

การปรับปรุงประสิทธิภาพการทำงานายคะแนคุณภาพชีวิตด้วยโมเดลแบบลำดับและการกั้นความรู้



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2565

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Improving QoL Score Prediction Performance by Sequence Models and Knowledge  
Distillation



A Thesis Submitted in Partial Fulfillment of the Requirements  
for the Degree of Master of Science in Computer Science  
Department of Computer Engineering  
FACULTY OF ENGINEERING  
Chulalongkorn University  
Academic Year 2022  
Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การปรับปรุงประสิทธิภาพการทำนายคะแนนคุณภาพชีวิต ด้วยโมเดลแบบลำดับและการกลั่นความรู้
โดย	นายธนสิทธิ์ ฤทธิธน์โสภณ
สาขาวิชา	วิทยาศาสตร์คอมพิวเตอร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ศาสตราจารย์บุญเสริม กิจศิริกุล
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม	ผู้ช่วยศาสตราจารย์ ดร.พิตติพล คันธวัฒน์

---

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้หัวข้อวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่ง  
ของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

..... คณะบดีคณะวิศวกรรมศาสตร์  
(ศาสตราจารย์ ดร.สุพจน์ เตชวรสินสกุล)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ  
(รองศาสตราจารย์ ดร.พีรพล เวทีกุล)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก  
(ศาสตราจารย์บุญเสริม กิจศิริกุล)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม  
(ผู้ช่วยศาสตราจารย์ ดร.พิตติพล คันธวัฒน์)

..... กรรมการภายนอกมหาวิทยาลัย  
(ผู้ช่วยศาสตราจารย์ ดร.อนันตพร หารรัชคุณาชัย)

ธนสิทธิ์ ฤทธิธโนโสภณ : การปรับปรุงประสิทธิภาพการทำนายคะแนนคุณภาพชีวิตด้วย  
โมเดลแบบลำดับและการกลั่นความรู้. ( Improving QoL Score Prediction  
Performance by Sequence Models and Knowledge Distillation) อ.ที่ปรึกษา  
หลัก : ศ.บุญเสริม กิจศิริกุล, อ.ที่ปรึกษาร่วม : ผศ. ดร.พิตติพล คັນธวัชนัน

ภายในเมืองใหญ่ความเป็นอยู่อาศัยที่ดีของผู้อยู่อาศัยเป็นสิ่งที่สำคัญที่สุด ซึ่งเป็นเหตุผล  
ว่าทำไมการออกแบบเมืองและการพัฒนาอย่างยั่งยืนจึงเป็นหัวข้อที่สำคัญอย่างมาก คุณภาพชีวิต  
ถูกใช้เป็นตัวชี้วัดความมีประสิทธิภาพของปัจจัยด้านปริมาณและคุณภาพของผังเมืองสำหรับผู้  
อาศัยภายในเมือง อีกทั้งคุณภาพชีวิตของผู้เดินทางเท้าก็เป็นสิ่งสำคัญเช่นกัน แนวคิดด้าน  
ความสามารถในการเดินจะเป็นตัวประเมินและวิเคราะห์คุณภาพชีวิตในฉากการเดิน อย่างไรก็ตาม  
การทำแบบสอบถามเป็นสิ่งที่มีความคุ้มค่าและเวลาที่ต้องใช้สูง รวมถึงข้อจำกัดต่างๆในการประเมิน  
พื้นที่ เพื่อที่จะจัดการกับข้อจำกัดเหล่านี้ งานวิจัยนี้ได้นำเสนอวิธีการที่นำเทคโนโลยี  
ปัญญาประดิษฐ์ (AI) มาใช้ในการช่วยประเมินวัดผลความสามารถในการเดิน โดยที่มีการเก็บข้อมูล  
ผ่านแบบสอบถามด้วยอุปกรณ์ความจริงเสมือน (VR) อีกทั้งงานวิจัยนี้ได้เสนอวิธีการถอดข้อมูล  
และการฝึกสอนแบบจำลองด้วยโครงข่ายประสาทเทียม convolutional เชิงลึก (Deep  
convolutional neural networks : DCNNs) ซึ่งใช้ในการทำนายผลคะแนนคุณภาพชีวิต รวมถึง  
นำวิธีการกลั่นข้อมูลเพื่อช่วยในการลดขนาดและเวลาในการประมวลผลของแบบจำลองลงทำให้  
ประสิทธิภาพของแบบจำลองดีขึ้น จากตัวอย่างผลการทดลองพบว่าวิธีการที่ได้นำเสนอนั้น  
สามารถนำไปใช้งานได้จริงและสามารถใช้เป็นหนึ่งในวิธีการทางเลือกที่สามารถช่วยเหลือในงาน  
ด้านคุณภาพชีวิตได้

จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY

สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์  
ปีการศึกษา 2565

ลายมือชื่อนิสิต .....  
ลายมือชื่อ อ.ที่ปรึกษาหลัก .....  
ลายมือชื่อ อ.ที่ปรึกษาร่วม .....

# # 6270111921 : MAJOR COMPUTER SCIENCE

KEYWORD:

Thanasit Rithanasophon : Improving QoL Score Prediction Performance by Sequence Models and Knowledge Distillation. Advisor: Prof. BOONSERM KIJIRIKUL Co-advisor: Asst. Prof. Dr. PITTIPOL KANTAVAT

The well-being of residents is a top priority for megacities, which is why urban design and sustainable development are crucial topics. Quality of Life (QoL) is used as an effective key performance index (KPI) to measure the efficiency of a city plan's quantity and quality factors. For city dwellers, QoL for pedestrians is also significant. The walkability concept evaluates and analyzes the QoL in a walking scene. However, the traditional questionnaire survey approach is costly, time-consuming, and limited in its evaluation area. To overcome these limitations, this research proposes using artificial intelligence (AI) technology to evaluate walkability data collected through a questionnaire survey using Virtual Reality tools. The proposed method involves knowledge extraction using deep convolutional neural networks (DCNNs) for information extraction and deep learning models to infer QoL scores. Knowledge distillation is also applied to reduce the model size and improve real-time performance. The experiment results demonstrate that the proposed approach is practical and can be considered an alternative method for acquiring QoL.

Field of Study: Computer Science

Academic Year: 2022

Student's Signature .....

Advisor's Signature .....

Co-advisor's Signature .....

## กิตติกรรมประกาศ

ขอขอบพระคุณ ศ. ดร. บุญเสริม กิจศิริกุล อาจารย์ที่ปรึกษาวิทยานิพนธ์ที่รับให้คำปรึกษา และคอยตรวจสอบงานวิจัย อีกทั้งติชมและคอยตรวจสอบแก้ไขวิทยานิพนธ์รวมถึงวารสารเชิงวิชาการในระดับนานาชาติ เพื่อให้เนื้อหาและผลทางการวิจัยมีความครบถ้วนสมบูรณ์

ขอขอบพระคุณ ผศ. ดร. พิตติพล คันธวัฒน์ อาจารย์ที่ปรึกษาร่วมที่คอยช่วยเหลือให้ข้อมูล และคอยช่วยเหลือให้งานวิจัยสามารถสำเร็จลุล่วงได้ตามเป้าหมาย อีกทั้งช่วยให้คำปรึกษา ตรวจสอบ แก้ไข และคอยช่วยเหลือด้านเนื้อหาสำหรับวิทยานิพนธ์รวมถึงวารสารเชิงวิชาการในระดับนานาชาติ

วิทยานิพนธ์นี้สำเร็จลุล่วงได้ด้วยความกรุณาจากอาจารย์ที่ปรึกษาวิทยานิพนธ์ทั้งสองท่านจึงขอขอบพระคุณอาจารย์ทั้งสองท่านเป็นอย่างสูงมา ณ ที่นี้

ขอขอบพระคุณ รศ. ดร. พีรพล เวทีกุล และ ผศ. ดร. อนันตพร หารัชชคุณาฒย์ ที่ให้เกียรติมาเป็นประธานกรรมการ และกรรมการภายนอก สำหรับการสอบวิทยานิพนธ์ในครั้งนี้ พร้อมทั้งให้คำติชม และคำแนะนำซึ่งเป็นประโยชน์ต่อวิทยานิพนธ์อย่างมาก

สุดท้ายนี้ขอขอบคุณครอบครัวที่คอยช่วยเหลือ รวมถึงสนับสนุนค่าใช้จ่ายในการเรียนระดับปริญญาโท และการทำวิทยานิพนธ์ รวมถึงเพื่อน ๆ พี่ ๆ น้อง ๆ ทุกคนที่คอยสนับสนุนและให้กำลังใจเสมอมา จนกระทั่งวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดี



ธนสิทธิ์ ฤทธิธนโสภณ

## สารบัญ

	หน้า
.....	ค
บทคัดย่อภาษาไทย.....	ค
.....	ง
บทคัดย่อภาษาอังกฤษ.....	ง
กิตติกรรมประกาศ.....	จ
สารบัญ.....	ฉ
สารบัญตาราง.....	ฅ
สารบัญรูปภาพ.....	ญ
บทที่ 1 บทนำ .....	1
1.1. ความเป็นมาและความสำคัญของปัญหา.....	1
1.2. วัตถุประสงค์ของงานวิจัย.....	3
1.3. ขอบเขตการวิจัย .....	3
1.4. ประโยชน์ที่คาดว่าจะได้รับ.....	3
บทที่ 2 ทฤษฎีที่เกี่ยวข้อง.....	4
2.1. การรู้จำภาพ (Image Recognition).....	4
2.1.1. การตรวจจับวัตถุ (Object Detection).....	4
2.1.2. การแบ่งส่วนความหมาย (Semantic Segmentation).....	5
2.2. การเรียนรู้เชิงลึก (Deep Learning) .....	5
2.2.1. การเรียนรู้เชิงลึกแบบมีผู้สอน (Deep Supervised Learning).....	5
2.2.2. การเรียนรู้เชิงลึกแบบไม่มีผู้สอน (Deep Unsupervised Learning).....	6
2.2.3. การเรียนรู้เชิงลึกแบบกึ่งมีผู้สอน (Deep Semi-Supervised Learning) .....	6

2.2.4. การเรียนรู้เชิงลึกแบบเสริมกำลัง (Deep Reinforcement Learning) .....	6
2.3. โครงข่ายประสาทเทียมเชิงลึก (Deep Neural Network) .....	6
2.4. โครงข่ายประสาทเทียมแบบวนกลับ (Recurrent Neural Network).....	7
2.4.1. Long Short-Term Memory (LSTM) และ Gated Recurrent Units (GRU).....	8
2.5. โครงข่ายแบบ Bidirectional.....	9
2.6. การกลั่นความรู้ (Knowledge Distillation).....	9
บทที่ 3 งานวิจัยที่เกี่ยวข้อง .....	11
3.1 คุณภาพชีวิตของการเดินเท้า (QoL of Walkability).....	11
3.2. การเรียนรู้เชิงลึก (deep learning) .....	12
3.3. การกลั่นความรู้ (Knowledge Distillation).....	13
บทที่ 4 วิธีการดำเนินงาน .....	14
4.1. ภาพรวมงานวิจัย (Research Overview) .....	14
4.2. ชุดข้อมูลที่ใช้ (Dataset) .....	15
4.3. การถอดคุณสมบัติจากภาพด้วยการรู้จำภาพ (Feature Extraction Using Image Recognition) .....	16
4.3.1. การตรวจจับวัตถุ (Object Detection).....	17
4.3.2. การแบ่งส่วนความหมาย (Semantic Segmentation) .....	18
4.4. ขั้นตอนการเตรียมข้อมูล (Preprocessing).....	20
4.5. การออกแบบสถาปัตยกรรมโครงข่ายประสาทเทียม (Neural Networks Architecture Design) .....	21
4.6. การแบ่งชุดข้อมูลและการฝึกสอนแบบจำลอง (Dataset Splitting and Model Training).....	25
4.7. การประเมินวัดผลแบบจำลอง (Model evaluation).....	27
4.8. วิธีการกลั่นความรู้ (Knowledge Distillation).....	27
บทที่ 5 ผลการทดลอง.....	30



5.1. ผลลัพธ์การฝึกสอนโมเดลด้วยวิธีการเรียนรู้เชิงลึก.....	30
5.2. ผลลัพธ์การฝึกสอนโมเดลด้วยวิธีการกลั่นความรู้.....	33
บทที่ 6 วิเคราะห์ผลการทดลอง .....	47
6.1. ผลการทดลองการฝึกสอนแบบจำลองด้วยวิธีการเรียนรู้เชิงลึก .....	47
6.2. ผลการทดลองการฝึกสอนแบบจำลองด้วยวิธีการกลั่นความรู้.....	47
บทที่ 7 สรุปผลงานวิจัย.....	48
บรรณานุกรม.....	49
ประวัติผู้เขียน .....	53



จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY

## สารบัญตาราง

	หน้า
ตารางที่ 1 องค์ประกอบที่ใช้ในการประเมินผ่านแบบสอบถามทั้ง 15 องค์ประกอบ .....	16
ตารางที่ 2 ประเภทของวัตถุที่ตรวจจับได้จากแบบจำลอง YOLOv4 ที่ใช้ในงานวิจัยนี้.....	17
ตารางที่ 3 ประเภทของวัตถุที่ถูกแบ่งส่วนได้จากแบบจำลอง DDRNet-23-Slim ที่ใช้ในงานวิจัยนี้..	19
ตารางที่ 4 ข้อมูลฝึกสอนและทดสอบของชุดข้อมูลกลุ่มภายในเมืองเดียวกัน .....	26
ตารางที่ 5 ข้อมูลฝึกสอนและทดสอบของชุดข้อมูลกลุ่มข้ามระหว่างเมือง.....	26
ตารางที่ 6 ผลลัพธ์การทดลองของชุดการทดลองภายในเมืองเดียวกัน .....	31
ตารางที่ 7 ผลลัพธ์การทดลองของชุดการทดลองข้ามระหว่างเมือง .....	33
ตารางที่ 8 ค่าผลลัพธ์ MSE ของชุดการทดลองด้วยชุดข้อมูลทดสอบภายในเมืองเดียวกัน .....	35
ตารางที่ 9 เปรียบเทียบค่า MSE ของแต่ละองค์ประกอบที่ใช้ในการประเมินผ่านแบบสอบถาม ของแบบจำลอง teacher และแบบจำลองที่ถูกกั้นความรู้ด้วยชุดข้อมูลทดสอบภายในเมืองเดียวกันที่ใช้ข้อมูลเมืองแคนเบอร์ราฉากที่ 3 เป็นชุดทดสอบ .....	39
ตารางที่ 10 เวลาที่ใช้ในการประมวลผลของแบบจำลองของชุดการทดลองด้วยชุดข้อมูลทดสอบภายในเมืองเดียวกัน .....	40
ตารางที่ 11 ค่าผลลัพธ์ MSE ของชุดการทดลองด้วยชุดข้อมูลทดสอบข้ามระหว่างเมือง .....	41
ตารางที่ 12 เปรียบเทียบค่า MSE ของแต่ละองค์ประกอบที่ใช้ในการประเมินผ่านแบบสอบถาม ของแบบจำลอง teacher และแบบจำลองที่ถูกกั้นความรู้ด้วยชุดข้อมูลทดสอบข้ามระหว่างเมืองที่ใช้ข้อมูลทุกฉากของเมืองแคนเบอร์ราเป็นชุดทดสอบ .....	45
ตารางที่ 13 เวลาที่ใช้ในการประมวลผลของแบบจำลองของชุดการทดลองด้วยชุดข้อมูลทดสอบข้ามระหว่างเมือง .....	46

## สารบัญรูปภาพ

	หน้า
รูปที่ 1 ตัวอย่างการตรวจจับวัตถุ [24].....	4
รูปที่ 2 ตัวอย่างของการแบ่งส่วนความหมาย [25].....	5
รูปที่ 3 เปรียบเทียบการทำงานของโครงข่ายประสาทเทียมกับประสาทของมนุษย์.....	7
รูปที่ 4 โครงข่ายประสาทเทียมแบบวนกลับ .....	7
รูปที่ 5 โครงสร้างของ LSTM (a) และ GRU (b).....	8
รูปที่ 6 การเรียนรู้ด้วยวิธีการกลั่นความรู้.....	9
รูปที่ 7 ภาพรวมของขั้นตอนฝึกสอนแบบจำลองด้วยวิธีการเรียนรู้เชิงลึก.....	14
รูปที่ 8 ภาพรวมขั้นตอนการกลั่นความรู้.....	15
รูปที่ 9 เปรียบเทียบค่า AP และ FPS ของแบบจำลองที่ใช้ในการตรวจจับวัตถุ [31].....	17
รูปที่ 10 ตัวอย่างผลลัพธ์จากแบบจำลอง YOLOv4.....	18
รูปที่ 11 เปรียบเทียบค่า accuracy และ inference speed ของแบบจำลองที่ใช้ในการแบ่งส่วนความหมาย [32].....	19
รูปที่ 12 ตัวอย่างผลลัพธ์จากแบบจำลอง DDRNet-23-Slim .....	20
รูปที่ 13 คุณสมบัติที่ถูกใช้เป็นข้อมูลขาเข้าของการฝึกสอนแบบจำลอง .....	21
รูปที่ 14 ภาพรวมสถาปัตยกรรมโครงข่ายประสาทเทียมเชิงลึกที่ใช้ในการฝึกสอนแบบจำลองทำนายผลคุณภาพชีวิต .....	22
รูปที่ 15 เวกลักษณะของโครงข่ายประสาทเทียม CONV1D.....	23
รูปที่ 16 เวกลักษณะของโครงข่ายประสาทเทียม fully connected.....	23
รูปที่ 17 สถาปัตยกรรมโครงข่ายประสาทเทียมรูปแบบ CONV1D LSTM หมายเลข 5 .....	24
รูปที่ 18 สถาปัตยกรรมโครงข่ายประสาทเทียมรูปแบบ fully connected LSTM หมายเลข 5.....	25
รูปที่ 19 ภาพรวมการฝึกสอนแบบจำลอง.....	27
รูปที่ 20 รูปแบบโครงสร้างสถาปัตยกรรมของแบบจำลองขนาดเล็ก .....	28

รูปที่ 21 ค่า hyperparameter ของโครงสร้างแบบจำลองขนาดเล็ก ..... 28

รูปที่ 22 รหัสเทียมของขั้นตอนการคำนวณหาค่า loss รวมทั้งใช้ในการปรับค่าน้ำหนักของแบบจำลอง  
..... 29

รูปที่ 23 ตัวอย่างของค่า loss ที่เปลี่ยนไปของการฝึกสอนแบบจำลองด้วยชุดข้อมูลภายในเมือง  
เดียวกัน..... 30

รูปที่ 24 ตัวอย่างของค่า loss ที่เปลี่ยนไปของการฝึกสอนแบบจำลองด้วยชุดข้อมูลข้ามระหว่างเมือง  
..... 31

รูปที่ 25 รูปแบบโครงสร้างสถาปัตยกรรมของแบบจำลอง fully connected tanh activation  
function ที่ใช้ร่วมกับโครงข่ายประสาทเทียมแบบย้อนกลับ bi-directional LSTM หมายเลข 1 ... 34

รูปที่ 26 ค่า hyperparameter ของแบบจำลอง fully connected tanh activation function ที่  
ใช้ร่วมกับโครงข่ายประสาทเทียมแบบย้อนกลับ bi-directional LSTM หมายเลข 1 ..... 34

