

การรู้จำใบหน้าแบบมีสิ่งบดบังโดยใช้โครงข่ายประสาทเทียมแบบสังวัตนาการ และการสร้างคีนภาพ
ด้วย 2DPCA และ 2DLDA



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมและเทคโนโลยีการป้องกันประเทศ ไม่สังกัดภาควิชา/เทียบเท่า

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2561

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Face Recognition with Occlusion Using Convolutional Neural Network and 2DPCA and
2DLDA Image Reconstruction



A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering in Defense Engineering and Technology

Common Course

Faculty of Engineering

Chulalongkorn University

Academic Year 2018

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การรู้จำใบหน้าแบบมีสิ่งบดบังโดยใช้โครงข่ายประสาทเทียมแบบสัง
	วัฒนาการ และการสร้างคุณภาพด้วย 2DPCA และ 2DLDA
โดย	นายสิทธิพันธุ์ สระภักดิ์
สาขาวิชา	วิศวกรรมและเทคโนโลยีการป้องกันประเทศ
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ผู้ช่วยศาสตราจารย์ ดร.ชาญชัย ปลื้มปิติวิริยะเวช
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม	พันเอก รองศาสตราจารย์ ดร.เผดิม หนังสือ

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้บัณฑิตวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

.....	คณบดีคณะวิศวกรรมศาสตร์
(ศาสตราจารย์ ดร.สุพจน์ เตชวรสินสกุล)	
คณะกรรมการสอบวิทยานิพนธ์	ประธานกรรมการ
.....	
(ผู้ช่วยศาสตราจารย์ ดร.วิทยากร อัครวิเศษ)	
.....	อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผู้ช่วยศาสตราจารย์ ดร.ชาญชัย ปลื้มปิติวิริยะเวช)	
.....	อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม
(พันเอก รองศาสตราจารย์ ดร.เผดิม หนังสือ)	
.....	กรรมการภายนอกมหาวิทยาลัย
(ผู้ช่วยศาสตราจารย์ ดร.สมเกียรติ วงศ์ศิริพิทักษ์)	

สิทธิบัตร สาระกัณฑ์ : การรู้จำใบหน้าแบบมีสิ่งบังบังโดยใช้โครงข่ายประสาทเทียมแบบสังวัตนาการ และการสร้างคืนภาพด้วย 2DPCA และ 2DLDA. (Face Recognition with Occlusion Using Convolutional Neural Network and 2DPCA and 2DLDA Image Reconstruction) อ.ที่ปรึกษาหลัก : ผศ. ดร.ชาญชัย ปลื้มปิติวิริยะเวช, อ.ที่ปรึกษาร่วม : พ.อ. รศ. ดร.เผดิม หนึ่งสี

แม้ว่าระบบจดจำใบหน้าส่วนใหญ่ต้องการส่วนประกอบของใบหน้าทั้งหมดเพื่อผลลัพธ์ที่แม่นยำ แต่ในโลกแห่งความเป็นจริงหลายคนรวมถึงผู้ก่อการร้ายมักอำพรางตัวเองจากกล้องวงจรปิดด้วยการสวมแว่นกันแดดหรือผ้าพันคอหรือหลีกเลี่ยงมุมกล้อง ทำให้การจดจำใบหน้ามีความท้าทายมากขึ้น ในวิทยานิพนธ์นี้ใช้การวิเคราะห์องค์ประกอบสองมิติ (2DPCA) และการวิเคราะห์แบ่งแยกเชิงเส้นสองมิติ (2DLDA) ประยุกต์กับการรวบรวมข้อมูลภาพแบบโครงข่ายประสาทแบบสังวัตนาการ (Convolutional Neural Network) โดยใช้แบบจำลองของ Alexnet เมื่อใช้ฐานข้อมูล AR กับ 227 มิติเวกเตอร์เฉพาะ (Eigenvector) สำหรับ 2DPCA และ 2DLDA ให้ความแม่นยำร้อยละ 53.06 และ 49.71 ตามลำดับ อย่างไรก็ตามเมื่อใช้เวกเตอร์เฉพาะสำหรับสร้างภาพใหม่ความแม่นยำจะเพิ่มขึ้นเป็นร้อยละ 81.55 และ 80.56 ตามลำดับ เมื่อใช้ฐานข้อมูล GTAV กับ 227 มิติเวกเตอร์เฉพาะ สำหรับ 2DPCA และ 2DLDA ให้ความแม่นยำร้อยละ 86.36 และ 78.54 ตามลำดับ อย่างไรก็ตามเมื่อใช้เวกเตอร์เฉพาะสำหรับสร้างภาพใหม่ความแม่นยำจะเพิ่มขึ้นเป็นร้อยละ 96.46 และ 91.92 ตามลำดับ

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สาขาวิชา	วิศวกรรมและเทคโนโลยีการป้องกัน	ลายมือชื่อนิสิต
	ประเทศ	
ปีการศึกษา	2561	ลายมือชื่อ อ.ที่ปรึกษาหลัก
		ลายมือชื่อ อ.ที่ปรึกษาร่วม

5970376721 : MAJOR DEFENSE ENGINEERING AND TECHNOLOGY

KEYWORD: two-dimensional linear discriminant analysis (2DLDA), two-dimensional principal component analysis (2DPCA), convolutional neural network (CNN)

Sittiphan Sarapakdi : Face Recognition with Occlusion Using Convolutional Neural Network and 2DPCA and 2DLDA Image Reconstruction. Advisor: Asst. Prof. Charnchai Pluempitiwiriyawej, Ph.D. Co-advisor: Col. Assoc. Prof. Phaderm Nangsue, Ph.D.

Although most face recognition systems require all face components for accurate results, in the real world, many people including the terrorists often camouflage themselves from CCTVs by wearing sunglasses or scarf or evade the camera angle. This makes face recognition more challenging. In this thesis, we are to recognize faces with occlusions using two-dimensional principal component analysis (2DPCA) and two-dimensional linear discriminant analysis (2DLDA), which are applied together with convolutional neural network using Alexnet model. When using the AR database with 227 eigenvectors for 2DPCA and 2DLDA, the recognition accuracy is 53.06% and 49.71%, respectively. However, when using the eigenvectors to reconstruct a new image, the accuracy is increased to 81.55% and 81.62%, respectively. When using the GTAV database with 227 eigenvectors for 2DPCA and 2DLDA, the recognition accuracy is 87.12% and 83.08%, respectively. When using the eigenvectors to reconstruct a new image, the accuracy is increased to 97.22% and 94.95%, respectively.



Field of Study: Defense Engineering and
Technology

Academic Year: 2018

Student's Signature

Advisor's Signature

Co-advisor's Signature

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้ได้สำเร็จลุล่วงไปด้วยดี โดยได้รับความแนะนำและความช่วยเหลือจากอาจารย์ที่ปรึกษาวิทยานิพนธ์ ผู้ช่วยศาสตราจารย์ ดร. ชานูชัย ปลื้มปิติวิริยะเวช และที่ปรึกษาร่วม พันเอก รองศาสตราจารย์ ดร. เผติมหนังสือ โดยทั้งสองท่านที่ได้ถ่ายทอดความรู้พื้นฐานวิชาการและด้านวิจัยที่นำไปประยุกต์ใช้เพื่อบูรณาการต่อการทำวิทยานิพนธ์ พร้อมทั้งแนวคิดและแรงผลักดันรวมทั้งกำลังใจในการทำวิทยานิพนธ์ฉบับนี้ให้สำเร็จลุล่วงไปด้วยดี ผู้วิจัยขอกราบขอบพระคุณไว้ ณ ที่นี้ นอกจากนี้ขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร. วิทยากร อัครวิเศษ ประธานกรรมการสอบวิทยานิพนธ์ และ ผู้ช่วยศาสตราจารย์ ดร. สมเกียรติ วังศิริพิทักษ์ กรรมการสอบวิทยานิพนธ์ที่ได้สละเวลาตรวจสอบ และให้คำแนะนำเพื่อให้วิทยานิพนธ์ฉบับนี้สมบูรณ์ยิ่งขึ้น

ขอขอบพระคุณ บัณฑิตวิทยาลัยที่สนับสนุนงานวิจัยในการเสนอผลงานทางวิชาการในประเทศระดับนานาชาติ รวมไปถึงสาขาวิศวกรรมและเทคโนโลยีการป้องกันประเทศ ที่ช่วยเหลือในด้านข้อมูลข่าวสารและประสานงานต่างๆ

ขอขอบคุณ เพื่อน รุ่นพี่ รุ่นน้องทุกคนของชมรมบัณฑิตศึกษา คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ที่คอยสนับสนุน และให้กำลังใจในการทำวิทยานิพนธ์เป็นอย่างดี

สุดท้ายนี้ขอขอบพระคุณบิดามารดา และพี่สาวที่ให้กำลังใจ เชื้อมั่นและสนับสนุนผู้วิจัยเสมอมาจนสำเร็จการศึกษา

สิทธิพันธุ์ สระภักดิ์

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ค
บทคัดย่อภาษาอังกฤษ.....	ง
กิตติกรรมประกาศ.....	จ
สารบัญ.....	ฉ
สารบัญตาราง.....	ญ
สารบัญรูปภาพ.....	ฎ
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์	3
1.3 ขอบเขตวิทยานิพนธ์	3
1.4 ขั้นตอนและวิธีการดำเนินการ	4
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	4
บทที่ 2 ทฤษฎีพื้นฐาน.....	5
2.1 กระบวนการรู้จำใบหน้า (Face recognition).....	5
2.1.1 วิธีการเปรียบเทียบกราฟ (Graph matching).....	5
2.1.2 ลักษณะทางเรขาคณิต (Geometrical feature).....	6
2.1.3 วิธีการโครงข่ายประสาท (Neural network approach).....	6
2.1.4 วิธีการทางสถิติ (Statistical approach).....	7
2.2 การวิเคราะห์ส่วนประกอบหลัก 2 มิติ (Two-dimensional Principal Component Analysis: 2DPCA).....	9
2.2.1 หลักการคำนวณ	9

2.2.2 การหาคุณลักษณะของภาพ (feature extraction).....	10
2.2.3 การสร้างคืนภาพ (reconstructed image).....	10
2.3 การวิเคราะห์การจำแนกเชิงเส้น 2 มิติ (Two-dimensional Linear Discriminant Analysis: 2DLDA).....	11
2.3.1 หลักการคำนวณ	11
2.3.2 การหาคุณลักษณะของภาพ (feature extraction).....	12
2.3.3 การสร้างคืนภาพ (reconstructed image).....	12
2.4 โครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Network: CNN).....	13
2.4.1 การทำงานของชั้นสังวัตนาการ (Convolution Layer).....	14
2.4.2 การทำงานของฟังก์ชันกระตุ้น (Activation Function).....	16
2.4.3 ชั้นพูลลิง (Pooling Layer)	18
2.4.4 ชั้นการเชื่อมโยงเต็มรูปแบบ (Fully Connected Layer).....	19
2.4.5 การปรับค่าไฮเพอร์พารามิเตอร์ (Fine tuning hyper-parameter).....	20
2.5.6 สัญญาณตกหาย (Dropout).....	20
2.6 โครงข่ายประสาทเทียมแบบสังวัตนาการของแบบจำลอง Alexnet.....	21
บทที่ 3 วิธีการดำเนินงานวิจัย.....	23
3.1 การได้มาของภาพ	23
3.1.1 ฐานข้อมูล AR.....	23
3.1.2 ฐานข้อมูล GTAV	25
3.2 กระบวนการปรับสีภาพ.....	26
3.3 โครงสร้างการทำงานของระบบรู้จำใบหน้าแบบมีสิ่งบดบัง.....	26
3.4 เครื่องมือที่ใช้พัฒนา.....	27
3.5 การทำงานของโครงข่ายประสาทเทียมแบบสังวัตนาการแบบจำลองของ Alexnet.....	27
3.6 โครงข่ายประสาทแบบจำลองของผู้วิจัย	29

บทที่ 4 ผลการทดลองและการวิเคราะห์ผล	31
4.1 ผลการทดลองการสร้างคืนภาพด้วยวิธี 2DPCA และ 2DLDA	31
4.1.1 ผลการทดลองการสร้างคืนภาพด้วยวิธี 2DPCA	31
4.1.2 ผลการทดลองการสร้างคืนภาพด้วยวิธี 2DLDA	34
4.2 ผลการทดลองการปรับค่าไฮเพอร์พารามิเตอร์สำหรับการฝึกสอนของโครงข่ายประสาทเทียม แบบสังวัตนาการ โดยใช้แบบจำลองของ Alexnet.....	36
4.2.1 ผลการทดลองการปรับค่าไฮเพอร์พารามิเตอร์ จากฐานข้อมูล AR.....	37
4.2.2 ผลการทดลองการปรับค่าไฮเพอร์พารามิเตอร์ จากฐานข้อมูล GTAV.....	39
4.3 ผลการทดลองการเปรียบเทียบแบบใช้และไม่ใช้วิธีเชิงสถิติร่วมกับโครงข่ายประสาท แบบสัง วัตนาการ โดยใช้แบบจำลองของ Alexnet.....	40
4.3.1 ผลการทดลองจากฐานข้อมูล AR.....	41
4.4 ผลการทดลองการเปรียบเทียบระหว่างวิธี 2DPCA และ 2DLDA ร่วมกับโครงข่ายประสาท แบบสังวัตนาการ โดยใช้แบบจำลองของ Alexnet กับผู้วิจัย	43
4.4.1 ผลการทดลองจากฐานข้อมูล AR.....	43
4.4.2 ผลการทดลองจากฐานข้อมูล GTAV	46
4.5 ผลการทดลองการเปรียบเทียบระหว่างการสร้างคืนภาพด้วยวิธี 2DPCA และ 2DLDA ร่วมกับ โครงข่ายประสาทเทียมแบบสังวัตนาการ โดยใช้แบบจำลอง Alexnet.....	48
4.5.1 ผลการทดลองจากฐานข้อมูล AR.....	48
4.5.2 ผลการทดลองจากฐานข้อมูล GTAV	51
4.6 ผลการทดลองนำระบบการรู้จำใบหน้าแบบมีสิ่งบดบังไปใช้ในชีวิตจริง	52
บทที่ 5 สรุป อภิปรายผล และข้อเสนอแนะ	56
5.1 เพื่อศึกษาทฤษฎีการรู้จำใบหน้าในการหาคุณลักษณะของภาพกรณีที่มีสิ่งบดบังใบหน้า เพื่อ นำไปใช้ในการวิจัย.....	56
5.2 เพื่อใช้วิธีเชิงสถิติการวิเคราะห์ส่วนประกอบหลัก 2 มิติ และการวิเคราะห์การจำแนก เชิงเส้น 2 มิติ เพื่อรู้จำใบหน้าที่มีสิ่งบดบัง	56

5.3 เพื่อวิเคราะห์และเปรียบเทียบความถูกต้องและความเร็วในการรู้จำจากการวิเคราะห์ ส่วนประกอบหลัก 2 มิติ และการวิเคราะห์การจำแนกเชิงเส้น 2 มิติ กรณีมีสิ่งบดบังบน ใบหน้า.....	57
5.4 สรุปผลการวิจัย.....	59
5.5 ข้อเสนอแนะเพื่อการนำไปใช้	60
5.6 ข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป	60
ภาคผนวก.....	61
บรรณานุกรม.....	70
ประวัติผู้เขียน.....	73



สารบัญตาราง

	หน้า
ตารางที่ 3.1 ตารางการทำงานแบบจำลอง Alexnet.....	28
ตารางที่ 3.2 ผลลัพธ์จากการคำนวณหาขนาดข้อมูลก่อนเข้าชั้นการเชื่อมโยงเต็มรูปแบบชั้นที่ 17... 30	30



สารบัญรูปภาพ

	หน้า
รูปที่ 2.1 ตัวอย่างชั้นโครงข่ายประสาทเทียมแบบสังวัตนาการที่มีหลายชั้น	13
รูปที่ 2.2 แสดงขั้นตอนการทำสังวัตนาการระหว่างภาพนำเข้า	15
รูปที่ 2.3 ตัวอย่างการขยับเคลื่อนของตัวกรองครึ่งละ 1 กับ 2 ช่อง	15
รูปที่ 2.4 การ padding size เพื่อเก็บรายละเอียดขอบภาพ	16
รูปที่ 2.5 กราฟฟังก์ชันซิกมอยด์ (Sigmoid Function).....	16
รูปที่ 2.6 กราฟฟังก์ชันแทนเจนไฮเพอร์โบลิก (Hyperbolic tangent function).....	17
รูปที่ 2.7 กราฟฟังก์ชันเรคตีไฟด์เชิงเส้น (Rectified Linear Unit: ReLU).....	18
รูปที่ 2.8 แสดงการทำพูลลิ่งเฉลี่ยและพูลลิ่งสูงสุด ขนาด 2 x 2.....	19
รูปที่ 2.9 แสดงการเชื่อมโยงเต็มรูปแบบโหนดแต่ละโหนด.....	19
รูปที่ 2.10 เปรียบเทียบสัญญาณตกหายของโครงข่ายประสาท (ก) แสดงโครงข่ายประสาทแบบปกติ (ข) แสดงการลดลงโหนดที่มีการตัดเส้นเชื่อมของโครงข่ายประสาท.....	20
รูปที่ 2.11 แสดงโครงข่ายประสาทเทียมแบบสังวัตนาการของแบบจำลอง Alexnet	21
รูปที่ 3.1 กลุ่มรูปฝึกสอนตัวอย่างของบุคคลหนึ่ง จากฐานข้อมูล AR.....	23
รูปที่ 3.2 กลุ่มรูปทดสอบตัวอย่างของบุคคลหนึ่ง จากฐานข้อมูล AR.....	24
รูปที่ 3.3 ตัวอย่างการเพิ่มการบดบังใบหน้าเพื่อช่วยในการฝึกสอน.....	24
รูปที่ 3.4 รูปกลุ่มฝึกสอนตัวอย่างจากฐานข้อมูล GTAV	25
รูปที่ 3.5 รูปกลุ่มฝึกสอนตัวอย่างจากฐานข้อมูล GTAV	25
รูปที่ 3.6 ตัวอย่างการปรับสีภาพในปริภูมิสีต่างๆ (ก) RGB (ข) Grayscale (ค) YCbCr (ง) YCbCr + Grayscale	26
รูปที่ 3.7 โครงสร้างการทำงานของระบบ	27
รูปที่ 4.1 กราฟแสดงการเรียงลำดับเวกเตอร์เฉพาะที่สมนัยกับค่าเฉพาะจากมากไปน้อย	32

รูปที่ 4.2 ตัวอย่างการสร้างคืนภาพของบุคคลหนึ่งด้วยวิธี 2DPCA (ก) ภาพต้นแบบ (ข) มิติ = 10 (ค) มิติ = 20 (ง) มิติ = 30 (จ) มิติ = 40.....	33
รูปที่ 4.3 กราฟแสดงเรียงลำดับเวกเตอร์เฉพาะที่สมนัยกับค่าเฉพาะจากมากไปน้อย 2DLDA	34
รูปที่ 4.4 ตัวอย่างการสร้างคืนภาพของบุคคลหนึ่งด้วยวิธี 2DLDA (ก) ภาพต้นแบบ (ข) มิติ = 222 (ค) มิติ = 223 (ง) มิติ = 224 (จ) มิติ = 225 (จ) มิติ = 226.....	35
รูปที่ 4.5 กราฟแสดงอัตราการสูญเสียระหว่างการฝึกสอนของทุกอัตราการเรียนรู้ต่อขนาดกลุ่มภาพฝึกสอน 32 ภาพต่อกลุ่มจำนวนในการฝึกสอน 20 รอบ จากฐานข้อมูล AR.....	36
รูปที่ 4.6 กราฟแสดงความแม่นยำของอัตราการเรียนรู้ต่อขนาดชุดข้อมูลจากฐานข้อมูล AR.....	37
รูปที่ 4.7 กราฟแสดงเวลาที่ใช้ในการฝึกสอนต่ออัตราเร็วการเรียนรู้ของทุกขนาดกลุ่มภาพฝึกสอน จากฐานข้อมูล AR.....	38
รูปที่ 4.8 กราฟแสดงความแม่นยำของอัตราการเรียนรู้ต่อขนาดชุดข้อมูลจากฐานข้อมูล GTAV	39
รูปที่ 4.9 กราฟแสดงเวลาในการฝึกสอนของอัตราการเรียนรู้ต่อขนาดชุดข้อมูลจากฐานข้อมูล GTAV	40
รูปที่ 4.10 กราฟแสดงความแม่นยำระหว่างการใช้และไม่ใช้วิธีเชิงสถิติร่วมกับโครงข่ายประสาทเทียมแบบสังวัตนาการ โดยใช้แบบจำลอง Alexnet จากฐานข้อมูล AR.....	41
รูปที่ 4.11 แสดงความแม่นยำระหว่างการใช้และไม่ใช้วิธีเชิงสถิติร่วมกันโครงข่ายประสาท แบบสังวัตนาการโดยใช้แบบจำลองของ Alexnet.....	42
รูปที่ 4.12 กราฟแสดงความแม่นยำระหว่างการใช้เชิงสถิติร่วมกันโครงข่ายประสาทแบบสังวัตนาการ โดยใช้แบบจำลองผู้วิจัย.....	44
รูปที่ 4.13 แสดงระยะเวลาการฝึกสอนระหว่างการใช้และไม่ใช้วิธีเชิงสถิติร่วมกันโครงข่ายประสาทแบบสังวัตนาการโดยใช้แบบจำลองผู้วิจัย	45
รูปที่ 4.14 กราฟแสดงความแม่นยำระหว่างการใช้วิธี 2DPCA และ 2DLDA ร่วมกันโครงข่ายประสาทแบบสังวัตนาการโดยใช้แบบจำลองของ Alexnet และ ผู้วิจัย	46
รูปที่ 4.15 กราฟแสดงระยะเวลาการฝึกสอนระหว่างการใช้วิธีเชิงสถิติร่วมกันโครงข่ายประสาทแบบสังวัตนาการโดยใช้แบบจำลองผู้วิจัย.....	47
รูปที่ 4.16 กราฟแสดงการเปรียบเทียบความแม่นยำในการสร้างคืนภาพของวิธีเชิงสถิติทุก 10 มิติของเวกเตอร์เฉพาะ จากฐานข้อมูล AR.....	49

รูปที่ 4.17 กราฟแสดงการเปรียบเทียบระหว่างบุคคลต่อการทำนายผล จากฐานข้อมูล AR ที่ร้อยละ 81.55..... 50

รูปที่ 4.18 กราฟแสดงการเปรียบเทียบความแม่นยำของการสร้างคืนภาพด้วยวิธีเชิงสถิติทุก 10 มิติของเวกเตอร์เฉพาะ จากฐานข้อมูล GTAV..... 51

รูปที่ 4.19 กราฟแสดงการเปรียบเทียบระหว่างบุคคลต่อการทำนายผล จากฐานข้อมูล GTAV ที่ร้อยละ 96.46 52

รูปที่ 4.20 ตัวอย่างการเก็บข้อมูลรูปภาพของบุคคลหนึ่งที่มีลักษณะแตกต่างกันออกไป 53

รูปที่ 4.21 แสดงการทดลองโดยใช้ใบหน้าปกติ..... 53

รูปที่ 4.22 แสดงการทดลองโดยสวมแว่นกันแดด 54

รูปที่ 4.23 แสดงการทดลองโดยสวมผ้าปิดปาก 54

รูปที่ 4.24 แสดงการทดลองโดยสวมแว่นกันแดดและผ้าปิดปาก..... 55



บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันนี้ เทคโนโลยีการประมวลผลภาพได้เข้ามามีบทบาทในชีวิตประจำวันมากขึ้น หนึ่งในเทคโนโลยีที่มีความโดดเด่นและน่าสนใจเป็นอย่างมากได้แก่ การรู้จำใบหน้า [1] ซึ่งเป็นงานวิเคราะห์ภาพใบหน้า และระบุตัวบุคคลในภาพ เนื่องจากเป็นเทคโนโลยีที่มีประโยชน์และใช้งานได้หลายประเภท เช่น การตั้งกล้องจับภาพคนเข้าเมืองเพื่อหาตัวผู้กระทำความผิดกฎหมายหรือผู้ก่อการร้าย งานรักษาความปลอดภัยในหน่วยงาน สถานที่ราชการ หรือสถานศึกษาต่างๆ เป็นต้น

ในการใช้จริงของเทคโนโลยีการประมวลผลภาพนั้นยังมีปัจจัยอีกหลายด้านที่ทำให้ความแม่นยำของกระบวนการรู้จำใบหน้าไม่ดีเท่าที่ควร เนื่องจากภาพที่นำมาประมวลผลนั้นอาจมีปัจจัยแปรผันหลายอย่าง เช่น อายุที่เปลี่ยนไปของแต่ละบุคคล (Ageing) การแปรผันทางตำแหน่งของใบหน้า (Pose) การแปรผันของแสง (Illumination) การแปรผันเนื่องจากการแสดงทางอารมณ์ (Expression) ภาพใบหน้าที่มีความละเอียดต่ำ (low resolution) [2] ดังนั้น การรู้จำใบหน้าในอดีตจำเป็นต้องใช้ส่วนประกอบของใบหน้าครบทุกส่วนจึงจะให้ประสิทธิภาพสูงสุด แต่ในโลกความเป็นจริงนั้น ยังมีอีกหลายกรณีที่ไม่สามารถใช้ส่วนประกอบของใบหน้าทั้งหมดได้ อันเนื่องมาจาก ยังมีผู้คนมากมายรวมถึงผู้ที่กระทำความผิดกฎหมายจะสวมใส่อุปกรณ์ปิดบังใบหน้า อาทิ แว่นกันแดด หรือผ้าปิดปาก รวมทั้งการเบี่ยงใบหน้าเพื่อที่จะอำพรางตนเองจากกล้องวงจรปิด ทำให้การรู้จำใบหน้าบุคคลยากยิ่งขึ้น

วิธีการรู้จำใบหน้ามีหลายวิธี [3, 4] แต่วิธีที่ได้รับความนิยมในอดีตคือวิธีเชิงสถิติ (Statistic Method) ซึ่งเป็นการเทียบค่าความสว่างของจุดภาพในภาพใบหน้าอ้างอิงกับภาพใบหน้าที่ต้องการรู้จำโดยการเปรียบเทียบแบบจุดภาพต่อจุดภาพ ทั้งนี้เพื่อให้ขนาดของข้อมูลที่ต้องการประมวลผลมีขนาดเล็กลง (dimension reduction) และใช้การวิเคราะห์ส่วนประกอบหลัก (principal component analysis: PCA) และการวิเคราะห์การจำแนกเชิงเส้น (linear discriminate analysis: LDA) เป็นหลักในการคำนวณเฉพาะเวกเตอร์ใน 1 มิติ เมื่อต้องการใช้วิธีดังกล่าวกับภาพ จึงจำเป็นต้องเปลี่ยนเมทริกซ์ภาพให้เป็นเวกเตอร์ภาพ จากนั้น จึงใช้พื้นฐานของเวกเตอร์มาคำนวณหาเมทริกซ์ความแปรปรวนร่วม (covariance matrix) เพื่อคำนวณหาเวกเตอร์ที่เหมาะสมที่สุดในการฉายภาพ (projection) หลักการดังกล่าวนี้ถูกเรียกว่า ไอเกนเฟส (Eigenface) [5] และ ฟิชเชอร์เฟส (Fisher face) [6] เมื่อเรานำหลักการดังกล่าวมาใช้ในงานด้านการรู้จำใบหน้าเป็นผลให้ได้กระบวนการรู้จำที่ดี แต่ว่าภาพใบหน้าที่ใช้มีมิติข้อมูลมาก เช่น ภาพใบหน้าที่มีขนาด $m \times n$ จุดภาพ

เมื่อเปลี่ยนเป็นเวกเตอร์ภาพจะมีขนาด $m \times n$ มิติ จากเหตุผลดังกล่าว จึงทำให้วิธีคำนวณด้วยไอเกนเฟสและพีชเซอร์เฟสใช้เวลามาก

Yang [7] ได้วิจัยและเสนอวิธีการที่ชื่อว่า “การวิเคราะห์ส่วนประกอบหลัก 2 มิติ” (two-dimensional principal component analysis: 2DPCA) สำหรับการรู้จำใบหน้า โดยได้ทำการศึกษาปรับปรุงประสิทธิภาพของเทคนิค PCA แบบเดิมและแก้ไขจุดบกพร่อง ผลการศึกษาพบว่า วิธีการใหม่นี้นอกจากจะพัฒนาให้มีความเร็วในการคำนวณแล้วยังมีความถูกต้องในการรู้จำที่ดีกว่าวิธีเดิมอีกด้วย

Kongsontana [4] ได้ศึกษาทดลองและเสนอวิธีการที่มีชื่อว่า การวิเคราะห์การจำแนกเชิงเส้น 2 มิติ (Two-dimensional Linear Discriminant Analysis: 2DLDA) โดยการหาคุณลักษณะของภาพเพื่อนำมาใช้ในงานด้านการจดจำใบหน้า วิธีการนี้พัฒนามาจากวิธีการแบ่งแยกเชิงเส้นของพีชเซอร์และวิเคราะห์ส่วนประกอบหลักของเมทริกซ์ภาพ ผลการทดลองพบว่าวิธีการใหม่นี้ให้อัตราความถูกต้องสูงที่สุดเมื่อเทียบกับเทคนิค PCA และ LDA ใน 1 มิติ

Wang [8] ได้ทำการศึกษาเปรียบเทียบประสิทธิภาพการรู้จำใบหน้าแบบไม่มีสิ่งบดบังบนใบหน้าด้วยวิธี 2DPCA และ 2DLDA ซึ่งเป็นการพัฒนามาจาก PCA และ LDA ใน 1 มิติ ผลการทดลองปรากฏว่าวิธี 2DPCA มีประสิทธิภาพที่ใกล้เคียงกันกับวิธี 2DLDA เมื่อเป็นการเปรียบเทียบข้อมูลขนาดไม่ใหญ่มาก แต่ถ้าหากเป็นการเปรียบเทียบข้อมูลขนาดใหญ่วิธี 2DPCA มีประสิทธิภาพที่ดีกว่า

Kim [9] ได้นำวิธีการวิเคราะห์ส่วนประกอบหลัก 2 มิติ มาใช้กับการรู้จำใบหน้าแบบมีสิ่งบดบังด้วยการใส่แว่นกันแดด และสวมผ้าคลุมปิดปาก โดยได้แบ่งการศึกษานี้ออกเป็นสองตอนคือการตรวจจับส่วนที่ปิดบังอยู่ด้วย และการใช้วิธีการเพื่อนบ้านใกล้ที่สุด (k-nearest neighbors: k-NN and 1-NN) เพื่อทำการลบส่วนที่บดบังออก ส่วนใบหน้าที่เหลือจะนำมาหาส่วนที่คล้ายคลึงกันระหว่างเมทริกซ์ โดยใช้วิธีการหาระยะทางยูคลิเดียน (Euclidean distance) ที่ผสมกับวิธีการเพื่อนบ้านใกล้ที่สุด ผลการศึกษาพบว่าวิธี k-NN และ 1-NN ดังกล่าวมีประสิทธิภาพดีที่สุดเมื่อเทียบกับวิธีอื่นๆ ที่เคยมีมา เช่น PCA แต่ยังไม่ได้ทดลองกับ LDA และ 2DLDA

