การลดความเบลอของภาพ

วิทยานิพนธ์

ของ

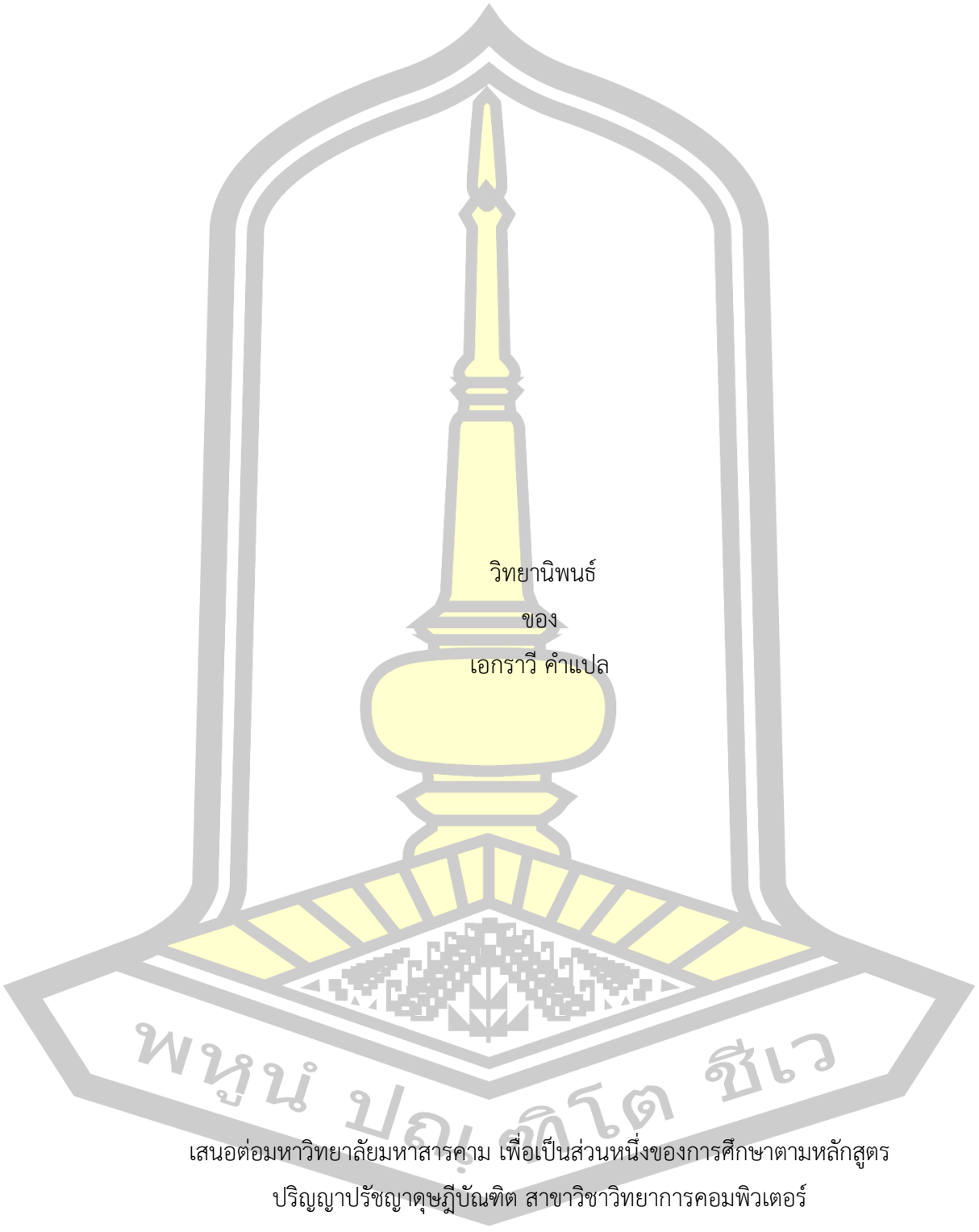
เอกราวี คำแปล

เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาปรัชญาดุษฎีบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์

มกราคม 2567

การลดความเบลอของภาพ



วิทยานิพนธ์

ของ

เอกราวี คำแปล

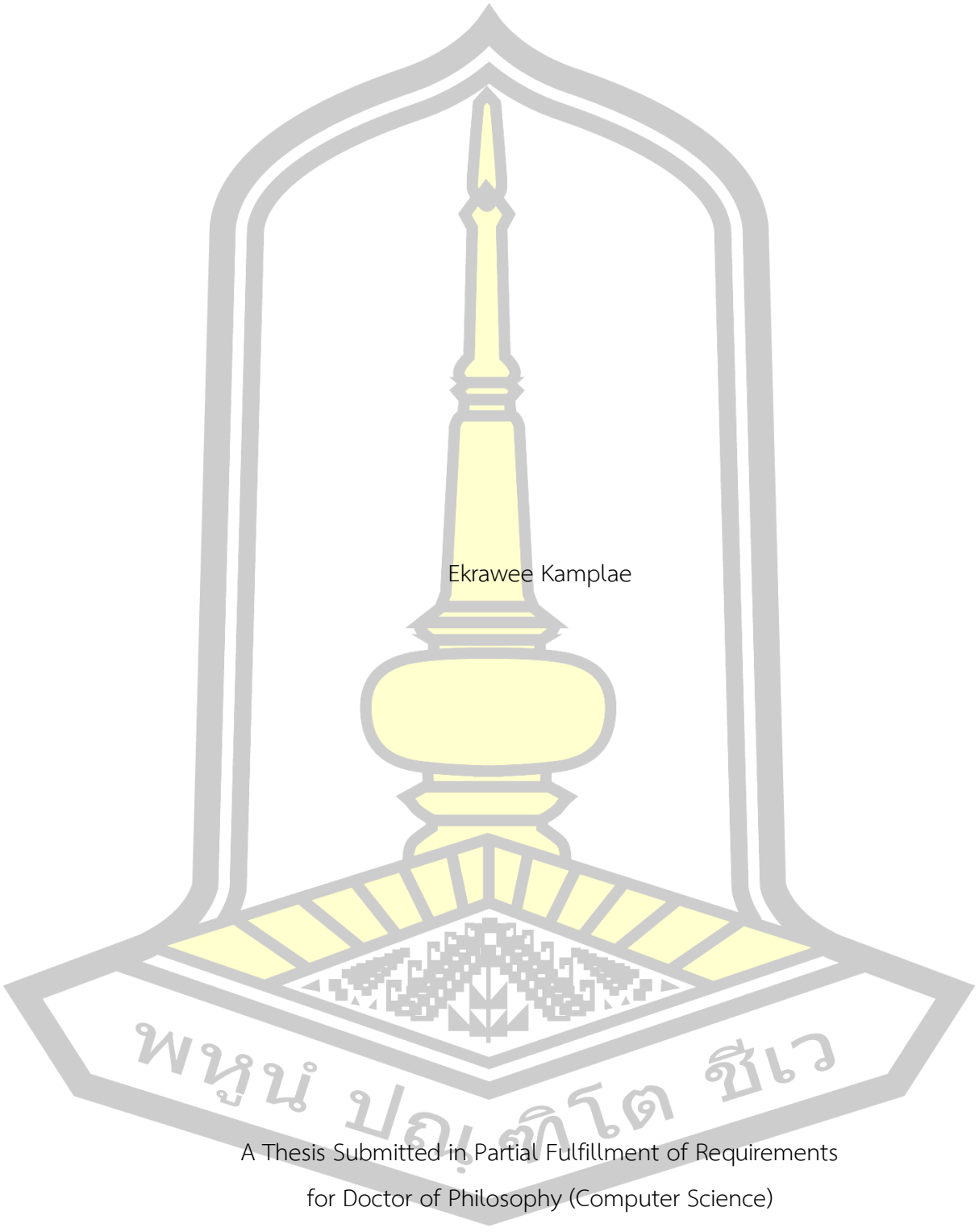
เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาปรัชญาดุษฎีบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์

มกราคม 2567

Image Deblurring

Ekrawee Kamplae



A Thesis Submitted in Partial Fulfillment of Requirements

for Doctor of Philosophy (Computer Science)

January 2024

คณะกรรมการสอบวิทยานิพนธ์ ได้พิจารณาวิทยานิพนธ์ของนายเอกราวี คำแปล แล้ว เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาปรัชญาดุษฎีบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ ของมหาวิทยาลัยมหาสารคาม

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ
(รศ. ดร. กริช สมกันธา)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผศ. ดร. พัฒนพงษ์ ชมภูวิเศษ)

.....กรรมการ
(ผศ. ดร. รพีพร ช่ำชอง)

.....กรรมการ
(ผศ. ดร. ฉัตรเกล้า เจริญผล)

.....กรรมการ
(รศ. ดร. พนิดา ทรงรัมย์)

มหาวิทยาลัยอนุมัติให้รับวิทยานิพนธ์ฉบับนี้ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญา ปรัชญาดุษฎีบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ ของมหาวิทยาลัยมหาสารคาม

.....
(รศ. ดร. จันทิมา พลพินิจ)
คณบดีคณะวิทยาการสารสนเทศ

.....
(รศ. ดร. กริสน์ ชัยมูล)
คณบดีบัณฑิตวิทยาลัย

ชื่อเรื่อง การลดความเบลอของภาพ
ผู้วิจัย เอกราวี คำแปล
อาจารย์ที่ปรึกษา ผู้ช่วยศาสตราจารย์ ดร. พัฒนพงษ์ ชมภูวิเศษ
ปริญญา ปรัชญาดุษฎีบัณฑิต **สาขาวิชา** วิทยาการคอมพิวเตอร์
มหาวิทยาลัย มหาวิทยาลัยมหาสารคาม **ปีที่พิมพ์** 2567

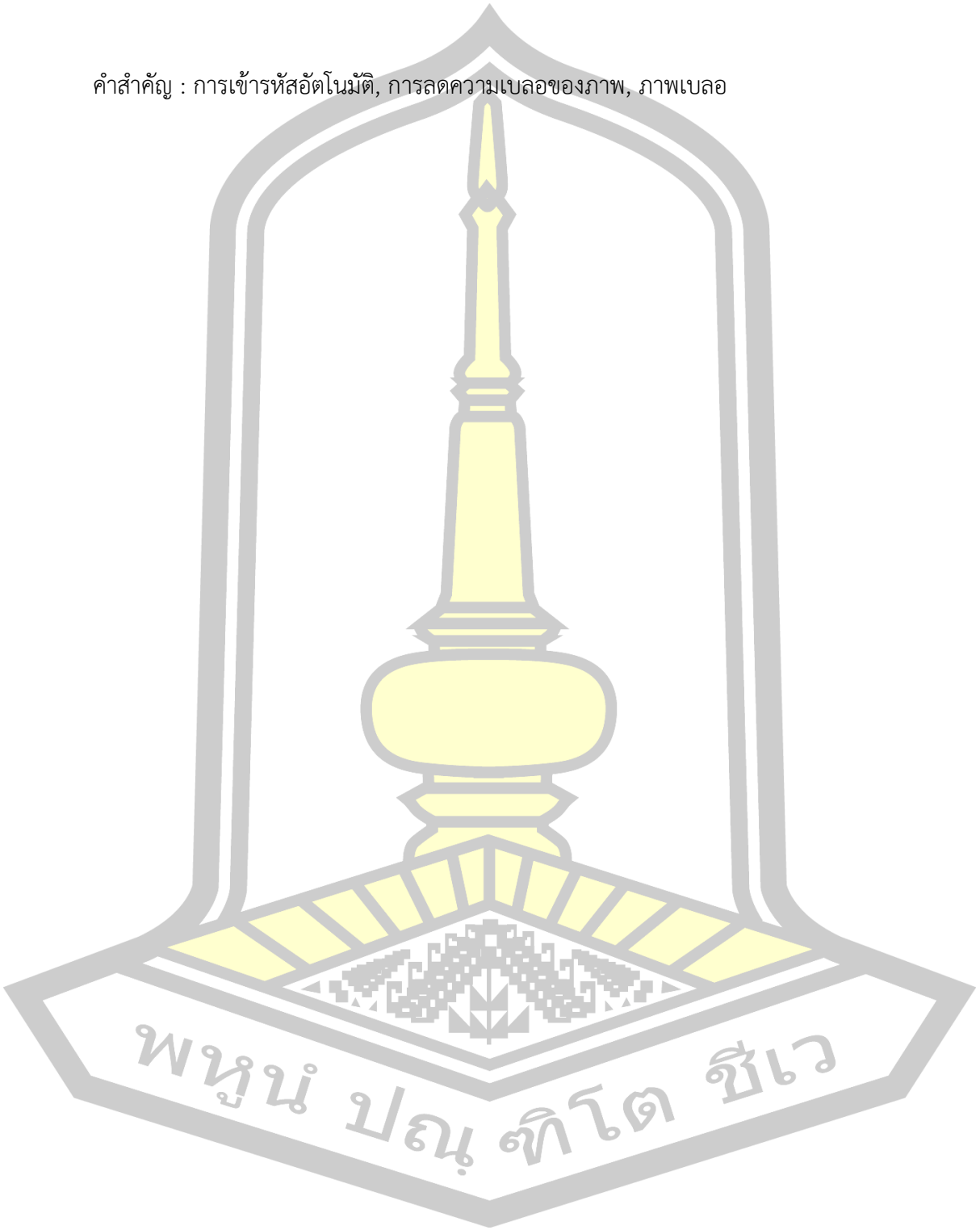
บทคัดย่อ

ข้อมูลภาพเป็นข้อมูลที่มีความสำคัญในปัจจุบัน จะเห็นได้ว่าแพลตฟอร์มทุกแพลตฟอร์มมีการนำส่วนข้อมูลภาพมาใช้เป็นส่วนประกอบ ดังนั้นข้อมูลภาพจึงจำเป็นสำหรับข้อมูลที่มีริชคอนเทนท์ (Rich Content) ในการนำมาใช้ประโยชน์ในหลายๆ ด้าน ไม่ว่าจะเป็นการนำมาใช้ในการตรวจสอบข้อมูลจากกล้อง การตรวจสอบภาพวิดีโอ บ่อยครั้งที่ข้อมูลเหล่านี้ไม่สามารถนำไปใช้ได้โดยตรงหรือใช้ได้งานได้อย่างมีประสิทธิภาพและถูกต้อง เนื่องจากปัญหาของคุณภาพของภาพ (Image Quality) ซึ่งอาจจะประกอบไปด้วยหลายสาเหตุ เช่น ภาพที่ถูกบิดเบือน ภาพที่โดนทำลาย ภาพที่เกิดจากการบีบอัดและภาพเบลอ ปัญหาเหล่านี้มักจะพบในภาพนิ่งและภาพเคลื่อนไหว โดยทั่วไปแล้วภาพที่เกิดในลักษณะนี้จะเป็นภาพที่มาจากกล้องวงจรปิดหรือกล้องถ่ายภาพนิ่งเป็นหลัก ซึ่งสาเหตุที่ทำให้ภาพจากกล้องวงจรปิดไม่มีความคมชัดจะมีข้อจำกัดในหลายรูปแบบ ไม่ว่าจะเป็นคุณภาพของกล้องวงจรปิดที่ยังคงใช้ระบบเก่า มีความละเอียดต่ำและมีจำนวนพิกเซลน้อย ระยะโฟกัสของกล้องและจุดที่ติดตั้งไม่มีความสัมพันธ์กัน สภาวะแสงน้อย หรือ Shutter Speed ต่ำ การบีบอัดไฟล์ภาพเพื่อประหยัดพื้นที่ในการจัดเก็บ อีกทั้งยังต้องเปิดทำงานงานตลอดเวลา รวมไปถึงการปรับความสมดุล (Balance) ระหว่างจำนวนเฟรมต่อวินาทีและระยะเวลาในการจัดเก็บ จากข้อจำกัดดังกล่าวจึงส่งผลให้ภาพที่ได้ไม่มีความคมชัด

สำหรับงานวิจัยนี้จึงได้ปรับปรุงภาพเบลอให้มีความชัดเจนและสามารถนำภาพเหล่านั้นกลับมาใช้งานได้อีก โดยมีวัตถุประสงค์ที่จะปรับปรุงภาพเบลอที่เกิดจากการสั่นให้มีความคมชัดด้วยวิธีการประมวลผลภาพผ่านกระบวนการเรียนรู้เชิงลึก (Deep learning) โดยใช้อัลกอริทึมของการเข้ารหัสอัตโนมัติ (Autoencoder) และ Generative adversarial Networks (GANs) เป็นโครงสร้างหลักในการวิจัย ซึ่งใช้ชุดข้อมูล 3 ชุดประกอบด้วย GoPro-Large, Real-world และ Kohler ซึ่งอัลกอริทึมของการเข้ารหัสอัตโนมัติ (Autoencoder) สามารถปรับภาพให้มีความเบลอลดลงและทำให้ภาพคมชัดขึ้น โดยเมื่อนำมาวัดประสิทธิภาพของภาพจะมีค่า SNR, PSNR และ SSIM เท่ากับ

32.64, 22.06, 0.777 ตามลำดับ ซึ่งสูงกว่า Baseline อย่างมีนัยสำคัญ

คำสำคัญ : การเข้ารหัสอัตโนมัติ, การลดความเบลอของภาพ, ภาพเบลอ



TITLE Image Deblurring

AUTHOR Ekrawee Kamplae

ADVISORS Assistant Professor Phatthanaphong Chompoowises , Ph.D.

DEGREE Doctor of Philosophy **MAJOR** Computer Science

UNIVERSITY Mahasarakham University **YEAR** 2024

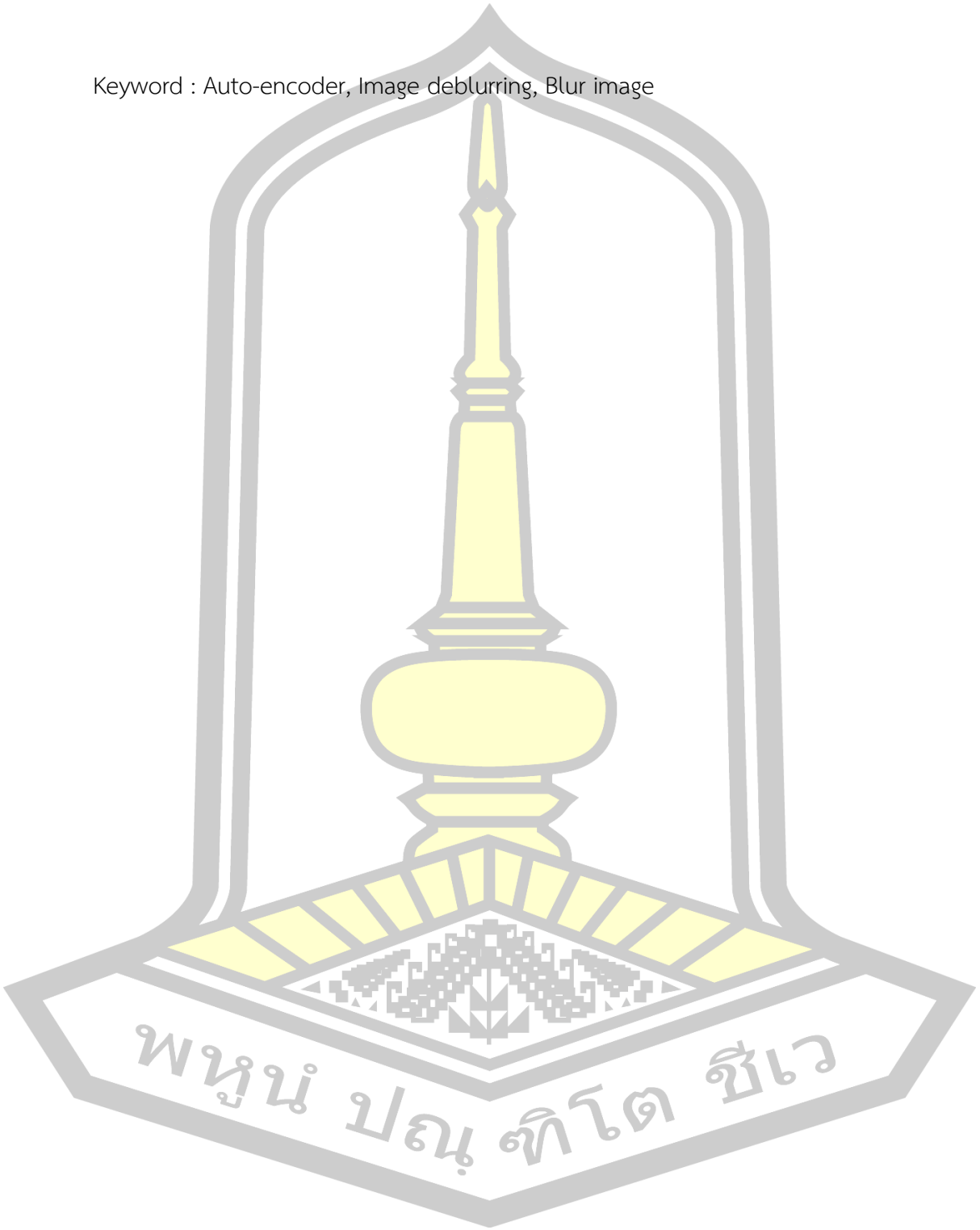
ABSTRACT

Pictorial information is important nowadays since every platform incorporates pictorial information as a component of the platform. Therefore, image data is essential for rich content and beneficial in various aspects, whether checking data from the camera or inspecting the video image. Often, this data can't be directly or efficiently used due to problems with image quality, which can stem from various causes such as distorted images, destroyed images, compressed images, and blurred images. These problems are commonly encountered in both still and moving images. Generally, images like these originate from CCTV or still cameras. The reasons for poor clarity in CCTV images can stem from various limitations, including outdated CCTV systems with low resolution and fewer pixels, unfocused camera distances and placements, low lighting conditions, slow shutter speeds, image file compression to save storage space, continuous operation at all times, balancing between frames per second, and storage duration. These limitations result in images lacking sharpness.

Thus, this study attempted to improve blurred images to enhance their clarity, making it possible to reuse these images effectively. The objective of this study was to improve blurry images caused by shaking through image processing using deep learning methods, employing Autoencoders and Generative Adversarial Networks (GANs) as the main structure of the research with three datasets: GoPro-Large, Real-World, and Kohler. The autoencoder algorithm successfully reduced blurriness in images, enhancing their clarity. When measuring the image performance, it has SNR, PSNR, and SSIM values of 32.64, 22.06, and 0.777, respectively, which are

significantly higher than baseline.

Keyword : Auto-encoder, Image deblurring, Blur image



กิตติกรรมประกาศ

วิทยานิพนธ์เรื่อง "การลดความเบลอของภาพ" ฉบับนี้ สามารถพัฒนาจนสำเร็จลุล่วงไปได้ด้วยดี โดยได้รับการดูแลเป็นอย่างดีจาก ผศ.ดร.พัฒนพงษ์ ชมภูวิเศษ ซึ่งเป็นที่ปรึกษาโครงการวิจัยและกรุณาสั่งสอน ให้คำแนะนำ ให้คำปรึกษา และช่วยชี้แนะแนวทาง รวมถึงการตรวจทานแก้ไขข้อบกพร่องต่างๆ ในการค้นคว้าหาความรู้ เพื่อนำมาใช้ในการเขียนโครงการวิจัยฉบับนี้ ผู้จัดทำรู้สึกซาบซึ้งในความกรุณาและอนุเคราะห์จากท่านอาจารย์ที่ปรึกษาเป็นอย่างมาก และขอกราบขอบพระคุณคณะกรรมการสอบทุกท่านที่ให้คำแนะนำ รวมไปถึงข้อเสนอแนะต่าง ๆ

ขอขอบพระคุณ คุณพ่อ คุณแม่ ญาติพี่น้อง รวมถึงเพื่อนๆ พี่ๆ และน้องๆ ทุกคน ที่ได้ให้ความช่วยเหลือพร้อมทั้งให้คำแนะนำ และเป็นกำลังใจทุกๆ อย่าง ให้สามารถดำเนินการในการทำโครงการวิจัยนี้ นอกจากนี้ผู้จัดทำขอขอบพระคุณคณาจารย์ในภาควิชาวิทยาการคอมพิวเตอร์ทุกท่าน ที่ได้ให้ความรู้ คำแนะนำ ที่สามารถนำมาปรับใช้ในระหว่างที่กำลังศึกษาเล่าเรียน ตลอดจนการนำไปใช้ประโยชน์ในการทำงานต่อไป อีกทั้งโครงการวิจัยนี้ยังได้รับทุนสนับสนุนงบประมาณจากสำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ (สวทช.) และมหาวิทยาลัยราชภัฏสุรินทร์ที่มอบโอกาสดีๆ อนุญาติให้ได้มาศึกษาต่อ สุดท้ายนี้ผลที่ได้รับและคุณค่าต่าง ๆ รวมไปถึงประโยชน์ที่เกิดขึ้นจากโครงงานวิจัยฉบับนี้ ผู้จัดทำขอมอบแด่ผู้มีพระคุณทุกท่าน

เอกราวี คำแปล

สารบัญ

หน้า

สารบัญภาพ

หน้า

สารบัญตาราง

หน้า

บทที่ 1
บทนำ

1.1 หลักการและเหตุผล

ข้อมูลภาพนับได้ว่าเป็นข้อมูลที่มีความสำคัญในปัจจุบันอีกอย่างหนึ่ง ซึ่งจะเห็นได้ว่าหลายๆ แพลตฟอร์มจะมีการนำส่วนข้อมูลภาพมาใช้เป็นส่วนประกอบในการนำมาใช้ประโยชน์ได้หลายๆ ด้าน ในการวิจัยนี้ได้แบ่งภาพออกเป็น 2 ประเภท ได้แก่ ภาพเคลื่อนไหวที่ได้จากกล้องวิดีโอหรือกล้องวงจรปิดและภาพที่ได้จากการถ่ายภาพนิ่ง ซึ่งหากนำมาใช้ในการตรวจสอบภาพวิดีโอหรือการตรวจสอบข้อมูลจากกล้อง บ่อยครั้งที่ข้อมูลเหล่านี้ไม่สามารถนำไปใช้ได้โดยตรงหรือใช้ได้งานได้อย่างมีประสิทธิภาพและถูกต้อง เนื่องจากปัญหาของคุณภาพของภาพ (Image Quality) ไม่ดีเท่าที่ควรจึงอาจจำไปใช้งานได้ไม่เต็มที่ ซึ่งอาจจะประกอบไปด้วยหลายสาเหตุ เช่น ภาพที่ถูกบิดเบือน ภาพที่โดนทำลาย ภาพที่เกิดจากการบีบอัดและภาพเบลอ ปัญหาเหล่านี้มักจะพบในภาพที่ได้จากกล้อง DSLR, กล้องมือถือ และภาพเคลื่อนไหวหรือวิดีโอที่มีอัตราเฟรมต่ำ [1] โดยทั่วไปแล้วภาพที่เกิดในลักษณะนี้จะเป็นภาพที่มาจากกล้องวงจรปิดหรือกล้องถ่ายภาพนิ่งเป็นหลัก

ภาพเคลื่อนไหวหรือภาพจากกล้องวงจรปิด ซึ่งสาเหตุที่ทำให้ภาพไม่มีความคมชัดจะมีข้อจำกัดในหลายรูปแบบดังนี้ ประการแรกเกิดจากคุณภาพของกล้องวงจรปิดที่ยังคงใช้ระบบเก่า มีความละเอียดต่ำและมีจำนวนพิกเซลน้อยจึงทำให้ภาพที่ได้ไม่มีความคมชัด ประการที่สองระยะโฟกัสของกล้องและจุดที่ติดตั้งไม่มีความสัมพันธ์กัน เช่น อาจจะติดตั้งในตำแหน่งที่ย้อนแสงหรือมีแสงสะท้อนเข้าไปได้น้อยจึงทำให้ภาพออกมามีคุณภาพไม่ดีหรือเบลอได้ ประการที่สามเกี่ยวกับการเคลื่อนไหวที่มีความสัมพัทธ์ระหว่างการเปิดรับแสงอาจจะอยู่ในสภาวะแสงน้อย ของกล้องกับฉากที่มีการเคลื่อนไหว [2] หรือ Shutter Speed ต่ำ โดยที่ Sensor จะเพิ่มค่า Shutter Speed เพื่อรักษาสภาพความสว่างของภาพ จึงทำให้เกิดสภาพ Motion Blur ทำให้ความชัดลดลง ประการที่สี่เกิดจากการบีบอัดไฟล์ภาพเพื่อประหยัดพื้นที่ในการจัดเก็บ อีกทั้งยังต้องเปิดทำงานงานตลอดเวลาจึงทำให้คุณภาพของภาพที่ได้ไม่มีความคมชัดเท่าที่ควร และประการสุดท้ายคือการปรับความสมดุล (Balance) ระหว่างจำนวนเฟรมต่อวินาทีและระยะเวลาในการจัดเก็บ ซึ่งอาจจะมีการปรับลด – เพิ่มบางอย่างใน Databases ให้สามารถที่จะบริหารจัดการได้ จากข้อจำกัดดังกล่าวจึงส่งผลให้ภาพที่ได้ไม่มีความคมชัด

นอกจากนี้ไฟล์ที่ได้จากการบันทึกของกล้องวงจรปิดจะมีขนาดใหญ่ ซึ่งจำเป็นที่จะต้องบันทึกลงในหน่วยความจำทั้งหมด จึงทำให้พื้นที่ในการจัดเก็บข้อมูลไม่เพียงพอ อย่างไรก็ตามในการแก้ปัญหาของผู้ดูแลระบบแต่ละที่อาจจะถูกกำหนดให้มีการลบข้อมูลออกจากหน่วยความจำทุกๆ 1, 6 หรือ 12 เดือนต่อครั้ง เพื่อเพิ่มพื้นที่ในหน่วยความจำให้กลับมาใช้งานได้อีก ดังนั้นภาพได้จากกล้อง

วงจรปิดจึงถูกลดคุณภาพของภาพลงโดยใช้วิธีการบีบอัดและลดค่าความละเอียด เพื่อให้สามารถเก็บลงในหน่วยความจำได้จึงทำให้ภาพที่ได้ไม่มีความละเอียดและคมชัด จากรายละเอียดดังกล่าวภาพที่ไม่ชัดนั้นอาจรวมไปถึงการโฟกัสภาพจากกล้องในบางมุมไม่ดีจึงส่งผลให้ภาพที่ได้กลายเป็นภาพเบลอ เช่น ภาพจากกล้องวงจรปิดเมื่อต้องการตรวจจับสิ่งผิดปกติในบางจุดหากซูมเข้าไปแล้วจะพบว่าภาพเหล่านั้นไม่มีความคมชัดทำให้ไม่สามารถเห็นความผิดปกติที่ต้องการได้

ภาพนิ่งเป็นภาพที่ได้จากการถ่ายด้วยกล้อง DSLR หรือกล้องมือถือ โดยทั่วไปแล้วภาพเบลอที่เกิดจากการถ่ายด้วยกล้องในลักษณะแบบนี้นั้นจะมีปัจจัยต่างๆ ที่ทำให้ภาพที่ได้ไม่มีความคมชัด เช่น การขาดโฟกัส กล้องสั่น หรือการเคลื่อนไหวของวัตถุที่เคลื่อนที่รวดเร็ว [3] ซึ่งอาจจะเกิดขึ้นได้ในหลายๆ งาน เช่น ภาพทางดาราศาสตร์ ภาพจากการสำรวจระยะไกล หรือสถานการณ์ที่อาจจะเกิดขึ้นได้เพียงครั้งเดียวและไม่สามารถจับภาพได้ชัดเจน ความเสื่อมสภาพของภาพนิ่งเหล่านี้ที่ไม่สามารถหลีกเลี่ยงได้อาจจะเกิดได้จากหลายปัจจัย เช่น ภาพถ่ายทางการแพทย์ซึ่งจะมีควบคุมระดับความเข้มของรังสีตกกระทบของเครื่องเอ็กซ์เรย์, CT scan, MRI ที่ใช้ในการรักษาเพื่อหลีกเลี่ยงความเสียหายต่ออวัยวะของผู้ป่วย ซึ่งภาพอาจจะไม่สามารถนำไปใช้งานได้ในบางส่วน ดังนั้นภาพที่ได้จึงจำเป็นต้องใช้การปรับปรุงภาพให้มีความคมชัดมากขึ้นเพื่อช่วยให้แพทย์สามารถนำภาพเหล่านี้ไปใช้ในการวินิจฉัยโรคต่อได้ [4]

จากสาเหตุดังกล่าวการลดความเบลอของภาพเริ่มได้รับความสนใจจากนักวิจัยเพิ่มมากขึ้น เนื่องจากความสำคัญและการลดสัญญาณรบกวนนี้ได้กลายมาเป็นพื้นฐานที่มีความสำคัญของงานวิจัย ดังนั้นการปรับปรุงภาพที่เสื่อมโทรมให้กลับมาใช้งานได้จึงได้รับความนิยมในการวิจัยมาจนถึงปัจจุบัน ด้วยเหตุนี้การฟื้นฟูภาพเบลอให้คมชัดนั้นจึงถือเป็นความท้าทายพื้นฐานของคอมพิวเตอร์วิทัศน์ในการประมวลผลภาพ ซึ่งวัตถุประสงค์ของการลดความเบลอของภาพ คือการกู้คืนภาพที่มีความคมชัดจากภาพเบลอที่อินพุตเข้าไปไม่ว่าจะเป็นภาพที่เกิดจากการเคลื่อนไหวของวัตถุ ความไม่สมบูรณ์ของแสงหรือสภาพบรรยากาศ ภาพนิ่งจากการสั่นของกล้อง [3, 5] เพื่อทำให้สามารถนำกลับมาใช้งานได้อีก ในยุคแรกๆ ของการปรับปรุงภาพนั้น เกิดจากการ Convolution ภาพชัดกับเคอร์เนล (Kernel) เพื่อให้ได้ภาพผลลัพธ์และทราบเคอร์เนลออกมา จากนั้นนำ Kernel ที่ได้มา Deconvolution กับภาพเบลอ เพื่อให้ได้ภาพที่คมชัดออกมา ซึ่งจะใช้วิธีการแบบ Linear และ Non-linear สำหรับการฟื้นฟูและกรองสัญญาณรบกวน (Noise) ออกจากภาพ [5] ซึ่งทำให้ภาพผลลัพธ์นั้นมีความชัดเจนเพิ่มขึ้นได้เพียงเล็กน้อย และในปัจจุบันนี้ได้มีนักวิจัยที่เสนอวิธีการปรับปรุงภาพในรูปแบบต่างๆ ไม่ว่าจะเป็นภาพเดี่ยว ภาพวิดีโอ โดยใช้เทคโนโลยีการเรียนรู้เชิงลึก (Deep learning) เข้ามาช่วยในการปรับปรุงภาพให้ดีขึ้น ด้วยเครือข่ายต่างๆ เช่น U-net, Autoencoder, GAN, Multi – scale Net, Reblurring Net และใช้วิธีการจัดการค่า Loss function ให้ลดลงเพื่อนำมาเปรียบเทียบกับภาพ

ต้นฉบับ อีกทั้งยังสามารถนำไปใช้ได้กับทุกแพลตฟอร์มเพื่อทำให้ความเบลอของภาพลดลงและทำให้ภาพชัดเจนขึ้นได้ [3]

จุดเด่นของงานวิจัยนี้จะใช้ความสามารถของอัลกอริทึมของแต่ละวิธีในการประมวลผล จะเริ่มต้นจากการแก้ปัญหาภาพเบลอที่มีความซับซ้อนและความแตกต่างกัน โดยจะใช้วิธีการเข้ารหัสอัตโนมัติ (Autoencoder) เข้ามาช่วยปรับปรุงกระบวนการลดการเบลอของภาพให้ดีขึ้น อย่างไรก็ตามเทคนิคต่างๆ ได้มีการพัฒนาเพิ่มมากขึ้นในปัจจุบัน โดยเฉพาะ CNNs ยังคงใช้งานได้อย่างมีประสิทธิภาพ เช่น การสร้างภาพที่ความละเอียดสูงและการสร้างภาพใหม่ ซึ่งวิธีการเข้ารหัสอัตโนมัติ (Autoencoder) กับชุดข้อมูลของ GoPro-large โดยจะนำเอาส่วนของการ Train มาประมวลผลเพื่อคำนวณหาค่า MSE โดยใช้วิธีการกำหนดค่า Loss function ในรูปแบบต่างๆ เพื่อทำให้ความแตกต่างระหว่างสองภาพนี้มีน้อยที่สุด จึงจะทำให้ภาพผลลัพธ์มีความคมชัดมากขึ้น แล้วจึงนำไปวัดประสิทธิภาพของภาพในด้านต่างๆ ซึ่งจะทำให้ค่า SNR, PSNR และ SSIM มีค่าสูงขึ้น

1.2 วัตถุประสงค์ของการวิจัย

เพื่อปรับปรุงภาพเบลอ (Blur) ที่เกิดจากการสั่นให้มีความคมชัดด้วยวิธีการประมวลผลภาพผ่านกระบวนการเรียนรู้เชิงลึก (Deep Learning)

1.3 ความสำคัญของการวิจัย

ภาพเบลอ (Blurring images) เป็นภาพที่อาจถูกมองว่าไม่สามารถนำไปใช้งานได้ ปัญหาเหล่านี้มักเกิดขึ้นกับภาพนิ่งหรือภาพเคลื่อนไหว หากมีความจำเป็นที่ต้องการนำภาพเหล่านี้มาใช้งานมักจะถูกนำมาปรับแก้ในกระบวนการต่าง ๆ เพื่อทำให้สามารถนำมาใช้งานได้ โดยทั่วไปแล้วข้อมูลที่ได้จากกล้องวงจรปิดเป็นข้อมูลที่มีขนาดใหญ่ เนื่องจากมีการจัดเก็บเป็นช่วงของเวลาจึงส่งผลให้ภาพที่ได้มีทั้งความคมชัดและเบลอ ซึ่งปัญหาที่ทำให้ภาพจากกล้องวงจรปิดไม่คมชัดหรือเบลอนั้นอาจเกิดจากหลายสาเหตุประกอบด้วย ตัวอุปกรณ์เก่าจึงทำให้มีประสิทธิภาพการทำงานไม่ดี มีความละเอียดของภาพต่ำหรือไม่มีความทันสมัย รวมถึงการบีบอัดไฟล์ภาพเพื่อให้สามารถบันทึกลงในหน่วยความจำที่มีอย่างจำกัด เป็นต้น

จากการวิเคราะห์เพื่อหาแนวทางในการแก้ไขภาพที่ไม่ชัดจากกล้องวงจรปิด พบว่าแนวโน้มที่จะสามารถแก้ไขภาพดังกล่าวได้แบ่งออกเป็น 2 แบบ ได้แก่ ด้านอุปกรณ์ (Hardware) และการแก้ไขจากไฟล์ภาพ (Image Files) ที่ได้จากกล้องวงจรปิด จากการวิเคราะห์แนวทางในการแก้ไขทั้ง 2 วิธี พบว่าหากมีการเปลี่ยนอุปกรณ์ใหม่ ยังคงมีปัญหาที่ยังแก้ไม่ได้คือการติดกล้องในตำแหน่งต่าง ๆ ที่ไม่สามารถโฟกัสได้ตรงตามความต้องการ เมื่อติดตั้งแล้วจะไม่สามารถเคลื่อนที่หรือปรับเปลี่ยนได้ง่าย และหากติดในที่ที่มีความถี่แสงน้อยภาพที่ได้จะไม่มีความคมชัดเหมือนเดิม

จากปัญหาดังกล่าวจะเห็นได้ว่าการไฟล์ภาพที่ได้จากกล้องวงจรปิดยังไม่มีคุณภาพ จึงจำเป็นที่จะต้อง ใช้เทคนิคการประมวลผลภาพเข้ามาช่วยปรับปรุงคุณภาพของภาพเพื่อทำให้ภาพชัดขึ้น (Image Deblurring) อย่างไรก็ตามในขั้นตอนการปรับปรุงภาพด้วยเทคนิคการประมวลผลภาพจะสามารถสกัดเอกลักษณ์ที่สำคัญของภาพให้อยู่ในรูปแบบของ Image Semantics และแปลงไปเก็บในรูปแบบของข้อความ (Text) เพื่อประหยัดและลดขนาดของพื้นที่ในการจัดเก็บข้อมูล อีกทั้งยังสามารถเก็บข้อมูลได้ตลอดเวลาซึ่งจะเป็นการลดต้นทุนในส่วนของพื้นที่การจัดเก็บข้อมูลอย่างคุ้มค่า