รูปที่ 27 เปรียบเทียบตัวอย่างผลลัพธ์ของแบบจำลอง teacher และแบบจำลองที่ถูกกั้นความรู้ด้วย  
ชุดข้อมูลทดสอบภายในเมืองเดียวกันที่ใช้ข้อมูลเมืองแคนเบอร์ราฉากที่ 3 เป็นชุดทดสอบ ..... 38

รูปที่ 28 เปรียบเทียบตัวอย่างผลลัพธ์ของแบบจำลอง teacher และแบบจำลองที่ถูกกั้นความรู้ด้วย  
ชุดข้อมูลทดสอบข้ามระหว่างเมืองที่ใช้ข้อมูลทุกฉากของเมืองแคนเบอร์ราเป็นชุดทดสอบ..... 44

# บทที่ 1

## บทนำ

### 1.1. ความเป็นมาและความสำคัญของปัญหา

การออกแบบผังเมืองมีความสำคัญอย่างมากต่อการพัฒนาเมืองอย่างยั่งยืนสำหรับเมืองใหญ่ การเติบโตที่รวดเร็วของเมืองใหญ่โดยปราศจากการออกแบบผังเมืองที่ดีส่งผลให้เกิดปัญหามากมายต่อผู้คนที่อยู่อาศัย เช่น ปัญหาด้านการจราจรติดขัด และ อุบัติเหตุจากการขนส่ง [1-5] ด้วยเหตุนี้ทำให้ผู้ที่กำหนดนโยบายควรจะคำนึงถึงความสำคัญของการพัฒนาเมืองอย่างยั่งยืนในการออกแบบผังเมือง โดยที่คำนึงถึงความเป็นมิตรต่อผู้พักอาศัยและสิ่งแวดล้อม

คุณภาพชีวิต (Quality of Life : QoL) คือหนึ่งในดัชนีที่ใช้ในการวัดคุณภาพในสิ่งที่ส่งผลต่อประสิทธิภาพของการดำเนินชีวิตของผู้อยู่อาศัย ผู้กำหนดนโยบายสามารถนำดัชนีนี้ไปใช้ในการชี้วัดประสิทธิภาพของการออกแบบเมืองได้ [6] แนวคิดของคุณภาพชีวิต [7] คือการประเมินชีวิตความเป็นอยู่ที่ดี หรือความสุขของผู้อยู่อาศัย การคำนึงถึงความต้องการส่วนบุคคล ซึ่งรวมถึงการมีปฏิสัมพันธ์ทางสังคมบนพื้นฐานของการพัฒนาแบบยั่งยืน [8-10] จากการศึกษาจำนวนมากผู้คนใช้คุณภาพชีวิตเป็นตัวชี้วัดประสิทธิภาพนโยบายสำหรับการขนส่งสาธารณะในเมืองใหญ่ การวางผังเมืองที่ไม่มีประสิทธิภาพส่งผลต่อความแออัดของการคมนาคม ตัวอย่างเช่น กรุงเทพมหานครที่มีนโยบายเน้นการใช้รถยนต์ส่วนตัวเป็นหลัก [1, 2, 11] ทำให้หลายๆงานวิจัยทางด้านคุณภาพชีวิตใช้กรุงเทพมหานครเป็นกรณีศึกษา เช่น Alonso [12] ซึ่งทำนายนโยบายการใช้พื้นที่ของการขนส่งสาธารณะตามความพึงพอใจของผู้อยู่อาศัย อีกทั้งงานวิจัยของ Hayashi et al. [13] และ Banister [14] ได้นำเสนอว่าประชาชนควรใช้แนวทางคุณภาพชีวิตโดยการส่งเสริมนโยบายการใช้รถยนต์ส่วนตัวให้น้อยลงและส่งเสริมการใช้รถสาธารณะมากยิ่งขึ้น

ผู้ใช้ทางเท้าเป็นปัจจัยที่สำคัญปัจจัยหนึ่งต่อความเป็นอยู่ที่ดีของผู้อยู่อาศัยภายในเมือง งานวิจัยในหลายงานได้นำเสนอแนวคิดความสามารถในการเดินเพื่อนำไปวิเคราะห์และประเมินผลคุณภาพชีวิตจากฉากการเดิน Vichiensan et al. [15] ได้ทำการเปรียบเทียบความต้องการของการเดินภายในกรุงเทพมหานครและเมืองนาโกย่า งานวิจัยนี้เน้นย้ำถึงความสำคัญของการพิจารณาปัจจัยด้านวัฒนธรรมและสิ่งแวดล้อมเมื่อทำการออกแบบพื้นที่เมืองที่ส่งเสริมการเดินและการขนส่งที่ใช้งานได้ Nakamura [16] ได้ทำการศึกษาความสัมพันธ์ระหว่างความสามารถในการเดินกับผลลัพธ์คุณภาพชีวิตของเมืองที่ใช้รถยนต์ส่วนตัวภายในภูมิภาคเอเชีย Vichiensan et al. [17] ได้สรุปว่าการเดินเป็นส่วนสำคัญต่อการเข้าถึงที่เท่าเทียมและความคล่องตัวในการเดินทางภายในเมือง Sou et al. [18] ทำการพัฒนาวิธีการประเมินพื้นที่ของถนนจากการพิจารณาอารมณ์และคุณค่าของมนุษย์ โดยมีเป้าหมาย

เพื่อปรับปรุงการสื่อสารระหว่างนักออกแบบและผู้มีส่วนได้ส่วนเสีย Nakamura [19, 20] ได้เน้นย้ำถึงความสำคัญของพื้นที่ทางเดินเท้าและประสิทธิภาพของการออกแบบ อีกทั้งนำเสนอวิธีการประเมินผลด้วยการใช้เทคโนโลยีความเป็นจริงเสมือน (Virtual Reality : VR) เพื่อเพิ่มประสิทธิภาพของการออกแบบและสร้างพื้นที่ทางเดินเท้าที่มีประสิทธิภาพ

เครื่องมือโดยทั่วไปที่ใช้ในการประเมินวัดผลคุณภาพชีวิตคือการทำแบบสอบถาม ซึ่งจำเป็นต้องใช้เวลาและเงิน โดยเฉพาะอย่างยิ่งในการทำแบบสอบถามจะมีข้อจำกัดทางด้านสถานที่และระยะเวลา และยิ่งไปกว่านั้นข้อจำกัดนี้ยังส่งผลต่อความยากลำบากและซับซ้อนต่อการนำไปปรับใช้กับระบบทางด้านเทคโนโลยีข้อมูลอื่น ๆ อีกด้วย บางงานวิจัยได้นำเสนอวิธีการใหม่โดยการนำเทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence : AI) มาแก้ปัญหาข้อจำกัดเหล่านี้ Kantavat et al. [21] ได้นำโครงข่ายประสาทเทียมเชิงลึกแบบคอนโวลูชัน (Deep Convolutional Neural Networks : DCNNs) มาใช้ซึ่งประกอบด้วย การตรวจจับวัตถุ (Object Detection) และการแบ่งส่วนความหมาย (Semantic Segmentation) เพื่อนำมาถอดปัจจัยสำคัญของความคล่องตัวของการขนส่งจากรูปภาพ Thitisiriwech et al. [22] ได้นำเสนอชุดข้อมูลที่มีชื่อว่า Bangkok Urbanscapes ที่เป็นชุดข้อมูลที่ถูกรับกับ label ชุดแรกในกรุงเทพมหานครและแบบจำลองที่มีประสิทธิภาพที่ดีเยี่ยมในส่วนของกระบวนการแบ่งส่วนความหมาย Thitisiriwech et al. [23] ได้ทำการสำรวจสภาพแวดล้อมที่สร้างขึ้นนั้นส่งผลกระทบต่อคุณภาพชีวิตที่เกี่ยวข้องกับการขนส่งอย่างไร ด้วยการใชระบบสารสนเทศภูมิศาสตร์ร่วมกับเทคโนโลยีการเรียนรู้เชิงลึก (Deep Learning) ภายในเขตสุขุมวิท กรุงเทพมหานคร

ดังนั้นงานวิจัยชิ้นนี้จึงได้นำเสนอวิธีการประเมินวัดผลคุณภาพชีวิตด้วยการใช้วิธีการเรียนรู้ของเครื่อง (Machine Learning) ซึ่งประกอบด้วย 2 ส่วน ได้แก่ การถอดคุณลักษณะจากภาพ และการฝึกสอนแบบจำลอง ซึ่งส่วนแรกจะทำการถอดข้อมูลโดยใช้วิธีการประมวลผลภาพที่ประกอบด้วย การตรวจจับวัตถุและการแบ่งส่วนความหมาย และในส่วนที่สองจะนำข้อมูลที่ได้จากส่วนแรกไปใช้ในการฝึกสอนแบบจำลองเพื่อนำไปใช้ในการทำนายผลคะแนนคุณภาพชีวิต อีกทั้งงานวิจัยนี้ได้นำวิธีการกลั่นความรู้ (Knowledge Distillation) มาใช้ในการลดระยะเวลาในการประมวลผลของการทำนายผลคะแนนคุณภาพชีวิต ซึ่งผลลัพธ์ของการทดลองแสดงให้เห็นว่าระบบที่ได้แนะนำเสนอนี้มีประสิทธิภาพ และสามารถนำไปใช้ในการพิจารณาร่วมกับการประเมินผลคุณภาพชีวิตของฉากเดินได้

ด้วยวิธีการประเมินวัดผลคุณภาพชีวิตนี้มีข้อดีมากกว่าวิธีการแบบดั้งเดิม โดยเฉพาะอย่างยิ่งสามารถแก้ปัญหาด้านเวลาที่ใช้ซึ่งใช้เวลาน้อยกว่าและลดค่าใช้จ่ายได้มากกว่า อีกทั้งยังสามารถนำไปปรับใช้กับกับสถานที่และในช่วงเวลาใดก็ได้ และง่ายต่อการนำไปประยุกต์ใช้กับระบบที่เกี่ยวข้องกับการประเมินวัดผลคุณภาพชีวิตอื่นอีกด้วย

## 1.2. วัตถุประสงค์ของงานวิจัย

เพื่อนำเสนอแนวทางการเพิ่มประสิทธิภาพให้กับระบบทำนายผลคะแนนคุณภาพชีวิต โดยมุ่งเน้นในการเพิ่มความแม่นยำให้กับการทำนายหรือลดความซับซ้อนของโมเดลแต่ยังคงประสิทธิภาพของการทำงาน

## 1.3. ขอบเขตการวิจัย

1.3.1. ข้อมูลที่ใช้ในงานวิจัยนี้ คือ ข้อมูลวิดีโอที่ถูกถ่ายในแต่ละเมือง 4 เมือง เมืองละ 3 ฉาก และข้อมูลจากแบบสอบถามที่ได้จากการให้ผู้เข้าร่วมควิดีโอในแต่ละสถานที่แล้วทำแบบสอบถามโดยให้คะแนนในแต่ละหัวข้อโดยคะแนนอยู่ที่ 0-10

1.3.2. ใช้เทคนิคการถอดคุณสมบัติจากรูปภาพเพื่อใช้ในการฝึกสอน 2 วิธีได้แก่ การแบ่งส่วนความหมาย และการตรวจจับวัตถุ

1.3.3. ออกแบบโครงสร้างแบบจำลองด้วยโครงข่ายประสาทเทียมเชิงลึก

1.3.3. ฝึกสอนแบบจำลองที่ใช้ในการทำนายผลคุณภาพชีวิต

1.3.4. ใช้วิธีการการกลั่นความรู้จากแบบจำลองไปสู่แบบจำลองที่มีความซับซ้อนน้อยกว่า

1.3.5. เปรียบเทียบผลลัพธ์ของขั้นตอนการฝึกสอนและการกลั่นความรู้

## 1.4. ประโยชน์ที่คาดว่าจะได้รับ

1.4.1. สามารถใช้วิธีการแยกคุณสมบัติจากรูปภาพเพื่อนำไปใช้ในการฝึกสอน

1.4.2. สามารถฝึกสอนแบบจำลองเพื่อใช้ในการทำนายคะแนนคุณภาพชีวิตด้วยโครงสร้างจากโครงข่ายประสาทเทียมเชิงลึก

1.4.3. สามารถใช้วิธีการการกลั่นความรู้จากแบบจำลองไปสู่แบบจำลองที่มีความซับซ้อนน้อยกว่า

1.4.4. ได้รับแบบจำลองที่ประมวลผลในระยะเวลาที่น้อยกว่าแบบจำลองเดิม

## บทที่ 2

### ทฤษฎีที่เกี่ยวข้อง

#### 2.1. การรู้จำภาพ (Image Recognition)

การรู้จำภาพ (image Recognition) เป็นการประมวลผลรูปภาพเพื่อที่จะเลียนแบบการทำงานของมนุษย์ที่ใช้สมองในการประมวลผลภาพจากสิ่งที่มองเห็นผ่านดวงตา โดยที่งานวิจัยนี้จะนำวิธีการในเทคโนโลยีนี้มาใช้ 2 วิธีการ ได้แก่ วิธีการตรวจจับวัตถุ (object detection) และ วิธีการแบ่งส่วนความหมาย (semantic segmentation)

##### 2.1.1. การตรวจจับวัตถุ (Object Detection)

การตรวจจับวัตถุ (object detection) เป็นวิธีการระบุวัตถุภายในรูปภาพหรือวิดีโอเพื่อทำการหาตำแหน่งของวัตถุและแยกแยะว่าวัตถุที่ตรวจพบนั้นเป็นอะไร ได้แก่ คน สัตว์ สิ่งของ ซึ่งรูปที่ 1 แสดงให้เห็นตัวอย่างของการตรวจจับวัตถุที่แยกตำแหน่งและประเภทภายใต้การตีกรอบและ label ของสิ่งที่ตรวจพบ



รูปที่ 1 ตัวอย่างการตรวจจับวัตถุ [24]



### 2.1.2. การแบ่งส่วนความหมาย (Semantic Segmentation)

การแบ่งส่วนความหมาย (semantic segmentation) เป็นวิธีการจำแนกวัตถุภายในรูปภาพในระดับ pixel โดยที่จะทำการระบุประเภทในทุกๆ pixel ของรูปภาพว่า ณ จุด pixel นั้นๆ คือส่วนประกอบของวัตถุชนิดไหน เช่น รถยนต์ ต้นไม้ ถนน ทางเดิน เป็นต้น รูปที่ 2 แสดงให้เห็นถึงการแบ่งประเภทโดยที่แต่ละประเภทก็จะใช้สีที่ต่างกันในแต่ละ pixel



รูปที่ 2 ตัวอย่างของการแบ่งส่วนความหมาย [25]

## 2.2. การเรียนรู้เชิงลึก (Deep Learning) มหาวิทยาลัย

การเรียนรู้เชิงลึกสามารถแยกประเภทออกมาได้ดังนี้ Deep supervised learning, Deep semi-supervised learning, Deep unsupervised learning และ Deep reinforcement learning

### 2.2.1. การเรียนรู้เชิงลึกแบบมีผู้สอน (Deep Supervised Learning)

การเรียนรู้เชิงลึกแบบมีผู้สอน (supervised learning) เป็นวิธีการที่ใช้ข้อมูลที่ถูกกำกับประเภทไว้แล้ว โดยที่ข้อมูลคุณลักษณะนั้นสอดคล้องกับประเภทของข้อมูลนั้นๆ ทำให้คอมพิวเตอร์สามารถเรียนรู้หาความสัมพันธ์ของข้อมูลคุณลักษณะและประเภทได้และนำไปสร้างเป็นแบบจำลองเพื่อนำไปใช้กับงานในประเภทเดียวกันต่อได้ ซึ่งสิ่งที่ทำกันเรียนรู้จะถูกเรียกว่า agent อีกทั้ง agent จะเรียนรู้ข้อมูลด้วยค่าความคาดเคลื่อนของข้อมูลที่ทำนายเปรียบเทียบกับข้อมูลจริง ตัวอย่างของการเรียนรู้แบบ Deep supervised learning ได้แก่ Deep Neural Networks (DNN), Convolutional

Neural Networks (CNN), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) และ Gated Recurrent Units (GRU) เป็นต้น [26]

### 2.2.2. การเรียนรู้เชิงลึกแบบไม่มีผู้สอน (Deep Unsupervised Learning)

การเรียนรู้เชิงลึกแบบไม่มีผู้สอน (deep unsupervised learning) เป็นวิธีการที่มีจุดประสงค์หลักในการหาความสัมพันธ์ของข้อมูลและใช้ในการจัดกลุ่มข้อมูลที่มีความใกล้เคียงกันเข้าด้วยกัน อีกทั้งยังสามารถใช้ในงานประเภทลตมิตติของข้อมูลได้อีกด้วย ตัวอย่างของแบบจำลองที่ใช้ในงานประเภทนี้ได้แก่ Auto-Encoders (AE), Restricted Boltzmann Machines (RBM), Generative Adversarial Networks (GAN), Long Short-Term Memory (LSTM) และ Gated Recurrent Units (GRU) เป็นต้น [26]

### 2.2.3. การเรียนรู้เชิงลึกแบบกึ่งมีผู้สอน (Deep Semi-Supervised Learning)

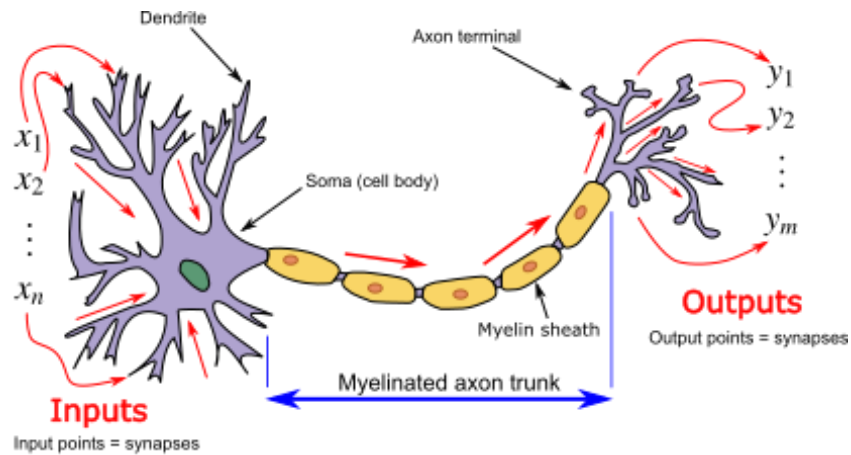
การเรียนรู้เชิงลึกแบบกึ่งมีผู้สอน (deep semi-supervised learning) เป็นวิธีการผสมระหว่างวิธีการเรียนรู้แบบ supervised และ unsupervised ที่ใช้ข้อมูลที่ถูกกำกับประเภทบางส่วนมาให้คอมพิวเตอร์เรียนรู้แบบ supervised ก่อนและหลังจากนั้นจะทำการจัดกลุ่มข้อมูลที่ไม่ถูกกำกับประเภทเข้ากับข้อมูลที่ถูกกำกับประเภทด้วยวิธีการเรียนรู้แบบ unsupervised ตัวอย่างของแบบจำลองที่ใช้ในงานประเภทนี้ได้แก่ Generative Adversarial Networks (GAN), Long Short-Term Memory (LSTM) และ Gated Recurrent Units (GRU) เป็นต้น [26]