Ren [10] ได้เสนอการวิเคราะห์ส่วนประกอบหลัก (PCA) เพื่อที่จะหาเวกเตอร์เฉพาะและใช้เป็นค่าเริ่มต้นของตัวกรองของสังวัตนาการ (convolutional kernel) ซึ่งจะรวมกับกระบวนการฝึกสอนของโครงข่ายประสาทเทียมแบบสังวัตนาการ (convolutional neural network) โดยค่าเริ่มต้นชนิดนี้ประกอบด้วยข้อมูลรูปภาพและลดผลกระทบของปัญหาการกระจาย (gradient diffusion) เนื่องจากพารามิเตอร์ไม่ถูกต้อง ทำการทดลองในชุดข้อมูล Mnist และ Cifar-10 โดยแบบจำลองที่นำมาเสนอนี้จะช่วยลดขั้นตอนการทำซ้ำและมีประสิทธิภาพมากกว่า มีโครงสร้างที่เรียบ

ง่ายรวมทั้งยังใช้เวลาฝึกอบรมที่น้อยลงเมื่อเทียบกับแบบจำลองของโครงข่ายประสาทเทียมแบบเดิมได้

วิทยานิพนธ์นี้ เป็นการประยุกต์วิธีเชิงสถิติเพื่อเปรียบเทียบประสิทธิภาพระหว่างการวิเคราะห์ส่วนประกอบหลัก 2 มิติ และ การวิเคราะห์การจำแนกเชิงเส้น 2 มิติ ซึ่งประยุกต์จาก PCA และ LDA ใน 1 มิติ มาใช้กับโครงข่ายประสาทเทียมแบบสังวัตนาการเพื่อใช้ในการเรียนรู้ และวัดประสิทธิภาพด้านความแม่นยำและเวลาในการเรียนรู้สำหรับการรู้จำใบหน้าแบบมีสิ่งบดบัง โดยใช้หลักการทำงานของระบบคือ เมื่อนำใบหน้าที่มีสิ่งบดบังประกอบด้วย แว่นกันแดด ผ้าคลุมปิดปาก และการบดบังด้วยตัวเอง จะสามารถตรวจสอบใบหน้าที่มีอยู่ในฐานข้อมูลได้หรือไม่ และเมื่อได้ผลลัพธ์จากการประมวลผลสามารถแสดงให้เห็นผู้ใช้งานทราบว่าบุคคลในฐานข้อมูลเป็นใคร

1.2 วัตถุประสงค์

1. เพื่อศึกษาทฤษฎีการรู้จำใบหน้า ที่เกี่ยวกับการหาคุณลักษณะของภาพกรณีที่มีสิ่งบดบังใบหน้าเพื่อนำไปใช้ในการวิจัย
2. เพื่อใช้วิธีเชิงสถิติการวิเคราะห์ส่วนประกอบหลัก 2 มิติ และการวิเคราะห์การจำแนกเชิงเส้น 2 มิติ เพื่อรู้จำใบหน้าที่มีสิ่งบดบัง
3. เพื่อวิเคราะห์และเปรียบเทียบความถูกต้องและความเร็วในการรู้จำจากการวิเคราะห์ส่วนประกอบหลัก 2 มิติ และการวิเคราะห์การจำแนกเชิงเส้น 2 มิติ กรณีมีสิ่งบดบังบนใบหน้า

1.3 ขอบเขตวิทยานิพนธ์

1. สร้างระบบการรู้จำใบหน้าที่มีการบดบัง และใบหน้าที่บดบังด้วยตนเอง โดยใช้การวิเคราะห์ส่วนประกอบหลัก 2 มิติ และ การวิเคราะห์การจำแนกเชิงเส้น 2 มิติ
2. สามารถจัดเก็บฐานข้อมูลรูปภาพบุคคลได้ไม่น้อยกว่า 100 คน โดยใช้ฐานข้อมูลภาพจากสองฐานข้อมูล คือ AR และ GTAV
3. เงื่อนไขการบดบังใบหน้าของวิทยานิพนธ์นี้ หมายถึง การสวมแว่นกันแดด การสวมผ้าคลุมปิดปาก และการบดบังด้วยตัวเอง

1.4 ขั้นตอนและวิธีการดำเนินการ

1. ศึกษาค้นคว้า รวบรวมทฤษฎีและงานวิจัยที่เกี่ยวกับรู้จำใบหน้าแบบมีสิ่งบดบัง
2. ดำเนินการศึกษาและทดลองตามวัตถุประสงค์ของการวิจัย
3. รวบรวมผลการทดลองและการวิเคราะห์ผล
4. สรุปและอภิปรายผล
5. เขียนรายงานวิทยานิพนธ์ตามรูปแบบของมหาวิทยาลัย
6. นำเสนอวิทยานิพนธ์ ตีพิมพ์ และเผยแพร่

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. สามารถเรียนรู้วิธีการรู้จำใบหน้าที่มีสิ่งบดบังด้วยวิธีการวิเคราะห์ส่วนประกอบหลัก 2 มิติ และการวิเคราะห์การจำแนกเชิงเส้น 2 มิติ ได้
2. สามารถวิเคราะห์และเปรียบเทียบการรู้จำด้วยการวิเคราะห์ส่วนประกอบหลัก 2 มิติ และการวิเคราะห์การจำแนกเชิงเส้น 2 มิติ กรณีที่มีสิ่งบดบังบนใบหน้าด้วยความถูกต้องและในเวลาที่รวดเร็ว
3. สามารถนำโปรแกรมการรู้จำใบหน้าที่มีสิ่งบดบังของผู้วิจัยไปเป็นต้นแบบของการระบุใบหน้า เมื่อมีการบดบังไปพัฒนาต่อไปได้
4. สามารถประยุกต์โปรแกรมการรู้จำใบหน้าที่มีสิ่งบดบังของผู้วิจัยไปใช้ในหน่วยงาน สถานที่ราชการ หรือสถานศึกษาต่างๆ ได้
5. สามารถตีพิมพ์บทความวิจัยในวารสารทางวิชาการระดับนานาชาติ

บทที่ 2

ทฤษฎีพื้นฐาน

ในบทนี้จะกล่าวถึงทฤษฎีพื้นฐานที่เกี่ยวข้องกับการรู้จำใบหน้ากรณีมีสิ่งบดบัง กล่าวถึงวิธีเชิงสถิติ ได้แก่ การวิเคราะห์ส่วนประกอบหลัก 2 มิติ การวิเคราะห์จำแนกเชิงเส้น 2 มิติ และ โครงข่ายประสาทเทียมแบบสังวัตนาการ

2.1 กระบวนการรู้จำใบหน้า (Face recognition)

กระบวนการรู้จำใบหน้ามีวิธีการมากมายที่มีการวิจัยในปัจจุบันสามารถแบ่งเป็นกลุ่มต่างๆ มีรายละเอียดดังต่อไปนี้

1. วิธีการเปรียบเทียบกราฟ (Graph matching)
2. ลักษณะทางเรขาคณิต (Geometrical feature)
3. วิธีการโครงข่ายประสาท (Neural network approach)
4. วิธีการทางสถิติ (Statistical approach)

2.1.1 วิธีการเปรียบเทียบกราฟ (Graph matching)

เป็นการมองภาพในลักษณะที่เป็นเวกเตอร์ของกราฟ ซึ่งมีจุดและเส้นในการเชื่อมต่อเป็นโครงร่างที่แนบไปบนใบหน้าตามลักษณะเด่นที่พิจารณา เช่น รูปร่างกราฟของตา ปาก คาง จมูก เป็นต้น

Lades [11] เสนอสถาปัตยกรรมการเชื่อมต่อแบบยืดหยุ่น (Dynamic Link Architecture) สำหรับการรู้จำใบหน้าแบบไม่ขึ้นอยู่กับความเพี้ยน โดยใช้กราฟแบบยืดหดได้ (Elastic bunch graph matching) เพื่อหากราฟพื้นฐานข้อมูลที่ใกล้เคียงที่สุด ทำการค้นหาในฐานข้อมูลทั้งหมด 87 คน โดยภาพที่นำมาทดสอบมีลักษณะที่แตกต่างกัน เช่น มีการหมุนของศีรษะ ผลการศึกษาพบว่า ใช้เวลาในการค้นหาเป็น 25 วินาที ต่อหนึ่งครั้ง แต่ไม่มีการแสดงผลถึงการรู้จำเลย

Hanmandlu [12] ได้ใช้การเทียบกราฟยืดหด (Elastic bunch graph matching) โดยใช้ความเหมือนเป็นเกณฑ์การจับคู่ สำหรับการรู้จำใบหน้าแบบไม่ขึ้นอยู่กับการส่องสว่างและการแสดงทางใบหน้า และใบหน้าที่หายไป 1/4 ของใบหน้า การศึกษาพบว่า สามารถรู้จำใบหน้าได้จริง แต่ยังพบปัญหาเมื่อมีการหลับตาอยู่ด้วย

ระบบการรู้จำใบหน้าที่มีการพิจารณาลักษณะที่เป็นเวกเตอร์ของกราฟ เป็นระบบที่มีประโยชน์ และสามารถนำไปใช้ได้ทั้งในทางปฏิบัติ แต่ระบบเหล่านี้จะต้องอ้างอิงตำแหน่งดวงตา ปาก จมูก โครงสร้างใบหน้าอยู่เสมอ จึงพบเห็นข้อเสียของวิธีเหล่านี้เมื่อนำไปประยุกต์กับการรู้จำใบหน้าแบบมีสิ่งบดบัง เพราะเป็นไปไม่ได้เลยที่จะเห็นองค์ประกอบของใบหน้าครบทุกส่วน

2.1.2 ลักษณะทางเรขาคณิต (Geometrical feature)

ระบบการรู้จำใบหน้าที่มีการพิจารณาลักษณะทางเรขาคณิต อาศัยอัตราส่วนหรือสัดส่วนของ ระยะทางจากตาซ้ายไปตาขวา จากตาไปจมูก จากจมูกไปปาก หรือใช้รูปร่างของตา รูปร่างของปาก ซึ่งเป็นลักษณะความสัมพันธ์ทางกายภาพของใบหน้าทั้งสิ้น

Song [13] ได้เพิ่มประสิทธิภาพของระบบการรู้จำโดยเอาหลักลักษณะทางเรขาคณิตผสมเข้ากับการวิเคราะห์องค์ประกอบหลัก (PCA) และการวิเคราะห์การแบ่งแยกเชิงเส้น (LDA) เพราะ Song มองเห็นว่า PCA และ LDA ที่มีอยู่ไม่สามารถสะท้อนเส้นขอบใบหน้าได้ เนื่องจากมีการวัดความคล้ายคลึงกันตามระดับของการกระจายตัวทางกายภาพ ผลการศึกษาพบว่า วิธีที่ Song พัฒนาขึ้นมานั้นให้อัตราการเรียนรู้สูงกว่าวิธีเดิมที่มีอยู่ แต่ไม่มากนัก

Tummala [14] ได้ประยุกต์การรู้จำใบหน้าระหว่างลักษณะทางเรขาคณิต (Geometrical feature) และการวิเคราะห์องค์ประกอบหลัก (PCA) ผสมเข้าด้วยกัน ผลการศึกษาพบว่า ให้อัตราความแม่นยำสูงมากเมื่อเทียบกับการวิเคราะห์องค์ประกอบหลักปกติ และอัตราส่วนทางเรขาคณิตปกติ

ระบบการรู้จำใบหน้าที่มีการพิจารณาลักษณะทางเรขาคณิตเป็นระบบที่มีประโยชน์และสามารถนำไปใช้ได้ทั้งในทางปฏิบัติ แต่ระบบเหล่านี้จะต้องอ้างอิงตำแหน่งดวงตา ปาก จมูก โครงสร้างใบหน้าอยู่เสมอ จึงพบเห็นข้อเสียของวิธีเหล่านี้เมื่อนำไปประยุกต์กับการรู้จำใบหน้าแบบมีสิ่งบดบัง เพราะเป็นไปไม่ได้เลยที่จะเห็นองค์ประกอบของใบหน้าครบทุกส่วน ระบบการรู้จำใบหน้าที่มีการพิจารณาลักษณะทางเรขาคณิตในปัจจุบันไม่เป็นที่นิยมมากนัก แต่เหมาะกับการสร้างแบบจำลอง 3 มิติแทน

2.1.3 วิธีการโครงข่ายประสาท (Neural network approach)

โครงข่ายประสาทเทียม (Artificial Neural Networks) หรือที่มักจะเรียกสั้น ๆ ว่า โครงข่ายประสาท (Neural Networks หรือ Neural Net) เป็นหนึ่งในเทคนิคของการทำเหมืองข้อมูล (Data

Mining) คือโมเดลทางคณิตศาสตร์ สำหรับประมวลผลสารสนเทศด้วยการคำนวณแบบคอนเนคชันนิสต์ (Connectionist) เพื่อจำลองการทำงานของเครือข่ายประสาทในสมองมนุษย์ ด้วยวัตถุประสงค์ที่จะสร้างเครื่องมือซึ่งมีความสามารถในการเรียนรู้การจดจำรูปแบบ (Pattern Recognition) และการสร้างความรู้ใหม่ (Knowledge Extraction) เช่นเดียวกับความสามารถที่มีในสมองมนุษย์ โครงข่ายประสาทเทียมมีโครงสร้างแตกต่างจากโครงข่ายในสมอง แต่ก็ยังเหมือนสมองในแง่ที่ว่าโครงข่ายประสาทเทียม คือการรวมกลุ่มแบบขนานของหน่วยประมวลผลย่อย ๆ และการเชื่อมต่อนี้เป็นส่วนสำคัญที่ทำให้เกิดสติปัญญาของโครงข่าย เมื่อพิจารณาขนาดแล้วสมองมีขนาดใหญ่กว่าโครงข่ายประสาทเทียมอย่างมาก

Patel [15] เพิ่มประสิทธิภาพการรู้จำใบหน้าแบบมีสิ่งบดบังด้วยวิธีโครงข่ายประสาท โดยใช้วิธีการหาใบหน้าที่มีสิ่งบดบังและการสร้างคืนภาพด้วยวิธีของ Viola-jones จากนั้นใช้วิธีการวิเคราะห์องค์ประกอบหลักแบบถ่วงน้ำหนักอย่างรวดเร็ว (fast weighted PCA) ผสมกลับโครงข่ายประสาท (Neural Network: NN) สำหรับการรู้จำใบหน้า ผลการทดลองพบว่า ให้อัตราการเรียนรู้สูงมากเมื่อเทียบกับกับการเรียนรู้แบบ PCA และ LDA ปกติ

อย่างไรก็ตามการรู้จำโดยโครงข่ายประสาทเทียมมีข้อดีคือ เป็นการเรียนรู้แบบมีคนสอน (supervise learning) สอนให้โครงข่ายให้ทำการรู้จำสิ่งที่ป้อนเข้ามากับสิ่งที่ต้องการที่ต้องการเป็นสิ่งเดียวกัน โครงข่ายก็จะทำการปรับค่าสิ่งต่างๆที่อยู่ในชั้นที่ซ่อนอยู่ภายในโครงข่าย (hidden layer) ส่วนข้อเสียคือ มีโครงสร้างที่ซับซ้อนมากถ้าหากมีจำนวนข้อมูลมากและ ใช้ในการเรียนรู้นานมากหากมีข้อมูลขนาดใหญ่ ข้อเสียอีกข้อคือจำนวนชั้นที่ซ่อนอยู่ภายใน หรือจำนวนโหนด (neural node) นั้น เป็นสิ่งที่คาดเดาไม่ได้เลยว่าจะมีจำนวนเท่าไรจึงจะเหมาะสมกับระบบหนึ่งระบบ

2.1.4 วิธีการทางสถิติ (Statistical approach)

วิธีการทางสถิติประกอบด้วยวิธีการเทียบเทมเพลต (Template Matching) ซึ่งเป็นพื้นฐานของระบบ โดยภาพที่ใช้สอนและภาพที่ทดสอบจะถูกเปรียบเทียบกัน (Match) โดยการวัดความสัมพันธ์ระหว่างภาพที่ใช้สอนกับภาพที่ใช้ทดสอบ อย่างไรก็ตามวิธีการทางสถิติยังประกอบไปด้วย การวิเคราะห์องค์ประกอบหลัก (PCA) วิเคราะห์การแบ่งแยกเชิงเส้น (LDA) ทั้ง 2 ทฤษฎีนี้เป็นทฤษฎีพื้นฐานที่สำคัญ ในการพัฒนาวิธีการรู้จำใบหน้าด้วยวิธีการสถิติมาจนถึงปัจจุบัน การวิเคราะห์องค์ประกอบหลัก 2 มิติ (2DPCA) การวิเคราะห์การแบ่งแยกเชิงเส้น 2 มิติ (2DLDA) เป็น

การประยุกต์มาจาก PCA และ LDA ในหนึ่งมิติ รายละเอียดของเนื้อหาและทฤษฎีจะกล่าวในหัวข้อต่อไป

วิธีการเปรียบเทียบเทมเพลต [16] นี้ทำงานโดยหาความสัมพันธ์ของภาพ 2 ภาพ โดยตรงเมื่อมีภาพขนาดเดียวกัน มืองค์ประกอบของใบหน้าวางอยู่ตรงกัน และมีการส่องสว่างของแสงในทิศเดียวกัน จะทำให้ผลการรู้จำอย่างมีประสิทธิภาพมากที่สุด แต่ถ้าภาพที่เข้ามาทดสอบนั้นเปลี่ยนแปลงไปจากเดิมจะทำให้ประสิทธิภาพนั้นเปลี่ยนไปจากเดิม ดังนั้นต้องมีการประมวลผลก่อน (Preprocessing) เพื่อให้ภาพเข้าสู่ตำแหน่งหรือการส่องสว่างใกล้เคียงกัน

เนื่องจากการเปรียบเทียบเทมเพลตนี้ทำงานโดยการหาความสัมพันธ์ของภาพ 2 ภาพ จึงมีข้อจำกัดมาก ดังนั้นวิธีการหาพีเจอร์ที่แท้จริงของภาพและลดขนาดของภาพได้คือวิธีที่ใช้การแปลงเชิงเส้น วิธีนี้สามารถตัดข้อมูลที่เกินผลกระทบบนใบหน้าอันได้แก่ การส่องสว่างและการแสดงออกทางอารมณ์บนใบหน้าได้

Kumeechai [3] ทำการวิจัยระบบรู้จำใบหน้าคนร้ายโดยด้วยการตรวจสอบสีผิวบุคคลด้วย ไอเก็นเฟส โดยนำรูปผู้ก่อการร้ายที่มีการนำรูปมาลงประกาศจับแล้ว นำมาใช้เป็นฐานข้อมูลและเป็นใบหน้าสีสองมิติระบบสี HSV และวิธีการ Haar-like feature เพื่อให้ได้ใบหน้าผู้ก่อการร้าย ต่อจากนั้นใช้วิธีการไอเก็นเฟสหรือ PCA เพื่อให้ได้คุณลักษณะเด่นบนใบหน้าเพื่อให้ได้เป็นไอเก็นเวกเตอร์ หลังจากได้ไอเก็นเวกเตอร์ของภาพทดลองและภาพทดสอบทั้งหมดนำมาหาระยะทางแบบยุคลิด สุดท้ายนำผลมาเปรียบเทียบกับวิธีการ LDA, ICA และ geometric based method ผลการศึกษาพบว่า การรู้จำใบหน้าด้วยไอเก็นเฟส ให้อัตราความถูกต้องมากที่สุด รองลงมาคือ LDA และ ICA ตามลำดับ ส่วนวิธีทางลักษณะเรขาคณิต (geometric based method) ให้อัตราความถูกต้องต่ำที่สุด

Kumar [17] ได้เปรียบเทียบข้อดีและข้อเสียของ PCA และ LDA ผลการศึกษาพบว่า PCA นั้นเป็นวิธีที่ง่ายที่สุด มีประสิทธิภาพที่สุด แต่ในการนำภาพไปฝึกสอนนั้นใช้เวลานานกว่า LDA และ PCA นั้นอ่อนไหวง่ายกับพวกแสง (Illumination) และตำแหน่งของใบหน้า (pose) เมื่อชุดฝึกอบรมมีขนาดเล็ก PCA สามารถทำงานได้ดีกว่า LDA แต่เมื่อจำนวนตัวอย่างมีขนาดใหญ่ LDA จะมีประสิทธิภาพดีกว่า PCA ส่วน LDA นั้นสามารถแก้ปัญหาเรื่องแสงได้ (Illumination) ดีกว่า และยังสรุปอีกว่า LDA กับ PCA นั้น มีผู้วิจัยหลายคนพยายามที่จะเปรียบเทียบทั้งที่สองวิธีนี้แต่ไม่มีความสัมพันธ์กันเลย ดังนั้นการนำสองวิธีนี้มาผสมผสานกันจึงจะเกิดผลลัพธ์ที่ดีกว่า

Karungaru [18] ได้ทำการวิจัยเกี่ยวกับการรู้จำใบหน้าโดยใช้วิธีการจับคู่กับแม่แบบ (Template Matching) เทียบกับการวัดระยะทาง (Distance measure) โดยมีหลักการคือ การนำภาพใบหน้ามาหาทีโซน (T-zone) ของใบหน้าก่อน จากนั้นจึงทำการเปรียบเทียบกับรูปในฐานข้อมูลจากการศึกษาพบว่า การเปรียบเทียบแบบเทมเพลตมีอัตราความถูกต้องสูงกว่าแบบวัดระยะทางในทางกลับกันการวัดแบบระยะทางนั้นทำงานได้เร็วกว่าในอัตราความถูกต้องที่พอกัน

2.2 การวิเคราะห์ส่วนประกอบหลัก 2 มิติ (Two-dimensional Principal Component Analysis: 2DPCA)

การวิเคราะห์ส่วนประกอบหลัก [4] เป็นวิธีเชิงสถิติ ที่ใช้เพื่อลดจำนวนของตัวแปรในชุดข้อมูล แต่ยังคงไว้ซึ่งเอกลักษณ์ของข้อมูล หลักการของ 2DPCA คล้ายกับ PCA เพราะเป็นอัลกอริทึมที่พัฒนามาจาก PCA แตกต่างกันตรงที่การทำงานในรูปแบบของเมทริกซ์ 2 มิติ ในขณะที่ PCA นั้นทำงานในรูปแบบของเวกเตอร์ 1 มิติ ฉะนั้น 2DPCA ไม่ต้องเสียเวลาแปลงรูปให้เป็นเวกเตอร์อีก

การคำนวณของวิธี 2DPCA คือ การคำนวณหาเมทริกซ์ความแปรปรวนร่วม (Covariance matrix) จากภาพที่ใช้ในการสอนทั้งหมดที่นำมาคำนวณหาเวกเตอร์เฉพาะ (Eigenvector) ที่สัมพันธ์กับค่าเฉพาะ (Eigenvalue) เรียงลำดับจากมากไปน้อยแล้วจึงนำมาฉายภาพ (Projection) ใบหน้าให้มีขนาดเล็กลง เช่น ภาพใบหน้าที่มีขนาด 227×227 จุดภาพ เมื่อเลือกเวกเตอร์เฉพาะ 27 มิติของสตรัม คุณสมบัติใหม่ของภาพใบหน้าที่ใหม่จะมีขนาดเป็น 227×27 จุดภาพ ซึ่งจะทำให้เห็นจำนวนมิติของสตรัมที่มีขนาดเล็กลง

2.2.1 หลักการคำนวณ

กำหนดให้ ภาพที่ใช้สอนมี M ภาพ และ A คือภาพแต่ละภาพมีขนาด $m \times n$ จุดภาพ จะได้ A_1, A_2, \dots, A_M จากนั้นคำนวณหาค่าเฉลี่ยของภาพใบหน้า (Mean Face) ดังสมการที่ 2.1

$$\bar{A} = \frac{1}{M} \sum_{i=1}^M A_i \quad (2.1)$$

กำหนดให้ C คือ เมทริกซ์ความแปรปรวนร่วม (Covariance matrix) สามารถคำนวณดังสมการที่ 2.2

$$C = \frac{1}{M-1} \sum_{i=1}^M (A_i - \bar{A})^T (A_i - \bar{A}) \quad (2.2)$$

คำนวณหาเวกเตอร์เฉพาะที่สมนัยกับค่าเฉพาะ (λ) จาก เมทริกซ์ความแปรปรวนร่วม สามารถคำนวณ ดังสมการที่ 2.3

$$Cv = \lambda v \quad (2.3)$$

เรียงลำดับเวกเตอร์เฉพาะที่สมนัยกับค่าเฉพาะ จากมากไปน้อยแล้วเลือกแกนการฉายมา d มิติ ดังนั้นจะได้เมทริกซ์ของเวกเตอร์เฉพาะ ดังสมการที่ 2.4

$$v = [v_1, v_2, v_3, \dots, v_d] \quad (2.4)$$

2.2.2 การหาคุณลักษณะของภาพ (feature extraction)

เวกเตอร์ที่เหมาะสมที่ใช้ในการฉายภาพคือ $v = [v_1, v_2, v_3, \dots, v_d]$ ถูกนำมาใช้หาคุณลักษณะของภาพ ฉะนั้นจึงเลือกแกนในการฉายภาพเพียง 1 ถึง d เท่านั้น เมื่อกำหนดให้ A เป็นภาพตัวอย่าง ดังนั้นการหาคุณลักษณะของภาพ A ดังสมการที่ 2.5

$$Y_d = Av_d \quad (2.5)$$

กำหนดให้ เวกเตอร์ใหม่ที่ได้จากการฉาย $Y = [Y_1, Y_2 \dots Y_d]$ มีชื่อว่าเวกเตอร์ส่วนประกอบหลัก (Principle Component Vector) ของภาพตัวอย่าง A เมื่อนำเวกเตอร์ของส่วนประกอบหลักมารวมกับเมทริกซ์ Y ที่มีขนาด $m \times d$ มิติ ดังนั้นเมทริกซ์ Y มีชื่อว่าเมทริกซ์คุณลักษณะ (Featured Matrix) ภาพ A

2.2.3 การสร้างคืนภาพ (reconstructed image)

การสร้างคืนภาพ สามารถนำเอาคุณลักษณะเมทริกซ์ (Featured Matrix) และแกนการฉายที่เหมาะสมที่สุดมาสร้างเป็นภาพใบหน้าใหม่ได้ตามขั้นตอนดังต่อไปนี้

กำหนดให้ A เป็นภาพ โดยที่ Y เป็นคุณลักษณะเมทริกซ์ (Featured Matrix) ภาพของภาพ A และ v_d เป็นแกนการฉายที่เหมาะสมที่สุด ดังนั้นจากสมการที่ 2.6 จะได้การสร้างคืนภาพ ดังสมการที่ 2.6

$$\tilde{A} = Yv_d^T \quad (2.6)$$

ในการสร้างคืนภาพใบหน้าใหม่ ถ้าเลือกจำนวนแกนการฉาย d เท่ากับ n จะได้ภาพที่สร้างใหม่เหมือนกับภาพ A แต่ถ้าเลือก d น้อยกว่า n ภาพ \tilde{A} จะมีความเหมือนกับภาพ A น้อยลงตามลำดับ

2.3 การวิเคราะห์การจำแนกเชิงเส้น 2 มิติ (Two-dimensional Linear Discriminant Analysis: 2DLDA)

การวิเคราะห์การจำแนกเชิงเส้นมีความเกี่ยวข้องกับการวิเคราะห์องค์ประกอบหลัก โดยทั้งสองวิธีมีพื้นฐานมาจากการใช้สมการเชิงเส้นหารูปแบบของข้อมูล โดยการวิเคราะห์ส่วนประกอบหลัก มีจุดมุ่งหมายเพื่อหาค่าที่น้อยที่สุดระหว่างเวกเตอร์ข้อมูลเดิมกับเวกเตอร์ข้อมูลใหม่ที่ลดจำนวนมิติลงแล้ว โดยการวิเคราะห์ส่วนประกอบหลักจะไม่ได้แยกแยะความแตกต่างของข้อมูลแต่ละประเภทแต่จะมองข้อมูลทุกตัวเป็นประเภทเดียวกันทั้งหมดจัดเป็นการเรียนรู้แบบไม่มีคนสอน (unsupervised learning) ประเภทหนึ่ง ในขณะเดียวกัน การวิเคราะห์การจำแนกเชิงเส้น เป็นวิธีการหาค่าสูงสุดของสัดส่วนระหว่าง ความแปรปรวนภายในกลุ่ม และ ความแปรปรวนระหว่างกลุ่ม ซึ่งมีจุดมุ่งหมายคือ จะหาจุดของข้อมูลประเภทเดียวกันให้มีความแตกต่างกันน้อยที่สุด จัดเป็นการเรียนรู้แบบมีคนสอนประเภทหนึ่ง (supervised learning)

การวิเคราะห์การจำแนกเชิงเส้น 2 มิติ [4] สามารถนำเมทริกซ์ภาพมาคำนวณหาเมทริกซ์การกระจายภายในกลุ่มและ เมทริกซ์การกระจายระหว่างกลุ่มแล้วนำมาคำนวณหาเวกเตอร์เฉพาะที่สมนัยกับค่าเฉพาะจากมากไปน้อยตามลำดับ จากนั้นนำภาพมาฉายลงบนเวกเตอร์เฉพาะจะได้คุณลักษณะมิติของคอลัมน์ภาพน้อยลง

2.3.1 หลักการคำนวณ

กำหนดให้ภาพที่ใช้ฝึกสอนทั้งหมดมี M ภาพ k คือจำนวนกลุ่ม A คือภาพที่มีขนาด $m \times n$ จุดภาพ กำหนดให้ \bar{A}_k คือภาพเฉลี่ยของแต่ละกลุ่ม ส่วน \bar{A} คือ ภาพเฉลี่ยของภาพที่ใช้ฝึกสอนทั้งหมด และ N_k คือ จำนวนภาพในแต่ละกลุ่ม

คำนวณเมทริกซ์การกระจายระหว่างกลุ่ม (between classes scatter matrix: S_B) สามารถคำนวณหาโดยการหาผลรวมเมทริกซ์ความแปรปรวนของความแตกต่างระหว่างค่าเฉลี่ยรวมกับค่าเฉลี่ยของแต่ละกลุ่ม ดังสมการที่ 2.7

$$S_B = \sum_{k=1}^k N_k (\bar{A}_k - \bar{A})^T (\bar{A}_k - \bar{A}) \quad (2.7)$$

การคำนวณเมทริกซ์การกระจายภายในกลุ่ม (within classes scatter matrix: S_W) เป็นการวัดค่าการกระจายระหว่างภาพภายในกลุ่มแต่ละกลุ่มเมื่อมีภาพทั้งหมด k กลุ่ม ดังนั้นการคำนวณจะใช้การหาผลรวมของเมทริกซ์ความแปรปรวนในแต่ละกลุ่ม ดังสมการที่ 2.8

$$S_W = \sum_{k=1}^k \sum_{A_i \in k} N_k (A_i - \bar{A}_k)^T (A_i - \bar{A}_k) \quad (2.8)$$

การคำนวณหาเวกเตอร์เฉพาะที่สมนัยกับค่าเฉพาะ (λ) จากเมทริกซ์ความแปรปรวนร่วม ดังสมการที่ 2.9

$$S_B v = \lambda S_W v \quad (2.9)$$

เมื่อเรียงลำดับเวกเตอร์เฉพาะที่สมนัยกับค่าเฉพาะจากมากไปน้อยแล้วเลือกเวกเตอร์เฉพาะที่สัมพันธ์กับค่าเฉพาะที่ไม่เท่ากับศูนย์ ดังนั้นจะได้เมทริกซ์ใหม่ของเวกเตอร์เฉพาะดังสมการที่ 2.10

$$v = [v_1, v_2, v_3, \dots, v_d] \quad (2.10)$$

2.3.2 การหาคุณลักษณะของภาพ (feature extraction)

