

การปรับปรุงประสิทธิภาพการตรวจจับใบหน้าภายใต้สภาวะแสงต่าง ๆ ด้วย YOLO และ Haar Cascade ร่วมกับ Canny Edge Detection

Enhancing Face Detection Performance under Various Illumination Conditions using YOLO and Haar Cascade with Canny Edge Detection

พิศณุ คุมีชัย¹ นภเดช อัดตเจริญวงศ์^{1*} จิตติ สัมภัตตะกุล¹ ธีรภัทร สาระวัน¹ กฤตณัย สิงห์เพือก¹
ศรัณย์ภัทร มะลิพรม¹ วิชชากร ชัยอรรวรรณ¹ และสืบพงษ์ วัฒนพฤกษาชาติ¹

Pisanu Kumeechai¹ Napadach Attachareanwong^{1*} Jitti Sampattakul¹ Teeraphat Sarawan¹
Krittanaai Singphueak¹ Saranphat Maliphrom¹ Witchakorn Chaorawan¹
and Sueppong Wattanaprukchat¹

¹กองวิชาวิศวกรรมศาสตร์ ฝ้ายศึกษา โรงเรียนนายเรือ

¹Department of Engineering, Education Branch, Royal Thai Naval Academy

*Corresponding author: napadach.att@hotmail.com

Received: October 30, 2025

Revised: December 19, 2025

Accepted: December 27, 2025

บทคัดย่อ

ปัญหาหลักของการตรวจจับใบหน้า คือ ความไม่เสถียรของประสิทธิภาพภายใต้สภาวะแสงที่แปรผัน โดยเฉพาะในสภาพแสงน้อย งานวิจัยนี้มีวัตถุประสงค์เพื่อ ปรับปรุงและเปรียบเทียบประสิทธิภาพ การตรวจจับใบหน้าภายใต้สภาวะแสงที่แตกต่าง (แสงมาก ปานกลาง และน้อย) โดยใช้การผสมผสานเทคนิค YOLO Haar Cascade และ Canny Edge Detection ซึ่ง Canny ถูกใช้เป็น Pre-processing เพื่อเน้นโครงร่างของใบหน้าและลดผลกระทบจากแสงก่อนเข้าโมเดล การประเมินผลใช้เมตริกมาตรฐาน (accuracy precision และ recall) จากชุดข้อมูล 1,500 ภาพ ผลการทดลองพบว่า วิธี YOLO ร่วมกับ Canny Edge Detection (YOLO+Canny) ให้ผลลัพธ์ที่เหนือกว่าทุกวิธี โดยมี Accuracy เฉลี่ยสูงสุด (92.0%) ในสภาวะแสงปานกลาง และยังคงประสิทธิภาพสูงในแสงมาก (91.0%) และแสงน้อย (87.2%) นอกจากนี้ YOLO+Canny ยังใช้เวลาประมวลผลต่ำที่สุด (<0.2 วินาทีต่อภาพ) สรุปได้ว่า การผสมผสาน YOLO และ Canny มีความทนทาน รวดเร็ว และแม่นยำสูงในการตรวจจับใบหน้าภายใต้สภาวะแสงที่แตกต่าง ซึ่งมีความสำคัญสูงในการประยุกต์ใช้ในงานจริง

คำสำคัญ: การตรวจจับใบหน้า YOLO Haar Cascade Canny Edge Detection สภาวะแสงแปรผัน

Abstract

The primary challenge in face detection is performance instability under varying illumination conditions, particularly in low-light environments. This research aims to enhance face detection accuracy under high, medium, and low lighting conditions by integrating YOLO, Haar Cascade, and Canny Edge Detection. Canny Edge Detection is employed as a pre-processing step to emphasize facial contours and mitigate illumination effects prior to model input. Performance evaluation was conducted using standard metrics, including Accuracy, Precision, and Recall, on a dataset of 1,500 images. Experimental results indicate that the integration of YOLO with Canny Edge Detection (YOLO + Canny) significantly outperformed the other methods, achieving the highest average Accuracy of (92.0%) under medium lighting conditions, while maintaining high performance under high light (91.0%) and low light (87.2%). Furthermore, the YOLO + Canny approach demonstrated the lowest processing time, at less than 0.2 seconds per image. In conclusion, the integration of YOLO and Canny Edge Detection provides high robustness, speed, and accuracy for face detection across varying illumination conditions, demonstrating strong potential for real-world applications.

Keywords: face detection, YOLO, Haar Cascade, Canny Edge Detection, varying illumination conditions



บทนำ

การตรวจจับใบหน้า (face detection) เป็นหนึ่งในหัวข้อที่มีความสำคัญและได้รับความสนใจอย่างต่อเนื่องในสาขาการประมวลผลภาพดิจิทัลและการเรียนรู้ของเครื่อง เนื่องจากการนำไปประยุกต์ใช้อย่างแพร่หลาย ไม่ว่าจะเป็นในงานด้านความปลอดภัย การเฝ้าระวัง การยืนยันตัวบุคคล หรือการพัฒนาาระบบปฏิสัมพันธ์ระหว่างมนุษย์และคอมพิวเตอร์ อย่างไรก็ตาม ความแม่นยำและความเสถียรของอัลกอริทึมตรวจจับใบหน้าที่ยังคงได้รับผลกระทบอย่างมีนัยสำคัญจากปัจจัยด้านสิ่งแวดล้อม โดยเฉพาะความแตกต่างของสภาวะแสง เช่น แสงมาก แสงปานกลาง และแสงน้อย ซึ่งอาจนำไปสู่ความคลาดเคลื่อน (false detection) การตรวจจับพลาด (miss) หรือความเชื่อมั่นต่ำของผลลัพธ์

งานวิจัยรุ่นแรก เช่น Haar Cascade Classifier ที่พัฒนาโดย Viola and Jones (2001) แสดงให้เห็นถึงประสิทธิภาพในการตรวจจับใบหน้าที่รวดเร็วและแม่นยำภายใต้สภาพแวดล้อมที่ควบคุมได้ แต่กลับมีข้อจำกัดเมื่อสภาวะแสงเปลี่ยนแปลงหรือมีความซับซ้อนมากขึ้น (Zhang & Zhang, 2010) ขณะที่การพัฒนาของโมเดลเชิงลึก (deep learning) เช่น YOLO: You Only Look Once สามารถตรวจจับวัตถุได้แบบเรียลไทม์และมีความทนทานต่อสภาพแวดล้อม

ที่ไม่แน่นอนมากกว่า (Redmon et al., 2016; Bochkovskiy et al., 2020)

เทคนิคการปรับปรุงคุณภาพภาพเชิงดั้งเดิมที่ใช้การปรับค่าความเข้ม (intensity-based) เช่น Histogram Equalization: HE และ Gamma Correction ถูกนำมาใช้บ่อยครั้งเพื่อจัดการกับภาพที่มีแสงไม่สม่ำเสมอ อย่างไรก็ตาม วิธีการเหล่านี้มีข้อจำกัดที่สำคัญ คือ มักจะนำไปสู่การขยายสัญญาณรบกวน (noise amplification) และการสร้างสิ่งแปลกปลอมเชิงภาพ (visual artifacts) ในทางตรงกันข้าม Canny Edge Detection ถูกเลือกใช้ในงานวิจัยนี้เนื่องจากสามารถ ลดสัญญาณรบกวนด้วย Gaussian Filter ก่อนการหาขอบ และให้ข้อมูลโครงสร้างที่บริสุทธิ์และแม่นยำสูง (high precision edges) ซึ่งเหมาะสมกว่าในการเป็นข้อมูลป้อนเข้าเสริมสำหรับโมเดลการเรียนรู้เชิงลึก นอกจากนี้ยังมีความพยายามปรับปรุง preprocessing เพื่อเพิ่มความเสถียรของโมเดล เช่น การใช้ Canny Edge Detection ซึ่งเป็นเทคนิคที่ทนทานต่อ noise และสามารถเน้นขอบเขตของวัตถุได้ดีกว่า Sobel หรือ Prewitt (Canny, 1986; Maini & Aggarwal, 2009) โดยการประมวลผลลักษณะขอบที่คมชัดช่วยให้การตรวจจับในขั้นตอนถัดไปมีประสิทธิภาพสูงขึ้น ซึ่งสอดคล้องกับ

แนวคิดที่ว่า preprocessing มีผลต่อความแม่นยำของระบบตรวจจับวัตถุ (Zareiforush et al., 2016)

ในช่วงไม่กี่ปีที่ผ่านมา มีงานวิจัยใหม่ ๆ ที่เสนอการพัฒนาโมเดลและเทคนิคเสริมเพื่อแก้ปัญหาสถานะแสงและการประมวลผลแบบ real-time ตัวอย่างเช่น Yisihak and Li (2024) ได้นำเสนอการบูรณาการ YOLOv8 เข้ากับโมเดล AI เพื่อตรวจจับใบหน้าอย่างมีประสิทธิภาพในระบบประยุกต์จริง Mittal (2024) ได้ทำการสำรวจโมเดลตรวจจับเชิงลึกแบบ lightweight ที่เหมาะสมกับ edge devices เพื่อให้รองรับการทำงานบนอุปกรณ์ที่มีข้อจำกัดด้านทรัพยากร และ Sun et al. (2022) ได้ทบทวนเทคนิค edge detection ทั้งแบบดั้งเดิมและที่ใช้ deep learning ซึ่งสามารถเชื่อมโยงกับ pipeline การตรวจจับใบหน้าได้โดยตรง

สำหรับการจัดการกับปัญหาสถานะแสงน้อย มีงานวิจัยที่ชี้ให้เห็นการพัฒนาโมเดลใหม่ที่ยืดหยุ่นมาเพื่อแก้ปัญหาโดยเฉพาะ เช่น Dark-YOLO ซึ่งบูรณาการกลไก attention หลายรูปแบบเพื่อเพิ่มประสิทธิภาพในการตรวจจับในภาพที่มีด (Liu et al., 2025) รวมถึงงานของ Hong et al. (2024) ที่เสนอการสร้างคุณลักษณะ (feature) ที่ไม่แปรผันตามความสว่าง (illumination-invariant) เพื่อเพิ่มความแม่นยำในการตรวจจับวัตถุในสภาพแสงต่ำ ทั้งหมดนี้สะท้อนถึงแนวโน้มงานวิจัยที่ให้ความสำคัญกับการออกแบบโมเดลที่สามารถทนต่อการเปลี่ยนแปลงของสถานะแสง

