

การสร้างข้อมูลภาพจาก GANS เพื่อนำไปต่อยอดในด้าน Vector Art

GENERATING PICTURE DATA FROM GANS TO EHHANCE

VECTOR ART

ชาลี ชีพพิมลชัย¹

ธนภัทร นังกะจิตร²

บทคัดย่อ

ธุรกิจด้านการออกแบบ graphic หรือ อาชีพ graphic design เป็นงานที่สร้างมูลค่าให้กับหลายธุรกิจเช่น ธุรกิจสื่อ , ธุรกิจ Event รวมทั้งยังสร้างอาชีพที่เกิดใหม่เช่น Vector specialist , การออกแบบ UX/UI ดังนั้นเพื่อปรับปรุงงานออกแบบให้มีคุณภาพมากขึ้น ควรอาศัยความก้าวหน้าทางด้านเทคโนโลยีทางด้านคอมพิวเตอร์วิชั่นซึ่งเทคนิคการเรียนรู้เชิงลึกนี้จะใช้ Python เพื่อสร้างผลลัพธ์ graphic ที่มีคุณภาพออกมา

ในปัจจุบันเทคโนโลยีการเรียนรู้เชิงลึกทางด้านคอมพิวเตอร์วิชั่นนี้ถูกนำมาใช้ประโยชน์ในหลาย ๆ ด้าน ไม่ว่าจะเป็นการสร้างรูปภาพเหมือนหน้าคน , การสร้าง Animation ในต้นทุนที่ไม่สูง , การออกแบบรถโดยสารหลักการออกแบบที่ดีที่สุด , การสร้างข้อมูลสำหรับ detection object ในหลากหลายมิติ ผู้วิจัยจึงได้ มองเห็นศักยภาพของระบบเทคโนโลยีในปัจจุบันที่สามารถนำมาประยุกต์ใช้ในงานออกแบบ และจาก ผลงานวิจัยนี้จึงทำให้สามารถสร้างวัตถุที่ต้องการได้อย่างถูกต้อง รวดเร็วและแม่นยำขึ้น

ABSTRACT

A graphic design business or graphic design career is a project that creates value for many companies such as media business, event business, and creating new careers such as Vector specialist, and UX / UI design, to improve the design to be of high quality. This depends on the advancement of computer vision technology. The deep learning technique was used with Python to produce graphic in the quality results.

¹นักศึกษาหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมข้อมูลขนาดใหญ่

²ที่ปรึกษาสารนิพนธ์หลัก

Today, this deep learning technology in computer vision is being utilized in many fields, such as creating human face images, creating low-cost animations, designing cars by the best design principles. Furthermore, it makes data for detecting objects in multiple dimensions. Therefore, the researcher sees the potential of today's technology systems that can be applied in design work, and from this research, it is possible to create the desired object correctly, more rapidly, and more accurately.

บทนำ

การพัฒนาให้คอมพิวเตอร์มีความสามารถเหมือนมนุษย์ เช่น การมองเห็น การพูด การตัดสินใจ ความรู้สึก เป็นต้น ในงานประชุม “International Conference on Neural Information” ปี 2014 ได้ให้กำเนิดงานที่เป็นรากฐานในการ Generate รูปภาพ สำหรับ computer vision โดย model นั้น มีชื่อ ว่า GANS จุดประสงค์คือ เพื่อสร้าง Framework รูปแบบใหม่ ในการเทรน 2 model คือ 1) Generative model เอาไว้จับ data distribution 2) Discriminative model เอาไว้ประมาณ ค่าความน่าจะเป็น ของตัวอย่างในการ training

Machine Learning เป็นการใช้อัลกอริทึมในการแยกแยะและวิเคราะห์ข้อมูล เรียนรู้จากข้อมูลนั้น เพื่อสร้างโมเดลในการตัดสินใจหรือคาดการณ์เกี่ยวกับบางสิ่งบางอย่าง แทนที่จะเขียนโค้ด เป็นชุดคำสั่งให้ปฏิบัติตามขั้นตอนที่กำหนดไว้ แต่จะทำการ “เทรนด” โดยใช้ชุดข้อมูลขนาดใหญ่ในการเรียนรู้ที่จะปฏิบัติงานนั้น ๆ การที่จะสอนให้เครื่องคอมพิวเตอร์สามารถเรียนรู้ได้นั้นก็มีเทคนิคต่าง ๆ มากมาย เช่น Supervised Learning, Unsupervised Learning, Reinforcement Learning หรือ Artificial Neuron Networks เป็นต้น และหนึ่งในเทคนิคที่ถูกพูดถึงกันเป็นอย่างมาก ก็คือ “Deep Learning”

