บทที่ 2
แนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

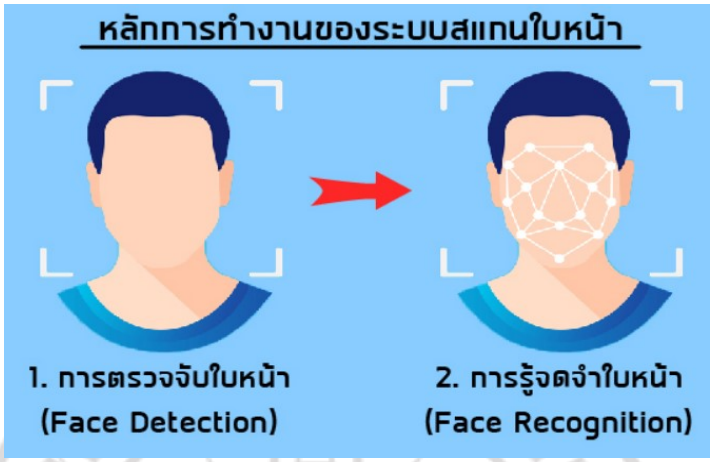
การประมวลภาพดิจิตอล

เทคโนโลยีในปัจจุบันมีการพัฒนาก้าวไกล และถูกนำมาใช้ประโยชน์ในชีวิตประจำวันหลายอย่าง หนึ่งในเทคโนโลยีเหล่าที่มีบทบาทกับชีวิตของมากที่สุดก็คือ เทคโนโลยีการรู้จดจำใบหน้าหรือที่รู้จักกันในชื่อ ระบบสแกนใบหน้า เทคโนโลยีที่นิยมนำมาใช้สร้างระบบรักษาความปลอดภัยด้วยการจดจำใบหน้า หลักการทำงานของ Face Recognition คือ การสร้างโมเดลการอ้างอิง ที่เรียกว่า “Face Print” ขึ้นมา โดยระบบจะวิเคราะห์จากลักษณะเฉพาะต่างๆ บนใบหน้า เช่น โครงหน้า ความกว้างของจมูก ระยะห่างระหว่างตาทั้งสองข้าง ขนาดของโหนกแก้ม ความลึกของเบ้าตา รวมถึงพื้นผิวบนใบหน้า (Facial texture) เป็นต้น จากนั้น ระบบจะทำการสร้างจุดเชื่อมโยงบนใบหน้า (Nodal points) เพื่อเปรียบเทียบกับรูปภาพที่ถูกเก็บไว้ในฐานข้อมูล (Database) ทั้งในลักษณะภาพนิ่งและภาพเคลื่อนไหว เพื่อความแม่นยำในการระบุตัวตนของผู้ที่ต้องเข้าสู่กระบวนการตรวจสอบ ซึ่งมีหลักการทำงานอยู่ 2 ขั้นตอน ดังนี้

1. การตรวจจับใบหน้า (Face Detection) คือ กระบวนการค้นหารูปใบหน้าของบุคคลจากภาพหรือวิดีโอ จากนั้นทำการประมวลผลภาพใบหน้าที่ได้สำหรับขั้นตอนถัดไปเพื่อให้ภาพใบหน้าที่ตรวจจับได้ง่ายต่อการจำแนก

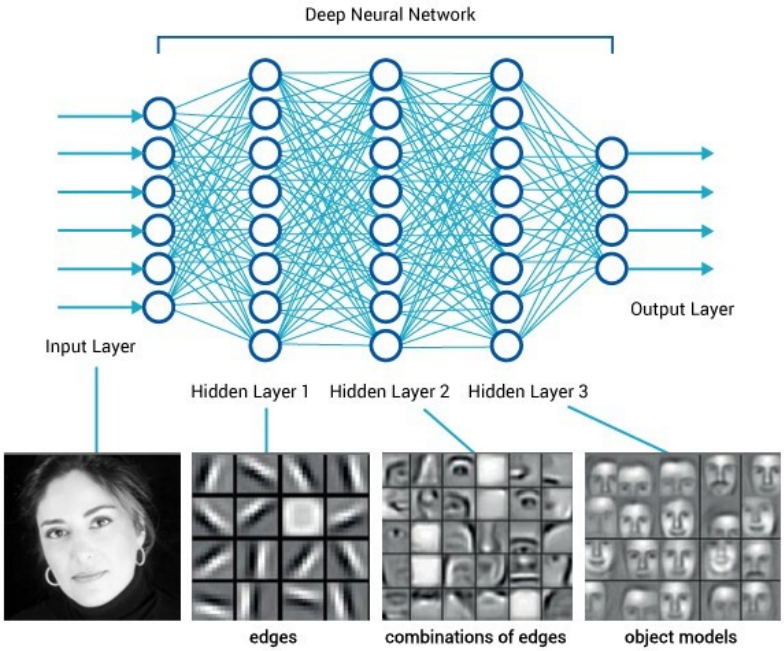
2. การรู้จดจำใบหน้า (Face Recognition) คือ กระบวนการที่นำภาพใบหน้าที่ตรวจจับได้ไปประมวลผลแล้วจากขั้นตอนการตรวจจับใบหน้านำข้อมูลมาเปรียบเทียบกับฐานข้อมูลของใบหน้าเพื่อระบุว่าใบหน้าที่ตรวจจับได้ตรงกับบุคคลใด แล้วจึงนำผลลัพธ์ที่ได้ ส่งไปให้ระบบหรือโปรแกรมเพื่อประมวลผลอื่น ๆ ต่อไป ดังภาพที่ 2.1

ลิขสิทธิ์ของมหาวิทยาลัยราชภัฏรำไพพรรณี



ภาพที่ 2.1 เปรียบเทียบการทำงานของระบบสแกนใบหน้า
ที่มา : (พี.ซีซีทีวี เน็ตเวิร์ค เอ็นจิเนียริ่ง, 2563)