นอกจากโมเดลที่กล่าวถึงข้างต้นแล้ว ในกลุ่มงานวิจัยที่เน้นความแม่นยำสูงและความทนทานต่อท่าทาง (pose) และขนาดที่หลากหลายของใบหน้า ก็มีการใช้สถาปัตยกรรมที่ซับซ้อนขึ้น เช่น MTCNN: Multi-Task Cascaded Convolutional Networks ซึ่งเป็นวิธีการแบบหลายขั้นตอนที่ใช้ Convolutional Neural Networks: CNNs สามส่วนเรียงต่อกันเพื่อหาตำแหน่งใบหน้าอย่างละเอียด (Zhang et al., 2016) วิธีนี้แสดงให้เห็นถึงความแม่นยำที่ยอดเยียมในการตรวจจับใบหน้าที่มีความท้าทายสูง แม้ว่าการวิจัยนี้จะมุ่งเน้นการเปรียบเทียบระหว่างการเชิงดั้งเดิม (Haar Cascade) และโมเดลเชิงลึกแบบ single-shot (YOLO) ที่มีการเสริม preprocessing แต่การรับรู้ถึงความก้าวหน้าของ MTCNN ก็เป็นสิ่งสำคัญในการวางตำแหน่งงานวิจัยในบริบทที่กว้างขึ้น

สรุปได้ว่า ปัญหาหลักของการตรวจจับใบหน้ายังคงอยู่ที่ความไม่เสถียรของอัลกอริทึมเมื่อเผชิญกับสถานะแสงที่เปลี่ยนแปลง โดยเฉพาะในสถานะแสงน้อยที่ Haar

Cascade แสดงข้อจำกัดชัดเจน ในขณะที่ YOLO และโมเดล deep learning ที่มีการเสริม preprocessing หรือกลไก attention/invariant feature สามารถรับมือได้ดีกว่า ดังนั้นงานวิจัยนี้จึงมุ่งเน้นไปที่การศึกษาและเปรียบเทียบประสิทธิภาพของการตรวจจับใบหน้าภายใต้สถานะแสงมาก แสงปานกลาง และแสงน้อย โดยใช้การผสาน YOLO ร่วมกับ Canny Edge Detection และ Haar Cascade Classifier เพื่อประเมินแนวทางที่เหมาะสมที่สุดสำหรับการประยุกต์ใช้งานจริง

เหตุผลทางทฤษฎีในการผสาน Canny Edge Detection เข้ากับ YOLO เนื่องจาก Canny เป็นเทคนิคที่สามารถดึงข้อมูลโครงสร้าง (structural information) ที่มีความคงทนต่อแสง (illumination invariance) ได้อย่างมีประสิทธิภาพและรวดเร็ว แม้ว่าสถาปัตยกรรม CNN ของ YOLO จะสามารถสกัดคุณลักษณะเด่นได้ด้วยตนเอง แต่การป้อนข้อมูลภาพที่ถูกระบุเฉพาะขอบเขตของใบหน้าเข้าไปยังโมเดล CNN จะช่วยลดความแปรปรวนของข้อมูล (variance) ที่เกิดจากแสงเงาที่รุนแรง และช่วยให้โมเดลสามารถโฟกัสไปที่รูปร่าง (contours) ของใบหน้าโดยตรง ซึ่งเป็นการเพิ่มคุณภาพของคุณลักษณะป้อนเข้า และอาจนำไปสู่การอนุมานผลที่แม่นยำยิ่งขึ้นและมีประสิทธิภาพในการประมวลผลสูงขึ้น แนวคิดนี้สอดคล้องกับงานวิจัยที่ชี้ให้เห็นว่า preprocessing ที่เหมาะสมสามารถเพิ่มความแม่นยำของระบบตรวจจับวัตถุได้ ดังนั้นการผสานเทคนิคนี้จึงเป็นแนวทางที่น่าสนใจในการปรับปรุงความเสถียรของ YOLO ต่อสถานะแสงแปรผันในเชิงปฏิบัติการ

วัตถุประสงค์การวิจัย

งานวิจัยนี้มีวัตถุประสงค์หลัก ดังต่อไปนี้

1. เพื่อพัฒนาและปรับปรุง กระบวนการตรวจจับใบหน้าโดยการผสานเทคนิค Canny Edge Detection เข้ากับวิธีการตรวจจับด้วย YOLO: You Only Look Once และ Haar Cascade Classifier

2. เพื่อศึกษาและเปรียบเทียบประสิทธิภาพ ของการตรวจจับใบหน้าด้วยวิธีการพื้นฐานและวิธีการผสาน (YOLO Haar Cascade และการผสานกับ Canny Edge Detection) ภายใต้สถานะแสงที่แตกต่างกัน (แสงมาก แสงปานกลาง และแสงน้อย)

3. เพื่อประเมินผลและระบุ วิธีการที่มีประสิทธิภาพสูงสุดในการจัดการกับความแปรผันของสภาวะแสง โดยใช้เมตริกมาตรฐาน ได้แก่ Accuracy Precision Recall และ F1-Score

4. เพื่อยืนยันความเป็นไปได้ ของการนำวิธีการที่ให้ผลลัพธ์ที่ดีที่สุดไปประยุกต์ใช้ในสภาพแวดล้อมจริงที่มีความท้าทายด้านแสงและความแปรผันของสภาพแวดล้อม

วิธีดำเนินการวิจัย

เพื่อให้ได้มาซึ่งผลการทดลองที่มีความถูกต้องและสามารถสะท้อนถึงประสิทธิภาพของการตรวจจับใบหน้าในสภาวะแสงที่แตกต่างกัน งานวิจัยนี้ได้ออกแบบกระบวนการดำเนินการอย่างเป็นระบบ ดังนี้

สถาปัตยกรรมและแนวคิดการวิจัย

งานวิจัยนี้ได้รับการออกแบบโดยมีจุดมุ่งหมายเพื่อปรับปรุงและเปรียบเทียบประสิทธิภาพของการตรวจจับใบหน้าภายใต้สภาวะแสงที่แตกต่างกัน ได้แก่ แสงมาก (bright light) แสงปานกลาง (normal light) และแสงน้อย (low light) ซึ่งเป็นปัจจัยที่ส่งผลโดยตรงต่อความแม่นยำของอัลกอริทึมการประมวลผลภาพ (Zhou et al., 2017) เพื่อบรรลุเป้าหมายดังกล่าว จึงได้มีการผสมผสานเทคนิคการตรวจจับใบหน้าที่หลากหลาย ได้แก่ YOLO: You Only Look Once ซึ่งเป็นวิธีการตรวจจับวัตถุแบบเรียลไทม์ที่มีความแม่นยำสูง (Redmon et al., 2016) Haar Cascade Classifier ที่อาศัยการเรียนรู้เชิงลำดับชั้นของพีเจอร์ Haar-like (Viola & Jones, 2001) และ Canny Edge Detection ที่มีความสามารถโดดเด่นในการดึงคุณลักษณะขอบภาพเพื่อเน้นโครงสร้างของวัตถุ (Canny, 1986)

เทคนิคการปรับปรุงคุณภาพภาพสำหรับสภาวะแสงต่าง ๆ (image enhancement techniques for various illumination) ปัญหาการตรวจจับใบหน้าภายใต้สภาวะแสงที่ไม่เหมาะสมได้นำไปสู่การพัฒนาเทคนิคการปรับปรุงคุณภาพภาพหลายรูปแบบ ซึ่งสามารถแบ่งออกเป็นกลุ่มหลัก ๆ ได้ดังนี้ วิธีการปรับปรุงค่าความเข้ม (intensity-based) ตัวอย่างวิธีการ Histogram Equalization: HE และ CLAHE หลักการทำงาน ปรับปรุงการกระจายตัว

ของค่าความเข้มของพิกเซล เพื่อเพิ่มความแปรปรวนในภาพ ข้อดีและข้อจำกัดช่วยเพิ่มความแปรปรวนได้ดี แต่เสี่ยงต่อการเกิดสัญญาณรบกวน (noise) หรือการเพิ่มความสว่างมากเกินไปในพื้นที่ที่ไม่ต้องการ วิธีการอ้างอิงทฤษฎีการสะท้อนแสง (retinex-based) ตัวอย่างวิธีการ MSR และ SSR หลักการทำงาน แยกองค์ประกอบของการส่องสว่าง (illumination) ออกจากองค์ประกอบการสะท้อนแสง (reflectance) ข้อดีและข้อจำกัดให้ผลลัพธ์ที่คงทนต่อการเปลี่ยนแปลงของแสงได้ดี แต่มีขั้นตอนการคำนวณที่ซับซ้อนและใช้เวลาในการประมวลผลสูงกว่า และวิธีการเน้นโครงสร้าง (structure-based) ตัวอย่างวิธีการ Canny Edge Detection หลักการทำงานแปลงภาพให้เป็นข้อมูลขอบภาพ ซึ่งเป็นคุณสมบัติที่ค่อนข้างคงทนต่อการเปลี่ยนแปลงของแสง ข้อดีและข้อจำกัดโครงสร้างที่ได้มีความชัดเจนและแม่นยำสูง (high precision edges) ช่วยให้โมเดลตรวจจับใบหน้าได้แม่นยำขึ้นที่รูปร่างของวัตถุโดยตรง

เครื่องมือที่ใช้ในการวิจัย

สภาพแวดล้อมและเครื่องมือที่ใช้ในการทดลองงานวิจัยนี้ดำเนินการและทดสอบบนสภาพแวดล้อมการประมวลผล Python 3.12 โดยใช้เฟรมเวิร์กและไลบรารีหลัก ดังนี้

การประมวลผลภาพและการใช้ Haar Cascade ใช้ไลบรารี OpenCV (cv2) เวอร์ชัน 4.9.0.80 การจัดการโมเดลและการฝึกฝน YOLO ใช้เฟรมเวิร์ก PyTorch เวอร์ชัน 2.2.2 และชุดเครื่องมือ YOLOv8x