เวกเตอร์ที่เหมาะสมที่ใช้ในการฉายภาพคือ $v = [v_1, v_2, v_3, \dots, v_d]$ ถูกนำมาใช้หาคุณลักษณะของภาพ ฉะนั้นจึงเลือกแกนในการฉายภาพเพียง 1 ถึง d เท่านั้น เมื่อกำหนดให้ A เป็นภาพตัวอย่าง ดังนั้นการหาคุณลักษณะของภาพ A ดังสมการที่ 2.11

$$Y_d = A v_d \quad (2.11)$$

2.3.3 การสร้างคืนภาพ (reconstructed image)

การสร้างคืนภาพ สามารถนำเอาคุณลักษณะเมทริกซ์ และแกนการฉายที่เหมาะสมที่สุดมาสร้างเป็นภาพใบหน้าใหม่ได้ตามขั้นตอนดังต่อไปนี้

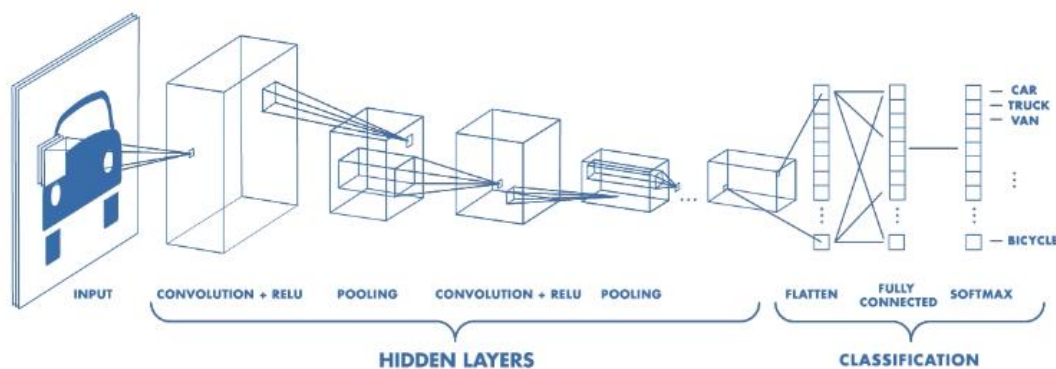
กำหนดให้ A เป็นภาพ โดยที่ Y เป็นคุณลักษณะเมทริกซ์ภาพ ของภาพ A และ v_d เป็นแกนการฉายที่เหมาะสมที่สุด ดังนั้นจากสมการที่ 2.12 จะได้การสร้างคืนภาพคือ

$$\tilde{A} = Yv_d^T \quad (2.12)$$

ในการสร้างคืนภาพใบหน้าใหม่ ถ้าเลือกจำนวนแกนการฉาย d เท่ากับ n จะได้ภาพที่สร้างใหม่เหมือนกับภาพ A แต่ถ้าเลือก d น้อยกว่า n ภาพ \tilde{A} จะมีความเหมือนกับภาพ A น้อยลงตามลำดับ

2.4 โครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Network: CNN)

โครงข่ายประสาทเทียมแบบสังวัตนาการ เป็นหนึ่งในโครงข่ายประสาทเชิงลึกที่ได้รับความนิยมสูงมาก [19] จัดเป็นการเรียนรู้แบบมีการสอน (supervisor learning) ด้วยการติดป้ายกำกับ (label) ทุกรูปฝึกสอน เป็นการจำลองการเรียนรู้ในมนุษย์ได้ว่าวัตถุที่เห็นอยู่ตรงหน้าคืออะไร จะต้องเกิดจากการจดจำซ้ำ ซึ่งสำหรับเครื่องประมวลผล (computer) แล้วการจดจำนั้นแตกต่างออกไป เพราะในโลกของเครื่องประมวลผลมีเพียงแค่ตัวเลขเท่านั้น ดังนั้นจะเห็นได้ว่าการที่จะสอนให้เครื่องประมวลผลจดจำและจำแนกภาพได้นั้นต้องมีอัลกอริทึมเข้ามาช่วย



รูปที่ 2.1 ตัวอย่างชั้นโครงข่ายประสาทเทียมแบบสังวัตนาการที่มีหลายชั้น

[ที่มา: <https://www.mathworks.com/solutions/deep-learning/convolutional-neural-network.html#howitworks> Accessed: July 4, 2018]

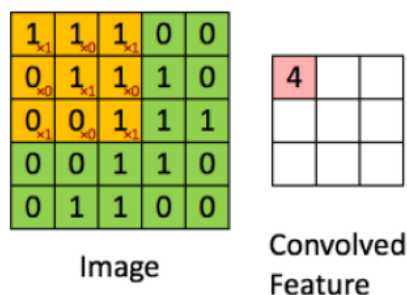
จากรูปที่ 2.1 แสดงโครงสร้างโครงข่ายประสาทเทียมแบบสังวัตนาการ มีส่วนประกอบหลักอยู่ 2 ส่วน ที่มีรายละเอียดดังต่อไปนี้

1. ส่วนของชั้นที่ซ่อนอยู่ (Hidden Layers) หรือ การแยกคุณสมบัติ (Feature extraction) โดยในส่วนนี้จะมีส่วนย่อยที่ทำงานร่วมกันอีก 3 ส่วน มีรายละเอียดดังต่อไปนี้

- ชั้นสังวัตนาการ (Convolution Layer) ซึ่งเป็นขั้นตอนการตรวจรูป (scan) จากรูปนำเข้าไปโดยจะสร้างตัวกรองขึ้นมาในรูปแบบของเมทริกซ์ จากนั้นจะทำการตรวจสอบเพื่อหาลักษณะหรือส่วนประกอบต่างๆ ของรูปภาพแต่ละรูปตามที่อธิบายไว้ในข้อที่ 2.4.1
 - ชั้นกระตุ้นฟังก์ชัน (Activation Function Layer) เป็นขั้นตอนที่ทำให้ CNN เกิดการเรียนรู้ข้อมูลที่ได้รับมาจากขั้นตอนการตรวจสอบรูปภาพก่อนหน้า โดยข้อมูลจะถูกแปลงมาอยู่ในรูปของเส้นที่ไม่เชิงเส้น (non-linear) โดยจะมีการใช้ตัวอย่างฟังก์ชันเรคตีไฟต์เชิงเส้น (Rectified Linear Unit: ReLU) เพื่อเข้ามาช่วยเพิ่มประสิทธิภาพของ CNN ตามที่ได้อธิบายการทำงานไว้ในข้อที่ 2.4.2
 - ชั้นพูลลิ่ง (Pooling Layer) เป็นการทำการลดขนาดของรูปภาพให้เล็กลง เพื่อให้การทำงานของ CNN มีความรวดเร็วขึ้น และยังลดการใช้พื้นที่ความจำลงได้อีกด้วย ตามที่ได้อธิบายการทำงานไว้ในข้อที่ 2.4.3
2. ชั้นการจำแนกประเภท (Classification Layer) หลังจากผ่านชั้นสังวัตนาการทั้งหมดมาแล้ว ต่อมาในส่วนของชั้นการจำแนกประเภท ที่ประกอบไปด้วยชั้นการเชื่อมโยงเต็มรูปแบบ (Fully Connected Layer) จะเชื่อมกับทุกโหนดของชั้นก่อนหน้า ซึ่งจะรับข้อมูลได้เพียง 1 มิติ โดยแปลงข้อมูลจากหลายมิติให้เป็น 1 มิติ (Flatten) เสียก่อน

2.4.1 การทำงานของชั้นสังวัตนาการ (Convolution Layer)

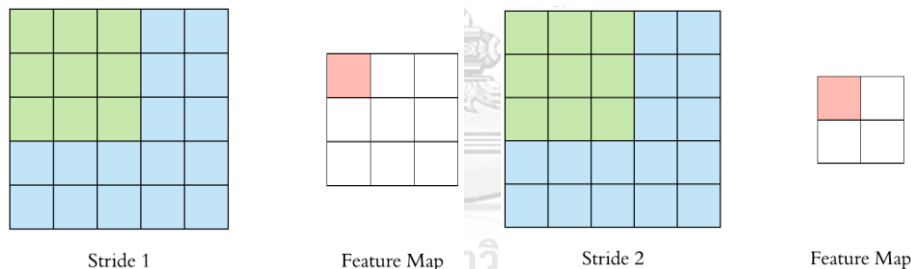
เป็นขั้นตอนการกราดตรวจรูปจากรูปนำเข้าไปโดยจะมีการสร้างตัวกรองขึ้นมาในรูปแบบของเมทริกซ์ จากนั้นจะทำการตรวจสอบเพื่อหาคุณลักษณะเด่นของรูปภาพแต่ละรูป ด้วยการขับเคลื่อนต่อไปเรื่อยๆ บนเมทริกซ์ของภาพต้นแบบ ดังรูปที่ 2.2 ผลลัพธ์จากรวบรวมข้อมูลภาพจะถูกเรียกว่า ผังคุณลักษณะ (Feature Map หรือ Convolved Feature) ในขั้นของรวบรวมข้อมูลภาพ มักจะตามด้วยฟังก์ชันกระตุ้น ซึ่งเป็นฟังก์ชันที่ไม่ใช่แบบเชิงเส้น (Non-linear Function)



รูปที่ 2.2 แสดงขั้นตอนการทำสังวัตนาการระหว่างภาพนำเข้า

[ที่มา: <https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148> Accessed: July 4, 2018]

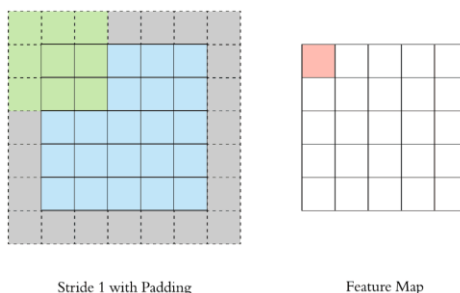
การขยับเคลื่อน (stride) ของตัวกรองจากรูปที่ 2.3 เป็นตัวกำหนดว่าจะเลื่อนตัวกรอง (filter) ไปด้วยครั้งละ 1 หรือ 2 ก้าว สามารถกำหนดค่าของการขยับเคลื่อนให้มากขึ้นได้ ถ้าหากต้องการให้การคำนวณหาคุณลักษณะมีพื้นที่ทับซ้อนกันน้อยลง อย่างไรก็ตามการกำหนดค่าของการขยับเคลื่อนที่มากขึ้นจะทำให้เราได้ฟังก์ชันลักษณะที่มีขนาดเล็กลง



รูปที่ 2.3 ตัวอย่างการขยับเคลื่อนของตัวกรองครั้งละ 1 กับ 2 ช่อง

[ที่มา: <https://towardsdatascience.com/applied-deep-learning-part-4-convolutional-neural-networks-584bc134c1e2> Accessed: July 4, 2018]

การขยายขอบภาพ (Padding) จากรูปที่ 2.4 พื้นที่สี่เเทรอบๆ เมทริกซ์ภาพพื้นที่เหล่านี้เป็นพื้นที่ที่เราเติมเข้าไป โดยอาจจะเติม 0 หรือค่าต่างๆ เข้าไป เพื่อให้เวลาในสังวัตนาการได้ฟังก์ชันลักษณะมีขนาดเดียวกับรูปต้นฉบับ



รูปที่ 2.4 การ padding size เพื่อเก็บรายละเอียดขอบภาพ

[ที่มา: <https://towardsdatascience.com/applied-deep-learning-part-4-convolutional-neural-networks-584bc134c1e2> Accessed: July 4, 2018]

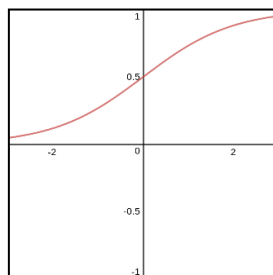
2.4.2 การทำงานของฟังก์ชันกระตุ้น (Activation Function)

เป็นขั้นตอนที่ทำให้โครงข่ายประสาทเทียมแบบสังวัตนาการเกิดการเรียนรู้ข้อมูลที่ได้รับมาจากขั้นตอนการกราดตรวจรูปภาพก่อนหน้า โดยข้อมูลจะถูกแปลงมาอยู่ในรูปที่ไม่ใช่แบบเชิงเส้น (non-linear) โดยจะใช้ฟังก์ชันกระตุ้นที่นิยมใช้ด้านโครงข่ายประสาทเทียมแบบสังวัตนาการ และการเรียนรู้เชิงลึก [20] เพื่อเข้ามาช่วยเพิ่มประสิทธิภาพของโครงข่ายประสาท ฟังก์ชันกระตุ้น 4 ตัวที่น่าสนใจมีรายละเอียด ดังต่อไปนี้

1. ฟังก์ชันซิกมอยด์ (Sigmoid Function) จากสมการที่ 2.13 ให้ค่าผลลัพธ์ตั้งแต่ 0 ถึง 1 เป็นค่าต่อเนื่องเหมาะสมกับการใช้ค่าความน่าจะเป็น (probability) จากรูปที่ 2.5 จะเห็นได้ว่าแกน y คือความน่าจะเป็นของคลาสนั้น โดยจุดกลางของกราฟคือ 0.5 ทำให้แบ่งได้ 2 คลาส คือ เป็นคลาสนั้น และไม่เป็นคลาสนั้น ซึ่งไม่เหมาะกับการแยกประเภทแบบหลายคลาส

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

$$f(x) = \frac{1}{1 + \exp^{-x}} \quad (2.13)$$



รูปที่ 2.5 กราฟฟังก์ชันซิกมอยด์ (Sigmoid Function)

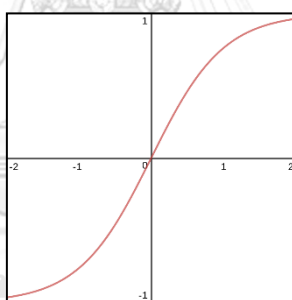
[ที่มา: <http://codeonthehill.com/machine-learning-2-neural-network/>

Accessed: July 4, 2018]

ปัญหาการหนึ่งของฟังก์ชันนี้คือ การเปลี่ยนแปลงค่า Y มีผลต่อค่า X มีค่าน้อยลงมาก ซึ่งทำให้การทำแกรเดียนต์หายไป (Gradient Descent) อาจส่งผลให้โครงข่ายประสาทติดขัดได้ขณะเวลาฝึกสอนแบบจำลอง การแก้ไขปัญหานี้แก้ด้วยฟังก์ชันกระตุ้นแทนเจนไฮเพอร์โบลิก (Hyperbolic tangent function)

2. ฟังก์ชันแทนเจนไฮเพอร์โบลิก (Hyperbolic tangent function) จากสมการที่ 2.14 เป็นฟังก์ชันที่มีลักษณะคล้ายฟังก์ชันมอยด์ที่ให้ค่าผลลัพธ์ที่ -1 ถึง 1 ทำให้การทำแกรเดียนต์หายไปน้อยกว่า แต่ยังคงผลลัพธ์ไว้ตลอดเวลา ส่งผลให้โครงข่ายประสาททำงานหนักตลอดเวลา ปัญหานี้แก้ไขด้วยฟังก์ชันเรกติไฟด์เชิงเส้น (Rectified Linear Unit: ReLU)

$$f(x) = \frac{e^x - e^{-x}}{(e^x + e^{-x})} \quad (2.14)$$



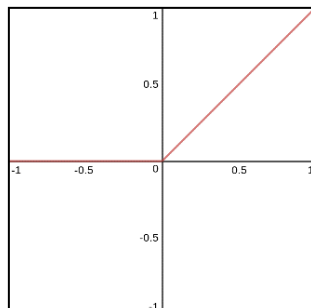
รูปที่ 2.6 กราฟฟังก์ชันแทนเจนไฮเพอร์โบลิก (Hyperbolic tangent function)

[ที่มา: <http://codeonthehill.com/machine-learning-2-neural-network/>

Accessed: July 4, 2018]

3. ฟังก์ชันเรกติไฟด์เชิงเส้น (Rectified Linear Unit: ReLU) จากสมการที่ 2.15 เป็นฟังก์ชันกระตุ้นที่ได้รับความนิยมสูงมากงานด้านโครงข่ายประสาทเทียมแบบสังวัตนาการ ที่นิยมใช้ในการแยกประเภทแบบหลายคลาส (multi-classification) ในขณะที่ฟังก์ชันซิงมอยด์จะนิยมใช้ในการแยกสองคลาส (binary classification) เมื่อผลลัพธ์ค่าเป็น ลบ ฟังก์ชันจะปรับให้เป็น 0 ซึ่งเป็นผลลัพธ์ที่ทำให้โครงข่ายประสาทไม่ทำงานตลอดเวลา ทำให้การฝึกรวดเร็ว ยกเว้นค่าเป็น บวก จะให้ผลลัพธ์เป็นไปตามค่าที่ส่งเข้ามา

$$f(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1, & \text{if } x \geq 0 \end{cases} \quad (2.15)$$



รูปที่ 2.7 กราฟฟังก์ชันเรกติไฟด์เชิงเส้น (Rectified Linear Unit: ReLU)

[ที่มา: <http://codeonthehill.com/machine-learning-2-neural-network/>

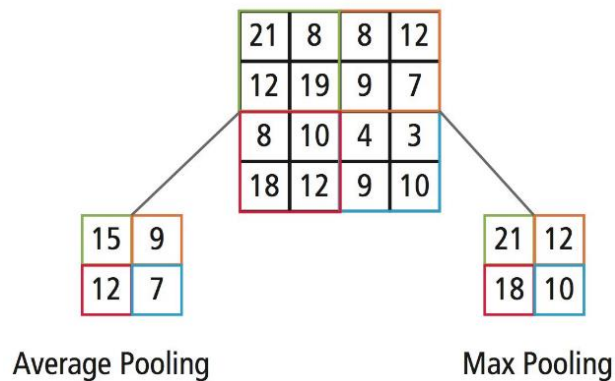
Accessed: July 4, 2018]

4. ฟังก์ชันค่าสูงสุดอย่างอ่อน (Softmax Function) จากสมการที่ 2.16 ที่เป็นประโยชน์หลักของการใช้ฟังก์ชันค่าสูงสุดอย่างอ่อน คือการนำค่าข้อมูลส่งออกจากชั้นการเชื่อมโยงเต็มรูปแบบ มาแปลงเป็นช่วงความน่าจะเป็นช่วงจะเป็น 0 ถึง 1 และผลรวมของความน่าจะเป็นทั้งหมดจะเท่ากับหนึ่ง หากฟังก์ชันค่าสูงสุดอย่างอ่อน ที่ใช้สำหรับการจำแนกหลายแบบจะส่งคืนความน่าจะเป็นของแต่ละคลาสและคลาสเป้าหมายจะมีความน่าจะเป็นสูง

$$f(x) = \frac{e^{x_j}}{\sum_{i=1}^k e^{x_i}} \quad (2.16)$$

2.4.3 ชั้นพูลลิง (Pooling Layer)

หลังจากการทำสังวัตนาการเสร็จแล้วส่วนใหญ่จะตามด้วยวิธีการพูลลิง ทำการลดขนาดของรูปภาพให้เล็กลงเพื่อให้การทำงานของ CNN มีความรวดเร็วมากขึ้น และยังลดการใช้พื้นที่ความจำลงได้อีกด้วย ซึ่งมีสองประเภทหลักที่นิยมกันคือ พูลลิงสูงสุด (max pooling) และ พูลลิงเฉลี่ย (average pooling) และการใช้ค่าพูลลิงสูงสุดของคุณลักษณะที่เด่นที่สุดจะถูกเก็บไว้คำนวณต่อ

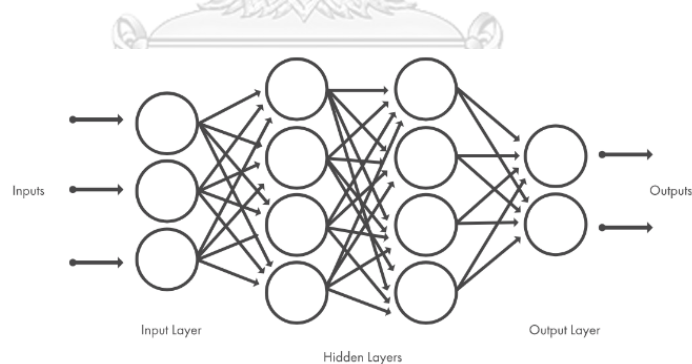


รูปที่ 2.8 แสดงการทำพูลลิ่งเฉลี่ยและพูลลิ่งสูงสุด ขนาด 2×2

[ที่มา: <http://www.davidsbatista.net/blog/2018/03/31/SentenceClassificationConvNets/>
Accessed: July 4, 2018]

2.4.4 ชั้นการเชื่อมโยงเต็มรูปแบบ (Fully Connected Layer)

เป็นชั้นสุดท้ายของสังวัตนาการ ที่จะเป็นการเชื่อมโยงเต็มรูปแบบ ดังรูปที่ 2.9 หลังจากการประกอบกันของชั้นสังวัตนาการและชั้นพูลลิ่ง โดยชั้นนี้ประกอบด้วยชั้นย่อยๆ ที่มีโหนดอยู่จำนวนหนึ่ง ซึ่งโหนดแต่ละตัวจะมีเส้นเชื่อมกับโหนดทุกตัวก่อนหน้า โดยแต่ละโหนดจะมีการจำแนกประเภทด้วยฟังก์ชันกระตุ้นที่ได้รับความนิยมคือฟังก์ชันค่าสูงสุดอย่างอ่อน (Softmax Function)



รูปที่ 2.9 แสดงการเชื่อมโยงเต็มรูปแบบโหนดแต่ละโหนด

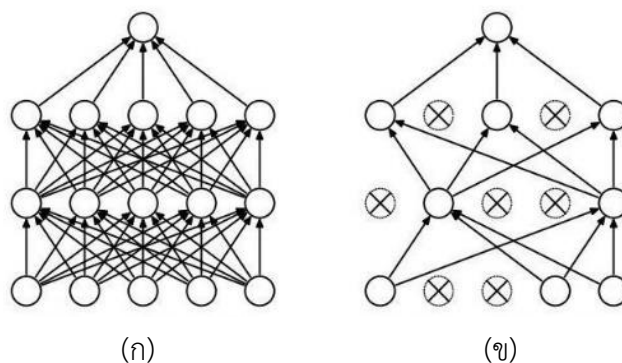
[ที่มา: <https://www.mathworks.com/solutions/deep-learning/convolutional-neural-network.html#howitworks> Accessed: July 4, 2018]

2.4.5 การปรับค่าไฮเพอร์พารามิเตอร์ (Fine tuning hyper-parameter)

การปรับค่าพารามิเตอร์ไฮเพอร์สำหรับโครงข่ายประสาท เพื่อให้เกินความแม่นยำสูงสุดเป็นเรื่องยาก เนื่องจากทำให้เวลาในการฝึกสอนช้าลงถ้าหากมีแบบจำลองขนาดใหญ่ ตัวอย่างของพารามิเตอร์ที่ต้องมีการปรับแต่งไม่ว่าจะเป็นแบบจำลองชนิดใดไหน ได้แก่ จำนวนรอบของการฝึกสอน (Number of epochs) จำนวนภาพต่อกลุ่มการฝึกสอน (Batch size) ความเร็วในการเรียนรู้ (Learning rate) ดังนั้นการปรับค่าพารามิเตอร์เหล่านี้สามารถหาได้จากการสุ่มที่ละค่า (randomized search) หรือทำตารางการค้นหา (Grid search)

2.5.6 สัญญาณตกหาย (Dropout)

เป็นเทคนิคหนึ่งซึ่งช่วยป้องกันการอิงกับข้อมูลเรียนรู้มากเกินไป (overfitting) กล่าวคือ ปัญหาการตรวจจับคุณลักษณะซ้ำๆ ของเทคนิคสัญญาณตกหายมาช่วยแก้ปัญหานี้ขึ้นอยู่กับหลักการสุ่มเพื่อที่จะกำจัดโหนดบางโหนดที่อยู่ในชั้นที่ซ่อนอยู่ เพื่อไม่ให้โหนดใดโหนดหนึ่งมีอิทธิพลมากเกินไป โดยเราสามารถปรับได้ตั้งแต่ 0.1–1 ถ้าหากสัญญาณตกหายร้อยละ 60 กล่าวคือเหลือโหนดที่ทำงานจริงอยู่ร้อยละ 40 ดังรูปที่ 2.10



รูปที่ 2.10 เปรียบเทียบสัญญาณตกหายของโครงข่ายประสาท (ก) แสดงโครงข่ายประสาทแบบปกติ

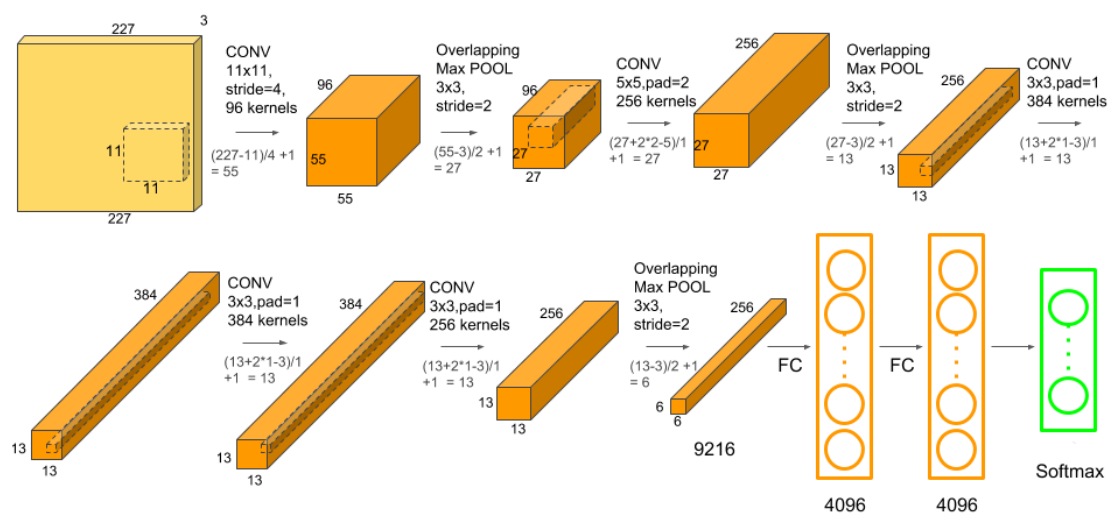
(ข) แสดงการลดลงโหนดที่มีการตัดเส้นเชื่อมของโครงข่ายประสาท

[ที่มา: <https://cv-tricks.com/cnn/understand-resnet-alexnet-vgg-inception/>

Accessed July 4, 2018]

2.6 โครงข่ายประสาทเทียมแบบสังวัตนาการของแบบจำลอง Alexnet

โครงข่ายประสาทเทียมแบบสังวัตนาการของแบบจำลอง Alexnet (Alexnet) [21] เป็นโครงข่ายประสาทเทียมแบบสังวัตนาการเชิงลึก (Deep Convolutional Neural Network) ที่มีการฝึกสอนรูปภาพมาแล้วกว่า 1.2 ล้านรูปจากฐานข้อมูลภาพที่ได้รับความนิยมคือฐานข้อมูล ImageNet แบบจำลอง Alexnet ประกอบด้วยชั้นสังวัตนาการ 5 ชั้น และชั้นการเชื่อมโยงเต็มรูปแบบ 3 ชั้น ในหนึ่งชั้นของการสังวัตนาการจะมีตัวกรองที่ไว้หาคุณลักษณะของรูปหลายตัว ดังรูปที่ 2.11



รูปที่ 2.11 แสดงโครงข่ายประสาทเทียมแบบสังวัตนาการของแบบจำลอง Alexnet

[ที่มา: <https://sushscience.wordpress.com/2016/12/04/understanding-alexnet/>

Accessed: July 10, 2019]

จากรูปที่ 2.11 แสดงโครงข่ายประสาทเทียมแบบสังวัตนาการของแบบจำลอง Alexnet ชั้นที่ 1 ของชั้นสังวัตนาการ ประกอบด้วย ตัวกรองขนาด 11 x 11 x 3 จุดภาพ จำนวน 96 ตัว ส่วนชั้นที่ 2 ตามด้วยการหาพหุคูณสูงสุด ส่วนชั้นที่ 4 และ 5 เป็นชั้นของสังวัตนาการแบบต่อเนื่อง และตามด้วยการหาพหุคูณสูงสุด ถัดมาเป็นชั้นการเชื่อมโยงเต็มรูปแบบ 2 ชั้น โดยที่ชั้นแรกจะเชื่อมกับทุกโหนดของชั้นก่อนหน้า ซึ่งจะรับข้อมูลได้เพียง 1 มิติ โดยแปลงข้อมูลจากหลายมิติ ให้กลายเป็น 1 มิติ (Flatten) จากนั้นจะถูกนำเข้าไปยังชั้นการเชื่อมโยงเต็มรูปแบบชั้นที่สอง โดยใช้ฟังก์ชันค่าสูงสุดอย่างอ่อนในการจำแนกข้อมูลส่งออกจากเดิมที่จำแนกวัตถุได้ 1,000 ชนิด แต่ผู้วิจัยได้ทำการประยุกต์กับฐานข้อมูล AR เป็น 117 คน และ ฐานข้อมูล GTAV เป็น 44 คน

หลักการคำนวณหาขนาดของข้อมูลส่งออกจากการส่งออกของแต่ละชั้นคำนวณได้ 2 กรณี คือ กรณีที่มีการขยายขอบภาพ และกรณีที่ไม่มีการขยายขอบภาพ

กรณีที่ไม่มีการขยายขอบภาพ กำหนดให้ตัวแปรส่งออก (output) คือ ขนาดหน้าต่างเมทริกซ์ของข้อมูลส่งออก ส่วนข้อมูลนำเข้า(input) คือขนาดหน้าต่างเมทริกซ์ของข้อมูลนำเข้า ส่วนตัวกรอง (filter) มีขนาดหน้าต่างเมทริกซ์ของตัวกรองข้อมูล และการขยับเคลื่อน (stride) คือ จำนวนของการขยับเคลื่อน ดังสมการที่ 2.17

$$\text{output} = \frac{(\text{input} - \text{filter})}{\text{stride}} + 1 \quad (2.17)$$

กรณีที่มีการขยายขอบภาพ กำหนดให้ตัวแปร output คือ ขนาดหน้าต่างเมทริกซ์ของข้อมูลส่งออก ส่วน input คือ ขนาดหน้าต่างเมทริกซ์ของข้อมูลนำเข้า โดย filter ขนาดหน้าต่างเมทริกซ์ของตัวกรองข้อมูล และ stride คือ จำนวนของการขยับเคลื่อน และ padding คือ จำนวนการขยายออกของภาพ ดังสมการที่ 2.18

$$\text{output} = \frac{(\text{input} + 2 \times \text{padding} - \text{filter})}{\text{stride}} + 1 \quad (2.18)$$

