy การดำเนินงานแสดงดัง Work Flow ในรูปที่ 3.64 การจดจำวัตถุจากการสัมผัสด้วยวิธี BoF สำหรับเซ็นเซอร์ถุงมือใช้วิธีทำการแบ่งส่วนของภาพที่ได้เป็นส่วนย่อยตามลักษณะมือ โดยใช้วิธีเดียวกับวิธีในหัวข้อที่ 3.1 โดยในส่วนของฝ่ามือแบ่งออกเป็น 4×4 แบบ Overlapping รวมเป็นส่วนย่อยทั้งหมด 30 ส่วนดังรูปที่ 3.65



รูปที่ 3.64 Work Flow สำหรับการจดจำวัตถุจากการสัมผัสด้วยวิธี BoF สำหรับ Glove Tactile Sensor



รูปที่ 3.65 การแบ่ง Sub Image ของ Full Hand Tactile Image สำหรับทำ Descriptor

3.5.4 การทดลองอัลกอริทึมแบบ DCNN สำหรับเซ็นเซอร์ถุงมือ

วิธี DCNN โดยมี Work Flow ดังรูปที่ 3.6 ในงานนี้ได้ทำการทดสอบด้วย DCNN จำนวน 7 Model ประกอบด้วย AlexNet, VGG19, GoogleNet, InceptionNetV3, InceptionResNetV2, XceptionNet และ NASNetMobile ซึ่งเป็น DCNN ที่นิยมใช้กันอย่างแพร่หลายในงาน Computer Vision และพบว่าให้ประสิทธิภาพสูงเมื่อนำมาประยุกต์ใช้กับการจดจำวัตถุจากการสัมผัสด้วย Tactile image ในงานนี้จึงได้นำมาใช้กับการจดจำวัตถุโดยการสัมผัสโดยใช้ Full Hand Tactile Image จากเซ็นเซอร์ถุงมือซึ่งมีจุดสัมผัสแบบเต็มมือ การทดลองทั้งหมดใช้วิธี Transfer Learning เช่นเดียวกับการทดลองในหัวข้อ 3.4.5 ในการทดสอบชุดข้อมูลด้วย DCNN พารามิเตอร์การฝึกอบรมที่เกี่ยวข้องจะเหมือนกันดังสรุปไว้ในตารางที่ 3.6 อัตราการเรียนรู้ Initial Learning Rate ที่เหมาะสมของแต่ละ Model แสดงในตารางที่ 3.7 ในการทดสอบกับ Network แต่ละ Model ได้รับการทดสอบ 10 ครั้งและหาค่าความถูกต้องเฉลี่ยเพื่อเปรียบเทียบผลที่ได้

ตารางที่ 3.6 Training Parameter for Glove Tactile Sensor

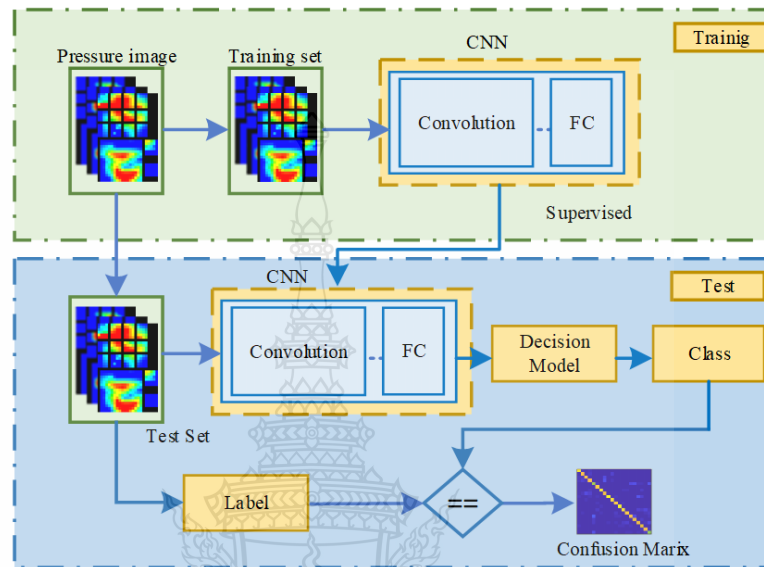
Parameters	Values
NumEpochs	30
NumBatchSize	8
Momentum	0.9
LearnRateDropFactor	0.1
LearnRateDropPeriod	8

ตารางที่ 3.7 ค่า Initial Learning Rate for Glove Tactile Sensor

Network	Values
AlexNet	0.001
VGG19	0.0001
GoogleNet	0.001
InceptionNetV3	0.01
InceptionResNetV2	0.01
XceptionNet	0.01

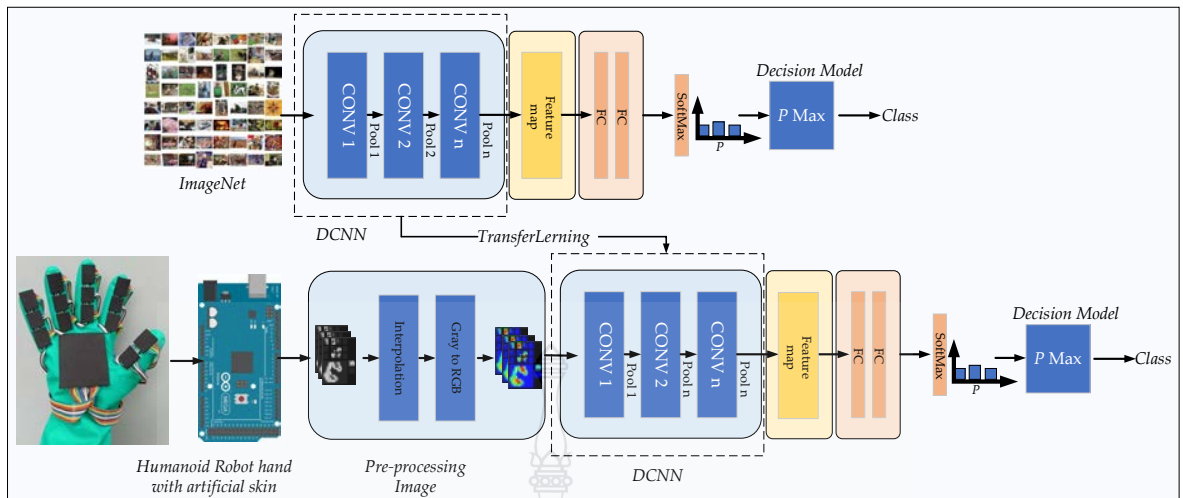
ตารางที่ 3.7 ค่า Initial Learning Rate for Glove Tactile Sensor (ต่อ)

Network	Values
NASNetMobile	0.01



รูปที่ 3.66 Work Flow สำหรับการจดจำวัตถุจากการสัมผัสด้วยวิธี DCNN สำหรับ Glove Tactile Sensor

องค์ประกอบที่เกี่ยวข้องกับ DCNN แสดงไว้ในรูปที่ 3.67 ซึ่งใช้กับการจดจำวัตถุสัมผัสโดยใช้ Transfer Learning Method จาก ImageNet DCNN ประกอบด้วยฟังก์ชัน SoftMax 20 เวกเตอร์ เป็นเลเยอร์สุดท้าย (Fully Connected Layer) เพื่อคำนวณการกระจายความน่าจะเป็นสำหรับ Class ของวัตถุ 20 Class และทำนายผลจากค่าความน่าจะเป็นสูงสุดที่ได้จากฟังก์ชัน SoftMax



รูปที่ 3.67 Transfer Learning Method สำหรับการจดจำวัตถุจาก Glove Tactile Sensor



บทที่ 4

การทดลองและผลการทดลอง

4.1 ผลการทดสอบสมมติฐานเบื้องต้น

4.1.1 Tactile Image จากถุงมือกระดาษ

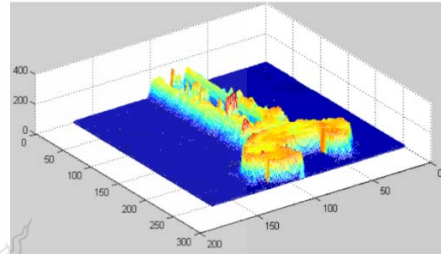
การสัมผัสวัตถุของมือมนุษย์แรงกดที่เกิดขึ้นจะมีลักษณะเหมือนรูปร่างของวัตถุ ตัวอย่างวัตถุและภาพที่ได้มีลักษณะดังรูปที่ 4.1 เมื่อนำมาพล็อตเพื่อดูการกระจายของแรงพบว่า มีลักษณะเหมือนกับรูปร่าง 3 มิติของวัตถุ ดังแสดงในรูปที่ 4.1(C)



(A) วัตถุ “ประแจ”



(B) Tactile Image



(C) รูปพล็อต Tactile Image 3 มิติ

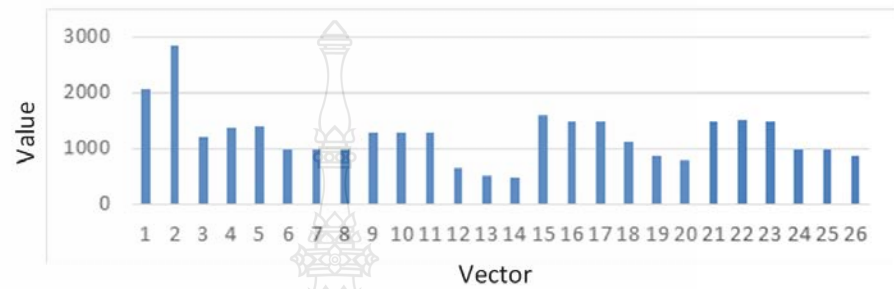
รูปที่ 4.1 วัตถุประแจ, ภาพ Tactile Image ของวัตถุประแจ และภาพ Tactile Image ที่นำมาพล็อตเป็น 3 มิติ

4.1.2 Descriptor

ภาพ Tactile Image ที่ได้มาแบ่งเป็นส่วนย่อยตามลักษณะของฝ่ามือสามแบบ คือ 15 ส่วน 20 ส่วน และ 26 ส่วน เมื่อทำการแบ่งจะได้ภาพย่อยๆ ที่ได้จากการแบ่งเป็น 15 ภาพ 20 ภาพ และ 26 ภาพตามลำดับแล้วทำการรวมพื้นที่ของภาพย่อยแต่ละภาพ และจัดเรียงค่าที่ได้เป็น Data Vector เพื่อเป็นตัวแทนของข้อมูลสำหรับนำไปทดสอบด้วย ANN รูปที่ 4.2 แสดงตัวอย่างของ Data Vector ของภาพวัตถุประแจที่แบ่งส่วนฝ่ามือเป็น 26 ส่วน ซึ่งเมื่อนำมาพล็อตเป็น Histogram จะพบว่าส่วนที่มีการสัมผัสโดนวัตถุจะมีผลรวมของพื้นที่ต่ำกว่าส่วนที่สัมผัสไม่โดนวัตถุ เนื่องจากภาพในส่วนสีขาวมีค่าตัวเลขของภาพมากกว่าภาพในส่วนสีดำ ดังนั้นในส่วนสีดำจึงมีค่าผลรวมต่ำกว่า



F11 F12 F21 F22 F23 F31 F32 F33 F41 F42 F43 F51 F52 F53 P11 P12 P13 P21 P22 P23 P31 P32 P33 P41 P42 P43



รูปที่ 4.2 Data Vector ของภาพวัตถุประแจที่แบ่งส่วนฝ่ามือเป็น 26 ส่วน

4.1.3 การจดจำวัตถุจากการแบ่งส่วนมือ

ในการทดสอบสมมติฐานเบื้องต้นนี้ ได้นำเสนอว่าหากทำการแบ่งส่วนของฝ่ามือในการนำมาใช้วิเคราะห์จดจำวัตถุจากภาพสัมผัส จะสามารถเพิ่มประสิทธิภาพในการจดจำวัตถุได้ โดยในการทดลองได้ทดสอบการแบ่งส่วนฝ่ามือเป็น 15 ส่วน, 20 ส่วนและ 26 ส่วน ในการทดลองได้ทำการลดค่า Signal to Noise Ratio เพื่อลดความคมชัดของภาพ เนื่องจากภาพได้จากการสแกนจึงมีความละเอียดสูง จึงให้ค่าความถูกต้องในการทดสอบ 100% จึงลดความคมชัดของภาพและทดสอบที่ Signal to Noise Ratio (SNR) 7 ระดับประกอบด้วย inf, 30, 20, 10, 5, 2.5, และ 1 และหาค่าความถูกต้องเฉลี่ย ผลการทดสอบการจดจำทดสอบกับวัตถุ 15 class ประกอบด้วย 1.ปลั๊กไฟฟ้า, 2.ลูกแบดมินตัน, 3.ตุ๊กตา, 4.ขาม, 5.หวี, 6.แปรง, 7.โคมก้นกระแตก, 8.ประแจ, 9.แปรงสีฟัน, 10.ตะกร้า, 11.กรวย, 12.รีโมทคอนโทรล, 13.แก้วกาแฟ1, 14.แก้วกาแฟ2 และ 15.กล่องกระดาษ ผลในการทดสอบแบ่งออกเป็นผลการทดสอบจากการแบ่งส่วนมือ 15 ส่วนแสดงดังตารางที่ 4.1 มีค่าความถูกต้องเฉลี่ยอยู่ที่ 71.14% ผลการทดสอบจากการแบ่งส่วนมือ 20 ส่วนแสดงดังตารางที่ 4.2 มีค่าความถูกต้องเฉลี่ยอยู่ที่ 73.09% และผลการทดสอบจากการแบ่งส่วนมือ 26 ส่วนแสดงดังตารางที่ 4.3 มีค่าความถูกต้องเฉลี่ยอยู่ที่ 78.38% ผลการทดสอบดังกล่าวแสดงให้เห็นว่าหากทำการแบ่งส่วนของมือหลายๆ ส่วนตามโครงสร้างของมือค่าความถูกต้องในการจดจำวัตถุจากการสัมผัสจะสูงขึ้น

ตารางที่ 4.1 ผลการจดจำจากการแบ่งส่วนมือเป็น 15 ส่วน

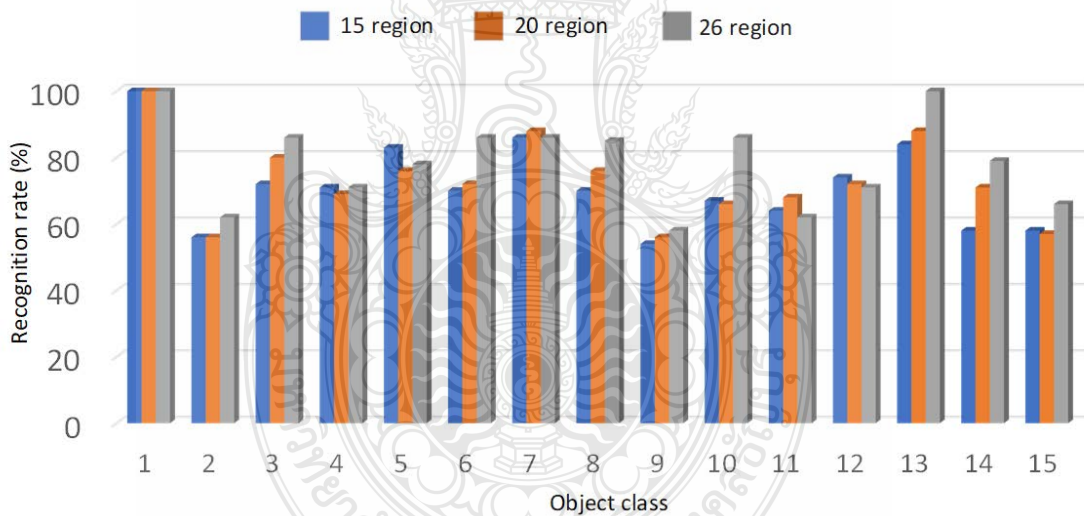
SNR	Classification accuracy															Mean
	Object															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Inf	100	100	100	100	100	100	100	90	100	100	100	100	100	100	100	99.33
30	100	100	100	100	100	100	100	90	100	100	100	100	100	100	100	99.33
20	100	100	100	100	100	100	100	90	100	100	100	100	100	100	100	99.33
10	100	90	100	100	100	100	100	90	80	100	100	100	100	100	97.33	
5	100	0	100	100	100	90	100	85	0	70	45	95	100	5	5	66.33
2.5	100	0	5	0	80	0	100	45	0	0	0	25	90	0	0	29.67
1	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6.67
Avg	100	56	72	71	83	70	86	86	54	67	74	74	84	58	58	71.14

ตารางที่ 4.2 ผลการจดจำจากการแบ่งส่วนมือเป็น 20 ส่วน

SNR	Classification accuracy															Mean
	Object															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Inf	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
30	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
20	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
10	100	90	100	100	100	100	100	100	95	100	100	100	100	100	100	99.00
5	100	0	100	85	100	100	100	95	0	60	75	95	100	95	0	73.67
2.5	100	0	60	0	30	5	100	40	0	0	0	10	100	0	0	29.67
1	100	0	0	0	0	0	15	0	0	0	0	0	15	0	0	8.67
Avg	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	73.06

ตารางที่ 4.3 ผลการจดจำจากการแบ่งส่วนมือเป็น 26 ส่วน

SNR	Classification accuracy															Mean
	Object															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Inf	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
30	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
20	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
10	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
5	100	35	100	100	95	100	100	100	5	100	35	95	100	100	65	82.00
2.5	100	0	100	0	50	100	100	80	0	100	0	0	100	50	0	52.00
1	100	0	5	0	0	0	0	15	0	0	0	0	100	0	0	14.67
Avg	100	62	86	71	78	86	86	85	58	86	62	71	100	79	66	78.38

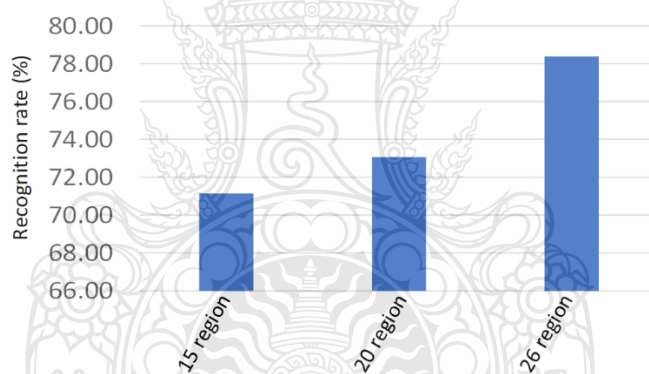


รูปที่ 4.3 ผลเปรียบเทียบความถูกต้องในการจดจำของวัตถุในแต่ละ Class เมื่อใช้การแบ่งส่วนฝ่ามือเป็น 15, 20 และ 26 ส่วน

ผลการทดลองในตารางที่ 4.1 นำเสนอผลการทดลอง เมื่อทำการแบ่งส่วนฝ่ามือเป็นจำนวน 15 ส่วน สำหรับการ Training และการ Testing ด้วยโครงข่ายประสาทเทียม ผลลัพธ์ความแม่นยำของการจดจำวัตถุในสภาพแวดล้อมที่ความคมชัดของภาพในสภาวะไร้สัญญาณรบกวน ($S/N = \text{inf}$) คือ 99.33% และได้ทดสอบต่อโดยเพิ่มสัญญาณรบกวนเข้าไปในระบบและหาค่าความถูกต้องเฉลี่ย ผลปรากฏว่าประสิทธิภาพลดลง ความถูกต้องที่ได้จากการทดสอบเมื่อแบ่งส่วนฝ่ามือเป็น 15 ส่วนคือ

71.14% จากนั้นในการทดลองแบ่งส่วนฝ่ามือออกเป็น 20 ส่วนและ 26 ส่วนและทำการทดสอบกับโครงข่ายประสาท ผลลัพธ์ของการทดสอบพบว่าเมื่อภาพมีความคมชัดความสามารถในการจดจำวัตถุเป็น 100% และเมื่อลดความชัดของภาพลงพบว่าประสิทธิภาพจะลดลง สัมพันธ์กับสัญญาณรบกวนที่เพิ่มขึ้น ความแม่นยำเมื่อมีการแบ่งส่วนฝ่ามือเป็น 20 ส่วนคือ 73.09% และสำหรับการแบ่งส่วนฝ่ามือเป็น 26 ส่วนคือ 78.38% ผลเปรียบเทียบในแต่ละวัตถุแสดงดังรูปที่ 4.3 ผลการทดสอบวัตถุแต่ละ Class พบว่าแทบทุก Class เมื่อมีการแบ่งส่วนฝ่ามือเพิ่มขึ้นค่าความถูกต้องจะเพิ่มขึ้นตามไปด้วย และผลเปรียบเทียบผลการจดจำเฉลี่ยในของการแบ่งส่วนทั้ง 3 แบบ แสดงดังรูปที่ 4.4 ซึ่งแสดงให้เห็นค่าความถูกต้องที่เพิ่มขึ้นเมื่อเพิ่มเพิ่มการแบ่งส่วนฝ่ามือเพิ่มขึ้น

ผลที่แสดงในรูปที่ 4.3 แสดงให้เห็นว่าวัตถุแทบทุกชนิดเมื่อเพิ่มการแบ่งส่วนภาพมากขึ้น ผลของ recognition rate จะเพิ่มขึ้น แต่อย่างไรก็ตามมีวัตถุบาง Class เช่นวัตถุ Class 5, 7, 11, 12 เมื่อเพิ่มจำนวนการแบ่งส่วนภาพมากขึ้นผล recognition rate กลับลดลงเล็กน้อย เนื่องจากวัตถุบาง Class เมื่อมีการแบ่งภาพเล็กลงพบว่า ภาพขนาดเล็กของวัตถุบาง Class มีความคล้ายคลึงกัน หรือมีพื้นที่ของรูปเท่ากัน จึงทำให้ผลการทำนายลดลง