สภาพแวดล้อมฮาร์ดแวร์ (hardware environment) การฝึกฝนโมเดลและการทดสอบประสิทธิภาพดำเนินการบนคอมพิวเตอร์ส่วนบุคคล โดยใช้หน่วยประมวลผลกลาง (CPU) แบบ AMD Ryzen 7 4800H with Radeon Graphics และการ์ดประมวลผลกราฟิก (GPU) แบบ NVIDIA GeForce (ตามรุ่นที่ติดตั้งในเครื่อง) โดยไม่มีการใช้บริการ Cloud-based Training ในกระบวนการฝึกฝนโมเดลทั้งหมด

ภาพ 1 แสดงแผนผังขั้นตอนการทำงานโดยรวมของระบบที่ใช้ในการวิจัย ซึ่งเริ่มต้นจากการรับภาพสแตนด์บายและผ่านกระบวนการปรับปรุงภาพด้วย Canny Edge Detection ก่อนจะนำไปป้อนเข้าสู่โมเดลตรวจจับทั้งสองเพื่อทำการประเมินและเปรียบเทียบผลลัพธ์



ภาพ 1 แผนผังแสดงขั้นตอนการทำงานของระบบตรวจจับใบหน้าภายใต้สภาวะแสงต่าง ๆ

โดยสถาปัตยกรรมของงานวิจัยนี้จึงประกอบด้วยกระบวนการหลัก 4 ขั้นตอน ดังนี้

1. การรวบรวมและเตรียมข้อมูลภาพ

การประเมินผลการทดลองได้ดำเนินการบนชุดข้อมูลภาพใบหน้าจำนวน 1,500 ภาพ ภายใต้สภาวะแสง 3 ระดับ ได้แก่ แสงมาก (high light) 500 ภาพ แสงปานกลาง (medium light) 500 ภาพ และแสงน้อย (low light) 500 ภาพ ข้อมูลภาพที่ใช้ในการทดลองถูกคัดเลือกและปรับสมดุลจำนวนภาพในแต่ละกลุ่มอย่างรอบคอบ เพื่อ

ให้ครอบคลุมและเป็นตัวแทนของสภาวะแสงที่ต่างกันอย่างหนึ่ง ซึ่งเป็นการหลีกเลี่ยงปัญหาความเอนเอียง (bias) ของแบบจำลอง ตามแนวทางที่แนะนำโดย Krizhevsky et al. (2017)

เกณฑ์การแบ่งสภาวะแสงใช้ค่าเฉลี่ยความสว่างของพิกเซล (average pixel intensity level) เป็นเกณฑ์หลักในการจำแนก โดยอิงตามหลักการของการวิเคราะห์การแจกแจงความสว่างเพื่อประเมินระดับแสงในภาพ (Pizer et al., 1987; Jobson et al., 1997) ภาพถูกแปลงเป็นภาพระดับสีเทาเพื่อคำนวณค่าความเข้มของพิกเซลเฉลี่ยในช่วง 0-255 ซึ่งนิยามเชิงปริมาณจากการปรับจูนกับชุดข้อมูล (tuning) ได้ดังนี้ แสงน้อย (low light) มีค่าเฉลี่ยความสว่างอยู่ในช่วง 50-100 แสงปานกลาง (medium light) มีค่าเฉลี่ยความสว่างอยู่ในช่วง 100-180 แสงมาก (high light) มีค่าเฉลี่ยความสว่าง มากกว่า 180

เพื่อให้สามารถทดสอบความทนทานของอัลกอริทึมได้อย่างมีประสิทธิภาพ โดยในแต่ละสภาวะแสงได้ออกแบบสัดส่วนของภาพให้มีความสมดุลระหว่าง ภาพที่มีใบหน้า (positive samples) และภาพที่ไม่มีใบหน้า (negative samples) เพื่อให้สามารถคำนวณค่าประสิทธิภาพมาตรฐาน ได้แก่ Accuracy Precision Recall และ F1-Score ได้อย่างเหมาะสม โดยกำหนดให้สัดส่วนอยู่ที่ 60:40 (ใบหน้า: ไม่มีใบหน้า) อันสอดคล้องกับแนวทางของงานวิจัยด้านการตรวจจับใบหน้าที่นิยมใช้ dataset ที่มีใบหน้าที่มากกว่าเล็กน้อยเพื่อสะท้อนสภาพการใช้งานจริง (Redmon et al., 2016; Li & Deng, 2020; Yisihak & Li, 2024)

ในเชิงปริมาณแต่ละกลุ่มย่อยจึงประกอบด้วย 300 ภาพที่มีใบหน้า และ 200 ภาพที่ไม่มีใบหน้า รวมทั้งสิ้น 900 ภาพที่มีใบหน้า (60%) และ 600 ภาพที่ไม่มีใบหน้า (40%) สำหรับทั้ง dataset ขนาด 1,500 ภาพ การมี negative samples ในสัดส่วนที่เพียงพอมีความสำคัญต่อการประเมินค่าความผิดพลาด (false positive) และ True Negative ซึ่งสะท้อนความเสถียรของโมเดลในสภาวะแวดล้อมที่มีสิ่งรบกวนหรือ background ที่ไม่ใช่ใบหน้า (Viola & Jones, 2001; Zhang & Zhang, 2010; Mittal, 2024) แม้ว่าจำนวนรวม 1,500 ภาพ อาจถือว่าอยู่ในระดับกลางเมื่อเทียบกับชุดข้อมูล benchmark ขนาดใหญ่ (เช่น wider face) แต่ชุดข้อมูลนี้ถูกออกแบบและ

คัดเลือกว่าจะมัวหรือสว่างเกินไปให้เกิดการกระจายที่สมดุล และเป็นตัวแทนของสภาวะแสงที่แตกต่างกันถึง 3 ระดับ (500 ภาพต่อระดับแสง) โดยเฉพาะซึ่งเป็นสิ่งสำคัญในการทดสอบความทนทานของอัลกอริทึมต่อความแปรผันของแสง โดยเฉพาะ ซึ่งสอดคล้องกับวัตถุประสงค์หลักของงานวิจัยนี้ ระยะเวลาภาพและขนาด ภาพถูกถ่ายในระยะ ใกล้-กลาง (close-to-mid range) และถูกปรับขนาดมาตรฐานเป็น 640x640 พิกเซล สำหรับการป้อนเข้าโมเดล YOLO และ Haar Cascade เกณฑ์การแบ่งสภาวะแสง การแบ่งกลุ่มสภาวะแสงใช้เกณฑ์ค่าเฉลี่ยความสว่างของพิกเซลเป็นหลัก โดยภาพที่มีค่าเฉลี่ยความสว่างสูงจัดอยู่ในกลุ่ม High Light และภาพที่มีค่าเฉลี่ยต่ำจัดอยู่ในกลุ่ม Low Light

2. การปรับปรุงคุณภาพภาพด้วย Canny Edge Detection

ภาพต้นฉบับถูกแปลงเป็นภาพระดับสีเทา (grayscale) ก่อน จากนั้นจึงเข้าสู่กระบวนการตรวจจับขอบด้วย Canny Edge Detection เพื่อเน้นโครงสร้างที่สำคัญของใบหน้า โดยผลลัพธ์จากการตรวจจับขอบจะถูกนำมาใช้เป็นตัวช่วยเสริมผ่านการผสมกับภาพ grayscale เดิมในสัดส่วนที่เหมาะสม เพื่อเพิ่ม contrast บริเวณที่มีโครงสร้างชัดเจน โดยไม่แทนที่ข้อมูล grayscale ดั้งเดิม วิธีการนี้ช่วยลดผลกระทบจากแสงที่ไม่สม่ำเสมอ ทำให้ขอบเขตของวัตถุชัดเจนขึ้น และส่งผลให้ขั้นตอนการตรวจจับใบหน้าในลำดับถัดไปสามารถทำงานได้แม่นยำมากยิ่งขึ้น (Canny, 1986)

ในการปรับปรุงประสิทธิภาพของวิธีการตรวจจับใบหน้า โดยเฉพาะในสภาวะแสงที่ท้าทาย งานวิจัยนี้ได้ผสมผสานเทคนิค Canny Edge Detection เข้ากับขั้นตอนการประมวลผลล่วงหน้า (preprocessing) ดังนี้

- การแปลงภาพและตรวจจับขอบ ภาพต้นฉบับจะถูกแปลงเป็น ภาพระดับสีเทา (grayscale) ก่อน จากนั้นจึงเข้าสู่กระบวนการตรวจจับขอบด้วย Canny Edge Detection (Canny, 1986) เพื่อเน้นโครงสร้างที่สำคัญของใบหน้า

- การเสริมข้อมูล (data augmentation) ผลลัพธ์จากการตรวจจับขอบจะถูกนำมาใช้เป็น ตัวช่วยเสริม ผ่านการผสมผสานกับภาพ Grayscale เดิมในสัดส่วนที่เหมาะสม เพื่อเพิ่ม Contrast บริเวณที่มีโครงสร้างชัดเจน โดยไม่แทนที่

ข้อมูลความเข้มของพิกเซลเดิม (grayscale data)

- วิธีการนี้มีวัตถุประสงค์เพื่อลดผลกระทบจากแสงที่ไม่สม่ำเสมอ ทำให้ขอบเขตของวัตถุชัดเจนขึ้น ซึ่งส่งผลให้ขั้นตอนการตรวจจับใบหน้าในลำดับถัดไป (ทั้ง Haar Cascade และ YOLO) สามารถทำงานได้แม่นยำมากยิ่งขึ้น