การเรียนรู้เชิงลึก (Deep Learning) เป็นการฝึกฝนคอมพิวเตอร์ให้สามารถทำงานได้เหมือนมนุษย์ เช่น การจดจำคำพูด การระบุภาพ หรือการคาดการณ์ แทนที่จะจัดระเบียบข้อมูลที่รันผ่านทางสมการที่กำหนดไว้ล่วงหน้า การเรียนรู้เชิงลึกจะกำหนดค่าพารามิเตอร์พื้นฐานเกี่ยวกับข้อมูลและฝึกให้คอมพิวเตอร์เรียนรู้ด้วยตัวเองโดยการจดจำรูปแบบวัตถุโดยการประมวลผลแบบหลายชั้น พื้นฐานการเรียนรู้เชิงลึกนำอัลกอริทึมที่พยายามจะสร้างแบบจำลองเพื่อแทนความหมายของข้อมูลในระดับสูงโดยการสร้างสถาปัตยกรรมข้อมูลขึ้นมาที่ประกอบไปด้วยโครงสร้างย่อย ที่เรียนแบบการทำงานของสมองมนุษย์ในกระบวนการประมวลผลข้อมูลและเป็นการสร้างรูปแบบ สำหรับใช้ในการตัดสินใจ การเรียนรู้เชิงลึกเป็นเซตย่อยของ Machine Learning หรือการเรียนรู้ด้วยเครื่อง ซึ่งเป็นลำดับชั้นของเครือข่ายประสามเทียม (Artificial Neural Network) โดยดำเนินการด้วย Machine Learning เนื่องจากเครือข่ายประสามเทียมถูกสร้างมาเหมือนสมองของมนุษย์มี Nodes เชื่อมต่อกันเหมือนเว็บไซต์ แม้ว่าโปรแกรมแบบเก่าจะสร้างการวิเคราะห์ข้อมูลเชิงเส้น โดยฟังก์ชันลำดับชั้นของระบบการเรียนรู้เชิงลึก ช่วยให้เครื่องประมวลผลข้อมูลด้วยวิธีการไม่ใช่เชิงเส้น ซึ่งเป็นเทคนิคการเรียนรู้เชิงลึก เพื่อกำจัดสิ่งแปลกปลอมในการทำธุรกรรมที่เสี่ยงต่อการหลอกลวงซึ่งจักประกอบด้วยเวลา ที่ตั้งทางภูมิศาสตร์ ตำแหน่งที่อยู่หมายเลขเครื่อง ประเภทของธุรกิจค้าปลีก และคุณลักษณะอื่น ๆ ดังภาพที่ 2.3 การเรียนรู้เชิงลึกปัญญาประดิษฐ์



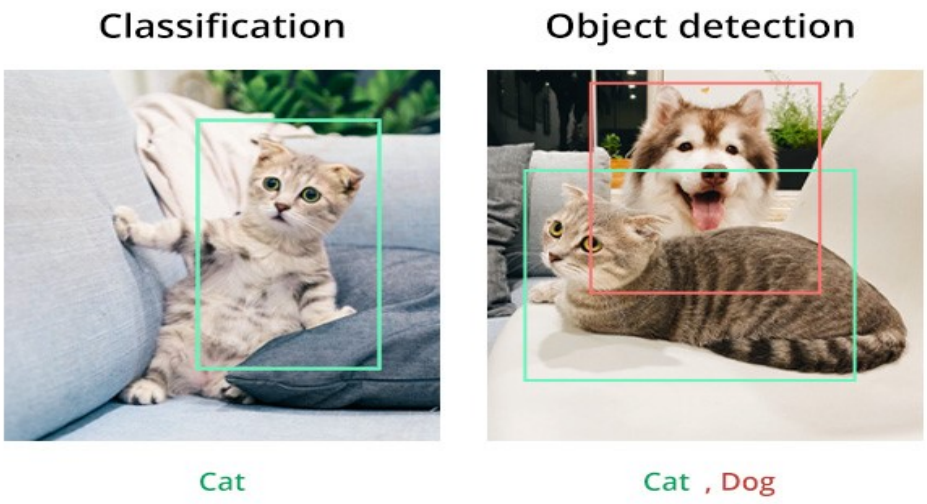
ภาพที่ 2.2 เทรนลักษณะใบหน้า
ที่มา : (ปัญญาประดิษฐ์, 2563)

ปัญญาประดิษฐ์

ปัญญาประดิษฐ์ ย่อมาจาก Artificial Intelligence (AI) โปรแกรมที่ถูกเขียนและพัฒนาให้มีความฉลาด มีความสามารถคิด วิเคราะห์ วางแผน และตัดสินใจได้ จากการประมวลผลของฐานข้อมูลขนาดใหญ่ และยังสามารถดัดแปลงการประมวลผล ประยุกต์ให้เป็นไปตามสถานการณ์ต่าง ๆ AI ยังถูกนำไปใช้ในอุตสาหกรรมอื่น ๆ อีกมากมาย อย่างเช่น เรื่องของการลงทุน ที่ให้ AI เป็นตัวช่วยในการประมวลผลเป็นต้น นอกจากนี้ในวงการแพทย์ การเกษตร ประกันภัย การขนส่ง การก่อสร้าง และการผลิต AI เป็นเทคโนโลยีที่จะเข้ามามีบทบาทกับหลากหลายอุตสาหกรรม และคาดว่าจะเป็นเทคโนโลยีที่เข้ามามีบทบาทและมีความสำคัญในการทำงานและการใช้ชีวิตในยุคดิจิทัลอย่างมาก ดังนั้นความสามารถของ AI ล้วนไม่ธรรมดาอย่างแน่นอน ทั้งเรื่องการประมวลผลข้อมูลที่รวดเร็วและแม่นยำ การเข้ามาช่วยเรื่องการลงทุน คิดแทนมนุษย์ในบางเรื่องช่วยตัดสินใจเป็นต้น หนึ่งในฟีเจอร์หลักของ AI (Artificial Intelligence) ที่ใช้กับกล้องวงจรปิด สามารถค้นหาสิ่งของโดยใช้ AI มาวิเคราะห์ข้อมูล จากการมองเห็นของคอมพิวเตอร์ (Computer Vision) และการประมวลผลภาพ (Image Processing) เพื่อตรวจจับวัตถุที่อยู่ในรูปหรือวิดีโอ เช่น มนุษย์ สัตว์ สิ่งของ รถยนต์ อาคาร และวัตถุอื่น ๆ ที่อยู่ในรูปภาพ หรือวิดีโอ โดยตามหลักแล้วก่อนที่จะพัฒนามาเป็นเทคโนโลยีตรวจจับ

สำนักวิทยบริการและเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏรำไพพรรณี

วัตถุ (Object Detection) จะต้องผ่านการจัดจำแนกหมวดของวัตถุ (Object Classification) มาก่อน โดยที่การจัดหมวดของวัตถุจะเป็นการจัดหมวดหมู่ของรูปภาพว่ารูปภาพนั้นคือภาพอะไร แต่เทคโนโลยีตรวจจับวัตถุจะเป็นการระบุเลยว่าในรูปภาพนั้นมีวัตถุอะไรบ้าง ซึ่งจุดนี้จะต้องอาศัยการทำงานของ AI เข้ามาช่วยในการวิเคราะห์ข้อมูลด้วยเช่นกัน ดังภาพที่ 2.2 การตรวจจำแนกวัตถุ



ภาพที่ 2.3 การตรวจจับวัตถุ
ที่มา : (ปัญญาประดิษฐ์, 2563)

ความสามารถของปัญญาประดิษฐ์ในการทำงานโดยรวบรวมข้อมูลปริมาณมหาศาลด้วยความเร็วประมวลผลซ้ำ ๆ ผ่านขั้นตอนการประมวลผลที่ชาญฉลาด อันช่วยให้ซอฟต์แวร์สามารถเรียนรู้จากรูปแบบและลักษณะของข้อมูลได้อย่างอัตโนมัติ AI เป็นแขนงของการศึกษาที่กว้างขวาง อันประกอบไปด้วย ทฤษฎีมากมาย วิธีการและเทคโนโลยี รวมถึงแขนงย่อยหลัก ๆ อันได้แก่ การเรียนรู้ของเครื่อง (Machine learning) ในการสร้างแบบจำลองการวิเคราะห์แบบอัตโนมัติ โดยใช้วิธีการจากโครงข่ายประสาทเทียม สถิติ การวิจัยดำเนินการ (operations research) และหลักฟิสิกส์ในการค้นหาข้อมูลเชิงลึกที่ซ่อนอยู่ในข้อมูลโดยไม่จำเป็นต้องเขียนโปรแกรมในการค้นหา โครงข่ายประสาทเทียม คือหนึ่งในระบบการเรียนรู้ของเครื่อง โดยใช้การเชื่อมโยงระหว่างยูนิตมีลักษณะเหมือนกับเซลล์ประสาท ทำหน้าที่ประมวลผลข้อมูล โดยการตอบสนองต่อข้อมูลภายนอก ถ่ายทอดข้อมูลซึ่งกันและกันระหว่างแต่ละหน่วยการประมวลผลจำเป็นต้องใช้ทางผ่านข้อมูลหลายทางเพื่อค้นหาความเชื่อมโยงและถ่ายทอดความหมายจากข้อมูลที่ไม่ชัดเจนเหล่านั้น การเรียนรู้เชิงลึก (Deep Learning) ใช้โครงข่ายประสาทเทียมขนาดใหญ่ที่มีหน่วยประมวลผลหลายชั้น