บทที่ 3

วิธีการดำเนินงานวิจัย

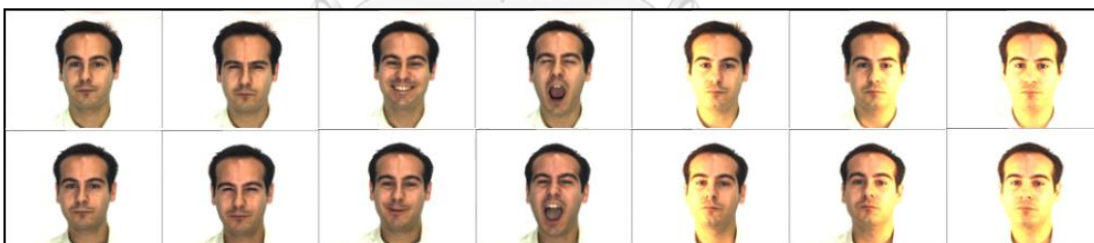
ในบทนี้จะกล่าวถึงการได้มาของภาพที่ใช้ในการทดลองของงานวิจัยนี้ และระบบรู้จำใบหน้าแบบมีสิ่งบดบังโดยใช้วิธีเชิงสถิติร่วมกับโครงข่ายประสาทเทียมแบบสังวัตนาการ

3.1 การได้มาของภาพ

ผู้วิจัยนำภาพที่ได้รับความนิยมใช้ในการวิจัยเกี่ยวกับการรู้จำใบหน้าบุคคลของฐานข้อมูล AR และ GTAV ซึ่งมีรายละเอียดดังต่อไปนี้

3.1.1 ฐานข้อมูล AR

ผู้วิจัยใช้ฐานข้อมูล AR ซึ่งเป็นข้อมูลรูปภาพที่ถูกสร้างโดย Alex Martinez และ Robert Benavente [22] ในกลุ่มของ Computer Vision Center ของ The Ohio State University ซึ่งเป็นฐานข้อมูลรูปภาพที่ได้รับความนิยมเกี่ยวกับการรู้จำใบหน้าเป็นอย่างมาก โดยมีฐานข้อมูลรูปภาพประกอบด้วย 117 คน คนละ 26 รูป มีขนาด 576 x 768 จุดภาพ



รูปที่ 3.1 กลุ่มรูปฝึกสอนตัวอย่างของบุคคลหนึ่ง จากฐานข้อมูล AR

จากรูปที่ 3.1 แบ่งเป็นกลุ่มรูปฝึกสอนออกเป็น 14 รูป ประกอบด้วยรูปใบหน้าตรงปกติ 2 รูป รูปที่แสดงออกทางอารมณ์ 6 รูป และรูปที่มีแสงส่องสว่าง 6 รูป รวมทั้งสิ้น 1,638 รูป



รูปที่ 3.2 กลุ่มรูปทดสอบตัวอย่างของบุคคลหนึ่ง จากฐานข้อมูล AR

จากรูปที่ 3.2 แสดงกลุ่มรูปทดสอบอีก 12 รูป ที่แบ่งเป็นรูปใส่แว่นกันแดด 6 รูป และผ้าคลุมปิดปาก 6 รูป รวมทั้งสิ้น 1,404 รูป

เนื่องจากรูปของฐานข้อมูลมีขนาดใหญ่ จึงปรับปรุงให้มีขนาดภาพที่ตรงกับความต้องการของโครงข่ายประสาทรูปแบบของ Alexnet โดยต้องปรับรูปให้มีขนาด 227×227 จุดภาพ ดังนั้น จึงต้องหาค่าเฉลี่ยภาพใบหน้าทั้งหมด เพื่อหาตำแหน่งใบหน้าจุดศูนย์กลางของใบหน้า แล้วทำการตัดส่วนเกินออกเพื่อให้เหลือเพียงขนาดดังกล่าวข้างต้น

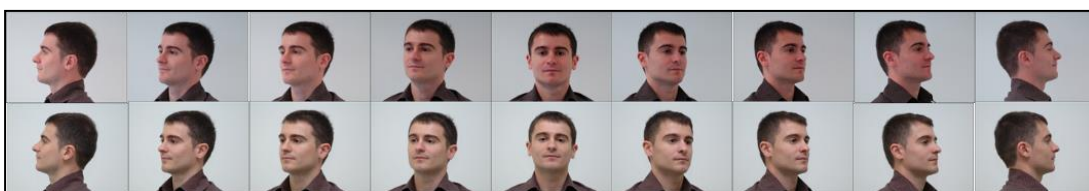


รูปที่ 3.3 ตัวอย่างการเพิ่มการบังบังใบหน้าเพื่อช่วยในการฝึกสอน

จากรูปที่ 3.3 แสดงภาพที่ทำการเพิ่มส่วนพื้นที่สีดำด้วยการบังบังที่บริเวณตาและปาก จากกลุ่มรูปฝึกสอนเดิมเพื่อที่จะฝึกสอนในกรณีที่อาจพบกลุ่มคนใส่แว่นกันแดดหรือคนที่มีผ้าปิดปาก จำนวนรูปฝึกสอนเดิมมี 1,638 รูป และรวมที่ปิดบริเวณตาและปากอีกอย่างละ 1,638 รูป รวมทั้งสิ้น 4,914 รูป ส่วนรูปทดสอบอีก 1,404 รูป โดยไม่มีการปรับปรุงเพิ่มเติมกับภาพทดสอบ เนื่องจากต้องการให้ใกล้เคียงสภาพความจริงให้มากที่สุด จากนั้นจึงเข้าสู่กระบวนการ 2DPCA และ 2DLDA ในการหาเวกเตอร์เฉพาะของฐานข้อมูลต่อไป

3.1.2 ฐานข้อมูล GTAV

โดยฐานข้อมูล GTAV นี้สร้างโดย Francesc Tarres และ Antonio Rama [23] ประกอบด้วย 44 คน คนละ 27 ภาพ ซึ่งจะมีลักษณะท่าทางที่แตกต่างกันออกไปจาก 0° , $\pm 30^\circ$, $\pm 45^\circ$, $\pm 60^\circ$ และ 90° ตามลำดับ หรือเรียกว่าการบดบังด้วยตัวเอง (self-occlusion) ภายใต้การส่องสว่างถึงสามแบบ มีขนาดภาพ 240×320 จุดภาพ โดยแบ่งกลุ่มเป็น 2 กลุ่ม



รูปที่ 3.4 รูปกลุ่มฝึกสอนตัวอย่างจากฐานข้อมูล GTAV

จากรูปที่ 3.4 แสดงกลุ่มรูปฝึกสอนที่ประกอบด้วยรูป 18 รูป รวมทั้งสิ้น 792 รูป ภายใต้การส่องสว่างแบบอ่อนและปานกลาง



รูปที่ 3.5 รูปกลุ่มฝึกสอนตัวอย่างจากฐานข้อมูล GTAV

จากรูปที่ 3.5 แสดงกลุ่มรูปทดสอบที่ประกอบด้วย 9 รูป ภายใต้การส่องสว่างแบบเข้ม รวมทั้งสิ้น 396 รูป เนื่องจากภาพฐานข้อมูลทั้งหมดมีขนาดใหญ่เกินไป จึงต้องย่อขนาดภาพให้มีขนาด 227×227 จุดภาพ เพื่อให้พอดีกับความต้องการของแบบจำลอง Alexnet

3.2 กระบวนการปรับสีภาพ

ในรูปภาพปกตินั้นย่อมมีสัญญาณรบกวน เช่น ความอึมัวตัวของแสง หรือสัญญาณรบกวนแบบเกลือและพริกไทย (salt and pepper) ดังนั้น ก่อนที่จะนำภาพไปประมวลผลจึงต้องทำความสะอาดข้อมูลก่อน



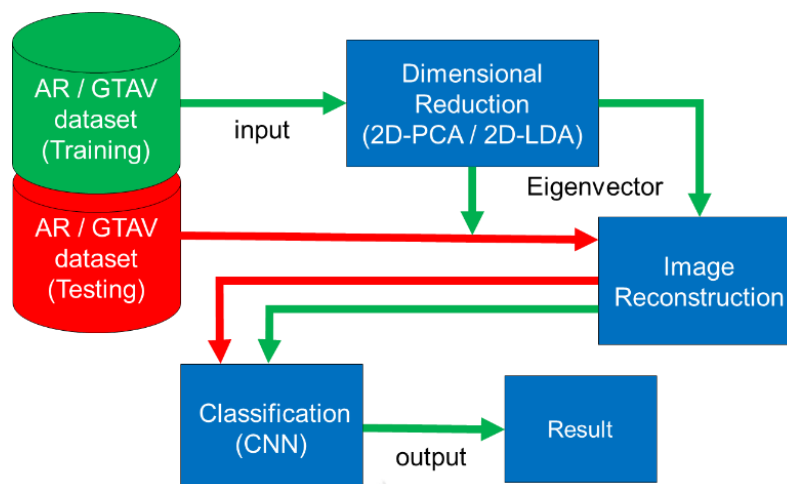
รูปที่ 3.6 ตัวอย่างการปรับสีภาพในปริภูมิสีต่างๆ (ก) RGB (ข) Grayscale
(ค) YCbCr (ง) YCbCr + Grayscale

จากรูปที่ 3.6 จะเห็นการแปลงจากภาพปริภูมิสีอาร์จีบีเป็นปริภูมิสีโทนเทาที่มักจะทำให้เกิดปัญหากับรูปที่มีการส่องสว่างมาก จึงควรทำการปรับเป็นปริภูมิสีวายซีบีซีอาร์ก่อนเพื่อตัดความอึมัวตัวของแสงที่ผสมมาในปริภูมิสีอาร์จีบีก่อน จากนั้นทำการแปลงเป็นภาพสีโทนเทาต่อไป เนื่องจากวิธี 2DPCA และ 2DLDA เหมาะกับการใช้ภาพขาว-ดำ จึงต้องมีการปรับสีของภาพเพื่อให้เกิดสะดวกต่อการประมวลผล

จุฬาลงกรณ์มหาวิทยาลัย

3.3 โครงสร้างการทำงานของระบบรู้จำใบหน้าแบบมีสิ่งบดบัง

ผู้วิจัยสร้างระบบการทำงานของระบบการรู้จำใบหน้าแบบมีสิ่งบดบังเป็นระบบปิด กล่าวคือ การรู้จำภายในฐานข้อมูลเท่านั้น การทดลองถูกแบ่งออกเป็น 2 ส่วน คือ ส่วนของฐานข้อมูล AR และ GTAV มีรายละเอียดดังรูปที่ 3.7 ดังต่อไปนี้



รูปที่ 3.7 โครงสร้างการทำงานของระบบ

จากรูปที่ 3.7 แสดงกระบวนการทั้งหมดที่เริ่มจากการนำรูปฝึกสอนทั้งหมดมาแปลงเป็นภาพโทนสีขาวซีบิซีอาร์ก่อนเพื่อลดความอึมตัวของแสง โดยปรับภาพเป็นโทนสีเทา (Grayscale) เนื่องจากวิธีเชิงสถิติเหมาะสมสำหรับภาพ 2 มิติและนำมาหาเวกเตอร์เฉพาะด้วยวิธี 2DPCA และ 2DLDA ของฐานข้อมูลนั้นๆ โดยนำเวกเตอร์เฉพาะมาคำนวณกับรูปฝึกสอนและรูปทดสอบเพื่อสร้างคืนภาพใหม่และเข้าสู่กระบวนการสังวัตนาการด้วย และจำแนกบุคคลแต่ละบุคคลโครงข่ายประสาทเทียมแบบสังวัตนาการโดยใช้แบบจำลองของ Alexnet กับแบบจำลองของผู้วิจัย ซึ่งจะกล่าวถึงผลการทดลองทั้งหมดจะกล่าวในบทที่ 4 ต่อไป

3.4 เครื่องมือที่ใช้พัฒนา จุฬาลงกรณ์มหาวิทยาลัย

การประมวลผลของงานวิจัยนี้ใช้โปรแกรม MATLAB2019a update 3 ใบอนุญาตที่ 40662904 ของจุฬาลงกรณ์มหาวิทยาลัย โดยเครื่องประมวลผลที่มีระบบปฏิบัติการ Windows 10 Home, Processor: Intel® Core™ i7-4710HQ CPU @ 2.50GHz, RAM 16.0 GB System type 64-bit Operating System

3.5 การทำงานของโครงข่ายประสาทเทียมแบบสังวัตนาการแบบจำลองของ Alexnet

โครงข่ายประสาทเทียมแบบสังวัตนาการของแบบจำลอง Alexnet มีทั้งหมด 8 ชั้นหลัก แบ่งเป็น ชั้นสังวัตนาการ 5 ชั้น และชั้นการเชื่อมเต็มรูปแบบ 3 ชั้น แต่ทั้งหมดจะมีการทำงานย่อยเป็น 25 ชั้น มีการแสดงการทำงานในส่วนของชั้นต่างๆ มีรายละเอียดดังตารางที่ 3.1 ต่อไปนี้

ตารางที่ 3.1 ตารางการทำงานแบบจำลอง Alexnet

ลำดับ	ชนิดของชั้น	การทำงาน
1	Image Input	รับข้อมูลที่มีขนาด $227 \times 227 \times 3$
2	Convolution	สร้างตัวกรองสังวัตนาการขนาด $11 \times 11 \times 3$ จำนวน 96 ตัว ขับเคลื่อนครั้งละ 4 ก้าว การขยายขอบภาพเป็น 0 คำนวณตามสมการที่ 2.13 จะได้ขนาดข้อมูลส่งออกเป็น $55 \times 55 \times 96$
3	ReLU	ฟังก์ชันกระตุ้น ReLU เพื่อปรับค่าที่ติดลบให้เป็นค่าศูนย์
4	Cross Channel Normalization	cross channel normalization with 5 channels per element
5	Max Pooling	สร้างตัวกรองพหุคูณสูงสุด ขนาด 3×3 ด้วยการเคลื่อนที่ครั้งละ 2 ก้าว และการขยายขอบภาพเป็น 0 คำนวณตามสมการที่ 2.17 จะได้ขนาดข้อมูลส่งออกเป็น $27 \times 27 \times 96$
6	Grouped Convolution	สร้างตัวกรองสังวัตนาการขนาด 5×5 จำนวน 256 ตัว ด้วยการเคลื่อนที่ครั้งละ 1 ก้าว และการขยายขอบภาพเป็น 2 คำนวณตามสมการที่ 2.18 จะได้ขนาดข้อมูลส่งออกเป็น $27 \times 27 \times 256$
7	ReLU	ฟังก์ชันกระตุ้น ReLU เพื่อปรับค่าที่ติดลบให้เป็นค่าศูนย์
8	Cross Channel Normalization	cross channel normalization with 5 channels per element
9	Max Pooling	สร้างตัวกรองพหุคูณสูงสุด 3×3 ด้วยการเคลื่อนที่ครั้งละ 2 ก้าว และการขยายขอบภาพเป็น 0 คำนวณตามสมการที่ 2.13 จะได้ขนาดข้อมูลส่งออกเป็น $13 \times 13 \times 256$
10	Convolution	สร้างตัวกรองสังวัตนาการขนาด 3×3 จำนวน 384 ตัว ด้วยการเคลื่อนที่ครั้งละ 1 ก้าว และการขยายขอบภาพเป็น 1 คำนวณตามสมการที่ 2.18 จะได้ขนาดข้อมูลส่งออกเป็น $13 \times 13 \times 384$
11	ReLU	ฟังก์ชันกระตุ้น ReLU เพื่อปรับค่าที่ติดลบให้เป็นค่าศูนย์
12	Grouped Convolution	สร้างตัวกรองสังวัตนาการขนาด 3×3 จำนวน 384 ตัว ด้วยการเคลื่อนที่ครั้งละ 1 ก้าว และการขยายขอบภาพเป็น 1 คำนวณตามสมการที่ 2.18 จะได้ขนาดข้อมูลส่งออกเป็น $13 \times 13 \times 384$
13	ReLU	ฟังก์ชันกระตุ้น ReLU เพื่อปรับค่าที่ติดลบให้เป็นค่าศูนย์

14	Grouped Convolution	สร้างตัวกรองสังวัตนาการขนาด 3x3 จำนวน 256 ตัว ด้วยการเคลื่อนที่ครั้งละ 1 ก้าว และการขยายขอบภาพเป็น 1 คำนวณตามสมการที่ 2.18 จะได้ขนาดข้อมูลส่งออกเป็น 13x13x256
15	ReLU	ฟังก์ชันกระตุ้น ReLU เพื่อปรับค่าที่ติดลบให้เป็นค่าศูนย์
16	Max Pooling	สร้างตัวกรองพหุคูณสูงสุด 3x3 max pooling ด้วยการเคลื่อนที่ครั้งละ 2 ก้าว และการขยายขอบภาพเป็น 0 คำนวณตามสมการที่ 2.17 จะได้ขนาดข้อมูลส่งออกเป็น 6x6x256
17	Fully Connected	เชื่อมโยงกับทุกโหนดในชั้นก่อนหน้าด้วยจำนวน 4096 จะได้ $(6 \times 6 \times 256) \times 4096$
18	ReLU	ฟังก์ชันกระตุ้น ReLU เพื่อปรับค่าที่ติดลบให้เป็นค่าศูนย์
19	Dropout	สุ่มปิดการทำงานชั้นการเชื่อมโยงเต็มรูปแบบร้อยละ 50
20	Fully Connected	เชื่อมโยงกับทุกโหนดในชั้นก่อนหน้าด้วยจำนวน 4096 จะได้ 4096×4096
21	ReLU	ฟังก์ชันกระตุ้น ReLU เพื่อปรับค่าที่ติดลบให้เป็นค่าศูนย์
22	Dropout	สุ่มปิดการทำงานชั้นการเชื่อมโยงเต็มรูปแบบร้อยละ 50
23	Fully Connected	เชื่อมโยงกับทุกโหนดในชั้นก่อนหน้า เข้ากับ 117 คลาสจากฐานข้อมูล AR และ 44 คลาสจากฐานข้อมูล GTAV จะได้ การเชื่อมโยงเป็น 4096×117 และ 4096×44 ตามลำดับ
24	Softmax	ฟังก์ชันกระตุ้น Softmax เพื่อหาความน่าจะเป็นของบุคคลที่นำมาเปรียบเทียบกับฐานข้อมูล
25	Classification Output	แสดงผลการเปรียบเทียบและแสดงความน่าจะเป็นของบุคคลที่นำมาเปรียบเทียบ

3.6 โครงข่ายประสาทแบบจำลองของผู้วิจัย

ผู้วิจัยได้ออกแบบโครงสร้างของแบบจำลองของ Alexnet ใหม่ โดยอ้างอิงจากโครงสร้างและการทำงานเดิมทั้งหมด จากเดิมแบบจำลองของ Alexnet ต้องการข้อมูลที่ขนาด 227×227 เท่านั้น เปลี่ยนโครงสร้างใหม่เป็นรับข้อมูลได้ $227 \times d$ มิติ โดยขยายขอบภาพออกไปอีก ที่ชั้นสังวัตนาการที่ 2, 5, 9 และ 16 จากตารางที่ 3.1 อย่างละ 1 ขอบภาพ เนื่องจากผู้วิจัยได้วางแผนการทดลองบันทึกค่าทุก 10 มิติโอแกนเวกเตอร์ จึงต้องใช้ค่าดังกล่าวและเปลี่ยนขนาดชั้นการเชื่อมโยงเต็มรูปแบบ ที่ชั้นที่ 17 จากเดิมมีที่มีขนาด 9612×4096 จุดภาพ ย่อขนาดภาพให้เหลือเปลี่ยนเป็น $d \times 4096$

จุดภาพ สามารถที่ d มิติ สามารถคำนวณได้สมการที่ 2.17 และ 2.18 ทำการคำนวณขนาดข้อมูลส่งออกตั้งแต่ชั้นที่ 1 ถึงชั้นที่ 5 ของชั้นสังวัตนาการซึ่งเป็นชั้นสุดท้ายก่อนส่งต่อให้กับชั้นการเชื่อมโยงเต็มรูปแบบ จะได้ผลลัพธ์ขนาดข้อมูลนำเข้าของชั้นการเชื่อมโยงเต็มรูปแบบของชั้นที่ 17 ดังตารางที่ 3.2 เพื่อปรับให้ใช้ได้ กับเวกเตอร์เฉพาะที่ลดขนาดมิติลง ด้วยวิธีเชิงสถิติได้ จากนั้นเปลี่ยนชั้นการเชื่อมโยงเต็มรูปแบบในชั้นที่ 23 จากเดิมแบบจำลอง Alexnet จำแนกวัตถุได้ 1,000 ประเภท จึงปรับให้สอดคล้องกับจำนวนคลาสของฐานข้อมูล AR 117 คลาส และ GTAV 44 คลาส

ตารางที่ 3.2 ผลลัพธ์จากการคำนวณหาขนาดข้อมูลก่อนเข้าชั้นการเชื่อมโยงเต็มรูปแบบชั้นที่ 17

ขนาดมิติเวกเตอร์เฉพาะ ทุก 10 มิติ	ขนาดเมทริกซ์ค่าถ่วงน้ำหนัก
227 × 10 ถึง 227 × 40	1792 × 4096
227 × 50 ถึง 227 × 70	3584 × 4096
227 × 80 ถึง 227 × 100	5376 × 4096
227 × 110 ถึง 227 × 130	7168 × 4096
227 × 140 ถึง 227 × 160	8960 × 4096
227 × 170 ถึง 227 × 200	10752 × 4096
227 × 210 ถึง 227 × 220	12544 × 4096

บทที่ 4

ผลการทดลองและการวิเคราะห์ผล

ผู้วิจัยดำเนินการศึกษาทดลองและวิเคราะห์ผลจาก 2 ฐานข้อมูล ได้แก่ ฐานข้อมูล AR และ ฐานข้อมูล GTAV มีผลการทดลองแบ่งเป็นออก 6 ส่วน ดังต่อไปนี้

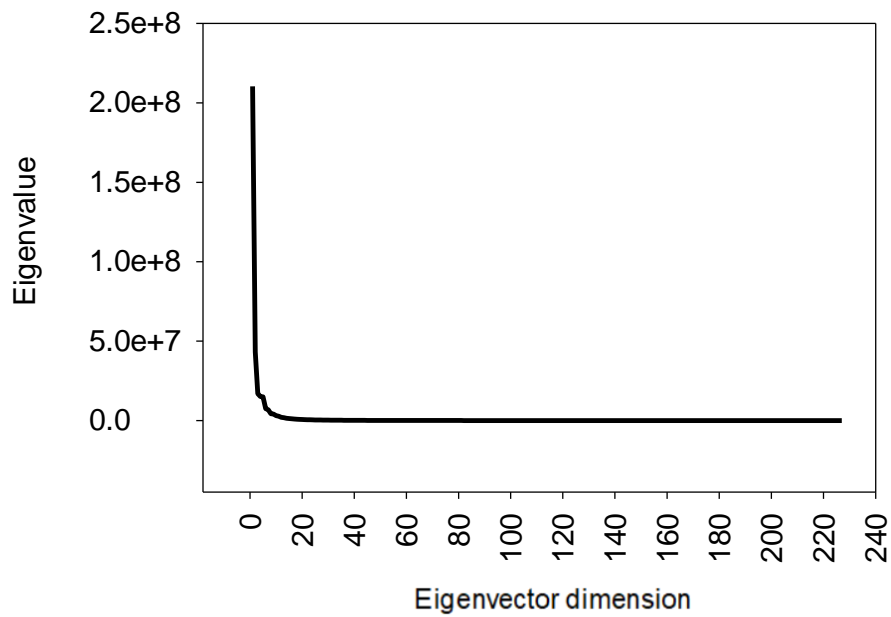
1. ผลการทดลองการสร้างคืนภาพด้วยวิธี 2DPCA และ 2DLDA
2. ผลการทดลองการปรับค่าไฮเพอร์พารามิเตอร์ สำหรับการฝึกสอนของโครงข่ายประสาทเทียมแบบสังวัตนาการ โดยใช้แบบจำลองของ Alexnet
3. ผลการทดลองการเปรียบเทียบแบบใช้และไม่ใช้วิธีเชิงสถิติร่วมกับโครงข่ายประสาทแบบสังวัตนาการ โดยใช้แบบจำลองของ Alexnet
4. ผลการทดลองการเปรียบเทียบแบบใช้วิธีเชิงสถิติร่วมกับโครงข่ายประสาทเทียมแบบสังวัตนาการ โดยใช้แบบจำลองของ Alexnet กับผู้วิจัย
5. ผลการทดลองการเปรียบเทียบระหว่างการสร้างคืนภาพด้วยวิธี 2DPCA และ 2DLDA ร่วมกับโครงข่ายประสาทเทียมแบบสังวัตนาการ โดยใช้แบบจำลอง Alexnet
6. ผลการทดลองนำระบบการรู้จำใบหน้าไปใช้ในชีวิตประจำวัน

4.1 ผลการทดลองการสร้างคืนภาพด้วยวิธี 2DPCA และ 2DLDA

จากการดำเนินการนำทฤษฎีของวิธี 2DPCA และ 2DLDA ซึ่งมีคุณสมบัติพื้นฐานอย่างหนึ่งที่เหมือนกันคือ สามารถสร้างคืนภาพใบหน้าใหม่ได้ ดังนั้นการที่จะทดสอบว่าทฤษฎีที่ได้เสนอในวิทยานิพนธ์นี้ว่ามีความถูกต้องหรือไม่ จึงทำการทดลองสร้างคืนภาพใบหน้าใหม่ มีรายละเอียดดังนี้

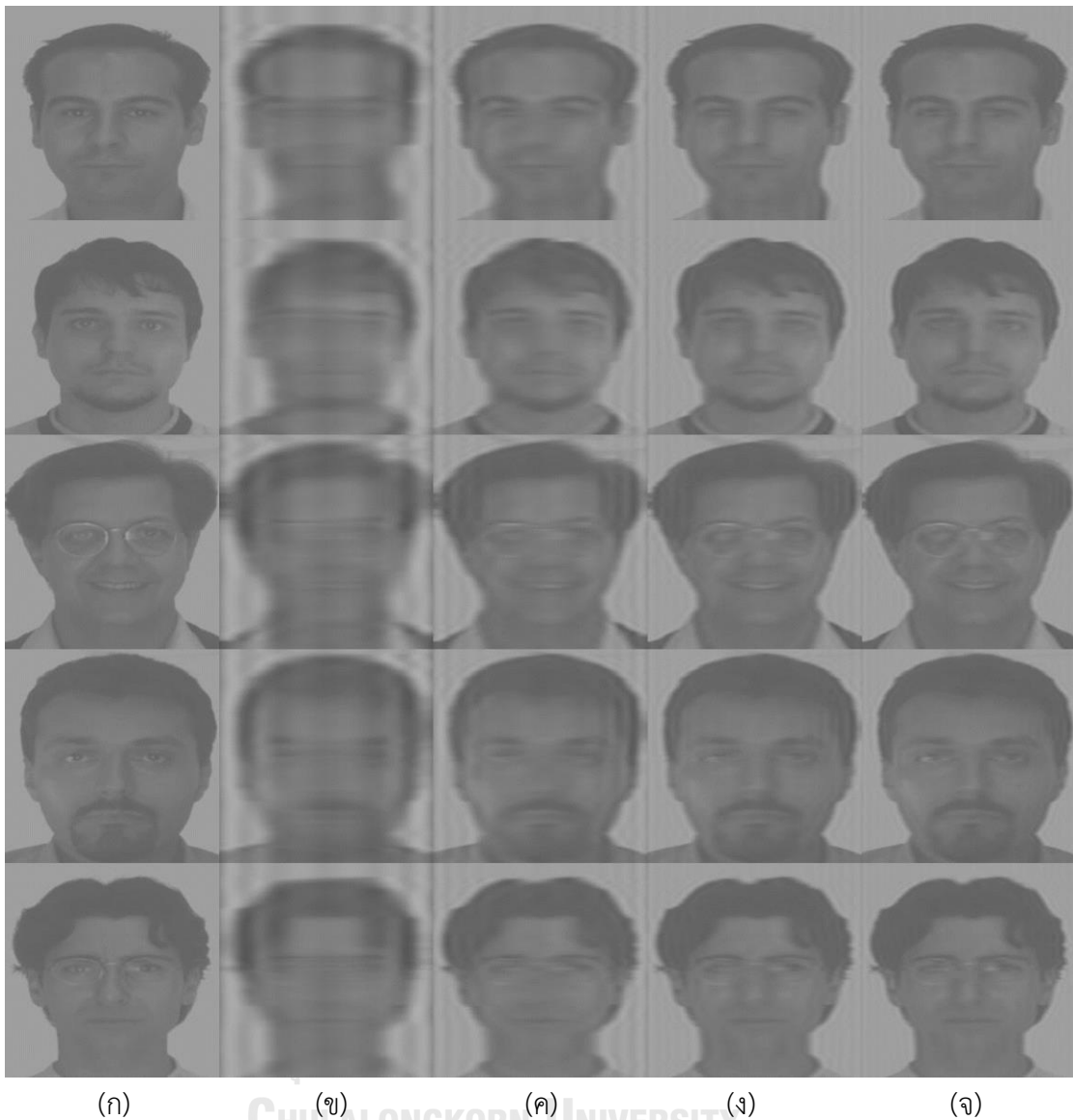
4.1.1 ผลการทดลองการสร้างคืนภาพด้วยวิธี 2DPCA

การทดลองการสร้างคืนภาพด้วยวิธี 2DPCA เริ่มจากการคำนวณหาค่าเฉพาะและเวกเตอร์เฉพาะตามสมการที่ 2.3 แล้ว โดยทำการเรียงลำดับเวกเตอร์เฉพาะที่สมนัยกับค่าเฉพาะจากมากไปน้อย จากนั้นเลือกแกนการฉายจำนวน d มิติ ดังรูปที่ 4.1



รูปที่ 4.1 กราฟแสดงการเรียงลำดับค่าเวกเตอร์เฉพาะที่สมนัยกับค่าเฉพาะจากมากไปน้อย
ด้วยวิธี 2DPCA

จากรูปที่ 4.1 กราฟแสดงการเรียงลำดับค่าเวกเตอร์เฉพาะที่สมนัยกับค่าเฉพาะจากมากไปน้อย ผลการทดลองพบว่า ค่าเฉพาะที่มีค่าสูงก็คือข้อมูลที่สำคัญและมีประโยชน์ต่อการแยกประเภทบุคคล ส่วนค่าเฉพาะที่มีค่าต่ำก็คือข้อมูลที่ไม่มีความสำคัญต่อการแยกประเภทบุคคล สามารถคัดออกได้โดยไม่เสียข้อมูลสำคัญ



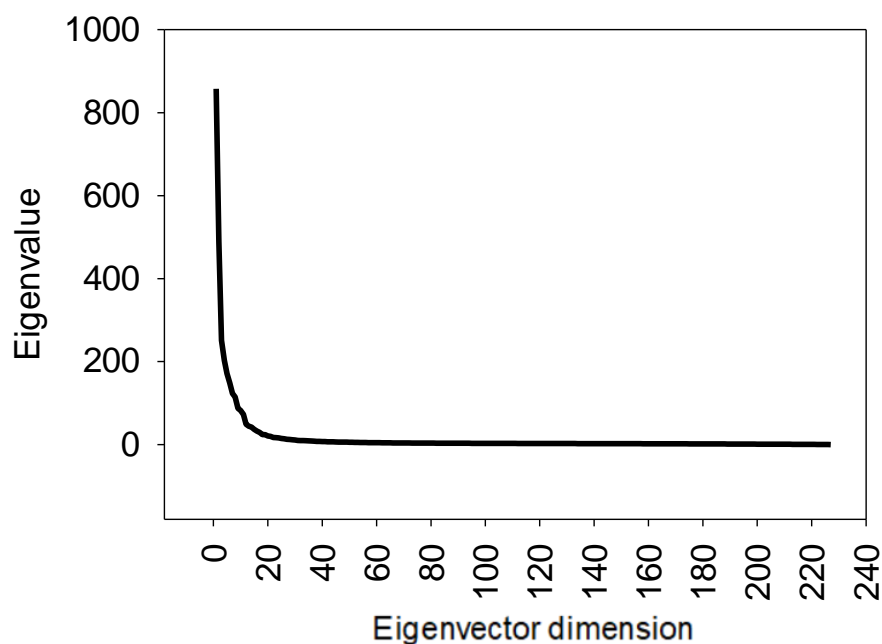
รูปที่ 4.2 ตัวอย่างการสร้างคืนภาพของบุคคลหนึ่งด้วยวิธี 2DPCA

