

เรื่อง การศึกษาและเปรียบเทียบประสิทธิภาพของการสร้างภาพ Depth
ด้วยกระบวนการทำงานของโครงข่ายประสาทเทียม

โดย 1. นายชนวัชร ปานมณี
2. นายสิปปภาส ปิติวรรณ
3. นายรชต ประทีป ณ ถลาง
โรงเรียนยุพราชวิทยาลัย

รายงานฉบับนี้เป็นส่วนประกอบของโครงการงานวิทยาศาสตร์ ระดับมัธยมศึกษา ตอนปลาย
ในงานเวทีวิชาการนวัตกรรมสะเต็มศึกษาขั้นพื้นฐานแห่งชาติ ครั้งที่ 1 (ออนไลน์)

The 1st National Basic STEM Innovation E-Forum 2021

วันที่ 18 – 19 กันยายน พ.ศ. 2564

เรื่อง การศึกษาและเปรียบเทียบประสิทธิภาพของการสร้างภาพ Depth
ด้วยกระบวนการทำงานของโครงข่ายประสาทเทียม

โดย 1. นายธนวัชร ปานมณี
2. นายสิปปภาส ปิติวรรณ
3. นายรชต ประทีป ฒกลาง
โรงเรียนยุพราชวิทยาลัย

อาจารย์ที่ปรึกษา นายศราวุธ วิริยะคุณานันท์
ที่ปรึกษาพิเศษ นางสาวรุ่งทิวา บุญมาโตน

ชื่อโครงการ	การเปรียบเทียบประสิทธิภาพของการสร้างภาพ Depth ด้วยกระบวนการทำงานของโครงข่ายประสาทเทียม
ผู้จัดทำโครงการ	1. นายธนวัชร ปานมณี 2. นายกฤตภัทร์ สมศรี 3. นายรชต ประทีป ณ ถลาง
อาจารย์ที่ปรึกษาโครงการ	นายศรายุทธ วิริยะคุณานันท์
ที่ปรึกษาพิเศษ	นางสาวรุ่งทิวา บุญมาโดน
โรงเรียน	ยุพราชวิทยาลัย
ที่อยู่	238 ถนนพระปกเกล้า ตำบลศรีภูมิ อำเภอเมืองเชียงใหม่ จังหวัดเชียงใหม่ โทรศัพท์ 0 – 5341 – 8673 – 5 โทรสาร 0 – 5341 – 8673 ต่อ 111
ระยะเวลาในการทำโครงการ	ตั้งแต่เดือน มกราคม – กรกฎาคม พ.ศ. 2564

บทคัดย่อ

ภาพ Depth เป็นภาพสองมิติที่แต่ละหน่วย (พิกเซล) ของภาพมีความสัมพันธ์กับระยะห่างระหว่างระนาบของภาพกับวัตถุของจริงในระบบ RGB เพื่อสะท้อนความเป็นสามมิติของภาพจริงบนภาพสองมิติให้ได้มากที่สุด

ผู้จัดทำต้องการสร้างภาพ Depth โดยใช้ชุดข้อมูลในการเทรนโครงข่ายประสาทเทียมให้น้อยที่สุด ด้วยการปรับปรุงแบบจำลอง Depth estimation model ของ Anindya Basu เนื่องจากแบบจำลองนี้ค่อนข้างมีปัญหา เมื่อมีชุดข้อมูลในการเทรนจำนวนน้อย ผู้จัดทำได้ปรับปรุงแบบจำลองของ Anindya โดยการเพิ่ม encoder ที่ช่วยในการแปลงข้อมูลเป็นเวกเตอร์ จากนั้นใช้หลักการการวัดระยะห่างของ Euclidean เปรียบเทียบความแตกต่างระหว่างข้อมูลนำเข้ากับข้อมูลส่งออก แล้วเพิ่ม discriminator ซึ่งเป็น encoder อีกตัวที่ถูกต่อด้วยฟังก์ชัน sigmoid 1-neural ที่ให้ค่าความน่าจะเป็นของการเป็นภาพ Depth ทั้งนี้เพื่อตรวจสอบว่าผลลัพธ์ที่ได้เป็นภาพ Depth จริงหรือไม่

ผู้จัดทำได้เตรียมไฟล์รูปภาพนามสกุล .jpeg ขนาด 256x256x1 ประกอบด้วย ข้อมูล X เป็นเซตของภาพระดับสีเทา (grayscale images) และ Y เป็นเซตของภาพ Depth ในการเทรน Depth estimation model เพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองที่ปรับปรุงจากแบบจำลองของ Anindya กับแบบจำลองเดิม จากการศึกษาพบว่า ด้วยชุดข้อมูลในการเทรน 100 ชุดเท่ากัน แบบจำลองที่ปรับปรุงใหม่ให้ค่าระยะห่างของ Euclidean ที่น้อยกว่าแบบจำลองเดิม นั่นหมายความว่าภาพ Depth ที่ได้มีความเหมือนกับรูปต้นแบบมากกว่า แต่แบบจำลองที่ปรับปรุงใหม่ให้ค่า discriminator score น้อยกว่า ทั้งนี้เป็นเพราะแบบจำลองของ Anindya นำภาพ Depth ที่ได้ไปเปรียบเทียบกับภาพ Depth ในชุดข้อมูลที่ใช้ในการเทรน ซึ่งเหตุการณ์นี้เรียกว่า overfitting