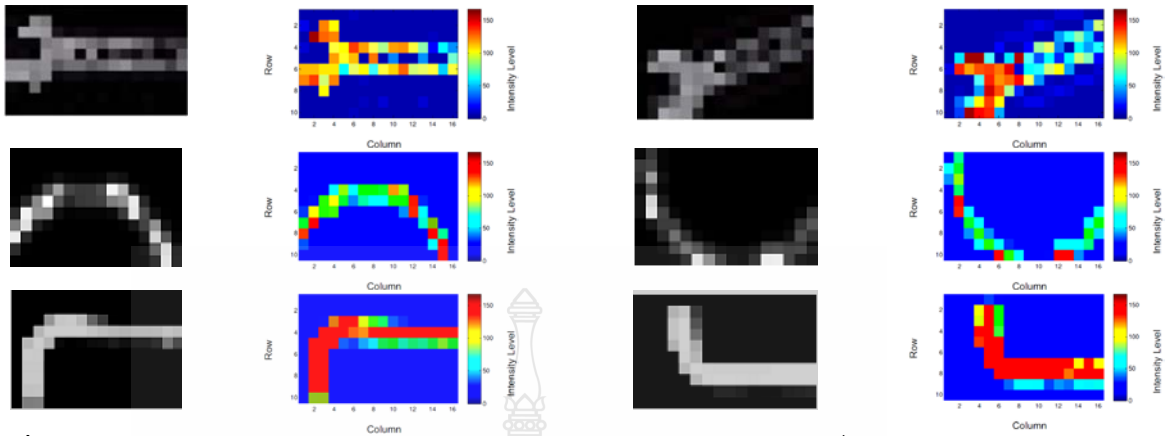
รูปที่ 4.4 เปรียบเทียบการจดจำวัตถุของการแบ่งส่วนทั้ง 3 แบบ

4.2 ผลการทดสอบ Tactile Sensor และ การจดจำวัตถุจาก Tactile Image

การเก็บข้อมูลภาพได้ทำการทดสอบระบบโดยการนำวัตถุ 10 ชนิดคือแก้วน้ำ, ลูกกอล์ฟ, ขาม, รีโมทคอลโทรล, ประแจ, ประแจหกเหลี่ยม, แม่เหล็กบาศก์, แม่เหล็กทรงสามเหลี่ยม, แม่เหล็กกลมยาว และแม่เหล็กทรงรี มาทำการทดสอบ

4.2.1 Tactile Image จาก Tactile Sensor

ผลจากการทดลองที่ได้ในส่วนของภาพที่ได้จากการสัมผัสจะเป็นภาพแสดงในรูปที่ 4.5 ซึ่งแสดงให้เห็นถึงการกระจายตัวของแรงกดที่กระทำต่อเซ็นเซอร์

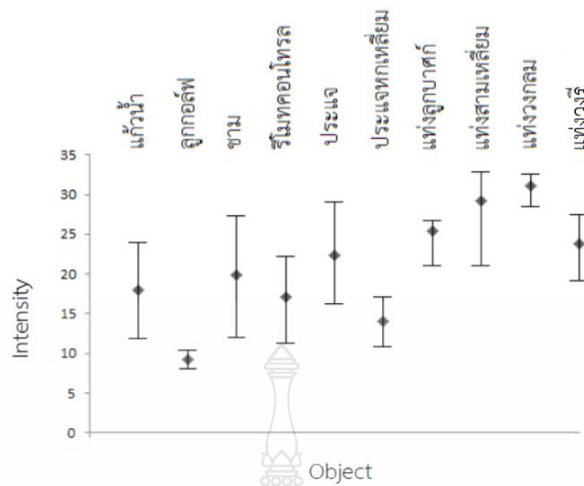


รูปที่ 4.5 ภาพ Tactile Image ของวัตถุประแจ แก้วกาแฟ และ ประแจหกเหลี่ยม

จากรูปที่ 4.5 เป็นรูปของตัวอย่างที่ได้จากการสัมผัสวัตถุประแจ แก้วกาแฟ และประแจหกเหลี่ยม เมื่อเซ็นเซอร์สัมผัสกับวัตถุจะเกิดรูปแบบแรงกดที่เกิดขึ้นกับเซ็นเซอร์เมื่อนำมาพล็อตเป็นรูปภาพ จะเห็นการกระจายตัวของแรงกดที่เกิดขึ้น ในรูปที่เป็นภาพที่เกิดจากการสัมผัสเมื่อนำประแจมาวางในมุมที่ต่างกันลักษณะการกระจายของแรงที่เกิดขึ้นบนตัวเซ็นเซอร์จะเปลี่ยนแปลงไปตามแนวหรือทิศทางของวัตถุ

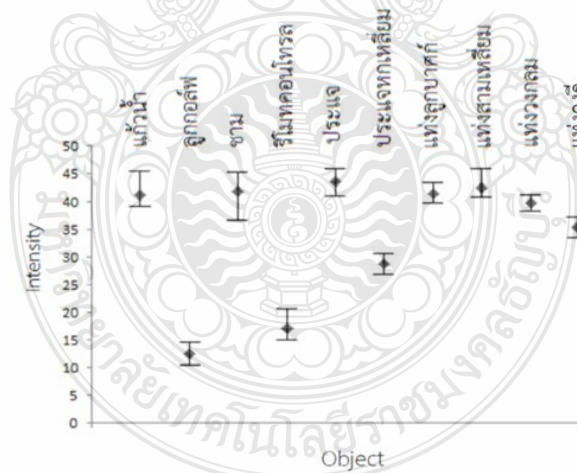
4.2.2 Descriptor

ภาพจากการสัมผัสที่ได้จะถูกดึง Feature ออกจากภาพเพื่อเป็นตัวแทนของภาพนั้น ในการทดลองนี้ใช้ Statistic Descriptor 2 วิธี ประกอบด้วยค่าเฉลี่ยของภาพ (μ) และค่าเบี่ยงเบนมาตรฐาน (σ) ค่าเฉลี่ยของแรงกดที่ได้แสดงดังรูปที่ 4.5 พบว่าแรงกดของแก้วน้ำ ขาม และ ประแจมีค่าเฉลี่ยในแต่ละภาพที่กว้างอยู่ในช่วงประมาณ 10-25 หน่วย ซึ่งเกิดจากลักษณะของตัววัตถุที่มีขนาดใหญ่กว่าเซ็นเซอร์จึงทำให้เมื่อนำเซ็นเซอร์มาสัมผัสภาพที่ได้จะมีความแตกต่างกันค่อนข้างมากแต่ค่าเฉลี่ยแรงกดของวัตถุ ลูกกอล์ฟ, ประแจหกเหลี่ยม, แท่งลูกบาศก์, แท่งสามเหลี่ยม, แท่งวงกลมและ แท่งวงรี มีค่าเฉลี่ยแรงกดของแต่ละภาพค่อนข้างเท่ากันเนื่องจากขนาดของวัตถุเล็กกว่าขนาดของเซ็นเซอร์แม้จะหมุนวัตถุอย่างไรรูปแบบแรงกดก็ยังคงมีลักษณะแบบเดิมจึงทำให้ค่าเฉลี่ยแรงกดคงที่



รูปที่ 4.6 ค่าเฉลี่ยของภาพที่ได้จากวัตถุแต่ละชนิด

จากรูปที่ 4.6 ค่าเบี่ยงเบนมาตรฐานของวัตถุชนิดเดียวกันมีค่าใกล้เคียงกันและวัตถุเกือบทุกชนิดจะมีค่าเบี่ยงเบนมาตรฐานอยู่ในช่วง 35-45 หน่วย แสดงให้เห็นว่าพื้นที่ของเซ็นเซอร์ถูกสัมผัสกับวัตถุเป็นบริเวณกว้างและจำนวน pixel ของเซ็นเซอร์มีหลาย pixel และกระจายรับแรงกดไปค่อนข้างเท่ากัน ยกเว้นลูกกอล์ฟ รีโมทคอนโทรลและประแจหกเหลี่ยม ที่มีค่าเบี่ยงเบนมาตรฐานต่ำแสดงว่าพื้นที่ส่วนใหญ่ของเซ็นเซอร์ไม่ถูกสัมผัส และในส่วนของที่ถูกสัมผัสมีแรงกดที่ถูกสัมผัสเท่ากัน



รูปที่ 4.7 ค่าเบี่ยงเบนมาตรฐานของภาพที่ได้จากวัตถุแต่ละชนิด

4.2.3 การจดจำวัตถุจากการใช้ Tactile Sensor

ในงานวิจัยนี้ได้ทำการทดสอบกับภาพจำนวนทั้งหมด 1000 ภาพจากการทดสอบกับวัตถุ 10 ชนิด ชนิดละ 100 ครั้งพบว่าหากใช้ค่าเฉลี่ยทำการ Training เพียงอย่างเดียว แล้วทำการ Testing พบว่า มีความถูกต้องเฉลี่ยอยู่ที่ 58.6% โดยวัตถุที่มีความถูกต้องในการจดจำมากที่สุดคือลูกกอล์ฟ มี

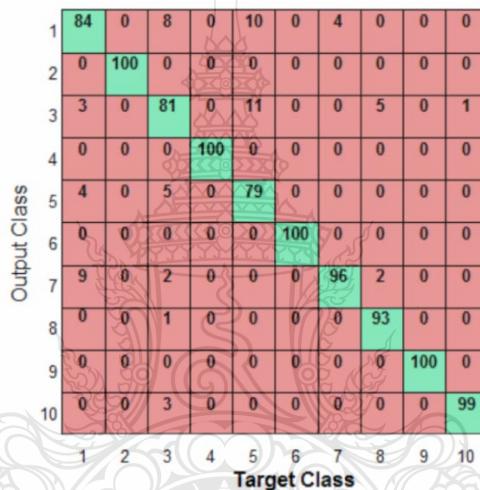
ค่าความถูกต้องในการจดจำถึง 100% และตัวที่มีค่าความถูกต้องในการจดจำ น้อยสุดคือประแจมีค่าความถูกต้อง 12% เท่านั้น เนื่องจากค่าเฉลี่ยของประแจมีค่าใกล้เคียงกับวัตถุอื่นจึงทำให้เกิดความผิดพลาด (Error) ในการจดจำสูง จากนั้นได้ทำการทดสอบระบบโดยการ Training ด้วยค่าเบี่ยงเบนมาตรฐานของรูปที่ได้จากการสัมผัสเพียงอย่างเดียว พบว่าความถูกต้องเฉลี่ยมากกว่าการสอนด้วยค่าเฉลี่ยโดยที่มีความถูกต้องเฉลี่ยเท่ากับ 86.9% โดยวัตถุที่มีค่าความถูกต้องในการจดจำมากที่สุดคือลูกกอล์ฟ รีโมทคอนโทรล และประแจหกเหลี่ยม มีค่าความถูกต้องในการจดจำที่ 100% และตัวที่มีค่าความถูกต้องในการจดจำน้อยสุดคือขาม มีค่าความถูกต้องในการจดจำที่ 54% เนื่องจากค่าเบี่ยงเบนมาตรฐานของขามมีความใกล้เคียงวัตถุอื่น และสุดท้ายคือการสอนด้วยคุณลักษณะทั้งสองตัวรวมกัน เป็นที่น่าสนใจว่าเมื่อใช้คุณลักษณะทั้งสองตัวรวมกันมาทำการสอนและทำการทดสอบพบว่าค่าเปอร์เซ็นต์ความถูกต้องเพิ่มขึ้น โดยมีค่าเปอร์เซ็นต์ความถูกต้องเฉลี่ยอยู่ที่ 93.2% โดยวัตถุที่มีค่าความถูกต้องในการจดจำมากที่สุดคือลูกกอล์ฟ, รีโมทคอนโทรล, ประแจหกเหลี่ยม และแท่งวงกลม มีค่าความถูกต้องในการจดจำถึง 100% และวัตถุที่มีค่าความถูกต้องในการจดจำน้อยสุดคือประแจ มีค่าเปอร์เซ็นต์ความถูกต้องที่ 79% ผลจากการทดลองพบว่าในกรณีที่ใช้ค่าเฉลี่ยเป็นข้อมูลอินพุตในการ Training ค่าความถูกต้องเฉลี่ยที่ได้จะต่ำสุดอยู่ที่ 58.6% เท่านั้น เนื่องจากการกระจายตัวของข้อมูลมีสูงจึงไม่สามารถเป็นตัวแทนของข้อมูลที่ดีพอ เมื่อใช้ค่าเบี่ยงเบนมาตรฐานมาเป็นค่าอินพุตในการ Training พบว่าความถูกต้องในการจดจำวัตถุมีค่าสูงกว่าการใช้ค่าเฉลี่ยมากเพราะมีค่าความถูกต้องเฉลี่ย 86.9% และเมื่อเราใช้ทั้งค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานมาใช้ในการ Training พบว่าค่าเฉลี่ยมีค่าสูงสุดถึง 93.2% ดังนั้นการ Training ระบบด้วยคุณลักษณะที่หลากหลายจะสามารถเพิ่มความถูกต้องในการจดจำวัตถุของระบบได้

ตารางที่ 4.4 ผลการจดจำวัตถุจากการใช้ Tactile Sensor ร่วมกับการทำ Statistic Descriptor

วัตถุที่ใช้ทดสอบ	% ความถูกต้อง		
	(μ)	(s)	(μ, s)
แก้วน้ำ	22	71	84
ลูกกอล์ฟ	100	100	100
ขาม	29	54	81
รีโมทคอนโทรล	43	100	100
ประแจ	12	55	79
ประแจหกเหลี่ยม	76	100	100

ตารางที่ 4.4 ผลการจดจำวัตถุจากการใช้ Tactile Sensor ร่วมกับการทำ Statistic Descriptor (ต่อ)

วัตถุที่ใช้ทดสอบ	% ความถูกต้อง		
	(μ)	(s)	(μ, s)
แท่งลูกบาศก์	70	93	96
แท่งสามเหลี่ยม	73	99	93
แท่งวงกลม	85	100	100
แท่งวงรี	76	97	99
% ความถูกต้องเฉลี่ย	58.6	86.9	93.2



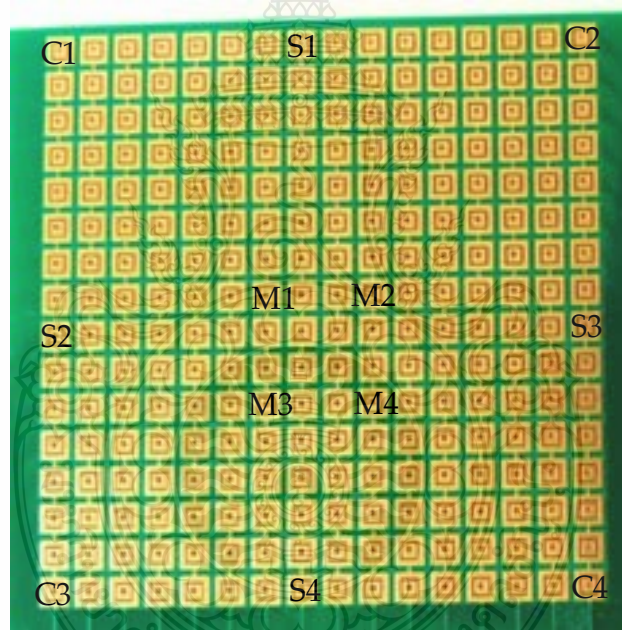
รูปที่ 4.8 ผลการทดสอบเมื่อ Training ระบบด้วยค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐาน

ในรูปที่ 4.6 เป็นผลที่ได้จากการทดสอบโดย Training ด้วยคุณลักษณะทั้งสองตัวร่วมกันค่าความผิดพลาดที่เกิดขึ้นเกิดจากตัวภาพจากการสัมผัสของประแจที่รายงานผลผลิตเป็นแก้วน้ำและชามเนื่องจากทั้งค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของประแจใกล้เคียงกับแก้วน้ำและชามจึงทำให้เกิดความผิดพลาดสูง ในทำนองเดียวกันผลของชามก็มีความผิดพลาดสูงรองลงมาเนื่องจากรายงานผลว่าเป็นประแจดังนั้นวัตถุที่มีรูปร่างต่างกันโดยสิ้นเชิงแต่มีโอกาสที่ค่าเฉลี่ยของความเข้มของแรงกดกับค่าเบี่ยงเบนมาตรฐานอาจจะใกล้เคียงกันเป็นผลให้การรายงานผลผิดพลาด และการรายงานผลของลูกกอล์ฟทั้งสามการทดสอบมีค่าความถูกต้องเป็น 100% เนื่องจากลักษณะการสัมผัสของลูกกอล์ฟกับเซ็นเซอร์จะเป็นเพียงจุดเล็กๆ จุดเดียว และลูกกอล์ฟเป็นบอลทรงกลมที่มีความสมมาตรในทุกทิศทางดังนั้นไม่ว่าจะหมุนลูกกอล์ฟอย่างไรภาพที่ได้ก็จะมีลักษณะเดิมเสมอจึงทำให้ค่าเฉลี่ยของภาพมีค่าต่ำแตกต่างจากตัวอื่นๆ มากจึงทำให้ระบบสามารถแยกแยะลูกกอล์ฟได้โดยง่ายค่าความถูกต้องจึงมีค่าสูงในทุกครั้งที่มีการทดสอบ

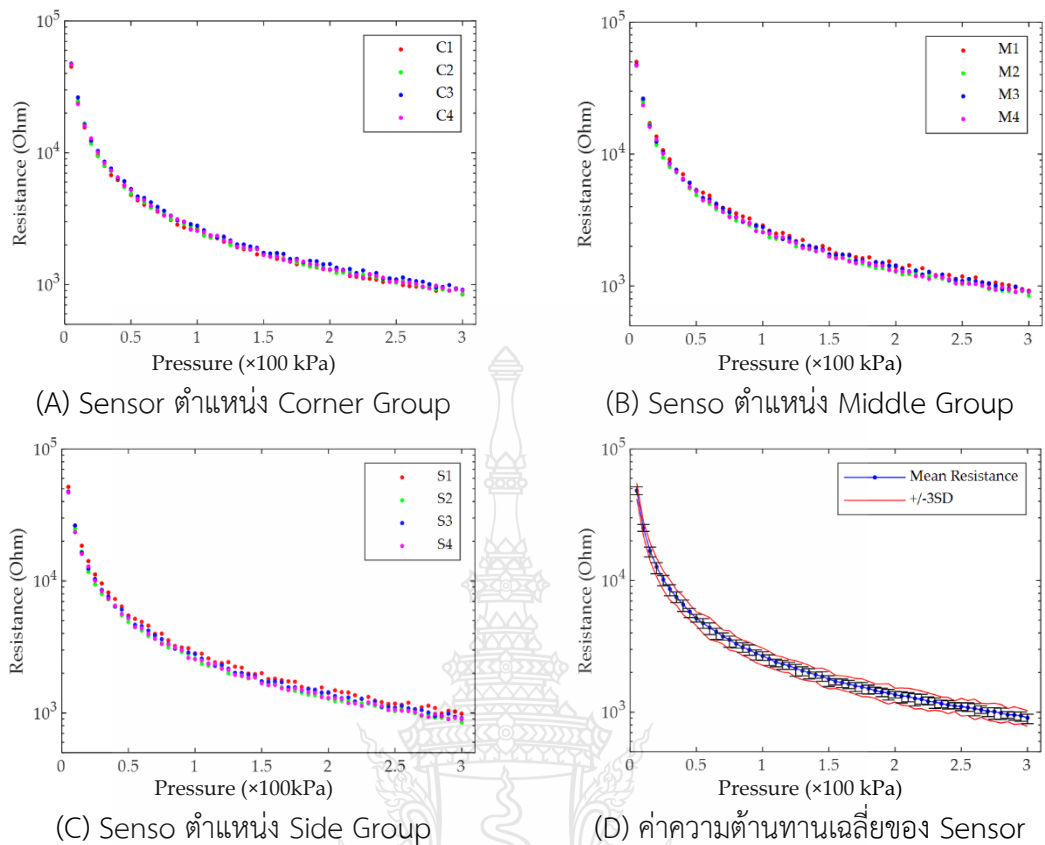
4.3 เซ็นเซอร์แบบใหม่

4.3.1 Sensor Sensitivity

Sensitivity เป็นการวัดคุณสมบัติด้านความต้านทานที่เกิดขึ้นต่อแรงที่มากระทำ ในการทดสอบนี้ได้แบ่งตำแหน่งที่ทดสอบเซ็นเซอร์ออกเป็น 3 กลุ่ม คือ 1) Sensor ตำแหน่ง Corner Group ประกอบด้วย C1, C2, C3 และ C4 2) ตำแหน่ง Side Group ประกอบด้วย S1, S2, S3 และ S4 3) ตำแหน่ง Middle Group ตรงกลางประกอบด้วย M1, M2, M3 และ M4 ดังรูปที่ 4.9 ซึ่งผลทดสอบแสดงดังรูปที่ 4.10 (A), (B) และ (C) ซึ่งผลการทดสอบที่ได้จาก Sensor กลุ่ม ตำแหน่ง Corner Group, Side Group, และ Middle Group ตามลำดับ ผลที่ได้จากรูปกราฟทั้ง 3 รูปจะพบว่าเมื่อเพิ่มแรงกด ค่าความต้านทานของ Sensor แต่ละตัวจะค่อยลดลง โดยในแต่ละตำแหน่งมีค่าความต้านทานที่ได้ใกล้เคียงกัน โดยในรูปที่ 4.9 (D) แสดงค่าความต้านทานของ Sensor เฉลี่ยของ Sensor ที่ทำการทดสอบจากทั้ง 3 Group