(ก) ภาพต้นแบบ (ข) มิติ = 10 (ค) มิติ = 20 (ง) มิติ = 30 (จ) มิติ = 40

จากรูปที่ 4.2 เป็นรูปตัวอย่างของบุคคลหนึ่งจากฐานข้อมูล AR ที่ถูกนำมาสร้างคืนภาพใหม่ด้วยวิธี 2DPCA เพื่อนำมาหาแผนการฉายที่เหมาะสมที่สุด จะได้คุณลักษณะเมทริกซ์ที่ใช้ในการสร้างคืนภาพ เริ่มจาก (ก) เป็นภาพต้นแบบ และภาพลำดับถัดมาจะเป็นการสร้างคืนภาพตามมิติสตรมภ์ของเวกเตอร์เฉพาะที่เพิ่มขึ้นตามลำดับ จากผลการทดลองพบว่า ยิ่งใช้มิติสตรมภ์ของเวกเตอร์เฉพาะเพิ่มขึ้นความคมชัดของรูปก็ยิ่งใกล้เคียงกับต้นฉบับมากขึ้นเช่นกัน เนื่องจากค่าเอกลักษณ์สำคัญจะถูกรวบรวมมาทั้งหมด จึงทำให้ภาพคมชัดขึ้น

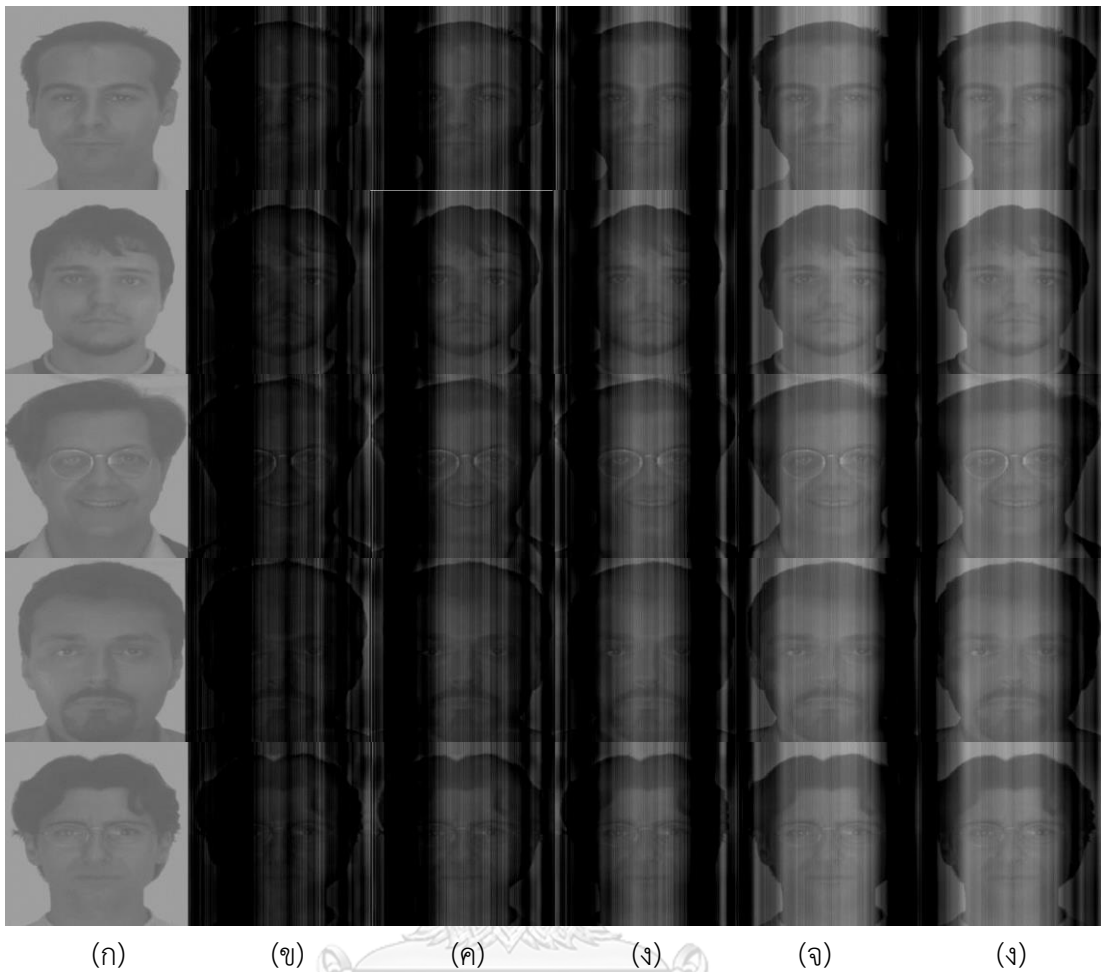
4.1.2 ผลการทดลองการสร้างคืนภาพด้วยวิธี 2DLDA

การทดลองการสร้างคืนภาพด้วยวิธี 2DLDA เริ่มจากการคำนวณหาค่าเฉพาะและเวกเตอร์เฉพาะตามสมการที่ 2.9 โดยเรียงลำดับเวกเตอร์เฉพาะที่สมนัยกับค่าเฉพาะจากมากไปน้อย และเลือกแกนการฉายจำนวน d มิติ ดังรูปที่ 4.3



รูปที่ 4.3 กราฟแสดงเรียงลำดับเวกเตอร์เฉพาะที่สมนัยกับค่าเฉพาะจากมากไปน้อย 2DLDA

จากรูปที่ 4.3 กราฟแสดงการเรียงลำดับค่าเวกเตอร์เฉพาะที่สมนัยกับค่าเฉพาะจากมากไปน้อย ผลการทดลองพบว่า ค่าเฉพาะที่มีค่าสูงก็คือข้อมูลที่สำคัญและมีประโยชน์ต่อการแยกประเภทบุคคล ส่วนค่าเฉพาะที่มีค่าต่ำคือข้อมูลที่ไม่มีความสำคัญต่อการแยกประเภทบุคคล สามารถคัดออกได้โดยไม่เสียข้อมูลสำคัญ



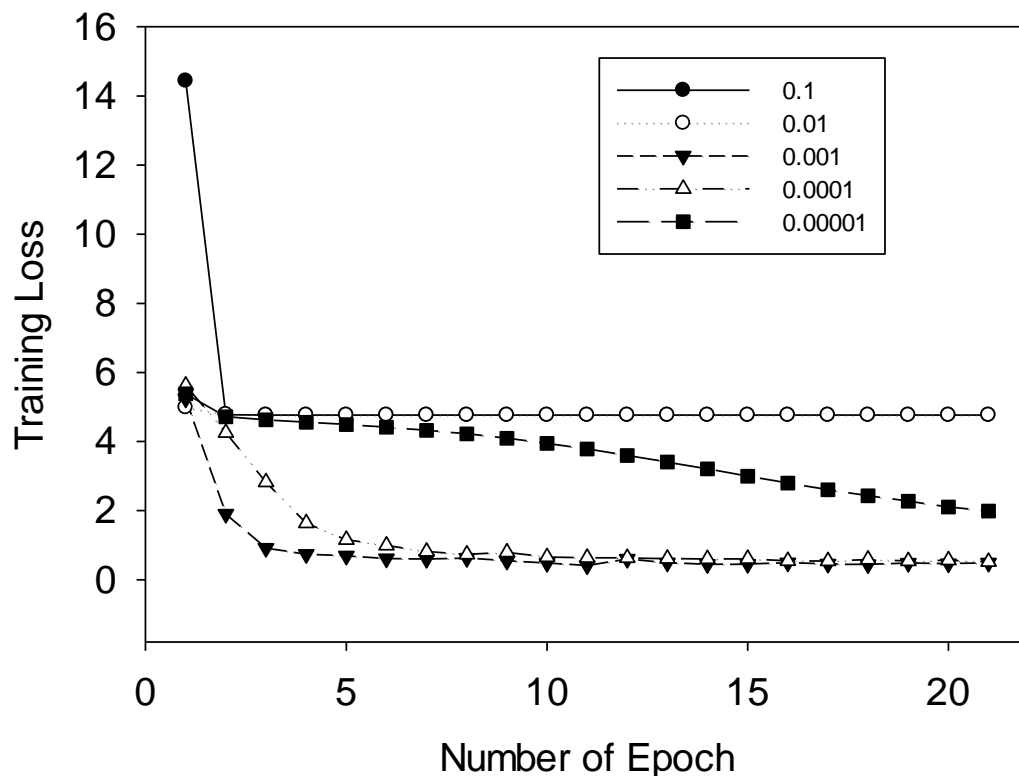
รูปที่ 4.4 ตัวอย่างการสร้างคืนภาพของบุคคลหนึ่งด้วยวิธี 2DLDA (ก) ภาพต้นแบบ (ข) มิติ = 222 (ค) มิติ = 223 (ง) มิติ = 224 (จ) มิติ = 225 (ฉ) มิติ = 226

รูปที่ 4.4 แสดงรูปตัวอย่างของบุคคลหนึ่งจากฐานข้อมูล AR ที่ถูกนำมาสร้างคืนภาพใหม่ด้วยวิธี 2DLDA เพื่อนำมาหาแกนการฉายที่เหมาะสมที่สุด จะได้คุณลักษณะเมทริกซ์ที่ใช้ในการสร้างคืนภาพ เริ่มจาก (ก) เป็นภาพต้นแบบ และภาพลำดับถัดมาจะเป็นการสร้างคืนภาพตามมิติสตรมภ์ของเวกเตอร์เฉพาะที่เพิ่มขึ้นตามลำดับ จากการทดลองพบว่า ยิ่งใช้มิติสตรมภ์ของเวกเตอร์เฉพาะเพิ่มขึ้น ความคมชัดของรูปก็ยิ่งใกล้เคียงกับต้นฉบับมากขึ้นเช่นกัน เนื่องจากค่าเอกลักษณะสำคัญจะถูกรวมรวมมาทั้งหมด จึงทำให้ภาพคมชัดขึ้น

4.2 ผลการทดลองการปรับค่าไฮเพอร์พารามิเตอร์สำหรับการฝึกสอนของโครงข่ายประสาทเทียมแบบสังวัตนาการ โดยใช้แบบจำลองของ Alexnet

ผลการทดลองการปรับค่าไฮเพอร์พารามิเตอร์สำหรับการฝึกสอนของโครงข่ายประสาทแบบสังวัตนาการ โดยใช้แบบจำลองของ Alexnet ที่แบ่งการทดลองออกเป็น 2 ส่วน คือฐานข้อมูล AR และ GTAV ตามลำดับ โดยกำหนดค่าไฮเพอร์พารามิเตอร์ 3 ตัวแปรที่สำคัญสำหรับการฝึกสอนเพื่อหาค่าที่มีประสิทธิภาพที่ดีที่สุด มีรายละเอียดดังต่อไปนี้

1. กำหนดอัตราเร็วของการเรียนรู้ 0.1, 0.01, 0.001, 0.0001, 0.00001
2. กำหนดขนาดกลุ่มภาพฝึกสอน ได้แก่ 16, 32, 64, 128 และ 256 ภาพต่อกลุ่ม
3. กำหนดจำนวนรอบการฝึกสอน 20 รอบ



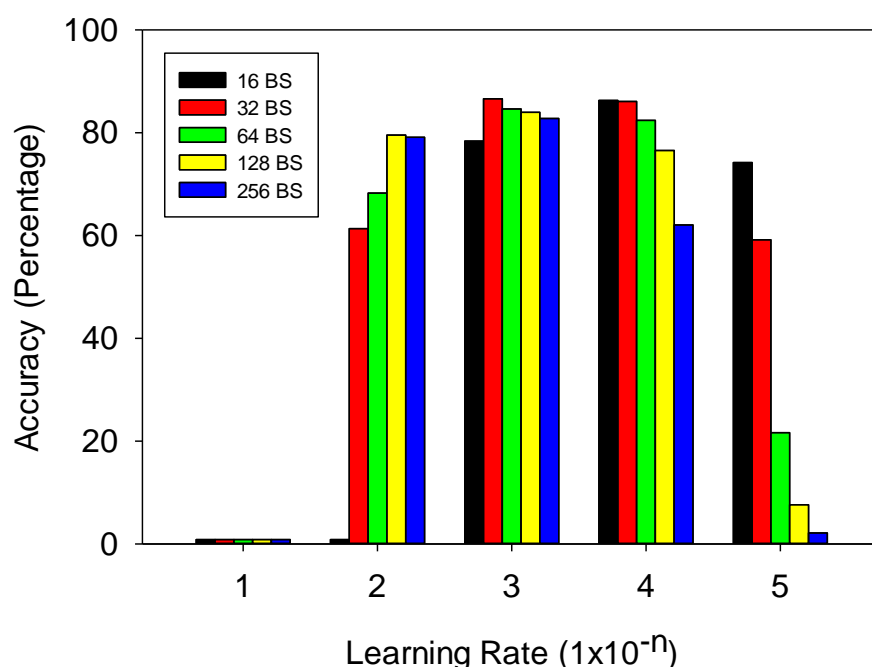
รูปที่ 4.5 กราฟแสดงอัตราการสูญเสียระหว่างการฝึกสอนของทุกอัตราการเรียนรู้ต่อขนาดกลุ่มภาพฝึกสอน 32 ภาพต่อกลุ่มจำนวนในการฝึกสอน 20 รอบ จากฐานข้อมูล AR

จากรูปที่ 4.5 แสดงแสดงอัตราการสูญเสียระหว่างการฝึกสอนของทุกอัตราการเรียนรู้ต่อขนาดกลุ่มภาพฝึกสอน 32 ภาพต่อกลุ่มจำนวนในการฝึกสอน 20 รอบ จากฐานข้อมูล AR พบว่า อัตราการเร็วของการเรียนรู้ที่ 0.1 เส้นกราฟกำลังตกลงมาและเป็นค่าคงที่ไม่ใกล้เคียงศูนย์ และ

อัตราเร็วของการเรียนรู้ที่ 0.01 เป็นค่าคงที่ไม่ใกล้เคียงศูนย์ตลอดจนการฝึกสอนรอบสุดท้าย แสดงให้เห็นว่า เกิดความเสียหายระหว่างการฝึกสอนกล่าวคือค่าอัตราเร็วของการเรียนรู้นั้นยังเป็นค่าที่สูงเกินไป ส่วนอัตราเร็วของการเรียนรู้ที่ 0.001 และ 0.0001 แสดงให้เห็นว่า กราฟกำลังตกลงเข้าใกล้ศูนย์ลงเรื่อยๆตลอดจนการฝึกสอนรอบสุดท้าย แสดงให้เห็นว่า เกิดความเสียหายระหว่างการฝึกสอนน้อยกล่าวคือค่าอัตราเร็วของการเรียนรู้เป็นค่าที่เหมาะสมของระบบ และท้ายสุดคืออัตราเร็วของการเรียนรู้ที่ 0.00001 แสดงให้เห็นว่า เส้นกราฟกำลังตกลงเรื่อยๆ แต่ยังไม่เข้าใกล้ค่าศูนย์ตลอดจนรอบสุดท้ายของการฝึก กล่าวคือค่าอัตราเร็วของการเรียนรู้นั้นเข้าไป

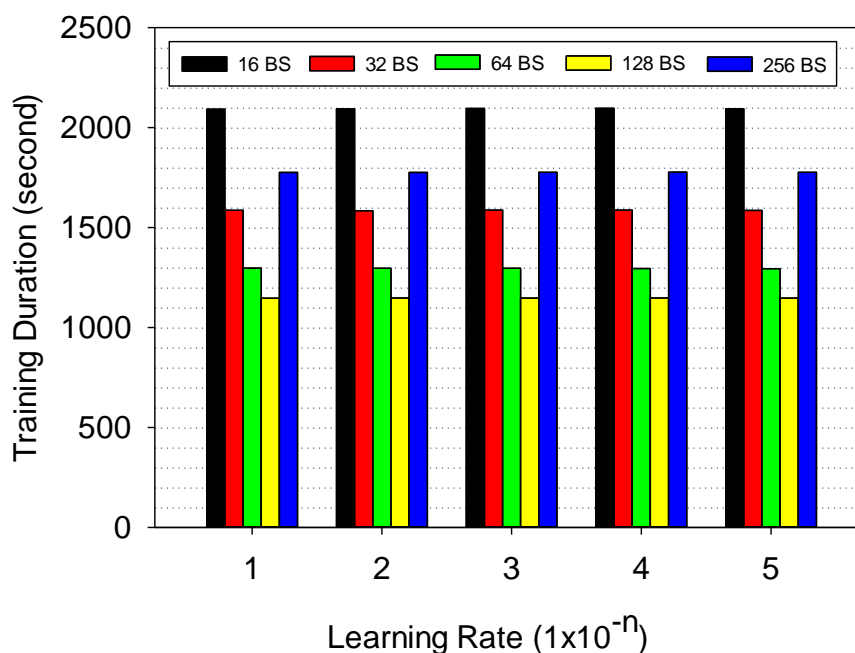
4.2.1 ผลการทดลองการปรับค่าไฮเพอร์พารามิเตอร์ จากฐานข้อมูล AR

ผลการทดลองการปรับค่าไฮเพอร์พารามิเตอร์ด้วยการสุ่มหาค่าที่เหมาะสม โดยทำการเพิ่มค่าทีละตัวแปรที่เริ่มจากกำหนดค่าอัตราการเรียนรู้คงที่ก่อน จากนั้นจะเพิ่มขนาดกลุ่มภาพฝึกสอนทุก 16, 32, 64, 128 และ 256 ตามลำดับ โดยใช้รูปภาพจากต้นฉบับสีโทนเทาในการฝึกสอนและทำการทดสอบโดยไม่ผ่านกระบวนการเชิงสถิติเพื่อให้ได้ค่าไฮเพอร์พารามิเตอร์ที่เป็นมาตรฐานของโครงข่ายประสาทกับฐานข้อมูล AR มีรายละเอียดดังรูปที่ 4.6



รูปที่ 4.6 กราฟแสดงความแม่นยำของอัตราการเรียนรู้ต่อขนาดชุดข้อมูลจากฐานข้อมูล AR

จากรูปที่ 4.6 แสดงความแม่นยำของอัตราการเรียนรู้ต่อขนาดชุดข้อมูลจากฐานข้อมูล AR โดยกราฟแท่งสีแดง ใช้อัตราเร็วของการเรียนรู้ที่ 0.001 และใช้ขนาดกลุ่มรูปฝึกสอน 32 ภาพต่อกลุ่ม ส่วนกราฟแท่งสีดำและสีแดง ใช้อัตราเร็วของการเรียนรู้ที่ 0.0001 และใช้ขนาดกลุ่มรูปฝึกสอน 16 และ 32 ภาพต่อกลุ่ม ตามลำดับ ผลการทดลองพบว่า ให้ความแม่นยำที่ใกล้เคียงกันที่ร้อยละ 85



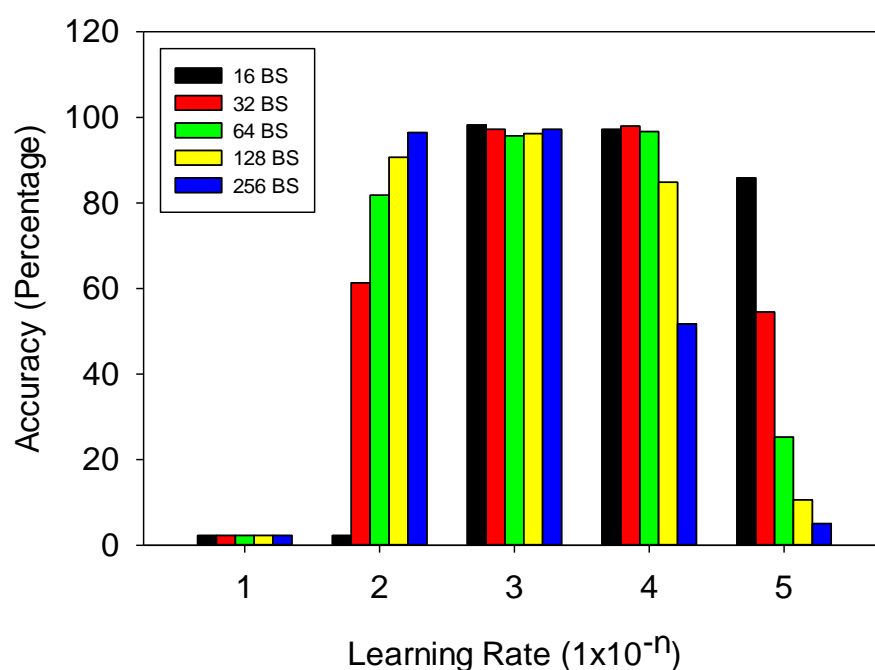
รูปที่ 4.7 กราฟแสดงเวลาที่ใช้ในการฝึกสอนต่ออัตราเร็วการเรียนรู้ของทุกขนาดกลุ่มภาพฝึกสอน จากฐานข้อมูล AR

จากรูปที่ 4.7 กราฟแสดงเวลาที่ใช้ในการฝึกสอนต่ออัตราเร็วของการเรียนรู้ของทุกขนาดกลุ่มภาพฝึกสอน จากฐานข้อมูล AR พบว่า ทุกอัตราเร็วการเรียนรู้โดยใช้ขนาดกลุ่มรูปฝึกสอนที่เท่ากัน จะใช้เวลาในการฝึกสอนที่ใกล้เคียงกัน แต่กลุ่มที่ใช้เวลาฝึกสอนน้อยที่สุด คือใช้ขนาดกลุ่มรูปฝึกสอน 128 ภาพต่อกลุ่ม ที่ใช้เวลาใกล้เคียง 1,150 วินาที ถึงแม้จะใช้ขนาดกลุ่มรูปฝึกสอน 128 ภาพต่อกลุ่ม โดยใช้เวลาในการฝึกสอนน้อยที่สุด แต่ไม่ให้ความแม่นยำสูงสุด

สรุปได้ว่า ค่าไฮเพอร์พารามิเตอร์ของโครงข่ายประสาทโดยใช้แบบจำลองของ Alexnet เพื่อการฝึกสอนฐานข้อมูล AR ที่ใช้ขนาดกลุ่มภาพฝึกสอน 32 ภาพต่อกลุ่ม และใช้อัตราเร็วของการเรียนรู้ที่ 0.001 หรือ 0.0001 ส่งผลต่อความแม่นยำในการฝึกสอนและใช้เวลาฝึกสอนได้เร็วกว่าการใช้ขนาดกลุ่มรูปฝึกสอน 16 ภาพต่อกลุ่ม ในขณะที่ให้ความแม่นยำใกล้เคียงกัน ซึ่งถือว่ามีประสิทธิภาพที่ดีที่สุด

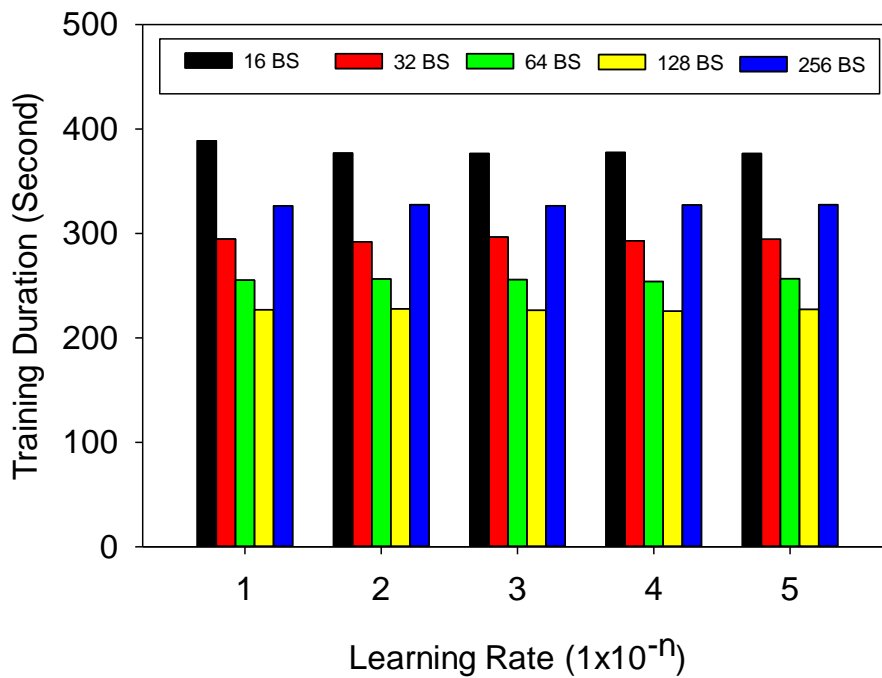
4.2.2 ผลการทดลองการปรับค่าไฮเปอร์พารามิเตอร์ จากฐานข้อมูล GTAV

ผลการทดลองการปรับค่าไฮเปอร์พารามิเตอร์ด้วยการสุ่มหาค่าที่เหมาะสมโดยการเพิ่มค่าที่ละตัวแปรที่เริ่มจากกำหนดค่าอัตราการเรียนรู้คงที่ก่อน จากนั้นเพิ่มขนาดกลุ่มภาพฝึกสอนทุก 16, 32, 64, 128 และ 256 ตามลำดับ การทดลองใช้รูปภาพจากต้นฉบับสีโทนเทาในการฝึกสอนและทดสอบ โดยที่ไม่ผ่านกระบวนการเชิงสถิติเพื่อให้ได้ค่าไฮเปอร์พารามิเตอร์ที่เป็นมาตรฐานของโครงข่ายประสาทกับฐานข้อมูล GTAV มีรายละเอียดดังรูปที่ 4.8



รูปที่ 4.8 กราฟแสดงความแม่นยำของอัตราการเรียนรู้ต่อขนาดชุดข้อมูลจากฐานข้อมูล GTAV

จากรูปที่ 4.8 แสดงความแม่นยำของอัตราการเรียนรู้ต่อขนาดชุดข้อมูลจากฐานข้อมูล GTAV โดยใช้อัตราเร็วของการเรียนรู้ที่ 0.0001 และใช้ขนาดกลุ่มรูปฝึกสอน 16, 32 และ 64 ภาพต่อกลุ่ม ผลการทดลองพบว่า ให้ความแม่นยำที่ใกล้เคียงกันที่ร้อยละ 96 แต่การใช้อัตราเร็วการเรียนรู้ที่ 0.001 กับทุกขนาดกลุ่มภาพฝึกสอนจะให้ความแม่นยำใกล้เคียงกันที่ร้อยละ 97



รูปที่ 4.9 กราฟแสดงเวลาในการฝึกสอนของอัตราการเรียนรู้ต่อขนาดชุดข้อมูลจากฐานข้อมูล GTAV

จากรูปที่ 4.9 แสดงเวลาในการฝึกสอนของอัตราการเรียนรู้ต่อขนาดชุดข้อมูลจากฐานข้อมูล GTAV โดยในทุกอัตราเร็วของการเรียนรู้ที่ใช้ขนาดกลุ่มรูปฝึกสอนที่เท่ากัน จะใช้เวลาฝึกสอนที่ใกล้เคียงกัน เมื่อพิจารณาจากกราฟแท่งสีเหลือง ที่ใช้อัตราเร็วการเรียนรู้ 0.001 และใช้ขนาดกลุ่มรูปฝึกสอน 128 ภาพต่อกลุ่ม ผลการทดลองพบว่า จะใช้เวลาฝึกสอนที่น้อยที่สุด (ใช้เวลาการฝึกสอนใกล้เคียง 226 วินาที) และในทุกขนาดกลุ่มรูปฝึกสอนจะให้ค่าความแม่นยำใกล้เคียงกันที่ร้อยละ 97

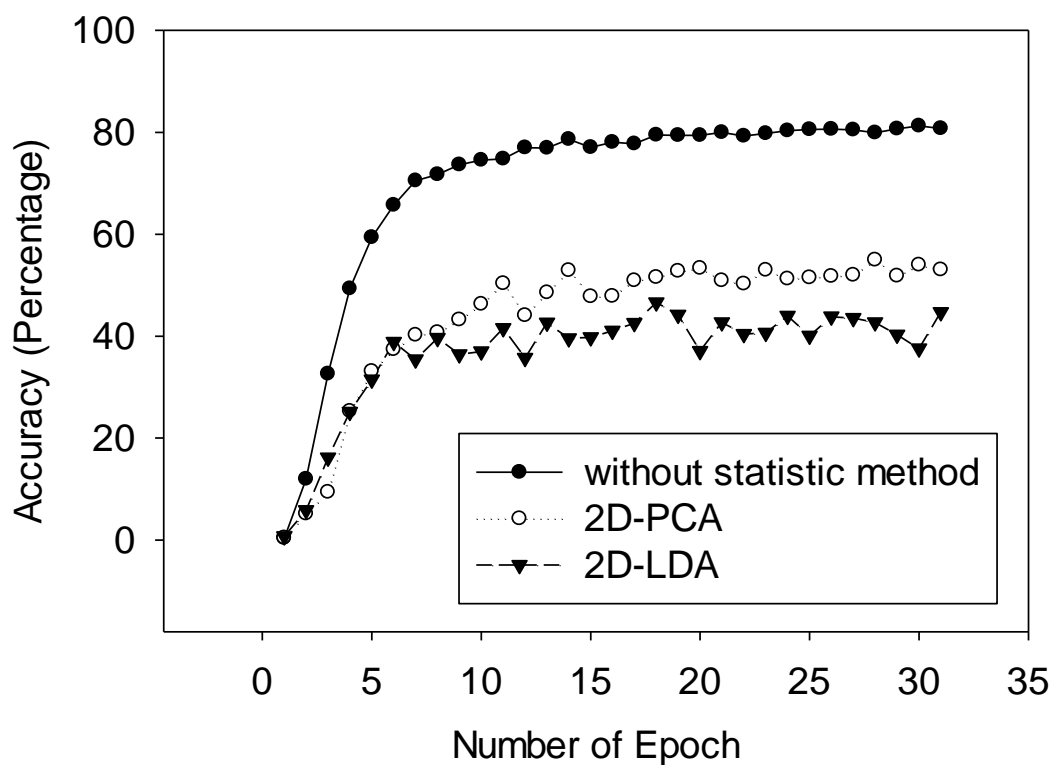
สรุปได้ว่า ค่าไฮเพอร์พารามิเตอร์ของโครงข่ายประสาทโดยใช้แบบจำลองของ Alexnet ในการฝึกสอนฐานข้อมูล GTAV ที่เหมาะสมกับการใช้ขนาดกลุ่มภาพฝึกสอน 128 ภาพต่อกลุ่ม และใช้อัตราการเรียนรู้ 0.001 จะมีประสิทธิภาพดีที่สุด