กิตติกรรมประกาศ

โครงการนี้สามารถสำเร็จลุล่วงตามเป้าหมายเพราะได้รับความช่วยเหลือและคำแนะนำที่เป็นประโยชน์อย่างยิ่งจากผู้มีพระคุณหลายท่าน จึงขอขอบคุณทุก ๆ ท่าน ดังนี้

กราบขอบพระคุณ คุณครูศรายุทธ วิริยะคุณานันท์ และ คุณครูนันทนัช นันทพงษ์ คุณครูที่ปรึกษาโครงการ คุณครูรุ่งทิวา บุญมาโตน คุณครูจากกลุ่มสาระการเรียนรู้คณิตศาสตร์ โรงเรียนยุพราชวิทยาลัย ผู้ที่ให้คำแนะนำและได้เมตตาให้ความช่วยเหลือในทุก ๆ ด้าน ตลอดจนให้คำปรึกษาเกี่ยวกับปัญหาที่เกิดขึ้นในระหว่างการทำโครงการนี้ ตรวจทานรูปเล่มรายงานจนโครงการสำเร็จลุล่วงไปด้วยดี

ท้ายที่สุด ขอกราบขอบพระคุณ คุณพ่อและคุณแม่ ผู้เป็นที่รัก ผู้ให้กำลังใจและให้โอกาสในการศึกษาอันมีค่ายิ่ง

คณะผู้จัดทำ

สารบัญ

	หน้า
บทคัดย่อ	ก
กิตติกรรมประกาศ	ข
สารบัญ	ค
สารบัญตาราง	ง
สารบัญรูปภาพ	จ
บทที่ 1 บทนำ	1
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง	3
2.1 ระบบโครงข่ายประสาทเทียม	3
2.2 Backpropagation	4
2.3 Loss Function	4
2.4 ระยะห่างของ Euclidean (Euclidean Distance)	5
บทที่ 3 วิธีดำเนินงาน	6
บทที่ 4 ผลการศึกษา	11
บทที่ 5 สรุปผลการศึกษา	12
บรรณานุกรม	12

สารบัญตาราง

	หน้า
ตารางที่ 4.1 ตารางเปรียบเทียบผลลัพธ์จากการสร้างภาพ Depth ของแบบจำลองที่ปรับปรุง Depth estimation model ของ Anindya กับแบบจำลองของ Anindya เดิม	3

สารบัญรูปภาพ

	หน้า
ภาพที่ 2.1 ส่วนประกอบของ Neural Network	3
ภาพที่ 3.1 แผนผังลำดับงานของ Depth estimation model ที่ปรับปรุงจากแบบจำลองของ Anindya	6
ภาพที่ 3.2 การเตรียมข้อมูลนำเข้าจากภาพสีในระบบ RGB เป็นเป็นเมทริกซ์ ขนาด 256x256x3	7
ภาพที่ 3.3 การปรับค่าสมาชิกในแต่ละตำแหน่งของเมทริกซ์ขนาด 256x256x1	7
ภาพที่ 3.4 แผนผังลำดับงานของ Depth Image Generator	8
ภาพที่ 3.5 แผนผังลำดับงานของ Encoder	8
ภาพที่ 3.6 แผนผังลำดับงานของ Discriminator	9
ภาพที่ 3.7 แผนผังลำดับงานการหาค่าคลาดเคลื่อนจากการดำเนินงานตามแบบจำลอง	9
ภาพที่ 3.8 แผนผังลำดับงานการเทรนแบบจำลอง Depth Image Generator	10

บทที่ 1 บทนำ

1.1 ที่มาและความสำคัญของโครงการ

Milton Guerry (2020) ประสานสหพันธ์หุ่นยนต์นานาชาติ ให้ข้อมูลว่ามีการใช้หุ่นยนต์ในโรงงานอุตสาหกรรมทั่วโลกเพิ่มขึ้นถึง 85% ภายใน 5 ปี ถือเป็นระดับสูงสุดในประวัติศาสตร์ของการใช้หุ่นยนต์อัจฉริยะและระบบอัตโนมัติ

ภาพ Depth ถูกใช้ในวงการหุ่นยนต์อย่างแพร่หลาย เช่น รถยนต์อัตโนมัติไร้คนขับ การสร้างฉากสามมิติขึ้นมาใหม่ รวมถึงเทคโนโลยีเสมือนจริง (AR) โดยใช้เซนเซอร์วัดระยะทาง แต่เนื่องจากเทคโนโลยีนี้ใช้ทรัพยากรและงบประมาณค่อนข้างสูง การจำลองภาพ Depth จึงได้ถูกสร้างขึ้นมาเพื่อทดแทนกระบวนการที่ต้องใช้งบประมาณสูง

เป้าหมายของการจำลองภาพ Depth คือการสร้างข้อมูลสามมิติ จากภาพสองมิติ เนื่องจากระบบการมองเห็นของมนุษย์และคอมพิวเตอร์ที่สามารถเห็นภาพได้แค่สองมิติเท่านั้น นอกจากนี้ภาพ Depth ที่ถูกสร้างขึ้นมาจะให้ข้อมูลความลึกดีและระยะห่างจากจุดสนใจกับตำแหน่งกล้องในภาพด้วยการไล่ระดับโทนสีจากเข้ม (อยู่ใกล้ตำแหน่งกล้อง) ไปสว่าง (อยู่ไกลตำแหน่งกล้อง)

