รหัสโครงการ SUT 7-709-60-12-24

รายงานการวิจัย

การสร้างคืนภาพความละเอียดสูงยิ่งสำหรับระบบกล้องโทรทัศน์วงจรปิด
(Super-resolution Image Reconstruction for CCTV System)

ได้รับทุนอุดหนุนการวิจัยจาก
มหาวิทยาลัยเทคโนโลยีสุรนารี

ผลงานวิจัยเป็นความรับผิดชอบของหัวหน้าโครงการวิจัยแต่เพียงผู้เดียว

รหัสโครงการ SUT 7-709-60-12-24

รายงานการวิจัย

การสร้างคืนภาพความละเอียดสูงยิ่งสำหรับระบบกล้องโทรทัศน์วงจรปิด
(Super-resolution Image Reconstruction for CCTV System)

คณะผู้วิจัย

หัวหน้าโครงการ
ผู้ช่วยศาสตราจารย์ เรืออากาศเอก ดร. ประโยชน์ คำสวัสดิ์
สาขาวิชาวิศวกรรมโทรคมนาคม
สำนักวิชาวิศวกรรมศาสตร์

มหาวิทยาลัยเทคโนโลยีสุรนารี

ได้รับทุนอุดหนุนการวิจัยจากมหาวิทยาลัยเทคโนโลยีสุรนารี ปีงบประมาณ พ.ศ. 2560
ผลงานวิจัยเป็นความรับผิดชอบของหัวหน้าโครงการวิจัยแต่เพียงผู้เดียว

สิงหาคม 2562

กิตติกรรมประกาศ

ผู้วิจัยขอขอบคุณ มหาวิทยาลัยเทคโนโลยีสุรนารี ที่ได้ให้ทุนอุดหนุนการวิจัย ประจำปีงบประมาณ 2560

ผู้วิจัยขอขอบคุณ เจ้าหน้าที่ของสถานวิจัยสำนักวิชาวิศวกรรมศาสตร์และสถาบันวิจัยพัฒนา มหาวิทยาลัยเทคโนโลยีสุรนารี ที่ได้ให้ความสะดวกในด้านข้อมูล งานเอกสารแบบฟอร์มต่าง ๆ ที่ต้องใช้ในการดำเนินโครงการวิจัย จนทำให้งานเอกสารที่เกี่ยวข้องกับงานวิจัยนี้ดำเนินไปได้ด้วยดี

ผู้วิจัยขอขอบคุณ นักศึกษาปริญญาโท สาขาวิชาวิศวกรรมโทรคมนาคม สำนักวิชาวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี ที่ได้ช่วยงานวิจัย เก็บข้อมูลการทดลอง จนทำให้งานวิจัยนี้สำเร็จลุล่วง

ประโยชน์ คำสวัสดิ์

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทคัดย่อ

รายงานวิจัยฉบับนี้นำเสนอ การสร้างคืนภาพความละเอียดสูงยิ่งสำหรับระบบกล้องโทรทัศน์วงจรปิดโดยใช้เครือข่ายประสาทเทียมแบบคอนโวลูชัน (SR-CNN) ในการทำงานของอัลกอริทึม SR-CNN จะถูกใช้ในการเรียนรู้เพื่อสร้างผังที่แสดงความสัมพันธ์ (Mapping) ของการจับคู่ระหว่างภาพความละเอียดต่ำกับภาพความละเอียดสูงหรือภาพอ้างอิง การ Mapping ดังกล่าวนั้น SR-CNN สามารถทำได้อย่างมีประสิทธิภาพเนื่องจากภาพความละเอียดต่ำกับภาพความละเอียดสูงมีองค์ประกอบในภาพที่คล้ายคลึงกันและแตกต่างกันเฉพาะรายละเอียดขององค์ประกอบความถี่สูงเท่านั้น ผู้วิจัยได้ประยุกต์ใช้เครือข่ายประสาทเทียมแบบคอนโวลูชันในการตรวจหาค่าเรซิดิว (Residual) ของภาพจากองค์ประกอบของภาพสี เมื่อเครือข่าย SR-CNN ได้ทำการเรียนรู้และประมาณภาพเรซิดิวได้แล้ว ภาพความละเอียดสูงก็สามารถสร้างได้ด้วยการรวมภาพเรซิดิวกับภาพความละเอียดต่ำที่ได้มีการปรับสเกลให้มีขนาดของภาพเพิ่มขึ้นจนเท่ากับภาพความละเอียดสูงในระดับสเกลที่ต้องการ ผลการทดลองได้แสดงให้เห็นว่าวิธีการที่นำเสนอมีประสิทธิภาพที่ดีกว่าวิธีการขยายภาพแบบไบคิวบิกทั้งในแง่ของการวัดค่าคุณภาพของภาพดิจิทัลด้วยมาตรวัดแบบต่าง ๆรวมถึงรายละเอียดที่ชัดเจนและขอบภาพที่คมชัดยิ่งขึ้น อย่างไรก็ตามเทคนิคที่นำเสนอยังมีข้อบกพร่องบางอย่างเช่น มีความเร็วในการฝึกสอนที่ช้ามากเกินไปและการติดปัญหาค่าที่ดีที่สุดเฉพาะที่ (Local optimum) ซึ่งเกิดขึ้นได้ง่ายในขั้นตอนของการเรียนรู้

มหาวิทยาลัยเทคโนโลยีสุรนารี

Abstract

This research project presents a new super-resolution image reconstruction technique for CCTV system using convolution neural network (SR-CNN). Based on the proposed SR-CNN, the convolution operation and non-linear mapping are employed to adapt images naturally and leaning the end-to-end mapping from low-resolution images and high-resolution images. This mapping is possible because low-resolution and high-resolution images have similar image content and differ primarily in high-frequency details. Therefore, SR-CNN is used to detects the residual image from the luminance of a color image. After the SR-CNN network learns to estimate the residual image, the high-resolution image can be reconstructed by adding the estimated residual image to the up-sampled low-resolution image. The experimental results demonstrate that the proposed technique outperforms state of-the-art bicubic image interpolation in terms of image quality metrics, clearer details and sharper edges. However, the proposed technique has some drawbacks. The network training speed is very slow, and local optimum entrapment is easy to occur.

มหาวิทยาลัยเทคโนโลยีสุรนารี

สารบัญ

เรื่อง **หน้า**

กิตติกรรมประกาศ ..... **ก**
บทคัดย่อ ..... **ข**
Abstract ..... **ค**
สารบัญ ..... **ง**
สารบัญตาราง ..... **จ**
สารบัญรูป ..... **ช**
บทที่ 1 บทนำ **1**
1.1 ความสำคัญและที่มาของปัญหา 1
1.2 วัตถุประสงค์ของโครงการวิจัย 2
1.3 ขอบเขตของงานวิจัย 2
1.4 วิธีการดำเนินงานวิจัยและสถานที่วิจัย 2
1.5 ประโยชน์ที่คาดว่าจะได้รับ 2
บทที่ 2 ทฤษฎีและหลักการที่เกี่ยวข้อง **3**
2.1 บทนำ 3
2.2 การสร้างภาพความละเอียดสูง 4
2.3 การทบทวนวรรณกรรม/สารสนเทศ (Information) ที่เกี่ยวข้อง 4
2.4 การแปลงเวฟเล็ต (Wavelet Transform) 9
2.4.1 การแปลงเวฟเล็ตแบบดีสครีต (Discrete wavelet transform; DWT) 10
2.4.2 สเกลลิงฟังชัน (Scaling function) 11
2.4.3 เวฟเล็ตฟังก์ชัน (Wavelet function) 11
2.4.4 การแปลงกลับเวฟเล็ตแบบดีสครีต 12
2.5 เครือข่ายประสาทแบบคอนโวลูชัน 13
2.6 สรุป 15
บทที่ 3 วิธีดำเนินการวิจัย **16**
3.1 บทนำ 16

มหาวิทยาลัยเทคโนโลยีสุรนารี

3.2 การสร้างภาพความละเอียดสูงด้วยการเรียนรู้เชิงลึก 16
3.2.1 เครือข่ายประสาทเทียมแบบคอนโวลูชัน 17
3.2.2 การเรียนรู้ของเครือข่ายประสาทเทียมแบบคอนโวลูชัน 18
3.3 การปรับปรุงคุณภาพของภาพความละเอียดสูงด้วยการใช้ตัวกรองแบบดิจิทัล 22
3.3.1 ตัวกรองดิจิตอลแบบ Guided filtering 23
3.3.2 ตัวกรองดิจิตอลแบบ Bilateral filtering 24
3.3.3 ตัวกรองดิจิตอลในโดเมนของการแปลงเวฟเล็ต 24
3.4 การประเมินคุณภาพของสัญญาณภาพดิจิทัล 24
3.5 สรุป 26
บทที่ 4 ผลการวิจัย 27
4.1 บทนำ 27
4.2 ผลการทดสอบการสร้างภาพความละเอียดสูงด้วย Deep learning 27
4.3 ผลการเปรียบเทียบคุณภาพสัญญาณภาพดิจิทัลด้วยการตัดและซูมภาพ 35
4.4 ผลการเปรียบเทียบคุณภาพสัญญาณภาพดิจิทัลที่ค่าอัตราการขยายภาพต่าง ๆ 41
4.5 การปรับปรุงคุณภาพด้วยการใช้ตัวกรองแบบดิจิทัล 48
4.6 สรุป 55
บทที่ 5 บทสรุป 57
5.1 สรุป 57
5.2 แนวทางในการวิจัยและพัฒนาต่อไปในอนาคต 58
บรรณานุกรม 59
ประวัตินักวิจัย 62

มหาวิทยาลัยเทคโนโลยีสุรนารี

สารบัญตาราง

ตารางที่ **หน้า**

3.1 ตัวอย่างการกำหนดค่าในแต่ละชั้นของเครือข่ายประสาทเทียมแบบคอนโวลูชัน 19
4.1 ผลการเปรียบเทียบคุณภาพระหว่างวิธี Bicubic และวิธี SR-CNN 32
4.2 ผลการเปรียบเทียบคุณภาพระหว่างวิธี Bicubic และวิธี SR-CNN 33
4.3 การเปรียบเทียบค่าเฉลี่ยจากการคำนวณ PSNR และ SSIM 41
4.4 การเปรียบเทียบค่าเฉลี่ยจากการคำนวณ NIQE และ BRISQUE 42
4.5 การเปรียบเทียบค่าเฉลี่ยของ PSNR และ SSIM 45
4.6 การเปรียบเทียบค่าเฉลี่ยของ NIQE และ BRISQUE 46
4.7 การเปรียบเทียบค่าเฉลี่ย NIQE และ BRISQUE จากตัวกรองแบบไกต์ฟิลเตอร์ 49
4.8 การเปรียบเทียบค่าเฉลี่ย NIQE และ BRISQUE จากการใช้ตัวกรองแบบไบเลเทอรอล 51
4.9 เปรียบเทียบค่าเฉลี่ย NIQE และ BRISQUE จากการใช้ตัวกรองในโดเมนเวฟเล็ต 52

มหาวิทยาลัยเทคโนโลยีสุรนารี

สารบัญรูป

รูปที่ **หน้า**

2.1 การสร้างภาพความละเอียดสูง ... 3
2.2 การทำอินเตอร์โพเลชันแบบไบคิวบิค ... 4
2.3 การลดช่วงกว้างของความถี่ทีละครึ่งหนึ่งสำหรับค่าแต่ละค่าสเกล ... 11
2.4 การแยกองค์ประกอบความถี่ของเวฟเล็ตและการรวมกลับของเวฟเล็ต ... 13
2.5 สถาปัตยกรรมของเครือข่ายประสาทแบบคอนโวลูชัน ... 14
2.6 ความแตกต่างระหว่าง Machine learning กับ Deep learning ... 14
3.1 แผนภาพการสร้างภาพความละเอียดสูงด้วยการเรียนรู้เชิงลึก ... 18
3.2 แผนภาพขั้นตอนการคำนวณหาภาพเรซิดิว ... 20
3.3 การเรียนรู้ของเครือข่าย SR-CNN ในช่วงเริ่มต้นการเรียนรู้ ... 22
3.4 การเรียนรู้ของเครือข่าย SR-CNN ที่รอบการเรียนรู้เท่ากับ 100 และสิ้นสุดการเรียนรู้ ... 22
3.5 รูปแบบเคอร์นอลของตัวกรองขนาด 3x3 และ 5x5 ... 23
3.6 กระบวนการปรับปรุงคุณภาพของภาพความละเอียดสูงด้วยการใช้ตัวกรองแบบดิจิทัล ... 24
4.1 ตัวอย่างภาพดิจิทัลที่ใช้ในการทดสอบ ... 28
4.2 ภาพเรซิดิวของ (ก) Lena (ข) Mandrill ... 29
4.3 ภาพเรซิดิวของ (ก) Airplane (ข) Peppers ... 29
4.4 ภาพเรซิดิวของ (ก) Fruits (ข) FangFha ... 30
4.5 ภาพเรซิดิวของ (ก) Tomato (ข) Lotus ... 30
4.6 ภาพเรซิดิวของ (ก) Green tea (ข) Cat ... 31
4.7ภาพเรซิดิวของ (ก) Temple (ข) Maple ... 31
4.8 เปรียบเทียบค่า PSNR ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN ... 32
4.9 เปรียบเทียบค่า SSIM ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN ... 33
4.10 เปรียบเทียบค่า NIQE ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN ... 34
4.11 เปรียบเทียบค่า BRISQUE ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN ... 34
4.12 เปรียบเทียบภาพ Lena จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา) ... 35
4.13 เปรียบเทียบภาพ Mandrill จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา) ... 35
4.14 เปรียบเทียบภาพ Airplane จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา) ... 36
4.15 เปรียบเทียบภาพ Peppers จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา) ... 36
4.16 เปรียบเทียบภาพ Fruits จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา) ... 37
4.17 เปรียบเทียบภาพ FangFha จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา) ... 37
4.18 เปรียบเทียบภาพ Tomato จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา) ... 38

สารบัญรูป (ต่อ)

รูปที่ **หน้า**

4.19 เปรียบเทียบภาพ Lotus จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา)..... 38
4.20 เปรียบเทียบภาพ Green tea จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา)..... 39
4.21 เปรียบเทียบภาพ Cat จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา)..... 39
4.22 เปรียบเทียบภาพ Temple จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา) 40
4.23 เปรียบเทียบภาพ Maple จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา) 40
4.24 เปรียบเทียบค่า PSNR ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN..... 41
4.25 เปรียบเทียบค่า SSIM ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN..... 42
4.26 เปรียบเทียบค่า NIQE ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN..... 43
4.27 เปรียบเทียบค่า BRISQUE ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN..... 43
4.28 ภาพตัวอย่างในการทดสอบการสร้างภาพความละเอียดสูงวิธี SR-CNN เทียบกับวิธี Bicubic.. 44
4.29 เปรียบเทียบค่า PSNR ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN..... 45
4.30 เปรียบเทียบค่า SSIM ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN..... 46
4.31 เปรียบเทียบค่า NIQE ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN..... 47
4.32 เปรียบเทียบค่า BRISQUE ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN..... 47
4.33 กระบวนการปรับปรุงคุณภาพของภาพความละเอียดสูงด้วยตัวกรองแบบดิจิทัล 49
4.34 เปรียบเทียบค่าเฉลี่ย NIQE จากการใช้ตัวกรองดิจิทัลแบบไกต์ฟิลเตอร์..... 50
4.35 เปรียบเทียบค่าเฉลี่ย BRISQUE จากการใช้ตัวกรองดิจิทัลแบบไกต์ฟิลเตอร์..... 50
4.36 เปรียบเทียบค่าเฉลี่ย NIQE จากการใช้ตัวกรองดิจิทัลแบบไบเลเทอรอล..... 51
4.37 เปรียบเทียบค่าเฉลี่ย BRISQUE จากการใช้ตัวกรองดิจิทัลแบบไบเลเทอรอล..... 52
4.38 เปรียบเทียบค่าเฉลี่ย NIQE จากการใช้ตัวกรองดิจิทัลในโดเมนเวฟเล็ต..... 53
4.39 เปรียบเทียบค่าเฉลี่ย BRISQUE จากการใช้ตัวกรองดิจิทัลในโดเมนเวฟเล็ต..... 54
4.40 เปรียบเทียบค่าเฉลี่ย NIQE จากการประยุกต์ใช้ตัวกรองดิจิทัล..... 54
4.41 เปรียบเทียบค่าเฉลี่ย BRISQUE จากการประยุกต์ใช้ตัวกรองดิจิทัล..... 55

บทที่ 1
บทนำ

1.1 ความสำคัญและที่มาของปัญหา

ในปี ค.ศ. 1970 อุปกรณ์เซนเซอร์ตรวจรับสัญญาณภาพที่เรียกว่า CCD (Charge-coupled device) และ CMOS (Complementary metal-oxide-semiconductor) ได้ถูกใช้งานอย่างกว้างขวางในการแคปเจอร์ (Capture) สัญญาณภาพสำหรับกล้องดิจิทัลและกล้องวีดิทัศน์ ถึงแม้ว่าเซนเซอร์ดังกล่าวจะเหมาะกับการประยุกต์ใช้งานโดยทั่วไป แต่ราคาของกล้องดิจิทัลที่มีความละเอียดสูงนั้นยังคงมีราคาแพง ซึ่งโดยทั่วไปแล้ว เรามักจะต้องการภาพจากกล้องดิจิทัลหรือกล้องวีดิทัศน์ที่มีความละเอียดสูงแต่มีราคาถูก ในขณะที่นักวิทยาศาสตร์นักวิจัยหรือผู้ปฏิบัติงานในภารกิจที่มีความเสี่ยงสูงมักจะต้องการภาพจากกล้องดิจิทัลหรือกล้องวีดิทัศน์ที่มีความละเอียดสูงมากในระดับที่ใกล้เคียงกับภาพที่ได้จากกล้องฟิล์มขนาด 35 mm ซึ่งจะไม่ทำให้เกิดการแตกของภาพจนสังเกตเห็นได้ (Visible artifact) เมื่อภาพนั้นถูกขยายให้มีขนาดใหญ่ขึ้น ดังนั้นการค้นหาแนวทางที่จะเพิ่มระดับความละเอียดของภาพให้เพิ่มสูงขึ้นจากในระดับปัจจุบันยังคงมีความจำเป็นต่อไป

แนวทางในการเพิ่มระดับความละเอียดของภาพดิจิทัลในเชิงพื้นที่ (Spatial resolution) ให้สูงขึ้นคือ การลดขนาดของพิกเซล (Pixels) หรือการเพิ่มจำนวนของพิกเซลต่อหนึ่งหน่วยพื้นที่ ซึ่งแนวทางนี้สามารถทำได้ในขั้นตอนของการผลิตเซนเซอร์ตรวจรับสัญญาณภาพ อย่างไรก็ตามเมื่อขนาดของพิกเซลลดลง ปริมาณของแสงที่จำเป็นก็จะลดลงตามไปด้วย ทำให้เกิดสัญญาณรบกวนแบบ Short noise ขึ้นในภาพซึ่งทำให้คุณภาพของภาพดิจิทัลที่ได้ลดลงเป็นอย่างมาก การลดขนาดของพิกเซลโดยไม่ทำให้สัญญาณรบกวนดังกล่าวนั้น จำเป็นจะต้องไม่ลดขนาดของพิกเซลให้เล็กเกินขีดจำกัดการลดพิกเซลโดยค่าที่เหมาะสมที่สุดจะอยู่ที่ระดับประมาณ 40 μm^2 สำหรับเทคโนโลยีการผลิต CMOS ที่ 0.35 ไมครอน [1] ซึ่งเทคโนโลยีในการสร้างเซนเซอร์ตรวจรับสัญญาณภาพในปัจจุบันก็ใกล้จะถึงระดับนี้ อีกแนวทางหนึ่งในการเพิ่มระดับความละเอียดของภาพดิจิทัลในเชิงพื้นที่คือ การเพิ่มขนาดของชิพ (Chip size) ซึ่งก็คือการเพิ่มจำนวนตัวเก็บประจุไฟฟ้า แต่เนื่องจากค่าความจุไฟฟ้าที่มากนั้นทำให้ยากต่อการเพิ่มความเร็วในการถ่ายโอนข้อมูลภาพ วิธีการดังกล่าวนี้จึงไม่ประสพผลสำเร็จเท่าที่ควร

ในปัจจุบัน การเพิ่มระดับความละเอียดของภาพดิจิทัลในเชิงพื้นที่โดยใช้เทคนิคการประมวลผลภาพดิจิทัล (Digital image processing) เพื่อให้ได้ภาพความละเอียดสูงยิ่งยวด (Super-resolution, SR) ได้รับความสนใจมากขึ้น เนื่องจากตัวประมวลผลของคอมพิวเตอร์มีความเร็วสูงมากขึ้น แนวทางนี้สามารถสร้างภาพความละเอียดสูง (High-resolution, HR) จากภาพความละเอียดต่ำ (Low-resolution, LR) เพียงภาพเดียว (Single-frame SR) หรือภาพความละเอียดต่ำจำนวนหลาย ๆภาพ (Multi-frame SR) แนวทางดังกล่าวนี้ ถือได้ว่าเป็นหนึ่งในงานวิจัยที่กำลังแอกทีฟอยู่ในปัจจุบัน

นอกจากการประยุกต์ใช้กับระบบ CCTV แล้ว การสร้างคืนภาพความละเอียดสูงยิ่งยวดในปัจจุบันนี้มีการประยุกต์ใช้งานอย่างกว้างขวาง เช่น ภาพทางการแพทย์ (Medical image) ที่มีความละเอียดสูงจะสามารถช่วยให้แพทย์ทำการวินิจฉัยโรคได้อย่างถูกต้อง ภาพถ่ายดาวเทียม (Satellite image) ที่มีความละเอียดสูงจะทำให้สามารถสังเกตเห็นวัตถุในภาพได้ง่ายกว่าภาพถ่ายดาวเทียมแบบธรรมดา ในขณะเดียวกันประสิทธิภาพในการรู้จำของระบบการมองเห็นด้วยคอมพิวเตอร์ (Computer vision) สามารถปรับปรุงให้ดีขึ้นด้วยการใช้ภาพที่มีความละเอียดสูงยิ่งยวด นอกจากนี้แล้วการสร้างคืน

ภาพความละเอียดสูงยิ่งยวดสามารถนำมาประยุกต์ใช้กับภาพจากการบินลาดตระเวณถ่ายภาพมุมสูงโดยใช้อากาศยานไร้คนขับ (Unmanned aerial vehicle, UAV) ในพื้นที่เสี่ยงอันตราย เป็นต้น

1.2 วัตถุประสงค์ของการวิจัย

1. เพื่อออกแบบและพัฒนาอัลกอริทึมการสร้างคืนภาพความละเอียดสูง
2. เพื่อพัฒนาอัลกอริทึมการกำจัดสัญญาณรบกวน
3. เพื่อพัฒนาอัลกอริทึมของกระบวนการสร้างคืนภาพสำหรับการสร้างคืนภาพความละเอียดสูงด้วยเทคนิคและวิธีการทางปัญญาประดิษฐ์

1.3 ขอบเขตของการวิจัย

1. การพัฒนาอัลกอริทึมการสร้างคืนภาพความละเอียดสูงยิ่งยวดในโดเมนเชิงพื้นที่หรือโดเมนการแปลงเวฟเล็ต
2. ทำการจำลองอัลกอริทึมการสร้างภาพความละเอียดสูงด้วยโปรแกรม MATLAB และโปรแกรมภาษา C/C++

1.4 วิธีการดำเนินงานวิจัยและสถานที่ทำการวิจัย

วิธีการดำเนินงานวิจัย: มีขั้นตอนดังต่อไปนี้

1. ศึกษาเทคโนโลยีจัดเก็บข้อมูลภาพดิจิทัลสำหรับใช้ในการทดสอบ
2. การพัฒนาอัลกอริทึมการลงทะเบียนภาพสำหรับการสร้างคืนภาพความละเอียดสูง
3. การออกแบบและพัฒนาอัลกอริทึมการสร้างคืนภาพความละเอียดสูง
4. การทดสอบอัลกอริทึมตามที่ได้ทำการออกแบบไว้
5. สรุปผลการวิจัยและจัดทำรายงานฉบับสมบูรณ์

สถานที่วิจัย:

- อาคารเครื่องมือ 11 ศูนย์เครื่องมือวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีสุรนารี

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. ได้อัลกอริทึมการสร้างคืนภาพความละเอียดสูง
2. ได้บทความวิจัยการสร้างคืนภาพความละเอียดสูง ตีพิมพ์ในการประชุมวิชาการระดับชาติหรือนานาชาติในแวดวงวิชาการทางด้านวิศวกรรมไฟฟ้า โทรคมนาคมและคอมพิวเตอร์
3. ได้แนวทางและหลักการในการออกแบบและพัฒนาโปรแกรมการสร้างคืนภาพความละเอียดสูง
4. เป็นแนวทางในการเพิ่มสมรรถนะ การพัฒนาศักยภาพและขีดความสามารถในการแข่งขันของประเทศในด้านเทคโนโลยีสารสนเทศและการสื่อสารของประเทศ
5. ได้นักวิจัยรุ่นใหม่ที่มีการเรียนรู้ในด้านของเทคโนโลยีการประมวลผลภาพดิจิทัล

มหาวิทยาลัยเทคโนโลยีสุรนารี

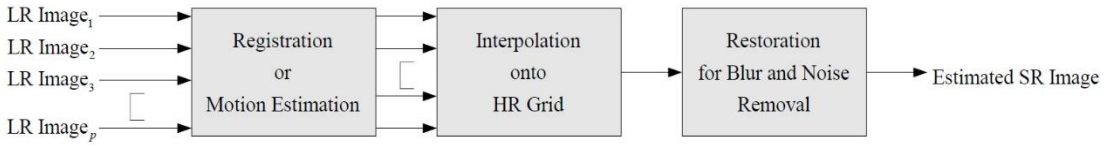
บทที่ 2
ทฤษฎีและหลักการที่เกี่ยวข้อง

2.1 บทนำ

ในบทนี้จะกล่าวถึง ทฤษฎีและหลักการที่เกี่ยวข้องกับงานวิจัย ซึ่งประกอบด้วย การสร้างภาพความละเอียดสูง การทบทวนวรรณกรรมที่เกี่ยวข้องในฐานข้อมูลต่าง ๆที่เกี่ยวข้องกับการสร้างภาพความละเอียดสูง จากนั้นจะได้กล่าวถึงการแปลงเวฟเล็ตโดยจะให้รายละเอียดของสมการการแปลงเวฟเล็ตแบบต่าง ๆที่นิยมใช้ในการประมวลสัญญาณภาพ และจะกล่าวถึงหลักการสำคัญของเครือข่ายประสาทแบบคอนโวลูชันซึ่งจัดได้ว่าเป็นการเรียนรู้เชิงลึกซึ่งผู้วิจัยจะนำมาใช้ในการสร้างภาพความละเอียดสูงในงานวิจัยนี้

2.2 การสร้างภาพความละเอียดสูง

ในปัจจุบัน การเพิ่มระดับความละเอียดของภาพดิจิทัลในเชิงพื้นที่โดยใช้เทคนิคการประมวลผลภาพดิจิทัล (Digital image processing) เพื่อให้ได้ภาพความละเอียดสูง (Super-resolution, SR) ได้รับความสนใจมากขึ้น ทั้งนี้เนื่องจากความสามารถในการประมวลผลของคอมพิวเตอร์สูงขึ้น โดยเฉพาะในส่วนของรีจีสเตอร์ (Register) ที่ใช้ในการประมวลผลที่มีจำนวนบิตมากขึ้น หน่วยความจำแคชระดับที่ 3 (Cache level 3) ที่มีความจุมากขึ้น ประกอบกับบัสในการรับส่งข้อมูล (Data bus) ที่มีความเร็วสูงมากยิ่งขึ้น โดยที่ราคาอาจจะปรับเพิ่มขึ้นบ้างแต่ก็ถือว่าไม่สูงมากเมื่อเทียบกับประสิทธิภาพที่ได้รับ ดังนั้นแนวทางในการสร้างภาพความละเอียดสูง (High-resolution, HR) จากภาพความละเอียดต่ำ (Low-resolution, LR) เพียงภาพเดียว (Single-frame SR) หรือใช้ภาพความละเอียดต่ำจำนวนหลาย ๆภาพ (Multi-frame SR) จึงเป็นแนวทางที่จะสามารถทำได้ และแนวทางดังกล่าวนี้ถือได้ว่าเป็นหนึ่งในงานวิจัยที่กำลังแอกทีฟอยู่ในปัจจุบัน [1] โดยนักวิจัยมักจะเรียกวิธีการดังกล่าวนี้ว่า การสร้างคืนภาพความละเอียดสูง หรืออาจเรียกอีกชื่อหนึ่งว่า การปรับเพิ่มระดับความละเอียด (Resolution enhancement) ข้อดีของเทคนิคดังกล่าวนี้คือ ใช้ต้นทุนที่ต่ำกว่าและสามารถนำไปประยุกต์ใช้กับระบบการสร้างภาพความละเอียดต่ำแบบเดิมที่ยังมีการใช้งานอยู่ แผนภาพการสร้างภาพความละเอียดสูง [1] แสดงดังรูปที่ 2.1



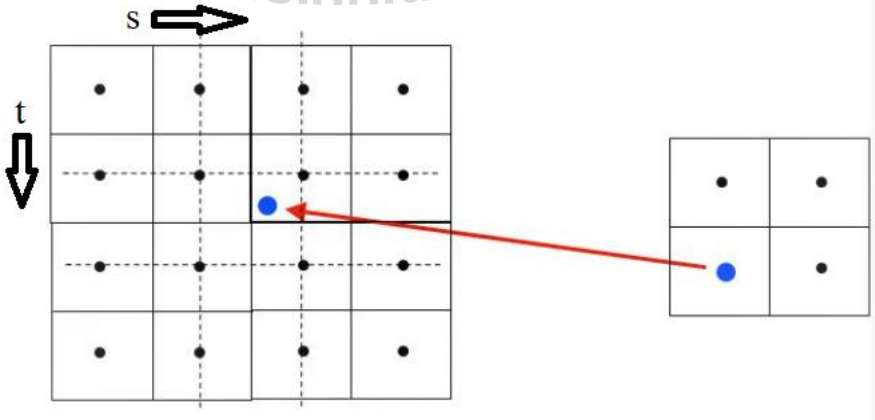
รูปที่ 2.1 การสร้างภาพความละเอียดสูง

การสร้างภาพให้มีความละเอียดสูงขึ้นหรือมีความชัดเจนมากยิ่งขึ้นวิธีการหนึ่งคือ การใช้เทคนิคการขยายภาพ (Resize) ให้ภาพเดิมมีขนาดของจุดภาพหรือพิกเซล (Pixel) ในภาพที่มากขึ้นด้วยการสุ่มตัวอย่างซ้ำ (Resampling) ซึ่งเป็นการสุ่มระดับสีจากจุดภาพเก่าที่อยู่ใกล้เคียงกับจุดภาพใหม่ การเพิ่มจำนวนจุดในภาพขึ้นมาใหม่จากการการสุ่มตัวอย่างซ้ำอาจเทียบได้กับการประมาณค่าหาจุดภาพใหม่จากจุดภาพรอบๆบริเวณข้างเคียงที่เรียกว่า (Neighborhood pixel) และมักจะเรียกวิธีการดังกล่าวว่าเป็นการทำอินเตอร์โพเลชัน (Interpolation)