4.3 ผลการทดลองการเปรียบเทียบแบบใช้และไม่ใช้วิธีเชิงสถิติร่วมกับโครงข่ายประสาทแบบสังวัตนาการ โดยใช้แบบจำลองของ Alexnet

ผลการทดลองการเปรียบเทียบระหว่างแบบใช้และไม่ใช้วิธี 2DPCA และ 2DLDA ร่วมกับโครงข่ายประสาทเทียมแบบสังวัตนาการ โดยใช้แบบจำลองของ Alexnet มีรายละเอียดดังต่อไปนี้

4.3.1 ผลการทดลองจากฐานข้อมูล AR

ผลการทดลองการเปรียบเทียบระหว่างแบบใช้และไม่ใช้วิธี 2DPCA และ 2DLDA ร่วมกับโครงข่ายประสาทเทียมแบบสังวัตนาการ โดยใช้แบบจำลองของ Alexnet จากฐานข้อมูล AR มีรายละเอียดดังรูปที่ 4.10



รูปที่ 4.10 กราฟแสดงความแม่นยำระหว่างการใช้และไม่ใช้วิธีเชิงสถิติร่วมกับโครงข่ายประสาทเทียมแบบสังวัตนาการ โดยใช้แบบจำลอง Alexnet จากฐานข้อมูล AR

จากรูปที่ 4.10 แสดงความแม่นยำระหว่างการใช้และไม่ใช้วิธีเชิงสถิติร่วมกับโครงข่ายประสาทเทียมแบบสังวัตนาการ โดยใช้แบบจำลอง Alexnet จากฐานข้อมูล AR โดยกำหนดค่าอัตราเร็วในการเรียนรู้ที่ 0.001 กำหนดขนาดกลุ่มภาพฝึกสอน 32 ภาพต่อกลุ่ม และกำหนดจำนวนรอบของการฝึกสอน 30 รอบ

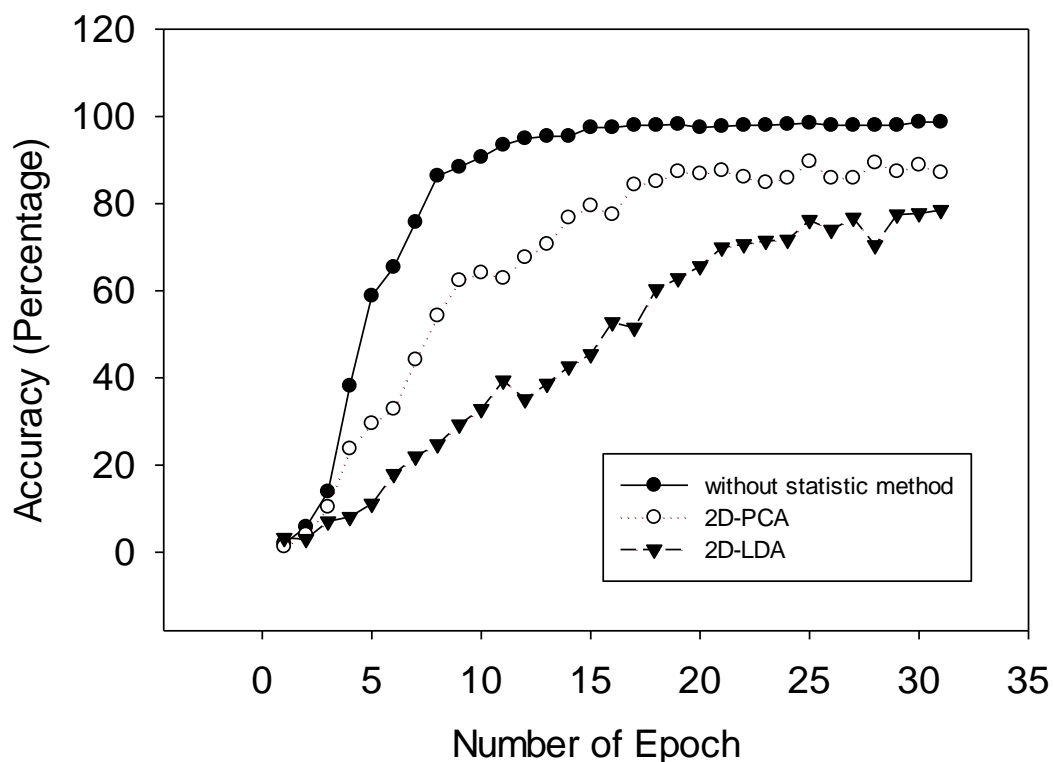
ผลการทดลองพบว่าการใช้วิธีเชิงสถิติจะต้องใช้ภาพต้นฉบับ ให้ความแม่นยำที่ร้อยละ 80.77 กราฟสีแดงเป็นการใช้วิธี 2DPCA ให้ความแม่นยำร้อยละ 53.06 กราฟสีเขียวเป็นการใช้วิธี 2DLDA

ให้ความแม่นยำร้อยละ 49.71 โดยที่การใช้และไม่ใช้วิธีเชิงสถิติใช้เวลาเฉลี่ย 2,175 วินาที ในการฝึกสอนที่ไม่แตกต่างกัน

สรุปได้ว่า เมื่อเทียบความแม่นยำระหว่าง 2DPCA และ 2DLDA มีค่าความแม่นยำที่ไม่แตกต่างกัน เนื่องจากค่าที่ใช้ในการฝึกสอนเป็นค่าของเวกเตอร์เฉพาะ ขนาด 227×227 จุดภาพ แต่เมื่อเทียบกับการไม่ใช้วิธีเชิงสถิติแล้วพบว่าความแม่นยำต่ำกว่าแบบไม่ใช้วิธีเชิงสถิติอยู่มาก เนื่องจากการไม่ใช้วิธีเชิงสถิติ คือการใช้รูปภาพบุคคลโดยตรง ขนาด 227×227 จุดภาพ และแบบจำลองของ Alexnet ซึ่งถูกออกแบบมาเพื่อฝึกสอนรูปโดยเฉพาะ

4.3.2 ผลการทดลองจากฐานข้อมูล GTAV

การทดลองการเปรียบเทียบระหว่างแบบใช้และไม่ใช้วิธี 2DPCA และ 2DLDA ร่วมกับโครงข่ายประสาทเทียมแบบสังวัตนาการโดยใช้แบบจำลองของ Alexnet จากฐานข้อมูล GTAV แสดงผลลัพธ์ดังรูปที่ 4.11 ดังต่อไปนี้



รูปที่ 4.11 แสดงความแม่นยำระหว่างการใช้และไม่ใช้วิธีเชิงสถิติร่วมกับโครงข่ายประสาทแบบสังวัตนาการโดยใช้แบบจำลองของ Alexnet

จากรูปที่ 4.11 กราฟแสดงความแม่นยำระหว่างการใช้และไม่ใช้วิธีเชิงสถิติร่วมกับโครงข่ายประสาทเทียมแบบสังวัตนาการโดยใช้แบบจำลอง Alexnet จากฐานข้อมูล GTAV โดยกำหนดค่าอัตราเร็วในการเรียนรู้ที่ 0.001 กำหนดขนาดกลุ่มภาพฝึกสอน 128 ภาพต่อกลุ่ม และกำหนดจำนวนรอบของการฝึกสอน 30 รอบ

ผลการทดลองพบว่า การไม่ใช้วิธีเชิงสถิติ ซึ่งต้องใช้ภาพต้นฉบับ ให้ความแม่นยำที่ร้อยละ 98.74 การใช้วิธี 2DPCA จะให้ความแม่นยำที่ร้อยละ 86.36 และการใช้วิธี 2DLDA จะให้ความแม่นยำที่ร้อยละ 78.54 ตามลำดับ โดยที่ทั้งสามวิธีใช้เวลาในการฝึกสอนเฉลี่ยที่ 345 วินาที

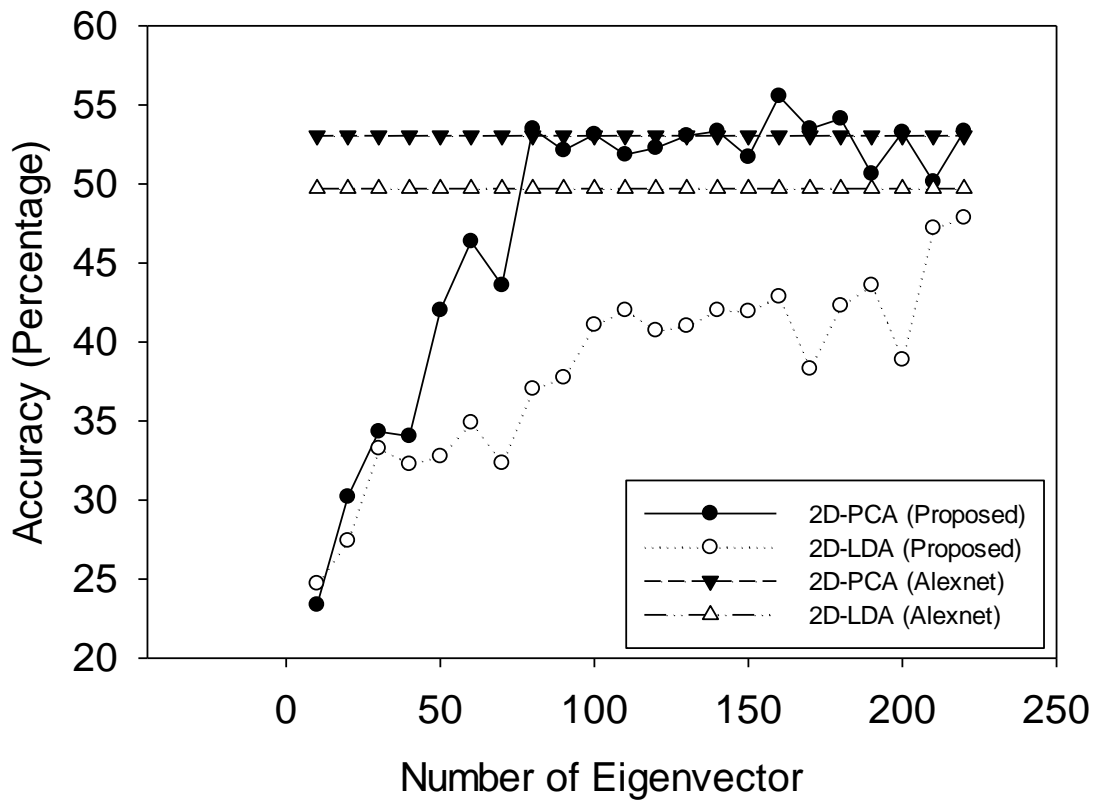
สรุปได้ว่า วิธี 2DPCA ให้ประสิทธิภาพที่ดีกว่า 2DLDA แต่ให้ความแม่นยำน้อยกว่าการไม่ใช้วิธีเชิงสถิติ เนื่องจากวิธีเชิงสถิติใช้เวกเตอร์เฉพาะขนาด 227×227 จุดภาพ แต่ในขณะที่การไม่ใช้วิธีเชิงสถิติต้องใช้รูปภาพต้นฉบับ ที่มีขนาด 227×227 จุดภาพ แทนการใช้เวกเตอร์เฉพาะตามแบบจำลองของ Alexnet ที่ถูกออกแบบมาเพื่อประมวลผลภาพโดยเฉพาะ จึงมีความแม่นยำที่สูงกว่า

4.4 ผลการทดลองการเปรียบเทียบระหว่างวิธี 2DPCA และ 2DLDA ร่วมกับโครงข่ายประสาทแบบสังวัตนาการ โดยใช้แบบจำลองของ Alexnet กับผู้วิจัย

การทดลองการเปรียบเทียบระหว่างวิธี 2DPCA และ 2DLDA ร่วมกับโครงข่ายประสาทแบบสังวัตนาการโดยใช้แบบจำลองของ Alexnet กับผู้วิจัย มีรายละเอียดดังต่อไปนี้

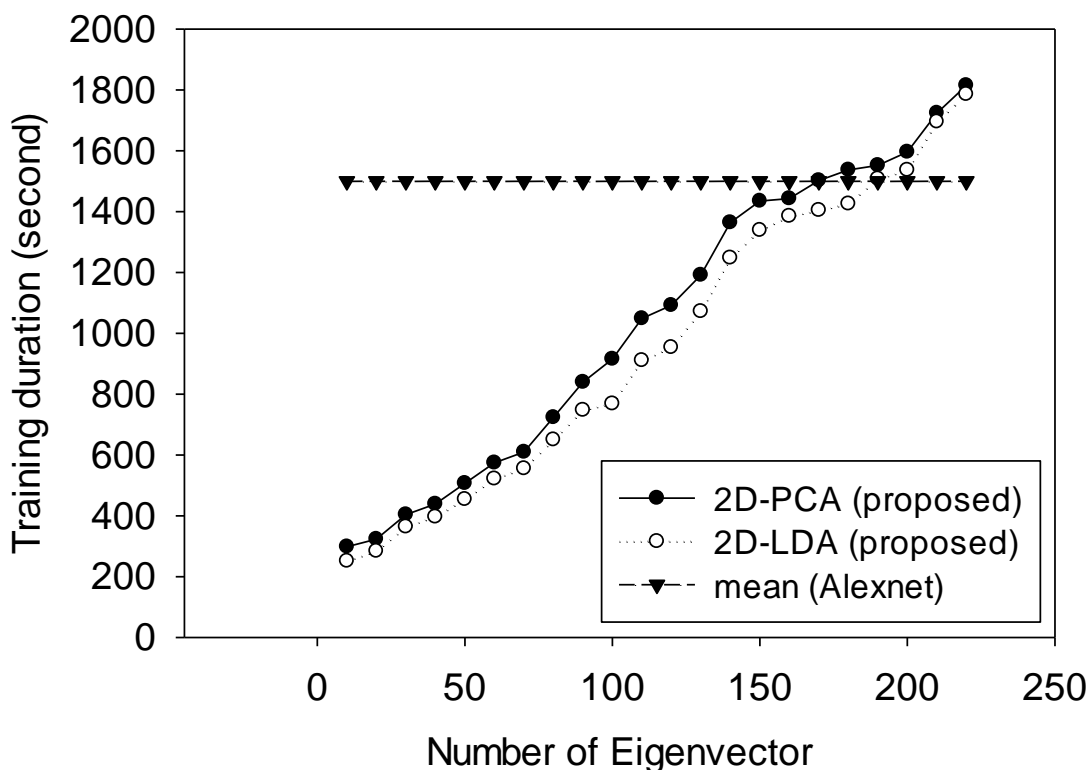
4.4.1 ผลการทดลองจากฐานข้อมูล AR

การทดลองระหว่าง 2DPCA และ 2DLDA โดยใช้ทุก 10 มิติของเวกเตอร์เฉพาะ จาก 10 ถึง 220 มิติ และแสดงกราฟเปรียบเทียบระหว่างวิธี 2DPCA และ 2DLDA ร่วมกับโครงข่ายประสาทแบบสังวัตนาการ โดยใช้แบบจำลองของ Alexnet และผู้วิจัยจากฐานข้อมูล AR มีรายละเอียดดังรูปที่ 4.14 ดังต่อไปนี้



รูปที่ 4.12 กราฟแสดงความแม่นยำระหว่างการใช้เชิงสถิติร่วมกันโครงข่ายประสาทแบบสังวัตนาการ โดยใช้แบบจำลองผู้วิจัย

จากรูปที่ 4.12 กราฟแสดงความแม่นยำระหว่างวิธี 2DPCA และ 2DLDA ร่วมกับโครงข่ายประสาทแบบสังวัตนาการโดยใช้แบบจำลองของ Alexnet และผู้วิจัย พบว่าการใช้วิธี 2DPCA ของผู้วิจัยที่ขนาดเวกเตอร์เฉพาะ 160 มิติ ให้ความแม่นยำร้อยละ 55.55 และการใช้วิธี 2DLDA ของผู้วิจัยที่ขนาดเวกเตอร์เฉพาะ 220 มิติ ให้ความแม่นยำร้อยละ 47.86 และการใช้ เปรียบเทียบกับการใช้วิธี 2DPCA ของ Alexnet ให้ความแม่นยำร้อยละ 53.06 และการใช้วิธี 2DLDA ของ Alexnet ที่ ให้ความแม่นยำร้อยละ 49.71



รูปที่ 4.13 แสดงระยะเวลาการฝึกสอนระหว่างการใช่และไม่ใช้วิธีเชิงสถิติร่วมกันโครงข่ายประสาทแบบสังวัตนาการโดยใช้แบบจำลองผู้วิจัย

จากรูปที่ 4.13 แสดงเวลาที่ใช้ในการฝึกสอนทั้งวิธี 2DPCA และ 2DLDA พบว่าใช้เวลาในการฝึกสอนไม่แตกต่างกัน แต่ถ้าใช้ขนาดมิติต่ำกว่า 160 ของมิติเวกเตอร์เฉพาะจะถือว่าใช้เวลาในการฝึกสอนน้อยกว่าแบบจำลองของ Alexnet

สาเหตุที่เป็นกราฟเส้นตรงของแบบจำลอง Alexnet เดิม เนื่องจากไม่สามารถใช้ข้อมูลที่ถูกลดขนาดมิติลงได้ จึงจำเป็นต้องใช้เวกเตอร์เฉพาะที่มีขนาด 227×227 จุดภาพ เท่านั้น ส่วนแบบจำลองของผู้วิจัยเป็นการดัดแปลงแบบจำลองของ Alexnet ให้สามารถรับเวกเตอร์เฉพาะได้ทุกขนาด

เมื่อเปรียบเทียบวิธี 2DPCA ของ Alexnet กับผู้วิจัย พบว่าที่ขนาด 80 มิติของเวกเตอร์เฉพาะให้ความแม่นยำใกล้เคียงกับแบบจำลองของ Alexnet เดิม และเมื่อเทียบเวลาที่ใช้ฝึกสอนจากรูปที่ 4.15 พบว่าวิธี 2DPCA ใช้เวลาฝึกสอน 722.93 วินาที และแบบจำลอง Alexnet ใช้เวลาฝึกสอนเฉลี่ยของทั้ง 2 วิธีใช้เวลาใกล้เคียง 1,500 วินาที ซึ่งถือว่าใช้เวลาในการฝึกสอนแตกต่างกันมาก

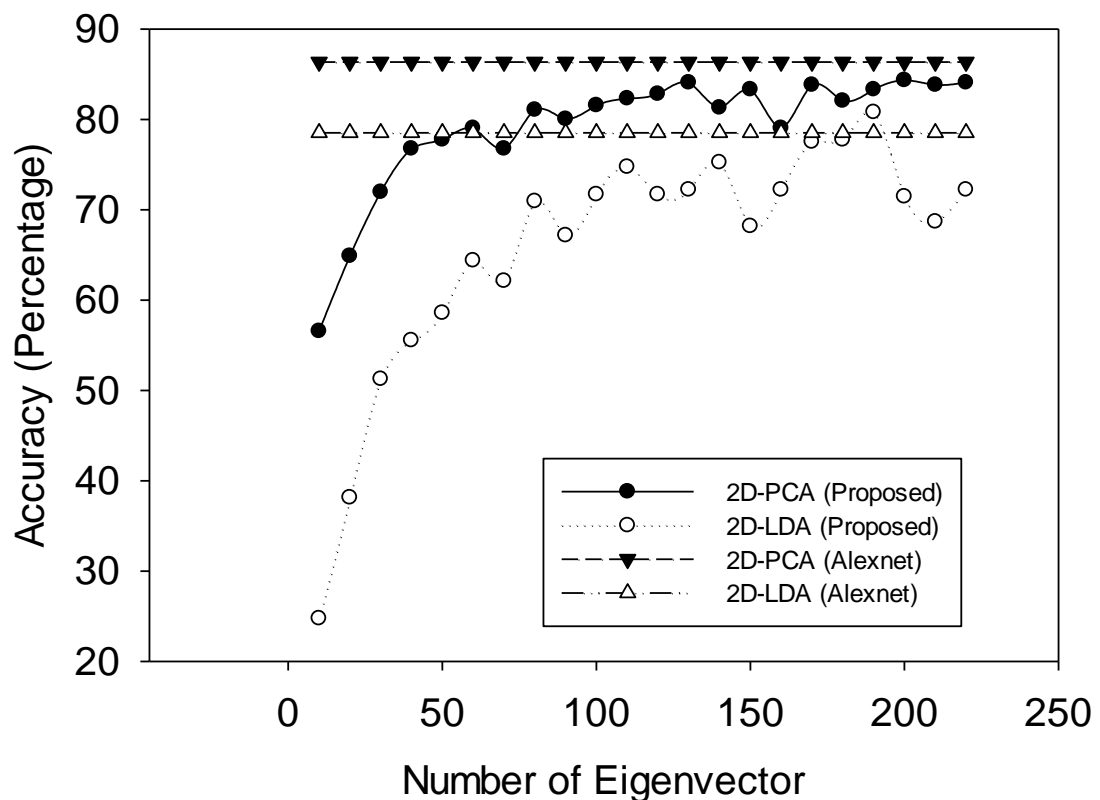
เมื่อเปรียบเทียบวิธี 2DLDA ของ Alexnet กับผู้วิจัย พบว่าการใช้วิธี 2DLDA แบบจำลองของผู้วิจัย ที่ขนาด 90 มิติของเวกเตอร์เฉพาะ ให้ความแม่นยำใกล้เคียงกับแบบจำลองของ Alexnet เดิม และเมื่อเทียบเวลาที่ใช้ฝึกสอนจากรูปที่ 4.15 พบว่าวิธี 2DPCA ใช้เวลาฝึกสอน 747.31 วินาที และ

แบบจำลอง Alexnet ใช้เวลาฝึกสอนเฉลี่ยของทั้ง 2 วิธีใช้เวลาใกล้เคียง 1,500 วินาที ซึ่งถือว่าใช้เวลาในการฝึกสอนแตกต่างกันมาก

สรุปได้ว่าการใช้เวกเตอร์คุณลักษณะด้วยวิธี 2DPCA และ 2DLDA ร่วมกับโครงข่ายประสาทแบบสังวัตนาการโดยใช้แบบจำลองของผู้วิจัย พบว่าวิธี 2DPCA ตั้งแต่ขนาด 70 มิติของเวกเตอร์เฉพาะขึ้นไปให้ความแม่นยำที่สูงกว่าวิธี 2DLDA และพบว่าทั้ง 2 วิธีให้ความแม่นยำไม่แตกต่างกันมากนักเมื่อเทียบกับแบบจำลองของ Alexnet เดิม แต่ใช้เวลาในการฝึกสอนน้อยกว่าแบบจำลองของ Alexnet มาก

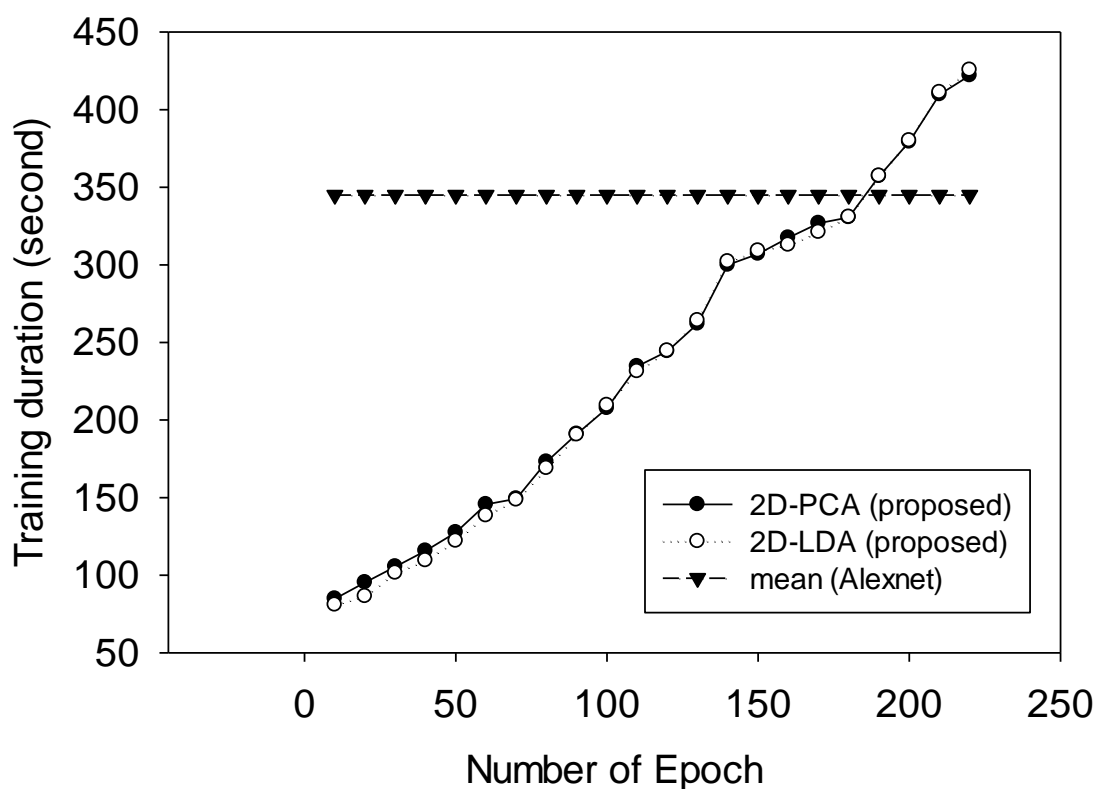
4.4.2 ผลการทดลองจากฐานข้อมูล GTAV

การทดลองการเปรียบเทียบระหว่างวิธี 2DPCA และ 2DLDA โดยใช้ทุกขนาด 10 มิติของเวกเตอร์เฉพาะ จาก 10 ถึง 220 มิติ ร่วมกับโครงข่ายประสาทแบบสังวัตนาการโดยใช้แบบจำลองของ Alexnet และผู้วิจัย จากฐานข้อมูล GTAV มีรายละเอียดดังรูปที่ 4.16 ดังต่อไปนี้



รูปที่ 4.14 กราฟแสดงความแม่นยำระหว่างการใช่วิธี 2DPCA และ 2DLDA ร่วมกับโครงข่ายประสาทแบบสังวัตนาการโดยใช้แบบจำลองของ Alexnet และ ผู้วิจัย

จากรูปที่ 4.14 แสดงความแม่นยำระหว่างการใช้ 2DPCA และ 2DLDA ร่วมกับโครงข่ายประสาทแบบสังวัตนาการโดยใช้แบบจำลองของ Alexnet และผู้วิจัย พบว่าการใช้วิธี 2DPCA ของผู้วิจัย ใช้ขนาด 200 มิติของเวกเตอร์เฉพาะ ให้ความแม่นยำร้อยละ 84.34 และการใช้วิธี 2DLDA ของผู้วิจัย ใช้ขนาด 190 มิติของเวกเตอร์เฉพาะ ให้ความแม่นยำร้อยละ 80.80 เปรียบเทียบกับการใช้วิธี 2DPCA ของ Alexnet ให้ความแม่นยำร้อยละ 86.36 และการใช้วิธี 2DLDA ของ Alexnet ที่ให้ความแม่นยำร้อยละ 78.54



รูปที่ 4.15 กราฟแสดงระยะเวลาการฝึกสอนระหว่างการใช้วิธีเชิงสถิติร่วมกันโครงข่ายประสาทแบบสังวัตนาการโดยใช้แบบจำลองผู้วิจัย

จากรูปที่ 4.15 แสดงเวลาที่ใช้ในการฝึกสอนทั้งวิธี 2DPCA และ 2DLDA พบว่าใช้เวลาในการฝึกสอนไม่แตกต่างกัน แต่ถ้าใช้ขนาดมิติต่ำกว่า 130 ของมิติเวกเตอร์เฉพาะจะถือว่าใช้เวลาในการฝึกสอนน้อยกว่าแบบจำลองของ Alexnet

ในการฝึกสอนทั้งวิธี 2DPCA และ 2DLDA ผลการทดลองพบว่า ทั้งสองวิธีใช้เวลาฝึกสอนไม่แตกต่างกัน แต่ถ้าใช้ขนาดมิติต่ำกว่า 130 ของมิติเวกเตอร์เฉพาะจะถือว่าใช้เวลาฝึกสอนน้อยกว่าแบบจำลองของ Alexnet

สาเหตุที่กราฟของแบบจำลอง Alexnet เดิม เป็นเส้นตรงในแนวนอน เนื่องจากไม่สามารถใช้ข้อมูลที่ถูกลดขนาดมิติลงได้ จึงจำเป็นต้องใช้เวกเตอร์เฉพาะที่มีขนาด 227×227 จุดภาพเท่านั้น ส่วนแบบจำลองของผู้วิจัยเป็นการดัดแปลงแบบจำลองของ Alexnet ให้สามารถรับเวกเตอร์เฉพาะได้ทุกขนาด

ดังนั้น เมื่อเปรียบเทียบวิธี 2DPCA ของ Alexnet กับผู้วิจัย พบว่า แบบจำลองของผู้วิจัยให้ความแม่นยำที่ร้อยละ 84.34 ในขณะที่แบบจำลองของอเล็กเน็ตให้ความแม่นยำที่ร้อยละ 86.36 ซึ่งถือว่าแบบจำลองของผู้วิจัยให้ความแม่นยำน้อยกว่าแบบจำลองของ Alexnet

เมื่อเปรียบเทียบวิธี 2DLDA ของ Alexnet กับผู้วิจัย พบว่า การใช้วิธี 2DLDA แบบจำลองของผู้วิจัย ที่ขนาด 170 มิติของเวกเตอร์เฉพาะ ให้ความแม่นยำใกล้เคียงกับแบบจำลองของ Alexnet เดิม และเมื่อเทียบเวลาที่ใช้ฝึกสอนจากรูปที่ 4.15 พบว่า วิธี 2DPCA ใช้เวลาฝึกสอน 209.48 วินาที และแบบจำลอง Alexnet ใช้เวลาฝึกสอนเฉลี่ยของทั้ง 2 วิธีใช้เวลาใกล้เคียง 275 วินาที ซึ่งถือว่าใช้เวลาในการฝึกสอนแตกต่างกัน