Anindya Basu (1995) ได้คิดค้นแบบจำลองในการสร้างภาพ Depth โดยใช้ชุดข้อมูลประมาณ 40,000 ชุด ไปเทรน Depth estimation model เพื่อให้ Depth estimation model สร้างภาพ Depth จากนั้นนำภาพ Depth ที่ได้มาเปรียบเทียบกับภาพ Depth ที่ถูกสร้างขึ้นจากเซนเซอร์ โดยใช้หลักการความคลาดเคลื่อนเฉลี่ยกำลังสอง (Mean square errors)

ปัญหาในงานของ Anindya คือ เมื่อใช้จำนวนชุดข้อมูลในการเทรนที่มีปริมาณน้อย จะทำให้ได้ภาพ Depth ที่ออกมาแตกต่างจากรูปต้นแบบในชุดข้อมูลขาเข้ามาก ซึ่งสามารถสังเกตเห็นได้ด้วยตาเปล่า อีกทั้งแบบจำลองของ Anindya ยังขาดการตรวจสอบว่า ภาพ Depth ที่ได้มานั้นเป็นภาพ Depth จริงหรือไม่

ผู้จัดทำจึงคิดปรับปรุงแบบจำลองของ Anindya ให้มีการตรวจสอบความเหมือนกันของรูปต้นแบบกับภาพ Depth ที่สร้างขึ้น โดยอาศัยหลักการการวัดระยะห่างของ Euclidean ถ้าระยะห่างของ Euclidean มีค่าน้อย หมายความว่าภาพ Depth ที่ได้มีความเหมือนกับรูปต้นแบบมาก นอกจากนี้ยังเพิ่มขั้นตอนที่ช่วยลดระยะห่างของ Euclidean ลงด้วย อีกประการหนึ่งผู้จัดทำต้องการเพิ่มกระบวนการในการตรวจสอบว่าภาพ Depth ที่ได้มานั้นมีความเป็นภาพ Depth มากน้อยเพียงใด โดยอาศัยฟังก์ชัน sigmoid 1-nueral ที่ให้ค่าความน่าจะเป็นของการเป็นภาพ Depth

1.2 วัตถุประสงค์ของโครงการ

เพื่อลดปริมาณชุดข้อมูลในการเทรน Depth estimation model

1.3 สมมติฐาน

ด้วยชุดข้อมูลในการเทรน 100 ชุด เท่ากัน Depth estimation model ที่ปรับปรุงจากแบบจำลองของ Anindya สามารถสร้างภาพ Depth ได้มีประสิทธิภาพมากกว่า Depth estimation model ของ Anindya Basu เพราะมีกระบวนการในการลดความแตกต่างของภาพ Depth ที่ได้กับรูปต้นแบบ อีกทั้งยังมีการตรวจสอบความเป็นภาพ Depth ด้วย

1.4 ตัวแปรที่ศึกษา

- 1.4.1 ตัวแปรต้น - Depth estimation model ที่ปรับปรุงจากแบบจำลองของ Anindya Basu
- Depth estimation model ของ Anindya Basu
- 1.4.2 ตัวแปรตาม - ค่าระยะห่างของ Euclidean ระหว่างรูปต้นแบบและภาพ Depth ที่สร้างขึ้น
- ค่า discriminator score
(ความน่าจะเป็นของการเป็นภาพ Depth มีค่าตั้งแต่ 0 – 1)
- 1.4.3 ตัวแปรควบคุม - ชุดข้อมูลในการเทรนจำนวน 100 ชุด
- ลักษณะของภาพของรูปต้นแบบในชุดข้อมูลนำเข้า

1.5 นิยามเชิงปฏิบัติการ

ประสิทธิภาพของภาพ Depth ที่ได้จาก Depth estimation model พิจารณาจากค่าระยะห่างของ Euclidean และค่า discriminator score (มีค่าตั้งแต่ 0 – 1) ดังนี้

- ค่า ระยะห่างของ Euclidean มีค่าน้อย หมายถึงภาพ Depth ที่ได้มีความเหมือนกับรูปต้นแบบมาก
- ค่า discriminator score มีค่าเข้าใกล้ 1 หมายถึงภาพ Depth ที่ได้ มีความเป็นภาพ Depth มากกว่า

1.6 ขอบเขตของการศึกษา

ด้วยข้อจำกัดด้านทรัพยากรและงบประมาณในการสร้างชุดข้อมูลที่เป็นภาพ Depth ที่ใช้เทรนแบบจำลอง ทำให้ผู้จัดทำไม่สามารถกำหนดตัวแปรต้นเป็นจำนวนชุดข้อมูลในการเทรนที่แตกต่างกันได้ ผู้จัดทำจึงใช้วิธีการเปรียบเทียบประสิทธิภาพของ Depth estimation model ที่ปรับปรุงจากแบบจำลองของ Anindya กับ Depth estimation model ของแบบจำลองเดิม ด้วยชุดข้อมูลในการเทรนจำนวน 100 ชุดเท่ากัน โดยใช้ชุดข้อมูลที่เป็นภาพ Depth ในการเทรนจากงานวิจัยในหัวข้อ Disentangling Task Transfer Learning ของ Zamir และคณะ ในปี ค.ศ. 2018 ที่สืบค้นจาก <http://taskonomy.stanford.edu/> เมื่อวันที่ 10 เมษายน พ.ศ. 2564 แล้วเปรียบเทียบประสิทธิภาพของภาพ Depth ที่ได้จากแบบลองทั้งสอง

