สารบัญ

หน้า

บทคัดย่อ ..... ก

กิตติกรรมประกาศ ..... ข

สารบัญตาราง ..... จ

สารบัญภาพประกอบ ..... ฉ

บทที่ 1 บทนำ 1

1.1 หลักการและเหตุผล 1
1.2 วัตถุประสงค์ของโครงงาน 1
1.3 ขอบเขตของโครงงาน 1
1.4 ภาพรวมของระบบ 3
1.5 ประโยชน์ที่คาดว่าจะได้รับ 6
1.6 อุปกรณ์และเครื่องมือที่ใช้ในการดำเนินงาน 7
1.7 แผนการดำเนินงาน 7

บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง 9

2.1 ทฤษฎีที่เกี่ยวข้อง 9
2.2 ระบบงานที่เกี่ยวข้อง 14
2.3 ตารางเปรียบเทียบ 15

บทที่ 3 ขั้นตอนการดำเนินงาน 17

3.1 กรอบการดำเนินงาน 17
3.2 การเตรียมชุดข้อมูล (Dataset) 18
3.3 ขั้นตอนการทำงานของ YOLO (You Only Look Once) 26
3.4 Detection 32
3.5 การระบุตำแหน่งวัตถุ 57
3.6 ขั้นตอนในการพัฒนา Mobile application 62
3.7 ขั้นตอนในการพัฒนา Household Model 80

บทที่ 4 ผลการทดลอง 83

Computer Science Department
Faculty of Informatics, Mahasarakham University

สารบัญ (ต่อ)

หน้า

4.1 การประเมินประสิทธิภาพโมเดล 83
4.2 ตัวอย่างการตรวจจับบน Mobile Application 91
4.3 สรุปและวิเคราะห์ผลการทดลอง 92
บทที่ 5 สรุปและอภิปรายผลการทดลอง 94
5.1 สรุปผลและอภิปรายผล 94
5.2 ปัญหาและอุปสรรคในการดำเนินงาน 94
5.3 ข้อเสนอแนะ 94
เอกสารอ้างอิง 95
ภาคผนวก 97
ภาคผนวก ก ตัวอย่าง Dataset 98
ภาคผนวก ข คู่มือการใช้งาน Mobile Application 103
บทความวิจัย 108
โปสเตอร์โครงงาน 123
ประวัติย่อผู้จัดทำโครงงาน 125

Computer Science Department
Faculty of Informatics, Mahasarakham University

สารบัญตาราง

หน้า

ตารางที่ 1.1 แผนการดำเนินงาน 8
ตารางที่ 2.1 ตารางเปรียบเทียบการทำงานของระบบ 15
ตารางที่ 3.1 รายละเอียดข้อมูลภาพ 19
ตารางที่ 4.1 ผลการประเมินประสิทธิภาพสถาปัตยกรรม YOLOv5n 84
ตารางที่ 4.2 ตัวอย่างผลการตรวจจับ 86
ตารางที่ 4.3 คลาสที่มีค่า mAP50 ต่ำกว่า 0.5 92
ตารางที่ ก-1 ตัวอย่าง dataset สิ่งของแต่ละอย่าง 99

Computer Science Department
Faculty of Informatics, Mahasarakham University

สารบัญภาพประกอบ

หน้า

ภาพประกอบที่ 1.1 การทำงานแอปพลิเคชัน 3
ภาพประกอบที่ 1.2 การหาตำแหน่งในรูปภาพ 5
ภาพประกอบที่ 1.3 หน้าโหมดสำรวจ 5
ภาพประกอบที่ 1.4 หน้าโหมดการค้นหา 6
ภาพประกอบที่ 2.1 หลักการการทำงาน Objective Detection 9
ภาพประกอบที่ 2.2 การทำงานของอัลกอริทึม YOLO 10
ภาพประกอบที่ 2.3 การทำงาน Anchor box 10
ภาพประกอบที่ 2.4 การทำงาน Non-max Suppression 10
ภาพประกอบที่ 2.5 ตัวอย่างแอปพลิเคชันที่สร้างโดยใช้ Flutter 11
ภาพประกอบที่ 2.6 รูปร่างเทนเซอร์ 12
ภาพประกอบที่ 2.7 กราฟที่มีการเชื่อมเทนเซอร์เข้าด้วยกัน 12
ภาพประกอบที่ 2.8 ตัวอย่างแอปพลิเคชัน Lookout 14
ภาพประกอบที่ 2.9 ตัวอย่างแอปพลิเคชัน Supersense 15
ภาพประกอบที่ 3.1 Dataset 18
ภาพประกอบที่ 3.2 Dataset test 18
ภาพประกอบที่ 3.3 การตีกรอบภาพสิ่งของ 21
ภาพประกอบที่ 3.4 การตีกรอบแบบ Polygon annotation 21
ภาพประกอบที่ 3.5 การเพิ่มปริมาณ dataset 22
ภาพประกอบที่ 3.6 dataset 22
ภาพประกอบที่ 3.7 ตัวอย่างไฟล์ images 23
ภาพประกอบที่ 3.8 ตัวอย่างไฟล์ labels 24
ภาพประกอบที่ 3.9 ตัวอย่างรูปภาพ 24
ภาพประกอบที่ 3.10 ค่าในไฟล์ label 24
ภาพประกอบที่ 3.11 ตัวอย่างรูปภาพ 25
ภาพประกอบที่ 3.12 ค่าในไฟล์ label 25
ภาพประกอบที่ 3.13 การทำ Convolution ขนาด 5 x 5 และ filter ขนาด 3 x 3 28
ภาพประกอบที่ 3.14 ตัวอย่างการทำ RELU 29
ภาพประกอบที่ 3.15 ตัวอย่างการทำ Max Pooling Layer 30
ภาพประกอบที่ 3.16 ตัวอย่างการทำ Flatten 31

