# การพัฒนาระบบตรวจจับใบหน้าสำหรับจอโฆษณาอัจฉริยะ

## Development of Facial Detection System for a Smart Digital Signage

อนุชา ตุงคัษฐาน[1] สุเทพ เชาว์สนิท[2] วิชิตา ตุงคัษฐาน[3] และ นิพัทธ์ จงสวัสดิ์[4]
สาขาเทคโนโลยีคอมพิวเตอร์ มทร.ธัญบุรี ปทุมธานี[1, 2]
สาขาวิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยราชภัฏหมู่บ้านจอมบึง ราชบุรี[3]
สาขาวิชาการวิเคราะห์และการจัดการข้อมูลขนาดใหญ่ มทร.ธัญบุรี ปทุมธานี[4]
E-mail: anucha_t@rmutt.ac.th[1], sutep_c@rmutt.ac.th[2], vichitatun@mcru.ac.th[3], nipat_j@rmutt.ac.th[4]

**บทคัดย่อ**

จอโฆษณาแอลอีดีขนาดใหญ่ในปัจจุบันนิยมนำมาติดตั้งในจุดที่มีผู้คนสัญจรเป็นจำนวนมาก ติดตั้งในเขตเมือง เส้นถนนหลัก เส้นทางรถไฟฟ้า ติดตั้งตามอาคารสูงเพื่อให้คนจำนวนมากได้เห็นภาพโฆษณาที่ปรากฏบนจอ งานวิจัยนี้ผู้วิจัยร่วมกับผู้ประกอบการพัฒนาจอโฆษณาอัจฉริยะที่มีความสามารถในการตรวจจับผู้ที่สนใจภาพโฆษณาที่ปรากฏบนจอแอลอีดีเพื่อนำข้อมูล เช่น จำนวนคน เพศ ช่วงอายุ คนที่สนใจโฆษณาในช่วงเวลาต่างๆ เพื่อนำข้อมูลมาวิเคราะห์สำหรับการปล่อยภาพโฆษณาให้ตรงกับช่วงเวลาและตรงกับกลุ่มเป้าหมาย รวมถึงการนำข้อมูลสำคัญจากในพื้นที่มาใช้ในการวางแผนการตลาด ผู้วิจัยได้พัฒนาคุณสมบัติใหม่นี้ให้กับจอโฆษณาแอลอีดีขนาดใหญ่ที่ติดตั้งอยู่ใจกลางเมืองให้เป็นจอโฆษณาอัจฉริยะ โดยติดตั้งกล้องและกล่องควบคุมและประมวลผลที่บริเวณจอ พัฒนาระบบปัญญาประดิษฐ์และการเรียนรู้ของเครื่อง (Machine Learning) ด้วยโครงข่ายประสาทเทียม การเรียนรู้เชิงลึก (Deep learning) แบบคอนโวลูชัน ร่วมกับระบบสมองกลฝังตัว (Embedded System) เพื่อเพิ่มความสามารถให้กับจอในการนับจำนวนคนในพื้นที่ แยกเพศชาย-หญิง ประมาณการช่วงอายุ และตรวจจับคนที่สนใจโฆษณาในช่วงเวลาต่างๆ ทำการบันทึกข้อมูลและส่งข้อมูลขึ้นประมวลผลบนระบบคลาวด์ผ่านเครือข่ายไร้สาย ผลจากการดำเนินงานจอโฆษณาอัจฉริยะที่พัฒนานั้นสามารถทำงานได้อย่างมีประสิทธิภาพและเป็นทรัพยากรและโครงสร้างพื้นฐานรองรับการพัฒนาเมืองอัจฉริยะต่อไป

**Abstract**

The smart LED digital signage is so popular. It is installed in a very crowded area, along the main road in the city, the sky train stations, or tall buildings in order to let more people see advertisements on advertising screens. In this research, researchers together with the company's developers have developed a smart digital signage that is capable of detecting people who are watching

advertisements. It can count the number of people by detecting human faces and classify gender and age from images. The data are captured and analyzed in order to launch advertisements or advertising campaigns at the right time, hit a target group, and make a marketing plan. Researchers developed new features for the digital signage system, which was installed in the city's area, to make it smarter. The digital signage is equipped with the camera and camera control unit. Several AI techniques such as machine Learning, deep learning in embedded systems with implementation of a convolutional neural network were developed to provide digital signage systems with built-in functions and advertising capabilities. It obtains an accurate count of the number of people, classifies gender, estimates people's ages, and classifies those who watch the advertisements at different time interval. The data were recorded and transferred from applications to cloud data centers through wireless channels. The smart digital signage can attract and direct the attention of people and it plays a significant role in developing a more efficient intelligent infrastructure in smart cities.

## 1. บทนำ

การพัฒนา "เมืองอัจฉริยะ" เป็นหัวข้อที่ทั้งภาครัฐและภาคเอกชนให้ความสำคัญและมีการกล่าวถึงอยู่เสมอในปัจจุบัน ทำให้องค์กรธุรกิจต้องมีการปรับตัวและมีการแข่งขันกันอย่างสูง ธุรกิจหนึ่งที่ได้รับผลกระทบและมีการปรับตัวอย่างเห็นได้ชัดคือธุรกิจจอโฆษณา โดยได้ปรับเปลี่ยนรูปแบบจากป้ายโฆษณาโปสเตอร์ที่ถูกใช้มาอย่างยาวนานและไม่ตอบโจทย์การเป็นเมืองอัจฉริยะมาเป็นจอแอลอีดี (LED) ขนาดใหญ่ที่มีความสามารถในการแสดงผลภาพวีดีโอโฆษณาที่มีความละเอียดสูง สามารถมองเห็นได้อย่างชัดเจนจากระยะไกล อีกทั้งยังสามารถปรับเปลี่ยนโฆษณาได้ง่ายและสะดวกเพื่อตอบโจทย์ความต้องการของลูกค้า ในทางสังคมจอโฆษณาแอลอีดีเหล่านี้ยังสามารถทำให้คนในพื้นที่ได้รับข่าวสารข้อมูล เช่น ข่าวสารด้านการเปลี่ยนแปลงของสภาพภูมิอากาศ ฝุ่นละอองขนาดเล็ก PM 2.5 ฝุ่นควันพิษในพื้นที่ สภาพการจราจร หรือความรู้ที่เป็นประโยชน์เพื่อปรับตัวและนำมาใช้ในชีวิตประจำวันได้อย่างทันท่วงที แต่สำหรับด้านธุรกิจนั้นบริษัทจอโฆษณาแอลอีดีเหล่านี้ต้องการสร้างจุดแข็งด้วยการนำเทคโนโลยีและนวัตกรรมด้านปัญญาประดิษฐ์ (Artificial Intelligence: AI) มาใช้กับจอแอลอีดีเพื่อพัฒนาให้เป็นจอโฆษณาอัจฉริยะที่ทำให้เกิดความได้เปรียบคู่แข่งขันทางธุรกิจ นำเสนอข้อมูลข่าวสารได้ตรงกับความต้องการและมีประโยชน์กับกลุ่มประชากร ทำให้มีความสามารถในด้านการมีปฏิสัมพันธ์กับคนในพื้นที่ ทั้งนี้คำถามสำคัญสำหรับผู้ให้บริการจอโฆษณา