1.4 ขอบเขตของการวิจัย

งานวิจัยนี้แบ่งขอบเขตออกเป็น 3 ส่วนสำคัญ ดังต่อไปนี้

1. สามารถปรับปรุงภาพเบลอให้คมชัด ผ่านกระบวนการประมวลผลภาพจากอัลกอริทึมที่สร้างขึ้น โดยการผสมผสานกับวิธีการประมวลผลภาพ (Image Processing) การเรียนรู้ของเครื่อง (Machin Learning) และการเรียนรู้เชิงลึก (Deep Learning)

2. ใช้ชุดข้อมูลภาพจากชุดข้อมูลมาตรฐานในการทดลอง โดยใช้ชุดข้อมูลภาพ 3 ชุด ประกอบด้วย GoPro Large Datasets, Real – World Datasets และ Kohler’s Datasets

3. สามารถนำข้อมูลภาพที่ได้มาเปรียบเทียบประสิทธิภาพด้วยการหาค่า Signal-to-Noise-Ratio (SNR), Peak-Signal-to-Noise-Ratio (PSNR) และ Structural Similarity Index Measure (SSIM) กับข้อมูลก่อนหน้านี้และแสดงผลลัพธ์ได้อย่างชัดเจน

1.5 ประโยชน์ที่ได้รับจากการวิจัย

1. เมื่อผ่านกระบวนการปรับปรุงภาพแล้วจะได้ภาพที่คมชัดและสามารถนำไปใช้ประโยชน์กับระบบดิจิทัลได้

2. ได้อัลกอริทึมสำหรับการประมวลผลภาพแบบใหม่ที่สามารถทำให้ภาพที่มีคุณภาพสูงขึ้น

1.6 นิยามศัพท์เฉพาะ

1. Image Deblurring คือ กระบวนการแปลงภาพเบลอให้มีความคมชัดมากขึ้น

2. Image Processing คือ การกระทำการอย่างใดอย่างหนึ่งกับภาพต้นฉบับ (Input Image) เพื่อให้ได้ภาพผลลัพธ์ (Output Image) มีลักษณะของภาพเป็นไปตามที่ต้องการ

3. Rich Content คือ ฟีเจอร์พิเศษที่ใช้สำหรับการส่งข้อความหรือเนื้อหาที่สมบูรณ์ประกอบด้วยรูปแบบสื่อต่าง ๆ เช่น รูปภาพ เสียง และวิดีโอ

4. Blurring images คือ ภาพที่เบลอ ไม่มีความคมชัดที่เกิดจากการนำเข้าข้อมูลและข้อจำกัดของอุปกรณ์ในด้านต่าง ๆ เช่น การ Focus, Light

5. Image Files คือ ไฟล์ภาพที่ได้จากกล้องวงจรปิดเพื่อนำมาใช้ในการแก้ไข

6. Shutter Speed คือ ค่าความเร็วชัตเตอร์ เป็นการทำงานของม่านที่เปิด-ปิดให้เซนเซอร์ภาพ (Image Sensor) รับแสงที่ผ่านมาจากเลนส์

7. Low Level คือ การประมวลผลในระดับต่ำ โดยจะจัดการข้อมูลกับ Pixel โดยตรงเพื่อใช้ในการปรับปรุงคุณภาพของภาพ

8. High Level คือ การประมวลผลในระดับสูง โดยใช้วิธีการเรียนรู้ของเครื่องเข้ามาช่วยในการปรับปรุงภาพ ตัวอย่างเช่น อาศัยหลักการของการเรียนรู้เชิงลึกในการปรับปรุงคุณภาพของภาพและเป็นการสร้างภาพขึ้นมาใหม่ (Deconstruct) จากการเรียนรู้เชิงลึกในการสร้างต้นแบบ (Model) ที่เหมาะสม

9. Noise คือ สัญญาณรบกวน มีลักษณะเป็นจุดหรือเส้นที่ปกปิดภาพทำให้ภาพไม่มีความคมชัด

10. Datasets คือ ชุดข้อมูลภาพที่เป็นแบบมาตรฐานที่ใช้ในการทดลองประกอบด้วยชุดข้อมูลภาพของ GoPro-Large Datasets, Real-World Datasets และ Kohler's Datasets

1.7 ภาพรวมของปริญญานิพนธ์

ปริญญานิพนธ์นี้มีจุดประสงค์เพื่อพัฒนาวิธีการในการปรับปรุงคุณภาพของภาพในการลดความเบลอของภาพ (Image Deblurring) โดยผู้วิจัยได้ศึกษาเทคนิคด้านการเรียนรู้ของเครื่องและการเรียนรู้เชิงลึกเพื่อนำใช้เป็น สำหรับรายละเอียดของการทำวิจัยในเล่มนี้จะประกอบไปด้วยเนื้อหาทั้งหมด 5 บท โดยมีรายละเอียดของการทำวิจัยในแต่ละบทดังต่อไปนี้

สำหรับในบทที่ 2 จะศึกษาเกี่ยวกับทฤษฎีและงานวิจัยที่เกี่ยวข้อง สำหรับการทำวิจัยเรื่องการลดความเบลอของภาพ (Image Deblurring) ผู้วิจัยได้ศึกษาและรวบรวมข้อมูลทั้งทฤษฎีและงานวิจัยต่างๆ เพื่อนำมาประกอบการทำวิจัย โดยเน้นเทคนิคทางด้านคอมพิวเตอร์วิทัศน์ (Computer Vision) การประมวลผลภาพ (Image Processing) การลดความเบลอของภาพ (Image Deblurring) การเรียนรู้เชิงลึก (Deep Learning) รวมไปถึงโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution Neural Network: CNN) ดังนั้นข้อมูลที่ผู้วิจัยได้ศึกษาเทคนิคเหล่านี้และรวบรวมทั้งทฤษฎีแบบเก่าและแบบใหม่เพื่อสร้างเป็นองค์ความรู้ (Knowledge) ก่อนที่จะนำไปออกแบบการดำเนินการวิจัยในลำดับถัดไป

สำหรับบทที่ 3 จะเป็นวิธีดำเนินการวิจัย ในบทนี้จะเป็นการอธิบายขั้นตอนการทำงานวิธีการออกแบบ (Design) รวมไปถึงการเลือกใช้ชุดข้อมูลและเครื่องมือที่ใช้ในการวัดประสิทธิภาพของการวิจัย (Performance) ในการทำวิจัยนี้จะเน้นในเรื่องของการใช้เทคนิคเกี่ยวกับการปรับปรุงค่า Loss Function เพื่อปรับฟังก์ชัน Mean Squared Error (MSE) ให้มีค่าต่ำลง ก่อนนำไปเทียบกับ

ภาพชัด และวัดประสิทธิภาพของภาพซึ่งผลลัพธ์ที่ได้จะถูกนำไปรายงานและสรุปผลการวิจัยในขั้นตอนต่อไป

บทที่ 4 ผลการวิจัย สำหรับบทนี้จะนำเอาผลลัพธ์ที่ได้จากการออกแบบและทดลองมารายงานผล รวมถึงการวิเคราะห์ผลการวิจัยของแต่ละวิธีด้วยเทคนิคต่างๆ และการนำไปสู่การต่อยอดของการทำวิจัยในครั้งถัดไป

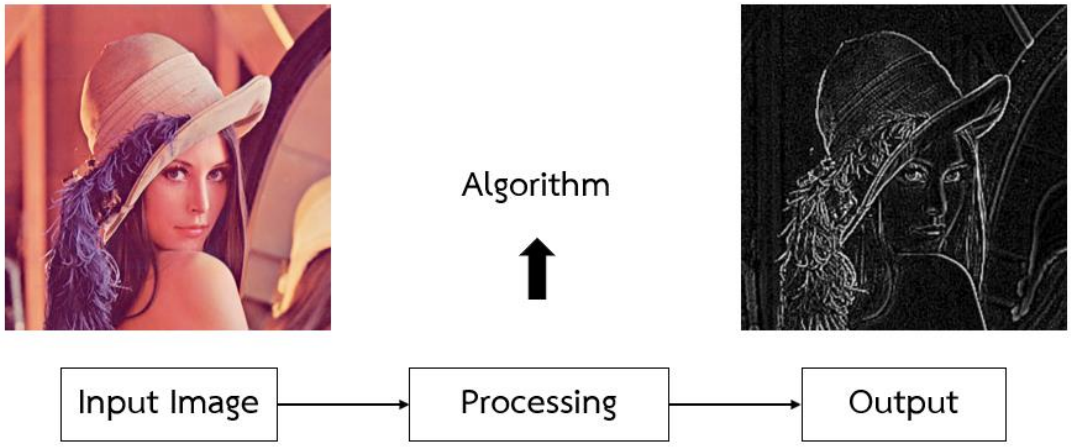
ในบทที่ 5 เป็นการสรุปผล การอภิปรายผล ข้อเสนอแนะ รวมไปถึงปัญหาที่เกิดขึ้นเกิดขึ้นกับการวิจัย

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในงานวิจัยนี้จะดำเนินการวิจัยเกี่ยวกับการปรับปรุงคุณภาพของภาพ ซึ่งจะต้องอาศัยเทคนิคต่างๆ ทางด้านคอมพิวเตอร์วิทัศน์ (Computer Vision) หรือการประมวลผลภาพเข้ามาช่วยในการทำภาพเบลอให้คมชัด ดังนั้นในบทนี้จึงได้อธิบายถึงทฤษฎีที่เกี่ยวข้องและจำเป็นทางด้านพื้นฐานการประมวลผลภาพ และอธิบายถึงการทำงานด้านการเรียนรู้เชิงลึก (Deep Learning) เพราะในปัจจุบันนี้จะเห็นได้ว่าการแก้ไขปัญหาเกี่ยวกับภาพโดยใช้ Deep Learning เข้ามาช่วยในการปรับปรุงภาพมีมากขึ้น จากนั้นจะได้อธิบายถึงสาเหตุที่ทำให้เกิดภาพเบลอรวมถึงเครื่องมือที่ใช้ในการวัดประสิทธิภาพและส่วนสุดท้ายจะได้กล่าวถึงงานวิจัยที่เกี่ยวข้องทางด้านการปรับปรุงภาพเบลอให้ชัดตั้งแต่เริ่มต้นจนถึงปัจจุบัน ซึ่งมีหัวข้อหลักๆ ดังต่อไปนี้

2.1 การประมวลผลภาพ (Image Processing)

การประมวลผลภาพ (Image Processing) เป็นกระบวนการประมวลผลภาพด้วยคอมพิวเตอร์ หรืออัลกอริทึมที่ใช้สำหรับแปลงภาพนำเข้า (Input Image) เพื่อสร้างภาพใหม่ให้ได้ภาพผลลัพธ์ (Output Image) ที่มีความละเอียด คมชัด หรือเป็นข้อมูลแบบดิจิทัลที่สามารถนำไปใช้งานในรูปแบบต่างๆ ตามความต้องการทั้งในเชิงปริมาณและคุณภาพ [6, 7] โดยใช้คอมพิวเตอร์ในการประมวลผล ซึ่งภาพที่นำมาใช้ในระบบการประมวลผลภาพดิจิทัลนั้นมีอยู่หลายรูปแบบ โดยจะพิจารณาจากปัจจัยต่างๆ เช่น ชนิดของภาพ ความละเอียดของภาพ โหมดโทนสีในระดับต่างๆ และความเข้มของสีให้อยู่ในระดับมาตรฐาน เพื่อจะช่วยให้ได้ภาพผลลัพธ์ของแต่ละแบบหรือข้อมูลที่ต้องการทั้งในเชิงปริมาณและเชิงคุณภาพ ดังภาพที่ 2.1 ซึ่งในปัจจุบันได้นำข้อมูลไปวิเคราะห์และสร้างเป็นระบบ เช่น ระบบดูแลการจราจรบนท้องถนนในการถ่ายภาพด้วยกล้องวงจรปิด ระบบตรวจสอบคุณภาพของผลิตภัณฑ์ในกระบวนการผลิต ระบบเก็บข้อมูลรถที่เข้า-ออก อาคารด้วยการถ่ายภาพป้ายทะเบียนรถเพื่อประโยชน์ในด้านความปลอดภัย ระบบตรวจจับใบหน้าโดยทำการวิเคราะห์ตรวจสอบและจดจำใบหน้า และการนำไปใช้งานทางด้านการแพทย์ เช่น นำเครื่อง MRI (Magnetic Resonance Imaging) การตรวจร่างกาย โดยใช้คลื่นสนามแม่เหล็กความเข้มสูงและคลื่นความถี่ในย่านความถี่วิทยุในการสร้างภาพเหมือนจริงของอวัยวะภายในต่างๆ ของร่างกาย และใช้ในการถ่ายภาพด้วยอัลตราซาวด์ (Ultrasound)

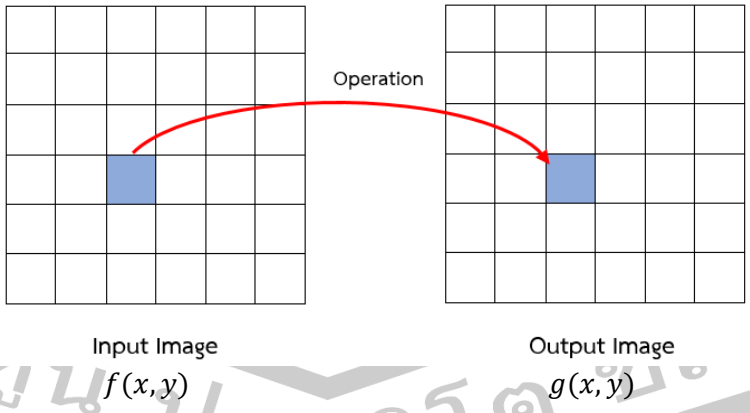


ภาพที่ 2.1 การประมวลผลภาพ (Image Processing)

สำหรับภาพดิจิทัลที่ถูกนำมาใช้ในการประมวลผล ซึ่งสามารถจำแนกได้ตามรูปแบบของการประมวลผลออกเป็น 2 แบบ คือ การประมวลผลภาพแบบจุด (Point Image Processing) และการประมวลภาพแบบบริเวณ (Local Image Processing) [8] ซึ่งมีรายละเอียดดังต่อไปนี้

2.1.1 การประมวลผลภาพแบบจุด (Point Image Processing)

สำหรับการประมวลผลภาพแบบจุด จะเป็นการนำภาพต้นฉบับที่ค่าระดับความเข้มเทา ในแต่ละพิกเซลของภาพผลลัพธ์ที่มีตำแหน่งตรงกันมาเปลี่ยนแปลงค่าพิกเซลของค่าผลลัพธ์ ซึ่งจะไม่ขึ้นกับค่าพิกเซลที่อยู่บริเวณใกล้เคียงของภาพต้นฉบับ ดังภาพที่ 2.2



ภาพที่ 2.2 ลักษณะการทำภาพแบบจุดต่อจุด

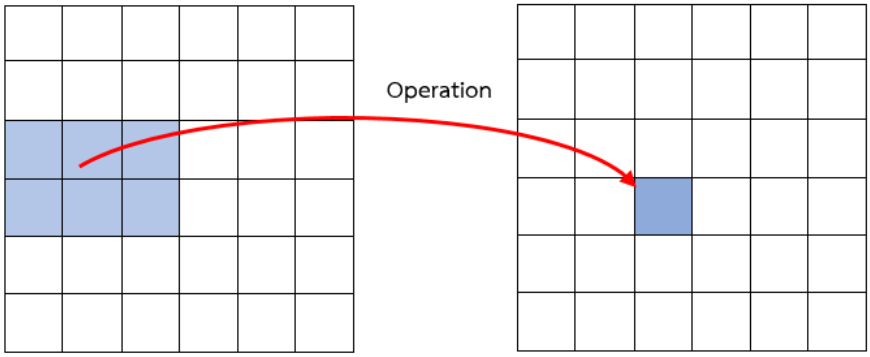
จากภาพที่ 2.2 กำหนดให้ $f(x,y)$ และ $g(x,y)$ เป็นภาพต้นฉบับและผลลัพธ์ตามลำดับ สมการที่ใช้ในการประมวลผลของภาพแบบจุดต่อจุด สามารถแสดงได้ดังสมการที่ (2.1)

$$g(x,y) = M[f(x,y)] \quad (2.1)$$

กำหนดให้ $M(.)$ เป็นฟังก์ชันแบบจุดหรือเป็นการแทนข้อมูลภาพด้วยวิธีการ Mapping Function โดยค่าระดับความเข้มเทาใหม่ที่ได้ในแต่ละพิกเซลของภาพจะถูกแทนที่ในแต่ละพิกเซลของภาพต้นฉบับที่พิกัด (x,y) เดิม รูปแบบของการประมวลผลภาพแบบดิจิทัลในลักษณะนี้ ได้แก่ การปรับค่าความสว่างและค่าคอนทราสต์ (Contrast) ของภาพดิจิทัล โดยใช้วิธีการบวก ลบ คูณ และหารด้วยจำนวนค่าใดๆ กับภาพดิจิทัลต้นแบบ

2.1.2 การประมวลผลภาพแบบบริเวณ (Local Image Processing)

ในกระบวนการประมวลผลภาพแบบบริเวณ เป็นการใช้ค่าระดับความเข้มเทาของพิกเซลในแต่ละจุดของภาพผลลัพธ์ ซึ่งจะขึ้นอยู่กับบริเวณข้างเคียงของกลุ่มพิกเซลนั้นๆ (Neighborhood Pixels) เพื่อนำมาใช้สำหรับการประมวลผล จากภาพที่ 2.3 แสดงให้เห็นถึงลักษณะของภาพเฉพาะบริเวณที่ใกล้เคียง ซึ่งจะเป็นรูปแบบของการประมวลผลภาพดิจิทัล ได้แก่ การกรองสัญญาณภาพในสเปเชียลโดเมน (Spatial Domain Filtering) หรือเรียกอีกอย่างหนึ่งว่า “Convolution”

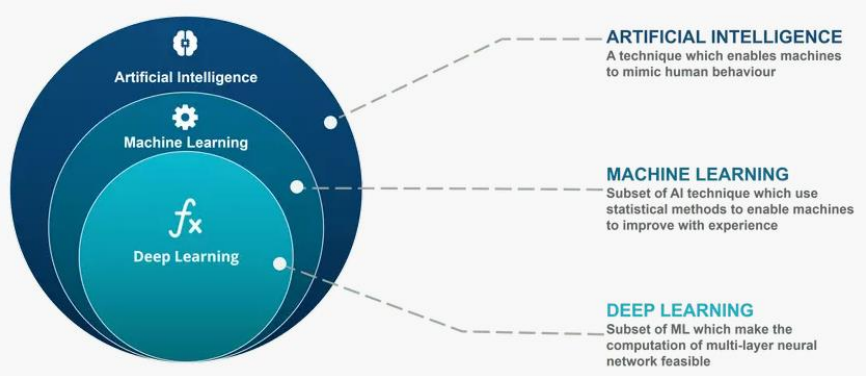


ภาพที่ 2.3 การประมวลผลภาพแบบบริเวณ

จากภาพที่ 2.3 กำหนดให้ค่า $f(x,y)$ และ $g(x,y)$ เป็นภาพต้นฉบับและภาพผลลัพธ์ที่จุดพิกัด (x,y) โดยจะนำค่าที่อยู่บริเวณใกล้เคียงของภาพ Input มาใช้ในการประมาณผลและจะได้ค่าที่เป็นผลลัพธ์ออกมา

2.2 การเรียนรู้เชิงลึก (Deep Learning)

การเรียนรู้เชิงลึก หรือ Deep Learning (DL) [9] เป็นวิธีการเรียนรู้แบบอัตโนมัติโดยอาศัยแนวคิดและเทคนิคการสร้างแบบปัญญาประดิษฐ์ (AI) ด้วยระบบโครงข่ายใยประสาทเสมือน (Artificial Neural Networks : ANN) แบบเดียวกับสมองมนุษย์ที่เชื่อมต่อกันเป็น "ระบบประสาท" สำหรับใช้ในการสื่อสารกัน ในการทำงานจะถูกสร้างขึ้นมาสำหรับการเรียนรู้ของเครื่องจักรหรือเครื่องคอมพิวเตอร์ ซึ่งชุดคำสั่งนี้จะสามารถนำไปใช้ในการประมวลผลข้อมูล โดยใช้วิธีประมวลผลแบบขนาน (Parallel Processing) ผ่านองค์ความรู้ (Knowledge) เพื่อทำให้เกิดการเรียนรู้จากข้อมูลที่ได้รับอย่างต่อเนื่องและสามารถเข้าใจในสิ่งต่างๆ ได้อย่างมีประสิทธิภาพ โดยในปี ค.ศ. 2006 Geoffrey Hinton [10] ได้เสนอโครงข่ายประสาทเทียมแบบหลายชั้นในการแปลงข้อมูลที่มีมิติระดับต่ำให้อยู่ในระดับกลางไปจนถึงระดับสูง ด้วยการไล่ระดับของสีและปรับระดับค่าน้ำหนักในเครือข่ายโดยใช้การเข้ารหัสแบบอัตโนมัติ ซึ่งการเรียนรู้เชิงลึก (DL) ได้เข้ามามีบทบาทมากขึ้น สำหรับกระบวนการประมวลผลภาพในช่วงระยะเวลาหลายปีที่ผ่านมา ดังนั้นการเรียนรู้เชิงลึกจึงเป็นที่นิยมนำไปใช้ในการประมวลผลในการทำงานด้านต่างๆ ดังภาพที่ 2.4



ภาพที่ 2.4 ลำดับชั้นของการเรียนรู้เชิงลึก [9]

จากภาพที่ 2.4 แสดงให้เห็นถึงลำดับชั้นของการเรียนรู้เชิงลึก ซึ่งประกอบด้วย Artificial Intelligence (AI) เป็นเทคโนโลยีที่อยู่ในกลุ่มของปัญญาประดิษฐ์ที่สามารถพบเห็นได้ในการใช้ในชีวิตประจำวัน โดยจะแบ่งออกเป็น 2 ประเภท ได้แก่ Machine Learning (ML) เป็นการรับข้อมูลที่มีจำนวนมาก ซึ่งทำหน้าที่ในการจดจำลักษณะเด่นและแบ่งข้อมูลออกเป็นกลุ่มๆ และในส่วนของ Deep Learning (DL) จะเป็นอีกวิธีหนึ่งที่สามารถจำลองรูปแบบการประมวลผลของสมองมนุษย์โดยใช้โครงข่ายเซลล์ประสาทในการประมวลผลภาพ

นอกจากนี้ Deep Learning ยังสามารถนำไปใช้งานในด้านต่างๆ ได้อีกมากมาย อย่างไรก็ตามรูปแบบการทำงานของ Deep Learning ยังคงมีข้อจำกัด เนื่องจาก Deep Learning มีรูปแบบ

การจดจำที่ซับซ้อน [11] โดยเฉพาะกับข้อมูลที่มีจำนวนมาก จึงจำเป็นที่จะต้องแบ่งข้อมูลออกเป็นสองส่วนเพื่อให้การเรียนรู้มีประสิทธิภาพมากขึ้น โดยใช้การแบ่งแบบสุ่ม เพื่อให้ได้ข้อมูลที่กระจายตัวทั้งชุดข้อมูล ได้แก่ Training Set เป็นการฝึกฝนชุดข้อมูลต้นแบบเพื่อให้สามารถเรียนรู้และจดจำข้อมูลได้อย่างแม่นยำ และในส่วนของ Test Set เป็นการนำโมเดลของ Training Set มาทดลองกับชุดข้อมูลที่แบ่งไว้ (Test Set) เพื่อทดสอบความแม่นยำ โดยทั่วไปแล้วในการแบ่งสัดส่วนของชุดข้อมูล Training Set และ Test Set สามารถแบ่งออกได้หลายแบบ เช่น (70:30, 80:20 หรือ 90:10) หรือขึ้นอยู่กับปริมาณข้อมูล ส่วนมากจะนิยมแบ่งเป็น 80:20

ตัวอย่างของการใช้งาน Deep Learning ที่สามารถนำมาใช้งานได้จริงและใช้ในปัจจุบันนี้ ได้แก่ รถยนต์อัจริยะหรือขับเคลื่อนอัตโนมัติ การวินิจฉัยโรค การแปลภาษา เกมส์ และการสร้างประโยคเพื่อโต้ตอบกับมนุษย์ (Siri) โดยมีรายละเอียดดังนี้

รถยนต์อัจฉริยะหรือขับเคลื่อนอัตโนมัติ (Autonomous Vehicles) สำหรับรูปแบบการทำงานของรถยนต์ไร้คนขับจะต้องใช้ (Computer Vision) และระบบนำทาง (Navigation) ในการรับข้อมูลเพื่อนำไปวิเคราะห์สภาพแวดล้อมต่างๆ ก่อนส่งมาประมวลผลที่หน่วยประมวลผลกลาง (Deep Learning) โดยใช้เทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence : AI) เข้ามาช่วยในการประมวลผลการวิเคราะห์สถานการณ์ต่างๆ โดยจะใช้เวลาที่น้อยที่สุด จากนั้นจะนำคำสั่งส่งต่อให้ระบบควบคุมรถ (Robotic) ที่มีความแม่นยำ เพื่อสั่งการให้รถยนต์ทำงานได้อย่างมีประสิทธิภาพ

การวินิจฉัยโรค ถูกนำมาเป็นเครื่องมือทางด้านการแพทย์เพื่ออำนวยความสะดวกและความรวดเร็วในการวินิจฉัยโรค เช่น ภาพเอ็กซ์เรย์ ภาพอัลตราซาวด์หรือภาพ MRI โดยนำ Deep Learning มาช่วยในการวิเคราะห์การประมวลผลภาพถ่ายของผู้ป่วย ก่อนนำไปเปรียบเทียบกับฐานข้อมูลเพื่อใช้ในการระบุตำแหน่งของความผิดปกติที่เกิดขึ้นกับอวัยวะต่างๆ

การแปลภาษา สำหรับระบบการแปลภาษาของ Google Translate จะใช้หลักการของ Deep Learning มาวิเคราะห์ข้อมูลจากการป้อนข้อมูลของผู้ใช้งาน ซึ่งจะมีหลายรูปแบบ เช่น ตัวอักษร รูปภาพหรือเสียง ก่อนจะนำข้อมูลเหล่านี้ไปเปรียบเทียบกับคำที่มีอยู่ในฐานข้อมูล จากนั้นจะหาความหมายที่มีความเหมาะสมและใกล้เคียงที่สุดมาแสดงผล

เกมส์ ในการพัฒนาเกมส์ได้ใช้ Deep Learning เข้ามาช่วยในการวิเคราะห์ โดยการส่งข้อมูลให้กับ AI เพื่อทำการเรียนรู้ด้วยการทำซ้ำจนกว่าจะได้ข้อมูลที่มีความแม่นยำ และสามารถสร้างรูปแบบการคิดให้ใกล้เคียงกับมนุษย์ก่อนจะประมวลผลออกมา และยังสามารถคิดวิธีการแก้ไขปัญหาและการเล่นกับมนุษย์ให้มีความสมจริงมากขึ้น

การสร้างประโยคหรือการโต้ตอบกับมนุษย์ (Chatbot) เป็นอีกหนึ่งรูปแบบที่สามารถนำ Deep Learning มาใช้ในการพูดคุยและเข้าใจภาษาของมนุษย์ผ่านข้อความ (Text) หรือข้อความเสียง (Voice) ก่อนประมวลผลแล้วตอบสนองในสิ่งที่ต้องการ เช่น Siri, Alexa หรือ Google

Assistant และในรูปแบบอื่นๆ โดยหลักการทำงานส่วนใหญ่จะใช้ Deep Learning ในการประมวลผล และนำข้อมูลจากการป้อนของผู้ใช้งานผ่านรูปแบบต่างๆ ส่งให้กับ AI ทำการเรียนรู้ ซึ่งจะเรียนรู้ว่าหากพบประโยคนี้แล้วควรตอบอย่างไร หรืออาจนำประโยคที่คล้ายกันมาตอบก็ได้ ซึ่งจะขึ้นอยู่กับภาษาที่มีการเปลี่ยนแปลงอยู่ตลอดเวลา

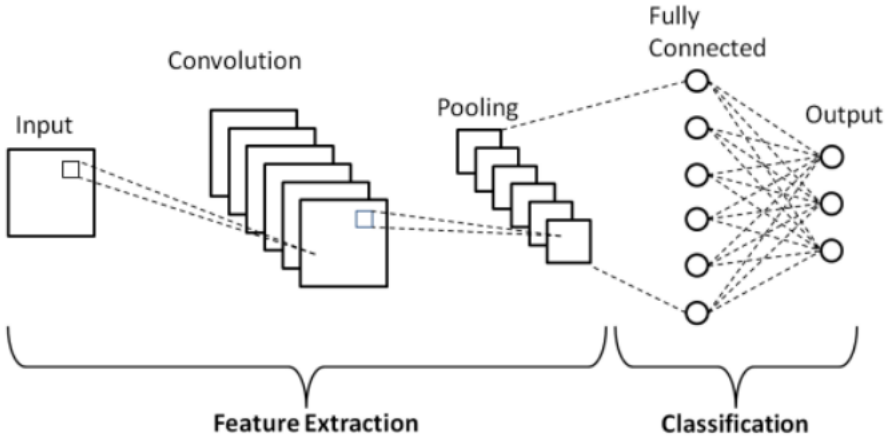
สำหรับช่วงระยะเวลาที่ผ่านมาในการเรียนรู้เชิงลึกได้รับความนิยมโดยนำมาใช้งานมากขึ้น อาทิเช่น งานทางด้านคอมพิวเตอร์วิชัน (Computer Vision) การเรียนรู้และจดจำใบหน้า (Face Recognition) และการรู้จำเสียง (Speech Recognition) เมื่อนำมาเปรียบเทียบกับ Machine Learning รูปแบบอื่น ๆ พบว่าข้อดีของ Deep Learning สามารถหาความสัมพันธ์ของข้อมูลที่มีรูปแบบที่แตกต่างกันได้ทันทีโดยไม่จำเป็นต้องแยกหมวดหมู่ ในส่วนของข้อเสียต้องใช้เวลาในการ “เรียนรู้” กับชุดข้อมูลที่มีความซับซ้อนเป็นเวลานาน รวมถึงต้องมีอุปกรณ์รองรับกับชุดข้อมูลจำนวนมากๆ จึงจำเป็นที่จะต้องมีผู้ดูแลระบบที่มีความรู้และความเข้าในการออกแบบการทำงานของ Deep Learning

2.2.1 โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution Neural Network: CNN)

โครงข่ายประสาทเทียมแบบคอนโวลูชัน เป็นโครงข่ายประสาทเทียมที่จำลองการมองเห็นของมนุษย์ที่มองพื้นที่เป็นส่วนย่อยๆ ก่อนนำกลุ่มพื้นที่ย่อยนั้นมาผสมผสานรวมกัน เพื่อตรวจสอบว่าสิ่งที่ต้องการนั้นคืออะไร และพื้นที่ย่อยนั้นจะมีการแยกคุณลักษณะ (Feature) ที่แตกต่างกัน ในการจำลองสามารถทำได้จากกระบวนการ Convolution เพื่อทำให้ภาพเล็กลงด้วยการกรอง ก่อนเพิ่มสัญญาณรบกวน (Noise) เข้าไป จากนั้นทำการหา Pattern ด้วยการเพิ่ม Pooling ในรูปแบบต่างๆ (Max Pooling หรือ Average) ให้กับภาพ ก่อนนำผลลัพธ์ที่ได้ (Output) ส่งต่อไปยังโมเดล Neural Network เพื่อสร้างกระบวนการเรียนรู้และการปรับค่าน้ำหนัก (Weight) และสุดท้ายนำ Feature แต่ละส่วนมาผสมผสานกัน ซึ่งกระบวนการเรียนรู้จะถูกทำซ้ำในหลายๆ รอบ เพื่อเป็นการปรับค่าพารามิเตอร์และลดค่าความผิดปกติ (Error) ในการทำนายของแต่ละรอบ

จากนั้นในปี 1989 LeChu และคณะ [12] ได้พบปัญหาที่เกิดขึ้นเกี่ยวกับตัวอักษรจึงได้นำ CNN เข้ามาช่วยแก้ปัญหาด้วยการจดจำตัวอักษรที่เป็นลายมือ เพื่อแยกลักษณะและขนาดของตัวอักษรรวมถึงการประมวลผลสัญญาณดิจิทัลจากการเรียนรู้แบบอัตโนมัติ หลังจากนั้น CNN ถูกนำเสนอในรูปแบบของสถาปัตยกรรมต่าง ๆ เนื่องจากมีการเรียนรู้คุณลักษณะที่มีประสิทธิภาพและความสามารถในการแสดงคุณลักษณะที่มากกว่าวิธีการเรียนรู้ของเครื่องแบบเดิม [13] จึงถูกนำไปประยุกต์ใช้ในงานด้านต่างๆ มากขึ้น โดยเฉพาะทางด้านการรู้จำรูปภาพ ตัวอักษร โดยมีผู้วิจัยที่ทำสนใจในเรื่องนี้ตั้งแต่อดีตจนถึงปัจจุบันประกอบด้วย งานวิจัยเกี่ยวกับการจดจำตัวอักษรโดยการเขียนด้วยลายมือ [14-16] การจดจำภาพ [17-19] และการรู้จำใบหน้า [20-22]

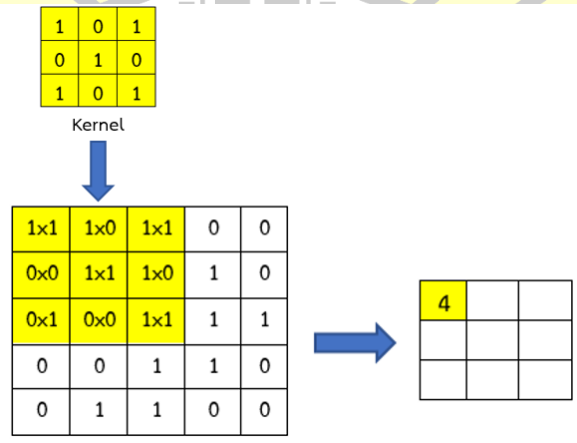
ในสถาปัตยกรรมของ CNN เป็นแนวคิดที่มีความซับซ้อน ดังนั้นจึงจำเป็นต้องใช้การคำนวณทางคณิตศาสตร์เข้ามารองรับ เพื่อให้เกิดความสอดคล้องกับแนวคิดนั้นๆ โดยสามารถแบ่งการทำงานออกเป็น 2 ส่วน ประกอบด้วย การสกัดคุณลักษณะ (Feature Extraction) และการจำแนกประเภท (Classification) ซึ่งแสดงให้เห็นในส่วนประกอบของสถาปัตยกรรมแบบ CNN ดังภาพที่ 2.5



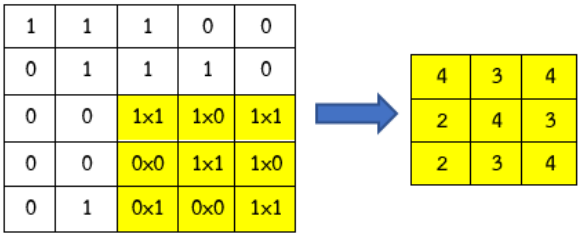
ภาพที่ 2.5 ส่วนประกอบของสถาปัตยกรรมแบบ Convolution Neural Network: CNN [23]

รูปแบบการทำงานของการสกัดคุณลักษณะของภาพประกอบด้วย 3 ส่วน ดังต่อไปนี้

1. Feature Extraction เป็นส่วนที่ช่วยในการกำหนดค่าในตัวกรอง (Filters) หรือ เคอร์เนล (Kernel) โดยที่ CNN จะใช้ตัวกรอง หรือ Filters ในการกรองส่วนที่ไม่สำคัญออกจากภาพ ดังแสดงให้เห็นจากตัวอย่างการทำงานของ Filters ซึ่งมีขนาด (K) 3x3 โดยใช้ Kernel ไปประมวลผลกับพิกเซลบนภาพทีละส่วน และเลื่อนไปต่อเรื่อย ๆ จนครบทึกส่วนในภาพ เพื่อสร้างชุดข้อมูลใหม่เรียกว่า ผังคุณลักษณะ (Feature Maps) ซึ่งแทนค่าเป็น $(I * K)$ ที่มีลักษณะต่างกันจากข้อมูลเดิม ดังภาพที่ 2.6 และ ภาพที่ 2.7



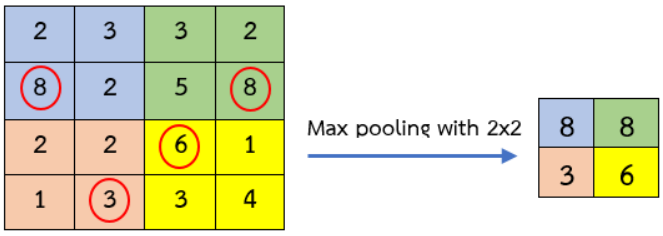
ภาพที่ 2.6 การทำงานของตัวกรองเพื่อสร้างชุดข้อมูลใหม่โดยการคำนวณทีละจุด [24]



ภาพที่ 2.7 ผลลัพธ์ที่ได้จากการนำค่า Kernel มาคำนวณเพื่อสร้างชุดข้อมูลใหม่ [24]

จากภาพที่ 2.6 และภาพที่ 2.7 ผลลัพธ์ที่ได้จากการนำค่า Kernel มาคำนวณเพื่อสร้างชุดข้อมูลใหม่แสดงผลลัพธ์ที่ได้จากการนำค่า Kernel มาประมวลผลกับภาพเบลอจะเห็นได้ว่ามีการกรองภาพบางส่วนออกจากภาพและเหลือส่วนที่ต้องการไว้ จึงทำให้ได้ภาพผลลัพธ์จากการประมวลผลมีขนาดเล็กลง

2. Pooling คือความสามารถในการย่อรูปแบบหนึ่ง ซึ่งมีอยู่สองประเภทที่นิยมนำมาใช้งาน คือ Max Pooling และ Average Pooling สำหรับ Max Pooling จะเป็นตัวกรองรูปแบบการหาค่าสูงสุดในบริเวณที่ตัวกรองทาบอยู่มาเป็นผลลัพธ์ เพื่อลดขนาดของ Feature MAPs ด้วยการหาค่าเฉลี่ย (Average Pool) และเลือกค่าที่สูงที่สุดบนตัวกรองนั้นมาเป็นผลลัพธ์ใหม่ และจะเลื่อนตัวกรองไปตาม Stride ที่กำหนดไว้ ซึ่งขนาดตัวกรองของการทำ Max Pooling เรียกกันว่า Pool Size

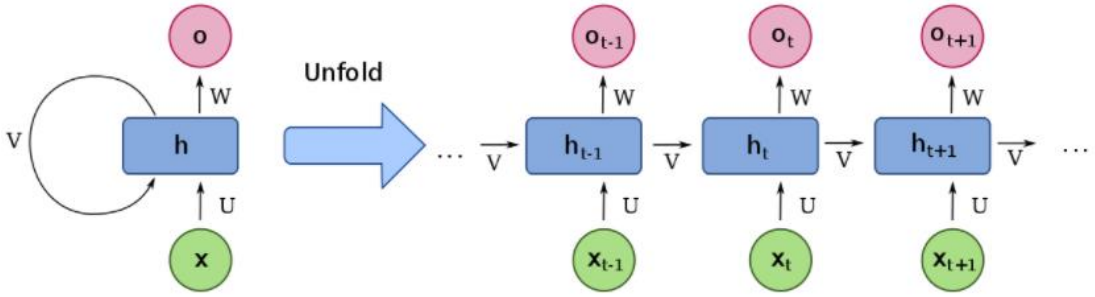


ภาพที่ 2.8 ตัวอย่างการทำ Pooling ด้วยวิธี Max Pooling [24]

จากภาพที่ 2.8 แสดงการทำ Max Pooling กำหนดให้แต่ละชุดสีใน Feature Maps มีขนาด 4x4 ซึ่งจะแทนบริเวณ Sub-Region โดยการเลือกค่าที่มากที่สุดในแต่ละ Sub-Region เพื่อนำมาสร้าง Feature Maps ใหม่ที่มีขนาด 2x2 พิกเซล