กระบวนการตรวจจับใบหน้าด้วย Haar Cascade ได้มีการเพิ่มขั้นตอนการปรับปรุงคุณภาพภาพโดยใช้ Canny Edge Detection เพื่อลดผลกระทบจากสภาวะแสงที่ไม่สม่ำเสมอและเพิ่มความคมชัดของโครงร่างใบหน้า ภาพต้นฉบับถูกแปลงเป็น grayscale จากนั้นผ่านการตรวจจับขอบด้วย Canny เพื่อเน้นเส้นขอบของวัตถุ ผลลัพธ์ดังกล่าวถูกใช้สร้าง mask และนำมารวมกับภาพ grayscale เดิม เพื่อเพิ่ม contrast ในบริเวณที่มีความสำคัญ กระบวนการนี้ช่วยให้ Haar Cascade สามารถระบุขอบเขตของใบหน้าได้แม่นยำมากขึ้น แม้ในกรณีที่สภาพแสงมีความแปรผันสูง ดังภาพ 2

จากภาพ 2 ประกอบด้วย (a) ภาพ grayscale (b) ผลลัพธ์การตรวจจับขอบด้วย Canny (c) ภาพที่ได้จากการเสริมขอบเขต (mask) โดยใช้ผลลัพธ์จาก Canny และ (d) ผลการตรวจจับใบหน้าที่ได้จาก Haar Cascade ซึ่ง bounding box ถูกแสดงด้วยสี่เหลี่ยม

ขั้นตอนการปรับปรุงคุณภาพภาพก่อนเข้าสู่กระบวนการตรวจจับใบหน้า ได้นำเสนอผลลัพธ์ในแต่ละขั้นตอน โดยเริ่มจากการแปลงภาพต้นฉบับเป็น grayscale จากนั้นทำการตรวจจับขอบด้วย Canny Edge Detection และนำผลลัพธ์มารวมกับภาพ grayscale เดิมเพื่อเพิ่ม contrast บริเวณโครงสร้างที่สำคัญ สุดท้ายภาพที่ผ่านการปรับปรุงจะถูกส่งเข้าสู่กระบวนการตรวจจับด้วย YOLO ซึ่งสามารถตรวจจับใบหน้าได้แม้ในสภาวะแสงที่มีความไม่สม่ำเสมอ ดังภาพ 3

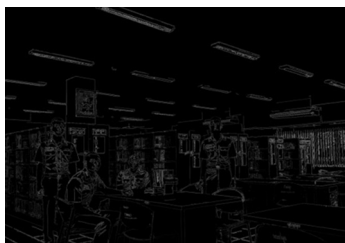
จากภาพ 3 แสดงกระบวนการตรวจจับใบหน้าด้วยวิธี YOLO ร่วมกับ Canny Edge Detection โดยประกอบด้วย (a) ภาพ grayscale, (b) ผลลัพธ์การตรวจจับขอบด้วย Canny, (c) ภาพที่ได้จากการรวม grayscale กับ edge mask เพื่อเพิ่ม contrast และ (d) ผลลัพธ์การตรวจจับใบหน้าด้วย YOLO แสดงในรูปของ bounding box สีเขียว



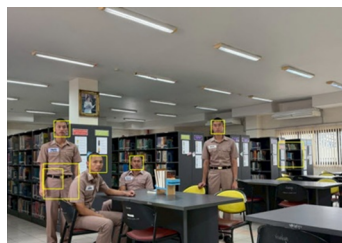
(a)



(b)



(c)

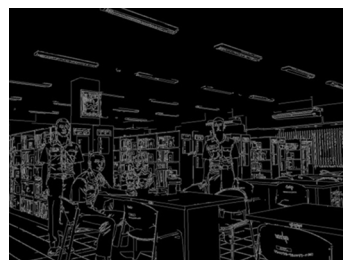


(d)

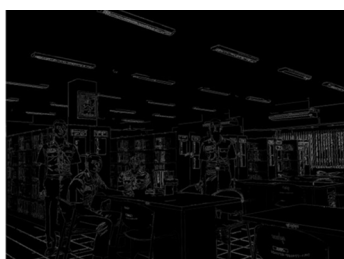
ภาพ 2 ขั้นตอนการประมวลผลภาพด้วย Canny Edge Detection และการตรวจจับใบหน้าด้วย Haar Cascade
(a) ภาพ grayscale (b) ผลลัพธ์การตรวจจับขอบด้วย Canny (c) ภาพที่ได้จากการเสริมขอบเขต (mask) โดยใช้ผลลัพธ์จาก Canny และ (d) ผลการตรวจจับใบหน้าที่ได้จาก Haar Cascade



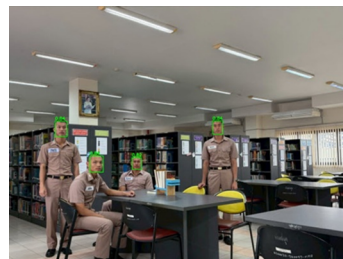
(a)



(b)



(c)



(d)

ภาพ 3 ขั้นตอนการประมวลผลภาพด้วย Canny Edge Detection และการตรวจจับใบหน้าด้วย YOLO
(a) ภาพ grayscale (b) ผลลัพธ์การตรวจจับขอบด้วย Canny (c) ภาพที่ได้จากการรวม grayscale กับ edge mask เพื่อเพิ่ม contrast และ (d) ผลลัพธ์การตรวจจับใบหน้าด้วย YOLO

3. วิธีการตรวจจับใบหน้าหลัก (detection methods)

งานวิจัยนี้ได้เปรียบเทียบประสิทธิภาพของการตรวจจับใบหน้าจากสองวิธีการหลัก คือ Haar Cascade Classifier: Haar วิธีนี้ใช้การเรียนรู้ฟีเจอร์เชิงสถิติผ่าน

กระบวนการ AdaBoost ซึ่งมีจุดเด่นด้านความเร็วและมีประสิทธิภาพในการตรวจจับใบหน้าที่มีขนาดและมุมมองมาตรฐาน (Viola & Jones, 2001) และ YOLO: You Only Look Once วิธีนี้ใช้สถาปัตยกรรมการเรียนรู้เชิงลึก (deep learning) เพื่อระบุตำแหน่ง (bounding

box) ของใบหน้าได้อย่างรวดเร็วและแม่นยำยิ่งขึ้น โดยแสดงความทนทานต่อสภาวะแสงที่ไม่เหมาะสมหรือฉากที่มีความซับซ้อน (Redmon et al., 2016) โดยมีวิธีการดังนี้

การประมวลผลภาพเบื้องต้น (preprocessing with canny edge detection)

การประมวลผลภาพเบื้องต้นด้วย Canny Edge Detection

กระบวนการปรับปรุงภาพด้วย Canny Edge Detection ถูกนำมาใช้เพื่อเน้นโครงสร้างของใบหน้าและลดผลกระทบจากแสงที่ไม่สม่ำเสมอ โดยมีขั้นตอนการประมวลผลที่ชัดเจน ดังนี้ แปลงภาพจากภาพสีต้นฉบับจะถูกแปลงเป็นภาพระดับสีเทา (Grayscale Image, $I_{grayscale}$) เพื่อลดความซับซ้อนของข้อมูลและเตรียมพร้อมสำหรับการหาขอบภาพ การหาขอบ นำภาพ $I_{grayscale}$ ไปใช้ในการตรวจจับขอบด้วย Canny Edge Detection ซึ่งประกอบด้วย การใช้ Gaussian Filter เพื่อลดสัญญาณรบกวน การคำนวณ Gradient Magnitude และ Orientation การทำในส่วน Non-Maximum Suppression และการใช้ Thresholds แบบ Hysteresis เพื่อให้ได้ขอบภาพ (Edge Mask, $I_{edge\ mask}$) ที่บางและมีความแม่นยำสูง และการผสานภาพ ผลลัพธ์จากการตรวจจับขอบ ($I_{edge\ mask}$) จะถูกนำมาผสานกลับเข้ากับภาพระดับสีเทาเดิม ($I_{edge\ mask}$) เพื่อสร้างภาพป้อนเข้า (I_{input}) ที่ถูกเสริมด้วยข้อมูลโครงสร้างของใบหน้า สูตรการผสานภาพแบบถ่วงน้ำหนัก (weighted sum) คือ

$$I_{input} = I_{grayscale} + I_{edge\ mask} \quad (1)$$

ในงานวิจัยนี้ กำหนดค่า $\alpha=0.7$ และ $\beta=0.3$ เพื่อให้ภาพสุดท้ายยังคงรักษาคุณสมบัติความเข้มของภาพเดิมไว้ในขณะที่เน้นขอบเขตสำคัญ ภาพที่นำไปใช้ในการฝึกฝน (train) และการทดสอบ (test) โมเดลตรวจจับ (ทั้ง Haar Cascade และ YOLO) จึงเป็นภาพระดับสีเทาที่ถูกเสริมด้วยข้อมูลขอบภาพ (enhanced grayscale image) จากขั้นตอนการผสานภาพ

ก่อนการตรวจจับใบหน้า ภาพทุกภาพในชุดข้อมูลถูกนำมาผ่านกระบวนการตรวจจับขอบ (edge detection) ด้วย Canny Edge Detection ซึ่งเป็นเทคนิคที่ได้รับความนิยมสูงเนื่องจากมีความสามารถในการระบุขอบเขตของวัตถุได้อย่างแม่นยำและทนต่อสัญญาณรบกวน (Canny, 1986) การประมวลผลนี้ช่วยเน้นโครงสร้างของใบหน้า เช่น

ขอบตา จมูก และกรอบหน้า ซึ่งมีความสำคัญต่อการเรียนรู้และการตรวจจับในขั้นตอนถัดไป นิยามทางคณิตศาสตร์ของ Canny Edge Detection สามารถอธิบายได้ดังสมการที่ (2)

$$E(x,y) = \begin{cases} 1 & \text{if } G(x,y) > T_{high} \\ 0 & \text{if } G(x,y) < T_{low} \\ \text{edge linked} & \text{if } T_{low} \leq T_{high} \end{cases} \quad (2)$$

โดยที่ $G(x,y)$ คือ ค่าความแรงของเกรเดียนต์ (gradient magnitude)

T_{high} , T_{low} คือ ค่าขอบเขตสำหรับการตัดสินใจจุดพิกเซลเป็นขอบหรือไม่

ในการทดลองนี้ ได้กำหนดค่า Thresholds ที่เหมาะสมจากการปรับจูน (tuning) สำหรับชุดข้อมูลนี้ โดยกำหนดให้ $T_{low}=50$ และ $T_{high}=150$ ซึ่งเป็นค่าที่ให้ขอบภาพที่มีความแม่นยำและช่วยลดสัญญาณรบกวนได้อย่างสมดุล