รูปที่ 4.9 ตำแหน่งวัดความต้านทานของเซ็นเซอร์



รูปที่ 4.10 ค่าความต้านทานของเซ็นเซอร์ต่อแรงที่กระทำ

เนื่องจากกระบวนการผลิต ค่าความต้านทานที่วัดได้จากเซ็นเซอร์แต่ละตัวมีการกระจายตัวแบบ Normal Distribution ซึ่งความไวที่ได้จากเซ็นเซอร์แต่ละตัวอาจไม่เท่ากัน จึงได้ทดสอบเพื่อวัดค่าที่ได้ของเซ็นเซอร์ โดยได้ทำการวัดทั้งหมด 12 จุดเพื่อเป็นตัวแทนของเซ็นเซอร์ ในรูปที่ 4.9 แสดงตำแหน่งที่ทำการวัดเซ็นเซอร์ 12 จุดประกอบด้วย C1, C2, C3, C4, S1, S2, S3, S4, M1, M2, M3, M4 โดยใช้แรงกดตั้งแต่ 10 kPa จนถึง 300 kPa ในรูปที่ 3.11 แสดงกราฟของค่า $\pm 3\sigma$ ซึ่งครอบคลุมความน่าจะเป็น 99.73% ของค่าความต้านทานเซ็นเซอร์ที่วัดได้ เมื่อสร้างความสัมพันธ์ระหว่างแรงและค่าความต้านทานของเซ็นเซอร์โดยใช้โมเดล สมการ $R = aP^b$, โดยที่ P และ R คือแรงดันและความต้านทานของเซ็นเซอร์ ตามลำดับ ค่าความต้านทานทั้งหมดคำนวณจากสมการ 4.1 ถึง 4.4 จากค่าความต้านทานที่วัดได้ในแต่ละจุดเมื่อนำมาสร้างสมการสอบเทียบค่าความต้านทานกับแรงกดจะได้

$$\text{ตำแหน่ง Sensor C1} \quad R = 216 \times P^{-0.97} \quad (4.1)$$

$$\text{ตำแหน่ง Sensor S1} \quad R = 254 \times P^{-0.97} \quad (4.2)$$

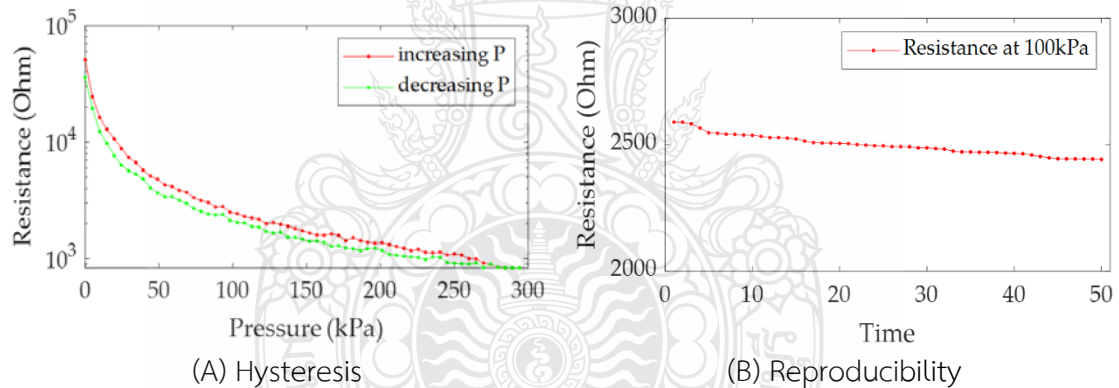
$$\text{ตำแหน่ง Sensor M1} \quad R = 244 \times P^{-0.97} \quad (4.3)$$

$$\text{ค่าความต้านทานเฉลี่ย} \quad R = 228 \times P^{-0.97} \quad (4.4)$$

เมื่อ R คือความต้านทานหน่วยเป็นโอห์ม $k\Omega$, และ P คือ แรงที่มากดเซ็นเซอร์ หน่วยเป็น kPa

4.3.2 Sensor Hysteresis and Reproducibility

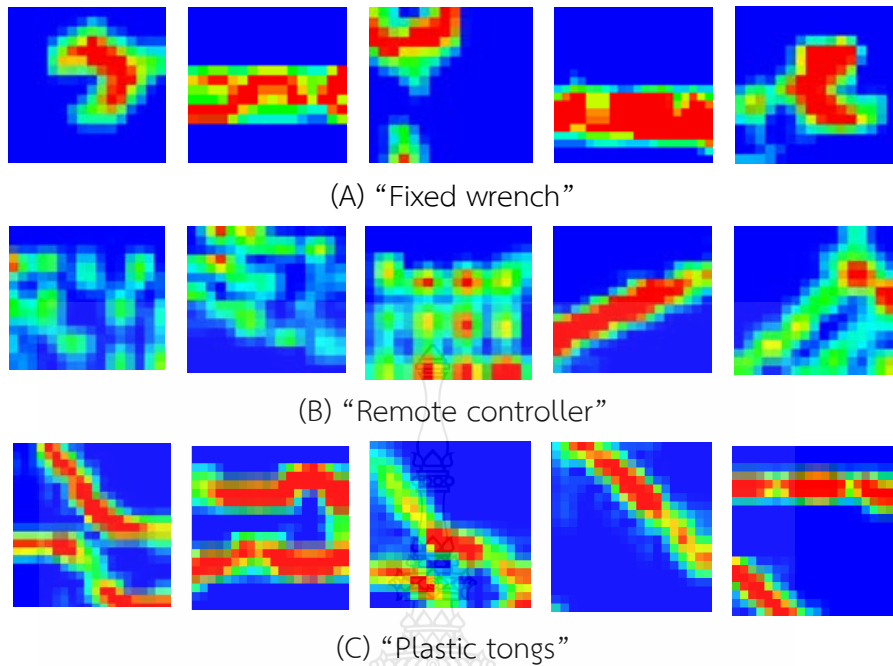
Hysteresis ของเซ็นเซอร์ได้รับการทดสอบด้วยแรงดันตั้งแต่ $0 kPa$ ถึง $300 kPa$ โดยเพิ่ม $10 kPa$ แต่ละขั้นตอนแล้วย้อนกลับเป็น $0 kPa$ โดยลดลง $10 kPa$ ในแต่ละขั้นตอน C1 คือ ตำแหน่งที่วัด Hysteresis พบว่าเซ็นเซอร์ได้รับเล็กน้อย พลังงานศักย์จากแรงดันก่อนหน้า ดังแสดงในรูปที่ 4.11 (A) ความสามารถในการทำซ้ำของเซ็นเซอร์ ทดสอบด้วยแรงดัน $100 kPa$ เซ็นเซอร์วัดความต้านทาน ถูกวัดโดยการกดทุกนาที พบว่าเซ็นเซอร์มีค่าความต้านทานลดลงเล็กน้อย ดังแสดงในรูปที่ 4.11 (B) ดังนั้นเซ็นเซอร์ที่พัฒนายังคงมีค่า Hysteresis อยู่ซึ่งส่งผลต่อ Reproducibility ของเซ็นเซอร์อยู่บ้าง แต่ค่าดังกล่าวไม่ส่งผลโดยตรงต่อรูปแบบของโปรไฟล์วัตถุ ซึ่งเป็นจุดประสงค์ของการใช้เซ็นเซอร์



รูปที่ 4.11 คุณสมบัติของเซ็นเซอร์

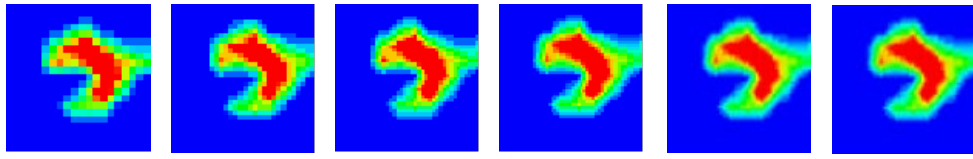
4.3.3 Tactile Image and Resolution Enhance

ในการทดลองนี้ได้สร้างชุดข้อมูลจากการจับวัตถุโดยจับวัตถุละ 200 ครั้ง จำนวน 20 วัตถุ ส่งผลให้ได้ภาพสำหรับชุดข้อมูล 4000 ภาพเป็นฐานข้อมูล เนื่องจากในการทดลองนี้ไม่ขึ้นกับทิศทาง และตำแหน่งภาพที่ได้จากการจับจึงมีการหมุนวัตถุจึงได้ภาพที่แตกต่างกันของวัตถุดังรูปที่ 4.12 แสดงภาพ Tactile Image จากวัตถุ 3 ชนิดในตำแหน่งต่างๆ ชุดข้อมูลนี้ถูกนำไปทดสอบในหัวข้อ ชุดข้อมูลนี้ถูกนำมาใช้ในการทดสอบในหัวข้อ 3.4.1 และ 3.4.2

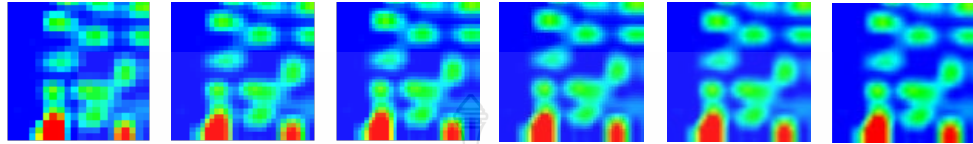


รูปที่ 4.12 ตัวอย่างภาพ Tactile Images (A) "Fixed wrench", (B) "Remote controller", (C) "Plastic tongs"

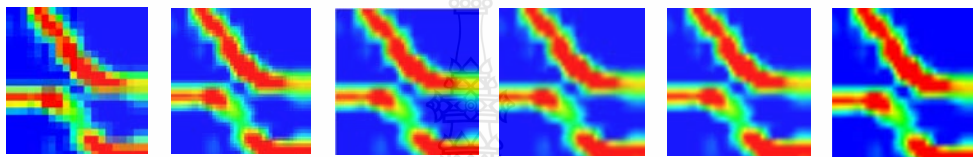
ในรูปที่ 4.13 แสดงการทำ Bicubic เพื่อเพิ่มความละเอียดของภาพ จากภาพที่ได้จาก Tactile Sensor Array ขนาด 16×16 pixels เป็นภาพขนาด 32×32 , 64×64 , 128×128 , 256×256 และ 512×512 pixels พบว่าคุณภาพของภาพดีขึ้นอย่างมาก เนื่องจากภาพความละเอียดต่ำการระบุโปรไฟล์ทางเรขาคณิตของวัตถุนั้นเป็นเรื่องยาก เมื่อความละเอียดสูงขึ้นรายละเอียดทางเรขาคณิตของภาพที่สัมผัสได้ก็ชัดเจนขึ้น ภาพที่ปรับปรุงแล้วเหล่านี้ถูกใช้เป็นข้อมูลป้อนเข้าสำหรับการพัฒนาระบบการจดจำ



(A) Tactile Image ของ “Fixed wrench”



(B) Tactile Image ของ “Remote controller”



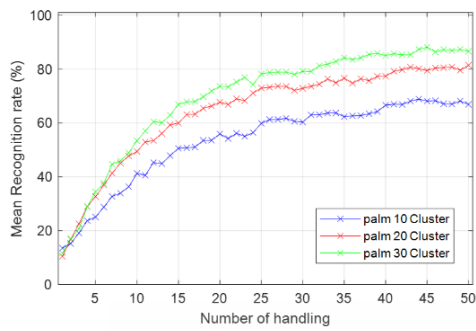
(C) Tactile Image ของ “Plastic tongs”

รูปที่ 4.13 ภาพที่ได้จากการทำ Bicubic Interpolation

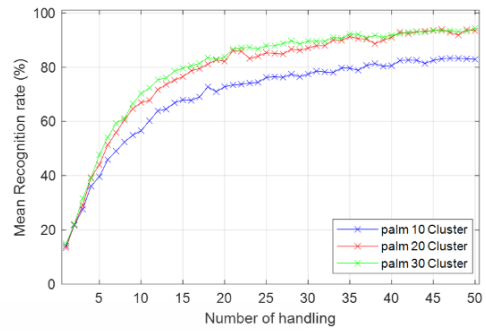
4.4 การจดจำวัตถุจากการสัมผัสสำหรับมือหุ่นยนต์

4.4.1 อัลกอริทึมแบบ BoF สำหรับมือหุ่นยนต์

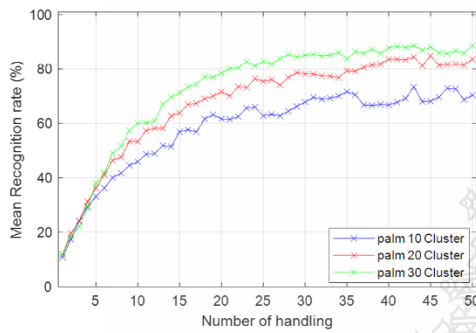
การทดสอบการจดจำด้วยวิธี BoF พบว่าความแม่นยำในการจดจำเพิ่มขึ้นเมื่อจำนวนการสัมผัสเพิ่มขึ้น ในการทดลองแบ่งตามวิธีที่ใช้ในการ Classify โดยได้ทำการทดลองด้วยวิธี Classify 3 วิธี คือ KNN, ANN และ SVM โดยได้ทดลองกับการทำ Descriptor 3 วิธี คือ Moment Analysis Descriptor, Polar Fourier Descriptor และ SIFT Descriptor และในการทำ Clustering ด้วย K-Mean Clustering ของวิธี BoF ได้ทดสอบกับ Cluster 3 ขนาดคือ 10 คลัสเตอร์ ($w=10$), 20 คลัสเตอร์ ($w=20$) และ 30 คลัสเตอร์ ($w=30$), ผลการทดลองด้วยวิธี KNN แสดงดังรูปที่ 4.14, ผลการทดลองด้วยวิธี ANN แสดงดังรูปที่ 4.15 และ ผลการทดลองด้วยวิธี SVM แสดงดังรูปที่ 4.16 เมื่อเปรียบเทียบผลทั้งหมดพบว่าการใช้วิธี Polar Fourier Descriptor ร่วมกับ SVM (รูปที่ 4.16) ให้ผลการจดจำสูงที่สุดเนื่องจากมีความแม่นยำมากกว่า 81.35% เมื่อสัมผัสวัตถุ 14 ครั้ง ($p=14$) ที่ขนาดขนาดคลัสเตอร์ 30 คลัสเตอร์ ($w=30$), การใช้วิธี Polar Fourier ร่วมกับ KNN ให้ความแม่นยำใกล้เคียงกันคือ ความแม่นยำมากกว่า 80% เมื่อสัมผัสวัตถุ 15 ครั้ง ($p=15$) ที่ขนาดขนาดคลัสเตอร์ 30 คลัสเตอร์ ($w=30$), ส่วนวิธี ANN ต้องจับวัตถุ 19 ครั้ง ($p=19$) ความแม่นยำจึงจะสูงกว่า 80%



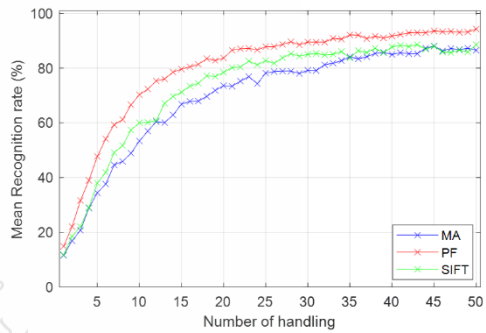
(A) Moment Analysis



(B) Polar Fourier

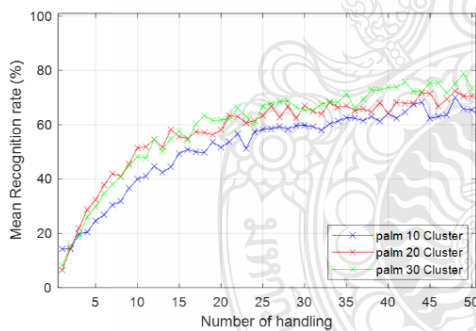


(C) SIFT

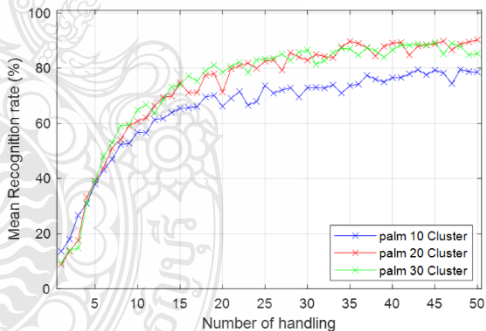


(D) KNN

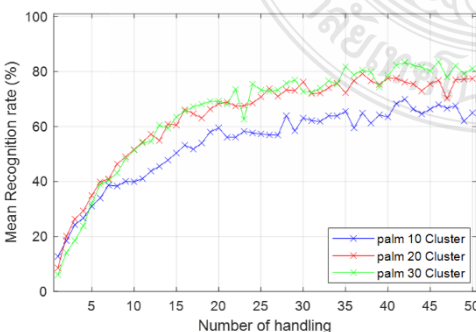
รูปที่ 4.14 ผลการจดจำเมื่อใช้ KNN



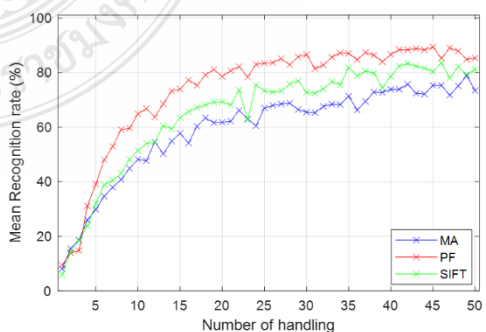
(A) Moment Analysis



(B) Polar Fourier

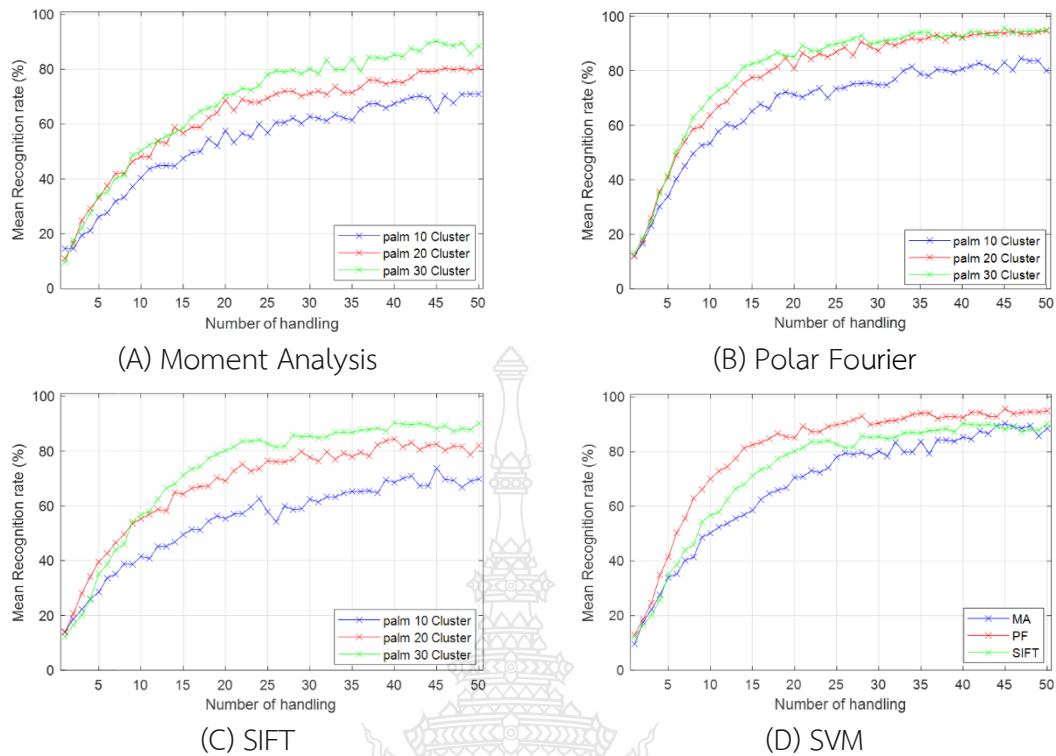


(C) SIFT



(D) ANN

รูปที่ 4.15 ผลการจดจำเมื่อใช้ ANN

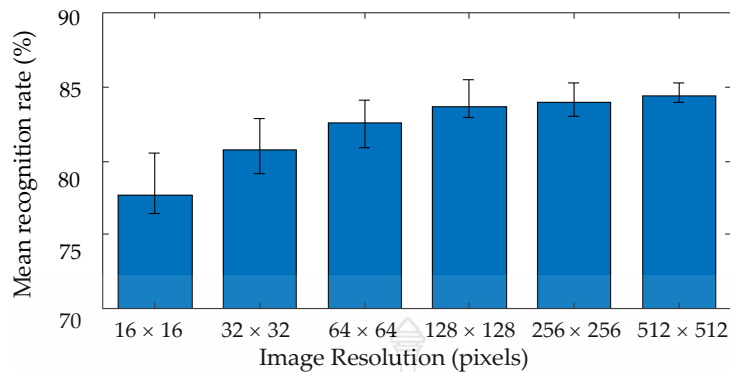


รูปที่ 4.16 ผลการจดจำเมื่อใช้ SVM

4.4.2 การทดลองอัลกอริทึมแบบ DCNN สำหรับมือหุ่นยนต์

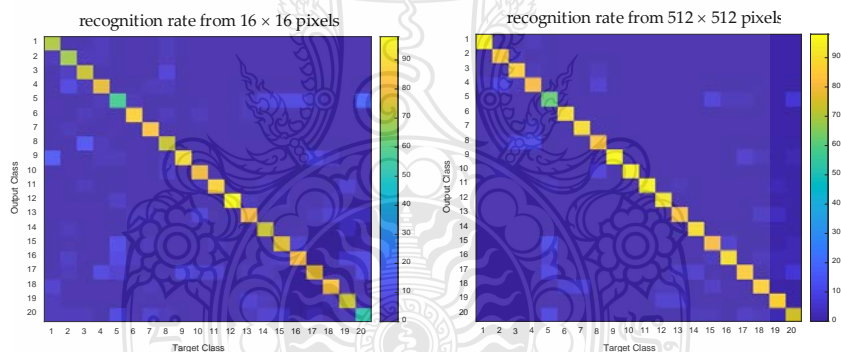
4.4.2.1 Recognition Rate from Resolution Enhanced