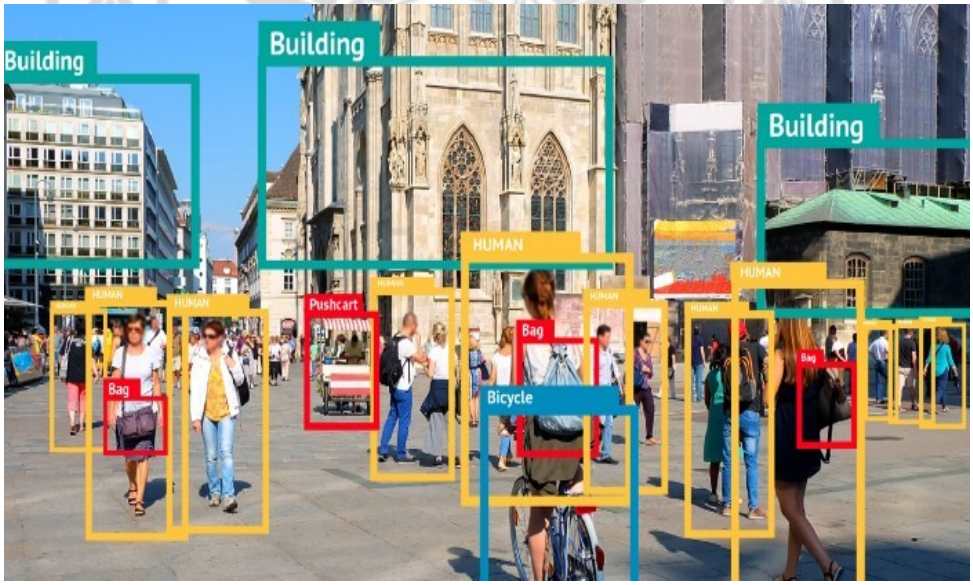
โดยอาศัยประโยชน์จากความก้าวหน้าในศักยภาพของคอมพิวเตอร์และเทคนิคในการเรียนรู้รูปแบบของข้อมูลปริมาณมหาศาลที่มีความซับซ้อนที่ได้รับการพัฒนาให้ดียิ่งขึ้นแล้ว แอปพลิเคชันแบบทั่วไปนั้นหมายถึงการจดจำภาพและคำพูดระบบการประมวลผลข้อมูลที่มีการเรียนรู้ (Cognitive Computing) เป็นแขนงย่อยหนึ่งของ AI ที่พยายามแสดงปฏิสัมพันธ์ให้เสมือนมนุษย์ผ่านเครื่องจักรกล การใช้ AI และการประมวลผลหน่วยความจำ มีเป้าหมายสูงสุดคือ การใช้เครื่องจักรกลในการเลียนแบบกระบวนการของมนุษย์ผ่านความสามารถในการตีความภาพและคำพูด และตอบสนองโดยทันที การประมวลผลภาพ (Computer Vision) ใช้การจดจำรูปแบบและการเรียนรู้เชิงลึกในการจดจำสิ่งที่อยู่ในภาพหรือวิดีโอ เมื่อเครื่องจักรกลสามารถประมวลผล วิเคราะห์และเข้าใจรูปภาพ มันจะสามารถจับภาพหรือวิดีโอได้แบบเรียลไทม์และตีความสภาพแวดล้อมได้ การประมวลผลภาษาธรรมชาติ (Natural Language Processing หรือ NLP) คือความสามารถของคอมพิวเตอร์ในการวิเคราะห์ ทำความเข้าใจและสร้างภาษามนุษย์ ซึ่งรวมถึงคำพูดด้วย ขั้นถัดไปของ NLP คือ การโต้ตอบด้วยภาษาธรรมชาติ ซึ่งช่วยให้มนุษย์สามารถสื่อสารกับคอมพิวเตอร์ได้โดยใช้ภาษาเพื่อดำเนินการงานต่าง ๆ นอกจากนี้ ยังมีหลายเทคโนโลยีที่ช่วยและส่งเสริมประสิทธิภาพของ AI เทคโนโลยีเหล่านั้นได้แก่ หน่วยประมวลผลกราฟิกเป็นกุญแจสำคัญของ AI เนื่องจากหน่วยประมวลผลจะช่วยเพิ่มพลังในการคำนวณอันจำเป็นต่อกระบวนการประมวลผลซ้ำไปมา การฝึกอบรมโครงข่ายประสาทจำเป็นต้องใช้ข้อมูลขนาดใหญ่และพลังงานในการคิดคำนวณ Internet of Things ก่อให้เกิดปริมาณข้อมูลมหาศาลจากอุปกรณ์ที่เชื่อมโยงอยู่ ซึ่งข้อมูลส่วนใหญ่มักไม่ผ่านการวิเคราะห์ แบบจำลองอัตโนมัติที่ใช้ AI จะช่วยให้ใช้ประโยชน์จากแบบจำลองได้อย่างเต็มที่อัลกอริธึมขั้นสูงกำลังได้รับการพัฒนาและผนวกรวมเป็นวิธีใหม่เพื่อทำการวิเคราะห์ข้อมูลที่รวดเร็วกว่าและได้หลายระดับข้อมูล กระบวนการอันชาญฉลาดนี้คือ กุญแจสำคัญในการระบุและพยากรณ์เหตุการณ์ที่เกิดขึ้นได้ยาก ทำความเข้าใจระบบที่ซับซ้อนและปรับเพื่อให้ได้มาซึ่งสถานการณ์ที่เหมาะสมที่สุด APIs หรือแอปพลิเคชันประมวลผลอินเตอร์เฟส เป็นแพ็คเกจของโค้ดคำสั่งที่สามารถพกพาได้ช่วยให้การเพิ่มเติมฟังก์ชันการทำงานของ AI ไปยังผลิตภัณฑ์ที่มีอยู่แล้วและแพ็คเกจซอฟต์แวร์สามารถเป็นไปได้ โดยมันสามารถเพิ่มความสามารถในการจดจำภาพ เพื่อจัดทำระบบความปลอดภัยและการตอบคำถาม ซึ่งสามารถอธิบายข้อมูล สร้างแคปชันและหัวเรื่อง หรือค้นหารูปแบบข้อมูลและเนื้อหาที่น่าสนใจได้ โดยสรุปเป้าหมายของ AI คือการมอบซอฟต์แวร์ที่สามารถหาคำตอบด้วยการคิดหาเหตุผลจากอินพุตที่ใส่เข้าไปและอธิบายคำตอบนั้นผ่านการแสดงผล AI จะแสดงปฏิสัมพันธ์เสมือนมนุษย์ผ่านซอฟต์แวร์และมอบเหตุผลสนับสนุนการตัดสินใจในงานเฉพาะ หาก AI ไม่ใช่สิ่งที่จะมาแทนที่มนุษย์ และจะไม่เป็นเช่นนั้นในระยะเวลาอันใกล้ ความสำคัญของการใช้ความสามารถของปัญญาประดิษฐ์ AI มีการเรียนรู้ซ้ำ ๆ ได้อย่างอัตโนมัติและศึกษาผ่านข้อมูลเหล่านั้น แต่ AI นั้นก็มีความแตกต่างจากหุ่นยนต์หรืออุปกรณ์อัตโนมัติ แทนที่จะประมวลผลงานแบบแมนนวล AI สามารถ