แอลอีดีและผู้ประกอบการที่เสียค่าใช้จ่ายในการลงโฆษณาคือมีความต้องการที่จะทราบว่าเมื่อปล่อยโฆษณาออกไปแล้วมีคนเห็นบ้างหรือไม่ และมีคนให้ความสนใจมากน้อยเพียงใด โดยคนที่สนใจนั้นเป็นเพศหญิงหรือเพศชายและอยู่ในช่วงอายุใด รวมไปถึงคนที่เห็นมีปฏิกิริยาอย่างไร เป็นต้น เพื่อนำข้อมูลมาวิเคราะห์ คำนวณช่วงเวลา และนำเสนอโฆษณาให้ตรงกับกลุ่มเป้าหมายให้ได้มากที่สุด เพื่อประโยชน์สูงสุดในทางธุรกิจของผู้ประกอบการที่เกี่ยวข้องทั้งหมด ตอบโจทย์ความต้องการของนักการตลาดที่มีกลุ่มเป้าหมายเป็นคนรุ่นใหม่กับจอแอลอีดีขนาดใหญ่ ทำให้บริษัทผู้พัฒนาสื่อโฆษณาสามารถสร้างสรรค์งานโฆษณาใหม่ ๆ ได้แบบไร้ขีดจำกัดอีกด้วย นอกจากนี้ในปัจจุบันยังพบปัญหาสำคัญว่าข่าวสารประชาสัมพันธ์ที่ทั้งจากหน่วยงานภาครัฐหรือเอกชนต้องการเผยแพร่นั้นมีการตอบรับเป็นอย่างไรบ้าง ซึ่งทำให้หน่วยงานไม่สามารถดำเนินการหรือจัดกิจกรรมในเชิงรุกได้ การพัฒนาจอโฆษณาอัจฉริยะจึงเป็นหนึ่งในระบบนิเวศน์ของเมืองอัจฉริยะ ที่สามารถตอบโจทย์ทางด้านธุรกิจ สังคม และชุมชนได้อย่างมีประสิทธิภาพ โดยบทความวิจัยฉบับนี้จะนำเสนอการประยุกต์ใช้ระบบปัญญาประดิษฐ์ในด้านการตรวจจับและรู้จำภาพใบหน้า (Facial Detection and Recognition) จากกล้องบันทึกภาพดิจิทัลและระบบสมองกลฝังตัว โดยใช้วิธีการเรียนรู้ของเครื่อง (Machine Learning) ด้วยเทคนิคการเรียนรู้เชิงลึก (Deep learning Technique) เพื่อพัฒนาจอโฆษณาแอลอีดีให้เป็นจอโฆษณาอัจฉริยะ ที่สามารถตรวจสอบจำนวนคนที่เห็นโฆษณา จำนวนคนที่ให้ความสนใจที่สามารถระบุเพศ และช่วงอายุได้ เพื่อนำข้อมูลเหล่านี้มาวิเคราะห์และประมวลผลทำให้จอโฆษณาสามารถโต้ตอบกับสภาพแวดล้อมและสามารถนำเสนอสื่อประชาสัมพันธ์ได้ตรงกับกลุ่มประชากรตามช่วงเวลาที่เหมาะสม โดยการทดสอบจะใช้วิธีการวัดประสิทธิภาพในด้านของความถูกต้องในการตรวจจับและรู้จำภาพใบหน้าจากภาพวิดีโอที่ได้จากกล้องติดไว้ที่บริเวณป้ายโฆษณาและการทดสอบประสิทธิภาพในด้านความเร็วในการประมวลผลบนอุปกรณ์ระบบสมองกลฝังตัวอีกด้วย

เนื้อหาในส่วนที่เหลือในบทความนี้ได้แก่ บทที่ 2 ทบทวนวรรณกรรม บทที่ 3 นำเสนอการพัฒนาจอโฆษณาอัจฉริยะ บทที่ 4 ผลลัพธ์ และบทสุดท้ายจะเป็นการสรุปเนื้อหาทั้งหมด

## 2. ทบทวนวรรณกรรมที่เกี่ยวข้อง

Christine Bauer และคณะ [1] นำเสนอแนวทางในการพัฒนาป้ายโฆษณาดิจิทัลที่ตอบโจทย์ด้านการตลาดและการเพิ่มยอดขายสินค้าโดยเน้นการสร้างเนื้อหาที่มีผลต่ออารมณ์ สร้างความรู้สึก และความพึงพอใจในการเลือกซื้อสินค้า และนำเสนอข้อมูลที่มีอิทธิพลต่อผู้บริโภค มีความสัมพันธ์กับพฤติกรรมของผู้บริโภค งานวิจัยนี้ได้ตั้งประเด็นคำถามวิจัย 2 ประเด็น ได้แก่ (1) มีความต้องการหลักใดบ้าง ที่จะนำมาสร้างเนื้อหาที่จะนำเสนอในป้ายโฆษณาดิจิทัลภายใต้วัตถุประสงค์เรื่อง “การเพิ่มยอดขายสินค้าขายปลีก” และ (2) จะนำข้อมูล