ในการทดสอบนี้ ใช้ AlexNet โดยมีอัตราการเรียนรู้เริ่มต้นที่ 0.001 การทดสอบได้ดำเนินการเพื่อแสดงผลการจดจำที่เกิดขึ้นเมื่อเปลี่ยนความละเอียดของภาพ ผลที่ได้แสดงในรูปที่ 4.17 ซึ่งพบว่าประสิทธิภาพการจดจำวัตถุเพิ่มขึ้นเมื่อเพิ่มความละเอียดของ Tactile Image ผลการจดจำมีค่า 77.66% เมื่อใช้ภาพที่มีความละเอียด 16 × 16 pixels เมื่อทำการเพิ่มความละเอียดของภาพโดยใช้การแก้ไขแบบ Bicubic ความละเอียดเป็น 32 × 32 pixels ผลการจดจำเพิ่มขึ้นเป็น 80.79% และผลการจดจำจะเพิ่มขึ้นในทุกระดับที่ทำการเพิ่มความละเอียดของภาพจนถึงค่าความละเอียด 128 128 pixels แต่มีความชันในการเพิ่มลดลง จากนั้นค่าการจดจำเพิ่มขึ้นเพียงเล็กน้อยเมื่อปรับความละเอียดจาก 128 × 128 pixels เป็น 512 × 512 pixels อัตราการจดจำสูงสุดอยู่ที่ 84.42% ผลลัพธ์เหล่านี้ชี้ให้เห็นว่าประสิทธิภาพของการจดจำวัตถุโดยใช้ DCNN สำหรับการจัดการวัตถุด้วย Tactile Sensor ความละเอียดต่ำสามารถเพิ่มขึ้นได้ด้วยการเพิ่มปรับปรุงความละเอียดของภาพผ่านวิธี Bicubic Interpolation ดังนั้น ความละเอียด 512 × 512 pixels จึงใช้สำหรับการทดสอบเพิ่มเติมกับเครือข่าย DCNN อื่นๆ



รูปที่ 4.17 ผลการจดจำเมื่อทำการปรับปรุงความละเอียดของภาพ

ผลลัพธ์การรู้จำที่ได้รับสามารถนำเสนอเป็น Confusion Matrix รูปที่ 4.18 แสดงการเปรียบเทียบ Confusion Matrix สองตัวที่มีความละเอียดของภาพที่นำมาทดสอบแตกต่างกัน เมื่อปรับปรุงภาพความละเอียดภาพความแม่นยำในการทำนายดีขึ้น เช่นผลการทำนาย class 1, 2 และ 20 ดีขึ้นอย่างชัดเจน และการทำนายของ class 9, 10 และ 11 ก็ดีขึ้นเล็กน้อยเช่นกัน

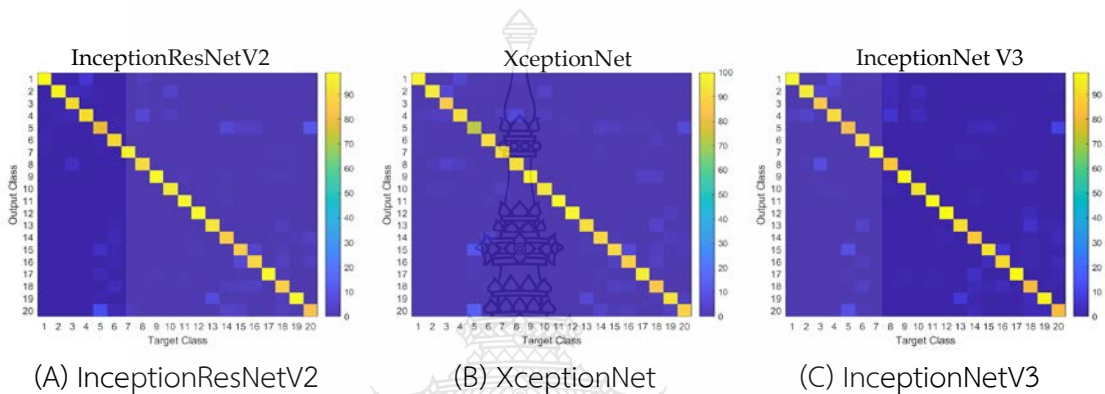


รูปที่ 4.18 ผลการจดจำเมื่อทำการปรับปรุงความละเอียดของภาพ 16 × 16 Pixels และ 512 × 512 Pixels

4.4.2.2 Recognition Results from DCNN

ในส่วนนี้เป็นผลการทดสอบที่ได้จาก DCNN 19 Model โดยได้ทำการทดสอบกับภาพขนาด 16 × 16 Pixels และ 512 × 512 Pixels พบว่า DCNN ที่ให้ผลลัพธ์ที่ดีที่สุดคือ InceptionResNetV2, XceptionNet และ InceptionNetV3 โดยใช้ Initial Learning Rate ที่ 0.01 ผลการทดสอบที่ได้จาก Network ทั้งสามนี้แสดงไว้ในรูปที่ 4.19 ซึ่งพบว่า InceptionResNetV2 ลดข้อผิดพลาดสำหรับ Class 5, 15 และ 20 ได้อย่างมีประสิทธิภาพมากกว่า Network อื่นๆ ผลที่ได้จาก DCNN ทั้ง 3 ตัวมีค่าใกล้เคียงกัน และให้ผลดีกว่า DCNN อื่นๆ มาก จึงเหมาะกับการนำมาใช้งานกับ

Tactile Image อย่างไรก็ตามข้อผิดพลาดที่เกิดขึ้นอยู่ที่ Class 5 และ Class 20 เป็นส่วนใหญ่ แต่ InceptionResNetV2 สามารถลดข้อผิดพลาดจาก 2 Class นี้ได้ดีกว่า DCNN ตัวอื่นๆ เนื่องจากโครงสร้างของ DCNN ของ Model InceptionResNetV2, XceptionNet และ InceptionNetV3 นี้มีพื้นฐานมาจาก GoogLeNet หรือ InceptionNetV1 ซึ่งให้ผลการทดสอบที่ดีต่อ Tactile Image แสดงให้เห็นว่าโครงสร้างของ Network แบบ Inception สามารถทำงานได้ดีกว่า Network แบบอื่นๆ เมื่อนำมาทดสอบกับภาพที่มีรายละเอียดต่ำ



รูปที่ 4.19 Confusion Matrix แสดงผลที่ได้จากการทดสอบด้วย InceptionResNetV2, XceptionNet, และ InceptionNetV3

ผลการทดสอบตามตารางที่ 4.5 DCNN ที่ให้ผลลัพธ์ที่ดีที่สุดคือ InceptionResNetV2, XceptionNet และ InceptionNetV3 ตามลำดับ และในผลการทดสอบในตารางที่ 4.5 ยังแสดงให้เห็นอีกด้วยว่า SqueezeNet, EfficienNetB0 and DarkNet 19 ให้ค่า Recognition Rate ที่ค่อนข้างต่ำจึงไม่เหมาะสมกับ Tactile Image ถึงแม้ว่า DCNN นี้จะให้ผลที่ดีที่สุดสำหรับ Image Net ก็ตาม เมื่อเปรียบเทียบโครงสร้างของ Network ชนิด SqueezeNet และ EfficienNetB0 พบว่ามีโครงสร้างในลักษณะที่ใช้ ตัวกรองขนาด 1×1 จำนวนมาก ดังนั้นโครงสร้างที่ใช้ตัวกรองขนาด 1×1 ไม่เหมาะกับการนำมาใช้กับภาพที่มีรายละเอียดต่ำ ซึ่งแตกต่างจาก Network ที่มีพื้นฐานมาจาก InceptionNet ที่มีตัวกรองขนาดใหญ่กว่า (3×3 , 5×5) ซึ่งให้ผลการทดสอบที่ดีกว่า ทั้งนี้ DarkNet 19 ให้ผลการทำนายที่ถูกต้องต่ำ เนื่องจาก Network มีขนาดเล็กและถูกปรับปรุงมาใช้สำหรับการตรวจจับวัตถุจึงไม่เหมาะสมกับการนำมาใช้กับการทำ Classification

ตารางที่ 4.5 ผลการทดสอบการจดจำวัตถุของมือหุ่นยนต์ด้วยวิธี DCNN

Model	Accuracy of 16 × 16 Pixels				Accuracy of 512 × 512 Pixels			
	Min	Max	Mean	SD	Min	Max	Mean	SD
AlexNet	76.40	80.55	77.66	1.27	84	85.2	84.42	0.40
VGG16	79.95	81.85	80.92	0.58	85.7	88.35	86.94	0.78
VGG19	80.10	84.20	81.84	1.11	86.4	88.6	87.62	0.66
GoogLeNet	85.15	86.30	85.57	0.36	89.2	91.01	90.09	0.59
ResNet18	80.05	83.00	81.36	0.99	85.25	87.25	86.5	0.61
ResNet50	81.55	84.60	82.87	0.95	86.1	88.55	87.59	0.82
ResNet101	81.75	85.40	83.61	1.12	86.75	89.65	88.07	0.85
Place365GoogLeNet	74.80	81.00	78.40	1.84	87.2	89.75	88.56	0.80
InceptionNetV3	85.25	87.45	86.47	0.75	90.1	91.8	91.06	0.50
EfficienNetB0	71.25	76.95	74.18	2.01	71.25	76.95	74.18	2.01
SqueezeNet	39.02	41.50	40.21	1.71	45.5	51.6	48.33	1.71
InceptionResNetV2	84.55	87.20	85.15	0.91	90.7	93.05	91.86	0.70
DarkNet19	61.60	68.05	65.01	1.74	75	78.8	77.36	1.17
DarkNet53	77.30	81.55	79.44	1.43	83.85	86.2	85.42	0.90
DenseNet201	82.40	86.10	83.79	1.15	87.9	89.6	88.69	0.56
XceptionNet	82.45	86.15	84.48	1.06	91	91.85	91.4	0.29
NASNetMobile	85.10	87.35	86.36	0.69	84.35	91.6	89.72	2.08
ShuffleNet	81.65	83.80	82.65	0.78	87.25	88.8	88.16	0.55
MobileNetV2	82.30	85.55	83.92	0.92	87.35	89.9	88.23	0.73

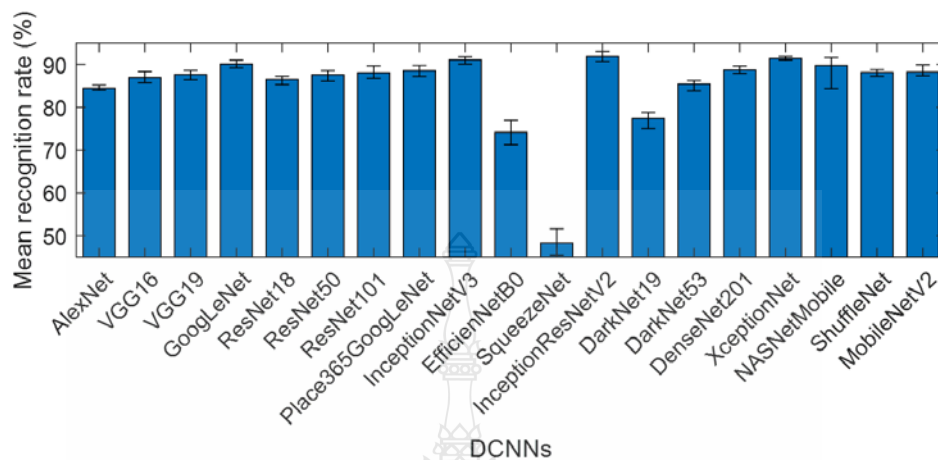
ผลการทดสอบตามตารางที่ 4.6 เป็นผลที่ได้จากการทดสอบการจดจำวัตถุจากการทดสอบด้วยภาพขนาด 16 × 16 pixels ซึ่งเป็นผลที่ได้จาก InceptionNetV3 และ ผลที่ได้จากการทดสอบด้วยภาพขนาด 512 × 512 pixels ซึ่งเป็นผลที่ได้จาก InceptionResNetV2 พบว่าวัตถุที่ระบบมีประสิทธิภาพสูงคือ Computer mouse มีความถูกต้อง 99.2% และวัตถุที่ให้ผลการทำนายถูกต้องน้อยที่สุดคือ Cup มีความถูกต้อง 78.5% เนื่องจากมีข้อมูลภาพ Tactile image คล้ายกับ Cream bottle

ตารางที่ 4.6 ผลการจดจำของวัตถุแต่ละ Class

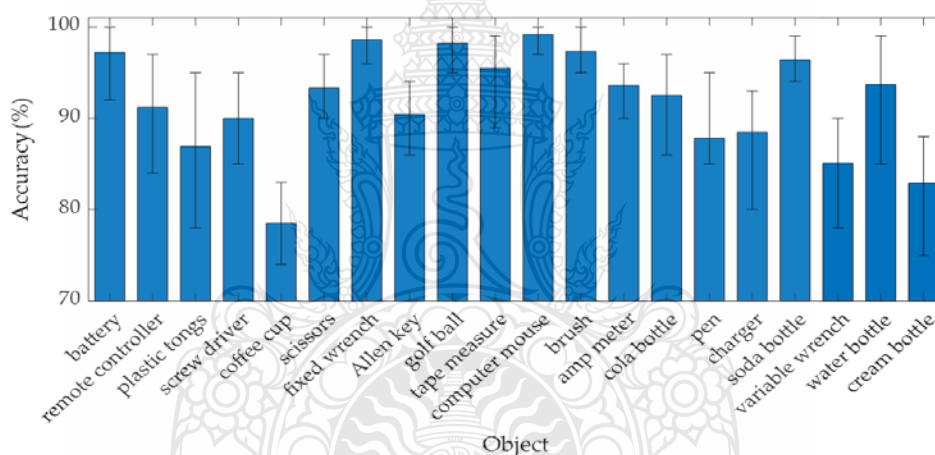
Class	Object	Accuracy of 16 × 16 Pixels (%)	Accuracy of 512 × 512 pixels (%)
1.	Battery	95.0	97.2
2.	Remote Controller	84.8	91.2
3.	Plastic tongs	79.6	86.9
4.	Screw driver	82.2	90.0
5.	Cup	67.6	78.5
6.	Scissors	92.2	93.3
7.	Fixed wrench	93.2	98.6
8.	Allen key	86.9	90.4
9.	Golf ball	96.9	98.2
10.	Tape measure	90.9	95.5
11.	Computer mouse	85.6	99.2
12.	Brush	93.6	97.3
13.	Amp meter	90.7	93.6
14.	Cola bottle	88.2	92.5
15.	Pen	82.7	87.8
16.	Charger	85.2	88.5
17.	Soda Bottle	87.3	96.4
18.	Variable wrench	82.5	85.1
19.	Water bottle	85.3	93.7
20.	Cream bottle	78.9	82.9
	average	86.47	91.86

ในรูปที่ 4.20 (A) แสดงผลการจดจำที่ได้จาก DCNN 19 Model และ ในรูปที่ 4.20 (B) แสดงผลการจดจำที่ได้จาก InceptionResNetV2 ของแต่ละ Class พบว่าวัตถุที่ระบบมีประสิทธิภาพสูง คือ Computer mouse, Golf ball and Fixed wrench ซึ่งมีค่าความถูกต้องสูงกว่า 98.2%, ส่วนวัตถุ Class ที่ให้ผลการจดจำต่ำสุดคือ Cup และ Cream bottle ซึ่งมีความถูกต้องต่ำกว่า 78.5% และ

82.9% ตามลำดับ ดังผลเปรียบเทียบในรูปที่ 4.18 และ รูปที่ 4.19 แสดงให้เห็นว่าความผิดพลาดที่เกิดขึ้นกับ Class 5 มีความสัมพันธ์กับ Class 20



(A) ผลการจดจำจาก DCNN 19 Model



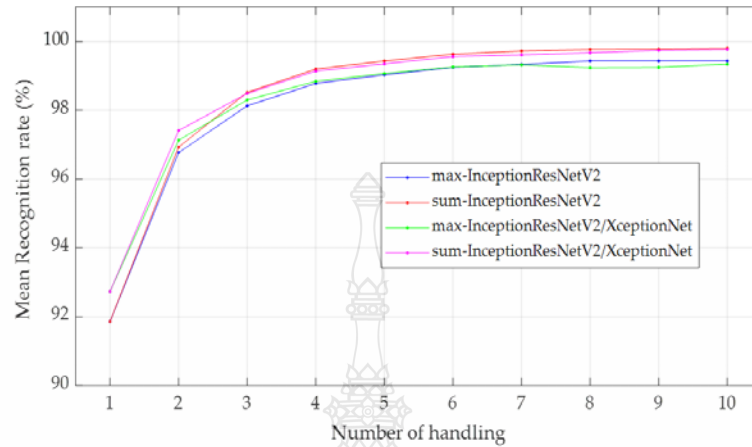
(B) ผลการจดจำวัตถุแต่ละ Class ด้วย InceptionResNetV2

รูปที่ 4.20 ผลการจดจำที่ทำการทดสอบด้วย DCNN ของภาพปรับปรุงขนาด 512 × 512 Pixels

4.4.2.3 การจดจำวัตถุจากการทำ Object Exploration

วิธีการของ BoF จะสำรวจวัตถุอย่างละเอียดโดยการจัดการกับวัตถุหลายครั้งเพื่อให้สามารถทำนายผลลัพธ์ได้อย่างถูกต้อง ดังนั้นจึงได้ทดลองสำรวจวัตถุโดยการจับวัตถุหลายครั้งและใช้วิธี DCNN ในการวิเคราะห์ การเพิ่มประสิทธิภาพในการจดจำสามารถทำได้โดยการจับวัตถุหลายครั้งแล้วหาค่า Maximum of Probability (P_{max}) หรือ Summation of Probability (P_{sum}) ผลการทดลองแสดงผลในรูปที่ 4.21 โดยการเปรียบเทียบผลโดยใช้ 2 วิธีในการวิเคราะห์คือ InceptionResNetV2 และการใช้ InceptionResNetV2 ร่วมกับ XceptionNet ผลที่ได้พบว่าเมื่อเพิ่มจำนวนครั้งในการจับวัตถุก่อนทำนายผลจะให้ค่าความถูกต้องสูงขึ้นอย่างมาก จากค่าความถูกต้อง 91.86 เป็น 96.93% เมื่อ

จับวัตถุเพิ่มขึ้นเป็น 2 ครั้ง อัตราความถูกต้องเริ่มคงที่และไม่เพิ่มขึ้นเมื่อมีการจับวัตถุมากกว่า 6 ครั้ง โดยมีค่าความถูกต้องที่ 99.63% สำหรับการจัดการมากกว่า 6 ครั้ง และหลังจากจับวัตถุประมาณ 10 ครั้งก่อนทำนายผลพบว่ามีค่าความถูกต้องในการทำนายผลถึง 99.9%



รูปที่ 4.21 ผลการจดจำเมื่อทำการสำรวจวัตถุก่อนทำนายผล

รูปที่ 4.22 เปรียบเทียบ Confusion Matrix เมื่อทำการสำรวจวัตถุหลายครั้งก่อนทำนายผล ผลการทำนายที่ได้จะมีความถูกต้องมากขึ้น ดังรูปเมื่อทำการจับวัตถุ 10 ครั้งก่อนทำนายผลพบว่ามีค่าความถูกต้องในการทำนายผลสูงมากเมื่อเปรียบเทียบกับตอนจับวัตถุครั้งเดียว

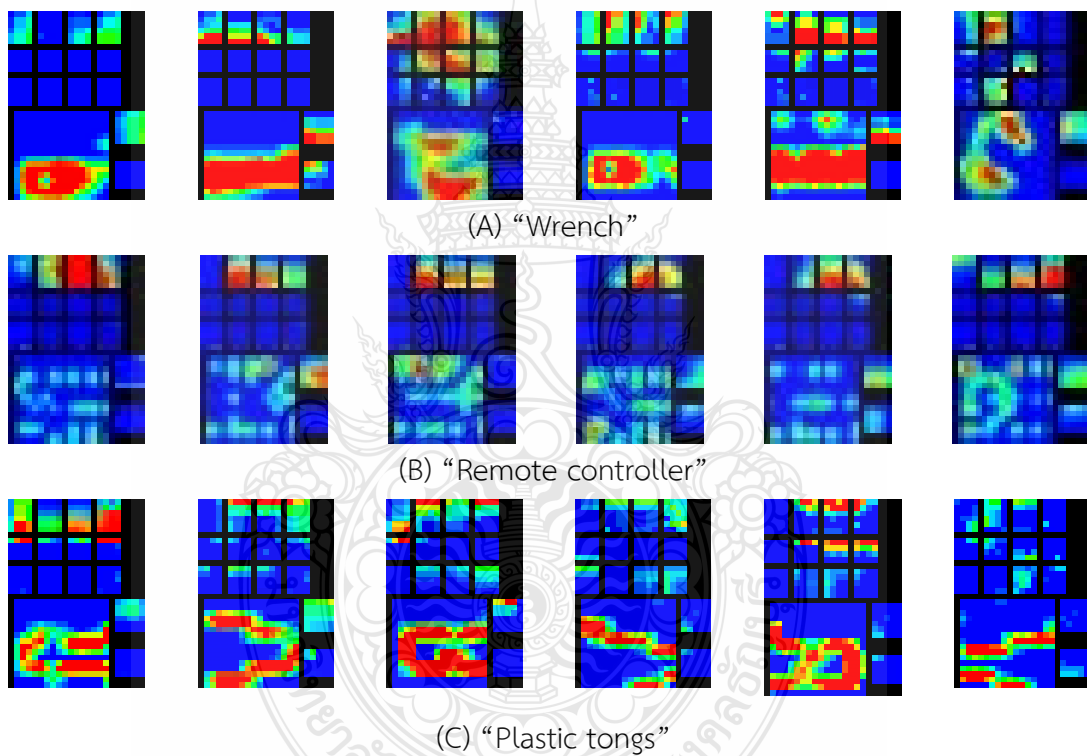


รูปที่ 4.22 Confusion Matrix แสดงผลการจดจำเมื่อทำการสำรวจวัตถุก่อนทำนายผล

4.5 การจดจำวัตถุจาก Glove Tactile Sensor

การทดลองนี้อธิบายการใช้วิธี Artificial Intelligence สำหรับการจดจำวัตถุจากการสัมผัส เมื่อรูปแบบของ Tactile Sensor มีลักษณะโครงสร้างแบบเดียวกับมือของมนุษย์ด้วยการนำเสนอการจดจำวัตถุจากการใช้ Glove Tactile Sensor ซึ่งมีการติดตั้ง Tactile Sensor กระจายอยู่ทั่วทั้งมือ มีจุดสัมผัสเหมือนกับมือของมนุษย์ มีลักษณะเป็น Tactile Sensor ขนาดเล็กกระจายตัวทั่วทั้งมือ Tactile Image ที่ได้จากการจับวัตถุถูกนำมาเปรียบเทียบประสิทธิภาพโดยการใช้ 2 วิธีในการวิเคราะห์