Artificial Neural Networks (ANN) เป็นเทคนิคที่เลียนแบบการทำงานของสมองมนุษย์ ซึ่ง ประกอบด้วยเซลล์ประสาท (Neuron) และแต่ละเซลล์จะถูกเชื่อมโยงกันเป็นโครงข่าย ซึ่งในซอฟต์แวร์ Neuron จะเรียกว่า “โหนด (Node)” และแต่ละโหนดจะถูกแบ่งออกเป็นชั้น (Layer) โดยหลักการ ของ Deep Learning ก็จะเป็น ANN ที่มีโหนดหลายๆชั้น ทำให้สามารถประมวลผลได้ครั้งละจำนวน มาก ช่วยให้การเรียนรู้ของเครื่องสามารถให้ผลลัพธ์ในการตัดสินใจและคาดการณ์ได้ดีมากยิ่งขึ้น

ปัจจุบันมีกลุ่มนักพัฒนางานด้าน Creating image เกิดขึ้นและแพร่ขยายไปทั่วโลก ทั้งในระดับ ทวีป หรือประเทศ โดยจะจัด Conference ทุกปี ใช้ชื่อว่า International computer vision (ICCV) ซึ่งเป็นองค์กรอิสระไม่หวัง ผลทางธุรกิจ เกิดจากการรวมกลุ่มของบริษัทต่างๆ ในภาคอุตสาหกรรมที่เกี่ยวข้องกับ Computer Vision และ ภาคการศึกษาที่มีการทำ Research ในด้าน Computer Vision

วัตถุประสงค์ของงานวิจัย

เพื่อสร้างผลลัพธ์ของงานรูปภาพและวิเคราะห์ให้ Generative Adversarial Networks model(GANs) มีความเสถียรมากขึ้น โดยจะนำ StyleGAN มาประยุกต์ใช้ในกระบวนการงานวิจัย และนำรูปภาพไปใช้ในธุรกิจ Graphic Design

ขอบเขตงานวิจัย

1. รูปแบบของชุดข้อมูลภาพที่ต้องการจะสร้างขึ้นใหม่ในการศึกษาอย่างชัดเจน
2. การนำเทคโนโลยีการเรียนรู้เชิงลึก (Deep Learning) มาช่วยในการสร้าง model ที่สามารถนำไปทำงานร่วมกับภาพที่เป็น ศิลปะ

ประโยชน์ที่คาดว่าจะได้รับ

1. เพื่อลดระยะเวลาในการสร้างงาน Art โดยใช้ Model ที่ถูกพัฒนาเพื่อให้เหมือนกับงานต้นแบบมากที่สุด
2. เพื่อให้ model สามารถสร้างภาพ vector เป็นภาพจริงได้ละเอียด และสมจริงขึ้น
3. นำไปสร้าง Sticker หรือ Graphic vector ทำได้โดยไม่จำเป็นต้องมีพื้นฐาน

ทฤษฎี และผลงานวิจัยที่เกี่ยวข้อง

1. เทคนิคการเรียนรู้ของเครื่อง (Machine Learning)

คือการใช้อัลกอริทึมในการแยกแยะและวิเคราะห์ข้อมูล เรียนรู้จาก ข้อมูลนั้น เพื่อสร้างโมเดลในการตัดสินใจหรือคาดการณ์เกี่ยวกับบางสิ่งบางอย่าง แทนที่จะเขียนโค้ด เป็นชุดคำสั่งให้ปฏิบัติตามขั้นตอนที่กำหนดไว้แต่จะใช้การ “เทรนด” โดยใช้ชุดข้อมูลขนาดใหญ่ในการเรียนรู้ที่จะปฏิบัติงานนั้น ๆ การที่จะสอนให้เครื่องคอมพิวเตอร์สามารถเรียนรู้ได้นั้นก็มีเทคนิคต่าง ๆ มากมาย เช่น Supervised Learning, Unsupervised Learning, Reinforcement Learning หรือ Artificial Neuron Networks เป็นต้น และหนึ่งในเทคนิคที่ถูกพูดถึงกันเป็นอย่างมาก ก็คือ “Deep Learning”

2. การเรียนรู้เชิงลึก (Deep Learning)

Deep Learning เป็นเซตย่อยของ Machine Learning หรือการเรียนรู้ด้วยเครื่อง ซึ่งเป็นลำดับขั้นของเครือข่ายประสาทเทียม (Artificial Neural Network) โดยดำเนินการด้วย Machine Learning เนื่องจากเครือข่ายประสาทเทียมถูกสร้างมาเหมือนสมองของมนุษย์มี Nodes เชื่อมต่อกันเหมือนเว็บไซต์ ระบบ Deep Learning ช่วยให้เครื่องประมวลผลข้อมูลด้วยวิธีการไม่ใช่เชิงเส้น ซึ่งเป็นเทคนิคการเรียนรู้เชิงลึก เพื่อกำจัดสิ่งแปลกปลอมในการทำธุรกรรมที่เสี่ยงต่อการหลอกลวงจะประกอบด้วยเวลา, ที่ตั้งทางภูมิศาสตร์, IP address, ประเภทของธุรกิจค้าปลีก และคุณลักษณะอื่น ๆ ที่ก่อให้เกิดวิธีการ หลอกลวงจากบุคคลที่ 3