พฤติกรรมการเลือกซื้อสินค้า ทัศนคติ และความพึงพอใจของผู้บริโภค มาเป็นส่วนหลักในการสร้างเนื้อหาข้อมูลอย่างไร และหาว่ามีผลกระทบต่อการเลือกซื้อสินค้ามากน้อยเพียงใด Swapnil Alase และ Vaibhavi Chinchur [2] นำ raspberry pi 3 มาใช้แทนระบบไมโครคอนโทรลเลอร์กับป้ายดิจิทัล ผลลัพธ์แสดงให้เห็นถึงประสิทธิภาพในการใช้งานและสามารถลดต้นทุนด้านการใช้พลังงานไฟฟ้าได้มาก ผู้วิจัยได้ออกแบบและพัฒนาป้ายโฆษณาที่สามารถแสดงข้อมูล เช่น สภาพการจราจร สภาพภูมิอากาศ ข่าวสารต่าง ๆ แบบ real-time และสามารถควบคุมได้จากระยะไกล ระบบใช้เว็บเซิร์ฟเวอร์ควบคุมทั้งในส่วนฟร้อนเอ็น (Front-End) และแบคเอ็น (Back-End) ส่วนแสดงผลถูกเชื่อมต่อเข้ากับบอร์ด raspberry pi 3 ผ่านพอร์ต HDMI ตัวบอร์ดเชื่อมต่อกับเร้าเตอร์เพื่อเชื่อมต่ออินเทอร์เน็ตและสามารถควบคุมได้จากระยะไกล Mohammad Arif Hossain และคณะ [3] นำเสนองานวิจัยเรื่องการวิเคราะห์ประสิทธิภาพของระบบป้ายดิจิทัลอัจฉริยะ (Smart Digital Signage System) โดยใช้ซอฟต์แวร์ที่เรียกว่า Software-Defined IoT (SD-IoT) และใช้ image sensor เป็นเทคโนโลยีการสื่อสาร ผู้วิจัยนำเสนอสถาปัตยกรรมของระบบที่ประกอบด้วย 3 เลเยอร์ (layer) หลัก ได้แก่ 1) เลเยอร์กายภาพ (physical layer) 2) เลเยอร์ควบคุม (control layer) และ 3) เลเยอร์แอปพลิเคชัน (application layer) ที่ทำงานประสานกัน เลเยอร์กายภาพประกอบด้วย อุปกรณ์ต่าง ๆ เช่น โทรศัพท์มือถือ ป้าย กล้องวงจรปิด เป็นต้น เชื่อมต่อข้อมูลจากอุปกรณ์เข้ากับระบบคลาวด์ เลเยอร์ควบคุมประกอบด้วย SDN เกตเวย์ และ SDN คอนโทรลเลอร์ที่เชื่อมต่อกับระบบคลาวด์โดยใช้ API เลเยอร์แอปพลิเคชัน เชื่อมต่อกับ SDN คอนโทรลเลอร์โดยใช้ API เพื่อดึงข้อมูลมาแสดงผลบนแอพพลิเคชั่นต่าง ๆ ที่แสดงผลบนป้ายดิจิทัล Euiseok Nahm [4] นำเสนองานวิจัยเรื่อง การใช้แพลตฟอร์ม Mobile CPU สำหรับการสร้างระบบ Context-aware Digital Signage ซึ่งโดยทั่วไประบบป้ายดิจิทัลจะใช้เครื่องคอมพิวเตอร์พีซีที่ลงระบบปฏิบัติการวินโดว์สำหรับควบคุมการแสดงผล ผู้วิจัยใช้สมาร์ทโฟนในตระกูล multi-core processor ยึดติดกับตัวจอ ซึ่งทำให้ประหยัดพลังงานลดต้นทุน และสามารถควบคุมการแสดงผลโฆษณาบนจอได้จากระยะไกลได้ผ่านเว็บเบราเซอร์ Taner Arsan และคณะ [5] ออกแบบและพัฒนาระบบควบคุมจอดิจิทัลจากระยะไกลโดยใช้ระบบสมองกลฝังตัวที่สามารถกำหนดกฎเกณฑ์ และคุณสมบัติต่าง ๆ ได้ จอดิจิทัลสามารถแสดงผลโฆษณาและข้อมูลต่าง ๆ เช่น รถไฟใต้ดิน รถประจำทาง ห้างสรรพสินค้า ได้สามารถควบคุมการแสดงผลได้จากระยะไกล ระบบสมองกลฝังตัวถูกพัฒนาด้วย Raspberry Pi ส่วนฟร้อนเอ็นใช้ภาษา PHP ในการพัฒนา และควบคุมผ่านเว็บเซิร์ฟเวอร์ ส่วนแบคเอ็นพัฒนาด้วย HTML5 bootstrap ใช้ Github ในการพัฒนาส่วนรหัสโปรแกรมของส่วนควบคุมต่าง ๆ Robert Ravnik และ Franc Solina [6] ใช้เทคนิค real-time computer vision สำหรับการพัฒนาป้ายดิจิทัลให้มีความสามารถแสดงผลแบบ interactive โดยใช้กล้องติดตั้งอยู่ใน

เฟรมของตัวจอ ใช้อัลกอริทึม computer vision แยกข้อมูลของวัตถุในแบบ temporal, spatial และ demographic features และดึงข้อมูลเหล่านั้นแสดงผลบนจอภาพ ใช้อัลกอริทึม Viola and Jones face detection สำหรับตรวจจับและวิเคราะห์ใบหน้า ใช้อัลกอริทึม SVM classifiers สำหรับวิเคราะห์เพศและช่วงอายุ Tungkasthan และคณะ [7] นำเสนอระบบการค้นคืนรูปภาพที่มีอยู่ในอินเทอร์เน็ตโดยใช้คำสำคัญร่วมกับการใช้เนื้อหาของภาพ โดยคำสำคัญจะถูกนำมาใช้ในกระบวนการ การเรียกคืนภาพด้วยข้อมูลระดับสูงเชิงความหมาย จากนั้นภาพสอบถามจะถูกนำมาใช้ในกระบวนการค้นคืนภาพจากข้อมูลระดับต่ำของภาพเพื่อเปรียบเทียบหาภาพที่ตรงกับความต้องการของผู้ค้นหา เนื่องจากระบบมีการทำงานแบบเรียลไทม์ ดังนั้น เพื่อให้ระบบสามารถค้นค้นภาพได้อย่างรวดเร็วและประสิทธิภาพ ผู้วิจัยได้นำเสนอวิธีการสกัดคุณสมบัติสำคัญของภาพโดยใช้สี ที่มีความรวดเร็วและมีประสิทธิภาพ ชื่อว่าวิธีการแบบออร์โตคัลเลอร์คอรีโลแกรมและความสัมพันธ์ของสี (Auto Color Correlogram and Correlation: ACCC) สำหรับการสร้างดัชนีภาพด้วยข้อมูลระดับต่ำ จากผลการทดสอบโปรแกรมการสืบค้นคืนรูปภาพโดยใช้คำและเนื้อหาของภาพเป็นหลัก โดยใช้คำในหมวดหมู่ที่แตกต่างกันจำนวน 49 คำ บนฐานข้อมูลภาพในอินเทอร์เน็ตของ Yahoo พบว่าระบบที่พัฒนาขึ้นจากวิธีการที่นำเสนอสามารถค้นหารูปภาพได้ตรงกับความต้องการมากขึ้นเมื่อเทียบกับการใช้คำสำคัญเพียงอย่างเดียว Tungkasthan และคณะ [8] ได้นำเสนอวิธีการหาส่วนวัตถุสำคัญของภาพ ผู้วิจัยใช้ geometric active contours เพื่อปรับระดับค่าของฟังก์ชั่นให้เข้าสู่ขอบของวัตถุที่อยู่ในภาพ ใช้อัลกอริทึม mean-shift ลดระดับค่าความไวของการเปลี่ยนแปลงของตัวแปรในสมการ ผลจากการทดสอบการหาส่วนวัตถุสำคัญของภาพให้ค่า precision และ recall ที่ระดับ 92.77% และ 88.95% ตามลำดับ ในภาพที่มีองค์ประกอบไม่ซับซ้อน สำหรับภาพที่มีองค์ประกอบซับซ้อน วิธีการที่นำเสนอสามารถให้ค่า precision และ recall ที่ระดับ 88.93% และ 89.10% ตามลำดับ Tungkasthan [9] ได้นำเสนอเทคนิค color feature extraction สำหรับกระบวนการค้นคืนภาพ และเสนอวิธีการปรับปรุงอัลกอริทึม autocorrelogram ให้มีประสิทธิภาพดีขึ้น เพื่อนำมาค้นหาภาพในฐานข้อมูลขนาดใหญ่ ซึ่งใช้งานได้อย่างมีประสิทธิภาพ Jongsawat และคณะ [10], [11] ได้นำเสนอวิธีการสร้างโมเดลเบย์เซียนเน็ตเวิร์กและแผนภาพอิทธิพลบนเว็บแอพลิ-เคชั่น โดยนำเสนอวิธีการป้อนข้อมูลเข้าสู่โมเดลแบบอัตโนมัติ การปรับเปลี่ยนของโมเดลโดยใช้ Learning Algorithm และกระบวนการเตรียมข้อมูลก่อนการประมวลผล เว็บแอปพลิเคชันที่พัฒนาขึ้นมานั้นใช้เครื่องมือ SMILE (Structural Modeling, Inference, and Learning Engine) สำหรับการประมวลผล และ ใช้ SMILE.NET Wrapper เป็นเครื่องมือเชื่อมต่อเพื่อรับ-ส่งข้อมูลกับ SMILE โมเดลและวิธีการที่นำเสนอนี้สามารถประยุกต์ใช้กับการวิเคราะห์ข้อมูลและความสัมพันธ์ของข้อมูลโดยใช้เทคนิค decision making และสามารถทำนายผลลัพธ์