2.2.2 Recurrent Neural Networks (RNNs)

แม้ว่า CNNs จะเป็นวิธีที่เหมาะสำหรับการแก้ไขปัญหาทางด้าน Computer Vision แต่ก็ยังมีอีกหลายวิธีที่ช่วยในการแก้ปัญหาดังกล่าว ดังนั้น Recurrent Neural Networks (RNNs) จึงเข้ามาเป็นอีกทางเลือกหนึ่งที่มีบทบาทในการแก้ปัญหาด้าน Computer Vision โดย RNNs เป็นโครงข่ายประสาทเทียมที่มีลักษณะแบบ Feedforward ที่มีหน่วยความจำภายใน ซึ่งออกแบบมาสำหรับการแก้ไขปัญหาข้อมูลที่มีลำดับ Sequence โดยใช้หลักการ Feed สถานะภายในของโมเดลให้กลับมาเป็น Input ใหม่คู่กับ Input ปกติ RNNs จะเกิดขึ้นซ้ำก็ต่อเมื่อทุก Input ของข้อมูลที่ทำหน้าที่เดียวกัน หลังจากสร้าง Output แล้วจะถูกคัดลอกและส่งกลับไปยังเครือข่ายที่เกิดซ้ำ เพื่อช่วยให้โมเดลสามารถรู้จำ Pattern ของลำดับ Input Sequence ได้



ภาพที่ 2.9 โครงข่ายประสาทเทียมแบบเกิดซ้ำ Recurrent Neural Networks (RNNs) [25]

จากภาพที่ 2.9 จะเห็นได้ว่าภาพทางซ้ายมีการบีบอัดรูปแบบการทำงานและเมื่อนำมากระจายออกทางด้านขวาจะแสดงให้เห็นถึงการทำงานของ RNNs ในแต่ละรอบเมื่อมีการทำซ้ำ โดยกำหนดให้ค่า U, V และ W เป็นน้ำหนักของเครือข่าย เริ่มจาก X_{t-1} เป็นลำดับของการ Input จากนั้นจึงส่ง O_{t-1} และ X_t ออกไปพร้อมกันเพื่อเป็น Input ในขั้นตอนต่อไป และในรอบต่อไปก็จะใช้ X_t เป็นตัว Input และส่ง O_t และ X_{t+1} เพื่อสร้างเป็น Input ไปเรื่อยๆ จนครบ ระบบก็จะหยุดทำงาน

นอกจากนี้ RNNs ยังมีประโยชน์ที่สามารถสร้างแบบจำลองให้ใช้ได้ทั้งในระยะสั้นและระยะยาวระหว่างพิกเซล เพื่อปรับปรุงการประมาณค่าการแบ่งกลุ่ม อีกทั้งยังสามารถจัดลำดับของข้อมูลเพื่อให้สามารถมองเห็นถึงลำดับก่อนหน้า และใช้ในการขยายขอบเขตของพิกเซลด้วยเลเยอร์แบบคอนโวลูชันได้

2.3 ภาพเบลอ (Blur Image)

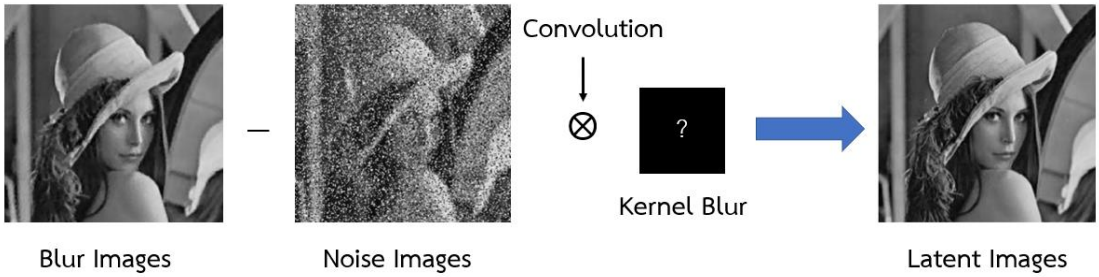
ภาพเบลอที่พบในปัจจุบันเกิดจากการถ่ายภาพด้วยกล้องในลักษณะต่างๆ ซึ่งเกิดจากกล้องขนาดเล็กและมีความละเอียดสูง (DSLR) หรือกล้องจากโทรศัพท์มือถือที่กำลังได้รับความนิยมในปัจจุบัน ซึ่งทั้งสองชนิดนี้จะมีลักษณะเบาและควบคุมค่อนข้างยาก หากไม่ได้รับการฝึกฝนให้มีการถ่ายภาพอย่างมืออาชีพ [26] ซึ่งหากการถ่ายภาพมีการเปิดรับแสงเป็นเวลานานๆ ภาพที่ได้จะออกมาในรูปแบบที่มืดและมีสัญญาณรบกวนปะปนเข้ามาในภาพ อีกทั้งการถ่ายภาพที่อยู่ในสภาวะแสงน้อยจะส่งผลให้ภาพที่ได้เป็นภาพเบลอและไม่สามารถใช้งานได้ตามที่ต้องการ [27] ดังนั้นความเบลอจากการสั่นของกล้องส่วนใหญ่มักเกิดจากการหมุนกล้องในรูปแบบ 3 มิติ ส่งผลให้ Kernel Blur ไม่มีความสม่ำเสมอทั่วทั้งภาพ ฉะนั้นจึงต้องปรับค่าภาพแฝง (Latent Images) ที่ซ่อนอยู่ในภาพชัด (Latent Sharp Images) ให้มีการเบลอลดลงเพื่อทำให้ภาพชัดเจนขึ้น โดยใช้วิธีการนำภาพเบลอ (Blur Images) มาลบสัญญาณรบกวน (Noise) ก่อนนำมา Convolution กับ Kernel ที่สร้างขึ้นเพื่อให้ได้ภาพแฝงที่มีความคมชัดขึ้น [28] ดังภาพที่ 2.10 ซึ่งสาเหตุที่ทำให้การเบลอของภาพอาจมาจากหลายสาเหตุดังนั้นจึงถูกจำแนกออกเป็นประเภทต่าง ๆ ได้ดังนี้

1. กล้องสั่น (Camera Shake) เกิดขึ้นในช่วงเวลาที่มีการเปิดรับแสง เช่น การเปิดรับแสงเป็นเวลานานๆ ก่อนกดชัตเตอร์จึงทำให้เกิดการสั่น หรืออาจเกิดจากตากล้อง เป็นต้น

2. การเคลื่อนไหวของวัตถุ (Object Movement) เกิดจากการเคลื่อนไหวของวัตถุในช่วงเวลาที่มีการเปิดรับแสงนานเกินไป เช่น การเคลื่อนที่ของนักกีฬา การเคลื่อนไหวของยานต์พาหนะ

3. การโฟกัสของกล้อง (Out of Focus or Defocus Blur) เกิดจากความชัดลึกของกล้องที่จำกัดจึงทำให้ไม่สามารถโฟกัสภาพได้ หรืออาจจะเกิดจากการผู้ใช้งานตั้งใจให้ได้ภาพเบลอเพื่อนำไปใช้งานในลักษณะอื่นๆ

4. การผสมผสาน ระหว่างการสั่นและการเคลื่อนไหว (Combinations between Vibration & Motion) เกิดจากการผสมผสานกันระหว่างกล้องสั่นหรือการเคลื่อนที่ของวัตถุ เช่น กล้องวงจรปิดบนแยกจราจร เมื่อมีรถวิ่งผ่านจุดที่ติดตั้งกล้องและเกิดการสั่นไหวของพื้นผิวจราจรทำให้กล้องจับภาพวัตถุที่ผ่านได้ไม่ชัด ส่งผลให้เกิดภาพเบลอเกิดขึ้น



ภาพที่ 2.10 วิธีการปรับปรุงภาพแฝง (Latent Images)

2.4 การลดสัญญาณรบกวนภาพ (Image Noise Reduction)

สำหรับภาพดิจิทัลที่ใช้งานจริงบางครั้งภาพต้นฉบับที่จัดเก็บมาได้อาจจะมีสัญญาณรบกวน (Noise) มาปรากฏทับซ้อนบนค่าระดับความเข้มเทาของจุดภาพ จึงมีความจำเป็นอย่างยิ่งที่จะต้องลดสัญญาณรบกวนออกจากภาพ ซึ่งเป็นอีกหนึ่งขั้นตอนในกระบวนการประมวลผลภาพเพื่อให้ภาพชัดเจนขึ้น นอกจากนี้สัญญาณรบกวนยังเป็นข้อมูลภาพที่อยู่ในช่วงความถี่สูง ดังนั้นในการขจัดสัญญาณรบกวนจะต้องใช้ตัวกรองที่มีความถี่ต่ำผ่าน (Low-Pass Filtering) มาทำการลดหรือขจัดสัญญาณรบกวนออกจากภาพ [26, 29-31] จากงานวิจัยดังกล่าวได้นำวิธีการลดสัญญาณรบกวนด้วยการกรองส่วนที่ไม่ต้องการออกจากภาพ ซึ่งวิธีการลดสัญญาณรบกวนที่นิยมใช้มีอยู่ 2 วิธีคือ การลดสัญญาณรบกวนแบบเชิงเส้นโดยอาศัยวิธีการคอนโวลูชัน (Convolution) และการลดสัญญาณรบกวนแบบไม่เชิงเส้น

2.4.1 การลดสัญญาณรบกวนแบบเป็นเชิงเส้น (Linear Filtering)

ในการลดสัญญาณรบกวนออกจากภาพประเภทที่มีความถี่สูง จะต้องอาศัยหลักการหาค่าเฉลี่ยความเข้มแสงเฉพาะบริเวณ อีกแบบหนึ่งคือการกรองสัญญาณความถี่ต่ำผ่าน (Low-Pass Filtering) ผลกระทบของการกรองด้วยสัญญาณความถี่ต่ำผ่านจะทำให้สัญญาณรบกวนลดลง ส่งผลให้ได้ภาพผลลัพธ์ที่มีความเรียบ (Smooth) การลดสัญญาณรบกวนออกจากภาพในบางกรณีที่ได้ผลลัพธ์เป็นภาพเบลอ (Blurring Effect) จะมีความคมชัดน้อยลง เนื่องจากขอบของวัตถุในภาพจะเป็นบริเวณที่มีการเปลี่ยนแปลงระดับค่าความเข้มแสง ดังนั้นสัญญาณความถี่สูงจึงจะถูกกรองออกไปจากเทคนิคการลดสัญญาณรบกวนส่วนมากจะเน้นไปในเรื่องของการลดสัญญาณรบกวนแต่จะไม่ทำลายขอบของวัตถุ มาร์คที่ใช้ในการลดสัญญาณรบกวนจึงมีลักษณะค่าสัมประสิทธิ์ทุกตำแหน่งของตัวกรองเป็นบวกทั้งหมด และผลรวมจะมีค่าเท่ากับหนึ่ง

ประเภทของตัวกรองที่ใช้ในการกรองสัญญาณรบกวนแบบเป็นเชิงเส้น ถูกแบ่งออกเป็น 2 ประเภท ได้แก่ ตัวกรองแบบค่าเฉลี่ย (Moving Averaging Filter) และตัวกรองแบบเกาส์เซียน (Gaussian Filter) ซึ่งมีรายละเอียดดังนี้

1. ตัวกรองแบบค่าเฉลี่ย (Moving Averaging Filter)

เป็นตัวกรองความถี่แบบค่าเฉลี่ยต่ำผ่านชนิดหนึ่งที่มีผลรวมของค่าสัมประสิทธิ์ของตัวกรองเป็นหนึ่ง โดนมาร์คของตัวกรองนี้จะกรองสัญญาณรบกวนแบบเกาส์เซียนติดมากับตัวกรองแบบเฉลี่ย โดยจะเลือกขนาดของมาส์คตังกรองที่แตกต่างกัน โดยการกรองงประเภทนี้จะมีทั้งข้อดีและข้อเสียที่แตกต่างกัน เช่น หากมีการเพิ่มขนาดจะสามารถช่วยลดสัญญาณรบกวนได้มากขึ้น และถ้าขนาดของมาส์คมีขนาดใหญ่ ปริมาณของสัญญาณรบกวนก็จะถูกลบได้มากขึ้นเช่นกัน แต่ภาพผลลัพธ์จะมีลักษณะการเบลอมากขึ้น

2. ตัวกรองแบบเกาส์เซียน (Gaussian Filter)

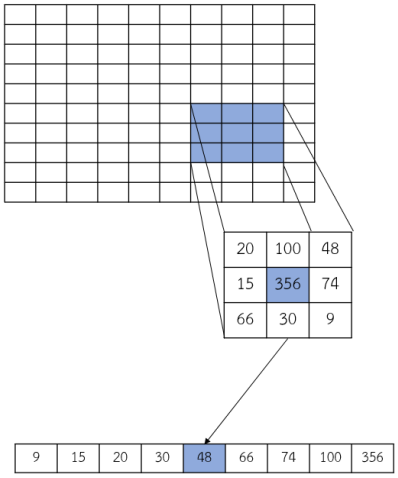
โดยทั่วไปแล้ว การลดสัญญาณรบกวนแบบเกาส์เซียนจะเป็นที่นิยมใช้งานมากกว่าตัวกรองแบบค่าเฉลี่ย เนื่องจากมีผลกระทบต่อการเบลอ (Blur) ของภาพต้นฉบับน้อยกว่า และยังสามารถออกแบบและหาค่าสัมประสิทธิ์ของตัวกรองได้หลายรูปแบบ วิธีการออกแบบตัวกรองชนิดนี้จะใช้ขนาดของมาส์คและค่า σ ที่แตกต่างกัน ส่งผลให้สัญญาณรบกวนลดน้อยลงและทำให้ได้ภาพที่มีความเรียบหรือการเบลอมากขึ้น เมื่อ σ มีค่าเพิ่มขึ้น ในขณะเดียวกันหากขนาดของตัวกรองแบบเกาส์เซียนมีขนาดใหญ่มากขึ้น จะส่งผลให้ความเรียบและความเบลอของภาพผลลัพธ์เพิ่มมากขึ้น [32]

2.4.2 การลดสัญญาณรบกวนแบบไม่เป็นเชิงเส้น (Nonlinear Filtering)

สัญญาณรบกวนประเภทนี้ไม่สามารถใช้ตัวกรองหรือฟิลเตอร์แบบเชิงเส้น (Linear Filtering) ในการลดสัญญาณรบกวนเพื่อให้ได้ผลลัพธ์ที่สมบูรณ์ได้ ดังนั้นจึงใช้การลดสัญญาณรบกวนแบบ Salt-and-Pepper-Noise กับภาพดิจิทัลด้วยตัวกรองแบบไม่เป็นเชิงเส้น (Nonlinear Filtering) ด้วยวิธีการกรองแบบมัธยฐาน (Median Filter) และตัวกรองแบบ Minimum and Maximum Filter โดยมีรายละเอียดดังต่อไปนี้

1. ตัวกรองแบบมัธยฐาน (Median Filter)

กรองแบบมัธยฐาน (Median Filter) เป็นตัวกรองที่อาศัยข้อมูลทางสถิติของข้อมูลภาพโดยใช้ค่ามัธยฐาน (Median) การหาค่ามัธยฐานทำได้โดยการนำข้อมูลระดับความเข้มเทาของภาพบริเวณที่ Mask ครอบคลุมอยู่ มาทำการการเรียงค่าจากน้อยไปหามากตามระดับความเข้มเทาของข้อมูลภาพ ซึ่งค่ามัธยฐานจะเป็นค่ากึ่งกลางของกลุ่มข้อมูล จากนั้นนำค่ามัธยฐานที่ได้แทนกลับไปยังตำแหน่งตรงกลางของ Mask ดังนั้นตัวกรองแบบนี้จะไม่ใช้เทคนิคการคอมโวลูชัน (Convolution) แต่จะใช้ Mask ขนาดต่างๆ ไปวางทาบกับภาพต้นฉบับ โดยนำจุดกึ่งกลางของ Mask ไปวางซ้อนทับกับจุดพิกัด (x, y) ของภาพต้นฉบับแล้วนำค่าระดับความเข้มเทาของภาพที่บริเวณ Mask ครอบคลุมอยู่มาทำการเรียงค่าจากน้อยไปหามากตามระดับความเข้มเทาของภาพก่อนนำค่าที่อยู่กึ่งกลางมาพิจารณา ดังภาพตัวอย่างต่อไปนี้

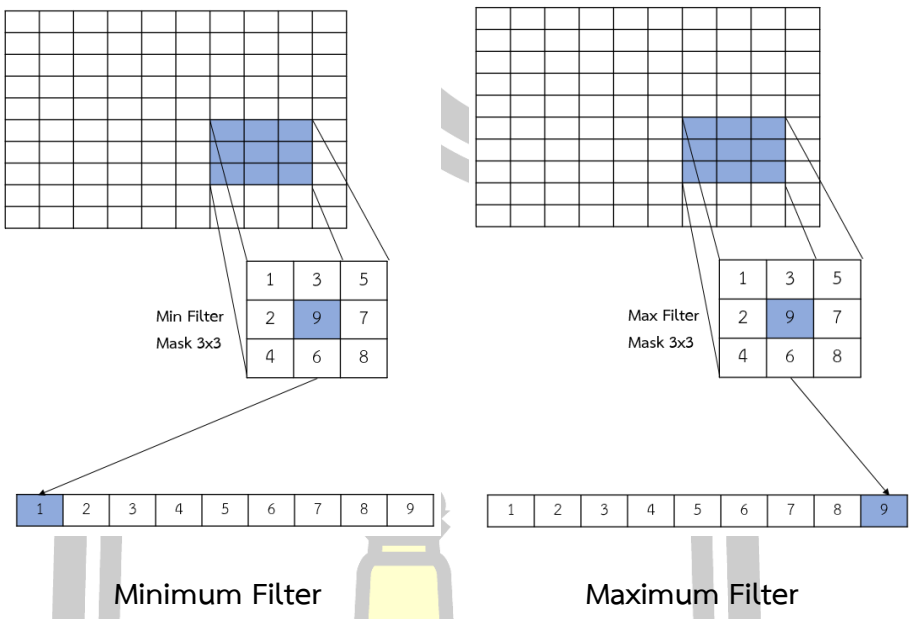


ภาพที่ 2.11 แสดงการกรองแบบมัธยฐาน (Median Filter) [33]

จากภาพเมื่อนำมาร์คที่มีขนาด 3x3 ไปวางทาบกับภาพต้นฉบับ โดยนำจุดกึ่งกลางของ Mask ไปวางซ้อนทับกับจุดพิกัด (x, y) แล้วนำค่าระดับความเข้มเทาของภาพที่บริเวณ Mask ครอบคลุมอยู่มาทำการเรียงค่าจากน้อยไปหามาก ในภาพแสดงให้เห็นว่าเมื่อนำข้อมูลมาเรียงจะได้ค่า 48 ซึ่งเป็นค่าที่ได้จากค่าเฉลี่ยการกรองแบบมัธยฐานแล้วจึงนำค่าที่ได้ไปใช้ในกระบวนการกรองต่อไป

2. ตัวกรองแบบ Minimum and Maximum Filter

Minimum และ Maximum Filter เป็นตัวกรองแบบไม่เป็นเชิงเส้นที่มีความคล้ายคลึงกับการกรองแบบมัธยฐาน โดยการกรองแบบ Minimum Filter และ Maximum Filter จะถูกแทนที่ต้องค่าต่ำสุดและสูงสุด โดยแทนค่าผลลัพธ์ลงไปในตำแหน่งตรงกลางของ Mask ซึ่งตัวกรองประเภทนี้จะนิยมนำไปใช้ในการลดสัญญาณรบกวนแบบอิมพัลส์ (Impulse Noise) โดยที่ Minimum Filter ลบจุดขาวของภาพ ส่วน Maximum Filter จะลบจุดดำของภาพ ทั้งนี้ในการใช้วิธีการกรองแบบ Minimum Filter และ Maximum Filter สามารถนำไปใช้กับภาพที่เป็นไดอะแกรมหรือรูปลายเส้น (แยกสีดำออกจากพื้นสีขาว) ซึ่งสามารถช่วยลดผลกระทบของสัญญาณรบกวนและทำให้รายละเอียดของลายเส้นเพิ่มมากขึ้น



ภาพที่ 2.12 ตัวกรองแบบ Minimum Filter และ Maximum Filter [33]

จากภาพจะเห็นได้ว่าตัวกรองทั้ง 2 แบบนั้นใช้การเรียงลำดับค่าความเข้มเทาคล้ายกับการกรองค่ามัธยฐาน แตกต่างกันตรงที่ Minimum Filter และ Maximum Filter ที่มีขนาด 3x3 จะใช้ค่าต่ำสุดและค่าสูงสุดจากการเรียงลำดับไปใช้ในการพิจารณา

2.5 รูปแบบของสัญญาณรบกวน (Noise Models)

การจำลองลักษณะของสัญญาณรบกวนสำหรับงานด้านการประมวลผลภาพ ส่วนใหญ่จะนิยมจำลองรูปแบบของสัญญาณรบกวนให้อยู่ในรูปแบบของฟังก์ชันความหนาแน่นของความน่าจะเป็น (Probability Density Function) ในทางสถิติต่างๆ สัญญาณรบกวนจะเกิดจากปรากฏการณ์ทางธรรมชาติในขั้นตอนการเปลี่ยนสัญญาณแสงเป็นสัญญาณไฟฟ้า ดังนั้นสมมติฐานของสัญญาณรบกวนที่นิยมนำมาใช้จึงถูกแบ่งออกเป็น 3 ประเภท ได้แก่ สัญญาณรบกวนแบบสม่ำเสมอ (Uniform Noise) สัญญาณรบกวนแบบอิมพัลส์ (Impulse Noise) [29] และสัญญาณรบกวนแบบเกาส์เซียน (Gaussian Noise) โดยมีรายละเอียดดังนี้

2.5.1 สัญญาณรบกวนแบบสม่ำเสมอ (Uniform Noise)

สัญญาณรบกวนประเภทนี้จะมีรูปแบบลักษณะการกระจายตัวของฟังก์ชันความหนาแน่นของความน่าจะเป็น (PDF)

$$P(z) = \begin{cases} \frac{1}{b-a}, & \text{ถ้า } a \le z \le b \\ 0, & \text{ที่เป็นอย่างอื่น} \end{cases} \quad (2.2)$$

โดยค่าเฉลี่ย (μ) ของฟังก์ชันความหนาแน่น (Density Function) จะถูกกำหนดด้วยสมการ (2.2) และค่าความแปรปรวน σ^2 ถูกกำหนดด้วยสมการที่ 2.4 รูปแบบของสัญญาณรบกวนแบบสม่ำเสมอซึ่งแสดงไว้ดังสมการต่อไปนี้

$$\mu = \frac{a+b}{2} \quad (2.3)$$

$$\sigma^2 = \frac{(b-a)^2}{2} \quad (2.4)$$

2.5.2 สัญญาณรบกวนแบบอิมพัลส์ (Impulse Noise)

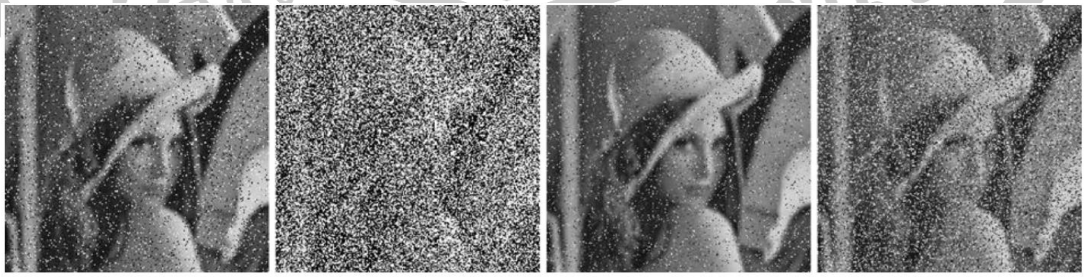
สัญญาณรบกวนประเภทนี้จะเป็นสัญญาณรบกวนที่กระจายอยู่บนพิกเซลของภาพเป็นจุดๆ ทั่วไป โดยเปลี่ยนค่าความเข้มแสงของพิกเซล ณ ตำแหน่งหนึ่งๆ ให้มีค่าความแตกต่างไปจากพิกเซลข้างเคียง [29] ซึ่งค่าความเข้มแสงที่จำลองส่วนมากจะเป็นความเข้มแสงสีขาวและสีดำหรือที่เรียกว่าแบบเกลือพริกไทย (Salt-and-Pepper-Noise) ดังนั้นภาพดิจิทัลบางตำแหน่งพิกเซลจะปนไปกับสัญญาณรบกวน คือพิกเซลบางตำแหน่งจะถูกเปลี่ยนเป็นจุดสีขาวหรือสีดำด้วยพิกเซลรบกวน (Noise Pixels) สัญญาณรบกวนแบบนี้จะมีการกระจายของฟังก์ชันความหนาแน่นของความน่าจะเป็น (PDF)

$$P(z)\begin{cases} Pa, & z = a \\ Pb, & z = b \end{cases} \quad (2.5)$$

เมื่อ a และ b คือ ระดับความเข้มเทาเดิม

Pa และ Pb คือ ค่าระดับความเข้มเทาใหม่

โดยปกติแล้วหากค่าของ $a > b$ ค่าระดับความเข้มเทาของ b จะถูกเปลี่ยนเป็นสัญญาณรบกวนที่มีจุดขาว และค่าความเข้มเทาเดิมของ a จะถูกเปลี่ยนเป็นสัญญาณรบกวนที่เป็นจุดดำ สำหรับภาพระดับความเข้มเทาแบบ 8 บิต (256 ระดับ) ค่า Pa จะมีค่าเท่ากับ 0 (สีดำ) และ Pb จะมีค่าเท่ากับ 255 (สีขาว)



ภาพที่ 2.13 แสดงผลการเบลอแบบ Salt-and-Pepper-Noise

ภาพที่ 2.14 แสดงผลการเบลอแบบ Random-Valued Impulse Noise

2.5.3 สัญญาณรบกวนแบบเกาส์เซียน (Gaussian Noise)

สัญญาณรบกวนประเภทนี้มีลักษณะการกระจายข้อมูลของสัญญาณรบกวนเป็นไปตามรูปแบบของฟังก์ชันความหนาแน่นของความน่าจะเป็นของฟังก์ชันเกาส์เซียน ซึ่งสัญญาณรบกวนนี้ได้รับความนิยมและถูกนำมาใช้จำลองในการเพิ่มสัญญาณรบกวนให้กับภาพ

$$\frac{1}{\sqrt{2\pi\sigma}}e - \frac{(z-\mu)}{2\sigma^2} \tag{2.6}$$

เมื่อ z แทนค่าระดับความเข้มเทา

μ แทนค่าเฉลี่ยทั้งหมดของ z

σ^2 แทนค่าความแปรปรวน (Variance)

s แทนค่าส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)

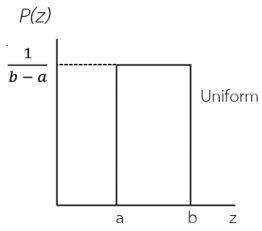
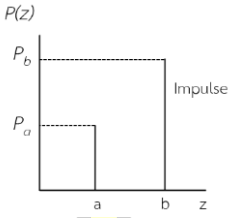
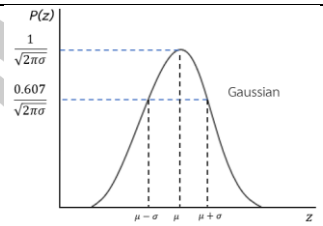
รูปแบบของสัญญาณรบกวนแบบเกาส์เซียนจะมีลักษณะของรูปสัญญาณคล้ายกับรูประฆังคว่ำโดยค่าสัญญาณรบกวนประมาณ 70% จะอยู่ในช่วง $[(\mu-\sigma),(\mu+\sigma)]$ และค่าสัญญาณรบกวนประมาณ 90% จะอยู่ในช่วง $[(\mu-2\sigma),(\mu+2\sigma)]$

สัญญาณรบกวนแบบเกาส์เซียน (Gaussian Noise) สามารถสร้างโมเดลการลดสัญญาณรบกวนจากสมการดังต่อไปนี้

$$B = D(I \otimes K) + N, \tag{2.7}$$

โดยที่ K เป็นเคอร์เนลเบลอ $N \sim N(0,\sigma^2)$ เป็นสัญญาณรบกวนและ σ^2 จะแปรผันตามพื้นที่ของภาพ เริ่มจากการสุ่มตัวอย่างของ $D(I)$ เพื่อทำการลดตัวอย่างภาพที่เกิดจากการสุ่มโดยที่ $I(m,n) = I(sm,sn)$ และ (s) เป็นอัตราการสุ่มตัวอย่างสำหรับพิกัดของพิกเซลจำนวนเต็ม (m,n) ดังนั้นฟังก์ชันของการสุ่มตัวอย่างจะช่วยให้สามารถหาค่าการสุ่มตัวอย่างขึ้นได้ โดยการแก้ไขบนตารางย่อยของพิกเซล ในการเบลอภาพปกติจะกำหนดค่า $(s=1)$ ส่วนการสุ่มตัวอย่างซ้ำจะกำหนดให้ $(s>1)$ จากสมการดังกล่าวจะถูกนำมาใช้ในการคำนวณเพื่อแก้ไขปัญหาการเบลอภาพแบบ

Deconvolution โดยใช้เฟรมเวิร์กของ Bayesian และค้นหาการประมาณค่าของภาพเบลอ อีกทั้งการลดระดับของสัญญาณรบกวนจะใช้เทคนิคของ Maximum a Posteriori (MAP) ในการหาผลลัพธ์ที่ได้แสดงให้เห็นว่ารูปแบบและวิธีการนี้สามารถปรับปรุงภาพและลดสัญญาณรบกวนได้ดี

		
สัญญาณรบกวนแบบสม่ำเสมอ (ก)	สัญญาณรบกวนแบบอิมพัลส์ (ข)	สัญญาณรบกวนแบบเกาส์เซียน (ค)

ภาพที่ 2.15 กราฟแสดงฟังก์ชันการกระจายข้อมูลของสัญญาณรบกวนแบบต่างๆ

2.6 เครื่องมือที่ในการวัดคุณภาพของภาพ (Performance Measurement)

ในการปรับปรุงภาพเพื่อให้เกิดภาพใหม่จำเป็นที่จะต้องมีการวัดและเปรียบเทียบคุณภาพของภาพเพื่อให้ได้ภาพที่มีมาตรฐานและมีประสิทธิภาพที่ดีขึ้น ซึ่งโดยทั่วไปแล้วเครื่องมือที่ใช้ในการวัดประสิทธิภาพของภาพนั้นมีมากมายหลายแบบ เช่น Contrast-to-Noise-Ratio (CNR), Mean Square Error (MSE), Signal-to-Noise-Ratio (SNR), Structural Similarity Index Measurement (SSIM), Peak Signal-to-Noise Ratio (PSNR) เป็นต้น สำหรับการวิจัยนี้ได้เลือกใช้เครื่องมือสำหรับ วัดประสิทธิภาพ 3 แบบได้แก่ Signal-to-Noise Ratio (SNR) [34], Peak Signal-to-Noise Ratio (PSNR) [7] และ Structural Similarity Index Measure (SSIM) [35] โดยมีรายละเอียดดังต่อไปนี้

1. Signal-to-Noise Ratio (SNR) เป็นเครื่องมือที่ใช้ในการวัดค่าอัตราส่วนสัญญาณต่อสัญญาณรบกวน คือ อัตราส่วนที่ใช้สำหรับวัดค่าระหว่างสัญญาณต้นฉบับกับสัญญาณรบกวน ซึ่งมีหน่วยที่ใช้วัดเป็นเดซิเบล Decibels (dB) โดยใช้เกณฑ์มาตรฐานในการวัดค่าสัญญาณรบกวนออกเป็นสัดส่วน ซึ่งสามารถคำนวณได้จากสมการต่อไปนี้

$$SNR_{dB} = 10\,log\,10\left(\frac{P_{signal}}{P_{noise}}\right) \quad (2.8)$$

2. Peak Signal-to-Noise Ratio (PSNR) เป็นเครื่องมือการประเมินประสิทธิภาพของเทคนิคการปรับปรุงคุณภาพของภาพที่ได้รับความนิยมมากที่สุด โดยค่าที่คำนวณได้นำไปเปรียบเทียบกับภาพต้นฉบับ ค่าที่ได้จะอยู่ในช่วง 0 – 100 ถ้าค่า PSNR สูงเข้าใกล้ 100 จะชี้ให้เห็นถึงคุณภาพ

ของภาพที่มีความใกล้เคียงกับภาพต้นฉบับ ซึ่งเป็นมาตรฐานที่นำมาใช้ในการเปรียบเทียบคุณภาพของรูปภาพดิจิทัล รวมทั้งการประเมินประสิทธิภาพของเทคนิคด้านความละเอียดและการกรองภาพ ซึ่งสามารถคำนวณได้จากสมการต่อไปนี้

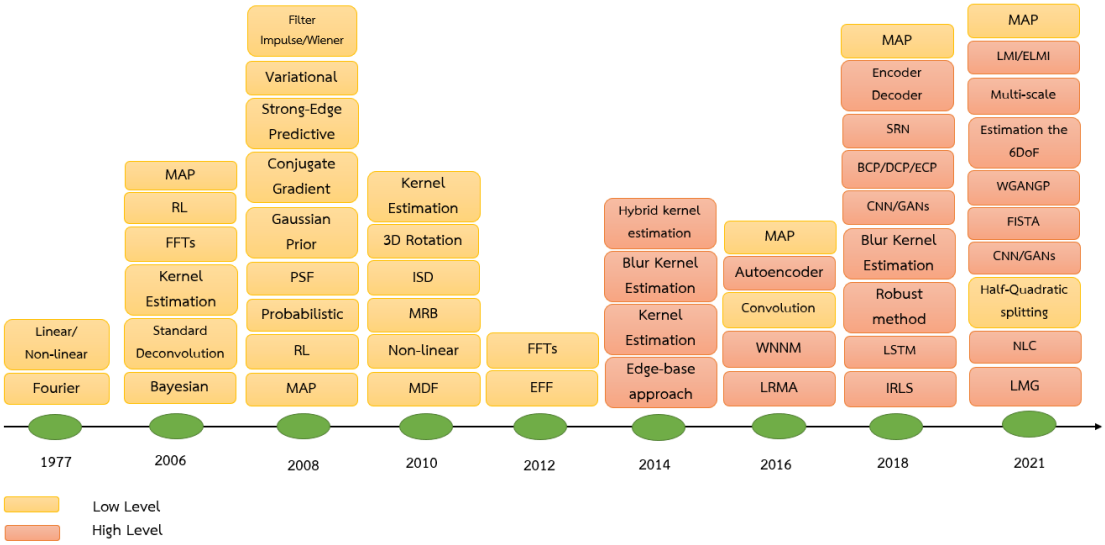
$$PSNR = 10\,log\,10\left(\frac{R^2}{MSE}\right) \tag{2.9}$$

3. Structural Similarity Index Measure (SSIM) เป็นเครื่องมือสำหรับการประเมินผลเพื่อใช้ในการวัดประสิทธิภาพของเทคนิคทางด้านภาพ โดยใช้ในการวัดโครงสร้างของภาพ และความคล้ายคลึงกันระหว่างสองภาพ ซึ่งค่าที่ได้จากการประมวลผลจะอยู่ระหว่าง 0 ถึง 1 ซึ่งสามารถคำนวณได้จากสมการต่อไปนี้

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{2.10}$$

2.7 งานวิจัยที่เกี่ยวข้อง

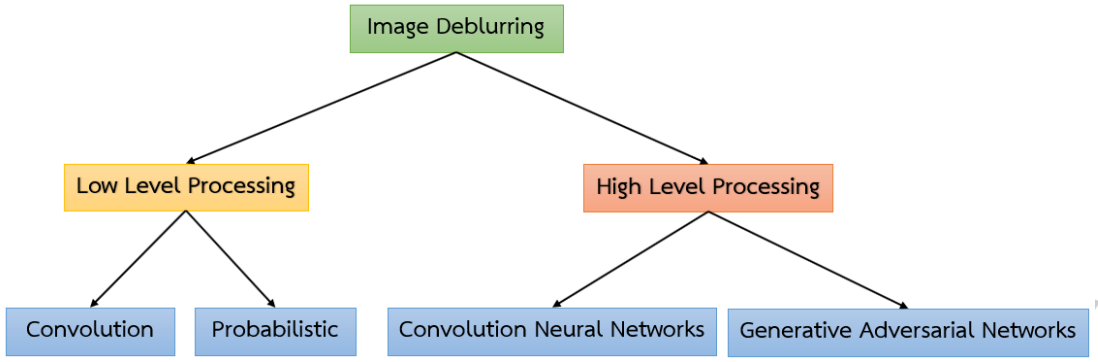
ในส่วนของงานวิจัยที่เกี่ยวข้องซึ่งเป็นการรวบรวมผลงานวิจัยที่เกี่ยวกับวิธีการลดความเบลอของภาพในลักษณะต่างๆ ที่เกิดขึ้นตั้งแต่อดีตจนถึงปัจจุบัน โดยใช้เทคนิคและอัลกอริทึมที่เกิดขึ้นจากการคิดค้นใหม่ๆ เข้ามาช่วยในการปรับปรุงภาพเบลอให้มีความชัดเจนยิ่งขึ้น ซึ่งในการปรับปรุงภาพเบลอ (Image Deblurring) จึงถือเป็นกระบวนการที่ใช้สำหรับลบสิ่งที่ผิดปกติออกจากภาพ ไม่ว่าจะเป็นจุดต่างๆ หรือแม้กระทั่งความเบลอที่เกิดจากการสั่นของกล้อง จากการสำรวจและศึกษาเกี่ยวกับงานวิจัยที่ผ่านมาในหลายๆ งานสามารถจำแนกได้เป็น 2 กลุ่มใหญ่ๆ คือ กระบวนการปรับปรุงภาพในระดับต่ำ (Low Level Processing) และกระบวนการปรับปรุงภาพในระดับสูง (High Level Processing) ดังภาพที่ 2.17



ภาพที่ 2.16 การปรับปรุงภาพเบลอตั้งแต่เริ่มต้นจนถึงปัจจุบัน

จากภาพที่ 2.16 แสดงให้เห็นถึงวิธีที่ใช้สำหรับการปรับปรุงภาพตั้งแต่เริ่มต้นจึงถึงปัจจุบัน โดยใช้เทคนิคและอัลกอริทึมต่างๆ เพื่อปรับปรุงภาพให้มีประสิทธิภาพมากขึ้น โดยใช้สีในการจำแนกกระบวนการในการปรับปรุงของแต่ละระดับ

ดังนั้นจึงนำกระบวนการปรับปรุงภาพในแต่ละระดับมาสร้างเป็นไดอะแกรมเพื่อให้เห็นถึงกระบวนการย่อยในแต่ละระดับที่ใช้วิธีการปรับปรุงภาพในรูปแบบที่แตกต่างกัน ดังภาพที่ 2.17



ภาพที่ 2.17 กระบวนการในการปรับปรุงภาพในระดับต่างๆ

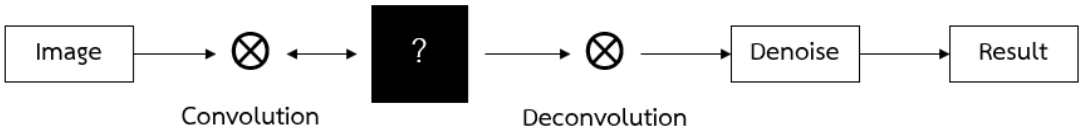
จากภาพที่ 2.17 แสดงให้เห็นถึงกระบวนการในการปรับปรุงภาพซึ่งในแต่ละระดับ โดยจะมีกระบวนการทำงานหรือวิธีการและเทคนิคต่างๆ ที่แตกต่างกัน มีรายละเอียดดังต่อไปนี้

กระบวนการปรับปรุงภาพระดับต่ำ Low Level Processing เป็นการประมวลผลในระดับต่ำ โดยจะใช้สำหรับจัดการข้อมูลกับ Pixel โดยตรง เพื่อใช้ในการปรับปรุงคุณภาพของภาพ ในการทำ Deblurring นั้นจะมีกระบวนการทำอยู่ 2 แบบ ได้แก่ กระบวนการทำ Convolutional เป็นการใช้