Computer Science Department
Faculty of Informatics, Mahasarakham University

สารบัญภาพประกอบ (ต่อ)

หน้า

ภาพประกอบที่ 3.17 ตัวอย่างการทำงาน Fully-Connected Layer....31
ภาพประกอบที่ 3.18 ตัวอย่างผลลัพธ์การทำ SoftMax....32
ภาพประกอบที่ 3.19 ตัวอย่างการแบ่งตาราง 3 x 3....33
ภาพประกอบที่ 3.20 ภาพตัวอย่างที่นำมาทำ k-mean....34
ภาพประกอบที่ 3.21 ตัวอย่างผลลัพธ์การจัดกลุ่มข้อมูลตัวอย่างที่ 1....35
ภาพประกอบที่ 3.22 ตัวอย่างผลลัพธ์การจัดกลุ่มข้อมูลตัวอย่างที่ 2....35
ภาพประกอบที่ 3.23 ตัวอย่างผลลัพธ์การจัดกลุ่มข้อมูล....36
ภาพประกอบที่ 3.24 ตัวอย่างผลลัพธ์การกำหนดขนาด anchor โดยใช้ k-mean....36
ภาพประกอบที่ 3.25 การแบ่งกลุ่มข้อมูลในแต่ละช่อง โดยแต่ละช่องจะมี 2 anchor....36
ภาพประกอบที่ 3.26 นำแต่ละ bounding box สีน้ำเงินสกัด Features ด้วย CNN และเข้าโมเดล....37
ภาพประกอบที่ 3.27 นำแต่ละ bounding box สกัด Features ด้วย CNN และเข้าโมเดล....37
ภาพประกอบที่ 3.28 ตัวอย่างข้อมูลของกรอบสีน้ำเงินและผลเฉลย....38
ภาพประกอบที่ 3.29 areaOverlap....38
ภาพประกอบที่ 3.30 ตัวอย่างหลังจากการหาค่า probity ทั้งหมด....39
ภาพประกอบที่ 3.31 ตัดข้อมูลที่มีความน่าจะเป็นน้อยกว่าที่กำหนดออก....40
ภาพประกอบที่ 3.32 ตัวอย่าง grid ที่ 5....41
ภาพประกอบที่ 3.33 ตัวอย่างข้อมูล bounding box สีน้ำเงิน และสีแดง....41
ภาพประกอบที่ 3.34 intersection_area ของ bounding box สีน้ำเงินและสีแดง....42
ภาพประกอบที่ 3.35 ตัวอย่าง bounding box ที่มีความน่าจะเป็นสูงสุด....43
ภาพประกอบที่ 3.36 ตัวอย่างการแบ่งตาราง 3x3....43
ภาพประกอบที่ 3.37 ภาพตัวอย่างที่นำมาทำ k-mean grid ที่ 2....45
ภาพประกอบที่ 3.38 ตัวอย่างการจัดกลุ่มข้อมูลตัวอย่างที่ 1....45
ภาพประกอบที่ 3.39 ตัวอย่างผลลัพธ์การจัดกลุ่มข้อมูลตัวอย่างที่ 2....46
ภาพประกอบที่ 3.40 ตัวอย่างผลลัพธ์การจัดกลุ่มข้อมูล grid ที่ 2....46
ภาพประกอบที่ 3.41 ตัวอย่างผลลัพธ์การกำหนดขนาด anchor โดยใช้ k-mean....46
ภาพประกอบที่ 3.42 ภาพตัวอย่างที่นำมาทำ k-mean grid ที่ 7....47
ภาพประกอบที่ 3.43 ตัวอย่างผลลัพธ์การจัดกลุ่มข้อมูลตัวอย่างที่ 2....47
ภาพประกอบที่ 3.44 ตัวอย่างผลลัพธ์การจัดกลุ่มข้อมูล grid ที่ 7....48
ภาพประกอบที่ 3.45 ตัวอย่างผลลัพธ์การกำหนดขนาด anchor โดยใช้ k-mean....48