ข้อมูลได้ ซึ่งสามารถประยุกต์ใช้ได้ทั้งในแพลตฟอร์มที่เป็น stand alone หรือ online ก็ได้ Jongsawat และคณะ [12] ได้พัฒนาโมเดลเบย์เซียนเน็ตเวิร์กที่สามารถดึงข้อมูลจากแหล่งข้อมูลออนไลน์ ผ่านกระบวนการ pre-processing และนำข้อมูลเข้าสู่โมเดลเพื่อทำการคำนวณตามหลักการและทฤษฎีของเบย์ และสามารถกำหนดระดับความเชื่อมั่นและคำนวณผลลัพธ์ของโมเดลใหม่เพื่อดูระดับของผลกระทบที่เกิดขึ้น (inference) และพัฒนาส่วนติดต่อในรูปแบบเว็บ ผลลัพธ์ข้อมูลจะแสดงผลผ่านเว็บในรูปแบบข้อมูลภาพ (visualized data) จากการทบทวนวรรณกรรมที่เกี่ยวข้องกับการพัฒนาอัลกอริทึมที่จะนำมาใช้ในการพัฒนาจอโฆษณาอัจฉริยะนั้น ผู้วิจัยในโครงการนี้มีอัลกอริทึมที่ใช้กับกระบวนการประมวลผลภาพที่ได้จากกล้องหรือภาพถ่าย ที่พัฒนาขึ้นมาเองและผ่านการทดสอบใช้งานกับฐานข้อมูลรูปภาพขนาดใหญ่แล้ว ซึ่งให้ผลลัพธ์ความถูกต้องอยู่ในเกณฑ์สูง ระดับที่แข่งขันกับอัลกอริทึมอื่น ๆ ได้ สามารถพัฒนาอัลกอริทึมให้มีความสามารถที่ดีมากยิ่งขึ้น มีการพัฒนาระบบในส่วนของฟร้อนเอ็นและแบคเอ็นในรูปแบบเว็บแอปพลิเคชันที่ติดตั้งบนเว็บเซิร์ฟเวอร์ และมีอัลกอริทึมที่ใช้ในกระบวนการ decision making ที่พัฒนาขึ้นมาเองและใช้เป็นเครื่องมือสำคัญในการประมวลผลข้อมูล เครื่องมือเหล่านี้สามารถนำไปใช้งานในระบบสมองกลฝังตัว ทำงานร่วมกับโมดูลอื่น ๆ ของจอเพื่อพัฒนาให้เป็นจอโฆษณาอัจฉริยะ ตามกรอบการวิจัยและพัฒนาระบบที่ได้นำเสนอ

## 3. นำเสนอการตรวจจับและจดจำใบหน้าหรับจอโฆษณาอัจฉริยะ

### 3.1 พื้นฐานทฤษฎี

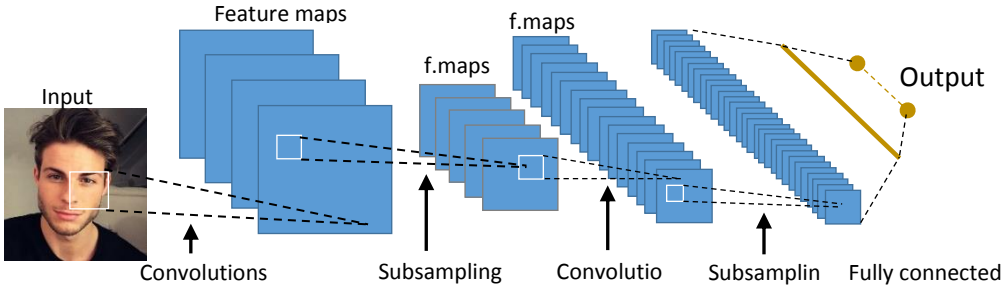
การตรวจจับและการรู้จำใบหน้าจากภาพดิจิทัลนั้นสามารถทำได้หลากหลายวิธีขึ้นอยู่กับสภาพแวดล้อม องค์ประกอบของภาพ และทรัพยากรที่ใช้ประมวลผล รวมถึงความเหมาะสมกับโจทย์ปัญหาวิจัย ทั้งนี้กระบวนการพื้นฐานโดยทั่วไปของการจดจำใบหน้าจากภาพถ่ายดิจิทัลมีกระบวนการหลักที่เกี่ยวข้องทั้งหมด 4 กระบวนการดังต่อไปนี้ [13]

1) การตรวจจับใบหน้า คือค้นหาใบหน้าตั้งแต่หนึ่งใบหน้าขึ้นไปภายในภาพและทำเครื่องหมายด้วย bounding box

2) การจัดตำแหน่งใบหน้า คือการทำให้ภาพใบหน้าอยู่ในมุมมองปกติและสอดคล้องกับฐานข้อมูลภาพใบหน้า

3) การแยกคุณลักษณะ คือการแยกคุณลักษณะ (features) จากใบหน้า และสามารถนำไปใช้สำหรับงานรู้จำภาพใบหน้าด้วยวิธีการที่แตกต่างกัน

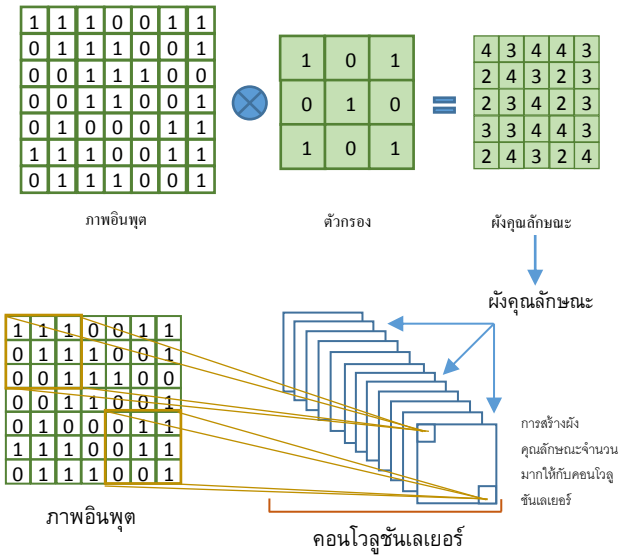
4) การจดจำใบหน้า ทำการจับคู่ระหว่างใบหน้าอินพุตกับใบหน้าที่รู้จักอย่างน้อยหนึ่งใบหน้าในฐานข้อมูลที่เตรียมไว้