### 2.2.4. การเรียนรู้เชิงลึกแบบเสริมกำลัง (Deep Reinforcement Learning)

การเรียนรู้เชิงลึกแบบเสริมกำลัง (deep reinforcement learning) เป็นวิธีการเรียนรู้แบบ reinforcement ซึ่งจะให้ agent นั้นเรียนรู้ด้วยตัวเองเพื่อหาคำตอบโดยปรับปรุงวิธีการจากข้อผิดพลาด ซึ่ง Deep reinforcement learning นั้นจะนำวิธีการเรียนรู้เชิงลึกเข้ามาใช้เพื่อแก้ปัญหาในโจทย์ที่มีความซับซ้อนสูงซึ่งวิธีการ reinforcement แบบทั่วไปไม่สามารถแก้ปัญหาได้ วิธีการนี้เริ่มต้นครั้งแรกในปี ค.ศ. 2013 และใช้ชื่อว่า Google Deep Mind [26]

## 2.3. โครงข่ายประสาทเทียมเชิงลึก (Deep Neural Network)

โครงข่ายประสาทเทียม (artificial neural network) คือสิ่งที่ได้ทำการเรียนแบบมาจากระบบประสาทของมนุษย์ ซึ่งจะรับข้อมูลขาเข้าเพื่อนำไปประมวลผลก่อนส่งไปที่ข้อมูลขาออก ทำให้โครงข่ายประสาทเทียมจะประกอบไปด้วยชั้นของข้อมูลขาเข้า (input layer) 1 ชั้น ชั้นซ่อน (hidden layer) 1 ชั้น และ ชั้นข้อมูลขาออก (output layer) 1 ชั้น ตามรูปที่ 3

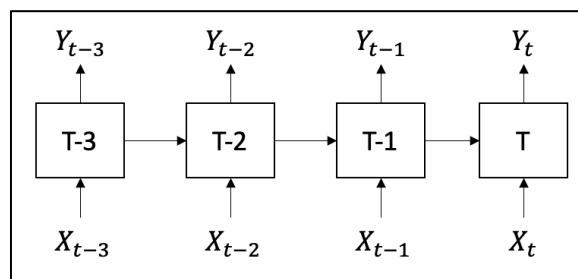


รูปที่ 3 เปรียบเทียบการทำงานของโครงข่ายประสาทเทียมกับประสาทของมนุษย์

โครงข่ายประสาทเทียมเชิงลึก (deep neural network) จะต่างจากโครงข่ายประสาทเทียมแบบทั่วไป โดยที่โครงข่ายประสาทเทียมเชิงลึกจะมีชั้นซ่อนมากกว่าแบบทั่วไปซึ่งช่วยให้แบบจำลองสามารถเรียนรู้ความสัมพันธ์ของข้อมูลที่ซับซ้อนมากยิ่งขึ้น ช่วยแก้ปัญหาที่มีความยากมากกว่าปกติได้

## 2.4. โครงข่ายประสาทเทียมแบบวนกลับ (Recurrent Neural Network)

โครงข่ายประสาทเทียมแบบวนกลับ (recurrent neural network) เป็นหนึ่งในโครงข่ายประสาทเทียมแบบดั้งเดิม โดยส่วนใหญ่จะใช้ข้อมูลแบบลำดับ (sequential data) มาใช้ในการเรียนรู้ ซึ่งตัวอย่างของข้อมูลประเภทนี้ได้แก่ ข้อมูลในรูปแบบข้อความ ข้อมูลในรูปแบบวิดีโอ และข้อมูลในรูปแบบอนุกรมเวลา (time series data) โดยที่ข้อมูลประเภทนี้จะให้ความสำคัญกับลำดับของข้อมูล หรือก็คือข้อมูลก่อนหน้าจะส่งผลต่อข้อมูลในปัจจุบัน ซึ่งตัวอย่างของแบบจำลองประเภทนี้ได้แก่ Long Short-Term Memory (LSTM), and Gated Recurrent Units (GRU) เป็นต้น



รูปที่ 4 โครงข่ายประสาทเทียมแบบวนกลับ

### 2.4.1. Long Short-Term Memory (LSTM) และ Gated Recurrent Units (GRU)

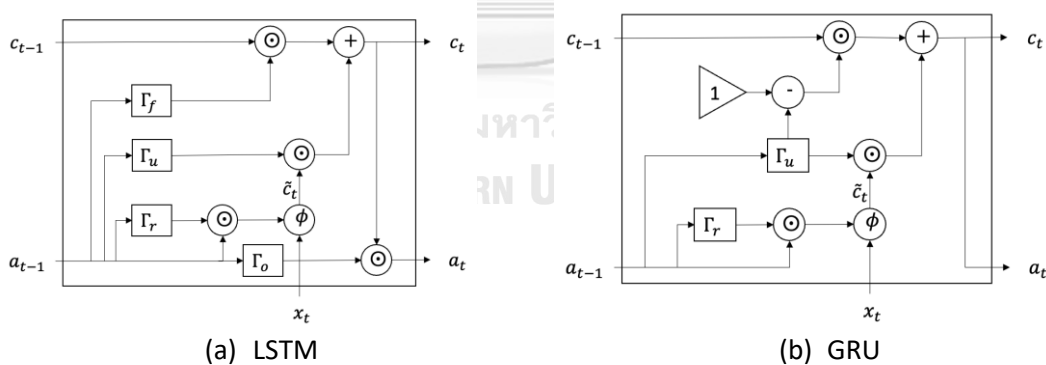
LSTM และ GRU เป็นแบบจำลองที่มีลักษณะเป็นโครงข่ายประสาทเทียมแบบวนกลับที่สามารถแก้ปัญหาด้าน Vanishing gradient ที่สามารถพบเจอได้ในโครงข่ายประสาทเทียมแบบวนกลับแบบทั่วไปได้ [27]

แบบจำลองทั้งสองจะประกอบด้วย gate ซึ่งแต่ละอันจะเป็นไปตามสมการที่ (1) และจะมีทั้งหมด 4 ประเภทได้แก่ update gate, relevance gate, forget gate และ output gate อีกทั้งแต่ละ gate จะมีหน้าที่ที่แตกต่างกัน

- Update gate ( $\Gamma_u$ ) - ทำหน้าที่เป็นตัวให้น้ำหนักความสำคัญของข้อมูลก่อนหน้า
- Relevance gate ( $\Gamma_r$ ) - ทำหน้าที่เป็นตัวตัดสินใจในการลบข้อมูลก่อนหน้า
- Forget gate ( $\Gamma_f$ ) - ทำหน้าที่เป็นตัวตัดสินใจว่าจะใช้ข้อมูลปัจจุบันหรือไม่
- Output gate ( $\Gamma_o$ ) - ทำหน้าที่เป็นตัวตัดสินใจว่าจะส่งข้อมูลปัจจุบันไปต่อหรือไม่

$$\Gamma = \sigma(Wx_t + Ua_{t-1} + b) \quad (1)$$

แบบจำลอง LSTM จะประกอบไปด้วย gate ทั้งสี่ แต่ GRU จะใช้เพียงแค่ update gate และ relevance gate เท่านั้น ซึ่งจะแสดงให้เห็นในรูปที่ 5



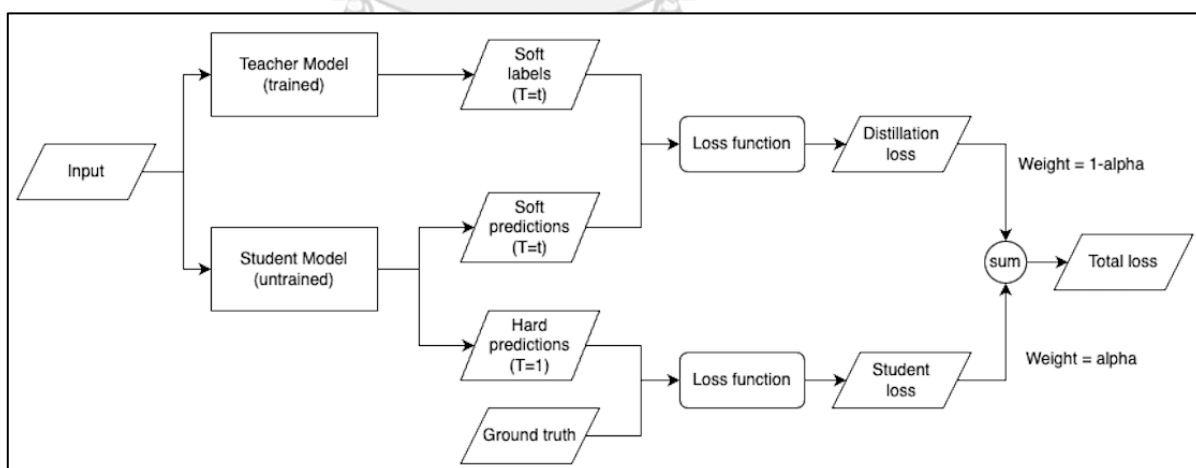
รูปที่ 5 โครงสร้างของ LSTM (a) และ GRU (b)

## 2.5. โครงข่ายแบบ Bidirectional

โครงข่ายแบบ Bidirectional นั้นสามารถนำไปใช้ร่วมกับโครงสร้างแบบจำลองประเภทอื่นๆ ได้ โดยที่เมื่อนำไปใช้จะก่อให้เกิดการเรียนรู้แบบสองทางขึ้น ซึ่งจะนำข้อมูลที่ใช้ในการฝึกสอนมาทำการเรียนรู้แบบเดินหน้าและถอยหลังก่อนนำทั้งมารวมกันเป็นข้อมูลใหม่ที่สามารถนำไปใช้เรียนรู้ต่อที่โครงสร้างชั้นอื่นได้ อีกทั้งวิธีการใช้นั้นไม่ได้ตายตัวและสามารถตั้งค่าการเรียนรู้ได้หลายแบบ ทำให้โครงข่ายนี้มีประโยชน์อย่างมากกับงานหลายประเภท เช่น ระบบแปลภาษา

## 2.6. การกลั่นความรู้ (Knowledge Distillation)

การกลั่นความรู้ (knowledge distillation) [28] เป็นวิธีการที่สามารถลดความซับซ้อนของแบบจำลองได้อีกทั้งยังช่วยแก้ปัญหาเรื่องระยะเวลาของการประมวลผลได้อีกด้วย แต่แลกมากับความแม่นยำที่ลดลง โดยการนำแบบจำลองที่เรียนรู้แล้วซึ่งมีความซับซ้อนสูงมาช่วยในการฝึกสอนให้กับแบบจำลองที่มีความซับซ้อนน้อยกว่า ดังนั้นแบบจำลองที่เรียนรู้แล้วจะถูกเรียกว่า แบบจำลอง teacher และ แบบจำลองที่มีขนาดเล็กกว่าจะถูกเรียกว่าแบบจำลอง student ซึ่งข้อดีของการเรียนรู้ด้วยวิธีนี้จะนำข้อดีของทั้งสองแบบจำลองมารวมกันนั่นก็คือ ความแม่นยำของแบบจำลอง teacher และ ความซับซ้อนต่ำความเร็วในการประมวลผลสูงของแบบจำลอง student โดยที่ผลลัพธ์ความแม่นยำของแบบจำลองที่ใช้วิธีการนี้จะมีค่าสูงกว่าแบบจำลอง student ที่เรียนรู้ด้วยตัวเอง ซึ่งลำดับวิธีการจะเป็นไปตามรูปที่ 6



รูปที่ 6 การเรียนรู้ด้วยวิธีการกลั่นความรู้

วิธีการกลั่นความรู้นั้นขั้นตอนแรกจะทำการฝึกสอนแบบจำลองด้วยโครงสร้างที่ซับซ้อนก่อน เพื่อให้ได้ค่าความแม่นยำที่สูง หลังจากที่ได้แบบจำลองมาแล้วจะกำหนดให้แบบจำลองนี้เป็นแบบจำลอง teacher หลังจากนั้นจะทำการเลือกโครงสร้างที่มีความซับซ้อนน้อยกว่าและมีความเร็วในการประมวลผลสูงมาใช้เป็นแบบจำลอง student เมื่อได้แบบจำลองทั้งสองแล้วจะให้แบบจำลอง teacher ทำนายผลจากข้อมูลออกมาเป็น soft label เพื่อใช้ในการเปรียบเทียบกับผลทำนายของแบบจำลอง student ซึ่งทั้งคู่จะถูกคำนวณร่วมกับค่า  $temperature (T)$  เพื่อนำไปเปรียบเทียบหาความคลาดเคลื่อน และนำผลทำนายที่ไม่ถูกคำนวณด้วยค่า  $temperature$  ของแบบจำลอง student ไปเปรียบเทียบกับผลลัพธ์จริงเพื่อหาความคลาดเคลื่อน เมื่อได้ค่าความคลาดเคลื่อนทั้งสองแล้วจะนำมารวมกันโดยที่มีค่า  $alpha$  ในการกำหนดน้ำหนักของความคลาดเคลื่อนของแต่ละอันก่อนนำมา รวมกันเพื่อนำค่าที่ได้ไปใช้ในการปรับค่าน้ำหนักของแบบจำลองสุดท้าย



## บทที่ 3

### งานวิจัยที่เกี่ยวข้อง

ระบบทำนายผลคะแนนคุณภาพชีวิตด้วยข้อมูลจากทางเดินร่วมกับโครงสร้างประสาทเทียมเชิงลึกและเพิ่มประสิทธิภาพด้วยการกลั่นความรู้ เป็นงานวิจัยที่มีความเกี่ยวข้องกับวิธีการ 3 วิธีการ ได้แก่ 1. คุณภาพชีวิตของการเดินเท้า (QoL of walkability) 2. การเรียนรู้เชิงลึก (deep learning) 3. การกลั่นความรู้ (knowledge distillation) อีกทั้งในหลายงานวิจัยทางด้านคุณภาพชีวิตสำหรับฉากการเดินได้แนะนำปัจจัยที่ส่งผลต่อการเดินเท้า

#### 3.1 คุณภาพชีวิตของการเดินเท้า (QoL of Walkability)

Vichiensan et al. [15] ได้ทำการเปรียบเทียบความต้องการในการเดินภายในกรุงเทพมหานครและเมืองนาโกย่า ซึ่งพบว่าความสะดวกสบายและความสุขเป็นสิ่งที่มีความต้องการอย่างมากของทั้ง 2 เมืองนี้ และความปลอดภัยเป็นอีกหนึ่งสิ่งที่มีความต้องการมากที่สุดสำหรับกรุงเทพมหานครเนื่องจากสภาพของถนน งานวิจัยนี้ได้แนะนำให้มีการปรับปรุงคุณภาพของถนนเพื่อส่งเสริมการเดินในกรุงเทพมหานครและทำการนำกิจกรรมข้างถนนที่หายไปของนาโกย่ากลับมา รวมถึงได้ทำการศึกษากิจกรรมที่ไม่เป็นทางการของการประเมินความสามารถในการเดินที่ทำให้ได้ข้อมูลเชิงลึกเกี่ยวกับการออกแบบถนนที่ใช้งานได้จริงสำหรับเมืองที่กำลังเติบโตในเอเชีย

Nakamura [16] ได้ทำการสำรวจความสัมพันธ์ระหว่างความสามารถในการเดินและผลลัพธ์คุณภาพชีวิต ซึ่งทำการสำรวจเมืองที่นิยมใช้รถส่วนตัวในเอเชีย งานวิจัยนี้ได้สร้างตัวบ่งชี้ของความสามารถในการเดินและทำการสอบถามผ่านแบบสอบถามกับผลเมือง 500 คนในเมืองนาโกย่า ประเทศญี่ปุ่น ผลลัพธ์แสดงให้เห็นถึงคุณภาพของถนนชุมชน โดยเฉพาะอย่างยิ่งสามารถสร้างความสุขที่ส่งผลต่อผลคุณภาพชีวิตจากการมีปฏิสัมพันธ์ ทำให้เห็นว่าคุณภาพถนนส่งผลอย่างยิ่งต่อการวางแผนการขนส่งทางบกเพื่อเพิ่มผลลัพธ์คุณภาพชีวิต

Vichiensan et al. [17] ได้ทำการสรุปผลว่าการเดินนั้นเป็นสิ่งสำคัญต่อระบบขนส่งแบบยั่งยืน เนื่องจากสามารถสร้างความเท่าเทียมต่อการเข้าถึงและความคล่องตัวในการเดินทางภายในเมือง อีกทั้งยังส่งผลต่อความสุข และความสะดวก รวมถึงช่วยประหยัดค่าใช้จ่าย ซึ่งทางเดินและถนนควรที่จะต้องมีความสวยงาม มีชีวิตชีวา ปลอดภัย และ มีการป้องกันจากการรถที่สัญจรบนถนนที่ดี

รวมถึงมีหลากหลายงานวิจัยที่ทำการศึกษาวិธีการและนำเทคโนโลยีที่ทันสมัยมาใช้ในการงานลักษณะนี้เช่น เทคโนโลยีปัญญาประดิษฐ์ และเทคโนโลยีความจริงเสมือน (VR) สำหรับงานด้านคุณภาพชีวิตและความสามารถในการเดิน

Sou et al. [18] ได้นำเสนอวิธีการ ปัญญาประดิษฐ์และวิธีการประเมินความร่วมมือของมนุษย์ (AI and human co-operative evaluation : AIHCE) ซึ่งอำนวยความสะดวกในการออกแบบการสื่อสารระหว่างนักออกแบบและผู้มีส่วนได้ส่วนเสียตามอารมณ์และค่านิยมของมนุษย์ สำหรับการประเมินพื้นที่ถนน งานวิจัยนี้ได้สรุปผลว่า วิธีการนี้สามารถช่วยส่งเสริมให้ผู้คนตระหนักว่าถนนเป็นของสาธารณะที่สะท้อนถึงหน้าที่สำคัญของพื้นที่สาธารณะและค่านิยมของผู้อยู่อาศัย รวมถึงลักษณะเฉพาะของภูมิภาค ซึ่งเพิ่มความยั่งยืนของเมือง

Nakamura [19, 20] ได้กล่าวถึงปัญหาของพื้นที่ทางเดินเท้าที่ต้องเน้นที่ฟังก์ชันแต่ละส่วนและประสิทธิภาพในการออกแบบขององค์ประกอบ งานวิจัยนี้นำเทคโนโลยีความจริงเสมือน (virtual reality : VR) มาใช้ในการประเมินผลวิเคราะห์ผลกระทบของการออกแบบแนวเขตของพื้นที่ทางเดินเท้า ผลสำรวจแสดงให้เห็นถึงความสัมพันธ์ที่ใกล้ชิดระหว่างการออกแบบแนวเขตของพื้นที่และความต้องการของทางเดินเท้า อีกทั้งงานวิจัยนี้ได้สำรวจความสัมพันธ์ระหว่างสภาพแวดล้อมของถนน การรับรู้และพฤติกรรมทางเดิน ซึ่งพบว่าวิธีการนี้สะท้อนให้เห็นความไวของการเดินที่ตอบสนองความต้องการในการเดิน

### 3.2. การเรียนรู้เชิงลึก (deep learning)

Thitisiriwech et al. [23] พบว่าวิธีการประเมินวัดผลคุณภาพชีวิตแบบดั้งเดิมนั้นใช้เวลาและค่าใช้จ่ายเป็นจำนวนมาก งานวิจัยนี้จึงได้นำเสนอวิธีการซึ่งนำเทคโนโลยีปัญญาประดิษฐ์ โดยใช้โครงข่ายประสาทเทียม deep convolutional neural networks (DCNNs) และ Linear regression ในการทำนายผลคะแนนคุณภาพชีวิตจากรูปภาพถนนขณะขับรถ รวมถึงใช้แบบจำลอง DeepLab-v3+ และ YOLO-v3 มาใช้ในการถอดคุณลักษณะจากภาพก่อนนำไปฝึกสอนด้วยโครงข่ายประสาทเทียม ผลลัพธ์เผยให้เห็นข้อมูลเชิงลึกว่าอะไรบ้างที่ทำให้คนกรุงเทพมหานครมีความสุขหรืออึดอัด เช่น ขนาดถนนที่กว้างและพื้นที่ของทางเดินเท้าซึ่งส่งผลต่อความสุขและความปลอดภัยที่มากขึ้น อีกทั้งการจราจรที่ติดขัดส่งผลต่อความสุขของผู้ขับขี่