Computer Science Department
Faculty of Informatics, Mahasarakham University

สารบัญภาพประกอบ (ต่อ)

หน้า

ภาพประกอบที่ 3.46 การแบ่งกลุ่มข้อมูลในแต่ละ grid โดยแต่ละ grid จะมี 2 anchor....48
ภาพประกอบที่ 3.47 นำแต่ละ bounding box สีน้ำเงิน สกัด Features ด้วย CNN และเข้าโมเดล...49
ภาพประกอบที่ 3.48 นำแต่ละ bounding box สีแดง สกัด Features ด้วย CNN และเข้าโมเดล 49
ภาพประกอบที่ 3.49 นำแต่ละ bounding box สีน้ำเงิน สกัด Features ด้วย CNN และเข้าโมเดล...50
ภาพประกอบที่ 3.50 นำแต่ละ bounding box สีแดง สกัด Features ด้วย CNN และเข้าโมเดล 50
ภาพประกอบที่ 3.51 ภาพตัวอย่างการคำนวณ IOU....51
ภาพประกอบที่ 3.52 ตัวอย่างข้อมูลของ bounding box สีน้ำเงินและผลเฉลย....51
ภาพประกอบที่ 3.53 intersection_area....52
ภาพประกอบที่ 3.54 ตัวอย่างหลังจากการหาค่า probity ทั้งหมด....53
ภาพประกอบที่ 3.55 ตัดข้อมูลที่มีความน่าจะเป็นน้อยกว่าที่กำหนดออก....54
ภาพประกอบที่ 3.56 ตัวอย่าง grid ที่ 2....54
ภาพประกอบที่ 3.57 ตัวอย่างข้อมูลของ bounding box สีน้ำเงินและ bounding box สีแดง....55
ภาพประกอบที่ 3.58 intersection_area ของ bounding box สีน้ำเงินและ bounding box สีแดง55
ภาพประกอบที่ 3.59 ตัวอย่าง bounding box ที่มีความน่าจะเป็นสูงสุด....56
ภาพประกอบที่ 3.60 ตัวอย่างผลลัพธ์ non-max suppression....56
ภาพประกอบที่ 3.61 การระบุตำแหน่งวัตถุ....57
ภาพประกอบที่ 3.62 ตัวอย่างการคำนวณหาตำแหน่ง....59
ภาพประกอบที่ 3.63 ตัวอย่างการคำนวณหาตำแหน่งหากวัตถุอยู่ทั้ง 3 ช่อง ในแนวตั้ง....59
ภาพประกอบที่ 3.64 ตัวอย่างการคำนวณหาตำแหน่งหากวัตถุอยู่ทั้ง 6 ช่อง ในแนวตั้ง....60
ภาพประกอบที่ 3.65 ตัวอย่างการคำนวณหาตำแหน่งหากวัตถุอยู่ทั้ง 9 ช่อง ในแนวตั้ง....60
ภาพประกอบที่ 3.66 ตัวอย่างการคำนวณหาตำแหน่งหากวัตถุอยู่ทั้ง 3 ช่อง ในแนวนอน....61
ภาพประกอบที่ 3.67 ตัวอย่างการคำนวณหาตำแหน่งหากวัตถุอยู่ทั้ง 6 ช่อง ในแนวนอน....61
ภาพประกอบที่ 3.68 ตัวอย่างการคำนวณหาตำแหน่งหากวัตถุอยู่ทั้ง 9 ช่อง ในแนวนอน....62
ภาพประกอบที่ 3.69 การเก็บไฟล์สำหรับ model....62
ภาพประกอบที่ 3.70 รายละเอียดไฟล์ labelmap.txt....63
ภาพประกอบที่ 3.71 การเพิ่มชื่อโฟลเดอร์ใน pubspec....63
ภาพประกอบที่ 3.72 import library tflite....63
ภาพประกอบที่ 3.73 ประกาศตัวแปร MODEL_FILE_NAME และLABEL_FILE_NAME....63
ภาพประกอบที่ 3.74 function สำหรับ load model....64