คือวิธี BoF และวิธี CNN ในการทดสอบได้ทดสอบกับการจับวัตถุ 20 Class ที่มีลักษณะแตกต่างกัน สำหรับการทดลองนี้ในการทดสอบระบบจดจำวัตถุได้สร้างชุดข้อมูลจากการจับวัตถุจริงจำนวน 20 Object หรือ 20 Class เป็นวัตถุที่มีรูปร่างมาตรฐาน 10 วัตถุและวัตถุในชีวิตประจำวันอีก 10 วัตถุ วัตถุทั้งหมดประกอบด้วย Battery, Cup, Plastic tongs, Remote controller, Golf ball, Scissors, Variable wrench, Allen key, Screwdriver, Measuring tape, Cube, Rectangular, Triangular prism, Pyramid, Tetrahedron, Large cylinder, Small Cylinder, Cone, Oval และ Sphere การสร้าง Data Set เกิดจากการจับวัตถุและบันทึกภาพ วัตถุละ 200 ภาพ รวมภาพทั้งหมด 4,000 เนื่องจากในการทดลองนี้ไม่ขึ้นกับทิศทางและตำแหน่งภาพที่ได้จากการจับจึงมีการหมุนวัตถุจึงได้ภาพที่แตกต่างกันดังตัวอย่างภาพจากการจับวัตถุในรูปที่ 4.23



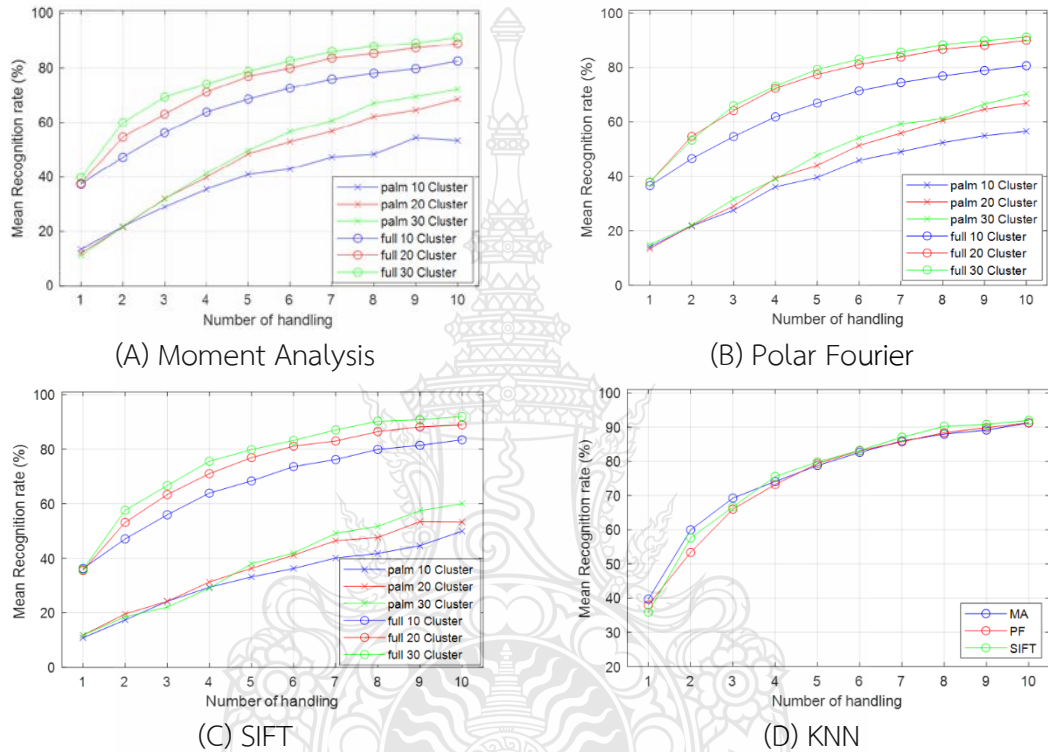
รูปที่ 4.23 ภาพ Tactile Image ที่ได้จากการจับวัตถุด้วยเซ็นเซอร์ถุงมือ

4.5.1 อัลกอริทึมแบบ BoF สำหรับ Glove Tactile Sensor

ในการทดลองที่ใช้ BoF นั้นได้ทดสอบกับ Classifier KNN, ANN และ SVM โดยใช้ Descriptor ด้วยวิธี Moment Analysis Descriptor, SIFT Descriptor และ Polar Fourier Descriptor และได้ทดสอบกับ Cluster 10, 20 และ 30 Cluster ผลการทดลองได้แบ่งตามวิธีการ Classifier คือวิธี KNN, ANN และ SVM ตามลำดับ โดยผลที่ได้เป็นดังนี้

4.5.1.1 KNN

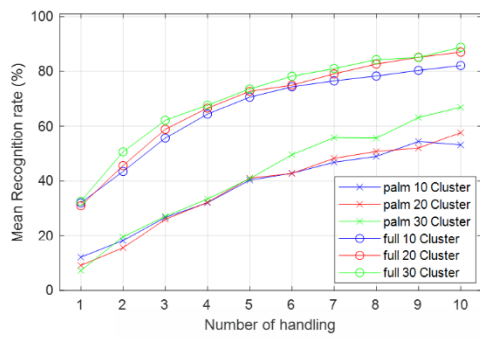
ในรูปที่ 4.24 แสดงผลการทดสอบด้วยวิธี BoF โดยใช้วิธี KNN เป็น Classifier โดยการ ใช้ Descriptor ทั้ง 3 วิธี พบว่าการแบ่ง Cluster เป็น 30 Cluster ให้ผลการทำนายถูกต้องสูงที่สุดในทุก Descriptor และเมื่อเปรียบเทียบในทุก Descriptor พบว่าการในช่วงแรกการใช้ Moment Analysis ให้ค่าความถูกต้องสูงกว่าวิธีอื่น แต่เมื่อจับวัตถุไปเรื่อยๆ จนถึง 4 ครั้งเป็นต้นไปพบว่า วิธี SIFT จะให้ค่าความถูกต้องสูงกว่า



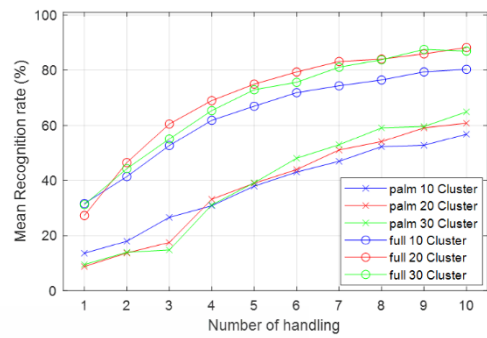
รูปที่ 4.24 ผลการทดสอบด้วยวิธี BoF โดยใช้วิธี KNN เป็น Classifier

4.5.1.2 ANN

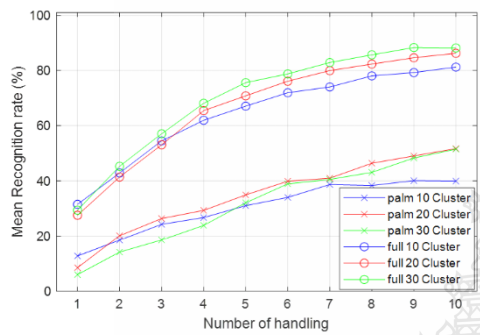
ในรูปที่ 4.25 แสดงผลการทดสอบด้วยวิธี BoF โดยใช้วิธี ANN เป็น Classifier โดยการ ใช้ Descriptor ทั้ง 3 วิธี พบว่าการแบ่ง Cluster เป็น 30 Cluster ให้ผลการทำนายถูกต้องสูงที่สุด สำหรับ Moment Analysis Descriptor และ SIFT Descriptor ส่วน Polar Fourier Descriptor พบว่าการแบ่ง Cluster เป็น 20 Cluster ให้ผลดีที่สุด และเมื่อเปรียบเทียบในทุก Descriptor พบว่าการในช่วงแรกการใช้ Moment Analysis ให้ค่าความถูกต้องสูงกว่าวิธีอื่น แต่เมื่อจับวัตถุไปเรื่อยๆ จนถึง 4 ครั้งเป็นต้นไป วิธี SIFT จะให้ค่าความถูกต้องสูงกว่า



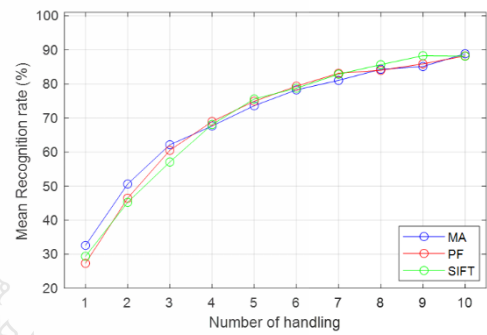
(A) Moment Analysis



(B) Polar Fourier



(C) SIFT

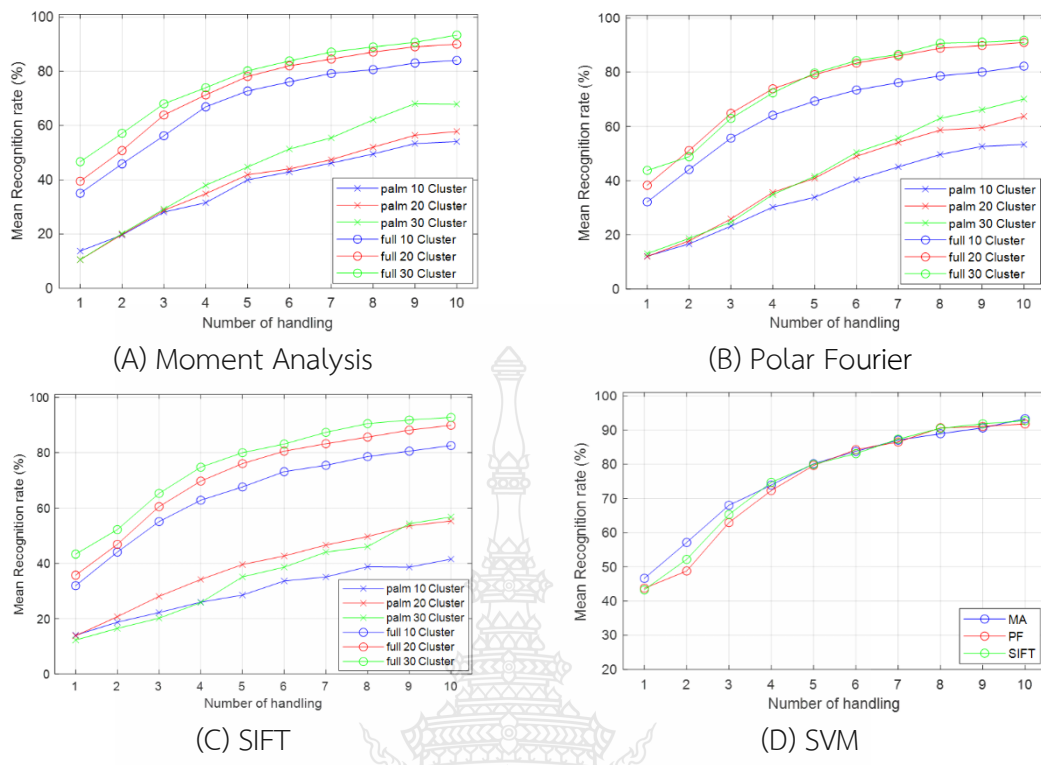


(D) ANN

รูปที่ 4.25 ผลการทดสอบด้วยวิธี BoF โดยใช้วิธี ANN เป็น Classifier

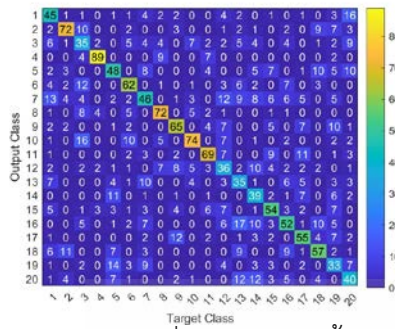
4.5.1.3 SVM

ในรูปที่ 4.26 แสดงผลการทดสอบด้วยวิธี BoF โดยใช้วิธี SVM เป็น Classifier โดยการใช้ Descriptor ทั้ง 3 วิธี พบว่าการแบ่ง Cluster เป็น 30 Cluster ให้ผลการทำนายถูกต้องสูงที่สุด สำหรับ Moment Analysis Descriptor และ SIFT Descriptor ส่วน Polar Fourier Descriptor พบว่าการแบ่ง Cluster เป็น 30 Cluster ให้ผลดีที่สุด และเมื่อเปรียบเทียบในทุก Descriptor พบว่าการในช่วงแรกการใช้ Moment Analysis ให้ค่าความถูกต้องสูงกว่าวิธีอื่น แต่เมื่อจับวัตถุไปเรื่อยๆ จนถึง 4 ครั้งเป็นต้นไปวิธี SIFT จะให้ค่าความถูกต้องสูงกว่าเช่นกัน เปรียบเทียบผลจาก Classifier ทั้ง 3 แบบพบว่าการใช้วิธี SVM เป็น Classifier ให้ผลที่ดีที่สุด เนื่องจากให้ค่าการจดจำสูงที่สุด

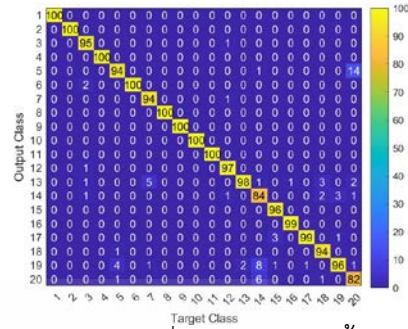


รูปที่ 4.26 ผลการทดสอบด้วยวิธี BoF โดยใช้วิธี SVM เป็น Classifier

จากรูปที่ 4.27 แสดง Confusion Matrix ของผลการจดจำที่ได้จากวิธี BoF เมื่อนำมาใช้กับ Tactile Sensor แบบ Glove Tactile Sensor โดยเป็นผลทดลองจาก Model SVM ร่วมกับ SIFT Descriptor ซึ่งวิธี BoF ให้ค่า Accuracy เมื่อจับวัตถุครั้งเดียวที่ 53.9% ในและเมื่อจับวัตถุที่ 10 ครั้งผลการจดจำเพิ่มเป็น 96.4% นั่นคือเมื่อใช้วิธี BoF สามารถให้ค่า Accuracy เพิ่มขึ้นเมื่อทำการสำรวจวัตถุหรือการจับวัตถุหลายครั้ง เมื่อเปรียบเทียบผลกับการทดลองในหัวข้อ 4.4.1 พบว่าการใช้ถุงมือแบบ Glove Tactile Sensor สามารถเพิ่มความถูกต้องในการจับวัตถุมากกว่าการติดตั้งเซ็นเซอร์ที่ฝ่ามือเพียงจุดเดียว จากรูปที่ 4.27 เมื่อจับวัตถุเพียงครั้งเดียวพบว่า Plastic tong, Rectangular, Triangular prism, Pyramid และ Oval มีความถูกต้องในการทำนายผลต่ำมาก กล่าวคือต่ำกว่า 40% แต่เมื่อเพิ่มจำนวนครั้งในการจับวัตถุความถูกต้องในการทำนายผลของ Class เหล่านี้สูงขึ้นมากอยู่ในช่วง 84% - 95% ส่วน Remote controller, Tape measure, Cup และ Allen key มีผลการทำนายที่มีความถูกต้องสูงตั้งแต่การจับเพียงครั้งเดียว คือสูงกว่า 70% ซึ่งพบว่าวัตถุให้ผลการทำนายสูงมักเป็นวัตถุที่ไม่ใช่วัตถุรูปทรงมาตรฐาน เนื่องจากวัตถุรูปทรงมาตรฐานมีบางส่วนของรูปทรงคล้ายคลึงกันเช่น Cube กับ Rectangular หรือ Oval กับ Sphere เป็นต้น ซึ่งแตกต่างจากวัตถุทั่วไปที่มักมีเอกลักษณ์ที่ชัดเจนกว่าเช่น Remote controller ที่ด้านหนึ่งเป็นปุ่มและอีกด้านเป็นสัน เป็นต้น



(A) ผลเมื่อจับวัตถุ 1 ครั้ง



(B) ผลเมื่อจับวัตถุ 10 ครั้ง

รูปที่ 4.27 ผลการจดจำวัตถุด้วยวิธี BoF เมื่อจับวัตถุ 1 ครั้งด้วยวิธี SVM ร่วมกับ SIFT Descriptor

4.5.2 การทดลองอัลกอริทึมแบบ DCNN สำหรับ Glove Tactile Sensor

เมื่อนำ Data Set ชุดเดียวกันกับวิธี BoF มาทดสอบโดยใช้วิธี DCNN 7 Model ประกอบด้วย AlexNet, VggNet19, GoogleNet, InceptionNetV3, InceptionRestNetV2, XceptionNet และ NASNetMobile ผลการทดสอบแสดงดังตารางที่ 4.7 ซึ่งผลที่ได้พบว่า DCNN ที่ให้ความถูกต้องในการทำนายผลสูงที่สุดคือ InceptionNetV3 โดยมีค่าความถูกต้องอยู่ที่ 98.28 % และ InceptionNetV3 ที่ให้ผลใกล้เคียงกันโดยมีค่าความถูกต้องอยู่ที่ 98.26 % ผลที่ได้แสดงให้เห็นได้อย่างชัดเจนว่าวิธี DCNN มีความถูกต้องสูงกว่าวิธี BoF ดังนั้นจึงเป็นวิธีที่เหมาะสมกับการนำไปประยุกต์ใช้กับระบบจดจำของหุ่นยนต์ Humanoid มากกว่า

ตารางที่ 4.7 ผลการจดจำจากการทดสอบ Glove Tactile Sensor ด้วย DCNN

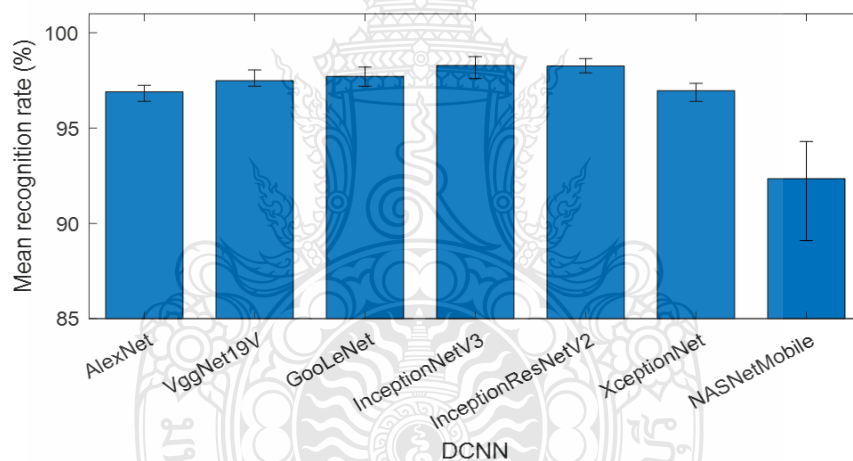
Model	Accuracy			
	Min	Max	Mean	SD
AlexNet	96.40	97.25	96.91	0.30
VggNet19	97.20	98.05	97.49	0.35
GoogleNet	97.20	98.20	97.71	0.33
InceptionNetV3	97.60	98.75	98.28	0.38
InceptionRestNetV2	97.90	98.65	98.26	0.35
XceptionNet	96.40	97.35	96.97	0.42
NASNetMobile	89.10	94.30	92.35	2.26

ผลการทดสอบการจดจำวัตถุแต่ละตัวแสดงตามตารางที่ 4.8 ซึ่งเป็นผลที่ได้จาก InceptionRestNetV2 และ InceptionNetV3 พบว่าผลที่ได้จาก InceptionNetV3 มีความถูกต้องสูงกว่า และวัตถุที่ระบบมีประสิทธิภาพสูงสุดคือ Cup และ Tape measure มีความถูกต้อง 100.0% และผลที่ได้จากทุกวัตถุที่ใช้วิธี InceptionNetV3 สูงกว่า 95.0% ซึ่งถือว่าสูงกว่าวิธี BoF มาก