Iamtrakul et al. [29] งานวิจัยนี้ได้ใช้ GIS และ การเรียนรู้เชิงลึกในการวิเคราะห์ผลกระทบของสภาพแวดล้อมที่สร้างขึ้นต่อคุณภาพชีวิตสำหรับการขนส่งภายในซอยสุขุมวิท กรุงเทพมหานคร ซึ่งพบว่าแต่ละคนรับรู้ถึงคุณภาพชีวิตที่แตกต่างกันขึ้นอยู่กับลักษณะทางกายภาพของสภาพแวดล้อมของแต่ละคน งานวิจัยเน้นย้ำความสำคัญของคุณภาพชีวิตสำหรับการขนส่งของการพัฒนาผังเมืองและการขนส่งเพื่อมุ่งสู่อนาคตที่ยั่งยืน



### 3.3. การกลั่นความรู้ (Knowledge Distillation)

Hayato, Futami et al. [30] งานวิจัยนี้เป็นงานวิจัยที่เกี่ยวกับการฝึกสอนแบบจำลองด้าน automatic speech recognition (ASR) หรือการตรวจจับคำพูดอัตโนมัติ โดยการใช้วิธีการกลั่นความรู้ (knowledge distillation) โดยวิธีการของงานวิจัยนี้คือการนำ แบบจำลองที่ฝึกสอนแล้ว (pre-trained model) ที่มีชื่อว่า Bidirectional Encoder Representations from Transformers (BERT) ซึ่งเป็นแบบจำลองสำหรับงานทางด้าน Natural language processing โดยวิธีการคือ นำข้อมูล input ผ่านแบบจำลอง BERT เพื่อให้แบบจำลองสร้าง soft labels เพื่อนำไปฝึกสอนใหม่ให้กับแบบจำลองแบบลำดับ (sequence model) ซึ่งแบบจำลองแบบลำดับและวัดผลด้วย Corpus of Spontaneous Japanese (CSJ) พบว่าได้ประสิทธิภาพที่สูงกว่าแบบจำลอง TrfLM(uni)



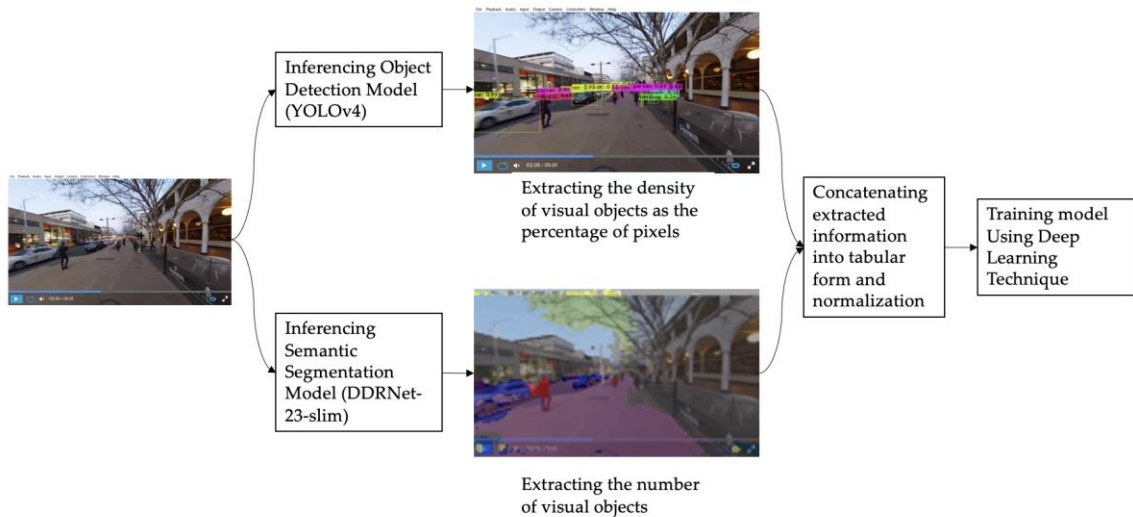
## บทที่ 4

### วิธีการดำเนินงาน

ในส่วนของวิธีการดำเนินงานจะอธิบายขั้นตอนต่างๆของการทำงานโดยจะแบ่งเป็น 2 ส่วน คือ การฝึกสอนแบบจำลองด้วยวิธีการเรียนรู้เชิงลึก และ การเพิ่มประสิทธิภาพแบบจำลองด้วยวิธีการกลั่นความรู้ ซึ่งถูกอธิบายในแต่ละหัวข้อดังนี้

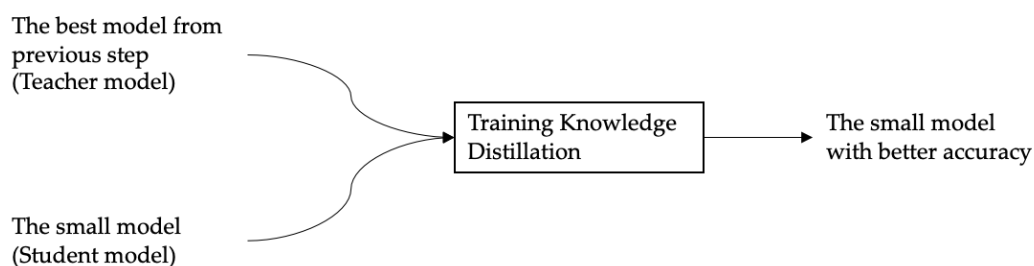
#### 4.1. ภาพรวมงานวิจัย (Research Overview)

ภาพรวมของงานวิจัยในส่วนแรกจะอธิบายถึงชุดข้อมูลที่นำมาใช้ การถอดคุณสมบัติจากภาพ ด้วยการรับรู้ภาพ การออกแบบสถาปัตยกรรมโครงข่ายประสาทเทียม ขั้นตอนการเตรียมข้อมูล การออกแบบสถาปัตยกรรมโครงข่ายประสาทเทียม การฝึกสอนแบบจำลอง การประเมินวัดผลแบบจำลอง วิธีการกลั่นความรู้ โดยรูปที่ 7 จะแสดงให้เห็นถึงขั้นตอนการประมวลผลข้อมูลและนำไปฝึกสอนด้วยวิธีการเรียนรู้เชิงลึก



รูปที่ 7 ภาพรวมของขั้นตอนฝึกสอนแบบจำลองด้วยวิธีการเรียนรู้เชิงลึก

ในส่วนที่สองจะเป็นขั้นตอนการเพิ่มประสิทธิภาพของแบบจำลองจากส่วนแรก โดยนำวิธีการกลั่นความรู้มาใช้ในส่วนนี้ ซึ่งในรูปที่ 8 คือภาพรวมของขั้นตอนในส่วนที่สอง



รูปที่ 8 ภาพรวมขั้นตอนการกลั่นความรู้

#### 4.2. ชุดข้อมูลที่ใช้ (Dataset)

ข้อมูลที่ใช้ในงานวิจัยนี้ได้นำวิดีโอที่ได้ทำการบันทึกของฉากเดินของเมือง 4 เมือง ได้แก่ เมืองแคนเบอร์รา (Canberra) เมืองบริสเบน (Brisbane) กรุงเทพมหานคร (Bangkok) และ เมืองซากาเอะ (Sakae) ซึ่งในแต่ละเมืองจะประกอบไปด้วยวิดีโอทั้งหมด 3 ฉากต่อเมือง มีความยาวของวิดีโอฉากละ 1 นาที เท่ากับวิดีโอที่ใช้มีทั้งหมด 12 ฉาก โดยที่วิดีโอถูกบันทึกด้วยกล้อง RICOH Theta S ที่ 24 fps สำหรับเมืองแคนเบอร์รา บริสเบน กรุงเทพมหานคร และที่ 30 fps สำหรับเมืองซากาเอะ

ในส่วนของคุณข้อมูลเป้าหมายจะนำมาจากข้อมูลแบบสอบถามจากผู้เข้าร่วมทั้งหมด 50 คน โดยให้ผู้เข้าร่วมใส่อุปกรณ์ VR Google เพื่อให้ดูวิดีโอก่อนหน้าและทำการให้คะแนนองค์ประกอบ 15 องค์ประกอบ และคะแนนจะอยู่ในช่วง 1-10 องค์ประกอบทั้งหมดมีดังนี้ อาคาร สิ่งอำนวยความสะดวก สะดวก หลังคา ม้านั่ง แผงลอย ผู้คนที่เดินหรือนั่งอยู่ ความกว้างของทางเท้า คนเดินเท้า ความสว่าง จุดจอดจักรยาน ต้นไม้ข้างทาง เสาไฟฟ้า ที่จอดรถ ทางม้าลาย เสียงของการจราจร ซึ่งถูกแสดงไว้ในตารางที่ 1

ตารางที่ 1 องค์ประกอบที่ใช้ในการประเมินผ่านแบบสอบถามทั้ง 15 องค์ประกอบ

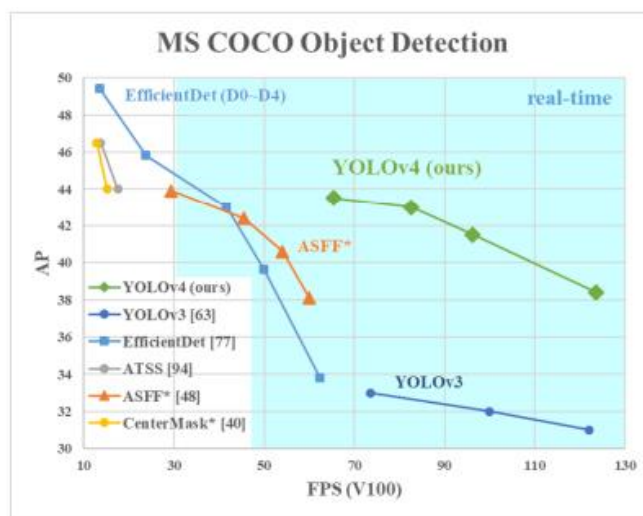
Building (อาคาร)	อาคาร (Building)
	สิ่งอำนวยความสะดวก (Facility)
	หลังคา (Roof)
Activity (กิจกรรม)	ม้านั่ง (Bench)
	แผงลอย (Stall)
	ผู้คนที่เดินหรือนั่งอยู่ (Walking or Sitting People)
Footpath (ทางเท้า)	ความกว้างของทางเท้า (Footpath Width)
	คนเดินเท้า (Pedestrian)
	ความสว่าง (Brightness)
Installation (การติดตั้ง)	จุดจอดจักรยาน (Bicycle Parking)
	ต้นไม้ข้างทาง (Roadside Tree)
	เสาไฟฟ้า (Electric Pole)
Roadway (ถนน)	ที่จอดรถ (Parking)
	ทางม้าลาย (Zebra Crossing)
	เสียงของการจราจร (Traffic Volume)

#### 4.3. การถอดคุณสมบัติจากภาพด้วยการรู้จำภาพ (Feature Extraction Using Image Recognition)

ในส่วนนี้จะทำการอธิบายวิธีการถอดคุณสมบัติจากภาพเพื่อใช้ในการดึงค่าคุณสมบัติจากภาพออกมาใช้เป็นข้อมูลขาเข้าของการฝึกสอนแบบจำลอง โดยขั้นตอนนี้จะประกอบไปด้วย 2 วิธีการ คือ การตรวจจับวัตถุ (object detection) และการแบ่งส่วนความหมาย (semantic segmentation) เนื่องจากผลลัพธ์ของงานวิจัย Thitisiriwech et al. [23] พบว่าการใช้วิธีการตรวจจับวัตถุร่วมกับวิธีการแบ่งส่วนความหมายให้ผลลัพธ์ความแม่นยำที่ดีกว่า หลังจากนั้นจะนำข้อมูลทั้งหมดที่ได้มารวมกันเพื่อเข้าสู่วิธีการถัดไป

#### 4.3.1. การตรวจจับวัตถุ (Object Detection)

ในขั้นตอนการตรวจจับวัตถุ งานวิจัยนี้ได้มีการนำแบบจำลองที่มีชื่อว่า YOLO (You Only Look Once) ซึ่งมีความสามารถในการตรวจจับวัตถุโดยข้อมูลที่ได้ออกมาเป็นรายจำนวน ประกอบด้วยประเภทวัตถุ และ ตำแหน่ง โดยรุ่นของแบบจำลองที่ใช้คือ YOLOv4 เนื่องจากมีค่า average precision ที่มากกว่า YOLOv3 โดยที่ค่า frame per second อยู่ในระดับที่ใกล้เคียงกันซึ่งอ้างอิงจากรูปที่ 9 อีกทั้งประเภทของวัตถุที่สามารถตรวจจับได้จะมีทั้งหมด 17 ประเภทซึ่งถูกแสดงให้เห็นตามตารางที่ 2

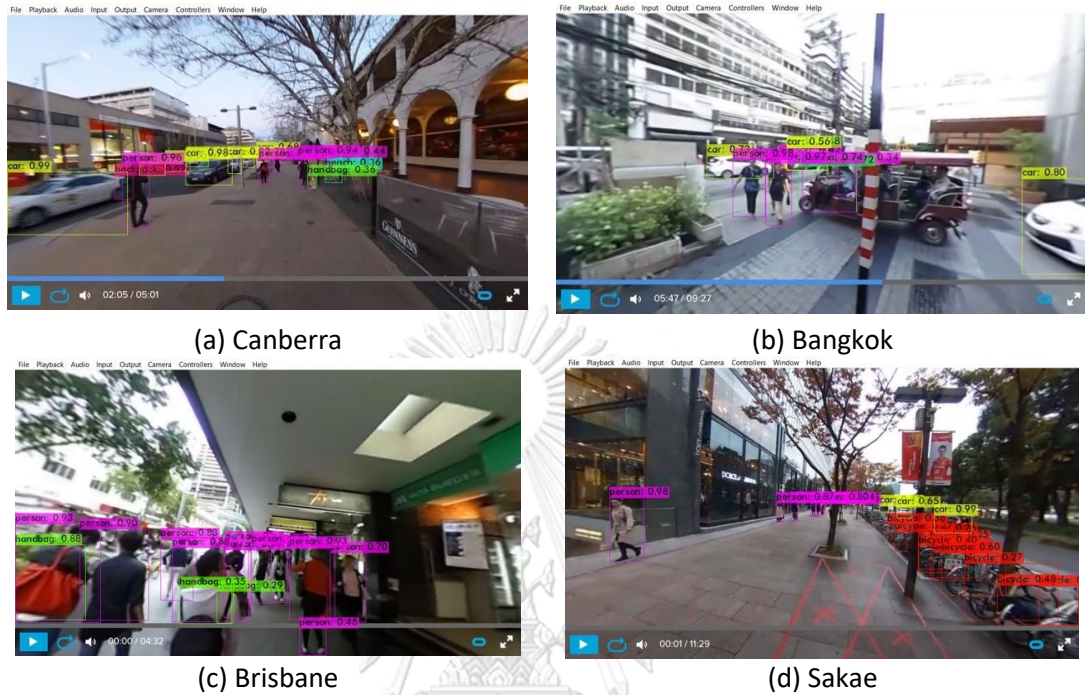


รูปที่ 9 เปรียบเทียบค่า AP และ FPS ของแบบจำลองที่ใช้ในการตรวจจับวัตถุ [31]

ตารางที่ 2 ประเภทของวัตถุที่ตรวจจับได้จากแบบจำลอง YOLOv4 ที่ใช้ในงานวิจัยนี้

1. บุคคล (Person)	10. ไฟจราจร (Traffic light)
2. รถจักรยาน (Bicycle)	11. หัวดับเพลิง (Fire hydrant)
3. รถยนต์ (Car)	12. ป้ายหยุด (Stop sign)
4. รถจักรยานยนต์ (Motorbike)	13. เสามิเตอร์จอดรถ (Parking meter)
5. เครื่องบิน (Aeroplane)	14. ม้านั่ง (Bench)
6. รถประจำทาง (Bus)	15. กระเป๋าเป้ (Backpack)
7. รถไฟ (Train)	16. ร่ม (Umbrella)
8. รถบรรทุก (Truck)	17. กระเป๋าถือ (Handbag)
9. เรือ (Boat)	

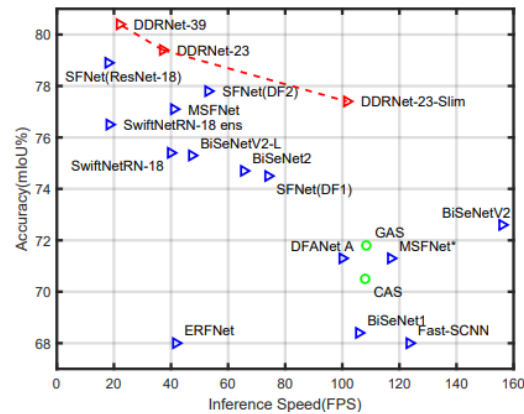
ผลลัพธ์จากแบบจำลอง YOLOv4 จะได้ออกมาเป็นประเภทและตำแหน่ง โดยรูปที่ 10 แสดงให้เห็นถึงวัตถุที่ตรวจจับได้และแสดงผลออกมาเป็นกรอบพร้อมด้วยชื่อของประเภทวัตถุ



รูปที่ 10 ตัวอย่างผลลัพธ์จากแบบจำลอง YOLOv4

#### 4.3.2. การแบ่งส่วนความหมาย (Semantic Segmentation)

ในส่วนของการแบ่งส่วนความหมาย งานวิจัยนี้ได้ใช้แบบจำลอง DDRNet-23-Slim เนื่องจากมีค่า accuracy อยู่ที่ 77.4% mIoU และ inference speed อยู่ที่ 102 FPS อ้างอิงจากรูปที่ 11 ซึ่งแบบจำลองนี้มีความสามารถในการระบุประเภทของวัตถุในระดับ pixel และผลลัพธ์ที่ได้จะอยู่ในรูปแบบเปอร์เซ็นต์ ดังนั้นผลรวมทั้งหมดจะเท่ากับ 100 โดยที่ประเภทของวัตถุที่ใช้ในงานวิจัยนี้มีทั้งหมด 19 ประเภทและแบ่งตามสีที่แสดงให้เห็นในตารางที่ 3



รูปที่ 11 เปรียบเทียบค่า accuracy และ inference speed ของแบบจำลองที่ใช้ในการแบ่งส่วน  
 ความหมาย [32]

ตารางที่ 3 ประเภทของวัตถุที่ถูกแบ่งส่วนได้จากแบบจำลอง DDRNet-23-Slim ที่ใช้ในงานวิจัยนี้

1. ถนน (Road)	
2. ทางเท้า (Sidewalk)	
3. อาคาร (Building)	
4. กำแพง (Wall)	
5. รั้ว (Fence)	
6. เสา (Pole)	
7. ไฟจราจร (Traffic light)	
8. ป้ายจราจร (Traffic sign)	
9. พืชพรรณ (Vegetation)	
10. พื้นดิน (Terrain)	
11. ท้องฟ้า (Sky)	
12. บุคคล (Person)	
13. คนขี่ (Rider)	
14. รถยนต์ (Car)	
15. รถบรรทุก (Truck)	
16. รถประจำทาง (Bus)	
17. รถไฟ (Train)	
18. รถจักรยานยนต์ (Moto)	
19. รถจักรยาน (Bicycle)	