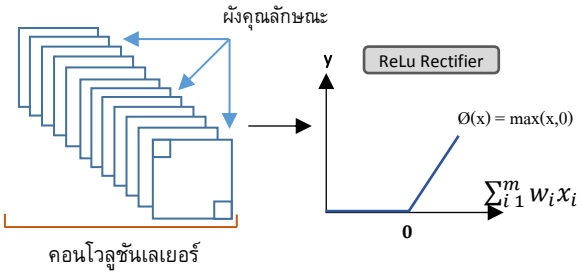

รูปที่ 1 สถาปัตยกรรมของทั่วไปของ Convolutional Neural Network

อย่างไรก็ตามหลังจากที่ Joseph Redmon และคณะ [14] ได้นำเสนอบทความการตรวจจับและรู้จำวัตถุภายในภาพแบบเรียลไทม์ในปี 2016 ก็ได้รับความสนใจอย่างมากจากนักวิจัยด้านการมองเห็นด้วยคอมพิวเตอร์ (computer vision) เพราะเป็นวิธีการที่ให้ผลลัพธ์ที่ดีทั้งด้านความถูกต้องและมีประสิทธิภาพด้านความเร็ว ยิ่งไปกว่านั้นยังเผยแพร่แบบจำลอง (model) การตรวจจับและรู้จำวัตถุภายในภาพแบบไม่มีค่าใช้จ่ายในชื่อว่า YOLO [14] ทั้งนี้ YOLO ใช้วิธีการตรวจจับวัตถุด้วยการใช้โครงข่ายประสาทเทียมแบบเรียนรู้เชิงลึก และใช้แบบจำลองโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network :CNN) แสดงดังรูปที่ 1 รายละเอียดในแต่ละขั้นตอนของการทำแบบจำลองโครงข่ายประสาทเทียมแบบคอนโวลูชันตามรูปที่ 1 มีดังต่อไปนี้

ขั้นตอนที่ 1 การทำคอนโวลูชัน (convolution) เป็นการคูณเมทริกซ์ระหว่างข้อมูลภาพอินพุตกับข้อมูลภาพตัวกรอง (feature detector) หรือตัวกรอง (filter) หรือบางครั้งถูกเรียกว่าเคอร์เนล (kernel) ซึ่งผลลัพธ์ทำให้ได้ภาพใหม่ที่เป็นข้อมูลลักษณะเด่นของภาพอินพุตที่ผ่านการกรองหรือเรียกว่าผังคุณลักษณะ (feature map) และเมื่อทำการคูณเมทริกซ์กับตัวกรองหลากหลายรูปแบบแล้วจะได้ผังคุณลักษณะจำนวนมากโดยจะเรียกทั้งหมดว่าคอนโวลูชันเลเยอร์ (convolutional layer) แสดงดังรูปที่ 2



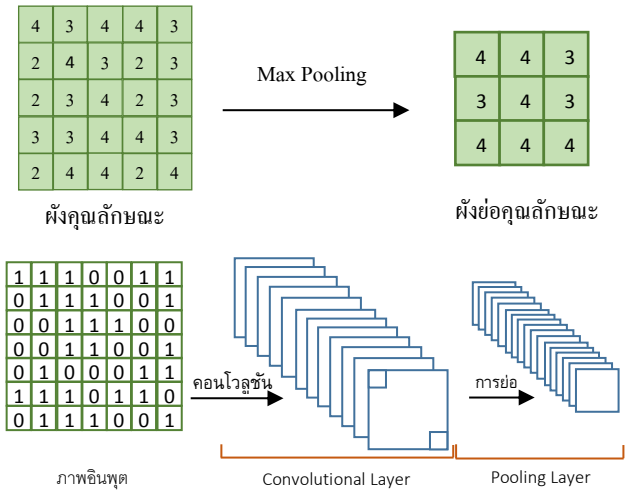

รูปที่ 2 ตัวอย่างการทำคอนโวลูชันด้วยวิธีการคูณเมทริกซ์และผังคุณลักษณะทั้งหมดในคอนโวลูชันเลเยอร์




รูปที่ 3 หลักการลดความซับซ้อนของข้อมูลในผังคุณลักษณะเด่นด้วยฟังก์ชั่น ReLu

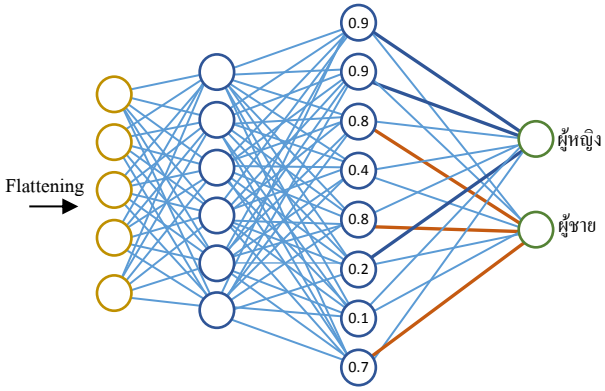
หลังจากได้คอนโวลูชันเลเยอร์แล้วจะทำการคำนวณค่าในผังคุณลักษณะใหม่เพื่อลดความซับซ้อนของข้อมูลทำให้สามารถฝึกโมเดลได้อย่างรวดเร็ว เรียกว่าการทำเร็คติไฟเออร์ (rectifier) ผ่านฟังก์ชัน activation ที่เรียบง่ายแต่มีประสิทธิภาพได้แก่ฟังก์ชัน ReLU โดยฟังก์ชัน ReLU มีคุณสมบัติสำคัญคือ ถ้าอินพุตเป็นบวกค่าความชัน (slope) จะเท่ากับ 1 เสมอ ทำให้ gradient ไม่หาย และไม่เกิด vanishing gradient โดยหลักการทำงานของเร็คติไฟเออร์แสดงได้ดังรูปที่ 3

ขั้นตอนที่ 2 ย่อผังคุณลักษณะ (Pooling) เป็นการลดปริมาณข้อมูลในการประมวลผลแต่พยามยามคงคุณลักษณะเด่นของผังคุณลักษณะจากขั้นตอนที่ 1 ไว้ ซึ่งปัจจุบันมีสองวิธีการหลักที่นิยมใช้คือ max pooling และ mean pooling อีกทั้งยังสามารถกำหนดขนาดตัวกรองได้หลายขนาดเพราะขนาดรูปที่ใส่เข้ามาทางอินพุตมีขนาดที่หลากหลาย ตัวอย่างการทำ max pooling แสดงดังรูปที่ 4




รูปที่ 4 ตัวอย่างการทำ max pooling

ขั้นตอนที่ 3 การเชื่อมต่อข้อมูลกับแบบจำลอง (Full Connection) ก่อนที่จะทำการเชื่อมต่อข้อมูลจากผังคุณลักษณะเข้ากับแบบจำลองจะต้องทำการแปลงผังคุณลักษณะ (flattening) ที่ย่อแล้วให้ข้อมูลอยู่ในรูปแบบเวคเตอร์มิติเดียวหรือทำให้เป็นคอลัมน์เดียวเพื่อความสะดวกรวดเร็วในการวิเคราะห์ข้อมูล ในขั้นตอนของการฝึก (train) การทำ fully connected เพื่อให้ได้โมเดลของการเรียนรู้เชิงลึกที่มีค่าน้ำหนักของแต่ละโหนดในแต่ละเลเยอร์ ทำให้สามารถแยกป้ายชื่อ (label) ที่เอาต์พุตเลเยอร์ได้แสดงได้ดังรูปที่ 5 ส่วนในขั้นตอนของการทดสอบหรือใช้งานจริงนั้นจะเป็นการใส่ค่าอินพุตในแต่ละโหนด โดยที่เอาต์พุตจะแสดงผลลัพธ์เป็นค่าน้ำหนักของแต่ละป้ายชื่อเพื่อใช้ในการหาคำตอบ