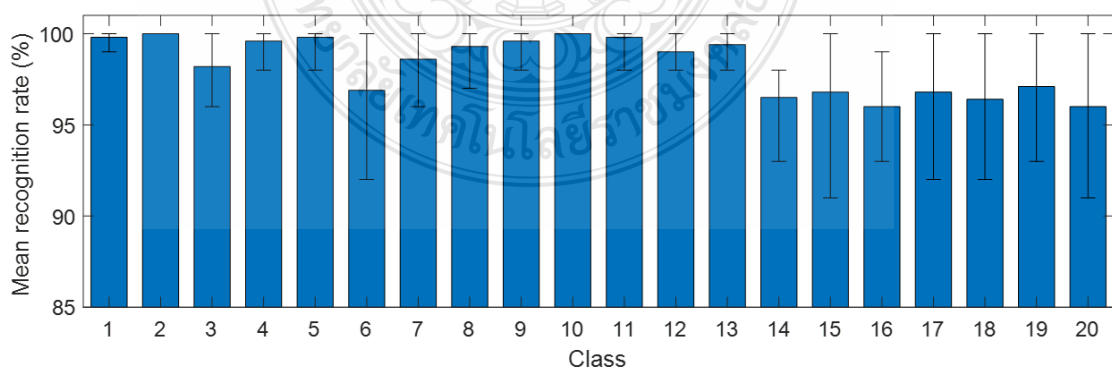
ตารางที่ 4.8 ผลการจดจำของวัตถุแต่ละ Class สำหรับ Glove Tactile Sensor

Class	Object	Accuracy of InceptionRestNetV2	Accuracy of InceptionNetV3
1.	Battery	99.7	99.8
2.	Cup	100.0	100.0
3.	Plastic tongs	94.1	98.2
4.	Remote controller	98.7	99.6
5.	Golf Ball	97.0	99.8
6.	Scissors	98.4	96.9
7.	Variable wrench	97.5	98.6
8.	Allen key	97.5	99.3
9.	Screwdriver	99.6	99.6
10.	Tape measure	99.8	100.0
11.	Cube	100.0	99.8
12.	Rectangular	99.2	99.0
13.	Triangular prism	99.1	99.4
14.	Pyramid	97.6	96.5
15.	Tetrahedron	98.0	96.8
16.	Large cylinder	99.6	96.0
17.	Small cylinder	98.3	96.8
18.	Cone	96.8	96.4
19.	Oval	98.8	97.1
20.	Sphere	95.9	96.0
	average	98.26	98.28

ในรูปที่ 4.28 (A) แสดงผลการจดจำที่ได้จาก DCNN 7 Model ซึ่งผลที่ได้จาก InceptionNetV3 ให้ค่าความถูกต้องสูงที่สุด และ InceptionResNetV2 ได้ผลการทดสอบใกล้เคียงกับ InceptionNetV3 ในรูปที่ 4.28 (B) แสดงผลการจดจำที่ได้จากการทดสอบด้วย InceptionNetV3 ซึ่งแสดงผลการจดจำในแต่ละ Class พบว่าวัตถุที่ระบบมีประสิทธิภาพสูงคือ Cub และ Tape measure ซึ่งมีค่าความถูกต้องสูงถึง 100.0%, ส่วนวัตถุ Class ที่ให้ผลการจดจำต่ำสุดคือ Sphere เนื่องจากภาพ tactile Image ของ Sphere คล้ายกับ golf ball จึงทำให้ถูกทำนายผลเป็น golf ball และพบว่า วัตถุที่มีรูปทรงมาตรฐานหลายอันมีความผิดพลาดมากกว่าวัตถุในชีวิตประจำวัน ส่วนหนึ่งอาจเนื่องมาจาก วัตถุที่มีรูปทรงมาตรฐานหลายอย่างมีลักษณะคล้ายคลึงกัน ผลเปรียบเทียบในรูปที่ 4.29 แสดงให้เห็นว่าความผิดพลาดที่เกิดขึ้นกับ Class 5 (golf ball) มีความสัมพันธ์กับ Class 20 (Sphere) เนื่องจาก ภาพ tactile Image ของ golf ball คล้ายกับ Sphere นั้นเอง และเมื่อเปรียบเทียบผลระหว่าง AlexNet กับ InceptionNetV3 จะพบว่า AlexNet มีการทำนายผลที่ผิดพลาดที่ Class 5 และ Class 20 สูงกว่าโดยเฉพาะอย่างเห็นได้ชัด ดังนั้นค่าความถูกต้องจึงต่ำกว่าวิธี InceptionNetV3

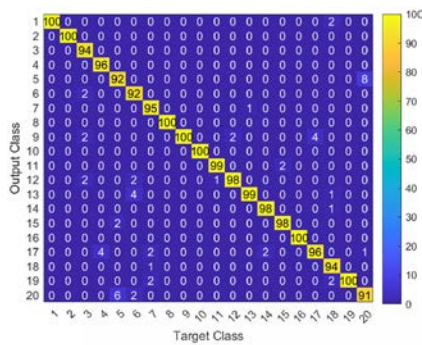


(A) ผลการจดจำจาก DCNN 7 Model

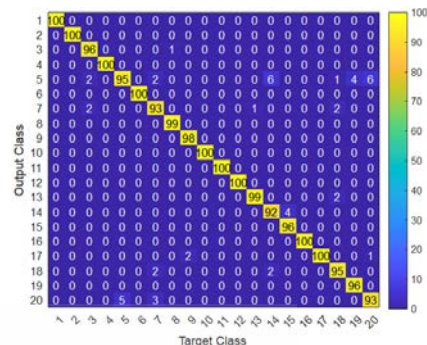


(B) ผลการจดจำวัตถุแต่ละ Class ด้วย InceptionNetV3

รูปที่ 4.28 ผลการจดจำวัตถุด้วย Glove Tactile Sensor ที่ได้รับการทดสอบด้วย DCNN



(A) AlexNet

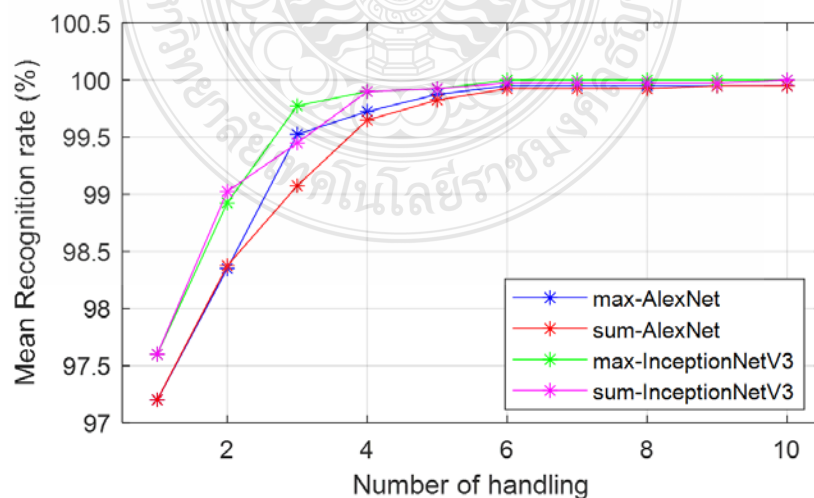


(B) InceptionNetV3

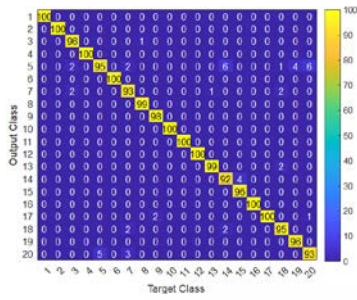
รูปที่ 4.29 Confusion Matrix แสดงผลการจัดจำวัตถุด้วยวิธี DCNN AlexNet และ InceptionNetV3

4.5.3 การจัดจำวัตถุจากการทำ Object Exploration ด้วย Glove Tactile Sensor

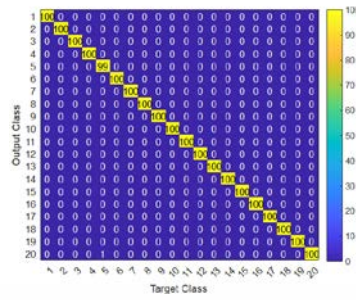
เช่นเดียวกับหัวข้อ 4.4.2.3 ที่จะสำรวจวัตถุอย่างละเอียดโดยการจัดการจับวัตถุหลายครั้ง เพื่อให้สามารถทำนายผลลัพธ์ได้อย่างถูกต้อง ดังนั้นจึงได้ทดลองสำรวจวัตถุโดยการจับวัตถุหลายครั้ง และใช้วิธี DCNN ในการวิเคราะห์แล้วหาค่า Maximum of Probability (P_{max}) และ Summation of Probability (P_{sum}) ผลการทดลองแสดงในรูปที่ 4.30 โดยการเปรียบเทียบผลโดยใช้ 2 วิธีในการวิเคราะห์คือ AlexNet และ InceptionNetV3 ผลที่ได้พบว่าเมื่อเพิ่มจำนวนครั้งในการจับวัตถุก่อนทำนายผลค่าความถูกต้องจะสูงขึ้นเมื่อใช้วิธี InceptionNetV3 จากค่าความถูกต้อง 97.2% เพิ่มเป็น 99.0% เมื่อจับวัตถุเพิ่มขึ้นเป็น 2 ครั้ง อัตราความถูกต้องเริ่มคงที่และไม่เพิ่มขึ้นเมื่อมีการจับวัตถุมากกว่า 6 ครั้ง โดยที่ InceptionNetV3 มีค่าความถูกต้องที่ 100% สำหรับการจัดการมากกว่า 6 ครั้ง และหลังจากจับวัตถุประมาณ 10 ครั้งก่อนทำนายผลพบว่าทั้งการใช้วิธี Maximum of Probability และ Summation of Probability มีความถูกต้องในการทำนายผลถึง 100% ซึ่งแสดงดังรูปที่ 4.31



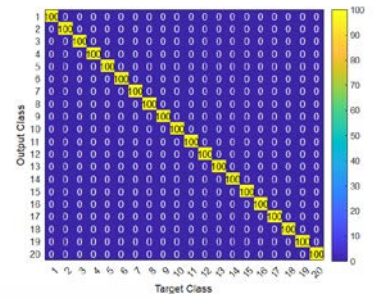
รูปที่ 4.30 ผลการจัดจำเมื่อทำการสำรวจวัตถุก่อนทำนายผล



(A) 1 image



(B) 10 image Maximum of Probability



(C) 10 image Summation of Probability

รูปที่ 4.31 Confusion Matrix แสดงผลการจดจำเมื่อทำการสำรวจวัตถุ 10 ครั้งก่อนทำนายผล



บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

ในบทนี้จะกล่าวถึงสรุปผลการวิจัยที่ทำมาทั้งหมด รวมไปถึงข้อเสนอนแนะในการนำงานวิจัยนี้ไปพัฒนาหรือประยุกต์ใช้ในอนาคตต่อไป

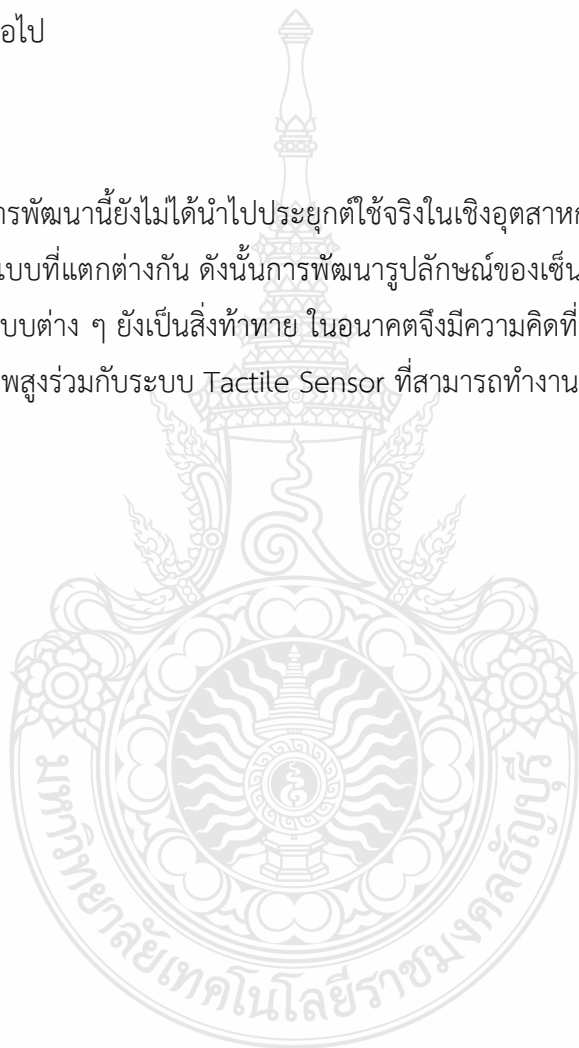
5.1 สรุปผลการวิจัย

งานวิจัยนี้ทำการศึกษาการจดจำวัตถุจากการประมวลผลภาพที่ได้จากการสัมผัส โดยการศึกษาเริ่มจากการศึกษาการจดจำวัตถุจากการจับวัตถุของมือมนุษย์ซึ่งในการทดลองได้ใช้วิธี Machine Learning ในการวิเคราะห์ผลในการจดจำ โดยใช้ ANN ในการดำเนินงานซึ่งพบว่าการแบ่งส่วนของภาพที่ได้จากการสัมผัสตามโครงสร้างของมือมนุษย์สามารถทำการจดจำวัตถุได้เป็นอย่างดี จากนั้นจึงได้ทำการพัฒนางานวิจัยโดยใช้ Tactile Sensor ร่วมกับการพัฒนาวงจรอ่านค่า และทำการวิเคราะห์การจดจำด้วยวิธี ANN ร่วมกับ Statistic Descriptor พบว่าระบบการจดจำสามารถทำงานได้เป็นอย่างดี แต่เนื่องจากขนาดของเซ็นเซอร์มีขนาดไม่เหมาะสมกับการติดตั้งที่ฝ่ามือ จึงได้ออกแบบ Tactile Sensor แบบ Piezoresistive โดยมีขนาด 16×16 Pixels มีขนาดกายภาพ 56.0×56.0 mm^2 โดยพัฒนาให้ Sensor มีขนาดที่เหมาะสมกับขนาดของฝ่ามือ และพัฒนาให้มีความไวในการวัดสูงขึ้นโดยการออกแบบ Electrode แบบ Square Electrode เพื่อลด Dead Area ของเซ็นเซอร์ จากนั้นนำมาติดตั้งเข้ากับฝ่ามือของหุ่นยนต์และทดสอบจับวัตถุเพื่อทำนายผลของวัตถุที่อยู่ในมือ การทดสอบพัฒนาบน Algorithm 2 วิธีคือ BoF และ DCNN โดยวิธี BoF ใช้ Descriptor 3 วิธีคือ Moment Analysis, Polar Fourier และ SIFT ร่วมกับวิธี Classify 3 วิธีคือ KNN, ANN และ SVM ร่วมกับการทำ Sub-Image พบว่าวิธีการใช้วิธี Polar Fourier Descriptor ร่วมกับ SVM ให้ผลการจดจำสูงที่สุดเนื่องจากมีความแม่นยำมากกว่า 80% เมื่อสัมผัสวัตถุ 14 ครั้ง ($p=14$) ที่ขนาดคลัสเตอร์ 30 คลัสเตอร์ ($w=30$) และการใช้วิธี Polar Fourier ร่วมกับ KNN ให้ความแม่นยำใกล้เคียงกันคือ มีความแม่นยำมากกว่า 80% เมื่อสัมผัสวัตถุ 15 ครั้ง ($p=15$) ที่ขนาดคลัสเตอร์ 30 คลัสเตอร์ ($w=30$) ส่วนวิธี ANN ต้องจับวัตถุ 19 ครั้งความแม่นยำจึงจะสูงกว่า 80% และผลการทดสอบด้วยวิธี DCNN 19 Model พบว่าวิธี InceptionNetV3 ให้ค่าความถูกต้องสูงสุดที่ 86.47% เมื่อจับวัตถุเพียงครั้งเดียว และเมื่อทำการปรับปรุงภาพด้วยวิธี Bicubic Interpolation เพิ่มขนาด Tactile Image เป็น 512×512 Pixels พบว่าวิธี InceptionResNetV2 ให้ค่าความถูกต้องสูงสุดที่ 91.86% จากนั้นได้พัฒนา Glove Tactile Sensor ซึ่งประกอบด้วย Tactile Sensor Array 15 จุดตามจุดสัมผัสของฝ่ามือ และทำการทดสอบด้วยวิธี BoF และ DCNN พบว่าการใช้ Glove Tactile Sensor ให้ความถูกต้องสูงกว่าการใช้ Tactile

Sensor ที่ฝ่ามือเพียงจุดเดียว โดยเมื่อใช้วิธีวิธี BoF พบว่าการใช้ SVM เป็น Classifier ร่วมกับ Moment Analysis Descriptor ให้ความถูกต้องสูงที่สุด โดยให้ความถูกต้องสูงกว่า 80% ในการจับวัตถุ 5 ครั้ง ($p=5$) เมื่อแบ่ง Cluster เป็น 30 Cluster ($w=30$) และวิธี DCNN พบว่าการใช้วิธี InceptionNetV3 ให้ความถูกต้องสูงที่สุดโดยให้ค่าความถูกต้องที่ 98.28% ในการจับวัตถุเพียงครั้งเดียว ดังนั้นการพัฒนาทั้งหมดนี้เป็นการพัฒนาการจดจำวัตถุจากการสัมผัสวัตถุซึ่งได้จากการประมวลผลข้อมูลที่อ่านได้จาก Tactile Sensor เพื่อเป็นพื้นฐานในการนำไปประยุกต์ใช้สำหรับระบบจดจำวัตถุของหุ่นยนต์เสมือนมนุษย์ต่อไป

5.2 ข้อเสนอแนะ

เนื่องด้วยการพัฒนานี้ยังไม่ได้นำไปประยุกต์ใช้จริงในเชิงอุตสาหกรรม เพราะลักษณะของมือหุ่นยนต์นั้นมีการออกแบบที่แตกต่างกัน ดังนั้นการพัฒนารูปลักษณะของเซ็นเซอร์ให้สามารถใช้ได้กับมือหุ่นยนต์เสมือนมนุษย์แบบต่าง ๆ ยังเป็นสิ่งท้าทาย ในอนาคตจึงมีความคิดที่จะมุ่งเน้นไปที่การพัฒนามือหุ่นยนต์ที่มีประสิทธิภาพสูงร่วมกับระบบ Tactile Sensor ที่สามารถทำงานร่วมกันได้อย่างเหมาะสม



บรรณานุกรม

- [1] Y. Matsuoka, "Embodiment and Manipulation Learning Process for a Humanoid Hand," Master of Science Thesis, Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Massachusetts, 1995
- [2] T. L. Baldi, S. Scheggi, L. Meli, M. Mohammadi and D. Prattichizzo, "GESTO: A Glove for Enhanced Sensing and Touching Based on Inertial and Magnetic Sensors for Hand Tracking and Cutaneous Feedback," *IEEE Transactions on Human-Machine Systems*, vol. 47, no. 6, pp. 1066-1076, Dec 2017. doi: 10.1109/THMS.2017.2720667.
- [3] S. J. Lederman, and R. L. Klatzky, "Haptic perception: A tutorial," *Attention, Perception, & Psychophysics*, vol. 71, No. 7, pp. 1439–1459, Oct 2009. doi:10.3758/APP.71.7.1439
- [4] A. Saudabayev, and H. A. Varol, "Sensors for Robotic Hands: A Survey of State of the Art," *IEEE Access*, vol. 3, pp. 1765-1782, 2015. doi:10.1109/ACCESS.2015.2482543
- [5] Z. Kappassov, J.A. Corrales, and V. Perdereau, "Tactile sensing in dexterous robot hands — Review," *Robotics and Autonomous Systems*, vol 74, pp. 195-220, Dec 2015. doi:10.1016/j.robot.2015.07.015
- [6] A. Schmitz, M. Maggiali, L. Natale, and G. Metta, "Touch sensors for humanoid hands," in *19th International Symposium in Robot and Human Interactive Communication*, 2010, pp. 691-697. doi:10.1109/ROMAN.2010.5598609
- [7] A. Schmitz, M. Maggiali, L. Natale, B. Bonino, and G. Metta, "A tactile sensor for the fingertips of the humanoid robot iCub," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2010, pp. 2212-2217. doi:10.1109/IROS.2010.5648838
- [8] M. Kaboli, R. Walker, and G. Cheng, "Re-using prior tactile experience by robotic hands to discriminate in-hand objects via texture properties," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 2242-2247. doi:10.1109/ICRA.2016.7487372

บรรณานุกรม (ต่อ)

- [9] D. Goger, N. Gorges, and H. Worn, "Tactile sensing for an anthropomorphic robotic hand: Hardware and signal processing," in *2009 IEEE International Conference on Robotics and Automation*, 2009, pp. 895-901.
doi:10.1109/ROBOT.2009.5152650
- [10] H. Liu, J. Greco, X. Song, J. Bimbo, L. Seneviratne, and K. Althoefer, "Tactile image based contact shape recognition using neural network," in *2012 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, 2012, pp. 138-143. doi:10.1109/MFI.2012.6343036
- [11] J. Schill, J. Laaksonen, M. Przybylski, V. Kyrki, T. Asfour, and R. Dillmann, "Learning continuous grasp stability for a humanoid robot hand based on tactile sensing," in *2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, 2012, pp. 1901-1906.
doi:10.1109/BioRob.2012.6290749
- [12] Z. Liao, W. Zhang, D. Kim, D. Im, K. Lim, and T. Miyoshi, "Recognition of the three-dimensional shape of objects grasped for PESA multi-fingered robot hand," in *2016 IEEE International Conference on Real-time Computing and Robotics (RCAR)*, 2016, pp. 472-476. doi:10.1109/RCAR.2016.7784075
- [13] J. M. Romano, K. Hsiao, G. Niemeyer, S. Chitta, and K. J. Kuchenbecker, "Human-Inspired Robotic Grasp Control With Tactile Sensing," *IEEE Transactions on Robotics*, vol. 27, No. 6, pp. 1067-1079, Dec. 2011.
doi:10.1109/TRO.2011.2162271
- [14] J. S. Son, R. Howe, J. Wang, and G. D. Hager, "Preliminary results on grasping with vision and touch," in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '96)*, 1996, pp. 1068-1075.
doi:10.1109/IROS.1996.568952.
- [15] R. Fernandez, A.S. Vazquez, I Payo, and A. Adan, "A Comparison of Tactile Sensors for In-Hand Object Location," *Journal of Sensors*, vol 2016 pp. 2943610. Jan 2016. doi:10.1155/2016/2943610