3. โครงข่ายประสาทแบบคอนโวลูชัน (CNN)

โครงข่ายประสาทแบบคอนโวลูชัน(CNN) เป็นโครงข่ายประสาทเทียมหนึ่งในกลุ่ม bio-inspired โดยที่ จะจำลองการมองเห็นของมนุษย์ที่มองพื้นที่เป็นที่ย่อยๆ และนำกลุ่มของพื้นที่ย่อยๆ มาพาสกัน การมองพื้นที่ย่อยของมนุษย์จะมีการแยกคุณลักษณะ (feature) ของพื้นที่ย่อยนั้น เช่น ลายเส้น และการตัดกันของสี ซึ่งการที่มนุษย์รู้ว่าพื้นที่ตรงนี้เป็นเส้นตรงหรือสีตัดกัน เพราะมนุษย์ดูทั้งจุดที่สนใจและบริเวณรอบ ๆ ประกอบกัน มีส่วน

1. Convolution state คือ ส่วนที่ทำการแยกองค์ประกอบของรูปออกมา เช่น ขอบ สี รูปทรง เป็นต้น โดยที่ CNN จะมีตัวกรอง หรือ Filter ในการตรวจสอบเพื่อแยกองค์ประกอบของรูป

2. Detector stage คือ ส่วนที่รับ output มาจาก convolution state และ แปลงให้อยู่ในรูปของ nonlinear และเป็นส่วนที่ต้องระบุ อัลกอริทึม (Activation) ที่จะใช้ในการจำแนกคุณลักษณะของข้อมูล

3. Pooling stage คือ ตัวกรองแบบหนึ่งที่หาค่าสูงสุดในบริเวณที่ตัวกรองทาบอยู่บนข้อมูลแล้วเลือกค่าที่สูงที่สุดบนตัวกรองนั้นมาเป็นผลลัพธ์ใหม่ และจะเลื่อนตัวกรองไปตาม Stride ที่กำหนดไว้

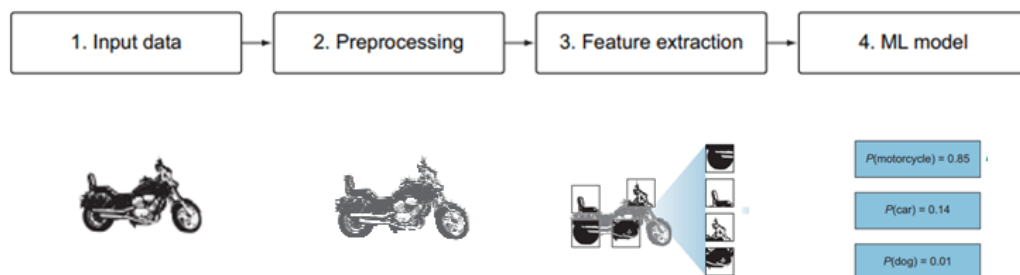
4. Computer vision

ศาสตร์ที่ให้ Computer สามารถเลียนแบบการมองเห็นของมนุษย์เพื่อจำแนกวัตถุหรือ สร้างขึ้นมาใหม่ โดยทำการรับ Input จาก อุปกรณ์ เช่น กล้อง หรือ เซนเซอร์ต่างๆเพื่อจำแนกวัตถุ หรือ ลักษณะต่างๆ โดยในศาสตร์นี้ จะจำแนกประเภทวัตถุหลักๆ เป็น 2 ประเภท 1) Low-Level จำแนกขอบ, เส้น , มุม ในรูปภาพ 2) High-Level จำแนกชิ้นส่วนต่างๆของรูปภาพ

ประเภทของ Computer vision ที่ใช้ในปัจจุบันมี 5 รูปแบบ

1. Image Classification การให้computer เรียนรู้และจำแนกภาพจาก Label(Supervise Learning) โดยงานที่นำมาประยุกต์ใช้ เช่น การจำแนกภาพปอดที่เป็นมะเร็ง
2. Object detection การที่คอมพิวเตอร์มีความสามารถเลียนแบบมนุษย์ในการการจำแนกภาพที่มองเห็นและแยก แยะ ได้ละเอียดมากขึ้น โดย Systems ที่ปัจจุบันใช้ เช่นYOLo(you only look once)
 - a. SSD(Single shot detector) , Faster R-CNN
3. Generating Art คอมพิวเตอร์สามารถcopy style จาก ภาพ A ไป B
4. Creating Images คอมพิวเตอร์สามารถสร้างภาพขึ้นมาจากภาพต้นแบบโดยใช้ machine learning
5. Face Recognition คอมพิวเตอร์สามารถระบุหรือติดตามใบหน้าคนโดยจะแบ่งเป็น Face Identification และ Face Varification