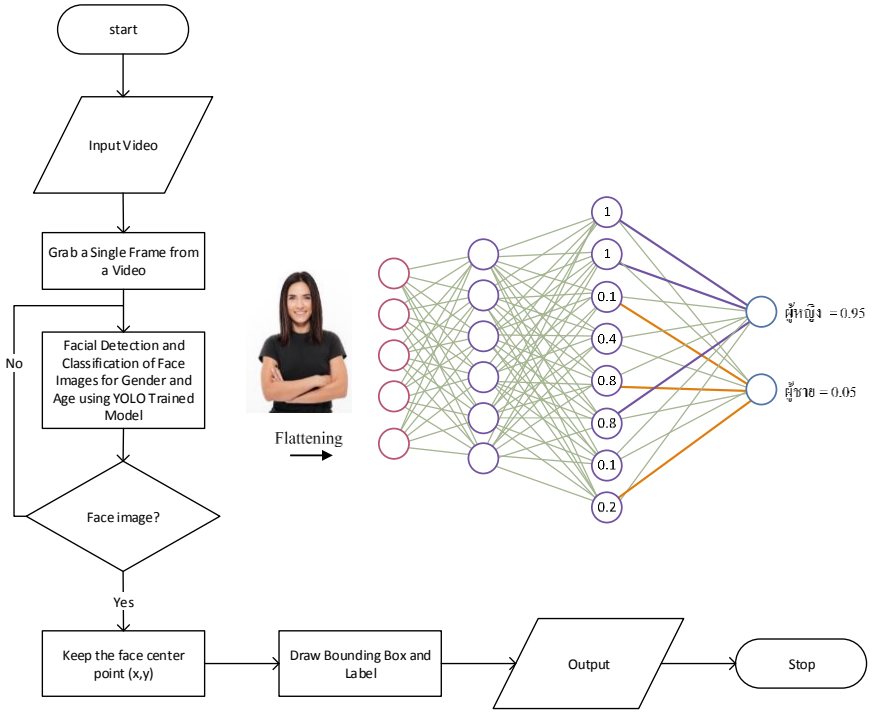



รูปที่ 5 การเชื่อมต่อข้อมูลคุณลักษณะเข้ากับโครงข่ายในขั้นตอนการฝึกเพื่อสร้างแบบจำลอง

### 3.2 การตรวจจับและการจำแนกใบหน้า

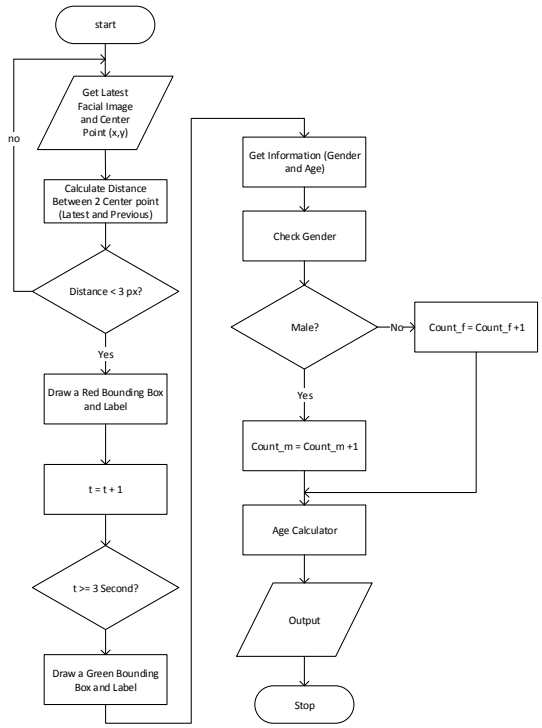
ในงานวิจัยนี้ไม่ได้ใช้แบบจำลองการตรวจจับวัตถุภายในภาพของ YOLO แต่พัฒนาขึ้นเองโดยใช้เครื่องมือการเรียนรู้ของเครื่องจาก library ของ OpenCV ทำให้เราสามารถพัฒนาแบบจำลอง

โครงข่ายประสาทเชิงลึกแบบคอนโวลูชันได้ง่ายและมีประสิทธิภาพ และทำให้มีความยืดหยุ่นในการใช้งานมากกว่า รวมถึงมีประสิทธิภาพด้านการประมวลผลที่เร็วกว่าการใช้ YOLO โดยรูปที่ 6 แสดงขั้นตอนการตรวจจับและจำแนกใบหน้าเพื่อนับจำนวนใบหน้าที่สามารถแยกเพศและช่วงอายุได้โดยใช้ฟังก์ชันใน library ของ OpenCV




รูปที่ 6 ผังไหลแสดงการทำงานของการตรวจจับและจำแนกใบหน้าด้วยโครงข่ายประสาทเทียมการเรียนรู้เชิงลึกแบบคอนโวลูชัน

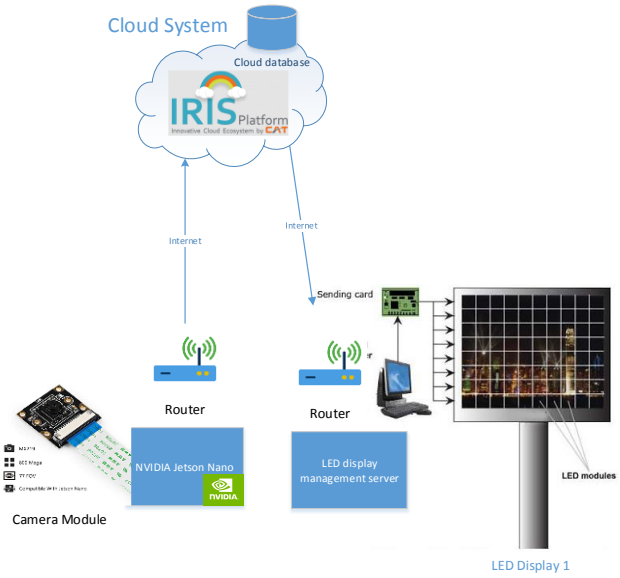
จากรูปที่ 6 เมื่อรับภาพวิดีโออินพุตเขามาจะทำการดึงภาพจากไฟล์วิดีโอจำนวน 1 เฟรม แล้วนำเข้าไปทดสอบยังแบบจำลองที่สร้างไว้แล้วเพื่อค้นหาภาพใบหน้า เมื่อพบภาพใบหน้าผลลัพธ์ที่ได้จากการทดสอบได้แก่ ภาพใบหน้า ข้อมูลเพศ และช่วงอายุ จากนั้นทำการวาดกรอบรูปสี่เหลี่ยมเฉพาะบริเวณภาพใบหน้าเพื่อคำนวณหาตำแหน่งศูนย์กลางของภาพใบหน้าและเก็บไว้ใช้ในกระบวนการตรวจสอบในขั้นถัดไป เมื่อได้ข้อมูลที่ผ่านกระบวนการในรูปที่ 7 แล้วกระบวนการถัดไปคือการตรวจสอบว่าภาพใบหน้าที่ตรวจจับได้นั้นเป็นภาพใบหน้าที่กำลังดูป้ายโฆษณาหรือไม่ โดยสร้างกฎในการตรวจสอบคือตำแหน่งของภาพใบหน้าตรงล่าสุดจะต้องไม่ต่างจากเทรสโฮลด์ (threshold) ที่กำหนดไว้ (threshold <= 3px โดยรอบรัศมี) เป็นเวลาอย่างน้อย 3 วินาที จึงถือว่าเป็นบุคคลที่สนใจและมองดูที่ป้ายโฆษณาอัจฉริยะ กระบวนการทำงานแสดงได้ดังรูปที่ 7