ในการทำอินเตอร์โพเลชันแบบไบลิเนียร์ (Bilinear interpolation) จะทำการเพิ่มจุดภาพใหม่ขึ้นมา 1 จุดภาพโดยใช้การคำนวณหาจุดภาพใหม่ดังกล่าวจากจุดภาพเก่าที่อยู่บริเวณรอบๆจำนวน 4 จุดภาพ ส่วนการทำอินเตอร์โพเลชันแบบไบคิวบิค (Bilinear interpolation) จะทำการเพิ่มจุดภาพใหม่ขึ้นมา 1 จุดภาพโดยใช้การคำนวณหาจุดภาพใหม่ดังกล่าวจากจุดภาพเก่าที่อยู่บริเวณรอบๆจำนวน 16 จุดภาพซึ่งมีความซับซ้อนในการคำนวณมากกว่าการทำอินเตอร์โพเลชันแบบไบลิเนียร์ อย่างไรก็ตามการทำอินเตอร์โพเลชันแบบไบคิวบิคจะให้ผลลัพธ์ที่ดีกว่า การคำนวณแบบไบคิวบิคดังกล่าวจะยังคงรักษารายละเอียดของภาพเอาไว้ได้ทำให้ภาพผลลัพธ์ที่ได้มีการไล่โทนสีบนภาพที่ดูสวยงามและราบเรียบมากกว่าจึงเป็นวิธีการที่ได้รับความนิยมมากกว่าการทำอินเตอร์โพเลชันแบบไบลิเนียร์ การคำนวณหาค่าจุดภาพใหม่ในขั้นตอนการทำอินเตอร์โพเลชันแบบไบคิวบิคแสดงในสมการที่ 2.1 และตัวอย่างแผนภาพการคำนวณอินเตอร์โพเลชันแบบไบคิวบิคแสดงในรูปที่ 2.2

$$v_{pq} = \sum_{i,j=0}^{3} a_{ij}\, s^i t^j \;\; ; \; p,q = \{0..3\}, s,t \in [-1,2] \tag{2.1}$$

เมื่อ v_{pq} คือค่าจุดภาพใหม่ และ a_{ij} คือค่าสัมประสิทธิ์ของสมการจำนวน 16 ค่า



รูปที่ 2.2 การทำอินเตอร์โพเลชันแบบไบคิวบิค

มหาวิทยาลัยเทคโนโลยีสุรนารี

โครงการวิจัยนี้จะดำเนินการพัฒนาอัลกอริทึมการสร้างคืนภาพความละเอียดสูง เพิ่มรายละเอียดเชิงพื้นที่ของภาพ (Spatial resolution) ซึ่งทำได้โดยการหลอมรวมข้อมูลองค์ประกอบภาพจากลำดับของภาพความละเอียดต่ำหรือภาพที่ได้จากแหล่งกำเนิดภาพหลาย ๆแหล่งกำเนิด (Multi-source images) โดยจะมีการประมวลผลทั้งในโดเมนเชิงพื้นที่ (Spatial domain) และโดเมนของการแปลงเวฟเล็ต (Wavelet domain) ทั้งนี้เนื่องจากเวฟเล็ตมีคุณลักษณะสมบัติเฉพาะที่ทำให้สามารถทำการวิเคราะห์ภาพดิจิทัลได้แบบหลายระดับความละเอียด (Multi-resolution analysis) ซึ่งจะทำให้สามารถลดระดับของสัญญาณรบกวนในภาพให้มีขนาดน้อยลง นอกจากนี้ผู้วิจัยจะนำเทคนิคและวิธีการทางปัญญาประดิษฐ์อย่างใดอย่างหนึ่ง เช่น จีนเนติกอัลกอริทึม (Genetic algorithms) เครือข่ายประสาทเทียม (Artificial neural networks) เครื่องเวกเตอร์เกื้อหนุน (Support vector machine) มาประยุกต์ใช้ในการค้นหาค่าพารามิเตอร์ที่จำเป็นและสำคัญสำหรับเทคนิคที่ออกแบบขึ้น ซึ่งงานวิจัยนี้จะเป็นแนวทางใหม่ในการสร้างคืนภาพความละเอียดสูงต่อไป แนวทางดังกล่าวนี้สามารถทำได้อย่างมีประสิทธิภาพมากขึ้นเนื่องจากตัวประมวลผลคอมพิวเตอร์ในปัจจุบันมีความเร็วสูงมากยิ่งขึ้น

2.3 การทบทวนวรรณกรรม/สารสนเทศ (Information) ที่เกี่ยวข้อง

จากการทบทวนวรรณกรรมที่เกี่ยวข้องในฐานข้อมูลต่าง ๆ พบว่านักวิจัยหลาย ๆ ท่านได้พยายามที่จะแก้ปัญหาหลักของการสร้างคืนภาพความละเอียดสูง นั่นคือ ปัญหาในการลงทะเบียนภาพ (Image registration) และปัญหาในการสร้างคืนภาพ (Image reconstruction) ไม่ว่าจะเป็นการสร้างคืนภาพจากภาพความละเอียดต่ำเพียงภาพเดียว หรือจากภาพที่มีความละเอียดต่ำหลาย ๆภาพก็ตาม

บทความที่นำเสนอโดย Tsai and Huang [2] ซึ่งเป็นนักวิจัยกลุ่มแรกที่นำเสนอแนวทางการปรับปรุงความละเอียดของภาพดิจิทัลในโดเมนความถี่ (Frequency domain) โดยใช้การแปลงฟูเรียร์ (Fourier Transform) ซึ่งได้นำภาพดิจิทัลมาทำการลดระดับของการสุ่มตัวอย่างลงเพื่อให้ได้ภาพความละเอียดต่ำ (LR) จากนั้นทำการสร้างกลับภาพจากชุดของภาพหลาย ๆ ภาพที่ถูกลดระดับการสุ่มตัวอย่างลง ทำให้ได้ภาพที่มีความละเอียดสูงขึ้น หลักการสำคัญพื้นฐานของการสร้างคืนภาพความละเอียดสูงในโดเมนความถี่คือ (1) คุณสมบัติในการเลื่อน (Shifting) ของการแปลงฟูเรียร์ (2) ความสัมพันธ์ของการเกิดเอเลียต (Aliasing) ระหว่างสัญญาณภาพความละเอียดสูงต้นแบบที่ผ่านการแปลงฟูเรียร์แบบต่อเนื่อง (Continuous Fourier transform , CFT) และสัญญาณภาพที่ได้จากการสังเกตซึ่งมีความละเอียดต่ำและผ่านการแปลงฟูเรียร์แบบดิสครีต (Discrete Fourier transform , DFT) และ (3) สมมุติฐานที่ว่าสัญญาณภาพความละเอียดสูงต้นแบบต้องเป็นสัญญาณแบนด์จำกัด (Bandlimited signal)

สำหรับงานวิจัยที่นำเสนอแนวทางการแก้ปัญญาการลงทะเบียนภาพ (Image registration) ในขั้นตอนของการสร้างคืนภาพความละเอียดสูงนั้น ในบทความที่นำเสนอโดย Stark and Oskoui [3] ได้กล่าวถึงวิธีการสร้างภาพจากการโปรเจ็คที่เรียกว่า Projection onto convex sets (POCS) ซึ่งสามารถสร้างคืนภาพความละเอียดสูงได้ดีกว่าวิธีการทำอินเตอร์โพเลชัน (Interpolation) ที่ใช้กันทั่วไป ข้อดีของวิธีการ POCS คือง่ายต่อการสร้างจริง อย่างไรก็ตามวิธีดังกล่าวก็มีข้อเสียคือ มีชุดคำตอบหลายค่า (Non-uniqueness solution) และมีต้นทุนการคำนวณสูง (High computation cost)

ทฤษฎีการแปลงสัญญาณ ทั้งเวฟเล็ต (Wavelet) และคอนทัวร์เล็ต (Contourlet) ถูกนำมาประยุกต์ใช้กับการสร้างคืนภาพความละเอียดสูง โดย Jiji and Chaudhuri [4] นำเสนอเทคนิคการสร้างคืนภาพความละเอียดสูงในโดเมนของการแปลงคอนทัวร์เล็ต (Contourlet transform) การแปลงดังกล่าวมีพื้นฐานเดียวกันกับการแปลงเวฟเล็ต (Wavelet transform) โดยสามารถให้รายละเอียดความเรียบ (Smooth) ของภาพในทิศทางใด ๆ ได้ โดยสัมประสิทธิ์ของการแปลงคอนทัวร์เล็ตที่ระดับละเอียด (Finer scale) ของภาพความละเอียดสูงที่ไม่ทราบค่า จะได้จากการฝึกสอน (Training) โดยใช้ชุดข้อมูลของภาพความละเอียดสูง ทั้งนี้เพื่อให้เกิดการเรียนรู้คุณลักษณะเฉพาะของภาพความละเอียดสูง จากนั้นทำการแปลงย้อนกลับ (Inverse contourlet transform) เพื่อสร้างภาพกลับคืนก็จะได้ภาพที่มีความละเอียดสูงขึ้น ซึ่งผลการทดสอบนั้นยืนยันว่าดีกว่าวิธีการสร้างคืนภาพความละเอียดสูงทั่วไป

Lie and Zhiming [5] ได้นำการแปลงเวฟเล็ต (Wavelet transform) มาใช้ในการแปลงสัญญาณเพื่อสร้างคืนภาพความละเอียดสูงสำหรับประยุกต์ใช้กับวิธีการ POCS ซึ่งอัลกอริทึมดังกล่าวเป็นการสร้างภาพความละเอียดสูง (HR) จากชุดของภาพที่มีความละเอียดต่ำ (LR) จำนวนหลาย ๆ ภาพ ด้วยคุณลักษณะสมบัติเฉพาะของการแปลงเวฟเล็ตที่สามารถทำการวิเคราะห์ภาพดิจิทัลได้ในหลายระดับความละเอียด (Multi-resolution analysis) จึงทำให้สามารถทำการคัดแยกข้อมูลภาพที่เป็นองค์ประกอบหลักของความถี่สูงและยังช่วยในการลดระดับสัญญาณรบกวน จึงส่งผลให้อัลกอริทึมสามารถสร้างภาพความละเอียดสูงได้ดีกว่าวิธีการ POCS แบบดั้งเดิม

งานวิจัยของ Ji and Fermuller [6] ได้นำเสนอการสร้างคืนภาพความละเอียดสูงจากภาพจำนวนหลาย ๆ ภาพที่ได้จากฉาก (Scene) เดียวกัน ซึ่งจัดเก็บในรูปแบบของสัญญาณวีดิทัศน์ (Video signal) งานวิจัยนี้ใช้อัลกอริทึมการสร้างคืนภาพความละเอียดสูงแบบวนซ้ำ (Iterative reconstruction algorithm) โดยมีการประมาณค่าการเปลี่ยนแปลงของจุดภาพแต่ละเฟลมของลำดับภาพที่เรียงกันไปซึ่งนักวิจัยเรียกว่า Homography เพื่อให้สามารถจัดวางภาพแต่ละเฟลมภาพให้อยู่ในแนวเดียวกัน ซึ่งจะทำให้ได้รายละเอียดของภาพสูงสุด งานวิจัยนี้มีการใช้เวฟเล็ตในการแปลงสัญญาณเพื่อลดระดับของสัญญาณรบกวน ผลการวิจัยนำเสนอว่าดีกว่าวิธีการที่ใช้ทั่วไป

งานวิจัยที่นำเสนอแนวทางการแก้ปัญญาการสร้างคืนภาพ (Image reconstruction) สำหรับขั้นตอนของการสร้างคืนภาพความละเอียดสูง ได้มีการนำเสนอทั้งในแนวทางของ Deterministic approach และแนวทาง Stochastic approach โดยหลักการแล้วก็คือ เป็นการทำเรกูไรซ์ (Regularized) สำหรับปัญหาสมการเชิงเส้นที่มีสภาวะเลว (III-posed) โดยงานวิจัยของ Wang et al. [7] นำเสนอการสร้างคืนภาพความละเอียดสูงแบบอัตโนมัติโดยใช้อัลกอริทึมลูกผสมระหว่างวิธีการสร้างภาพความละเอียดสูงด้วยเทคนิค MAP (Maximum a posteriori) และ POCS ซึ่งเป็นแนวทาง Stochastic approach โดยเริ่มจากการประมาณค่าฟังก์ชันการแผ่กระจาย Point spread function (PSF) ของอุปกรณ์การรับสัญญาณภาพ ขั้นตอนต่อมาได้ทำการสร้างแบบจำลองภาพจาก HMRF (Huber-Markov random field) โดยทำการควบคุมค่าพารามิเตอร์ที่ใช้ด้วยเทคนิคการประมาณค่าที่เรียกว่า Maximum likelihood (ML) estimation จากนั้นจึงทำการสร้างคืนภาพความละเอียดสูงแบบอัตโนมัติ ด้วยเทคนิคการหาคำตอบแบบวนซ้ำที่เรียกว่า Two-phase iterative solution ซึ่งเป็นการสร้างคืนภาพภายใต้สภาวะเลว (Ill-posed reconstruction) ผลการวิจัยได้แสดงให้เห็นว่าวิธีการดังกล่าวนี้ให้ผลดีกว่าการสร้างคืนภาพความละเอียดสูงด้วยเทคนิคไบลิเนียร์อินเตอร์โพเลชัน (Bilinear interpolation) และวิธีการ ML-POCS

Mendez et al. [8] นำเสนอแนวทางใหม่ในการสร้างคืนภาพความละเอียดสูงจากภาพที่มีความละเอียดต่ำเพียงภาพเดียว วิธีการดังกล่าวใช้หลักการเรียนรู้จากค่าทางสถิติของภาพโดยเฉพาะองค์ประกอบของภาพบริเวณความถี่สูง ทั้งนี้เพื่อช่วยแยกแยะให้เห็นถึงความแตกต่างที่ชัดเจนระหว่างภาพที่มีความละเอียดสูงและภาพที่มีความละเอียดต่ำ ซึ่งงานวิจัยนี้ได้ใช้แบบจำลอง Markov random fields ในสกัดคุณลักษณะสมบัติเฉพาะในทางสถิติของภาพความละเอียดต่ำและภาพความละเอียดสูง

งานวิจัยของ Gajjar and Joshi [9] นำเสนอการสร้างคืนภาพความละเอียดสูงจากภาพที่มีความละเอียดในเชิงพื้นที่ต่ำซึ่งใช้แนวทางการเรียนรู้จากชุดของข้อมูลภาพที่มีทั้งของภาพความละเอียดสูงและภาพความละเอียดต่ำ โดยเริ่มจากการสร้างคืนภาพความละเอียดสูงด้วยการประมาณค่าที่ได้จากการเรียนรู้ในรายละเอียดขององค์ประกอบความถี่สูงของภาพที่ใช้แบบจำลอง Inhomogeneous Gaussian Markov random field (IGMRF) และนักวิจัยได้ใช้วิธีการ MAP (Maximum a posteriori estimation) ในการประมาณค่าพารามิเตอร์สำหรับแบบจำลองดังกล่าว สำหรับขั้นตอนการเรียนรู้นั้นงานวิจัยนี้ได้ใช้การแปลงเวฟเล็ตในการแปลงสัญญาณภาพทั้งนี้เพื่อที่จะทำการจับคู่ (Matching) สัมประสิทธิ์ของการแปลงเวฟเล็ตของภาพ LR และภาพ HR จากฐานข้อมูลภาพ นักวิจัยได้ทำการทดสอบอัลกอริทึมกับภาพสี (Color image) และภาพระดับเทา (Gray level image) โดยภาพความละเอียดสูงที่ได้พบว่ามีความคมชัดสูงและมีสัญญาณรบกวนต่ำ

การสร้างคืนภาพความละเอียดสูง สามารถมองได้ว่าเป็นปัญหาการหาค่าเหมาะที่สุด (Optimization problems) โดยในงานวิจัยของ Ellis and Walker [10] ได้นำเสนอเทคนิคการลดเวลาการคำนวณในการสร้างคืนภาพความละเอียดสูงที่เรียกว่า QTONE (Quick time domain optimized near-field estimator) โดยเทคนิคดังกล่าวได้ทำการสุ่มตัวอย่าง (Sampling) สัญญาณภาพแบบละเอียดเฉพาะในบริเวณที่สนใจ (Region of interest, ROI) เท่านั้น จากนั้นจึงทำการปรับปรุงประสิทธิภาพของระบบด้วยการออปติไมเซชันแบบวงกว้าง (Global optimization) ซึ่งทำให้สามารถลดเวลาในการคำนวณได้ถึง 8 เท่าเมื่อเทียบกับวิธีการเดิมก่อนการปรับปรุง

งานวิจัยของ Wong et al. [11] ได้นำเสนองานวิจัยการสร้างคืนภาพความละเอียดสูงสำหรับภาพ Mammogram image โดยใช้เทคนิคการวิเคราะห์ค่าทางสถิติของโมเมนต์ (Statistical moment analysis, SMA) ซึ่งงานวิจัยนี้มองปัญหาของการสร้างภาพ SR ว่าเป็นการทำออฟติไมเซชันแบบค่าคงตัว (Constrained optimization problem) โดยประยุกต์ใช้แบบจำลองแบบปรับตัวได้อันดับสามของมาร์คอฟ (Adaptive third-order Markov prior model) จากนั้นทำการแก้ปัญหาโดยใช้แนวทางคอนจุเกตการ์เดียน (Conjugate gradient approach) ซึ่งผลที่ได้พบว่าดีกว่าวิธีการขยายภาพด้วยวิธีไบคูบิคอินเตอร์โพเลชัน (Bicubic interpolation)

นอกจากนี้ยังพบว่ามีงานวิจัยที่นำเทคนิคและวิธีการทางปัญญาประดิษฐ์ (Artificial intelligence) มาประยุกต์ใช้ในการหาคำตอบที่เหมาะที่สุด เช่น งานวิจัยของ Mei et al. [12] นำเสนอการสร้างคืนภาพความละเอียดสูงด้วยเทคนิคทางปัญญาประดิษฐ์ โดยได้นำจีนเนติกอัลกอริทึม (Genetic algorithm) มาใช้ในการหาค่าพารามิเตอร์ที่เหมาะสมที่สุดของการสร้างภาพความละเอียดสูง นักวิจัยได้นำเสนอการปรับปรุงการทำงานของจีนเนติกอัลกอริทึมโดยเพิ่มตัวปฏิบัติการ Gradient descent operator ในขั้นตอนการปฏิบัติการทางสายพันธุ์ (Genetic operation) ซึ่งพบว่าได้ผลลัพธ์ที่ดีกว่าวิธีการเดิม

Ni et al. [13] นำเสนอการสร้างคืนภาพความละเอียดสูงในโดเมนเชิงพื้นที่ (Spatial domain) และโดเมนการแปลงโคไซน์แบบไม่ต่อเนื่อง (Discrete cosine transform, DCT) โดยใช้ภาพความละเอียดต่ำเพียงภาพเดียว ผู้วิจัยได้ใช้เทคนิคการทำออปติไมเซชัน (Optimization) ที่เรียกว่า Support vector machine (SVM) ทั้งนี้เพื่อหาค่าพารามิเตอร์ที่เหมาะที่สุดในสร้างภาพความละเอียดสูง ผู้วิจัยกล่าวว่าด้วยโครงสร้างของการแปลงโคไซน์แบบไม่ต่อเนื่องทำให้ได้ผลลัพธ์ของการสร้างคืนภาพความละเอียดสูงที่ดีกว่าผลลัพธ์ในโดเมนเชิงพื้นที่

Zhang and Lu [14] นำเสนอวิธีการสร้างคืนภาพความละเอียดสูงด้วยเทคนิคทางปัญญาประดิษฐ์ โดยฟังก์ชันวัตถุประสงค์ถูกสร้างขึ้นจากแบบจำลองการสังเกตที่ใกล้ชิดมากกับกระบวนการจัดเก็บข้อมูลภาพดิจิทัลในทางฟิสิกส์ โดยโครงข่ายประสาทเทียมแบบ Hopfield Neural

network ถูกนำใช้ในการหาคำตอบที่เหมาะที่สุด งานวิจัยได้นำฟังก์ชันการแผ่กระจาย PSF ในการเบลอของภาพและการเกิดของสัญญาณรบกวน มาพิจารณาในการสร้างภาพความละเอียดสูงด้วย ซึ่งการวิจัยนำเสนอว่าดีกว่าวิธีการพื้นฐานที่ใช้โดยทั่วไปเช่น เทคนิคไบลีเนียร์และไบคูบิคอินเตอร์โพเลชัน (Bicubic interpolation)

การสร้างคืนภาพความละเอียดสูง สามารถประยุกต์ใช้ได้หลายแนวทางเช่น การเพิ่มความละเอียดของภาพสำหรับระบบรู้จำใบหน้าและเป้าหมายอัตโนมัติ [15] การเพิ่มความละเอียดของภาพถ่ายดาวเทียม [16] ตลอดจนภาพทางการแพทย์ (Medical image) เช่น งานวิจัยของ Gompel et al. [17] ได้นำเสนอการสร้างคืนภาพความละเอียดสูงด้วยเทคนิคที่เรียกว่า DART (Discrete algebraic reconstruction technique) ซึ่งประยุกต์ใช้กับภาพจากกล้อง Micro CT scan ซึ่งงานวิจัยนี้สามารถลดปัญหาที่เกิดจากข้อจำกัดของอุปกรณ์ เช่น ขนาดโดสของรังสี (X-ray dose limitations) และเวลาในการสแกนภาพ (Scanning times) ทำให้สามารถสร้างคืนภาพความละเอียดสูงได้ชัดเจนมากขึ้นเมื่อเทียบกับวิธีสร้างคืนภาพความละเอียดสูงทั่วไป

งานวิจัยของ R. Xi [18] ได้นำเสนอเทคนิคการสร้างภาพความละเอียดสูงยิ่งสำหรับภาพ SAR (synthetic aperture radar) โดยใช้วิธีการ Matching pursuit ซึ่งถือว่าเป็นวิธีการสร้างภาพความละเอียดสูงยิ่งที่มีประสิทธิภาพแต่ทว่ามีการคำนวณที่ค่อนข้างจะซับซ้อนมาก ผู้วิจัยจึงได้นำจีนเนติกอัลกอริทึมซึ่งเป็นเทคนิคทางปัญญาประดิษฐ์ มาใช้ในการค้นหาค่าพารามิเตอร์ที่เหมาะสมที่สุดของวิธีการดังกล่าวทำให้ได้ภาพผลลัพธ์ที่ดีมากขึ้นและใช้เวลาในการประมวลผลน้อยลงกว่าเดิม

งานวิจัยของ H. Okuhata et al. [19] ได้นำเสนอเทคนิคการสร้างภาพความละเอียดสูงยิ่งสำหรับภาพทางการแพทย์ (Medical image) โดยใช้วิธีการขยาย Dynamic range ในภาพดิจิทัลซึ่งเป็นวิธีการที่ใช้กับภาพจากกล้องดิจิทัลทั่วไป นอกจากนี้ผู้วิจัยยังได้ประยุกต์ใช้เทคนิคการสร้างภาพความละเอียดสูงแบบเฟรมเดี่ยวซึ่งถือว่าเป็นวิธีการสร้างภาพความละเอียดสูงยิ่งที่มีประสิทธิภาพแต่ทว่ามีการคำนวณที่น้อยมากผู้วิจัยจึงสามารถทำการประมวลผลได้แบบเวลาจริง

ในงานวิจัยของ J. Xu et al. [20] ได้นำเสนอเทคนิคการสร้างภาพความละเอียดสูงยิ่งโดยใช้เครื่องเวกเตอร์เกื้อหนุน (Support vector machine) ในการจำแนกและจัดกลุ่มคลาสของพิกเซลในแต่ละบริเวณของภาพ โดยนักวิจัยได้ทำการสร้างแบบจำลองสำหรับเครื่องเวกเตอร์เกื้อหนุน (Support vector machine) ที่เรียกว่า Multi-layer support vector regression model ซึ่งใช้แทนความสัมพันธ์ระหว่างภาพความละเอียดสูงและภาพความละเอียดต่ำ ซึ่งผู้วิจัยกล่าวว่าวิธีการที่นำเสนอให้ผลลัพธ์ที่ดีกว่าวิธีการพื้นฐานทั่วไป

2.4 การแปลงเวฟเล็ต (Wavelet Transform)

การแปลงเวฟเล็ต ได้ถูกคิดค้นขึ้นมาเพื่อใช้กับสัญญาณหรือสัญญาณภาพที่ต้องการข้อมูลทางด้านเวลาและความถี่ โดยใช้หลักในการแทนคลื่นใด ๆ ด้วยคลื่นเล็ก ๆ ที่มีความยาวและพลังงานจำกัดทำให้ได้ทั้งเวลาและความถี่ของคลื่นใด ๆ นั้น [21] การแปลงเวฟเล็ตได้ถูกพัฒนาขึ้นเพื่อใช้ในการแก้ไขปัญหาที่ไม่สามารถกระทำได้ในการแปลงฟูเรียร์ (Fourier Transform) บ่อยครั้งที่การแปลงเวฟเล็ตถูกนำไปเปรียบเทียบกับการแปลงฟูริเยร์ จุดแตกต่างหลักของการแปลงทั้งสอง คือ การแปลงเวฟเล็ตเป็นการพิจารณาความถี่ที่เกิดขึ้นในแต่ละช่วงเวลา ทำให้ผลลัพธ์การแปลงที่ได้ให้ข้อมูลทั้งในแกนเวลาและแกนความถี่ ในขณะที่การแปลงฟูริเยร์เป็นการพิจารณาความถี่โดยรวมตลอดสัญญาณอินพุต อย่างไรก็ดีมีความพยายามที่จะเปรียบเทียบการแปลงเวฟเล็ตกับการแปลงฟูริเยร์ในช่วงเวลาสั้นๆ (Short Time Fourier Transform; STFT) ซึ่งพิจารณาข้อมูลทั้งเวลาและความถี่โดยอาศัยการเลื่อนหน้าต่างเพื่อทำการแปลง เมื่อเปรียบเทียบผลการแปลงสัมประสิทธิ์ด้วยเทคนิคทั้งสองแล้ว การแปลงเวฟเล็ตจะให้ผลลัพธ์การเป็นตัวแทนสัญญาณได้ดีกว่าโดยผ่านทางการวิเคราะห์สัญญาณแบบหลายความละเอียด (Multiresolution Analysis)

เวฟเล็ต เป็นการวิเคราะห์สัญญาณแบบ Multiresolution โดยการแปลงเวฟเล็ตจะได้ผลลัพธ์เป็นค่าสัมประสิทธิ์เวฟเล็ต (Wavelet Coefficient) การแปลงเวฟเล็ตจะเป็นการแปลงค่าสัมประสิทธิ์ผ่านฟังก์ชันหน้าต่าง แต่ฟังก์ชันหน้าต่างนี้สามารถเลื่อนตำแน่งและปรับเปลี่ยนช่วงเวลาให้เหมาะสมกับความถี่ที่จะวิเคราะห์ โดยเรียกฟังก์ชันหน้าต่างนี้ว่า ฟังก์ชันเวฟเล็ต (Wavelet Function)

ทฤษฎีเวฟเล็ตใช้ในการอธิบายสิ่งใดสิ่งหนึ่งที่เสมือนกับการแยกสิ่งเหล่านั้นออกเป็นส่วนประกอบเล็กๆที่มีความสัมพันธ์กัน ซึ่งอยู่ในรูปของเวฟเล็ตที่ถูกปรับขนาดและเลื่อนตำแหน่งและมีค่าถ่วงน้ำหนัก (สัมประสิทธิ์เวฟเล็ต) ต่างๆ กันคูณอยู่กับเวฟเล็ตแต่ละตัว การทำส่วนการแปลงกลับเวฟเล็ต (Inverse Wavelet Transform; IWT) หรือเรียกอีกอย่างหนึ่งว่า การแปลงกลับเวฟเล็ต (Wavelet Reconstruction)

ให้ $f(t)$ เป็นสัญญาณที่ประกอบด้วยหลายความถี่ ถ้าเราพิจารณาในรูปแบบ 2 มิติ โดยแกนหนึ่งเป็นแกนเวลา และอีกแกนหนึ่งเป็นแกนของความถี่ ดังนั้นสามารถเขียนสมการสำหรับการแปลงเวฟเล็ตแบบต่อเนื่อง (CWT) ที่ใช้ในการแตกกระจายเวฟเล็ตของสัญญาณของฟังก์ชันใดๆ $f(t)$ ด้วยเวฟเล็ตแม่ (Mother wavelet) $M(t)$ เป็นดังนี้

$$W_M(a,b) = |a|^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(t) M^* \left(\frac{t-b}{a}\right) dt \qquad (2.2)$$

โดย * คือ Complex Conjugate ของ $M(t)$

$M(t)$ คือ ค่าสัมประสิทธิ์เวฟเล็ตที่ได้เป็นผลลัพธ์

ค่า $M(t)$ เป็นค่าสัมประสิทธิ์ที่ได้เป็นผลลัพธ์ ซึ่งอาจเทียบได้ว่าเป็นการวิเคราะห์ความสัมพันธ์ (Cross-correlation) ระหว่างสัญญาณที่พิจารณา $f(t)$ และฟังก์ชันเวฟเล็ตแม่ (Mother Wavelet) โดย $M(t)$ ที่ถูกเลื่อนตำแหน่งด้วยพารามิเตอร์ b และสเกลความถี่ด้วยพารามิเตอร์ a ซึ่งจะ

มีลักษณะการเปลี่ยนแปลงเป็นช่วงเวลาและความถี่ที่สัมพันธ์กันคือ ในช่วงเวลาความถี่สูงจะใช้ช่วงเวลาการวิเคราะห์สั้นและในช่วงความถี่ต่ำจะมีช่วงเวลาในการวิเคราะห์ที่ยาวนานขึ้น

2.4.1 การแปลงเวฟเล็ตแบบดิสครีต (Discrete Wavelet Transform; DWT)

การแปลงเวฟเล็ตสามารถทำได้ทั้งแบบต่อเนื่อง (Continuous Wavelet Transform; CWT) และแบบไม่ต่อเนื่อง แต่ในที่นี้กล่าวเฉพาะการแปลงเวฟเล็ตแบบไม่ต่อเนื่องเท่านั้น ฟังก์ชันการแปลงเวฟเล็ตแบบไม่ต่อเนื่องนี้เป็นการปรับสเกลความถี่ (Scale) และค่าการเลื่อนตำแหน่ง (Translate) ที่ไม่ต่อเนื่องด้วย โดยการแปลงเวฟเล็ตจะทำการคำนวณหาค่าสัมประสิทธิ์การแปลง $DWT_{j,k}$ จากการหาความสัมพันธ์ของสัญญาณอินพุต $f(x)$ กับฟังก์ชันเวฟเล็ตแม่ $M(x)$ ที่เลือกใช้ในแต่ละสเกล (j) และการเลื่อนตำแหน่ง (k) ดังสมการที่ (2.3) และสมการที่ (2.3)