Computer vision pipeline



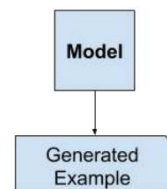
ภาพที่ 1: แสดง computer vision pipeline

1. ขั้นตอนแรก คอมพิวเตอร์ จะรับ input เข้ามา จาก อุปกรณ์เช่น กล้อง ที่จับภาพ หรือ วิดีโอ
2. ขั้นตอนที่สอง แต่ละภาพต้องทำการเปลี่ยนรูปแบบ(preprocessing) จุดประสงค์เพื่อให้ภาพมีพื้นฐานเหมือนกัน โดยส่วนใหญ่ จะทำ การปรับขนาดภาพ(resizing) , การทำให้ภาพชัดขึ้น(blurring) , การหมุนภาพ(rotating) , การเปลี่ยนรูปร่าง หรือ การเปลี่ยนจากภาพสีให้เป็นขาวดำ
3. ขั้นตอนที่สาม การจับคุณลักษณะ(extract feature) กระบวนการนี้จะทำการจับรูปแบบคุณลักษณะของภาพ เช่น รูปทรง , สี โดยจะตรวจสอบและ list คุณลักษณะที่มีความเฉพาะ ออกมา

4. ขั้นตอนที่สี่ Feature จะถูกนำเข้าไปใน การจัดกลุ่ม(classification model) เพื่อทำกระบวนการแยกแยะว่าภาพนี้คือ อะไร โดยจะใช้ feature ในการตัดสินใจภาพนั้นๆ

5. GANS

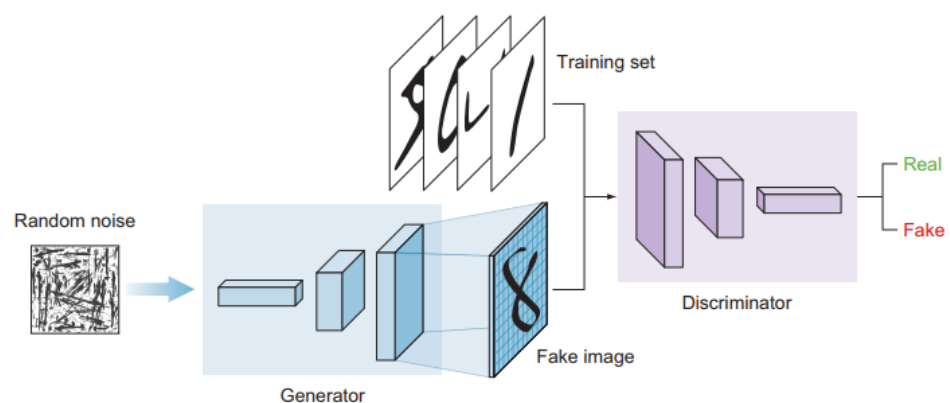
Generative Adversarial Networks (GANs) คือ generative model ที่ใช้ deep learning ในการเรียนรู้ โดย GANs จัดเป็น unsupervised learning โดยจะหา รูปแบบในการเรียนรู้ด้วยตัวเอง โดยหน้าที่ของ model นี้คือ สร้างผลลัพธ์จาก input และเกิดเป็น ผลลัพธ์ใหม่ขึ้นมา ภายใน model จะมีการสรุปการกระจายตัว ของ input ทำให้สามารถสร้าง ตัวอย่างขึ้นมาใหม่ ภายใน input distribution ได้



ภาพที่ 2: แสดงภาพรวมของ model gans

ใน GANs มีพื้นฐานมาจาก Generative model โดรนรูปแบบโครงสร้างใช้เหมือนกับ Generative model โดยจะมีองค์ประกอบอยู่ 2 ส่วน

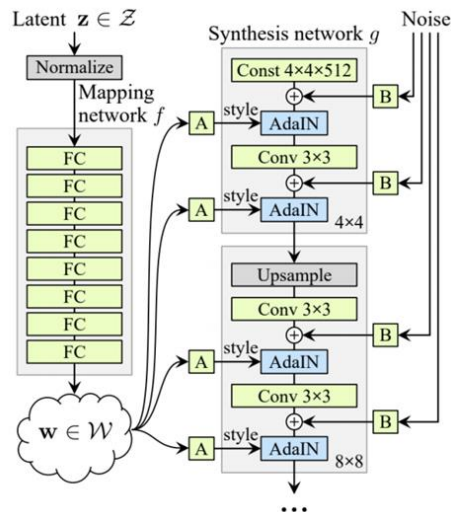
- 1.) Generator คือ โมเดลที่ใช้ในการสร้างตัวอย่างขึ้นมาใหม่ โดนครู่มความแปรปรวน จากข้อมูล ตั้งต้น
- 2.) Discriminator คือ โมเดลที่ใช้ในการจำแนกตัวอย่าง จริง หรือ ตัวอย่างปลอม