ประมวลผลในงานซ้ำ ๆ ที่มีปริมาณมากด้วยความเที่ยงตรงและมีประสิทธิภาพผ่านระบบคอมพิวเตอร์ สำหรับการประมวลผลการทำงานอัตโนมัติด้วยวิธีนี้ ยังคงจำเป็นต้องใช้มนุษย์ในการติดตั้งระบบและป้อนคำสั่งที่เหมาะสม AI เพิ่มความชาญฉลาด แก่ผลิตภัณฑ์ที่มีอยู่เดิม โดยทั่วไป จะไม่มีการจำหน่าย AI ในรูปแบบแอปพลิเคชันเดี่ยว หากแต่จะใช้ประสิทธิภาพของ AI ในการพัฒนาผลิตภัณฑ์ที่มีอยู่เดิม ซึ่งประสิทธิภาพดังกล่าวมีความเหมือนอย่างมากกับ Siri ที่ได้รับการติดตั้งเพิ่มในผลิตภัณฑ์รุ่นใหม่ ๆ ของ Apple เครื่องจักรอัตโนมัติ (Auto machine) เครื่องจักรที่โต้ตอบกับมนุษย์ได้ (conversational platform) โปรแกรมบอท (Bot) และเครื่องจักรอัจฉริยะ (Smart Machine) จะได้รับการผสานเข้ากับข้อมูลปริมาณมหาศาลเพื่อพัฒนาเทคโนโลยีไม่ว่าจะเป็นที่บ้านหรือที่ทำงาน จากระบบความมั่นคงอัจฉริยะ (Security Intelligence) สู่การวิเคราะห์การลงทุน (Investment Analysis) AI เรียนรู้จากอัลกอริธึมการเรียนรู้แบบก้าวหน้า (Progressive) ในการนำข้อมูลมาเขียนคำสั่งโปรแกรม AI ค้นหาโครงสร้างและความสม่ำเสมอของข้อมูล เพื่อกำหนดอัลกอริธึมทักษะด้านต่าง ๆ กล่าวคือ อัลกอริธึมจะกลายเป็นตัวแยกประเภทหรือตัวพยากรณ์ ดังนั้นอัลกอริธึมจะสามารถเรียนรู้วิธีการเล่นหมากรุก และเรียนรู้ว่าควรจะเดินหมากตัวใดในตาถัดไป ซึ่งแบบจำลองประเภทนี้จะได้รับการปรับให้ดีขึ้นเมื่อได้รับข้อมูลใหม่ กระบวนการส่งค่าย้อนกลับ (Back Propagation) คือเทคนิคหนึ่งของ AI ในการปรับแต่งแบบจำลองผ่านการฝึกฝนและข้อมูลเพิ่ม เมื่อผลลัพธ์ครั้งแรกยังไม่ถูกต้องนัก AI จะวิเคราะห์ข้อมูลมากกว่าและลึกกว่า โดยใช้เครือข่ายประสาทเทียม (Neural Network) ที่มีหลายชั้น การสร้างระบบตรวจจับการทุจริตที่มี 5 ลำดับชั้นนั้นแทบจะเป็นไปไม่ได้เลยในไม่กี่ปีที่ผ่านมา แต่ทั้งหมดนั้นได้เปลี่ยนแปลงไปด้วยประสิทธิภาพอันเหลือเชื่อของคอมพิวเตอร์และข้อมูลขนาดใหญ่จำเป็นต้องใช้ข้อมูลปริมาณมากในการพัฒนาด้านการเรียนรู้เชิงลึกของแบบจำลอง เนื่องจากแบบจำลองเหล่านี้จะเรียนรู้จากข้อมูลโดยตรง ยิ่งสามารถป้อนข้อมูลปริมาณมากขึ้นเท่าใด แบบจำลองจะยิ่งก่อให้เกิดความแม่นยำมากขึ้นเท่านั้น AI สามารถสร้างความแม่นยำอย่างเหลือเชื่อผ่านเครือข่ายประสาทเทียมซึ่งที่ผ่านมาไม่สามารถเป็นไปได้เลย ยกตัวอย่างเช่น การโต้ตอบกับ Google Search และ Google Photos ล้วนใช้เทคนิคการเรียนรู้เชิงลึก (Deep Learning) ทั้งนั้น และนับวันโปรแกรมเหล่านี้ยิ่งมีความแม่นยำมากยิ่งขึ้นตามการใช้งานที่เพิ่มขึ้น ในทางการแพทย์ เทคนิคของ AI ด้านการเรียนรู้เชิงลึก การจำแนกภาพ (Image Classification) และการจดจำวัตถุ (Object Recognition) ได้รับการใช้ค้นหามะเร็งผ่านเครื่อง MRIs ด้วยความแม่นยำจากรังสีแพทย์ที่ได้รับการฝึกฝน AI สามารถใช้ประโยชน์อย่างสูงสุดจากข้อมูลที่มี เมื่ออัลกอริธึมสามารถเรียนรู้ได้ด้วยตัวเอง ข้อมูลก็จะกลายเป็นทรัพย์สินทางปัญญาอันมีค่า ความลับซ่อนอยู่ในข้อมูลนั่นเอง เพียงแค่สามารถประยุกต์ใช้ AI เพื่อดึงเอาความลับนั้นออกมา เนื่องจากบทบาทของข้อมูลนับว่ามีความสำคัญมากกว่าที่เคยเป็นมา มันสามารถก่อให้เกิดความได้เปรียบทางการแข่งขัน หากมีข้อมูลที่ดีที่สุดใน

อุตสาหกรรมที่มีการแข่งขันกัน แม้ว่าต่างคนจะใช้เทคนิคกลวิธีที่เหมือนกัน ผู้ซึ่งมีข้อมูลที่ดีที่สุดย่อมเป็นผู้ชนะ

การตรวจจับวัตถุ

การตรวจจับวัตถุ (Object Detection) มีเป้าหมายเพื่อต้องการตรวจหาตำแหน่งของวัตถุพร้อมทั้งระบุตำแหน่งของวัตถุที่ทำการตรวจจับพบว่าวัตถุนั้นอยู่ในตำแหน่งไหนของภาพหรือวิดีโอ โดยที่วัตถุนั้นสามารถเป็นได้ทั้ง คน สัตว์ หรือสิ่งของ ซึ่งเป็นระเบียบวิธีที่พัฒนาขึ้นเพื่อให้เครื่องจักรค้นหาสิ่งของในภาพหรือวีดีโอแทนคน ก่อนที่จะนำมาเป็นโมเดลของการตรวจจับวัตถุที่แสดงผลลัพธ์ดังภาพที่ 2.3 ได้ จำเป็นต้องให้คอมพิวเตอร์เรียนรู้ภาพจำนวนหนึ่ง (Dataset) ของวัตถุที่ต้องการจะตรวจจับพร้อมทั้งบอกให้คอมพิวเตอร์รู้ว่าวัตถุที่กำหนดให้ระบบเรียนรู้อยู่ตอนนี้คือวัตถุอะไร และตำแหน่งอยู่ตรงไหน ซึ่งจะต่างจากการทำจำแนกประเภทของรูปภาพ (Image Classification) เหมือนอย่างทั่วไปที่ต้องการรู้แค่เพียงชนิดของรูปภาพเพียงอย่างเดียว