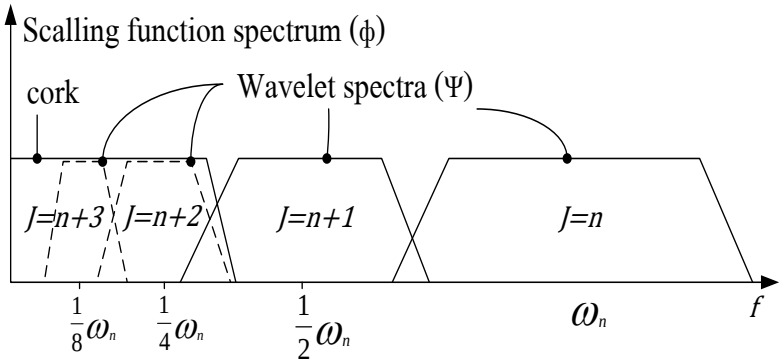
การเลื่อนตำแหน่ง (k) ของฟังก์ชันเวฟเล็ตแม่ ซึ่งแน่นอนย่อมถูกจำกัดด้วยช่วงของสัญญาณที่ทำการวิเคราะห์ $[0, N-1]$ แต่เมื่อพิจารณาขอบเขตการปรับความถี่ของฟังก์ชันเวฟเล็ตแม่ จะพบว่าเมื่อเราทำการปรับขนาดของฟังก์ชันเวฟเล็ตแม่ในแกนเวลาด้วยน้ำหนักกำลังสอง (2^j) เทียบได้กับการลดขอบเขตช่วงของความถี่ลงไปครึ่งหนึ่ง จากคุณสมบัติฟูริเยร์ในสมการที่ (2.5)

$$DWT_{j,k} = \frac{1}{\sqrt{N}}\left[2^j \sum_{x=0}^{N-1} f(x) W_{j,k}(x)\right] \tag{2.3}$$

$$= \frac{1}{\sqrt{N}}\left[2^j \sum_{x=0}^{N-1} f(x) M(2^j x - k)\right] \tag{2.4}$$

$$F\{f(at)\} = \frac{1}{|a|} F\left(\frac{\omega}{a}\right) \tag{2.5}$$

ดังนั้น การยืดขยายฟังก์ชันเวฟเล็ตแม่ไปสองเท่าจะเป็นการลดช่วงกว้างของความถี่ที่วิเคราะห์ลงไปครึ่งหนึ่ง แสดงดังรูปที่ 2.3



รูปที่ 2.3 การลดช่วงกว้างของความถี่ทีละครึ่งหนึ่งสำหรับค่าแต่ละค่าสเกล (อรฉัตร จิตต์ภักตร์, 2552)

ดังนั้นหากต้องการวิเคราะห์สัมประสิทธิ์ความถี่เวฟเล็ตย่อยลงไปในลักษณะที่ความถี่ถูกปรับลดลงทีละครึ่ง และต้องการวิเคราะห์ความถี่ต่ำที่ความถี่ศูนย์ $f = 0$ จำเป็นต้องทำการแบ่งช่วงความถี่ย่อยลงไปถึงจำนวนอนันต์ค่า จึงจะสามารถวิเคราะห์ความถี่ต่ำลงไปจนครอบคลุมถึงความถี่ศูนย์ได้ ซึ่งในทางปฏิบัตินั้นใช้เวลานานเกินไป และถ้าเป็นในกรณีของสัญญาณไม่ต่อเนื่องนั้นเป็นไปไม่ได้เลยที่จะทำการวิเคราะห์ความถี่ศูนย์ได้ ดังนั้นสามารถวิเคราะห์ความถี่ต่ำที่ผ่านตำแหน่งความถี่ศูนย์ (Low Pass Spectrum) แยกออกจากช่วงความถี่สูง (High Pass Spectrum) แทนการพยายามวิเคราะห์ความถี่ย่อยลงเป็นจำนวนอนันต์ครั้ง โดยฟังก์ชันที่วิเคราะห์ช่วงความถี่ต่ำเรียกว่า สเกลลิ่งฟังก์ชัน (Scaling Function; $S(f)$) และฟังก์ชันที่วิเคราะห์ช่วงความถี่สูงจะเรียกว่า เวฟเล็ตฟังก์ชัน (Wavelet Function; $W(t)$)

2.4.2 สเกลลิ่งฟังก์ชัน (Scaling Function)

คุณสมบัติของสเกลลิ่งฟังก์ชันค่าโมเมนต์ที่ศูนย์จะต้องไม่เป็นศูนย์ ดังแสดงในสมการที่ (2.6)

$$\int_{-\infty}^{\infty} S(t)dt = 1 \tag{2.6}$$

เราสามารถทำการแปลงสัมประสิทธิ์ความถี่ต่ำของเวฟเล็ต (Scaling Coefficients; DWT_S) ด้วยสเกลลิ่งฟังก์ชัน $S(t)$ ดังแสดงในสมการที่ (2.7)

$$DWT_S(j_0,k) = \frac{1}{\sqrt{N}}\left[2^{j_0/2}\sum_{x=0}^{N-1} f(x)S(2^{j_0}x-k)\right] \tag{2.7}$$

โดย j_0 คือ ค่าเสกลความถี่สิ้นสุดของสเกลลิ่งฟังก์ชัน
k คือ ค่าตำแหน่งการเลื่อน

2.4.3 เวฟเล็ตฟังก์ชัน (Wavelet Function)

เวฟเล็ตฟังชันแสดงด้วยสมการดังต่อไปนี้

$$DWT_w(j,k) = \frac{1}{\sqrt{N}}\left[2^{j/2}\right]\sum_{x=0}^{N-1} f(x)W\left(2^j x-k\right) \tag{2.8}$$

โดย j คือ ค่าเสกลความถี่เริ่มต้นที่ j_0
k คือ ค่าตำแหน่งการเลื่อน

การแปลงสัมประสิทธิ์เวฟเล็ตด้วยสเกลลิ่งฟังก์ชันและเวฟเล็ตฟังก์ชัน และบางครั้งถูกเรียกว่าการแยกองค์ประกอบความถี่ของเวฟเล็ต (Wavelet Decomposition) ในบางครั้งสเกลลิ่งฟังก์ชันอาจเรียกว่า ฟังก์ชันการประมาณค่า (Approximation Function) ในขณะที่เวฟเล็ตฟังก์ชันอาจเรียกว่าฟังก์ชันรายละเอียด (Detail Function) ได้เช่นกัน

มหาวิทยาลัยเทคโนโลยีสุรนารี

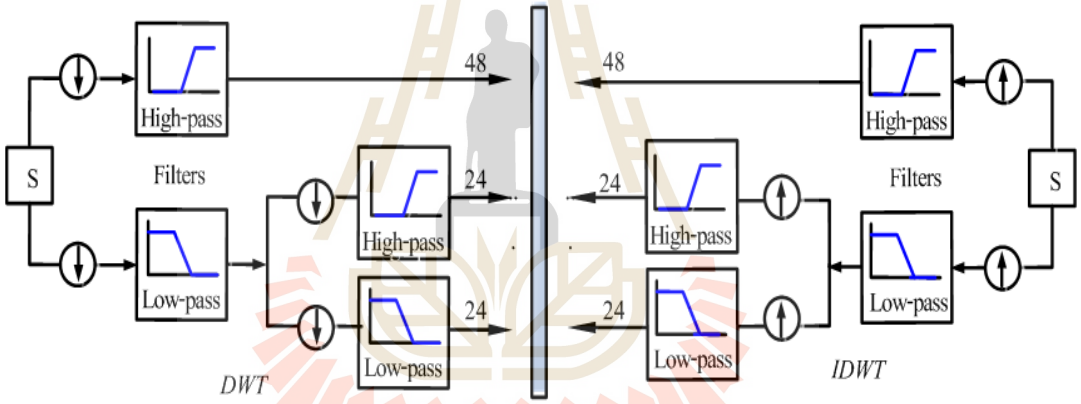
2.4.4 การแปลงกลับเวฟเล็ตแบบดีสครีต (Inverse Discrete Wavelet Transform; IDWT)

การแปลงเวฟเล็ตเพื่อทำการสร้างคืน (Reconstruction) สัญญาณอินพุตต้นฉบับ $f(x)$ โดยการคำนวณผลรวมของความสัมพันธ์ของค่าสัมประสิทธิ์เวฟเล็ตและฟังก์ชันการแปลงสัมประสิทธิ์จากทางสเกลลิ่งฟังก์ชันและเวฟเล็ตฟังก์ชันที่แต่ละสเกล (j) และการเลื่อนตำแหน่ง (k) ดังแสดงในสมการที่ (2.9)

$$f(x) = \frac{1}{\sqrt{N}} \sum_{k} DWT_s(j_0, k) 2^{j_0/2} S_{j_0,k}(2^{j_0}x - k)$$

$$+ \frac{1}{\sqrt{N}} \sum_{j=j_0}^{\infty} \sum_{k} DWT_w(j, k) 2^{j/2} W_{j,k}(2^{j}x - k) \qquad (2.9)$$

การแยกองค์ประกอบความถี่ของเวฟเล็ตและการแปลงกลับเวฟเล็ตแบบดีสครีตแสดงในรูปที่ 2.4



รูปที่ 2.4 การแยกองค์ประกอบความถี่ของเวฟเล็ตและการแปลงกลับเวฟเล็ตแบบดีสครีต

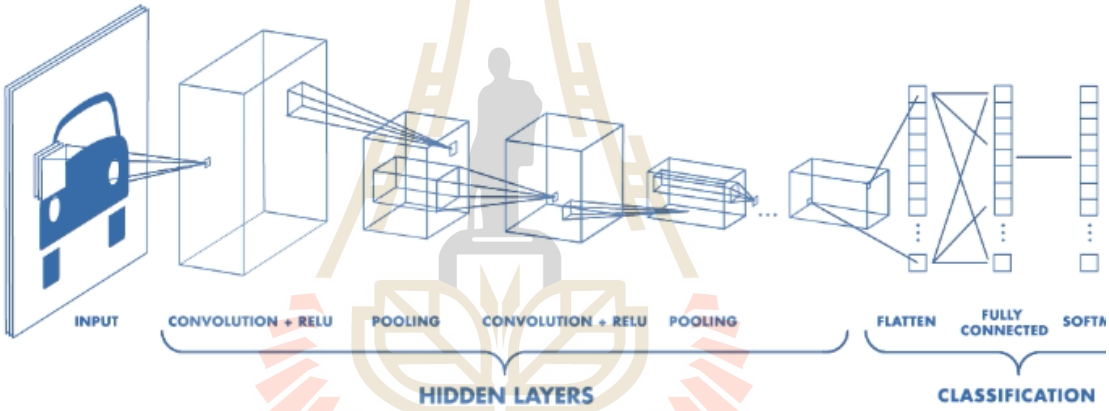
มหาวิทยาลัยเทคโนโลยีสุรนารี

2.5 เครือข่ายประสาทแบบคอนโวลูชัน (Convolutional Neural Network, CNN)

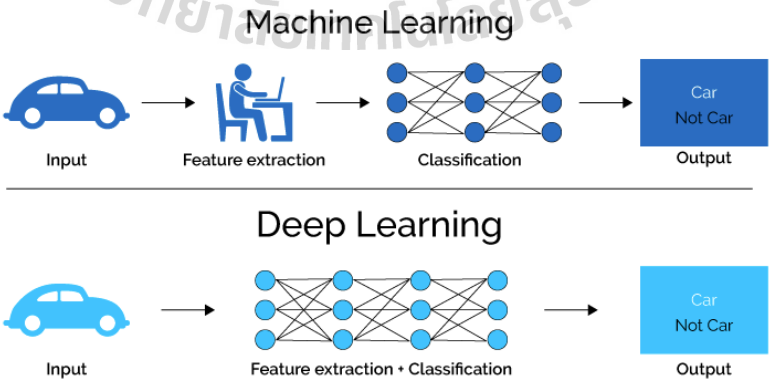
เครือข่ายประสาทแบบคอนโวลูชัน (Convolutional Neural Network) หรือ CNN จัดเป็นเครือข่ายประสาทเทียมประเภทหนึ่งที่อยู่ในกลุ่มวิธีการค้นหาคำตอบที่เหมาะสมที่สุด (Optimizations) ที่ได้รับแรงบันดาลใจจากสิ่งมีชีวิตจากธรรมชาติ (Bio-inspired) [22] โดยที่ CNN จะจำลองการมองเห็นของมนุษย์ที่มองพื้นที่เป็นส่วนที่ย่อย ๆจากนั้นนำกลุ่มของพื้นที่ย่อย ๆเหล่านั้นมาผสานกันเพื่อตรวจสอบดูว่าสิ่งที่กำลังมองอยู่ในพื้นที่หรือภาพดังกล่าวหรือไม่ หรืออาจกล่าวได้ว่าเป็นการตรวจจับหรือการสกัดลักษณะเด่น (Feature extraction) ของภาพ และเรียกว่าเป็น “Local feature” ของภาพดิจิทัลนั้น ๆ ค่า Local feature ของภาพที่ตรวจจับได้นั้นถือเป็นส่วนหนึ่งของเครือข่ายประสาทเทียม (Artificial neural networks, ANN) ที่จะถูกปรับไปพร้อม ๆ กับตัวที่ใช้ในการแบ่งแยกประเภท (Classifier) ดังนั้นเมื่อทำการสอนให้เครือข่ายทำการเรียนรู้ (Training) เสร็จสิ้นแล้วจะได้ลักษณะของ

เครือข่ายประสาทเทียมที่สามารถสกัดและจำแนกลักษณะเด่น (Feature extractor) ที่เหมาะสมกับงาน ซึ่งถือว่าเป็นจุดเด่นของเครือข่าย CNN

เครือข่ายประสาทแบบคอนโวลูชันจัดว่าเป็นการเรียนรู้เชิงลึก (Deep learning) ซึ่งมีความแตกต่างไปจากการเรียนรู้ของเครื่องจักร (Machine learning) ทั่ว ๆไป ที่ผู้ใช้งานจะต้องทำการสกัดลักษณะเด่นด้วยตนเองก่อนจะป้อนเป็นอินพุตให้กับเครือข่ายประสาทเทียมใช้ในการเรียนรู้ ส่วนการเรียนรู้เชิงลึกนั้นมีการใช้เครือข่าย ANN ที่มีชั้นซ่อนเร้น (Hidden layers) หลายชั้นทั้งนี้เพื่อเพิ่มความสามารถในการคิดที่มากกว่าปกติ ทำให้สามารถทำการคำนวณโจทย์ปัญหาซับซ้อนได้ สามารถใช้เทคนิคต่าง ๆได้มากขึ้น และที่สำคัญที่สุดคือ CNN สามารถทำการคิดอย่างเป็นขั้นเป็นตอนได้ซึ่งสามารถลอกเลียนแบบการทำงานของสมองมนุษย์ได้ดีขึ้น [23] แผนภาพตัวอย่างสถาปัตยกรรมของเครือข่ายประสาทแบบคอนโวลูชันและแผนภาพแสดงความแตกต่างระหว่าง Machine learning กับ Deep learning แสดงในรูปที่ 2.5 และรูปที่ 2.6 ตามลำดับ [32]



รูปที่ 2.5 สถาปัตยกรรมของเครือข่ายประสาทแบบคอนโวลูชัน (Mathworks, 2019)



รูปที่ 2.6 ความแตกต่างระหว่าง Machine learning กับ Deep learning (Mathworks, 2019)

สถาปัตยกรรมของเครือข่ายประสาทแบบคอนโวลูชันแสดงในรูปที่ 2.5 ซึ่งแสดงให้เห็นว่า CNN คือ Neural networks (NN) ที่มีหลายเลเยอร์ (Layers) และโครงสร้างเฉพาะตัวซึ่งถูกออกแบบมาเพื่อการเพิ่มความสามารถในการสกัดเอาลักษณะเด่น (Features) ที่มีความซับซ้อนมากยิ่งขึ้นจากข้อมูล ในปัจจุบัน CNN มักจะถูกใช้เพื่อการสกัดลักษณะเด่นจากข้อมูลประเภทที่ไม่ค่อยเป็นระเบียบหรือข้อมูลที่ไม่ได้มีโครงสร้างเป็นรูปแบบเฉพาะตัว (Unstructured data) เช่น ภาพดิจิทัล (Digital image) เป็นต้น

การคำนวณตามสถาปัตยกรรรมของ CNN มีขั้นตอนการคำนวณที่แบ่งออกได้ 3 ขั้นตอนคือ ขั้นตอนการคอนโวลูชัน (Convolution stage) ขั้นตอนการตรวจจับ (Detector stage) และขั้นตอนการพูลลิ่ง (Pooling stage)

1.) ขั้นตอนการคอนโวลูชัน (Convolution stage) การคำนวณในขั้นตอนนี้จะใช้หลักการเดียวกันกับการคำนวณคอนโวลูชันเชิงพื้นที่ (Spatial convolution) ที่ใช้ในงานในด้านการประมวลผลภาพดิจิทัล (Digital image processing) จุดมุ่งหมายของการคำนวณคอนโวลูชันเชิงพื้นที่กับรูปภาพดิจิทัลก็คือ การสกัดลักษณะเด่นจากรูปภาพอินพุตแบบดิจิทัล โดยการคำนวณคอนโวลูชันทำให้เกิดการแปลงเชิงเส้น (Linear transformation) ของรูปภาพอินพุตที่สอดคล้องกับในข้อมูลเชิงพื้นที่จากตัวกรอง (Filters) โดยมีค่าถ่วงน้ำหนัก (Weight) ของแต่ละ layer จะเป็นตัวกำหนดรายละเอียดของคอนโวลูชันเคอร์นอล (Convolution kernel) ดังนั้น Convolution kernel สามารถทำการเทรน (Training) หรือทำการฝึกสอนได้และขึ้นอยู่กับอินพุตของเครือข่ายประสาทเทียบแบบ CNN

ขั้นตอนการคอนโวลูชันจะเริ่มจากการกำหนดจำนวนของตัวกรอง (Filters) สำหรับใช้ในการแยกลักษณะเด่นหรือ Feature ของรูปภาพโดยปกติตัวกรอง 1 ตัวกรองจะสามารถทำการคัดแยกได้ 1 Feature โดยต้องมีการกำหนดขนาดของหน้าต่าง Sliding window หรือขนาดของเคอร์นอล (Kernel size) ที่ใช้ของตัวกรองนั้น ๆไว้ด้วย

ในขั้นตอนนี้จะมีการคำนวณคอนโวลูชันเชิงพื้นที่ (Spatial convolution) ระหว่างตัวกรองกับรูปภาพอินพุตโดยใช้การเทคนิค Sliding window หรือการสแกนรูปอินพุตเพื่อสร้างผังคุณลักษณะ (Feature map) โดยมีการกำหนดค่า Strides เพื่อให้ Sliding window ทำการเลื่อนตำแหน่งไปครั้งละกี่พิกเซลภาพ (Pixels) ในขั้นตอนของการสแกนภาพอินพุตตามแกนแนวตั้งและแนวแกนนอนให้ครอบคลุมตลอดทั้งรูปภาพอินพุต

2.) ขั้นตอนการตรวจจับ (Detector stage) ในขั้นตอนนี้จะทำหน้าที่รับข้อมูลที่ได้จากขั้นตอน Convolution stage มาแปลงให้อยู่ในรูปแบบที่ไม่เป็นเชิงเส้น (Nonlinear) โดยใช้ฟังก์ชันการกระตุ้น (Activation function) เช่น Rectified Linear Units (ReLU) โดยผลลัพธ์ที่ได้จากการทำคอนโวลูชันในแต่ละตำแหน่งจะผ่านการแปลงค่าด้วยฟังก์ชัน ReLU ที่เป็นการแปลงแบบไม่เป็นเชิงเส้น เพื่อความง่ายในการคำนวณและประสิทธิภาพของผลลัพธ์

3.) ขั้นตอนการพูลลิ่ง (Pooling stage) การคำนวณ Pooling เป็นการประมวลผลที่ทำให้เกิดการลดขนาดข้อมูลหรือการสุ่มต่ำ (Subsampling) ข้อมูลโดยทำให้ข้อมูลที่ได้ทางด้านเอาต์พุตมีขนาดเล็กลงโดยที่รายละเอียดของข้อมูลที่ป้อนเข้ามายังคงครบถ้วนเหมือนเดิม การพูลลิ่งแบบค่าสูงสุด (Max Pooling) เป็นตัวกรองแบบหนึ่งที่ค้นหาค่าสูงสุด (Maximum) ในบริเวณที่ตัวกรองทาบอยู่แล้วนำมาเป็นผลลัพธ์ โดยจะเตรียมตัวกรองในลักษณะเดียวกับขั้นตอนการทำ Feature extraction ของ CNN มาทาบบนข้อมูลแล้วเลือกค่าสูงสุดบนตัวกรองนั้นมาเป็นผลลัพธ์ใหม่ จากนั้นทำการเลื่อนตัวกรองไป

ตาม Stride ที่กำหนดไว้ การ Pooling มีประโยชน์ในเรื่องของการเพิ่มความไวในการคำนวณและยังช่วยในการแก้ปัญหาการเกิด Overfitting ในขั้นตอนของการเรียนรู้

วิธีการ Pooling ที่นิยมใช้ทั่วไปคือ วิธีการ MAX Pooling กับวิธีการ L2 โดยในสถาปัตยกรรมของ CNN โดยทั่วไปแล้ว การคำนวณ Pooling จะใช้ขนาดของหน้าต่างในการคำนวณเท่ากับ 2x2 และใช้ขนาดของ Stride เท่ากับ 2 โดยที่ไม่ต้องมีการเพิ่มพิกเซลภาพ (Padding) ที่บริเวณขอบของภาพ ซึ่งเป็นวิธีการคำนวณที่แตกต่างไปจากการคำนวณ Convolution โดยทั่วไป

2.6 สรุป

บทนี้ได้กล่าวถึง ทฤษฎีและหลักการที่เกี่ยวข้องกับงานวิจัย ซึ่งประกอบด้วย การสร้างภาพความละเอียดสูง การทบทวนวรรณกรรมที่เกี่ยวข้องในฐานข้อมูลต่าง ๆที่เกี่ยวข้องกับการสร้างภาพความละเอียดสูง จากนั้นจะได้กล่าวถึงการแปลงเวฟเล็ตโดยจะให้รายละเอียดของสมการการแปลงเวฟเล็ตแบบต่าง ๆที่นิยมใช้ในการประมวลสัญญาณภาพ และจะกล่าวถึงหลักการสำคัญของเครือข่ายประสาทแบบคอนโวลูชันซึ่งจัดได้ว่าเป็นการเรียนรู้เชิงลึกซึ่งผู้วิจัยจะนำมาใช้ในการสร้างภาพความละเอียดสูงในงานวิจัยนี้ วิธีการดำเนินงานวิจัยและผลการวิจัยจะได้นำเสนอในบทถัดไป

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 3
วิธีดำเนินการวิจัย

3.1 บทนำ

บทนี้จะนำเสนอวิธีดำเนินการวิจัยโดยจะกล่าวถึง การสร้างภาพความละเอียดสูงด้วยการเรียนรู้เชิงลึก (Single Image Super-Resolution Using Deep Learning) ซึ่งเรียกสั้น ๆว่า SR-CNN โดยจะให้รายละเอียดของการออกแบบโครงสร้างของเครือข่ายประสาทเทียมแบบคอนโวลูชันที่ใช้ในการสร้างภาพความละเอียดสูง และขั้นตอนในการเรียนรู้ของเครือข่ายประสาทเทียมแบบคอนโวลูชันที่ได้นำมาประยุกต์ใช้กับการสร้างภาพความละเอียดสูง จากนั้นผู้วิจัยจะได้กล่าวถึงการปรับปรุงคุณภาพของภาพความละเอียดสูงด้วยการใช้ตัวกรองแบบดิจิทัล และกล่าวถึงการประเมินคุณภาพของสัญญาณภาพดิจิทัลเป็นส่วนสุดท้าย

3.2 การสร้างภาพความละเอียดสูงด้วยการเรียนรู้เชิงลึก

(Single Image Super-Resolution Using Deep Learning)

Super-resolution คือการสร้างภาพความละเอียดสูง (High-resolution images) จากภาพความละเอียดต่ำ (Low-resolution images) โดยมีวัตถุประสงค์เพื่อให้ภาพดิจิทัลมีคุณภาพที่ดีมากขึ้นทั้งในด้านของความละเอียดของภาพและความสวยงาม ความชัดเจนของเฉดสีที่อยู่ในภาพ งานวิจัยนี้จะกล่าวถึงการสร้างภาพความละเอียดสูงจากภาพความเอียดต่ำเพียงภาพเดียวซึ่งเรียกว่า Single image super-resolution (SISR)

การสร้างภาพความละเอียดสูงด้วย SISR เป็นวิธีการที่มีความท้าทายมาก เนื่องจากโดยปกติทั่วไปแล้วองค์ประกอบความถี่สูงของสัญญาณภาพจะไม่สามารถกู้คืนได้จากภาพความละ เอียดต่ำจึงทำให้คุณภาพของสัญญาณภาพที่ได้จากการสร้างภาพความละเอียดสูงในวิธีดังกล่าวมีขีดจำกัด นอกจากนี้แล้ว วิธีการ SSIR จัดว่าเป็นการแก้ปัญหาที่มีคำตอบได้มากกว่าหนึ่งคำตอบ (Ill-posed problem) เพราะว่าภาพความละเอียดต่ำสามารถสร้างเป็นภาพความละเอียดสูงได้หลากหลายรูปแบบ

ในงานวิจัยนี้เป็นการประยุกต์ใช้เครือข่ายประสาทเทียมแบบคอนโวลูชันในการสร้างภาพความละเอียดสูง โดยจะเรียกว่าเป็นวิธีการ SR-CNN (Superresolution image reconstruction using convolutional neural networks) และเนื่องจาก CNN เป็นสถาปัตยกรรมของเครือข่ายประสาทเทียมที่สามารถทำการเรียนรู้ได้ในรูปแบบที่เรียกว่า การเรียนรู้เชิงลึก (Deep learning) ผู้ใช้งานจึงไม่จำเป็นต้องทำการสกัดลักษณะเด่น (Feature extraction) หรือคัดแยกลักษณะเด่นของภาพอินพุต (Image input) ก่อนป้อนให้กับ CNN แต่อย่างใด จุดนี้ช่วยทำให้การประมวลเพื่อการเรียนรู้ทำได้อย่างรวดเร็วและสะดวกต่อการใช้งานเป็นอย่างมาก และมีความแตกต่างไปจากการประยุกต์ใช้เครือข่ายประสาทเทียมในรูปแบบเดิมที่เรียกว่า การเรียนรู้ของเครื่องจักร (Machine learning) ที่ผู้ใช้งานจำเป็นจะต้องเป็นผู้ทำการสกัดลักษณะเด่นของสัญญาณภาพอินพุตก่อนทำการป้อนให้เครือข่ายประสาทเทียมทำการประมวลผลเพื่อการเรียนรู้

3.2.1 เครือข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution neural networks)

ในหัวข้อนี้จะนำเสนอ การออกแบบโครงสร้างของเครือข่ายประสาทเทียมแบบคอนโวลูชันเพื่อการใช้ในการสร้างภาพความละเอียดสูงโดยเฉพาะ SR-CNN จะถูกใช้ในการเรียนรู้เพื่อสร้างผังที่แสดงความสัมพันธ์ (Mapping) เพื่อการจับคู่ระหว่างภาพความละเอียดต่ำกับภาพความละเอียดสูง การ Mapping ดังกล่าว CNN สามารถทำได้อย่างมีประสิทธิภาพ เนื่องจากภาพความละเอียดต่ำกับภาพความละเอียดสูงมีส่วนขององค์ประกอบภาพที่คล้ายคลึงกัน (Similar content) และองค์ประกอบหลักที่แตกต่างกันเฉพาะรายละเอียดความถี่สูง (High-frequency) เท่านั้น

วิธีการสร้างภาพความละเอียดสูง (High-resolution images) จากภาพความละเอียดต่ำ (Low-resolution images) โดยในงานวิจัยนี้ได้ประยุกต์มาจากงานวิจัยที่นำเสนอโดย Kim et. al. 2016 [24] ซึ่งได้ทำการประยุกต์ใช้สถาปัตยกรรมเครือข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional neural networks) หรือเรียกสั้น ๆว่า CNN ซึ่งเป็นอัลกอริทึมที่ถูกออกแบบขึ้นมาเพื่อการประยุกต์ใช้งานกับรูปภาพดิจิทัล CNN สามารถดึงลักษณะเด่นของภาพนั้น ๆออกมาเพื่อการจำแนก (Classification) ว่าภาพดังกล่าวอยู่ในคลาสใด เป็นคลาสที่เรากำลังให้ความสนใจอยู่หรือไม่

ในขั้นตอนการเรียนรู้นั้นเครือข่าย SR-CNN จะใช้เทคนิคการเรียนรู้เรซิดิว (Residual) โดยเครือข่ายประสาทเทียม CNN จะถูกใช้ในการเรียนรู้เพื่อการประมาณค่าภาพเรซิดิว (Residual image) ซึ่งภาพเรซิดิวคือภาพที่เป็นผลต่างระหว่างภาพความละเอียดสูงต้นแบบหรือภาพอ้างอิง (High-resolution reference image) กับภาพความละเอียดต่ำ (Low-resolution image) ซึ่งได้ทำการปรับเพิ่มขนาดของภาพให้มีขนาดเท่ากับภาพอ้างอิงด้วยวิธีการไบคิวบิค (Bicubic interpolation) ดังนั้นแล้วภาพเรซิดิวจึงเป็นองค์ประกอบภาพที่ช่วยเพิ่มรายละเอียดที่ครบถ้วนในด้านความถี่สูงของภาพดิจิทัลนั่นเอง

ในอัลกอริทึมที่นำเสนอ เครือข่ายประสาทเทียม CNN จะทำการตรวจหาค่าเรซิดิวของภาพจากองค์ประกอบภาพสีในปริภูมิสี YCbCr ที่เรียกว่า Luminance ของภาพ โดยองค์ประกอบ Luminance ของภาพ (Y-Channel) จะเป็นส่วนที่แสดงรายละเอียดเกี่ยวข้องกับค่าความสว่าง (Brightness) ของแต่ละพิกเซลในภาพที่มีการรวมกันแบบเป็นเชิงเส้น (Linear combination) ของแต่ละองค์ประกอบค่าพิกเซลภาพทั้งสามองค์ประกอบของสีคือ สีแดง (Red) สีเขียว (Green) และ สีน้ำเงิน (Blue) ในทางตรงกันข้ามองค์ประกอบของภาพสีในส่วนที่เป็น Chrominance ของภาพคือ Cb และ Cr ซึ่งเป็นผลต่างเชิงเส้น (Difference linear combination) ของสีแดง (Red) สีเขียว (Green) และสีน้ำเงิน (Blue) ซึ่งเป็นค่าที่แทนข้อมูลผลต่างของสีหลัก (Color difference information) อย่างไรก็ตาม สายตามนุษย์จะมีความไวต่อการรับรู้ถึงการเปลี่ยนแปลงของค่าความสว่างมากกว่าการเปลี่ยนแปลงค่าของสีในภาพ ดังนั้นในขั้นตอนการเรียนรู้ของเครือข่ายประสาทเทียม SR-CNN จะถูกฝึกสอน (Training) โดยใช้ค่า Luminance หรือ Y-Channel เท่านั้น