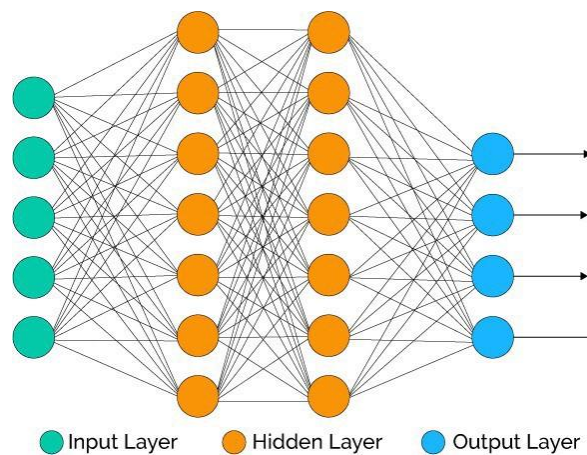
1.7 ประโยชน์ที่คาดว่าจะได้รับ

ได้ Depth estimation model ที่สามารถทำงานได้อย่างมีประสิทธิภาพในชุดข้อมูลในการเทรนที่มีปริมาณเพียงแค่ 100 ชุด

บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง

2.1 ระบบโครงข่ายประสาทเทียม

โครงข่ายประสาทเทียม (Neural Network) คือ แบบจำลองทางคณิตศาสตร์หรือแบบจำลองทางคอมพิวเตอร์สำหรับประมวลผลสารสนเทศด้วยการคำนวณแบบคอนเนกชันนิสต์ (connectionist) แนวคิดเริ่มต้นของเทคนิคนี้ได้มาจากการศึกษาโครงข่ายไฟฟ้าชีวภาพ (bioelectric network) ในสมอง ซึ่งประกอบด้วย เซลล์ประสาท (neurons) และ จุดประสานประสาท (synapses) จากแบบจำลองนี้ ข่ายงานประสาทเกิดจากการเชื่อมต่อระหว่างเซลล์ประสาท จนเป็นเครือข่ายที่ทำงานร่วมกัน ส่วนประกอบของโครงข่ายประสาทเทียมจะเป็นดังภาพที่ 2.1



ภาพที่ 2.1 ส่วนประกอบของโครงข่ายประสาทเทียม

ส่วนประกอบของโครงข่ายประสาทเทียมมีรายละเอียดดังต่อไปนี้

2.1.1 Neuron คือ ส่วนที่เป็นรูปวงกลมในภาพที่ 2.1 ข้างใน Neuron จะต่างกันตาม layer ที่มันอยู่ ถ้า Neuron อยู่ในส่วนที่เป็น Input Layer ข้างในตัวมันก็จะมีข้อมูลนำเข้า แต่ถ้า Neuron อยู่ในส่วนที่เป็น Hidden Layer ข้างในตัวมันจะมีสมการที่ช่วยในการคำนวณเพื่อทำนายว่าเป็นคลาสอะไร หรือคำนวณแบบถดถอย (Regression) เป็นต้น

2.1.2 Input Layer (วงกลมสีเขียว) มีหน้าที่ในการรับข้อมูลนำเข้ามาในโครงข่ายประสาทเทียมโดย Input Layer จะมีเพียงชั้นเดียวเท่านั้นและมีหน้าส่งข้อมูลไปยังชั้นถัดไปซึ่งก็คือ Hidden Layer

2.1.3 Hidden Layer (วงกลมสีส้ม) มีหน้าที่รับข้อมูลจาก Input Layer จะสังเกตว่า Hidden Layer สามารถมีจำนวนมากกว่า 1 ชั้นได้ และโดยพื้นฐาน ถ้ายังต้องการความแม่นยำที่มากขึ้น ก็ควรจะเพิ่มจำนวนชั้นของ Hidden Layer และจำนวน Neurons ให้มากขึ้น แต่ไม่เป็นเช่นนั้นเสมอไป โดยทุกๆ Hidden Layer จะมี bias เชื่อมต่ออยู่เพื่อให้ทุกๆ การคำนวณที่มีการส่งต่อมีความเท่าเทียมกัน ดีความหมายง่าย ๆ คือ ทุก ๆ neurons ใน Hidden Layer จะต้องมี Bias เข้าไปคำนวณเพื่อให้ decision boundary ไม่จำเป็นต้องผ่านจุด

origin ส่วน **Weight** จะเป็นน้ำหนักซึ่งมันจะส่งผลให้ทุก ๆ Neurons มีค่า Output ที่ไม่เท่ากัน ทำให้แต่ละคลาสมีน้ำหนักไม่เท่ากัน ทำให้เราสามารถแยกข้อมูลนี้ได้ว่าอยู่ในคลาสใดได้ด้วยการดูผลลัพธ์ที่ Output

2.1.4 Weight และ Bias สองตัวแปรนี้เป็นตัวแปรที่จะเปลี่ยนแปลงค่าตัวเองไปเรื่อย ๆ ระหว่างการเทรนเพื่อให้ Fit กับข้อมูลเรียกว่า Learnable Parameters หรือ Trainable Parameters ซึ่งในทุก ๆ Node ใน Neural Network จะมีตัวแปรสองตัวนี้เพื่อคอยปรับค่าของข้อมูลนำเข้า ซึ่งตอนเริ่มต้นของการเทรน Weight และ Bias จะถูกสุ่มค่ามาก่อน

2.1.5 Output Layer เป็น Layer ที่อยู่ท้ายสุดรอรับค่าจาก Hidden Layer ชั้นสุดท้าย โดยในชั้น Output นั้นแต่ละ Neurons จะมีค่าน้ำหนักของคลาสอยู่เช่น เรามีประเภทของ Output ทั้งหมด 2 รูปแบบ คือ แมว กับ สุนัข เพราะฉะนั้น Output Layer ของเราจะมี Neurons สองตัว โดยที่ Neuron ตัวแรกอาจจะเป็นสุนัข และ Neuron ตัวที่สองอาจจะเป็นแมว โดยเมื่อข้อมูลผ่าน Hidden Layer ไปสู่ Output Layer แล้ว Neurons ทั้งสองตัวจะมีค่าข้างในไม่เท่ากัน