รูปที่ 7 ผังไหลการตรวจจับภาพใบหน้าที่ดูป้ายโฆษณาและการนับจำนวนแยกตามประเภทของใบหน้า

### 3.3 โครงสร้างและสถาปัตยกรรมระบบ

ในการตรวจจับใบหน้าบุคคลที่เดินผ่านบริเวณที่สนใจ ข้อมูลทั้งหมดที่ประมวลผลแล้วจะถูกส่งไปเก็บบนระบบคลาวด์เพื่อทำการวิเคราะห์ข้อมูลและส่งผลลัพธ์ไปยังสถานที่ติดตั้งป้ายโฆษณาในแต่ละจุดผ่านระบบอินเทอร์เน็ตเพื่อปรับเปลี่ยนโฆษณาให้เหมาะสมกับกลุ่มประชากร มีขั้นตอนเริ่มจากโมดูลกล้องที่ถูกเชื่อมต่อกับบอร์ดประมวลผลความเร็วสูง NVIDIA Jetson Nano จะทำการบันทึกข้อมูลวิดีโออินพุต ข้อมูลภาพวิดีโอจะถูกส่งไปยังส่วนประมวลผลของบอร์ด Jetson Nano ซึ่งจะทำการประมวลผลโดยใช้วิธีการเรียนรู้ของเครื่องจักร (Machine Learning) ด้วยเทคนิคการเรียนรู้เชิงลึก ตามที่อธิบายในข้อ 3.1 เพื่อตรวจจับใบหน้าตามแบบจำลองที่พัฒนาขึ้นโดยใช้เครื่องมือ OpenCV เป็นเครื่องมือหลักในการพัฒนาเพราะมีฟังก์ชันเกี่ยวกับการเรียนรู้ของเครื่องจักร โดยข้อมูลที่ถูกส่งขึ้นไปบนระบบคลาวด์จะเป็นข้อมูลที่ประมวลผลแล้ว ได้แก่ ภาพใบหน้าที่ตรวจจับได้ เวลา จำนวน ข้อมูลบุคคล เช่น เพศ อายุ เป็นต้น ทั้งนี้สถาปัตยกรรมโดยทั่วไปของระบบแสดงได้ดังรูปที่ 8




รูปที่ 8 สถาปัตยกรรมโดยทั่วไปของระบบ

## 4. ผลลัพธ์

### 4.1 ผลการดำเนินงาน

ทีมวิจัยและพัฒนาได้ทำการทดสอบการทำงานของระบบและอุปกรณ์ในห้องปฏิบัติการ แสดงดัง รูปที่ 9 โดยภาพทางด้านซ้ายสุดจะแสดงกรอบสีขาว (Unknown) บนใบหน้าคนแสดงถึง ระบบสามารถตรวจจับใบหน้าได้ในเฟรมภาพ แต่ตรวจสอบได้ว่าคนนั้นไม่ได้มองมาที่กล้อง กระบวนการประมวลผลจึงยังไม่เข้ากระบวนการตรวจสอบและบันทึกเวลา ภาพกลางแสดงกรอบสีแดง (Detecting) บนใบหน้าคนแสดงถึงระบบตรวจสอบได้ว่าคนนั้นมองมาที่กล้อง

หรือภาพโฆษณา กระบวนการประมวลผลจะเริ่มขึ้นและทำการบันทึกเวลา เพื่อหาว่าคนนั้นสนใจภาพโฆษณาอยู่หรือไม่ ภาพทางด้านขวาสุดจะแสดงกรอบสีเขียว (Logged) บนใบหน้าคนแสดงถึง ระบบสามารถตรวจจับใบหน้าได้ในเฟรมภาพ ซึ่งตรวจสอบได้ว่าคนนั้นได้มองมาที่กล้องหรือภาพโฆษณาบนจอแอลอีดีเกินกว่าค่าระยะเวลาที่กำหนด (Time configuration) ระบบจะบันทึกเวลาและข้อมูลต่าง ๆ ที่เกี่ยวข้องทั้งหมด

รูปที่ 9 การทดสอบการตรวจจับภาพใบหน้าคนในห้องปฏิบัติการจากโมดูลกล้องที่ถูกเชื่อมต่อกับบอร์ดประมวลผลความเร็วสูง NVIDIA Jetson Nano

รูปที่ 10 แสดงผลการตรวจจับภาพใบหน้าคนแบบ real-time จากภาพวีดีโอที่บันทึกได้จากสภาพแวดล้อมจริง โปรแกรมจะตรวจจับใบหน้าบุคคลที่อยู่ในเฟรมภาพ เมื่อคนเดินเข้ามาอยู่ในระยะที่สามารถตรวจจับได้ โดย “กรอบสีขาว” แสดงส่วนที่คาดว่าจะเป็นภาพใบหน้าคน และ “กรอบสีแดง” แสดงส่วนที่ยืนยันว่าเป็นภาพใบหน้าคนและได้มองมาที่กล้องหรือภาพโฆษณา

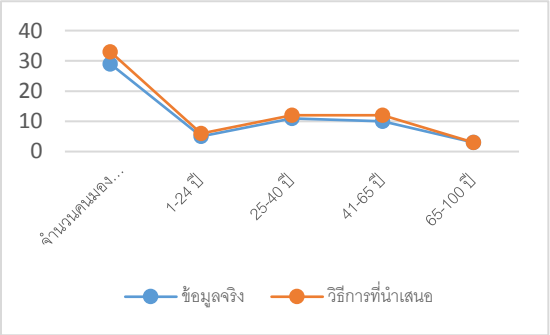
รูปที่ 10 ผลลัพธ์การตรวจจับภาพใบหน้าคนที่มองภาพโฆษณา

เมื่อทำการประมวลผลแล้วโปรแกรมจะบันทึกข้อมูล เช่น เพศชาย-หญิง ช่วงอายุ จำนวนคนที่พบ จำนวนคนที่สนใจภาพโฆษณาบนจอโฆษณา ผ่านระบบเครือข่ายเพื่อจัดเก็บข้องมูลลงในเครื่องคอมพิวเตอร์เซิร์ฟเวอร์ นอกจากนั้นข้อมูลที่ detect ได้ยังแสดงผลบนภาพแบบ real-time แสดงดังรูปที่ 11

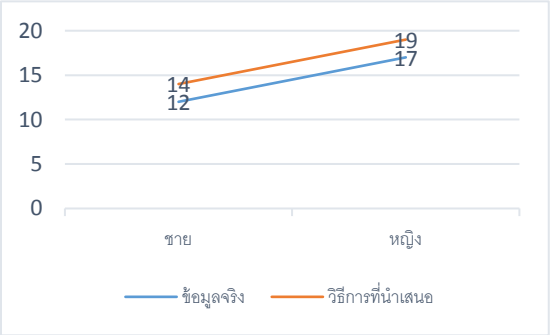
รูปที่ 11 ข้อมูลที่ตรวจจับได้บนภาพแสดงแบบ real-time