ค่า T_{low} และ T_{high} เหล่านี้มีความสำคัญในการควบคุมคุณภาพของขอบภาพที่ถูกส่งเข้าสู่โมเดลตรวจจับ โดย T_{high} ใช้เพื่อระบุขอบที่ชัดเจน และ T_{low} ใช้เพื่อเชื่อมต่อขอบที่ยังไม่สมบูรณ์ที่อยู่ใกล้เคียง

Canny Edge Detection ถูกเลือกใช้ในงานวิจัยนี้ เพื่อทำหน้าที่เป็นกระบวนการปรับปรุงภาพเบื้องต้น (preprocessing) ก่อนส่งภาพเข้าสู่โมเดลตรวจจับ (Haar Cascade และ YOLO) แทนที่จะเลือกใช้เทคนิคการปรับปรุงความสว่างโดยตรง (เช่น HE หรือ retinex) เนื่องจากมีข้อได้เปรียบที่สำคัญดังต่อไปนี้ ความคงทนต่อแสง (illumination invariance) ข้อมูลขอบภาพ (edge features) เป็นคุณสมบัติที่ทนทานต่อการเปลี่ยนแปลงของสภาวะแสงได้ดีกว่าค่าความเข้มของพิกเซลโดยรวม (intensity values) การใช้ Canny ช่วยลดความแปรปรวนที่เกิดจากแสงเงาที่รุนแรงหรือแสงจ้า (saturation) ทำให้สามารถระบุขอบเขตของใบหน้าได้อย่างเสถียยิ่งขึ้น ความแม่นยำและลดสัญญาณรบกวน (high precision and noise reduction) Canny Edge Detection มีการใช้ตัวกรองแบบ Gaussian Filter ในขั้นตอนแรก ซึ่งช่วยลดสัญญาณรบกวน (noise) ก่อนการหาขอบและมีขั้นตอน Non-Maximum Suppression ทำให้ได้ขอบภาพที่บางและแม่นยำสูง (single-pixel edge) ส่งผลให้โครงสร้างของใบหน้ามีความคมชัดและแยกแยะได้ง่ายขึ้นเมื่อนำไป

ผลงานกับภาพ Grayscale เดิม และประสิทธิภาพการทำงาน (computational efficiency) เมื่อเทียบกับวิธีการอ้างอิง Retinex ที่ต้องมีการคำนวณที่ซับซ้อน Canny เป็นวิธีที่มีขั้นตอนการคำนวณที่ตรงไปตรงมาและรวดเร็ว เหมาะสำหรับการนำไปประยุกต์ใช้ในระบบตรวจจับแบบเรียลไทม์ ซึ่งสอดคล้องกับข้อกำหนดด้านความเร็วของงานวิจัยนี้

วิธีการปรับปรุงภาพด้วย Canny Edge Detection มีความเหมาะสมและให้ผลลัพธ์ที่ดีกว่าเมื่อเปรียบเทียบกับเทคนิคอื่น ๆ ในบริบทของการตรวจจับใบหน้าภายใต้สภาวะแสงที่ไม่คงที่ ดังนี้ เหนือกว่าวิธีการปรับปรุงค่าความเข้ม (intensity-based methods) เช่น Histogram Equalization: HE และ CLAHE: Contrast-Limited Adaptive Histogram Equalization HE/CLAHE มุ่งเน้นการยืดช่วงค่าความเข้ม ซึ่งมักจะนำไปสู่การขยายสัญญาณรบกวน (noise amplification) ในบริเวณที่มีดีหรือก่อให้เกิด สิ่งแปลกปลอมเชิงภาพ (visual artifacts) ที่เป็นอุปสรรคต่อการเรียนรู้ของโมเดล Canny Edge Detection ใช้ Gaussian Filter เพื่อลดสัญญาณรบกวนก่อนการหาขอบ ทำให้โครงสร้างของใบหน้ามีความบริสุทธิ์ (pure structure) และแม่นยำสูง (high precision) ซึ่งเป็นข้อได้เปรียบที่มีคุณภาพสูงกว่าและเหนือกว่าวิธีการอ้างอิงทฤษฎีการสะท้อนแสง (Retinex-based Methods) Retinex-based Methods ให้ผลลัพธ์ที่ดีในเชิงคุณภาพของภาพแต่มีความซับซ้อนในการคำนวณสูง (high computational cost) และใช้เวลาในการประมวลผลที่นาน (higher latency) ซึ่งไม่สอดคล้องกับข้อกำหนดของงานวิจัยที่ต้องการประสิทธิภาพในการทำงานแบบเรียลไทม์ (real-time detection) Canny Edge Detection มีประสิทธิภาพการประมวลผลสูง (high computational efficiency) ในขณะเดียวกันก็ให้คุณสมบัติที่สำคัญที่สุดสำหรับการแก้ไขปัญหากลาง คือ ข้อมูลโครงสร้างที่คงทนต่อแสง (illumination-invariant structural information) การใช้ Canny Edge Detection เป็นการเลือกใช้กลยุทธ์ที่ให้ความสมดุลที่ดีที่สุด (optimal balance) ระหว่างความแม่นยำ (accuracy) ความทนทานต่อสภาวะแสง (robustness) และความเร็วในการประมวลผล (speed) เมื่อเทียบกับเทคนิคการปรับปรุงภาพอื่น ๆ ทำให้โมเดลตรวจจับใบหน้าสามารถทำงานได้อย่างมีประสิทธิภาพภายใต้เงื่อนไขแสงที่หลากหลาย

Canny Edge Detection ในงานวิจัยนี้มีจุดมุ่งหมายเพื่อปรับปรุงคุณลักษณะเชิงโครงสร้าง (structural features) ของใบหน้า ก่อนที่จะนำเข้าสู่โมเดลตรวจจับ (YOLO และ

Haar Cascade) ซึ่งคาดว่าจะช่วยเพิ่มความแม่นยำของการตรวจจับโดยเฉพาะในสภาวะแสงที่มีความแตกต่างสูง

การตรวจจับใบหน้า (face detection methods)

เพื่อตอบโจทย์การประเมินประสิทธิภาพในสภาวะแสงที่แตกต่างกัน งานวิจัยนี้เลือกใช้สองวิธีการตรวจจับใบหน้าหลัก ได้แก่ Haar Cascade Classifier และ YOLO: You Only Look Once ซึ่งมีพื้นฐานทางทฤษฎีและแนวคิดที่แตกต่างกัน ดังนี้

เพื่อตอบโจทย์การประเมินประสิทธิภาพภายใต้สภาวะแสงที่แตกต่างกัน งานวิจัยนี้ได้เลือกใช้สองวิธีการตรวจจับใบหน้าหลัก ซึ่งมีพื้นฐานทางทฤษฎีและแนวคิดที่แตกต่างกัน ได้แก่ Haar Cascade Classifier และ YOLO: You Only Look Once ดังรายละเอียดต่อไปนี้

Haar Cascade Classifier

Haar Cascade Classifier พัฒนาขึ้นจาก Viola-Jones Algorithm (Viola & Jones, 2001) ซึ่งเป็นแนวคิดด้านการเรียนรู้ของเครื่องเชิงสถิติ (statistical machine learning) หลักการทำงานประกอบด้วย

- คุณลักษณะเด่น (features) ใช้วิธี Haar-like Features ในการคำนวณความแตกต่างของความเข้มพิกเซลในบริเวณต่าง ๆ ของภาพ เพื่อค้นหาคุณลักษณะของใบหน้า เช่น ขอบ ตา หรือจมูก

- การจัดจำแนก (classification) ใช้ AdaBoost Classifier ในการรวม Weak Classifiers จำนวนมากเข้าด้วยกัน เพื่อสร้างเป็น Cascade Structure (โครงสร้างแบบเรียงลำดับ) ซึ่งช่วยลดเวลาการประมวลผลโดยการคัดกรองพื้นที่ที่ไม่ใช่ใบหน้าออกอย่างรวดเร็ว

- ข้อดีและข้อจำกัด ข้อดีของวิธีนี้ คือ มีความเร็วในการตรวจจับสูง และเหมาะกับงาน Real-Time อย่างไรก็ตาม Haar Cascade มัก ไวต่อสภาวะแสงและคุณภาพของภาพ ซึ่งเป็นข้อจำกัดหลักที่ส่งผลต่อความแม่นยำของการตรวจจับ

สำหรับการทดลองด้วย Haar Cascade Classifier ได้ใช้ชุดคุณลักษณะใบหน้ามาตรฐานของ OpenCV (haarcascade_frontalface_default.xml) และกำหนดค่าพารามิเตอร์หลัก ดังนี้ scaleFactor=1.1 minNeighbors=5 และ minSize= 30, 30) ซึ่งเป็นค่าที่ให้ความสมดุลระหว่างความแม่นยำและความเร็วในการทำงานภายใต้สภาพแวดล้อมการทดลองนี้

YOLO (You Only Look Once)

YOLO เป็นโมเดลตรวจจับวัตถุเชิงลึก (deep learning-based object detection) ที่ได้รับการพัฒนาโดย Redmon et al. (2016) ซึ่งมีแนวคิดหลัก คือ การมองการตรวจจับวัตถุเป็นปัญหาการถดถอย (regression problem) โดยคาดการณ์ bounding box และ class probability ของวัตถุในครั้งเดียว (single forward pass) จึงมีความเร็วและความแม่นยำสูงกว่าวิธีการแบบดั้งเดิม ฟังก์ชันที่ใช้ในการทำนาย bounding box สามารถแสดงได้ดังสมการที่ (3)

$$b_x = \sigma(t_x) + c_x, b_y = \sigma(t_y) + c_y, b_w = p_w e^{t_w}, b_h = p_h e^{t_h} \quad (3)$$