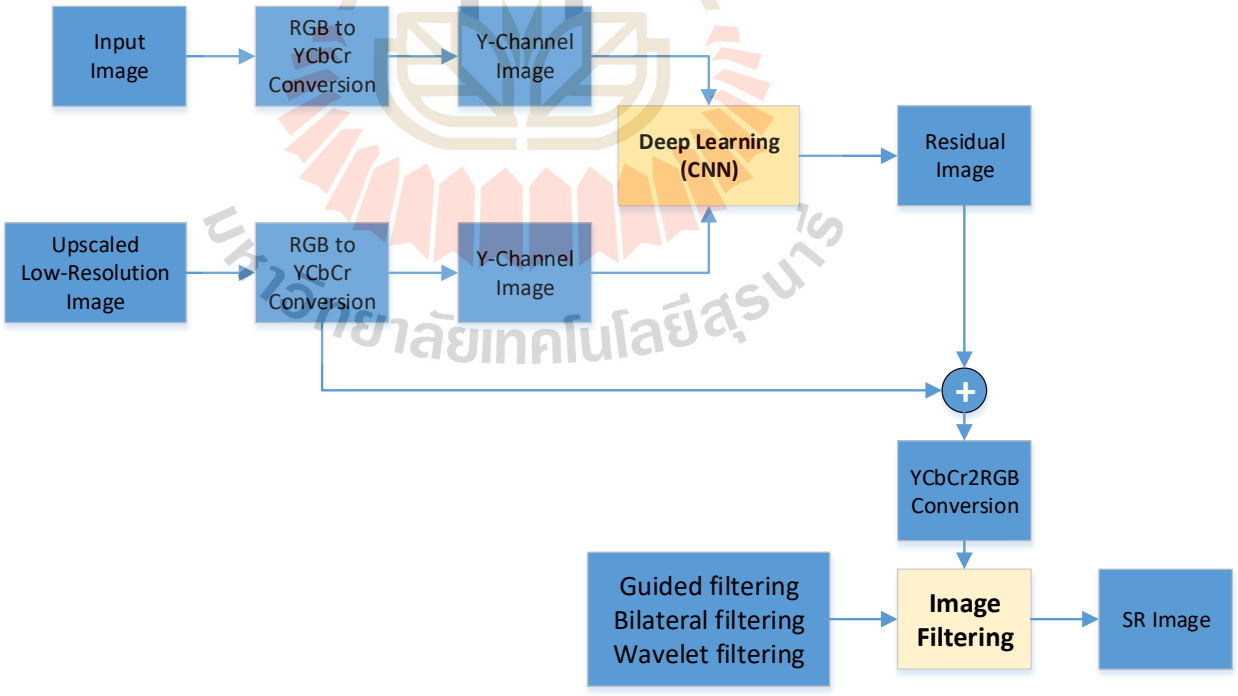
หลังจากที่เครือข่าย CNN ได้ทำการประมาณค่าภาพเรซิดิวแล้วก็จะสามารถสร้างภาพความละเอียดสูงได้ด้วยการรวมภาพเรซิดิวกับภาพความละเอียดต่ำที่ได้มีการปรับสเกลให้มีขนาดของภาพเพิ่มขึ้นจนเท่ากับขนาดของภาพความละเอียดสูงในระดับสเกลที่ต้องการ จากนั้นจะทำการแปลงข้อมูลภาพจากปริภูมิสีจาก YCbCr เป็นปริภูมิสี RGB และทำการบันทึกภาพความละเอียดสูงเพื่อใช้ในการประเมินผลคุณภาพของภาพดิจิทัลในขั้นตอนต่อไป แผนภาพแสดงประยุกต์ใช้การเรียนรู้ของเครือข่ายประสาทเทียมแบบ CNN แสดงในรูปที่ 3.1

ขั้นตอนการสร้างภาพความละเอียดสูงด้วยการเรียนรู้เชิงลึก ที่ประยุกต์ใช้การเรียนรู้ของเครือข่ายประสาทเทียมแบบ CNN แสดงในรูปที่ 3.1 สามารถสรุปได้ดังนี้

1. ทำการแปลงภาพจากปริภูมิสี (Color space) RGB เป็นปริภูมิสี YCbCr เพื่อทำการแยกองค์ประกอบภาพ Y-Channel ออกจากองค์ประกอบอื่น ๆ ซึ่ง Y-Channel นั้นจัดว่าช่องสัญญาณภาพที่สามารถให้รายละเอียดขององค์ประกอบภาพในด้านความถี่สูงมากกว่าองค์ประกอบอื่น ซึ่งจะมีผลอย่างมากที่จะสามารถช่วยให้ภาพดิจิทัลมีรายละเอียดของภาพที่ชัดเจนมากขึ้น โดยขั้นตอนการแปลงภาพสี RGB แสดงในแผนภาพดังรูปที่ 3.2

2. ในขั้นตอนการเรียนรู้ของเครือข่ายประสาทเทียมแบบ CNN ผู้วิจัยประยุกต์ใช้เครือข่ายดังกล่าวเพื่อการเรียนรู้เรซิดิว (Residual) ผลลัพธ์จากการเรียนรู้เรซิดิวจะทำให้เครือข่ายประสาทเทียม CNN สามารถทำการประมาณค่าภาพเรซิดิว (Residual image) โดยภาพเรซิดิวเป็นภาพที่เป็นผลต่างระหว่างภาพความละเอียดสูงต้นแบบหรือภาพอ้างอิง รายละเอียดการเรียนรู้ของเครือข่ายประสาทเทียมแบบ CNN แสดงในหัวข้อที่ 3.2.2 และขั้นตอนการประมาณค่าภาพเรซิดิวโดยเครือข่ายดังกล่าวแสดงในรูปที่ 3.2

3. เมื่อเครือข่าย CNN ทำการประมาณค่าภาพเรซิดิวแล้วก็จะสามารถสร้างภาพความละเอียดสูงได้ด้วยการรวมภาพเรซิดิวกับภาพความละเอียดต่ำที่ได้มีการปรับสเกลให้มีขนาดของภาพเพิ่มขึ้นจนมีขนาดที่เท่ากับขนาดของภาพความละเอียดสูงในระดับสเกลที่ต้องการ จากนั้นนำไปผ่านกระบวนการกรองสัญญาณภาพดิจิทัลเพื่อปรับปรุงคุณภาพของสัญญาณภาพให้ดีขึ้น (Image enhancement) ก่อนจะทำการแปลงย้อนกลับจากปริภูมิสี YCbCr เป็นปริภูมิสี RGB เพื่อการแสดงผลและจัดเก็บภาพต่อไป รายละเอียดการปรับปรุงคุณภาพของภาพความละเอียดสูงด้วยการใช้ตัวกรองแบบดิจิทัลแสดงในหัวข้อที่ 3.3 และขั้นตอนการสร้างภาพความละเอียดสูงด้วยการเรียนรู้เชิงลึกที่ประยุกต์ใช้การเรียนรู้ของเครือข่ายประสาทเทียมแบบ CNN แสดงในรูปที่ 3.6



รูปที่ 3.1 แผนภาพการสร้างภาพความละเอียดสูงด้วยการเรียนรู้เชิงลึก

3.2.2 การเรียนรู้ของเครือข่ายประสาทเทียมแบบคอนโวลูชัน

ขั้นตอนการเรียนรู้ของเครือข่ายประสาทเทียมแบบคอนโวลูชัน ที่ได้นำมาประยุกต์ใช้กับการสร้างภาพความละเอียดสูงด้วยการประยุกต์ใช้เครือข่ายประสาทเทียมแบบคอนโวลูชัน (SR-CNN) สามารถสรุปได้ดังนี้

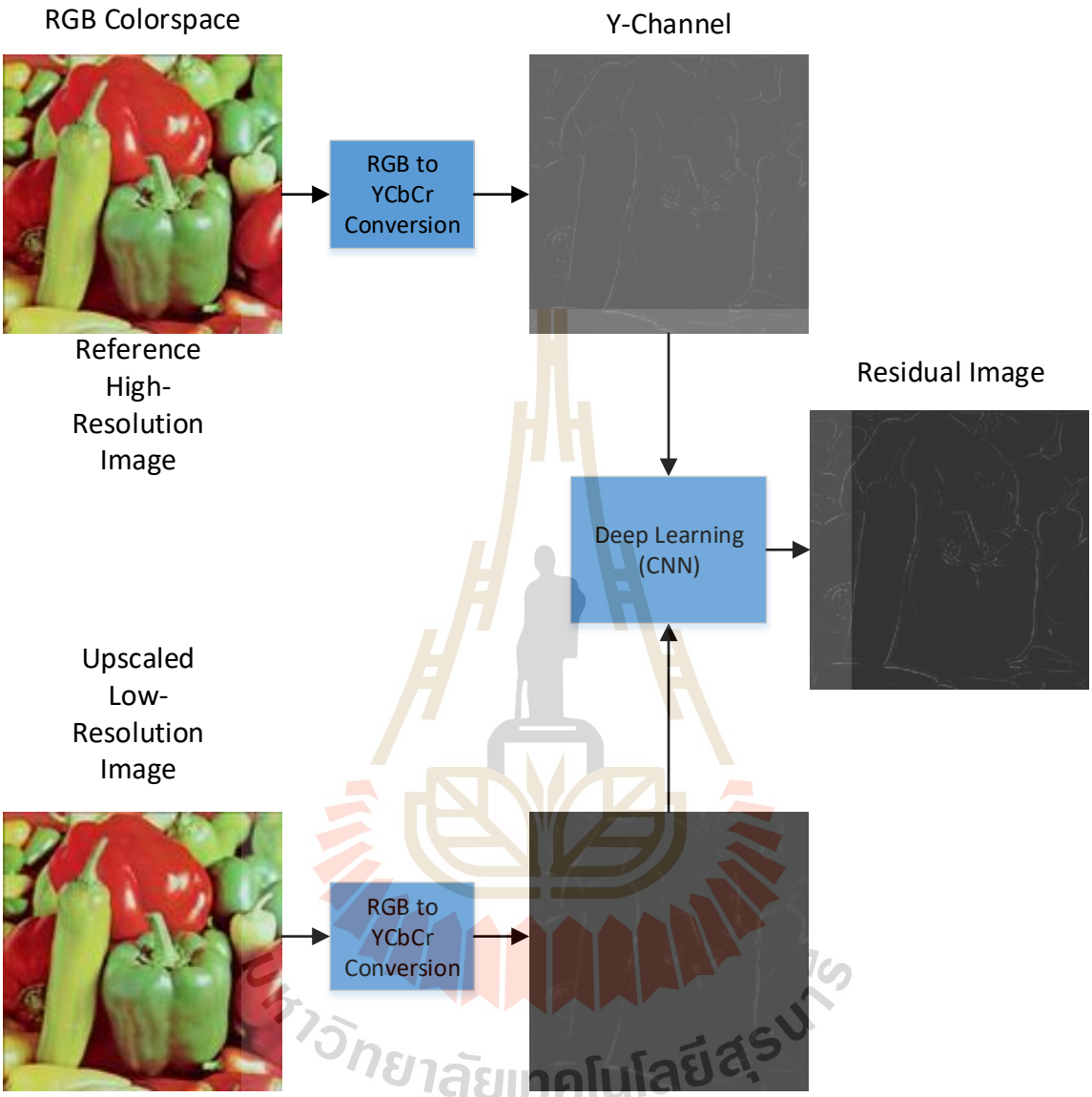
1. ทำการจัดเตรียมฐานข้อมูลสำหรับการเรียนรู้ให้กับเครือข่ายประสาทเทียมแบบคอนโวลูชัน ซึ่งเป็นข้อมูลภาพที่ไม่ผ่านการปรับแต่งหรือผ่านการประมวลผลภาพใด ๆ (Natural images) จำนวน 20,000 รูปจากฐานข้อมูล IAPR TC-12 Benchmark [25] ซึ่งเป็นฐานข้อมูลที่ไม่มีค่าใช้จ่ายในการใช้งาน นอกจากนี้ผู้วิจัยยังได้ทำการเพิ่มรูปภาพจากการเก็บรวบรวมเข้าไปในฐานข้อมูลอีกจำนวน 50 รูป

2. ทำการกำหนดจำนวนกลุ่มภาพย่อยขนาดเล็ก (Mini-batch) สำหรับการเรียนรู้ในแต่ละรอบของเครือข่ายประสาทเทียม และกำหนดขนาดพิกเซลของภาพเพื่อใช้ในการดึงหรือสกัดลักษณะเด่นข้อมูลภาพสำหรับใช้ในการเรียนรู้ ซึ่งโดยทั่วไปจะใช้กลุ่มภาพย่อยขนาดเล็กเท่ากับ 64 กลุ่ม และภาพย่อยดังกล่าวจะมีขนาด 41x41 พิกเซล และสามารถปรับเปลี่ยนการตั้งค่าได้ อย่างไรก็ตามในขั้นตอนการเรียนรู้แต่ละรอบนั้น จะมีกลุ่มของภาพย่อยเพียงกลุ่มเดียวเท่านั้นที่จะถูกสกัดลักษณะเด่นเพื่อการเรียนรู้ และทุกภาพย่อยจะใช้วิธีการสุ่มโดยสุ่มเลือกที่ตำแหน่งต่าง ๆของภาพ สามารถกำหนดค่าอัตราการขยายภาพแบบหลายอัตรา (Multiscale factor) ได้ในขั้นตอนนี้

3. กำหนดจำนวนชั้น (Layers) ของเครือข่ายประสาทเทียมแบบคอนโวลูชัน ผู้วิจัยได้กำหนดจำนวนชั้นของเครือข่ายเท่ากับ 41 และใช้การคำนวณคอนโวลูชันแบบ 2 มิติ (2-D convolution layer) โดยใช้กล่องเครื่องมือ Neural Network Toolbox™ ในโปรแกรม MATLAB เป็นตัวช่วยในการพัฒนาโปรแกรม รายละเอียดการตั้งค่ากำหนดในตารางที่ 3.1 ดังนี้

ตารางที่ 3.1 ตัวอย่างการกำหนดค่าในแต่ละชั้นของเครือข่ายประสาทเทียมแบบคอนโวลูชัน

ชั้นที่	ชื่อเรียก	รายละเอียด
1	ImageInputLayer	อินพุตเป็นสัญญาณภาพ 32x32x3 with zerocenter normalization
2	Convolutional	32 3x3 Convolutions with stride [1 1] and padding [1 1 1 1]
3	ReLu	Rectified linear unit (ReLU) layer
4	Convolution	32 3x3 Convolutions with stride [1 1] and padding [1 1 1 1]
5	ReLu	Rectified linear unit (ReLU) layer
6	Max Pooling	3x3 Max pooling with stride [2 2] and padding [0 0 0 0]
7	Fully Connected	64 Fully connected layer
8	ReLu	Rectified linear unit (ReLU) layer
9	Fully Connected	2 Fully connected layer
10	Softmax	Softmax
11	Classification Output	Crossentropyex



รูปที่ 3.2 แผนภาพขั้นตอนการคำนวณหาภาพเรซิดิว

4. ทำการกำหนดตัวเลือก (Options) ในการเรียนรู้ของเครือข่ายประสาทเทียมแบบคอนโวลูชัน โดยผู้วิจัยเลือกใช้วิธีการหาค่าที่เหมาะที่สุด (Optimization) ให้กับเครือข่ายประสาทเทียมแบบแบบคอนโวลูชันด้วยวิธี Stochastic gradient descent with momentum (SGDM) optimization โดยในตอนเริ่มต้นได้กำหนดอัตราการเรียนรู้ (Learning rate) ให้มีค่าเริ่มต้นเท่ากับ 0.1 และลดลงเป็นจำนวนเท่าของ 10 ในทุกๆรอบของการคำนวณ (Epochs) อย่างไรก็ตาม ขั้นตอนในการเรียนรู้ดังกล่าวถือว่าเป็นขั้นตอนที่ใช้เวลาในการประมวลผลมาก

การกำหนดค่าอัตราการเรียนรู้ที่สูงขึ้นก็จะช่วยลดเวลาในการคำนวณแต่ละรอบลง แต่อาจจะทำให้ค่าแกรเดียน์ของเครือข่ายประสาทเทียมแบบคอนโวลูชันไม่มีการลู่เข้า โดยอาจจะโตขึ้นเรื่อย ๆจนไม่สามารถควบคุมได้ทำให้การเรียนรู้ไม่ประสบผลสำเร็จ ดังนั้นจึงต้องเพิ่มเงื่อนไขควบคุมการเพิ่มขึ้นของค่าแกร์เดียน์ไม่ให้มากไปกว่าที่กำหนดไว้ด้วยวิธีการควบคุมค่าขีดเริ่มเปลี่ยนของแกร์เดียน์ (Gradient threshold) โดยได้กำหนดเงื่อนไขในการเรียนรู้ GradientThreshold Method โดยใช้ค่า L2-norm ในการควบคุมค่าแกร์เดียน์

5. เมื่อทำการตั้งค่าการเรียนรู้ให้เครือข่ายประสาทเทียมแบบคอนโวลูชันตามเงื่อนไขที่ได้กล่าวถึงข้างต้นเรียบร้อยแล้ว ก็สามารถทำการรันโปรแกรมเพื่อให้เครือข่าย CNN ดังกล่าวทำการเรียนรู้โดยใช้ฐานข้อมูลภาพจำนวน 20,000 ภาพ ในขั้นตอนดังกล่าวนี้ ผู้วิจัยได้ทำการประมวลผลด้วยคอมพิวเตอร์แบบตั้งโต๊ะ (Personal computer) ที่มีตัวประมวลผลเป็น Intel processor Core i5 มีหน่วยความจำ (DDR RAM) ขนาด 8 GB มีตัวเก็บข้อมูลแบบโซลิดสเตท (SSD Drive) ขนาด 256 GB ผู้วิจัยได้ใช้ระบบปฏิบัติการ Windows 10 พร้อมโปรแกรม MATLAB ที่ได้ติดตั้งกล่องเครื่องมือ CNN Toolbox ซึ่งมีองค์ประกอบที่สำคัญในการใช้งาน CNN หรือไรบรารี (Library) ของเครือข่ายประสาทเทียมแบบคอนโวลูชันที่ต้องใช้ Graphic processing unit (GPU) ในการประมวลผล

เนื่องจากข้อกำหนดในการใช้งานไรบรารีดังกล่าวกำหนดไว้ว่าตัวประมวลผลต้องเป็น CUDA-capable NVIDIA™ GPU with compute capability 3.0 or higher 3.0 ผู้วิจัยจึงได้เลือกใช้การ์ดจอภาพรุ่น VGA ASUS GTX 1050TI ซึ่งเป็นการ์ดที่มีจำนวนแกน (Core) ของ CUDA GPU เท่ากับ 768 แกน มีขนาดของหน่วยความจำบนการ์ดเท่ากับ 8 GB โดยได้ใช้เวลาในการประมวลผลเพื่อการเรียนรู้ของเครือข่ายประสาทเทียมแบบคอนโวลูชันประมาณ 7 ชั่วโมงซึ่งถือว่าใช้เวลาในการเรียนรู้นานมาก ผลลัพธ์ที่ได้จากการเรียนรู้จะเป็นเครือข่าย CNN ที่พร้อมสำหรับการสร้างภาพความละเอียดสูง (Pretrained SR-CRR) ที่ค่าสเกล 2 3 และ 4 อย่างไรก็ตาม เมื่อได้ทำการเรียนรู้เสร็จแล้ว เราสามารถนำผลลัพธ์จากการเรียนรู้ซึ่งเรียกว่าเครือข่าย SR-CNN ไปใช้งานได้ทันทีโดยไม่ต้องทำการเรียนรู้อีกนอกจากจะมีการเปลี่ยนแปลงฐานข้อมูลรูปภาพที่ใช้ในการเรียนรู้

ผลการเรียนรู้ของเครือข่าย SR-CNN ที่รอบของการเรียนรู้ (Epoch) เท่ากับ 100 แสดงในรูปที่ 3.3 และรูปที่ 3.4 ตามลำดับ

6. ทำการสร้างภาพความละเอียดสูงด้วยเทคนิคการเรียนรู้เชิงลึก (Deep learning) ซึ่งมีการประยุกต์ใช้เครือข่ายประสาทเทียมแบบคอนโวลูชัน ผู้วิจัยเลือกใช้การสร้างภาพความละเอียดสูงด้วยวิธีการ SISR (Single image super-resolution) โดยเทคนิคการเรียนรู้เชิงลึกจะใช้กลยุทธที่มีการเรียนรู้เรซิดิว (Residual) หมายความว่าเครือข่ายประสาทเทียมจะเรียนรู้เพื่อการประมาณค่าภาพเรซิดิว (Residual image) สำหรับใช้เพื่อการภาพความละเอียดสูง การเรียนรู้ของเครือข่ายประสาทเทียม SR-CNN ถือว่าเป็นกลยุทธหรือเทคนิคที่จะช่วยทำให้ภาพความละเอียดสูงมีความชัดเจนในรายละเอียดต่าง ๆของภาพมากยิ่งขึ้น โดยเฉพาะในด้านองค์ประกอบความถี่สูงของสัญญาณภาพดิจิทัล และเมื่อทำการสังเกตภาพเรซิดิวจะสามารถมองเห็นรายละเอียดของภาพในด้านความถี่สูงหรือขอบของภาพได้อย่างชัดเจน

7. ทำการทดสอบคุณภาพของภาพความละเอียดสูงด้วยเทคนิคการเรียนรู้เชิงลึกที่ได้สร้างขึ้น โดยการคำนวณค่า PNSR SSIM NIQE และ BRISQUE เพื่อตรวจสอบคุณภาพของสัญญาณภาพดิจิทัลเทียบกับภาพจากวิธีการปรับเพิ่มขนาดของภาพ (Resize) ด้วยการคำนวณไบคูบิกอินเตอร์โพเลชัน (Bicubic interpolation) ซึ่งเป็นวิธีการที่ใช้กันทั่วไป

```
Training on single GPU.
|=========================================================================================|
|  Epoch  |  Iteration  |  Time Elapsed  |  Mini-batch  |  Mini-batch  |  Base Learning  |
|         |             |   (hh:mm:ss)   |     RMSE     |     Loss     |      Rate       |
|=========================================================================================|
|       1 |           1 |       00:00:02 |         1.78 |          1.6 |          0.1000 |
|       1 |          50 |       00:00:23 |         1.27 |          0.8 |          0.1000 |
|       1 |         100 |       00:00:44 |         1.62 |          1.3 |          0.1000 |
|       1 |         150 |       00:01:05 |         2.36 |          2.8 |          0.1000 |
|       1 |         200 |       00:01:26 |         2.81 |          3.9 |          0.1000 |
|       1 |         250 |       00:01:47 |         1.03 |          0.5 |          0.1000 |
|       1 |         300 |       00:02:08 |         1.71 |          1.5 |          0.1000 |
|       1 |         350 |       00:02:29 |         2.66 |          3.5 |          0.1000 |
|       1 |         400 |       00:02:50 |         1.81 |          1.6 |          0.1000 |
```

รูปที่ 3.3 การเรียนรู้ของเครือข่าย SR-CNN ในช่วงเริ่มต้นการเรียนรู้

```
|      99 |       60850 |       07:06:05 |         1.62 |          1.3 |      1.0000e-10 |
|      99 |       60900 |       07:06:26 |         1.75 |          1.5 |      1.0000e-10 |
|      99 |       60950 |       07:06:47 |         1.64 |          1.3 |      1.0000e-10 |
|     100 |       61000 |       07:07:08 |         1.48 |          1.1 |      1.0000e-10 |
|     100 |       61050 |       07:07:29 |         0.90 |          0.4 |      1.0000e-10 |
|     100 |       61100 |       07:07:50 |         1.76 |          1.6 |      1.0000e-10 |
|     100 |       61150 |       07:08:11 |         2.09 |          2.2 |      1.0000e-10 |
|     100 |       61200 |       07:08:32 |         2.60 |          3.4 |      1.0000e-10 |
|     100 |       61250 |       07:08:53 |         2.11 |          2.2 |      1.0000e-10 |
|     100 |       61300 |       07:09:14 |         0.80 |          0.3 |      1.0000e-10 |
|     100 |       61350 |       07:09:34 |         2.34 |          2.7 |      1.0000e-10 |
|     100 |       61400 |       07:09:55 |         1.47 |          1.1 |      1.0000e-10 |
|     100 |       61450 |       07:10:16 |         4.31 |          9.3 |      1.0000e-10 |
|     100 |       61500 |       07:10:37 |         1.43 |          1.0 |      1.0000e-10 |
|     100 |       61550 |       07:10:58 |         1.88 |          1.8 |      1.0000e-10 |
|     100 |       61600 |       07:11:19 |         0.72 |          0.3 |      1.0000e-10 |
|=========================================================================================|
>>
```

รูปที่ 3.4 การเรียนรู้ของเครือข่าย SR-CNN ที่รอบการเรียนรู้เท่ากับ 100 และสิ้นสุดการเรียนรู้

มหาวิทยาลัยเทคโนโลยีสุรนารี

3.3 การปรับปรุงคุณภาพของภาพความละเอียดสูงด้วยการใช้ตัวกรองแบบดิจิทัล (Image enhancement)

เทคนิคที่นิยมใช้ในการปรับเพิ่มคุณภาพของภาพดิจิทัล (Image enhancement) คือ การประยุกต์ใช้ตัวกรองสำหรับสัญญาณภาพดิจิทัล (Digital image filtering) ในการออกแบบตัวกรองดิจิทัลเพื่อใช้ในการกรองสัญญาณภาพดิจิทัลโดยปกติทั่วไปแล้ว สามารถเลือกทำการออกแบบตัวกรองเพื่อการประมวลผลสัญญาณภาพได้ใน 2 โดเมนหลัก ๆ คือ ตัวกรองเพื่อการประมวลผลภาพในโดเมนเชิงพื้นที่ (Spatial domain filtering) และตัวกรองเพื่อการประมวลผลภาพในโดเมนความถี่ (Frequency domain filtering) งานวิจัยนี้จะประยุกต์ใช้ตัวกรองเพื่อการประมวลผลภาพในโดเมนเชิงพื้นที่ซึ่งเป็นการประมวลผลกับค่าพิกเซลในภาพโดยตรงทำให้สามารถทำการประมวลผลได้อย่างรวดเร็ว และมีความซับซ้อนของอัลกอริทึมการประมวลผลน้อยกว่าการประมวลผลภาพในโดเมนความถี่

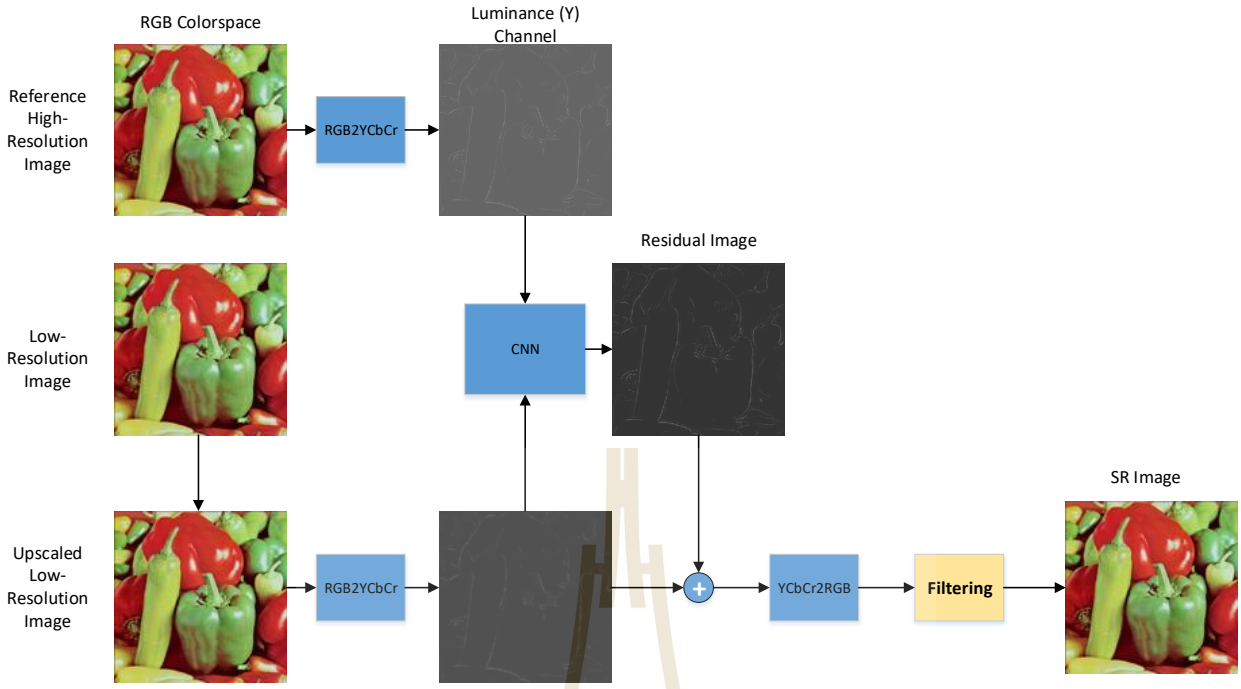
ในการประมวลผลภาพในโดเมนเชิงพื้นที่ โดยทั่วไปแล้วจะเป็นการคำนวณที่ต้องอาศัยพิกเซลข้างเคียง (Neighborhood processing) ในการประมวลผลเพื่อคำนวณค่าและให้คำตอบเป็นค่าพิกเซลของศูนย์กลางพิกเซลจากการกำหนดขนาดขึ้นในแต่ละกรอบของพิกเซลนั้น ๆ ซึ่งกรอบในการคำนวณจะเรียกว่าหน้าต่าง (Windows) หรือแมส (Mask) หรืออาจจะเรียกว่าเป็นเคอร์นอลของตัวกรอง (Kernal filter) โดยปกติทั่วไปการประมวลผลภาพในโดเมนเชิงพื้นที่ รูปของเคอร์นอลจะถูกกำหนดให้มีลักษณะเป็นหน้าต่างที่มีจุดศูนย์กลาง เช่น ขนาด 3x3 5x5 7x7 9x9 เป็นต้น ในรูปที่ 3.5 แสดงภาพเคอร์นอลของตัวกรองขนาด 3x3 และ 5x5 และในรูปที่ 3.6 แสดงแผนภาพการประยุกต์ใช้ตัวกรองดิจิทัลในการปรับปรุงคุณภาพของภาพความละเอียดสูง