ผลลัพธ์จากแบบจำลอง DDRNet-23-Slim จะได้ออกมาเป็นประเภทในแต่ละ pixel ซึ่งเมื่อทำการระบุสีให้ในแต่ละประเภทก็จะได้ผลลัพธ์ตามรูปที่ 12 ที่แสดงให้เห็นถึงวัตถุที่ถูกหาสีด้วยสีของประเภทนั้นๆ



(a) Canberra



(b) Bangkok



(c) Brisbane



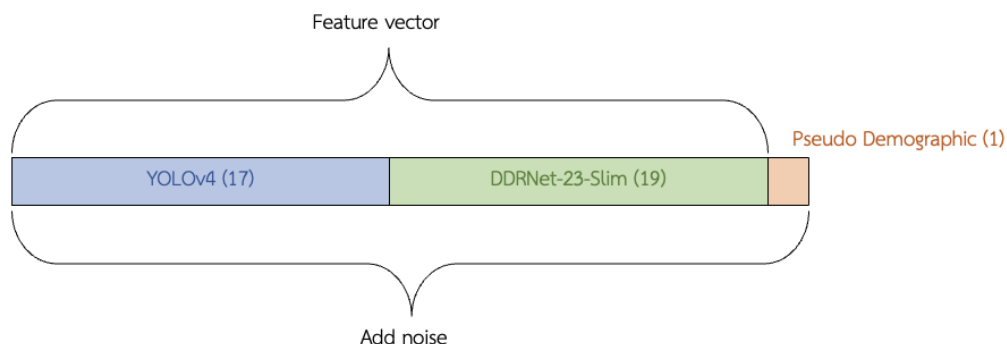
(d) Sakae

รูปที่ 12 ตัวอย่างผลลัพธ์จากแบบจำลอง DDRNet-23-Slim

#### 4.4. ขั้นตอนการเตรียมข้อมูล (Preprocessing)

ในขั้นตอนแรกจะนำวิดีโอมาทำการตัดแบ่งเพื่อลดปริมาณของข้อมูลที่ใช้ในการประมวลผล โดยจะแบ่งเป็น 30 รูปภาพต่อ 1 ฉาก ซึ่งจะถูกประมวลผลด้วยวิธีการถอดคุณสมบัติจากภาพ อีกทั้งข้อมูลที่ได้จากขั้นตอนการตรวจจับวัตถุจะถูกนำไปนับก่อนว่าพบวัตถุที่สนใจจำนวนเท่าไร และหลังจากนั้นจะนำข้อมูลที่ได้มารวมกับข้อมูลของขั้นตอนการแบ่งส่วนความหมายและทำการ Normalize เพื่อให้ข้อมูลอยู่ในช่วงตัวเลขเดียวกัน เมื่อได้ช่วงตัวเลขเดียวกันแล้วจะทำการเพิ่มข้อมูล Pseudo Demographic ซึ่งเป็นค่าที่ถูกสุ่มจากช่วงตัวเลขที่ถูก Normalize ก่อนหน้าเพื่อทำให้ข้อมูลมีความแตกต่างกันเล็กน้อยต่อ 1 ตัวอย่าง ดังนั้นค่าคุณสมบัติที่ได้ต่อ 1 รูปภาพจะมีจำนวนเท่ากับ 37 คุณสมบัติ ดังนั้นขนาดของข้อมูลขาเข้าจะเท่ากับ  $30 \times 37$





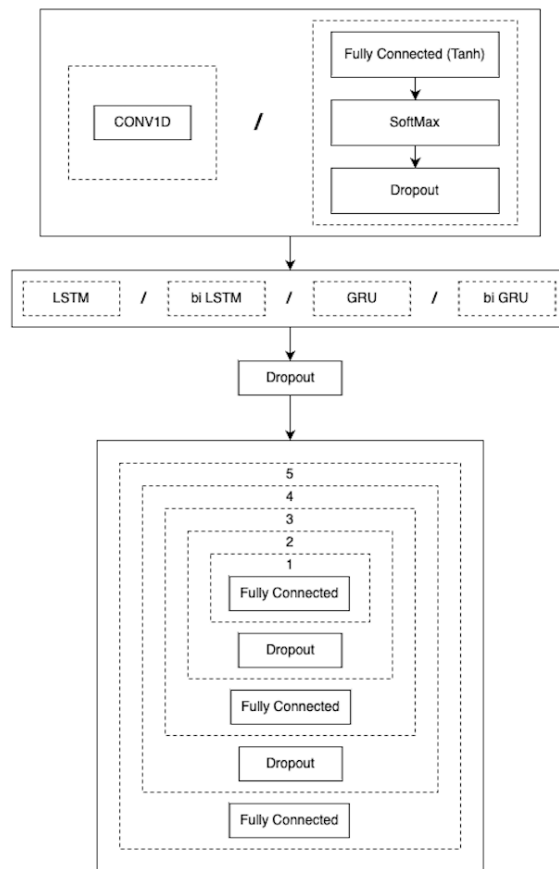
รูปที่ 13 คุณสมบัติที่ถูกใช้เป็นส่วนข้อมูลเข้าของการฝึกสอนแบบจำลอง

เมื่อได้ค่าคุณสมบัติต่อ 1 รูปภาพแล้วจะทำการจับคู่คุณสมบัติกับข้อมูลจากแบบสอบถามของผู้เข้าร่วมทั้ง 50 คนเข้าด้วยกันเพื่อใช้เป็นข้อมูลเป้าหมายซึ่งมีทั้งหมด 15 องค์ประกอบตั้งนั้นขนาดของข้อมูลจะเท่ากับ  $50 \times 15$  และเมื่อนำมาจับคู่กันทำให้ขนาดของข้อมูลเข้าจะกลายเป็น  $50 \times 30 \times 37$  ต่อ 1 ฉาก

ในขั้นตอนสุดท้ายก่อนทำการฝึกสอนจะทำการแบ่งข้อมูลออกเป็นชุดก่อนนำไปเข้าสู่กระบวนการฝึกสอน โดยจะแบ่งข้อมูลออกเป็น 2 ชุดต่อ 1 ฉากเพื่อลดขนาดของโครงสร้างของโครงข่ายประสาทเทียม ดังนั้นขนาดของข้อมูลต่อ 1 ชุดการฝึกสอนของข้อมูลเข้าจะเท่ากับ  $100 \times 15 \times 37$  ทำให้ต้องใช้ข้อมูลเป้าหมาย 2 ชุดด้วยดังนั้นขนาดของข้อมูลเป้าหมายจะเท่ากับ  $100 \times 15$  ต่อ 1 ฉาก

#### 4.5. การออกแบบสถาปัตยกรรมโครงข่ายประสาทเทียม (Neural Networks Architecture Design)

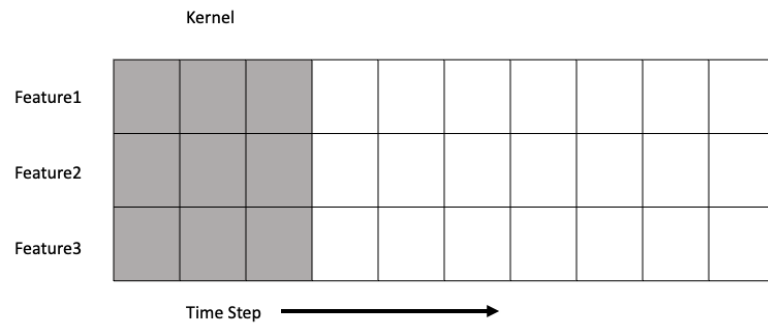
ในขั้นตอนการออกแบบนั้นได้เฟรมเวิร์ค TensorFlow มาใช้ ซึ่งเป็นเฟรมเวิร์คที่เขียนด้วยภาษาคอมพิวเตอร์ Python โดยใช้ในการพัฒนาระบบการเรียนรู้ของเครื่อง งานวิจัยนี้ได้ทำการออกแบบโครงข่ายประสาทเทียมเชิงลึกสำหรับการฝึกสอนแบบจำลองทำนายผลคุณภาพชีวิตตามรูปที่ 14



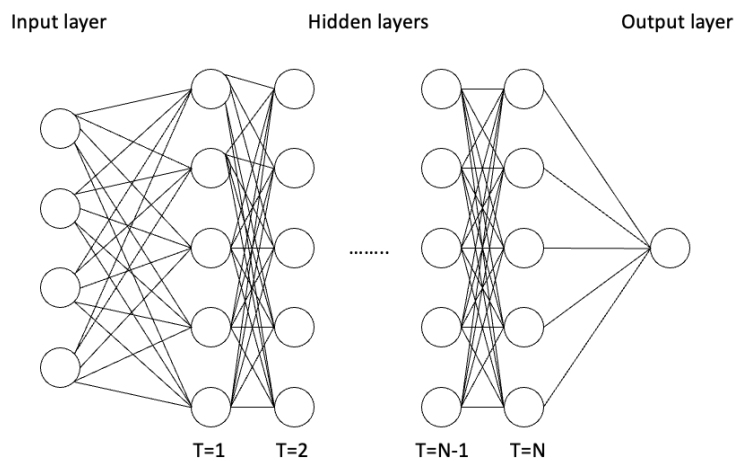
รูปที่ 14 ภาพรวมสถาปัตยกรรมโครงข่ายประสาทเทียมเชิงลึกที่ใช้ในการฝึกสอนแบบจำลองทำนาย

ผลคุณภาพชีวิต

จากรูปที่ 14 จะพบว่ามีการนำโครงสร้างของ CONV1D มาใช้ซึ่งเป็นโครงข่ายประสาทเทียมประเภท Convolutional ประเภทหนึ่ง โดยที่การทำงานจะทำงานเหมือน Convolutional ประเภทอื่นแต่จะแตกต่างที่การขยับของ Kernel จะทำการขยับในทิศทางเดียวเท่านั้น เพื่อที่จะจับคู่คุณลักษณะและสร้าง feature map ด้วยอัลกอริทึมต่างๆตามที่เราเลือกไว้ต่อการประมวลผลในหนึ่งช่วงเวลา ซึ่งจะแสดงให้เห็นตามรูปที่ 15 รวมถึงโครงข่ายประสาทเทียมแบบ fully connected คือโครงข่ายประสาทเทียมที่จุดทุกจุดจะเชื่อมต่อถึงกันเพื่อหาความสัมพันธ์ระหว่างจุดโดยขึ้นอยู่กับค่าน้ำหนักที่ให้ โดยจะแสดงให้เห็นในรูปที่ 16 โดยที่มีการใช้ activation function - tangent hyperbolic (Tanh) ซึ่งเป็นอัลกอริทึมที่ใช้คำนวณค่าในแกน Y ให้อยู่ในช่วงของ -1 ถึง 1 โดยที่มีสมการตามสมการที่ (2) อีกทั้งยังมีการนำอัลกอริทึม SoftMax มาใช้ในการคำนวณเปรียบเทียบสัดส่วนของคุณลักษณะซึ่งผลรวมทั้งหมดจะต้องเท่ากับ 1 โดยที่มีสมการตามสมการที่ (3) รวมถึงมีการใช้ Dropout เพื่อช่วยแก้ปัญหา overfitting อีกด้วย ซึ่งค่า hyperparameter ที่ใช้กับ Dropout ทั้งหมดเท่ากับ 0.2



รูปที่ 15 เวกลักษณะของโครงข่ายประสาทเทียม CONV1D



รูปที่ 16 เวกลักษณะของโครงข่ายประสาทเทียม fully connected

$$\text{Tanh} = \frac{(e^x - e^{-x})}{(e^x + e^{-x})} \quad (2)$$

$$\text{SoftMax} = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad (3)$$

$z_i$  คือข้อมูลขาเข้า ณ ตำแหน่งปัจจุบัน

$e^{z_i}$  คือฟังก์ชัน standard exponential สำหรับข้อมูลขาเข้า ณ ตำแหน่งปัจจุบัน

$K$  คือจำนวนข้อมูลทั้งหมดในรูปแบบ vector

$e^{z_j}$  คือฟังก์ชัน standard exponential สำหรับข้อมูลขาเข้า ณ ตำแหน่งที่  $j$

จากรูปที่ 14 จะอธิบายโครงสร้างได้ดังนี้ ภายในกล่องทึบจะประกอบไปด้วยกล่องเส้นประซึ่งกล่องเส้นประคือโครงสร้างที่สามารถเลือกมาใช้ได้เพียง 1 อย่างภายในกล่องทึบต่อการฝึกสอน 1 ครั้ง ตัวอย่างเช่น ในกล่องทึบแรกได้ทำการเลือกโครงสร้าง CONV1D ทางซ้ายก็จะไม่สามารถเลือก fully

connected ฟังก์ชันได้ ณ รอบนั้น และในกล่องที่ปิดไปก็จะสามารถเลือกโครงข่ายประสาทเทียมแบบวนกลับได้แค่ 1 ประเภทเท่านั้น และในกล่องสุดท้ายจะระบุเลขกำกับไว้หมายความว่าถ้าเลือกกล่องเส้นประที่ 5 จะเท่ากับเลือกใช้โครงสร้างของกล่องหมายเลข 1 ถึงหมายเลข 5 แต่ถ้าเลือกหมายเลข 4 ก็จะเท่ากับเลือกใช้โครงสร้างของกล่องหมายเลข 1 ถึงหมายเลข 4 เป็นต้น จึงสรุปได้ว่าจะมีการฝึกสอนด้วยโครงสร้างทั้งหมด 40 รูปแบบ และในทุกชั้นของโครงสร้างใช้ค่า hyperparameter เท่ากับ 128 ในรูปที่ 17 แสดงตัวอย่างโครงสร้างรูปแบบ CONV1D โครงข่ายประสาทเทียมแบบวนกลับ LSTM และกล่องที่ปิดสุดท้ายหมายเลข 5 ในรูปที่ 18 แสดงตัวอย่างโครงสร้างรูปแบบ fully connected โครงข่ายประสาทเทียมแบบวนกลับ LSTM และกล่องที่ปิดสุดท้ายหมายเลข 5



Layer (type)	Output Shape	Param #
input (InputLayer)	[(None, 15, 37)]	0
conv1d (Conv1D)	(None, 13, 64)	7168
lstm (LSTM)	(None, 128)	98816
dropout (Dropout)	(None, 128)	0
dense (Dense)	(None, 128)	16512
dropout_1 (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 128)	16512
dropout_2 (Dropout)	(None, 128)	0
dense_2 (Dense)	(None, 128)	16512
dense_3 (Dense)	(None, 15)	1935
=====		
Total params: 157,455		
Trainable params: 157,455		
Non-trainable params: 0		

รูปที่ 17 สถาปัตยกรรมโครงข่ายประสาทเทียมรูปแบบ CONV1D LSTM หมายเลข 5

Layer (type)	Output Shape	Param #
input (InputLayer)	[(None, 15, 37)]	0
dense (Dense)	(None, 15, 128)	4736
tf.nn.softmax (TFOpLambda)	(None, 15, 128)	0
dropout (Dropout)	(None, 15, 128)	0
lstm (LSTM)	(None, 128)	131584
dropout_1 (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 128)	16512
dropout_2 (Dropout)	(None, 128)	0
dense_2 (Dense)	(None, 128)	16512
dropout_3 (Dropout)	(None, 128)	0
dense_3 (Dense)	(None, 128)	16512
dense_4 (Dense)	(None, 15)	1935
=====		
Total params: 187,791		
Trainable params: 187,791		
Non-trainable params: 0		

รูปที่ 18 สถาปัตยกรรมโครงข่ายประสาทเทียมรูปแบบ fully connected LSTM หมายเลข 5

#### 4.6. การแบ่งชุดข้อมูลและการฝึกสอนแบบจำลอง (Dataset Splitting and Model Training)

ในส่วนนี้จะทำการแบ่งชุดข้อมูลออกเป็น 2 กลุ่ม ได้แก่ ภายในเมืองเดียวกัน (Within-the-city group) และ ข้ามระหว่างเมือง (Across-the-cities group) ซึ่งหมายความว่ากลุ่มภายในเมืองเดียวกันจะทำการฝึกสอนโดยการสลับฉากภายในเมืองในการฝึกสอนและทดสอบ และกลุ่มข้ามระหว่างเมืองจะทำการฝึกสอนด้วยฉากทั้งหมดของเมืองนั้นๆ และสลับข้อมูลทดสอบของเมืองอื่นๆ โดยที่แสดงให้เห็นในตารางที่ 4 และ ตารางที่ 5 ตามลำดับ

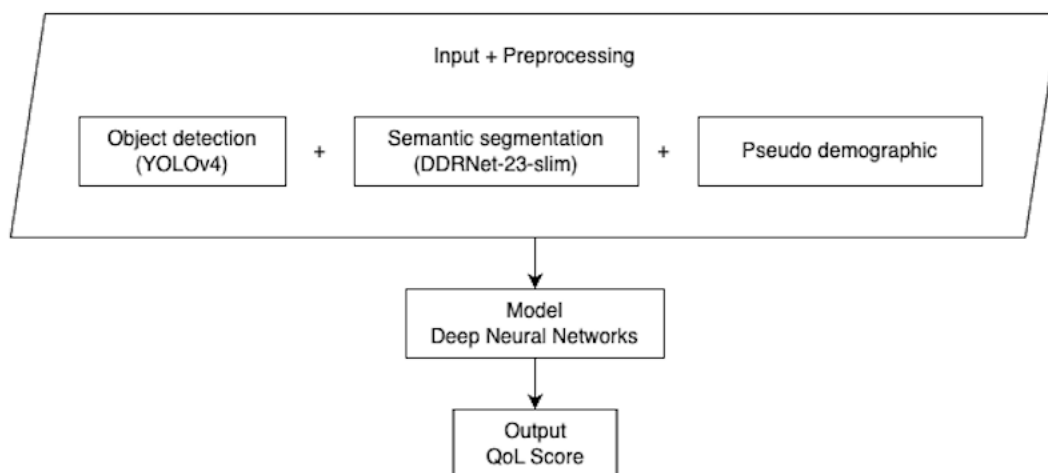
ตารางที่ 4 ข้อมูลฝึกสอนและทดสอบของชุดข้อมูลกลุ่มภายในเมืองเดียวกัน

Experiment Set	Training Data	Testing Data
1	Canberra Scene2, Canberra Scene3	Canberra Scene1
2	Canberra Scene1, Canberra Scene3	Canberra Scene2
3	Canberra Scene1, Canberra Scene2	Canberra Scene3
4	Bangkok Scene2, Bangkok Scene3	Bangkok Scene1
5	Bangkok Scene1, Bangkok Scene3	Bangkok Scene2
6	Bangkok Scene1, Bangkok Scene2	Bangkok Scene3
7	Brisbane Scene2, Brisbane Scene3	Brisbane Scene1
8	Brisbane Scene1, Brisbane Scene3	Brisbane Scene2
9	Brisbane Scene1, Brisbane Scene2	Brisbane Scene3
10	Sakae Scene2, Sakae Scene3	Sakae Scene1
11	Sakae Scene1, Sakae Scene3	Sakae Scene2
12	Sakae Scene1, Sakae Scene2	Sakae Scene3

ตารางที่ 5 ข้อมูลฝึกสอนและทดสอบของชุดข้อมูลกลุ่มข้ามระหว่างเมือง

Experiment Set	Training Data	Testing Data
1	Bangkok Scene all, Brisbane Scene all, Sakae Scene all	Canberra Scene all
2	Canberra Scene all, Brisbane Scene all, Sakae Scene all	Bangkok Scene all
3	Canberra Scene all, Bangkok Scene all, Sakae Scene all	Brisbane Scene all
4	Canberra Scene all, Bangkok Scene all, Brisbane Scene all	Sakae Scene all