โดยที่ b_x และ b_y คือ จุดกึ่งกลางของ bounding box
 b_w และ b_h คือ ความกว้างและความสูงของ bounding box

c_x และ c_y คือ ตำแหน่งของ grid cell

p_w และ p_h คือ ค่า prior anchor box

σ คือ sigmoid function

YOLO จึงสามารถเรียนรู้และตรวจจับใบหน้าได้ แม้ในสภาพแสงที่ซับซ้อนกว่าวิธีการดั้งเดิม ซึ่งเหมาะสมกับวัตถุประสงค์ของงานวิจัยนี้ที่ต้องการเปรียบเทียบความทนทานของโมเดลต่อความแตกต่างของสภาวะแสง

เพื่อให้เห็นความแตกต่างอย่างชัดเจนระหว่างสองเทคนิคที่ใช้ในการตรวจจับใบหน้า ซึ่งสามารถแสดงการเปรียบเทียบเชิงแนวคิดระหว่าง Haar Cascade Classifier และ YOLO (You Only Look Once) โดยเน้นจุดเด่นและข้อจำกัดที่เกี่ยวข้องกับงานวิจัยนี้ดังตาราง 1

4. การประเมินผลการตรวจจับด้วยเมตริกมาตรฐาน

ผลการตรวจจับจากวิธีการทั้งหมดจะถูกนำมาวิเคราะห์โดยใช้ชุดของเมตริกมาตรฐานที่ได้รับการยอมรับในงานด้าน Computer Vision (Powers, 2011) เพื่อเปรียบเทียบประสิทธิภาพภายใต้สภาวะแสงที่แตกต่างกันอย่างเป็นระบบ เมตริกที่ใช้ในการประเมินประกอบด้วย

Accuracy (ความแม่นยำรวม)

Precision (ความเที่ยงตรง)

Recall (ความไว/ความสมบูรณ์)

F1-Score (ค่าเฉลี่ยแบบฮาร์โมนิกของ Precision และ Recall)

Processing Time (เวลาในการประมวลผล)

ซึ่งเป็นตัวชี้วัดที่สะท้อนทั้งความถูกต้องและความสามารถในการตรวจจับวัตถุอย่างครอบคลุม (Powers, 2011)

โดยนิยามและสูตรคำนวณแสดงดังสมการที่ (4) (5) (6) และ (7)

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$F1 = 2 \frac{Precision \cdot Recall}{Precision + Recall} \quad (7)$$

โดยที่ True Positive: TP คือ จำนวนครั้งที่ระบบตรวจจับใบหน้าได้ถูกต้อง (ตรวจเจอและเป็นใบหน้าจริง)

True Negative: TN คือ จำนวนครั้งที่ระบบตรวจไม่พบใบหน้า และในภาพนั้นไม่มีใบหน้าจริง

False Positive: FP คือ จำนวนครั้งที่ระบบตรวจจับผิดพลาด โดยตรวจว่าเป็นใบหน้าแต่จริง ๆ แล้วไม่มีใบหน้า

False Negative: FN คือ จำนวนครั้งที่ระบบตรวจไม่พบใบหน้า ทั้งที่ในภาพมีใบหน้าจริง

ความหมายของเมตริก

Accuracy แสดงสัดส่วนความถูกต้องโดยรวมของการตรวจจับ

Precision บ่งบอกความถูกต้องของการตรวจจับเมื่อระบบบอกว่ามีใบหน้า (ลด false detection)

Recall สะท้อนความสามารถของระบบในการตรวจจับใบหน้าที่มีอยู่จริงทั้งหมด (ลด missed detection)

F1-Score เป็นค่าเฉลี่ยเชิงฮาร์โมนิกระหว่าง Precision และ Recall ใช้เมื่อต้องการวัดประสิทธิภาพโดยรวมในกรณีที่ต้องการทั้งความถูกต้องและความครอบคลุม

จากภาพ 4 และ 5 แสดงตัวอย่างกรณี False Negative: FN ซึ่งเกิดขึ้นเมื่อโมเดลไม่สามารถตรวจจับวัตถุเป้าหมายได้ ทั้งที่วัตถุ (ใบหน้า) มีอยู่ในภาพจริง กรอบสีแดง แสดงถึงผลลัพธ์การตรวจจับ ขอบเขตของใบหน้าที่โมเดลตรวจจับได้ถูกต้อง และกรอบสีน้ำเงิน แสดงถึงตำแหน่งของใบหน้าจริง ที่โมเดล ตรวจไม่พบ (missed detection) ซึ่งเท่ากับค่า False Negative: FN ข้อผิดพลาด FN มักเกิดจากคุณภาพของภาพที่ต่ำ ความเปรียบต่าง (contrast) ระหว่างใบหน้าและพื้นหลังที่ไม่ชัดเจน หรือสัญญาณรบกวน (noise) ที่เกิดจากสภาวะแสงน้อย ซึ่งส่งผลให้โมเดลไม่สามารถแยกแยะคุณลักษณะเด่นของใบหน้าได้

ตาราง 1

การเปรียบเทียบคุณลักษณะของ Haar Cascade และ YOLO

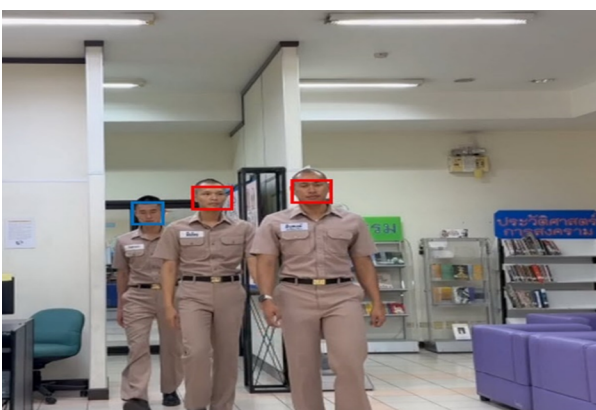
คุณลักษณะ	Haar Cascade Classifier	YOLO (You Only Look Once)
หลักการทำงาน	ใช้ Haar-like features และ AdaBoost classifier จัดเรียงเป็น cascade ของ weak classifiers หลายชั้น (Viola & Jones, 2001)	ใช้ Convolutional Neural Networks (CNNs) เพื่อตรวจจับวัตถุใน single forward pass (Redmon et al., 2016)
ความเร็ว	เร็วมาก เหมาะกับ real-time บนอุปกรณ์ประสิทธิภาพต่ำ	เร็วมาก โดยเฉพาะรุ่นใหม่ (YOLOv3-YOLOv8) ที่รองรับ real-time
ความแม่นยำ	ขึ้นอยู่กับสภาพแสงและคุณภาพของภาพ มักลดลงในสภาวะแสงน้อย (Lienhart & Maydt, 2002)	มีความแม่นยำสูงและทนต่อสภาวะแสงที่หลากหลายมากกว่า (Bochkovskiy et al., 2020)
ความยืดหยุ่น	ใช้ได้เฉพาะใบหน้าหรือวัตถุที่ผ่านการ train ด้วย features ที่กำหนด	ตรวจจับได้หลายประเภทของวัตถุพร้อมกัน และสามารถ fine-tune ได้
การประยุกต์ใช้งาน	เหมาะกับงานที่ต้องการความเร็วสูงแต่ไม่ซับซ้อน เช่น face tracking เบื้องต้น	เหมาะกับงานที่ต้องการความแม่นยำสูง เช่น surveillance biometrics และ autonomous systems



ภาพ 4 ภาพตัวอย่าง False Negative: FN



ภาพ 6 False Positive: FP



ภาพ 5 ภาพตัวอย่าง False Negative: FN



ภาพ 7 False Positive: FP

จากภาพ 6 และ 7 แสดงตัวอย่างข้อผิดพลาดแบบ False Positive: FP ซึ่งเกิดขึ้นเมื่อโมเดลตรวจจับวัตถุที่ไม่ใช่ใบหน้าและตีความว่าเป็นใบหน้า กรอบสีน้ำเงินในภาพ 6 แสดงถึงการตรวจจับใบหน้าที่ผิดพลาดมักเกิดขึ้นในภาพที่มีสภาวะแสงมาก (high light) หรือมีเงาจัดบนพื้นผิว ภาพ 7 เนื่องจากความเปรียบต่าง (contrast) ที่รุนแรงของแสงและเงา (เช่น ขอบเสื้อ มุมวัตถุ หรือบริเวณแสงสะท้อนบนผิว) สร้างรูปแบบขอบ (edge patterns) การใช้เมตริกทั้งสองตัวร่วมกันช่วยให้การประเมินผลครอบคลุมทั้งความแม่นยำโดยรวมและประสิทธิภาพในการจัดการกับข้อผิดพลาดที่เกิดขึ้นในแต่ละกรณี

การประเมินเวลาในการประมวลผล (processing time evaluation)

นอกจากเมตริกความแม่นยำแล้ว งานวิจัยยังประเมินประสิทธิภาพเชิงความเร็วของระบบ โดยวัดเวลาในการประมวลผล (processing time) ในหน่วยวินาทีต่อภาพ (seconds per image) ซึ่งวัดตั้งแต่การรับภาพเข้าสู่ระบบจนกระทั่งได้ผลลัพธ์การตรวจจับ (bounding box) ตามสูตรการวัดเวลาดังสมการที่ (8)

$$\text{Processing Time} = \text{Time}(\text{Pre-processing} + \text{Detection}) \quad (8)$$

โดยที่ $\text{Time}(\text{Pre-processing})$ คือ เวลาที่ใช้ในการแปลงภาพสีเป็น Grayscale การทำ Canny Edge Detection และการผสมภาพ (weighted sum)

$\text{Time}(\text{Detection})$ คือ เวลาที่โมเดล (Haar Cascade หรือ YOLO) ใช้ในการอนุมานผลลัพธ์ (inference time)

ค่าที่รายงานในส่วนผลการวิจัยเป็น ค่าเฉลี่ยของเวลาในการประมวลผลต่อภาพ ซึ่งได้จากการทดสอบภาพ

ทั้งหมด 1,500 ภาพ บนสภาพแวดล้อมฮาร์ดแวร์ที่ระบุ

ผลการวิจัยและการอภิปรายผล

ชุดข้อมูลและการแบ่งระดับสภาวะแสง

การประเมินผลการทดลองได้ดำเนินการบนชุดข้อมูลภาพใบหน้าจำนวน 1,500 ภาพ ภายใต้สภาวะแสง 3 ระดับ ได้แก่ แสงมาก (high light) แสงปานกลาง (medium light) และแสงน้อย (low light)