### 4.2 การทดสอบประสิทธิภาพ

ในการทดสอบประสิทธิภาพในงานวิจัยนี้จะมุ่งเน้นที่การทดสอบด้านความถูกต้องในการตรวจจับและนับจำนวนใบหน้าที่มองกล้องรวมถึงความสามารถในการแยกเพศและช่วงอายุ โดยทำการเปรียบเทียบกับความถูกต้องโดยคนนับและจำแนกภาพใบหน้า ในการทดสอบวิดีโอความยาว1.34 นาที จะถูกเล่นและใช้คนตรวจสอบแล้วบันทึกค่าเก็บไว้ จากนั้นใช้ระบบประมวลผลบนบอร์ด NVIDIA Jetson Nano ทำการตรวจสอบภาพใบหน้าจากข้อมูลอินพุตวิดีโอ โดยข้อมูลที่บันทึกได้จะถูกเก็บลงฐานข้อมูลบนระบบคลาวด์




รูปที่ 12 แสดงผลการเปรียบเทียบประสิทธิภาพจำแนกตามช่วงอายุ




รูปที่ 13 แสดงผลการเปรียบเทียบประสิทธิภาพจำแนกตามช่วงอายุ

จากนั้นข้อมูลทั้งสองจะถูกนำมาเปรียบเทียบเพื่อหาประสิทธิภาพ โดยผลลัพธ์ของการเปรียบแสดงดังรูปที่ 12 และรูปที่ 13

## 5. สรุปผล

งานวิจัยนี้ ผู้วิจัยได้พัฒนาจอโฆษณาแอลอีดีให้มีความสามารถนับจำนวนคนในพื้นที่ แยกเพศชาย-หญิง ประมาณการช่วงอายุ และตรวจจับคนที่สนใจโฆษณาในช่วงเวลาต่าง ๆ โดยใช้ระบบปัญญาประดิษฐ์และการเรียนรู้ของเครื่อง ด้วยโครงข่ายประสาทเทียม พัฒนาวิธีการตรวจจับวัตถุด้วยการใช้โครงข่ายประสาทเทียมแบบเรียนรู้เชิงลึกร่วมกับระบบสมองกลฝังตัว และใช้แบบจำลองโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network : CNN) โดยพัฒนาวงจรควบคุมที่ทำงานร่วมกับกล้องและติดตั้งอยู่กับจอแอลอีดี เพื่อบันทึกข้อมูลในพื้นที่และส่งข้อมูลขึ้นประมวลผลบนระบบคลาวด์ผ่านเครือข่ายอินเทอร์เน็ต โมดูลกล้องจะเชื่อมต่อกับบอร์ดประมวลผลความเร็วสูง NVIDIA Jetson Nano ซึ่งทำการบันทึกข้อมูลวิดีโออินพุตจากในพื้นที่ ข้อมูลภาพวิดีโอจะถูกส่งไปยังส่วนประมวลผลของบอร์ด Jetson Nano จากนั้นจะทำการประมวลผลโดยใช้วิธีการเรียนรู้ของเครื่อง เพื่อตรวจจับใบหน้าตามแบบจำลองที่พัฒนาขึ้นโดยใช้เครื่องมือ OpenCV โดยข้อมูลภาพใบหน้าที่ตรวจจับได้ เวลา จำนวนคน เพศ ช่วงอายุ จะถูกส่งขึ้นไปบนระบบคลาวด์ ผลจากการดำเนินงานสามารถเปลี่ยนจอโฆษณาแอลอีดีให้กลายเป็นจอโฆษณาอัจฉริยะได้

อย่างมีประสิทธิภาพ เพื่อเป็นโครงสร้างพื้นฐานที่จะรองรับการพัฒนาเมืองอัจฉริยะต่อไปในอนาคต

## กิตติกรรมประกาศ

คณะผู้วิจัยขอขอบคุณสำนักงานคณะกรรมการส่งเสริมวิทยาศาสตร์ วิจัยและนวัตกรรม (สกสว.) และบริษัท กสท. โทรคมนาคม จำกัด (มหาชน) ที่ได้ให้ทุนสนับสนุนการวิจัยและพัฒนาในโครงการการพัฒนาจอโฆษณาอัจฉริยะเพื่อตอบโจทย์ความสำเร็จเมืองอัจฉริยะ ในปีงบประมาณ พ.ศ. 2562 สำหรับการพัฒนาโครงสร้างพื้นฐานของเมืองอัจฉริยะ

## เอกสารอ้างอิง

[1] Christine Bauer et al. (2018), “Research Directions for Digital Signage Systems in Retail,” Procedia Computer Science, 141, 503–506.

[2] Swapnil Alase and Vaibhavi Chinchur (2017), “IoT Based Digital Signage Board using Raspberry PI 3,” International Research Journal of Engineering and Technology (IRJET), 4(5), 310–313.

[3] Mohammad Arif Hossain et al. (2016), “Performance analysis of smart digital signage system based on software defined IoT and invisible image sensor communication,” International Journal of Distributed Sensor Networks, 2(7), 1–14.

[4] Euiseok Nahm (2015), “Implementation of Context-aware Digital Signage System based on Mobile CPU Platform,” Indian Journal of Science and Technology, 8(14), 1–6.

[5] Taner Arsan, Alp Parkan and Hakkı Konu (2014), “Design and Implementation of Remotely Managed Embedded Digital Signage System,” International Journal of Computer Science, Engineering and Applications, 4(3), 1–15.

[6] Robert Ravnik และ Franc Solina (2013), “Interactive and Audience Adaptive Digital Signage Using Real-Time Computer Vision,” International Journal of Advanced Robotic Systems, 10, 1–7.

[7] Tungkasthan et al. (2010), “Micro level Attacks In Real-Time Image Processing for An On-Line CBIT Systems,” The 2010 International Conference of Computer Science and Engineering, London, United Kingdom, 30 June- 2 July, 2010.

[8] Tungkasthan et al. (2011), “Automatic Region of Interest detection in natural images,” Proceedings of the 15th WSEAS international conference on Computers, July 15 – 17, 437-444.

[9] Tungkasthan (2017), “Limitations of autocorrelogram technique in CBIR system,” Proceedings of the International Conference

on ICT and Knowledge Engineering, November 22 – 24, 37-45.

[10] Jongsawat et al., (2008), “Dynamic Data Feed to Bayesian Network Model and SMILE Web Application,” In Proceedings of IEEE on Software Engineering: Artificial Intelligence, Networking, and Parallel/Distributed Computing, Phuket, Bangkok, Thailand, 931-936.

[11] Jongsawat et al., (2009), “A SMILE Web-Based Interface for Learning the Causal Structure and Performing a Diagnosis of a Bayesian Network,” Proceedings of IEEE on Systems, Man, and Cybernetic, Systems Science & Engineering, San Antonio, Texas, USA.

[12] Jongsawat et al., (2010), “Automatically Building Diagnostic Bayesian Networks from On-line Data Sources and the SMILE Web-based Interface,” Decision Support Systems, Book edited by: Chiang S. Jao, ISBN 978-953-7619-64-0, 321-334.

[13] Thai Hoang Le, “Applying Artificial Neural for Face Recognition,” Hindawi , Advances in Artificial Neural Systems, Volume 2011, Article ID 673016, 16 pages doi:10.1155/2011/673016

[14] Joseph Redmon et al. (2016), "You Only Look Once: Unified, Real-Time Object Detection," Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA
doi: 10.1109/CVPR.2016.91