งานวิจัยนี้ได้ใช้ Google Collaboratory Pro ในการฝึกสอนแบบจำลอง โดยกำหนดค่าจำนวนรอบในการฝึกสอนทั้งหมด 10,000 รอบ ด้วยอัตราการเรียนรู้เท่ากับ 0.001 โดยภาพรวมของขั้นตอนการฝึกสอนเป็นไปตามรูปที่ 19



รูปที่ 19 ภาพรวมการฝึกสอนแบบจำลอง

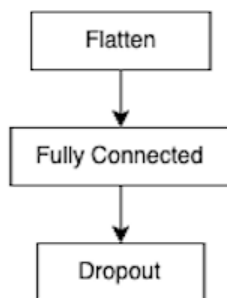
#### 4.7. การประเมินวัดผลแบบจำลอง (Model evaluation)

ในงานวิจัยนี้ใช้ Mean Square Error (MSE) เป็นสมการที่ใช้ในการคำนวณค่าความคลาดเคลื่อนระหว่างผลลัพธ์จริงและผลลัพธ์ที่ทำนายจากแบบจำลอง ซึ่งสมการที่ใช้เป็นดังสมการที่ (4) โดยที่  $n$  เท่ากับจำนวนของข้อมูลเป้าหมาย

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (4)$$

#### 4.8. วิธีการกลั่นความรู้ (Knowledge Distillation)

หลังจากฝึกสอนข้อมูลเสร็จแล้วจะทำการเลือกแบบจำลองที่มีค่าความแม่นยำมากที่สุดมาใช้เป็นแบบจำลอง teacher และจะออกแบบโครงสร้างของโครงข่ายประสาทเทียมขนาดเล็กเพื่อใช้เป็นแบบจำลอง student หลังจากนั้นจะทำการเปรียบเทียบระหว่างแบบจำลองโครงสร้างขนาดเล็กที่ฝึกสอนด้วยตัวเอง และ แบบจำลองโครงสร้างขนาดเล็กที่ฝึกสอนด้วยการกลั่นความรู้ ซึ่งโครงสร้างและ hyperparameter แสดงให้เห็นในรูปที่ 20 และ รูปที่ 21 ตามลำดับ



รูปที่ 20 รูปแบบโครงสร้างสถาปัตยกรรมของแบบจำลองขนาดเล็ก

Layer (type)	Output Shape	Param #
input (InputLayer)	[(None, 15, 37)]	0
flatten (Flatten)	(None, 555)	0
dense (Dense)	(None, 256)	142336
dropout (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 15)	3855

=====  
 Total params: 146,191  
 Trainable params: 146,191  
 Non-trainable params: 0  
 =====

รูปที่ 21 ค่า hyperparameter ของโครงสร้างแบบจำลองขนาดเล็ก

ขั้นตอนในการฝึกสอนด้วยวิธีการกลั่นความรู้ได้จะต้องกำหนดค่า *temperature* และ *alpha* ซึ่งการทดลองนี้ใช้ค่าเท่ากับ 10 และ 0.1 ตามลำดับ อีกทั้งรูปที่ 22 แสดงให้เห็นถึงรหัสเทียมที่ใช้ในการคำนวณหาค่า *loss* ที่ใช้ในการคำนวณปรับน้ำหนักของแบบจำลอง ซึ่ง *loss function* นั้นใช้ MSE ทั้งในส่วนของ *student loss function* และ *distillation loss function*



```

Constant value is temperature and alpha

Function distillationLoss (Argument teacher output, Argument student output) {
    Divide argument teacher output by temperature to get teacher soft label
    Divide argument student output by temperature to get student soft label

    Return Compare teacher soft label with student soft label by distillation loss function
}

Function calculateTotalLoss (Argument student loss result, Argument distillation loss result) {
    Return alpha x argument student loss result + (1 - alpha) x argument distillation loss result
}

Function knowledgeDistillation (Argument input, Argument target) {
    Predict student output by student model with argument input
    Compare student output with argument target by student loss function to get student loss result

    Predict teacher output by pre-trained teacher model with argument input
    Send predicted student output and predicted teacher output to the distillationLoss function to get the distillation loss result

    Send student loss result and distillation loss result to the calculateTotalLoss function to get total loss result
    Use total loss result to adjust weights of student model
}

```

รูปที่ 22 รหัสเทียมของขั้นตอนการคำนวณหาค่า loss รวมที่ใช้ในการปรับค่าน้ำหนักของแบบจำลอง



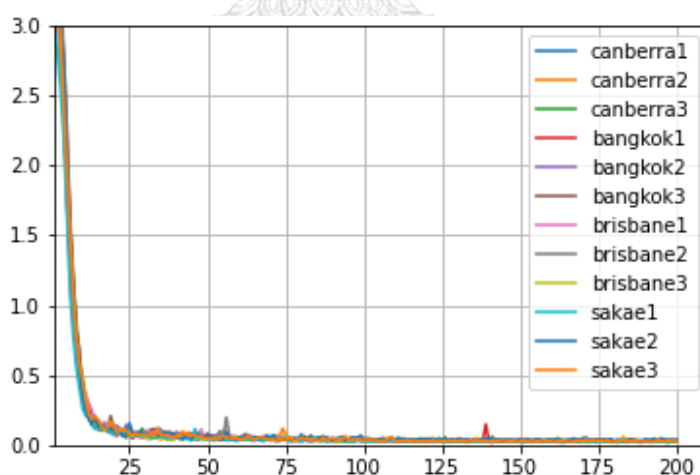
## บทที่ 5

### ผลการทดลอง

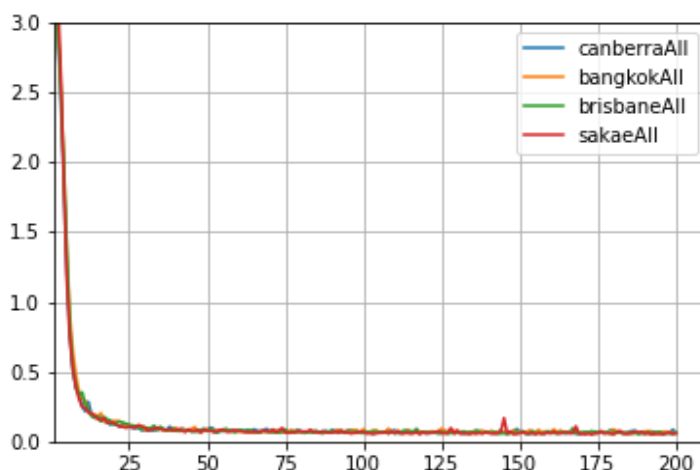
ในบทนี้จะกล่าวถึงผลการทดลองที่แบ่งเป็นสองส่วน ประกอบด้วย 1. ผลลัพธ์การฝึกสอนโมเดลด้วยวิธีการเรียนรู้เชิงลึก ซึ่งจะเปรียบเทียบและแสดงผลโครงสร้างที่ดีที่สุดของแต่ละชุดการทดลองทั้งแบบภายในเมืองเดียวกัน (Within-the-city group) และข้ามระหว่างเมือง (Across-the-cities group) 2. ผลลัพธ์ของการฝึกสอนด้วยวิธีการกลั่นความรู้ ซึ่งจะทำการเปรียบเทียบระหว่างแบบจำลองโครงสร้างขนาดเล็กที่ฝึกสอนด้วยตัวเองกับฝึกสอนด้วยวิธีการกลั่นความรู้

#### 5.1. ผลลัพธ์การฝึกสอนโมเดลด้วยวิธีการเรียนรู้เชิงลึก

ขั้นตอนนี้จะทำการฝึกสอนด้วยจำนวนรอบเท่ากับ 10,000 รอบต่อหนึ่งรูปแบบการทดลอง รูปที่ 23 และ รูปที่ 24 แสดงให้เห็นถึงตัวอย่างรูปแบบของค่า loss ที่ลดลง ของการฝึกสอนด้วยชุดการทดลองภายในเมืองเดียวกันและข้ามระหว่างเมืองตามลำดับ ซึ่งจะทำการบันทึกผลทุกๆ 50 รอบ ดังนั้นจะได้จุดของกราฟทั้งหมด 200 จุด โดยที่ค่าที่แสดงผลได้จากค่า loss ที่คำนวณผลต่างของค่าจริงเทียบกับค่าที่ทำนายได้ด้วย loss function MSE



รูปที่ 23 ตัวอย่างของค่า loss ที่เปลี่ยนไปของการฝึกสอนแบบจำลองด้วยชุดข้อมูลภายในเมืองเดียวกัน



รูปที่ 24 ตัวอย่างของค่า loss ที่เปลี่ยนไปของการฝึกสอนแบบจำลองด้วยชุดข้อมูลข้ามระหว่างเมือง

ผลลัพธ์แบบจำลองที่ดีที่สุดซึ่งมีค่า loss ต่ำที่สุดของแต่ละชุดทดสอบของชุดข้อมูลภายในเมืองเดียวกันจะแสดงให้เห็นตามตารางที่ 6 และชุดข้อมูลข้ามระหว่างเมืองจะแสดงให้เห็นตามตารางที่ 7

จากตารางที่ 6 แสดงให้เห็นว่าแบบจำลองที่ใช้โครงสร้าง fully connected tanh activation function ที่ใช้ร่วมกับโครงข่ายประสาทเทียมแบบย้อนกลับ bi-directional LSTM หมายเลข 1 คือแบบจำลองที่ดีที่สุดเนื่องจากผลลัพธ์ที่ได้มันดีที่สุด 8 ชุดจากทั้งหมด 12 ชุดการทดลอง รวมถึงตารางที่ 7 ก็แสดงให้เห็นเช่นเดียวกันว่าโครงสร้าง fully connected tanh activation function ที่ใช้ร่วมกับโครงข่ายประสาทเทียมแบบย้อนกลับ bi-directional LSTM หมายเลข 1 ได้ผลลัพธ์ที่ดีที่สุด

ตารางที่ 6 ผลลัพธ์การทดลองของชุดการทดลองภายในเมืองเดียวกัน

Training Dataset	Testing Dataset	Model Name	MSE
Canberra Scene2, Canberra Scene3	Canberra Scene1	Tanh bi LSTM 1	$7.41 \times 10^{-3}$
Canberra Scene1, Canberra Scene3	Canberra Scene2	Tanh bi LSTM 1	$8.35 \times 10^{-3}$
Canberra Scene1, Canberra Scene2	Canberra Scene3	Tanh bi LSTM 1	$7.92 \times 10^{-3}$
Bangkok Scene2, Bangkok Scene3	Bangkok Scene1	CONV1D bi LSTM 1	$8.11 \times 10^{-3}$

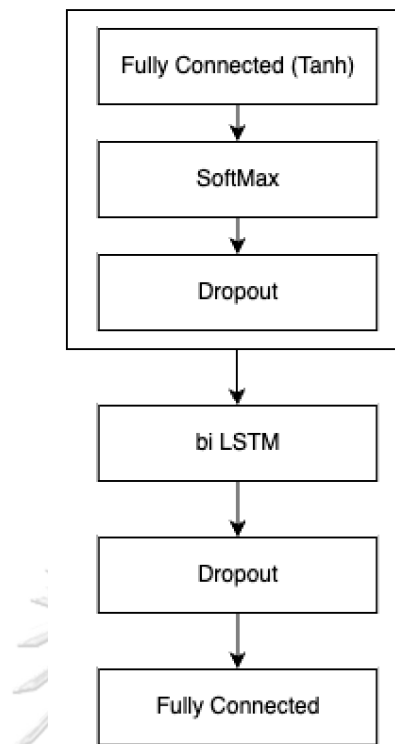
Bangkok Scene1, Bangkok Scene3	Bangkok Scene2	Tanh bi LSTM 1	$7.55 \times 10^{-3}$
Bangkok Scene1, Bangkok Scene2	Bangkok Scene3	CONV1D bi LSTM 1	$8.62 \times 10^{-3}$
Brisbane Scene2, Brisbane Scene3	Brisbane Scene1	Tanh bi LSTM 1	$7.19 \times 10^{-3}$
Brisbane Scene1, Brisbane Scene3	Brisbane Scene2	Tanh bi GRU 1	$9.27 \times 10^{-3}$
Brisbane Scene1, Brisbane Scene2	Brisbane Scene3	Tanh bi LSTM 1	$7.55 \times 10^{-3}$
Sakae Scene2, Sakae Scene3	Sakae Scene1	Tanh bi GRU 1	$6.80 \times 10^{-3}$
Sakae Scene1, Sakae Scene3	Sakae Scene2	Tanh bi LSTM 1	$1.14 \times 10^{-2}$
Sakae Scene1, Sakae Scene2	Sakae Scene3	Tanh bi LSTM 1	$8.90 \times 10^{-3}$

ตารางที่ 7 ผลลัพธ์การทดลองของชุดการทดลองข้ามระหว่างเมือง

Training Dataset	Testing Dataset	Model Name	MSE
Bangkok Scene all, Brisbane Scene all, Sakae Scene all	Canberra Scene all	Tanh bi LSTM 1	$9.73 \times 10^{-3}$
Canberra Scene all, Brisbane Scene all, Sakae Scene all	Bangkok Scene all	Tanh bi LSTM 1	$1.20 \times 10^{-2}$
Canberra Scene all, Bangkok Scene all, Sakae Scene all	Brisbane Scene all	Tanh bi LSTM 1	$1.16 \times 10^{-2}$
Canberra Scene all, Bangkok Scene all, Brisbane Scene all	Sakae Scene all	Tanh bi LSTM 1	$1.17 \times 10^{-2}$

## 5.2. ผลลัพธ์การฝึกสอนโมเดลด้วยวิธีการกั้นความรู้

จากขั้นตอนก่อนหน้าสรุปได้ว่าแบบจำลอง fully connected tanh activation function ที่ใช้ร่วมกับโครงข่ายประสาทเทียมแบบย้อนกลับ bi-directional LSTM หมายเลข 1 คือแบบจำลองที่ดีที่สุดเนื่องจากมีค่า loss ที่ต่ำที่สุดเกือบทุกชุดการทดลอง ดังนั้นในขั้นตอนนี้จะนำแบบจำลองนี้มาใช้ในการเป็นแบบจำลอง teacher เพื่อใช้ในการฝึกสอนให้กับแบบจำลองโครงสร้างขนาดเล็ก รูปที่ 25 แสดงให้เห็นถึงโครงสร้างสุดท้ายที่ถูกนำมาใช้เป็นแบบจำลอง teacher ซึ่งใช้ค่า hyperparameter ตามรูปที่ 26



รูปที่ 25 รูปแบบโครงสร้างสถาปัตยกรรมของแบบจำลอง fully connected tanh activation function ที่ใช้ร่วมกับโครงข่ายประสาทเทียมแบบย้อนกลับ bi-directional LSTM หมายเลข 1

Layer (type)	Output Shape	Param #
input (InputLayer)	[(None, 15, 37)]	0
dense (Dense)	(None, 15, 128)	4736
tf.nn.softmax (TFOpLambda)	(None, 15, 128)	0
dropout (Dropout)	(None, 15, 128)	0
bidirectional (Bidirectional)	(None, 256)	263168
dropout_1 (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 128)	32896
dense_2 (Dense)	(None, 15)	1935
=====		
Total params: 302,735		
Trainable params: 302,735		
Non-trainable params: 0		

รูปที่ 26 ค่า hyperparameter ของแบบจำลอง fully connected tanh activation function ที่ใช้ร่วมกับโครงข่ายประสาทเทียมแบบย้อนกลับ bi-directional LSTM หมายเลข 1

ในการทดลองส่วนนี้ได้้นำแบบจำลอง fully connected tanh activation function ที่ใช้ร่วมกับโครงข่ายประสาทเทียมแบบย้อนกลับ bi-directional LSTM หมายเลข 1 มาใช้ในการฝึกสอนด้วยวิธีการกลั่นความรู้ทั้งชุดทดสอบรูปแบบภายในเมืองเดียวกันและข้ามระหว่างเมือง โดยตารางที่ 8 และ ตารางที่ 10 คือผลลัพธ์ของการฝึกสอนด้วยชุดทดสอบภายในเมืองเดียวกัน ตารางที่ 11 และ ตารางที่ 13 คือผลลัพธ์ของการฝึกสอนด้วยชุดข้อมูลทดสอบข้ามระหว่างเมือง

ตารางที่ 8 แสดงให้เห็นถึงผลลัพธ์ของค่า MSE loss ในแต่ละชุดการทดลองด้วยชุดข้อมูลทดสอบภายในเมืองเดียวกัน ผลลัพธ์แสดงให้เห็นว่าแบบจำลองที่ถูกกลั่นความรู้ให้ผลลัพธ์ของค่า MSE loss ที่ต่ำกว่า แบบจำลองโครงสร้างขนาดเล็กที่ฝึกสอนด้วยตัวเอง สรุปได้ว่าประสิทธิภาพของวิธีการกลั่นความรู้เป็นไปตามที่คาดหวัง