สำหรับตัวกรองที่ผู้วิจัยได้นำมาประยุกต์ใช้เพื่อการปรับปรุงคุณภาพของภาพความละเอียดสูง ได้แก่ ตัวกรองดิจิทัลแบบ Guided filtering ตัวกรองดิจิทัลแบบ Bilateral filtering และตัวกรองดิจิทัลในโดเมนของการแปลงเวฟเล็ต โดยมีรายละเอียดดังต่อไปนี้

$1/9 \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ $1/25 \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$

มหาวิทยาลัยเทคโนโลยีสุรนารี

รูปที่ 3.5 รูปแบบเคอร์นอลของตัวกรองขนาด 3x3 และ 5x5



รูปที่ 3.6 กระบวนการปรับปรุงคุณภาพของภาพความละเอียดสูงด้วยการใช้ตัวกรองแบบดิจิทัล

3.3.1 ตัวกรองดิจิทัลแบบ Guided filtering

ตัวกรองดิจิทัลแบบไกด์ (Guided filtering) [26] เป็นตัวกรองดิจิทัลที่ถูกออกแบบมาเพื่อใช้ในการกรองภาพดิจิทัลเพื่อการปรับให้ภาพเนียนหรือมีความราบเรียบมากยิ่งขึ้นโดยที่ยังคงไว้ซึ่งความคมบริเวณที่เป็นขอบของภาพเอาไว้ (Edge-preserving smoothing) ทำให้ภาพดังกล่าวยังคงมีความชัดเจนดังเดิม การประมวลผลของตัวกรองดังกล่าวต้องอาศัยภาพอ้างอิงที่นำไปสู่ผลลัพธ์ที่ต้องการเพื่อใช้เป็นแนวทางในการประมวลผลของตัวกรอง หรือเรียกสั้น ๆว่า ภาพไกด์ (Guidance image) ซึ่งภาพดังกล่าวจะถูกใช้ในการคำนวณค่าความแปรปรวนในขั้นตอนของการประมวลผล ดังนั้นภาพไกด์จึงมีอิทธิพลต่อการประมวลผลของตัวกรองเป็นอย่างมาก ภาพไกด์ที่ใช้ในการประมวลผลอาจเป็นภาพในอีกรูปแบบหนึ่งของภาพที่กำลังประมวลผล หรืออาจเป็นภาพอื่นที่เราต้องการให้ภาพเอาต์พุตจากการประมวลผลมีคุณลักษณะตามภาพไกด์นั้น

3.3.2 ตัวกรองดิจิทัลแบบ Bilateral filtering

ตัวกรองดิจิทัลแบบไบเลเทอรอล (Bilateral filtering) [27] เป็นตัวกรองดิจิทัลที่ถูกออกแบบมาเพื่อใช้ในการกรองภาพดิจิทัลเพื่อการปรับให้ส่วนพื้นผิว (Texture) ที่ทำให้เห็นลวดลายในภาพมีความเรียบมากขึ้นโดยที่ยังคงบริเวณที่เป็นขอบของภาพเอาไว้ (Edge-preserving) ทำให้ภาพดังกล่าวมีความชัดเจนและน่าสนใจมากขึ้น ตัวกรองดังกล่าวถูกออกแบบขึ้นมาเพื่อการใช้งานกับภาพระดับเทา (Grayscale image) และภาพสี (Color image) ทั้งในปริภูมิสี RGB และปริภูมิสีอื่น ๆ

3.3.3 ตัวกรองดิจิทัลในโดเมนของการแปลงเวฟเล็ต (Wavelet filtering)

ตัวกรองดิจิทัลในโดเมนของการแปลงเวฟเล็ต (Wavelet transform) [28] เป็นตัวกรองดิจิทัลที่ใช้ในการกรองภาพดิจิทัลเพื่อทำการกำจัดสัญญาณรบกวนที่เกิดขึ้นในภาพโดยที่ยังคงบริเวณที่เป็นขอบของภาพเอาไว้ (Edge-preserving) ทำให้ภาพดังกล่าวมีความชัดเจนและน่าสนใจมากขึ้น อัลกอริทึมการกำจัดสัญญาณรบกวนในโดเมนของการแปลงเวฟเล็ตแบบดีสครีต (Discrete wavelet transform) ซึ่งต่อไปจะเรียกว่า การแปลง DWT ผู้วิจัยจะทำการแปลงสัญญาณภาพด้วยการแปลงเวฟเล็ตเพื่อแยกองค์ประกอบของความถี่ในสัญญาณภาพออกเป็นแบนด์ย่อยความถี่สูง (Highpass subband) และแบนด์ย่อยความถี่ต่ำ (Approximation subband) โดยมีระดับชั้น (Level) ของการแปลง DWT เป็นตัวกำหนดจำนวนภาพแบนด์ย่อย (Subband image) ซึ่งแสดงถึงระดับความละเอียดในการวิเคราะห์สัญญาณ

ในการกำจัดสัญญาณรบกวน อัลกอริทึมจะทำกำหนดค่าขีดเริ่มเปลี่ยน (Threshold) ในการกำจัดสัญญาณรบกวนด้วยวิธีการที่นำเสนอโดย Donoho *et.al.* [28] ซึ่งใช้วิธีการประมาณจากค่าความแปรปรวนของสัญญาณรบกวน วิธีการนี้จึงเป็นการตั้งค่าขีดเริ่มเปลี่ยนที่เหมาะสมกับระดับความรุนแรงของสัญญาณรบกวน จากนั้นทำการแปลงย้อนกลับเวฟเล็ตหรือ (Inverse discrete wavelet transform, IDWT) เพื่อทำการสร้างกลับสัญญาณภาพดิจิทัลภายหลังการกำจัดสัญญาณรบกวนแล้ว

มหาวิทยาลัยเทคโนโลยีสุรนารี

3.4 การประเมินคุณภาพของสัญญาณภาพดิจิทัล

ในการประมวลผลสัญญาณภาพดิจิทัลที่ให้เอาต์พุตของการประมวลผลเป็นภาพดิจิทัลนั้น โดยปกติทั่วไปจะต้องมีการประเมินคุณภาพของสัญญาณภาพดิจิทัลเอาต์พุตเสมอ ทั้งนี้เพื่อเป็นการตรวจสอบความสามารถของอัลกอริทึมในการประมวลผลและตรวจสอบคุณภาพของภาพเอาต์พุตที่ได้จากการประมวลผลนั้น ๆว่ามีความเหมาะสมสำหรับการใช้งานตามวัตถุประสงค์หรือไม่ ในการสร้างภาพความละเอียดสูง ผู้วิจัยได้ใช้มาตรวัดคุณภาพ (Quality metrics) หรือตัวประเมินคุณภาพของสัญญาณที่แตกต่างกัน 2 รูปแบบคือ แบบแรกใช้การประเมินคุณภาพที่ใช้การคำนวณค่าความแตกต่างของพิกเซล (Pixels) ในภาพระหว่างภาพต้นฉบับกับภาพที่ผ่านการประมวลผลสัญญาณ ซึ่งได้แก่ การประเมินคุณภาพด้วยค่า PSNR และ SSIM และแบบที่สองจะใช้วิธีการประเมินคุณภาพของสัญญาณภาพด้วยการใช้แบบจำลองทางคณิตศาสตร์ของสัญญาณภาพซึ่งไม่จำเป็นต้องใช้ภาพต้นฉบับ

แต่อย่างใด ซึ่งถือว่าเป็นวิธีการใหม่ในการประเมินคุณภาพของสัญญาณภาพดิจิทัล ซึ่งได้แก่ การประเมินด้วยค่า NIQE และ BRISQUE รายละเอียดของวิธีการประเมินคุณภาพในวิธีการดังกล่าวข้างต้น สามารถสรุปได้ดังนี้

PSNR (Peak signal to noise ratio)

PSNR (Peak signal to noise ratio) เป็นค่าที่นิยมใช้ในการเปรียบเทียบคุณภาพของภาพดิจิทัลระหว่างภาพต้นฉบับกับภาพที่ถูกทำการประมวลผลด้วยวิธีการประมวลผลภาพแบบต่าง ๆเพื่อใช้ในการประเมินคุณภาพเพื่อบ่งชี้ว่าภาพดิจิทัลที่ผ่านการประมวลผลมีคุณภาพเพิ่มขึ้นหรือลดลงอย่างไร และโดยทั่วไปถ้าพบว่าภาพดิจิทัลมีค่า PSNR มากกว่า 30 dB จะถือว่าภาพนั้นมีคุณภาพที่ดี อย่างไรก็ตามภาพดิจิทัลที่มีค่า PSNR สูง ๆเมื่อมองดูด้วยสายตาอาจจะพบว่ามันมีคุณภาพและความละเอียดที่ต่ำกว่าภาพที่มี PSNR ต่ำ ๆก็ได้ เพราะการคำนวณค่า PSNR ไม่ได้คำนึงถึงความสามารถในการมองเห็นของสายตามนุษย์ (Visual model) ที่มีความไวต่อการสังเกตและเปลี่ยนแปลงของแต่ละองค์ประกอบในภาพแต่อย่างใด

SSIM (Structural Similarity Index)

SSIM (Structural Similarity Index) [29] เป็นการคำนวณค่าที่ใช้ในการบ่งชี้คุณภาพของสัญญาณภาพดิจิทัลโดยใช้การคำนวณค่าที่มีผลกระทบต่อการมองเห็นภาพ เช่น ความสว่าง (Luminance) ความแตกต่างระหว่างสีขาวกับสีดำ (Contrast) และเค้าโครงของภาพ (Structure) เพื่อใช้ในการเปรียบเทียบความเหมือนหรือความคล้ายกันของภาพดิจิทัลต้นฉบับกับภาพดิจิทัลที่ผ่านการประมวลผลภาพแบบต่าง ๆเพื่อประเมินว่าภาพที่ผ่านการประมวลผลดังกล่าวมีค่าคุณภาพเพิ่มขึ้นหรือลดลงอย่างไรโดยจะมีช่วงของค่าการประเมินผลอยู่ระหว่าง 0 - 1 โดยในการเปรียบเทียบกับภาพอ้างอิงนั้นภาพที่มีค่า SSIM ใกล้เคียง 1 มากกว่าจะมีคุณภาพที่ดีกว่า

NIQE (Naturalness Image Quality Evaluator)

NIQE (Naturalness Image Quality Evaluator) [30] เป็นการคำนวณเพื่อประเมินคุณภาพความละเอียดของสัญญาณภาพดิจิทัลแบบที่ไม่ต้องใช้การอ้างอิงหรือเปรียบเทียบกับภาพต้นฉบับ ซึ่งทำให้สะดวกต่อการใช้งานมากกว่าสองแบบแรกที่ได้กล่าวถึงคือ PSNR และ SSIM ในการคำนวณ NIQE จะใช้การเปรียบเทียบค่าการลดทอนภาพดิจิทัลกับแบบจำลองภาพที่ใช้การกระจายตัวแบบ Multidimensional Gaussian distributions ค่า NIQE ที่คำนวณได้จะถูกกำหนดให้อยู่ในช่วง 0.0 – 1.0 ผลการเปรียบเทียบภาพดิจิทัลที่ให้ค่า NIQE ที่ต่ำกว่าแสดงว่าภาพดิจิทัลนั้นมีคุณภาพหรือมีความละเอียดของสัญญาณภาพดิจิทัลที่ดีกว่า

BRISQUE (Blind/Referenceless Image Spatial Quality Evaluator)

BRISQUE (Blind/Referenceless Image Spatial Quality Evaluator) [31] เป็นการประเมินคุณภาพความละเอียดของสัญญาณภาพดิจิทัลโดยที่ไม่ต้องใช้สัญญาณภาพอ้างอิง โดย BRISQUE วิธีการประเมิน ที่ถูกพัฒนามาจาก NIQE โดยในการคำนวณหาค่า BRISQUE นั้นจะใช้การเปรียบเทียบภาพดิจิทัลดังกล่าวกับโมเดลรูปแบบเริ่มต้น (Default model) ที่มีการคำนวณจากภาพดิจิทัลแบบธรรมชาติ (Natural scenes) ซึ่งเป็นภาพที่ได้จากการถ่ายภาพโดยกล้องดิจิทัลและไม่มีการประมวลผลภาพใด ๆเพิ่มเติม จากนั้นนำมาทำการเปรียบเทียบที่การลดทอนแบบเดียวกัน ค่า BRISQUE ที่คำนวณได้จะถูกกำหนดให้อยู่ในช่วง 0 – 100 โดยในการเปรียบเทียบนั้นภาพที่มีค่า BRISQUE ต่ำกว่าจะมีคุณภาพที่ดีกว่า

3.5 สรุป

บทนี้ได้นำเสนอวิธีดำเนินการวิจัย โดยได้กล่าวถึง การสร้างภาพความละเอียดสูงด้วยการเรียนรู้เชิงลึกซึ่งได้ให้รายละเอียดของการออกแบบโครงสร้างของเครือข่ายประสาทเทียมแบบคอนโวลูชันเพื่อการใช้ในการสร้างภาพความละเอียดสูง (SR-CNN) รวมถึงขั้นตอนในการเรียนรู้ของเครือข่ายประสาทเทียมแบบคอนโวลูชันที่ได้นำมาประยุกต์ใช้กับการสร้างภาพความละเอียดสูง จากนั้นได้กล่าวถึงการปรับปรุงคุณภาพของภาพความละเอียดสูงด้วยการใช้ตัวกรองแบบดิจิทัล และการประเมินคุณภาพของสัญญาณภาพดิจิทัลในการสร้างภาพความละเอียดสูง ทั้งนี้เพื่อเป็นการตรวจสอบความสามารถของอัลกอริทึมในการประมวลผลและตรวจสอบคุณภาพของภาพเอาต์พุตที่ได้จากกระบวนการนั้น ๆว่ามีคุณภาพดีหรือไม่อย่างไร การทดสอบอัลกอรึทึมโดยใช้ฐานข้อมูลภาพมาตรฐานและผลการทดลองต่าง ๆจะได้นำเสนอในบทถัดไป

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 4
ผลการทดลอง

4.1 บทนำ

บทนี้จะนำเสนอ ผลการทดสอบการสร้างภาพความละเอียดสูงด้วยเทคนิค Deep learning โดยจะนำเสนอขั้นตอนและวิธีการในการทำงานของอัลกอริทึมที่พัฒนาขึ้นทั้งในส่วนของการสร้างภาพความละเอียดสูงและการปรับเพิ่มคุณภาพของสัญญาณด้วยตัวกรองแบบดิจิทัล จากนั้นผู้วิจัยจะนำเสนอ ผลการทดสอบการทำงานของอัลกอริทึมการสร้างภาพความละเอียดสูงด้วยเทคนิค Deep learning โดยชุดข้อมูลสัญญาณภาพดิจิทัลจากฐานข้อมูล IAPR TC-12 Benchmark [25] จากนั้นจะนำเสนอผลการวิเคราะห์และสรุปผลการทดลองเป็นส่วนสุดท้าย

4.2 ผลการทดสอบการสร้างภาพความละเอียดสูงด้วย Deep learning (Single Image Super-Resolution Using Deep Learning)

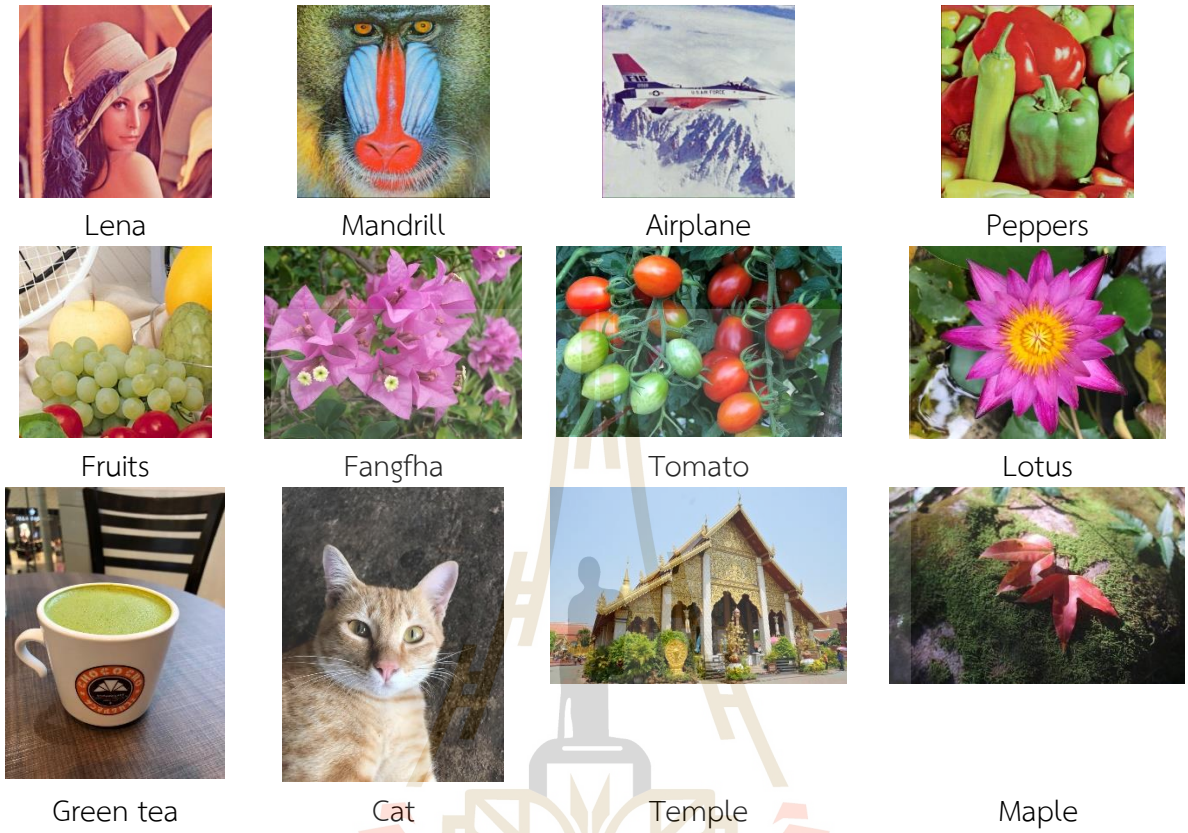
ในบทก่อนหน้านี้ผู้วิจัยได้นำเสนอ การออกแบบโครงสร้างของเครือข่ายประสาทเทียมแบบคอนโวลูชันเพื่อการใช้ในการสร้างภาพความละเอียดสูงโดยเฉพาะ ซึ่งในหัวข้อนี้ผู้วิจัยจะนำเสนอ ผลการทดสอบการสร้างภาพความละเอียดสูงด้วย Deep learning SR-CNN จะถูกใช้ในการเรียนรู้เพื่อสร้างผังที่แสดงความสัมพันธ์ (Mapping) เพื่อการจับคู่ระหว่างภาพความละเอียดต่ำกับภาพความละเอียดสูง การ Mapping ดังกล่าว CNN สามารถทำได้อย่างมีประสิทธิภาพ เนื่องจากภาพความละเอียดต่ำกับภาพความละเอียดสูงมีส่วนขององค์ประกอบภาพที่คล้ายคลึงกัน (Similar content) และองค์ประกอบหลักที่แตกต่างกันเฉพาะรายละเอียดความถี่สูง (High-frequency) เท่านั้น

เครือข่ายประสาทเทียมที่มีสถาปัตยกรรมของการเรียนรู้เชิงลึก SR-CNN จะใช้กลยุทธการเรียนรู้เรซิดิวเพื่อทำการประมาณค่าภาพเรซิดิว (Residual image) ที่เป็นผลต่างระหว่างภาพความละเอียดสูงต้นแบบหรืออาจเรียกว่าภาพอ้างอิง (High-resolution reference image) กับภาพความละเอียดต่ำ (Low-resolution image) ซึ่งได้ทำการปรับเพิ่มขนาดของภาพให้มีขนาดเท่ากับภาพอ้างอิงด้วยวิธีการไบคิวบิค (Bicubic interpolation) ดังนั้นแล้วภาพเรซิดิวจึงเป็นภาพที่จะช่วยเพิ่มรายละเอียดในด้านความถี่สูงของภาพให้ครบถ้วน

ในการประเมินคุณภาพของสัญญาณภาพดิจิทัลจากอัลกอริทึมการสร้างภาพความละเอียดสูง ผู้วิจัยได้ใช้ตัวประเมินคุณภาพของสัญญาณที่แตกต่างกัน 2 วิธีคือ วิธีแรกใช้วิธีการประเมินคุณภาพที่ใช้การคำนวณค่าความแตกต่างของค่าพิกเซล (Pixel) ในภาพระหว่างภาพต้นฉบับหรือภาพอ้างอิงกับภาพที่ผ่านการประมวลผลสัญญาณ ซึ่งได้แก่ การประเมินคุณภาพด้วยค่า PSNR และ SSIM และแบบที่สองจะใช้วิธีการประเมินคุณภาพของสัญญาณภาพด้วยการใช้แบบจำลองทางคณิตศาสตร์ของสัญญาณภาพซึ่งไม่จำเป็นต้องใช้ภาพต้นฉบับหรือภาพอ้างอิงแต่อย่างใด (No-reference image quality assessment) ซึ่งถือว่าเป็นวิธีการใหม่ในการประเมินคุณภาพของสัญญาณภาพดิจิทัล ซึ่งได้แก่ การประเมินด้วยค่า NIQE และ BRISQUE

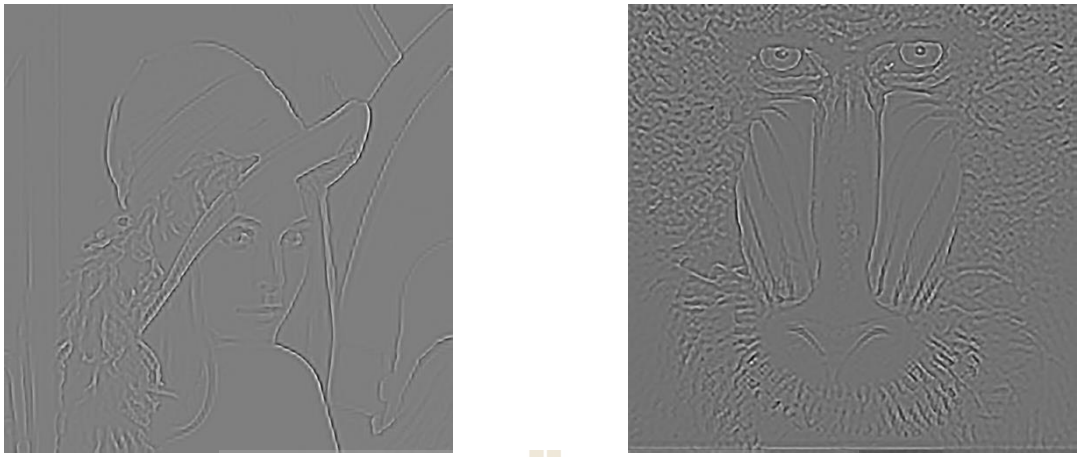
ในวิธีการแรกจะเปรียบเทียบระหว่างวิธีการที่นำเสนอที่เรียกว่า วิธีการสร้างภาพความละเอียดสูงด้วยวิธีการเรียนรู้เชิงลึก (Deep learning; SR-CNN) กับการขยายภาพให้มีความละเอียดสูงด้วยวิธีการไบคิวบิก

อินเตอร์โพเลชัน (Bicubic interpolation; Bicubic) โดยใช้ภาพมาตรฐานจากฐานข้อมูลและภาพจากกล้องดิจิทัล ตัวอย่างของภาพดิจิทัลมาตรฐานที่ใช้ในการทดสอบเบื้องต้น แสดงในรูปที่ 4.1



รูปที่ 4.1 ตัวอย่างภาพดิจิทัลที่ใช้ในการทดสอบ

รูปที่ 4.2 ถึงรูปที่ 4.7 แสดงภาพเรซิดิว (Residual image) ที่ได้จากขั้นตอนการเรียนรู้ของเครือข่ายประสาทเทียมแบบคอนโวลูชัน (CNN) ซึ่งถือว่าเป็นกลยุทธหรือเทคนิคที่จะช่วยทำให้ภาพความละเอียดสูงมีความชัดเจนในรายละเอียดต่าง ๆของภาพมากยิ่งขึ้น ในขั้นตอนดังกล่าวเครือข่ายประสาทเทียมจะเรียนรู้เพื่อการประมาณค่าภาพเรซิดิว (Residual image) ซึ่งเป็นผลต่างระหว่างภาพความละเอียดสูงกับภาพความละเอียดต่ำ (Low-resolution image) ที่ได้ทำการปรับเพิ่มขนาดของภาพให้มีขนาดเท่ากันก่อนการคำนวณหาผลต่างดังกล่าว ภาพเรซิดิวจึงเป็นภาพที่จะช่วยให้รายละเอียดที่ครบถ้วนโดยเฉพาะในด้านองค์ประกอบความถี่สูงของสัญญาณภาพดิจิทัล โดยจากการสังเกตภาพเรซิดิวในรูปที่ 4.2 ถึงรูปที่ 4.7 เราจะสามารถมองเห็นรายละเอียดของภาพในด้านความถี่สูงหรือขอบของภาพได้อย่างชัดเจน



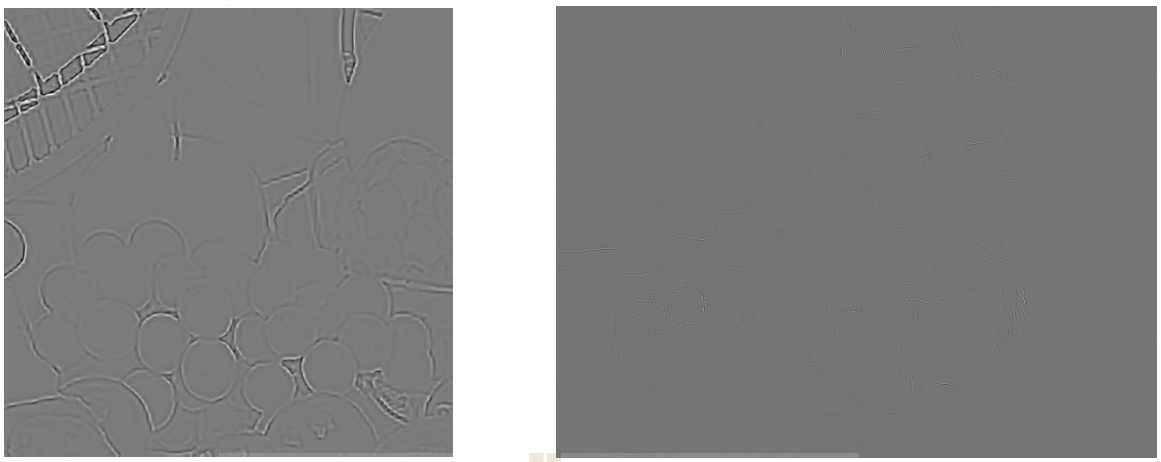
(ก) (ข)

รูปที่ 4.2 ภาพเรซิดิวของ (ก) Lena (ข) Mandrill

(ก) (ข)

รูปที่ 4.3 ภาพเรซิดิวของ (ก) Airplane (ข) Peppers

มหาวิทยาลัยเทคโนโลยีสุรนารี



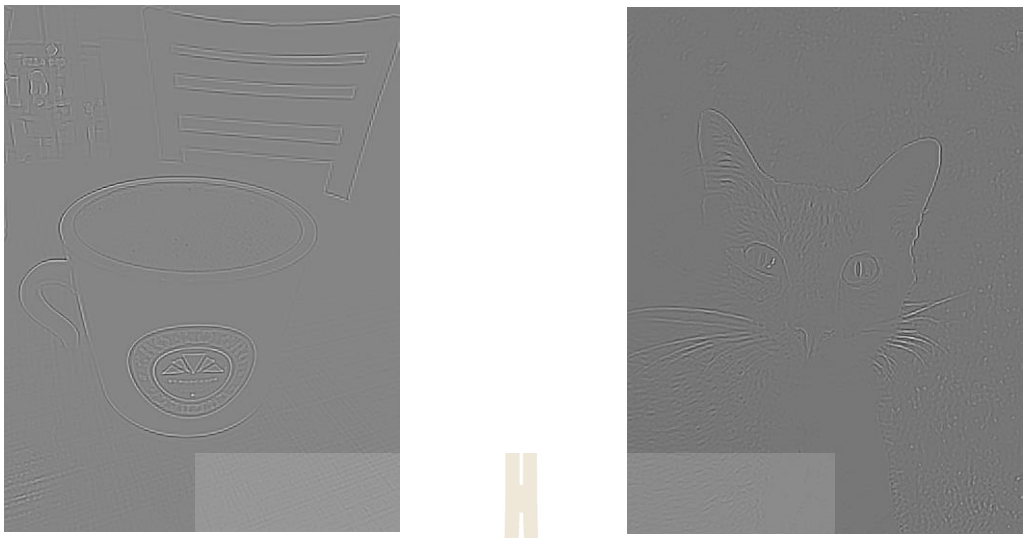
(ก) (ข)

รูปที่ 4.4 ภาพเรซิดิวของ (ก) Fruits (ข) FangFha

(ก) (ข)

รูปที่ 4.5 ภาพเรซิดิวของ (ก) Tomato (ข) Lotus

มหาวิทยาลัยเทคโนโลยีสุรนารี



(ก) (ข)

รูปที่ 4.6 ภาพเรซิดิวของ (ก) Green tea (ข) Cat

(ก) (ข)

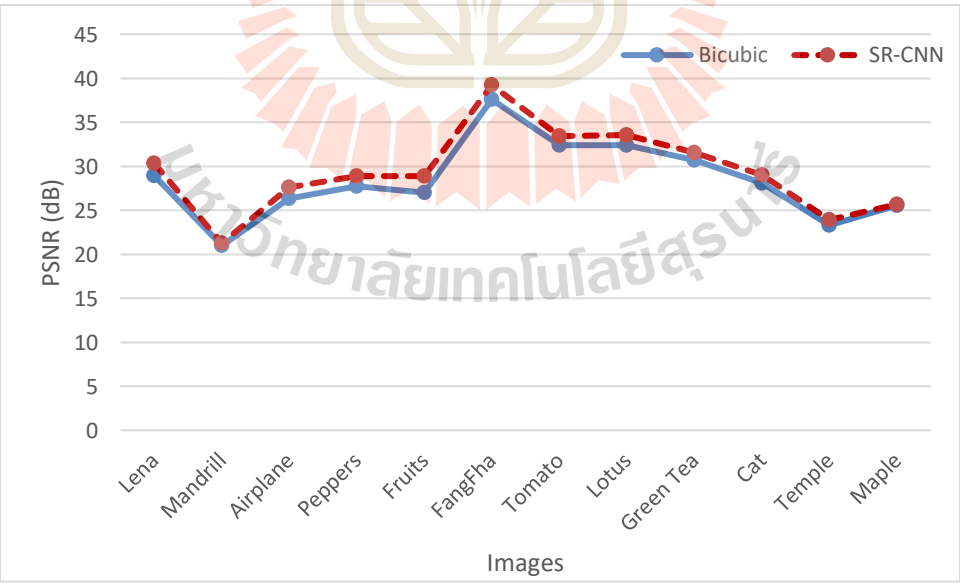
รูปที่ 4.7 ภาพเรซิดิวของ (ก) Temple (ข) Maple