อัลกอริทึมในการปรับปรุงสัญญาณจากข้อมูลที่บันทึกไว้ด้วยโอเปอร์เรเตอร์ทางคณิตศาสตร์ จากนั้นใช้กระบวนการ Deconvolution ถอดรหัสเพื่อคืนค่าสัญญาณจากการปรับปรุงข้อมูล และในกระบวนการประมวลผลด้วยความน่าจะเป็น (Probabilistic) จะใช้อัลกอริทึมทางคณิตศาสตร์มาใช้ในการคำนวณทางด้านความน่าจะเป็น เพื่อทำให้การประมวณผลมีความรวดเร็วและแม่นยำขึ้น

กระบวนการปรับปรุงภาพระดับสูง High Level Processing เป็นกระบวนการที่ในการปรับปรุงภาพระดับสูง ซึ่งใช้วิธีการเรียนรู้ของเครื่อง (Machin Learning) เข้ามาช่วยในการปรับปรุงภาพ โดยอาศัยหลักการของการเรียนรู้เชิงลึก (Deep Learning) ในการปรับปรุงคุณภาพของภาพและเป็นการสร้างภาพขึ้นมาใหม่ (Deconstruct) จากการเรียนรู้เชิงลึกในการสร้างต้นแบบ (Model) ที่เหมาะสม โดยนำการประมวลผลด้วยโครงข่ายประสาทเทียม Convolutional Neural Networks (CNNs) และเครือข่ายแบบ Generative Adversarial Networks (GANs) มาใช้ในการประมวลผลเพื่อให้ได้ผลลัพธ์ที่มีประสิทธิภาพและสมบูรณ์มากขึ้น

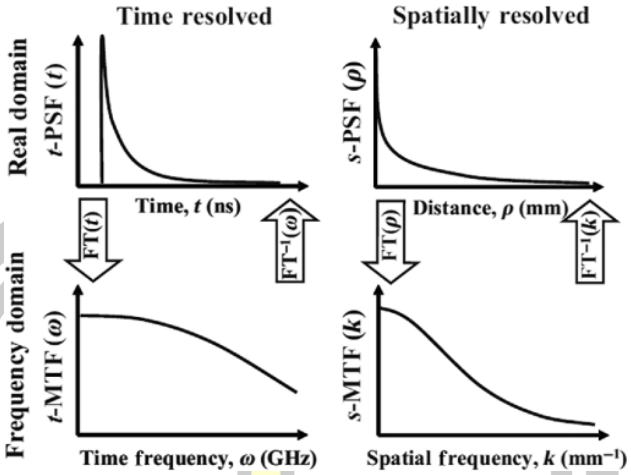
ในการปรับปรุงคุณภาพของภาพให้มีความคมชัดในกระบวนการทำงานแบบต่ำ ซึ่งหลักการทำงานดังกล่าวสามารถดำเนินการได้หลายวิธี โดยมีตัวอย่างงานวิจัยบางงานใช้วิธีการปรับปรุงขอบของวัตถุมีความเด่นชัดมากขึ้น หรือทำให้ส่วนของวัตถุที่เป็นพื้นผิวเรียบให้มีความเรียบเพิ่มมากขึ้น และในส่วนของการจำแนกภาพ [27] แต่หลักการนี้จะอยู่ภายใต้หลักการปรับปรุงคุณภาพของภาพ [30] โดยทั่วไปแล้วจะมีกระบวนการทำงานอยู่ 2 ขั้นตอน คือ กระบวนการในการสร้าง Kernel และกระบวนการทำ Convolution และ Deconvolution ดังภาพที่ 2.18



ภาพที่ 2.18 กระบวนการปรับปรุงภาพระดับต่ำ (Low Level Processing)

จากภาพที่ 2.18 แสดงให้เห็นถึงการปรับปรุงภาพระดับต่ำ Low Level เมื่อภาพที่นำเข้า (Input Image) ถูกนำมากรอง (Filters) สิ่งที่ไม่ต้องการออกจากตัวดำเนินการด้วยการ Convolution ภาพกับ Kernel จนครบทุกส่วนของภาพ จากนั้นนำผลลัพธ์ที่ได้มารวมกันด้วยวิธีการทำ Deconvolution กับภาพและนำค่าที่ได้มาลดสัญญาณรบกวนแล้วจึงได้ผลลัพธ์

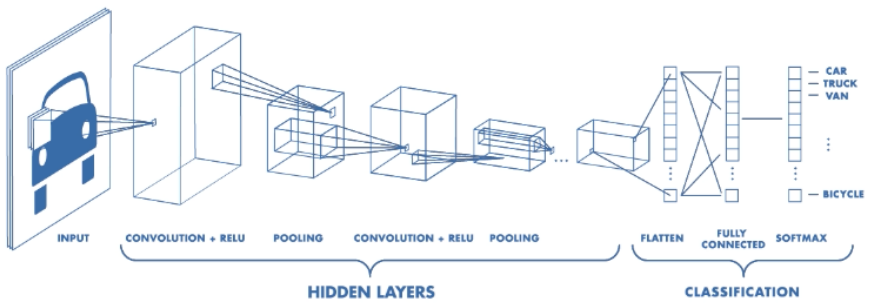
ในกระบวนการทำ Convolution และ Deconvolution จะมีกระบวนการทำงานย่อยออกเป็น 2 โดเมนประกอบด้วย Spatial/Discrete Domain และ Frequency Domain ดังภาพที่ 2.19



ภาพที่ 2.19 รูปแบบการทำงานของ Spatial/Discrete Domain และ Frequency Domain [36]

จากภาพที่ 2.19 แสดงการทำงานของ Discrete Domain เป็นการแปลงภาพออกมาให้อยู่ในรูปของตัวเลขก่อนนำมาคูณการ Kernel แล้วจะได้ผลลัพธ์ออกมา และในส่วนของ Frequency Domain เป็นการนำภาพที่ได้มาแปลงให้อยู่ในรูปสัญญาณคลื่นก่อนนำมา Convolution แล้วจึงได้ผลลัพธ์ออกมา

ในส่วนของการปรับปรุงคุณภาพของภาพให้มีความคมชัดในกระบวนการปรับปรุงภาพระดับสูง จะใช้กระบวนการเรียนรู้เชิงลึก (Deep Learning) เช่น CNN และ GAN ดังภาพที่ 2.17 โดยใช้อัลกอริทึมและวิธีการต่างๆ ซึ่งชุดข้อมูลที่นำมาใช้จะถูกแบ่งออกเป็น 2 ส่วน คือ ชุดข้อมูลสำหรับการฝึกฝน (Training Set) และชุดข้อมูลสำหรับการทดสอบ (Test Set) ตามอัตราส่วนที่ต้องการ อีกทั้งยังนำเครื่องมือที่ใช้ในการวัดประสิทธิภาพของภาพมาช่วยในการประเมินคุณภาพของภาพทั้งในเชิงคุณภาพและเชิงปริมาณ



ภาพที่ 2.20 กระบวนการทำงานแบบสูง (High Level) ด้วยวิธีการแบบ Convolutional Neural Networks : CNNs [24]

จากภาพที่ 2.20 แสดงให้เห็นถึงโครงสร้างการประมวลผลด้วยโครงข่ายประสาทแบบคอนโวลูชัน (CNN) เป็นกระบวนการทำงานแบบสูง โดยที่ CNN จะแยกภาพที่นำเข้า (Input Images)ออกเป็นพิกเซล (Pixel) ก่อนนำมาแยกออกเป็นเลเยอร์ต่างๆ 4 เลเยอร์ ประกอบด้วยเลเยอร์ที่ใช้ในการเก็บโทนสีและข้อมูลอื่นๆ 3 เลเยอร์และเก็บภาพขาวดำ 1 เลเยอร์ เพื่อใช้ในการ Training จากนั้นจะทำการเรียนรู้และจดจำข้อมูลจากเลเยอร์ต่างๆ ของโมเดลหรือที่เรียกว่า Parameter Estimationด้วยวิธีการทำซ้ำหลายๆ รอบก่อนนำไปเปรียบเทียบกับโมเดล Feature Map ของสีเพื่อให้เกิดความแม่นยำ เมื่อ Training เสร็จจะนำโมเดลที่ได้มาใช้ในการจำแนกหาค่าว่าตรงกันกับภาพที่นำเข้าหรือไม่ ดังนั้นในปัจจุบันจึงได้นำเอาวิธีการเรียนรู้ของเครื่อง (Machine Learning) โครงข่ายประสาทเทียม (Naturel Network: NN) การเรียนรู้เชิงลึก (Deep Learning) เข้ามาช่วยในกระบวนการเรียนรู้และสร้างแบบจำลองเพื่อพัฒนาให้เกิดประสิทธิภาพที่สูงขึ้น ซึ่งวิธีการเรียนรู้

จากวรรณกรรมที่ได้ศึกษาในงานวิจัยนี้จะเน้นในเรื่องของการออกแบบ Kernel ที่เหมาะสมเพื่อทำภาพให้คมชัด (Image Deblurring) โดยแบ่งออกเป็น 2 กลุ่ม หลักๆ ประกอบด้วย

1. กระบวนการปรับปรุงภาพระดับต่ำ (Low Level Processing) โดยใช้การประมาณค่าแบบ Convolutional และกระบวนการประมวลผลด้วยความน่าจะเป็น (Probabilistic)

2. กระบวนการปรับปรุงภาพระดับสูง (High Level Processing) โดยใช้การประมวลผลด้วยโครงข่ายประสาทเทียม Convolutional Neural Networks (CNNs) และเครือข่ายแบบ Generative Adversarial Networks (GANs ซึ่งจะมีงานวิจัยที่โดดเด่นหลายงานที่ใช้รูปแบบและวิธีดังกล่าว โดยมีรายละเอียดดังต่อไปนี้

2.7.1 กระบวนการปรับปรุงภาพระดับต่ำ (Low Level Processing) โดยใช้การประมาณค่าแบบ Convolutional และการปรับปรุงภาพด้วยความน่าจะเป็น (Probabilistic)

1. การปรับปรุงภาพโดยใช้การประมาณค่าแบบ Convolutional

ภาพเบลอยังคงเป็นปัญหาสำคัญที่มักเกิดขึ้นจากหลายสาเหตุ เช่น ภาพถ่ายที่อยู่ภายใต้สภาวะแสงน้อยหรือสถาณการณ์ที่มีแสงไม่เพียงพอ รวมถึงการเปิดรับแสงนานๆ [27], [30] มีการเคลื่อนไหวสัมพัทธ์ที่เกิดจากกล้องและฉาก [30, 32] การเคลื่อนที่ของวัตถุและอาจเกิดการสั่นของกล้อง โดยเฉพาะอย่างยิ่งการถ่ายภาพจากกล้องขนาดเล็กที่มีความละเอียดสูงและมีน้ำหนักเบา [26] จึงส่งผลให้ได้ภาพเบลอและเป็นที่น่าผิดหวัง รวมไปถึงการทำให้สูญเสียข้อมูลที่ไม่สามารถหลีกเลี่ยงได้ [37] ดังนั้นจึงได้มีการปรับปรุงภาพถ่ายหรือภาพวิดีโอเหล่านี้ให้มีความคมชัดขึ้น โดยใช้ทฤษฎีการฟื้นฟูภาพดิจิทัล (Image Restoration) ซึ่งเป็นส่วนหนึ่งของคอมพิวเตอร์วิชัน (Computer Visions) ซึ่งเป็นที่รู้จักกันอย่างแพร่หลายในระบบประมวลผลรูปภาพดิจิทัล (Digital Image Processing

System) มาใช้ในการปรับปรุงคุณภาพของภาพให้สูงขึ้นตามเป้าหมายที่ต้องการ ซึ่งได้รับความสนใจเป็นอย่างมากในการทำวิจัยเกี่ยวกับคอมพิวเตอร์วิทัศน์มาจนถึงปัจจุบันนี้

สำหรับวิธีการปรับปรุงภาพเบลอในช่วงแรกส่วนใหญ่จะใช้เทคนิคเชิงตัวเลขเป็นเครื่องมือในการกู้คืนภาพดิจิทัลกับวัตถุที่เสื่อมสภาพ และใช้กระบวนการสร้างภาพด้วยแบบจำลองการเสื่อมสภาพก่อนทำการฟื้นฟูภาพด้วยเทคนิคแบบ Fourier, Linear Algebraic และเทคนิค Non-Linear Algebraic ซึ่งเกิดขึ้นในปี ค.ศ.1977 ในงานวิจัยของ Andrews [38] ต่อมาพบว่าการปรับปรุงภาพเบลอยังมีอีกหลายรูปแบบที่สามารถนำมาใช้กับภาพได้หลายประเภท เช่น ภาพนิ่ง ภาพเคลื่อนไหวหรือวิดีโอ ภาพจากกล้องวงจรปิดและกล้องขนาดเล็ก อย่างไรก็ตามภาพเบลอที่เกิดขึ้นส่วนใหญ่จะเกิดได้หลายสาเหตุ อาทิเช่น เกิดจากอุปกรณ์ เกิดจากความตั้งใจเพื่อนำไปใช้งานในรูปแบบต่างๆ หรือเกิดจากสภาวะแวดล้อมต่างๆ ดังนั้นจึงถูกนำมาแบ่งประเภทออกเป็นลักษณะต่างๆ ได้ดังนี้ [39]

1. ภาพเบลอจากการสั่นของกล้อง Camera Shake (Camera Motion Blur)
2. ภาพเบลอจากการเคลื่อนไหวของวัตถุ Object Movement (Object Motion Blur)
3. ภาพเบลอที่ไม่อยู่ในโฟกัส Out of Focus (Defocus Blur)
4. ภาพเบลอที่เกิดจากการผสมกันทั้งการสั่นสะเทือนและการเคลื่อนไหว Combinations (Vibration & Motion)

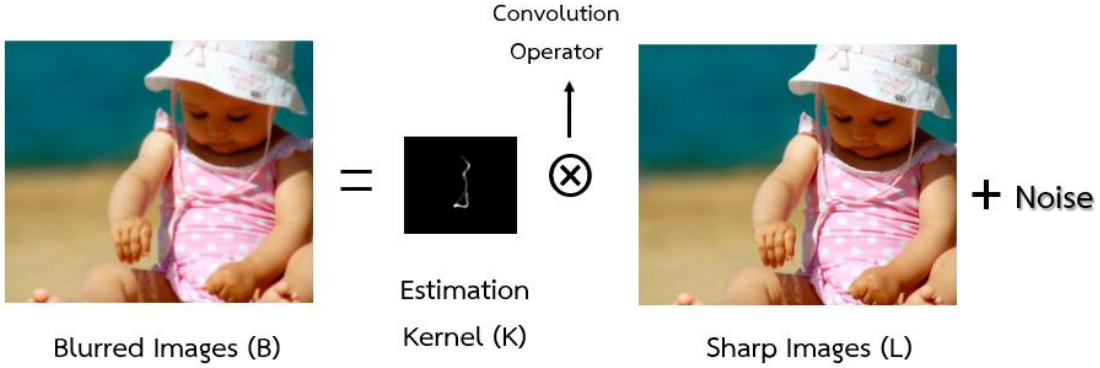
ภาพที่ 2.21 ลักษณะการเบลอประเภทต่างๆ

จากภาพที่ 2.21 แสดงให้เห็นถึงลักษณะของการเบลอ [39] ในแต่ละประเภทที่มีอยู่ในปัจจุบัน ดังนั้นรูปแบบของการเบลอถูกแบ่งออกเป็น 2 แบบ คือ การเบลอแบบสม่ำเสมอ (Uniform) และการเบลอแบบไม่สม่ำเสมอ (Non-uniform) โดยมีรายละเอียดดังนี้

การเบลอแบบสม่ำเสมอ (Uniform) คือ ทุกๆ พิกเซลของภาพจะมีการเบลอที่เท่าๆ กันและไปในทิศทางเดียวกัน

การเบลอแบบไม่สม่ำเสมอ (Non-uniform) คือ พิกเซลเบลอจะมีความแตกต่างกันในแต่ละพื้นที่ และเป็นไปคนละทิศทางซึ่งอาจจะเกิดจากการสั่นของกล้องหรือสภาวะอากาศ และสัญญาณรบกวนที่ไม่สม่ำเสมอ ซึ่งเป็นที่นิยมในการนำไปใช้สำหรับงานวิจัยเกี่ยวกับการเบลอภาพเดี่ยวแบบ Blind Deblurring [40-43]

ด้วยเหตุนี้จึงมีนักวิจัยหลายๆ คน จึงได้นำแบบจำลองของภาพเบลอมาเพื่อใช้สำหรับการปรับปรุงภาพให้คมชัด ด้วยแบบจำลองพื้นฐานที่ใช้กันทั่วไป



ภาพที่ 2.22 โครงสร้างแบบจำลองภาพเบลอและการปรับปรุงภาพเบลอทั่วไป

จากภาพที่ 2.22 โครงสร้างแบบจำลองภาพเบลอและการปรับปรุงภาพเบลอทั่วไปซึ่งสามารถนำมาสร้างเป็นแบบจำลองภาพเบลอในการหาสมการเชิงเส้นได้ดังต่อไปนี้

$$B = K \otimes L + N \tag{2.11}$$

จากสมการที่ 2.11 กำหนดให้ B คือภาพเบลอ ส่วน K เป็นเคอร์เนล L คือภาพแฝง และ N เป็นสัญญาณรบกวน ในส่วนของ $\otimes$ เป็นตัวดำเนินการในกระบวนการ Convolution จากกระบวนการข้างต้นสามารถหาภาพเบลอได้จากการประมาณค่าเคอร์เนล K มา Convolution กับภาพแฝง L โดยเพิ่มสัญญาณรบกวน N เข้าไปในกระบวนการเบลอ แล้วจึงได้ภาพเบลอ B

จากสมการข้างต้นจึงได้ถูกนำมาใช้ในการปรับปรุงภาพเบลอกันอย่างแพร่หลายจนถึงปัจจุบัน โดยแบ่งออกเป็น 2 แบบ ได้แก่ วิธีการปรับปรุงภาพเดี่ยวแบบ Blind Deconvolution และแบบ Non-Blind Deconvolution [44, 45] ซึ่งเป็นวิธีที่นักวิจัยหลายคนได้ให้ความสนใจโดยนำรูปแบบการเบลอมาตรฐานไปใช้ในการสร้างอัลกอริทึมหรือเทคนิคต่างๆ สำหรับการกู้คืนภาพเบลอให้มีความชัดเจนยิ่งขึ้น

ในงานวิจัยของ Fergus ปี 2006 [26] ได้นำเสนอวิธีการลดการสั่นของกล้องจากภาพเดี่ยว โดยใช้เทคนิคใหม่ในการลบเอฟเฟกต์การสั่นของกล้องออกจากภาพ ซึ่งเป็นผลมาจากการปรับปรุงจากงานวิจัยก่อนหน้านี้ โดยใช้ประโยชน์จากการวิจัยล่าสุดที่เกี่ยวกับสถิติของภาพธรรมชาติด้วยการไล่ระดับสีของภาพที่เฉพาะเจาะจงมากขึ้น จากนั้นทำการต่อยอดงานวิจัยของ Miskin และ MacKay ในปี 2000 โดยใช้ Bayesian ในการหาความน่าจะเป็นของค่าเคอร์เนล โดยใช้อัลกอริทึมสำหรับสร้างภาพใหม่ซึ่งแบ่งออกเป็น 2 ขั้นตอนหลักๆ ได้แก่ การประมาณค่าเคอร์เนลเบลอจากภาพที่นำเข้ามาด้วยกระบวนการประมาณค่าแบบหยาบไปถึงละเอียดเพื่อเป็นการหลีกเลี่ยงค่าต่ำสุด และใช้การประมาณค่าเคอร์เนลด้วยอัลกอริธึมมาตรฐานแบบ Deconvolution เพื่อประเมินค่าภาพแฝง ซึ่งได้นำมาชุดข้อมูลภาพเบลอมาจาก Omar Khan และคณะ ผลลัพธ์ที่ได้แสดงให้เห็นว่าวิธีการที่นำเสนอสามารถปรับปรุงภาพเบลอที่ได้ทั้งขนาดเล็กและขนาดใหญ่ในเวลาไม่กี่วินาที อีกทั้งยังมีความคมชัดกว่าวิธีก่อนหน้านี้อย่างมีนัยสำคัญ

ในปี ค.ศ. 2016 Ren และคณะ [46] ได้ศึกษาการประมาณค่าเมทริกซ์ระดับต่ำเกี่ยวกับการกู้คืนภาพเบลอแบบภาพเดี่ยว ซึ่งถูกนำมาใช้แก้ปัญหาแบบ Bild Deconvolution จึงได้เสนออัลกอริทึมใหม่สำหรับการปรับปรุงภาพในระดับต่ำก่อนการกู้คืนภาพเบลอ โดยการวิเคราะห์ผลการประมาณค่าเมทริกซ์ระดับต่ำ (LRMA) ของภาพเบลอแบบ Blind และใช้อัลกอริทึมแบบใหม่ โดยใช้คุณสมบัติของความเข้มภาพและการไล่ระดับสีในระดับต่ำของแพทช์ภาพ สำหรับการกู้คืนภาพระดับกลางด้วยการประมาณค่าเคอร์เนล โดยการย่อขนาดมาตรฐานแบบถ่วงน้ำหนัก (WNNM) จึงทำให้สามารถลดรายละเอียดพื้นผิวและจุดเล็กๆ ได้ โดยใช้ชุดข้อมูลมาตรฐานของ Levin ปี 2009 และ Sun ปี 2013 ผลการทดลองเมื่อนำอัลกอริทึมไปใช้กับชุดข้อมูลมาตรฐานทั้ง 2 ชุด โดยการวัดค่า PSNR ค่า SSIM และค่า KSIM กับภาพในรูปแบบต่างๆ พบว่าอัลกอริทึมที่นำเสนอนี้สามารถทำงานได้ดีกว่าวิธีอื่นเมื่อเทียบกับวิธีการลบเลือนที่ล้ำสมัย โดยมีค่าเป็น 22.56, 0.68 และ 0.60 ตามลำดับ

และในปี ค.ศ. 2018 Pan และคณะ [28] ได้เสนออัลกอริธึมที่ใช้สำหรับการกู้คืนภาพผ่านช่องสัญญาณมืด (Dark Channel) เพื่อจัดการกับภาพธรรมชาติ ข้อความ ภาพใบหน้าและภาพที่มีแสงน้อย โดยใช้วิธีการทางคณิตศาสตร์ในการแบ่งครึ่งกำลังสอง (Half-Quadratic Splitting) เพื่อเพิ่มค่าพิกเซลช่องสัญญาณมืด (Dark Channel) โดยใช้ชุดข้อมูลภาพมาตรฐาน 3 ชุด ได้แก่ ชุดข้อมูลของ Levin ปี 2009, ชุดข้อมูลของ Kohler ปี 2012 และชุดข้อมูลของ Sun ปี 2013 ผลลัพธ์ที่ได้จากการทดลองพบว่าเมื่อนำอัลกอริทึมไปใช้กับชุดข้อมูลมาตรฐานทั้ง 3 ด้วยการวัดค่า PSNR พบว่ามีค่าเฉลี่ยโดยรวมสูงกว่าวิธีการกู้คืนภาพด้วยวิธีการอื่นๆ อีกทั้งยังสามารถนำไปใช้ในการปรับปรุงภาพเบลอที่ไม่สม่ำเสมอ

2. กระบวนการประมวลผลด้วยความน่าจะเป็น (Probabilistic)

ในกระบวนการปรับปรุงภาพมีรูปแบบและเทคนิคที่ใช้ในหลากหลายวิธี ที่จะทำให้ภาพเบลอกลับมามีคุณภาพและสามารถนำไปใช้งานได้การปรับปรุงภาพแบบ Blind Deconvolution และ Non-blind Deconvolution ยังคงเป็นอีกวิธีที่ได้รับความนิยมในการปรับปรุงภาพเบลอทั้งที่เป็นภาพนิ่งและภาพเคลื่อนไหว สำหรับวิธีที่นิยมใช้มากที่สุดคือ MAP และ Bayesian

การหาค่าสูงสุดแบบ Maximum a Posteriori (MAP) [47-49]

วิธีการแก้ปัญหาด้วย Blind Deconvolution ยังคงมีข้อจำกัดที่เกี่ยวกับโดเมนความถี่ในภาพหรือรูปแบบการเรียนรู้โมเดลด้วยการประมาณค่าพารามิเตอร์ (Parameter Estimation) ดังนั้นในการเรียนรู้โมเดลต่างๆ จึงจำเป็นต้องหาค่าถ่วงน้ำหนัก Weight หรือ Parameter ที่มีความเหมาะสมกับการประมาณค่าข้อมูลที่ต้องการเรียนรู้มากที่สุด [46] โดยใช้ทฤษฎีของเบย์ (Bayesian) มาใช้ในการคำนวณหาความน่าจะเป็นจากสูตรการคำนวณดังต่อไปนี้

$$P(A/B) = P(B/A)P(A)/P(B) \tag{2.12}$$

โดยที่ $P(.)$ คือค่าความน่าจะเป็น และ A คือค่าของ W (Weight) แล้ว B คือชุดข้อมูล X จากนั้นนำมาเขียนเป็นสมการใหม่ได้ดังนี้

$$P(W/X) = P(X/W)P(W)/P(X) \tag{2.13}$$

ซึ่งในทฤษฎี Bayesian ค่าของ $P(.)$ จะแทนค่าด้วย

- $P(W|X)$ เรียกว่า Posterior
- $P(X|W)$ เรียกว่า Likelihood
- $P(W)$ เรียกว่า Prior
- (X) เรียกว่า Evidence

จากสมการที่ 2.13 แสดงให้เห็นถึงวิธีการหาค่าสูงสุดของ Posterior ซึ่งจะแทนค่าในการคำนวณเท่ากับ $P(X/W)P(W)/P(X)$ โดยที่ W มีค่าเป็น

$$W = argmax\,(W)P(X/W)P(W)/P(X) \tag{2.14}$$

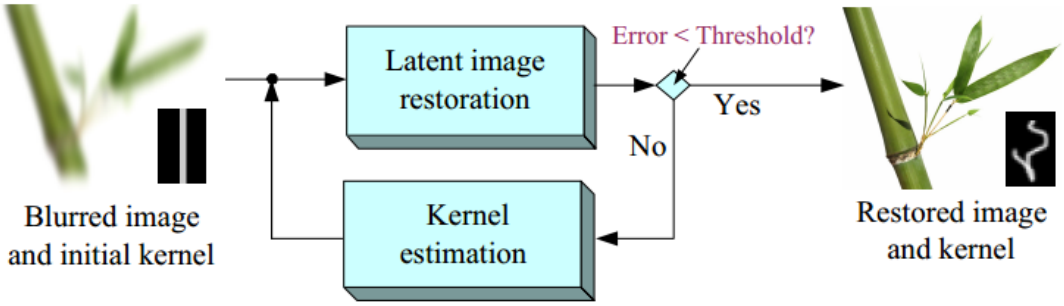
เนื่องจากค่า $P(X)$ เป็นค่าบวกและไม่ขึ้นอยู่กับค่า W จึงทำให้การหาค่าอาร์กิวเมนต์ของค่าสูงสุด $argmax$ สามารถตัดค่า $P(X)$ ออกได้ ดังนั้นจึงมีค่าเป็น

$$W = argmax\,(W)P(X/W)P(W) \tag{2.15}$$

จากนั้นจึงจะสามารถหาค่าอาร์กิวเมนต์สูงสุดของ Posterior ได้ ดังนั้นในการเรียนรู้โมเดลด้วยการหาค่าสูงสุดแบบ Maximum a Posteriori (MAP) จึงถูกพัฒนาขึ้นเพื่อนำเฟรมเวิร์ก MAP มาช่วยเพิ่มประสิทธิภาพในระดับหยาบไปจนถึงละเอียด [26, 28, 30, 47, 48, 50-52]

ในปี ค.ศ. 2007 Jia และคณะ [47] ได้เสนอวิธีการจำแนกการปรับปรุงภาพเบลอด้วยการประมาณค่าการกรองและกระบวนการ Deconvolution ภาพ โดยได้เสนออัลกอริทึมใหม่สำหรับการประมาณค่ากรองภาพเบลอจากการเคลื่อนไหวของภาพเดี่ยว เพื่อกู้คืนทั้งภาพเบลอที่เกิดจากการเคลื่อนไหวของกล้องและการเคลื่อนที่ของวัตถุ ซึ่งใช้วิธีวิเคราะห์การเสื่อมโทรมของภาพและความโปร่งใส่ที่เกี่ยวข้อง โดยใช้ตัวกรองเบลอมากำหนดความโปร่งใสในขอบเขตของวัตถุ อีกทั้งยังใช้วิธีการประมาณค่าการกรอง โดยใช้ Maximum a Posteriori (MAP) ในการคำนวณเพื่อเพิ่มประสิทธิภาพให้กับพื้นที่ที่โปร่งใส จากนั้นใช้วิธีการของ Lucy – Richardson (L-R) ในการคืนค่าภาพเบลอ (Deconvolution) จากผลการทดลองพบว่า ภาพที่มีความสว่างน้อยเมื่อนำไปคำนวณด้วยค่าความโปร่งใสด้วยวิธีของ L-R สามารถทำงานได้ดี และในส่วนของภาพเคลื่อนไหว เมื่อนำไปคำนวณด้วยค่า Alpha Matte บนพื้นที่ที่โปร่งใสจะสามารถทำให้ภาพมีความชัดเจนเพิ่มขึ้น

และในปีต่อมา Shan และคณะ [30] ได้นำเสนออัลกอริทึมใหม่ที่ใช้สำหรับลบภาพเบลอที่เกิดจากการเคลื่อนไหวออกจากภาพเดี่ยวที่มีคุณภาพสูง โดยใช้เทคนิคที่เพิ่มประสิทธิภาพและเป็นประโยชน์อยู่ 3 ประการ ได้แก่ ประการแรกจะช่วยให้สร้างรูปแบบใหม่โดยใชฟังก์ชันการกระจายจุดคงที่ของการเปลี่ยนแปลงเชิงเส้น Point Spread Function (PSF) ซึ่งจะช่วยในการแยกข้อผิดพลาดที่เกิดขึ้นระหว่างการประเมินสัญญาณรบกวนของภาพและการประมาณค่าเคอร์เนลเบลอ ประการที่สอง สามารถจำกัดความเรียบเนียนแบบใหม่ที่สามารถกำหนดให้กับภาพที่มีค่าความคมชัดต่ำ ซึ่งจะมีประสิทธิภาพมากในการลดจุดที่เกิดขึ้นในภาพและบริเวณใกล้เคียง แต่อาจจะมีผลกระทบในขั้นตอนของการปรับแต่งเคอร์เนล และประการสุดท้าย อัลกอริทึมจะถูกปรับให้มีความเหมาะสมโดยใช้ Maximum a Posteriori (MAP) สำหรับเพิ่มประสิทธิภาพโดยใช้รูปแบบและเทคนิคที่มีประสิทธิภาพขั้นสูงโดยเน้นเกี่ยวกับการคำนวณในโดเมนความถี่ โดยใช้ชุดข้อมูลของ Levin และคณะ ปี 2007 จากการทดลองสามารถสร้างผลลัพธ์ที่มีการเบลอคุณภาพสูงได้โดยใช้เวลาในการประมวลผลต่ำกว่าเมื่อเทียบกับวิธีการของ RL และ Levin และคณะ ในปี 2007



ภาพที่ 2.23 ภาพรวมของอัลกอริทึมที่ใช้ในการกู้คืนภาพเบลอ [30]

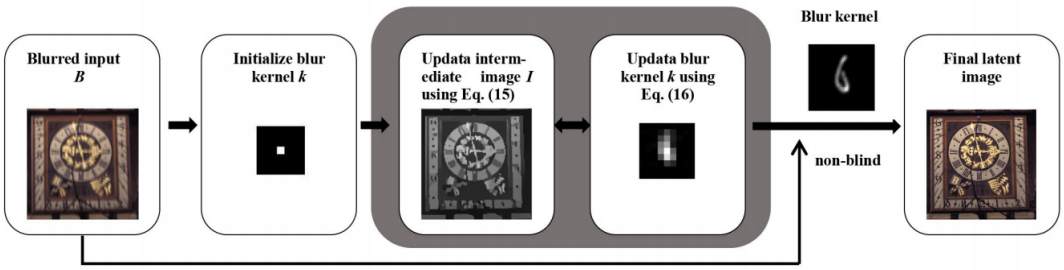
จากภาพที่ 2.23 ด้วยภาพเบลอและการประมาณค่าเคอร์เนลเริ่มต้น วิธีการที่นำเสนอจะเป็นการประเมินค่าภาพแฝง โดยการตรวจสอบค่าความผิดปกติ (Error) ที่เกิดขึ้นภายในภาพหากมีค่าน้อยกว่าค่า Threshold จะถูกนำกลับไปประมาณค่าใหม่ด้วยวิธีการทำซ้ำและปรับแต่งเคอร์เนลเบลอจนกว่าจะครอบคลุมทั่วทั้งภาพ ก่อนจะได้ภาพที่ถูกการเบลอหรือภาพที่ชัดเจน

นอกจากนี้ยังมีวิธีการหาส่วนต่างสูงสุดของพื้นที่แบบ Local Maximum Difference (LMD) [53] โดยหาค่าผลรวมสูงสุดของพิกเซลที่มีความแตกต่างกันระหว่างความเข้มของพิกเซลกับพื้นที่โดยรอบของแพทช์ภาพ ด้วยการประมาณค่าเคอร์เนลเบลอก่อนนำมาลดค่าความเข้มตามกระบวนการเบลอ จากนั้นได้นำมาตัวดำเนินการเชิงเส้นมาคำนวณหาค่า LMD และใช้ข้อจำกัดของ L1 norm ที่เกี่ยวข้องกับ LMD มาช่วยเพิ่มประสิทธิภาพด้วยวิธีการ Half-Quadratic Splitting เพื่อลดปัญหาการนูนของภาพให้น้อยที่สุด [49, 53]

ในปี ค.ศ. 2020 Liu และคณะ [53] ได้เสนออัลกอริทึมสำหรับการปรับปรุงภาพเบลอแบบ Blind จากภาพธรรมชาติผ่านส่วนต่างสูงสุดในพื้นที่ Local Maximum Difference (LMD) โดยหาค่าผลรวมสูงสุดของพิกเซลที่มีความแตกต่างกันระหว่างความเข้มของพิกเซลนั้นกับ 8 พิกเซลโดยรอบในแพทช์ภาพด้วยการประมาณค่าเคอร์เนลเบลอ ก่อนนำมาลดค่าความเข้มตามกระบวนการเบลอ

จากนั้นได้นำมาตัวดำเนินการเชิงเส้นมาคำนวณหาค่า LMD และใช้ข้อจำกัดของ L1 norm ที่เกี่ยวข้องกับ LMD มาช่วยเพิ่มประสิทธิภาพด้วยวิธีการ Half-quadratic splitting ในการจัดการกับภาพเพื่อลดปัญหาการนูนของภาพให้น้อยที่สุด และใช้วิธีกู้คืนภาพแบบ Non-blind เพื่อให้ได้ภาพที่ชัดเจนที่สุดในขั้นตอนสุดท้าย โดยใช้ชุดข้อมูลทั้งภาพสังเคราะห์และภาพจริงจำนวน 4 ชุดข้อมูล ประกอบด้วย ชุดข้อมูลของ Levin 2009, Kohler 2012, Pan 2016 และ Lai 2016 จากผลการทดลองเมื่อนำวิธีที่นำเสนอไปประเมินประสิทธิภาพมาตรฐาน โดยการเปรียบเทียบค่า Error Ratio มีค่าประมาณ 1.20 ซึ่งมีค่าเฉลี่ยต่ำกว่าวิธีการอื่นๆ ทำให้ภาพที่ได้ใกล้เคียงกับภาพที่คมชัด และการหา

ค่าเฉลี่ย PSNR และค่าเฉลี่ย SSIM ซึ่งมีค่าเป็น 31.97 และ 0.91 ตามลำดับ พบว่ามีค่าเฉลี่ยสูงกว่าวิธีการอื่นๆ อย่างมีนัยสำคัญ



ภาพที่ 2.24 การปรับปรุงภาพด้วย LMD [53]

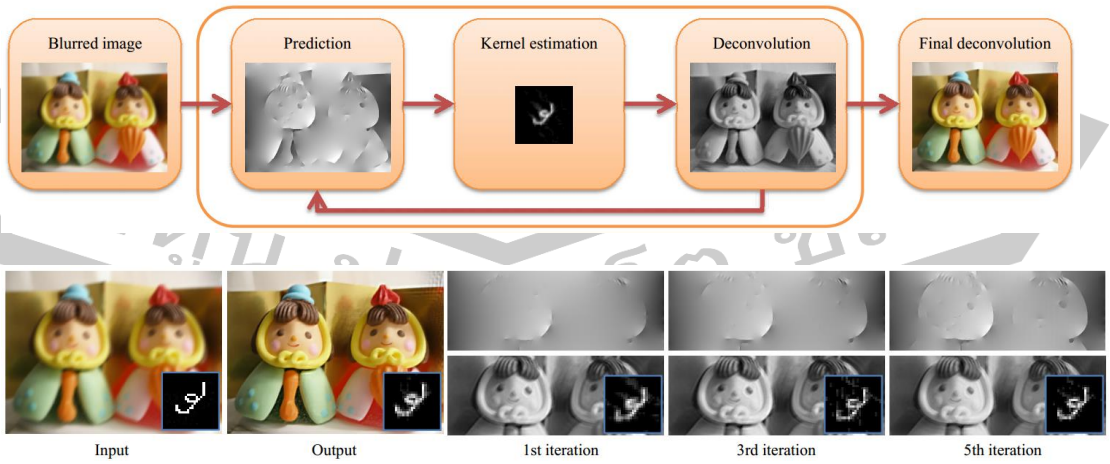
จากภาพที่ 2.24 เป็นการปรับปรุงภาพด้วยวิธี LMD เริ่มจาก Input ภาพเบลอเข้าสู่อัลกอริทึมที่สร้างขึ้น โดยในกรอบสีเทาจะเป็นการประมาณค่าเคอร์เนลเบลอ หลังจากผ่านการประมาณค่าเบลอที่ดีที่สุด จะใช้วิธีการเบลอภาพแบบ Non-Blind เพื่อกู้คืนภาพเบลอให้ชัดเจนในขั้นตอนสุดท้าย

จากนั้น Hu และคณะ [49] ได้ศึกษาเกี่ยวกับปัญหาการปรับปรุงภาพเบลอโดยถูกนำมาแก้ไขด้วยวิธีการต่าง ๆ อย่างเหมาะสม ซึ่งพบว่าผลงานก่อนหน้านี้อาจจะใช้ Sparse Priors สำหรับเคอร์เนลเบลอหรือภาพแฝง รวมถึงข้อจำกัดเพียงเล็กน้อย (Sparse Constraints) และ Normalized Sparsity Prior ในการปรับปรุงภาพ โดยใช้เฟรมเวิร์กของ Maximum a Posteriori Framework (MAP) ดังนั้นจึงได้เสนอวิธีการปรับปรุงภาพเบลอผ่านความเข้มสูงสุดของส่วนที่ได้รับการปรับปรุง Enhanced Local Maximum Intensity Prior (ELMI) โดยใช้การผสมผสานระหว่างความเข้มสูงสุดในพื้น Local Maximum Intensity (LMI) และการไล่ระดับสีสูงสุดของพื้นที่ Local Maximum Gradient (LMG) จากนั้นทำการตรวจสอบภาพเบลอที่สอดคล้องกันซึ่งค่า ELMI จะลดลงตามกระบวนการเบลอ ซึ่งเคอร์เนลเบลอจะถูกประมาณค่าก่อนและใช้วิธีการปรับปรุงภาพแบบ Non-Blind กับภาพที่ชัดเจน โดยใช้ชุดข้อมูลสังเคราะห์และภาพจริงทั้งหมด 4 ชุด ประกอบด้วยชุดข้อมูลของ Pan (2017), Levin (2011), Kohler (2012) และ Lai (2014) จากผลการทดลองได้นำวิธีที่เสนอไปเปรียบเทียบกับอัลกอริทึมที่ล้ำสมัยด้วยการวัดค่าเฉลี่ย Error Ratio, PSNR และ SSIM กับชุดข้อมูลสังเคราะห์ (Synthetic Datasets) พบว่าวิธีที่นำเสนอเมื่อเปรียบเทียบกับชุดข้อมูลดังกล่าวจะมีค่าเฉลี่ย PSNR และค่าเฉลี่ย SSIM สูงกว่าวิธีอื่นๆ อีกทั้งยังลดจุดต่างๆได้น้อยลง ในส่วนของชุดข้อมูลภาพจริง (Real Images) เมื่อเปรียบเทียบกับวิธีการที่นำเสนอพบว่าสามารถปรับภาพได้ชัดเจนและมีรายละเอียดที่คมชัดขึ้นกว่าวิธีอื่นๆ อีกทั้งยังทำให้จุดในภาพลดลงด้วย