ตารางที่ 8 ค่าผลลัพธ์ MSE ของชุดการทดลองด้วยชุดข้อมูลทดสอบภายในเมืองเดียวกัน

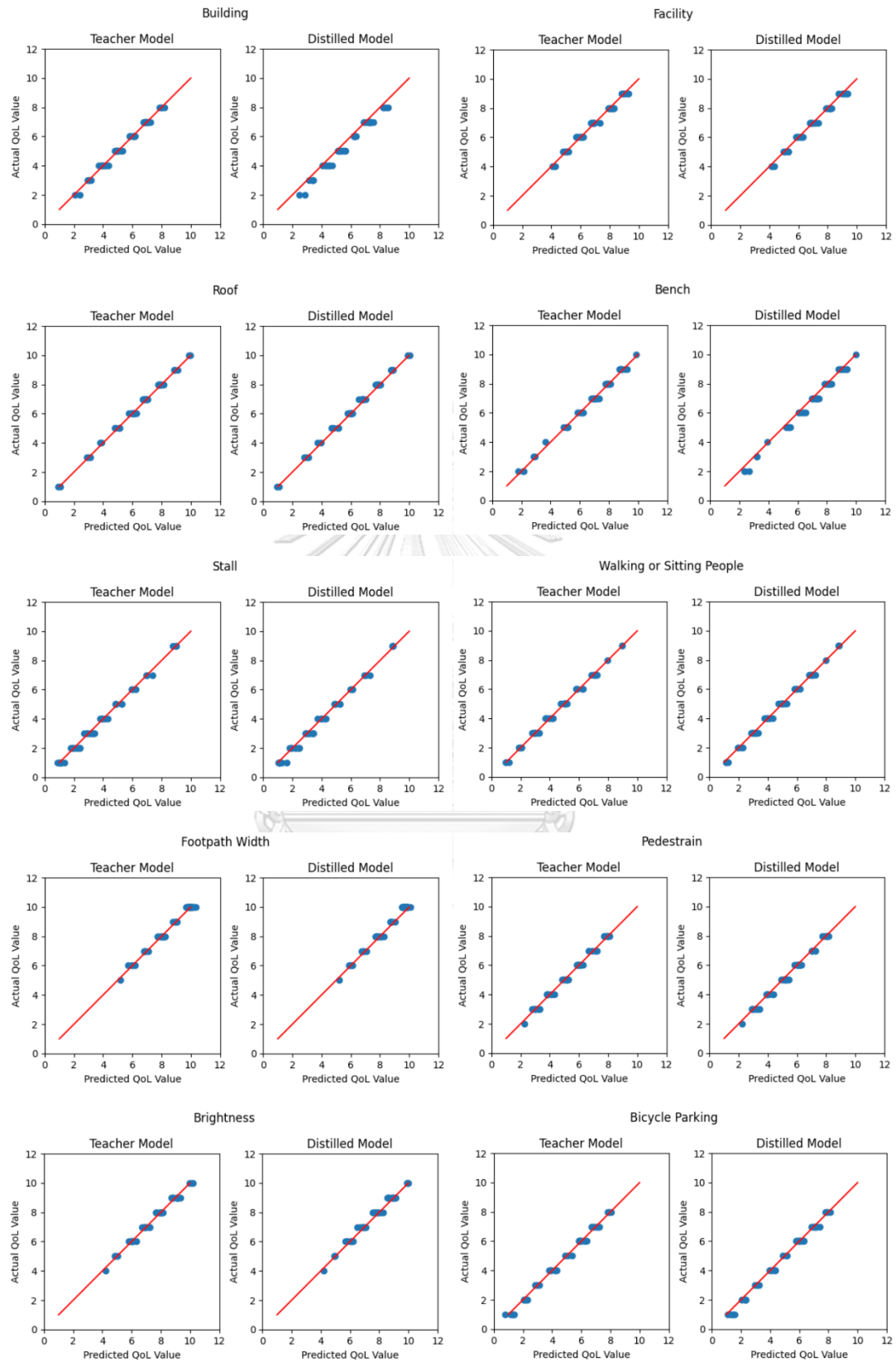
Training Dataset	Testing Dataset	Student Model MSE	Teacher Model MSE	Distilled Model MSE
Canberra Scene2, Canberra Scene3	Canberra Scene1	2.92	$7.16 \times 10^{-3}$	$6.16 \times 10^{-2}$
Canberra Scene1, Canberra Scene3	Canberra Scene2	3.50	$9.08 \times 10^{-3}$	$5.24 \times 10^{-2}$
Canberra Scene1, Canberra Scene2	Canberra Scene3	3.62	$8.26 \times 10^{-3}$	$2.77 \times 10^{-2}$
Bangkok Scene2, Bangkok Scene3	Bangkok Scene1	3.40	$6.09 \times 10^{-3}$	$3.30 \times 10^{-2}$
Bangkok Scene1, Bangkok Scene3	Bangkok Scene2	3.49	$8.46 \times 10^{-3}$	$6.86 \times 10^{-2}$
Bangkok Scene1,	Bangkok	4.04	$6.93 \times 10^{-3}$	$8.91 \times 10^{-2}$

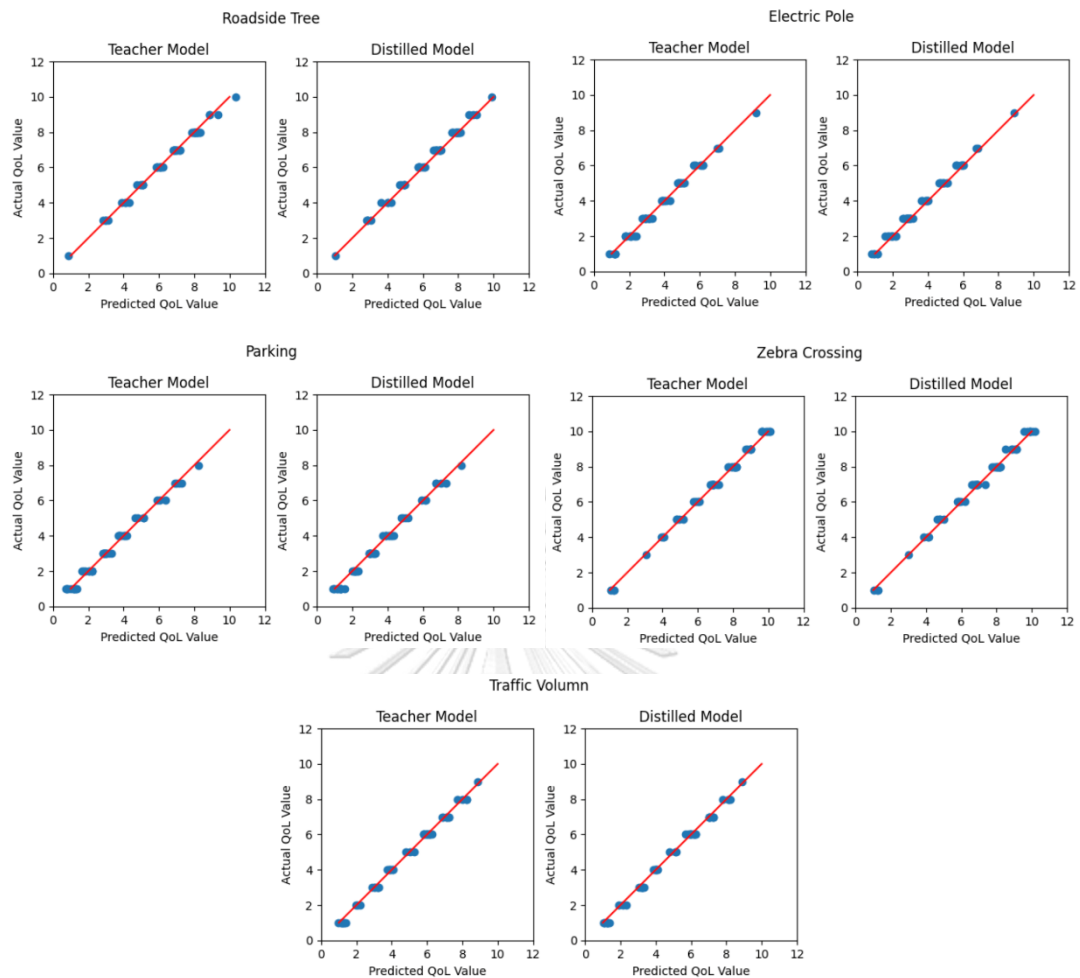
Bangkok Scene2	Scene3			
Brisbane Scene2, Brisbane Scene3	Brisbane Scene1	<b>3.93</b>	$7.74 \times 10^{-3}$	$5.07 \times 10^{-2}$
Brisbane Scene1, Brisbane Scene3	Brisbane Scene2	<b>3.67</b>	$7.26 \times 10^{-3}$	$5.99 \times 10^{-2}$
Brisbane Scene1, Brisbane Scene2	Brisbane Scene3	<b>3.57</b>	$6.53 \times 10^{-3}$	$5.34 \times 10^{-2}$
Sakae Scene2, Sakae Scene3	Sakae Scene1	<b>2.97</b>	$8.73 \times 10^{-3}$	$4.80 \times 10^{-2}$
Sakae Scene1, Sakae Scene3	Sakae Scene2	<b>3.63</b>	$9.45 \times 10^{-3}$	$5.08 \times 10^{-2}$
Sakae Scene1, Sakae Scene2	Sakae Scene3	<b>3.64</b>	$5.14 \times 10^{-3}$	$2.93 \times 10^{-2}$
Average MSE		<b>3.53</b>	$7.57 \times 10^{-3}$	$5.20 \times 10^{-2}$

### จุฬาลงกรณ์มหาวิทยาลัย

ผลลัพธ์จากแบบจำลอง teacher และ แบบจำลองที่ถูกกลั่นความรู้แสดงให้เห็นในรูปที่ 27 โดยที่นำแบบจำลองที่ใช้ชุดข้อมูลเมืองแคนเบอร์ราฉากที่ 1 และ 2 เป็นชุดข้อมูลฝึกสอนและใช้ชุดข้อมูลเมืองแคนเบอร์ราฉากที่ 3 เป็นชุดข้อมูลทดสอบ ซึ่งจากรูปแสดงให้เห็นว่าผลทำนายเป็นไปในทิศทางเดียวกัน แต่แบบจำลองที่ถูกกลั่นความรู้มีขนาดที่เล็กกว่าและใช้เวลาในการประมวลผลน้อยกว่า ซึ่งสามารถใช้แทนแบบจำลอง teacher ได้ อีกทั้งยังมีประสิทธิภาพที่ดีกว่า







รูปที่ 27 เปรียบเทียบตัวอย่างผลลัพธ์ของแบบจำลอง teacher และแบบจำลองที่ถูกกลั่นความรู้อยู่ด้วยชุดข้อมูลทดสอบภายในเมืองเดียวกันที่ใช้ข้อมูลเมืองแคนเบอร์ราฉากที่ 3 เป็นชุดทดสอบ

จากรูปที่ 27 เมื่อนำผลลัพธ์ที่ได้มาทำการคำนวณหาค่า MSE ก็จะแสดงให้เห็นว่าองค์ประกอบใดที่แบบจำลองทำนายผลได้ค่าความแม่นยำที่สูงและองค์ประกอบใดที่แบบจำลองทำนายผลได้ค่าความแม่นยำต่ำ อีกทั้งยังเปรียบเทียบว่าแบบจำลองที่ถูกกลั่นความรู้นั้นสามารถทำนายผลที่องค์ประกอบต่างๆแตกต่างจากแบบจำลอง teacher เท่าไหร่ ซึ่งแสดงให้เห็นในตารางที่

ตารางที่ 9 เปรียบเทียบค่า MSE ของแต่ละองค์ประกอบที่ใช้ในการประเมินผ่านแบบสอบถาม ของแบบจำลอง teacher และแบบจำลองที่ถูกกลั่นความรู้ด้วยชุดข้อมูลทดสอบภายในเมืองเดียวกันที่ใช้ข้อมูลเมืองแคนเบอร์ราจากที่ 3 เป็นชุดทดสอบ

QoL Factor	Teacher Model MSE	Distilled Model MSE
Building	0.02543	0.13402
Facility	0.03127	0.03624
Roof	0.02014	0.02908
Bench	0.02794	0.07245
Stall	0.04166	0.05144
Walking or Sitting People	0.02340	0.02461
Footpath Width	0.02537	0.05869
Pedestrian	0.02919	0.03759
Brightness	0.02867	0.04574
Bicycle Parking	0.03830	0.04777
Roadside Tree	0.02384	0.03634
Electric Pole	0.03089	0.04556
Parking	0.03772	0.04361
Zebra Crossing	0.02850	0.04087
Traffic Volume	0.03100	0.03202

ตารางที่ 10 แสดงให้เห็นถึงผลลัพธ์ของเวลาที่ใช้ในการประมวลผลในแต่ละชุดการทดลอง ด้วยชุดข้อมูลทดสอบภายในเมืองเดียวกัน ผลลัพธ์การเปรียบเทียบแสดงให้เห็นว่าแบบจำลองที่ถูกกลั่นความรู้ให้ผลลัพธ์ของเวลาในการประมวลผลที่ลดลงจากแบบจำลอง teacher สรุปได้ว่าประสิทธิภาพของวิธีการกลั่นความรู้เป็นไปตามที่คาดหวัง

ตารางที่ 10 เวลาที่ใช้ในการประมวลผลของแบบจำลองของชุดการทดลองด้วยชุดข้อมูลทดสอบ  
ภายในเมืองเดียวกัน

Training Dataset	Testing Dataset	Teacher Model Computational Time (ms)	Distilled Model Computational Time (ms)
Canberra Scene2, Canberra Scene3	Canberra Scene1	7.518	1.271
Canberra Scene1, Canberra Scene3	Canberra Scene2	7.076	0.861
Canberra Scene1, Canberra Scene2	Canberra Scene3	6.804	0.893
Bangkok Scene2, Bangkok Scene3	Bangkok Scene1	6.830	0.908
Bangkok Scene1, Bangkok Scene3	Bangkok Scene2	7.100	0.868
Bangkok Scene1, Bangkok Scene2	Bangkok Scene3	7.023	0.836
Brisbane Scene2, Brisbane Scene3	Brisbane Scene1	7.051	1.306
Brisbane Scene1, Brisbane Scene3	Brisbane Scene2	6.867	0.893
Brisbane Scene1, Brisbane Scene2	Brisbane Scene3	6.991	0.948
Sakae Scene2,	Sakae Scene1	10.182	1.295

Sakae Scene3			
Sakae Scene1, Sakae Scene3	Sakae Scene2	<b>7.208</b>	<b>0.882</b>
Sakae Scene1, Sakae Scene2	Sakae Scene3	<b>8.152</b>	<b>1.203</b>
Average MSE		<b>7.400</b>	<b>1.014</b>

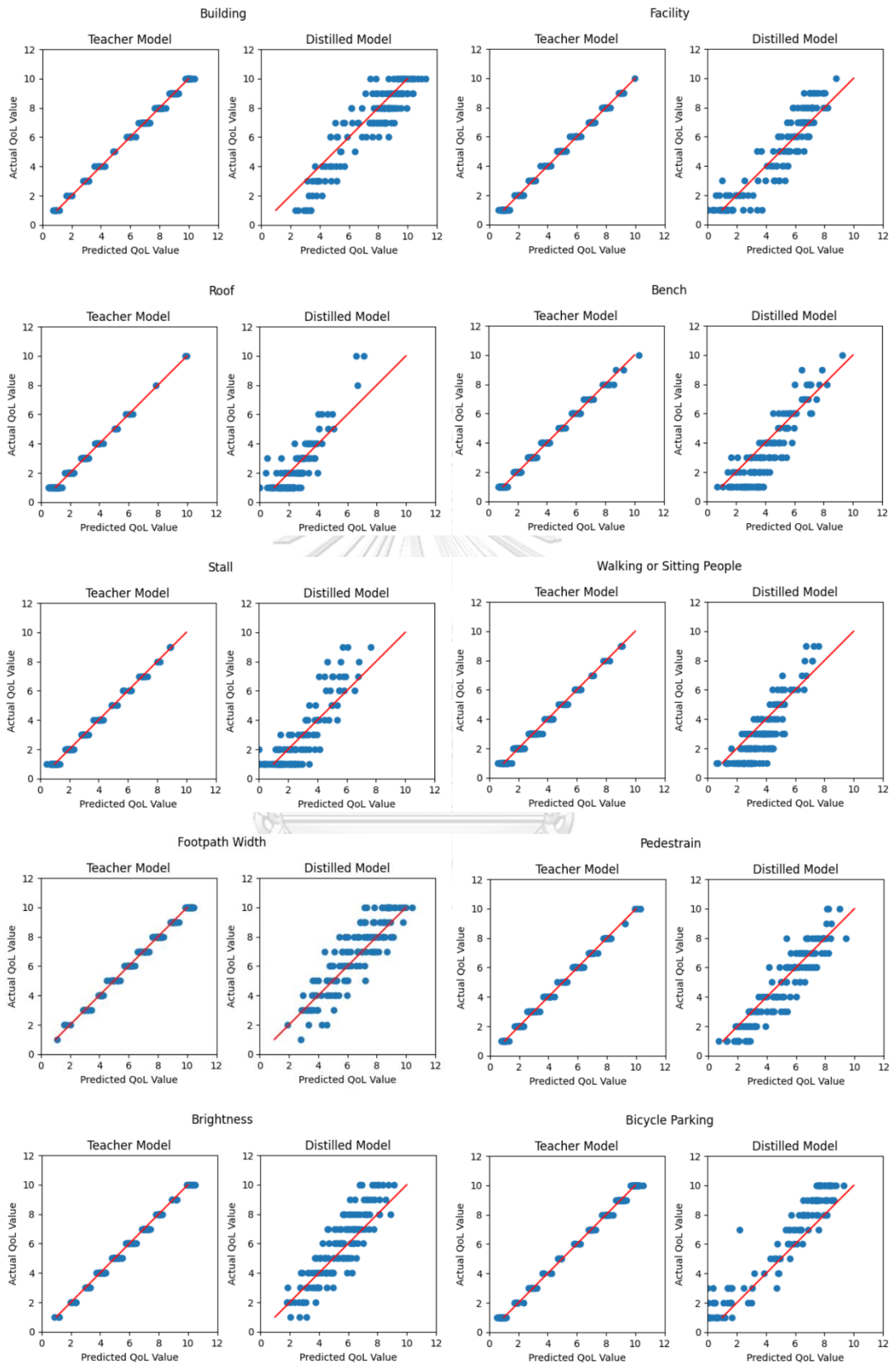
ตารางที่ 11 แสดงให้เห็นถึงผลลัพธ์ของค่า MSE loss ในแต่ละชุดการทดลองด้วยชุดข้อมูลทดสอบข้ามระหว่างเมือง ผลลัพธ์แสดงให้เห็นว่าแบบจำลองโครงสร้างขนาดเล็กที่ถูกฝึกสอนด้วยวิธีการกลั่นความรู้มีค่า MSE loss ต่ำกว่าการฝึกสอนด้วยตัวเองแต่สูงกว่าแบบจำลอง teacher สรุปได้ว่าประสิทธิภาพของวิธีการกลั่นความรู้เป็นไปตามที่คาดหวัง

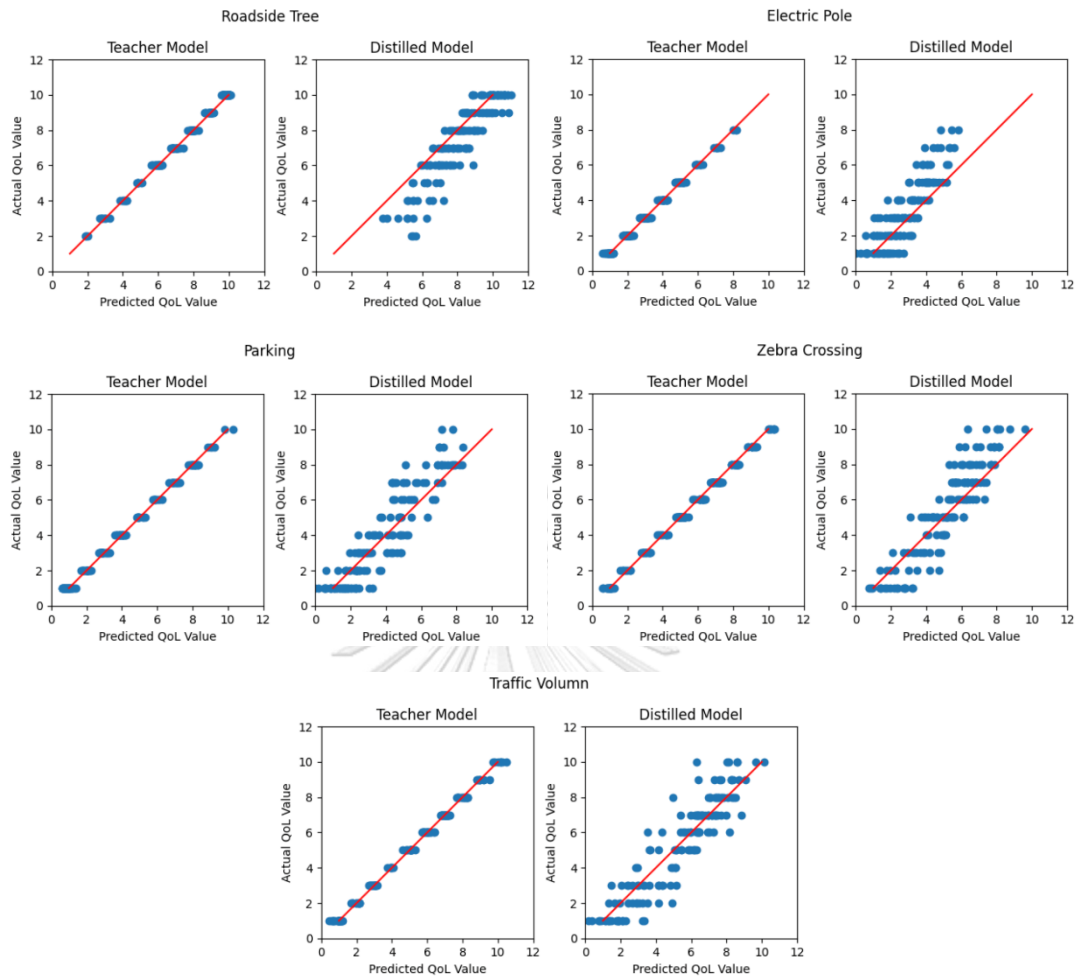
ตารางที่ 11 ค่าผลลัพธ์ MSE ของชุดการทดลองด้วยชุดข้อมูลทดสอบข้ามระหว่างเมือง