ภาพที่ 3: แสดงกระบวนการภายในของ gans

6. Style GAN

กระบวนการ GAN ชนิดนี้ เป็น กระบวนการส่วนที่ขยาย มาจาก progressive growing GAN วิธีการคือ generator model สังเคราะห์ ภาพขนาดใหญ่ คุณภาพสูง โดยเริ่มต้นจากทั้ง model discriminator และ generator โดยขนาดภาพนั้น จะเริ่มจากเล็กไปใหญ่ ในกระบวนการ training ใน ส่วน ของ generator จะไม่ได้ใช้จุดใดจุดหนึ่งใน latent space แต่จะใช้การสุ่มเพื่อสร้างภาพ แบบ mapping network แบบเดียว และ noise layers ผลลัพธ์ที่เกิดจากการทำ mapping network คือ vector ที่มีการกำหนดรูปแบบที่รวมไว้ในแต่ละจุดของ generator โดย layer แบบใหม่ เรียกว่า adaptive instance normalization(AdaIN)



ภาพที่ 4: แสดงกระบวนการภายในของ style gan

Mapping Network and AdaIN

การ standalone mapping network เป็น การใช้จุดสุ่มตัวอย่างจาก latent space as input และ สร้างในรูปแบบ vector ดังนั้น Mapping Network คือ ประกอบ fully connected layers 8 ชั้น หลังจากนั้น vector จะ ถูกแปลง และ รวมเข้ากับแต่ละ block ของ model โดยทำหลังจาก กระบวนการที่เรียกว่า adaptive instance normalization(AdaIN)

AdaIN layers คือ การสร้างมาตรฐาน output ของ feature map เพื่อทำ standard Gaussian และเพิ่ม vector ในรูปแบบ ของ bias

$$\text{AdaIN}(x_i, y) = y_{s,i} \frac{x_i - \mu(x_i)}{\sigma(x_i)} + y_{b,i}$$

ภาพที่ 6: สูตรของ AdaIN layers

การเพิ่ม รูปแบบใหม่ของ mapping network ใน โครงสร้าง ส่งผลต่อผลลัพธ์ใน Generator model ใน การ synthesis network

7. Pix2Pix

Image-to-image translation คือการควบคุมของรูปภาพต้นฉบับที่กำหนดเป็นรูปภาพเป้าหมาย เช่นการเปลี่ยนภาพขาวดำเป็นภาพสี การทำ image translation เป็นเรื่องที่ยากเนื่องจากต้องกำหนด model โดยเฉพาะ และ ค่า loss function ที่จะทำให้ผลลัพธ์ออกมาสมบูรณ์

Pix2Pix GAN เป็นประเภทของ conditional gan โดย discriminator จะเตรียม รูปภาพจริง และ รูปภาพเป้าหมาย โดยเป้าหมายใน model คือการพิจารณาความเป็นไปได้ในการเปลี่ยนผลลัพธ์ ส่วน generator มีหน้าที่ในการสร้างภาพเพื่อให้สามารถทำให้ใกล้เคียง กับ target domain ได้

8. CycleGAN

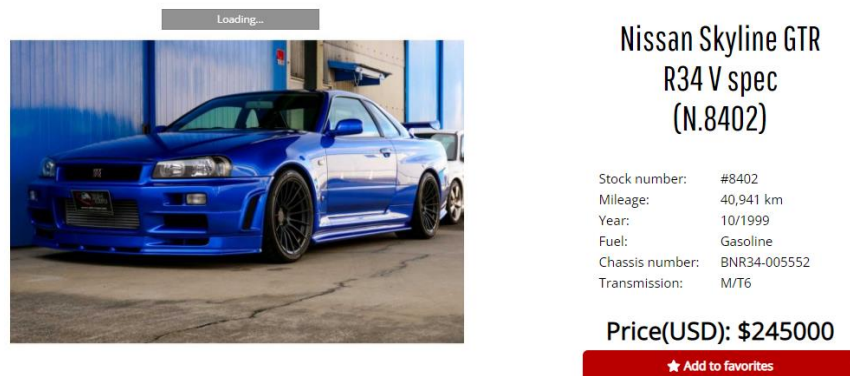
กระบวนการ GAN ชนิดนี้ เป็นเทคนิค การ training ของ image translation โดยไม่ต้องมีภาพตัวอย่าง CycleGAN จะถูก train ในลักษณะ unsupervised โดยจะใช้ข้อมูลตั้งต้น และ target domain ที่ไม่ได้มีความเกี่ยวข้องกัน ตัวอย่างที่นำเทคนิคนี้ไปใช้คือ การแปลงภาพม้าเป็นม้าลาย และ ม้าลายเป็นม้า

CycleGAN จะมีส่วนขยายเพิ่ม จาก GANs ปกติคือ cycle consistency แนวคิดคือการนำเอาผลลัพธ์อันแรกที่ได้จากการ Train generator ครั้งที่หนึ่ง มาใช้ใน generator ครั้งที่สอง และ เอาผลลัพธ์ครั้งที่สอง มาดูความสัมพันธ์กับภาพจริง