2.2 Backpropagation

Backpropagation คือการ Feed Forward หรือ Forward Pass ซึ่งเป็นกระบวนการแบบไปข้างหน้า หมายความว่าปกติแล้ว โครงข่ายประสาทเทียมจะรับข้อมูลนำเข้าจาก Input Layer สร้างน้ำหนัก (W) ประมวลผลผ่าน Hidden Layer แล้วแสดงผลลัพธ์ที่ Output Layer แต่กระบวนการนี้สามารถย้อนกลับจาก Output Layer กลับมาที่ Input Layer ได้ โดยที่ Backpropagation จะทำหน้าที่ในการปรับค่าน้ำหนัก (W) ในแต่ละเส้นอีกครั้งโดยดูจาก Error หรือ Cost ที่เกิดขึ้นในแต่ละ Neurons

2.3 Loss Function

Loss Function คือ ฟังก์ชันที่กำหนดเป้าหมายให้กับโครงข่ายประสาทเทียม ในการศึกษาครั้งนี้ใช้ Loss Function แบบ Binary Cross Entropy (Log Loss) ซึ่งสามารถหาได้ด้วยสมการต่อไปนี้

$$J(\mathbf{w}) = \frac{1}{N} \sum_{n=1}^N H(p_n, q_n) = - \frac{1}{N} \sum_{n=1}^N \left[y_n \log \hat{y}_n + (1 - y_n) \log(1 - \hat{y}_n) \right]$$

ซึ่ง Binary Cross Entropy คือ ค่าเฉลี่ยของ Cross Entropy ที่เกิดจากการแจกแจงความน่าจะเป็นสองรูปแบบ คือ การแจกแจงความน่าจะเป็นที่อยากได้ (Actual) กับการแจกแจงความน่าจะเป็นที่ถูกประมาณโดยแบบจำลอง (Predicted) ของ **Class 0** และ **Class 1** จากสมการด้านบน การได้ค่าเฉลี่ยน้อยนั้นดีกว่า การได้ค่าเฉลี่ยมาก โดย Binary Cross Entropy Loss จะให้ค่าเฉลี่ยต่ำที่สุด เท่ากับ 0

2.4 ระยะห่างของ Euclidean (Euclidean Distance)

ระยะห่างของ **Euclidean** คือ ระยะทางปกติระหว่างจุดสองจุดในแนวเส้นตรงเดียวกัน มีที่มาจากทฤษฎีบทพีทาโกรัส เหตุที่เรียกว่า แบบยูคลิด เนื่องจากการวัดระยะทางในปริภูมิแบบยูคลิด คือไม่มีความโค้งและไม่สามารถทำให้โค้งงอ และการใช้สูตรนี้การวัดระยะทางทำให้กลายเป็นปริภูมิอิงระยะทาง ค่าประจำ (norm) ที่เกี่ยวข้องก็จะเรียกว่าเป็น **ค่าประจำแบบยูคลิด** (Euclidean norm) ซึ่งสามารถหาค่าระยะห่างของ Euclidean ในปริภูมิ n มิติ ได้จากสมการต่อไปนี้

$$d(\mathbf{p}, \mathbf{q}) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_n - q_n)^2} = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}$$