ลิขสิทธิ์ของมหาวิทยาลัยราชภัฏรำไพพรรณี

ภาพที่ 2.4 การตรวจจับภาพ

ที่มา : (Markus Hafellner, 2019)

หลักการทำ AI ตรวจจับวัตถุการตรวจจับวัตถุในรูปภาพ ด้วย TensorFlow.js โดยใช้โมเดลสำเร็จรูป COCO-SSD ซึ่งเป็นโมเดลขนาดเล็ก ไม่สิ้นเปลืองหน่วยความจำมาก และเหมาะสำหรับใช้งานบน Web Browser และอุปกรณ์ขนาดเล็กพัฒนาด้วยภาษา JavaScript งานการตรวจจับวัตถุในรูปภาพ สามารถเจาะลึกลงไปได้อีกหลายแขนง เช่น การทำตรวจจับหน้าคนในพื้นที่จุดเสี่ยงเกิดเหตุ

อาชญากรรม ตรวจจับคนเดินถนน สามารถประยุกต์ใช้ได้หลากหลาย เช่น ใช้ในงานรักษาความปลอดภัย เป็นต้น AI ตรวจจับวัตถุ ในงานทางด้านมุมมองคอมพิวเตอร์ที่ใช้จำแนกและตรวจจับวัตถุที่อยู่ในรูปภาพ การตรวจจับ ระบุจุด ระบุพื้นที่ โดยหลักการสามารถทำได้หลายวิธี การทำระบุพื้นที่ที่นิยมได้แก่ วาดกล่องรอบวัตถุ (Bounding Box) หรือเติมสีให้ทุกพิกเซลของวัตถุนั้น เรียกว่า Segmentation ส่วนการทำ COCO-SSD คือ โมเดลสำหรับ AI ตรวจจับวัตถุ Object Detection ที่จะตรวจหา และจำแนก วัตถุทั้งหมดทุกชิ้น ที่อยู่ภายในภาพ 1 ภาพโมเดลที่จะใช้นี้คือ โมเดล COCO-SSD เวอร์ชันที่ถูกแปลง เพื่อมาให้ใช้กับ TensorFlow.js ผ่าน Tensorflow object detection API โมเดลนี้จะตรวจจับวัตถุ จำนวน 90 Class ตามใน COCO Dataset ซึ่งเป็น Dataset ขนาดใหญ่ที่ใช้ในการ Object Detection Segmentation และ Captioning SSD ย่อมาจาก Single Shot Multi-Box Detection ตัวอย่างการกำหนดค่าใส่รหัสโปรแกรมด้านล่างไว้ในส่วนระหว่าง HTML tag head และ body โค้ดนี้เป็นการโหลด TensorFlow.js และ โหลดโมเดล COCO- SSD ที่ใช้ในการทำการตรวจจับวัตถุ

```
<script src="https://cdn.jsdelivr.net/npm/@tensorflow/tfjs@latest"> </script>
<script src="https://cdn.jsdelivr.net/npm/@tensorflow-models/coco-ssd">
</script>
```

กำหนดรูปภาพ Input ใน canvas สำหรับวาดรูปและ Bounding Box และ div ข้อความ Output

```
<div id="output"></div>
<canvas id="canvas" width="320" height="320"></canvas>
<img id="canvasimg" src="img/img01.jpg" styel="display:none;"/>
```

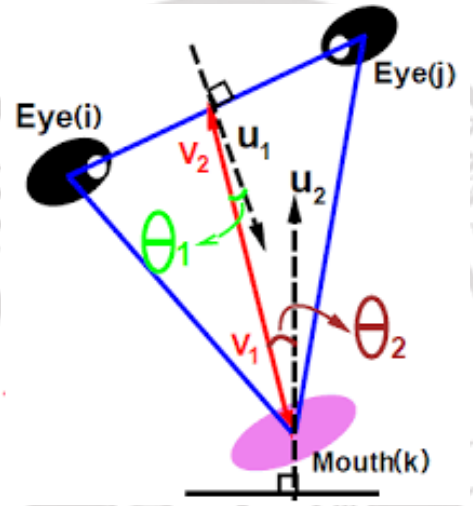
ใน JavaScript ให้ getElementById ของ 3 Element ด้านบนมาเก็บไว้ก่อน ดึง context จาก canvas แล้วนำรูปที่ซ่อนไว้มาวาดลงไป

สิงห์เกียรติมหาวิทยาลัยราชภัฏรำไพพรรณี

```
var canvas = document.getElementById("canvas");
var img = document.getElementById("canvasimg");
const outp = document.getElementById('output');
```

แบบจำลองเชิงเรขาคณิตสามเหลี่ยมของตาและปาก

แบบจำลองเชิงเรขาคณิตสามเหลี่ยมของตาและปาก ประกอบด้วยสามจุดคือ ตาซ้าย (i) ตาขวา (j) และปาก (k) ดังแสดงในภาพที่ 2.4 ซึ่งใช้ในการหาตำแหน่งของตาและปากโดยเลือกสามเหลี่ยมที่มีค่าน้ำหนักความเอียงน้อยที่สุดจากสมการที่ (1)



ภาพที่ 2.5 แสดงแบบจำลองเชิงเรขาคณิตสามเหลี่ยมของตาและปาก
ที่มา : (สมปอง เวฬุวนาธร และ สุพจน์ นิตย์สุวัฒน์, 2562)

$$ow(i,j,k)=\prod_{r=1}^{2}e^{-3(1-\cos^2(\theta_r(i,j,k)))} \qquad (1)$$

$$\cos(\theta_r(i,j,k))=\frac{\vec{u_r}\cdot\vec{v_r}}{\left\|\vec{v_r}\right\|} \qquad (2)$$