มหาวิทยาลัยเทคโนโลยีสุรนารี

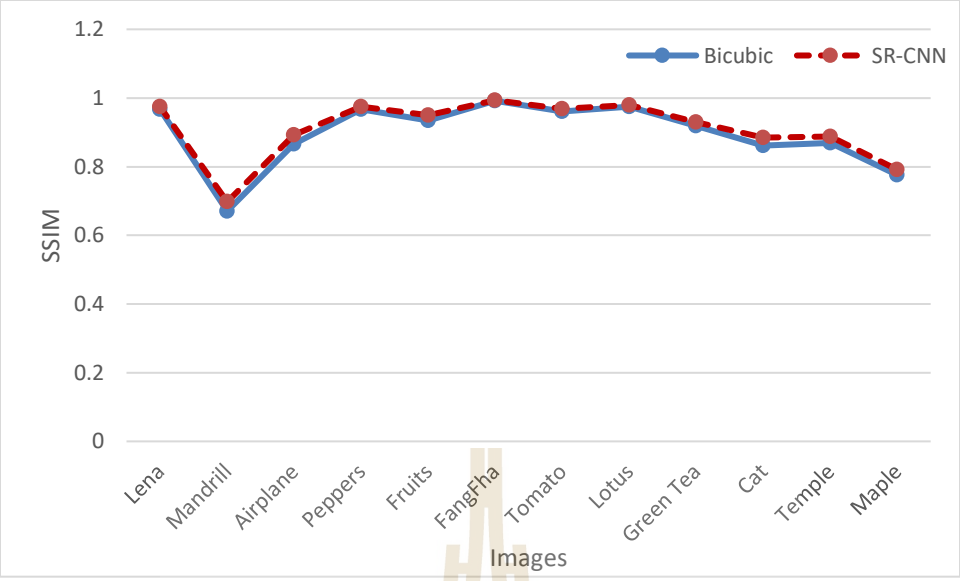
ในตารางที่ 4.1 และตารางที่ 4.2 แสดงผลการเปรียบเทียบคุณภาพสัญญาณภาพดิจิทัลระหว่างการขยายภาพ (Resize image) ด้วยวิธีการไบคิวบิกอินเตอร์โพเลชัน (Bicubic interpolation; วิธี Bicubic) ซึ่งเป็นวิธีการที่นิยมใช้กันทั่วไปในการขยายภาพให้มีขนาดของภาพใหญ่ขึ้น และวิธีการสร้างภาพความละเอียดสูงด้วยการเรียนรู้เชิงลึก (Deep learning; วิธี SR-CNN) โดยผู้วิจัยได้ใช้ภาพมาตรฐานและภาพจากกล้องดิจิทัลในการทดสอบ

ตารางที่ 4.1 ผลการเปรียบเทียบคุณภาพระหว่างวิธี Bicubic และวิธี SR-CNN

ภาพดิจิทัล	PSNR (dB)		SSIM	
	Bicubic	SR-CNN	Bicubic	SR-CNN
Lena	28.9119	30.3557	0.9677	0.9753
Mandrill	20.9856	21.2423	0.6706	0.6976
Airplane	26.3157	27.6159	0.8663	0.8925
Peppers	27.7558	28.8753	0.9675	0.9743
Fruits	27.0162	28.8897	0.9343	0.9502
FangFha	37.5879	39.2344	0.9919	0.9934
Tomato	32.4042	33.4445	0.9612	0.9689
Lotus	32.4045	33.5696	0.9751	0.9792
Green Tea	30.6874	31.5326	0.9185	0.9295
Cat	28.1132	28.9775	0.8621	0.8845
Temple	23.3106	23.9266	0.8692	0.8874



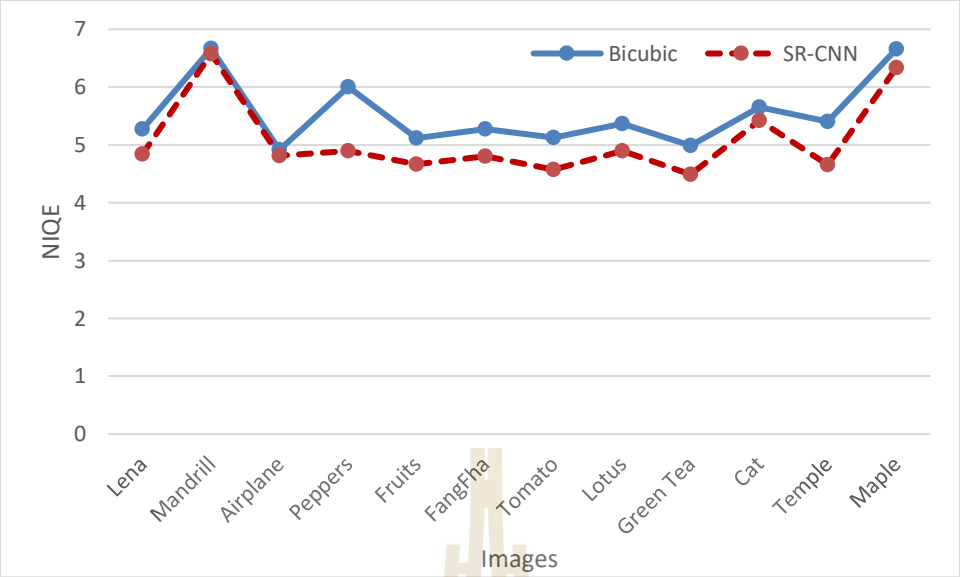
รูปที่ 4.8 เปรียบเทียบค่า PSNR ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN



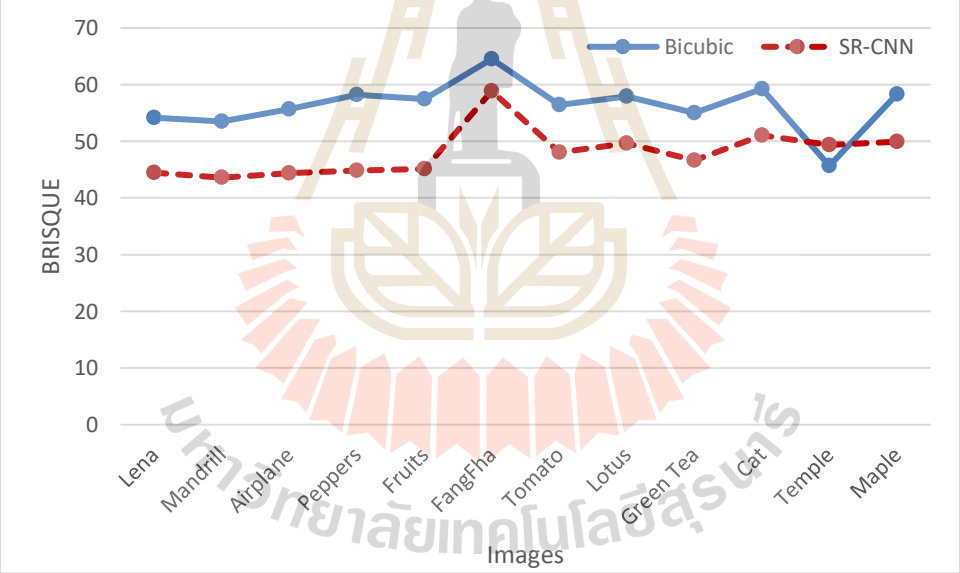
รูปที่ 4.9 เปรียบเทียบค่า SSIM ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN

ตารางที่ 4.2 ผลการเปรียบเทียบคุณภาพระหว่างวิธี Bicubic และวิธี SR-CNN

ภาพดิจิทัล	NIQE		BRISQUE	
	Bicubic	SR-CNN	Bicubic	SR-CNN
Lena	5.2719	4.8457	54.1639	44.5154
Mandrill	6.6661	6.5789	53.5418	43.5839
Airplane	4.9152	4.8125	55.6840	44.4182
Peppers	6.0039	4.8947	58.1775	44.8080
Fruits	5.1185	4.6629	57.4455	45.1409
FangFha	5.2729	4.8080	64.5222	58.8600
Tomato	5.1304	4.5767	56.4118	48.0434
Lotus	5.3664	4.8981	57.9553	49.6078
Green Tea	4.9882	4.4900	54.9921	46.5905
Cat	5.6533	5.4241	59.2148	51.0308
Temple	5.4030	4.6558	45.6743	49.4037



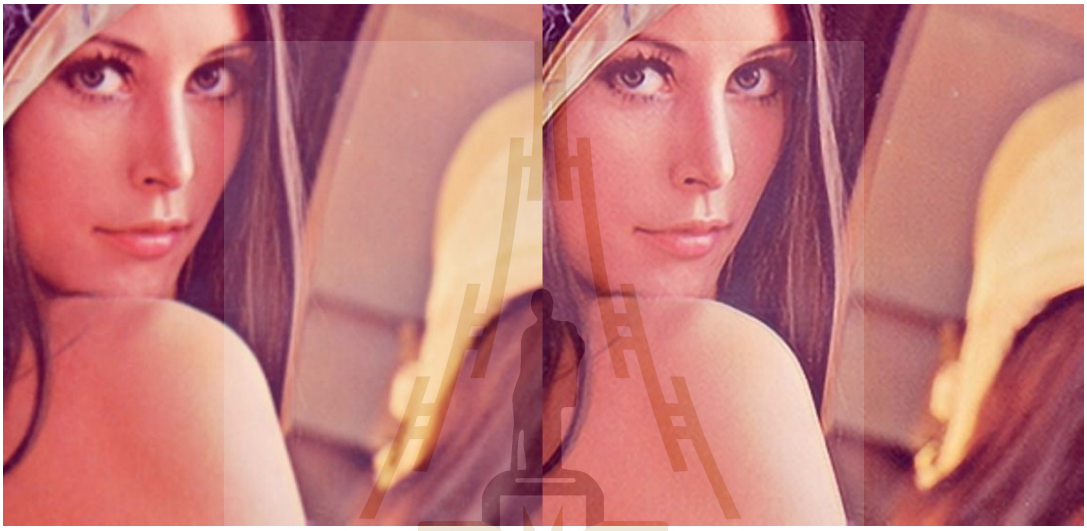
รูปที่ 4.10 เปรียบเทียบค่า NIQE ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN



รูปที่ 4.11 เปรียบเทียบค่า BRISQUE ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN

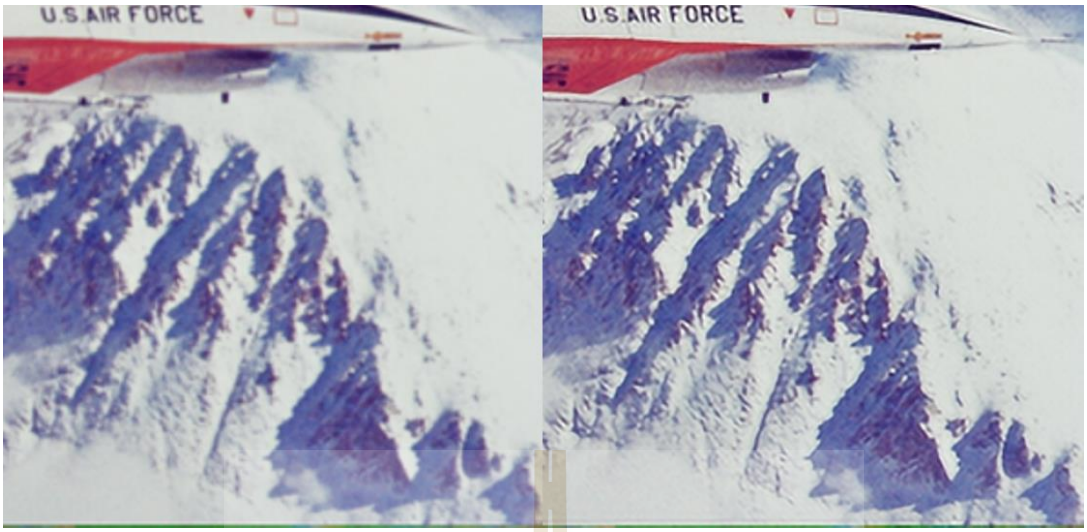
4.3 ผลการเปรียบเทียบคุณภาพสัญญาณภาพดิจิทัลด้วยการตัดและซูมภาพ

ผลการเปรียบเทียบคุณภาพสัญญาณภาพดิจิทัลระหว่างวิธีการขยายภาพให้มีความละเอียดสูง โดยภาพทางซ้ายมือเป็นภาพความละเอียดสูงจากวิธีการไบคิวบิกอินเตอร์โพเลชัน (Bicubic interpolation; Bicubic) และภาพทางขวามือเป็นการสร้างภาพความละเอียดสูงจากวิธีการเรียนรู้เชิงลึก (SR-CNN) ผู้วิจัยได้ทำการทดสอบโดยใช้ภาพมาตรฐานจากฐานข้อมูล และภาพจากกล้องดิจิทัล โดยได้ทำการตัดและซูมภาพที่ได้จากการทดลอง ทั้งนี้เพื่อแสดงให้เห็นถึงความแตกต่างที่สังเกตได้อย่างชัดเจนว่า วิธีการที่นำเสนอ SR-CNN ให้รายละเอียดของภาพที่ชัดเจนและมีขอบภาพที่คมชัดยิ่งขึ้น ผลการทดสอบดังกล่าวได้ใช้ภาพจำนวน 10 ภาพ ดังปรากฏรายชื่อภาพในตารางที่ 1 และแสดงภาพผลการเปรียบเทียบในรูปที่ 4.12 ถึงรูปที่ 4.23 ตามลำดับ



รูปที่ 4.12 เปรียบเทียบภาพ Lena จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา)

รูปที่ 4.13 เปรียบเทียบภาพ Mandrill จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา)



รูปที่ 4.14 เปรียบเทียบภาพ Airplane จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา)

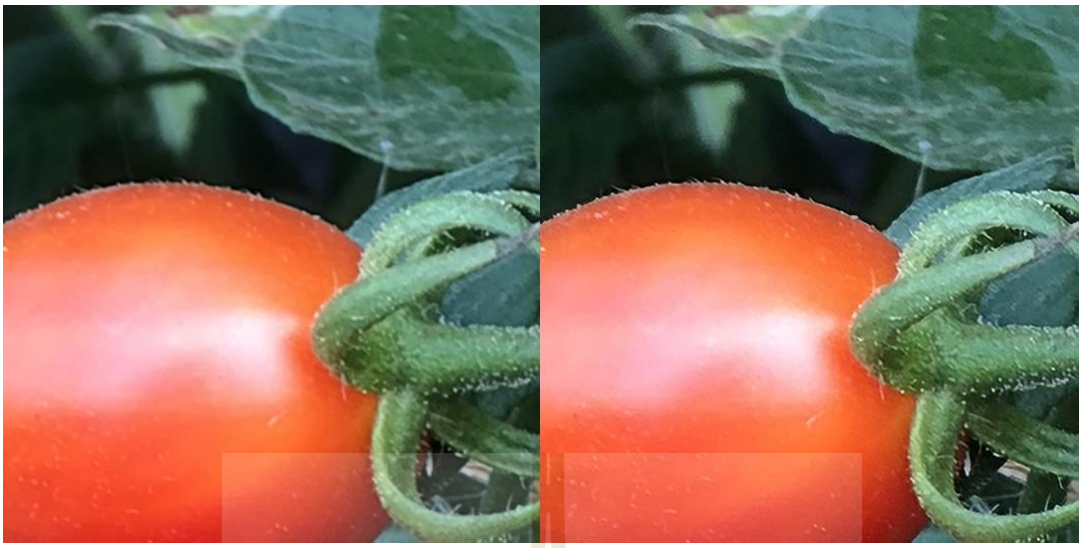
รูปที่ 4.15 เปรียบเทียบภาพ Peppers จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา)

มหาวิทยาลัยเทคโนโลยีสุรนารี

รูปที่ 4.16 เปรียบเทียบภาพ Fruits จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา)

มหาวิทยาลัยเทคโนโลยีสุรนารี

รูปที่ 4.17 เปรียบเทียบภาพ FangFha จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา)



รูปที่ 4.18 เปรียบเทียบภาพ Tomato จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา)

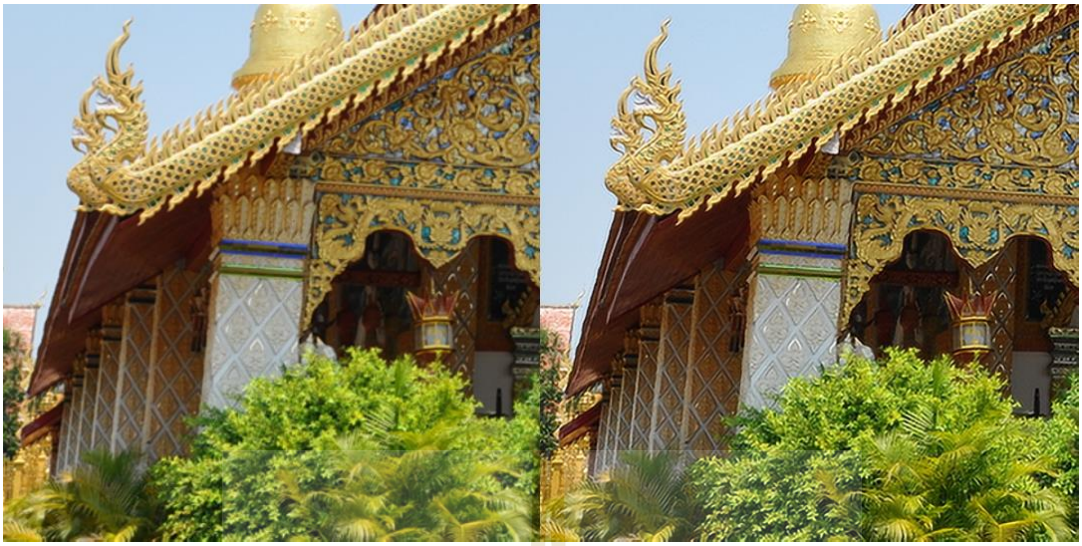
รูปที่ 4.19 เปรียบเทียบภาพ Lotus จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา)

มหาวิทยาลัยเทคโนโลยีสุรนารี

รูปที่ 4.20 เปรียบเทียบภาพ Green tea จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา)

รูปที่ 4.21 เปรียบเทียบภาพ Cat จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา)

มหาวิทยาลัยเทคโนโลยีสุรนารี



รูปที่ 4.22 เปรียบเทียบภาพ Temple จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา)

รูปที่ 4.23 เปรียบเทียบภาพ Maple จากวิธี Bicubic (ภาพซ้าย) และ SR-CNN (ภาพขวา)

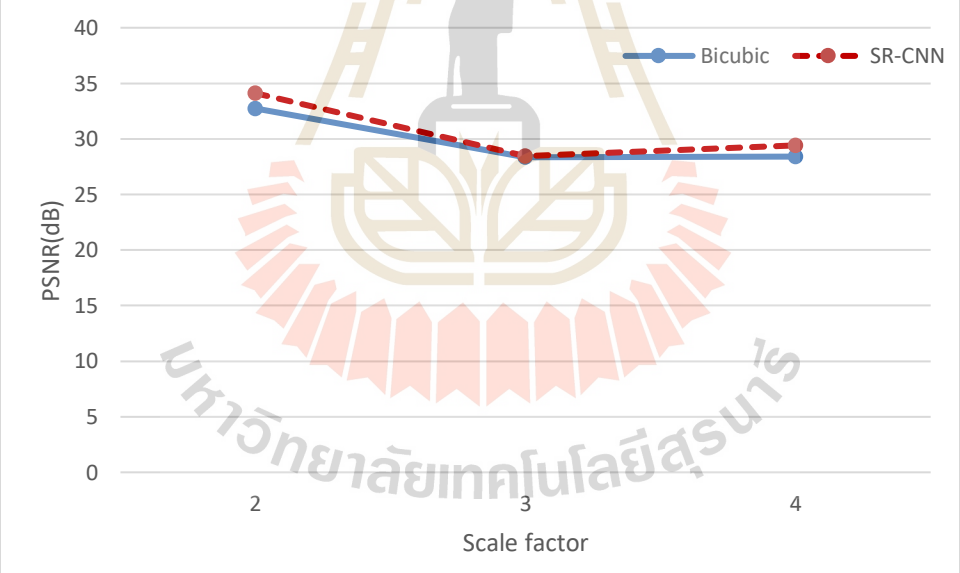
มหาวิทยาลัยเทคโนโลยีสุรนารี

4.4 ผลการเปรียบเทียบคุณภาพสัญญาณภาพดิจิทัลที่ค่าอัตราการขยายภาพต่าง ๆ

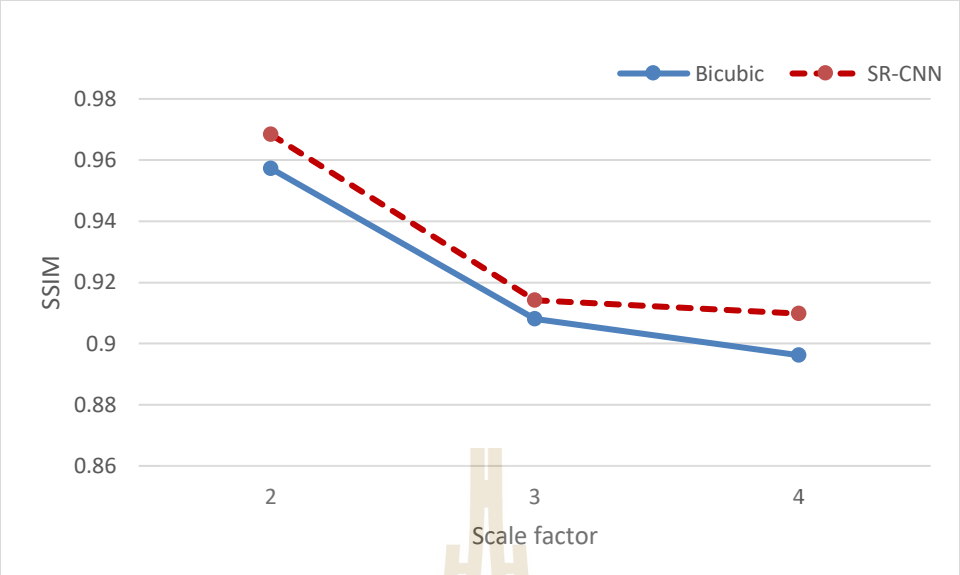
ในการทดสอบเบื้องต้น ผู้วิจัยได้ทำการทดสอบเทคนิคการสร้างภาพความละเอียดสูง SR-CNN เทียบกับวิธีการ Bicubic interpolation โดยใช้ภาพดิจิทัลในการทดสอบจำนวน 12 ภาพ โดยทดสอบที่ค่าอัตราการขยายภาพ (Scale factor) ที่มีค่ามากขึ้นเพื่อทดสอบความสามารถของอัลกอริทึมที่ได้นำเสนอ ผลการทดสอบที่ค่าอัตราการขยายภาพ 2 3 และ 4 โดยเปรียบเทียบการประเมินคุณภาพของสัญญาณภาพดิจิทัลโดยใช้ค่า PSNR SSIM NIQE และ BRISQUE แสดงในตารางที่ 4.3 และตารางที่ 4.4 และรูปที่ 4.24 ถึงรูปที่ 4.27 ตามลำดับ โดยพบว่าวิธีการที่นำเสนอ (SR-CNN) ให้ผลการทดสอบที่ดีกว่าในทุกกรณี

ตารางที่ 4.3 การเปรียบเทียบค่าเฉลี่ยจากการคำนวณ PSNR และ SSIM

Scale factor	PSNR (dB)		SSIM	
	Bicubic	SR-CNN	Bicubic	SR-CNN
2	32.7356	34.1073	0.9573	0.9684
3	28.3483	28.4525	0.9081	0.9142
4	28.4025	29.4190	0.8962	0.9098



รูปที่ 4.24 เปรียบเทียบค่า PSNR ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN

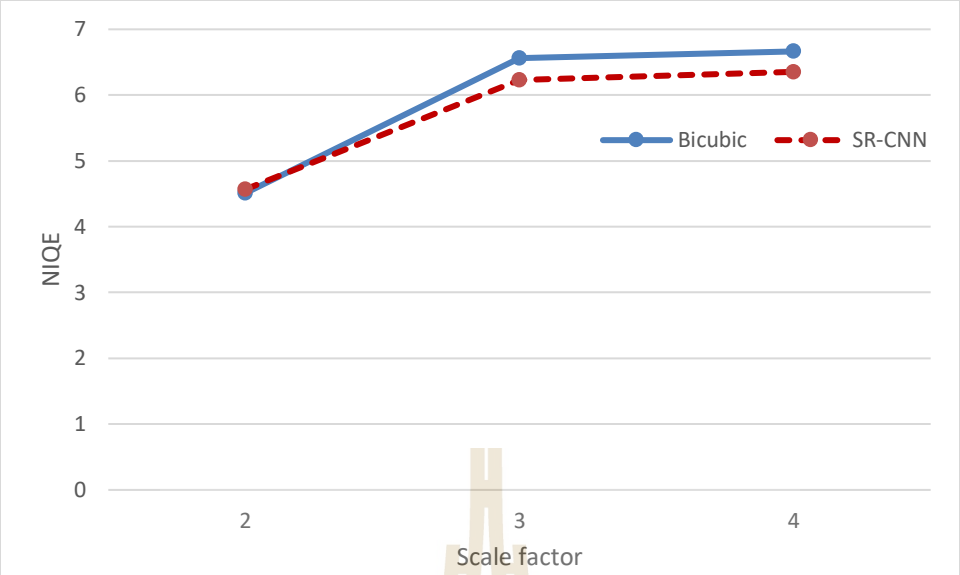


รูปที่ 4.25 เปรียบเทียบค่า SSIM ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN

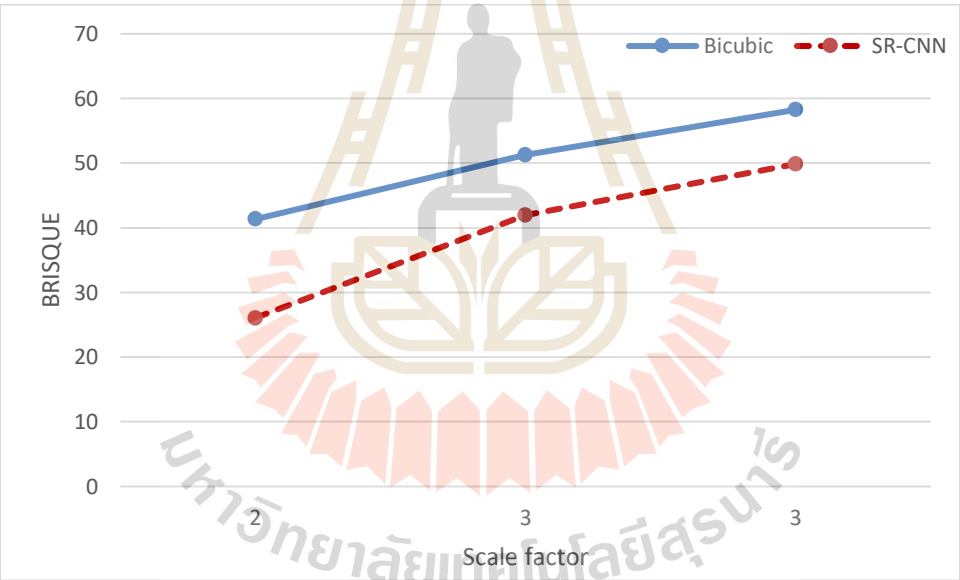
ตารางที่ 4.4 การเปรียบเทียบค่าเฉลี่ยจากการคำนวณ NIQE และ BRISQUE

Scale factor	NIQE		BRISQUE	
	Bicubic	SR-CNN	Bicubic	SR-CNN
2	4.5103	4.5680	41.3376	26.0702
3	6.5625	6.2313	51.2595	41.9395
4	6.6639	6.3536	58.2552	49.8657

มหาวิทยาลัยเทคโนโลยีสุรนารี



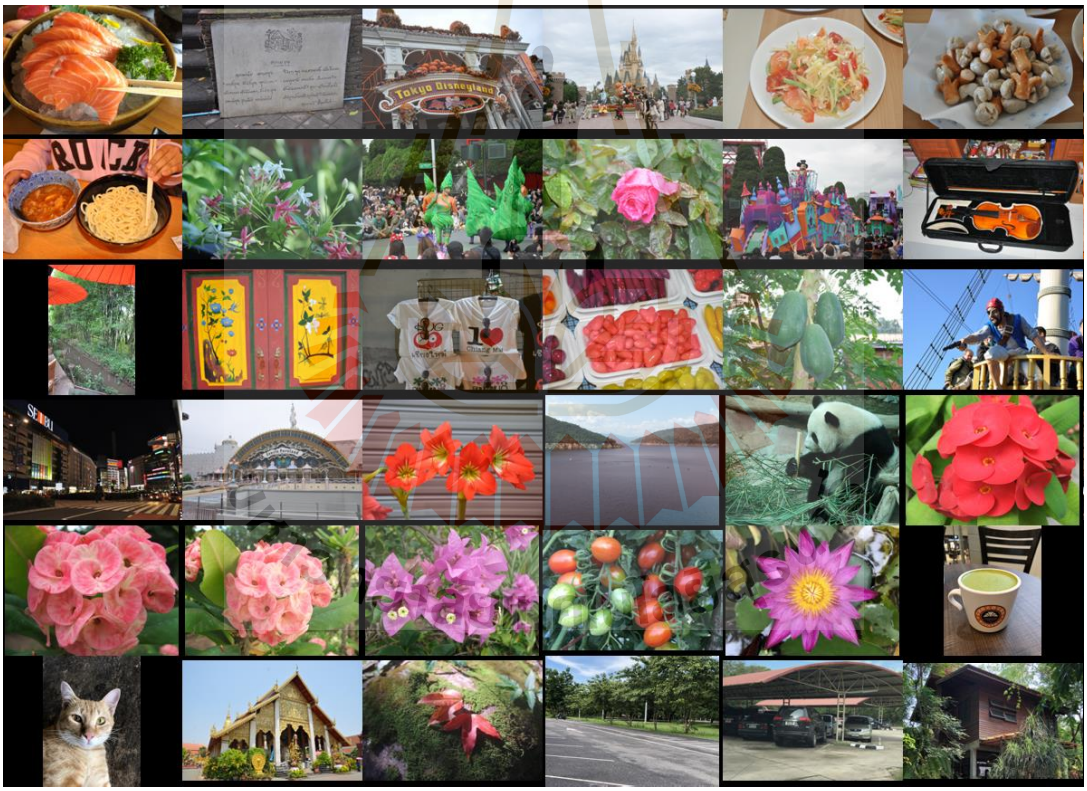
รูปที่ 4.26 เปรียบเทียบค่า NIQE ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN



รูปที่ 4.27 เปรียบเทียบค่า BRISQUE ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN

ในการทดสอบอัลกอริทึมในลำดับต่อมา ผู้วิจัยได้ทำการทดสอบเทคนิคการสร้างภาพความละเอียดสูง SR-CNN เทียบกับวิธีการ Bicubic interpolation โดยใช้ภาพดิจิทัลในการทดสอบจำนวน 50 ภาพ โดยภาพที่ใช้ในการทดสอบดังกล่าวเป็นภาพที่มีความหลากหลายรูปแบบมากยิ่งขึ้นทำให้อัลกอริทึม SR-CNN มีการเรียนรู้ที่มากยิ่งขึ้นกว่าเดิมในการหาอัลกอริทึมที่มีผลลัพธ์ที่เหมาะสมสำหรับข้อมูลโดยรวม

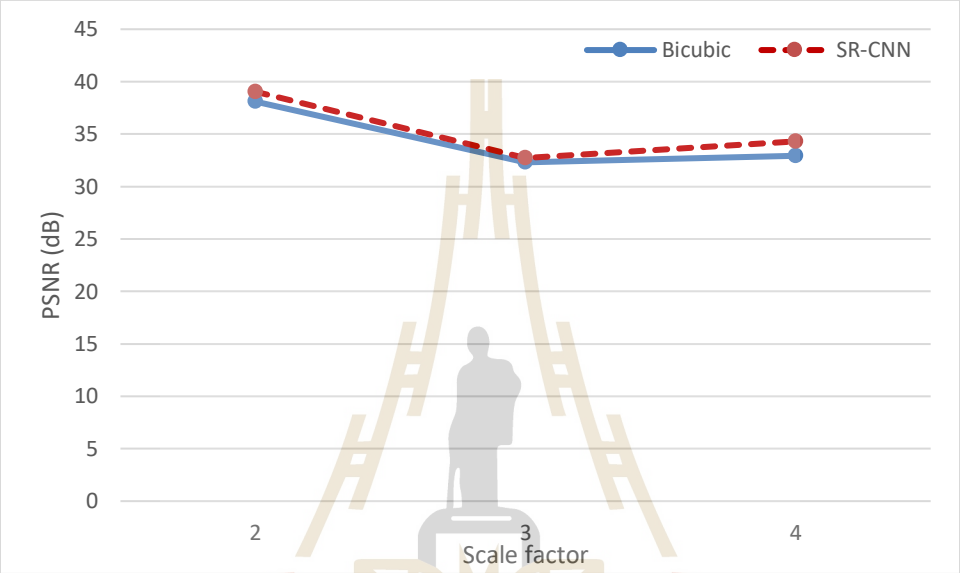
ในรูปที่ 4.28 แสดงภาพตัวอย่างบางส่วนที่ใช้ในการทดสอบจำนวน 36 ภาพ โดยผู้วิจัยได้ทำการทดสอบวิธีการสร้างภาพความละเอียดสูง SR-CNN เทียบกับวิธีการ Bicubic โดยได้ทดสอบที่ค่าอัตราการขยายภาพ (Scale factor) จำนวน 3 ค่าทั้งนี้เพื่อทำการทดสอบประสิทธิภาพของอัลกอริทึมที่ผู้วิจัยนำเสนอในเงื่อนไขที่แตกต่างกัน ผลการทดสอบที่ค่าอัตราการขยายภาพ 2 3 และ 4 โดยเปรียบเทียบการประเมินคุณภาพของสัญญาณภาพดิจิทัลโดยใช้ค่า PSNR SSIM NIQE และ BRISQUE แสดงในตารางที่ 4.5 และตารางที่ 4.6 และแสดงกราฟของการเปรียบเทียบค่าดังกล่าวในรูปที่ 4.29 ถึงรูปที่ 4.32 ตามลำดับ โดยพบว่าวิธีการที่นำเสนอ (SR-CNN) ให้ผลการทดสอบค่าคุณภาพที่ดีกว่าในทุกกรณี



รูปที่ 4.28 ภาพตัวอย่างในการทดสอบการสร้างภาพความละเอียดสูงวิธี SR-CNN เทียบกับวิธี Bicubic

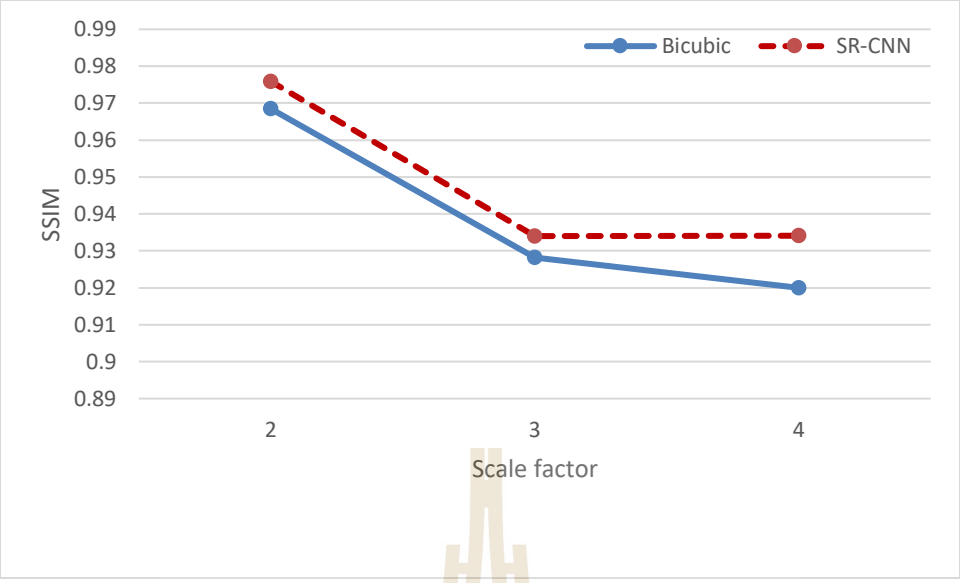
ตารางที่ 4.5 การเปรียบเทียบค่าเฉลี่ยของ PSNR และ SSIM

Scale factor	PSNR (dB)		SSIM	
	Bicubic	SR-CNN	Bicubic	SR-CNN
2	38.0973	39.0422	0.9685	0.9759
3	32.3056	32.7202	0.9282	0.9340
4	32.9528	34.3182	0.9200	0.9341



รูปที่ 4.29 เปรียบเทียบค่า PSNR ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN

มหาวิทยาลัยเทคโนโลยีสุรนารี

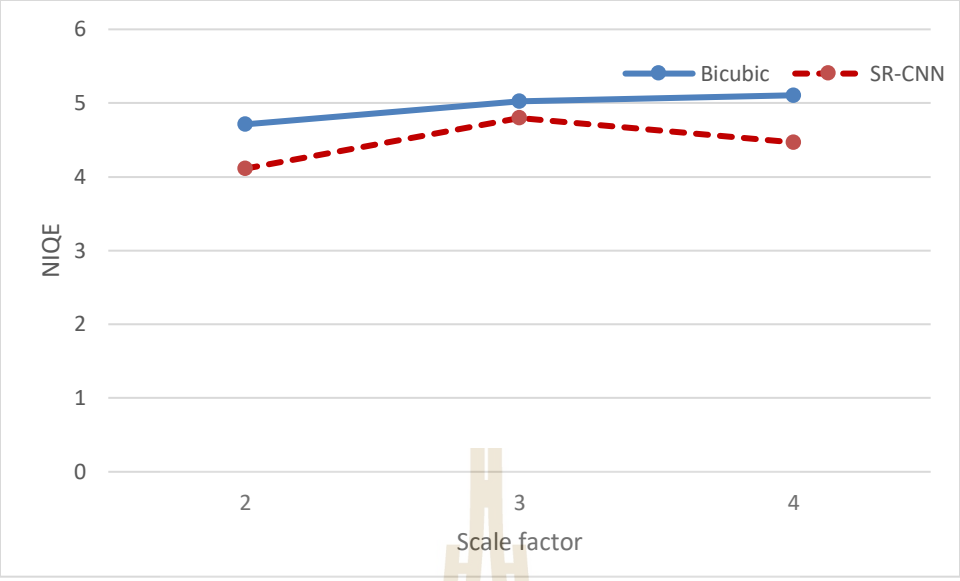


รูปที่ 4.30 เปรียบเทียบค่า SSIM ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN

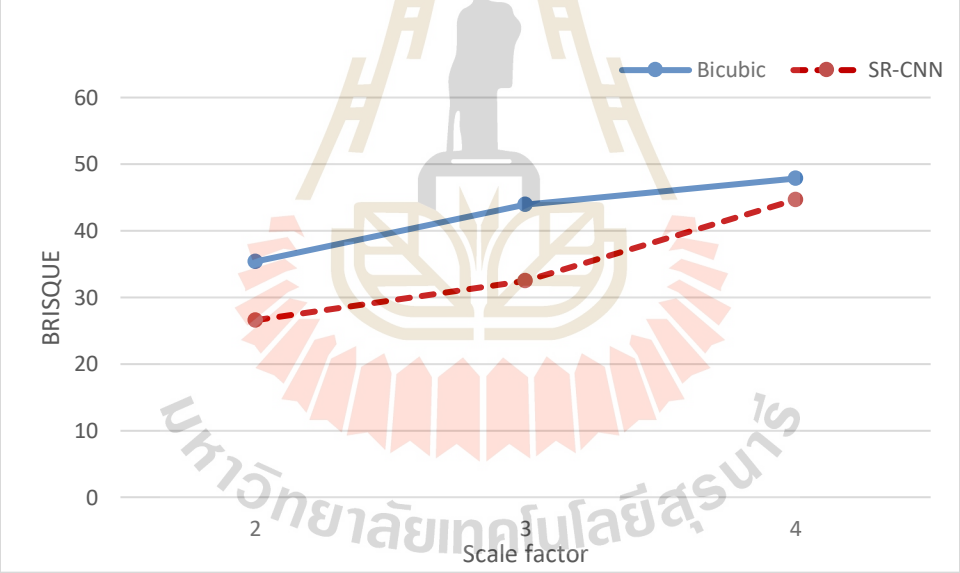
ตารางที่ 4.6 การเปรียบเทียบค่าเฉลี่ยของ NIQE และ BRISQUE

Scale factor	NIQE		BRISQUE	
	Bicubic	SR-CNN	Bicubic	SR-CNN
2	4.7119	4.1117	35.3857	26.5742
3	5.0247	4.7978	43.8827	32.4630
4	5.1063	4.4658	47.8454	44.6626

มหาวิทยาลัยเทคโนโลยีสุรนารี



รูปที่ 4.31 เปรียบเทียบค่า NIQE ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN



รูปที่ 4.32 เปรียบเทียบค่า BRISQUE ของภาพดิจิทัลระหว่างวิธี Bicubic และ SR-CNN

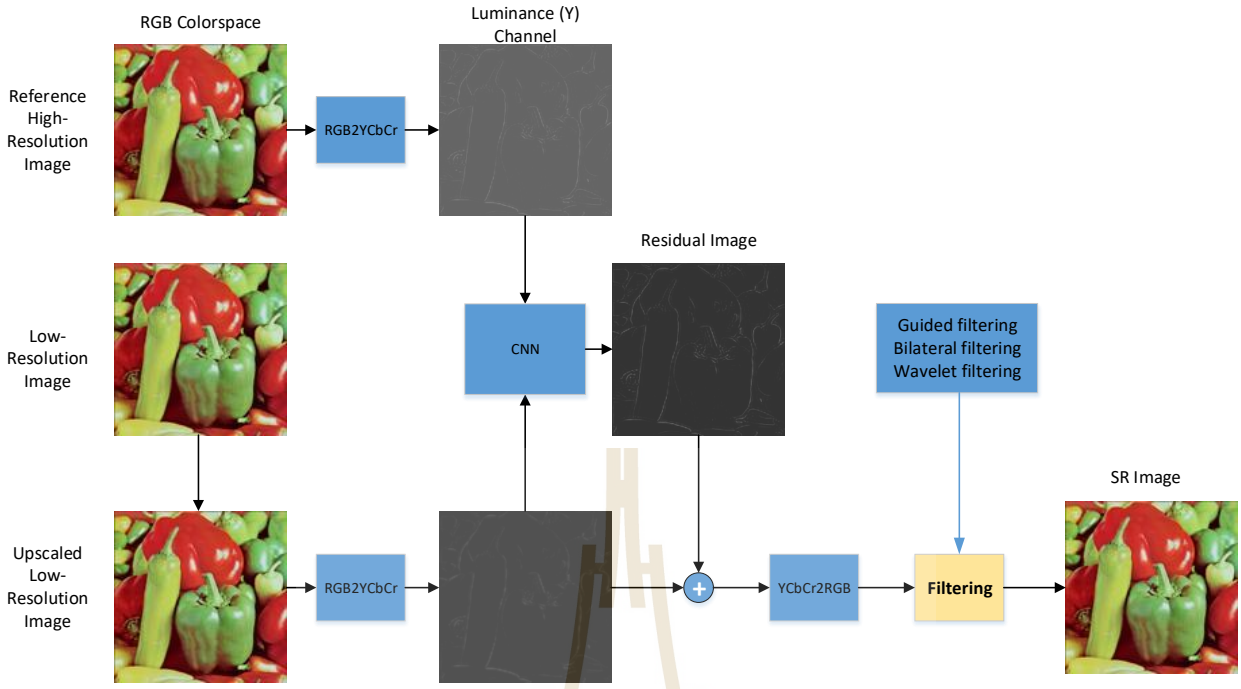
มหาวิทยาลัยเทคโนโลยีสุรนารี

ในตารางที่ 4.5 และตารางที่ 4.6 และรูปกราฟที่นำเสนอในรูปที่ 4.29 ถึงรูปที่ 4.32 เป็นการนำเสนอผลการทดสอบการสร้างภาพความละเอียดสูงด้วยวิธีการที่นำเสนอ (วิธี SR-CNN) เทียบกับวิธีการมาตรฐานทั่วไปที่นิยมใช้โดยทั่วไปในการขยายภาพที่เรียกว่าวิธีการ Bicubic interpolation (วิธี Bicubic) จากผลการทดลองดังกล่าวจะสังเกตเห็นรูปกราฟได้อย่างชัดเจนว่า ผลการทดสอบการสร้างภาพความละเอียดสูงด้วยวิธีการที่นำเสนอ (วิธี SR-CNN) มีค่าคุณภาพที่ดีกว่าวิธีการมาตรฐานทั่วไป (วิธี Bicubic) ไม่ว่าจะประเมินด้วยวิธีการแรกซึ่งใช้การคำนวณค่าความแตกต่างของค่าพิกเซล (Pixel) ของภาพระหว่างภาพต้นแบบกับภาพที่ผ่านการประมวลผลสัญญาณ เช่น การคำนวณ PSNR SSIM และในแบบที่สองที่ใช้วิธีการประเมินคุณภาพของสัญญาณภาพด้วยการใช้แบบจำลองทางคณิตศาสตร์ โดยเราจะสังเกตเห็นค่า NIQE และ BRISQUE ของวิธี SR-CNN ต่ำกว่าวิธี Bicubic ซึ่งแสดงให้เห็นว่าวิธีการที่นำเสนอให้คุณภาพของภาพความละเอียดที่ดีกว่าในทุกกรณี

4.5 การปรับปรุงคุณภาพด้วยการใช้ตัวกรองแบบดิจิทัล

ในหัวข้อนี้จะนำเสนอ การปรับเพิ่มความละเอียดของภาพ (Image enhancement) ความละเอียดสูงด้วยการใช้ตัวกรองแบบดิจิทัล (Digital filtering) โดยผู้วิจัยได้นำเสนอผลการทดสอบการประยุกต์ใช้ตัวกรองดิจิทัล 3 แบบ ดังนี้คือ แบบแรกเป็นตัวกรองดิจิตอลแบบไกด์ (Guided filtering) เป็นตัวกรองดิจิตอลที่ใช้ในการกรองภาพดิจิทัลเพื่อทำการปรับให้ภาพเนียนหรือมีความราบเรียบมากยิ่งขึ้นโดยที่ยังคงบริเวณที่เป็นขอบของภาพเอาไว้ (Edge-preserving smoothing) ทำให้ภาพดังกล่าวยังคงมีความชัดเจนดังเดิม แบบที่สองเป็นตัวกรองดิจิตอลแบบไบเลเทอรอล (Bilateral filtering) เป็นตัวกรองดิจิตอลที่ใช้ในการกรองภาพดิจิทัลเพื่อทำการปรับให้ส่วนพื้นผิว (Texture) ที่ทำให้เห็นลวดลายในภาพมีความเรียบมากขึ้นโดยที่ยังคงบริเวณที่เป็นขอบของภาพเอาไว้ (Edge-preserving) ทำให้ภาพดังกล่าวมีความชัดเจนและน่าสนใจมากขึ้น โดยมีการทดสอบการใช้งานในปริภูมิสี RGB และแบบสุดท้ายเป็นตัวกรองดิจิทัลในโดเมนของการแปลงเวฟเล็ต (Wavelet transform) ซึ่งเป็นตัวกรองดิจิทัลที่ใช้ในการกรองเพื่อทำการกำจัดสัญญาณรบกวนที่เกิดขึ้นในภาพโดยที่ยังคงบริเวณที่เป็นขอบของภาพเอาไว้ ในรูปที่ 4.33 แสดงแผนภาพกระบวนการปรับปรุงคุณภาพของภาพความละเอียดสูงด้วยการใช้ตัวกรองแบบดิจิทัล

มหาวิทยาลัยเทคโนโลยีสุรนารี

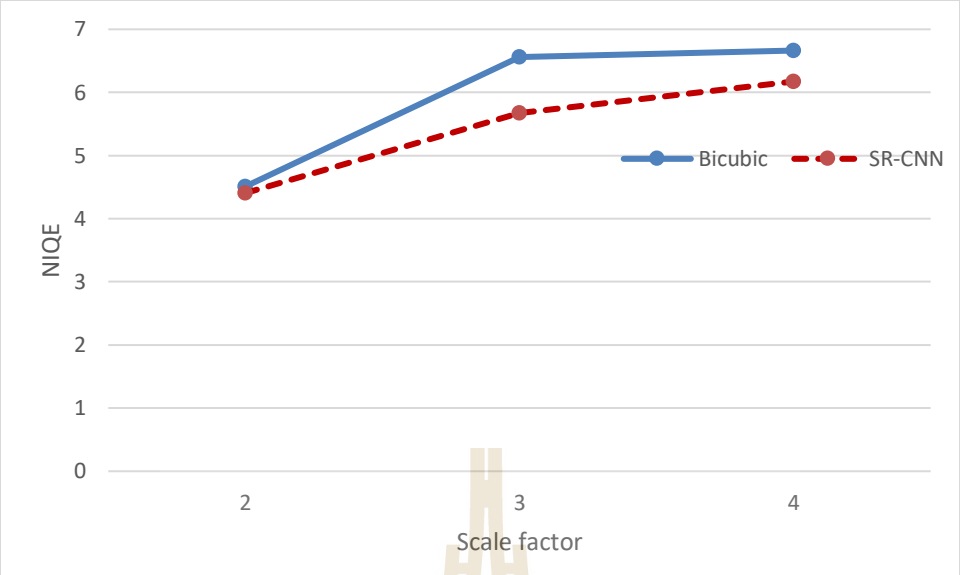


รูปที่ 4.33 กระบวนการปรับปรุงคุณภาพของภาพความละเอียดสูงด้วยตัวกรองแบบดิจิทัล

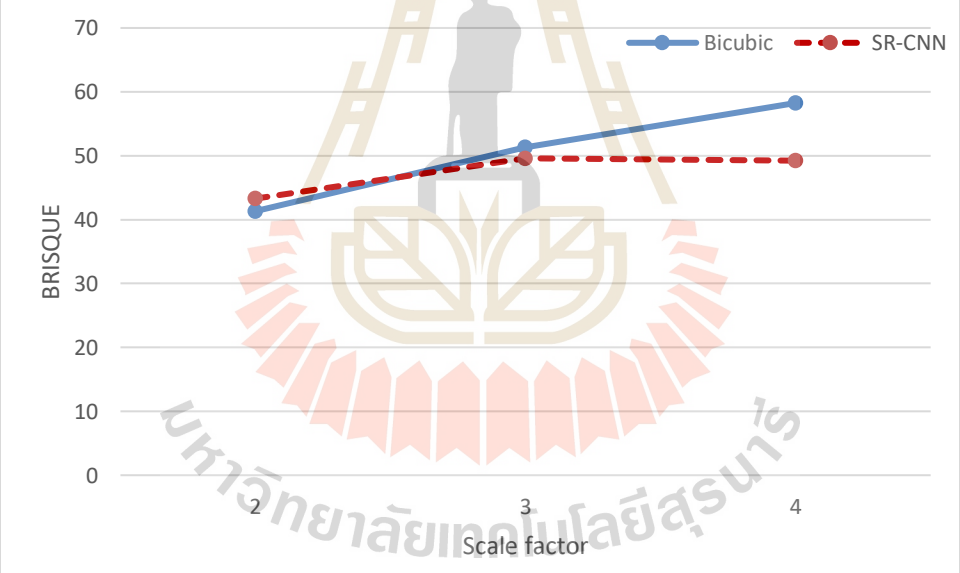
ในตารางที่ 4.7 แสดงการเปรียบเทียบค่าเฉลี่ย NIQE และ BRISQUE จากการทดสอบอัลกอริทึมการสร้างภาพความละเอียดสูงที่มีการประยุกต์ใช้ตัวกรองดิจิทัลแบบไกด์ฟิลเตอร์ (Guide digital filtering) และแสดงผลการเปรียบเทียบค่าเฉลี่ย NIQE และ BRISQUE จากการประยุกต์ใช้ตัวกรองดิจิทัลแบบไกด์ฟิลเตอร์ ในรูปที่ 4.34 และรูปที่ 4.35 ตามลำดับ

ตารางที่ 4.7 การเปรียบเทียบค่าเฉลี่ย NIQE และ BRISQUE จากตัวกรองแบบไกด์ฟิลเตอร์

Scale factor	NIQE		BRISQUE	
	Bicubic	SR-CNN	Bicubic	SR-CNN
2	4.5103	4.4063	41.3376	43.2926
3	6.5625	5.6767	51.2954	49.6013
4	6.6639	6.1736	58.2552	49.2683



รูปที่ 4.34 เปรียบเทียบค่าเฉลี่ย NIQE จากการใช้ตัวกรองดิจิทัลแบบไกด์ฟิลเตอร์



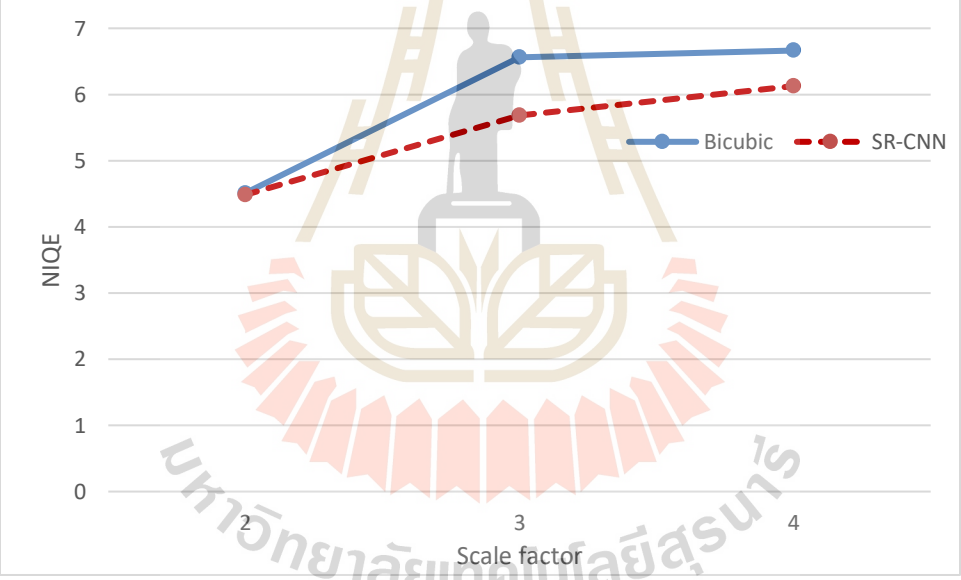
รูปที่ 4.35 เปรียบเทียบค่าเฉลี่ย BRISQUE จากการใช้ตัวกรองดิจิทัลแบบไกด์ฟิลเตอร์

มหาวิทยาลัยเทคโนโลยีสุรนารี

ในตารางที่ 4.8 แสดงการเปรียบเทียบค่าเฉลี่ย NIQE และ BRISQUE จากการทดสอบอัลกอริทึมการสร้างภาพความละเอียดสูงที่มีการประยุกต์ใช้ตัวกรองแบบไบเลเทอรอล (Bilateral filtering) ในปริภูมิสี RGB และแสดงผลการเปรียบเทียบค่าเฉลี่ย NIQE และ BRISQUE จากการประยุกต์ใช้ตัวกรองดิจิทัลแบบไบเลเทอรอล ในรูปที่ 4.36 และรูปที่ 4.37 ตามลำดับ

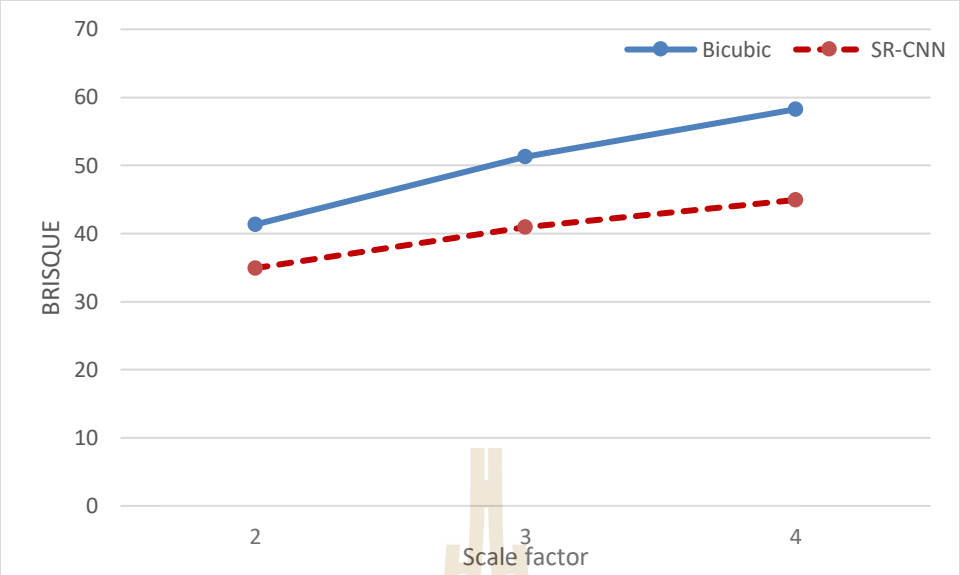
ตารางที่ 4.8 การเปรียบเทียบค่าเฉลี่ย NIQE และ BRISQUE จากการใช้ตัวกรองแบบไบเลเทอรอล

Scale factor	NIQE		BRISQUE	
	Bicubic	SR-CNN	Bicubic	SR-CNN
2	4.5103	4.4839	41.3376	34.9338
3	6.5625	5.6847	51.2954	40.9269
4	6.6639	6.1269	58.2552	44.9415



รูปที่ 4.36 เปรียบเทียบค่าเฉลี่ย NIQE จากการใช้ตัวกรองดิจิทัลแบบไบเลเทอรอล

มหาวิทยาลัยเทคโนโลยีสุรนารี

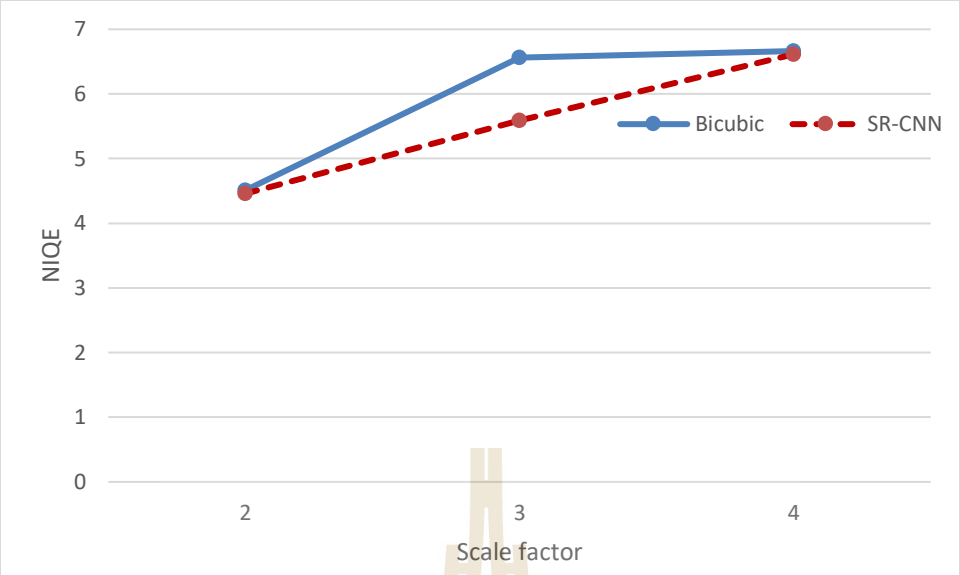


รูปที่ 4.37 เปรียบเทียบค่าเฉลี่ย BRISQUE จากการใช้ตัวกรองดิจิทัลแบบไบเลเทอรอล

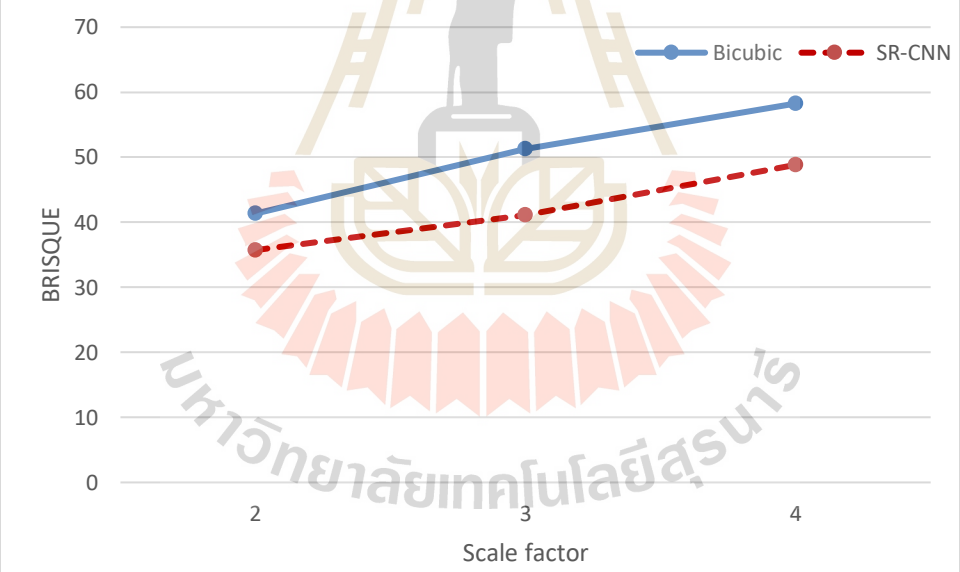
ในตารางที่ 4.9 แสดงการเปรียบเทียบค่าเฉลี่ย NIQE และ BRISQUE จากการทดสอบอัลกอริทึมการสร้างภาพความละเอียดสูงที่มีการประยุกต์ใช้ตัวกรองในโดเมนของการแปลงเวฟเล็ตในปริภูมิสี RGB และแสดงผลการเปรียบเทียบค่าเฉลี่ย NIQE และ BRISQUE จากการประยุกต์ใช้ตัวกรองดิจิทัลดังกล่าวตามรูปที่ 4.38 และรูปที่ 4.39 ตามลำดับ

