



การใช้เทคโนโลยีสารสนเทศร่วมกับปัญญาประดิษฐ์ในการตรวจสอบการเคลื่อนไหวและติดตามเชิงพื้นที่:
การปรับใช้ในฉากทัศน์การแพร่ระบาดของโรคติดต่อโควิด-19

EXPLOITING INDOOR GEOSPATIAL INFORMATION TECHNOLOGY AND ARTIFICIAL
INTELLIGENCE FOR MOVEMENT MONITORING AND SPATIAL TRACKING: AN ADAPTIVE
RESPONSE TO COVID-19 PANDEMIC SCENARIOS

อนวัช เกริญเกษม¹, เทพนรินทร์ นาหนองตูม¹, พงศกร พรณรัตน์ศิลป์², ชาติชาย ไวยสุระสิงห์^{2*} และชุตินา ไวยสุระสิงห์³

¹นักศึกษาหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต, สาขาวิชาวิศวกรรมโยธา คณะวิศวกรรมศาสตร์ มหาวิทยาลัยขอนแก่น

²รองศาสตราจารย์, ศูนย์วิจัยและพัฒนาโครงสร้างมูลฐานอย่างยั่งยืน คณะวิศวกรรมศาสตร์ มหาวิทยาลัยขอนแก่น

³อาจารย์, คณะศิลปศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

*Corresponding Author, E-Mail: fcecw@kku.ac.th

บทคัดย่อ

ความก้าวหน้าของระบบปัญญาประดิษฐ์ในการตรวจติดตามการเคลื่อนไหวของวัตถุใด ๆ ผ่านกล้องวงจรปิดที่มีอยู่ในปัจจุบันเมื่อนำมาผสมผสานเข้ากับเทคโนโลยีสารสนเทศนับเป็นประเด็นที่น่าสนใจและมีความสำคัญอย่างยิ่งต่อการวิศวกรรมสำรวจที่สามารถใช้ในการบริหารจัดการความปลอดภัยทั้งในมิติของทรัพย์สินและสุขภาพ โดยเฉพาะอย่างยิ่งในช่วงเวลาที่มีการแพร่กระจายของเชื้อโควิด-19 ซึ่งเป็นการแพร่ระบาดที่ก่อให้เกิดความเสียหายต่อมวลมนุษยชาติอย่างมหาศาล แนวทางหนึ่งในการป้องกันคือการรักษาระยะห่างระหว่างบุคคลที่ควรอยู่ห่างกัน 2 เมตร ดังนั้น การระบุตำแหน่งของบุคคลในสถานที่ต่าง ๆ จึงเป็นสิ่งสำคัญในการช่วยการสืบสวน โรค สิ่งนี้นับเป็นโจทย์วิจัยที่สำคัญในการนำเอากล้องวงจรปิดมาผสมผสานกับเทคโนโลยีสารสนเทศภายในอาคารร่วมกับการใช้ระบบปัญญาประดิษฐ์ในการตรวจติดตามการเคลื่อนไหวมาใช้ในการระบุตำแหน่ง และสิ่งสำคัญที่เข้ามามีบทบาทในการเชื่อมโยงเทคโนโลยีเหล่านี้เข้าไว้ด้วยกันคือ การแปลงค่าพิกัด จากที่กล่าวมาทั้งหมดนี้งานวิจัยนี้จึงมีวัตถุประสงค์ในการแปลงค่าพิกัดแบบสมการสมมาตรร่วมกับระบบปัญญาประดิษฐ์ในการตรวจติดตามการเคลื่อนไหวของบุคคลสำหรับใช้ในการกำหนดตำแหน่งในระบบพิกัดของห้องภายในอาคาร โดยงานวิจัยนี้ได้มีการวิเคราะห์การเชื่อมโยงกล้องที่มีจำนวนมากกว่า 1 ตัว ผลการศึกษาได้แสดงให้เห็นถึงตำแหน่งของบุคคลที่ได้จากการประยุกต์ใช้ระบบปัญญาประดิษฐ์ในการตรวจติดตามการเคลื่อนไหวมาใช้ในการระบุตำแหน่งควบคู่กับการแปลงค่าพิกัดแบบสมการสมมาตร โดยมีประสิทธิภาพในการระบุตำแหน่งที่ความคลาดเคลื่อนไม่เกิน 0.5 เมตร ซึ่งเป็นระดับความคลาดเคลื่อนที่อยู่ในเกณฑ์ที่ยอมรับได้เนื่องจากมีขนาดน้อยกว่าเส้นผ่านศูนย์กลางของบุคคลทั่วไป และถึงแม้ว่าในปัจจุบันโรคโควิด-19 จะมีการบรรเทามาตรการรักษาระยะห่างระหว่างบุคคลลงไปแล้วแต่ผลการศึกษาได้ชี้ให้เห็นว่า ในอนาคตหากมีการระบาดของโรคอื่น ๆ ที่ต้องมีการบังคับใช้มาตรการรักษาระยะห่างเช่นเดียวกันนี้ ผลการศึกษาจากการศึกษาวิจัยนี้ก็นับได้ว่าเป็นผลการศึกษาที่สามารถนำไปปรับใช้ได้เลยในทันที

คำสำคัญ: การตรวจจับวัตถุ; โควิด-19; การรักษาระยะห่างทางสังคม; ภูมิสารสนเทศภายในอาคาร

ABSTRACT

The advancement of artificial intelligence systems in monitoring the movements of any object using current closed-circuit cameras in conjunction with geospatial information technology is an intriguing and essential topic in surveying engineering. It can be used to manage property and health security, particularly during the COVID-19 pandemic, which has caused significant injury to humanity. The concept of "social distancing" refers to maintaining a distance of more than 2 meters between individuals as a prevention method. Disease surveillance necessitates, therefore, the precise identification of the location of individuals in various locations. This research seeks to incorporate closed-circuit television cameras with geospatial information technology within structures, using artificial intelligence systems to monitor and identify movements and locations. This integration's primary function is coordinate transformation. Based on the preceding, this research aims to convert coordinates using an affine transformation equation and an artificial intelligence system to track individuals' movements for positioning within the indoor coordinate system. The research analyzed the interconnection of multiple cameras. It demonstrated the position of individuals derived by applying artificial intelligence systems for motion tracking and converting coordinates using a joint equation system. The accuracy of position identification was attained within a tolerance of 0.5 meters, which is an acceptable error level because it is smaller than the average diameter of a person. Even though the current COVID-19 pandemic has necessitated the enforcement of physical separation, the study's results indicated that if another disease outbreak necessitating similar separation measures occurs in the future, the findings of this research can be applied immediately.

KEYWORDS: Object Detection; COVID-19; Social Distancing; In-door geographic information systems

1. บทนำ

1.1 ที่มาและความสำคัญของปัญหา

การบูรณาการระบบปัญญาประดิษฐ์เข้ากับระบบภูมิสารสนเทศที่เป็นการพัฒนาระบบตรวจติดตามการเคลื่อนไหวของวัตถุใด ๆ ผ่านกล้องวงจรปิด (Closed-Circuit Television: CCTV) แล้วเชื่อมโยงตำแหน่งเข้ากับระบบภูมิสารสนเทศภายในอาคาร นับเป็นอีกหนึ่งประเด็นที่กำลังเป็นที่สนใจอย่างยิ่ง โดยเฉพาะอย่างยิ่งภาวะการระบาดของโรคติดเชื้อไวรัสโคโรนา 2019 หรือโรคโควิด-19 (Coronavirus Disease 2019: COVID-19) ของมณฑลอุฮั่นที่เริ่มต้นเมื่อปลายปี ค.ศ. 2019 และมีการกระจายไปทั่วโลกภายในระยะเวลาไม่ถึง 3 เดือน โดยวันที่ 11 มีนาคม ค.ศ. 2020 องค์การอนามัยโลก (World Health Organization [WHO]) ออกแถลงการถึงสถานะการระบาดของโรคโควิด-19 ที่นับเป็นการระบาดครั้งใหญ่ที่สุดของมวลมนุษยชาติในระดับรุนแรง (Pandemic) [1] หลังจากนั้นจึงเริ่มมีการมาตรการการรักษาระยะห่างทางสังคม ซึ่งหมายถึงการสร้างระยะห่างระหว่างตัวเราเองกับคนอื่น ๆ ในสังคม รวมถึงการลดการออกไปนอกบ้านโดยไม่จำเป็น หลีกเลี่ยงการใช้ขนส่งสาธารณะ การไม่เข้าร่วมกิจกรรมร่วมกับผู้อื่น และการทำงานอยู่ที่บ้าน เพื่อช่วยลดอัตราการแพร่ระบาดของไวรัสจากคนสู่คน ประธานสาขาวิทยาศาสตร์สุขภาพประชากร (Population Health Sciences) จาก Georgia State University ได้เผยว่า ทุก ๆ การลดจำนวนการติดต่อกันระหว่างคนต่อวัน ไม่ว่าจะเป็นการติดต่อกันระหว่างเครือญาติ เพื่อน เพื่อนร่วมงาน หรือที่โรงเรียนนั้นจะช่วยลดการแพร่เชื้อไวรัสในสังคมเป็นอย่างมาก และการใช้มาตรการ Social Distancing นั้นยังถูกนำมาใช้ในเหตุการณ์การระบาดทั่วโลกในหลายๆ ครั้ง ตั้งแต่การระบาดของไข้หวัดใหญ่สเปน (Spanish Flu) ในระหว่างปี ค.ศ. 1918 ถึง ค.ศ. 1920 โดยมีอัตราการเสียชีวิตสูงถึง 50-100 ล้านคนทั่วโลก และไข้หวัดใหญ่สายพันธุ์ใหม่ 2009 ที่ระบาดจากประเทศเม็กซิโกไปยังทั่วโลก ซึ่งการทำ Social Distancing นั้นก็ได้พิสูจน์ว่ามาตรการป้องกันการ