ระเบียบวิธีวิจัย

1. ข้อมูลที่ใช้ในงานวิจัย

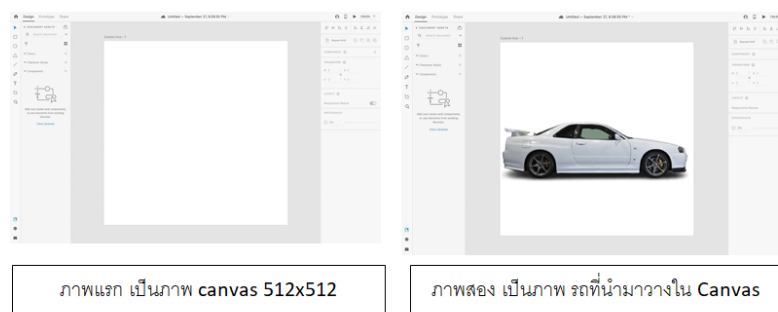
ข้อมูลที่ใช้ในงานวิจัยเป็นข้อมูลภาพถ่ายรถ GTR R33,R32,R34,R35 ซึ่งใช้ภาพทั้งหมด 300 ภาพ มีขนาด 512 x 512 pixel โดยได้จากแหล่งข้อมูล jdm-expo และ Nissan Japan official



ภาพที่ 7: แสดงตัวอย่างของแหล่งข้อมูลที่ใช้ในงานวิจัย

2. การเตรียมข้อมูล (Prepare Data)

- 1) ลบภาพพื้นหลังของภาพโดยใช้ tool ของ remove.bg โดย tool จะทำหน้าที่เอาแต่ object ที่เราต้องการ
- 2) ปรับแก้ขนาดของภาพ ให้มีขนาดเท่ากับ 512 x 512 pixel
- 3) ใช้เทคนิคจำลองภาพพื้นหลังให้เหมือนกันทั้งหมดโดยใช้พื้นหลังสีขาว



ภาพที่ 8: แสดงวิธีการใช้เทคนิคจำลองพื้นหลัง

3. เตรียม Environment ก่อนสร้าง model style gan

ในงานวิจัย GAN ฉบับนี้ได้เลือกใช้ Environments จาก google colab เพื่อให้ผู้ที่สนใจสามารถนำไปประยุกต์หรือศึกษาต่อได้โดย google colab ที่ใช้จะเป็น google colab ชนิด pro ซึ่งมี

ค่าใช้จ่ายในการสมัครซึ่งได้รับความนิยมเป็นอย่างมากในหมู่ นักพัฒนาโปรแกรม รวมถึงผู้สนใจด้าน Deep learning เพราะติดตั้งได้ง่าย รวมถึง การติดตั้ง package ต่าง ๆ ที่สำคัญ โดยเฉพาะ TensorFlow , ml4a(Machine Learning for Artists) , และอื่น ๆ อีกมากมาย

นอกจากนี้ google colab ยังมีการจำลองมาให้ใช้แบบไม่คิดค่าบริการสำหรับงานtrain ที่ไม่ได้มีความละเอียดสูงและ Dashboard TensorFlow ช่วยบริหารจัดการ ram และ การ์ดจออัตโนมัติ โดยมีขั้นตอนดังนี้

- 1) เปิด Colab เข้าไปที่เว็บไซต์ <https://colab.research.google.com> File > New notebook
- 2) ให้กดที่ Edit เปลี่ยน Notebook setting เป็น GPU และ High-RAM
- 3) การติดตั้ง Python Library ที่จำเป็นต้องใช้งานในโปรเจก ได้แก่ TensorFlow version แรก และ ชุด Code จาก stylegan2-ada

4. Preprocess model training

ให้นำภาพทั้งหมดที่ทำเป็นขนาด 512x512 PX มาเปลี่ยนชนิดของfile โดยขั้นตอนนี้ผู้วิจัยพบว่าในกรณีที่ถ้าไม่เปลี่ยนชนิดของไฟล์ให้เหมาะสม เช่น ชื่อไฟล์ไม่ได้เรียงตามเลข หรือ ชนิดไฟล์เป็นไฟล์ PNG จะไม่สามารถแปลงไฟล์ได้

```
import os
import glob

os.chdir(r"c://Users//chalee//Desktop//art_gtr//gtr//images")
for index, oldfile in enumerate(glob.glob("*.png"), start = 0 ):
    newfile = '00000{}.png'.format(index)
    os.rename (oldfile,newfile)
```

ภาพที่ 9: ภาพ code แสดงการเปลี่ยนชื่อไฟล์

หลังจากนั้น ให้นำข้อมูลไฟล์ภาพ ไปแปลงให้อยู่ในรูปของformat tfrecords หรือ large numpy arrays โดยทำหน้าที่เก็บหลายๆ sample