บทที่ 3 วิธีดำเนินงาน

3.1 โปรแกรมพิเศษที่ใช้

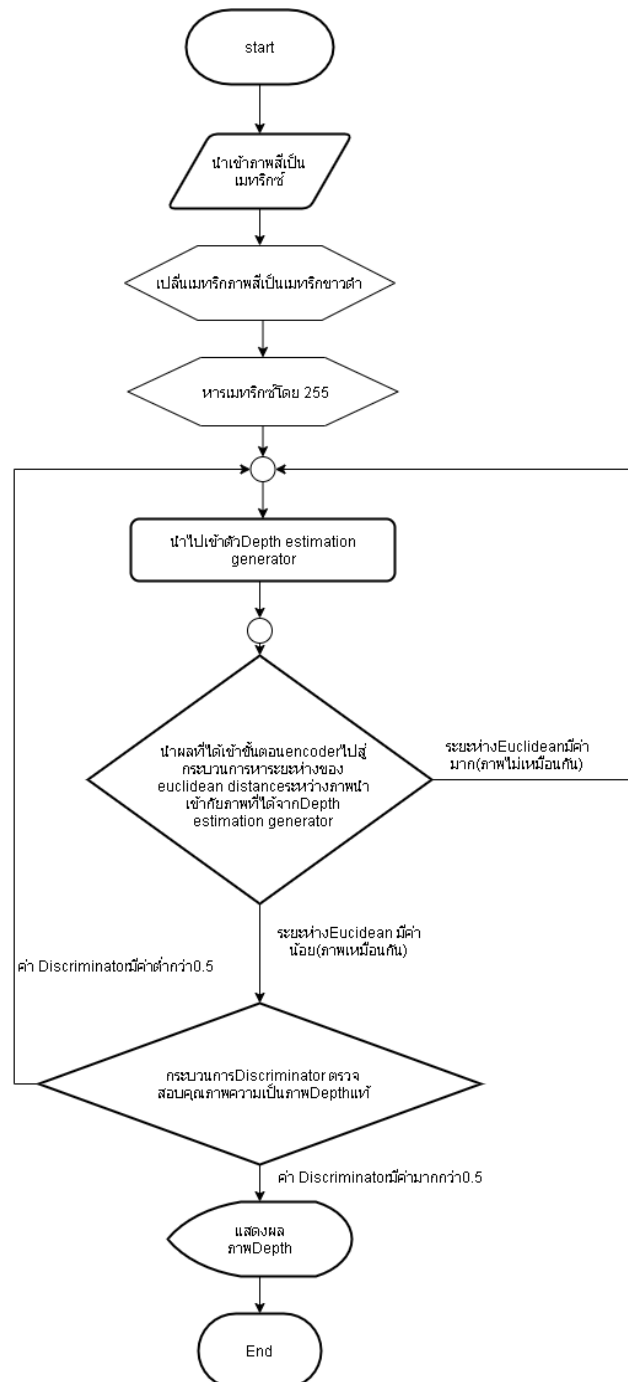
3.1.1 Python

3.1.2 Google Colab

3.1.3 TensorFlow

3.2 วิธีดำเนินการศึกษา

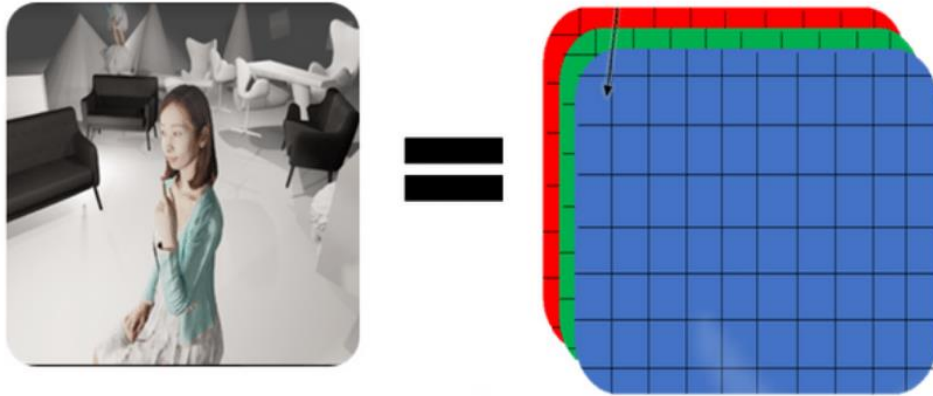
การดำเนินการศึกษาในโครงงานนี้เป็นไปตามแผนผังลำดับงาน ดังภาพที่ 3.1



ภาพที่ 3.1 แผนผังลำดับงานของ Depth estimation model ที่ปรับปรุงจากแบบจำลองของ Anindya

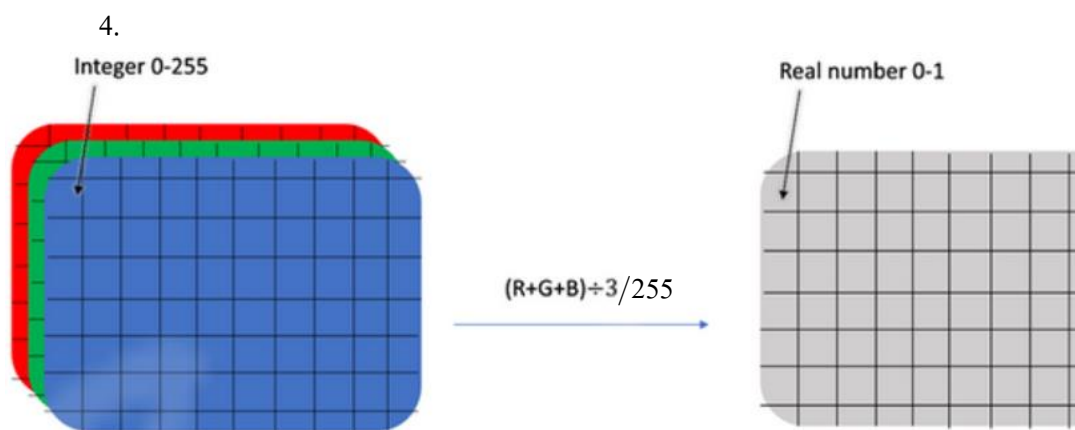
ขั้นตอนที่ 1 เตรียมข้อมูลนำเข้า

1. นำเข้าข้อมูลภาพสีที่มีเมทริกซ์ ขนาด $256 \times 256 \times 3$ เพราะคอมพิวเตอร์เห็นภาพเป็นเมทริกซ์
- 2.



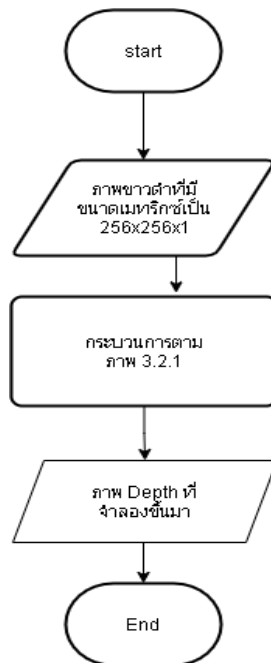
ภาพที่ 3.2 การเตรียมข้อมูลนำเข้าจากภาพสีในระบบ RGB เป็นเมทริกซ์ ขนาด $256 \times 256 \times 3$