ลิขสิทธิ์ของมหาวิทยาลัยราชภัฏรำไพพรรณี

เมื่อ ow คือค่าน้ำหนักความเอียงของสามเหลี่ยม

v1 และ v2 คือเวคเตอร์ระหว่างจุดกึ่งกลางของ ij กับ k

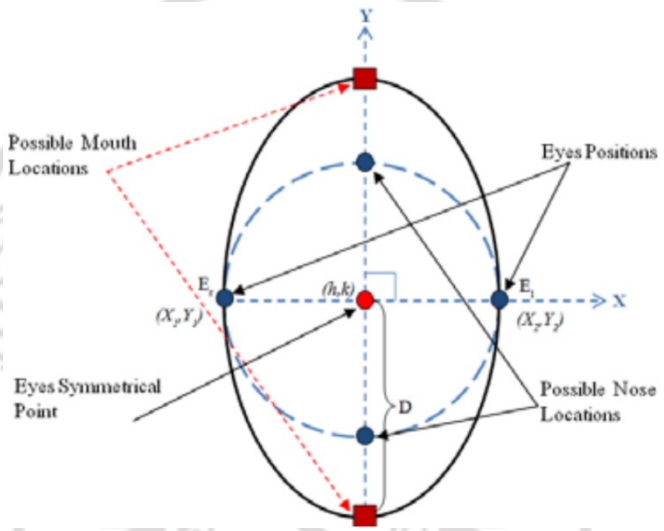
θ1 คือ มุมระหว่างเวคเตอร์ u1 และ v1

θ2 คือ มุมระหว่างเวคเตอร์ u2 และ v2

u1 คือเวคเตอร์ที่มีต้นกำเนิดที่จุดกึ่งกลางของ ij ที่ตั้งฉากกับเวคเตอร์ ij

u2 เป็นเวคเตอร์ที่มีตนกำเนิดที่ k ที่ตั้งฉากกับแนวราบ

จากภาพที่ 2.4 ถ้า θ_1 มีขนาดเข้าใกล้ 0 มุมที่จุดยอด i และ j จะมีขนาดใกล้เคียงกันและจะเท่ากันเมื่อ θ_1 เท่ากับ 0 และถ้า θ_1 และ θ_2 มีขนาดเข้าใกล้ 0 สามเหลี่ยมจะค่อย ๆ มีสมมาตรจนเป็นสามเหลี่ยมหน้าจั่วเมื่อ $\theta_1 = \theta_2 = 0$ และภาพจำลองรูปแบบใบหน้าดังภาพที่ 2.5 (สมปอง เวฬุวนาธร และ สุพจน์ นิตย์สุวัฒน์, 2562)



ภาพที่ 2.6 แสดงแบบจำลองใบหน้าเชิงเรขาคณิต
ที่มา : (สมปอง เวฬุวนาธรและ สุพจน์ นิตย์สุวัฒน์, 2562)

แบบจำลองใบหน้าเชิงเรขาคณิต เป็นอีกวิธีหนึ่งที่นิยมนำมาประยุกต์ใช้ในการช่วยหาค่าของตำแหน่งต่าง ๆ บนภาพใบหน้าซึ่ง ประกอบด้วยวงกลมและวงรีดังแสดงในภาพที่ 2-6 จุดตัดของวงกลมและวงรี จะเป็นสัดส่วนในการหาตำแหน่งของตาจมูกและปากดังนี้

ระยะห่างระหว่างจุดตัดบนแกน X ของวงกลมและวงรีทั้งสองจุด (Er และ El) หาได้จากสมการ

จุดกึ่งกลางระหว่างจุดตัดบนแกน X ของวงกลมและวงรี (h, k) หาได้จากสมการ

$$\overline{E_r E_l} = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2} \qquad (3)$$

$$(h,k) = \left(\frac{(X_2 - X_1)}{2}\right), \left(\frac{(Y_2 - Y_1)}{2}\right) \qquad (4)$$

จุดตัดบนแกน X ของวงกลมและวงรีคือตำแหน่งของตาซ้าย และตาขวา

จุดตัดบนแกน Y ของวงรีที่ตั้งฉากกับแกน X ด้วยระยะห่างเท่ากับ D คือ ตำแหน่งของปาก ซึ่งมีโอกาสเป็นได้ทั้งจุดบนและจุดล่าง

จุดตัดบนแกน Y ของวงกลม ด้วยระยะห่าง 0.6D คือ ตำแหน่งของจมูก ซึ่งมีโอกาสเป็นได้ทั้งจุดบนและจุดล่าง

จุดด้อยของแบบจำลองใบหน้าเชิงเรขาคณิตคือ ต้องทราบตำแหน่ง ของตาทั้งสองข้างมาก่อนแล้ว จึงจะสามารถคำนวณหาตำแหน่งของจมูกและปากได้ (สมปอง เวฬุวนาธรและ สุพจน์ นิตย์สุวัฒน์, 2562)

งานวิจัยที่เกี่ยวข้อง

ปฏิญญา ตันฑวิวัฒน์ และคนอื่นๆ . (2564) การวิเคราะห์ภาพคนและสัมภาระสำหรับการตรวจจับวัตถุที่ปราศจากเจ้าของซึ่งงานวิจัยนี้ใช้ชุดข้อมูลในรูปแบบเปิดสาธารณะสำหรับการฝึกสอนระบบและทดสอบระบบ โดยมีการประเมินผลในสี่ส่วน ประกอบด้วย 1) ประสิทธิภาพแบบจำลอง YOLOv3 เปรียบเทียบกับ YOLOv3-tiny 2) ประสิทธิภาพการตรวจจับวัตถุแต่ละชนิด 3) ประสิทธิภาพการระบุความเป็นเจ้าของ และ 4) ประสิทธิภาพการระบุ การละทิ้งกระเป๋า ข้อมูลที่นำใช้ในงานวิจัยนี้มี 2 ส่วนนั้นคือ ส่วนแรกเป็นข้อมูลสำหรับฝึกสอน การรู้จำคนและกระเป๋า และส่วนที่สองคือข้อมูล i-Lids AVSS-AB สำหรับการวิเคราะห์ความเป็นเจ้าของและการละทิ้งกระเป๋าใช้แบบจำลอง YOLOv3 สำหรับการตรวจจับวัตถุซึ่งประมวลผลได้แบบเรียลไทม์ และมีประเมินผลด้านประสิทธิภาพความถูกต้องเท่ากับร้อยละ 98 ส่วนการวิเคราะห์ความเป็นเจ้าของและการละทิ้งกระเป๋าจะใช้วิธีการพิจารณาความสัมพันธ์ในเชิงตำแหน่งและการเคลื่อนที่ของคนและกระเป๋า ซึ่งประสิทธิภาพการระบุความเป็นเจ้าของมีค่าความถูกต้องเฉลี่ยอยู่ที่ร้อยละ 65.1 และระบบมีความสามารถในการวิเคราะห์ประสิทธิภาพของการระบุการละทิ้งวัตถุที่ร้อยละ 66.6