การแปลงฟูเรียร์อย่างรวดเร็ว (Fast Fourier Transform: FFT)

ในการประมวลผลภาพทางคอมพิวเตอร์วิชัน (Computer Vision) จำเป็นที่จะต้องใช้การวิธีคำนวณเข้ามาช่วยในการประมวลผลเพื่อทำให้ได้ค่าที่มีความถูกต้องและรวดเร็ว ในการประมวลผลจะต้องใช้เวลาในการคำนวณที่มีความแม่นยำและรูปแบบที่ซับซ้อน ซึ่งอาจขึ้นอยู่กับปัจจัยหลายอย่างเพื่อทำให้ได้ผลลัพธ์และใช้เวลาในการประมวลผลน้อย เช่น ขนาดของภาพที่นำมาประมวลผลและความยาวของชุดข้อมูล ดังนั้นอัลกอริทึมที่ใช้ในการแปลงสัญญาณโดเมนความถี่จึงมีอยู่หลายวิธีที่นิยมนำมาใช้งาน สำหรับการแปลงฟูเรียร์อย่างรวดเร็ว (Fast Fourier Transform: FFT) [37, 41] เป็นอัลกอริทึมในการการแปลงสัญญาณแบบไม่ต่อเนื่อง (Discrete Fourier Transform : DFT) ที่มีความซับซ้อนซึ่งจะช่วยให้การคำนวณเร็วขึ้นและจะทำให้ความผิดพลาดสะสมในการคำนวณลดลง จึงเป็นอีกวิธีหนึ่งที่ผู้วิจัยนิยมนำมาใช้คำนวณ

Cho และ Lee [37] ได้ใช้วิธีการปรับปรุงภาพเบลอจากภาพเดียวที่มีขนาดกลาง เพื่อให้ได้ภาพผลลัพธ์ที่รวดเร็วในเวลาไม่กี่วินาที โดยใช้วิธีการเร่งความเร็วทั้งการประมาณค่าภาพแฝงและการประมาณค่าเคอร์เนลในกระบวนการทำซ้ำ ด้วยขั้นตอนการทำนายแบบใหม่ที่ทำงานกับอนุพันธ์ของภาพมากกว่าค่าพิกเซล ซึ่งในขั้นตอนการทำนายได้ใช้เทคนิคในการทำนายขอบภาพจากการประมาณค่าภาพแฝงใช้วิธีการคำนวณแบบ Gaussian Prior สำหรับขั้นตอนการประมาณค่าเคอร์เนล และได้กำหนดฟังก์ชันการแปลงแบบ Fourier Transforms ในการคำนวณที่เหมาะสมกับอนุพันธ์ของภาพเพื่อเร่งความเร็วในการประมวลผล จากนั้นใช้วิธีการไล่ระดับสีแบบ Conjugate โดยใช้ชุดข้อมูลของ Levin และคณะ ในปี ค.ศ. 2009 [50] ผลการทดลองแสดงให้เห็นว่าวิธีการที่นำเสนอสามารถทำงานตามลำดับความสำคัญได้ดีและใช้เวลาในการประมวลผลน้อยกว่าวิธีที่ผ่านมา ในขณะที่คุณภาพของการเบลอนั้นสามารถเทียบเคียงได้กับการใช้งานจริง



ภาพที่ 2.25 ภาพรวมของกระบวนการปรับปรุงภาพเบลอและขั้นตอนการทำซ้ำ [37]

จากภาพที่ 2.25 แสดงให้เห็นถึงกระบวนการปรับปรุงภาพโดยเริ่มจากการนำภาพเบลอมาผ่านขั้นตอนการทำนายขอบภาพ ก่อนนำไปประมาณค่าเคอร์เนลและทำการ Deconvolution ภาพ และในส่วนของการทำซ้ำจากภาพจะเห็นได้ว่ามีการทำซ้ำถึง 5 ครั้งกว่าจะได้ภาพ Output ตามที่ต้องการ

ต่อมาในปี ค.ศ. 2011 M. Hirsch และคณะ [41] ได้เสนออัลกอริทึมในการเบลอภาพเดี่ยวแบบ Blind Deblurring สำหรับลบภาพเบลอจากการเคลื่อนไหวที่ไม่สม่ำเสมอที่เกิดจากการสั่นของกล้อง โดยใช้เฟรมเวิร์กของ Efficient Filter Flow (EFF) ที่มีข้อจำกัดในการสั่นของกล้องจากโครงสร้างแบบจำลอง Projective Motion Path Blur (PMPB) และทำการ Convolution ของแต่ละแพทช์ภาพด้วยการแปลงฟูเรียร์เร็วแบบสั้น (Fast Fourier Transform: FFT) เพื่อเร่งความเร็วในการคำนวณ จากนั้นประมาณค่าเคอร์เนลเพื่อทำนายภาพเบลอและนำไปไล่ระดับสีด้วย R-map ก่อนนำมาผ่านกระบวนการหมุนด้วยอัลกอริทึมที่สร้างขึ้นและ Deconvolution ภาพเพื่อให้ได้ภาพสุดท้ายที่มีความคมชัด โดยใช้ชุดข้อมูลจาก Real-World Dataset ผลลัพธ์ที่ได้จากการทดลองพบว่าอัลกอริทึมที่สร้างขึ้นมาใหม่มีการประมวลผลเร็วขึ้นและยังประสิทธิภาพมากกว่าวิธีที่ล้ำสมัยก่อนหน้านี้อย่างมีนัยสำคัญ

2.7.2 กระบวนการปรับปรุงภาพระดับสูง (High Level Processing)โดยใช้การประมวลผลภาพด้วยโครงข่ายประสาทเทียม Convolutional Neural Networks (CNNs) และเครือข่ายแบบ Generative Adversarial Networks (GANs)

การปรับปรุงภาพด้วยการเรียนรู้เชิงลึก (Deep Learning) ได้เข้ามามีบทบาทในการประมวลผลภาพซึ่งแสดงประสิทธิภาพได้ดีและมีความซับซ้อนมากขึ้น จึงได้นำมาช่วยในการปรับปรุงภาพให้มีความแม่นยำและถูกต้อง ซึ่งได้อธิบายไว้ในหัวข้อที่ 2.2 อีกทั้งยังมีงานวิจัยอื่นๆ ที่ได้นำโครงข่ายประสาทเทียม Neural Network (NN) และ Generative Adversarial Networks (GANs) [54-57] สำหรับการปรับปรุงคุณภาพของภาพในงานวิจัย

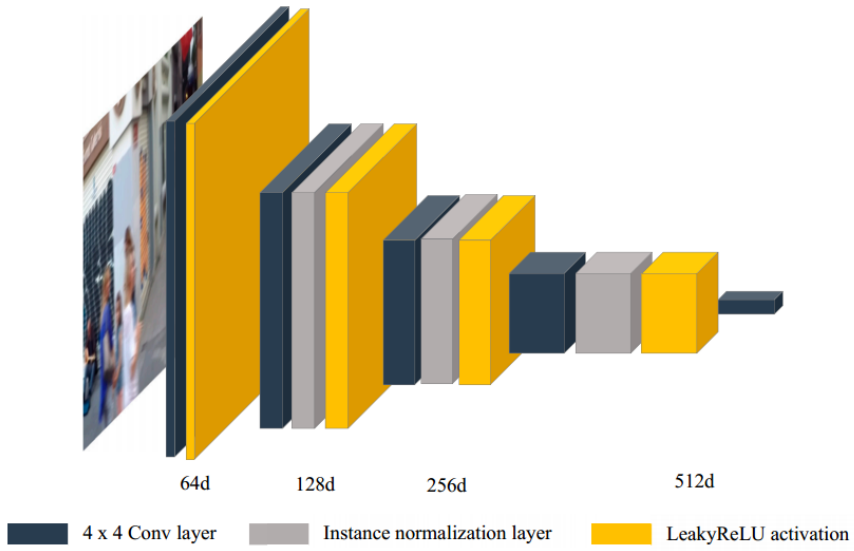
1. การประมวลผลภาพด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution Neural Networks: CNN)

โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution Neural Network: CNN) ถูกนำมาใช้อย่างแพร่หลายทั้งในด้านการคำนวณทางคณิตศาสตร์ สถิติ การประมวลผลสัญญาณ และการประมวลผลภาพ (Computer Vision) ซึ่งการคอนโวลูชันจะเป็นการเปลี่ยนแปลงฟังก์ชัน (f) เมื่อมีฟังก์ชันอื่นๆ เข้ามา ดังนั้น CNN จะเลือกเอาเฉพาะลักษณะเด่น (Feature Extraction) จากโมเดลที่ผ่านการเรียนรู้ของชุดข้อมูลได้ด้วยตัวเองมาใช้งาน ด้วยเหตุนี้ CNN จึงเป็นที่นิยมนำมาใช้สำหรับการประมวลผลในงานวิจัยต่างๆ ซึ่งมีงานวิจัยที่โดดเด่นที่นำเอา CNN มาใช้ดังต่อไปนี้

T. M. Nimisha และคณะ [54] ได้เสนอวิธีการปรับปรุงภาพเบลอจากค่าไม่คงที่แบบ Blind Deblurring ด้วยกระบวนการเรียนรู้เชิงลึก (Deep learning) โดยใช้ CNN ช่วยอำนวยความสะดวกกับการปรับปรุงภาพเบลอแบบ End-to-End ซึ่งไม่จำเป็นต้องมีการสร้างแบบจำลอง โดยใช้วิธีการฝังลงในกระบวนการเรียนรู้ ด้วยการเข้ารหัสแบบอัตโนมัติ (Autoencoder) และ GAN และใช้ชุดข้อมูลมาตรฐานจาก L. Sun 2013 และ Pascal VOC 2012 Dataset ผลลัพธ์ที่ได้จากการทดลองถูกนำมาเปรียบเทียบในเชิงคุณภาพ เมื่อนำมาวัดค่า PSNR และ MSSIM (Mean SSIM) มีค่าเป็น 30.54 และ 0.9553 ตามลำดับ และใช้เวลาในการทำงาน 2 นาที ซึ่งเวลาที่ใช้ดีกว่าอัลกอริทึมที่ล้ำสมัยอื่นๆ ในเชิงปริมาณพบว่าวิธีนี้สามารถหลีกเลี่ยงข้อผิดพลาดได้ดีกว่าและทำให้ภาพชัดเจนขึ้นอย่างเห็นได้ชัดเมื่อเปรียบเทียบกับวิธีอื่นๆ

X. Tao และคณะ [55] ได้ศึกษาวิธีการปรับปรุงภาพเบลอด้วยเรียนรู้แบบ CNN โดยได้เสนอเครือข่ายที่เกิดซ้ำตามขนาด Scale-recurrent Network (SRN) สำหรับการทำให้ภาพเบลอหลายมาตราส่วน ด้วยการเรียนรู้แบบ CNN เข้ามาช่วยในการแก้ปัญหา ด้วยวิธีการออกแบบโครงสร้างแบบสเกลซ้ำ Scale-recurrent Structure และการเข้าและถอดรหัสแบบ Encoder-decoder ResBlock Network โดยใช้ชุดข้อมูลของ GoPro Datasets จำนวน 3,214 ภาพ และ Kohler Datasets จำนวน 48 ภาพ ผลการทดลองได้นำวิธีที่นำเสนอไปใช้กับชุดข้อมูลมาตรฐานทั้ง 2 ชุด และได้นำมาวัดประสิทธิภาพของภาพ (PSNR) และความคล้ายคลึง (SSIM) พบว่ามีค่าเฉลี่ยมากกว่าวิธีการอื่นๆ อีกทั้งยังใช้เวลาในการงานน้อยกว่าอย่างมีนัยสำคัญ ดังนั้นผลลัพธ์ที่เกิดจากการทดลองนี้มีความล้ำสมัยว่าวิธีอื่นๆ ทั้งเชิงคุณภาพและเชิงปริมาณ

Hong และ Choe และคณะ [56] ได้ศึกษาวิธีการปรับปรุงภาพด้วยการเรียนรู้เชิงลึก (Deep Learning) เพื่อใช้สำหรับแก้ไขสัญญาณรบกวนซึ่งมีคุณภาพสูงในการสร้างข้อมูลพื้นผิวของรูปภาพ ดังนั้นจึงได้เสนอวิธีการลบภาพเบลอแบบ Wasserstein Generative Adversarial Network with Gradient Penalty (WGANGP) โดยใช้ความคล้ายคลึงในการรับรู้ ด้วยวิธีการเรียนรู้เชิงลึก (Deep Learning) และใช้วิธีการแทนที่ Loss Function ด้วย Style Loss Function เพื่อรักษารายละเอียดข้อมูลขอบภาพ โดยใช้ชุดข้อมูลของ GoPro Large Dataset จำนวน 3,124 ภาพ และ Kohler Datasets จำนวน 48 ภาพ ซึ่งทั้งสองชุดได้แบ่งข้อมูลออกเป็นชุดการฝึกฝน (Training Set) และชุดการทดสอบ (Test Set) โดยปรับค่าความละเอียดลดลงให้ภาพอยู่ในระดับกลางทั้งหมด จากการทดลองเมื่อใช้วิธีที่นำเสนอมาประเมินประสิทธิภาพกับชุดข้อมูลของ GoPro Large Dataset ด้วยการหาค่าเฉลี่ยของ PSNR และ SSIM พบว่ามีค่าเป็น 33.29 และ 0.98 ซึ่งแสดงให้เห็นว่าค่าเฉลี่ยสูงกว่าวิธีการที่ล้ำสมัยอื่น ๆ และในชุดข้อมูลของ Kohler Datasets มีค่า PSNR เฉลี่ยน้อยกว่าแต่ยังมีค่า SSIM ที่สูงกว่าวิธีอื่นๆ



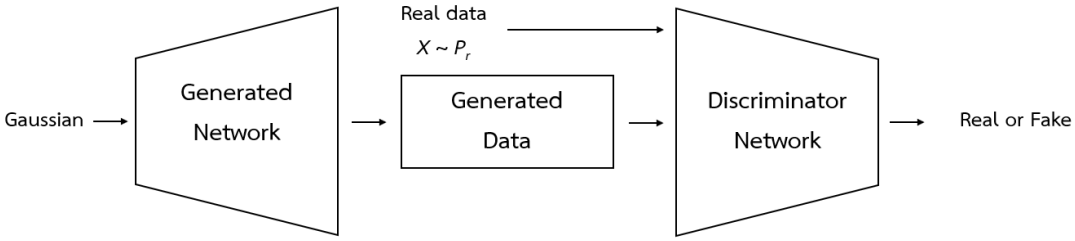
ภาพที่ 2.26 สถาปัตยกรรมของ Discriminator ใน WGAN-GP [56]

จากภาพที่ 2.26 เป็นสถาปัตยกรรมเดียวกันกับ Patch-GAN โดยจำแนกว่าแต่ละแพทช์ $N \times N$ ในภาพเป็นของจริงหรือของปลอม ซึ่งจะประกอบด้วยชั้น Convolutional จำนวน 5 ชั้น และชั้น Normalization จำนวน 3 ชั้น ซึ่งจะมีความแตกต่างกัน และในส่วนของตัว LeakyReLU Activation จะถูกนำไปใช้กับ Convolutional Layer

ในปี ค.ศ. 2019 L. Li และคณะ [57] ได้ศึกษาเกี่ยวกับวิธีการปรับปรุงภาพเบลอโดยใช้การเรียนรู้เชิงลึก (Deep Learning) จึงได้เสนออัลกอริทึมที่มีประสิทธิภาพในการปรับปรุงภาพเบลอแบบ Blind ที่ขึ้นอยู่กับข้อมูลที่นำเข้า ซึ่งใช้วิธีการคำนวณและจำแนกประเภทของภาพที่มีขนาดแตกต่างกันด้วย CNN จากนั้นนำโมเดลที่กำหนดมาฝึกฝนข้อมูลเพื่อให้สามารถจดจำภาพ ด้วยเฟรมเวิร์กแบบหยาบไปถึงละเอียด (MAP) และใช้วิธีการคำนวณทางคณิตศาสตร์แบบ (Half-Quadratic Splitting) ในการประมาณค่าเคอร์เนล โดยใช้ชุดข้อมูลภาพมาตรฐาน 4 ประเภท จากผลการทดลองพบว่าเมื่อนำมาประเมินประสิทธิภาพในเชิงปริมาณและคุณภาพ วิธีที่นำเสนอโดยรวมมีค่าเฉลี่ยที่สูงกว่าวิธีการอื่นๆ ที่มีความล้ำสมัยอย่างมีนัยสำคัญ

2. เครือข่ายแบบ Generative Adversarial Networks (GANs)

เครือข่ายแบบ Generative Adversarial Networks (GANs) [54, 56, 58] ถูกนำเสนอโดย Goodfellow ในปี ค.ศ. 2014 โดยใช้สำหรับงานสร้างภาพและเรียนรู้จากการแข่งขันระหว่างโมเดล โครงข่ายประสาทเทียมสองรุ่น โมเดลโครงข่ายประสาทเทียมสองแบบเรียกว่าตัวสร้าง G และตัวแบ่งแยก D เป้าหมายของ G คือการสร้างภาพที่ D ไม่สามารถแยกแยะจากภาพจริงได้ และเป้าหมายของ D คือการแยกความแตกต่างของภาพจริงและภาพที่สร้างขึ้น



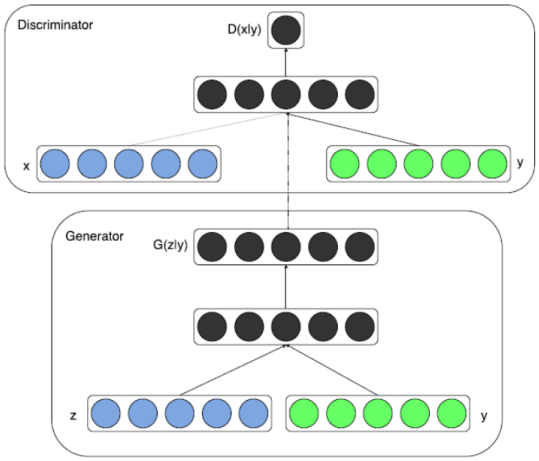
ภาพที่ 2.27 โครงสร้างของ Generative Adversarial Networks (GANs) [56]

จากงานวิจัยที่ผ่านมา [54-56, 58] ได้ใช้หลักการของ CNN และ GANs มาช่วยปรับปรุงภาพที่เสื่อมโทรมด้วยการจำแนกประเภทจากกระบวนการเรียนรู้เชิงลึก [22, 56] โดยใช้อัลกอริทึมและฟังก์ชันที่มีความเหมาะสมมาใช้ในการปรับปรุงภาพ เช่น Loss Function, Scale-Recurrent Structure และ Encoder-Decoder ResBlock Network [55] สำหรับการทดลองใช้กับชุดข้อมูลมาตรฐานที่แบ่งออกเป็นชุดข้อมูลสำหรับการเรียนรู้ (Training) และสำหรับการทดสอบ (Test) พบว่าวิธีที่ใช้ CNN และ GANs เมื่อประเมินประสิทธิภาพด้วยค่า PSNR, SSIM และเวลาในการงาน ผลลัพธ์ที่ได้สามารถทำงานได้ดีเมื่อเปรียบเทียบกับวิธีที่ล้ำสมัย ทั้งเชิงคุณภาพและเชิงปริมาณ

ในปี ค.ศ. 2017 S. Ramakrishnan และคณะ [58] ได้เสนอวิธี Generative Adversarial Network (GAN) ตามสถาปัตยกรรมของการกรองเชิงลึกแบบใหม่ร่วมกับสถาปัตยกรรมที่มีความหนาแน่นสำหรับปรับปรุงภาพเคลื่อนไหว ซึ่งในกระบวนการทำงานจะถูกแบ่งออกเป็น 2 ส่วน ได้แก่ ในส่วนแรกจะใช้ในการออกแบบโครงสร้างและสร้างเฟรมเวิร์กของ GAN และในส่วนที่สองเป็นฟังก์ชันในการลดค่าความเสียหายที่เกิดในภาพ เป็นการเรียนรู้จากภาพที่เสื่อมโทรมโดยใช้ L1 และ L2 ที่เสียหายจากภาพคมชัดและภาพที่แก้ไขเป็นหลัก ก่อนนำค่าที่ได้มาเปรียบเทียบหาค่าประสิทธิภาพ ใช้ชุดข้อมูลทั้งหมด 4 ชุด ประกอบด้วย Places Datasets (2014), GoPro Dataset (2017), Lai et al. Dataset (2016) และ Kohler et al. Dataset (2012) ซึ่งใช้เป็นข้อมูลในการทดสอบ (Test) และในส่วนของข้อมูลที่ใช้ในการเรียนรู้ (Train) ใช้ชุดข้อมูล 3 ชุด ประกอบด้วย GoPro Dataset รวมกับ MS-COCO และ ImageNet Dataset โดยถูกกำหนดขนาดให้เป็น 256x256 พิกเซล จากผลการทดลองเมื่อนำวิธีการที่นำเสนอไปใช้กับชุดข้อมูลมาตรฐานดังกล่าว พบว่าแบบจำลองและเฟรมเวิร์กที่นำเสนอมีประสิทธิภาพเหนือว่าอัลกอริทึมการเบลอที่ล้ำสมัยที่ใช้สำหรับการกู้คืนภาพแบบไม่สม่ำเสมอ ทั้งเชิงปริมาณและคุณภาพ

เครือข่ายแบบ Convolutional Generative Adversarial Networks (CGANs)

เป็นโครงข่ายประสาทเทียมที่มีความซับซ้อนแบบหนึ่งหรือ CNNs ในระยะสั้น ซึ่งเป็นเครือข่ายประสาทเทียมที่มีการเรียนรู้เชิงลึกที่ได้รับการออกแบบมาโดยเฉพาะสำหรับการวิเคราะห์รูปภาพ และเครือข่ายเหล่านี้สามารถจดจำรูปแบบการประมวลผลที่มีความซับซ้อนต่างๆ ภายในรูปภาพได้อย่างแม่นยำ โดยโครงสร้างของการทำงานจะเป็นการนำภาพที่ถูกต้องและไม่ถูกต้องมาทำการฝึกฝน (Training) เพื่อให้สามารถจดจำภาพที่นำเข้ามาได้อย่างแม่นยำ ก่อนจะทำการคัดแยกภาพที่ไม่ต้องการออกให้เหลือเฉพาะภาพที่ต้องการเท่านั้น ดังภาพที่ 2.28 เป็นการแสดงตัวอย่างโครงสร้างของเครือข่ายแบบง่ายเพื่อให้สามารถเข้าใจถึงกระบวนการทำงานได้ดังต่อไปนี้



ภาพที่ 2.28 โครงสร้างของเครือข่ายแบบ Convolutional Generative Adversarial Networks (CGANs)

3. การประมาณค่าเคอร์เนลแบบไฮบริด (Hybrid Kernel Estimation)

รูปแบบของการประมาณค่าเคอร์เนลแบบไฮบริดจะเป็นเคอร์เนลแบบผสมซึ่งเป็นเทคนิคที่ใช้ในการรวมการประมาณค่าในรูปแบบต่างๆ ที่ใช้สำหรับการปรับปรุงภาพ โดยนำมารวมไว้ที่เดียวเพื่อทำการประมวลผล ก่อนนำค่าที่ได้ไปใช้ในกระบวนการอื่นๆ ต่อไป โดยในงานวิจัยที่ผ่านมาจะมีการนำเทคนิคต่างๆ ที่ได้รับความนิยมเข้ามาใช้ในการวิจัย เช่น การแยกส่วนของภาพ (Image Segmentation) ซึ่งจะมีรายละเอียดดังต่อไปนี้

การแยกส่วนของภาพ (Image Segmentation)

สำหรับวิธีการแยกส่วนของภาพ (Image Segmentation) เป็นอีกหนึ่งเทคนิคในการประมวลผลภาพโดยมีเป้าหมายในการแยกวัตถุ (Object) หรือลักษณะจุดเด่น (Feature) ออกจากภาพพื้นหลัง (Background) ซึ่งเป็นวิธีการพื้นฐานที่ได้รับความนิยมในการแยกวัตถุออกจากภาพแบ่งออกเป็น 2 วิธีคือ การตรวจจับขอบภาพ (Edge Detection) และการกระทำแบบเทรชโฮลด์

(Threshold) นอกจากนี้ ในการวิเคราะห์หรือการประมวลผลของข้อมูลภาพจะเน้นไปถึงการแยกวัตถุหรือสิ่งที่สนใจในภาพออกจากพื้นหลัง เช่น การแยกตัวอักษรสีดำออกจากพื้นหลังสีขาวในงานทางด้านการรู้จำตัวอักษร (Character Recognition) หรือประยุกต์ใช้การประมวลผลภาพทางด้านการรู้จำใบหน้าคน (Face Recognition) ดังนั้นจึงจำเป็นที่จะต้องแยกส่วนของวัตถุออกจากพื้นหลังก่อนนำไปวิเคราะห์หรือประมวลผลภาพในขั้นตอนต่อไป เพื่อให้สามารถเรียนรู้เทคนิคและวิธีการทั้งสองจะได้กล่าวถึงรายละเอียดดังต่อไปนี้

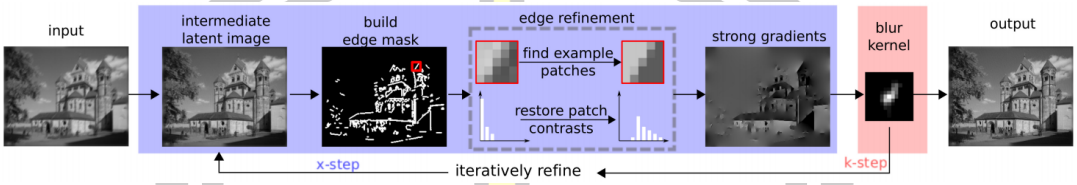
การตรวจจับขอบภาพ (Edge Detection)

ขอบของภาพ (Edge) เป็นส่วนของข้อมูลที่แสดงถึงโครงร่างของวัตถุภายในภาพ ซึ่งประกอบด้วยข้อมูลของภาพที่มีความสำคัญและสามารถนำไปใช้ประโยชน์ในงานด้านต่างๆ เช่น สามารถนำไปใช้ในการระบุถึงขนาดของวัตถุที่อยู่ในภาพ หรือประยุกต์ใช้ในการแยกแยะระหว่างวัตถุหรือข้อมูลในภาพกับส่วนของพื้นหลังของภาพ (Background) หรือการนำไปใช้ในการระบุวัตถุที่อยู่ภายในภาพ อีกทั้งยังสามารถรักษารายละเอียดของขอบภาพให้มีความแข็งแกร่งขึ้นอีกด้วย [56]

ในหลายงานวิจัยได้ใช้วิธีการตรวจจับขอบภาพ (Edge Detection) มาใช้ในการกู้คืนภาพที่เสื่อมโทรมให้กลายมาเป็นภาพที่คมชัดได้ ซึ่งเป็นวิธีที่ใช้ในการตรวจสอบหาความแตกต่างของแต่ละพิกเซล (Pixel) เพื่อหาค่าความแข็งแกร่งในบริเวณขอบภาพ โดยการตรวจสอบค่าความเข้มของสี (Intensity) ในแต่ละพิกเซลเพื่อหาขอบภาพ ก่อนใช้กระบวนการปรุงปรุงขอบให้คมชัดด้วยการเพิ่มค่าคอนทราสต์ (Contrast) ในพื้นที่ของแพทช์ภาพและการไล่ระดับสีของภาพให้เข้มขึ้น (Gradient Image) [44] จากวิธีการดังกล่าวในงานวิจัยของ [45, 52] จึงได้มีการเพิ่มประสิทธิภาพของการปรับปรุงขอบภาพ โดยใช้การประมาณค่าสเปกตรัมของเคอร์เนลเบลอ และการประมาณค่าเบลอแบบไฮบริด (Hybrid) จากข้อมูลที่ได้ทั้งหมดของขอบภาพและค่าสเปกตรัม ในงานวิจัยของ [45] โดยประเมินประสิทธิภาพจากค่าเฉลี่ยของ PSNR, SSIM และค่า Error Ratio ทำให้ภาพมีความคมชัด

ในปี ค.ศ. 2013 L.Sun และคณะ [44] ได้ศึกษาเกี่ยวกับการปรับปรุงภาพเบลอแบบ Blind Deconvolution ที่มีการแยกส่วน เช่น การประมาณเคอร์เนลเบลอและภาพแฝง จากภาพที่เบลออินพุตเข้ามาถูกมองเป็นปัญหาร้ายแรงที่ควรได้รับการแก้ไข ดังนั้นจึงได้เสนอวิธีการปรับปรุงขอบภาพด้วยการประมาณค่าเคอร์เนลเบลอโดยใช้ Patch Priors บนขอบภาพแฝง สำหรับอัลกอริทึมที่นำเสนอจะใช้วิธีการทำซ้ำระหว่างภาพแฝงและเคอร์เนลเบลอ โดยใช้ Patch Priors ช่วยในกระบวนการปรับปรุงขอบ และปรับขอบภาพให้คมชัดด้วยการเพิ่มค่าคอนทราสต์ (Contrast) ในพื้นที่แพทช์ของขอบภาพ จากนั้นอัพเดตเคอร์เนลเบลอด้วยการไล่ระดับสีให้เข้มขึ้นจากภาพแฝงที่กู้คืน หลังจากการประมาณค่าเคอร์เนลด้วยวิธีของ D. Zoran และ Y. Weiss แล้ว จึงใช้การกู้คืนภาพแบบ Non-Blind Deconvolution ชุดข้อมูลมาตรฐานของ Levin 2009 และ 2011ซึ่งประกอบด้วยชุดข้อมูล 2 ประเภท คือ ข้อมูลภาพธรรมชาติและข้อมูลสังเคราะห์ จากการทดลองพบว่าเมื่อนำไป

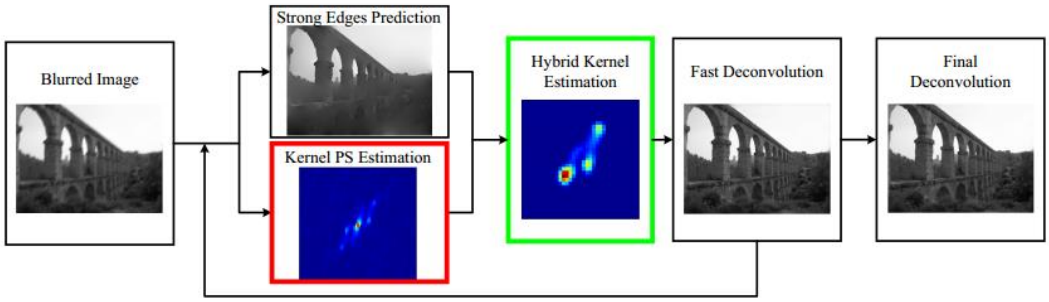
เปรียบเทียบกับชุดข้อมูลทั้ง 2 ประเภท โดยใช้การวิเคราะห์เชิงปริมาณ ประกอบด้วย ค่าเฉลี่ย PSNR, ค่า SSIM และค่าเฉลี่ยเรขาคณิตของอัตราส่วนความผิดพลาด (Geometric Mean) โดยรวมวิธีที่นำเสนอมีการปรับปรุงประสิทธิภาพที่มีความล้ำสมัยและมีความแข็งแกร่งซึ่งมีผลลัพธ์ที่เทียบได้กับวิธีการที่ทันสมัยที่สุด



ภาพที่ 2.29 การปรับปรุงภาพแบบ X-Step และ K-Step [44]

จากภาพที่ 2.29 แสดงให้เห็นถึงขั้นตอนในการปรับปรุงภาพจากการอินพุตภาพเบลอเข้าสู่ระบบ จากนั้นปรับภาพแฝงให้เป็นภาพระดับกลางและสร้าง Mask ให้กับขอบภาพ ก่อนนำมาแต่งขอบภาพ ซึ่งเป็นขั้นตอนของ X-Step จากนั้นนำภาพที่ได้ไปสู่ขั้นตอนของ K-Step ด้วยการปรับค่า Blur Kernel แล้วจึงจะได้ภาพสุดท้ายที่คมชัดออกมา

Yue และคณะ [45] ได้เสนอวิธีการกู้คืนภาพแบบไฮบริดโดยใช้ทั้งขอบภาพและข้อมูลสเปกตรัมจากภาพอินพุต โดยแบ่งออกเป็น 2 องค์ประกอบ คือการประมาณค่าเคอร์เนลและทำนายขอบภาพ ด้วยกระบวนการที่ความเหมาะสม ซึ่งแบ่งออกเป็น 2 วิธี ได้แก่ วิธีการแรก ปรับปรุงการประมาณค่าสเปกตรัมของเคอร์เนลเบลอ โดยกำจัดผลกระทบเชิงลบออกจากโครงสร้างของขอบภาพ และวิธีการที่สอง ประมาณค่าเบลอแบบไฮบริดที่รวบรวมข้อมูลของขอบภาพและค่าสเปกตรัมที่ได้รับการเพิ่มประสิทธิภาพ โดยใช้ชุดข้อมูลมาตรฐาน 3 ชุด ได้แก่ ชุดข้อมูลของ Sun (2013) ชุดข้อมูลของ Levin (2009) และชุดข้อมูลของ Kohler (2012) มาใช้ในการทดลอง จากผลการทดลองเมื่อนำวิธีที่นำเสนอไปใช้กับชุดข้อมูลดังกล่าว มาใช้สำหรับการเปรียบเทียบข้อมูลสังเคราะห์และการเปรียบเทียบตัวอย่างจริงด้วยวิธีที่ล้ำสมัย ด้วยการวัดค่า PSNR ค่า SSIM และค่า Error Ratio พบว่าผลการทดลองแสดงให้เห็นถึงวิธีการที่นำเสนอมีประสิทธิภาพดีกว่าวิธีการทดลองก่อนหน้านี้ที่อิงกับการปรับปรุงขอบและการประมาณค่าสเปกตรัม โดยใช้ชุดข้อมูลเดียวกันอย่างมีนัยสำคัญ



ภาพที่ 2.30 การปรับปรุงขอบภาพแบบไฮบริด (Hybrid) [45]

จากภาพที่ 2.30 แสดงให้เห็นวิธีการปรับปรุงภาพเบลอด้วยการทำขอบภาพให้ชัดด้วยวิธี (Edge Detection) และการประมาณค่าเคอร์เนลเบลอ จากนั้นนำทั้งสองค่าไปประมาณค่าแบบไฮบริด (Hybrid) แล้วทำการ Deconvolution ภาพเพื่อแปลงกลับคืนและสุดท้ายจะได้รับภาพที่มีความคมชัด

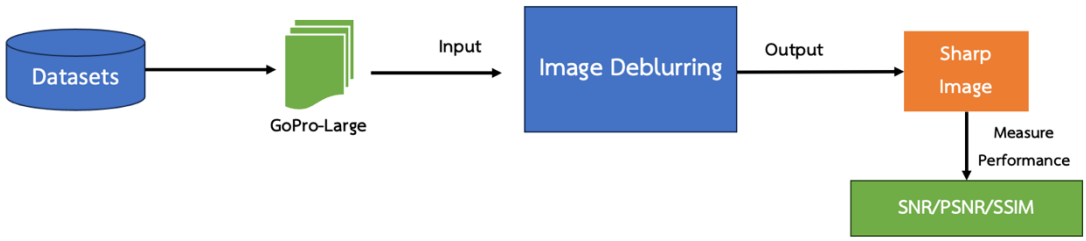
บทที่ 3
วิธีดำเนินการวิจัย

การวิจัยเรื่อง การลดความเบลอของภาพ ในครั้งนี้ ผู้วิจัยได้ทำการศึกษางานวิจัยและเอกสารต่างๆ ที่เกี่ยวข้องกับการเรียนรู้เชิงลึก (Deep Learning) เพื่อนำข้อมูลมาใช้ในการพัฒนาอัลกอริทึมเพื่อปรับปรุงภาพเบลอให้มีความชัดเจนมากยิ่งขึ้น โดยในการวิจัยครั้งนี้ผู้วิจัยทำการแบ่งขั้นตอนการดำเนินการวิจัยออกเป็น 3 ขั้นตอน ตามแนวความของ S. Ramakrishnan และคณะ [58] มีรายละเอียดดังต่อไปนี้

3.1 ขั้นตอนการเตรียมชุดข้อมูล (Dataset Preprocessing) เป็นการทำงานในส่วนของการเตรียมชุดข้อมูลมาตรฐาน โดยใช้ชุดข้อมูล GoPro-Large [59]

3.2 ขั้นตอนการประมวลผล (Processing) เป็นการนำข้อมูลทั้งหมดเข้ามาทำงานในกระบวนการของ Deep Learning ประกอบด้วย Autoencoders กับ GANs

3.3 ขั้นตอนการวัดประสิทธิภาพ (Evaluation) เป็นการเปรียบเทียบการทำงานที่นำเสนอและวัดประสิทธิภาพของภาพที่ผ่านการประมวลผล โดยใช้วิธีการวัดแบบ SNR, PSNR, SSIM



ภาพที่ 3.1 โครงสร้างและขั้นตอนการปรับปรุงภาพ

3.1 ขั้นตอนการเตรียมชุดข้อมูล (Dataset Preprocessing)

ในการเตรียมชุดข้อมูลเพื่อทำการทำวิจัยครั้งนี้ ผู้วิจัยได้นำชุดข้อมูลที่เป็นแบบมาตรฐานและสามารถให้นำมาใช้งานโดยไม่มีค่าใช้จ่าย อีกทั้งยังเป็นที่นิยมนำมาใช้งานเพื่อนำมาเปรียบเทียบผลลัพธ์ที่ได้จากการวิจัยกันอย่างแพร่หลาย โดยในการวิจัยนี้ได้ใช้ชุดข้อมูลของ GoPro Dataset [59] ซึ่งชุดข้อมูลนี้จะมีลักษณะเป็นภาพนิ่งที่เกิดจากกล้องวิดีโอ โดยสามารถแยกรายละเอียดของชุดข้อมูลแต่ละส่วนออกได้ดังต่อไปนี้

GoPro Dataset [59] เป็นชุดข้อมูลที่ได้มาจากกล้องวิดีโอ ซึ่งเป็นภาพชุดที่มีลักษณะเรียงต่อทีละภาพ เป็นการเคลื่อนที่ทีละเฟรม โดยแยกออกเป็น 2 ประเภท ได้แก่ Blur, Blur_gamma ในแต่ละประเภทจะมีความเบลอที่แตกต่างกัน อีกทั้งยังแบ่งชุดข้อมูลออกเป็นการชุดข้อมูลฝึก (Train) จำนวน 6,309 ภาพ และชุดข้อมูลทดสอบ (Test) จำนวน 3,333 ภาพ รวมทั้งหมด ภาพ 9,642 ภาพ