บรรณานุกรม (ต่อ)

- [16] J. M. Gandarias, A. J. García-Cerezo, and J. M. Gómez-de-Gabriel, "CNN-Based Methods for Object Recognition With High-Resolution Tactile Sensors," *IEEE Sensors Journal*, vol. 19, No. 16, pp. 6872-6882, Aug 2019. doi:10.1109/JSEN.2019.2912968
- [17] F. Pastor, J. M. Gandarias, A. J. García-Cerezo, and J. M. Gómez-de-Gabriel, "Using 3D Convolutional Neural Networks for Tactile Object Recognition with Robotic Palpation," *Sensors*, vol. 19, pp. 5356, Dec 2019. doi:10.3390/s19245356
- [18] S. Huang, and H. Wu, "Texture Recognition Based on Perception Data from a Bionic Tactile Sensor," *Sensors*, vol. 21, pp 5224, Aug 2021. doi:10.3390/s21155224
- [19] S. Luo, W. Mou, K. Althoefer, and H. Liu, "Novel Tactile-SIFT Descriptor for Object Shape Recognition," *IEEE Sensors Journal*, vol. 15, No. 9, pp. 5001-5009, Sept 2015. doi:10.1109/JSEN.2015.2432127
- [20] T. Matsubara, and K. Shibata, "Active tactile exploration with uncertainty and travel cost for fast shape estimation of unknown objects," *Robotics and Autonomous Systems*, vol. 91, pp. 314-326 Feb 2017. doi:10.1016/j.robot.2017.01.014
- [21] A. M. Almassri, W. Z. Wan Hasan, S. A. Ahmad, A. J. Ishak, A. M. Ghazali, D. N. Talib, and C. Wada, "Pressure Sensor: State of the Art, Design, and Application for Robotic Hand," *Journal of Sensors*, vol. 2015, pp. 846487 Dec 2014. doi:10.1155/2015/846487
- [22] Pressure Sensor Catalogue-PPS (*online*), 2021, Available: <https://pressureprofile.com/catalog> (17 October 2021).
- [23] A. Schmitz, P. Maiolino, M. Maggiali, L. Natale, G. Cannata, and G. Metta, "Methods and Technologies for the Implementation of Large-Scale Robot Tactile Sensors," *IEEE Transactions on Robotics*, vol. 27, No. 3, pp. 389-400, June 2011. doi:10.1109/TRO.2011.2132930

บรรณานุกรม (ต่อ)

- [24] T. D. Nguyena, H. S. Hana, H. Y. Shina, C. T. Nguyena, H. Phunga, H. V. Hoangb, and H. R. Choia, "Highly sensitive flexible proximity tactile array sensor by using carbon micro coils," *Sensors and Actuators A: Physical*, vol. 266, pp. 166–177, Sep 2017. doi:10.1016/j.sna.2017.09.013
- [25] Z. Ji, H. Zhu, H. Liu, N. Liu, T. Chen, Z. Yang, and L. Sun, "The Design and Characterization of a Flexible Tactile Sensing Array for Robot Skin," *Sensors*, vol. 16, pp. 2001, Nov 2016. doi:10.3390/s16122001
<https://doi.org/10.1016/j.nanoen.2018.05.061>
- [26] T. S. Ramadoss, Y. Ishii, A. Chinnappan, M. H. Ang, and S. Ramakrishna, "Fabrication of Pressure Sensor Using Electrospinning Method for Robotic Tactile Sensing Application," *Nanomaterials*, vol. 2021, No. 11, pp.1320. May 2021.
doi:10.3390/nano11051320
- [27] W. Ding, L. Lu, Y. Chen, J. Liu, and B. Yang, "Flexible P(VDF-TrFE) Shared Bottom Electrode Sensor Array Assisted with Machine Learning for Motion Detection," *Coatings*, vol. 2020, No. 10, pp. 1094 Nov 2020. doi:10.3390/coatings10111094
- [28] X. X. Zhu, Z. B. Li, X. S. Li, L. Su, X. Y. Wei, S. Y. Kuang, B. W. Su, J. Yang, Z. L. Wang, and G. Zhu, "Triboelectrification-enabled thin-film tactile matrix for self-powered high-resolution imaging." *Nano Energy*, vol. 50, pp. 497-503, Aug 2018. doi:10.1016/j.nanoen.2018.05.061
- [29] H. Xie, H. Liu, S. Luo, L. D. Seneviratne, and K. Althoefer, "Fiber optics tactile array probe for tissue palpation during minimally invasive surgery," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013, pp. 2539-2544. doi:10.1109/IROS.2013.6696714
- [30] A. Cirillo, P. Cirillo, G. De Maria, C. Natale, and S. Pirozzi, "A Distributed Tactile Sensor for Intuitive Human-Robot Interfacing," *Journal of Sensors*, vol. 2017, pp. 1357061, April 2017. doi:10.1155/2017/1357061

บรรณานุกรม (ต่อ)

- [31] J. Konstantinova, A. Stilli, and K. Althoefer, "Fingertip Fiber Optical Tactile Array with Two-Level Spring Structure," *Sensors*, vol. 17, pp. 2337. doi:10.3390/s17102337
- [32] K. Weiss, and H. Worn, "The working principle of resistive tactile sensor cells," in *IEEE International Conference Mechatronics and Automation 2005*, 2005, pp. 471-476. doi:10.1109/ICMA.2005.1626593
- [33] F. Pizarro, P. Villavicencio, D. Yunge, M. Rodríguez, G. Hermosilla, and A. Leiva, "Easy-to-Build Textile Pressure Sensor," *Sensors*, vol. 18, pp. 1190, April 2018. doi:10.3390/s18041190
- [34] H. Nakamoto, F. Kobayashi, I. Nobuaki, S. Hidenori, and F. Kojima, "Universal robot hand equipped with tactile and joint torque sensors development and experiments on stiffness and object recognition," *Sytemics, Cybernetics and Informatic*, vol. 5, pp. 79-84, 2007.
- [35] Y. W. R. Amarasinghe, A. L. Kulasekera, and T. G. P. Priyadarshana, "Quantum Tunneling Composite (QTC) based tactile sensor array for dynamic pressure distribution measurement," in *2013 Seventh International Conference on Sensing Technology (ICST)*, 2013, pp. 1-4. doi:10.1109/ICSensT.2013.6727794
- [36] S. Ando, and H. Shinoda, "Ultrasonic emission tactile sensing," *IEEE Control Systems Magazine*, vol. 15, No. 1, pp. 61-69, Feb 1995. doi:10.1109/37.341866
- [37] C. S. Chang, and Y. C. Lee, "Ultrasonic Touch Sensing System Based on Lamb Waves and Convolutional Neural Network," *Sensors*, vol. 20, pp. 2619, May 2020. doi:10.3390/s20092619
- [38] H. Wang, G. de Boer, J. Kow, A. Alazmani, M. Ghajari, R. Hewson, and P. Culmer, "Design Methodology for Magnetic Field-Based Soft Tri-Axis Tactile Sensors," *Sensors*, vol. 16, pp. 1356, Aug2016. doi:10.3390/s16091356

บรรณานุกรม (ต่อ)

- [39] H. Nakamoto, T. Wakabayashi, F. Kobayashi, and F. Kojima, "Estimation of Displacement and Rotation by Magnetic Tactile Sensor Using Stepwise Regression Analysis," *Journal of Sensors*, vol. 2014, pp. 459059, Oct 2014. doi:10.1155/2014/459059
- [40] C. Sferrazza, and R. D'Andrea, "Design, Motivation and Evaluation of a Full-Resolution Optical Tactile Sensor," *Sensors*, vol. 19, pp. 928, Feb 2019. doi:10.3390/s19040928
- [41] W. Yuan, S. Dong, and E. H. Adelson, "GelSight: High-Resolution Robot Tactile Sensors for Estimating Geometry and Force," *Sensors*, vol. 17, pp. 2762, Nov 2017. doi:10.3390/s17122762
- [42] H. Jiang, Y. Yan, X. Zhu, and C. Zhang, "A 3-D Surface Reconstruction with Shadow Processing for Optical Tactile Sensors," *Sensors*, vol. 18, pp. 2785, Aug 2018. doi:10.3390/s18092785
- [43] C. A. Gutierrez, and E. Meng, "A dual function Parylene-based biomimetic tactile sensor and actuator for next generation mechanically responsive microelectrode arrays," in *Transducers 2009 - 2009 International Solid-State Sensors, Actuators and Microsystems Conference*, 2009, pp. 2194-2197. doi:10.1109/SENSOR.2009.5285612
- [44] H. Shinoda, and H. Oasa, "Wireless tactile sensing element using stress-sensitive resonator," *IEEE/ASME Transactions on Mechatronics*, vol. 5, No. 3, pp. 258-265, Sept 2000. doi:10.1109/3516.868917
- [45] M. M. Iskarous, and N. V. Thakor, "E-Skins: Biomimetic Sensing and Encoding for Upper Limb Prostheses," *Proceedings of the IEEE*, vol. 107, No. 10, pp. 2052-2064, Oct 2019. doi:10.1109/JPROC.2019.2939369
- [46] Y. Al-Handarish, O. M. Omisore, T. Igbe, S. Han, H. Li, W. Du, J. Zhang, and L. Wang, "A Survey of Tactile-Sensing Systems and Their Applications in Biomedical Engineering," *Advances in Materials Science and Engineering*, vol. 2020, pp. 4047937, Jan 2020. doi:10.1155/2020/4047937

บรรณานุกรม (ต่อ)

- [47] G. Palli, and S. Pirozzi, "A Tactile-Based Wire Manipulation System for Manufacturing Applications," *Robotics*, vol. 8, pp. 46, June 2019. doi:10.3390/robotics8020046
- [48] Z. You, and Y. Chen, "The Use of Tactile Sensors and PIV Analysis for Understanding the Bearing Mechanism of Pile Groups," *Sensors*, vol. 18, pp. 476, Feb 2018. doi:10.3390/s18020476
- [49] D. Giovanelli, and E. Farella, "Force Sensing Resistor and Evaluation of Technology for Wearable Body Pressure Sensing," *Journal of Sensors*, vol. 2016, pp. 9391850, Jan 2016. doi:10.1155/2016/9391850
- [50] L. Zou, C. Ge, J. Wang, E. Cretu, and X. Li, "Novel Tactile Sensor Technology and Smart Tactile Sensing Systems: A Review," *Sensors*, vol. 17, pp. 2653, Nov 2017. doi:10.3390/s17112653
- [51] W. Y. Chang, T. H. Fang, S. H. Yeh, and Y. C. Lin, "Flexible Electronics Sensors for Tactile Multi-Touching," *Sensors*, vol. 9, pp. 1188-1203, Feb 2009. doi:10.3390/s9021188
- [52] D. J. van den Heever, K. Schreve, and C. Scheffer, "Tactile Sensing Using Force Sensing Resistors and a Super-Resolution Algorithm," *IEEE Sensors Journal*, vol. 9, No. 1, pp. 29-35, Jan 2009. doi:10.1109/JSEN.2008.2008891
- [53] J. Castellanos-Ramosa, R. Navas-González, H. Maciciorb, T. Sikorab, E. Ochotecob, and F. Vidal-Verdúa, "Tactile Sensors Based on Conductive Polymers," *Microsystem Technologies*, vol. 16, No. 5, pp. 765-776, May 2010. doi:10.1007/s00542-009-0958-3
- [54] A. Drimus, G. Kootstra, A. Bilberg, and D. Kragic, "Classification of rigid and deformable objects using a novel tactile sensor," in *2011 15th International Conference on Advanced Robotics (ICAR)*, 2011, pp. 427-434. doi:10.1109/ICAR.2011.6088622.
- [55] H. Wang, D. Zhou, and J. Cao, "Development of a Skin-Like Tactile Sensor Array for Curved Surface," *IEEE Sensors Journal*, vol. 14, No. 1, pp. 55-61, Jan. 2014, doi:10.1109/JSEN.2013.2279394.

บรรณานุกรม (ต่อ)

- [56] S. Khan, S. Tinku, L. Lorenzelli, and R. S. Dahiya, "Flexible Tactile Sensors Using Screen-Printed P(VDF-TrFE) and MWCNT/PDMS Composites," *IEEE Sensors Journal*, vol. 15, no. 6, pp. 3146-3155, June 2015. doi:10.1109/JSEN.2014.2368989.
- [57] Y. H. Liu, Y. T. Hsiao, W. T. Cheng, Y. C. Liu, and J. Y. Su, "Low-Resolution Tactile Image Recognition for Automated Robotic Assembly Using Kernel PCA-Based Feature Fusion and Multiple Kernel Learning-Based Support Vector Machine," *Mathematical Problems in Engineering*, vol. 2014, pp. 497275, Feb 2014. doi:10.1155/2014/497275
- [58] F. Wang, Y. Song, Z. Zhang, and W. Chen, "Structure Analysis and Decoupling Research of a Novel Flexible Tactile Sensor Array," *Journal of Sensors*, vol. 2015, pp. 476403, May 2015. doi:10.1155/2015/476403
- [59] Y. C. Tsai, C. W. Ma, Y. H. Lin, and Y. J. Yang, "Development of a Large-Area 8 × 8 Tactile Sensing Array with High Sensitivity," *Sensors and Materials*, vol. 29, No. 3, pp. 303–309, Feb 2017. doi:10.18494/SAM.2017.1456
- [60] C. Gerlach, A. Sanli, R. Ramalingame, and O. Kanoun, "Flexible, dynamic piezoresistive sensor matrix based on carbon nanotube polymer composite for pressure distribution measurement," *Journal of Sensors and Sensor Systems*, vol. 8, No. 1, pp. 1-7, Jan 2019. doi:10.5194/jsss-8-1-2019
- [61] R. Ramalingame, Z. Hu, C. Gerlach, D. Rajendran, T. Zubkova, R. Baumann, and O. Kanoun, "Flexible piezoresistive sensor matrix based on a carbon nanotube PDMS composite for dynamic pressure distribution measurement," *Journal of Sensors and Sensor Systems*, vol. 8, pp. 1–7, Jan 2019 doi:10.5194/jsss-8-1-2019
- [62] S. Luo "Object Perception Through Tactile Images," Doctor of Philosophy Thesis, Department of Informatics, King's College, London, 2016.

บรรณานุกรม (ต่อ)

- [63] A. Schneider, J. Sturm, C. Stachniss, M. Reisert, H. Burkhardt, and W. Burgard, "Object identification with tactile sensors using bag-of-features," in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, St. Louis, MO, USA, 2009, pp. 243-248. doi:10.1109/IROS.2009.5354648
- [64] N. Gorges, S. Escalda Navarro, D. Göger, and H. Wörn, "Haptic object recognition using passive joints and haptic key features," in *2010 IEEE International Conference on Robotics and Automation*, Anchorage, AK, USA, 2010, pp. 2349-2355. doi:10.1109/ROBOT.2010.5509553
- [65] A. Drimus, G. Kootstra, A. Bilberg and D. Kragic, "Classification of rigid and deformable objects using a novel tactile sensor," in *2011 15th International Conference on Advanced Robotics (ICAR)*, Tallinn, Estonia, 2011, pp. 427-434. doi:10.1109/ICAR.2011.6088622
- [66] Z. Pezzementi, E. Plaku, C. Reyda and G. D. Hager, "Tactile-Object Recognition From Appearance Information," *IEEE Transactions on Robotics*, vol. 27, No. 3, pp. 473-487, June 2011. doi:10.1109/TRO.2011.2125350.
- [67] T. Bhattacharjee, J. M. Rehg, and C. C. Kemp, "Haptic classification and recognition of objects using a tactile sensing forearm," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vilamoura, Algarve, Portugal, 2012, pp. 4090-4097. doi:10.1109/IROS.2012.6386142
- [68] S. Dattaa, A. Khasnobishb, A. Konara, D.N. Tibarewalab, and R.Janarthananc, "Performance Analysis of Object Shape Classification and Matching from Tactile Images using Wavelet Energy Features," *Procedia Technology*, vol. 10 pp. 805–812, Dec 2013, doi:10.1016/j.protcy.2013.12.425
- [69] A. M. Cretu, T. E. A. de Oliveira, V. Prado da Fonseca, B. Tawbe, E. M. Petriu, and V. Z. Groza, "Computational intelligence and mechatronics solutions for robotic tactile object recognition," in *2015 IEEE 9th International Symposium on Intelligent Signal Processing (WISP) Proceedings*, Siena, Italy, 2015, pp. 1-6. doi:10.1109/WISP.2015.7139165.

บรรณานุกรม (ต่อ)

- [70] A. Schmitz, Y. Bansho, K. Noda, H. Iwata, T. Ogata, and S. Sugano, "Tactile object recognition using deep learning and dropout," in *2014 IEEE-RAS International Conference on Humanoid Robots*, Madrid, Spain, 2014, pp. 1044-1050.
doi:10.1109/HUMANOIDS.2014.7041493
- [71] H. Liu, D. Guo, and F. Sun, "Object Recognition Using Tactile Measurements: Kernel Sparse Coding Methods," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, No. 3, pp. 656-665, March 2016,
doi:10.1109/TIM.2016.2514779
- [72] S. Luo, W. Mou, K. Althoefer and H. Liu, "Iterative Closest Labeled Point for tactile object shape recognition," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Daejeon, Korea, 2016, pp. 3137-3142.
doi:10.1109/IROS.2016.7759485
- [73] P. Falco, S. Lu, A. Cirillo, C. Natale, S. Pirozzi, and D. Lee, "Cross-modal visuo-tactile object recognition using robotic active exploration," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, Singapore, 2017, pp. 5273-5280. doi:10.1109/ICRA.2017.7989619
- [74] A. Albin, S. Denei, and G. Cannata, "Human hand recognition from robotic skin measurements in human-robot physical interactions," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Vancouver, BC, Canada, 2017, pp. 4348-4353. doi:10.1109/IROS.2017.8206300
- [75] J. M. Gandarias, J. M. Gómez-de-Gabriel, and A. García-Cerezo, "Human and object recognition with a high-resolution tactile sensor," in *2017 IEEE SENSORS*, Glasgow, UK, 2017, pp. 1-3. doi:10.1109/ICSENS.2017.8234203
- [76] J. M. Gandarias, A. J. García-Cerezo, and J. M. Gómez-de-Gabriel, "CNN-Based Methods for Object Recognition With High-Resolution Tactile Sensors," *IEEE Sensors Journal*, vol. 19, No. 16, pp. 6872-6882, Aug 2019.
doi:10.1109/JSEN.2019.2912968

บรรณานุกรม (ต่อ)

- [77] Y. Le Cun, B. Boser, J. S. Denker, R. E. Howard, W. Hubbard, L. D. Jackel, and D. Henderson, "Handwritten digit recognition with a back-propagation network," in *Advances in neural information processing systems*, Morgan, San Francisco, CA, United States, Kaufmann Publishers Inc., 1990. doi:10.5555/109230.109279
- [78] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, No. 11, pp. 2278-2324, Nov 1998. doi:10.1109/5.726791
- [79] I. Tabian, H. Fu, and Z. S. Khodaei, "A Convolutional Neural Network for Impact Detection and Characterization of Complex Composite Structures," *Sensors*, vol. 19, pp. 4933, Nov 2019. doi:10.3390/s19224933
- [80] Softmax Activation (*online*), 2021, Available: <http://rinterested.github.io/statistics/softmax.html> (13 January 2022).
- [81] M. K. Hu, "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory*, vol. 8, No. 2, pp. 179-187, Feb 1962. doi:10.1109/TIT.1962.1057692
- [82] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, vol. 60, pp. 91-110, Jan 2004. doi:10.1023/B:VISI.0000029664.99615.94
- [83] S. Kanai, "Content-based 3D mesh model retrieval from hand-written sketch," *International Journal on Interactive Design and Manufacturing*, vol. 2, No. 2, pp. 87-98, May 2008, doi:10.1007/s12008-008-0038-4
- [84] ปริญญา สงวนสัตย์, *Artificial Intelligence with Machine Learning*, พิมพ์ครั้งที่ 1, 2562, นนทบุรี : ไอดีซี พรีเมียร์, 2562.
- [85] C. Cortes, and V. Vapnik, "Support-Vector Networks," *Machine Learning*, vol. 20, pp. 273-297, 1995.
- [86] C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford : Clarendon Press, 1995.