3. เปลี่ยนเมทริกซ์ภาพสี ในระบบ RGB ให้เป็นเมทริกซ์ระดับสีเทาโดยการหาค่าเฉลี่ยของเมทริกซ์ แดง (R) เขียว (G) น้ำเงิน (B) จะทำให้เมทริกซ์ของภาพสีในระบบ RGB ที่มีขนาดเท่ากับ $256 \times 256 \times 3$ เป็นเมทริกซ์ของภาพระดับสีเทาที่มีเมทริกซ์ขนาด $256 \times 256 \times 1$ จากนั้นปรับค่าสมาชิกในแต่ละตำแหน่งของเมทริกซ์ขนาด $256 \times 256 \times 1$ ให้มีค่าเป็นจำนวนจริงตั้งแต่ 0 ถึง 1 ด้วยการนำสมาชิกทุกตัวไปหารด้วย 255 เพื่อลดการกระจายของข้อมูล



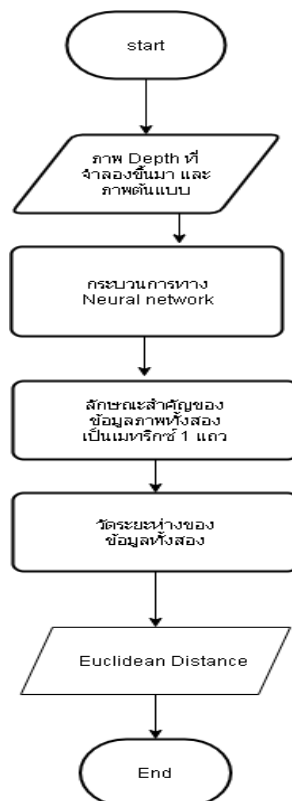
ภาพที่ 3.3 การปรับค่าสมาชิกในแต่ละตำแหน่งของเมทริกซ์ขนาด $256 \times 256 \times 1$

ขั้นตอนที่ 2 นำผลลัพธ์จากขั้นตอนที่ 1 มาเข้าแบบจำลอง Depth image Generator เพื่อสร้างภาพ Depth ซึ่งแบบจำลองนี้จะทำงานตามแผนผังลำดับงานดังภาพที่ 3.4



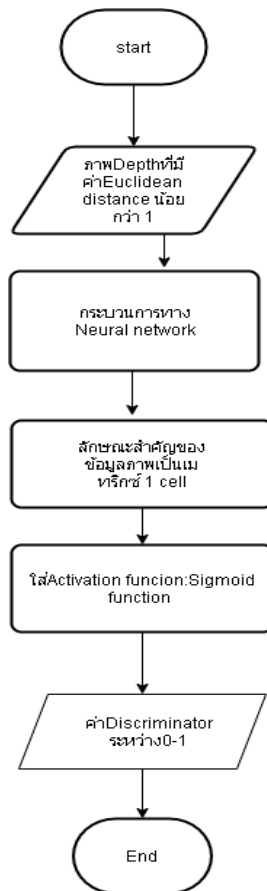
ภาพที่ 3.4 แผนผังลำดับงานของ Depth Image Generator

ขั้นตอนที่ 3 นำผลลัพธ์จากขั้นตอนที่ 2 มาเข้าแบบจำลอง Encoder ซึ่งแบบจำลองนี้จะทำงานตามแผนผังลำดับงานดังภาพที่ 3.5



ภาพที่ 3.5 แผนผังลำดับงานของ Encoder

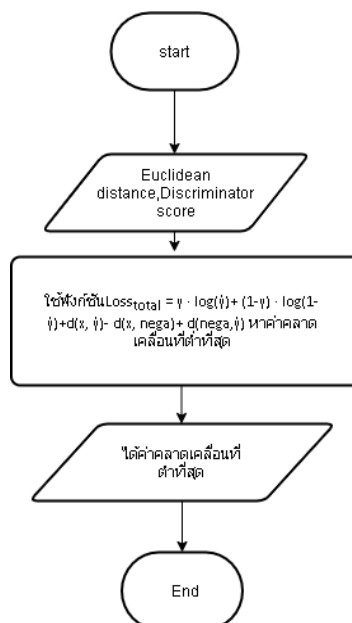
ขั้นตอนที่ 4 นำผลลัพธ์จากขั้นตอนที่ 3 ที่มีค่าระยะห่างของ Euclidean น้อยกว่า 1 มาเข้าแบบจำลอง Discriminator ซึ่งแบบจำลองนี้จะทำงานตามแผนผังลำดับงานดังภาพที่ 3.6



ภาพที่ 3.6 แผนผังลำดับงานของ Discriminator

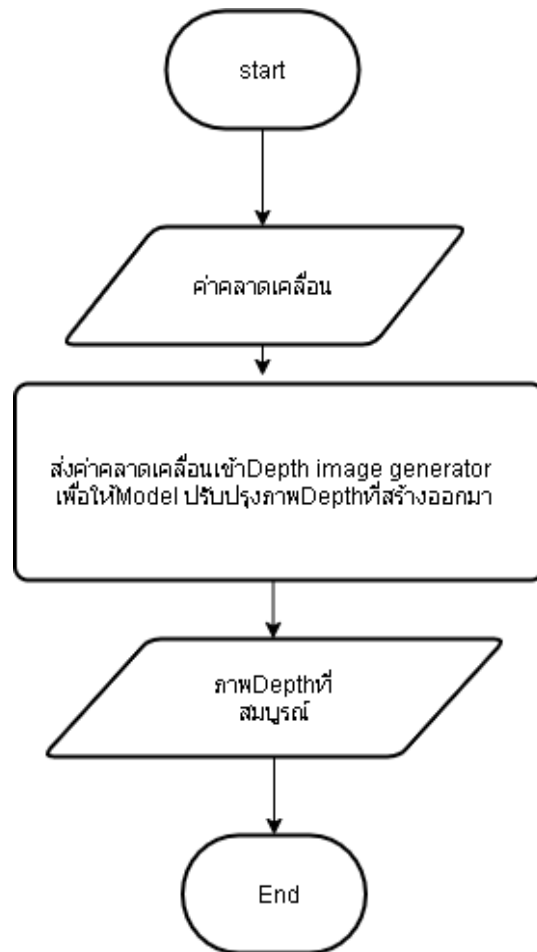
ขั้นตอนที่ 5 ปรับปรุงการทำงานของแบบจำลอง ดังนี้