เกณฑ์ในการแบ่งระดับสภาวะแสง พิจารณาจากค่าเฉลี่ยของความสว่าง (brightness) ของภาพ ซึ่งคำนวณจากค่าเฉลี่ยของความเข้มพิกเซล (pixel intensity average) ในช่วงค่า 0–255 (โดยที่ 0 คือ มืดที่สุด และ 255 คือ สว่างที่สุด) โดยมีเกณฑ์การแบ่งสภาวะแสง ดังตาราง 2

วิธีการทดสอบและตัวชี้วัดประสิทธิภาพ

ภาพในแต่ละสภาวะแสงถูกนำมาทดสอบด้วย 4 วิธีการตรวจจับหลัก เพื่อเปรียบเทียบประสิทธิภาพ ได้แก่

- 1) Haar Cascade Classifier (Haar)
- 2) YOLO (You Only Look Once)
- 3) Haar Cascade + Canny Edge Detection (Haar + Canny)
- 4) YOLO + Canny Edge Detection (YOLO + Canny)

การประเมินผลใช้ตัวชี้วัดมาตรฐานที่ได้รับการยอมรับอย่างแพร่หลายในการวิจัยด้านการตรวจจับวัตถุและใบหน้า (Powers, 2011) ประกอบด้วย Accuracy Precision Recall และ F1-Score นอกจากนี้ ยังมีการวัดเวลาในการประมวลผล (time) เพื่อประเมินประสิทธิภาพเชิงความเร็วของแต่ละวิธีการด้วย

ตาราง 2

สภาวะแสงและลักษณะทั่วไป

สภาวะแสง	ช่วงค่าเฉลี่ยความสว่าง (intensity level)	ลักษณะภาพโดยทั่วไป
แสงน้อย	50 - 100	ภาพมืด รายละเอียดบางส่วนอาจสูญหาย
แสงปานกลาง	100 - 180	ระดับแสงปกติ เหมาะสมที่สุดสำหรับการตรวจจับใบหน้า
แสงมาก	มากกว่า 180	ภาพสว่างจ้า อาจเกิดการสูญเสียรายละเอียดจากการโอเวอร์เอ็กซ์โพส

ตาราง 3

ประสิทธิภาพของ Haar Cascade Haar + Canny YOLO และ YOLO + Canny ภายใต้สภาวะแสงต่าง ๆ

สภาวะแสง	Model	TP	FP	FN	TN	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Time (s)
มาก	Haar Cascade	254	44	46	156	82.0	85.2	84.7	85.0	1.502
	Haar + Canny	243	53	57	147	78.0	82.1	80.9	81.5	1.179
	YOLO	273	25	27	175	89.6	91.6	90.9	91.2	2.128
	YOLO + Canny	276	21	24	179	91.0	92.8	92.1	92.4	0.133
ปานกลาง	Haar Cascade	248	47	52	153	80.2	83.9	82.5	83.2	1.378
	Haar + Canny	239	58	61	142	76.2	80.5	79.6	80.1	1.218
	YOLO	278	19	22	181	91.8	93.5	92.7	93.1	2.175
	YOLO + Canny	278	18	22	182	92.0	94.1	92.8	93.4	0.162
น้อย	Haar Cascade	213	80	87	120	66.6	72.6	70.9	71.7	1.087
	Haar + Canny	200	91	100	109	61.8	68.7	66.8	67.7	0.808
	YOLO	260	36	40	164	84.8	87.9	86.5	87.2	2.020
	YOLO + Canny	266	30	34	170	87.2	89.7	88.5	89.1	0.157

การอภิปรายผล

จากตาราง 3 สามารถสรุปและอภิปรายผลได้ ดังนี้

1. YOLO + Canny

วิธีการ YOLO ร่วมกับ Canny Edge Detection ให้ผลลัพธ์ที่ดีที่สุดในทุกสภาวะแสง ทั้งในด้าน Accuracy Precision Recall และ F1-Score โดยเฉพาะในสภาวะแสงปานกลาง ซึ่งได้ค่า Accuracy สูงสุดถึง 92.0% และยังคงรักษาความแม่นยำในสภาวะแสงมาก (91.0%) และสภาวะแสงน้อย (87.2%) ได้อย่างต่อเนื่อง จุดเด่นสำคัญ คือ ใช้เวลาในการประมวลผลเฉลี่ยเพียง 0.13–0.16 วินาที/ภาพ เร็วกว่าวิธีอื่นทั้งหมดอย่างมาก แสดงให้เห็นว่า preprocessing ด้วย Canny ช่วยเน้นขอบเขตของใบหน้าและลดผลกระทบ

จากแสงที่ไม่สม่ำเสมอ ทำให้โมเดล YOLO สามารถตรวจจับใบหน้าได้แม่นยำและรวดเร็วยิ่งขึ้น

2. YOLO (ปกติ)

วิธีการ YOLO เพียงอย่างเดียวมีประสิทธิภาพสูงเช่นกัน โดยได้ Accuracy เฉลี่ยสูงกว่า 84% แม้ในสภาวะแสงน้อย ซึ่งยังคงเหนือกว่าวิธี Haar Cascade อย่างมีนัยสำคัญ อย่างไรก็ตามเวลาที่ใช้ในการประมวลผลค่อนข้างสูง (~2.0–2.2 วินาที/ภาพ) เมื่อเทียบกับ YOLO + Canny การทำงานที่ช้าลงสะท้อนถึงการที่ YOLO ต้องประมวลผลจากภาพดิบโดยตรงโดยไม่ได้รับการปรับปรุงคุณภาพภาพก่อนการตรวจจับ

3. Haar Cascade

วิธี Haar Cascade แสดงให้เห็นถึงความสามารถในการทำงานที่รวดเร็วปานกลาง ($\approx 1.1-1.5$ วินาที/ภาพ) แต่มีข้อจำกัดด้านความแม่นยำ โดยเฉพาะในสภาวะแสงน้อย ที่ค่า Accuracy ลดลงเหลือเพียง 66.6% เนื่องจากเป็นวิธีการตรวจจับเชิงคุณลักษณะ (feature-based detection) ซึ่งมีความอ่อนไหวต่อการเปลี่ยนแปลงของแสงและสภาพแวดล้อมที่ไม่สม่ำเสมอ

4. Haar + Canny

การนำ Canny Edge Detection มาผสานกับ Haar Cascade ไม่ได้ช่วยเพิ่มประสิทธิภาพการตรวจจับ กลับทำให้ค่าความแม่นยำลดลงในหลายสภาวะ เช่น ในสภาวะแสงน้อย Accuracy เหลือเพียง 61.8% สาเหตุเนื่องจากการตรวจจับขอบภาพ (edges) มากเกินไปก่อให้เกิด noise ที่รบกวนการทำงานของ cascade classifier และทำให้การแยกแยะใบหน้าไม่แม่นยำเท่าที่ควร

5. การเปรียบเทียบกับงานวิจัยที่เกี่ยวข้อง (comparison with related work)

เพื่อเติมเต็มช่องว่างที่ขาดการเปรียบเทียบกับงานวิจัยอื่น และเพื่อให้ผลการวิจัยนี้มีความสมบูรณ์ยิ่งขึ้น จึงได้นำผลลัพธ์ของวิธี YOLO + Canny มาเปรียบเทียบกับเชิงแนวคิดกับประสิทธิภาพของโมเดลตรวจจับใบหน้าเชิงลึกที่รายงานในงานวิจัยอื่น ๆ ที่เน้นการจัดการกับสภาพแสงที่ท้าทาย (แม้ว่าจะใช้ชุดข้อมูลและสภาพแวดล้อมการทดลองที่แตกต่างกัน) โดยเปรียบเทียบในเชิงประสิทธิภาพโดยรวมและคุณสมบัติเด่น

YOLOv8 (Yisihak & Li, 2024) งานวิจัยที่รายงานประสิทธิภาพของ YOLOv8 ในระบบประยุกต์จริง โดยให้ค่าความแม่นยำสูงถึง 98.1% บนชุดข้อมูลที่แตกต่างกัน. ผลลัพธ์นี้แสดงให้เห็นว่ายังมีโอกาสในการพัฒนาโมเดล YOLO ให้ดีขึ้นได้อีกหากมีการใช้เวอร์ชันที่ใหม่กว่า (เช่น YOLOv8) และมีการปรับแต่ง (fine-tuning) พารามิเตอร์ของโมเดลเชิงลึกโดยเฉพาะ Dark-YOLO (Liu et al., 2025) โมเดลนี้ถูกออกแบบมาเพื่อรับมือกับภาพมืดโดยเฉพาะ รายงานค่า F1-Score สูงกว่า 90% ในสภาวะแสงน้อย ซึ่งใกล้เคียงกับ F1-Score ของ YOLO + Canny (89.1%) ในงานวิจัยนี้

บทสรุปจากการเปรียบเทียบวิธีการ YOLO + Canny ที่เสนอในงานวิจัยนี้ ให้ผลลัพธ์ด้านความแม่นยำ (สูงสุด 92.0%) และความทนทานต่อสภาวะแสงที่แปรผันได้อย่างยอดเยี่ยม เมื่อพิจารณาพร้อมกับจุดเด่นด้านความเร็วในการประมวลผลที่สั้นที่สุด (< 0.2 วินาที/ภาพ) พบว่า วิธีการนี้มีความได้เปรียบในเชิงของ Trade-off ระหว่างความแม่นยำ