ภาพที่ 10: ภาพแสดงการเก็บข้อมูลของ TFRecord file

5. Process กระบวนการเรียนรู้ GANS

ในขั้นตอนการวิจัยในส่วนนี้ เป็นการทำงานในกระบวนการของ GANS ซึ่งการทำงานของ GANS จะมีขั้นตอนหลักๆอยู่ 2 ขั้นตอน คือ การทำงานในส่วน ของโมเดล และ การนำโมเดลที่ได้ไปใช้งานโดยการทำงานในส่วน โมเดลนั้นจะมีขั้นตอนการทำงาน หลักมีอยู่ 5 ขั้นตอน ดังนี้

- 1) Load library
- 2) Load Data
- 3) Load code
- 4) Save target model
- 5) Generate

6. เครื่องมือที่ใช้ในการวิจัย

- 6.1 TensorFlow เป็นไลบรารีที่ใช้ในการพัฒนา Machine Learning พัฒนาโดย Google
- 6.2 PyTorch เป็น Deep learning library ของภาษา Python พัฒนาโดย Facebook
- 6.3 NumPy เป็นโมดูลเสริมใน Python มีฟังก์ชันเกี่ยวกับคณิตศาสตร์ แลพ การคำนวณ
- 6.4 Colab เป็นบริการ Software as a Service (Saas) โฮสต์โปรแกรม Jupyter Notebook บน Cloud จาก Google ใช้เขียนโปรแกรมภาษา Python และมี GPU และ TPU มาให้

7. การนำแบบจำลองไปประยุกต์ใช้กับข้อมูลจริง

ขั้นตอนการนำ GANSไปใช้งานมีความคล้ายคลึงกับขั้นตอนการเตรียมข้อมูลเพื่อนำมาสร้างมา GANS จะแตกต่างกันเวลาดึงค่าชนิดสกุลไฟล์ หรือ เวอร์ชันของ package ที่ใช้ในการ train ทั้งนี้ขึ้นอยู่กับจุดประสงค์ของผู้ใช้งาน รวมทั้งข้อมูลที่จะนำมา train ควรจะไปในทางเดียวกัน

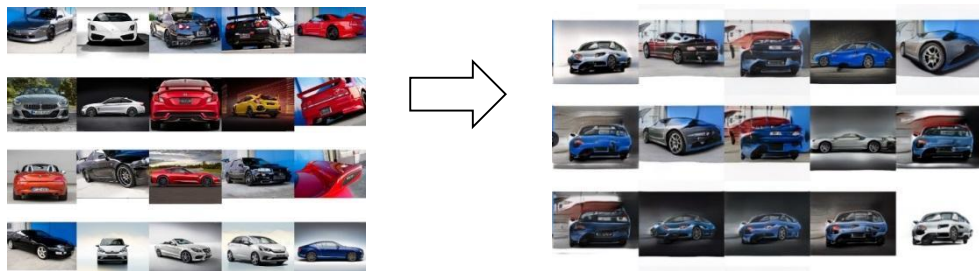
เพื่อให้ผลลัพธ์เกิดความเสถียรมากที่สุด โดยผู้วิจัยได้นำไปประยุกต์ใช้กับงาน Few-shot Image Generation via Cross-domain Correspondence ที่ทำให้ภาพจริงเป็นภาพศิลปะ



ภาพที่ 11: ภาพแสดงการนำ pretrain จาก GANS ไปประยุกต์ใช้งาน

ผลการศึกษา

1. ผลลัพธ์ของข้อมูลที่ไม่ได้ถูกการตัด Background



ภาพที่ 12 : ผลลัพธ์ของข้อมูลที่ไม่ได้ถูกการตัด Background

2. ผลลัพธ์ของข้อมูลที่ถูกการตัด Background และไปในทิศทางเดียวกัน



ภาพที่ 13 : ผลลัพธ์ ของข้อมูลที่ถูกการตัด Background และ ไปในทิศทางเดียวกัน

3 ผลลัพธ์จากการสร้างภาพจาก pretrain

ตารางที่ 1 แสดงวัดผลลัพธ์ที่ได้จากการวัดค่า Frèchet inception distance (FID)

Method	GTR_DATASET
A Baseline Progressive GAN	8.05
B +Tuning (incl. bilinear up/down)	5.5
C +Add mapping and styles	4.85
D +Remove traditional input	4.9
E +Add noise inputs	4.45
F +Mixing regularization	4.2

จากตารางที่ 1 แสดงวัดผลลัพธ์ที่ได้จากการวัดค่า Frèchet inception distance (FID) ยิ่งค่า Frèchet inception distance (FID) มีค่าที่เข้าใกล้ 0 แสดงว่าภาพนั้นสามารถสร้างได้เทียบเท่ากับของจริง