บรรณานุกรม (ต่อ)

- [87] สุรินทร์ คำฝอย, *โครงข่ายประสาทเทียมเชิงปฏิบัติสำหรับงานวิศวกรรมไฟฟ้า*, พิมพ์ครั้งที่ 1, 2554, กรุงเทพฯ : มินิ เซอร์วิส ซัพพลาย, 2554.
- [88] A.Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems*, Lake Tahoe, NV, USA, Dec 2012, pp. 1097–1105. doi:10.1145/3065386
- [89] MS9723 FSR Matrix Array Sensor (16x10 Rows and Columns / 80mm x 50mm Active Sensing Area) (*online*), 2021, https://www.kitronyx.com/store/p30/%5BMS9723%5D_FSR_Matrix_Array_Sensor_%2816x10_Rows_and_Columns_%2F_80mm_x_50mm_Active_Sensing_Area%29.html (13 January 2022).
- [90] K. Simonyan, and A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition. *In conference paper at ICLR 2015*, Sep 2014.
- [91] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, June 2015. doi:10.1109/CVPR.2015.7298594
- [92] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, June 2016. doi:10.1109/CVPR.2016.90
- [93] B. Zhou, A. Khosla, A. Lapedriza, A. Torralba, and A. Oliva, "Places: An Image Database for Deep Scene Understanding," *J. Vis*, vol. 17, 2016.
- [94] C. Szegedy, V. Vanhoucke, S. Ioffe, Z. Wojna, and J. Shlens, "Rethinking the Inception Architecture for Computer Vision," in *Proceedings of the Computer Vision and Pattern Recognition 2016*, Las Vegas, NV, USA, June 2016. doi:10.1167/17.10.296

บรรณานุกรม (ต่อ)

- [95] M. Tan, and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in *Proceedings of the 36th International Conference on Machine Learning*, Long Beach, CA, USA, June 2019.
- [96] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size," in *ICLR conference 2017*, 2016,
- [97] C. Szegedy, I. Sergey, V. Vanhoucke, and A. Alemi, "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning," in *Proceedings of the AAAI Conference on Artificial Intelligence*, Phoenix, AZ, USA, Feb 2016.
- [98] G. Huang, H. Liu, M. van der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, July 2017. doi:10.1109/CVPR.2017.243
- [99] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, June 2016. doi:10.1109/CVPR.2016.91
- [100] J. Redmon, and A. Farhadi, "YOLO9000: Better, Faster, Stronger," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, July 2017. doi:10.1109/CVPR.2017.690
- [101] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, July 2017. doi:10.1109/CVPR.2017.195
- [102] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning Transferable Architectures for Scalable Image Recognition," in *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, June 2018. doi:10.1109/CVPR.2018.00907

บรรณานุกรม (ต่อ)

- [103] N. Ma, X. Zhang, H. Zheng, and J. Sun, "ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design," in *Proceedings of the European Conference on Computer Vision ECCV 2018*, Munich, Germany, Sep 2018. doi:10.1007/978-3-030-01264-9_8
- [104] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Lake City, UT, USA, June 2018. doi:10.1109/CVPR.2018.00474
- [105] S. J. Pan, and Q. A. Yang, "Survey on Transfer Learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, pp. 1345–1359, 2010. doi:10.1109/TKDE.2009.191
- [106] A. S. Rao, T. Nguyen, M. Palaniswami, and T. Ngo, "Vision-based automated crack detection using convolutional neural networks for condition assessment of infrastructure," *Structural Health Monitoring*, vol. 20, pp. 2124–2142, 2010. doi:10.1177/1475921720965445
- [107] จักรินทร์ สนุกแสน, และ โอลาริก สุรินตะ, "โครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึกสำหรับการจำแนกพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติ," *วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยมหาสารคาม*, vol. 38, pp. 113-124, 2019.
- [108] X. Ou, P. Yan, Y. Zhang, B. Tu, G. Zhang, J. Wu, and W. Li, "Moving Object Detection Method via ResNet-18 With Encoder–Decoder Structure in Complex Scenes," *IEEE Access*, vol. 7, pp. 108152-108160, 2019. doi:10.1109/ACCESS.2019.2931922
- [109] T. A. Putra, S. I. Rufaida and J. S. Leu, "Enhanced Skin Condition Prediction Through Machine Learning Using Dynamic Training and Testing Augmentation," *IEEE Access*, vol. 8, pp. 40536-40546, 2020. doi: 10.1109/ACCESS.2020.2976045
- [110] C. Chousangsuntorn, T. Tongloy, S. Chuwongin, and S. Boonsang, "A Deep Learning System for Recognizing and Recovering Contaminated Slider Serial Numbers in Hard Disk Manufacturing Processes," *Sensors*, Vol. 21, pp. 6261, 2021. doi:10.3390/s21186261

บรรณานุกรม (ต่อ)

[111] Darknet53 (*online*), 2021, Available:

<https://github.com/developer0hye/PyTorch-Darknet53> (13 January 2022).

[112] U. Seidaliyeva, D. Akhmetov, L. Ilipbayeva, and E. T. Matson, "Real-Time and

Accurate Drone Detection in a Video with a Static Background," *Sensors*, Vol.

20, pp. 3856, 2020. doi:10.3390/s20143856





ภาคผนวก ก

วารสารทางวิชาการและบทความการประชุมวิชาการ

- [1] S. Pohtongkam, and J. Srinonchat, "Object Recognition from Human Tactile Image Using Artificial Neural Network," in *2016 13th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, Chiang Mai, Thailand, June 28 2016–July 1 2016, pp. 1–6. doi:10.1109/ECTICon.2016.7561402
- [2] สมชาย เปาะทองคำ, และ จักรี ศรีนนท์ฉัตร, "การจดจำวัตถุจากการสัมผัสของหุ่นยนต์เสมือนมนุษย์ด้วยวิธีการประมวลผลภาพจากการสัมผัสร่วมกับโครงข่ายประสาทเทียม," *วิศวกรรมสาร มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี*, vol. 13, No. 2, pp. 66–79, Dec 2018.
- [3] S. Pohtongkam, and J. Srinonchat, "Tactile Object Recognition for Humanoid Robots Using New Designed Piezoresistive Tactile Sensor and DCNN," *Sensors*, vol. 21, No. 18, pp. 1–26, Sep 2021. doi:10.3390/s21186024
- [4] S. Pohtongkam, and J. Srinonchat, "Object Recognition Using Glove Tactile Sensor," in *2022 International Electrical Engineering Congress (IEECON2022)*, Khon Kaen, Thailand, March 9 2022–March 11 2022, pp. 1–4. doi:10.1109/IEECON53204.2022.9741672

Object Recognition from Human Tactile Image Using Artificial Neural Network

Somchai Pohtongkam

Signal Processing Research Laboratory, Faculty of
Engineering, Rajamangala University of Technology
Thanyaburi, Pathumthani, Thailand
somchai_po@mail.rmutt.ac.th

Jakkree Srinonchat

Signal Processing Research Laboratory, Faculty of
Engineering, Rajamangala University of Technology
Thanyaburi, Pathumthani, Thailand
jakkree.s@en.rmutt.ac.th

Abstract— This paper presents a recognition of images objects that are out of touch by capturing the texture of objects that existed in everyday life which sub-divided according to the different parts of the human hand palm. They are divided into sections based on the physiology of the human hand by dividing the exposure of 15,20 and 26 proportions and then analyzed. This work is the basis design of human-like hand for a robot to work effectively. These pictures of touch present the genuine sensory system of a human hand. They will be processed to characterize the touched object as the different area for the exposed surface of the object will create different for the palm's pressure. After that, we extract the features of the pressure from pressing the object and use an artificial neural network to distinguish different types of objects. The results of tests on 15 types of objects show the accurate result of the analysis of at the average maximum of 78.38% with the 26 proportions palm. This research will form the basis design of human-like hand for a robot that can recognize objects robot caught.

Keywords—Human Palm, Robot hand, Hand Segmentation, Tactile Image, ANN

I. INTRODUCTION

Humanoid Robot has been constantly developed to be able to move, walk and have appearance like a human. Moreover, it can see and recognize faces and objects and also be able to handle objects [1] which are the research that has been constantly evolving. This article aims to present a study of mechanism of the process for learn and remember the human touch to apply for designing of the robot arm as a human. In related research depicted, it has presented the memorable objects of touch, which concluded that there are two learning process is 1. The memorization of the shape and angel of the fingers that hold object 2. From the images caused by pressure on the object against censorship In 2003 The authors presented a method to design a sensitive skin for robot [2]. The Topic is "Tactile Perception in Robots: From the Somatic Alphabet to the Realization of a Fully 'Sensitive Skin' ". It is an artificial robot palm with the capability of haptic perception has been designed. In [3] the authors are proposed a new technique for object classification of robot hand and finger with distributed touch sensor. In 2011 [4][5] the authors used self-organizing map is proposed for object recognition based on tactile form

perception. A robot hand with three fingers, with the same number of degrees of freedom as the human hand but in this paper presented the tactile image. In 2012 [6] the authors present the technique of shape recognition from tactile images. they can classify 4 type of object shape. [7] The authors are proposed a method to recognize an objects from human tactile image. They can recognize 4 object with the mean classification accuracy of classifier is 88%. 2013 [8] the authors are proposed a method to recognize an objects from tactile image. They used image from tactile sensor pad and four different statistic classification method classifiers are kNN, Naïve Bayes classifier, LDA and Ensemble.

This work has studied the working mechanism of the human hand as a basis for the design of human-like robot hand that can work as close as possible to human physiological studies of human hands with the collaboration of the skin and the bone that holds your hand to the skin together. It can improve the classification accuracy of the robot palm hand was design like a human palm.

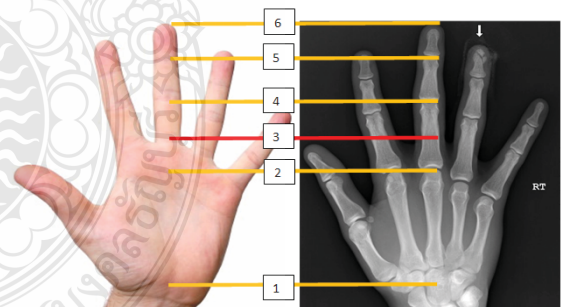


Figure 1. The structure of a human palm.

Figure 1 is a segmentation of the hand which divided into two parts: the palm and the fingers which having the palm's bone linked the tissue by a section of the palm to hold the finger bone sections with about half of the long bones (from point 1 to point. 3) As for the thumb to the bone, there are three pieces of bones embedded in the palm one-piece and two-piece on the stroke. As for the rest fingers, there are four pieces of bones. The first piece is embedded in palm. The second piece is half-embedded in the palm and another half is the finger. The rest are the portion of the fingertip. The seizure by the skin and

การจดจำวัตถุจากการสัมผัสของหุ่นยนต์เสมือนมนุษย์ด้วยวิธีการประมวลผลภาพจาก
การสัมผัสร่วมกับโครงข่ายประสาทเทียม
Object Recognition for Humanoid Robot using Tactile image processing and
Artificial Neural Network

สมชาย เปาะทองคำ, จักรี ศรีนนท์ฉัตร*
S. Pohtongkam, J. Srinonchat

ห้องปฏิบัติการและวิจัยทางการประมวลผลสัญญาณ ภาควิชาวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม
คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
Signal Processing Research Laboratory Department of Electronics and Telecommunication Engineering
Faculty of Engineering Rajamangala University of Technology Thanyaburi
*Corresponding author E-mail : jakkree.s@en.rmutt.ac.th

(Received: September 30, 2017, Revised: April 17, 2018, Accepted: July 2, 2018)

บทคัดย่อ

บทความนี้นำเสนอการพัฒนาาระบบสำหรับจดจำวัตถุจากการสัมผัสบริเวณพื้นผิววัตถุโดยใช้ประสาทสัมผัสทางกายเทียม ทั้งนี้ งานวิจัยนี้ได้ออกแบบระบบการทดสอบเพื่อเป็นพื้นฐานของมือหุ่นยนต์ที่ติดตั้งระบบเซ็นเซอร์วัดแรงกดแบบอาร์เรย์ ระบบได้รับการออกแบบโดยใช้เซ็นเซอร์สัมผัสขนาด 16x10 พิกเซล เมื่อระบบสัมผัสวัตถุจะทำการจัดเรียงข้อมูลให้เป็นรูปภาพ ภาพที่ได้จะถูกนำมาประมวลผลเพื่ออธิบายลักษณะของวัตถุที่สัมผัสที่แตกต่างกันสำหรับผิวสัมผัสของวัตถุจะสร้างความแตกต่างของแรงกดให้มือหุ่นยนต์ หลังจากนั้นได้จะแยกลักษณะของวัตถุโดยใช้โครงข่ายประสาทเทียม ผลการทดสอบวัตถุ 10 ชนิดแสดงผลการวิเคราะห์ที่ถูกต้องโดยเฉลี่ยสูงสุด 93.2% โดยใช้การหาค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของแรงกดที่กระทำบนเซ็นเซอร์

คำสำคัญ: เซ็นเซอร์วัดแรงกด, ระบบประสาทสัมผัสทางกาย, ภาพจากการสัมผัส, โครงข่ายประสาทเทียม

Abstract

This paper presents an improvement recognition objects system, in the term of touching surface of objects, by using Artificial Neural Networks (ANNs). The experiment of this article is implemented and designed base on the robot hand for installs a sensor array on the robot hand. The sensor array is designed as 16X10 pixels. When the robot hand touches the object, the data is then organized into the image form for processing. This is to identify the characterize of object surface which different objects will provides a different images. This also depends on the force to touch objects. Then the Artificial Neural Networks technique is used to classify the objects. This system is tested with 10 different objects. The experiment results shown that it provides the accuracy approximately 93.2% based on average and standard deviation values.

Keywords: Tactile Sensor, Nervous System, Tactile image, Artificial Neural Networks

Article

Tactile Object Recognition for Humanoid Robots Using New Designed Piezoresistive Tactile Sensor and DCNN

Somchai Pohtongkam and Jakkree Srinonchat * 

Department of Electronics and Telecommunication Engineering, Rajamangala University of Technology Thanyaburi, Khlong Luang 12110, Thailand; somchai_po@mail.rmutt.ac.th

* Correspondence: jakkree.s@en.rmutt.ac.th; Tel.: +66-897-775-038

Abstract: A tactile sensor array is a crucial component for applying physical sensors to a humanoid robot. This work focused on developing a palm-size tactile sensor array (56.0 mm × 56.0 mm) to apply object recognition for the humanoid robot hand. This sensor was based on a PCB technology operating with the piezoresistive principle. A conductive polymer composites sheet was used as a sensing element and the matrix array of this sensor was 16 × 16 pixels. The sensitivity of this sensor was evaluated and the sensor was installed on the robot hand. The tactile images, with resolution enhancement using bicubic interpolation obtained from 20 classes, were used to train and test 19 different DCNNs. InceptionResNetV2 provided superior performance with 91.82% accuracy. However, using the multimodal learning method that included InceptionResNetV2 and XceptionNet, the highest recognition rate of 92.73% was achieved. Moreover, this recognition rate improved when the object exploration was applied to demonstrate.

Keywords: tactile sensor; tactile object recognition; DCNN; humanoid robot; transfer learning



Citation: Pohtongkam, S.; Srinonchat, J. Tactile Object Recognition for Humanoid Robots Using New Designed Piezoresistive Tactile Sensor and DCNN. *Sensors* **2021**, *21*, 6024. <https://doi.org/10.3390/s21186024>

Academic Editor: Nunzio Cennamo

Received: 4 August 2021

Accepted: 6 September 2021

Published: 8 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Unlike humans who can identify objects by touching, humanoid robots do not have this capability due to the lack of suitable tactile sensors and efficient recognition processing systems. The critical development of humanoid robot technology can be divided into two parts: (1) robot anatomy [1]; (2) the robot nervous system [2]. Developing a physical structure and human-like learning ability is necessary to enable robots to operate in a home or office environment. In addition, the development of robots having a human-like hand structure is desirable [3–5]. This study examines a humanoid robot's physical sensory system that can recognize objects by touch. Its essential function is developed based on the human physical sensory system [6]. In object learning and recognition systems of humanoid robots employed artificial haptic perception [7–11], pressure sensors or tactile sensors are utilized [7–11], and the obtained information is sent to a computer for analysis [10]. Object learning and recognition systems are similar to the human sensory system where nerve-ending receptors (e.g., Ruffini endings and Pacinian receptors) obtain information sent to the brain for interpretation. There have been numerous studies describing the development of robotic hands. These studies focus on tactile sensor arrays for robot hand artificial skin application [7–11]. Human sensory recognition is a complicated action resulting from the biosensor system in the body, which includes three modes of data perception [6]. The first mode is tactile perception where contact with the skin of the fingers or palm provides information on the contact geometry or pressure profile.

A tactile sensor array produces this mode of data perception for robots and presents data in a 2D format or tactile image [10]. The second perception mode is kinesthetic perception, a perception from motion such as rubbing or scrubbing objects. For robots, this mode of data perception is produced by tactile sensors on the fingertips or palm from dragging the sensor onto the object and presents data in a 1D format [12]. The third perception mode is global object shape, where perception data is gathered through the

Object Recognition Using Glove Tactile Sensor

Somchai Pohtongkam
department of Electronics and Telecommunication
Rajamangala University of Technology Thanyaburi
Pathum Thani, Thailand
somchai_po@mail.rmutt.ac.th

Jakkree Srinonchat*
department of Electronics and Telecommunication
Rajamangala University of Technology Thanyaburi
Pathum Thani, Thailand
jakkree_s@rmutt.ac.th

Abstract— This article presents the artificial intelligence for object recognition by touching method with the tactile sensor, which has similar structural characteristics as the human hand. The object recognition uses a tactile glove sensor with small tactile sensors to spread throughout the glove, and there are touching points similar to human hands. Using the glove sensor to grasp the object that will obtain an image, there is tactile image. These images were used to compare recognition performance with two techniques: the Bag of Word (BoW) and the Convolutional Neural Network (CNN). In the experiment, the tactile glove performs 20 different objects. The BoW technique, the Scale Invariant Feature Transform (SIFT), has been used to feature extraction and then classified with K Nearest Neighbors (KNN) to evaluate the performance. The results illustrate that the accuracy of BoW and CNN is 70.80% and 97.20%, respectively. When comparing both techniques, it was found that the CNN had higher accuracy of about 26.40%. The results clearly show that CNN was more performance than BoW. Therefore, it is suitable to analyze tactile glove recognition, which can be applied to the recognition system of humanoid robots.

Keywords—(tactile object recognition, tactile image, glove tactile sensor, BoW, CNN, K-mean clustering, K-NN)

I. INTRODUCTION

Humans can recognize objects by touching them, which is the brain's interpretation, and humans need this ability to live and work. However, this complex operation requires two capacities: First, data retrieval from objects through touch. Second, intellectual ability to process this information [1]. Human-like robotics has been developed in appearance and artificial intelligence (AI), so these robots look more similar to humans. For hardware development, the structure of the robot hands can be similar to the human hand [2] but lack only the artificial sensory system, which is inappropriate robotic hand characteristics. Even tactile sensor has been applied. However, it does not cover whole areas of the hands [3]. The recognition analysis system is still being developed to fit only the part of the tactile sensor [4].

In literature reviews have continued to develop sensors for the humanoid robot hands. Most of them focus on mounting some parts of the robot hands, such as fingertip [5], finger [6], palm [7], fingertip, and palm [8]. These sensor installations were very different when compared to the fundamental human hands. However, there is research to install sensors throughout the robot's hand [9], but there are not all five fingers like humans, or the resolutions of the sensors are too low [10]. The development of sensors for the humanoid robotic hands equipped with a palm-full hand and the five fingers sensor was proposed by Liu et al. [11], but data analysis was ineffective.

Researchers developed algorithms for tactile sensors that can be divided into two groups [12]. First, they used small tactile sensors to test the objects. If the sensors are much

smaller than the objects, we call the tactile image a local shape tactile image. The recognition method is local shape recognition, which employs the Bag of Word Model (BoW) [13] to analyze the segments. On the other hand, Second, they applied tactile sensors whose sizes are large or similar to the objects. It calls a tactile image a global shape tactile image. This recognition is recognized as Global shape recognition, which is operated by Machine Learning (ML) [14] and Deep Learning (DL) for analysis [15].

In this article, a tactile glove sensor is presented. In order to retrieve information from the objects, these data have touching points like human hands to recognize the objects as human hands. The information is the local shape tactile image, which has several images joined together until it looks like a discrete Global shape tactile image. Therefore, we can analyze by employing both local shape recognition and global shape recognition techniques. Both techniques will be applied and compared to compare the performance to optimize the other suitable method.

II. PROPOSED METHOD

A. Glove Tactile Sensor

The tactile glove sensor used in this work was designed according to the shape of the hand palm and then installed on the robot hand. In the palm of the hands as Table 1, sensors are 16×16 pixels. The robotic hands suitable for this sensor should be a flat palm and have five fingers, four finger must be divided into three parts according to the knuckles, and thump finger must be divided into 2 parts. For the four fingers except for the thumps, the sensors were 4×6 pixels, 4×4 pixels, and 4×5 pixels, respectively. The sensors of the thumps were 5×6 pixels and 5×5 pixels as shown in Fig. 1.

TABLE I. TACTILE SENSOR FOR HUMANOID ROBOT HAND SENSORS

Sensor	Number of Set	Resolution (Pixels)	Size (mm)	Total Pixel (Pixel)
Palm	1	16×16	56.0×56.0	256
Thump	1	5×6 , 5×5	17.5×21.0 17.5×17.5	55
Finger	4	4×6 4×4 4×5	14.0×21.0 14.0×14.0 14.0×17.5	240

ประวัติผู้เขียน

ชื่อ-สกุล	นายสมชาย เปาะทองคำ
วัน เดือน ปีเกิด	วันที่ 12 เดือน มีนาคม พ.ศ. 2523
ที่อยู่	432/360 หมู่ที่ 4 ตำบลหลักหก อำเภอเมืองปทุมธานี จังหวัดปทุมธานี 12000
การศึกษา	- ปริญญาตรี วศ.บ.(วิศวกรรมไฟฟ้า - โทรคมนาคม) คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีราชมงคล - ปริญญาตรี สบ. อาชีวอนามัยและความปลอดภัย สาขาวิชาวิทยาศาสตร์สุขภาพ มหาวิทยาลัยสุโขทัยธรรมาธิราช - ปริญญาโท วศ.ม.(วิศวกรรมไฟฟ้า) คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ประสบการณ์ทำงาน	วิศวกร, เจ้าหน้าที่ความปลอดภัยทางรังสี บริษัทไทยยูนิคจำกัด ตั้งแต่ พ.ศ. 2546 จนถึงปัจจุบัน
เบอร์โทรศัพท์	08-1614-5101
อีเมลล์	somchai_po@mail.rmutt.ac.th