ตารางที่ 3.1 ลักษณะของชุดข้อมูล (Dataset) ของ GoPro-Large

GoPro-Large		
Type	Train	Test
จำนวน	6,309	3,333
รวม	9,642	

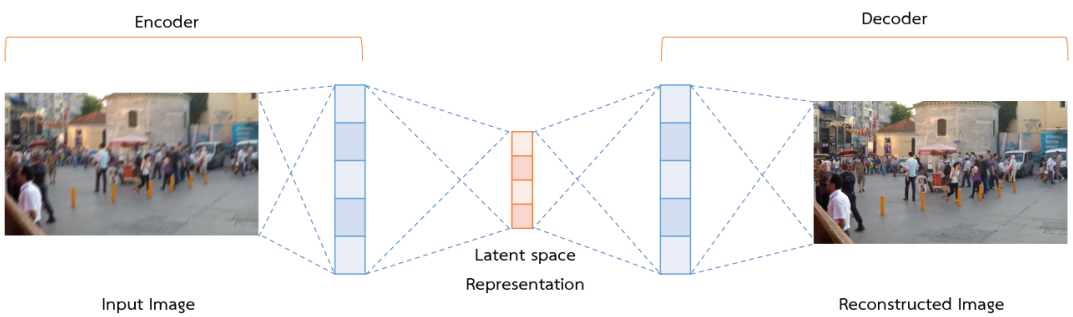
จากตารางที่ 3.1 แสดงให้เห็นรายละเอียดของชุดข้อมูลของ GoPro-Large ทั้ง 2 ประเภท ที่ได้นำมาใช้ในการทดลอง

3.2 ขั้นตอนการประมวลผล (Processing)

ในขั้นตอนการประมวลผลได้ใช้วิธีการทดลองผ่านกระบวนการเรียนรู้เชิงลึก (Deep learning) โดยจะใช้วิธีการเข้ารหัสอัตโนมัติ (Autoencoder) และ Generative Adversarial Networks: GANs

3.2.1 Autoencoder

Autoencoder เป็น Neural Network ประเภทหนึ่งที่ประกอบไปด้วยการทำงานอยู่ 3 ส่วน ได้แก่ การเข้ารหัส (Encoder) ส่วนที่จัดการกับข้อมูลที่ถูกบีบอัด (หรือคอขวด) และการถอดรหัส (Decoder) โดยมีหลักการทำงาน คือ การรับข้อมูลเข้ามา (Input) แล้วทำการเข้ารหัสข้อมูลด้วยการบีบอัดข้อมูล ซึ่งจะไม่สามารถดูออกว่าข้อมูลที่ถูกเข้ารหัสนั้นมีหน้าตาเหมือนกับต้นฉบับหรือไม่ จากนั้นจะทำการลดสัญญาณรบกวน (Noise) ที่ไม่ต้องการออกจากภาพ โดยจะแบ่งออกเป็นเลเยอร์ต่างๆ เพื่อเพิ่มความซับซ้อนและบีบอัดข้อมูลให้มีขนาดเล็กลง ก่อนที่จะทำการแปลงค่ากลับหรือการถอดรหัส (Decode) ให้มีค่าเท่ากับขนาดของภาพเดิมที่ Input เข้ามา โดยข้อมูลที่ถูกถอดรหัสนั้นจะสูญเสียข้อมูลบางอย่างไป ก่อนจะนำภาพผลลัพธ์ (Output) มาเปรียบเทียบกับภาพต้นฉบับ โดยจะวิธีการเทียบภาพแบบ Pixel ต่อ Pixel ด้วย Loss Function สำหรับวิธีการดังกล่าวนี้จะทำให้การทำงานเร็วขึ้นหรือช้าลงก็ขึ้นอยู่กับจำนวนของเลเยอร์ (Layer) ที่ถูกสร้างขึ้น



ภาพที่ 3.2 โครงสร้างของการทำ Autoencoder Architecture

จากภาพที่ 3.2 จะแสดงให้เห็นถึงโครงสร้างและขั้นตอนการทำงาน Autoencoder ตั้งแต่การเข้ารหัสไปจนถึงการถอดรหัสข้อมูลเพื่อให้ได้ภาพผลลัพธ์

จากโครงสร้างการทำงานของ Autoencoder ในการวิจัยนี้จะเน้นการปรับค่า Loss Function ให้มีค่าน้อยที่สุดโดยการปรับค่าพารามิเตอร์และฟังก์ชันให้มีความเหมาะสม เพื่อทำให้มีค่าใกล้เคียงกับภาพชัด ดังนั้นจึงได้นำเอาวิธีการต่างๆ มาทำงานร่วมกับโครงสร้าง Autoencoder โดยแบ่งออกเป็น 7 วิธี และการทำ GANs จำนวน 1 วิธี ดังต่อไปนี้

1. การทำ Autoencoder
2. การทำ Autoencoder ด้วย VGG16
3. การทำ Autoencoder ด้วย MSE รวมกับ SSIM
4. การทำ Autoencoder ด้วย MSE รวมกับ Total Variance (TV)
5. การทำ Autoencoder ด้วย MSE รวมกับ SSIM และ Total Variance (TV)
6. การทำ Autoencoder ด้วย MSE รวมกับ SSIM และ VGG16
7. การทำ Autoencoder ด้วย MSE รวมกับ SSIM รวมกับ Total Variance (TV) และ VGG16
8. การปรับปรุงภาพด้วย Generative Adversarial Networks: GANs

วิธีการทดลองที่ 1 การทำ Autoencoder จะใช้วิธีการเข้ารหัส (Encoder) ภาพที่อินพุตเพื่อลบส่วนที่ไม่สำคัญออกจากภาพก่อนจะเหลือส่วนที่มีลักษณะที่ใกล้เคียงกันมากที่สุดแล้วนำไปสร้างภาพใหม่ จากนั้นจึงนำไปถอดรหัส (Decoder) เพื่อแปลงภาพที่ได้ให้มีขนาดเท่ากับภาพต้นฉบับ โดยจะมีฟังก์ชัน Mean Squared Error (MSE) ไว้สำหรับปรับค่า Loss Function ให้มีค่าต่ำลงเพื่อให้ภาพที่ได้ใกล้เคียงกันมากที่สุด โดยสามารถเขียนเป็นสมการเพื่อนำไปใช้ในการคำนวณ ได้ดังต่อไปนี้

การทำ Encoder

เริ่มจากการ Input: X (ข้อมูลที่เป็นอินพุต) เข้าสู่ระบบ

$$\text{Encoding: } Z = fencoder(X) \tag{3.1}$$

จากสมการนี้ $fencoder$ แสดงถึงฟังก์ชันในการเข้ารหัส ซึ่งจะทำ Feature map ค่า X กับ Z ให้ตรงกัน ฟังก์ชันจะมีตัวเลือกการเข้ารหัสที่แตกต่างกันไป โดยทั่วไปแล้วจะเป็นเลเยอร์โครงข่ายประสาทเทียมที่มีการ Activation Functions

การ Decoder

$$\text{Decoding: } \hat{X} = fdecoder(Z) \tag{3.2}$$

จากสมการนี้ $fdecoder$ เป็นฟังก์ชันการถอดรหัสข้อมูล ด้วยการ Map ภาพให้ตรงกันก่อนส่งกลับไปยังข้อมูลที่สร้างขึ้นใหม่

Output $\hat{X}$ (ข้อมูลที่สร้างขึ้นใหม่)

$$L(\hat{X}, X) = \frac{1}{n}\sum_{i=1}^{n}(\hat{X}_i - X_i)^2 \quad (3.3)$$

กำหนดให้

$L(\hat{X}, X)$ คือ ค่า Loss

$\hat{X}_i$ คือ ส่วนประกอบที่อยู่ในข้อมูลที่สร้างขึ้นใหม่ $\hat{X}$

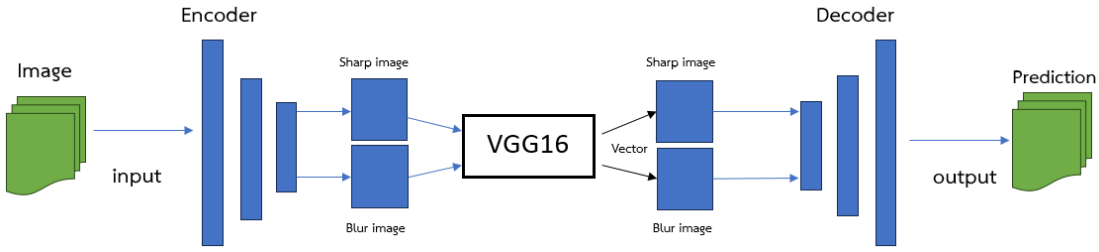
X_i คือ ส่วนประกอบที่อยู่ในการอินพุตข้อมูล X

n คือ จำนวนจุดของข้อมูล

MSE จะเป็นตัวที่ช่วยในการวัดปริมาณความแตกต่างของค่าเฉลี่ยที่เกิดขึ้นระหว่างภาพที่ถูกสร้างขึ้นใหม่และภาพอินพุต และยังช่วยลด Noise ให้เหลือน้อยที่สุดระหว่างการฝึก (Train) รวมถึงยังช่วยส่งเสริมให้การทำ Autoencoder สามารถประมวลผลได้อย่างแม่นยำ

วิธีการทดลองที่ 2 การทำ Autoencoder ด้วย VGG16 ในวิธีการนี้จะเอาใช้ภาพที่ Input เข้ามากระบวนการเข้ารหัส การบีบอัดภาพ หลังจากนั้นจะได้ภาพ Predict ออกมา ก่อนจะนำภาพชัดกับภาพที่ถูก Predict ส่งเข้าไปในเลเยอร์ของ VGG16 เพื่อแปลงภาพนั้นให้เป็น Vector ก่อนที่จะนำไปคำนวณและส่งไปยังขั้นตอนการถอดรหัส แล้วจึงได้ภาพออกมา โดยสมการของ VGG16 มีดังต่อไปนี้

Perceptual VGG16 เป็นการนำภาพที่ได้จากการทำนาย (Prediction) และภาพชัด เข้าสู่ VGG16 เพื่อทำการสร้าง Feature map ให้กับภาพ โดยจะกรองสิ่งที่ไม่สำคัญออกจากภาพจนเหลือแต่ส่วนที่สำคัญ (Perceptual) ก่อนที่จะนำไปคำนวณหาค่า loss ซึ่งเครือข่ายนี้สามารถใช้เพื่อวัดความคล้ายคลึงกันระหว่างภาพได้ VGG16 มักใช้ในการถ่ายโอนรูปแบบประสาทและงานการสร้างภาพ อีกทั้งยังใช้วัดความแตกต่างในการแสดงคุณลักษณะระหว่างภาพต้นฉบับและภาพที่สร้างขึ้นใหม่ โดยมีสูตรที่ใช้ดังสมการต่อไปนี้



ภาพที่ 3.3 วิธีการทำงานของการทดลองด้วย Autoencoder รวมกับ VGG16

$$L_{VGG}(X, \hat{X}) = \frac{1}{N} \sum_{i=1}^{N} \| f_i(X) - f_i(\hat{X}) \|^2 \quad (3.4)$$

กำหนดให้

$L_{VGG}(X, \hat{X})$ คือ VGG16 perceptual loss

X คือ ภาพต้นฉบับ

$\hat{X}$ คือ ภาพที่ถูกสร้างขึ้นมาใหม่

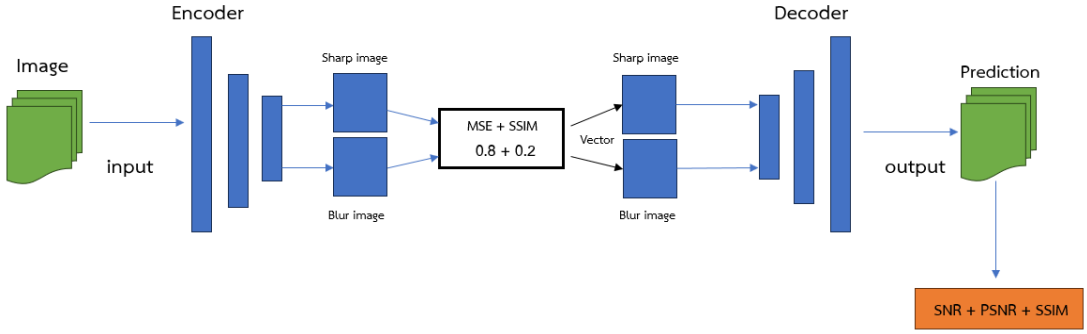
N คือจำนวนเลเยอร์ที่ใช้ในโมเดล VGG16 สำหรับการแยกคุณลักษณะ

$f_i(X)$ คือ การแสดงคุณลักษณะของภาพต้นฉบับ X ที่เลเยอร์ i

$f_i(\hat{X})$ คือ การนำเสนอคุณลักษณะของรูปภาพ X ที่สร้างขึ้นใหม่ที่ชั้น i

วิธีการทดลองที่ 3 การทำ Autoencoder ด้วยการปรับค่าแบบ Customize โดยใช้ MSE รวมกับดัชนีความคล้ายคลึงกันของโครงสร้างภาพ (SSIM)

วิธีนี้จะสามารถปรับค่าพารามิเตอร์ได้โดยการใช้ Customized Loss ซึ่งมีการเรียกใช้ฟังก์ชันโดยเพิ่มการเปรียบเทียบความสว่างของภาพ เปรียบเทียบค่าคอนทราสต์ และโครงสร้างความคล้ายคลึงของภาพ (SSIM) เข้ามา เพื่อเน้นให้การทำปรับโครงสร้างของภาพมีความชัดเจนขึ้น



ภาพที่ 3.4 แสดงวิธีการทำงานของการทำ Autoencoder ด้วย MSE รวมกับ SSIM

ซึ่งดัชนีความคล้ายคลึงกันของโครงสร้าง (SSIM) เป็นอีกหนึ่งตัวชี้วัดที่ถูกนำมาใช้กันอย่างแพร่หลายในการเปรียบเทียบความคล้ายคลึงกันของโครงสร้างระหว่างภาพสองภาพ โดยทั่วไปจะใช้สำหรับการประเมินคุณภาพของภาพ โดยจะวัดความคล้ายคลึงกันในเรื่องของความสว่าง คอนทราสต์ และโครงสร้างระหว่างภาพต้นฉบับกับภาพที่ถูกสร้างขึ้นใหม่ อย่างไรก็ตาม SSIM ไม่ใช่ฟังก์ชันการสูญเสียมาตรฐานสำหรับการฝึกตัวเข้ารหัสอัตโนมัติ เนื่องจากโดยทั่วไปจะใช้สำหรับการประเมินมากกว่าการฝึก

สูตร SSIM นั้นค่อนข้างซับซ้อนเนื่องจากจะเกี่ยวข้องกับองค์ประกอบของการเปรียบเทียบความสว่าง การเปรียบเทียบคอนทราสต์ และการเปรียบเทียบโครงสร้างภาพ โดยมีสูตรที่ใช้ในการคำนวณดังต่อไปนี้

$$SSIM(X, \hat{X}) = \mathrm{l}(X, \hat{X}) \cdot \mathrm{c}(X, \hat{X}) \cdot \mathrm{s}(X, \hat{X}) \quad (3.5)$$

กำหนดให้

X คือ ภาพต้นฉบับ

$\hat{X}$ คือ ภาพที่ถูกสร้างขึ้นมาใหม่

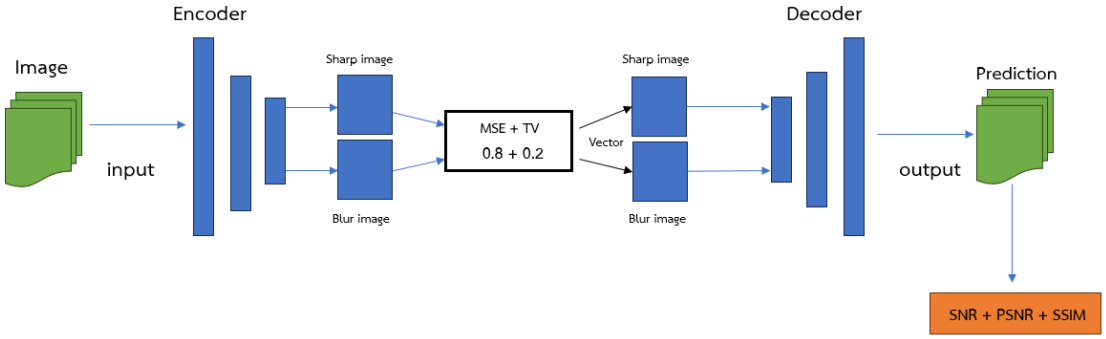
$\mathrm{l}(X, \hat{X})$ คือ การเปรียบเทียบความสว่าง

$\mathrm{c}(X, \hat{X})$ คือ การเปรียบเทียบค่าคอนทราสต์

$\mathrm{s}(X, \hat{X})$ คือ การเปรียบเทียบโครงสร้างภาพ

วิธีการทดลองที่ 4 การทำ Autoencoder ด้วยการปรับจูนค่าแบบ Manual โดยใช้ MSE รวมกับ Total Variance (TV)

ในวิธีนี้จะสามารถปรับค่าพารามิเตอร์ได้โดยการใช้ Customized Loss ซึ่งจะนำค่าความแปรปรวนรวม (TV) ที่มีลักษณะของการหาค่าเฉลี่ยของแต่ละพิกเซลก่อนนำมารวมกันเพื่อหาค่าเฉลี่ยรวม มาทำงานรวมกับค่า MSE ในการปรับจูนค่าพารามิเตอร์ให้มีค่าที่ดีที่สุด



ภาพที่ 3.5 แสดงวิธีการทำงานของการทำ Autoencoder ด้วย MSE รวมกับ Total Variance (TV)

Total Variance หรือค่าความแปรปรวนรวม โดยทั่วไปแล้วมักจะไม่ใช้ Loss Function ในการ Train ข้อมูล สำหรับการทำ Autoencoder แต่จะใช้เป็นการวัดความแปรปรวนของแต่ละพิกเซลเพื่อให้ทราบถึงค่าเฉลี่ยโดยรวม ซึ่งเหมาะสำหรับการวิเคราะห์ข้อมูล โดยมีสูตรที่ใช้ในการคำนวณดังต่อไปนี้

$$TV = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2 \quad (3.6)$$

กำหนดให้

TV คือ ค่าความแปรปรวนรวม

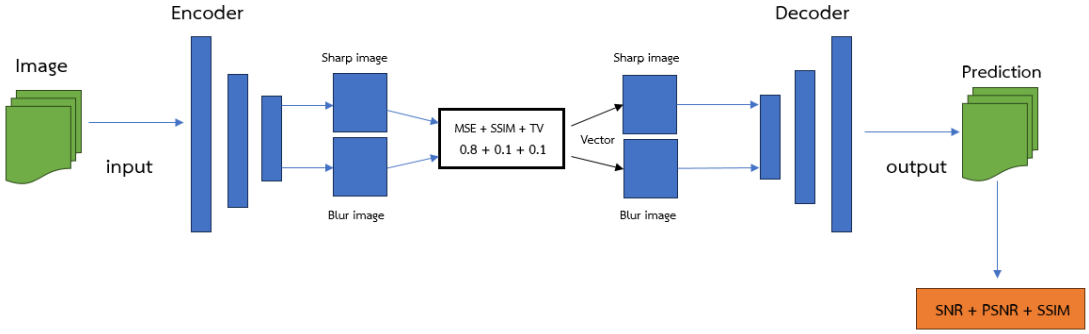
n คือ จำนวนจุดของข้อมูล

X_i คือ จุดของข้อมูลเริ่มต้นตั้งแต่ i-th

μ คือค่าเฉลี่ยของจุดข้อมูล

วิธีการทดลองที่ 5 การทำ Autoencoder ด้วยการปรับจูนค่าแบบ Manual โดยใช้ MSE รวมกับ SSIM และ Total Variance (TV)

ในวิธีการนี้ได้เพิ่มจำนวนสมการเข้ามาใช้งานเพื่อทดลองว่าผลลัพธ์ที่ได้นั้นจะมีค่าสูงขึ้นหรือไม่ เมื่อผ่านการวัดประสิทธิภาพ โดยวิธีนี้จะใช้การถ่วงน้ำหนักเพื่อปรับความสมดุลในการ Train ข้อมูลให้ได้โมเดลที่ดีที่สุดก่อนที่จะนำไป Test เพื่อให้ได้ผลลัพธ์ที่ดีขึ้น สำหรับการปรับจูนค่าพารามิเตอร์ (Customized Loss) ได้นำสมการของ SSIM และ TV เข้ามาเพื่อจะใช้ในการปรับโครงสร้างของภาพและค่าเฉลี่ยของภาพในแต่ละพิกเซลให้ดีขึ้น โดยมีสูตรที่ใช้ในการคำนวณดังต่อไปนี้



ภาพที่ 3.6 แสดงวิธีการทำงานของการทำ Autoencoder ด้วย MSE รวมกับ SSIM และ Total Variance (TV)

โดยจะใช้สูตรในการคำนวณของ MSE รวมกับ SSIM และ Total Variance (TV) มีรายละเอียดดังนี้

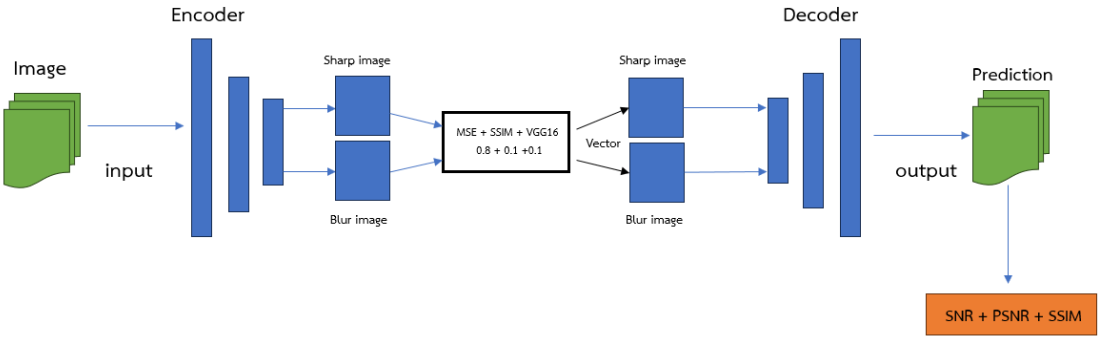
$$L(\hat{X}, X) = \frac{1}{n}\sum_{i=1}^{n}\left(\widehat{X_i} - X_i\right)^2 \quad (3.3)$$

$$SSIM(X, \hat{X}) = l(X, \hat{X}) \cdot c(X, \hat{X}) \cdot s(X, \hat{X}) \quad (3.5)$$

$$TV = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \quad (3.6)$$

วิธีการทดลองที่ 6 การทำ Autoencoder ด้วยการปรับจูนค่าแบบ Manual โดย MSE + SSIM + VGG16

สำหรับวิธีนี้จะใช้ MSE ในการควบคุมค่า Loss เป็นหลักและใช้ SSIM ในการเปรียบเทียบโครงสร้างภาพ ก่อนจะใช้ VGG16 ทำหน้าที่กรองสิ่งที่ไม่สำคัญออกจากภาพ โดยกำหนดค่าพารามิเตอร์ที่ดีที่สุด



ภาพที่ 3.7 แสดงวิธีการทำงานของการทำ Autoencoder ด้วย MSE รวมกับ SSIM และ VGG16

โดยจะใช้สูตรในการคำนวณของ MSE รวมกับ SSIM และ VGG16 มีรายละเอียดดังนี้

$$L(\hat{X}, X) = \frac{1}{n}\sum_{i=1}^{n}(\widehat{X_i} - X_i)^2 \quad (3.3)$$

$$SSIM(X, \hat{X}) = l(X, \hat{X}) \cdot c(X, \hat{X}) \cdot s(X, \hat{X}) \quad (3.5)$$

$$L_{VGG}(X, \hat{X}) = \frac{1}{N}\sum_{i=1}^{N}\|f_i(X) - f_i(\hat{X})\|^2 \quad (3.4)$$

วิธีการทดลองที่ 7 การทำ Autoencoder ด้วยการปรับจูนค่าแบบ Manual โดย MSE + SSIM + TV + VGG16

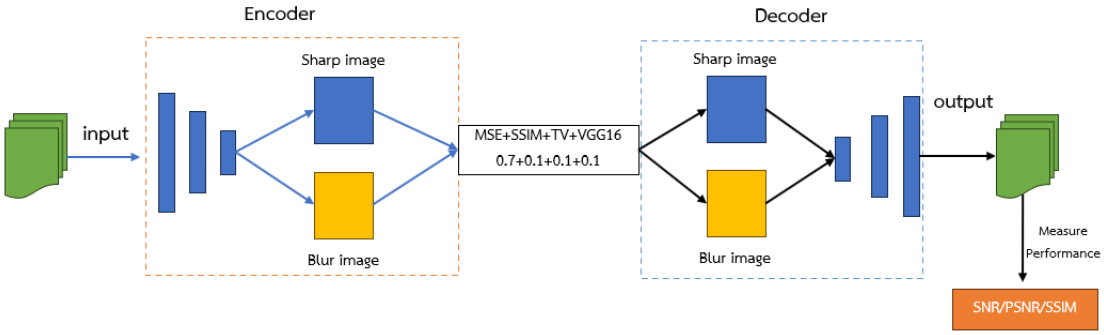
สำหรับวิธีนี้จะใช้วิธีการปรับปรุงภาพในมิติต่างๆ โดยใช้สมการทั้งหมดที่เคยทดลองมารวมกันเพื่อทำให้ภาพที่ได้มีความคมชัดมากขึ้น ไม่ว่าจะเป็นการลดค่า Loss ให้มีค่าน้อยลง การปรับปรุงโครงสร้างภาพ ค่าContrast ของภาพ ค่าเฉลี่ยของพิกเซลโดยรวม รวมไปถึงการแบ่งส่วนเลเยอร์ในการคำนวณ โดยสามารถนำมาเขียนเป็นสมการได้ดังต่อไปนี้

$$L(\hat{X}, X) = \frac{1}{n}\sum_{i=1}^{n}(\widehat{X_i} - X_i)^2 \quad (3.3)$$

$$SSIM(X, \hat{X}) = l(X, \hat{X}) \cdot c(X, \hat{X}) \cdot s(X, \hat{X}) \quad (3.5)$$

$$TV = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \quad (3.6)$$

$$L_{VGG}(X, \hat{X}) = \frac{1}{N}\sum_{i=1}^{N}\|f_i(X) - f_i(\hat{X})\|^2 \quad (3.4)$$



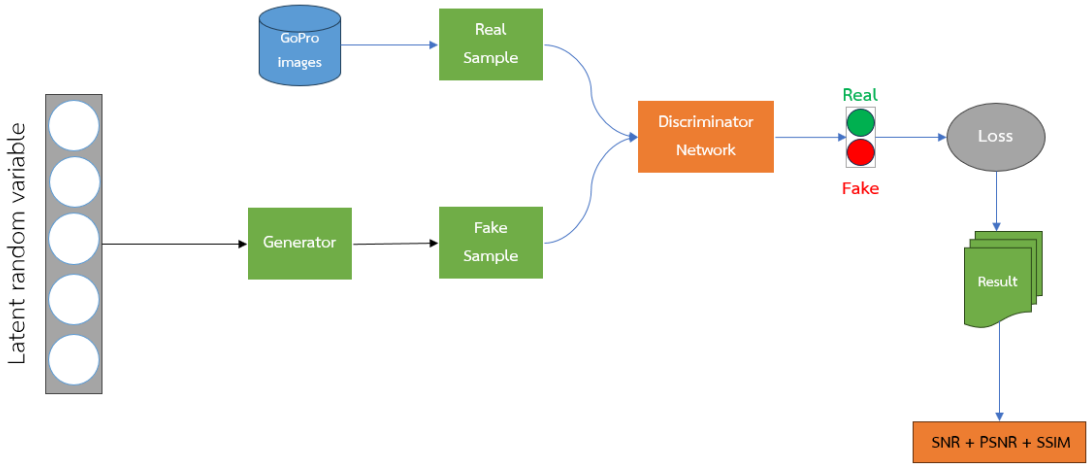
ภาพที่ 3.8 โครงสร้างการทำงานของสมการทั้งหมดที่ใช้ในการทดลอง

จากสมการดังกล่าวสามารถนำมาเขียนเป็นโครงสร้างการทำงานได้ดังต่อไปนี้

จากภาพที่ 3.8 เป็นการรวมสมการทั้งหมดเพื่อใช้ในการประมลผลภาพ ซึ่งจะทำให้ผลลัพธ์ที่ได้มีภาพที่มีความเบลอลดน้อยลงและมีความคมชัดมากขึ้น

3.2.2 Generative Adversarial Networks: GANs

ในส่วนของ Generative Adversarial Networks : GANs จะถูกแบ่งการทำงานออกเป็น 2 ส่วน คือ ตัวที่ใช้ในการสร้างภาพหรือที่เรียกว่า Generator และอีกอันคือ Discriminator ดังนั้นในการทำงานของทั้งสองนี้ก็จะทำงานรวมกัน ดังนั้นวิธีการทดลองที่ 8 จะใช้การสร้างข้อมูลที่คล้ายกับชุดข้อมูลที่กำหนดซึ่งเริ่มจาก Generator จะ Random noise แล้วนำมาสร้างภาพ แล้วส่งไปยังตัว Discriminator เพื่อทำการตรวจสอบ โดยจะทำการ Test กับภาพที่เป็นภาพจริง ก่อนนำมาเปรียบเทียบกับภาพที่ถูกสร้างขึ้นมาใหม่เพื่อให้รู้ว่าภาพนั้นเป็นภาพจริงหรือภาพเท็จ ซึ่งจะมีสมการและวิธีการออกแบบดังต่อไปนี้



ภาพที่ 3.9 แสดงวิธีการทำงานของ Generative Adversarial Networks: GANs

$$\text{Generator (G): } G(z;\theta_g) \rightarrow X \tag{3.8}$$

กำหนดให้

G คือ การสร้างข้อมูล

z คือ เวกเตอร์สัญญาณรบกวนที่อินพุต

θ_g คือ ค่าพารามิเตอร์

X คือ การสร้างข้อมูลตัวอย่าง

$$\text{Discriminator (D): } D(x;\theta_d) \rightarrow Probability \tag{3.9}$$

กำหนดให้

D คือ การสร้างข้อมูลออก

x คือ ข้อมูลเข้า

θ_d คือ ค่าพารามิเตอร์

Probability คือ ความน่าจะเป็นที่เป็นตัวอย่างข้อมูลจริง

3.3 ผลการวัดประสิทธิภาพของภาพ (Measurement Performance)

หลังจากได้ออกแบบวิธีที่ใช้สำหรับการปรับปรุงภาพเบลอด้วยอัลกอริทึมของ Autoencoder และ Generative Adversarial Networks: GAN ด้วยวิธีการทดลองทั้งหมด 8 วิธี ในส่วนนี้จะเป็นการอธิบายถึงวิธีการทดลอง การเซ็ตค่าพารามิเตอร์ และการวัดประสิทธิภาพ

วิธีที่ 1 การทำ Autoencoder

ในวิธีนี้จะเป็นการนำโครงสร้างของ Autoencoder ที่ประกอบไปด้วย การเข้ารหัสภาพที่อินพุต (Encoder) เข้ามา ก่อนจะแบ่งภาพออกเป็นเลเยอร์ย่อยๆ จากนั้นทำการลดสัญญาณรบกวน (Noise) ที่อยู่ในภาพออก และลบค่าที่ไม่จำเป็นออกเพื่อเก็บ Feature map ที่สำคัญของภาพไว้ จากนั้นทำการบีบอัดภาพแล้วสร้างภาพใหม่ที่มีขนาดเล็กขึ้นมา หลังจากนั้นทำการถอดรหัสภาพ (Decoder) ด้วยการแปลงภาพที่ถูกเข้ารหัส (Encoder) ให้กลับให้มีขนาดเท่าเดิม แล้วจึงได้ภาพผลลัพธ์ที่มีความชัดเจนขึ้น โดยมีสูตรที่ใช้ในการคำนวณจากสมการที่ 3.1-3.3

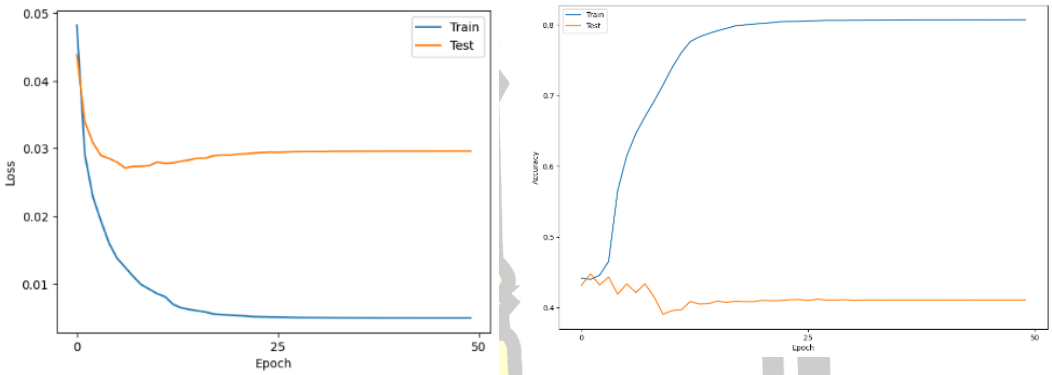
ในการทดลองนี้ผู้วิจัยได้ใช้ชุดข้อมูลของ GoPro-Large ซึ่งแบ่งเป็นภาพ 2 ประเภทคือ Blur และ Blur_gamma การกำหนดค่าพารามิเตอร์จะกำหนดค่า Loss Function เท่ากับ MSE จากนั้นนำชุดข้อมูลภาพที่ใช้สำหรับการ Train จำนวน 2,301 ภาพ มา Train จำนวน 50 รอบ หลังจากนั้นนำโมเดลที่ได้ไป Test กับชุดข้อมูลภาพที่ใช้ในการ Test จำนวน 1,111 ภาพ จนได้ภาพผลลัพธ์ที่เกิดจากการทำนาย (Prediction) โดยใช้เวลาประมาณ 2.35 ชั่วโมง หลังจากนั้นนำภาพที่ได้มาวัดประสิทธิภาพของภาพ ทั้ง 3 ด้าน ได้แก่ ค่าอัตราส่วนสัญญาณต่อสัญญาณรบกวน: Signal-to-Noise Ratio (SNR), หาค่าอัตราส่วนของสัญญาณรบกวนสูงสุด: Peak Signal-to-Noise Ratio (PSNR) และวัดประสิทธิภาพของเทคนิคทางด้านภาพ: Structural Similarity Index Measure (SSIM) โดยมีรายละเอียดดังต่อไปนี้

ตารางที่ 3.2 แสดงผลลัพธ์ที่เกิดจากวิธีที่ 1 การทำ Autoencoder

ประเภทของภาพเบลอ	การวัดประสิทธิภาพของภาพ		
	SNR	PSNR	SSIM
Blur	32.53	22.00	0.771
Blur_gamma	32.91	22.20	0.770

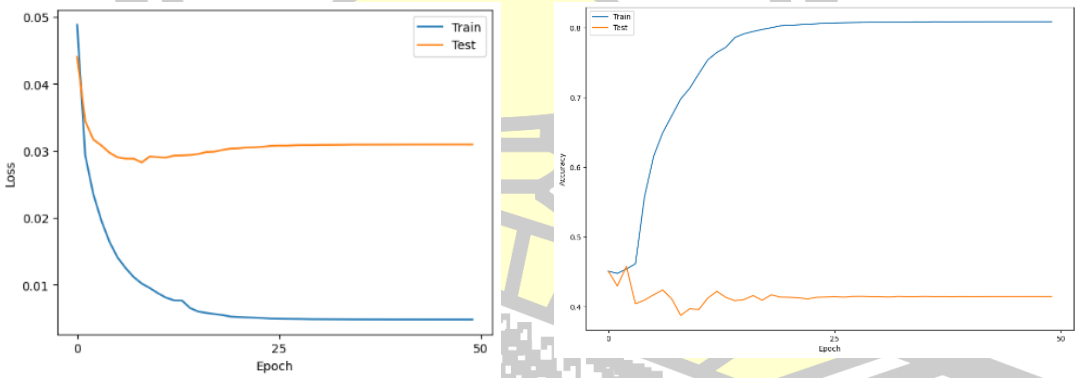
ผลลัพธ์ที่เกิดจากการทดลองโดยวิธีการของ Autoencoder จะใช้เป็นฐาน (Baseline) ในการเปรียบเทียบการวัดประสิทธิภาพกับวิธีการทดลองที่เหลืออีกทั้ง 6 วิธี

กราฟแสดงค่า Loss Function และค่า Accuracy ที่เกิดจากการ Train และ Test ในการทดลองในวิธีที่ 1



ภาพที่ 3.10 แสดงค่า Loss Function และค่า Accuracy ของภาพประเภท Blur

จากภาพที่ 3.10 แสดงให้เห็นถึงการทำงานของฟังก์กัน เมื่อเริ่มต้น Train ค่า Loss ในแต่ละรอบจะเริ่มลดลงจนถึงรอบที่ 50 ค่า Loss ถ้าต่ำลงมากเท่าไหร่ ภาพที่ได้ก็จะมีค่าใกล้เคียงกันมากเท่านั้น ในส่วนของค่า Accuracy จะเป็นการค่าความถูกต้องของโมเดล



ภาพที่ 3.11 แสดงค่า Loss Function และค่า Accuracy ของภาพประเภท Blur_gamma

จากภาพที่ 3.10 และภาพที่ 3.11 แสดงให้เห็นว่ากราฟของทั้ง 2 ประเภทมีความคล้ายคลึงกัน เนื่องจากผลที่เกิดจากการวัดประสิทธิภาพของภาพทั้ง 2 ประเภทมีความใกล้เคียงกัน กราฟที่ได้จึงออกมาในรูปแบบนี้

วิธีที่ 2 การทำ Autoencoder รวมกับ VGG16

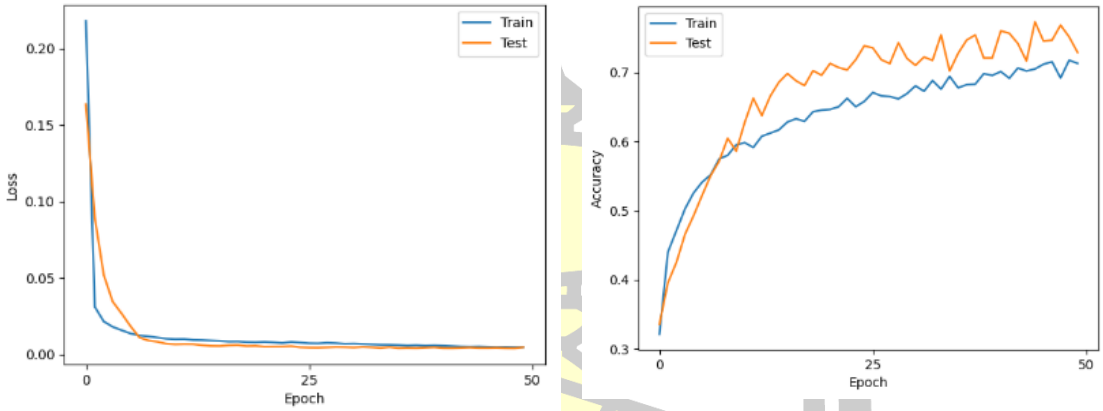
ในวิธีการนี้จะใช้วิธีการทำงานของ Autoencoder เป็นหลัก โดยจะนำภาพชัด (Sharp) กับภาพเบลอ (Blur) ที่ผ่านการ Encoder และกระบวนการบีบอัดภาพแล้ว ก่อนส่งภาพทั้ง 2 ไปยัง

ฟังก์ชันการทำงานของ VGG16 ที่แทรกเข้ามา ขั้นตอนนี้ VGG16 จะแปลงภาพให้เป็น Vector เพื่อนำมาคำนวณหาค่า Loss หลังจากนั้นจึงส่งไปยังการ Decoder เพื่อแปลงภาพผลลัพธ์ให้มีขนาดเท่ากับภาพที่อินพุตเข้ามา โดยใช้สูตรการคำนวณของ Autoencoder รวมกับสมการที่ 3.4 ก่อนที่จะนำมาภาพผลลัพธ์มาวัดประสิทธิภาพของภาพทั้ง 3 ด้าน (SNR, PSNR, SSIM) โดยมีวิธีที่ใช้ในการคำนวณดังภาพต่อไปนี้