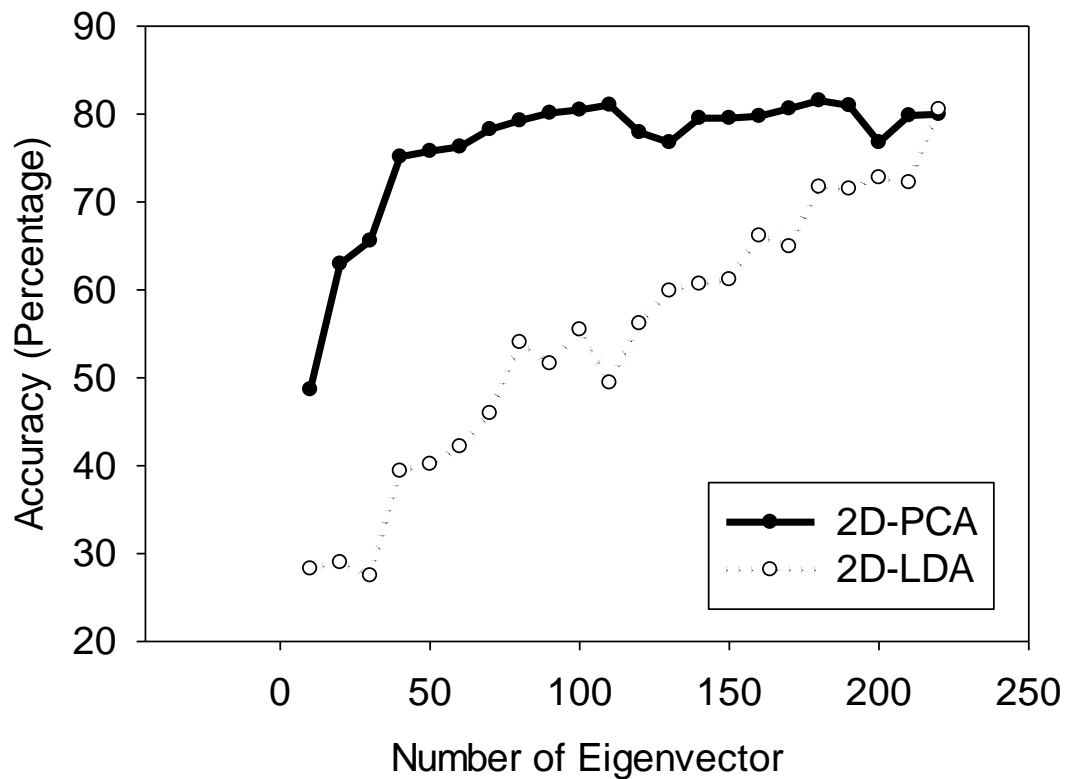
สรุปได้ว่า การใช้เวกเตอร์คุณลักษณะด้วยวิธี 2DPCA และ 2DLDA ร่วมกับโครงข่ายประสาทเทียมแบบสังวัตนาการ โดยใช้แบบจำลองของผู้วิจัย พบว่าวิธี 2DPCA ให้ความแม่นยำที่สูงกว่าวิธี 2DLDA ในทุกมิติเวกเตอร์เฉพาะ

4.5 ผลการทดลองการเปรียบเทียบระหว่างการสร้างคืนภาพด้วยวิธี 2DPCA และ 2DLDA ร่วมกับโครงข่ายประสาทเทียมแบบสังวัตนาการ โดยใช้แบบจำลอง Alexnet

ผลการทดลองการเปรียบเทียบการสร้างคืนภาพด้วยวิธี 2DPCA และ 2DLDA ร่วมกับโครงข่ายประสาทเทียมแบบสังวัตนาการจากฐานข้อมูล AR และ GTAV มีรายละเอียดดังต่อไปนี้

4.5.1 ผลการทดลองจากฐานข้อมูล AR

ผลการทดลองการสร้างคืนภาพด้วยวิธี 2DPCA และ 2DLDA โดยใช้ทุก 10 มิติของเวกเตอร์เฉพาะ จาก 10 ถึง 220 มิติ ร่วมกับโครงข่ายประสาทเทียมแบบสังวัตนาการ จากฐานข้อมูล AR มีรายละเอียดดังรูปที่ 4.16

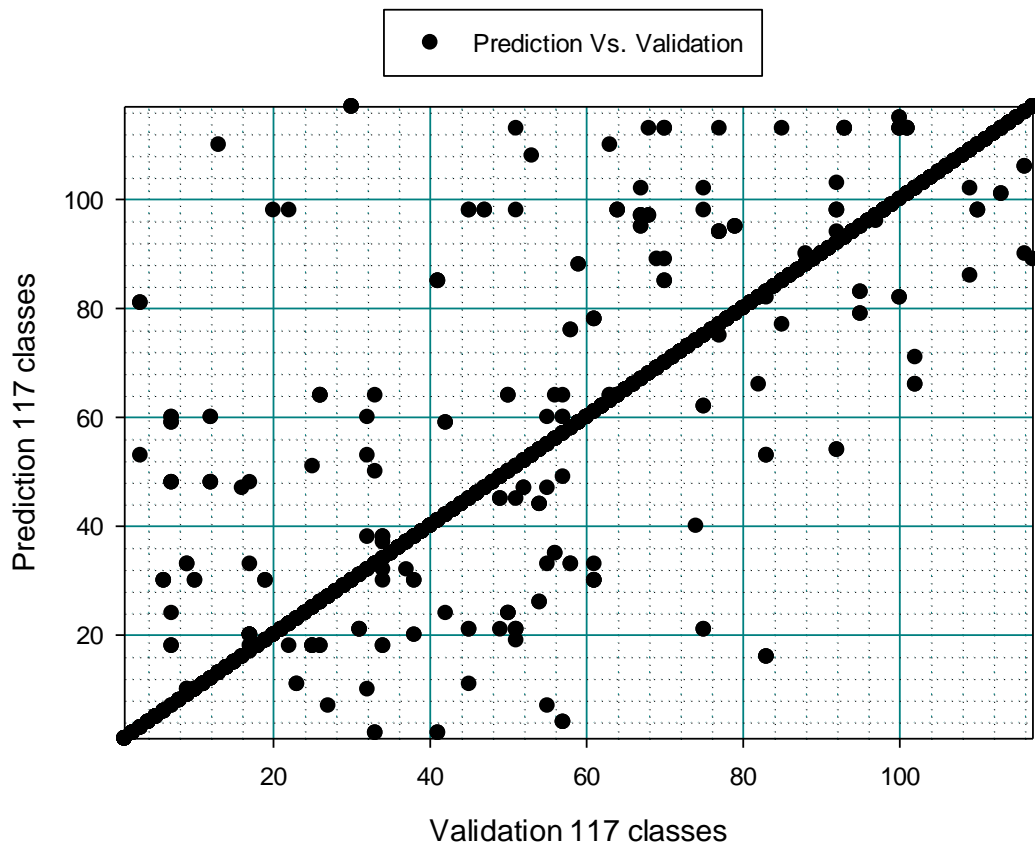


รูปที่ 4.16 กราฟแสดงการเปรียบเทียบความแม่นยำในการสร้างคืนภาพของวิธีเชิงสถิติทุก 10 มิติของเวกเตอร์เฉพาะ จากฐานข้อมูล AR

จากรูปที่ 4.16 แสดงความแม่นยำของการสร้างคืนภาพจากฐานข้อมูล AR ที่ทำการบันทึกผลทุก 10 มิติของเวกเตอร์เฉพาะจาก 10 ถึง 220 มิติ ที่ใช้ 1,404 รูปทดสอบ โดยที่กำหนดอัตราเร็วในการเรียนรู้ที่ 0.001 กำหนดขนาดกลุ่มภาพฝึกสอนจำนวน 32 ภาพต่อกลุ่ม และกำหนดรอบของจำนวนการฝึกสอน 30 รอบ

ผลการทดลองพบว่าวิธี 2DPCA ที่ใช้ 180 มิติของเวกเตอร์เฉพาะให้ความแม่นยำที่ร้อยละ 81.55 และ วิธี 2DLDA ที่ใช้ 220 มิติของเวกเตอร์เฉพาะให้ความแม่นยำที่ร้อยละ 80.56 เนื่องจากรูปที่เกิดจากการสร้างคืนภาพ มีขนาด 227×227 จุดภาพ ดังนั้น จึงใช้เวลาฝึกสอนเฉลี่ย 2,175 วินาที ในทุกมิติของเวกเตอร์เฉพาะ

สรุปได้ว่า วิธี 2DPCA ให้ความแม่นยำโดยรวมสูงกว่าวิธี 2DLDA แต่เมื่อมิติเวกเตอร์เฉพาะที่มีค่าสูง จะให้ความแม่นยำใกล้เคียงกัน

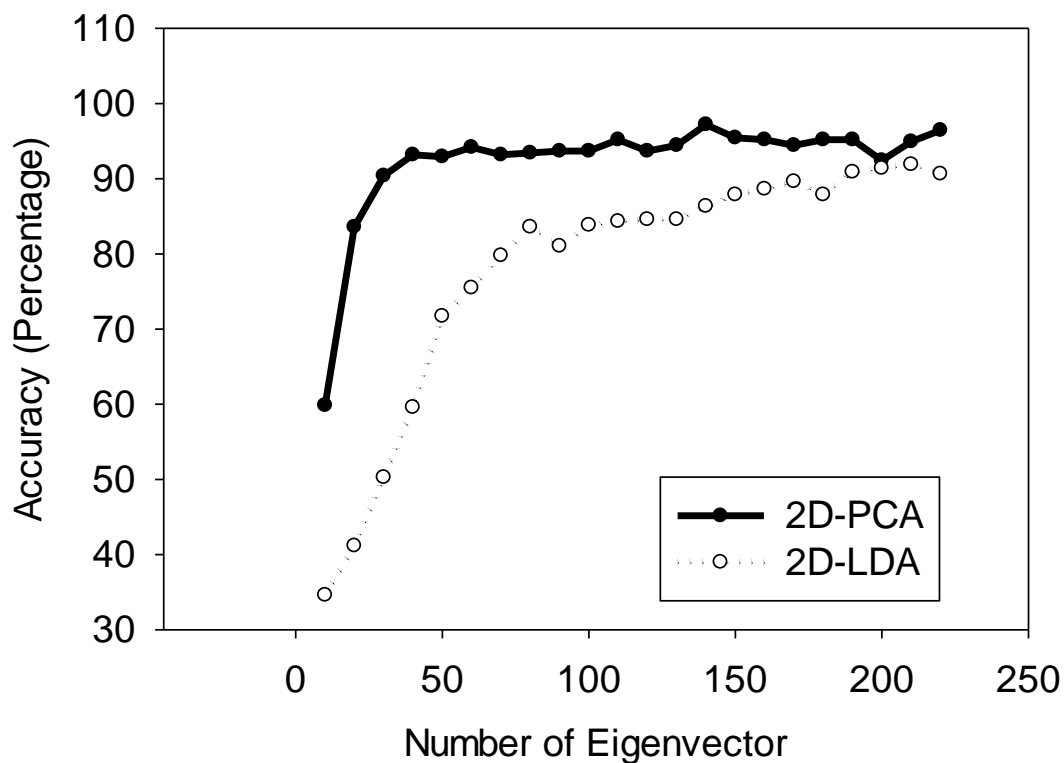


รูปที่ 4.17 กราฟแสดงการเปรียบเทียบระหว่างบุคคลต่อการทำนายผล จากฐานข้อมูล AR ที่ร้อยละ 81.55

จากรูปที่ 4.17 แสดงอัตราส่วนระหว่างจำนวนคนที่ทดสอบกับจำนวนคนที่ทำนายผล เมื่อใช้รูปทดสอบจำนวน 1,404 รูป กับวิธี 2DPCA พบว่าเมื่อนำมาสร้างกราฟและผลการทำนายถูกคนจุดดำจะเรียงกันเป็นเส้นทะแยงมุม และมีผลความแม่นยำที่ร้อยละ 81.55

4.5.2 ผลการทดลองจากฐานข้อมูล GTAV

ผลการทดลองการสร้างคืนภาพด้วยวิธี 2DPCA และ 2DLDA โดยใช้ทุก 10 มิติของเวกเตอร์เฉพาะจาก 10 ถึง 220 มิติ ร่วมกับโครงข่ายประสาทเทียมแบบสั่งวัฒนาการ จากฐานข้อมูล GTAV มีรายละเอียดดังรูปที่ 4.18 ต่อไปนี้

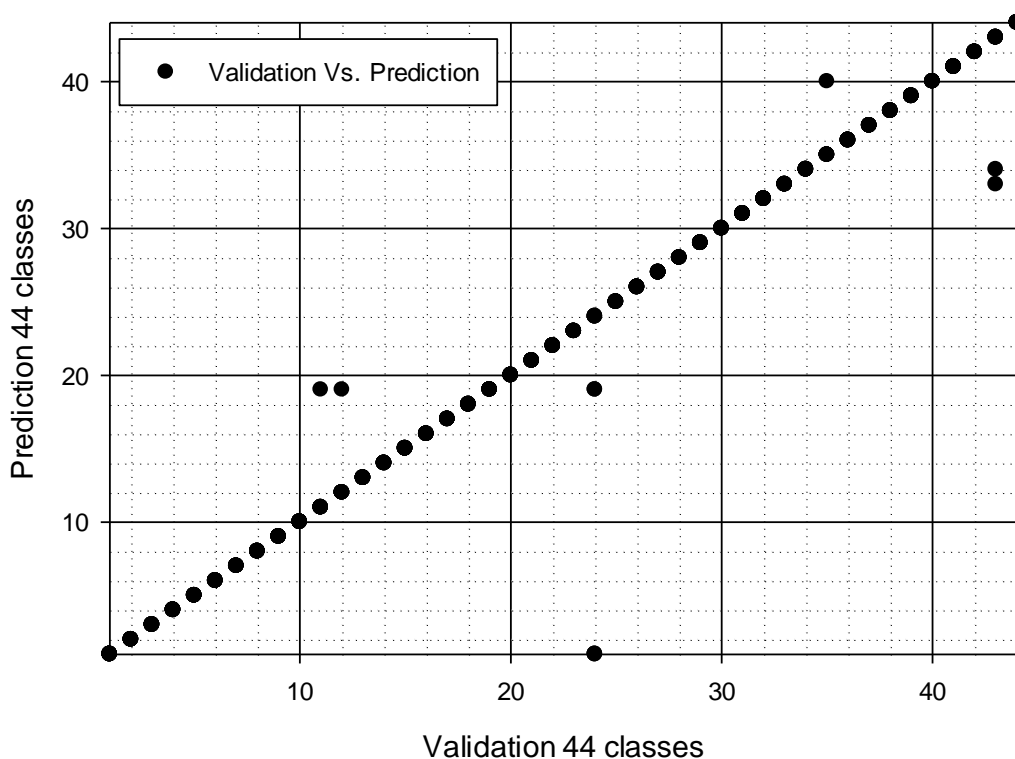


รูปที่ 4.18 กราฟแสดงการเปรียบเทียบความแม่นยำของการสร้างคืนภาพด้วยวิธีเชิงสถิติทุก 10 มิติของเวกเตอร์เฉพาะ จากฐานข้อมูล GTAV

จากรูปที่ 4.18 แสดงการเปรียบเทียบความแม่นยำทั้งสองวิธีของการสร้างคืนภาพจากรูปที่มีขนาด 227×227 พิกเซล โดยที่กำหนดอัตราเร็วในการเรียนรู้ที่ 0.001 กำหนดขนาดกลุ่มภาพฝึกสอนจำนวน 128 ภาพต่อกลุ่ม และกำหนดรอบของจำนวนการฝึกสอน 30 รอบ

ผลการทดลองพบว่า วิธี 2DPCA ที่เวกเตอร์เฉพาะ 220 มิติ ให้ความแม่นยำสูงสุดที่ร้อยละ 96.46 และวิธี 2DLDA ที่เวกเตอร์เฉพาะ 210 มิติ ให้ความแม่นยำสูงสุดที่ร้อยละ 91.92 โดยทั้งสองใช้เวลาเฉลี่ย 345 วินาที

สรุปได้ว่า วิธี 2DPCA ให้ความแม่นยำโดยรวมสูงกว่าวิธี 2DLDA แต่เมื่อมิติเวกเตอร์เฉพาะที่มีค่าสูง จะให้ความแม่นยำใกล้เคียงกัน



รูปที่ 4.19 กราฟแสดงการเปรียบเทียบระหว่างบุคคลต่อการทำนายผล จากฐานข้อมูล GTAV ที่ร้อยละ 96.46

จากรูปที่ 4.19 แสดงอัตราส่วนระหว่างจำนวนคนที่ทดสอบกับจำนวนคนที่ทำนายผล จากการใช้รูปทดสอบ จำนวน 396 รูป และนำมาสร้างกราฟ พบว่าผลการทำนายถูกคนจุดดำจะเรียงกัน เส้นตรงตามแนวทะแยงมุม และจะให้ความแม่นยำที่ร้อยละ 96.46

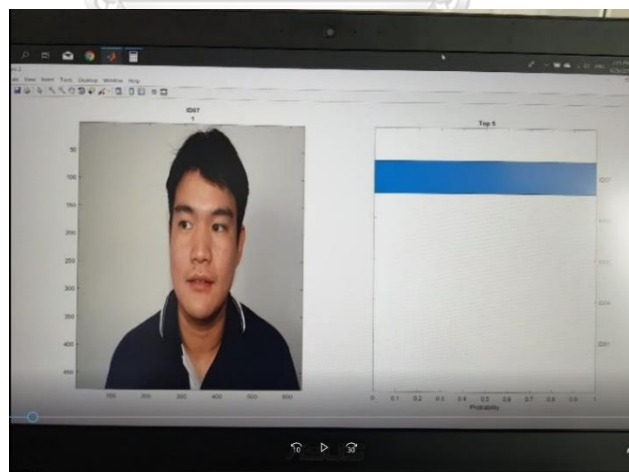
4.6 ผลการทดลองนำระบบการรู้จำใบหน้าแบบมีสิ่งบดบังไปใช้ในชีวิตจริง

ผู้วิจัยได้ทำการทดลองโดยนำระบบการรู้จำใบหน้าไปใช้ในชีวิตจริงกับโปรแกรมรู้จำใบหน้าแบบมีสิ่งบดบัง โดยใช้โครงข่ายประสาทเทียมแบบสังวัตนาการของแบบจำลองของ Alexnet และได้เก็บรูปภาพบุคคลจากชมรมบัณฑิตศึกษา คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย จำนวน 9 คน คนละ 45 ภาพ ที่ถ่ายในมุม 0° , $\pm 30^\circ$, $\pm 45^\circ$, $\pm 60^\circ$ และ 90° ตามลำดับ มีรายละเอียดดังรูปที่ 4.25 ต่อไปนี้



รูปที่ 4.20 ตัวอย่างการเก็บข้อมูลรูปภาพของบุคคลหนึ่งที่มีลักษณะแตกต่างกันออกไป

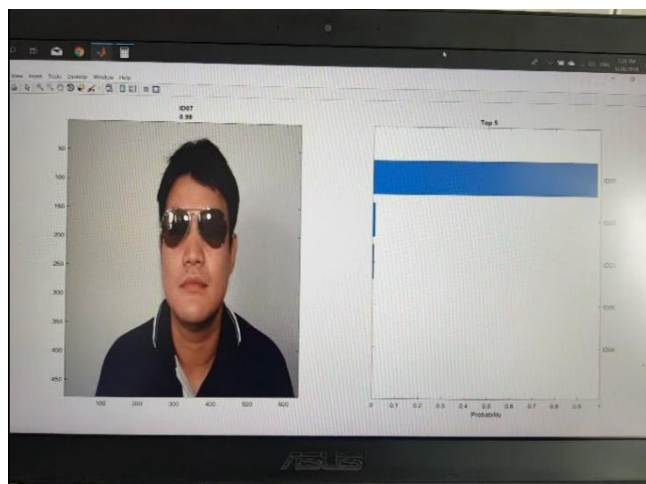
จากรูปที่ 4.20 แสดงการแบ่งกลุ่มตัวอย่างออกเป็น 5 กลุ่ม โดยกลุ่มแรกกับกลุ่มที่สองจะเป็นภาพที่มีลักษณะใกล้เคียงกัน กลุ่มที่สามมีการส่องสว่างที่บริเวณใบหน้า กลุ่มที่สี่มีการสวมแว่นกันแดด กลุ่มที่ห้ามีสวมผ้าปิดปาก จากนั้นนำภาพสองกลุ่มแรกไปฝึกสอน และ สามกลุ่มสุดท้ายไปทดสอบ ด้วยการติดตั้งกล้อง (web cam) และถ่ายรูปรูปประมวลผลแบบเวลาจริง (real-time) และเรียงความเป็นไปได้จากค่าแม่นยำสูงสุด 5 คนแรก



รูปที่ 4.21 แสดงการทดลองโดยใช้ใบหน้าปกติ

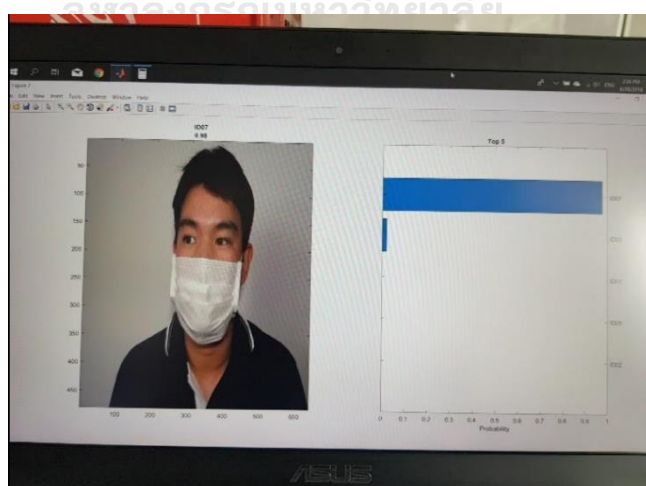
จากรูปที่ 4.21 แสดงผลการทดลองด้วยใบหน้าปกติ ด้านซ้ายคือรูปที่ทำการถ่ายภาพปัจจุบัน ในเวลาจริงและนำไปเปรียบเทียบกับโครงข่ายประสาทที่ผ่านการฝึกสอนแล้ว และทำการวัด

ความเป็นไปได้ของบุคคลใดในฐานข้อมูล พบว่า ในขณะที่ใช้ใบหน้าปกติโปรแกรมให้ผลลัพธ์ที่ตรงกับตัวบุคคลจริง



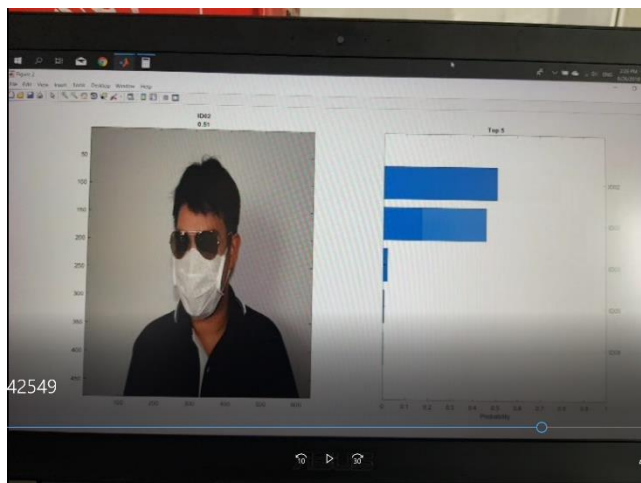
รูปที่ 4.22 แสดงการทดลองโดยสวมแว่นกันแดด

จากรูปที่ 4.22 ผลการทดลองการสวมแว่นกันแดด โดยด้านซ้ายคือรูปที่ทำการถ่ายภาพปัจจุบันในเวลาจริงและนำไปเปรียบเทียบกับโครงข่ายประสาทที่ผ่านการฝึกสอนแล้ว ทำการวัดความเป็นไปได้ของบุคคลใดในฐานข้อมูล พบว่า ในขณะที่ใช้ใบหน้าปกติโปรแกรมให้ผลลัพธ์ที่ตรงกับตัวบุคคลจริง



รูปที่ 4.23 แสดงการทดลองโดยสวมผ้าปิดปาก

จากรูปที่ 4.23 ผลการทดลองการสวมผ้าปิดปาก โดยภาพทางด้านซ้ายคือรูปที่ทำการถ่ายภาพ ปัจจุบันในเวลาจริงและนำไปเปรียบเทียบกับโครงข่ายประสาทที่ผ่านการฝึกสอนแล้ว ทำการวัดความเป็นไปได้ของบุคคลใดในฐานข้อมูล พบว่าในขณะที่ใช้ใบหน้าปกติโปรแกรมให้ผลลัพธ์ที่ตรงกับตัวบุคคลจริง



รูปที่ 4.24 แสดงการทดลองโดยสวมแว่นกันแดดและผ้าปิดปาก

จากรูปที่ 4.24 ผลการทดลองการสวมผ้าปิดปากและแว่นกันแดด โดยภาพทางด้านซ้ายคือรูปที่ทำการถ่ายภาพปัจจุบันในเวลาจริงและนำไปเปรียบเทียบกับโครงข่ายประสาทที่ผ่านการฝึกสอนแล้ว ทำการวัดความเป็นไปได้ของบุคคลใดในฐานข้อมูล พบว่า ในขณะที่ใช้ใบหน้าปกติโปรแกรมให้ผลลัพธ์ที่ไม่ตรงกับตัวบุคคลจริง

บทที่ 5

สรุป อภิปรายผล และข้อเสนอแนะ

วิทยานิพนธ์ฉบับนี้ ทำการศึกษา ทดลองและเปรียบเทียบประสิทธิภาพระหว่าง 2DPCA และ 2DLDA ในการรู้จำใบหน้าแบบมีสิ่งบดบัง โดยมีวัตถุประสงค์ดังต่อไปนี้

1. เพื่อศึกษาทฤษฎีการรู้จำใบหน้าที่เกี่ยวข้องกับการหาคุณลักษณะของภาพกรณีที่มีสิ่งบดบังใบหน้าเพื่อนำไปใช้ในการวิจัย
2. เพื่อใช้วิธีเชิงสถิติการวิเคราะห์ส่วนประกอบหลัก 2 มิติ และการวิเคราะห์การจำแนกเชิงเส้น 2 มิติ เพื่อรู้จำใบหน้าที่มีสิ่งบดบัง
3. เพื่อวิเคราะห์และเปรียบเทียบความถูกต้องและความเร็วในการรู้จำจากการวิเคราะห์ส่วนประกอบหลัก 2 มิติ และการวิเคราะห์การจำแนกเชิงเส้น 2 มิติ กรณีมีสิ่งบดบังบนใบหน้า

5.1 เพื่อศึกษาทฤษฎีการรู้จำใบหน้าในการหาคุณลักษณะของภาพกรณีที่มีสิ่งบดบังใบหน้า เพื่อนำไปใช้ในการวิจัย

ผู้วิจัยได้ศึกษาทฤษฎีการรู้จำใบหน้าในการหาคุณลักษณะของภาพกรณีที่มีสิ่งบดบังใบหน้าและความรู้ที่เกี่ยวข้อง ได้แก่ วิธีการเปรียบเทียบกราฟ วิธีทางลักษณะทางเรขาคณิต แต่วิธีเหล่านี้มีข้อเสียเรื่องของการอ้างอิงตำแหน่งบนใบหน้าต้องมียอดประกอบครบทุกส่วนในการรู้จำ จึงใช้วิธีเชิงสถิติซึ่งเป็นลักษณะของการเปรียบเทียบภาพต้นฉบับและภาพทดสอบ ซึ่งมีสถิติที่นำมาใช้คือ 2DPCA และ 2DLDA เพื่อหาเมตริกซ์คุณลักษณะในมิติที่เล็กลง และใช้วิธีโครงข่ายประสาทซึ่งมีโครงข่ายประสาทที่นำมาใช้คือ โครงข่ายประสาทเทียมแบบสังวัตนาการ เพื่อใช้ในจำแนกตัวบุคคล

5.2 เพื่อใช้วิธีเชิงสถิติการวิเคราะห์ส่วนประกอบหลัก 2 มิติ และการวิเคราะห์การจำแนกเชิงเส้น 2 มิติ เพื่อรู้จำใบหน้าที่มีสิ่งบดบัง

การสร้างคินภาพด้วยวิธี 2DPCA และ 2DLDA เป็นการคำนวณหาเวกเตอร์เฉพาะที่เรียงลำดับเวกเตอร์เฉพาะที่สมนัยกับค่าเฉพาะจากมากไปน้อย พบว่าค่าเฉพาะที่มีค่าสูงที่ได้เป็นข้อมูลที่สำคัญและมีประโยชน์ต่อการแยกประเภทบุคคล ส่วนค่าเฉพาะที่มีค่าต่ำเป็นข้อมูลที่ไม่มีความสำคัญต่อการจำแนกบุคคล ซึ่งสามารถคัดออกได้โดยไม่เสียข้อมูลสำคัญ ดังนั้นเมื่อนำเวกเตอร์เฉพาะมาสร้างคินภาพใหม่จำนวน d มิติสมมภ์ พบว่ายิ่งใช้มิติสมมภ์ของเวกเตอร์เฉพาะเพิ่มมากขึ้น ความคมชัดของรูปก็ยิ่งใกล้เคียงกับต้นฉบับมากขึ้นเช่นกัน เนื่องจากการใช้ขนาดมิติของเวกเตอร์เฉพาะที่มีค่าน้อยจะ

เพียงเอกลักษณ์เฉพาะที่สำคัญของตัวบุคคลนั้น แต่ถ้าใช้ขนาดมิติของเวกเตอร์เฉพาะมีค่ามากคือ การใช้ข้อมูลเอกลักษณ์ทั้งหมดของบุคคลนั้น ขนาดข้อมูลที่ใช้ในการบันทึกจะมากขึ้นตามลำดับ

5.3 เพื่อวิเคราะห์และเปรียบเทียบความถูกต้องและความเร็วในการรู้จำจากการวิเคราะห์ ส่วนประกอบหลัก 2 มิติ และการวิเคราะห์การจำแนกเชิงเส้น 2 มิติ กรณีมีสิ่งบดบังบนใบหน้า

ผู้วิจัย วิเคราะห์และเปรียบเทียบความถูกต้องและความเร็วในการรู้จำและการฝึกสอน สรุปได้ดังต่อไปนี้

1. ผลการทดลองการเปรียบเทียบไฮเพอร์พารามิเตอร์ของโครงข่ายประสาทมีตัวแปรที่สำคัญประกอบด้วย อัตราเร็วในการเรียนรู้ ขนาดกลุ่มภาพฝึกสอน และ จำนวนรอบในการฝึกสอน โดยการเลือกค่าต่างๆ จากการสุ่มค่าใดค่าหนึ่งก่อนเพื่อหาค่าที่เหมาะสมของตัวแปรตัวอื่นๆ ต่อไป ซึ่งตัวแปรที่มีผลต่อเวลาในการฝึกสอนมากที่สุดคือ ขนาดกลุ่มภาพฝึกสอน ผลการทดลองพบว่า ขนาดกลุ่มภาพที่น้อยหรือมากเกินไปจะใช้เวลาฝึกสอนนานเช่นเดียวกัน ส่วนการเลือกอัตราเร็วในการเรียนรู้ที่มีค่าน้อยเกินไป โครงข่ายประสาทจะเรียนรู้ได้ช้า แต่ถ้าเลือกอัตราเร็วในการเรียนรู้ที่มีค่ามากไปก็จะเรียนรู้ไม่ได้เลย และถ้าเลือกจำนวนรอบฝึกสอนที่มีค่าน้อยเกินไป การฝึกสอนของโครงข่ายก็จะได้ประสิทธิภาพเต็มที่ และถ้าจำนวนรอบการฝึกสอนมีมากเกินไปก็เสียเวลาในการฝึกซ้ำมากขึ้น ซึ่งได้ประโยชน์น้อย

2. ผลการทดลองจากฐานข้อมูล AR ด้วยการใช่วิธีเชิงสถิติร่วมกับโครงข่ายประสาทเทียมแบบสังวัตนาการ โดยใช้แบบจำลองของ Alexnet ที่ขนาดเวกเตอร์เฉพาะ 227 มิติ โดยวิธี 2DPCA และ 2DLDA ให้ความแม่นยำอยู่ที่ร้อยละ 53.06 และ 49.71 ตามลำดับ แต่เมื่อนำเวกเตอร์เฉพาะมาสร้างคืนภาพใหม่จะให้ความแม่นยำเพิ่มขึ้นเป็นร้อยละ 81.55 และ 80.56 ตามลำดับ เนื่องจากข้อจำกัดประการแรกของแบบจำลอง Alexnet ที่ต้องการข้อมูลที่มีขนาด 227×227 มิติ เท่านั้น จึงไม่สามารถลดขนาดเวกเตอร์เฉพาะได้ และประการที่สอง การใช้เวกเตอร์เฉพาะเป็นเพียงเมทริกซ์คุณลักษณะของบุคคลเท่านั้น ซึ่งไม่ใช่เมทริกซ์ภาพ จึงให้ผลความแม่นยำต่ำกว่าการสร้างคืนภาพใหม่ เพราะแบบจำลอง Alexnet ถูกออกแบบมาเพื่อการจำแนกรูปภาพเป็นการเฉพาะ