ระบอบแบบเว้นระยะห่างนั้นมีประสิทธิภาพ และสามารถช่วยทุเลาระดับความรุนแรงของการระบาดได้จริง [2] โดยการเว้นระยะห่างทางสังคมนั้นแนะนำให้ห่างจากคนอื่น 6 ฟุต หรือ 1.83 เมตรจากบุคคลอื่น [3] แต่ทางที่ดีที่สุดคือการอยู่ให้ห่างกันมากที่สุด เนื่องจากเชื้อไวรัสโคโรนาเมื่อแรงลมเพิ่มขึ้นจะสามารถแพร่กระจายได้ถึง 6 เมตร [4] การศึกษาจาก WHO พบว่าความเสี่ยงจากการติดเชื้อจะลดลงอย่างมากเมื่อห่างจากผู้ป่วย 1 เมตรขึ้นไป [5] ซึ่งในเวลาต่อมาได้มีงานวิจัย [6] เกิดขึ้นและได้เสนอแนะการใช้กล้องวงจรปิด CCTV เพื่อพิจารณาการเว้นระยะห่างทางสังคมผ่านการระบุตำแหน่งของบุคคลต่าง ๆ ภายในกล้อง CCTV ซึ่งนับเป็นเทคโนโลยีที่น่าสนใจและมีความเป็นไปได้ในการช่วยควบคุมและลดการแพร่กระจายของเชื้อโรคได้แล้ว ยิ่งไปกว่านั้นยังถือเป็นต้นแบบเทคโนโลยีที่สามารถบรรเทาความเสียหายที่จะเกิดขึ้นจากทางเศรษฐกิจให้ลดขนาดลงได้อย่างมาก ทั้งนี้เพราะหากมีเทคโนโลยีที่ช่วยระบุตำแหน่งหรือเส้นทางการเดินของผู้ที่มีความเสี่ยงต่อการติดเชื้อโรคโควิด-19 ที่ดีพอ อาทิเช่น ในวันที่ตรวจพบว่า มีผู้ที่มีความเสี่ยงต่อการติดเชื้อโรคโควิด-19 ฝ่าฝืนมาตรการในการควบคุมโรคติดต่อโดยแอบไปเที่ยวห้างสรรพสินค้าที่จังหวัดระยองตามที่ปรากฏเป็นข่าวในช่วงกลางเดือนกรกฎาคม ค.ศ. 2020 [7] ห้างสรรพสินค้านั้นอาจจะไม่จำเป็นต้องปิดห้างลงเพียงเพราะบุคคลที่เป็นบุคคลเสี่ยงเข้ามาเพียงแค่บริเวณใดบริเวณหนึ่งของห้างสรรพสินค้าก็เป็นได้ โดยเฉพาะอย่างยิ่งถ้าหากมีเทคโนโลยีที่สามารถระบุตำแหน่งและเส้นทางการเดินของผู้ที่มีความเสี่ยงต่อการติดเชื้อโรคโควิด-19 โดยจุดเด่นของงานศึกษา [6] คือ การใช้เทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligent: AI) ที่ในปัจจุบันนั้นสามารถจะนำมาใช้ได้อย่างง่ายดายรวมถึงการใช้กล้องวงจรปิดในการเฝ้าระวังพื้นที่ต่าง ๆ นั้นเป็นเรื่องที่พบเห็นได้ทั่วไปในมิติของการรักษาความปลอดภัยในพื้นที่ต่าง ๆ ซึ่งนำมาประยุกต์ใช้ได้ง่าย โดยได้มีการจับภาพบุคคลผ่านการใช้อุปกรณ์ปัญญาประดิษฐ์และมีการระบุตำแหน่งของบุคคล งานศึกษาชิ้นนี้ถือเป็นงานที่น่าสนใจอย่างยิ่ง อย่างไรก็ตาม การศึกษาครั้งนี้จะมีประสิทธิภาพอย่างมากหากมีการนำมาผสมผสานกับเทคโนโลยีด้านภูมิสารสนเทศภายในอาคาร (Indoor Geographic Information System: IGIS) พร้อมกับการระบุแนวเส้นทางการเดิน ด้วยเหตุนี้เองการพัฒนางานวิจัยที่ช่วยสนับสนุนแนวคิดดังกล่าวจึงเป็นสิ่งจำเป็น ซึ่งผู้วิจัยได้เล็งเห็นถึงความสำคัญของการพัฒนาเทคโนโลยีฉลาด (Smart technology) ผ่านการพัฒนาต่อยอดงานวิจัยในมิติของการเชื่อมต่อเข้ากับระบบ IGIS อันจะเป็นเครื่องมือสำคัญที่ช่วยสนับสนุนมาตรการการป้องกันการระบาดของโรคโควิด-19 ได้ต่อไป

1.2 วัตถุประสงค์การวิจัย

เพื่อศึกษาระบบการแปลงค่าพิกัดที่ใช้ในการเชื่อมโยงและระบุตำแหน่งจากระบบการค้นคืนวิดีโอเข้ากับเทคโนโลยีด้านภูมิสารสนเทศภายในอาคาร (Indoor Geographic Information System: IGIS) พร้อมกับการระบุแนวเส้นทางการเดิน สำหรับใช้เป็นเทคโนโลยีเสริมในการติดตามบุคคลเพื่อสนับสนุนมาตรการการป้องกันการระบาดของโรคโควิด-19

1.3 พื้นที่สำหรับศึกษาวิจัยและอาสาสมัคร

การศึกษานี้อาศัยสถานการณ์จำลองการเดินทางและใช้ชีวิตภายในศูนย์วิจัยและพัฒนาโครงสร้างมูลฐานอย่างยั่งยืนมหาวิทยาลัยขอนแก่น ผ่านการจำลองสถานการณ์ด้วยการใช้บุคคลที่เป็นคณะผู้วิจัยซึ่งได้มีพันธบัตรความนี้ขึ้นมา ทำให้มั่นใจได้ว่าเก็บข้อมูลสำหรับการวิจัยเฉพาะบุคคลที่ได้ผ่านการยินยอมเท่านั้น รวมทั้งป้องกันการละเมิดข้อมูลส่วนบุคคลจากบุคคลที่ไม่ได้รับการยินยอม โดยบุคคลดังกล่าวนี้ถือเป็นอาสาสมัครที่ร่วมออกแบบงานวิจัย และต้องได้รับการชี้แจงและขอความยินยอมกับอาสาสมัครทุกคนโดยอาสาสมัครมีสิทธิ์ถอนตัวออกจากโครงการนี้เมื่อใดก็ได้ และข้อมูลของอาสาสมัครทั้งหมดจะถูกเก็บไว้เป็นความลับ รวมทั้งทำลายทิ้งหลังเสร็จสิ้นการวิจัย

2. ทบทวนวรรณกรรมการระบุตำแหน่งจากระบบการสืบค้นวิดีโอเข้ากับเทคโนโลยีด้านภูมิสารสนเทศ

ในปี ค.ศ. 2019 ได้มีการศึกษาของ Shirsat [8] ได้ทำการศึกษาระบบตรวจจับอาชญากรรมด้วยกล้อง CCTV โดยใช้ระบบตรวจจับใบหน้าผ่านการใช้โปรแกรม OpenCV ซึ่งยังมีปัญหาเรื่องการคุณภาพของความแม่นยำและค่าใช้จ่าย เนื่องจากหากต้องการความแม่นยำที่สูงจำเป็นต้องมีค่าใช้จ่ายที่สูงมากเช่นกัน อย่างไรก็ตาม งานวิจัยของ Shirsat [8] นับเป็นปฐมบทแห่งการเปิดทางการใช้ AI ผสานเข้ากับกล้อง CCTV ที่เริ่มมีความน่าสนใจในขณะนั้น ต่อมาในปี ค.ศ. 2021 จึงมีการศึกษาของ Yang [6] ที่ค้นพบว่าการระบาดของเชื้อไวรัสโคโรนา 2019 ทำให้การรักษาระยะห่างทางสังคมกลับมาเป็นบทบาทอีกครั้งโดยการรักษาระยะห่างในสถานที่สาธารณะ อย่างไรก็ตามการรักษาระยะห่างนั้นมีแนวโน้มที่จะถูกฝ่าฝืนเนื่องจากความไม่คุ้นเคยในการติดตอสื่อสารที่ระยะห่างเกินกว่า 2 เมตร งานศึกษาของ Yang [6] จึงได้ใช้มุมมองเกี่ยวกับการใช้ระบบสัญญาณเตือนภัยเมื่อมีผู้ที่ฝ่าฝืนการรักษาระยะห่างและยังระบุความหนาแน่นประชากรวิกฤติเพื่อควบคุมการเข้าและออกของฝูงชน ซึ่งทั้งหมดจำเป็นต้องเป็นระบบอัตโนมัติ โดยใช้การตรวจจับและการควบคุมอัตโนมัติแบบวิสัยทัศน์เป็นฐาน (Vision-based) การที่จะทำให้ประหยัดและมีประสิทธิภาพสูงสุดในส่วนของการออกแบบระบบในพื้นที่สาธารณะจำเป็นต้องใช้ระบบอัจฉริยะ โดยมีการแจ้งสัญญาณเตือนหากมีผู้ฝ่าฝืนการรักษาระยะห่าง ซึ่งเป็นงานวิจัยที่เริ่มใช้ AI ในการช่วยป้องกันโควิด-19 หลังจากนั้นก็ได้มีงานวิจัย Pi [9] ที่เล็งเห็นถึงความสำคัญของการป้องกันโรคโควิด-19 ออกมาเช่นเดียวกัน งานวิจัยชิ้นนี้ได้สร้างข้อมูลที่เรียกว่า Xiamen Dataset ที่ถูกสร้างมาจากวิดีโอที่ถูกบันทึกที่เมือง Xiamen ประเทศจีนในวันที่ 16 เมษายน ปี ค.ศ. 2020 ในช่วงสถานการณ์โควิด-19 โดยมีจุดประสงค์เพื่อตามรอยบุคคลจากการศึกษาผ่านการข้ามถนนของผู้คนโดยดึงข้อมูลดิบมาจากวิดีโอ และใช้การเปลี่ยนภาพจากมุมมองเปอร์สเปกทีฟ (Perspective) ไปสู่ภาพออร์ทोगอนอล (Orthogonal) โดยกำหนดจุดไว้ 4 จุดเพื่อเป็นจุดอ้างอิง โดยใช้การแปลงภาพด้วยระบบโฮโมกราฟี (Homography Transformation) ในปีเดียวกันงานวิจัยของ Manoharan [10] ก็ได้ทำการออกแบบเพื่อพัฒนาวิธีการประมาณระยะห่าง ตามมาตรการการรักษาระยะห่างโดยใช้เมตริกซ์ เมื่อคนสองคนยืนในเฟรมเดียวกันและรักษาระยะห่างแก่กัน ซอร์ฟแวร์จะระบุว่าคุณคนนั้น ๆ ได้รักษาระยะห่างตามคำแนะนำของ WHO หรือไม่ โดยใช้โมเดลของชุดคำสั่ง YOLO เวอร์ชัน 3 ซึ่งเป็นปัญญาประดิษฐ์ที่มีจุดประสงค์เพื่อใช้ตรวจจับวัตถุแต่ผลลัพธ์กลับแสดงให้เห็นว่าการใช้ฐานข้อมูล (Dataset) สำเร็จรูปที่ไม่ได้ฝึกฝนเพิ่มเติมจะให้ผลลัพธ์ที่แย่กว่าฐานข้อมูล (Dataset) ที่ทำการฝึกฝนเพิ่มเติมเนื่องจากการบิดเบี้ยวของภาพซึ่งนี่ถือเป็นประเด็นสำคัญที่ต้องพิจารณาต่อไป อีกทั้งงานวิจัยทั้ง 2 ยังใช้ชุดคำสั่ง YOLO ในการตรวจจับวัตถุอีกด้วย นั่นแสดงให้เห็นว่าชุดคำสั่ง YOLO มีศักยภาพเพียงพอต่อการใช้ตรวจจับวัตถุ

จากงานวิจัย [6,8-10] ที่กล่าวมานี้จะเห็นได้ว่าหัวใจหลักของการเชื่อมโยงระหว่างการสืบค้นวิดีโอกับเทคโนโลยีด้านภูมิสารสนเทศจะประกอบไปด้วยกระบวนการตรวจจับวัตถุ (Object Detection) ที่ให้ได้มาซึ่งค่าพิกัด โดยเหตุนี้เองคณะผู้วิจัยจึงได้เลือกใช้ชุดคำสั่ง YOLO ที่ผ่านการพิสูจน์จากวรรณกรรมที่เกี่ยวข้องโดยได้แสดงให้เห็นถึงศักยภาพสูงในการตรวจจับวัตถุควบคู่ไปกับการมีความรวดเร็วเหมาะสำหรับการทำงานแบบทันทีทันใด (Real Time) อย่างที่สองคือกระบวนการแปลงค่าพิกัดจากรูปแบบของชุดคำสั่ง YOLO (YOLO Coordinate System) ให้เข้าสู่ค่าพิกัดของภาคพื้นดินหรือที่เรียกว่าค่าพิกัดภายในของห้องภายในอาคาร (Room Coordinate System) โดยเหตุนี้เอง คณะผู้วิจัยจึงขออธิบายรายละเอียดและกระบวนการทำงานที่เกี่ยวข้องกับชุดคำสั่ง YOLO ซึ่งรายละเอียดดังนี้

2.1 ชุดคำสั่งโยโล (YOLO)

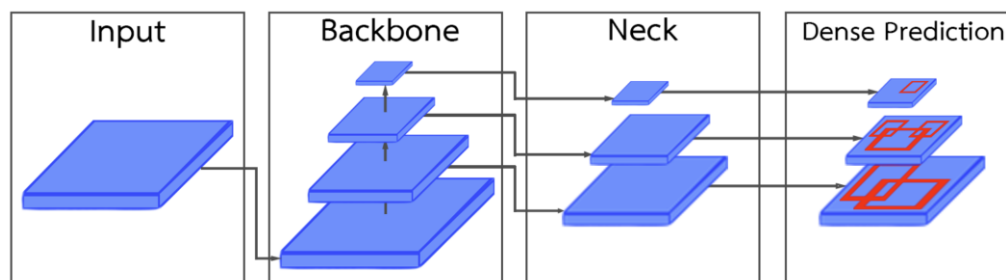
ชุดคำสั่ง YOLO (ย่อมาจาก You Only Look Once) คือชุดคำสั่งที่ประยุกต์ใช้โครงข่ายประสาทเทียม (Convolutional Neural Network) มาจำลองการมองเห็นของมนุษย์และตรวจจับรูปภาพ ซึ่งชุดคำสั่ง YOLO เป็นโอเพนซอร์สปัญญาประดิษฐ์โดยเริ่มแรกพัฒนาผ่านภาษา C++ โดย Redmon [11] ตัวของชุดคำสั่ง YOLO นั้นมีความสามารถในการประมวลผลผ่านกราฟิกการ์ดได้อย่างดี

จึงเหมาะเป็นอย่างยิ่งกับการใช้งานที่ต้องการตรวจจับภาพจากกล้องวิดีโอหรือภาพถ่ายแบบที่ต้องการใช้งานแบบทันทีทันใด ต่อมาจึงได้สิ้นสุดการพัฒนาไว้ที่ชุดคำสั่ง YOLO เวอร์ชัน 3 แต่ด้วยการเป็นโอเพนซอร์สจึงให้ชุมชนเข้ามามีส่วนร่วมในการพัฒนาได้ จึงได้แบ่งย่อยออกมาเป็นชุดคำสั่ง YOLO เวอร์ชัน 4 พัฒนาต่อยอดโดย Bochkovskiy [12] ในขณะที่ชุดคำสั่ง YOLO เวอร์ชัน 5 พัฒนาต่อยอดโดย Joscher [13]

ชุดคำสั่ง YOLO เวอร์ชัน 5 จะใช้สถาปัตยกรรม PyTorch ที่มีความได้เปรียบเรื่องชุมชนที่มีขนาดใหญ่ มีขั้นตอนการติดตั้งง่ายและไม่ยุ่งยาก โดยทุกอย่างยังอยู่บนฐานของชุดคำสั่ง YOLO เวอร์ชัน 3 ประกอบด้วยแกนหลักที่เปรียบเสมือนกระดูกสันหลัง (Backbone) ของโมเดลจะใช้เป็น CSPDarknet53 ซึ่งเป็นสถาปัตยกรรมของโครงข่ายประสาทเทียมที่ใช้สำหรับการตรวจจับวัตถุที่ใช้สถาปัตยกรรม Darknet ที่เป็นสถาปัตยกรรมโอเพนซอร์สของโครงข่ายประสาทเทียม มีจุดประสงค์มุ่งเน้นการตรวจจับวัตถุเป็นหลัก โดยเดิมทีถูกเขียนมาให้รองรับกับภาษา C และรองรับการทำงานของ CUDA [12,14] แต่จะมีการใช้ CSPNet เข้าช่วยในการแบ่งรูปออกเป็นสองส่วนแล้วรวมเข้ากันซึ่งช่วยให้การไล่ระดับผ่านเครือข่ายตื้นมากขึ้น, คอ (Neck) ซึ่งเป็นตัวเชื่อมระหว่าง Backbone และ Head จะใช้ PANet (Path Aggregation Network) และหัว (Head) หรือเลเยอร์ผลลัพธ์ของการตรวจจับครั้งสุดท้ายจะใช้แบบเดียวกับชุดคำสั่ง YOLO เวอร์ชัน 3 ดังสมการที่ 1 แต่ในด้านสถาปัตยกรรมเชิงลึกทางผู้พัฒนาชุดคำสั่ง YOLO เวอร์ชัน 5 ไม่ได้มีการตีพิมพ์งานวิจัยออกมาเปิดเผยในส่วนนี้ [15]

$$\text{จำนวนเลเยอร์} = B(5+C) \quad (1)$$

โดย B คือจำนวนของกรอบสี่เหลี่ยม
และ C คือจำนวนของประเภทวัตถุ

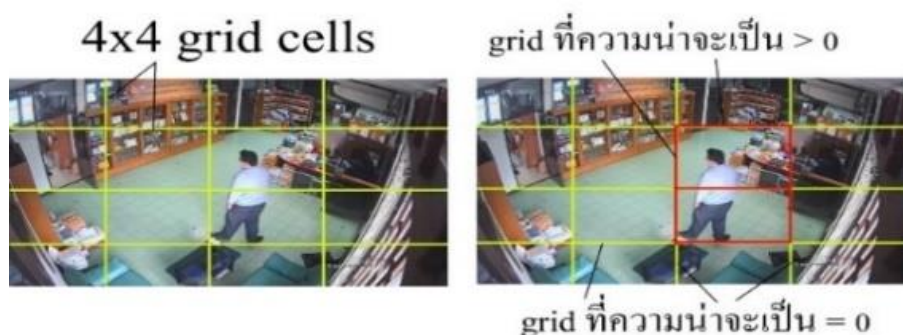


รูปที่ 1 สถาปัตยกรรมการทำงานของชุดคำสั่ง YOLO

โดยการทำงานของชุดคำสั่ง YOLO จะประกอบด้วยขั้นตอนต่าง ๆ ดังต่อไปนี้

2.1.1 บล็อกตัวต้านทาน (Residual blocks)

ขั้นตอน Residual blocks เป็นการแบ่งเป็นภาพตั้งต้น A ให้เป็นตารางกริด (grid cells) ขนาดที่เท่า ๆ กันโดยจำนวน NxN grid cells โดยในรูปที่ 2 ได้มีการแบ่งภาพถ่ายออกจำนวน $4 \times 4 = 16$ grid cells ซึ่งแต่ละช่องในตารางกริดจะใช้ในการคำนวณเพื่อทำนายถึงความน่าจะเป็นวัตถุในภาพซึ่งเป็นประเภท (Class) ของวัตถุที่ให้ความสนใจ กล่าวคือ ในรูปที่ 2 จะเห็นว่า วัตถุที่ต้องการค้นหาคือภาพของบุคคล โดยภาพของบุคคลในกรอบสีแดงจำนวน 2 grid cells ที่มีค่าความน่าจะเป็นมากกว่า 0 ในขณะที่ grid cells อื่นๆจะมีความน่าจะเป็นเท่ากับ 0



รูปที่ 2 แปลงภาพต้นฉบับเป็น grid cells และจำลอง probability score ในแต่ละ grid cell

2.1.2 การปรับตำแหน่งกรอบสี่เหลี่ยม (Bounding Box Regression)

จากนั้นวัตถุในกรอบสีแดงของเซลล์ทั้ง 2 ในภาพดังรูปที่ 2 จะมีการปรับตำแหน่งกรอบสี่เหลี่ยมเป็นการสร้างกรอบ Bounding Box เพื่อล้อมวัตถุซึ่งในที่นี้คือภาพบุคคลโดยขึ้นอยู่กับกรอบสี่เหลี่ยมที่ล้อมรอบวัตถุ โดยแสดงผ่านเวกเตอร์ Y ในแต่ละ Bounding Box (รูปที่ 3) โดยมีรูปแบบดังสมการที่ 2 ที่เป็นพารามิเตอร์แสดงรายละเอียดที่เกี่ยวข้องกับตำแหน่งของกรอบสีแดงดังนี้

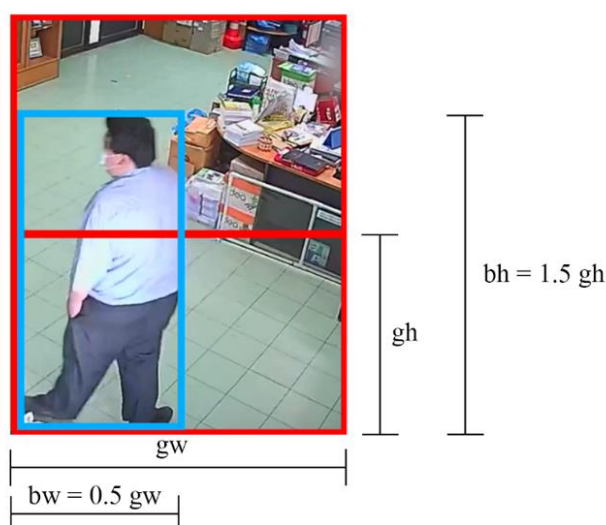
$$Y = [pc, bx, by, bh, bw, c1] \quad (2)$$

โดย pc คือ probability score หรือความน่าจะเป็นของแต่ละตารางที่มีวัตถุ วัตถุที่อยู่ในตารางสีแดงจะมีค่ามากกว่า 0 ในขณะที่ภายในกรอบสี่เหลี่ยมจะมีค่าเท่ากับ 0 ดังแสดงในรูปที่ 1

bx และ by คือค่าของ x, y หรือก็คือตำแหน่งตรงกลาง Bounding Box โดยพิจารณาจาก grid cells

bh และ bw คือค่าของความกว้างและความสูงของ Bounding Box โดยพิจารณาจาก grid cells

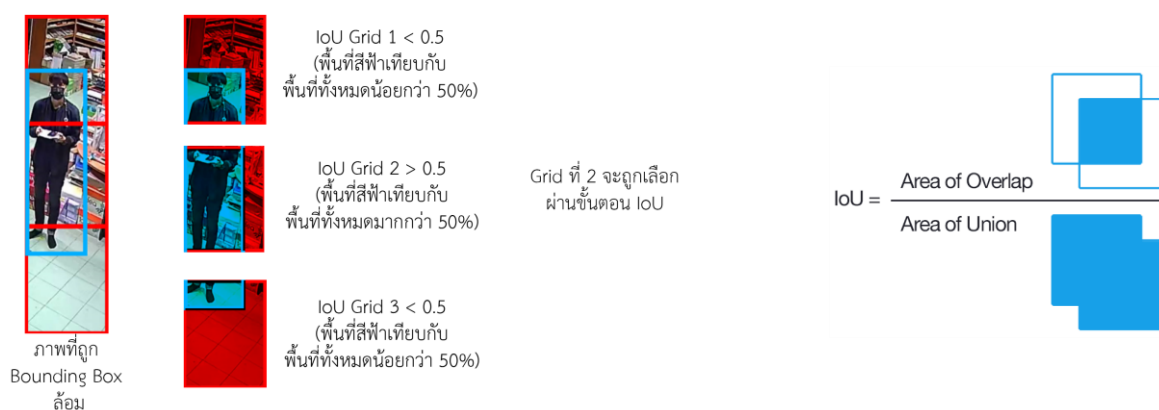
$c1$ คือ class ของวัตถุในที่นี้คือบุคคล



รูปที่ 3 ค่าต่าง ๆ ของ Bounding Box

2.1.3 กระบวนการอินเตอร์เซกชันโอเวอร์ยูเนียน (Intersection Over Unions: IoU)

เนื่องจากวัตถุ 1 ชิ้นจะสามารถอยู่ในกล่องตารางได้มากกว่า 1 ประเภท ดังนั้น การพิจารณาขอบเขตของวัตถุพิจารณาจากค่า IoU ที่มีค่าตั้งแต่ 0 ถึง 1 ในการระบุว่าวัตถุดังกล่าวควรเป็นวัตถุประเภทใด โดยเป็นการระบุเทรชโฮลด์ (Threshold) ของค่า IoU ทั้งนี้ค่า Threshold ที่นิยมใช้กันคือ 0.5 โดยจะนำมาใช้ร่วมกับผลลัพธ์ที่ได้จากชุดคำสั่ง YOLO ที่ทำการคำนวณค่า IoU ในแต่ละ grid cell ด้วยวิธีการใช้พื้นที่ที่มีการเกิด Intersection หาคด้วยพื้นที่ที่มีการเกิด Union แล้วคัดเลือกเฉพาะพื้นที่ที่มีค่า IoU ที่มากกว่าค่า Threshold ยกตัวอย่างเช่น ภาพบุคคลในรูปที่ 4 พื้นที่สีฟ้ามีค่า Threshold เกินกว่า 0.5 ก็จะถูกละเลือก ส่วนพื้นที่อื่น ๆ ซึ่งเป็นพื้นที่สีแดงมีค่า IOU น้อยกว่าค่า Threshold ก็จะไม่นำมาวิเคราะห์ ซึ่งจะเป็นกรองพื้นที่ให้



รูปที่ 4 สรุปหลักการคำนวณค่า IoU

2.1.4. การลดค่าสูงสุดที่ไม่เกี่ยวข้อง (Non-Maximum Suppression: NMS)

ในบางกรณีการนำเฉพาะค่า IoU มาพิจารณาอาจจะให้ผลลัพธ์ที่ไม่ถูกต้องเพียงพอ ทั้งนี้เพราะบ่อยครั้งที่วัตถุจะมีหลาย grid cells ที่มีค่า IoU มากกว่าค่า Threshold ดังนั้นหลักการลดค่าสูงสุดที่ไม่เกี่ยวข้อง (Non-Maximum Suppression: NMS) จึงถูกนำมาประยุกต์ใช้ โดยวิธีการ NMS นั้นเป็นการเก็บเฉพาะกล่องที่มีค่าคะแนนความน่าจะเป็นสูงที่สุดไว้เป็นผลลัพธ์ของการตรวจจับวัตถุ ดังสมการที่ 3

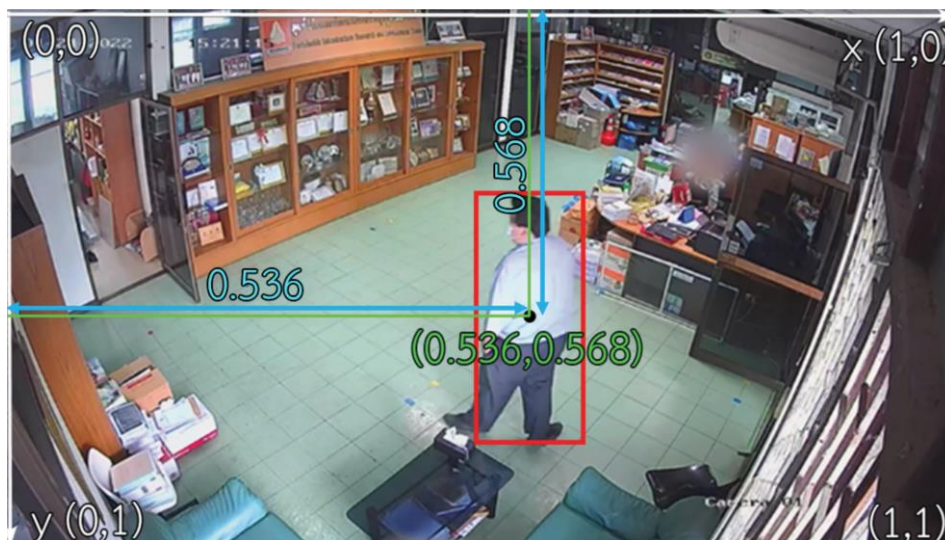
$$P_0 = P_{object} \cdot IoU \quad (3)$$

เมื่อ P_{object} = ความน่าจะเป็นที่กล่องจะครอบวัตถุได้

$$\text{และ } IoU = \frac{\text{กล่องที่ทำงานไว้} \cap \text{กล่องที่ถูกต้อง}}{\text{กล่องที่ทำงานไว้} \cup \text{กล่องที่ถูกต้อง}}$$

โดย P_0 จะมีค่าได้ตั้งแต่ 0 ถึง 1

เมื่อจับความสัมพันธ์ระหว่าง x และ y ที่ตำแหน่งใด ๆ ของรูปภาพจะสามารถวิเคราะห์สร้างแกนของตำแหน่งกึ่งกลาง (x,y) ของวัตถุใด ๆ บนรูปแบบค่าพิกัดที่ได้จากการประมวลผลด้วยชุดคำสั่ง YOLO ได้ดังรูปที่ 5



รูปที่ 5 ระบบพิกัดภาพของ ชุดคำสั่ง YOLO ที่ตำแหน่ง x,y ใด ๆ

หลังจากได้ค่าพิกัดที่ได้จากการประมวลผลด้วยชุดคำสั่ง YOLO ซึ่งในที่นี้จะอ้างถึงเป็นระบบพิกัด YOLO (YOLO Coordinate System) ซึ่งการจะแปลงค่าพิกัดให้เป็นค่าพิกัดห้อง (Room Coordinate) จะสามารถทำได้ด้วยระบบโฮโมกราฟีเช่นเดียวกันกับงานวิจัยของ Pi [9] แต่ถึงแม้ว่าการแปลงค่าพิกัดแบบโฮโมกราฟีจะได้ผลลัพธ์ออกมาเป็นออร์โธโกนอล (Orthogonal) แต่จะเห็นได้ว่ารูปแบบของสมการมีรูปแบบอยู่ในรูปเศษส่วน ทำให้การประมาณค่าเริ่มต้นที่มักจะประมาณค่าออกมาเป็น 0 ไม่สามารถทำได้ นั่นจึงทำให้สมการโฮโมกราฟีนั้นมีความยากมากขึ้นไปอีก แต่ก็ยังมีการแปลงค่าพิกัดในอีกระบบที่ทำได้ในเชิงปฏิบัติมากกว่านั้นคือสมการการแปลงค่าพิกัดแบบสมพรรค (Affine Transformation) ที่สามารถประมาณค่าเริ่มต้นที่ 0 ได้ ทำให้สามารถทำงานได้เร็วมากยิ่งขึ้น จึงเหมาะกับการใช้งานแบบทันทีทันใด

2.2 การแปลงค่าพิกัด

กระบวนการแปลงค่าพิกัดนับเป็นหัวใจสำคัญของการศึกษาวิจัยนี้ โดยจะเป็นการแปลงค่าพิกัดจากระบบพิกัดภาพที่ได้จากกล้อง CCTV ซึ่งเป็นค่าพิกัดที่อยู่ในระบบพิกัดจากชุดคำสั่ง YOLO (YOLO Coordinate System) ให้เป็นค่าพิกัดภายในของระบบพิกัดห้องภายในอาคาร (Room Coordinate System) โดยการศึกษาของ Pi [9] นั้นได้เสนอแนะให้ใช้การแปลงค่าพิกัดแบบโฮโมกราฟี (Homography Transformation) ที่เป็นการแปลงค่าพิกัดระหว่าง 2 ระนาบโดยใช้การถ่ายภาพลง 2 ระนาบลงบนเมตริกซ์ขนาด 3×3 ดังแสดงในสมการที่ 4

$$(X,Y)=\begin{pmatrix} \frac{p_1x+p_2y+p_3}{p_7x+p_8y+p_9} & \frac{p_4x+p_5y+p_6}{p_7x+p_8y+p_9} \end{pmatrix} \quad (4)$$

เมื่อ X คือค่าพิกัดแกน x ของระบบพิกัดห้องภายในอาคาร (Room Coordinate System)

Y คือค่าพิกัดแกน y ของระบบพิกัดห้องภายในอาคาร (Room Coordinate System)

x คือค่าพิกัดแกน x ของระบบพิกัดจากชุดคำสั่ง YOLO (YOLO Coordinate System)

y คือค่าพิกัดแกน y ของระบบพิกัดจากชุดคำสั่ง YOLO (YOLO Coordinate System)

$P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8, P_9$ คือพารามิเตอร์การแปลงค่าพิกัด

จากสมการที่ 4 จะเห็น สมการนี้เป็นสมการที่ไม่เป็นเชิงเส้น (Nonlinear equation) ซึ่งเมื่อนำมาวิเคราะห์หาคำนวณปรับแก้ด้วยวิธีการ Least Squares Adjustment ในการแก้ปัญหาเพื่อหาพารามิเตอร์การแปลงค่าพิกัดนั้นจะค่อนข้างยุ่งยากทั้งนี้เพราะไม่สามารถที่จะกำหนดค่าเริ่มต้นให้เป็น 0 ทั้งนี้เพราะหากค่าเริ่มต้นเป็น 0 ผลลัพธ์ที่ตามมาคือ สมการนี้จะมีตัวหารเท่ากับ 0 ซึ่งไม่สามารถนำมาใช้ในการคำนวณปรับแก้เพื่อหาพารามิเตอร์การแปลงค่าพิกัด และกระบวนการพิจารณาค่าเริ่มต้นให้กับระบบสมการนี้ค่อนข้างยุ่งยาก ด้วยเหตุนี้ คณะผู้วิจัยจึงได้เสนอแนะกระบวนการแปลงค่าพิกัดด้วยสมการสมพรรค (Affine Transformation) ที่เป็นการแปลงค่าพิกัดโดยอาศัยหลักการเคลื่อนย้ายแกนในสองทิศทาง, การทำมุมระหว่างแกนเป็น θ , มาตรฐานสองมาตรฐานสำหรับในทั้งสองแกนคือ S_x และ S_y และการบิดเบี้ยวของระบบพิกัด (Angular Affinity: ϵ) สามารถมีได้เล็กน้อย [16] จึงทำให้ได้ความสัมพันธ์ดังสมการที่ 5 และสมการที่ 6 โดยมีรายละเอียดดังนี้

$$X = S_x \cdot \cos(\theta) \cdot x + S_y \cdot \sin(\theta) \cdot y + \Delta x \quad (5)$$

$$Y = S_x \cdot \sin(\theta + \epsilon) \cdot x + S_y \cdot \cos(\theta + \epsilon) \cdot y + \Delta y \quad (6)$$

เมื่อ X คือค่าพิกัดแกน x ของภาคห้อง (ภาคพื้นดิน)

Y คือค่าพิกัดแกน y ของภาคห้อง (ภาคพื้นดิน)

x คือค่าพิกัดแกน x ของภาพแกน ชุดคำสั่ง YOLO

y คือค่าพิกัดแกน y ของภาพแกน ชุดคำสั่ง YOLO

Δx คือค่าคงที่ของการย้ายแกน x

Δy คือค่าคงที่ของการย้ายแกน y

จากสมการที่ (5) และ (6) นำทั้งสองสมการนี้สามารถจัดรูปใหม่โดยกำหนดให้ $a = S_x \cdot \cos\theta$, $b = S_y \cdot \sin\theta$, $c = \Delta x$, $d = -S_x \cdot \sin(\theta + \epsilon)$, $e = S_y \cdot \cos(\theta + \epsilon)$, และ $f = \Delta y$ จะสามารถจัดรูปสมการได้เป็นสมการที่ 7 และสมการที่ 8 ซึ่งอยู่ในรูปแบบที่ง่ายขึ้นต่อการนำไปประยุกต์ใช้งานได้ดังนี้

$$X = a \cdot x + b \cdot y + c \quad (7)$$

$$Y = d \cdot x + e \cdot y + f \quad (8)$$

เมื่อ X คือค่าพิกัดแกน x ของระบบพิกัดห้องภายในอาคาร (Room Coordinate System)

Y คือค่าพิกัดแกน y ของระบบพิกัดห้องภายในอาคาร (Room Coordinate System)

x คือค่าพิกัดแกน x ของระบบพิกัดจากชุดคำสั่ง YOLO (YOLO Coordinate System)

y คือค่าพิกัดแกน y ของระบบพิกัดจากชุดคำสั่ง YOLO (YOLO Coordinate System)

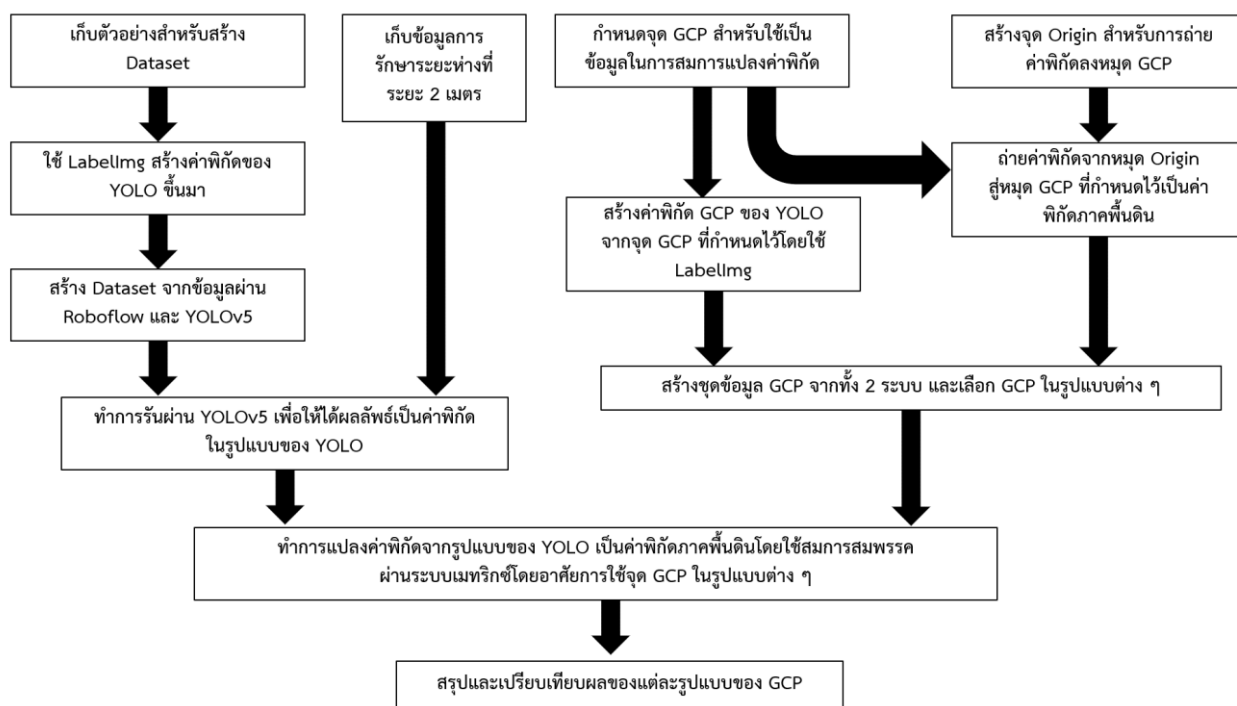
ทั้งนี้จากสมการที่ (7) และ (8) จะเห็นว่า เป็นรูปแบบของสมการเชิงเส้น (Linear Equation) ที่สามารถจะกำหนดค่าเริ่มต้นให้เป็น 0 ได้ และนับเป็นจุดเด่นในการนำไปประยุกต์ใช้กับการคำนวณปรับแก้ที่สามารถวิเคราะห์ค่าพารามิเตอร์การแปลงค่าพิกัดสำหรับการแปลงค่าพิกัดจากระบบพิกัดภาพที่ได้จากกล้อง CCTV ซึ่งเป็นค่าพิกัดที่อยู่ในระบบพิกัดจากชุดคำสั่ง YOLO (YOLO Coordinate System) ให้เป็นค่าพิกัดภายในของระบบพิกัดห้องภายในอาคาร (Room Coordinate System) ได้โดยสะดวกและรวดเร็วต่อการคำนวณ ซึ่งหลักคิดนี้นับเป็นแกนหลักที่งานวิจัยชิ้นนี้มุ่งที่จะนำเสนอเพื่อเป็นประเด็นสำคัญในการศึกษาความถูกต้องเชิงตำแหน่งที่ได้จากการแปลงค่าพิกัดด้วยสมการสมพรรค (Affine Transformation) ซึ่งนำไปเป็นประเด็นที่น่าสนใจในการต่อยอดการศึกษาจากการเรียนรู้เชิงลึกด้วยชุดคำสั่ง YOLO ที่น่าสนใจอย่างมาก

3. วิธีดำเนินงานวิจัย

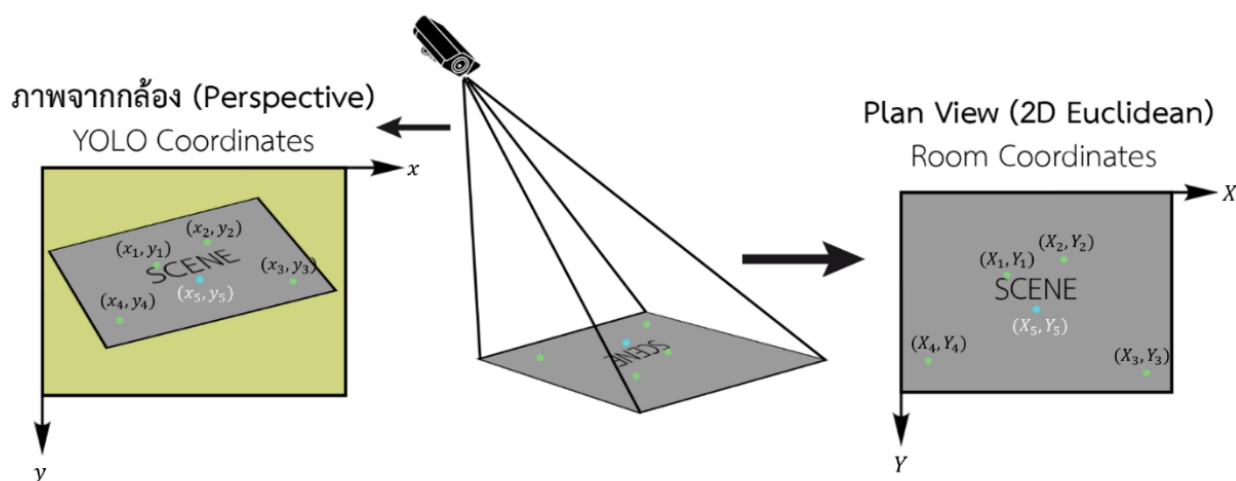
ระเบียบวิธีในการศึกษาครั้งนี้ได้มีการออกแบบกรอบแนวคิดในการศึกษาวิจัยในภาพรวมไว้ในรูปที่ 6 ที่เป็นการแสดงขั้นตอนความเชื่อมโยงระหว่างกระบวนการวิเคราะห์ในแต่ละส่วนเข้าไว้ด้วยกัน ซึ่งมีหลักการสำคัญของงานวิจัยคือ การประยุกต์ใช้สมการสมพรรคในการช่วยวิเคราะห์หาความสัมพันธ์ทางคณิตศาสตร์ที่ใช้ในการแปลงค่าพิกัด โดยเป็นการแปลงค่าพิกัดของวัตถุที่ได้จากภาพถ่ายในกล้อง CCTV ซึ่งเป็นค่าพิกัดที่เป็นภาพถ่ายจากภาพถ่ายเฉียงต่ำฉายแบบศูนย์ทิวทัศน์ (low-oblique picture with perspective projection) ซึ่งในการศึกษาวิจัยนี้จากค่าพิกัดภาพถ่ายที่ได้จากการรังวัดบนภาพถ่ายด้วยชุดคำสั่ง YOLO นี้น่า ระบบพิกัด YOLO (YOLO Coordinate System) ทั้งนี้การแปลงค่าพิกัดจากระบบพิกัด YOLO ให้เป็นค่าพิกัดภายในระบบพิกัดห้องภายในอาคาร (Room Coordinate System) ดังภาพสรุปกรอบแนวการแปลงระบบพิกัดในรูปที่ 7

ในขั้นตอนต่อไป คณะผู้วิจัยทำการแปลงค่าพิกัดจากของชุดคำสั่ง YOLO มุมมองจาก Perspective Camera System สู่ Room Coordinate System ที่เป็น 2D Euclidean (รูปที่ 7) โดยการใช้กล้องวงจรปิดเป็นอุปกรณ์บันทึกภาพและใช้ ชุดคำสั่ง YOLO ในการตรวจจับบุคคลเพื่อประยุกต์ใช้กับโรคติดต่อโควิด-19 ที่มีการระบาด รวมถึงสามารถประยุกต์ใช้กับโรคติดต่อในอนาคตได้ โดยวางแผนการสร้างโมเดลของการเรียนรู้เชิงลึกสำหรับการตรวจจับคนเดินเท้าและปัญญาประดิษฐ์หรือ AI อย่าง ชุดคำสั่ง YOLO เวอร์ชัน 5 นั้นจำเป็นต้องใช้ข้อมูลที่มีจำนวนมาก ยิ่งจำนวนของข้อมูลที่น่ามาฝึกฝนมากขึ้นก็จะส่งผลให้ปัญญาประดิษฐ์นั้นมีความแม่นยำที่มากขึ้นด้วยเช่นกัน รวมทั้งข้อมูลจากงานวิจัย [10] ได้แสดงให้เห็นว่าการใช้ชุดข้อมูลสำเร็จรูปมีโอกาสที่จะทำให้ปัญญาประดิษฐ์นั้นตรวจจับวัตถุไม่ได้ โดยเฉพาะกับกล้องที่ใช้ซึ่งเป็นกล้องที่มีความเอนของเลนส์ที่สูงด้วยแล้ว ทางผู้วิจัยจึงเลือกที่จะสร้างชุดข้อมูลขึ้นมาเอง เริ่มจากการเก็บข้อมูลตัวอย่างสำหรับใช้ในการทำชุดข้อมูลสำหรับการเรียนรู้ด้วย AI ซึ่งในที่นี้จะเรียกว่า Survey-KKU Dataset ในการทำการฝึกฝนปัญญาประดิษฐ์ หลังจากนั้นจึงเริ่มจากการฝึกฝนปัญญาประดิษฐ์ ด้วยการ Labelling รูปที่ทำการเก็บมาแต่ละรูปเพื่อให้ได้ค่าพิกัดในรูปแบบของ ชุดคำสั่ง YOLO หลังจากนั้นจึงนำเข้า RoboFlow [17] ที่เป็นเว็บไซต์ที่ช่วยในการแปลงรูปแบบของการตรวจจับวัตถุเพื่อเพิ่มรูปแบบของภาพที่จะนำไปฝึกฝนเช่นการเพิ่มสัญญาณรบกวนแล้วจึงทำการฝึกฝนปัญญาประดิษฐ์ผ่าน ชุดคำสั่ง YOLO เพื่อให้ได้ฐานข้อมูล (Dataset) แล้วนำไปใช้รันเพื่อจับภาพบุคคลตามวัตถุประสงค์ที่ต้องการ เมื่อผลลัพธ์ออกมาแล้วจะได้เป็นแผนที่ที่มีค่าพิกัดของชุดคำสั่ง YOLO (รูปที่ 7) หลังจากนั้นจะกำหนดจุดที่

ต้องการใช้เป็นหมวด GCP และใช้การถ่ายค่าพิกัดจากจุดที่ทราบค่า (5000, 5000) แล้วจึงถ่ายค่าพิกัดไปยังจุด GCP อื่น ๆ เพื่อสร้าง GCP ภาคพื้นดิน (รูปที่ 8 และรูปที่ 9)



รูปที่ 6 กรอบแนวคิดในการวิจัย

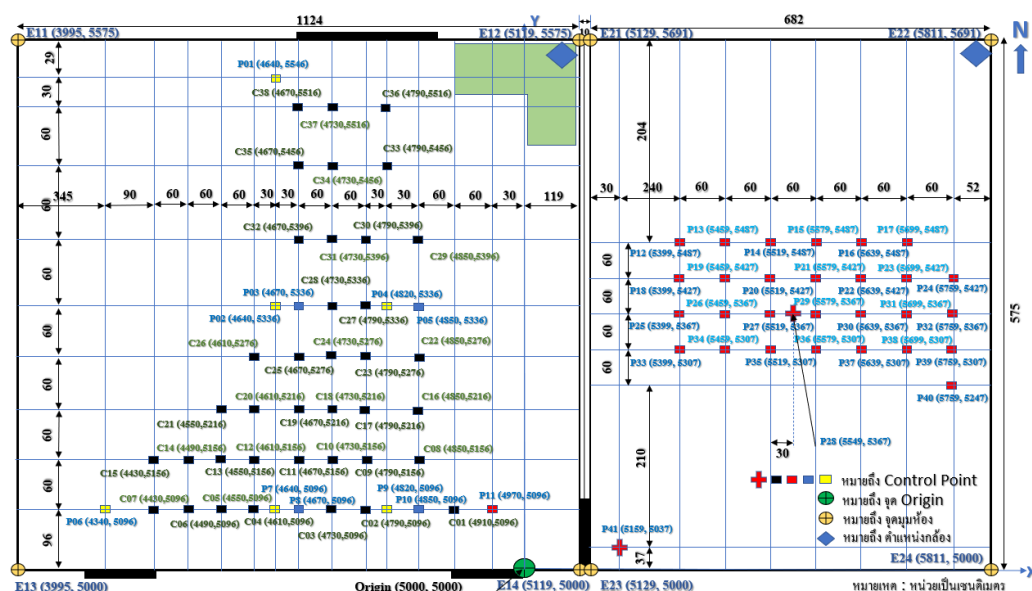


รูปที่ 7 สรุปกรอบแนวการแปลงระบบพิกัดการแปลงค่าพิกัดในระบบพิกัด YOLO (YOLO Coordinate System) สู่ระบบพิกัดห้องภายในอาคาร (Room Coordinate System)



รูปที่ 8 ภาพถ่ายจากกล้อง CCTV ที่มีการร่างระบบแกนของระบบพิกัดห้องซึ่งมีจุดศูนย์กลางที่กำหนดขึ้นเองโดยกำหนดที่มุมของวงกบประตูให้มีค่าพิกัดจุดสี่เหลี่ยม และมีตำแหน่งมุมบังคับภาพถ่ายและมุมตรวจสอบกระจายอยู่ทั่วทั้งห้อง (ทั้งนี้ห้องทั้งสองนี้มีใช้ระบบพิกัดร่วมกัน)

หลังจากนั้นจึงทำการสร้างค่าพิกัด GCP ของ ชุดคำสั่ง YOLO โดยใช้โปรแกรม LabelImg ซึ่งเป็น Freeware แล้วสร้างความสัมพันธ์ระหว่างค่าพิกัดทั้ง 2 ระบบ จากทั้งพิกัด GCP ทั้งจากภาคพื้นดินและ ชุดคำสั่ง YOLO เก็บไว้เป็นชุดข้อมูลของ GCP เมื่อเตรียมข้อมูลเสร็จสิ้นแล้วจึงทำการเก็บข้อมูลการรักษาระยะห่างที่ 200 เซนติเมตร (2 เมตร) โดยใช้เชือกผูกและดึงให้ตึงระหว่างคน 2 คนที่ระยะ 200 เซนติเมตร หลังจากนั้นทำการแปลงค่าพิกัดจากของชุดคำสั่ง YOLO มุมมองจาก Perspective Camera System สู่ Room Coordinate System โดยใช้สมการสมพรรคผ่านระบบเมทริกซ์ในซอฟต์แวร์ Mathematica เพื่อให้ได้ภาพแผนที่ภาคพื้นดิน โดยเลือกรูปแบบต่าง ๆ ของ GCP เพื่อให้ข้อมูลมีความหลากหลาย และเพื่อศึกษาว่ากรณีใดจะให้ผลลัพธ์ได้ดีที่สุดโดยเปรียบเทียบความถูกต้องกับระยะทาง 200 เซนติเมตร



รูปที่ 9 ระบบพิกัดห้องซึ่งมีจุดศูนย์กลางที่กำหนดขึ้นเองของห้องทั้งสองที่อ้างอิงอยู่ในระบบพิกัดเดียวกัน โดยมีมุมบังคับภาพ (GCP) ที่กระจายอยู่ทั่วทั้งห้องในรูปแบบของแปลนอาคารที่ลงรายละเอียดค่าพิกัด (Top View)

4. ผลการวิจัยและการอภิปรายผลงานวิจัย

ผลลัพธ์ที่ได้จากการศึกษาวิจัยนี้ได้มีการประยุกต์ใช้สมการการแปลงค่าพิกัดแบบสมพรรค (Affine Transformation) ซึ่งได้แสดงความสามารถในการเชื่อมโยงและระบุตำแหน่งจากระบบการค้นคืนวิดีโอเข้ากับเทคโนโลยีด้านภูมิสารสนเทศภายในอาคาร (Indoor Geographic Information System: IGIS) โดยเป็นการดำเนินการแปลงค่าพิกัดระหว่างค่าพิกัดจากของชุดคำสั่ง YOLO ที่เป็นภาพในมุมมองศูนย์ทวิทัศนจากกล้อง CCTV ที่เป็น Perspective Camera System ดังรูปที่ 10 (ก) ผ่านการใช้ค่าพิกัดของหมุดบังคับภาพ (Ground Control Point: GCP) ซึ่งเป็นค่าพิกัดในระบบพิกัดห้องภายในอาคาร Room Coordinate System ที่เป็นค่าพิกัดในโลกความเป็นจริงที่ฝังไว้ในหน่วยเซนติเมตร ทั้งนี้ได้กำหนดให้มีการกระจายตัวของหมุดบังคับภาพดังรูปที่ 10 (ข) จำนวน 14 หมุด (ในวงกลมสีแดง) เพื่อนำไปผ่านกระบวนการคำนวณปรับแก้ด้วยวิธีกำลังสองน้อยที่สุดแบบสมการค่าสังเกต ซึ่งให้ผลลัพธ์ที่ได้ออกมาเป็นค่าสัมประสิทธิ์ A, B, C, D, E, และ F โดยเมื่อนำมาเขียนเป็นระบบสมการได้ดังนี้

$$X = 753.13x - 843.35y + 5378.84 \quad (10)$$

$$Y = -581.84x - 630.19y + 5464.97 \quad (11)$$

เมื่อ A = 753.13, B = 843.35, C = 5378.84, D = -581.84, E = -630.19, F = 5464.97

X, Y แทนค่าพิกัดของหมุดบังคับภาพในระบบพิกัด Room Coordinate System

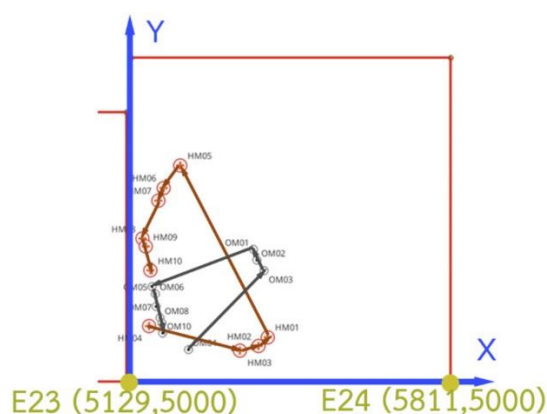
x, y แทนค่าพิกัดภาพของหมุดบังคับภาพในระบบพิกัด YOLO Coordinate System



รูปที่ 10 การประยุกต์ใช้ AI ในการตรวจหาแนวเส้นทางการเดินโดย (ก) การวิเคราะห์ภาพถ่ายจาก CCTV ซึ่งเป็นระบบพิกัด YOLO (YOLO Coordinate System) ให้เป็น (ข) ค่าพิกัดในระบบพิกัดห้องภายในอาคาร (Room Coordinate System) โดยใช้ตำแหน่งของหมุดบังคับภาพตามวงกลมสีแดงที่วางไว้

นำสมการที่ (10) และ (11) มาใช้ในการแปลงค่าพิกัดจากของบุคคลที่อยู่ในระบบ YOLO Coordinate System ให้เป็นค่าพิกัดของบุคคลที่อยู่ในระบบ Room Coordinate System พร้อมด้วยความคลาดเคลื่อนอยู่ที่ ± 47 เซนติเมตร เมื่อเทียบกับระยะทางสัมพัทธ์ (Relative Distance) ที่ 2 เมตรที่เป็นเกณฑ์สำหรับการรักษาระยะห่างที่กำหนดไว้ก่อนเก็บข้อมูล ผลลัพธ์นี้มีความสำคัญอย่างยิ่งในมิติของการเพิ่มขีดความสามารถในการระบุตำแหน่งของบุคคลที่สอดคล้องกับโลกของความเป็นจริงให้กับภาพที่ได้จากกล้อง CCTV ได้ ยิ่งไปกว่านั้น จุดเด่นของภาพที่ได้จากการใช้กล้อง CCTV บันทึกข้อมูลคือ ข้อมูลเวลา ณ ขณะที่ทำการบันทึกภาพ ซึ่งสามารถใช้ในการเชื่อมโยงตำแหน่งของบุคคลเข้ากับเวลาได้ ดังนั้น เมื่อผสมผสานข้อมูลค่าพิกัดของบุคคลในระบบ Room

Coordinate System ร่วมกับเวลาที่สอดคล้องกัน ผลลัพธ์ที่ได้จากทำให้ได้รูปแบบของแนวเส้นทางการเดินดังแสดงในรูปที่ 11 โดยเส้นสีส้มเป็นเส้นทางการเดินของบุคคลที่ 1 และเส้นสีดำเป็นเส้นทางการเดินของบุคคลที่ 2 ซึ่งผลลัพธ์ที่ได้แสดงให้เห็นนี้ได้แสดงให้เห็นถึงผลลัพธ์ได้จากการประยุกต์ใช้สมการการแปลงค่าพิกัดแบบ Affine Transformation ซึ่งได้แสดงความสามารถในการเชื่อมโยงและระบุตำแหน่งจากระบบการค้นคืนวิดีโอเข้ากับ Indoor GIS ที่สามารถนำมาใช้เป็นเทคโนโลยีเสริมในการติดตามบุคคลที่นับเป็นหนึ่งในเครื่องมือในการสนับสนุนกระบวนการติดตามสอบสวนโรคภายในมาตรการการเว้นระยะห่างทางสังคม (Social Distancing) อย่างการป้องกันการระบาดของโรคโควิด-19 ในระหว่างปี ค.ศ. 2020 – 2022



รูปที่ 11 ผลลัพธ์ที่ได้จากการแปลงค่าพิกัดให้อยู่ในระบบพิกัด Room Coordinate System ซึ่งเมื่อนำพิกัดของบุคคล ณ แต่ละจุดที่แต่ละระยะเวลาจะทำให้เกิดการเชื่อมต่อกันเป็นทางเดินซึ่งเป็นการเคลื่อนของบุคคลในอาคาร

อนึ่ง การประยุกต์ใช้สมการการแปลงค่าพิกัดแบบ Affine Transformation ที่มีความเป็นสมการเชิงเส้นนั้นมีข้อได้เปรียบมากกว่าสมการการแปลงค่าพิกัดแบบ Homography Transformation ที่เสนอแนะโดย Pi [9] ทั้งนี้เพราะสมการการแปลงค่าพิกัดแบบ Homography Transformation เป็นสมการที่ไม่เป็นเชิงเส้น โดยมีรูปแบบของสมการเป็นเศษส่วนดังที่ได้แสดงไว้ในสมการที่ (4) จากการนี้เอง ผลที่ตามมาคือ การประมาณค่าเริ่มต้นเพื่อใช้ในการคำนวณปรับแก้ไม่สามารถประมาณการได้อย่างสะดวกเหมือนสมการการแปลงค่าพิกัดแบบ Affine Transformation ที่สามารถกำหนดค่าเริ่มต้นที่ 0 ได้เลย ส่งผลให้สมการการแปลงค่าพิกัดแบบ Affine Transformation ใช้แบบทันทีทันใด (Real Time) ได้เลย นอกจากนี้ ความคลาดเคลื่อนอยู่ที่ได้อยู่ในระดับ ± 47 เซนติเมตร เป็นค่าความคลาดเคลื่อนที่ไม่เกินขนาดของตัวคนโดยทั่วไปจึงอยู่ในเกณฑ์ที่ยอมรับได้ ส่งผลให้การนำมาใช้ในการติดตามพฤติกรรมของบุคคลที่อยู่ในข่ายไม่ปลอดภัยในทางสาธารณสุข (200 เซนติเมตร) ที่เป็นหนึ่งมาตรการทางสาธารณสุขระหว่างปี ค.ศ. 2020 – 2022 นั้น สามารถกระทำได้

งานวิจัยในอนาคตยังสามารถนำงานวิจัยนี้ไปต่อยอดได้อีกหลาย ๆ ด้าน อย่างเช่นการประยุกต์ใช้กับโรคติดต่อในอนาคต หรือจากงานวิจัยของ Shirsat [8] การตรวจจับใบหน้าสำหรับอาชญากรรมนั้น ถ้าหากต้องการความแม่นยำที่สูงจำเป็นต้องใช้ต้นทุนที่สูงหรือถ้ามีต้นทุนต่ำจะทำให้ได้ผลลัพธ์ที่ต่ำมาก ซึ่งการใช้ค่าพิกัดเข้ามาช่วยในการติดตามและระบุตำแหน่งหลังจากตรวจจับใบหน้าแล้วจะสามารถลดภาระของการตรวจจับใบหน้าได้ เป็นการเพิ่มประสิทธิภาพของการตรวจจับใบหน้าไปอีกขั้น หรือแม้กระทั่งการใช้ระบบดังกล่าวก็ยังจะสามารถช่วยติดตามบุคคลสูญหายได้เช่นกัน อย่างไรก็ตาม การเลือกหมุดบังคับภาพ (GCP) ยังส่งผลต่อความ

แม่นยำของผลลัพธ์อีกด้วย การศึกษาต่อยอดในอนาคตรวอย่างยิ่งที่จะต้องมีการลดทอนจำนวนหมวดบังคับภาพให้อยู่ในจำนวนที่เหมาะสมสอดคล้องกับสถานการณ์ในฉากทัศน์ของห้องที่หลากหลาย

5. สรุปผลการวิจัย

ในช่วงเวลาที่มีการแพร่กระจายของเชื้อโควิด-19 ได้แสดงศักยภาพให้เห็นว่าความสามารถในการตามรอยผู้ที่ติดเชื้อและผู้เสี่ยงติดเชื้อมีความสำคัญต่อการลดความสูญเสียได้มากเพียงใด ดังนั้นงานวิจัยชิ้นนี้จึงใช้เทคโนโลยีปัญญาประดิษฐ์ด้วยชุดคำสั่ง YOLO เข้ามาช่วยตรวจจับบุคคลสำหรับช่วยสนับสนุนมาตรการหลักในการป้องกันโรคโควิด-19 นั่นก็คือการควบคุมการรักษาระยะห่างระหว่างบุคคลที่ควรมีระยะห่างมากกว่า 2 เมตร โดยอาศัยภาพวิดีโอที่ได้จากกล้อง CCTV ที่ปัจจุบันพบเห็นได้ทั่วไปผสมผสานเข้ากับเทคโนโลยีด้านภูมิสารสนเทศภายในอาคารเพื่อสร้างเส้นทางการเดินของบุคคลแต่ละคน โดยมีแก่นหลักคือการแปลงค่าพิกัดด้วยสมการสมพรรค ซึ่งประสิทธิภาพในการระบุตำแหน่งอยู่ที่ไม่เกิน ± 47 เซนติเมตร ซึ่งเป็นระดับความคลาดเคลื่อนที่อยู่ในเกณฑ์ที่ยอมรับได้เนื่องจากมีขนาดน้อยกว่าเส้นผ่านศูนย์กลางของบุคคลโดยทั่วไป อีกทั้งหากในอนาคตมีการระบาดของโรคอื่นๆ ก็จะสามารถนำไปประยุกต์ใช้ทั้งกับการรักษาระยะห่างและการติดตามตัวผู้ป่วยเพื่อลดความเสียหายได้เช่นกัน การศึกษาในอนาคตสามารถปรับปรุงโดยการใช้รุ่นปัญญาประดิษฐ์ให้ทันสมัยมากขึ้น รวมทั้งการใช้ข้อมูลสำหรับฝึกฝนฐานข้อมูล (Dataset) ให้มากขึ้นจะสามารถลดปัญหาไม่สามารถตรวจจับบุคคลได้ ทำให้งานวิจัยสมบูรณ์แบบยิ่งขึ้น แต่แตกต่างกับการต้องใช้ทรัพยากรการวิจัยที่มากขึ้นซึ่งเป็นข้อจำกัดของงานวิจัยชิ้นนี้

ผลประโยชน์ทับซ้อน

ผู้เขียนขอประกาศว่าบทความนี้ไม่มีผลประโยชน์ทับซ้อน

กิตติกรรมประกาศ

ผู้วิจัยขอขอบคุณศูนย์วิจัยพัฒนาฐานข้อมูลอย่างยั่งยืน คณะวิศวกรรมศาสตร์ มหาวิทยาลัยขอนแก่น ที่ให้ความอนุเคราะห์พื้นที่ศึกษา วัสดุ อุปกรณ์และเครื่องมือในการดำเนินงานวิจัยในครั้งนี้

เอกสารอ้างอิง

- [1] World Health Organization. *Epidemiological update: Coronavirus disease (COVID-19)*, 11 March 2021. Available form: <https://reliefweb.int/report/costa-rica/epidemiological-update-coronavirus-disease-covid-19-11-march-2021> [Accessed 26 August 2023].
- [2] Nakhon Ratchasima Rajanagarindra Psychiatric Hospital. “Social distancing” what is it and how to apply, 19 August 2020. Available form: <https://www.jvkorat.go.th/th/?p=2114> [Accessed 24 April 2021].
- [3] Centers for Disease Control and Prevention. *COVID-19*, 26 January 2021. Available form: <https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/prevention.html> [Accessed 08 July 2021].
- [4] Zhao, T. *et al.* Is one- or two-meters social distancing enough for COVID-19. *Public health*, 2020, 185: 87. DOI: 10.1016/j.puhe.2020.06.005.
- [5] Ciric L. *One metre or two? the science behind social distancing*, 18 June 2020. Available from: <https://theconversation.com/one-metre-or-two-the-science-behind-social-distancing-139929> [Accessed 24 April 2021].
- [6] Yang, D. *et al.* A vision-based social distancing and critical density detection system for COVID-19. *Image and Video Processing Computer Vision and Pattern Recognition Sensors*, 2021, 21(13): 4608. DOI: 10.3390/s21134608.

- [7] Thairath online. *Timeline 31 Egyptian navy unit visits 2 malls in Rayong*, 14 July 2020. Available from: <https://www.thairath.co.th/news/politic/1888919> [Accessed 28 April 2021].
- [8] Shirsat, S. *et al.* Proposed system for criminal detection and recognition on CCTV data using cloud and machine learning. In: *international conference on vision towards emerging trends in communication and networking (ViTECoN)*, Vellore, India, 30-31 March 2019.
- [9] Pi, Y. *et al.* Deep learning for visual analytics of the spread of COVID-19 infection in crowded urban environments. *American Society of Civil Engineers (ASCE)*, 2021, 0000492: 1527-6996. DOI: 10.1061/(ASCE)NH.1527-6996.0000492.
- [10] Manoharan, J. S. Finding an optimal distance of social distancing for COVID 19, *Journal of ISMAC*, 2021, 3 (3), pp.206-220. DOI: 10.36548/jismac.2021.3.003.
- [11] Redmon, J. *YOLO: real-time object detection*, n.d. Available from: <https://pjreddie.com/darknet/yolo/> [Accessed 07 June 2023]
- [12] Bochkovskiy, A. *YOLOv4 / scaled-YOLOv4 / YOLO - neural networks for object detection*, 23 April 2020. Available from: <https://github.com/AlexeyAB/darknet> [Accessed 07 June 2023].
- [13] Joscher, G. *YOLOv5 in PyTorch*, 26 June 2020. Available from: <https://github.com/ultralytics/yolov5> [Accessed 07 June 2023]
- [14] Bochkovskiy, A. *Darknet*, 21 June 2017. Available from: <https://github.com/pjreddie/darknet> [Accessed 07 June 2023].
- [15] Kukil. *Object detection using YOLOv5 OpenCV DNN in C++ and Python*, 12 April 2022. Available from: <https://learnopencv.com/object-detection-using-yolov5-and-opencv-dnn-in-c-and-python/> [Accessed 17 August 2022].
- [16] Waisurasingha, C. *Photogrammetry*. SE-EDUCATION Public Company Limited: SE-EDUCATION, 2020.
- [17] Roboflow. *Roboflow*, 2020. Available from: <https://roboflow.com/> [Accessed 10 December 2021].