ตารางที่ 3.3 แสดงผลลัพธ์ที่เกิดจากวิธีที่ 2 การทำ Autoencoder รวมกับ VGG16

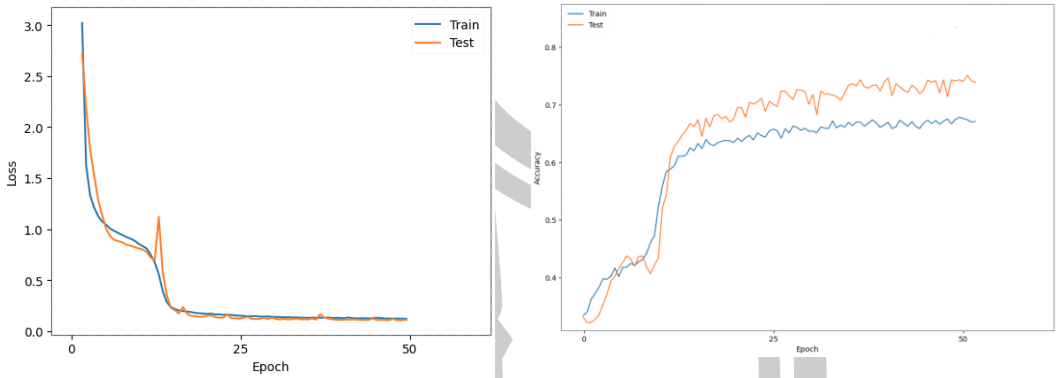
ประเภทของภาพเบลอ	การวัดประสิทธิภาพของภาพ		
	SNR	PSNR	SSIM
Blur	32.55	22.02	0.757
Blur_gamma	31.82	21.65	0.753

กราฟแสดงค่า Loss Function และค่า Accuracy ที่เกิดจากการ Train และ Test ของการทดลองในวิธีที่ 2



ภาพที่ 3.12 แสดงค่า Loss Function และค่า Accuracy ของภาพประเภท Blur ของการทำ Autoencoder รวมกับ VGG16

จากภาพที่ 3.12 กราฟที่แสดงค่า Loss มีลดลงทั้งการ Train และการ Test โดยมีค่า Loss เท่ากับ 0.051 แสดงว่าภาพที่ได้จะมีค่าใกล้เคียงกันกันมากขึ้น ส่วนในกราฟของ Accuracy จะมีค่าความถูกต้องของโมเดลที่อยู่ที่ 0.721



ภาพที่ 3.13 แสดงค่า Loss Function และค่า Accuracy ของภาพประเภท Blur_gamma โดยใช้การทำ Autoencoder รวมกับ VGG16

จากภาพที่ 3.13 กราฟที่แสดงค่า Loss เมื่อ Train Model ได้ประมาณรอบที่ 18 แสดงให้เห็นว่ากราฟมีแนวโน้มลดลงจนทำให้ค่า Loss อยู่ระหว่าง 0 - 0.5 และในส่วนของกราฟ Accuracy ก็ส่งผลให้ค่าความถูกต้องของโมเดลอยู่ระหว่าง 0.7 - 0.8

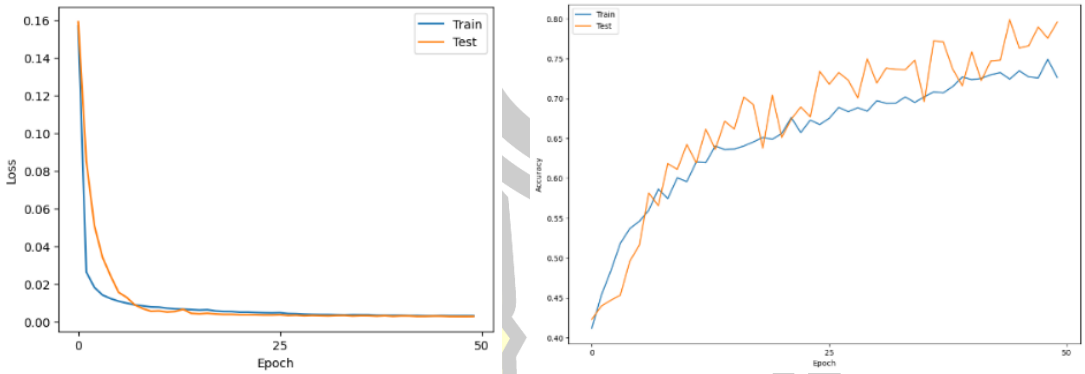
วิธีที่ 3 การทำ Autoencoder ด้วย MSE รวมกับ SSIM

ในวิธีการนี้จะใช้การกำหนดค่าฟังก์ชันแบบ Customized Loss เพื่อปรับค่าพารามิเตอร์ของ MSE โดยเพิ่มค่า SSIM เข้ามาช่วยในการคำนวณ และใช้การให้มีความเหมาะสมซึ่งการทดลองนี้จะ Weighting น้ำหนักเพื่อปรับให้ค่า Loss ลดลง ซึ่งค่าที่ดีที่สุดคือ MSE = 0.8 และ SSIM = 0.2 ก่อนที่จะนำไป Train Model โดยเพิ่มสูตรการคำนวณจากสมการที่ 3.6 เข้ามาจึงสามารถแสดงผลลัพธ์ที่ได้ดังต่อไปนี้

ตารางที่ 3.4 ผลลัพธ์ที่ได้จากการทดลองด้วยวิธีการทดลองที่ 3 ด้วย MSE + SSIM

ประเภทของภาพเบลอ	การวัดประสิทธิภาพของภาพ		
	SNR	PSNR	SSIM
Blur	31.87	21.71	0.742
Blur_gamma	31.67	21.57	0.763

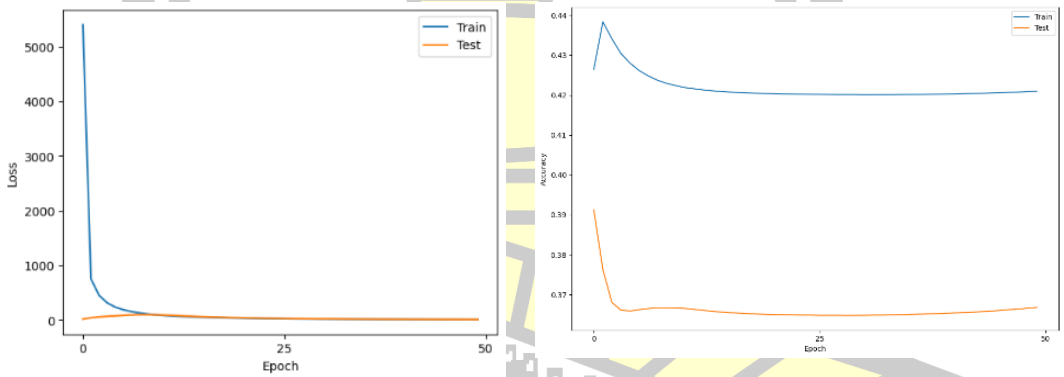
กราฟแสดงค่า Loss Function และค่า Accuracy ที่เกิดจากการ Train และ Test ของการทดลองในวิธีที่ 3 โดยใช้ภาพประเภท Blur



ภาพที่ 3.14 แสดงค่า Loss Function และค่า Accuracy ของภาพประเภท Blur โดยการกำหนดค่า MSE และ SSIM

จากภาพที่ 3.14 กราฟแสดงค่า Loss ของภาพประเภท Blur ทำให้เห็นว่าค่า Loss มีค่าลดลงตามลำดับ ในการ Train ทั้งหมด 50 รอบ หากใช้จำนวนการ Train มากขึ้นมีแนวโน้มที่ค่า Loss จะลดลงอย่างมีนัยสำคัญ สำหรับกราฟที่แสดงค่า Accuracy แสดงให้เห็นค่าความถูกต้องของการ Train Model ซึ่งในการทดลองนี้จะมีแนวโน้มสูงถ้ามีการ Train เพิ่มมากขึ้น

กราฟแสดงค่า Loss Function และค่า Accuracy ที่เกิดจากการ Train และ Test ของการทดลองในวิธีที่ 3 โดยใช้ภาพประเภท Blur_gamma



ภาพที่ 3.15 แสดงค่า Loss Function และค่า Accuracy ของภาพประเภท Blur_gamma โดยการกำหนดค่า MSE และ SSIM

จากภาพที่ 3.15 กราฟค่า Loss มีค่าลดลงเมื่อใช้วิธีการทดลองแบบนี้ ส่วนค่า Accuracy มีความถูกต้องของโมเดลอยู่ในระดับคงที่

วิธีที่ 4 การทำ Autoencoder ด้วย MSE รวมกับ Total Variance (TV)

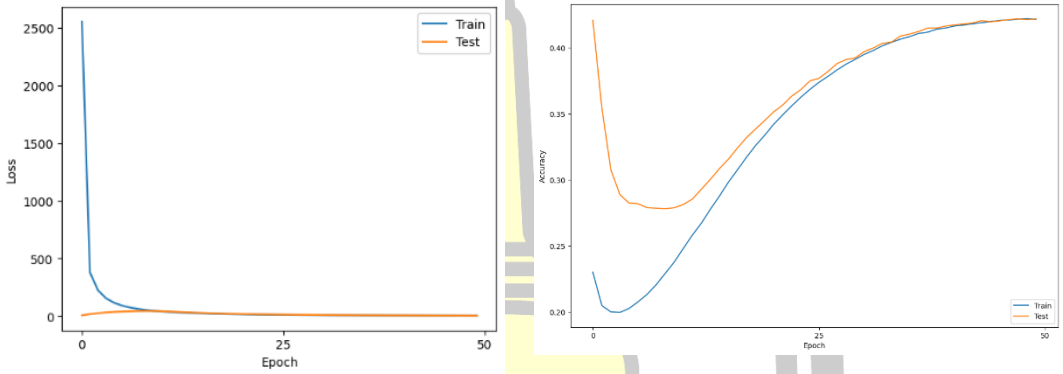
ในวิธีการนี้จะใช้การกำหนดค่าฟังก์ชันแบบ Customized Loss เพื่อปรับค่าพารามิเตอร์ของ MSE และค่า Total Variance (TV) ซึ่งการทดลองนี้จะ Weighting น้ำหนักให้มีความสมดุล เพื่อทำ

ให้ค่า Loss ลดลง ซึ่งค่าที่ปรับได้ดีที่สุดคือ MSE = 0.8 และ TV = 0.2 ก่อนที่จะนำไป Train Model โดยเพิ่มสูตรการคำนวณจากสมการที่ 3.7 เข้ามาจึงสามารถแสดงผลลัพธ์ที่ได้ดังต่อไปนี้

ตารางที่ 3.5 ผลลัพธ์ที่ได้จากวิธีที่ 4 ด้วย MSE + Total Variance (TV)

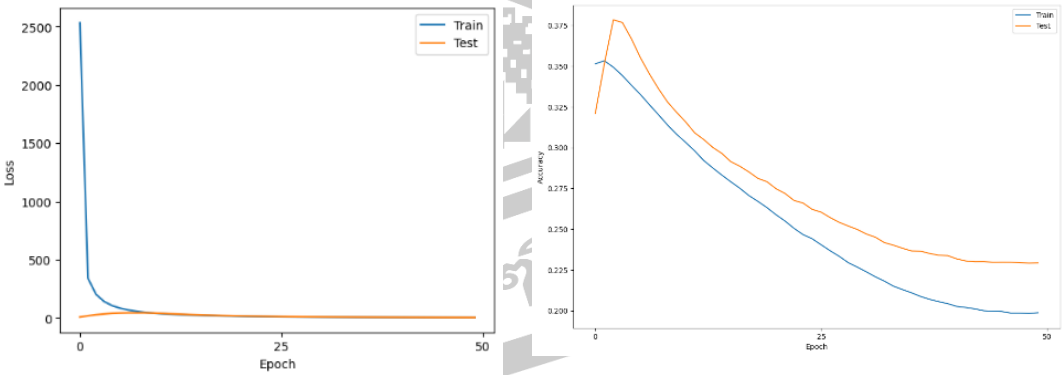
ประเภทของภาพเบลอ	การวัดประสิทธิภาพของภาพ		
	SNR	PSNR	SSIM
Blur	14.05	12.76	0.575
Blur_gamma	13.99	12.57	0.561

กราฟแสดงค่า Loss Function และค่า Accuracy ที่เกิดจากการ Train และ Test ของการทดลองในวิธีที่ 4 โดยใช้ภาพประเภท Blur



ภาพที่ 3.16 แสดงค่า Loss Function และค่า Accuracy ของภาพประเภท Blur โดยการการกำหนดค่า MSE และ Total Variance

กราฟแสดงค่า Loss Function และค่า Accuracy ที่เกิดจากการ Train และ Test ของการทดลองในวิธีที่ 4 โดยใช้ภาพประเภท Blur_gamma



ภาพที่ 3.17 แสดงค่า Loss Function และค่า Accuracy ของภาพประเภท Blur_gamma โดยการกำหนดค่า MSE และ Total Variance

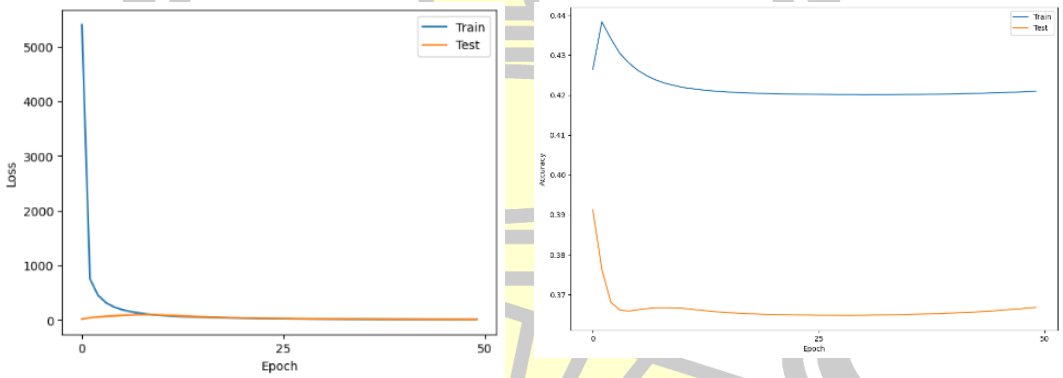
วิธีที่ 5 การทำ Autoencoder ด้วย MSE รวมกับ SSIM และ Total Variance (TV)

ในวิธีการนี้จะใช้การกำหนดค่าฟังก์ชันแบบ Customized Loss เพื่อปรับค่าพารามิเตอร์ของ MSE ซึ่งรวมทั้ง SSIM และค่า Total Variance (TV) ในการทดลองนี้จะ Weighting น้ำหนักให้มีความสมดุล เพื่อทำให้ค่า Loss ลดลง ซึ่งค่าที่ปรับได้ดีที่สุดคือ MSE = 0.8 SSIM = 0.1 TV = 0.1 ก่อนที่จะนำไป Train Model โดยใช้สูตรการคำนวณจากสมการที่ 3.5 – 3.7 เข้ามาจึงสามารถแสดงผลลัพธ์ที่ได้ดังต่อไปนี้

ตารางที่ 3.6 ผลลัพธ์ที่ได้จากวิธีที่ 5 ด้วย MSE รวมกับ SSIM และ Total Variance (TV)

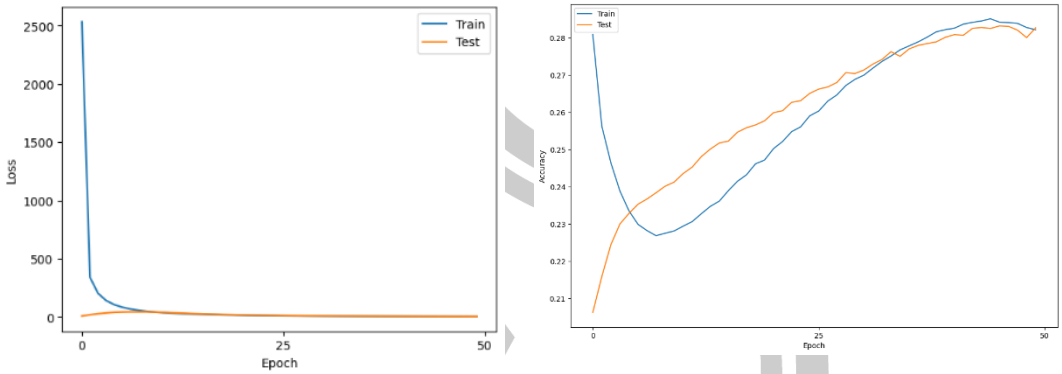
ประเภทของภาพเบลอ	การวัดประสิทธิภาพของภาพ		
	SNR	PSNR	SSIM
Blur	10.26	11.36	0.600
Blur_gamma	11.07	12.76	0.609

กราฟแสดงค่า Loss Function และค่า Accuracy ที่เกิดจากการ Train และ Test ของการทดลองในวิธีที่ 5 โดยใช้ภาพประเภท Blur



ภาพที่ 3.18 แสดงค่า Loss Function และค่า Accuracy ของภาพประเภท Blur โดยการกำหนดค่า MSE และ SSIM และTotal Variance (TV)

กราฟแสดงค่า Loss Function และค่า Accuracy ที่เกิดจากการ Train และ Test ของการทดลองในวิธีที่ 5 โดยใช้ภาพประเภท Blur_gamma



ภาพที่ 3.19 แสดงค่า Loss Function และค่า Accuracy ของภาพประเภท Blur_gamma โดยการกำหนดค่า MSE และ SSIM และTotal Variance (TV)

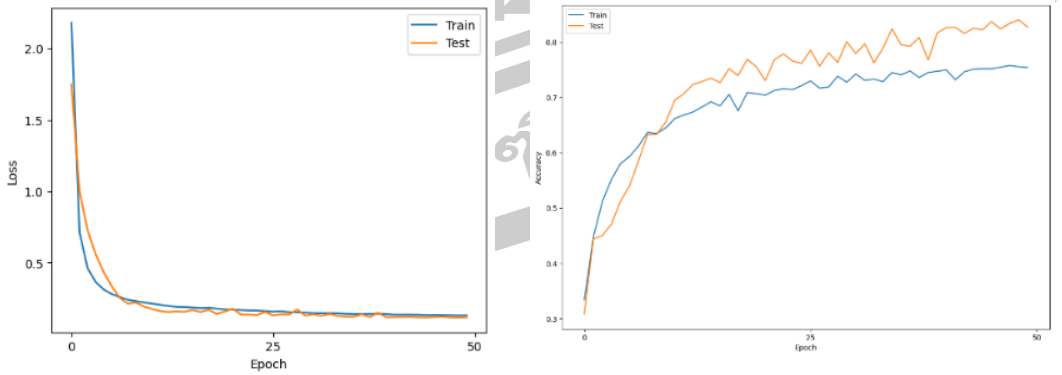
วิธีที่ 6 การทำ Autoencoder ด้วย MSE รวมกับ SSIM และ VGG16

ในวิธีการนี้จะใช้การกำหนดค่าฟังก์ชันแบบ Customized Loss เพื่อปรับค่าพารามิเตอร์ของ MSE ซึ่งรวมทั้ง SSIM และค่า VGG16 ในการทดลองนี้จะ Weighting น้ำหนักให้มีความสมดุล เพื่อทำให้ค่า Loss ลดลง ซึ่งค่าที่ปรับได้ดีที่สุดคือ MSE = 0.8 SSIM = 0.1 VGG16 = 0.1 ก่อนที่จะนำไป Train Model โดยใช้สูตรการคำนวณจากสมการที่ 3.4 รวมกับ 3.5 และ 3.6 เข้ามารวมกันจึงสามารถแสดงผลลัพธ์ที่ได้ดังต่อไปนี้

ตารางที่ 3.7 ผลลัพธ์ที่ได้จากวิธีที่ 6 ด้วย MSE + SSIM + VGG16

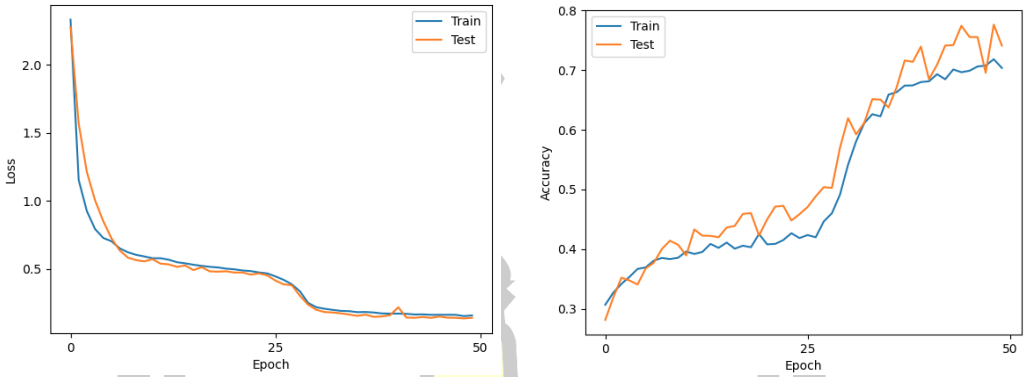
ประเภทของภาพเบลอ	การวัดประสิทธิภาพของภาพ		
	SNR	PSNR	SSIM
Blur	32.64	22.06	0.777
Blur_gamma	31.84	21.66	0.766

กราฟแสดงค่า Loss Function และค่า Accuracy ที่เกิดจากการ Train และ Test ของการทดลองในวิธีที่ 6 โดยใช้ภาพประเภท Blur



ภาพที่ 3.20 แสดงค่า Loss Function และค่า Accuracy ของภาพประเภท Blur โดยการกำหนดค่า MSE และ SSIM และVGG16

กราฟแสดงค่า Loss Function และค่า Accuracy ที่เกิดจากการ Train และ Test ของการทดลองในวิธีที่ 6 โดยใช้ภาพประเภท Blur_gamma



ภาพที่ 3.21 แสดงค่า Loss Function และค่า Accuracy ของภาพประเภท Blur_gamma โดยการกำหนดค่า MSE และ SSIM และVGG16

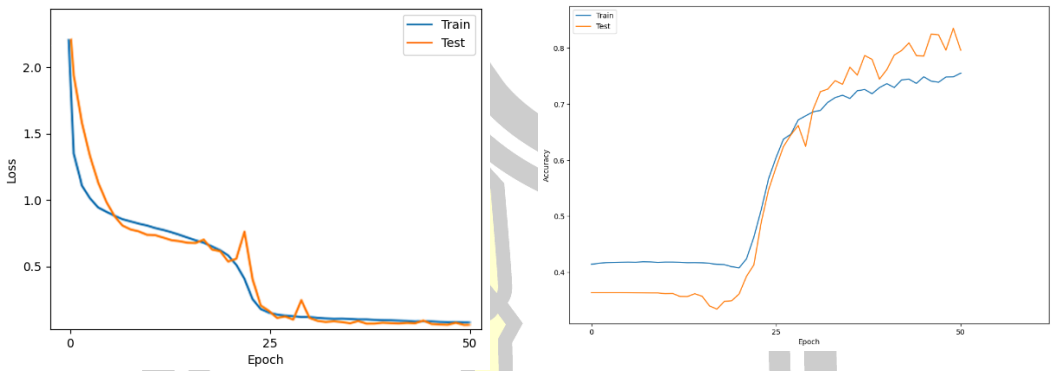
วิธีที่ 7 การทำ Autoencoder ด้วย MSE รวมกับ SSIM รวมกับ TV และ VGG16

ในวิธีการนี้จะใช้การกำหนดค่าฟังก์ชันแบบ Customized Loss เพื่อปรับค่าพารามิเตอร์ของ MSE ซึ่งรวมทั้ง SSIM, TV และค่า VGG16 ในการทดลองนี้จะ Weighting น้ำหนักให้มีความสมดุล เพื่อทำให้ค่า Loss ลดลง ซึ่งค่าที่ปรับได้ดีที่สุดคือ MSE = 0.7 SSIM = 0.1 TV = 0.1 VGG16 = 0.1 ก่อนที่จะนำไป Train Model โดยใช้สูตรการคำนวณจากสมการที่ 3.4, 3.5, 3.6 และ 3.7 เข้ามารวมกันจึงสามารถแสดงผลลัพธ์ที่ได้ดังต่อไปนี้

ตารางที่ 3.8 ผลลัพธ์ที่ได้จากวิธีที่ 7 ด้วย MSE + SSIM + TV + VGG16

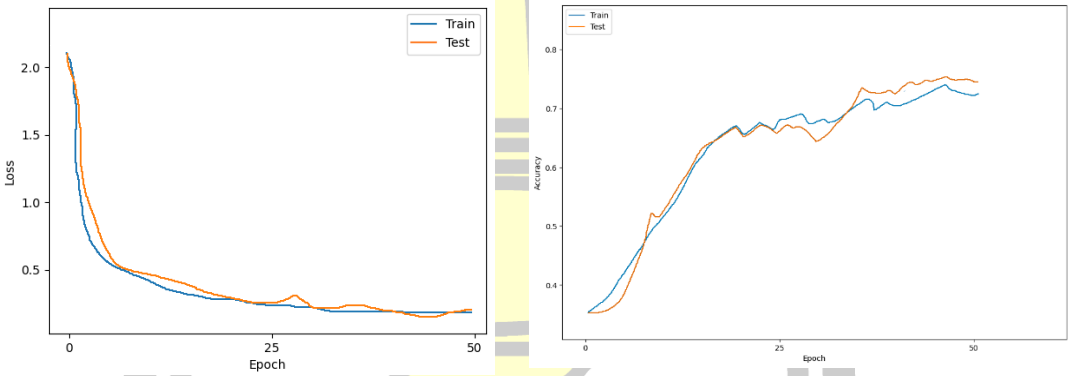
ประเภทของภาพเบลอ	การวัดประสิทธิภาพของภาพ		
	SNR	PSNR	SSIM
Blur	15.14	15.79	0.633
Blur_gamma	15.47	16.05	0.689

กราฟแสดงค่า Loss Function และค่า Accuracy ที่เกิดจากการ Train และ Test ของการทดลองในวิธีที่ 7 โดยใช้ภาพประเภท Blur



ภาพที่ 3.22 แสดงค่า Loss Function และค่า Accuracy ของภาพประเภท Blur โดยการทำกำหนดค่า MSE รวมกับ SSIM และ TV และ VGG16

กราฟแสดงค่า Loss Function และค่า Accuracy ที่เกิดจากการ Train และ Test ของการทดลองในวิธีที่ 7 โดยใช้ภาพประเภท Blur_gamma



ภาพที่ 3.23 แสดงค่า Loss Function และค่า Accuracy ของภาพประเภท Blur_gamma โดยการกำหนดค่า MSE รวมกับ SSIM และ TV และ VGG16

วิธีที่ 8 การทดลองด้วย Generative Adversarial Networks: GANs

สำหรับวิธีการนี้จะใช้วิธีการของ GANs โดยการทำงานนี้แบ่งออกเป็น 2 ส่วน คือการสร้างภาพ Generator และ การทำนายภาพ Discriminator เป็นหลักเข้ามาใช้ในการทดลอง ซึ่งการทดลองนี้ได้ใช้ชุดข้อมูลของ GoPro-Large เช่นเดียวกัน

ตารางที่ 3.9 ผลลัพธ์ที่ได้จากวิธีที่ 8 ด้วย GANs

ประเภทของภาพเบลอ	การวัดประสิทธิภาพของภาพ		
	SNR	PSNR	SSIM
Blur	11.64	11.70	0.580
Blur_gamma	11.27	10.46	0.479

ในการทดลองด้วยวิธีของ GANs กับชุดข้อมูลของ GoPro-Large พบว่าผลลัพธ์ที่ไม่ดีเมื่อเปรียบเทียบกับการทำ Autoencoder

บทที่ 4
ผลการวิจัย

จากการออกแบบวิธีการวิจัยในรูปแบบต่างๆ ในบทที่ 3 จึงส่งผลให้นำผลการวิจัยมารายงานในบทที่ 4 และในบทนี้จะได้อธิบายถึงผลการวิจัยในวิธีการต่างๆ รวมถึงความเป็นไปได้ที่จะส่งผลให้งานวิจัยนี้สำเร็จ โดยมีรายละเอียดดังต่อไปนี้

4.1 ผลการวิจัยด้วยอัลกอริทึมของ Autoencoder ในรูปแบบต่างๆ กับภาพประเภท Blur

ในการออกแบบวิธีการวิจัยด้วยวิธีการต่างๆ โดยใช้ชุดข้อมูล (Dataset) ของ GoPro-Large ซึ่งมีรูปแบบของการเบลอ 2 ประเภท ประกอบด้วย Blur และ Blur_gamma ดังนั้นจึงสามารถสรุปผลการวิจัยจากวิธีการทั้งหมด 8 วิธี แบ่งออกเป็นการเข้ารหัสแบบอัตโนมัติ (Autoencoder) จำนวน 7 วิธี และ การทำ Generative Adversarial Networks: GANs จำนวน 1 วิธี ซึ่งมีรายละเอียดดังต่อไปนี้

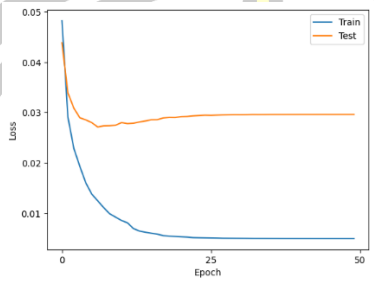
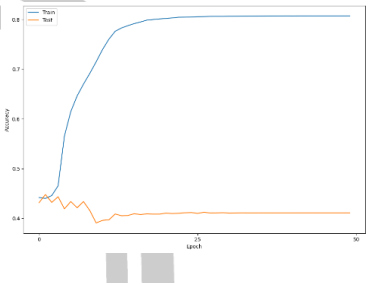
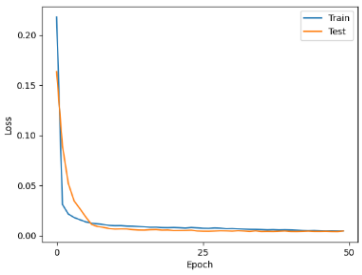
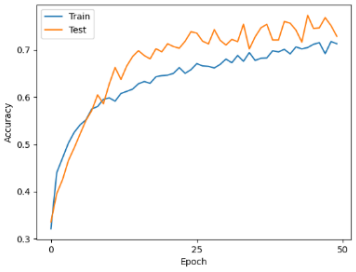
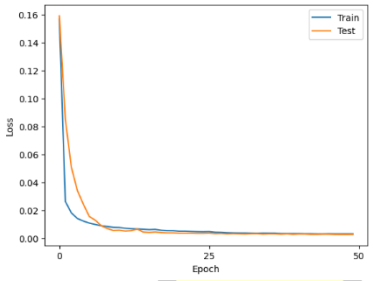
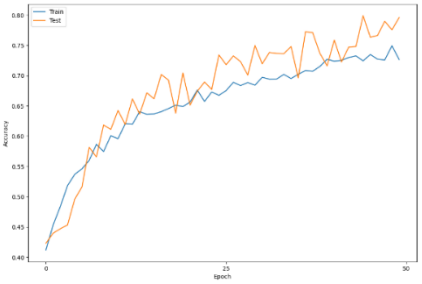
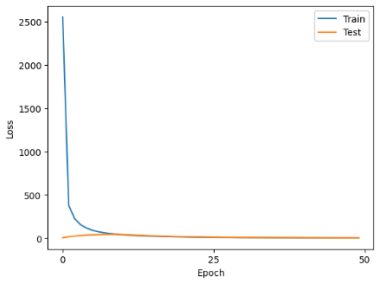
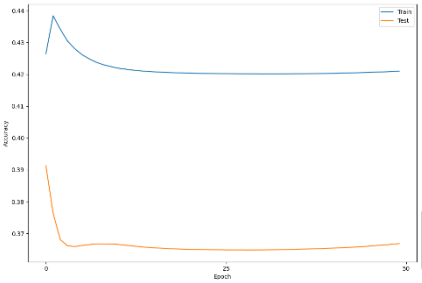
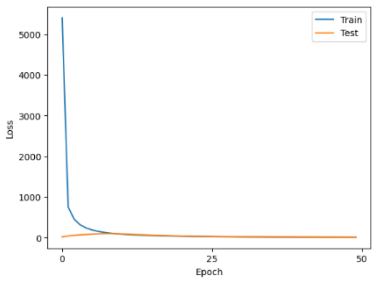
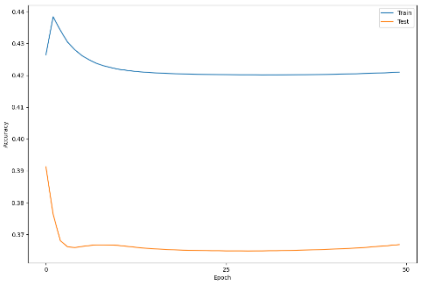
ตารางที่ 4.1 แสดงผลการทดลองด้วยการเข้ารหัสแบบอัตโนมัติ (Autoencoder) จำนวน 7 วิธี กับภาพประเภท Blur

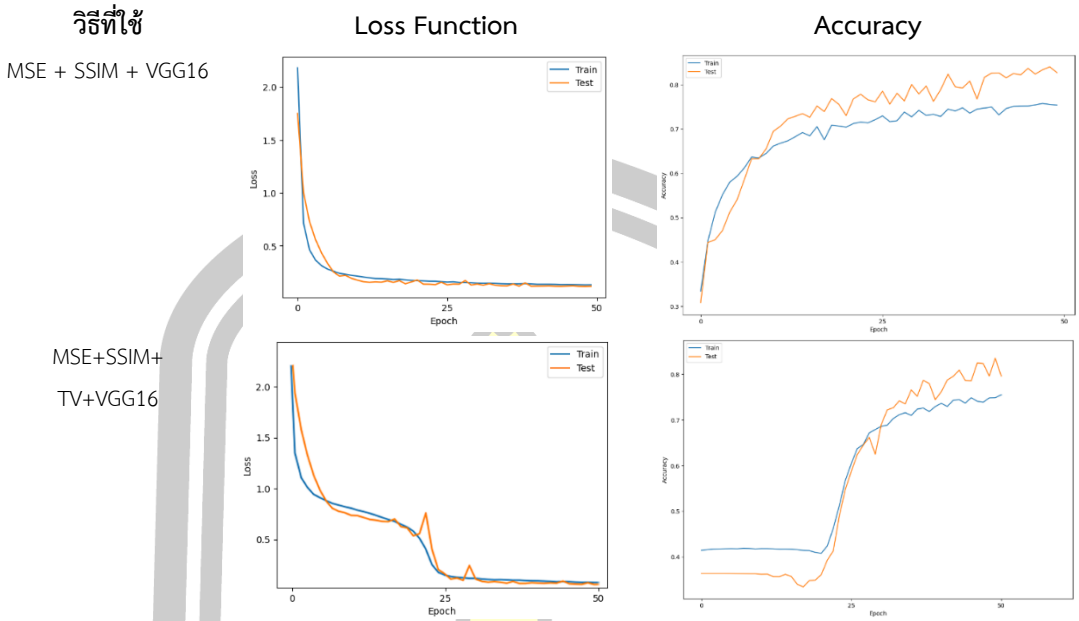
วิธีที่ใช้ทดลอง Blur	SNR	PSNR	SSIM	Time
วิธีที่ 1 Autoencoder	32.53	22.00	0.771	2.10
วิธีที่ 2 Perceptual (VGG16)	31.82	21.65	0.753	2.50
วิธีที่ 3 MSE + SSIM	31.87	21.71	0.742	2.28
วิธีที่ 4 MSE + Total Variance (TV)	14.05	12.76	0.575	3.30
วิธีที่ 5 MSE + SSIM + Total Variance (TV)	10.26	11.36	0.600	2.45
วิธีที่ 6 MSE + SSIM + VGG16	**32.64**	**22.06**	**0.777**	**3.16**
วิธีที่ 7 MSE + SSIM + Total Variance (TV) + VGG16	15.14	15.79	0.633	4.52

จากตารางที่ 4.1 แสดงให้เห็นถึงผลลัพธ์ที่มีค่าเฉลี่ยในแต่ละด้านของการทดลองแต่ละวิธีโดยใช้ภาพประเภท Blur

จากตารางสรุปผลการวิจัยที่เกิดจากการออกแบบ โดยใช้วิธีการเข้ารหัสอัตโนมัติ (Autoencoder) ในแต่ละแบบกับภาพประเภท Blur พบว่าการออกแบบในวิธีที่ 6 ซึ่งได้ใช้การปรับปรุงค่า Loss function โดยนำสมการของค่า MSE รวมกับ SSIM และ VGG16 มาทำงานร่วมกัน ซึ่งส่งผลให้ภาพผลลัพธ์มีความคมชัดมากกว่าวิธีอื่นๆ เมื่อนำภาพผลลัพธ์ที่ได้มาเปรียบเทียบกับวิธีการเข้ารหัสอัตโนมัติแบบปกติหรือเกณฑ์พื้นฐาน (Base line) ในการวัดประสิทธิภาพของภาพ

จากการออกแบบการวิจัยด้วยการเข้ารหัสแบบอัตโนมัติ (Autoencoder) ได้นำการเปรียบเทียบค่า Loss และ Accuracy ที่เกิดขึ้นในการเรียนรู้เพื่อสร้างโมเดล ประเภท Blur โดยมีรายละเอียดดังต่อไปนี้

วิธีที่ใช้	**Loss Function**	**Accuracy**
Autoencoder		
Perceptual (VGG16)		
MSE + SSIM		
MSE + TV		
MSE + SSIM + TV		



ภาพที่ 4.1 กราฟแสดงค่า Loss function และค่า Accuracy ของการทดลองแต่ละแบบของภาพประเภท Blur

4.2 ผลการวิจัยด้วยอัลกอริทึมของการเข้ารหัสแบบอัตโนมัติ Autoencoder ในรูปแบบต่างๆ กับภาพประเภท Blur_gamma

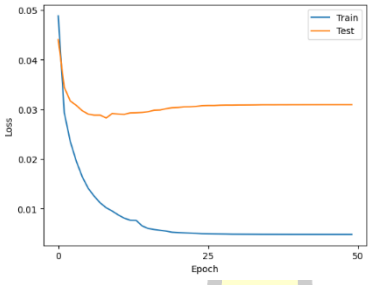
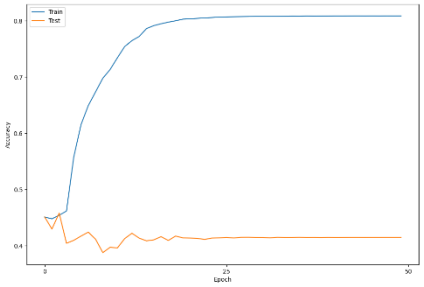
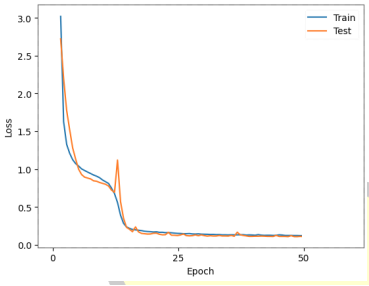
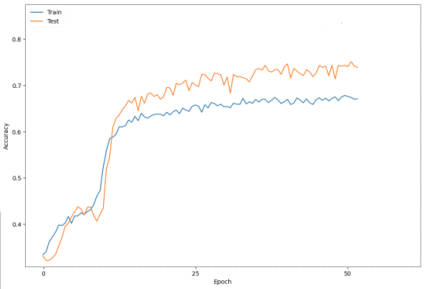
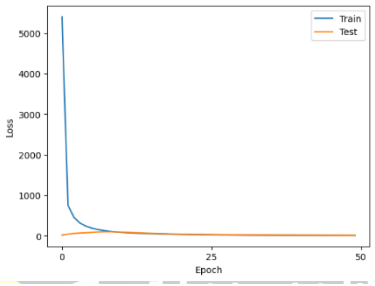
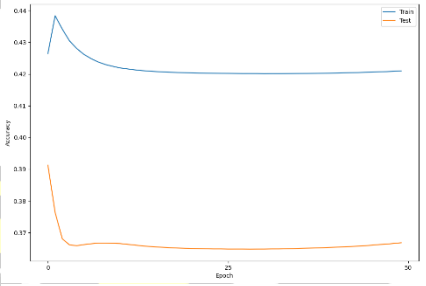
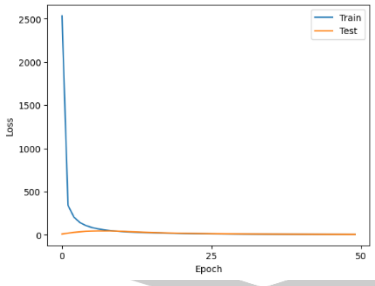
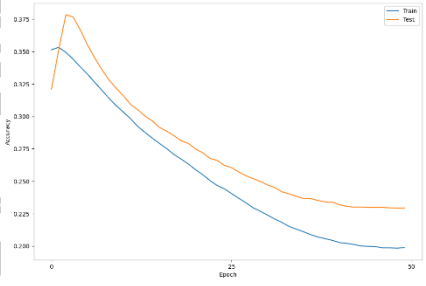
จากการออกแบบโดยใช้อัลกอริทึมของการเข้ารหัสแบบอัตโนมัติ (Autoencoder) โดยนำมาใช้กับชุดข้อมูลของ GoPro-Large และใช้ภาพประเภท Blur_gamma ในการทดลอง จึงสามารถสรุปผลได้ดังต่อไปนี้