3. ผลการทดลองจากฐานข้อมูล AR การใช่วิธีเชิงสถิติร่วมกับโครงข่ายประสาทเทียมแบบสังวัตนาการ โดยใช้แบบจำลองของผู้วิจัย

เมื่อเปรียบเทียบวิธี 2DPCA ของ Alexnet กับผู้วิจัย ที่ใช้ขนาด 80 มิติของเวกเตอร์เฉพาะ ให้ความแม่นยำใกล้เคียงกับแบบจำลองของ Alexnet เดิม และเมื่อเทียบเวลาที่ใช้ฝึกสอน พบว่าวิธี 2DPCA ใช้เวลาฝึกสอน 722.93 วินาที และแบบจำลองของ Alexnet ใช้เวลาฝึกสอนเฉลี่ยของทั้ง 2 วิธีใช้เวลาใกล้เคียง 1,500 วินาที ซึ่งถือว่าใช้เวลาในการฝึกสอนแตกต่างกันมาก

เมื่อเปรียบเทียบวิธี 2DLDA ของ Alexnet กับผู้วิจัย พบว่าการใช้วิธี 2DLDA แบบจำลองของผู้วิจัย ที่ขนาด 90 มิติของเวกเตอร์เฉพาะ ให้ความแม่นยำใกล้เคียงกับแบบจำลองของ Alexnet เดิม และเมื่อเทียบเวลาที่ใช้ฝึกสอน พบว่าวิธี 2DPCA ใช้เวลาฝึกสอน 747.31 วินาที และแบบจำลอง Alexnet ใช้เวลาฝึกสอนเฉลี่ยของทั้ง 2 วิธีใช้เวลาใกล้เคียง 1,500 วินาที ซึ่งถือว่าใช้เวลาในการฝึกสอนแตกต่างกันมาก

สรุปได้ว่า การใช้เวกเตอร์คุณลักษณะด้วยวิธี 2DPCA และ 2DLDA ร่วมกับโครงข่ายประสาทเทียมแบบสังวัตนาการ โดยใช้แบบจำลองของผู้วิจัย พบว่าวิธี 2DPCA ตั้งแต่ขนาด 70 มิติของเวกเตอร์เฉพาะขึ้นไป วิธี 2DPCA ให้ความแม่นยำที่สูงกว่าวิธี 2DLDA และพบว่าทั้ง 2 วิธีให้ความแม่นยำไม่แตกต่างกัน เมื่อเทียบกับแบบจำลองของ Alexnet เดิม แต่ใช้เวลาในการฝึกสอนน้อยกว่าแบบจำลองของ Alexnet มาก

4. ผลการทดลองจากฐานข้อมูล GTAV ที่ใช้วิธีเชิงสถิติร่วมกับโครงข่ายประสาทเทียมแบบสังวัตนาการ โดยใช้แบบจำลองของ Alexnet พบว่า การใช้ขนาดเวกเตอร์เฉพาะ 227 มิติ โดยวิธี 2DPCA และ 2DLCA จะให้ความแม่นยำอยู่ที่ร้อยละ 86.36 และ 78.54 ตามลำดับ แต่เมื่อนำเวกเตอร์เฉพาะมาสร้างคืนภาพใหม่จะให้ความแม่นยำเพิ่มขึ้นเป็นร้อยละ 96.46 และ 91.92 ตามลำดับ เนื่องจากประการที่หนึ่ง แบบจำลอง Alexnet มีข้อจำกัดที่ต้องการข้อมูลที่มีขนาด 227 x 227 มิติ เท่านั้น จึงไม่สามารถลดขนาดเวกเตอร์เฉพาะได้ และประการที่สอง การใช้เวกเตอร์เฉพาะเป็นเพียงเมทริกซ์คุณลักษณะของตัวบุคคลเท่านั้น ยังไม่ใช่เมทริกซ์ภาพ จึงให้ผลความแม่นยำน้อยกว่าการสร้างคืนภาพใหม่ เพราะแบบจำลองของ Alexnet ถูกออกแบบมาเพื่อการจำแนกรูปภาพเป็นการเฉพาะ

5. ผลการทดลองจากฐานข้อมูล GTAV ที่ใช้วิธีเชิงสถิติร่วมกับโครงข่ายประสาทเทียมแบบสังวัตนาการ โดยใช้แบบจำลองของผู้วิจัย

เมื่อเปรียบเทียบวิธี 2DPCA ของ Alexnet กับผู้วิจัย พบว่าแบบจำลองของผู้วิจัยให้ความแม่นยำที่ร้อยละ 84.34 ในขณะที่แบบจำลองของ Alexnet ให้ความแม่นยำที่ร้อยละ 86.36 ซึ่งถือว่าน้อยกว่าแบบจำลองของ Alexnet

เมื่อเปรียบเทียบวิธี 2DLDA ของ Alexnet กับผู้วิจัย พบว่าการใช้วิธี 2DLDA แบบจำลองของผู้วิจัย ที่ขนาด 100 มิติของเวกเตอร์เฉพาะ ให้ความแม่นยำใกล้เคียงกับแบบจำลองของ Alexnet เดิม และเมื่อเทียบเวลาที่ใช้ฝึกสอนกับผู้วิจัย พบว่าวิธี 2DPCA ใช้เวลาฝึกสอน 209.48 วินาที และแบบจำลองของ Alexnet ทั้ง 2 วิธี ใช้เวลาฝึกสอนเฉลี่ยใกล้เคียง 275 วินาที ซึ่งถือว่าใช้เวลาฝึกสอนแตกต่างกัน

สรุปได้ว่า การใช้เวกเตอร์คุณลักษณะด้วยวิธี 2DPCA และ 2DLDA ร่วมกับโครงข่ายประสาทเทียมแบบสังวัตนาการ โดยใช้แบบจำลองของผู้วิจัย พบว่า วิธี 2DPCA ที่ตั้งแต่ขนาด 110 มิติของเวกเตอร์เฉพาะขึ้นไปให้ความแม่นยำที่สูงกว่าวิธี 2DLDA

6. ผลการทดลองการนำระบบรู้จำใบหน้าบุคคลแบบมีสิ่งบดบังมาใช้จริงในชีวิตประจำวัน โดยใช้แบบจำลองของ Alexnet พบว่ามีความสอดคล้องกับการฐานข้อมูล AR ที่มีการสวมแว่นกันแดด และสวมผ้าปิดปาก และสอดคล้องกับฐานข้อมูล GTAV ในลักษณะของการบดบังด้วยตนเอง และสามารถนำมาใช้ในการตรวจสอบลักษณะบุคคลตามที่ต้องการได้

5.4 สรุปผลการวิจัย

วิทยานิพนธ์นี้ เสนอการเปรียบเทียบประสิทธิภาพระหว่าง 2DPCA และ 2DLDA สำหรับการรู้จำใบหน้าแบบมีสิ่งบดบัง ซึ่งประกอบด้วย

การวิเคราะห์ส่วนประกอบหลัก ซึ่งมีจุดมุ่งหมายเพื่อหาค่าที่น้อยที่สุดระหว่างเวกเตอร์ข้อมูลเดิมกับเวกเตอร์ข้อมูลใหม่ที่ลดจำนวนมิติลง โดยการวิเคราะห์ส่วนประกอบหลักจะไม่ได้จำแนกความแตกต่างของข้อมูลแต่ละประเภท แต่จะมองข้อมูลทุกตัวเป็นประเภทเดียวกันทั้งหมดในขณะเดียวกัน การวิเคราะห์การจำแนกเชิงเส้น (LDA) เป็นวิธีการหาค่าสูงสุดของสัดส่วนระหว่างความแปรปรวนภายในกลุ่มและความแปรปรวนระหว่างกลุ่ม ซึ่งมีจุดมุ่งหมายเพื่อหาจุดของข้อมูลประเภทเดียวกันให้มีความแตกต่างกันน้อยที่สุด จึงทำให้วิธีทั้ง 2 วิธีนี้ให้ผลลัพธ์แตกต่างกัน

จากผลการทดลองการเปรียบเทียบประสิทธิภาพระหว่างวิธี 2DPCA และ 2DLDA สำหรับการรู้จำใบหน้าแบบมีสิ่งบดบัง พบว่า วิธี 2DPCA ให้ประสิทธิภาพดีกว่าในด้านความแม่นยำ และ การใช้วิธีเชิงสถิติร่วมกับโครงข่ายประสาทเทียมแบบสังวัตนาการอาจไม่ส่งผลต่อการเพิ่มประสิทธิภาพโดยรวมมีความแม่นยำสูงและใช้เวลาฝึกสอนไม่มากนัก และยังคงรักษาระดับความแม่นยำในขณะที่ขนาดมิติข้อมูลลดลง แต่เมื่อนำโครงข่ายประสาทเทียมแบบสังวัตนาการ โดยใช้แบบจำลองของ Alexnet มาดัดแปลงใช้กับข้อมูลที่มีมิติลดลงจะยังคงรักษาความแม่นยำในระดับใกล้เคียงกับแบบเดิมแม้จะไม่มาก แต่มีพัฒนาการด้านความเร็วในการฝึกสอน

5.5 ข้อเสนอแนะเพื่อนำไปใช้

1. สามารถนำผลการเรียนรู้วิธีการรู้จำใบหน้าที่มีสิ่งบดบังด้วยวิธีการวิเคราะห์ส่วนประกอบหลัก 2 มิติ และการวิเคราะห์การจำแนกเชิงเส้น 2 มิติ ไปใช้ในการลดขนาดข้อมูลฝึกสอนให้เล็กลงแต่ยังคงรักษาข้อมูลสำคัญได้
2. สามารถนำระบบการรู้จำใบหน้ากรณีที่มีสิ่งบดบังบนใบหน้าไปใช้ได้จริงและแสดงผลในเวลาที่รวดเร็วด้วยความถูกต้องและแม่นยำ
3. สามารถนำระบบการรู้จำใบหน้าที่มีสิ่งบดบังของผู้วิจัยไปเป็นต้นแบบของการระบุใบหน้าเมื่อมีการบดบังให้พัฒนาต่อไปได้
4. สามารถประยุกต์ระบบการรู้จำใบหน้าที่มีสิ่งบดบังของผู้วิจัยไปใช้ในหน่วยงาน สถานศึกษา ราชการ หรือสถานศึกษาต่างๆ ได้

5.6 ข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป

1. ควรศึกษาเพื่อพัฒนาระบบการรู้จำใบหน้าที่มีสิ่งบดบังแบบอื่นๆ
2. ควรศึกษาระบบการรู้จำใบหน้าที่มีสิ่งบดบังไปเป็นต้นแบบของการระบุใบหน้าเมื่อมีการบดบังและพัฒนาต่อไปในระบบ 3 มิติ
3. ควรสร้างระบบการรู้จำใบหน้าและระบบการยืนยันตัวบุคคลให้เป็นแบบระบบเปิด เพื่อพัฒนาระบบให้มีความสมบูรณ์แบบยิ่งขึ้น



ภาคผนวก

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ภาคผนวก ก
โปรแกรมที่ใช้ในการคำนวณ

1. ฟังก์ชันคำนวณหาเวกเตอร์เฉพาะด้วยวิธี 2DPCA

```

1  %----- 2D-PCA -----
2  %----- General Information -----
3  % Program: TwoDPCA.m
4  % Author: SS
5  %----- Define Variable -----
6  % XTrain          - 4D gray training image table
7  % rows            - rows of image
8  % columns         - columns of image
9  % TotalTrainImage - total of training image
10 % mn              - mean image of training set
11 % C               - covariance matrix
12 % A               - current image subtract mean image
13 % V               - eigenvectors
14 % D               - eigencvalues
15 % eigval_st      - ordered eigenvectors(eigenspace)
16 % eigvec_st      - ordered eigenvalues
17
18 function [eigvec_st, eigval_st] = TwoDPCA(XTrain)
19     [rows, columns, ~, TotalTrain] = size(XTrain);
20
21     %Finding mean image of training dataset
22     mn = zeros(rows,columns);
23     for i = 1:TotalTrain
24         mn = mn + double(XTrain(:,:, :, i));
25     end
26     mn = mn/TotalTrain;
27
28     %Finding covaraince matrix of current image minus mean
29     C = zeros(columns,columns);
30     for n = 1:TotalTrain
31         A = double(XTrain(:,:, :, n)) - mn;
32         C = C + A' * A;
33     end
34     C = C/TotalTrain;
35
36     %Finding Eigenvectors(V) and Eigenvalues(D)
37     [V,D] = eig(C);
38
39     %Order the eigenvectors according to the eigenvalues
40     [D, idx] = sort(diag(D));
41     eigval_st = D(end:-1:1)';
42     eigvec_st = V(:, idx(end:-1:1));
43
44 end

```

2. ฟังก์ชันคำนวณหาเวกเตอร์เฉพาะด้วยวิธี 2DLDA

```
1  %----- 2DLDA -----
2  %----- General Information -----
3  % Program: TwoDLDA.m
4  % Author: SS
5  %----- Define Variable -----
6  % Train - training image
7  % mn     - mean of Train
8  % msc    - mean subtracted traing images
9  % sw     - within class scatter matrix
10 % sb     - between class scatter matrix
11 % U      - Projection Vectors
12 %-----
13
14 function [U] = TwoDLDA(Train,NumClasses,ImgsPerClass)
15     fprintf('.....2DLDA.....\n');
16     [N1, N2, ~, P] = size(Train);
17
18     % compute mean of each class
19     fprintf('compute mean of each class\n')
20     mn = zeros(N1,N2);
21     for i = 1:P
22         mn = mn + double(Train(:,:,i));
23     end
24     mn = mn/P;
25
26     ma = zeros(N1,N2);
27     msm = zeros(N1,N2);
28     for i = 1:NumClasses
29         for j = 1:ImgsPerClass
30             ma = ma + double(Train(:,:,:(i-1)*ImgsPerClass+j));
31         end
32         m(:,:,i) = ma/ImgsPerClass;
33         ma = zeros(N1,N2);
34     end
35
36     % Mean subtract all the images
37     fprintf('Mean subtract all the images\n')
38     for i = 1:NumClasses*ImgsPerClass
39         msc(:,:,i) = double(Train(:,:,i)) - m(:,:,fix((i-
40 1)/ImgsPerClass)+1);
41     end
42
43     % calculate the sw matrix (within class scatter)
44     fprintf('calculate within class scatter\n')
45     sw = zeros(N2,N2);
46     for i = 1:NumClasses * ImgsPerClass
47         sw = sw + msc(:,:,i)' * msc(:,:,i);
48     end
49
50     % calculate the sb matrix (within class scatter)
51     fprintf('calculate within class scatter\n')
52     sb = zeros(N2,N2);
53     for i = 1:NumClasses
54         sb = sb + msc(:,:,i)' * msc(:,:,i);
55     end
56
```

```
57     % solved the generalized eigenvalue problem sb*w = d*sw*w
58     fprintf('Calculate eigenvalue\n');
59     [V, D, U] = svd(sb*inv(sw));
60     szd = size(D);
61     for i = szd(1)
62         evals(i) = D(i,i);
63     end
64
65 end
```



3. โปรแกรมการทำงานหลักของระบบโดยใช้แบบจำลองของ Alexnet

```
1  clc; clear all; close all;
2  addpath(genpath(pwd));
3  load('F:\thesis\dataset\dataset-gtav.mat');
4
5  [XTrain,XValidation] = rgb2ycbcr2gray(XTrain,XValidation);
6  numClasses = numel(categories(YTrain))
7  ImgsPerClass = size(XTrain,4)/numClasses
8
9  %[u,v] = TwoDPCA(XTrain);
10 u = TwoDLDA(XTrain,numClasses,ImgsPerClass);
11
12 net = alexnet;
13 layerTransform = net.Layers(1:end-3);
14
15 for dim = 227
16     for i = 1 : size(XTrain,4)
17         proj(:,:,,i) = real(double(XTrain(:,:,,i)) * u(:,1:dim));
18     end
19
20     for i = 1 : size(XValidation,4)
21         projt(:,:,,i) = real(double(XValidation(:,:,,i)) *
22 u(:,1:dim));
23     end
24
25     augimdsTrain = augmentedImageDatastore([227 227],proj,YTrain, ...
26 'ColorPreprocessing','gray2rgb');
27     augimdsValidation = augmentedImageDatastore([227
28 227],projt,YValidation, ...
29 'ColorPreprocessing','gray2rgb');
30
31     layer = [
32         layerTransform
33         fullyConnectedLayer(numClasses)
34         softmaxLayer
35         classificationLayer];
36
37     miniBatchSize = 128;
38     InitialLearnRate = 1e-3;
39     ValidationFrequency = floor(size(XTrain,4)/miniBatchSize);
40     MaxEpochs = 30;
41
42     options = trainingOptions('sgdm', ...
43         'InitialLearnRate',InitialLearnRate, ...
44         'MaxEpochs',MaxEpochs, ...
45         'ExecutionEnvironment','gpu', ...
46         'Shuffle','every-epoch', ...
47         'MiniBatchSize',miniBatchSize, ...
48         'ValidationData',augimdsValidation, ...
49         'ValidationFrequency',ValidationFrequency, ...
50         'Plots','training-progress', ...
51         'Verbose',false);
52
53     tic;
54     [netTransfer,netInfo] = trainNetwork(augimdsTrain,layer,options);
55     TrainingTime = toc;
56
```

```
57     path = pwd;
58     folder = '\log\gtav\TwoDPCAvsTwoDLDA (eigenvector) -
59 alexnet (cnn)\TwoDLDA';
60     fullpath = strcat(path, folder);
61     str =
62 sprintf('%s\gtav_2dlda_eigenvector_alexnet (cnn)_%03ddim.mat', fullpat
63 h, dim);
64     save(str, 'netInfo', 'TrainingTime');
65     fprintf('saved acc = %.2f timing =
66 %.2f\n', netInfo.ValidationAccuracy(end), TrainingTime);
67
68     proj = [];
69     projt = [];
70
71     netInfo.ValidationAccuracy(isnan(netInfo.ValidationAccuracy)) =
72 0;
73     idx = find(netInfo.ValidationAccuracy>0);
74     Vacc = netInfo.ValidationAccuracy(idx);
75 end
```



4.โปรแกรมการทำงานหลักของระบบโดยใช้แบบจำลองของผู้วิจัย

```
1  addpath(genpath(pwd));
2  load('F:\thesis\dataset\dataset-ar-modify.mat');
3
4  [XTrain,XValidation] = rgb2ycbcr2gray(XTrain,XValidation);
5  numClasses = numel(categories(YTrain))
6  ImgsPerClass = size(XTrain,4)/numClasses
7
8  net = alexnet;
9
10 p1=1;p2=1;p3=0;p4=1;
11 p5=0;p6=0;p7=0;p8=1;
12
13 for dim = 150:10: 210
14
15     if dim==10 || dim==20 || dim==30 || dim==40
16         resizefully = 1792;
17     elseif dim==50 || dim==60 || dim==70
18         resizefully = 3584;
19     elseif dim==80 || dim==90 || dim==100
20         resizefully = 5376;
21     elseif dim==110 || dim==120 || dim==130
22         resizefully = 7168;
23     elseif dim==140 || dim==150 || dim==160
24         resizefully = 8960;
25     elseif dim==170 || dim==180 || dim==190 || dim==200
26         resizefully = 10752;
27     elseif dim==210 || dim==220
28         resizefully = 12544;
29     end
30
31     Weightsfc1 = imresize(net.Layers(17, 1).Weights,[4096
32 resizefully]);
33     Biasfc1 = imresize(net.Layers(17, 1).Bias,[4096 1]);
34
35     Weightsfc2 = imresize(net.Layers(20, 1).Weights,[4096 4096]);
36     Biasfc2 = imresize(net.Layers(20, 1).Bias,[4096 1]);
37
38     %[u,v] = TwoDPCA(XTrain);
39     u = TwoDLDA(XTrain,numClasses,ImgsPerClass);
40
41     for i = 1: size(XTrain,4)
42         proj(:, :, :, i) = real(double(XTrain(:, :, :, i)) * u(:, 1:dim));
43
44     end
45
46     for i = 1: size(XValidation,4)
47         projt(:, :, :, i) = real(double(XValidation(:, :, :, i)) *
48 u(:, 1:dim));
49
50     end
51
52     AvgImage = imresize(net.Layers(1, 1).AverageImage,[227 dim]);
53
54     layer = [
55         imageInputLayer([227, dim, 3], 'AverageImage', AvgImage)
```

```

56
57 convolution2dLayer(11,96,'Stride',4,'Padding',0+p1,'NumChannels',3,'W
58 eights',net.Layers(2,1).Weights,'Bias',net.Layers(2,
59 1).Bias,'WeightLearnRateFactor',1,'BiasLearnRateFactor',2)
60     reluLayer
61     crossChannelNormalizationLayer(5,'K',1)
62     maxPooling2dLayer(3,'Stride',2,'Padding',0+p2)
63
64
65 groupedConvolution2dLayer(5,128,2,'Stride',1,'Padding',2+p3,'WeightLe
66 arnRateFactor',1,'BiasLearnRateFactor',2,'Weights',net.Layers(6,
67 1).Weights,'Bias',net.Layers(6,1).Bias)
68     reluLayer
69     crossChannelNormalizationLayer(5,'K',1)
70     maxPooling2dLayer(3,'Stride',2,'Padding',0+p4)
71
72 convolution2dLayer(3,384,'Stride',1,'Padding',1+p5,'NumChannels',256,
73 'Weights',net.Layers(10,1).Weights,'Bias',net.Layers(10,
74 1).Bias,'WeightLearnRateFactor',1,'BiasLearnRateFactor',2)
75
76     reluLayer
77
78 groupedConvolution2dLayer(3,192,2,'Stride',1,'Padding',1+p6,'WeightLe
79 arnRateFactor',1,'BiasLearnRateFactor',2,'Weights',net.Layers(12,
80 1).Weights,'Bias',net.Layers(12,1).Bias)
81     reluLayer
82
83 groupedConvolution2dLayer(3,128,2,'Stride',1,'Padding',1+p7,'WeightLe
84 arnRateFactor',1,'BiasLearnRateFactor',2,'Weights',net.Layers(14,
85 1).Weights,'Bias',net.Layers(14,1).Bias)
86     reluLayer
87
88     maxPooling2dLayer(3,'Stride',2,'Padding',0+p8)
89
90 fullyConnectedLayer(4096,'Weights',Weightsfc1,'Bias',Biasfc1,'WeightL
91 earnRateFactor',1,'BiasLearnRateFactor',2)
92     reluLayer
93     dropoutLayer(0.5)
94
95 fullyConnectedLayer(4096,'Weights',Weightsfc2,'Bias',Biasfc2,'WeightL
96 earnRateFactor',1,'BiasLearnRateFactor',2)
97
98     reluLayer
99     dropoutLayer(0.5)
100
101 fullyConnectedLayer(numClasses,'WeightLearnRateFactor',1,'BiasLearnRa
102 teFactor',2)
103     softmaxLayer
104     classificationLayer
105 ];
106 %analyzeNetwork(layer)
107
108     augimdsTrain = augmentedImageDatastore([227 dim],proj,YTrain, ...
109     'ColorPreprocessing','gray2rgb');
110     augimdsValidation = augmentedImageDatastore([227
111 dim],projt,YValidation, ...
112     'ColorPreprocessing','gray2rgb');
113

```

```

114
115     MaxEpochs = 30;
116     miniBatchSize = 128;
117     InitialLearnRate = 1e-3;
118     ValidationFrequency = floor(size(XTrain,4)/miniBatchSize);
119
120
121     %delete(gcp('nocreate'));
122     %numberOfWorkers = 4;
123     %parpool(numberOfWorkers);
124
125     options = trainingOptions('sgdm', ...
126         'ExecutionEnvironment','gpu', ...
127         'MiniBatchSize',miniBatchSize, ...
128         'MaxEpochs',MaxEpochs, ...
129         'InitialLearnRate',InitialLearnRate, ...
130         'Shuffle','every-epoch', ...
131         'ValidationData',augimdsValidation, ...
132         'ValidationFrequency',ValidationFrequency, ...
133         'Verbose',false);
134     %'Plots','training-progress'
135     %     'WorkerLoad',1, ...
136
137     tic;
138     [netTransfer,netInfo] =
139     trainNetwork(augimdsTrain,layer,options);
140     TrainingTime = toc;
141     %delete(gcp('nocreate'));
142
143
144     path = pwd;
145     %folder = '\log\ar\alexnet-newModify-(Eigenvector)\2dpca';
146     folder = '\log\ar\alexnet-newModify-(Eigenvector)\2dlda';
147
148     fullpath = strcat(path,folder);
149
150     %str = จุฬาลงกรณ์มหาวิทยาลัย
151     sprintf('%s\\ar_alexnet_newModify_(2dpca_Eigenvector)_%03ddim.mat',fu
152     llpath,dim);
153     str =
154     sprintf('%s\\ar_alexnet_newModify_(2dlda_Eigenvector)_%03ddim.mat',fu
155     llpath,dim);
156
157     %fprintf('saved2dpca acc = %.2f timing =
158     %.2f\n',netInfo.ValidationAccuracy(end),TrainingTime);
159     fprintf('saved2dlda acc = %.2f timing =
160     %.2f\n',netInfo.ValidationAccuracy(end),TrainingTime);
161
162     save(str,'netInfo','TrainingTime');
163
164     proj = [];
165     projt = [];
166     end

```

บรรณานุกรม

1. Montra Buaped and Charnchai Pluempittiwiriyawej, "Expression-invariant 3D face recognition using shape context", *Department of Electronic Engineering*. 2007, Chulalongkorn University. p. 90.
2. Wilmar Fernando and Moya Rueda, "Deep Face Recognition", *Faculty of Computer Science 2017*, Technical University of Dortmund. p. 88.
3. Pitsanu Kumeechai, "Terrorists face recognition system with color skin detection and Eigen face", *Defence Engineering*. 2014: King Mongkut's Institute of Technology Ladkrabang, Central Library of KMITL. p. 92.
4. Sittinon Kongsontana, "Face recognition using 2DLDA algorithm", *engineering of telecommunication*. 2005, king mongkut's institute of technology ladkrabang. p. 125.
5. Matthew Turk and Alex Pentland, "Eigenfaces for recognition". *Journal of cognitive neuroscience*, 1991. 3(1): p. 71-86.
6. Peter N Belhumeur, João P Hespanha, and David J Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection". *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 1997(7): p. 711-720.
7. Jian Yang, David Zhang, Alejandro F Frangi, and Jing-yu Yang, "Two-dimensional PCA: a new approach to appearance-based face representation and recognition". *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2004. 26(1): p. 131-137.
8. Xiao-ming Wang, Chang Huang, Xiao-yingFang, and Jin-gao Liu. "2DPCA vs. 2DLDA: Face recognition using two-dimensional method". *Artificial Intelligence and Computational Intelligence*. 2009. IEEE.
9. Tae Young Kim, Kyoung Mu Lee, Sang Uk Lee, and Chung-Hyuk Yim, "Occlusion invariant face recognition using two-dimensional PCA", *Advances in Computer Graphics and Computer Vision*. 2007, Springer. p. 305-315.
10. Xu-Die Ren, Hao-Nan Guo, Guan-Chen He, Xu Xu, Chong Di, and Sheng-Hong Li. "Convolutional neural network based on principal component analysis

- initialization for image classification". *2016 IEEE First International Conference on Data Science in Cyberspace (DSC)*. 2016. IEEE.
11. Martin Lades, Jan C Vorbruggen, Joachim Buhmann, Jörg Lange, Christoph Von Der Malsburg, Rolf P Wurtz, and Wolfgang Konen, "Distortion invariant object recognition in the dynamic link architecture". *IEEE Transactions on computers*, 1993. 42(3): p. 300-311.
 12. Madasu Hanmandlu, Divya Gupta, and Shantaram Vasikarla. "Face recognition using Elastic bunch graph matching". *Applied Imagery Pattern Recognition Workshop (AIPR): Sensing for Control and Augmentation, 2013 IEEE*. 2013. IEEE.
 13. Young-Jun Song, Young-Gil Kim, Nam Kim, and Jae-Hyeong Ahn. "Face recognition using both geometric features and pca/lda". *Advanced Language Processing and Web Information Technology, 2007. ALPIT 2007. Sixth International Conference on*. 2007. IEEE.
 14. Navya Sushma Tummala and Pnrl Chandra Sekhar. "Face recognition using PCA and geometric approach". *Computing Methodologies and Communication (ICCMC), 2017 International Conference on*. 2017. IEEE.
 15. Tanvi B Patel and Jalpa T Patel. "Occlusion detection and recognizing human face using neural network". *Intelligent Computing and Control (I2C2), 2017 International Conference on*. 2017. IEEE.
 16. Roberto Brunelli and Tomaso Poggio, "Face recognition: Features versus templates". *IEEE transactions on pattern analysis and machine intelligence*, 1993. 15(10): p. 1042-1052.
 17. Sanjeev Kumar and Harpreet Kaur, "Face recognition techniques: Classification and comparisons". *International Journal of Information Technology and Knowledge Management*, 2012. 5(2): p. 361-363.
 18. Stephen Karungaru, Minoru Fukumi, and Norio Akamatsu. "Face recognition using genetic algorithm based template matching". *Communications and Information Technology, 2004. ISCIT 2004. IEEE International Symposium on*. 2004. IEEE.
 19. Manassakan Sanayha and Peerapon Vateekul, "Remaining Useful Life Prediction Using Enhanced Convolutional Neural Network on Multivariate Time Series Sensor Data.". *Walailak Journal of Science and Technology (WJST)*, 2019.

20. Chigozie Nwankpa, Winifred Ijomah, Anthony Gachagan, and Stephen Marshall, "Activation Functions: Comparison of trends in Practice and Research for Deep Learning". *arXiv preprint arXiv:1811.03378.*, 2018.
21. Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. "Imagenet classification with deep convolutional neural networks". *Advances in neural information processing systems*. 2012.
22. Aleix Martínez and Robert Benavente. "The AR face database". 1998; Available from: <http://www2.ece.ohio-state.edu/~aleix/ARdatabase.html>.
23. F Tarrés and A Rama. "GTAV Face Database". 2011; Available from: <https://gtav.upc.edu/en/research-areas/face-database>.



ประวัติผู้เขียน

ชื่อ-สกุล	นายสิทธิพันธุ์ สระภักดิ์
วัน เดือน ปี เกิด	29 มกราคม 2536
สถานที่เกิด	ชุมพร
วุฒิการศึกษา	ปริญญาบัณฑิต คณะวิศวกรรมศาสตร์ สาขาวิศวกรรมคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีมหานคร
ที่อยู่ปัจจุบัน	บ้านเลขที่ 6/2 หมู่ 2 ตำบลดอนกำยาน อำเภอเมือง จังหวัดสุพรรณบุรี รหัสไปรษณีย์ 72000
ผลงานตีพิมพ์	Sittiphan Sarapakdi, Phaderm Nangsue and Charnchai Pluempitiwiriyawej, "Occluded Facial Recognition with 2DPCA based Convolutional Neural Network," IEEE International Conference on Consumer Electronics (ICCE) Asia, pp 135-138, 2019.