ราเมศวร์ และคนอื่น ๆ (2564). การตรวจจับการสวมหน้ากากอนามัยโดยประยุกต์ใช้โครงข่ายประสาทเทียมแบบเชิงลึกและโครงข่ายประสาทเทียมแบบคอนโวลูชันโดยมีการดำเนินงานของการวิจัยบรรลุวัตถุประสงค์ที่ได้กำหนดไว้คือเพื่อพัฒนาระบบตรวจจับการสวมใส่หน้ากากภาพ การพัฒนาระบบตรวจสอบการสวมหน้ากากอนามัยมีขั้นตอนการพัฒนา 3 ขั้นตอน ประกอบด้วย 1) การเก็บรวบรวมข้อมูล 2) การสร้างแบบจำลอง และ 3) การประยุกต์ใช้แบบจำลอง ทีมวิจัยได้เก็บข้อมูลใบหน้าภาพแบบสวมหน้ากาก 1,919 ภาพ และไม่สวมหน้ากาก 1,916 ภาพ มีลวดลาย ก้มหน้า เอียงหน้า และสิ่งปกปิดอัตลักษณ์บนใบหน้าทั้งหมด 3,835 ภาพ ใช้เครื่องมือคือ Confusion Matrix ทำให้ผลการวิจัยแบบ Per Frame จากการทดลองมีผลลัพธ์ของค่าความถูกต้องมากสุดอยู่ที่ร้อยละ

97.00 และการทดลองแบบ Real Time ใช้เครื่องมือคือ Intersection over Union (IoU) ความถูกต้องอยู่ที่ร้อยละ 98.00 และแบบค่าความถูกต้องแบบตามเหตุการณ์อยู่ที่ร้อยละ 96.00 ใช้ Mobile NetV2 ร่วมกับกระบวนการโครงข่ายประสาทเทียม (Artificial Neuron Network : ANN)

พุฒิพงศ์ จันทร์แจ่ม. (2563) การปรับปรุงประสิทธิภาพของกระบวนการบริการลูกค้าโดยใช้การวิเคราะห์จากกล้องวงจรปิด โดยงานวิจัยมีการประเมินประสิทธิภาพกระบวนการบริการลูกค้าของร้านอาหารทั่วไป (Casual Dining) มีหลายส่วนซึ่งหนึ่งในนั้นคือการประเมินความรวดเร็วในการบริการโดยจะวัดจากระยะเวลาที่ลูกค้าใช้ในการรอรับบริการในแต่ละช่วงของการบริการ หากทำการประเมินด้วยมนุษย์ในวิธีต่าง ๆ เช่น ใช้เครื่องจับเวลามีโอกาสเกิดความคาดเคลื่อนได้สูง ซึ่งในปัจจุบันระบบตรวจจับวัตถุด้วยเทคนิคการเรียนรู้เชิงลึกมีการพัฒนาอย่างต่อเนื่องและมีความแม่นยำสูง รวมถึงร้านอาหารโดยทั่วไปมีการติดตั้งกล้องวงจรปิดมากขึ้นจึงเป็นอีกทางเลือกหนึ่งในการนำข้อมูลดังกล่าวมาประยุกต์ใช้กับเทคนิคการตรวจจับการให้บริการของพนักงาน งานวิจัยนี้ได้นำเสนอแนวทางในการพัฒนาและปรับปรุงระบบการประเมินระยะเวลาที่ลูกค้าใช้ในการรอรับการบริการจากพนักงานทั้งหมด 3 ช่วง ได้แก่ การต้อนรับ (Greeting) รับคำสั่ง (Order) และจัดส่งมอบอาหาร (Serve) จากการฝึกแบบจำลองในการตรวจจับการเข้าออกของพนักงานผ่านกล้องวงจรปิดด้วยการเรียนรู้เชิงลึก YOLOv4 ผ่านไลบรารี Darknet ซึ่งมีค่า Average IoU เท่ากับร้อยละ 71.86 และ mAP เท่ากับร้อยละ 96.56

ศุภกิตติ โสภาส. (2560) การพัฒนาเทคนิคการตรวจจับพื้นที่ใบหน้าและวัตถุบริเวณดวงตาโดยใช้การประมวลผลภาพ ซึ่งงานวิจัยการค้นหาและเปรียบเทียบภาพลักษณะของใบหน้าโดยใช้เทคนิคการประมวลผลภาพเป็นขั้นตอนสำคัญในการออกแบบระบบการจดจำใบหน้า แต่การตรวจจับใบหน้ายังเกิดโจทย์รูปร่างและลักษณะแตกต่างกันโดยการทดลองใช้ภาพจำนวน 200 ภาพ จำแนกภาพด้วยเทคนิคการใช้โมเดลสีและใช้เทคนิควิธีการ Sobel Edge Detection มาใช้ในการตรวจจับใบหน้าและตรวจจับพื้นที่บริเวณรอบดวงตา ความผิดพลาดส่วนใหญ่เกิดจากใบหน้ามีความสว่างน้อยหรือมากเกินไปและสีของผิวคน

ลิขสิทธิ์ของมหาวิทยาลัยราชภัฏรำไพพรรณี