Computer Science Department
Faculty of Informatics, Mahasarakham University

สารบัญภาพประกอบ (ต่อ)

หน้า

ภาพประกอบที่ 3.75 function สำหรับ load label 64
ภาพประกอบที่ 3.76 function ปรับขนาดและประมวลผลรูปภาพ....65
ภาพประกอบที่ 3.77 บันทึกเวลาเริ่มต้นของกระบวนการทำนายผล....65
ภาพประกอบที่ 3.78 ตรวจสอบ Interpreter....65
ภาพประกอบที่ 3.79 แปลง image เป็น TensorImage และ ปรับขนาดและประมวลผลภาพ....65
ภาพประกอบที่ 3.80 สร้าง TensorBuffer สำหรับเก็บผลลัพธ์ของ output tensors....66
ภาพประกอบที่ 3.81 สร้างรายการ inputs และแผนที่ outputs....66
ภาพประกอบที่ 3.82 ทำนายผลจากโมเดล และเก็บผลลัพธ์....66
ภาพประกอบที่ 3.83 คำนวณจำนวนผลลัพธ์ และกำหนดค่า labelOffset....66
ภาพประกอบที่ 3.84 แปลง TensorBuffer เป็น List....67
ภาพประกอบที่ 3.85 ค้นหาผลลัพธ์ และเพิ่มลงในรายการ recognitions....67
ภาพประกอบที่ 3.86 class TextToSpeech....68
ภาพประกอบที่ 3.87 สร้าง Map เก็บคำศัพท์ในการทำ text-to-speech....69
ภาพประกอบที่ 3.88 สร้าง instance ของ SpeechToText และตัวแปร....69
ภาพประกอบที่ 3.89 สร้าง State ใน StatefulWidget และเรียกใช้งาน TextToSpeech....70
ภาพประกอบที่ 3.90 สร้างปุ่มที่เรียกใช้งานการจับเสียงพูด Speech to Text....70
ภาพประกอบที่ 3.91 ตรวจสอบคำที่ถูกจับเป็นเสียง....71
ภาพประกอบที่ 3.92 หยุดการรับเสียง และหยุดการทำงาน Speech to Text....71
ภาพประกอบที่ 3.93 เปรียบเทียบข้อมูลวัตถุใหม่ กับข้อมูลวัตถุเดิม....71
ภาพประกอบที่ 3.94 ประกาศตัวแปรภายในคลาส....72
ภาพประกอบที่ 3.95 Method renderLocation....73
ภาพประกอบที่ 3.96 override Method....74
ภาพประกอบที่ 3.97 ประกาศตัวแปร checkText และ recognitionLabel....74
ภาพประกอบที่ 3.98 กำหนดค่าต่างๆที่ใช้ในการคำนวณระบุตำแหน่ง....75
ภาพประกอบที่ 3.99 นำ checkText และ recognitionLabel มาตรวจสอบ....76
ภาพประกอบที่ 3.100 โค้ดภายในหน้าหลัก Application....77
ภาพประกอบที่ 3.101 โค้ดภายในหน้าหลัก Application....78
ภาพประกอบที่ 3.102 clone yolov5....80
ภาพประกอบที่ 3.103 สร้าง Virtual Environment....80

Computer Science Department
Faculty of Informatics, Mahasarakham University

สารบัญภาพประกอบ (ต่อ)

หน้า

ภาพประกอบที่ 3.104 ติดตั้ง requirements 80
ภาพประกอบที่ 3.105 กำหนดพารามิเตอร์และการตั้งค่าชุดข้อมูล....81
ภาพประกอบที่ 3.106 script ที่ใช้ในการ train....82
ภาพประกอบที่ 4.1 ผลการประเมินประสิทธิภาพสถาปัตยกรรม YOLOv5n....84
ภาพประกอบที่ 4.2 การประเมินด้วย confusion matrix ของ YOLOv5n....86
ภาพประกอบที่ 4.3 ตัวอย่างการตรวจจับบน Mobile Application ในโหมด Explore....91
ภาพประกอบที่ 4.4 ตัวอย่างการตรวจจับบน Mobile Application ในโหมด Find....92
ภาพประกอบที่ ข-1 หน้าแรกของ Application เป็นหน้าโหมดการสำรวจ(Explore).... 104
ภาพประกอบที่ ข-2 โหมดค้นหา (Find).... 105
ภาพประกอบที่ ข-3 วิธีเปลี่ยนโหมด.... 106
ภาพประกอบที่ ข-4 การใช้โหมดค้นหา (Find).... 107

Computer Science Department
Faculty of Informatics, Mahasarakham University