บทสรุปและข้อเสนอแนะ

1. สรุปผลการศึกษา

ในการเตรียมข้อมูลรูปภาพได้มีการนำ canvas เทียม มาประยุกต์ใช้งาน เนื่องจากภาพถ่ายที่มีหลายมุม เป็นภาพที่เมื่อเปลี่ยนองศาของภาพจะทำให้ภาพเปลี่ยนรูปร่างไปด้วยโดยจะเอาเฉพาะมุมที่ต้องการมาประยุกต์ใช้เท่านั้น รวมถึงการตัดพื้นหลังที่ไม่เกี่ยวข้องทำให้ผลลัพธ์ที่ได้จากการวิเคราะห์ ข้อมูลภาพมีความละเอียดถูกต้อง และมีความสะดวก รวดเร็วในการเตรียมข้อมูลภาพมากยิ่งขึ้น

ในส่วนของการประยุกต์ใช้เทคโนโลยี GANS กับภาพถ่าย รถยนต์ ได้เลือกใช้ ภาษาไพธอนในการดำเนินงานเนื่องจากเป็นภาษาคอมพิวเตอร์ระดับสูงที่สามารถนำไปประยุกต์ใช้ ได้กับงานในหลายๆด้าน และสามารถนำไปประยุกต์ใช้กับระบบปฏิบัติการอื่นๆได้ ส่วนไลบรารีที่ใช้ในการพัฒนาจะใช้ไลบรารี TensorFlow กับ PyTorch เป็นหลัก เนื่องจากเป็น deep learning library ที่

ใช้งานง่ายและมีประสิทธิภาพสูงในการรันโมเดล ได้รับความนิยมในปัจจุบันในการใช้พัฒนา Machine Learning โดยแนวคิดของเทคโนโลยีการเรียนรู้เชิงลึกนั้นเป็นการนำข้อมูลไม่ว่าจะเป็น ข้อมูลภาพตัวอักษร หรืออื่นๆ มาสอนหรือ train ให้โมเดลเข้าใจว่าสิ่งๆ นั้นคืออะไร โดยในการ train โมเดลนั้นใน บางครั้งอาจจะต้องใช้ข้อมูลจำนวนมากเพื่อให้โมเดลสามารถเรียนรู้และเข้าใจได้ เมื่อ ทำการ train โมเดลแล้วจะนำโมเดลที่ได้ไปทดสอบกับภาพใหม่เพื่อให้ได้ผลลัพธ์ออกมา ซึ่ง เทคโนโลยีการเรียนรู้เชิงลึกที่นำมาใช้คือ GANS โดยโมเดลสามารถนำไปสร้างข้อมูลภาพขึ้นมา ใหม่ได้จากจุดเริ่มต้น โดยจะมีคุณภาพมากกว่าการทำ data augmentation ทั่วไปคือ การสามารถ style transfer ในสิ่งที่ต้องการได้และสามารถสร้างภาพซ้ำได้ในปริมาณมาก

2. ข้อเสนอแนะ

2.1 สามารถนำไปพัฒนาต่อยอดให้กับธุรกิจต่างๆ ได้เช่น ธุรกิจการออกแบบ, ธุรกิจสื่อ เป็นต้น

2.2 สามารถใช้ภาพถ่ายประเภทอื่นได้เช่น ยานพาหนะ และอื่นๆ มาประยุกต์กับโมเดล ได้ แต่ต้องมีการวิเคราะห์และตรวจสอบรายละเอียดของภาพให้ชัดเจน เนื่องจากถ้าเป็นภาพที่คนละ ประเภทกันและมุมที่ต่างกันการนำ pretrain ที่ไม่ได้เกี่ยวข้องมา train ภาพจะทำให้ภาพไม่สามารถ สร้างภาพขึ้นมาได้

2.3 สามารถนำไปประยุกต์ใช้ในงานด้านอื่น ๆ ได้ แต่ต้องมีการปรับโมเดลให้เข้ากับงาน นั้น ๆ ด้วย เนื่องจากโมเดล Deep Learning เป็น โมเดลที่ทำงานเฉพาะด้าน

2.4 ในการเตรียมข้อมูลภาพที่ใช้เทรนโมเดลควรเก็บข้อมูลให้ถูกต้องทุกครั้งและตัดสิ่ง ที่ไม่ต้องการจะ train ออก เพื่อให้ผลลัพธ์ที่สร้างออกมามีความถูกต้องและลดระยะเวลาการทำงาน

บรรณานุกรม

Jakub Langr and Vladimir Bok. *GANs in Action*. manning September 2019

Jason Brownlee. *Generative Adversarial Networks with Python*. manning July 2019

Ran Yi, Yong-Jin Liu, *APDrawingGAN: Generating Artistic Portrait Drawings from Face Photos with Hierarchical GANs*. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019

Tero Karras , Miika Aittala ,Samuli Laine. *Training Generative Adversarial Networks with Limited Data*. arXiv:2006.06676v2 [cs.CV] 7 Oct 2020

Tero Karras , Miika Aittala ,Samuli Laine. *Analyzing and Improving the Image Quality of StyleGAN*. arXiv:1912.04958v2 [cs.CV] 23 Mar 2020