Training Dataset	Testing Dataset	Student Model MSE	Teacher Model MSE	Distilled Model MSE
Bangkok Scene all, Brisbane Scene all, Sakae Scene all	Canberra Scene all	<b>3.81</b>	$9.65 \times 10^{-3}$	<b>1.23</b>
Canberra Scene all, Brisbane Scene all, Sakae Scene all	Bangkok Scene all	<b>3.55</b>	$1.13 \times 10^{-2}$	<b>1.28</b>
Canberra Scene all, Bangkok Scene all,	Brisbane Scene all	<b>3.51</b>	$1.14 \times 10^{-2}$	<b>1.14</b>

Sakae Scene all				
Canberra Scene all, Bangkok Scene all, Brisbane Scene all	Sakae Scene all	<b>3.87</b>	<b><math>1.04 \times 10^{-2}</math></b>	<b>1.26</b>
Average MSE		<b>3.69</b>	<b><math>1.07 \times 10^{-2}</math></b>	<b>1.23</b>

ผลลัพธ์จากแบบจำลอง teacher และ แบบจำลองที่ถูกกั้นความรู้แสดงให้เห็นในรูปที่ 28 โดยที่เลือกองค์ประกอบตามการเปรียบเทียบด้วยชุดข้อมูลทดสอบภายในเมืองเดียวกัน และนำแบบจำลองที่ใช้ชุดข้อมูลทุกฉากของกรุงเทพมหานคร เมืองบริสเบน เมืองซากาเอะเป็นชุดข้อมูลฝึกสอน และใช้ชุดข้อมูลทุกฉากของเมืองแคนเบอร์ราเป็นชุดข้อมูลทดสอบ ซึ่งจากรูปแสดงให้เห็นว่าผลทำนายของแบบจำลองที่ถูกกั้นความรู้มีความเสถียรและความแม่นยำน้อยกว่าแบบจำลอง teacher ซึ่งแลกมากับความเร็วในการประมวลผลที่มากกว่าและขนาดแบบจำลองที่เล็กกว่า





รูปที่ 28 เปรียบเทียบตัวอย่างผลลัพธ์ของแบบจำลอง teacher และแบบจำลองที่ถูกกลั่นความรู้ด้วยชุดข้อมูลทดสอบข้ามระหว่างเมืองที่ใช้ข้อมูลทุกภาคของเมืองแคนเบอร์ราเป็นชุดทดสอบ

CHULALONGKORN UNIVERSITY

จากรูปที่ 28 เมื่อนำผลลัพธ์ที่ได้มาทำการคำนวณหาค่า MSE เช่นเดียวกับตารางที่ 9 ก็จะพบว่าแสดงให้เห็นว่าแบบจำลองนั้นสามารถทำนายองค์ประกอบไหนได้แม่นยำมากกว่า และองค์ประกอบไหนมีความแม่นยำน้อยกว่า อีกทั้งค่า MSE ที่คำนวณได้ทำให้เห็นว่าแบบจำลองที่ถูกกลั่นความรู้มีค่าความคาดเคลื่อนมากกว่าแบบจำลอง teacher ในทุกองค์ประกอบ ซึ่งแสดงให้เห็นในตารางที่ 12



ตารางที่ 12 เปรียบเทียบค่า MSE ของแต่ละองค์ประกอบที่ใช้ในการประเมินผ่านแบบสอบถาม ของแบบจำลอง teacher และแบบจำลองที่ถูกกลั่นความรู้ด้วยชุดข้อมูลทดสอบข้ามระหว่างเมืองที่ใช้ข้อมูลทุกฉากของเมืองแคนเบอร์ราเป็นชุดทดสอบ

QoL Factor	Teacher Model MSE	Distilled Model MSE
Building	0.03144	1.26911
Facility	0.02819	0.99365
Roof	0.03177	0.81009
Bench	0.03356	1.67097
Stall	0.02984	1.23812
Walking or Sitting People	0.03891	1.65389
Footpath Width	0.04261	1.26393
Pedestrian	0.03639	0.88205
Brightness	0.04128	1.41395
Bicycle Parking	0.03659	1.86472
Roadside Tree	0.03021	1.36591
Electric Pole	0.02704	1.25414
Parking	0.03466	1.17904
Zebra Crossing	0.04012	1.40958
Traffic Volume	0.03675	1.11398

ตารางที่ 13 แสดงให้เห็นถึงผลลัพธ์ของเวลาที่ใช้ในการประมวลผลในแต่ละชุดการทดลอง ด้วยชุดข้อมูลทดสอบภายในเมืองเดียวกัน ผลลัพธ์การเปรียบเทียบแสดงให้เห็นว่าแบบจำลองที่ถูกกลั่นความรู้ให้ผลลัพธ์ของเวลาในการประมวลผลที่ดีกว่าแบบจำลอง teacher เนื่องจากมีค่าการใช้เวลาในการประมวลผลลดลงเกือบ 10 เท่าโดยประมาณ สรุปได้ว่าประสิทธิภาพของวิธีการกลั่นความรู้เป็นไปตามที่คาดหวัง

ตารางที่ 13 เวลาที่ใช้ในการประมวลผลของแบบจำลองของชุดการทดลองด้วยชุดข้อมูลทดสอบข้ามระหว่างเมือง

Training Dataset	Testing Dataset	Teacher Model Computational Time (ms)	Distilled Model Computational Time (ms)
Bangkok Scene all, Brisbane Scene all, Sakae Scene all	Canberra Scene all	2.330	0.299
Canberra Scene all, Brisbane Scene all, Sakae Scene all	Bangkok Scene all	2.308	0.291
Canberra Scene all, Bangkok Scene all, Sakae Scene all	Brisbane Scene all	2.392	0.292
Canberra Scene all, Bangkok Scene all, Brisbane Scene all	Sakae Scene all	2.302	0.322
Average MSE		2.333	0.301

## บทที่ 6

### วิเคราะห์ผลการทดลอง

#### 6.1. ผลการทดลองการฝึกสอนแบบจำลองด้วยวิธีการเรียนรู้เชิงลึก

จากผลลัพธ์แสดงให้เห็นว่าแบบจำลองโครงสร้าง fully connected tanh activation function ร่วมกับ bi-directional LSTM หมายเลข 1 ซึ่งมีโครงสร้างตามรูปที่ 25 โดยประกอบไปด้วยชั้นซ่อนทั้งหมด 6 ชั้น ได้แก่ 1. ชั้น fully connected tanh activation function 2. ชั้น อัลกอริทึม SoftMax 3. ชั้น Dropout 4. ชั้น bi-directional LSTM 5. ชั้น Dropout 6. ชั้น fully connected เป็นโครงสร้างแบบจำลองที่ได้ผลลัพธ์ดีที่สุด ซึ่งให้ค่า MSE loss ต่ำที่สุดเกือบทุกชุดการทดลอง ซึ่งต่ำที่สุดของชุดการทดลองภายในเมืองเดียวกันมีค่าเท่ากับ  $7.19 \times 10^{-3}$  ด้วยชุดข้อมูลทดสอบเมืองบริสเบน ฉากที่ 3 ที่ใช้ข้อมูลเมืองบริสเบน ฉากที่ 1 และ 2 ในการฝึกสอน ซึ่งแสดงให้เห็นในตารางที่ 6 อีกทั้ง ต่ำที่สุดของชุดการทดลองข้ามระหว่างเมืองมีค่าเท่ากับ  $9.73 \times 10^{-3}$  ด้วยชุดข้อมูลทดสอบเมืองแคนเบอร์รา ทั้ง 3 ฉาก ที่ใช้ข้อมูลในการฝึกสอนด้วย กรุงเทพมหานครทั้ง 3 ฉาก เมืองบริสเบนทั้ง 3 ฉาก และ เมืองซากาเอะทั้ง 3 ฉาก ซึ่งแสดงให้เห็นในตารางที่ 7

#### 6.2. ผลการทดลองการฝึกสอนแบบจำลองด้วยวิธีการกลั่นความรู้

วิธีการกลั่นความรู้ได้ถูกใช้ด้วยการนำแบบจำลองที่ดีที่สุดของขั้นตอนก่อนหน้ามาใช้ ซึ่งก็คือแบบจำลองโครงสร้าง fully connected tanh activation function ร่วมกับ bi-directional LSTM หมายเลข 1 โดยสามารถลดความซับซ้อนได้ด้วยการถ่ายทอดความรู้ให้กับแบบจำลองที่มีโครงสร้างขนาดเล็กกว่า ผลลัพธ์จากการทดลองพบว่าแบบจำลองที่ผ่านขั้นตอนการกลั่นความรู้ให้ผลลัพธ์ด้านเวลาที่ใช้ในการประมวลผลดีกว่าแบบจำลองดั้งเดิม แต่แลกมากับความแม่นยำที่ลดลง ซึ่งผลลัพธ์ของชุดการทดลองภายในเมืองเดียวกันมีค่าเฉลี่ยของค่า MSE loss อยู่ที่  $5.20 \times 10^{-2}$  ตามตารางที่ 8 และผลลัพธ์ของชุดการทดลองข้ามระหว่างเมืองมีค่าเฉลี่ยของค่า MSE loss อยู่ที่ 1.23 ตามตารางที่ 11 ยิ่งไปกว่านั้นผลลัพธ์ของเวลาที่ใช้ในการประมวลผลนั้นลดลงอย่างเห็นได้ชัดจากตารางที่ 10 และ ตารางที่ 13

## บทที่ 7

### สรุปผลงานวิจัย

ในเมืองใหญ่นั้นให้ความสำคัญกับความเป็นอยู่ที่ดีของผู้อยู่อาศัยได้ด้วยการออกแบบผังเมืองและพัฒนาสิ่งที่มีความสำคัญอย่างยั่งยืน คุณภาพชีวิตเป็นดัชนีหนึ่งที่มีความสำคัญและมีประสิทธิภาพที่สามารถใช้ในการวัดผลประสิทธิภาพของผังเมือง ความสามารถในการเดินเป็นสิ่งสำคัญต่อผู้อยู่อาศัยภายในเมือง อีกทั้งการทำแบบสอบถามเป็นวิธีการที่ใช้ต้นทุนและเวลาสูง ซึ่งเป็นข้อจำกัดหนึ่งในการประเมินวัดผลพื้นที่ เพื่อช่วยเหลือในข้อจำกัดนั้นงานวิจัยนี้จึงได้นำเสนอวิธีการใช้เทคโนโลยีปัญญาประดิษฐ์ที่ช่วยในการประเมินวัดผลซึ่งใช้ร่วมกับข้อมูลจากแบบสอบถามที่เก็บรวบรวมด้วยการใช้งานร่วมกับเครื่องมือเทคโนโลยีความจริงเสมือน

อีกทั้งวิธีการที่นำเสนอนี้ได้ใช้โครงข่ายประสาทเทียมเชิงลึกมาช่วยในการถอดคุณลักษณะของรูปภาพและใช้ในการฝึกสอนแบบจำลองทำนายผลอีกด้วย รวมถึงได้นำวิธีการกลั่นความรู้มาช่วยในการลดความซับซ้อนของแบบจำลองที่มีขนาดใหญ่ เพื่อช่วยเหลือในเรื่องของเวลาที่ใช้ในการประมวลผล ซึ่งจากผลการทดลองพบว่าวิธีการกลั่นความรู้ให้ผลเป็นไปตามที่คาดหวังโดยที่แบบจำลองสุดท้ายมีขนาดเล็กและประมวลผลได้รวดเร็วมากยิ่งขึ้นเมื่อเทียบกับแบบจำลองที่ฝึกสอนด้วยวิธีการเรียนรู้เชิงลึกซึ่งแลกมากับความแม่นยำที่ลดลง สุดท้ายนี้วิธีการที่ได้นำเสนอเป็นเพียงวิธีการหนึ่งที่ใช้ในการนำเสนอวิธีการแก้ปัญหาเท่านั้น ซึ่งสามารถใช้วิธีการอื่นมาทดแทนในแต่ละส่วนของงานได้เพื่อแก้ไขปัญหาให้ผลลัพธ์ของการทำนายผลคะแนนคุณภาพชีวิตที่ดียิ่งขึ้น

## บรรณานุกรม

1. Nakamura, K., et al., *Evaluation for Low-carbon Land-use Transport Development with QOL Indexes in Asian Developing Megacities: a Case Study of Bangkok*. Journal of the Eastern Asia Society for Transportation Studies, 2015: p. 1047-1063.
2. Nakamura, K., et al., *Comparative Analysis of QOL in Station Areas between Cities at Different Development Stages, Bangkok and Nagoya*. Transportation Research Procedia, 2017. **25**: p. 3188-3202.
3. Besleme, K. and M. Mullin, *Community indicators and healthy communities*. National Civic Review, 1997. **86**: p. 43+.
4. Diener, E. and E. Suh, *MEASURING QUALITY OF LIFE: ECONOMIC, SOCIAL, AND SUBJECTIVE INDICATORS*. Social Indicators Research, 1997. **40**(1): p. 189-216.
5. Kachi, N., H. Yoshitsugu, and H. Kato, *Making cities more compact by improving transport and amenity and reducing hazard risk*. Journal of the Eastern Asia Society for Transportation Studies, 2008.
6. Briassoulis, H., *Sustainable Development and its Indicators: Through a (Planner's) Glass Darkly*. Journal of Environmental Planning and Management, 2001. **44**(3): p. 409-427.
7. Felce, D. and J. Perry, *Quality of life: its definition and measurement*. Res Dev Disabil, 1995. **16**(1): p. 51-74.
8. Veenhoven, R., *Freedom and happiness: A comparative study in forty-four nations in the early 1990s*, in *Culture and subjective well-being*. 2000, The MIT Press: Cambridge, MA, US. p. 257-288.
9. Hayashi, Y. and I. Sugiyama, *Dual Strategies for the Environmental and Financial Goals of Sustainable Cities: De-Suburbanization and Social Capitalization*. Built Environment (1978-), 2003. **29**(1): p. 8-15.
10. Doi, K., M. Kii, and H. Nakanishi, *An Integrated Evaluation Method of Accessibility, Quality of Life, and Social Interaction*. Environment and Planning B: Planning and Design, 2008. **35**(6): p. 1098-1116.

11. Wasuntarasook, V.H., Y, *A historic review on consequences of critical events leading revolution in Mass Rapid Transit in Bangkok*, in *Proceedings of the 10th International Conference of Eastern Asia Society for Transportation Studies*. 2013: Taipei.
12. Alonso, W., *Location and Land Use*. Toward a General Theory of Land Rent. 1964: Harvard University Press.
13. Yoshitsugu, H., X. Mai, and H. Kato, *The Role of Rail Transport for Sustainable Urban Transport*. 2011. p. 161-174.
14. Banister, D., *The sustainable mobility paradigm*. *Transport Policy*, 2008. **15**(2): p. 73-80.
15. Vichiensan, V. and K. Nakamura *Walkability Perception in Asian Cities: A Comparative Study in Bangkok and Nagoya*. *Sustainability*, 2021. **13**, DOI: 10.3390/su13126825.
16. Nakamura, K., *The relationship between walkability and QOL outcomes in residential evaluation*. *Cities*, 2022. **131**: p. 104008.
17. Vichiensan, V., A. Fukuda, and S. Malaitham. *Perceived Walkability Evaluation in Transit Station Catchment Areas in Bangkok*. in *2022 Moratuwa Engineering Research Conference (MERCOn)*. 2022.
18. Sou, K., et al. *Street Design for Hedonistic Sustainability through AI and Human Co-Operative Evaluation*. *Sustainability*, 2021. **13**, DOI: 10.3390/su13169066.
19. 守田, 賢. and 一. 中村, *歩道境界空間デザインを考慮した VR 歩行空間評価*. *土木学会論文集 D3 (土木計画学)*, 2020. **75**(6): p. I\_137-I\_143.
20. Nakamura, K., *Experimental analysis of walkability evaluation using virtual reality application*. *Environment and Planning B: Urban Analytics and City Science*, 2020. **48**(8): p. 2481-2496.
21. Kantavat, P., et al. *Transportation Mobility Factor Extraction Using Image Recognition Techniques*. in *2019 First International Conference on Smart Technology & Urban Development (STUD)*. 2019.
22. Thitisiriwech, K., et al., *The Bangkok Urbanscapes Dataset for Semantic Urban*

- Scene Understanding Using Enhanced Encoder-Decoder With Atrous Depthwise Separable A1 Convolutional Neural Networks*. IEEE Access, 2022. **10**: p. 59327-59349.
23. Thitisiriwech, K., et al. *Quality of Life Prediction in Driving Scenes on Thailand Roads Using Information Extraction from Deep Convolutional Neural Networks*. Sustainability, 2023. **15**, DOI: 10.3390/su15032847.
  24. *Object Detection Guide*. 2021; Available from: <https://www.fritz.ai/object-detection/>.
  25. Zhou, B., et al., *Semantic understanding of scenes through the ade20k dataset*. International Journal on Computer Vision, 2018.
  26. Alom, M.Z., et al. *A State-of-the-Art Survey on Deep Learning Theory and Architectures*. Electronics, 2019. **8**, DOI: 10.3390/electronics8030292.
  27. Afshine, A.S., Amidi. *Recurrent Neural Networks cheatsheet*. Available from: <https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-recurrent-neural-networks>.
  28. Hinton, G., O. Vinyals, and J. Dean, *Distilling the knowledge in a neural network*. arXiv preprint arXiv:1503.02531, 2015.
  29. lamtrakul, P., et al. *Exploring the Spatial Effects of Built Environment on Quality of Life Related Transportation by Integrating GIS and Deep Learning Approaches*. Sustainability, 2023. **15**, DOI: 10.3390/su15032785.
  30. Futami, H., et al., *Distilling the knowledge of BERT for sequence-to-sequence ASR*. arXiv preprint arXiv:2008.03822, 2020.
  31. Bochkovskiy, A., C.-Y. Wang, and H.-Y.M. Liao, *Yolov4: Optimal speed and accuracy of object detection*. arXiv preprint arXiv:2004.10934, 2020.
  32. Hong, Y., et al., *Deep dual-resolution networks for real-time and accurate semantic segmentation of road scenes*. arXiv preprint arXiv:2101.06085, 2021.



จุฬาลงกรณ์มหาวิทยาลัย  
**CHULALONGKORN UNIVERSITY**



## ประวัติผู้เขียน

ชื่อ-สกุล	ชนสิทธิ์ ฤทธิธน์โสภณ
วัน เดือน ปี เกิด	23 กันยายน 2538
สถานที่เกิด	กรุงเทพมหานคร
วุฒิการศึกษา	มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขต กำแพงแสน
ที่อยู่ปัจจุบัน	234/23 ตรอกวัดราชสิทธิาราม แขวงวัดท่าพระ เขตบางกอกใหญ่ กรุงเทพมหานคร 10600



จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY