

บทที่ 2

ปรัทัศนัวรรณกรรมและงานวิจัยที่เกี่ยวข้อง

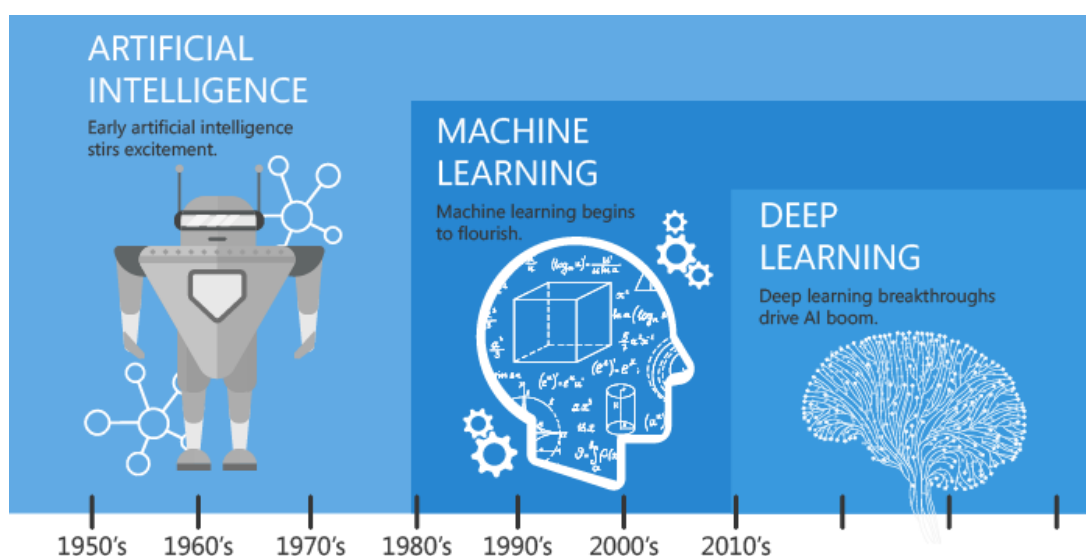
2.1 ปัญญาประดิษฐ์ (Artificial Intelligence : AI)

ปัญญาประดิษฐ์ (Artificial Intelligence) หรือเรียกว่า Ai คือ เป็นวิวัฒนาการด้านเทคโนโลยีที่พัฒนาวิธีการวิเคราะห์ข้อมูล ประมวลผลข้อมูล เพื่อเลียนแบบความสามารถของมนุษย์ในด้านการคิด การเรียนรู้ การแก้ปัญหาให้กับระบบต่าง ๆ โดยอาศัยเครื่องคอมพิวเตอร์และซอฟต์แวร์ที่ซับซ้อน ซึ่งทำให้มีความสามารถในการคิดและการตัดสินใจ ได้ใกล้เคียงกับคน โดยใช้วิธีการจากการศึกษากระบวนการคิดและตัดสินใจหรือหลักการให้เหตุผลของคน นำไปใช้กับกระบวนการพัฒนาศักยภาพของคอมพิวเตอร์ให้ทำการประมวลผลและแสดงผลลัพธ์ที่มากกว่าโปรแกรมทั่วไป เริ่มต้นจากแนวความคิดนี้ จึงกำหนดกระบวนการทำงานให้เป็นลำดับขั้นตอนเพื่อให้เครื่องคอมพิวเตอร์เริ่มต้นจากรับชุดข้อมูล (Data set) ส่งไปยังอัลกอริทึม (Algorithms) ชุดคำสั่งบอกให้ทำการประมวลผล (Computational power) คอมพิวเตอร์ทำการคิด แก้ไขปัญหา ตัดสินใจ เรียนรู้ได้ด้วยตนเอง คล้ายกับคน (Human) ส่งผลให้คอมพิวเตอร์มีความฉลาดมากขึ้น เพิ่มขีดความสามารถให้กับระบบที่มีความสลับซับซ้อนได้อย่างมีประสิทธิภาพ ซึ่งไม่ต้องรอคอยการป้อนข้อมูลจากคนเพียงอย่างเดียว

ปัญญาประดิษฐ์เป็นวิธีการอีกแนวทางหนึ่งของวิทยาการด้านคอมพิวเตอร์ ที่ต้องการปรับปรุงพัฒนาในด้านต่าง ๆ เช่น เครื่องจักรกลอัตโนมัติ ยานยนต์อัตโนมัติ หุ่นยนต์อัตโนมัติ เป็นต้น การตอบสนองการปฏิบัติคล้ายกับคนมากที่สุด รับข้อมูล แก้ปัญหา หาเหตุผล ในการตัดสินใจ หรือการฝึกสอนโมเดลนำมาวิเคราะห์ ประมวลผลแปลความหมายจากปัญหาให้ได้ ทำนายผลที่ถูกต้องแม่นยำและเที่ยงตรง ซึ่งใกล้เคียงกับคนที่ใช้เหตุผลในการตัดสินใจ ดังนั้นปัญญาประดิษฐ์จึงช่วยสนับสนุนการเรียนรู้ของเครื่องจักร (Machine Learning) การเรียนรู้เชิงลึก (Deep Learning) การปรับปรุงกระบวนการที่เหมาะสมและมีประสิทธิภาพ ในงานต่าง ๆ เพื่อให้เกิดความผิดพลาดน้อยที่สุด มีการรับรู้ปัญหา คิดหาวิธีแก้ไข และตัดสินใจ ที่รวดเร็วกว่าการใช้สมองของคน เนื่องจากคนใช้ร่างกายและสมอง ทำงานช้า ๆ กัน ในระยะเวลานาน ๆ เกิดความเมื่อยล้า อาจตัดสินใจผิดพลาดได้

2.1.1 ชนิดของปัญญาประดิษฐ์แบ่งออก 3 ประเภท

- 1) ปัญญาประดิษฐ์ (Artificial Intelligence : AI)
- 2) การเรียนรู้ของเครื่องจักร (Machine Learning : ML)
- 3) การเรียนรู้เชิงลึก (Deep Learning : DL)



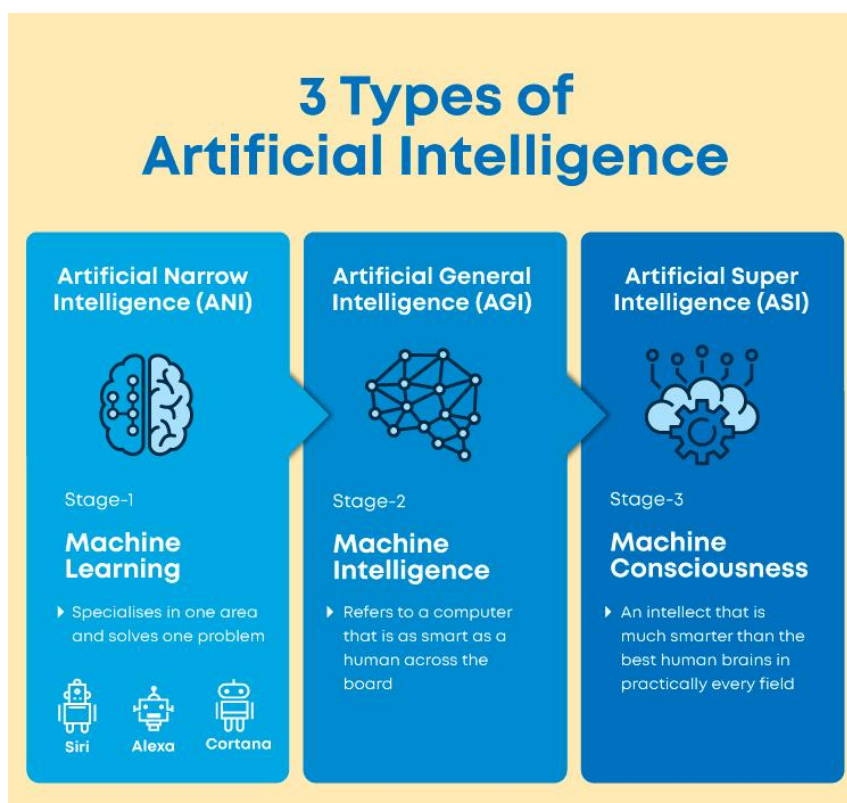
รูปที่ 2.1 ชนิดของปัญญาประดิษฐ์ 3 ประเภท (อ้างอิง <https://www.linkedin.com/pulse/biggest-confusion-cleared-ai-vs-ml-dl-ashwin-kumar>)

2.1.2 ประเภทของปัญญาประดิษฐ์ 3 ระดับ

1) ปัญญาประดิษฐ์แบบเชิงแคบ (Artificial Narrow Intelligence : ANI) หมายถึง ระบบปฏิบัติการหรือเครื่องมือที่ได้ออกแบบการทำงานไว้เพียงอย่างเดียว สามารถทำงานได้อย่างเดียว ANI ถูกจำกัดไว้ที่โดเมนแคบ ๆ ขึ้นพื้นฐาน ได้แก่ ผู้ช่วยด้านสนทนา เช่น Siri หรือ Alexa ระบบประชาสัมพันธ์ติดต่อสอบถาม และระบบจดจำเอกลักษณ์ส่วนบุคคล

2) ปัญญาประดิษฐ์แบบเชิงทั่วไป (Artificial General Intelligence : AGI) หมายถึง ประเภทของปัญญาประดิษฐ์เป็นระบบเครื่องจักรที่มีความสามารถทำงานได้ชาญฉลาดรับรู้ข้อมูลและการคิด อยู่ในระดับเดียวกับคน หมายถึง ระบบ Ai ที่มีความเข้าใจได้ เรียนรู้ได้ และเอาไปใช้งานกับโดเมนด้านต่าง ๆ AGI จะมีทักษะในการคิด ในเชิงนามธรรม กระบวนการความคิดที่มีเหตุผล จัดเรียงข้อมูลในปัญหา สามารถปรับปรุงพัฒนาตัวเอง ให้เข้ากับสถานการณ์ในปัจจุบัน

3) ปัญญาประดิษฐ์แบบเข้มข้น (Strong Artificial Intelligence : ASI) คือ เป็นรูปแบบสมมุติฐานของ Ai ที่มีความสามารถที่เหนือกว่าสติปัญญาของคนในแทบทุกด้าน แสดงถึง ระบบ Ai ที่มีความสามารถเฉลียวฉลาดและมีประสิทธิภาพมากกว่าสมองของคน ASI มีความสามารถเข้าใจปัญหาอย่างลึกซึ้งในเรื่องที่ซับซ้อน แสดงความคิดที่ล้ำหน้า และมีวิวัฒนาการปรับปรุงตัวเองได้อย่างต่อเนื่องไม่หยุดยั้ง นำไปสู่การเจริญเติบโตด้านสติปัญญาแบบทวีคูณ

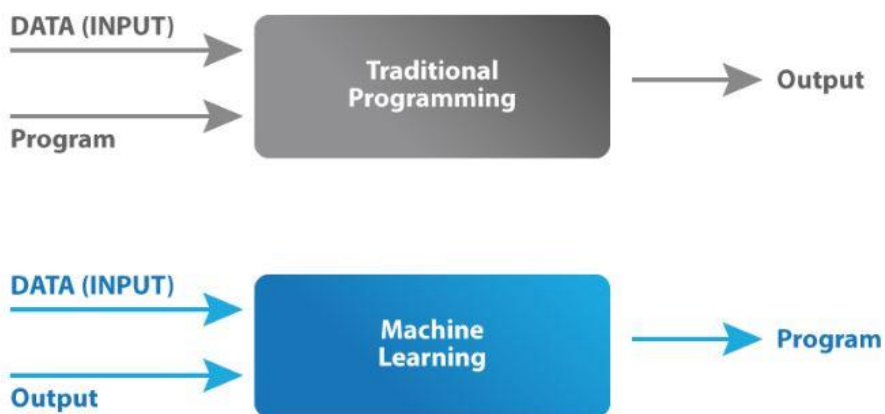


รูปที่ 2.2 ประเภทของปัญญาประดิษฐ์ 3 ระดับ (อ้างอิง <https://www.mygreatlearning.com/blog/what-is-artificial-intelligence/>)

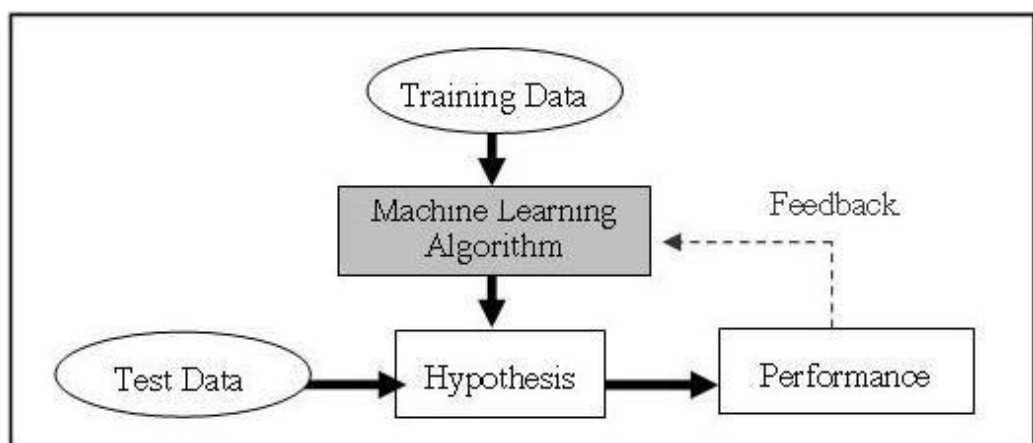
2.2 การเรียนรู้ของเครื่องจักร (Machine Learning : ML)

การเรียนรู้ของเครื่องจักร (Machine Learning) คือ การที่คนทำให้ระบบหรือคอมพิวเตอร์เพิ่มทักษะขีดความสามารถในการเรียนรู้สิ่งต่าง ๆ ลดข้อจำกัดและทำการปรับปรุงด้วยตัวเอง จากชุดข้อมูล (Data set) การฝึกสอนโมเดลให้จดจำ โดยคนไม่ต้องกำกับดูแลหรือเขียนโปรแกรมเพิ่มเติม แม้ว่าในอนาคตจะมีข้อมูลใหม่ ๆ เข้ามาเพิ่มเติม คนไม่ต้องเขียนโปรแกรมเข้าไปใหม่ เพราะคอมพิวเตอร์มีความสามารถตีความหมายและปรับปรุงกระบวนการเพื่อประมวผลได้ด้วยตัวเอง ธุรกิจหรืออุตสาหกรรมใดนำเทคโนโลยีปรับปรุงได้ถูกช่องทาง จะทำให้ได้เปรียบในเชิงกลยุทธ์การแข่งขันทางธุรกิจเป็นอย่างมาก เพราะช่วยลดเวลาการวิเคราะห์และประมวลผลข้อมูลต่าง ๆ การเรียนรู้ของเครื่องจักร (Machine Learning) มีหลักการทำงาน คล้ายการคิดและการตัดสินใจของคน ที่จำเป็นต้องเรียนรู้จากประสบการณ์ในอดีต การป้อนข้อมูล (Data) และอัลกอริทึม (Algorithms) เพื่อให้คอมพิวเตอร์ มีความรับรู้ จดจำ จัดหมวดหมู่ คัดแยก จำแนกประเภท รวมถึง คน สัตว์ สิ่งของ เป็นต้น และเพื่อให้มีความถูกต้อง แม่นยำและเที่ยงตรงยิ่งขึ้น ต้องเพิ่มชุดข้อมูลใหม่ ๆ และ

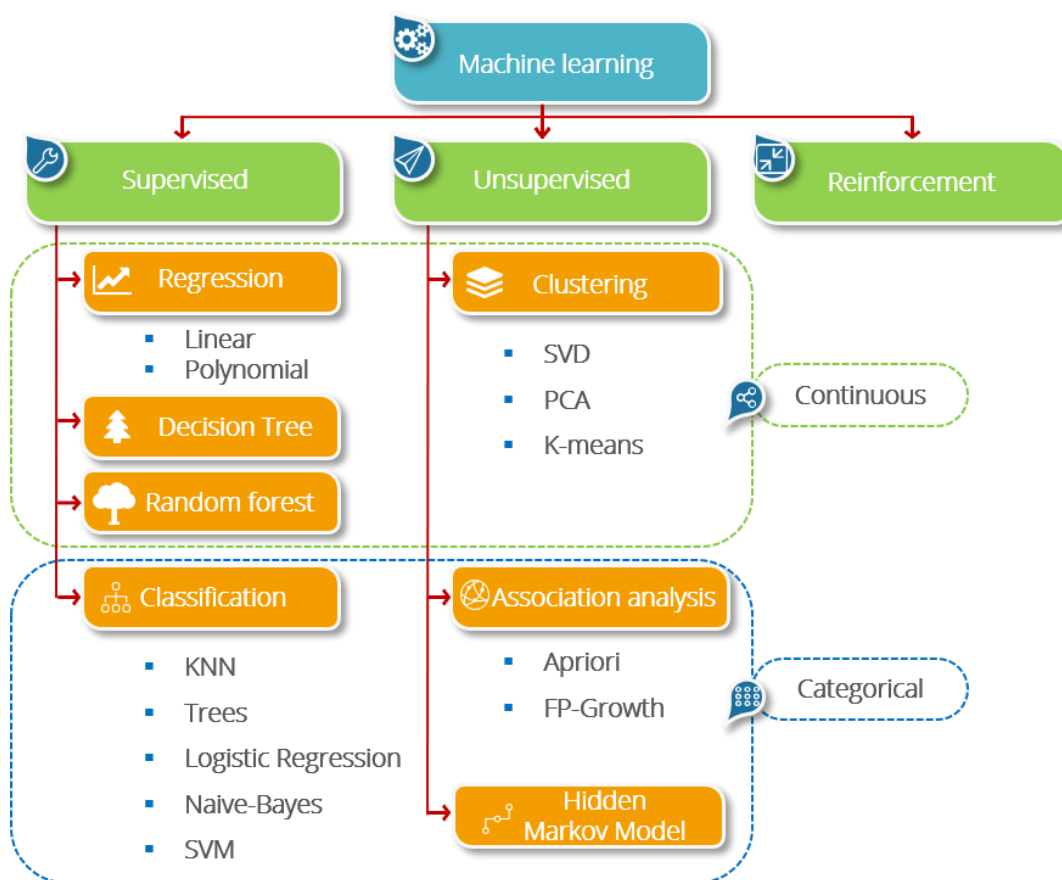
ฝึกสอนโมเดลอยู่เสมอ เพื่อให้ระบบปรับปรุงข้อมูลเป็นปัจจุบัน สามารถคิดได้ด้วยตัวเอง เป็นเทคนิคการเรียนรู้ของเครื่องจักร ดังรูปที่ 2.3



รูปที่ 2.3 ความแตกต่างระหว่างการเรียนรู้ของเครื่องจักรและโปรแกรม (อ้างอิง <https://www.mygreatlearning.com/blog/what-is-machine-learning/>)



รูปที่ 2.4 เทคนิคการเรียนรู้ของเครื่องจักร (อ้างอิง <https://th.wikipedia.org/wiki/การเรียนรู้ของเครื่อง>)



รูปที่ 2.5 การเรียนรู้ของเครื่องจักร แบ่งได้เป็น 3 ประเภท (อ้างอิง <https://4foxsolutions.com/blog/machine-learning-in-mobile-applications>)

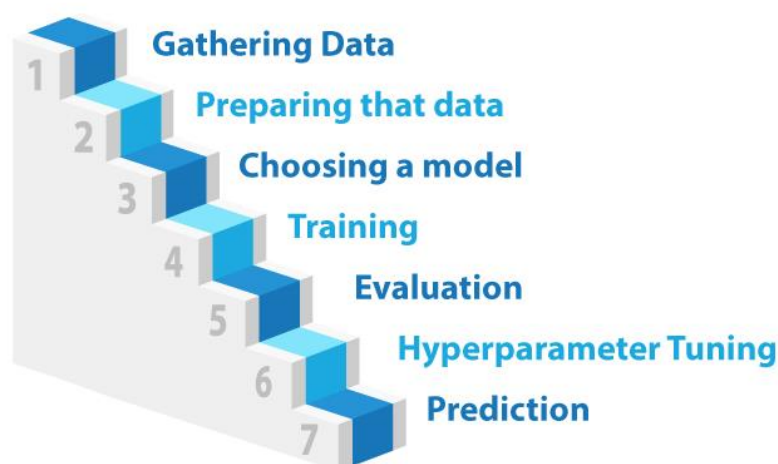
2.2.1 การเรียนรู้ของเครื่องจักร แบ่งได้เป็น 3 ประเภท

1) การเรียนรู้แบบมีผู้สอน (Supervised learning) คือ การเรียนรู้แบบมีผู้สอน การเรียนรู้แบบขึ้นนำด้วยข้อมูล เป็นการฝึกสอนเครื่องคอมพิวเตอร์โดยการใส่ข้อมูล (Input) เข้าไปแล้วมีผลลัพธ์ (Output) ออกมา แบ่งออกได้ 2 ประเภท ได้แก่ การจำแนกประเภท (Classification) และการจำแนกประเภทแบบถดถอย (Regression Classification) การจำแนกประเภท (Classification) คือการจำแนกประเภทของข้อมูล 2 กลุ่มหรือชุดข้อมูลที่ไม่มีความต่อเนื่องโดยใช้ Confusion Matrix เป็นตัววัดค่าความแม่นยำจะเป็นการจัดกลุ่มข้อมูล เช่น การคัดแยกคุณภาพของผลิตภัณฑ์ การคัดแยกประเภทของวัตถุ คน สัตว์ สิ่งของ เป็นต้น การจำแนกประเภทแบบถดถอย (Regression Classification) หรือการถดถอยจะใช้กับข้อมูลที่มีความต่อเนื่อง ไม่ได้รวมกันเป็นกลุ่ม โดยใช้ค่า Root Mean Square Error (RMSE) ในการวัดความแม่นยำ เช่น การทำนายยอดขายสินค้าล่วงหน้า การพยากรณ์อากาศ เป็นต้น

2) การเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning) คือการเรียนรู้ได้ด้วยตนเอง แบบไม่มีผู้สอนหรือไม่ชี้แนะโดยไม่มีชุดข้อมูล เป็นการฝึกสอนเครื่องคอมพิวเตอร์โดยการป้อนข้อมูล (Input) ให้เรียนรู้และค้นพบรูปแบบด้วยตนเอง การเรียนรู้แบบนี้มักนำไปใช้เพื่อการแยกกลุ่ม (Clustering) กลุ่มลูกค้าสำหรับ (Target Advertisement) การลดขนาดมิติข้อมูล (Dimensionality Reduction) สำหรับการทำ Dashboard หรือเตรียมข้อมูลสำหรับทำ Supervised learning ต่อไป

3) การเรียนรู้แบบเสริมกำลัง (Reinforcement learning) คือ การเรียนรู้จากการกระทำ การเรียนรู้เกิดมาจากการปฏิสัมพันธ์ (Interaction) ระหว่างผู้เรียนรู้ (Agent) กับสภาพแวดล้อม (Environment) ภายใต้กระบวนการเลือกปฏิบัติในสิ่งต่าง ๆ ให้ได้ผลลัพธ์ที่ดีที่สุดผ่านการกระทำลองผิดลองถูกปรับปรุงความสามารถจากผลลัพธ์ภายใต้ระบบเสมือนจริงหรือการจำลองเหตุการณ์ ที่พัฒนากระบวนการคิดและการตัดสินใจดีขึ้นเรื่อย ๆ คือ การที่กำหนดเงื่อนไขบางอย่าง แล้วทำให้ระบบหรือคอมพิวเตอร์สามารถทำงานบรรลุหรือทำตามเงื่อนไขได้ผลสำเร็จ โดยผู้ทำการทดลอง ตั้งเป้าหมาย Feedback Loop และเงื่อนไขความสำเร็จ รางวัล การลงโทษ (Reward-Penalty) เช่น Gaming Robot, Autonomous Vehicle, Alpha Go เป็นต้น

7 steps of Machine Learning



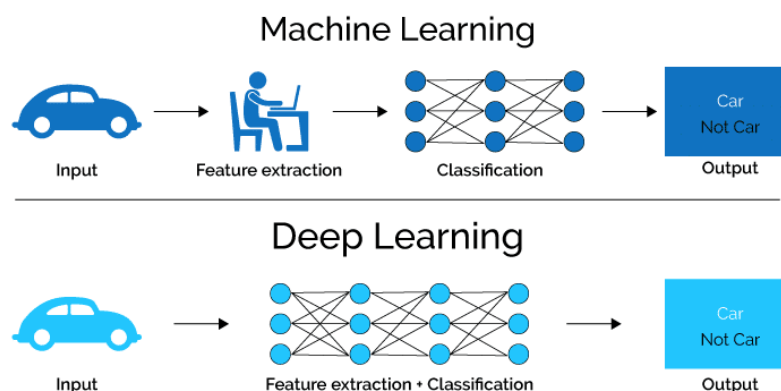
รูปที่ 2.6 การเรียนรู้ของเครื่องจักร 7 ขั้นตอน (อ้างอิง <https://www.mygreatlearning.com/blog/what-is-machine-learning/>)

2.2.2 การเรียนรู้ของเครื่องจักร 7 ขั้นตอน

- 1) การรวบรวม (Gathering Data) ปริมาณและคุณภาพของข้อมูลเป็นตัวกำหนดความแม่นยำของแบบจำลอง
- 2) การจัดเตรียมข้อมูล (Preparing that data) สำหรับข้อมูลที่ใช้ในการฝึกให้จดจำ การเคลียร์ข้อมูลที่ไม่จำเป็น แบ่งเป็นชุดข้อมูลที่ทำาการฝึกให้จดจำและการทดสอบประเมินผล
- 3) การเลือกโมเดล (Choosing a model) การเลือกอัลกอริทึมที่แตกต่างกัน สำหรับงานที่แตกต่างและที่ถูกต้อง
- 4) การฝึกสอนโมเดล (Training) การฝึกโมเดลให้จดจำ คือ การตอบคำถามหรือคำทำนายที่ถูกต้องบ่อยที่สุด การทำซ้ำแต่ละกระบวนการเป็นขั้นตอนการฝึกโมเดลให้จดจำ
- 5) ประเมินการจำลอง (Evaluation) การวัดประสิทธิภาพของกระบวนการตามวัตถุประสงค์ของแบบจำลอง ทดสอบโมเดล ที่ไม่เคยนำมาใช้งาน ข้อมูลที่ไม่เคยป้อนเป็นตัววัดประสิทธิภาพของโมเดล
- 6) การปรับปรุงพารามิเตอร์ (Hypermeter Tuning) ปรับค่าต่าง ๆ ของพารามิเตอร์ เพื่อประสิทธิภาพที่ดีขึ้น เช่น จำนวนรอบการฝึกสอนโมเดลให้จดจำ อัตราการเรียนรู้ ค่าเริ่มต้น และการแจกแจง เป็นต้น
- 7) การคาดการณ์ (Prediction) นำชุดข้อมูลมาใช้งาน ขั้นตอนนี้จะระงับการใช้งานแบบจำลอง เมื่อได้ผลการทดสอบแบบจำลองที่ดีที่สุด

2.3 การเรียนรู้เชิงลึก (Deep Learning : DL)

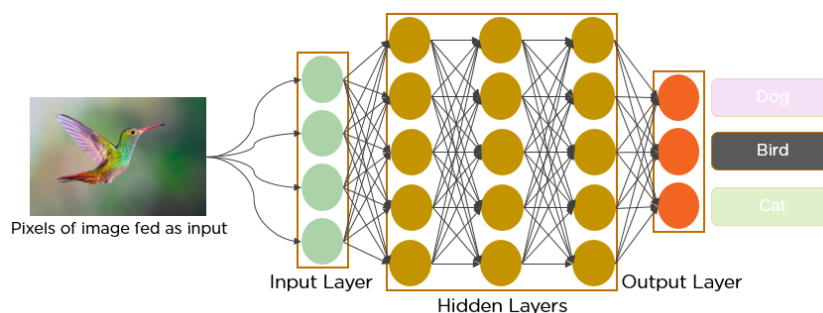
การเรียนรู้เชิงลึก (Deep learning : DL) วิธีการเรียนรู้แบบอัตโนมัติด้วยการเลียนแบบการทำงานของโครงข่ายประสาทของมนุษย์ (Neural Network) มาซ้อนกันหลายชั้น (Layer) และทำการเรียนรู้ชุดข้อมูลตัวอย่าง ซึ่งข้อมูลดังกล่าวจะถูกนำไปในการตรวจจับรูปแบบ (Pattern) หรือจัดหมวดหมู่ชุดข้อมูล (Classify the data) โครงข่ายที่มีความซับซ้อนและลึกมากขึ้น เป็นที่มาของคำว่า Deep learning จะทำการรับข้อมูลทันทีและทำการประมวลผลอัตโนมัติเพื่อหาข้อมูลตัวอย่างที่จำเป็นในการตรวจจับ รูปแบบหรือจัดหมวดหมู่ข้อมูล ความสามารถในการเรียนรู้ คุณลักษณะอัตโนมัติทำให้ Deep learning เป็นประโยชน์อย่างยิ่ง สำหรับการใช้งานในสถานะการณ์ต่าง ๆ คือ การหาโครงข่ายประสาทเทียมที่เหมาะสมกับตัวแปรและค้นหาตัวแปรที่มีผลต่อสมรรถนะในการสอน (Training Performance) คือ ยังไม่สามารถจัดการข้อมูลรับเข้าที่มีความละเอียดเฉพาะทาง (Carefully Crafted Input) อาจทำให้โมเดล เกิดการอนุมานผิดพลาด (Wrong Inferences)



รูปที่ 2.7 เปรียบเทียบการเรียนรู้ของเครื่องจักรและการเรียนรู้เชิงลึก (อ้างอิง <https://fastwork.co/user/yamntd/data-science-44404281>)

2.4 โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution Neural Network : CNN)

การเรียนรู้เชิงลึกโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution Neural Network) ส่วนใหญ่ใช้ในการวิเคราะห์ภาพ 2 มิติ ใช้เทคนิคพิเศษที่เรียกว่า Convolution แทนที่จะพึ่งพาการคูณเมทริกซ์เพียงอย่างเดียว เหมือนโครงข่ายประสาทเทียมแบบเดิม โครงข่าย Convolutional ใช้วิธีการที่เรียกว่า Convolution ซึ่งรวม 2 ฟังก์ชันเข้าด้วยกัน เพื่อแสดงให้เห็นว่าฟังก์ชันหนึ่งเปลี่ยนรูปร่างของอีกฟังก์ชันหนึ่งเป็นอย่างไร สิ่งที่สำคัญ คือ การลดรูปภาพให้เป็นรูปแบบที่ประมวลผลได้ง่ายขึ้น โดยไม่สูญเสียลักษณะเด่นของรูปภาพที่สำคัญสำหรับการทำนายที่ดี Convolutional (CNN) มีสามารถในการจดจำตัวเลขที่เขียนด้วยลายมือ อ่านตัวอักษร สิ่งสำคัญเกี่ยวกับการเรียนรู้เชิงลึก คือ ต้องใช้ข้อมูลจำนวนมากในการฝึกและยังต้องใช้ทรัพยากรคอมพิวเตอร์จำนวนมาก ต้องทำความเข้าใจพื้นฐาน เช่น รูปภาพ RGB เมทริกซ์ของค่าพิกเซลที่มีสามระนาบ รูปภาพระดับสีเทาจะมีระนาบเดียว



รูปที่ 2.8 สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบคอนโวลูชัน (อ้างอิง <https://www.analyticsvidhya.com/blog/2021/05/convolutional-neural-networks-cnn/>)

2.4.1 การแยกคุณลักษณะเด่น (Feature Extraction)

การแยกคุณลักษณะเป็นกระบวนการสำคัญในการเรียนรู้ของเครื่องจักรที่ช่วยลดความซับซ้อนของข้อมูลในขณะที่รักษาข้อมูลสำคัญไว้ ช่วยลดจำนวนทรัพยากรที่จำเป็นสำหรับการประมวลผล ทำให้การวิเคราะห์ข้อมูลมีประสิทธิภาพมากขึ้นยังคงความถูกต้องและแม่นยำไว้ เป็นระบบคำนวณที่สอดคล้องกับกระบวนการคิด (Concept) และต้องคำนวณตัวเลขมารองรับ โดยทำการคำนวณใช้หลักการเดียวกันกับคอนโวลูชันเชิงพื้นที่ การประมวลผลภาพ เริ่มจากกำหนดค่าตัวกรอง (Filter) หรือ เคอร์เนล (Kernel) ทำหน้าที่ช่วยลดขนาดมิติของข้อมูลที่ไม่ต้องการออก สามารถลดข้อมูลออกมาได้ 1 อย่าง ดังนี้ ต้องมีตัวกรองหลาย ๆ ตัว เพื่อใช้หามิติข้อมูลหรือคุณลักษณะพื้นที่ประกอบกัน

1	1	1	0	0
0	1	1	1	0
0	0	1	1	1
0	0	1	1	0
0	1	1	0	0

Input

1	0	1
0	1	0
1	0	1

Filter / Kernel

รูปที่ 2.9 ตัวกรอง (Filter) หรือ เคอร์เนล (Kernel) (อ้างอิง <https://towardsdatascience.com/applied-deep-learning-part-4-convolutional-neural-networks-584bc134c1e2>)

1) ลักษณะของตัวกรอง (Filter) รูปภาพปกติ ลักษณะจะเป็นตาราง 2 มิติ ด้านกว้างและด้านยาว โดยมีการแบ่งตารางออกเป็นพื้นที่ย่อย ๆ ทำการพิจารณาในตำแหน่งกึ่งกลางของตาราง Input สีฟ้า คือ Anchor นำไปหาบบนพิกเซลของข้อมูลรูปภาพ 2 มิติ ที่นำเข้ามา ตัวกรอง (Filter) จะนำไปหาบบนพิกเซลแรกของข้อมูลรูปภาพที่นำเข้ามา (Input) จากนั้นจะถูกเลื่อน ไปหาบบนพิกเซลอื่น ๆ จนครบทุกพิกเซล ภายในรูปภาพเดียวกันเรียกว่า ผังคุณลักษณะ (Feature map) การกรองภาพเป็นเทคนิคที่ใช้ในการประมวลผลภาพเพื่อปรับปรุงหรือแก้ไขลักษณะของภาพโดยใช้ตัวกรอง (Filter) หรือ เคอร์เนล (Kernel) กับแต่ละพิกเซล กระบวนการนี้จะปรับค่าพิกเซลตามพิกเซลข้างเคียงเพื่อให้ได้เอฟเฟกต์ที่ต้องการ เช่น การลดสัญญาณรบกวน การตรวจจับขอบ หรือการ

ทำให้ภาพคมชัดขึ้น ตัวกรอง (Filter) จะกำหนดน้ำหนักที่ใช้กับพิกเซลข้างเคียง ซึ่งจะส่งผลต่อการแปลงภาพในระหว่างกระบวนการกรอง

1	1x1	1x0	0x1	0
0	1x0	1x1	1x0	0
0	0x1	1x0	1x1	1
0	0	1	1	0
0	1	1	0	0

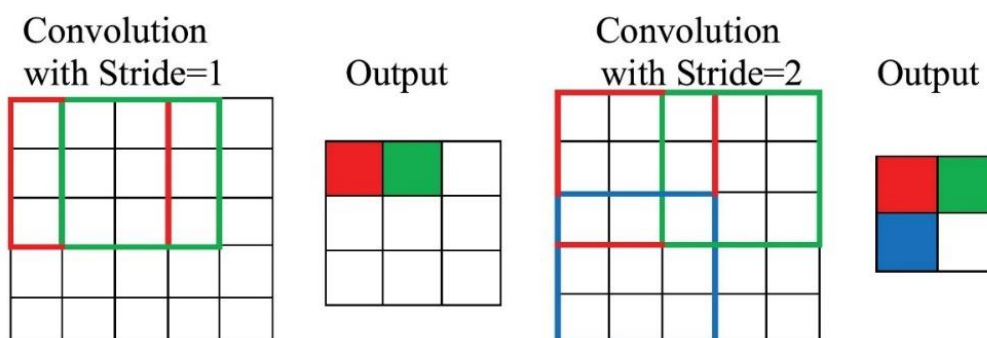
Input x Filter

4	3	

Feature Map

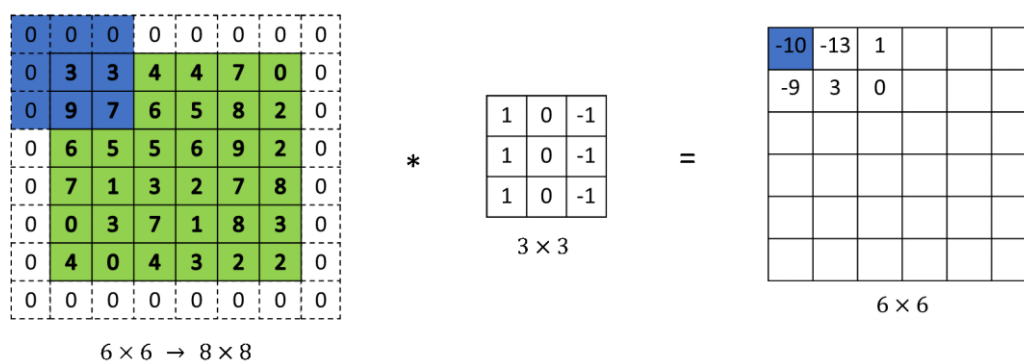
รูปที่ 2.10 ผังคุณลักษณะ (Feature map) (อ้างอิง <https://towardsdatascience.com/applied-deep-learning-part-4-convolutional-neural-networks-584bc134c1e2>)

2) Stride และ Padding ความหมายของ Stride คือ เจื่อนไขในการเลื่อนตัวกรอง (Filter) เป็นตัวกำหนดให้เลื่อนไปจำนวนกี่ช่อง หรือ Step เท่าไร ตัวอย่างที่ รูปที่ 2.11 Stride เช่น ด้านล่างกำหนด Stride เท่ากับ 1 จะทำการเลื่อนไปที่ละ 1 ช่อง ผลลัพธ์คือลดขนาดภาพลง 1 pixel หากกำหนด Stride เท่ากับ 2 จะทำการเลื่อนไปที่ละ 2 ช่อง ผลลัพธ์คือลดขนาดภาพลง 2 pixel พื้นที่ที่มีทับซ้อนกันมีค่าลดลง ขนาดรูปภาพก็จะเล็กลง



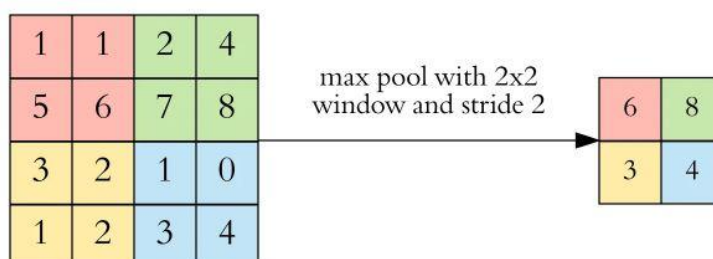
รูปที่ 2.11 Stride (อ้างอิง <https://www.analyticsvidhya.com/blog/2022/03/basics-of-cnn-in-deep-learning/>)

Padding เป็นการเพิ่มพื้นที่โดยรอบของรูปภาพโดยอาจจะเติม ศูนย์ หรือ ตัวเลขอื่น ๆ เข้าไป เพื่อให้กระบวนการทำขั้นตอนของ CNN ส่งผลให้ฟังก์ชันลักษณะ ได้ขนาดเท่ากับ Input



รูปที่ 2.12 Padding (อ้างอิง <https://datahacker.rs/what-is-padding-cnn/>)

3) Pooling หลังจากการปรับขนาดข้อมูลหรือการบิด จะทำการรวมเพื่อทำการปรับปรุงปริมาณจำนวนชุดข้อมูลที่ผ่านมาขั้นตอน Convolution Layer ให้ปริมาณลดลง ผ่านไปยัง Layer ถัดไป ให้คงคุณลักษณะเด่น การทำ Pooling มีอยู่ 2 รูปแบบ ได้แก่ Max Pooling และ Average Pooling, Max Pooling เป็นตัวกรองที่ใช้ค้นหาค่ามากที่สุดในบริเวณที่ตัวกรองทำการทาบบนข้อมูล มาเป็นผลลัพธ์ ส่วน Average Pooling เป็นตัวกรองที่ใช้ค้นหาค่าเฉลี่ยในบริเวณที่ตัวกรองทำการทาบบนข้อมูล มาเป็นผลลัพธ์ใหม่ และทำการเลื่อนตัวกรองไปตาม Stride ที่กำหนดค่าสำหรับทำการทดสอบ ซึ่งขนาดของตัวกรองในการทำ Max pooling เรียกว่า Pool size และรูปแบบ Mean Pooling จะเป็นตัวกรองที่ค่าเฉลี่ยในบริเวณที่ตัวกรองทาบบนข้อมูลมาเป็นผลลัพธ์



รูปที่ 2.13 การทำ Max pooling (อ้างอิง <https://towardsdatascience.com/applied-deep-learning-part-4-convolutional-neural-networks-584bc134c1e2>)

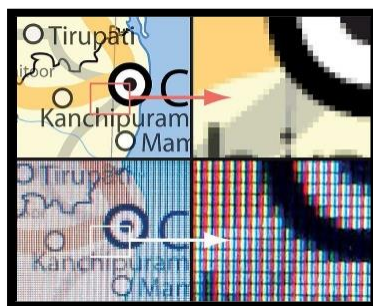
2.5 การประมวลผลภาพ (Image Processing)

การประมวลผลภาพ (Image Processing) คือ กระบวนการแปลงภาพเป็นรูปแบบดิจิทัล และดำเนินการบางอย่างเพื่อรับข้อมูลที่มีประโยชน์จากภาพ โดยทั่วไประบบประมวลผลภาพ 2 มิติ การประมวลผลภาพมี 5 ประเภท ได้แก่ การสร้างภาพค้นหาวัตถุที่ไม่ปรากฏในภาพ การจดจำ แยกแยะหรือตรวจจับวัตถุในภาพ การทำให้คมชัดและฟื้นฟูสร้างภาพที่ได้รับการปรับปรุงจากภาพ ต้นฉบับ การจดจำรูปแบบ วัตถุรูปแบบต่าง ๆ รอบ ๆ วัตถุในภาพ การดึงข้อมูลเรียกดูและค้นหาภาพ จากฐานข้อมูลขนาดใหญ่ของภาพดิจิทัลที่คล้ายกับภาพต้นฉบับ

ขั้นตอนการประมวลผลภาพพื้นฐาน

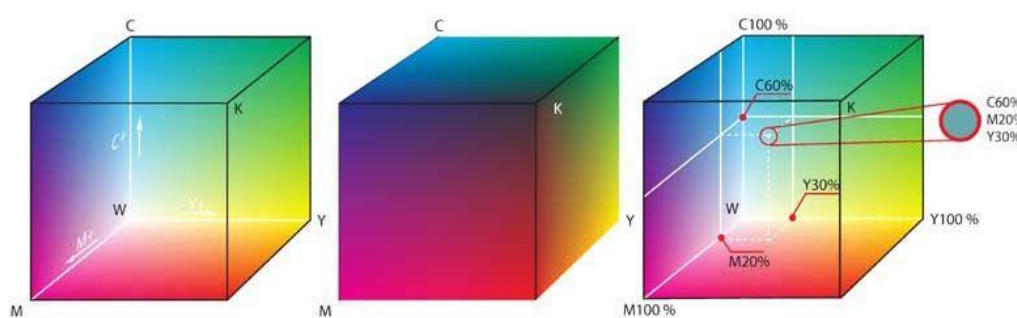
- 1) การรับภาพ (Image acquisition) ขั้นตอนแรกของการประมวลผลภาพ เรียกว่า การประมวลผลล่วงหน้าในการประมวลผลภาพ ซึ่งเกี่ยวข้องกับการดึงภาพจากแหล่งฮาร์ดแวร์
- 2) การปรับปรุงภาพ (Image enhancement) คือกระบวนการนำเสนอและเน้น คุณลักษณะบางอย่างที่น่าสนใจในภาพที่ถูกบดบัง เช่น การปรับความสว่าง ความคมชัด ฯลฯ
- 3) การฟื้นฟูภาพ (Image restoration) คือกระบวนการปรับปรุงลักษณะของภาพ แตกต่างจากการปรับปรุงภาพ การฟื้นฟูภาพโดยใช้แบบจำลองทางคณิตศาสตร์และความน่าจะเป็น
- 4) การประมวลผลภาพสี (Color image processing) เทคนิคการสร้างแบบจำลองสี จำนวนหนึ่งในโดเมนดิจิทัล มีความสำคัญเนื่องจากการใช้ภาพดิจิทัลบนอินเทอร์เน็ตเป็นจำนวนมาก
- 5) การประมวลผลเวฟเลต (Wavelets) และความละเอียดหลายระดับเวฟเลตใช้เพื่อ แสดงภาพในระดับความละเอียดต่าง ๆ ภาพจะถูกแบ่งย่อยเป็นเวฟเลตหรือพื้นที่ที่เล็กกว่าสำหรับการ บีบอัดข้อมูลและสำหรับการแสดงแบบพีระมิด
- 6) การบีบอัด (Compression) คือกระบวนการที่ใช้เพื่อลดพื้นที่จัดเก็บที่จำเป็นในการ บันทึกภาพหรือแบนด์วิดท์ในการส่งข้อมูล
- 7) การประมวลผลทางสัณฐานวิทยา (Morphological processing) คือการ ดำเนินการประมวลผลสำหรับการแปลงภาพตามรูปร่าง
- 8) การแบ่งส่วน (Segmentation) เป็นขั้นตอนที่ยากที่สุดของการประมวลผลภาพ ซึ่ง เกี่ยวข้องกับการแบ่งภาพออกเป็นส่วนประกอบหรือวัตถุ
- 9) การแสดงและคำอธิบาย (Representation and Description) หลังจากที่ภาพถูก แบ่งส่วนแล้ว แต่พื้นที่จะถูกแสดงและอธิบายในรูปแบบที่เหมาะสมสำหรับการประมวลผลด้วย คอมพิวเตอร์ การแสดงผลลักษณะของภาพและคุณสมบัติในพื้นที่ต่างๆ สกัดข้อมูลเชิงปริมาณช่วย แยกแยะวัตถุประเภทหนึ่งจากอีกประเภทหนึ่ง
- 10) การจดจำ (Recognition) กำหนดป้ายกำกับให้กับวัตถุและคำอธิบายของวัตถุ

2.5.1 จุดภาพ หรือ พิกเซล (Pixel) พิกเซลเป็นหน่วยที่เล็กที่สุดของภาพดิจิทัลหรือกราฟิก แสดงบนจอแสดงผลดิจิทัล พิกเซลทำหน้าที่เป็นส่วนประกอบพื้นฐานของภาพดิจิทัล ซึ่งใช้แสดงข้อความไปจนถึงกราฟิกและภาพถ่ายที่ซับซ้อน ภาพดิจิทัลทุกภาพประกอบด้วยพิกเซล โดยแต่ละพิกเซลจะแสดงจุดสีเดียว เช่น จุดหรือสีเหลี่ยมบนจอแสดงผล ความหนาแน่นของพิกเซลตอนนี้ (PPI) จะกำหนดความละเอียด โดย PPI ที่สูงขึ้นจะทำให้ภาพมีความคมชัดมากขึ้น จะระบุจำนวนพิกเซลแนวนอน x แนวตั้ง เช่น 1366 x 768 pixel



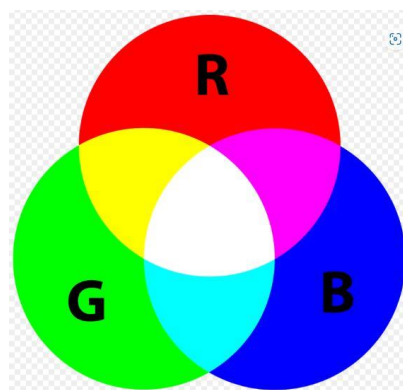
รูปที่ 2.14 จุดภาพ หรือ พิกเซล (pixel) (อ้างอิง <https://th.wikipedia.org/wiki/พิกเซล>)

2.5.2 ระบบสี CMYK คือ ระบบสีที่ใช้งานกับเครื่องพิมพ์ (Printer) ความหมายของระบบสี CMYK ย่อมาจาก cyan (สีฟ้าอมเขียว) magenta (สีแดงอมม่วง) yellow (สีเหลือง) key (สีดำ) ซึ่งเป็นชื่อของสี ที่นำมาใช้ในการผสมทั้ง 4 สี ทำให้เกิดสีได้อีกหลาย 100 สี หากเลือกค่าสี RGB ให้กับเครื่องพิมพ์ สีที่ออกมาจะผิดเพี้ยนไปจากหน้าจอ ต้องกำหนดระบบสี CMYK เช่น C=100% M=0% Y=100% K=0% จะได้สีถูกต้องที่สำหรับงานพิมพ์ สามารถกำหนดค่าสีของ CMYK ตั้งแต่ 0% – 100% หรือใช้รหัสสีตัวอย่างจาก Chat สี โดยกระบวนการพิมพ์สี CMYK จะเป็นการใช้กระดาษวิ่งผ่านแท่นพิมพ์แต่ละสีเรียงกัน สีจะไหลลงมามาติดกับแผ่นสังกะสี ที่ศัพท์ทางโรงพิมพ์ เรียกว่า เพจ และติดกับลูกยางในเครื่องพิมพ์ จากนั้นกระดาษจะวิ่งผ่านแล้วเกิดรอยพิมพ์แต่ละสี ก็จะเกิดสีตามที่ต้องการ หรือเรียกว่าการพิมพ์ 4 สี หากพิมพ์ 1 สี จะผ่านเพียง 1 แท่น หรือ 2 สี จะผ่านเพียง 2 แท่น แต่ละแท่นคือสีนั่นเอง โปรแกรมที่นิยมใช้กับงานพิมพ์ ได้แก่ Adobe Photoshop Adobe Illustrator หรือ Adobe InDesign



รูปที่ 2.15 ระบบสี CMYK (อ้างอิง https://th.wikipedia.org/wiki/ระบบสี_CMYK)

RGB คือ สีแดง (Red) สีเขียว (Green) และสีน้ำเงิน (Blue) โมเดลสีพื้นฐานที่ใช้ในการออกแบบดิจิทัล โดยการปรับความเข้มของสีทั้งสาม นักออกแบบ จึงสามารถสร้างภาพที่สดใสที่เห็นบนหน้าจอ ไม่ว่าจะเป็นทีวี แล็ปท็อป หรือโทรศัพท์มือถือ การเชี่ยวชาญโมเดลสี RGB จะช่วยให้ได้สีที่แม่นยำและสม่ำเสมอ สีแต่ละช่องมีช่วงความเข้มตั้งแต่ 0 (สีดำ) ถึง 255 (ความสว่างสูงสุด) เมื่อตั้งค่าความเข้มของสีทั้งสามเป็นระดับต่ำสุด R=0 G=0 B=0 ผลลัพธ์ที่ได้จะเป็นสีดำ เมื่อทั้งสามสีมีความเข้มสูงสุด R=255 G=255 B=255 ผลลัพธ์ที่ได้จะเป็นสีขาว เมื่อตั้งค่าสี R=255 G=0 B=0 ผลลัพธ์ที่ได้จะเป็นสีแดง เนื่องจาก RGB ขึ้นอยู่กับแสง จึงใช้กันอย่างแพร่หลายในการออกแบบและเทคโนโลยีดิจิทัล จอแสดงผลใด ๆ ที่เปล่งแสงจะใช้โมเดลสี RGB



รูปที่ 2.16 ระบบสี RGB (อ้างอิง <https://en.wiktionary.org/wiki/RGB>)

2.6 การตรวจจับวัตถุ (Object detection)

การตรวจจับวัตถุ (Object Detection) คือ กระบวนการระบุและระบุตำแหน่งของวัตถุเฉพาะภายในภาพหรือวิดีโอ โดยทำงานโดยตรวจจับวัตถุใช้กรอบล้อมรอบและกำหนดป้ายกำกับ

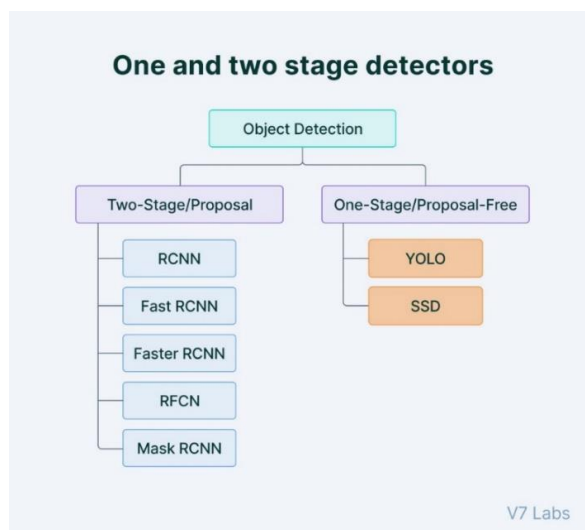
คลาสให้กับวัตถุที่ตรวจพบแต่ละรายการ วิธีนี้ช่วยให้แบบจำลองสามารถระบุได้ว่าวัตถุคืออะไรและอยู่ตำแหน่งใดภายในภาพ ส่วนประกอบหลักของการตรวจจับวัตถุ การตรวจจับวัตถุประกอบด้วยงานหลักสองอย่าง ได้แก่ การจำแนกประเภทภาพ (Image Classification) และการระบุตำแหน่งของวัตถุ (Object Localization), การจำแนกประเภทภาพ (Image Classification) กำหนดประเภทหรือคลาสของวัตถุภายในภาพอินพุตคือภาพที่มีวัตถุชิ้นเดียว (เช่น แมว) เอาต์พุตคือป้ายกำกับคลาสพร้อมคะแนนความน่าจะเป็นที่แสดงถึงความมั่นใจในการทำนาย การระบุตำแหน่งของวัตถุ (Object Localization) ผลลัพธ์ประกอบด้วยกรอบขอบเขตที่กำหนดตำแหน่งวัตถุ โดยใช้ตำแหน่ง ความสูง และความกว้างทำงานกับรูปภาพที่มีวัตถุหนึ่งชิ้นหรือหลายชิ้น การรวมงานทั้งสองนี้เข้าด้วยกัน ทำให้การตรวจจับวัตถุช่วยให้การเรียนรู้ของเครื่องจักรสามารถวิเคราะห์ภาพได้แบบเรียลไทม์ ทำให้เป็นเทคโนโลยีที่จำเป็นในแอปพลิเคชันต่างๆ เช่น การเฝ้าระวัง การขับเคลื่อนอัตโนมัติ และการถ่ายภาพทางการแพทย์ เป็นต้น

วิธีการตรวจจับวัตถุสามารถแบ่งได้ 2 ประเภท วิธีการที่ใช้โครงข่ายประสาทและไม่ใช่โครงข่ายประสาท

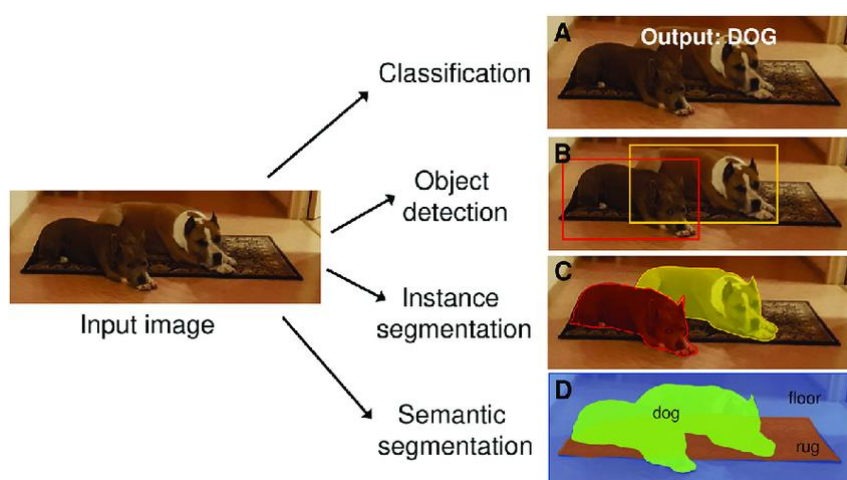
1) วิธีการที่ไม่ใช่โครงข่ายประสาทตรวจจับวัตถุ ต้องใช้การออกแบบคุณลักษณะ ซึ่งลักษณะเฉพาะของวัตถุจะถูกกำหนดขึ้นด้วยตนเองก่อนที่จะใช้อัลกอริทึมการจำแนกประเภท เช่น Support Vector Machines (SVM) วิธีการเหล่านี้มีความแม่นยำน้อยกว่าวิธีการเรียนรู้เชิงลึก แต่ยังคงใช้ในแอปพลิเคชันบางประเภท วิธีการที่ไม่ใช่โครงข่ายประสาท ได้แก่ การตรวจจับวัตถุของ Viola-Jones ใช้คุณลักษณะ Haar สำหรับการตรวจจับใบหน้าและวัตถุ การแปลงคุณลักษณะที่ไม่แปรผันตามมาตราส่วน (SIFT) ระบุและจับคู่จุดสำคัญในภาพโดยไม่คำนึงถึงมาตราส่วนและการหมุน-ฮิสโทแกรมของการไล่ระดับสีตามทิศทาง (HOG) ดึงคุณลักษณะการไล่ระดับสีจากภาพ ซึ่งมักใช้เพื่อตรวจจับบุคคลการจับคู่เทมเพลต รูปร่าง หรือสี เปรียบเทียบวัตถุกับเทมเพลตหรือรูปแบบที่กำหนดไว้ล่วงหน้า

2) วิธีการที่ใช้โครงข่ายประสาทสามารถตรวจจับวัตถุได้แบบครบวงจรโดยไม่ต้องกำหนดคุณลักษณะด้วยตนเอง วิธีการเหล่านี้มีความแม่นยำมากกว่าอย่างเห็นได้ชัดและอาศัยโครงข่ายประสาทเทียมแบบคอนโวลูชัน (CNN) เพื่อแยกคุณลักษณะจากภาพโดยอัตโนมัติ แบบจำลองการตรวจจับวัตถุที่ใช้การเรียนรู้เชิงลึกที่ได้รับความนิยม แบบ 2 ขั้นตอน ได้แก่ R-CNN, Fast R-CNN, Faster R-CNN, Mask R-CNN แบบจำลองเหล่านี้ใช้เทคนิคการเสนอพื้นที่เพื่อระบุวัตถุและจัดประเภทด้วย CNN ตัวตรวจจับ แบบครั้งเดียว ได้แก่ (SSD) ตัวตรวจจับแบบขั้นตอนเดียวที่รวดเร็วและมีประสิทธิภาพซึ่งทำนายกล่องขอบเขตและป้ายชื่อคลาสในครั้งเดียว, RetinaNet ใช้ฟังก์ชันการสูญเสียเฟสเพื่อปรับปรุงการตรวจจับวัตถุขนาดเล็กและตรวจจับได้ยาก You Only Look Once (YOLO) แบบจำลองการตรวจจับวัตถุแบบเรียลไทม์ที่ประมวลผลภาพในครั้งเดียว ทำให้

มีประสิทธิภาพสูงสำหรับแอปพลิเคชัน เช่น การเฝ้าระวังวิดีโอและการขับเคลื่อนอัตโนมัติ โครงข่ายประสาทเทียมได้รับการใช้กันอย่างแพร่หลาย เนื่องจากมีความแม่นยำ ความเร็ว

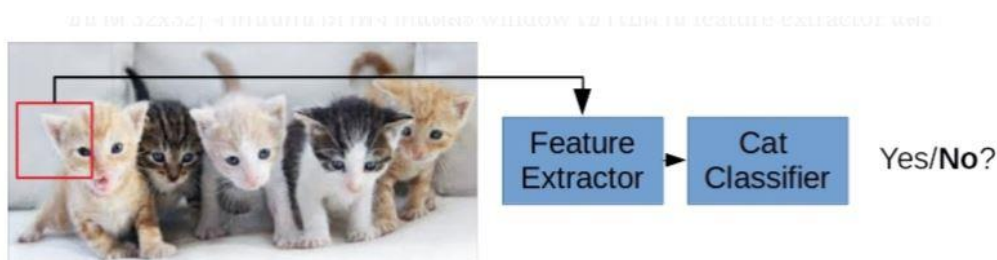


รูปที่ 2.17 One and Two stage Objectors (อ้างอิง <https://www.v7labs.com/blog/object-detection-guide>)



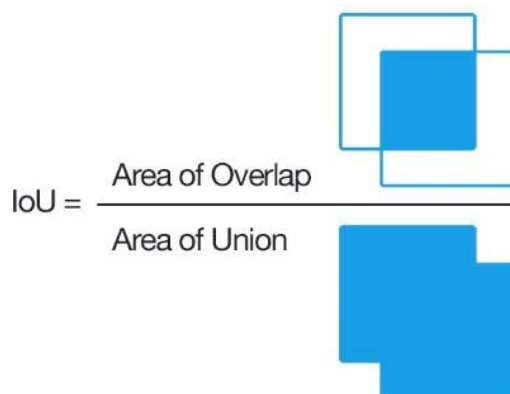
รูปที่ 2.18 Object Classification และ Object detection (อ้างอิง https://www.researchgate.net/figure/Comparison-of-image-classification-object-detection-instanceand-semantic-segmentation_fig5_366553651)

2.6.1 Sliding Window การตรวจจับวัตถุมีความนิยมมากในเมื่อสิบปีที่ผ่านมา วิธีการจะนำผลลัพธ์ในจากขั้นตอนของการจำแนกประเภทของวัตถุ (Object Classification) ด้วยวิธีการทำการตีกรอบพื้นที่สี่เหลี่ยมแล้วขยายกรอบสี่เหลี่ยมไปเรื่อย ๆ จากภาพที่ต้องการค้นหา จะทำการเลื่อน Sliding Window ที่มีขนาดเท่ากับการจำแนกประเภทของภาพ Classifier ที่ต้องการ เช่น ฝึกโมเดล Cat classifier ด้วยรูปภาพแนว 32x32 pixel ก็ใช้กรอบพื้นที่ window ขนาด 32x32 pixel จากนั้นนำรูปภาพ แต่ละกรอบพื้นที่ window เข้าไปผ่าน feature extractor และตรวจสอบกับการจำแนกประเภท classifier ที่ละกรอบพื้นที่สี่เหลี่ยมว่า ตำแหน่งนี้คือรูปภาพแมวหรือไม่ แล้วขยายไปเรื่อย ๆ อาจจะเลื่อนไปที่ละ 1 pixel จนครบทั้งรูปภาพ ซึ่งจะทำให้การตรวจสอบรูปภาพที่ต้องการค้นหาว่า เป็นรูปภาพแมว ใช่ (Yes) หรือ ไม่ใช่ (No)

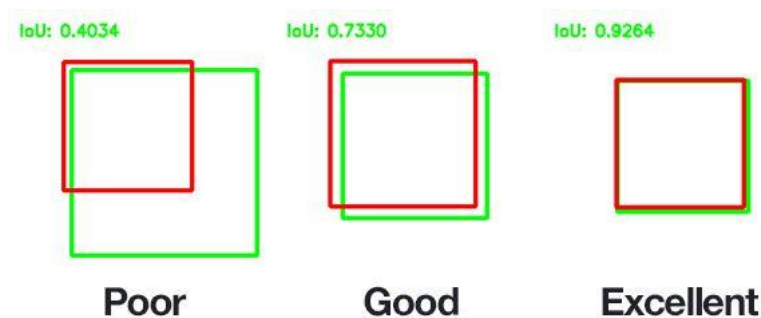


รูปที่ 2.19 Sliding Window (อ้างอิง <https://medium.com/@sapjunior/object-detection-part-1-37bf39dc557f>)

2.6.2 การวัดผล Intersection over Union (IoU) เป็นวิธีการวัดค่าประสิทธิภาพของรูปภาพการตรวจจับวัตถุ (Object Detection) เป็นการใช้อัตราส่วนระหว่างพื้นที่ของวัตถุภายในรูปภาพ (Area) ที่เป็นจุดตัดของวัตถุ (Intersection) ที่ประกอบกันอยู่ในรูปภาพ ของ กรอบสี่เหลี่ยมขอบเขตของวัตถุภายในรูปภาพ (2 bounding box) หารด้วย ผลรวมของพื้นที่ (Area) รวมของกรอบสี่เหลี่ยมทั้ง 2 กรอบ หรือ เรียกว่า ดัชนี (Jaccard) วิธีการคำนวณหาจำนวนเปอร์เซ็นต์พื้นที่รูปภาพที่ทับกันระหว่างผลเฉลย (Ground Truth) และผลลัพธ์ของทำนายผล (Predict) หากได้ค่า IoU ที่มีค่ามากกว่า 0.5 ถึงถือว่ายอมรับได้ ทำการคำนวณพื้นที่ที่ทับซ้อนกันระหว่างกรอบสี่เหลี่ยมขอบเขตที่คาดการณ์ไว้กับกรอบสี่เหลี่ยมขอบเขตที่เป็นจริงของรูปภาพ ถ้าได้ค่า IoU ที่มีค่าเท่ากับ 1.0 แสดงว่ากรอบสี่เหลี่ยมขอบเขตพื้นที่ที่ทับซ้อนกันหมดและเป็นรูปภาพเดียวกัน



รูปที่ 2.20 การวัดผล Intersection over Union (IoU) (อ้างอิง <https://pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/>)



รูปที่ 2.21 ตัวอย่างการวัดผล (IoU) (อ้างอิง <https://pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/>)

2.6.3 Confusion Matrix เป็นวิธีการแสดงผลลัพธ์ของการทำนายผล (Prediction) เป็นการทำนายผลลัพธ์จากโมเดล (Model) ทำการสร้างชุดข้อมูล (Data set) ใช้วิธีการเรียนรู้ของเครื่องจักร (Machine learning) การทำนายวัดผลของโมเดล (Model) กับ สิ่งที่เกิดขึ้นจริง มีสัดส่วนเป็นอย่างไร

True Positive (TP)	หมายถึง สิ่งที่โปรแกรมทำนายว่าจริง และเป็นความจริง คือจริง
True Negative (TN)	หมายถึง สิ่งที่โปรแกรมทำนายว่าไม่จริง และเป็นความจริง คือไม่จริง
False Positive (FP)	หมายถึง สิ่งที่โปรแกรมทำนายว่าจริง แต่ความเป็นจริง คือไม่จริง
False Negative (FN)	หมายถึง สิ่งที่โปรแกรมทำนายว่าไม่จริง แต่ความเป็นจริง คือจริง

Confusion Matrix

	Actually Positive (1)	Actually Negative (0)
Predicted Positive (1)	True Positives (TPs)	False Positives (FPs)
Predicted Negative (0)	False Negatives (FNs)	True Negatives (TNs)

รูปที่ 2.22 Confusion Matrix (อ้างอิง <https://medium.com/@pagongatchalee/confusion-matrix-เครื่องมือสำคัญในการประเมินผลลัพธ์ของการทำนาย-ในmachine-learning-fba6e3f9508c>)

Accuracy คือ ค่าความแม่นยำ เป็นการวิเคราะห์ที่มีค่าใกล้เคียงระหว่างค่าที่ทำการทดสอบ กับค่าจริง ดังสมการที่ 2.1

$$\text{Accuracy} = \frac{(\text{TPs} + \text{TNs})}{(\text{TPs} + \text{TNs} + \text{FPs} + \text{FNs})} \quad (2.1)$$

Precision คือ ค่าความเที่ยงตรงเป็นผลการวิเคราะห์ซึ่งทำการทดลองหลาย ๆ ครั้ง ในสถานะเดียวกัน และได้ค่าใกล้เคียงกัน หรือคือความแม่นยำของการวัดผลซ้ำหลาย ๆ ครั้ง ผลการทำนายว่าจริงและเกิดขึ้นจริง (TP) กับ การทำนายว่า จริง แต่สิ่งที่เกิดขึ้น คือ ไม่จริง (FP) ดังสมการที่ 2.2

$$\text{Precision} = \frac{\text{TPs}}{(\text{TPs} + \text{FPs})} \quad (2.2)$$

Recall หรือ Sensitivity คือ ค่าความถูกต้องของการทำนายผลลัพธ์ว่าจะเป็นจริง เทียบกับจำนวนครั้งของเหตุการณ์ที่ทำนายผลลัพธ์และเกิดขึ้นว่า เป็นจริง ดังสมการที่ 2.3

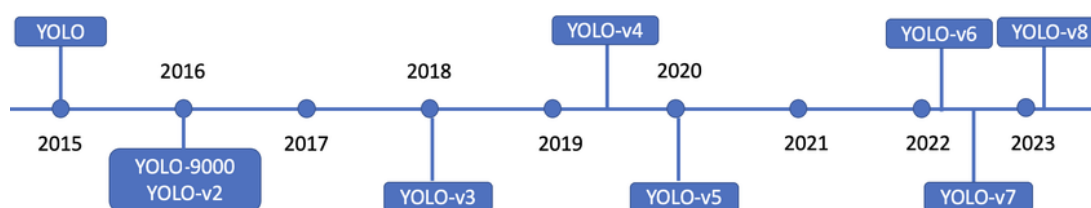
$$\text{Recall} = \frac{2 \times (\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (2.3)$$

F1-Score คือ การนำค่าความเที่ยงตรง (Precision) กับค่าความถูกต้อง (Recall) มาคำนวณรวมกันเป็นค่าเฉลี่ยแบบ Harmonic mean สร้างขึ้นเพื่อเป็น Single metric ที่วัดความสามารถของโมเดล (Model) ดังสมการที่ 2.4

$$F1 = \frac{TPs}{(TPs+FNs)} \quad (2.4)$$

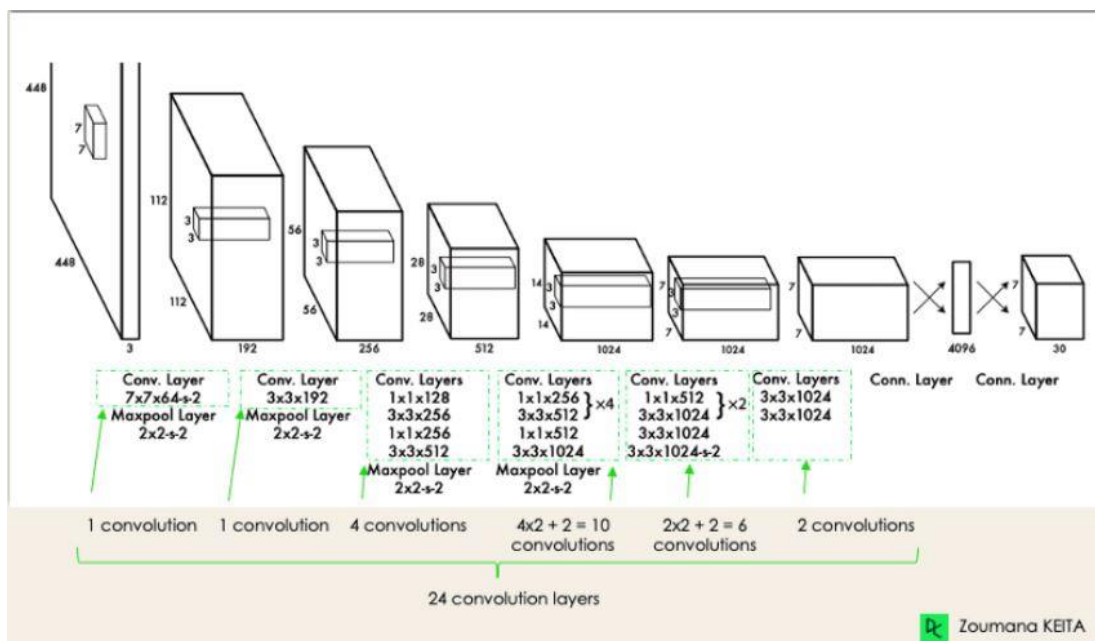
2.7 You Only Look Once (YOLO)

You Only Look Once หรือ YOLO คือ อัลกอริธึม ที่ทำการปรับปรุงมาจาก CNN ซึ่งเป็น อัลกอริธึมของการเรียนรู้เชิงลึก มีการพัฒนาความสามารถในการตรวจจับรูปภาพได้อย่างมีความแม่นยำสูง ความเร็วในการประมวลผล เป็น อัลกอริธึมตรวจจับวัตถุแบบครั้งเดียว (One-Stage Object Detection Algorithms) ที่สามารถตรวจจับรูปภาพได้เร็วที่สุดที่แสดงให้เห็นถึงความมีศักยภาพในการตรวจจับวัตถุทำนายคลาสของวัตถุและกรอบขอบเขตที่กำหนดตำแหน่งของวัตถุบนภาพอินพุต ทำการเทียบกับวิธีการอื่น ที่สามารถตรวจจับวัตถุได้แบบเรียลไทม์ ซึ่งมีการปรับปรุงและพัฒนา ดังรูปที่ 2.23



รูปที่ 2.23 Timeline of You Look Once (YOLO) variants (อ้างอิง https://www.researchgate.net/figure/Timeline-of-You-Only-Look-Once-YOLO-variants_fig1_369379818)

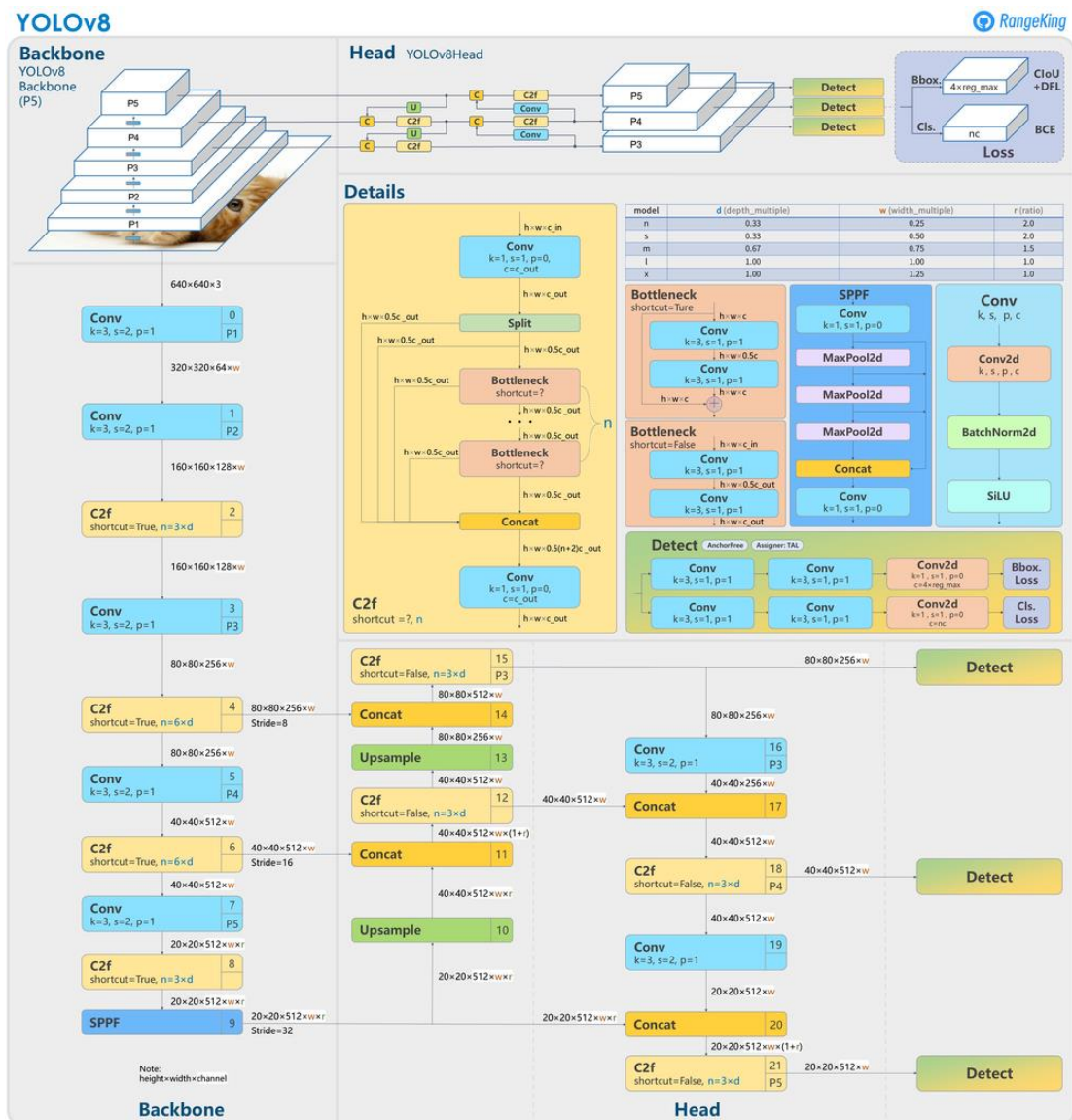
โครงสร้างของ You only look once (YOLO) ที่เป็นชั้น ๆ แบบ Network แต่ละชั้นมีกระบวนการที่หลากหลายมีความถี่เพิ่มขึ้นเรื่อย ๆ สามารถตรวจจับภาพวัตถุที่ซ้อนกัน แต่ละชั้นที่มีขนาดเล็กลงไปเรื่อย ๆ ในแต่ละ Layers ดังรูปที่ 2.24



รูปที่ 2.24 สถาปัตยกรรมของ YOLO (อ้างอิง <https://www.datacamp.com/blog/yolo-object-detection-explained>)

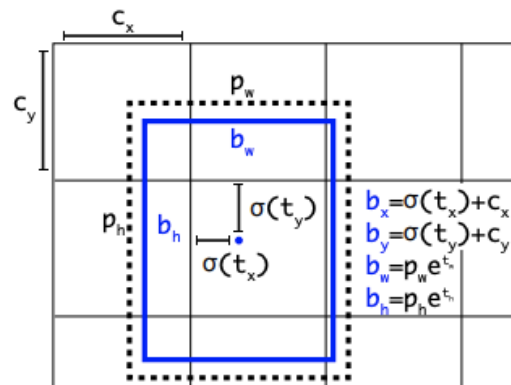
ในปัจจุบัน YOLO ได้พัฒนามาถึงเวอร์ชัน 8 แล้ว ซึ่งพัฒนามาจาก YOLO เวอร์ชัน 5 สามารถใช้ในการตรวจจับวัตถุ การจัดกลุ่ม จัดหมวดหมู่ของรูปภาพ มีพัฒนาประสิทธิภาพการให้ดีขึ้นจากเวอร์ชันที่ผ่านมา แต่ยังคงสามารถทำงานที่มีเอกลักษณ์ในความรวดเร็วของการประมวลผลตรวจจับวัตถุแบบเรียลไทม์ เหตุผลที่ YOLO นั้นความสามารถในการตรวจจับวัตถุที่รวดเร็ว มากกว่าโมเดลประเภท Region based อื่น ๆ เนื่องจากโครงสร้างโมเดลของ YOLO เป็นแบบ Pass through image ส่งผ่านในรอบเดียวโดยผ่าน CNN และทำนายผลลัพธ์ค้นหาตำแหน่งการตรวจจับวัตถุที่สนใจ โดยที่ YOLOv8 จะมีอัลกอริทึม ที่เหมาะสมกับการใช้งานของโครงการการทดลอง

- 1) YOLOv8 มีความแม่นยำสูง โดยการวัดจาก COCO และ Roboflow
- 2) YOLOv8 มีฟีเจอร์สำหรับนักพัฒนา ตั้งแต่ CLI ที่ง่าย ๆ และใช้งานได้กับ Python
- 3) YOLOv8 นิยมกับผู้ใช้งานคอมพิวเตอร์วิทัศน์ (Computer Vision)



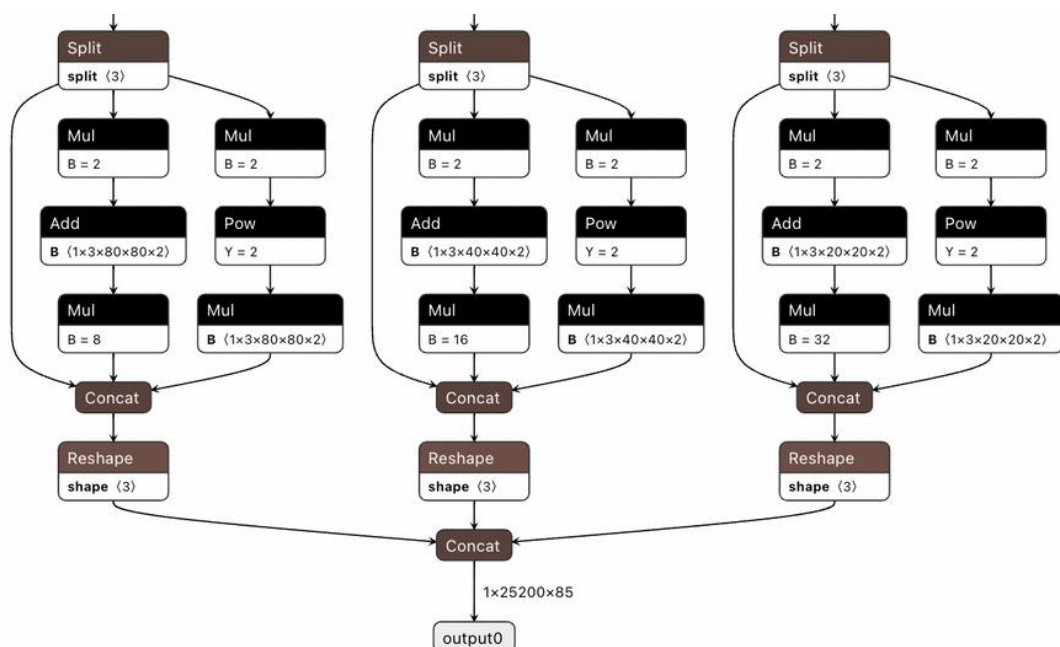
รูปที่ 2.25 สถาปัตยกรรมของ YOLOv8 (อ้างอิง <https://blog.roboflow.com/whats-new-in-yolov8/>)

ระบบ YOLOv8 การตรวจจับแบบไม่มีจุดในการยึด จะทำการคาดการณ์จุดศูนย์กลางของวัตถุโดยตรง



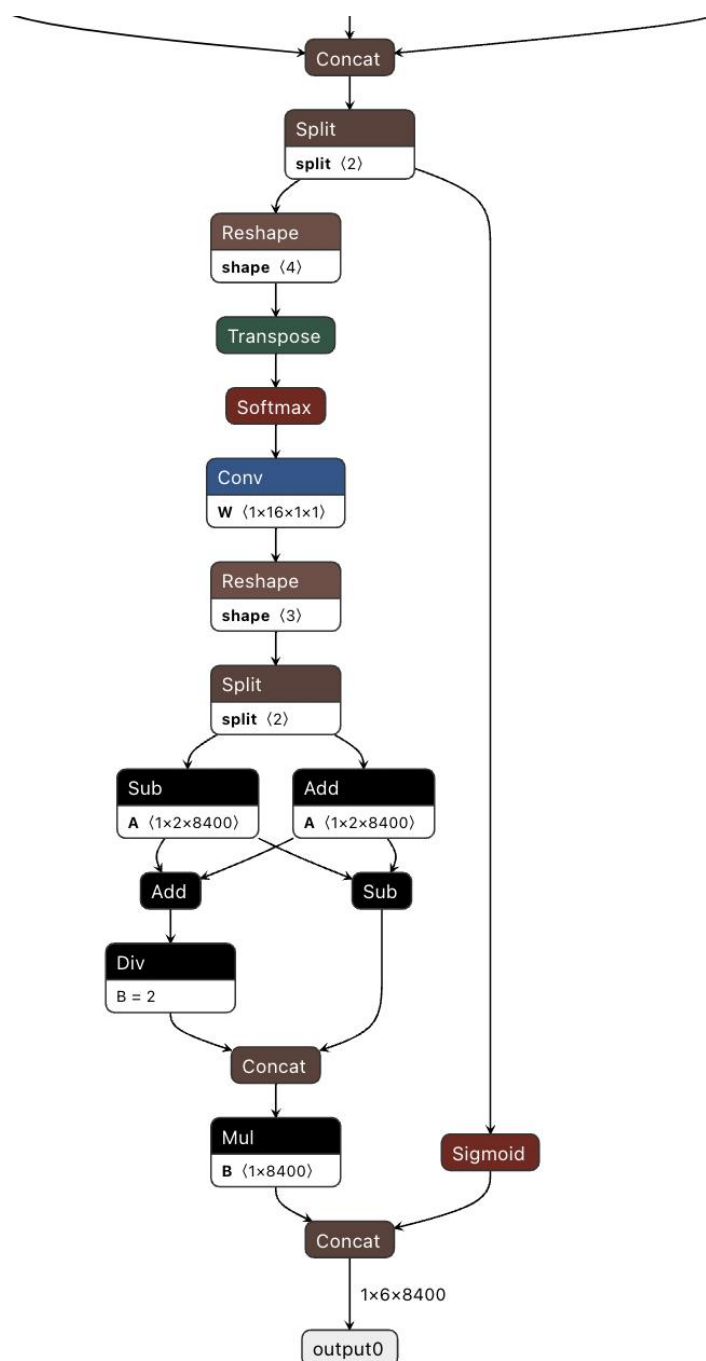
รูปที่ 2.26 Anchor Free Detection (อ้างอิง <https://blog.roboflow.com/whats-new-in-yolov8/>)

Anchor Box เป็นส่วนที่เป็นจุดบกพร่องของ YOLO รุ่นก่อน เนื่องจากชุดข้อมูลไม่สามารถกระจายตัวเป็นไปตามเกณฑ์มาตรฐานที่กำหนด



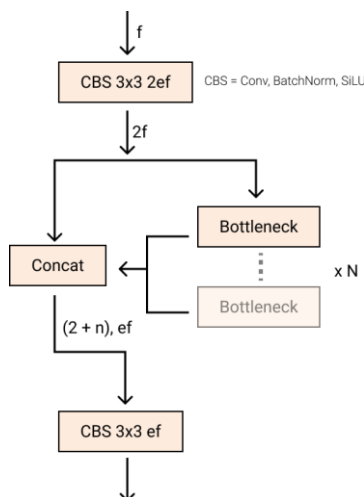
รูปที่ 2.27 The detection head of YOLOv5 (อ้างอิง <https://blog.roboflow.com/whats-new-in-yolov8/>)

การตรวจจับแบบไม่มียี่ดเกาะ (Anchor free detection reduces) ช่วยลดจำนวนการ
คาดการณ์ Number of box predictions ช่วยเพิ่มความเร็ว Non-Maximum Suppression (NMS)
ขั้นตอนการประมวลผลที่ซับซ้อน ผ่านการกรองและตรวจจับ



รูปที่ 2.28 Anchor free detection reduces (อ้างอิง <https://blog.roboflow.com/whats-new-in-yolov8/>)

คอนโวลูชันใหม่ (New Convolution) ทำการเริ่มต้นด้วย Conv 6x6 และถูกแทนที่ด้วย 3x3 โครงสร้างหลักเปลี่ยนไป และ C2f แทนที่ด้วย C3 โดยที่ “f” หมายถึงเลขลักษณะ “e” คือ อัตราการขยายตัว และ CBS ในบล็อกที่ประกอบด้วย Conv, BatchNorm and SiLU ดังรูปที่ 2.29



รูปที่ 2.29 คอนโวลูชันใหม่ (New Convolution) (อ้างอิง <https://blog.roboflow.com/whats-new-in-yolov8/>)

Closing the Mosaic Augmentation การวิจัยการเรียนรู้เชิงลึกมีแนวโน้มที่มุ่งเน้นสถาปัตยกรรมแบบจำลอง แต่การฝึกให้ทำการจดจำใน YOLOv5 และ YOLOv8 เป็นส่วนที่สำคัญ YOLOv8 สามารถเพิ่มรูปภาพระหว่างการฝึกสอนโมเดลแบบออนไลน์ได้ YOLO ในแต่ละรุ่น โมเดลจะมองเห็นรูปภาพที่แตกต่างกันเล็กน้อยจากการเสริมแบบหนึ่ง เรียกว่าการเสริมแบบโมเสก เกี่ยวข้องกับการต่อภาพสี่ภาพเข้าด้วยกัน เพื่อให้โมเดลเรียนรู้วัตถุในตำแหน่งใหม่ การบดบังบางส่วน และเทียบกับพิกเซลบริเวณโดยรอบที่แตกต่างกัน



รูปที่ 2.30 การเสริมแบบโมเสก (อ้างอิง <https://blog.roboflow.com/whats-new-in-yolov8/>)

การวิจัยของ YOLOv8 ได้รับการประเมินตามเกณฑ์มาตรฐาน COCO ได้รับการฝึกสอนโมเดล ได้รับการปรับปรุง การทดลองใหม่ สามารถทำการตรวจสอบความถูกต้องของการปรับปรุงที่มีผลต่อการสร้างแบบจำลอง COCO ความแม่นยำ YOLOv8 COCO (Common object in Context) คือ เกณฑ์มาตรฐานอุตสาหกรรมสำหรับการประเมินโมเดลการตรวจจับวัตถุ เมื่อเปรียบเทียบโมเดลบน COCO จากค่า mAP และการวัด FPS สำหรับความเร็วรอบ การเปรียบเทียบกับแบบจำลองด้วยความเร็วรอบที่ใกล้เคียงกันความแม่นยำของ YOLOv8 บน COCO

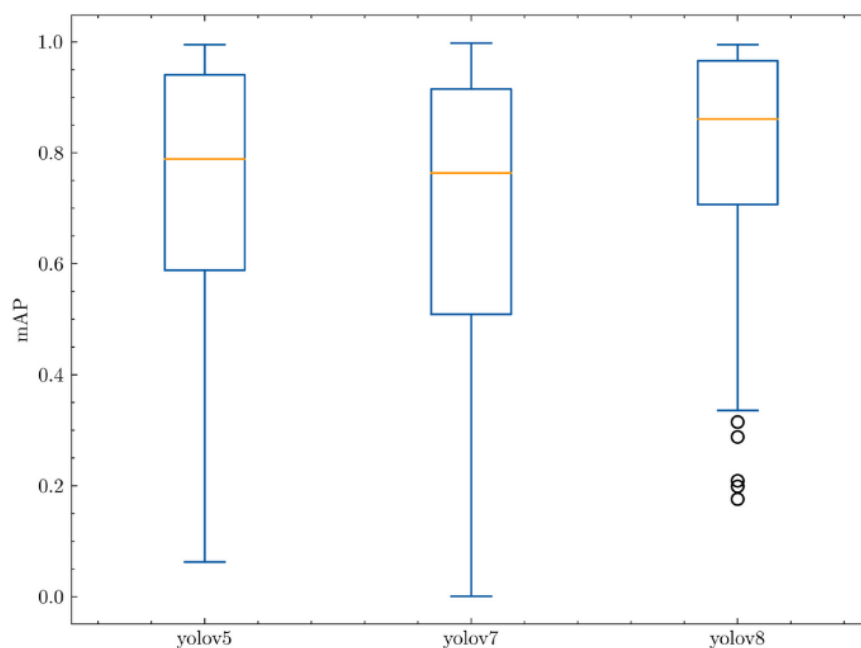
▼ Detection

Model	size (pixels)	mAP ^{val} 50-95	Speed CPU (ms)	Speed T4 GPU (ms)	params (M)	FLOPs (B)
YOLOv8n	640	37.3	-	-	3.2	8.7
YOLOv8s	640	44.9	-	-	11.2	28.6
YOLOv8m	640	50.2	-	-	25.9	78.9
YOLOv8l	640	52.9	-	-	43.7	165.2
YOLOv8x	640	53.9	-	-	68.2	257.8

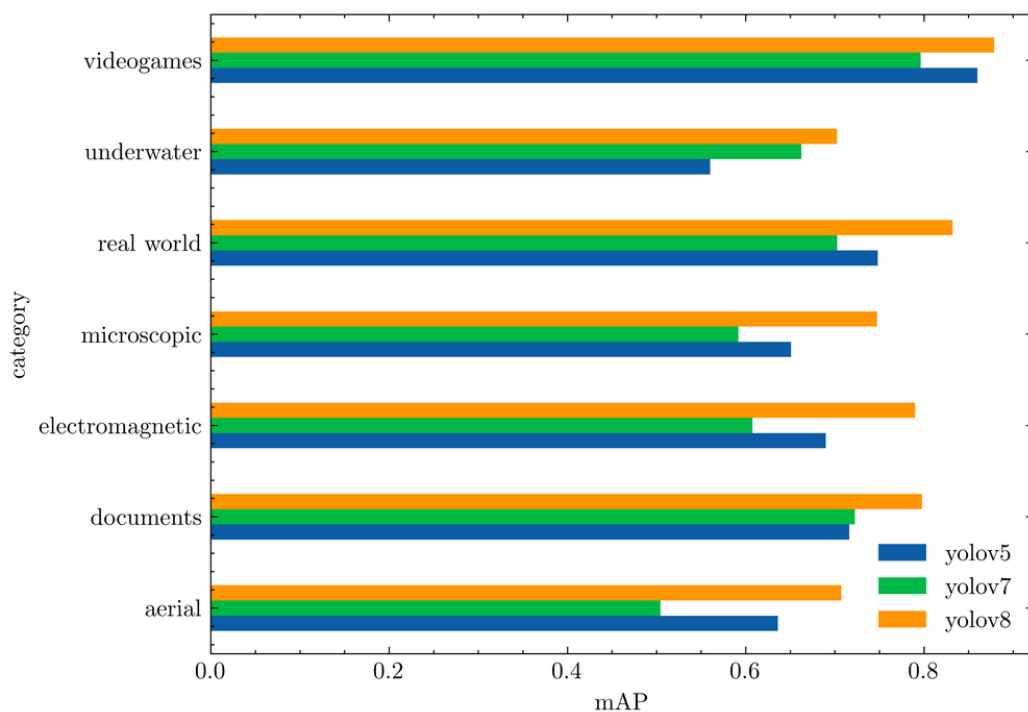
- mAP^{val} values are for single-model single-scale on [COCO val2017](#) dataset.
Reproduce by `yolo mode=val task=detect data=coco.yaml device=0`
- Speed averaged over COCO val images using an [Amazon EC2 P4d](#) instance.
Reproduce by `yolo mode=val task=detect data=coco128.yaml batch=1 device=0/cpu`

รูปที่ 2.31 ความแม่นยำ YOLOv8 (อ้างอิง <https://blog.roboflow.com/whats-new-in-yolov8/>)

ความแม่นยำของ RF100 Roboflow ชุดข้อมูลตัวอย่าง 100 ชุดจาก Roboflow University เป็นที่เก็บข้อมูลมากกว่า 100,000 ชุดข้อมูล เพื่อประเมินโมเดล สามารถสรุปผลภาพรวมของโมเดลตามเกณฑ์มาตรฐานที่ได้รับการสนับสนุนของ Intel เป็นเกณฑ์มาตรฐานสำหรับผู้ปฏิบัติงานด้าน Computer Vision ประเมิน YOLOv8 ตามเกณฑ์มาตรฐาน RF100 ควบคู่ไปกับ YOLOv5 และ YOLOv7 โมเดลแสดงค่า mAP@50



รูปที่ 2.32 YOLOs $mAP@.50$ against RF100 (อ้างอิง <https://blog.roboflow.com/whats-new-in-yolov8/>)



รูปที่ 2.33 YOLOs average $mAP@.50$ against RF100 categories (อ้างอิง <https://blog.roboflow.com/whats-new-in-yolov8/>)

2.8 อากาศยานไร้คนขับ (Unmanned Aerial Vehicle: UAV)

อากาศยานไร้คนขับ (Unmanned Aerial Vehicle: UAV) คือ เครื่องบินที่ไม่มีนักบินประจำประจำการอยู่บนเครื่องบิน และมีรูปลักษณะ UAV ที่แตกต่างกันออกไป โดยใช้หลักอากาศพลศาสตร์ และขับเคลื่อนด้วยมอเตอร์ไฟฟ้าหรือเครื่องยนต์ที่ขับเคลื่อนด้วยก๊าซ ต่างจากอากาศยานที่มีคนขับ อากาศยานไร้คนขับออกแบบและผลิตให้มีความคล่องตัวและปลอดภัย ไม่จำเป็นต้องมีสนามบินเฉพาะสำหรับการขึ้นและลงจอด ควบคุมจากระยะไกลสามารถทำการควบคุมได้ 2 ลักษณะ คือ การควบคุมแบบบังคับเอง และการควบคุมแบบอัตโนมัติ ซึ่งการควบคุมแบบอัตโนมัติจากระยะไกลโดยใช้ระบบการบินด้วยตนเองต้องอาศัยโปรแกรมจากเครื่องคอมพิวเตอร์ที่มีระบบที่ซับซ้อน และมีการติดตั้งอุปกรณ์สำหรับรับส่งสัญญาณไว้ในอากาศยานไร้คนขับ จึงเครื่องบินที่สามารถบินได้ด้วยระบบอัตโนมัติ โดยไม่ต้องใช้นักบินอยู่บนอากาศยาน อาจมีการติดตั้งกล้องสำหรับการบันทึกภาพที่มีคุณภาพสูง สามารถบันทึกภาพได้ในเวลากลางวัน (Electro Optical) และกล้องชนิดอินฟราเรด (Infrared Sensor) สำหรับบันทึกภาพในเวลากลางคืน ทำการรับส่งสัญญาณภาพทำการบันทึกภาพระยะไกลได้ แล้วแพร่ภาพสัญญาณภาพมายังจอแสดงผล ที่สถานีภาคพื้นดิน ในเวลาที่ใกล้เคียงจริงมากที่สุด (Near Real Time : NRT) ทำให้ผู้บังคับอากาศยานสามารถมองเห็นภาพในเวลาใกล้เคียงความเป็นจริงมากที่สุด นอกจากนั้นอากาศยานไร้คนขับ ยังสามารถปฏิบัติการกิจด้านบันทึกภาพจากมุมสูง การบินสำรวจพื้นที่ สำรวจป่าไม้ สำรวจน้ำท่วม การค้นหาเป้าหมาย และการบินลาดตระเวนหรือเรียกว่า ISTAR (Intelligence, Surveillance, Target Acquisition, Reconnaissance) เป็นต้น

ประเภทระบบอากาศยานไร้คนขับ สามารถกำหนดรูปแบบ รูปร่าง ขนาด ได้หลายลักษณะ ทั้งนี้ ขึ้นอยู่กับวัตถุประสงค์ในการนำไปใช้ แต่ละภารกิจ การออกแบบและพัฒนาขึ้นสำหรับการใช้งาน สำหรับภารกิจใด ภารกิจหนึ่ง หรือ สภาภูมิประเทศ การนำไปใช้ นอกจากนั้นทำการพิจารณาถึงอากาศยานไร้คนขับ ดังกล่าวผู้นำไปใช้ในหน่วยงานใด การใช้เพื่อวัตถุประสงค์ใด เป้าหมาย และเหตุผลใด โดยสามารถที่จะกำหนดแนวทางการแบ่งประเภทของอากาศยานไร้คนขับได้ ดังนี้

- 1) เป้าหมายและเป้าล่อ เป็นเป้าฝึกให้กับพลปืนต่อต้านอากาศยานหรือขีปนาวุธ
- 2) ข่าวกกรอง เป็นหน่วยข่าวกรองในสมรภูมิ
- 3) โจมตี ทำภารกิจโจมตี
- 4) ลำเลียง เป็นยูเอวี (UAV) ที่ออกแบบมาเป็นพิเศษเพื่อการขนส่ง
- 5) วิจัยและพัฒนา ใช้เพื่อการพัฒนาเทคโนโลยีของ UAV เพื่อนำไปใช้จริง
- 6) พลเรือนและการตลาด เป็นอากาศยานไร้คนขับหรือยูเอวีที่ถูกออกแบบมา

2.8.1 ส่วนประกอบของระบบอากาศยานไร้คนขับ

วัสดุอุปกรณ์ส่วนประกอบต่าง ๆ ของอากาศยานไร้คนขับและระบบควบคุม กำหนดได้จากการการคำนวณ การออกแบบ ระบบควบคุมการบิน ระบบรับส่งสัญญาณ และระบบสนับสนุนภาคพื้นดิน ความต้องการหลัก 5 ประการ ได้แก่ ระยะเวลาในการบิน ความเร็วในการบิน รัศมีการบิน ความสูงในการบินและน้ำหนักรวม เมื่อพิจารณาโดยรวมแล้วระบบอากาศยานไร้คนขับสามารถแยกได้ 10 ส่วน คือ

- 1) โครงเครื่องบิน (Air frame) โครงสร้าง ลักษณะรูปร่าง เช่น อากาศยานไร้คนขับรุ่น Pioneer รูปร่างคล้ายกล่องสี่เหลี่ยมหรือรุ่น Cypher รูปร่างคล้ายโดรน ส่วนวัสดุใช้งานก็มีหลายแบบ เช่น โฟม โลหะ พลาสติก คาร์บอนผสมไฟเบอร์ และวัสดุที่สามารถดูดกลืนคลื่นค้นหา เป็นต้น
- 2) ระบบขับเคลื่อนและเครื่องยนต์ (Propulsion System) ระบบขับเคลื่อนที่ใช้กับอากาศยานไร้คนขับ มีหลากหลาย เช่น เครื่องยนต์โรตารี เครื่องยนต์ 2 หรือ 4 จังหวะ มอเตอร์ไฟฟ้า มอเตอร์ไฟฟ้าความเร็วสูง เป็นต้น
- 3) ระบบบังคับ (Control System) แบบใช้วิทยุจากพื้นดิน ติดต่อสื่อสารไร้สายหรือการใช้โปรแกรมหรือซอฟต์แวร์ควบคุมการบินด้วยเครื่องคอมพิวเตอร์
- 4) ระบบขึ้นบินและลงจอด (Launch and Recovery System) การส่งอากาศยานไร้คนขับหรือยูเอวีขึ้นไปทำได้หลายวิธี เช่น การยิงจากเครื่องส่ง (Launch) ขึ้นบินจากทางวิ่ง หรือการปล่อยจากอากาศยานขนาดใหญ่ เช่น C-130 การลงจอดกลับคืนฐานที่ตั้งก็สามารถทำให้หลากหลาย เช่น การจับด้วยตาข่าย การใช้ร่มชูชีพ และการบังคับลงบนรันเวย์ด้วยใช้วิทยุจากพื้นดิน
- 5) ระบบนำวิถีและนำร่อง (Navigation and Guidance System) เป็นระบบที่มีความสำคัญของอากาศยานไร้คนขับ ในปัจจุบันนิยมใช้ จีพีเอส (GPS) เป็นอุปกรณ์ในการช่วยเหลือของระบบนี้ โดยปกติแล้ว UAV ใช้เครื่องคอมพิวเตอร์ทำงานด้านระบบนำวิถีและนำร่องโดยเฉพาะทำการแยกออกมาจากระบบควบคุมการบินอัตโนมัติ (Autopilot)
- 6) ระบบสนับสนุนภาคพื้นดิน (Ground Control Station) เป็นระบบควบคุมและสนับสนุนภาคพื้นดินปฏิบัติหน้าที่คล้ายกับ ระบบควบคุมภาคพื้นดินของอากาศยานทั่ว ๆ ไป โดยทำปฏิบัติหน้าที่ตรวจสอบ อุปกรณ์ตรวจวัดต่าง ๆ รับส่งข้อมูลแบบไร้
- 7) ระบบบรรทุก (Payload) อากาศยานไร้คนขับที่ปฏิบัติการภารกิจต่าง ๆ ต้องทำการบรรทุกและติดตั้งเช่น กล้องความละเอียดสูง กล้องตรวจจับความร้อน กล้องอินฟราเรด และเรดาร์ ในปัจจุบันมีการพัฒนาอากาศยานไร้คนขับ ในภารกิจการสอดแนม และโจมตี มีการติดตั้งจรวดหรือระเบิดขนาดต่าง ๆ ตามภารกิจ
- 8) ระบบจัดเก็บข้อมูลและเชื่อมต่อ (Data Link and Storage System) เป็นระบบรับส่งข้อมูลระหว่างอากาศยานไร้คนขับ เข้ากับระบบสนับสนุนภาคพื้นดิน สามารถใช้งานได้หลายย่าน

ความถี่ เช่น ย่านความถี่สูง (HF) ย่านความถี่สูงมาก (VHF) และย่านไมโครเวฟ ถ้าระบบเครือข่ายนี้เกิดระบบขัดข้องจะส่งต่อไปยังเครือข่ายอื่น ๆ เช่น GPS Satellites เป็นต้น เพื่อส่งกลับภาคพื้นดิน

9) ระบบปกป้องตนเอง (Self Protection System) ใช้วัสดุอุปกรณ์ที่มีคุณลักษณะดูดกลืนการตรวจจับ คลื่นเรดาร์แบบเครื่องบินขับไล่ เพื่อป้องกันการตรวจจับได้ง่าย ได้แก่ ภารกิจอากาศยานไร้คนขับ (UAV) บรรทุกระเบิด ขีปนาวุธ

10) ผู้ปฏิบัติการบิน (Operating Personnel) ที่จะใช้งานปัจจุบันผู้ปฏิบัติการบินอากาศยานไร้คนขับจะเป็นผู้ที่ ได้รับความศึกษาอบรมฝึกฝน มาเป็นอย่างดีและมีประสบการณ์สูง

2.9 อากาศยานไร้คนขับขนาดเล็ก

อากาศยานไร้คนขับแบ่งตามประเภทได้ดังนี้ ประเภทปีกหมุน (Multirotor) ประเภทปีกตรึง (Fixed wing) ประเภทไฮบริด (Vertical Takeoff and Landing) ซึ่งแต่ละประเภทมีข้อดีข้อเสียและวัตถุประสงค์ในการใช้งานที่แตกต่างกัน ประเภทปีกตรึง ระยะเวลาบินยาวนาน ใช้ทำแผนที่ทางอากาศ แต่ทำการลงจอดยาก ประเภทไฮบริด ระยะเวลาบินยาวนาน ใช้งานสำรวจ การเฝ้าระวัง การค้นหา ควบคุมยากต้องมีการฝึกอบรม ประเภทปีกหมุน ควบคุมและบังคับง่าย ใช้พื้นที่น้อยในการบินขึ้นลงในแนวดิ่ง ใช้กับงานภาพถ่ายทางอากาศ เกษตรกรรม ระยะเวลาบินจำกัด

อากาศยานไร้คนขับขนาดเล็ก ประเภทปีกหมุน (Multirotor) อุปกรณ์ที่ใช้ใช้งานประกอบด้วย รีโมตควบคุม (Controller) สายรับส่งข้อมูล (Data Link and Storage System) แท็บเล็ต คอมพิวเตอร์ (iPad or Tablet) หรือ โทรศัพท์เคลื่อนที่แบบสมาร์ทโฟน (Smart phone) ทำการติดตั้งแอปพลิเคชัน (Application) “DJI GO 4” สำหรับใช้เชื่อมต่อเข้ากับรีโมตควบคุม (Controller) รับส่งข้อมูลแบบไร้สายกับอากาศยานไร้คนขับ ยี่ห้อ DJI รุ่น Mavic Pro ทำการควบคุมแบบบังคับเองและแบบอัตโนมัติหรือซอฟต์แวร์สำหรับวางแผนการบินในส่วนหลักการทำงานของอากาศยานไร้คนขับ เริ่มต้นจากนำ Smart phone ใช้ Data Link กับ Controller ทำการเชื่อมต่อแบบไร้สายไปยังตัวเครื่องอากาศยาน (Platform) แล้วเปิดแอปพลิเคชัน (Application) “DJI GO 4” ที่ได้ดาวน์โหลดติดตั้งในอุปกรณ์ควบคุมอากาศยานไว้แล้ว (Controller) เพื่อเชื่อมต่อตัวเครื่องอากาศยาน (Platform) และสามารถใช้เครื่องคอมพิวเตอร์ส่วนบุคคล เครื่องคอมพิวเตอร์โน้ตบุ๊ก ควบคุมการใช้งานได้ โดยทำการติดตั้งโปรแกรมหรือซอฟต์แวร์ สำหรับแสดงภาพแบบเรียลไทม์



รูปที่ 2.34 อากาศยานไร้คนขับ แบบสี่ใบพัด ยี่ห้อ Dji รุ่น Mavic Pro (อ้างอิง [https://www.djibangkok.com/shop/dji-mavic-pro/#iLightbox\[product-gallery\]/0](https://www.djibangkok.com/shop/dji-mavic-pro/#iLightbox[product-gallery]/0))

คุณลักษณะของ Drone ยี่ห้อ Dji รุ่น Mavic Pro

- 1) น้ำหนัก 734 กรัม
- 2) ขนาด (ความกว้างxยาวxความสูง) 83 x 198 x 83 มิลลิเมตร
- 3) ความยาวทแยงมุม 335 มิลลิเมตร
- 4) ความเร็วสูงสุดในการบินขึ้น 16.4 ฟุตต่อวินาที หรือ 5 เมตรต่อวินาที (ในโหมด Sport)
- 5) ความเร็วสูงสุดในการบินขึ้น 9.8 ฟุตต่อวินาที หรือ 3 เมตรต่อวินาที
- 6) ความเร็วในการบินสูงสุด 65 กิโลเมตรต่อชั่วโมง หรือ 40.4 ไมล์ต่อชั่วโมง (ในโหมด Sport)
- 7) ระดับความสูงมากที่สุดในการบิน 164,404 ฟุต หรือ 5,000 เมตร
- 8) ระยะเวลาในการบินสูงสุด 27 นาที (ในสภาวะลมสงบ)
- 9) ระยะเวลาในการบินเฉลี่ย 21 นาที (ในการบินปกติ แบตเตอรี่จะเหลือ 15 เปอร์เซ็นต์)
- 10) สามารถควบคุมได้ไกลสูงสุด 8 ไมล์ หรือ 13 กิโลเมตร (ในสภาวะลมสงบ)

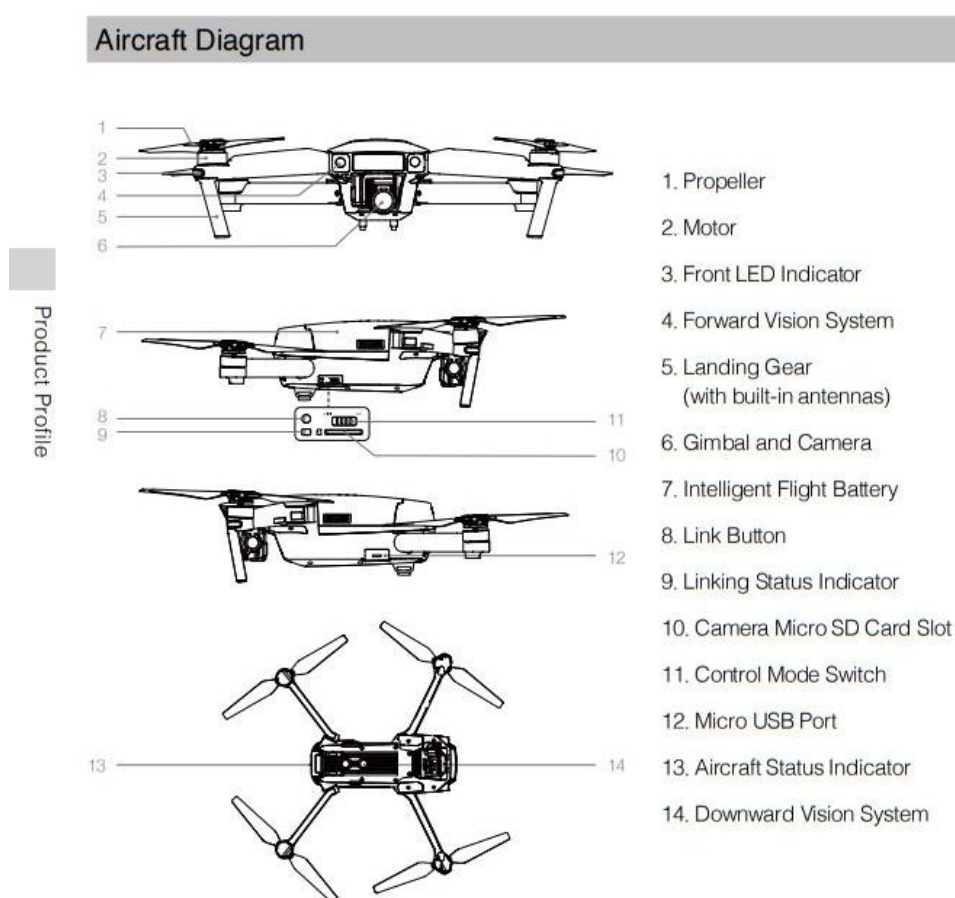
- 11) อุณหภูมิที่สามารถทำการบินได้ 32 ถึง 104 องศาฟาเรนไฮต์ (0 ถึง 40 องศาเซลเซียส)
- 12) ระบุตำแหน่งด้วย GPS/GLONASS
- 13) คลื่นความถี่ในการใช้งาน FCC : 2.4-2.4835 GHz; 5.150-5.250 GHz; 5.725-5.850 GHz
CE : 2.4-2.4835 GHz; 5.725-5.850 GHz
SRRC : 2.4-2.4835 GHz; 5.725-5.850 GHz
- 14) กำลังส่งสัญญาณวิทยุบังคับ 2.4 GHz
FCC : <26 dBm; CE : <20 dBm; SRRC <20 dBm
5.2 GHz
FCC: <23 dBm
5.8 GHz
FCC: <23 dBm; CE <13 dBm; SRRC: <23 dBm
- 15) กล้องถ่ายภาพ 1/2.3" CMOS ความละเอียด 12.35 Megapixels
- 16) เลนส์ 78.8 FOV, 28 มิลลิเมตร (35 มิลลิเมตร รูปแบบเทียบเท่า) f2.2 ความคลาดเคลื่อน <15% โฟกัส จาก 0.5 ถึง ไม่จำกัด
- 17) พิสัย ISO 100 – 3200 (วิดีโอ), 100-1600 (รูปภาพ)
- 18) ความเร็วชัตเตอร์ 8 วินาที ถึง 1/8000 วินาที
- 19) ขนาดของรูป 4000 x 3000 ถ่ายครั้งเดียว ถ่ายภาพต่อเนื่อง 3/5/7 เฟรม ตัดแสงอัตโนมัติ
- 20) ถ่ายวิดีโอ C4k : 4096x2160 42p, 4K: 3840x2160 24/25/30p
2.7k : 2720x1530 24/25/30p
FHD : 1920x1080 24/25/30/48/50/60/96p
HD : 1280x720 24/25/30/48/50/60/120p
- 21) อัตราการส่งภาพวิดีโอ 60 Mbps
- 22) ไฟล์ภาพชนิด JPEG, DNG
- 23) ไฟล์วิดีโอ MP4, MOV (MPEG-4AVC/H.264)
- 24) รองรับ SD Card microSD 64GB Max Class 10 or UHS-1 rating required
- 25) การชาร์จ แรงดัน 13.5 โวลต์ พลังงาน 50 วัตต์

แบตเตอรี่ ความจุ 3,830 มิลลิแอมป์ แรงดัน 11.4 โวลต์ ชนิด LiPo 3S พลังงาน 43.6 วัตต์ชั่วโมง น้ำหนัก 240 กรัม

2.10 การวางแผนการบินแบบบังคับเองและแบบอัตโนมัติ

2.10.1 การวางแผนการบินแบบบังคับเอง อากาศยานประเภทปีกหมุน แบบหลายใบพัด (Multirotor) บินขึ้นลงแนวดิ่ง ยี่ห้อ DJI รุ่น Mavic Pro สามารถเคลื่อนตัวได้อย่างรวดเร็วและคล่องตัว ระยะเวลาในการบินประมาณ 27 นาทีต่อการชาร์จ 1 ครั้ง รัศมีทำการบินในพื้นที่กว้าง การขึ้นบินและลงจอดไม่ต้องอาศัยพื้นที่โล่งกว้างมาก เนื่องจากสามารถบินขึ้นลงในแนวดิ่ง ซึ่งเป็นข้อดีของชนิดปีกหมุน (Multirotor)

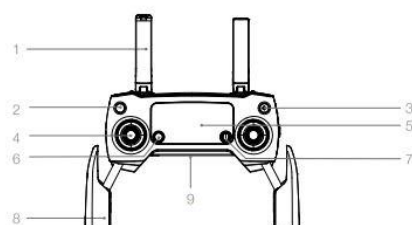
1) ตัวเครื่องอากาศยาน (Platform) ยี่ห้อ DJI รุ่น Mavic Pro ส่วนประกอบต่าง ๆ



รูปที่ 2.35 ส่วนประกอบของอากาศยานไร้คนขับ ยี่ห้อ DJI รุ่น Mavic Pro (อ้างอิง <https://dl.djicdn.com/downloads/mavic/20171219/Mavic%20Pro%20User%20Manual%20V2.0.pdf>)

2) อุปกรณ์รีโมตควบคุมอากาศยาน (Controller) ยี่ห้อ DJI

Remote Controller Diagram

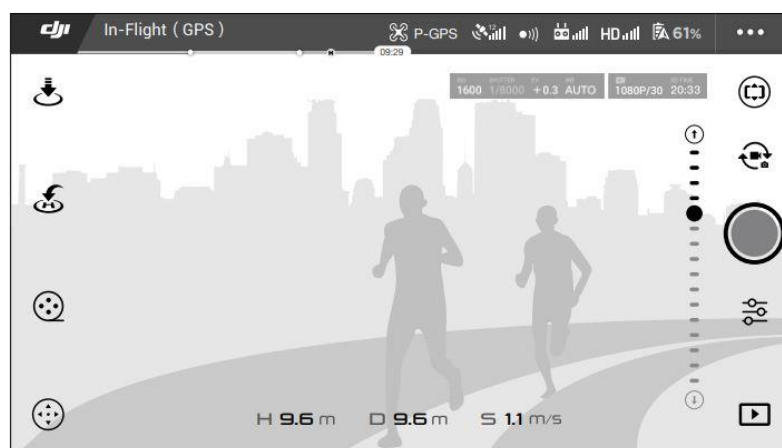


1. **Antennas**
Relays aircraft control and video signal.
2. **Return to Home (RTH) Button**
Press and hold the button to initiate Return to Home (RTH). Press again to cancel RTH.
3. **Power Button**
Used to turn the remote controller on and off.
4. **Control Stick**
Controls the orientation and movement of the aircraft.

8 © 2017 DJI All Rights Reserved.

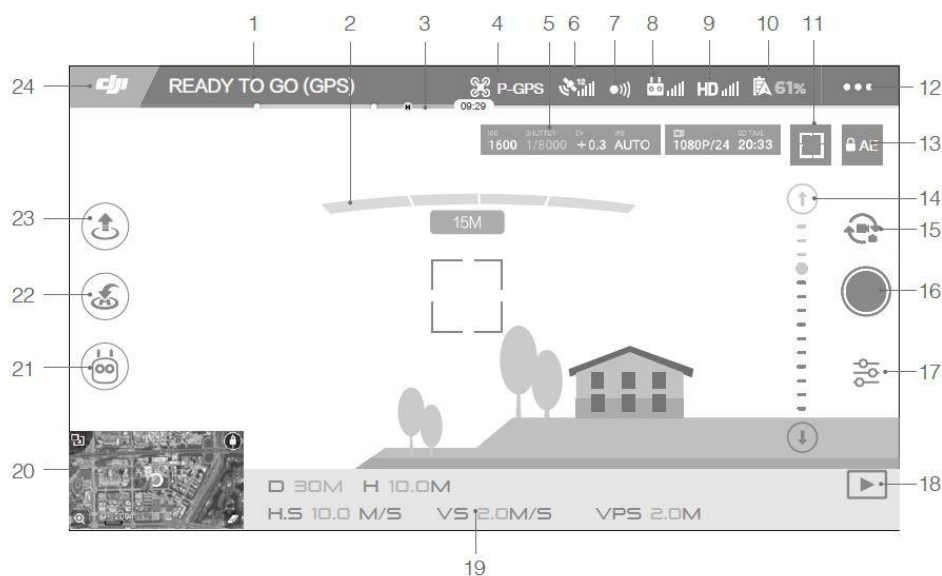
รูปที่ 2.36 อุปกรณ์รีโมตควบคุมอากาศยาน (Controller) (อ้างอิง <https://dl.djicdn.com/downloads/mavic/20171219/Mavic%20Pro%20User%20Manual%20V2.0.pdf>)

3) แอปพลิเคชัน DJI GO 4 เป็นแอปพลิเคชัน ใช้สำหรับตั้งค่าการใช้งานและฟังก์ชันต่าง ๆ ของอากาศยานไร้คนขับ และสามารถมองเห็นภาพเรียลไทม์จากมุมมองของกล้องที่ติดตั้งผ่านทางแอปพลิเคชันบนหน้าจอสมาร์ทโฟน ได้ด้วยสำหรับใช้ควบคุมการบินบนสมาร์ทโฟนแทนรีโมทคอนโทรล สำหรับ Drone รุ่น Mavic Pro สามารถนำไปเชื่อมต่อหรือใช้งานร่วมกับรุ่นอื่น ๆ ได้



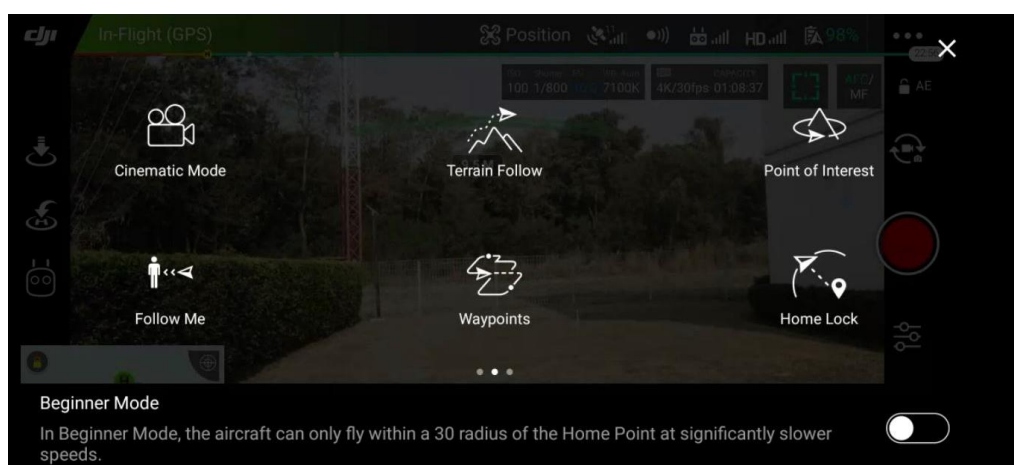
รูปที่ 2.37 แอปพลิเคชัน DJI GO 4 (อ้างอิง <https://dl.djicdn.com/downloads/mavic/20171219/Mavic%20Pro%20User%20Manual%20V2.0.pdf>)

2.10.2 การวางแผนการบินแบบอัตโนมัติ อากาศยานไร้คนขับโดยใช้แอปพลิเคชัน DJI GO 4 ฟังก์ชัน Intelligent Flight Mode วางแผนการเส้นทางการบินอัตโนมัติ

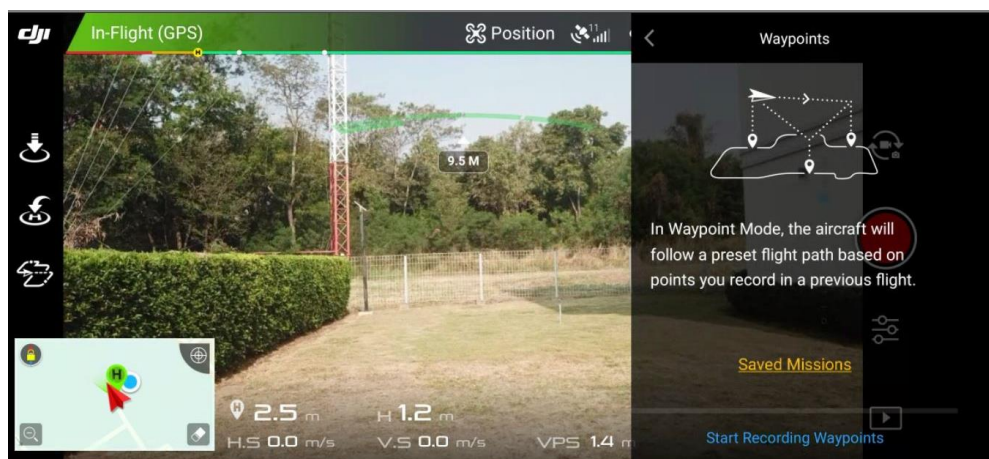


รูปที่ 2.38 แอปพลิเคชัน DJI GO 4 วางแผนการบิน (อ้างอิง <https://www.dji.com/hk-en/downloads/products/mavic>)

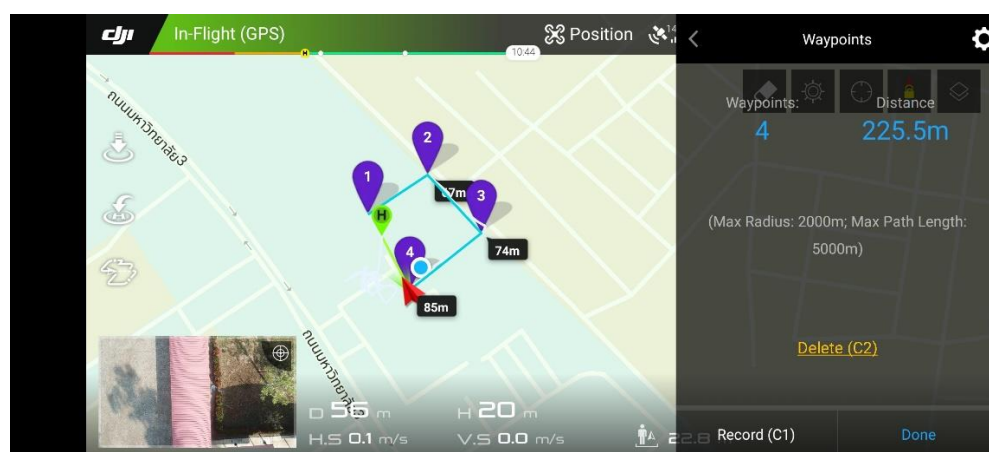
การใช้งานแอปพลิเคชัน DJIGO 4 สำหรับการวางแผนเส้นทางการบิน (Path) ของอากาศยานไร้คนขับ ยี่ห้อ DJI รุ่น Mavic ใช้งานในโทรศัพท์เคลื่อนที่แบบสมาร์ทโฟนหรือแท็บเล็ต (Tablet) รองรับระบบปฏิบัติการ IOS และ Android กำหนดภารกิจการเคลื่อนที่ Waypoint ไปตำแหน่งที่ต้องการ



รูปที่ 2.39 ฟังก์ชันภารกิจการเคลื่อนที่ Waypoint



รูปที่ 2.40 ทำการวางแผนเส้นทางการบิน (Path) จุดบินขึ้นลง และตั้งค่าความสูงในการบิน



รูปที่ 2.41 ทำการวางแผนเส้นทางการบิน (Path) เส้นทางการบินจุดที่ 1 ถึง 4

2.11 ปรัชญ์นวัตกรรมและงานวิจัยที่เกี่ยวข้อง

การใช้งานอากาศยานไร้คนขับ ทำการบิน ถ่ายภาพหรือวิดีโอ และส่งข้อมูลภาพแบบไร้สาย เพื่อทำการตรวจจับตำแหน่งของคนทีแนวรั้วหอพักโรงเรียนสุรวิวัฒน์ โดยการใช้การประมวลผลภาพ (Image processing) เป็นการหาพื้นที่พิกเซล (Pixel) เป็นหน่วยพื้นฐานที่เล็กที่สุดของภาพดิจิทัล เทียบได้กับจุดสีของภาพ 1 จุด เพื่อทำการจำแนกประเภทของตรวจจับคนหาคน และทำการจัดกลุ่มตามประเภทของรูปแบบต่าง ๆ จากนั้นนำไปประยุกต์ใช้งานกับเทคนิคการเรียนรู้ของเครื่องจักร (Machine Learning) โดยใช้งานโครงข่ายประสาทเทียมในการประมวลผลภาพและวิเคราะห์เพื่อทำการตรวจจับคนหาคนบุกรุกป็นรั้ว

งานวิจัยที่เกี่ยวข้องกับอากาศยานไร้คนขับและการตรวจจับวัตถุ

ในปี ค.ศ.2021 Aswini N และ Uma S V [1] ได้นำเสนอ การตรวจจับวัตถุด้วยอากาศยานไร้คนขับโดยใช้อัลกอริธึมเรียนรู้เชิงลึก โดยทั่วไปอากาศยานไร้คนขับ (UAV) เรียกว่า “โดรน” เป็นเครื่องบินที่ไร้นักบิน สามารถนำมาใช้งานในพื้นที่เสี่ยงหรือยากในการเข้าถึงในการปฏิบัติงาน โดยพื้นฐานแล้ว คือ หุ่นยนต์บินได้ สามารถควบคุมด้วยตนเองหรือแบบอัตโนมัติ ตามแผนการเคลื่อนที่ด้วยซอฟต์แวร์ การเคลื่อนที่ที่ได้รับการช่วยเหลือด้วยโปรเซสเซอร์แบบฝังตัวจากบอร์ดเซ็นเซอร์ และระบบกำหนดตำแหน่งภาคพื้นดิน (GPS) การใช้งานอากาศยานไร้คนขับมีความต้องการมากขึ้นเรื่อยๆ ในแต่ละปี เนื่องจากสามารถใช้งานได้จริงและหลากหลาย โดยแอปพลิเคชัน โดยมีความท้าทายในการใช้แอปพลิเคชันกับอากาศยานไร้คนขับ หนึ่งในความท้าทายที่เผชิญ คือ การตรวจจับและหลบหลีกวัตถุเข้ามาในเส้นทางการเคลื่อนที่ ในบทความนี้นำเสนอการเรียนรู้เชิงลึก โดยใช้พื้นฐานการตรวจจับวัตถุด้านหน้า โดยใช้โครงข่ายประสาทเทียมที่ได้รับการฝึกสอนโมเดลมาล่วงหน้า ภาพแต่ละเฟรมที่ได้จากกล้องของอากาศยานไร้คนขับ ประมวลผลและป้อนเข้าไปในเครือข่ายการเรียนรู้เชิงลึกสำหรับการตรวจจับวัตถุ ภาพรวมและการเปรียบเทียบอัลกอริธึมการเรียนรู้เชิงลึกทั้งสามแบบคำนึงถึงความแม่นยำและรวดเร็ว

ทำการทดลองภายนอกอาคาร ใช้ภาพวิดีโอจากกล้องของ โดรน ยี่ห้อ DJI รุ่น Phantom ความเร็วในการเคลื่อนที่ 1-2 เมตรต่อวินาที ระดับความสูง 25-30 เมตร บันทึกภาพ คน จักรยานยนต์ และรถยนต์ การเรียนรู้เชิงลึก ทดลองด้วย Faster RCNN, SSD และ YOLOv3

จากผลการทดลองในระดับความสูง 25 เมตร วิธีการ Faster RCNN. ตรวจจับวัตถุได้ทั้งหมด วิธีการ SSD ตรวจจับวัตถุไม่ได้ 2 กรณี และ วิธีการ YOLOv3 ตรวจจับวัตถุได้ทั้งหมด ประโยชน์-สามารถนำไปใช้กับแอปพลิเคชันแบบเรียลไทม์ เช่น การค้นหา การกู้ภัย ตรวจสอบการจราจร ตรวจสอบฝูงชน ตรวจสอบโครงสร้างพื้นฐาน เป็นต้น

ในปี ค.ศ.2021 Aswini N และ Uma S V [2] ได้นำเสนอการตรวจจับวัตถุสิ่งกีดขวาง แบบกำหนดเอง โดยใช้ YOLOv3 ด้วยอากาศยานไร้คนขับบินต่ำ อากาศยานบินต่ำเป็นที่ต้องการมากในขณะนี้ เนื่องจากมีการใช้งานที่หลากหลาย ในด้านการรักษาความปลอดภัย การเฝ้าระวัง สถานที่ก่อสร้าง การเข้าช่วยเหลือเหตุฉุกเฉิน การขนส่ง การดับเพลิง การเกษตรกรรมที่แม่นยำ การตรวจสอบการจราจร การตรวจสอบฝูงชน เป็นต้น ข้อที่ดีที่สุด คือไม่มีความเสี่ยงของชีวิตมนุษย์ อากาศยานไร้คนขับอำนวยความสะดวกและความปลอดภัย ในการตรวจสอบที่ยากในการเข้าถึงและของสภาพแวดล้อมที่อันตราย อากาศยานอัตโนมัติเต็มรูปแบบ ต้องติดตั้ง เซ็นเซอร์ต่าง ๆ เช่น Lidar, Radars, Ultrasonic sensors, Inertial Measurement Unit, ระบบกำหนดตำแหน่งภาคพื้นดิน และกล้องสำหรับตรวจจับและการหลบหลีกสิ่งกีดขวางในเส้นทางการเคลื่อนที่ ในบทความนี้

ทำการค้นคว้าข้อมูลความเป็นไปได้ โซลูชันในการตรวจจับสิ่งกีดขวางแบบกำหนดเองโดยใช้กล้องเป็นส่วนคำสำคัญหลักในการตรวจจับวัตถุ

วิธีการทดลองโดยการปรับขนาดภาพเป็น 1024x720 ใช้ Nearest Neighbour (NN) แก้ไขพิกเซลที่ใกล้ที่สุดเฉลี่ยแทนที่ ใช้ YOLOv3 จากข้อมูลภาพทั้งหมด จำนวน 80% สำหรับฝึกสอนโมเดล อีก 20% สำหรับการทดสอบ และทำการลดขนาดรูปภาพให้มีมิติ 416x416 ฝึกสอนโมเดลของโดรน ใช้งานกับ Google Colab เพื่อให้ทำงานได้รวดเร็วและมีประสิทธิภาพ ค่าเริ่มต้น คือ 0.01 และลดลงเหลือ 0.0005 ทำการฝึกสอนโมเดล 900 epochs การทดสอบภาพและวิดีโอ การตรวจจับวัตถุทั้ง 5 กลุ่ม ต้นไม้ เนินเขา หิน พุ่มไม้ และเสาบ้าน แสดงผลใน Google Colab ทำการเพิ่มโมเดลและปรับปรุงความแม่นยำ หลังจากการฝึกสอนโมเดล 900 epochs สามารถตรวจจับวัตถุจากโดรนที่ถ่ายธรรมชาติ ภัยพิบัติ ได้ ด้วยชุดข้อมูลภาพที่กำหนดเองมากกว่า 2,000 ภาพ

ในปี ค.ศ.2022 Rohit และคณะ [3] ได้นำเสนอการตรวจจับวัตถุด้วยโดรน โดยใช้ปัญญาประดิษฐ์ การเฝ้าระวังทางอากาศ มีบทบาทสำคัญ แอปพลิเคชันสำหรับการรักษาความปลอดภัย สามารถช่วยในการสอดส่องดูแลตรวจสอบพื้นที่ที่ชายแดน เขตหวงห้าม และโครงสร้างพื้นฐานที่สำคัญ โดรนสามารถในการช่วยเหลือ เฝ้าระวังและทราบตำแหน่งที่แน่นอน ของวัตถุต่าง ๆ การตรวจจับวัตถุทางอากาศมีหลายอย่าง ความท้าทาย เช่น วัตถุที่มีขนาดเล็กถึง 20 x 20 พิกเซล ภาพถ่ายจากดาวเทียมมีความละเอียดหลายร้อยล้านพิกเซล ใช้วิธีการดั้งเดิม เช่น วิธีการฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ (HOG) และวิธีค้นหาและกำหนดลักษณะเด่นของภาพ (SIFT) นำมาใช้ในการแยกคุณสมบัติของวัตถุ จากนั้นได้ส่งคุณสมบัตินี้ไป ทำการแยกประเภทของการเรียนรู้ของเครื่องจักร (Machine Learning) เช่น วิเคราะห์การถดถอย (Regression Analysis) วิธี Support vector machine (SVM) วิธีการแก้ปัญหาด้วยการลดข้อผิดพลาด และ Random forest (RF) การฝึกสอนโมเดลที่เหมือนกันหลายครั้งบนชุดข้อมูลเดียวกัน สำหรับการตรวจจับวัตถุและการจำแนกประเภทอย่างไรก็ตามปัญหาของวิธีต่าง ๆ นี้ คือ มีความคลาดเคลื่อนอย่างมาก และสร้างการตรวจจับที่ผิดพลาดมากมาย และการคัดแยกประเภทที่ผิด การประมวลผลกราฟิกหน่วย (GPU) และการนำระบบโครงข่ายประสาทเทียมแบบ Convolution (CNN) อัลกอริธึมการเรียนรู้เชิงลึก (Deep Learning) ไปสู่การเปลี่ยนแปลง สามารถดึงข้อมูลเพิ่มเติมได้และความแม่นยำที่มากขึ้น โบบทความนี้ศึกษาการตรวจจับวัตถุ You only look one เวอร์ชัน 4 (YOLOv4) โดยใช้ Darknet 53 เป็นประเภทของ CNN เป็น backbone สำหรับในการแยกคุณลักษณะ ในงานนี้ ใช้ YOLOv4 เป็นระบบตรวจจับและตำแหน่งของยานพาหนะที่อยู่ในเขตพื้นที่หวงห้าม แล้วติดแท็กตำแหน่ง

การใช้ YOLOv4 ในการตรวจจับวัตถุ ใช้ชุดข้อมูล VisDrone 2019 จำนวน 8,599 ภาพ กล้องติดโดรนบันทึกภาพจากมุมสูง สถานที่ สภาพการจราจรที่หนาแน่น ชุดข้อมูลภาพ ได้แก่ รถตู้ รถบัส รถบรรทุก จักรยานยนต์ หลังคารถสามล้อ รถสามล้อ จักรยาน คนเดินเท้า รถยนต์ และ

รถสามล้อ ทำการทดสอบ YOLOv4 ฝึกสอนโมเดล ความกว้าง 768 ความสูง 64 batch size อัตราการเรียนรู้ 0.0001 สำหรับ 8,000 iteration เวลาในการฝึกสอนโมเดล 80 ชั่วโมง ประมวลผลด้วย Nvidia RTX 2080 GPU with 12GB RAM.ผลการทดลอง YOLOv4 ฝึกสอนโมเดล ใช้ GPU ช่วยให้ค่าเฉลี่ยความแม่นยำประมวลผลเร็วขึ้น 68.06%

ในปี ค.ศ.2021 Oyku Sahin and Sedat Ozer [4] ได้นำเสนอ YOLODrone การปรับปรุงสถาปัตยกรรม YOLO สำหรับการใช้โดรนตรวจจับวัตถุภาพ ความก้าวหน้าในด้านวิทยาการหุ่นยนต์และคอมพิวเตอร์วิชั่น ทำให้เกิดแอปพลิเคชันใหม่ๆ สำหรับกล้องที่ติดตั้งที่โดรนแอปพลิเคชันหนึ่ง คือการตรวจจับวัตถุจากภาพถ่ายทางอากาศ จะมีความก้าวหน้าในวรรณกรรมที่เกี่ยวข้องกับการตรวจจับวัตถุยังคงเป็นงานที่ท้าทายในคอมพิวเตอร์วิชั่น อัลกอริธึมการตรวจจับวัตถุที่มีอยู่ แสดงให้เห็นภาพถ่ายจากโดรนมีประสิทธิภาพที่ต่ำกว่า (ภาพถ่ายทางอากาศ) เนื่องจากปัญหาในการตรวจจับวัตถุภาพถ่ายทางอากาศเป็นปัญหาที่ท้าทายมากขึ้น เมื่อเปรียบเทียบกับงานตรวจจับในภาพถ่ายภาคพื้นดิน มีสาเหตุหลายประการได้แก่ (i) การขาดแคลนภาพขนาดใหญ่ ชุดข้อมูลของโดรนที่มีความแปรปรวนของวัตถุขนาดใหญ่ (ii) สิ่งแปรปรวนที่เพิ่มมากขึ้นทั้งขนาดและการวางแนวภาพของโดรน และ (iii) ความแตกต่างในด้านรูปทรงและลักษณะพื้นดินกับภาพถ่ายที่อากาศ ในบทความนี้ ปรับปรุงด้วยอัลกอริธึม YOLODrone สำหรับการใช้โดรนตรวจจับวัตถุภาพ ประเมินอัลกอริธึมด้วยชุดข้อมูล VisDrone2019 และรายงานผลลัพธ์ที่ดีขึ้นเมื่อเปรียบเทียบกับ อัลกอริธึม YOLOv3

ในปี ค.ศ.2022 Teddy Surya Gunawan และ Islam Mohamoud Ismail [5] ได้นำเสนอการเปรียบเทียบประสิทธิภาพของ YOLO ต่าง ๆ สถาปัตยกรรมการตรวจจับวัตถุภาพถ่ายจากอากาศยานไร้คนขับ ปัจจุบันมีการพัฒนาอย่างรวดเร็วของการเรียนรู้เชิงลึก (Deep Learning) การยกระดับประสิทธิภาพและประสิทธิผลของอุตสาหกรรมต่าง ๆ ได้แก่ ธุรกิจ การทหาร การแพทย์ และการขนส่ง การใช้อัลกอริธึมอย่างล้ำลึกการเรียนรู้ในอุตสาหกรรมขนส่ง เช่น ทำให้อากาศยานไร้คนขับมีความสำคัญมากและมีประสิทธิภาพต่อภาคอุตสาหกรรม การใช้งานอากาศยานไร้คนขับ (UAV) ส่งเสริมการพัฒนาระบบขนส่ง เพื่อรวบรวมข้อมูลการจราจรแบบเรียลไทม์โดยใช้อากาศยานไร้คนขับเนื่องจากความคล่องตัวในการพกพาอากาศยานไร้คนขับ โดยเฉพาะบุคคลทั่วไป ระบุวัตถุแบบเรียลไทม์ เพื่อการทหาร ใช้การสังเกตการณ์ รักษาความปลอดภัยและการป้องกัน จุดสูงสุดในการวิวัฒนาการเทคโนโลยีคอมพิวเตอร์วิชั่น เป็นการพัฒนาอัลกอริธึมที่ซับซ้อนมีศูนย์กลางอยู่ที่การฝึกสอนโมเดล ชุดข้อมูลและการทดสอบที่ครอบคลุม งานวิจัยนี้มีวัตถุประสงค์ เพื่อเปรียบเทียบประสิทธิภาพของการตรวจจับวัตถุของภาพจากอากาศยานไร้คนขับ โดยใช้ YOLO ต่าง ๆ ใช้โมเดลขนาดเล็ก YOLOv3 และ YOLOv5 เพื่อแยกข้อมูลคุณสมบัติของวัตถุในภาพและจำแนกออกเป็นหลายชุดข้อมูลในแต่ละกลุ่ม ในบทความนี้เลือกชุดข้อมูลจาก VisDrone2019 เป็นประเภทของวัตถุ เช่น คนเดินเท้า บุคคล จักรยาน รถยนต์ รถตู้ รถบรรทุก รถสามล้อ กันสาด รถบัส และจักรยานยนต์

ผลลัพธ์แสดงให้เห็นว่า YOLOv5 มีค่าความแม่นยำเป็นที่ยอมรับได้ และมีประสิทธิภาพประมวลผลที่รวดเร็ว

2.12 สรุป

การศึกษาข้อมูล ทฤษฎีที่เกี่ยวข้องกับงานวิจัยและปรัทัศน์วรรณกรรม เพื่อมาทำงานวิจัยระบบตรวจการณ์อัจฉริยะสำหรับหอพักโรงเรียนสุรวิวัฒน์โดยใช้อากาศยานไร้คนขับ ให้มีประสิทธิภาพ ความแม่นยำ (Accuracy) และความเที่ยงตรง (Precision) ในการค้นหาคน โดยจะพบตัวอย่างจากงานวิจัยอื่น ๆ ที่เกี่ยวกับการตรวจจับวัตถุ รถจักรยานยนต์ รถยนต์ อาคาร บ้านเรือน ต้นไม้ คน สัตว์ สิ่งของ ใช้ข้อมูลจากภาพถ่ายทางอากาศนั้น มีการใช้งานอย่างแพร่หลายในปัจจุบัน ผู้วิจัยได้เสนอใช้เทคนิคในการตรวจสอบ การประมวลผลภาพ (Image processing) การตรวจจับวัตถุ (Object detection) การเรียนรู้ของเครื่องจักร (Machine Learning) การเรียนรู้เชิงลึก (Deep learning) โดยใช้อัลกอริธึม โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution neural network) และ You only look once (YOLOv8) ใช้สำหรับการตรวจจับวัตถุระบุองค์ประกอบภายในภาพถ่ายทางอากาศ สุดท้ายก็เพื่อแจ้งเตือน เมื่อตรวจพบคนบุกรุกป็นรั้ว