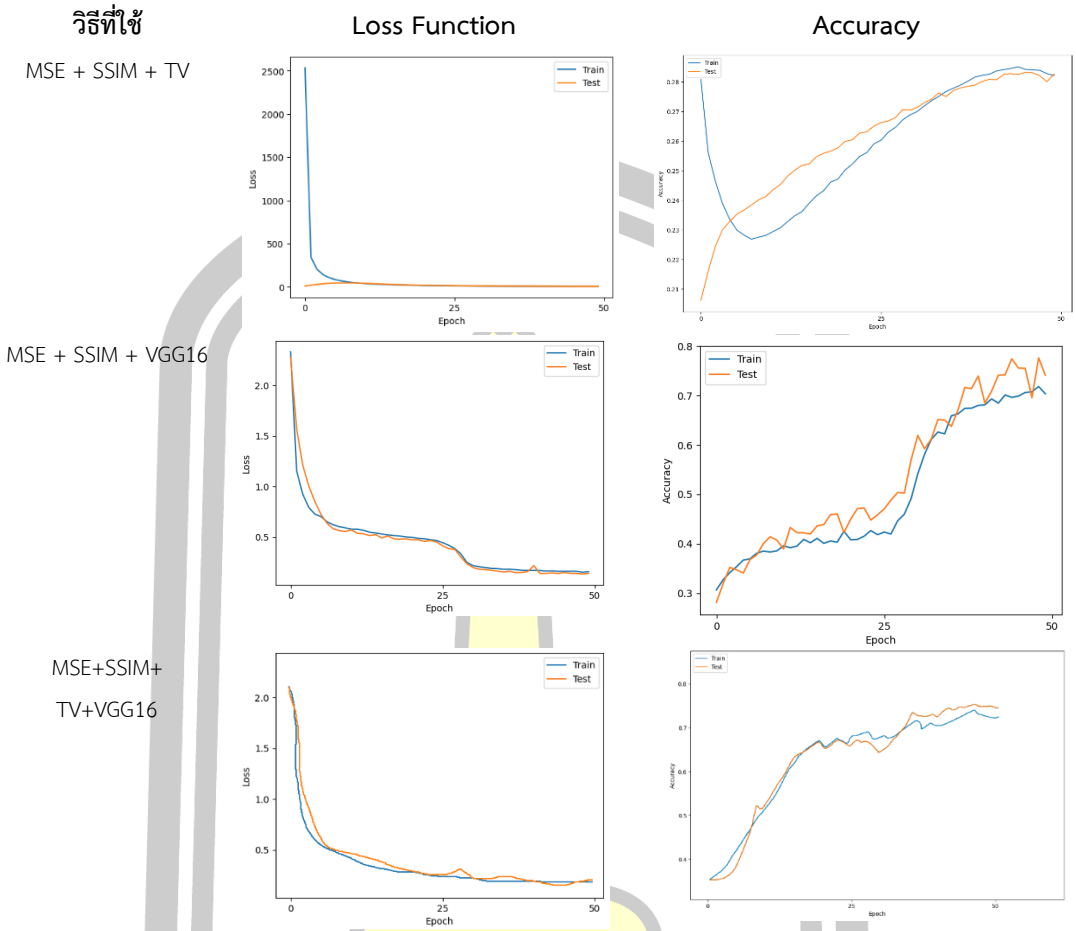
ตารางที่ 4.2 แสดงผลการทดลองด้วยการเข้ารหัสแบบอัตโนมัติ (Autoencoder) จำนวน 7 วิธี กับภาพประเภท Blur_gamma

วิธีที่ใช้ทดลอง Blur_gamma	SNR	PSNR	SSIM	Time
วิธีที่ 1 Autoencoder	32.91	22.20	0.770	2.26
วิธีที่ 2 Perceptual (VGG16)	32.55	22.02	0.757	2.59
วิธีที่ 3 MSE + SSIM	31.67	21.57	0.763	2.33
วิธีที่ 4 MSE + Total Variance (TV)	13.99	12.57	0.561	3.45
วิธีที่ 5 MSE + SSIM + Total Variance (TV)	11.07	12.76	0.609	2.56
วิธีที่ 6 MSE + SSIM + VGG16	31.84	21.66	0.766	3.30
วิธีที่ 7 MSE + SSIM + Total Variance (TV) + VGG16	15.47	16.05	0.689	5.12

จากตารางสรุปผลการวิจัยที่เกิดจากการออกแบบ โดยใช้วิธีการเข้ารหัสอัตโนมัติ (Autoencoder) ในแต่ละแบบกับภาพประเภท Blur_gamma พบว่าในแต่ละวิธีที่ออกแบบไม่สามารถทำให้ภาพเบลอประเภทนี้มีความคมชัดเพิ่มขึ้นได้ เนื่องจากภาพประเภทนี้มีรูปแบบการเบลอที่แตกต่างจากภาพประเภท Blur จึงทำให้ผลลัพธ์ที่ได้ไม่ดีขึ้น

ผลการออกแบบในการวิจัยนี้มีการเปรียบเทียบค่า Loss และ Accuracy ที่เกิดขึ้นในการเรียนรู้เพื่อสร้างโมเดล ประเภท Blur_gamma โดยมีรายละเอียดดังต่อไปนี้

วิธีที่ใช้	Loss Function	Accuracy
Autoencoder		
Perceptual (VGG16)		
MSE + SSIM		
MSE + TV		



ภาพที่ 4.2 กราฟแสดงค่า Loss function และค่า Accuracy ของการทดลองแต่ละแบบของภาพประเภท Blur_gamma

4.3 ผลจากการออกแบบด้วยอัลกอริทึมของ Generative Adversarial Networks: GANs

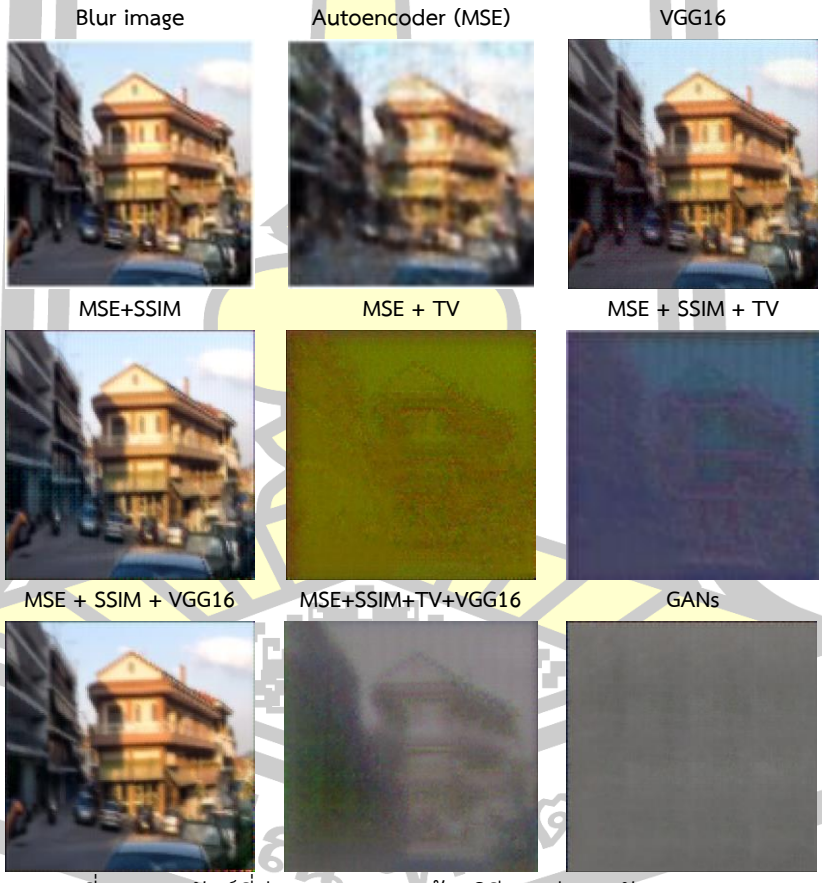
ในส่วนของการออกแบบการวิจัยด้วย Generative Adversarial Networks: GANs เป็นอีก 1 วิธีที่ได้ออกแบบมาเพื่อใช้ในการปรับปรุงภาพให้มีประสิทธิภาพดีขึ้น โดยใช้ชุดข้อมูลเดียวกับกับการทดลองทั้ง 7 วิธี ซึ่งผลลัพธ์ที่ได้ไม่ดีขึ้นเมื่อเปรียบเทียบกับภาพเบลอปกติในแต่ละประเภท ดังนั้นจึงสามารถสรุปได้ว่าวิธีการปรับปรุงภาพด้วย Generative Adversarial Networks: GANs ไม่สามารถปรับปรุงคุณภาพของภาพในลักษณะแบบนี้ได้ โดยผลลัพธ์ที่ได้เมื่อวัดประสิทธิภาพของภาพโดยมีค่า SNR, PSNR และ SSIM เท่ากับ 11.64, 11.70, 0.580 ตามลำดับ และใช้เวลาในการประมวลผลประมาณ 4-5 ชั่วโมง อีกทั้งยังทำให้ภาพผลลัพธ์ไม่มีความคมชัด เมื่อเทียบกับวิธีการของ Autoencoder

จากการออกแบบการวิจัยด้วยอัลกอริทึมทั้ง 2 แบบ ได้แก่การทำ Autoencoder และ Generative Adversarial Networks: GANs ซึ่งใช้การออกแบบทั้งหมด 8 วิธี โดยใช้ชุดข้อมูล

GoPro-Large ซึ่งแบ่งภาพเป็น 2 ประเภทได้แก่ แบบ Blur และ Blur_gamma พบว่า วิธีการทดลองโดยใช้ Autoencoder ที่มีการปรับค่าพารามิเตอร์ด้วยการทำ Customized loss function โดยนำสมการของค่า Mean Squared Error (MSE) ทำงานร่วมกับ SSIM และ VGG16 โดยใช้การ Train จำนวน 50 รอบ สามารถลดความเบลอของภาพได้สูงกว่า Baseline ซึ่งใช้ได้ดีกับภาพประเภท Blur โดยมีค่า SNR, PSNR และ SSIM เท่ากับ 32.64, 22.06, 0.777 ตามลำดับ และใช้เวลาประมาณ 3.16 นาที ซึ่งถือว่ายังใช้เวลานานมาก

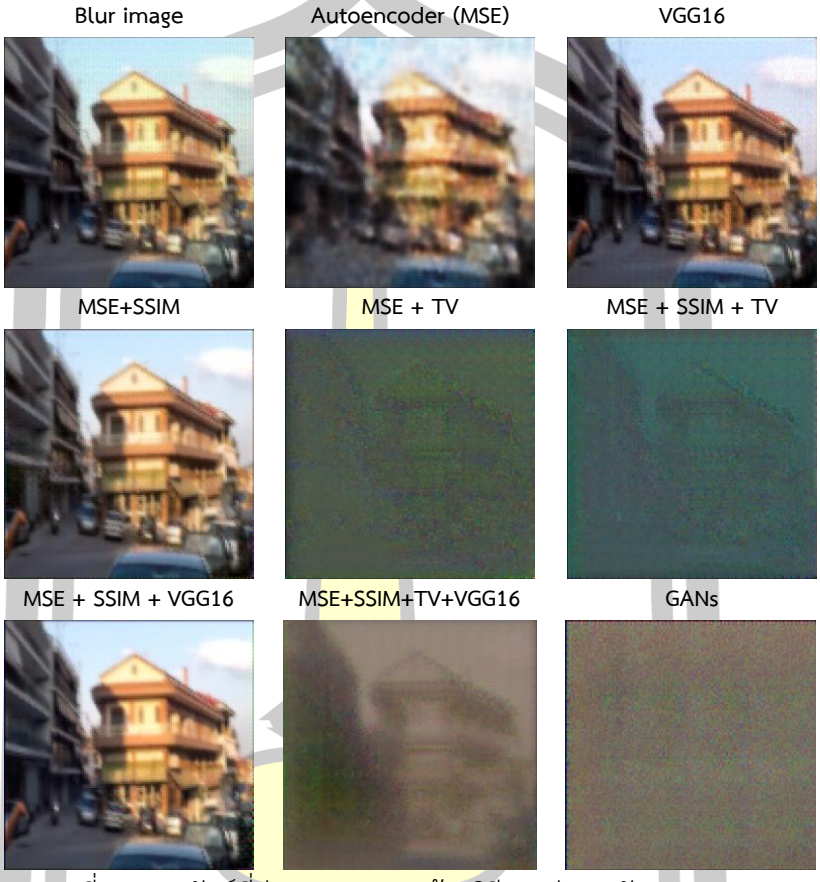
4.4 ภาพผลลัพธ์ที่เกิดจากการใช้อัลกอริทึมทั้ง 2 แบบ โดยใช้ภาพทั้ง 2 ประเภท

ภาพผลลัพธ์ที่เกิดจากการออกแบบการวิจัยด้วยอัลกอริทึม 2 วิธี จำนวนทั้งหมด 8 วิธี จากภาพประเภท Blur โดยใช้การ Train จำนวน 50 รอบ



ภาพที่ 4.3 ผลลัพธ์ที่ผ่านการทดลองด้วยวิธีการต่างๆ กับภาพประเภท Blur

ภาพผลลัพธ์ที่เกิดจากการออกแบบการวิจัยด้วยอัลกอริทึม 2 วิธี จำนวนทั้งหมด 8 วิธี จากภาพประเภท Blur_gamma โดยใช้การ Train จำนวน 50 รอบ



ภาพที่ 4.4 ผลลัพธ์ที่ผ่านการทดลองด้วยวิธีการต่างๆ กับภาพประเภท Blur_gamma

บทที่ 5
สรุปผล อภิปรายผล และข้อเสนอแนะ

การวิจัยเรื่อง การลดความเบลอของภาพ เป็นการวิจัยเชิงทดลอง (Experiment Research) โดยสร้างขึ้นมาเพื่อใช้สำหรับการปรับปรุงภาพให้มีความเบลอลดลงและมีความชัดเจนมากขึ้น ซึ่งเป็นการประยุกต์การทำงานของการเรียนรู้เชิงลึก (Deep learning) การประมวลผลภาพ (Image Processing) และ Perceptual VGG16 รวมกับการออกแบบอัลกอริทึมที่มีโครงสร้างการทำงานเพื่อให้สามารถลดความเบลอของภาพ และเมื่อได้ทดลองเป็นที่เรียบร้อยแล้ว ในบทนี้จึงขอสรุปผล อภิปรายผล รวมถึงปัญหาและอุปสรรคในการดำเนินงาน ตลอดจนข้อสนอแนะต่างๆ ดังต่อไปนี้

5.1 สรุปผล

ในการวิจัยนี้จะขอสรุปผลการทดลองออกเป็น 3 ด้าน โดยมีรายละเอียดดังต่อไปนี้

1. ด้านการใช้ชุดข้อมูล

ในการวิจัยนี้ได้เลือกใช้ชุดข้อมูลทั้งหมด 3 ชุด ประกอบด้วย GoPro-large, Real-world และ Kohler แต่เนื่องจากชุดข้อมูลของ Real-world และ Kohler มีตำแหน่งของภาพระหว่างภาพเบลอ (Blur image) กับภาพชัด (Sharp image) ไม่ตรงกันจึงไม่สามารถนำมาประมวลผลภาพได้ ซึ่งจากการทดลองในหลายวิธีเมื่อวัดประสิทธิภาพของภาพจากหลายๆ วิธีแล้ว พบว่ามีค่าเฉลี่ยสูงกว่าค่าเบลอมาตรฐานมากเกินไป เมื่อตรวจข้อเท็จจริง จึงทำให้ทราบว่าชุดข้อมูลภาพที่นำมามีตำแหน่งไม่ตรงกัน ดังนั้นจึงทำให้ไม่สามารถนำชุดข้อมูลทั้งสองชุดนี้มาใช้ในการวิจัยได้

ดังนั้นในการทดลองตลอดการวิจัยเรื่องการลดความเบลอของภาพ (Image deblurring) จึงเลือกใช้ชุดข้อมูลของ GoPro-Large เพียงชุดเดียว ซึ่งเป็นชุดที่มีข้อมูลภาพจำนวนมาก โดยมีจำนวน 9,642 ภาพ แบ่งออกเป็น 2 ประเภท คือ Blur และ Blur_gamma อีกทั้งยังแบ่งออกเป็นชุดข้อมูลที่ใช้สำหรับการ Train 6,309 ภาพ และชุดข้อมูลที่ใช้สำหรับการ Test จำนวน 3,333 ภาพ ดังนั้นจึงสามารถนำมาใช้ในการวิจัยได้

2. การออกแบบการทดลอง

สำหรับการออกแบบการทดลองในการวิจัยนี้ได้ใช้อัลกอริทึมหลักๆ อยู่ 2 วิธีคือ Autoencoder และ Generative Adversarial Networks: GANs โดยแบ่งการทดลองออกเป็น 8 วิธี ซึ่งใช้อัลกอริทึมของ Autoencoder จำนวน 7 วิธี ประกอบไปด้วย

- วิธีที่ 1 Autoencoder
- วิธีที่ 2 Perceptual (VGG16)
- วิธีที่ 3 MSE + SSIM
- วิธีที่ 4 MSE + Total Variance (TV)

- วิธีที่ 5 MSE + SSIM + Total Variance (TV)
- วิธีที่ 6 MSE + SSIM + VGG16
- วิธีที่ 7 MSE + SSIM + Total Variance (TV) + VGG16

จากการออกแบบการทดลองด้วยอัลกอริทึมของ การเข้ารหัสอัตโนมัติ (Autoencoder) สามารถลดความเบลอของภาพได้ดีและมีประสิทธิภาพสูงขึ้น ส่วนวิธีการในการปรับปรุงภาพโดยใช้อัลกอริทึมของ Generative Adversarial Networks: GANs เป็นวิธีที่ 8 นั้น พบว่าอัลกอริทึมนี้ยังไม่สามารถปรับปรุงภาพให้มีความคมชัดขึ้นได้

3. ผลที่ได้จากการวิจัย

จากการออกแบบวิธีการวิจัยโดยใช้อัลกอริทึม 2 แบบ ซึ่งมีทั้งหมดทั้ง 8 วิธี พบว่า วิธีที่ 6 โดยการใช้ฟังก์ชันของ MSE แบบ Custom Loss ที่ใช้ MSE + SSIM + VGG16 ส่งผลให้การปรับปรุงภาพมีประสิทธิภาพเพิ่มขึ้น สำหรับการทดลองนี้ใช้การ Train ข้อมูลจำนวน 50 รอบ และสามารถลดความเบลอของภาพได้สูงกว่า Baseline กับภาพประเภท Blur โดยมีค่า SNR, PSNR, SSIM เท่ากับ 32.74, 22.11, 0.780 ตามลำดับ ซึ่งจะทำให้การปรับปรุงภาพประเภท Blur มีความเบลอลดลงตามวัตถุประสงค์อย่างมีนัยสำคัญ ดังนั้นอัลกอริทึมของ Autoencoder สามารถปรับปรุงภาพเบลอได้ดีกับชุดข้อมูลของ GoPro-Large กับภาพประเภท Blur ในส่วนของภาพประเภท Blur_gamma ยังไม่สามารถปรับให้ชัดเจนขึ้นได้ซึ่งอาจจะต้องใช้วิธีการอื่นๆเพื่อทำให้ภาพชัดเจนขึ้นได้

5.2 อภิปรายผล

จากการทดลองทั้ง 8 วิธี พบว่าการทดลองในวิธี Image Deblurring Autoencoder Network โดยใช้อัลกอริทึมนี้ โดยปรับมาใช้วิธีการทดลองที่ผ่านมาส่งผลให้ผลลัพธ์บางวิธีดีขึ้น แต่ยังมีค่า SNR, PSNR และ SSIM สูงกว่า Baseline เพียงเล็กน้อย ซึ่งอาจเกิดจากการทดลองดังกล่าวไม่ได้ใช้จำนวนการ Train ชุดข้อมูลในจำนวนมากเท่าที่ควร เนื่องจากมีข้อจำกัดของเครื่องที่ใช้ในการประมวลผล

5.3 ข้อเสนอแนะ

จากการทดลองในการวิจัยนี้ ผู้วิจัยได้พบปัญหาที่เกิดขึ้นสำหรับการทดลองนี้อยู่หลายอย่าง โดยสรุปได้ดังต่อไปนี้

1. ชุดข้อมูลที่ใช้ จากการที่ผู้วิจัยได้เลือกชุดข้อมูลมา 3 ชุด ได้แก่ GoPro-Large, Real-world และ Kohler ในการทดลองพบว่าชุดข้อมูลของ Real-world และ Kohler ไม่สามารถนำมาใช้งานได้เนื่องจากมีตำแหน่งของภาพที่ไม่ตรงกัน ดังนั้นจึงทำให้ภาพและการวัดประสิทธิภาพของภาพมี

ความไม่ถูกต้องจึงไม่สามารถนำมาใช้งานได้ ดังนั้นในการวิจัยนี้จึงใช้ชุดข้อมูลของ GoPro-Large ได้เพียงชุดเดียวเท่านั้น

2. การประมวลผล ในการใช้เครื่องคอมพิวเตอร์ในการประมวลผลนั้น สำหรับการทำ Image Processing จำเป็นอย่างยิ่งที่จะต้องใช้เครื่องความพิวเตอร์ที่การประมวลผลสูงๆ แต่ในการวิจัยนี้ผู้วิจัยได้ใช้เครื่องที่มีประสิทธิภาพของเครื่องไม่สูงพอ โดยผู้วิจัยได้ใช้เครื่องคอมพิวเตอร์ Notebook ระบบปฏิบัติการแบบ Mac OS ซึ่งมีรายละเอียดดังต่อไปนี้ CPU : 2Ghz Quad-Core Intel Core i5, Graphics Intel Iris Plus Graphic 1,536 MB, Memory : 16 GB ซึ่งในการประมวลผลยังถือว่าช้า อีกทั้งชุดข้อมูลของ GoPro-Large มีขนาดใหญ่และมีจำนวนภาพที่เยอะ ดังนั้นหากเป็นไปได้จึงขอให้ใช้เครื่อง PC ที่สามารถเพิ่ม หรือปรับแต่งให้เครื่องมีประสิทธิภาพสูงขึ้น จึงจะทำให้การประมวลผลเร็วขึ้นและการทำงานได้ผลดีขึ้น

3. โปรแกรมที่ใช้ในการประมวลผล ทางผู้วิจัยได้เลือก Google Colaboratory ซึ่งเป็นโปรแกรมที่สามารถใช้งานได้ฟรีและสามารถใช้งานได้ง่าย และได้ผลลัพธ์ที่ดี แต่เนื่องจากการประมวลผลของโปรแกรมมีขีดจำกัดในการใช้ทรัพยากรของโปรแกรม จึงจำเป็นต้องซื้อเพิ่ม เพื่อที่จะทำให้การประมวลผลเร็วขึ้น และปัญหาที่ตามมาคือโปรแกรม Google Colaboratory นี้จะหลุด Session บ่อย เนื่องจากการประมวลจะต้องใช้เวลานานและใช้ทรัพยากรของเครื่องเยอะจึงทำให้ต้องมารันโปรแกรมใหม่ตลอดเวลา ดังนั้นผู้วิจัยจึงของเสนอให้ใช้โปรแกรมอื่นเช่น Visual Studio Code (VScode) ในการเขียนโปรแกรมแทน

4. การ Train ข้อมูล สำหรับการทดลองทั้งหมดในการวิจัยนี้ได้ใช้การ Train ข้อมูลเพียง 50 รอบ เนื่องจากมีข้อจำกัดทางเครื่องที่ใช้ในการประมวสลผล ดังนั้นผลลัพธ์ที่ได้อาจจะไม่สูงมากและผู้วิจัยขอแนะนำให้ผู้ที่สนใจในเรื่องนี้ทำการ Train ข้อมูลในปริมาณที่เยอะกว่านี้ เช่น 100, 500 หรือ 1,000 รอบ เพื่อที่จะได้ข้อมูลที่มีประสิทธิภาพมากขึ้น แต่ข้อเสียคือจะต้องใช้เวลาในการประมวลผลค่อนข้างนาน

5. สำหรับวิธีการปรับปรุงภาพด้วยวิธีการของการเข้ารหัสอัตโนมัติ (Autoencoder) อาจจะนำไปใช้กับชุดข้อมูลที่เป็นภาพเกี่ยวกับทางการแพทย์ เพื่อปรับภาพที่เกิดจากการ X- ray หรือภาพจากการทำ CT scan, MRI ให้มีความชัดเจนและมีประสิทธิภาพมากขึ้น และจะทำให้เกิดประโยชน์แก่การวินิจฉัยโรคได้ดีขึ้น

บรรณานุกรม

[1] Kupyn O, Martyniuk T, Wu J, Wang Z. Deblurgan-v2: Deblurring (orders-of-magnitude) faster and better. Proceedings of the IEEE/CVF international conference on computer vision; 8878-8887.

[2] Levin A. Blind motion deblurring using image statistics. Advances in Neural Information Processing Systems 2006; 19

[3] Zhang K, Ren W, Luo W, Lai W-S, Stenger B, Yang M-H, et al. Deep image deblurring: A survey. International Journal of Computer Vision 2022; 130[9]: 2103-2130.

[4] Satish P, Srikantaswamy M, Ramaswamy NK. A Comprehensive Review of Blind Deconvolution Techniques for Image Deblurring. Traitement du Signal 2020; 37[3]:

[5] Solovyeva E, Abdullah A. Dual Autoencoder Network with Separable Convolutional Layers for Denoising and Deblurring Images. Journal of Imaging 2022; 8[9]: 250.

[6] Gonzalez RC. Morphological image processing. Digital image processing 2002;

[7] Wikipedia. Peak Signal-to-Noise Ratio (PSNR) [18 oct 2020]; https://hmong.in.th/wiki/Peak_signal-to-noise_ratio

[8] McAndrew A. An introduction to digital image processing with matlab notes for scm2511 image processing. School of Computer Science and Mathematics, Victoria University of Technology 2004; 264[1]: 1-264.

[9] Sites T. AI vs Machine Learning vs Deep Learning. 2015; [14 November 2020]; https://www.edureka.co/blog/ai-vs-machine-learning-vs-deep-learning/

[10] Hinton GE, Osindero S, Teh Y-W. A fast learning algorithm for deep belief nets. Neural computation 2006; 18[7]: 1527-1554.

[11] LeCun Y, Bengio Y, Hinton G. Deep learning. nature 2015; 521[7553]: 436-444.

[12] Le Cun Y, Jackel LD, Boser B, Denker JS, Graf HP, Guyon I, et al. Handwritten digit recognition: Applications of neural network chips and automatic learning. IEEE Communications Magazine 1989; 27[11]: 41-46.

[13] Al-Saffar AAM, Tao H, Talab MA. Review of deep convolution neural network in image classification. 2017 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET); IEEE; 26-31.

[14] El-Sawy A, Loey M, El-Bakry H. Arabic handwritten characters recognition using convolutional neural network. WSEAS Transactions on Computer Research 2017; 511-19.

[15] Alizadehashraf B, Roohi S. Persian handwritten character recognition using convolutional neural network. 2017 10th Iranian Conference on Machine Vision and Image Processing (MVIP); IEEE 2018; 247-251.

[16] Inunganbi S, Choudhary P, Manglem K. Manipuri Handwritten Character Recognition by Convolutional Neural Network. International Conference on Computer Vision and Image Processing; Springer; 307-318.

[17] Kagaya H, Aizawa K, Ogawa M. Food detection and recognition using convolutional neural network. Proceedings of the 22nd ACM international conference on Multimedia; 1085-1088.

[18] Hijazi S, Kumar R, Rowen C. Using convolutional neural networks for image recognition. Cadence Design Systems Inc: San Jose, CA, USA 2015; 1-12.

[19] Liu YH. Feature extraction and image recognition with convolutional neural networks. Journal of Physics: Conference Series; 062032.

[20] Lawrence S, Giles CL, Tsoi AC, Back AD. Face recognition: A convolutional neural-network approach. IEEE transactions on neural networks 1997; 8[1]: 98-113.

[21] Coşkun M, Uçar A, Yildirim Ö, Demir Y. Face recognition based on convolutional neural network. 2017 International Conference on Modern Electrical and Energy Systems (MEES); IEEE; 376-379.

[22] Li Y, Wang G, Nie L, Wang Q, Tan W. Distance metric optimization driven convolutional neural network for age invariant face recognition. Pattern Recognition 2018; 7551-62.

[23] Phung VH, Rhee EJ. A High-Accuracy Model Average Ensemble of Convolutional Neural Networks for Classification of Cloud Image Patches on Small Datasets. Applied Sciences 2019; 9[21]: 4500.

[24] Dertat A. Applied Deep Learning - Part 4: Convolutional Neural Networks. 2017; [8 September 2021]; https://towardsdatascience.com/applied-deep-learning-part-4-convolutional-neural-networks-584bc134c1e2

[25] Surapong K. Recurrent Neural Network (RNN) คืออะไร Gated Recurrent Unit (GRU) คืออะไร สอนสร้าง RNN ถึง GRU ด้วยภาษา Python – NLP ep.9. 2019; [17 September 2020]; https://commons.wikimedia.org/wiki/File:Recurrent_neural_network_unfold.svg

[26] Fergus R, Singh B, Hertzmann A, Roweis ST, Freeman WT. Removing camera shake from a single photograph. *ACM SIGGRAPH 2006 Papers* 2006:787-794.

[27] Yuan L, Sun J, Quan L, Shum H-Y. Image deblurring with blurred/noisy image pairs. *ACM SIGGRAPH 2007 papers* 2007:1-es.

[28] Pan J, Sun D, Pfister H, Yang M-H. Deblurring images via dark channel prior. IEEE transactions on pattern analysis and machine intelligence 2017; 40[10]: 2315-2328.

[29] Cai J-F, Chan RH, Nikolova M. Two-phase approach for deblurring images corrupted by impulse plus Gaussian noise. Inverse Problems & Imaging 2008; 2[2]: 187.

[30] Shan Q, Jia J, Agarwala A. High-quality motion deblurring from a single image. Acm transactions on graphics (tog) 2008; 27[3]: 1-10.

[31] Joshi N, Zitnick CL, Szeliski R, Kriegman DJ. Image deblurring and denoising using color priors. 2009 IEEE Conference on Computer Vision and Pattern Recognition; IEEE; 1550-1557.

[32] Chen F, Ma J. An empirical identification method of Gaussian blur parameter for image deblurring. IEEE Transactions on signal processing 2009; 57[7]: 2467-2478.

[33] Aguado MNA. Feature Extraction & Image Processing. 2002; [8 September 2021]; https://www.southampton.ac.uk/~msn/book/new_demo/median

[34] ITIGIC. Signal-to-Noise Ratio. 2021; [20 June 2020]; https://itigic.com/th/signal-to-noise-ratio-or-snr-in-audio-what-is-it-for/

[35] Wikipedia. Structural Similarity Index Measure (SSIM). [18 oct 2020]; https://hmong.in.th/wiki/Structural_similarity

[36] Gioux S, Mazhar A, Cuccia DJ. Spatial frequency domain imaging in 2019: principles, applications, and perspectives. Journal of biomedical optics 2019; 24[7]: 071613.

[37] Cho S, Lee S. Fast motion deblurring. *ACM SIGGRAPH Asia* 2009 *papers* 2009:1-8.

[38] Andrews HC, Hunt BR. Digital image restoration. 1977;

[39] Lee S, Cho S. Recent advances in image deblurring. SIGGRAPH Asia 2013 Courses 2013; 1-108.

[40] Gupta A, Joshi N, Zitnick CL, Cohen M, Curless B. Single image deblurring using motion density functions. European conference on computer vision; Springer; 171-184.

[41] Hirsch M, Schuler CJ, Harmeling S, Schölkopf B. Fast removal of non-uniform camera shake. 2011 International Conference on Computer Vision; IEEE; 463-470.

[42] Whyte O, Sivic J, Zisserman A, Ponce J. Non-uniform deblurring for shaken images. International journal of computer vision 2012; 98[2]: 168-186.

[43] Gong D, Tan M, Zhang Y, Van den Hengel A, Shi Q. Blind image deconvolution by automatic gradient activation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 1827-1836.

[44] Sun L, Cho S, Wang J, Hays J. Edge-based blur kernel estimation using patch priors. IEEE International Conference on Computational Photography (ICCP); IEEE 2013; 1-8.

[45] Yue T, Cho S, Wang J, Dai Q. Hybrid image deblurring by fusing edge and power spectrum information. European conference on computer vision; Springer; 79-93.

[46] Ren W, Cao X, Pan J, Guo X, Zuo W, Yang M-H. Image deblurring via enhanced low-rank prior. IEEE Transactions on Image Processing 2016; 25[7]: 3426-3437.

[47] Jia J. Single image motion deblurring using transparency. 2007 IEEE Conference on Computer Vision and Pattern Recognition; IEEE; 1-8.

[48] Chen L, Fang F, Wang T, Zhang G. Blind image deblurring with local maximum gradient prior. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 1742-1750.

[49] Hu D, Tan J, Zhang L, Ge X, Liu J. Image deblurring via enhanced local maximum intensity prior. Signal Processing: Image Communication 2021; 96116311.

[50] Levin A, Weiss Y, Durand F, Freeman WT. Understanding and evaluating blind deconvolution algorithms. 2009 IEEE Conference on Computer Vision and Pattern Recognition; IEEE; 1964-1971.

[51] Levin A, Fergus R, Durand F, Freeman WT. Image and depth from a conventional camera with a coded aperture. ACM transactions on graphics (TOG) 2007; 26[3]: 70-es.

[52] Pan J, Sun D, Pfister H, Yang M-H. Blind image deblurring using dark channel prior. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 1628-1636.

[53] Liu J, Tan J, He L, Ge X, Hu D. Blind Image Deblurring via Local Maximum Difference Prior. IEEE Access 2020; 8219295-219307.

[54] Nimisha TM, Kumar Singh A, Rajagopalan AN. Blur-invariant deep learning for blind-deblurring. Proceedings of the IEEE International Conference on Computer Vision; 4752-4760.

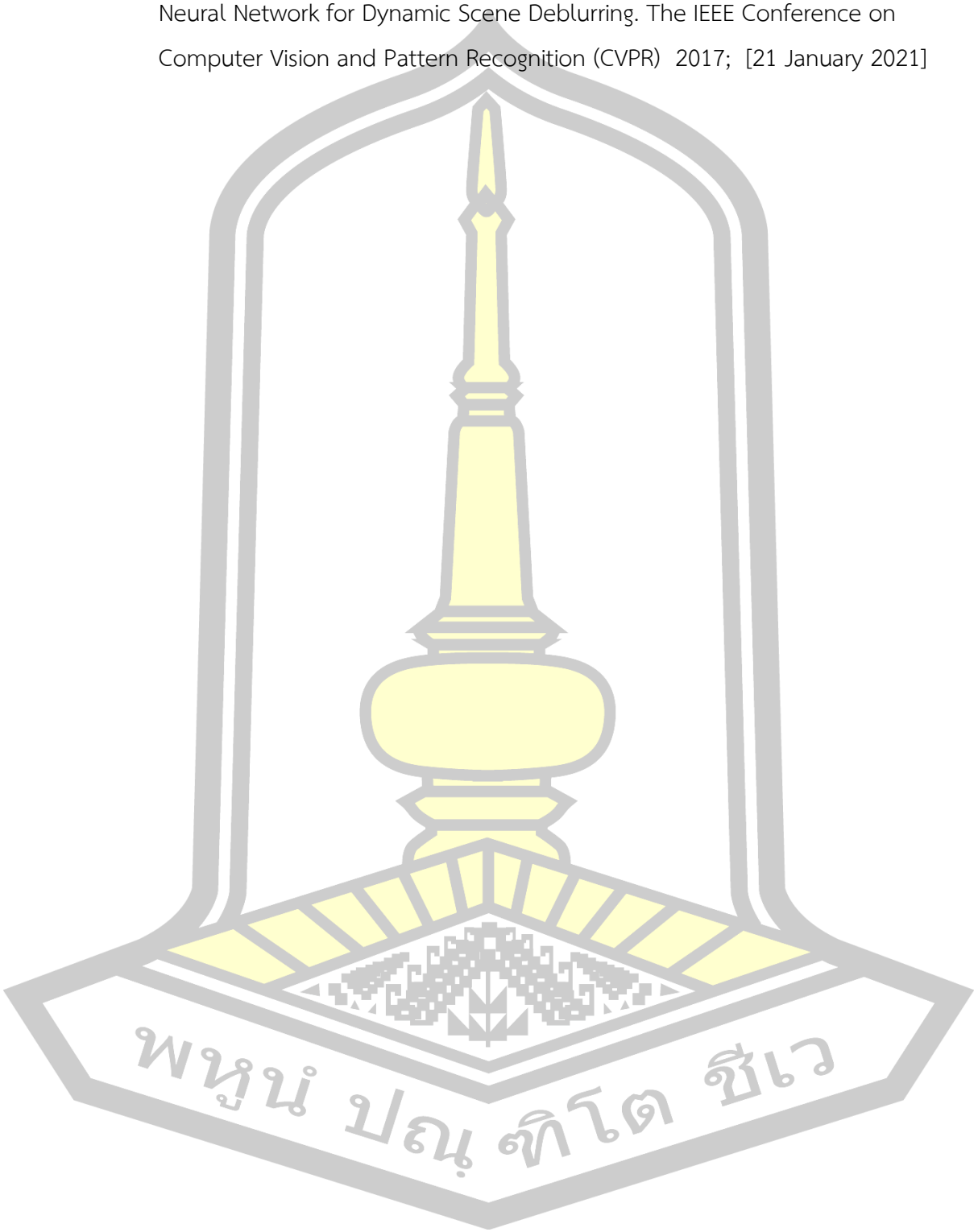
[55] Tao X, Gao H, Shen X, Wang J, Jia J. Scale-recurrent network for deep image deblurring. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 8174-8182.

[56] Hong M, Choe Y. Wasserstein Generative Adversarial Network Based De-Blurring Using Perceptual Similarity. Applied Sciences 2019; 9[11]: 2358.

[57] Li L, Pan J, Lai W-S, Gao C, Sang N, Yang M-H. Blind image deblurring via deep discriminative priors. International journal of computer vision 2019; 127[8]: 1025-1043.

[58] Ramakrishnan S, Pachori S, Gangopadhyay A, Raman S. Deep generative filter for motion deblurring. Proceedings of the IEEE international conference on computer vision workshops; 2993-3000.

[59] Nah SaK, Tae Hyun and Lee, Kyoung Mu. Deep Multi-Scale Convolutional Neural Network for Dynamic Scene Deblurring. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2017; [21 January 2021]



ประวัติผู้เขียน

ชื่อ	นายเอกราวี คำแปล
วันเกิด	3 พฤศจิกายน 2527
สถานที่เกิด	จังหวัดศรีสะเกษ
สถานที่อยู่ปัจจุบัน	583/59 ซ.20 มิถุนา 11 แขวงห้วยขวาง เขตห้วยขวาง กรุงเทพมหานคร 10310
ตำแหน่งหน้าที่การงาน	อาจารย์
สถานที่ทำงานปัจจุบัน	
ประวัติการศึกษา	-2550 ระดับปริญญาตรี มหาวิทยาลัยราชภัฏสวนดุสิต สาขาวิทยาการคอมพิวเตอร์ -2554 ระดับปริญญาโท มหาวิทยาลัยศรีปทุม สาขาวิทยาการสารสนเทศ -2566 ระดับปริญญาเอก มหาวิทยาลัยมหาสารคาม สาขาวิทยาการคอมพิวเตอร์
ทุนวิจัย	กระทรวงวิทยาศาสตร์และเทคโนโลยี