ตารางที่ 4.9 เปรียบเทียบค่าเฉลี่ย NIQE และ BRISQUE จากการใช้ตัวกรองในโดเมนเวฟเล็ต

Scale factor	NIQE		BRISQUE	
	Bicubic	SR-CNN	Bicubic	SR-CNN
2	4.5103	4.4592	41.3376	35.7265
3	6.5625	5.5864	51.2954	41.1058
4	6.6639	6.6134	58.2552	48.8318

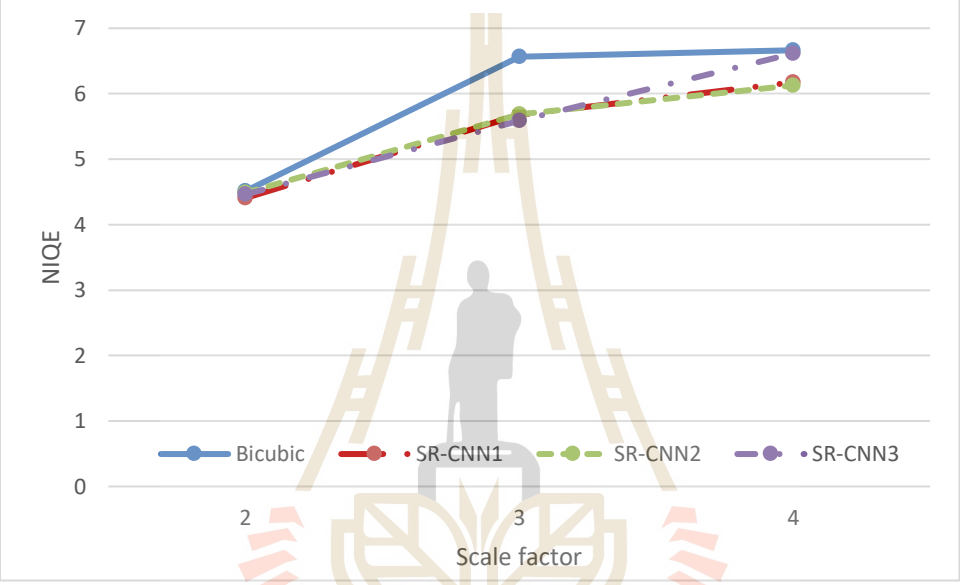


รูปที่ 4.38 เปรียบเทียบค่าเฉลี่ย NIQE จากการใช้ตัวกรองดิจิทัลในโดเมนเวฟเล็ต



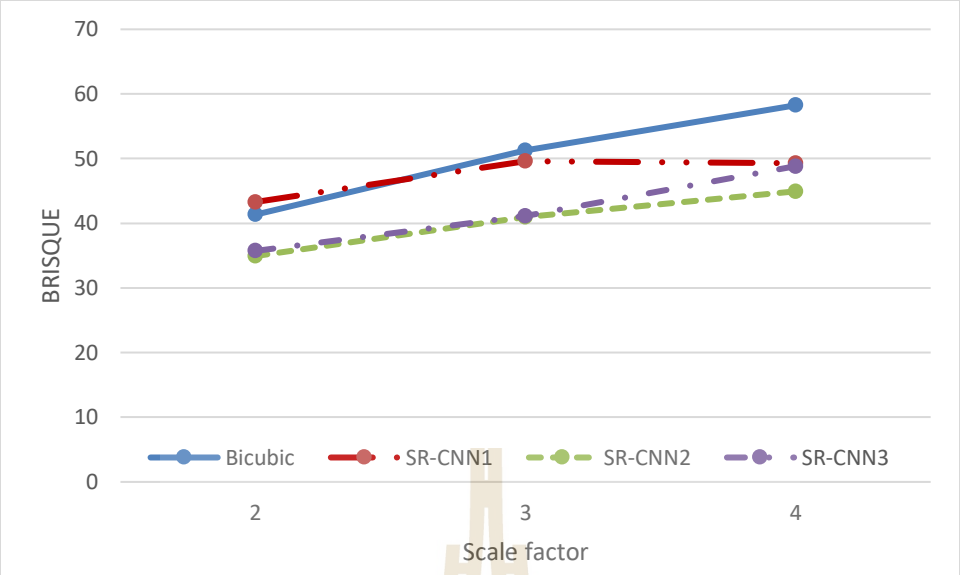
รูปที่ 4.39 เปรียบเทียบค่าเฉลี่ย BRISQUE จากการใช้ตัวกรองดิจิทัลในโดเมนเวฟเล็ต

ในรูปที่ 4.40 และรูปที่ 4.41 แสดงผลการทดลองเพื่อทำการเปรียบเทียบค่าเฉลี่ย NIQE และ BRISQUE ในการประยุกต์ใช้ตัวกรองดิจิทัล แบบไกต์ฟิลเตอร์ (Guide digital filtering) แบบไบเลเทอรอล (Bilateral filtering) ในปริภูมิสี RGB และตัวกรองดิจิทัลในโดเมนของการแปลงเวฟเล็ต (Wavelet filtering) โดยตัวกรองทั้ง 3 ตัวกรองที่ได้ประยุกต์ใช้จะใช้ชื่อเรียกสั้น ๆเพื่อให้ง่ายต่อการแสดงผลบนกราฟที่นำเสนอ โดยจะเรียกว่า SR-CNN1 SR-CNN2 และ SR-CNN3 ตามลำดับ และผู้วิจัยได้ทำการทดสอบโดยใช้ภาพตัวอย่างจำนวน 50 ภาพ ผลการทดสอบดังกล่าวได้แสดงให้เห็นว่าเมื่อมีการประยุกต์ใช้ตัวกรองดิจิทัลดังกล่าวแล้วสามารถทำให้ภาพดิจิทัลมีคุณภาพสูงมากยิ่งขึ้นและมีคุณภาพที่ดีกว่าวิธี Bicubic ในทุกกรณี โดยที่ตัวกรอง SR-CNN2 ให้ผลลัพธ์ที่ดีที่สุด



รูปที่ 4.40 เปรียบเทียบค่าเฉลี่ย NIQE จากการประยุกต์ใช้ตัวกรองดิจิทัล

มหาวิทยาลัยเทคโนโลยีสุรนารี



รูปที่ 4.41 เปรียบเทียบค่าเฉลี่ย BRISQUE จากการประยุกต์ใช้ตัวกรองดิจิทัล

4.6 สรุป

บทนี้ได้นำเสนอ ผลการทดสอบการสร้างภาพความละเอียดสูงด้วยเทคนิค Deep learning โดยในขั้นตอนของการสร้างภาพความละเอียดสูงดังกล่าว เทคนิค Deep learning ที่ออกแบบขึ้นโดยใช้เครือข่าย SR-CNN เพื่อใช้ในการเรียนรู้การสร้างผังที่แสดงความสัมพันธ์ (Mapping) เพื่อการจับคู่ระหว่างภาพความละเอียดต่ำกับภาพความละเอียดสูง การ Mapping ดังกล่าว CNN สามารถทำได้อย่างมีประสิทธิภาพ เนื่องจากภาพความละเอียดต่ำกับภาพความละเอียดสูงมีส่วนขององค์ประกอบภาพที่คล้ายคลึงกัน (Similar content) และองค์ประกอบหลักที่แตกต่างกันเฉพาะรายละเอียดความถี่สูง (High-frequency) เท่านั้น เครือข่ายประสาทเทียม SR-CNN ที่มีสถาปัตยกรรมของการเรียนรู้เชิงลึกจะใช้กลยุทธการเรียนรู้เรซิดิว เพื่อทำการประมาณค่าภาพเรซิดิว (Residual image) ที่เป็นผลต่างระหว่างภาพความละเอียดสูงต้นแบบหรืออาจเรียกว่าภาพอ้างอิง ดังนั้นแล้วภาพเรซิดิวจึงเป็นภาพที่จะช่วยเพิ่มรายละเอียดในด้านความถี่สูงของภาพให้ครบถ้วน

ในการประเมินคุณภาพของสัญญาณภาพดิจิทัล ผู้วิจัยได้ใช้ตัวประเมินคุณภาพของสัญญาณที่แตกต่างกัน 2 วิธีคือ วิธีแรกใช้วิธีการประเมินคุณภาพที่ใช้การคำนวณค่าความแตกต่างของค่าพิกเซลในภาพระหว่างภาพต้นฉบับหรือภาพอ้างอิงกับภาพที่ผ่านการประมวลผลสัญญาณ ซึ่งได้แก่ การประเมินคุณภาพด้วยค่า PSNR และ SSIM และแบบที่สองจะใช้วิธีการประเมินคุณภาพของสัญญาณภาพด้วยการใช้แบบจำลองทางคณิตศาสตร์ของสัญญาณภาพซึ่งไม่จำเป็นต้องใช้ภาพต้นฉบับหรือภาพอ้างอิงแต่อย่างใด ซึ่งได้แก่ การประเมินด้วยค่า NIQE และ BRISQUE

ผู้วิจัยได้ทำการทดสอบเทคนิคการสร้างภาพความละเอียดสูงด้วยเทคนิค Deep learning เทียบกับวิธีการ Bicubic interpolation โดยใช้ภาพดิจิทัลในการทดสอบที่ค่าอัตราการขยายภาพ (Scale factor) ที่มีค่ามากขึ้นเพื่อทดสอบความสามารถของอัลกอริทึมที่ได้นำเสนอ และทำการทดสอบที่อัตราการขยายภาพ 2 3 และ 4 ผลการทดสอบการสร้างภาพความละเอียดสูงด้วยวิธีการที่นำเสนอ (วิธี SR-CNN) เทียบกับวิธีการ Bicubic

interpolation (วิธี Bicubic) แสดงให้เห็นว่าวิธีการสร้างภาพความละเอียดสูงด้วยวิธีการที่นำเสนอ (วิธี SR-CNN) มีค่าคุณภาพที่ดีกว่าวิธีการมาตรฐานทั่วไป (วิธี Bicubic) ไม่ว่าจะประเมินด้วยวิธีการแรกซึ่งใช้การคำนวณค่า PSNR และ SSIM และในแบบที่สองที่ใช้วิธีการประเมินคุณภาพของสัญญาณภาพด้วยการใช้แบบจำลองทางคณิตศาสตร์ โดยจะสังเกตเห็นว่าค่าการคำนวณ NIQE และ BRISQUE จากวิธี SR-CNN ต่ำกว่าวิธี Bicubic ซึ่งแสดงให้เห็นว่าวิธีการที่นำเสนอให้คุณภาพของภาพความละเอียดสูงที่ดีกว่าในทุกกรณี

นอกจากนี้ ผู้วิจัยยังได้ทำการทดสอบการประยุกต์ใช้ตัวกรองดิจิทัลกับเทคนิคการสร้างภาพความละเอียดสูง ผลการทดลองที่นำเสนอได้ทำการเปรียบเทียบค่าเฉลี่ย NIQE และ BRISQUE จากการประยุกต์ใช้ตัวกรองดิจิทัลแบบไกต์ฟิลเตอร์ (Guide digital filtering) ตัวกรองดิจิทัลแบบไบเลเทอรอล (Bilateral filtering) และตัวกรองดิจิทัลในโดเมนของการแปลงเวฟเล็ต (Wavelet filtering) โดยเรียกตัวกรองทั้ง 3 แบบว่า SR-CNN1 SR-CNN2 และ SR-CNN3 ตามลำดับ ผลการทดสอบดังกล่าวได้แสดงให้เห็นว่า เมื่อมีการประยุกต์ใช้ตัวกรองดิจิทัลดังกล่าวแล้วสามารถทำให้ภาพดิจิทัลมีคุณภาพสูงมากยิ่งขึ้นและมีคุณภาพที่ดีกว่าวิธี Bicubic ในทุกกรณี โดยที่ตัวกรอง SR-CNN2 ให้ผลลัพธ์การกรองสัญญาณที่ดีที่สุด ส่วนตัวกรองดิจิทัลในโดเมนของการแปลงเวฟเล็ตในงานวิจัยนี้ใช้เวลาในการประมวลผลมากที่สุด

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 5
บทสรุป

5.1 สรุปผลงานวิจัย

งานวิจัยนี้นำเสนอ การสร้างภาพความละเอียดสูงด้วยการเรียนรู้เชิงลึก (Single Image Super-Resolution Using Deep Learning, SR-CNN) โดยจะนำเสนอขั้นตอนและวิธีการในการทำงานของอัลกอริทึมที่พัฒนาขึ้นทั้งในส่วนของการสร้างภาพความละเอียดสูงและการปรับเพิ่มคุณภาพของสัญญาณด้วยตัวกรองแบบดิจิทัลแบบต่าง ๆ จากนั้นผู้วิจัยจะนำเสนอผลการทดสอบการทำงานของอัลกอริทึมการสร้างภาพความละเอียดสูงด้วยเทคนิค Deep learning โดยชุดข้อมูลสัญญาณภาพดิจิทัลจากฐานข้อมูล SIPA image database ผลการวิจัยสรุปได้ดังนี้

1.) การสร้างภาพความละเอียดสูงด้วย Deep learning เครือข่ายประสาทเทียมแบบคอนโวลูชันที่มีสถาปัตยกรรมของการเรียนรู้เชิงลึก (SR-CNN) ได้ถูกใช้ในการเรียนรู้เพื่อสร้างผังที่แสดงความสัมพันธ์ (Mapping) ในการจับคู่ระหว่างภาพความละเอียดต่ำกับภาพความละเอียดสูง การ Mapping ดังกล่าว CNN สามารถทำได้อย่างมีประสิทธิภาพ เนื่องจากภาพความละเอียดต่ำกับภาพความละเอียดสูงมีส่วนขององค์ประกอบภาพที่คล้ายคลึงกัน (Similar content) และองค์ประกอบหลักที่แตกต่างกันเฉพาะรายละเอียดความถี่สูง (High-frequency) เท่านั้น โดยเครือข่าย SR-CNN จะใช้กลยุทธการเรียนรู้เรซิดิว (Residual) เพื่อทำการประมาณค่าภาพเรซิดิว (Residual image) ซึ่งเป็นผลต่างระหว่างภาพความละเอียดสูงต้นแบบหรืออาจเรียกว่าภาพอ้างอิง (High-resolution reference image) กับภาพความละเอียดต่ำ (Low-resolution image) ซึ่งได้ทำการปรับเพิ่มขนาดของภาพให้มีขนาดเท่ากับภาพอ้างอิงด้วยวิธีการไบคิวบิก (Bicubic interpolation) ดังนั้นแล้วภาพเรซิดิวจึงเป็นส่วนที่จะช่วยเพิ่มรายละเอียดของภาพในด้านความถี่สูงให้ครบถ้วน

2.) ในการประเมินคุณภาพของสัญญาณภาพดิจิทัล ผู้วิจัยได้ประยุกต์ใช้วิธีการประเมินคุณภาพของสัญญาณภาพดิจิทัล 2 วิธีที่แตกต่างกันคือ วิธีแรกใช้วิธีการประเมินคุณภาพที่อาศัยการคำนวณค่าความแตกต่างของพิกเซลในภาพ ระหว่างภาพต้นฉบับหรือภาพอ้างอิงกับภาพที่ผ่านการประมวลผลสัญญาณ ซึ่งได้แก่ การประเมินคุณภาพด้วยค่า PSNR และ SSIM และแบบที่สองจะใช้วิธีการประเมินคุณภาพของสัญญาณภาพด้วยการใช้แบบจำลองทางคณิตศาสตร์ซึ่งไม่จำเป็นต้องใช้ภาพต้นฉบับหรือภาพอ้างอิงแต่อย่างใด (No-reference image quality assessment) ซึ่งถือว่าเป็นวิธีการใหม่ในการประเมินคุณภาพของสัญญาณภาพดิจิทัล ซึ่งได้แก่ การประเมินด้วยค่า NIQE และ BRISQUE

3.) ผู้วิจัยได้ทำการทดสอบเทคนิคการสร้างภาพความละเอียดสูงด้วยเทคนิค Deep learning เปรียบเทียบกับวิธีการ Bicubic interpolation โดยใช้ภาพดิจิทัลในการทดสอบที่ค่าอัตราการขยายภาพ (Scale factor) ที่มีค่ามากขึ้นเพื่อทดสอบความสามารถของอัลกอริทึมที่ได้นำเสนอ และทำการทดสอบที่อัตราการขยายภาพ 2 3 และ 4 ผลการทดสอบการสร้างภาพความละเอียดสูงด้วยวิธีการที่นำเสนอ (วิธี SR-CNN) เทียบกับวิธีการ Bicubic interpolation (วิธี Bicubic) แสดงให้เห็นว่าวิธีการสร้างภาพความละเอียดสูงด้วยวิธีการที่นำเสนอ (วิธี SR-CNN) มีค่าคุณภาพที่ดีกว่าวิธีการมาตรฐานทั่วไป (วิธี Bicubic) ไม่ว่าจะประเมินด้วยวิธีการแรกที่ใช้การคำนวณค่า PSNR และ SSIM และในแบบที่สองที่ใช้วิธีการประเมินคุณภาพของสัญญาณภาพด้วยการใช้แบบจำลองทางคณิตศาสตร์ โดยจะสังเกตเห็นว่าค่าการคำนวณ NIQE และ BRISQUE จากวิธี SR-CNN ต่ำกว่าวิธี Bicubic ซึ่งแสดงว่าวิธีการที่นำเสนอให้คุณภาพของภาพความละเอียดสูงที่ดีกว่าในทุก

กรณี ทั้งในแง่ของการวัดค่าคุณภาพของภาพดิจิทัลด้วยมาตรวัดแบบต่าง ๆรวมถึงรายละเอียดของภาพความละเอียดสูงที่มีความชัดเจนและมีขอบภาพที่คมชัดยิ่งขึ้น

4.) ผู้วิจัยยังได้ทำการทดสอบการประยุกต์ใช้ตัวกรองดิจิทัลกับเทคนิคการสร้างภาพความละเอียดสูงจากผลการทดลองที่นำเสนอ ได้ทำการเปรียบเทียบค่าเฉลี่ย NIQE และ BRISQUE จากการประยุกต์ใช้ตัวกรองดิจิทัลแบบไกต์ฟิลเตอร์ (Guide digital filtering) ตัวกรองดิจิทัลแบบไบเลเทอรอล (Bilateral filtering) และตัวกรองดิจิทัลในโดเมนของการแปลงเวฟเล็ต (Wavelet filtering) โดยเรียกตัวกรองทั้ง 3 แบบว่า SR-CNN1 SR-CNN2 และ SR-CNN3 ตามลำดับ ผลการทดลองได้แสดงให้เห็นว่า เมื่อมีการประยุกต์ใช้ตัวกรองดิจิทัลดังกล่าวแล้วสามารถทำให้ภาพดิจิทัลมีคุณภาพดีมากขึ้น และมีคุณภาพที่ดีกว่าวิธี Bicubic โดยที่ตัวกรอง SR-CNN2 ให้ผลลัพธ์การกรองสัญญาณที่ดีที่สุด ส่วนตัวกรองดิจิทัลในโดเมนของการแปลงเวฟเล็ตในงานวิจัยนี้ SR-CNN3 ใช้เวลาในการประมวลผลมากที่สุด เพราะมีการคำนวณในโดเมนของการแปลงสัญญาณจึงยังไม่เหมาะกับการประยุกต์ใช้งาน

5.2 แนวทางในการวิจัยและพัฒนาต่อไปในอนาคต

1. การสร้างภาพความละเอียดสูงด้วยการเรียนรู้เชิงลึก โดยใช้เครือข่ายประสาทเทียมแบบคอนโวลูชันที่มีสถาปัตยกรรมของการเรียนรู้เชิงลึก (SR-CNN) ใช้ระยะเวลาในการประมวลผลเพื่อการเรียนรู้ที่ยาวนานมาก จึงควรมีการปรับเปลี่ยนวิธีการประมวลผลให้อยู่ในรูปของการประมวลผลแบบขนาน (Parallel processing) เพื่อลดระยะเวลาในการคำนวณ หรืออาจมีการประยุกต์ใช้ฮาร์ดแวร์ประมวลผลที่มีสมรรถนะสูงมากยิ่งขึ้นโดยเฉพาะในส่วนของ Graphic processing unit (GPU) ที่ใช้ในการประมวลผลซึ่งจำเป็นต้องใช้การ์ดที่มีจำนวนแกน (Core) ของ CUDA GPU ที่มากขึ้น

2. ทำการพัฒนาเทคนิคการสร้างภาพความละเอียดสูงให้อยู่ในรูปของแอพพลิเคชันบนอุปกรณ์เคลื่อนที่ (Mobile applications) โดยการประยุกต์ใช้ตัวประมวลผลกราฟิก CUDA GPU ที่สามารถใช้งานได้ผ่านระบบประมวลผลแบบกลุ่มเมฆ (Cloud computing) ในขั้นตอนของการประมวลผลและในการเรียนรู้ของเครือข่ายประสาทเทียมแบบคอนโวลูชันที่มีสถาปัตยกรรมของการเรียนรู้เชิงลึก เพื่อให้สะดวกต่อการใช้งานจริงมากยิ่งขึ้น ทำให้เครื่องคอมพิวเตอร์ที่มีฮาร์ดแวร์การ์ด CUDA GPU ก็สามารถใช้งานได้อย่างสมบูรณ์

บรรณานุกรม

[1] S. C. Park, M. K. Park, and M. G. Kang, “Super-resolution Image Reconstruction: A Technical Review,” IEEE Signal Processing Magazine, Vol. 20, pp. 21–36, May 2003.

[2] R. Y. Tsai and T. S. Huang, “Multipleframe Image Restoration and Registration”, Advances in Computer Vision and Image Processing, pp. 1984 :JAI Press Inc..

[3] H. Stark and P. Oskoui, “High Resolution Image Recovery from Image-plane Arrays Using Convex Projections,” J. Opt. Soc. Am. A, Vol. 6, pp. 1715-1726, 1989.

[4] C. V. Jiji and S. Chaudhuri, “Single-frame Image Super-resolution through Contourlet Learning,” EURASIP Journal on Applied Signal Processing, Vol. 2006, pp 1–11, 2006.

[5] G. Lei and H. Zhiming, “A Projection on Convex Sets Super-Resolution Algorithm Using Wavelet Transform,” Proc. International Conference on Signal Processing, pp. 1039 – 1041, 2008.

[6] H. Ji and C. Fermuller, “Robust Wavelet-Based Super-resolution Reconstruction: Theory and Algorithm,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 31, no. 4, pp. 649–660, April 2009.

[7] T. Wang, Y. Zhang, Y. S. Zhang and L.X. Lin, “Automatic Super-resolution Image Reconstruction Based on Hybrid MAP-POCS” Proc. International Conference on Wavelet Analysis and Pattern Recognition, Vol. 1, pp. 426 – 431, 2007.

[8] L. A. Méndez, M. I. R. S. Morán and M. Castelán, “A Single-frame Super-resolution Innovative Approach,” Lecture Notes in Computer Science, Vol. 4827, pp. 640-649, 2007.

[9] P. P. Gajjar and M. V. Joshi, “New Learning Based Super-resolution: Use of DWT and IGMRF Prior” IEEE Transactions on Image Processing, Vol. 19, No. 5, pp. 1201-1213, May 2010.

[10] M. A. Ellis and W. F. Walker, “Super-resolution Image Reconstruction with Reduced Computational Complexity,” Proc. IEEE International Symposium on Ultrasonics, pp. 2351 – 2354, 2009.

[11] A. Wong, A. Mishra, D.A. Clausi and P. Fieguth, “Mammogram Image Super-resolution Based on Statistical Moment Analysis,” Proc. Canadian Conference on Computer and Robot Vision (CRV), pp. 339 – 346, 2010.

[12] L. Mei, L. W. Dong, S. D. Qing, C. Guoqiao and L. Huinian, “A New Super-resolution Image Reconstruction Method Based on Hybrid Genetic Algorithm,” Proc. IEEE International Conference on Control Applications, pp. 211 - 216, 2004.

[13] Ni, K.S.; Nguyen, T.Q.; “Image Super-resolution Using Support Vector Regression,” IEEE Transactions on Image Processing, Vol. 16(6), pp. 1596 – 1610, March 2007.

[14] S. Zhang and Y. Lu, “Image Resolution Enhancement Using a Hopfield Neural Network,” Proc. International Conference on Information Technology, pp. 224 – 228, 2007.

[15] วิทยากร อัศดรวิเศษ “การรู้จำใบหน้าและเป้าหมายอัตโนมัติและการสร้างคืนภาพความละเอียดสูงยิ่งยวด” วารสารช่างพูด ฉบับที่ 3/50 คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2552.

[16] M. T. and J. Nunez , “Merino Super-Resolution of Remotely Sensed Images With Variable-Pixel Linear Reconstruction,” IEEE Transactions on Geoscience and Remote Sensing, Vol. 45, No. 5, pp. 1446 – 1457, May 2007.

[17] G. V. Gompel, K. J. Batenburg, E. V. Casteele, W. V. Aarle, J. Sijbers, “A Discrete Tomography Approach for Super-resolution Micro-CT Images: Application to Bone,” Proc. IEEE International Symposium on Biomedical Imaging, pp. 816 – 819, 2010.

[18] R. Xi, “Super Resolution Processing of SAR Images by Matching Pursuit Method Based on Genetic Algorithm,” In Proc.: 2010 3rd International Congress on Image and Signal Processing (CISP2010), Vol. 5, pp. 2066 - 2070 , 16-18 October 2010.

[19] H. Okuhata, R. Imai, M. Ise, R. Y. Omaki, H. Nakamura, S. Hara, and I. Shirakawa, “Implementation of Dynamic-Range Enhancement and Super-Resolution Algorithms for Medical Image Processing,” In Proc.: 2014 IEEE International Conference on Consumer Electronics (ICCE), Vol. 1, pp. 181-184, 10-13 January 2014.

[20] J. Xu, C. Deng, X. Gao, D. Tao, and X. Li “Image Super-resolution Using Multi-Layer Support Vector Regression,” In Proc.: 2014 IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP), Vol. 5, pp. 5799-5803, 4-9 May 2014.

[21] อรฉัตร จิตต์ภักตร์. “Digital Image Processing,” พิมพ์ครั้งที่ 1, บริษัท สงวนกิจ พริ้นท์ แอนด์ มีเดีย, หน้า 200-222, 2552.

[22] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition”, CoRR abs/1409-1556, 2014.

[23] Introduction to Deep Learning, Available from: http://introtodeeplearning.com/

[24] Kim, J., J. K. Lee, and K. M. Lee. "Accurate Image Super-Resolution Using Very Deep Convolutional Networks." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016, pp. 1646-1654.

[25] Grubinger, M., P. Clough, H. Müller, and T. Deselaers. "The IAPR TC-12 Benchmark: A New Evaluation Resource for Visual Information Systems." Proceedings of the OntoImage 2006 Language Resources for Content-Based Image Retrieval. Genoa, Italy. Vol. 5, May 2006, p. 10.

[26] Kaiming He, Jian Sun, Xiaoou Tang, Guided Image Filtering. IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 35, Issue 6, pp. 1397-1409, June 2013

[27] Tomasi, C., and R. Manduchi. "Bilateral Filtering for Gray and Color Images". Proceedings of the 1998 IEEE International Conference on Computer Vision. Bombay, India. Jan 1998, pp. 836–846.

[28] Donoho, D.L.; I.M. Johnstone, “Ideal de-noising in an orthonormal basis chosen from a library of bases,” C.R.A.S. Paris, Ser. I, t. 319, pp. 1317–1322, 1994.

[29] Zhou, W., A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. "Image Qualifty Assessment: From Error Visibility to Structural Similarity." IEEE Transactions on Image Processing. Vol. 13, Issue 4, April 2004, pp. 600–612.

[30] Mittal, A., R. Soundararajan, and A. C. Bovik. "Making a Completely Blind Image Quality

Analyzer." IEEE Signal Processing Letters. Vol. 22, Number 3, March 2013, pp. 209–212.
[31] Mittal, A., A. K. Moorthy, and A. C. Bovik. "No-Reference Image Quality Assessment in the Spatial Domain." IEEE Transactions on Image Processing. Vol. 21, Number 12, December 2012, pp. 4695–4708.
[32] Mathworks (2019) Convolution neural networks, Available at: https://www.mathworks.com/solutions/deep-learning/convolutional-neural-network.html

ประวัตินักวิจัย

หัวหน้าโครงการวิจัย

ชื่อ (ภาษาไทย) ผู้ช่วยศาสตราจารย์ เรืออากาศเอก ดร.ประโยชน์ คำสวัสดิ์

(ภาษาอังกฤษ) Asst.Prof.Flt.Lt.Dr.Prayoth Kumsawat

ตำแหน่งทางวิชาการ ผู้ช่วยศาสตราจารย์

ที่อยู่ สาขาวิชาวิศวกรรมโทรคมนาคม สำนักวิชาวิศวกรรมศาสตร์
มหาวิทยาลัยเทคโนโลยีสุรนารี
111 ถนนมหาวิทยาลัย ตำบลสุรนารี อำเภอเมือง จังหวัดนครราชสีมา 30000
โทรศัพท์ 044 224392 โทรสาร 044 224603
E-mail: prayoth@sut.ac.th

ประวัติการศึกษา

2549 วิศวกรรมศาสตรดุษฎีบัณฑิต (วิศวกรรมไฟฟ้า)	มหาวิทยาลัยเทคโนโลยีสุรนารี
2541 วิศวกรรมศาสตรมหาบัณฑิต (วิศวกรรมไฟฟ้า)	มหาวิทยาลัยเกษตรศาสตร์
2536 วิศวกรรมศาสตรบัณฑิต (วิศวกรรมไฟฟ้า)	โรงเรียนนายเรืออากาศ

สาขาวิชาการที่มีความชำนาญพิเศษ

- Digital signal and image processing
- Artificial intelligence and applications
- Microcontroller and embedded systems
- Wireless sensor networks and applications
- Internet of Things (IoT) applications

มหาวิทยาลัยเทคโนโลยีสุรนารี