5.1 หาค่าความคลาดเคลื่อน ซึ่งแบบจำลองจะดำเนินการตามแผนผังลำดับงานดังภาพที่ 3.7



ภาพที่ 3.7 แผนผังลำดับงานการหาค่าคลาดเคลื่อนจากการดำเนินงานตามแบบจำลอง

5.2 เทรนแบบจำลอง Depth Image Generator ซึ่งจะดำเนินการตามแผนผังลำดับงานดังภาพที่ 3.8









ภาพที่ 3.8 แผนผังลำดับงานการเทรนแบบจำลอง Depth Image Generator

บทที่ 4 ผลการศึกษา

จากการศึกษาและเปรียบเทียบประสิทธิภาพของการสร้างภาพ Depth ด้วยกระบวนการทำงานของโครงข่ายประสาทเทียม ได้ผลการศึกษาดังนี้

ตารางที่ 4.1 ตารางเปรียบเทียบผลลัพธ์จากการสร้างภาพ Depth ของแบบจำลองที่ปรับปรุง Depth estimation model ของ Anindya กับแบบจำลองของ Anindya เดิม

		แบบจำลอง ของ Anindya เดิม	แบบจำลองที่ปรับปรุง Depth estimation model ของ Anindya
Input	รูปต้นแบบ รูปแบบ RGB (256 x 256 x 3)		
	รูปต้นแบบ ภาพระดับสีเทา (256 x 256 x 1)		
Output	ภาพ Depth ที่ได้		
	ค่าระยะห่างของ Euclidean	22.00	0.50
	ค่า discriminator score	0.70	0.58

จากตารางที่ 4.1 ค่าระยะห่างของ Euclidean ของภาพ Depth จากแบบจำลองที่ปรับปรุง Depth estimation model ของ Anindya มีค่าน้อยกว่า หมายความว่า ภาพ Depth ที่ได้จาก แบบจำลองนี้มีความเหมือนกับรูปต้นแบบมากกว่า

และจากตารางที่ 4.1 ค่า discriminator score ของภาพ Depth จากแบบจำลองของ Anindya มีค่าใกล้เคียงกับ 1 มากกว่า ทั้งนี้เป็นเพราะแบบจำลองของ Anindya นำภาพ Depth ที่ได้ไปเปรียบเทียบกับภาพ Depth ในชุดข้อมูลที่ใช้ในการเทรน ซึ่งเหตุการณ์นี้เรียกว่า overfitting แต่อย่างไรก็ตามค่า discriminator score ของภาพ Depth จากแบบจำลองที่ปรับปรุง Depth estimation model ของ Anindya ก็มีค่ามากกว่า 0.5 ซึ่งถือว่ามีความเป็นภาพ Depth ค่อนข้างสูง

บทที่ 5 สรุปผลการศึกษา

5.1 สรุปผลการศึกษา

จากการศึกษาได้ข้อสรุปว่าแบบจำลองที่ปรับปรุง Depth estimation model ของ Anindya Basu สามารถลดปริมาณข้อมูลนำเข้าที่ใช้ในการเทรน Depth estimation model ได้ สังเกตได้จากด้วยชุดข้อมูลที่ใช้ในการเทรนแบบจำลอง 100 ชุด เท่ากัน แบบจำลองที่ปรับปรุง Depth estimation model ของ Anindya Basu ให้ภาพ Depth ที่มีความเหมือนกับรูปต้นแบบมากกว่า และมีความเป็นภาพ Depth ด้วยความน่าจะเป็นที่สูงกว่า 0.5

5.2 ข้อเสนอแนะ

ควรหาแหล่งข้อมูลที่รวบรวมภาพ Depth ที่จะใช้ในการเทรนแบบจำลอง ที่มากกว่า 100 ชุดข้อมูล ทั้งนี้เพื่อสามารถหาจำนวนชุดข้อมูลที่น้อยที่สุดในการสร้างภาพ Depth ที่ให้ค่าระยะห่าง Euclidean น้อยที่สุด และมีค่า discriminator score เข้าใกล้ 1 มากที่สุด

บรรณานุกรม

IBM Cloud Education. (2020). **Neural Networks**. สืบค้นจาก <https://www.ibm.com/>
วันที่สืบค้น 20 กรกฎาคม พ.ศ. 2564

KHYATI MAHENDRU. (2019). **A Detailed Guide to 7 Loss Functions for Machine Learning Algorithms with Python Code**. สืบค้นจาก <https://www.analyticsvidhya.com/>
วันที่สืบค้น 20 กรกฎาคม พ.ศ. 2564

Paul Barrett. (2005). **Euclidean Distance raw, normalized, and double-scaled coefficients**.
สืบค้นจาก www.pbarrett.net/techpapers/euclid.pdf
วันที่สืบค้น 20 กรกฎาคม พ.ศ. 2564