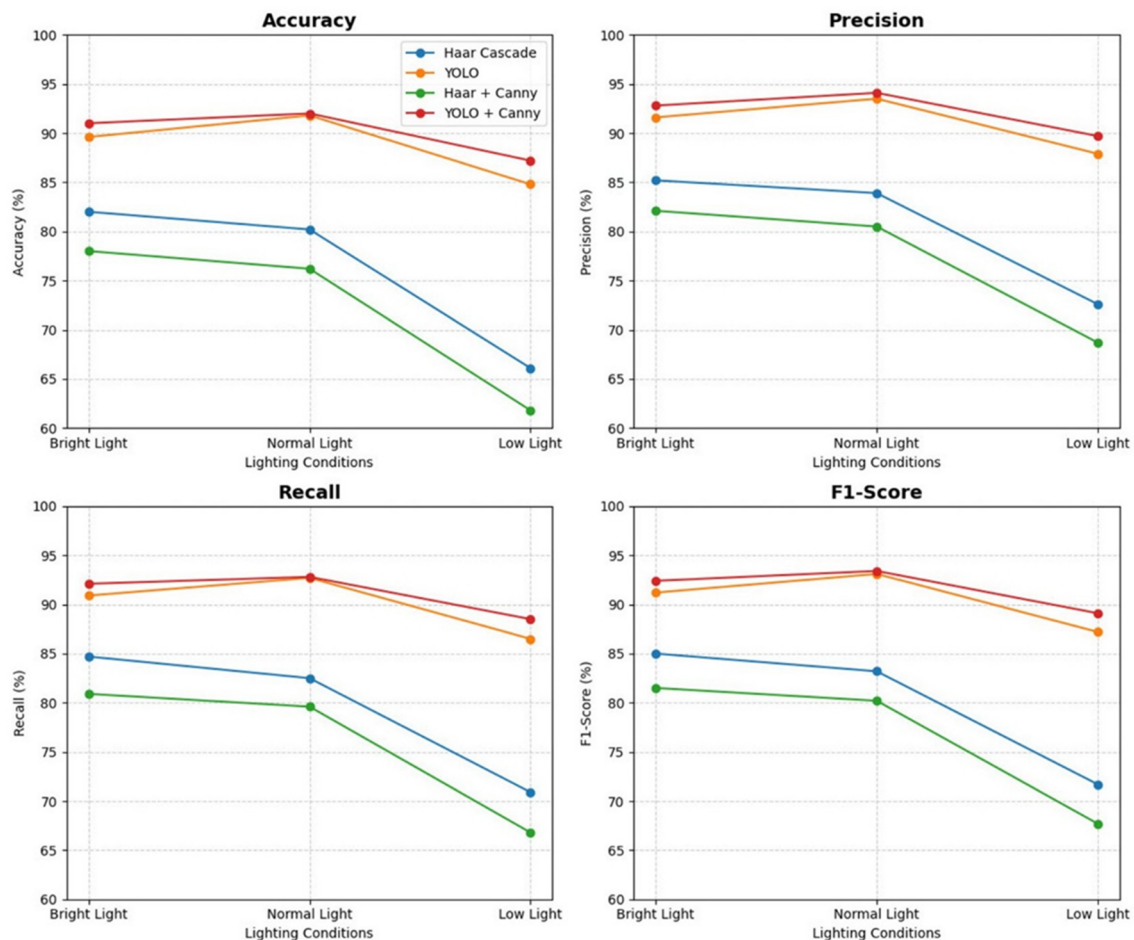
และความเร็ว ซึ่งถือเป็นทางเลือกที่มีประสิทธิภาพสูงสำหรับการประยุกต์ใช้งานแบบเรียลไทม์บนอุปกรณ์ที่มีทรัพยากรจำกัด เมื่อเทียบกับโมเดลเชิงลึกที่ซับซ้อนกว่าซึ่งมักแลกมาด้วยเวลาประมวลผลที่สูงกว่าอย่างมีนัยสำคัญ

โดยสรุป YOLO + Canny แสดงให้เห็นถึงความเหนือกว่าในทุกด้าน ทั้งด้านความแม่นยำและเวลาในการประมวลผล โดยสามารถตรวจจับใบหน้าได้อย่างมีประสิทธิภาพภายใต้สภาวะแสงที่แตกต่างกัน ขณะที่ YOLO แบบปกติ แม้จะมีความแม่นยำสูง แต่มีข้อจำกัดเรื่องเวลา ส่วน Haar Cascade และ Haar + Canny มีจุดเด่นด้านความเร็วในระดับหนึ่ง แต่ความแม่นยำต่ำและไม่เหมาะสมในสภาวะแสงที่ท้าทาย โดยเฉพาะแสงน้อย

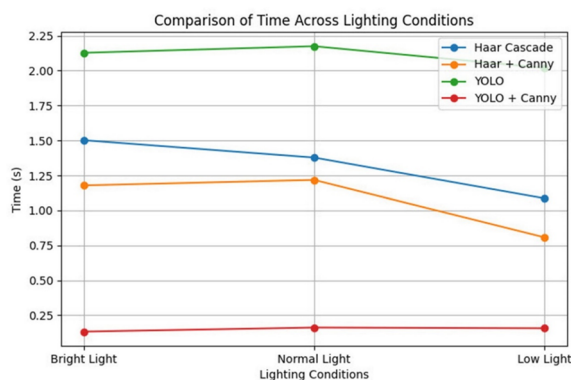
YOLO + Canny ใช้เวลาในการประมวลผลเฉลี่ยน้อยกว่า YOLO ปกติ แม้จะเพิ่มขั้นตอน preprocessing ก็ตามสาเหตุที่เป็นไปได้คือการประมวลผลด้วย Canny Edge Detection ช่วยลดรายละเอียดและข้อมูลที่ไม่จำเป็น ทำให้ input ของโมเดลมีความเด่นชัดของโครงร่างมากขึ้น ส่งผลให้ YOLO สามารถสร้าง feature map ได้รวดเร็วขึ้นและลดภาระการคำนวณในบาง layer นอกจากนี้งานวิจัยที่ผ่านมา เช่น Liu et al. (2025), Sun et al. (2022) และ Yisihak and Li (2024) ยังรายงานผลลัพธ์ในทิศทางเดียวกันว่า preprocessing สามารถช่วยเพิ่มทั้งความแม่นยำและความเร็วในการตรวจจับได้ อย่างไรก็ตามควรระบุว่า ประสิทธิภาพนี้อาจขึ้นอยู่กับลักษณะของ dataset และสถาปัตยกรรมฮาร์ดแวร์ที่ใช้ ซึ่งควรมีการทดสอบเพิ่มเติมในอนาคต

เพื่อให้เห็นแนวโน้มประสิทธิภาพของอัลกอริทึมแต่ละวิธีชัดเจนมากขึ้น งานวิจัยนี้ได้นำผลการจำลองจากตารางไปแสดงในรูปของกราฟเส้น โดยแยกออกเป็น 2 ส่วน ได้แก่ กราฟแสดงค่าความแม่นยำเชิงเมตริกมาตรฐาน (Accuracy Precision Recall และ F1-Score) และกราฟเปรียบเทียบเวลาในการประมวลผล (time) ภายใต้สภาวะแสงที่แตกต่างกัน ทั้งแสงมาก แสงปานกลาง และแสงน้อย

จากภาพ 8 พบว่า YOLO + Canny ให้ผลลัพธ์ที่ดีที่สุดในทุกสภาวะแสง โดยมีค่าของทุกเมตริกสูงกว่า YOLO ปกติ และสูงกว่า Haar Cascade อย่างมีนัยสำคัญ โดยเฉพาะในสภาวะแสงปานกลางที่ Accuracy สูงสุดถึง 92.0% ขณะที่ Haar Cascade และ Haar + Canny มีแนวโน้มค่าลดลงชัดเจนเมื่อความสว่างของภาพต่ำลง โดย Haar + Canny ให้ผลด้อยที่สุดในสภาวะแสงน้อย (Accuracy เพียง 61.8%) แสดงให้เห็นว่า YOLO มีความเสถียรมากกว่าในการจัดการกับการเปลี่ยนแปลงของแสง และ preprocessing แบบ Canny ช่วยเพิ่มประสิทธิภาพของ YOLO ได้ แต่กลับไม่ส่งผลบวกต่อ Haar



ภาพ 8 กราฟ Accuracy Precision Recall และ F1-Score



ภาพ 9 กราฟ Comparison of Time Across Lighting Conditions (s)

จากผลการทดลองที่แสดงในภาพ 9 สามารถสรุปได้อย่างชัดเจนว่า YOLO + Canny Edge Detection เป็นวิธีที่มีประสิทธิภาพเหนือกว่าอัลกอริทึมอื่น ๆ ในทุกสภาวะแสงที่ทำการทดสอบ ทั้งด้านความแม่นยำ (accuracy) ความถูกต้อง (precision) ความครอบคลุม (recall) และ

ความสมดุลระหว่าง Precision และ Recall (F1-Score)

ประสิทธิภาพสูงสุด วิธี YOLO + Canny แสดงค่า Accuracy สูงสุดถึง 92.0% ในสภาวะแสงปานกลาง และยังคงรักษาประสิทธิภาพได้ดีอย่างยิ่งแม้ในสภาวะแสงน้อย ซึ่งโดยปกติเป็นเงื่อนไขที่ท้าทายต่อการตรวจจับใบหน้า

ข้อจำกัดของ Haar Cascade แม้ว่า Haar Cascade จะมีการประมวลผลที่ค่อนข้างรวดเร็ว แต่มีข้อจำกัดด้านความแม่นยำอย่างชัดเจน โดยเฉพาะในสภาวะแสงน้อย ที่ค่า Accuracy ลดลงเหลือเพียง 66.6%

เมื่อพิจารณาด้านเวลาในการประมวลผลต่อภาพ (time/image) ซึ่งแสดงในภาพ 9 พบข้อสรุปที่สำคัญเกี่ยวกับผลกระทบของ preprocessing

YOLO (ปกติ) ใช้เวลาในการประมวลผลสูงที่สุด (≈ 2 วินาที/ภาพ)

Haar Cascade และ Haar + Canny ใช้เวลาใกล้เคียงกัน (1.1–1.5 วินาที/ภาพ)

YOLO + Canny กลับใช้เวลาต่ำที่สุด (< 0.2 วินาที/ภาพ) ในทุกสภาวะแสง

การลดลงของเวลาในการประมวลผลนี้เป็นผลโดยตรงจากการเปลี่ยนรูปแบบอินพุต (input format) โดยสำหรับวิธี YOLO ปกติ ภาพป้อนเข้า คือ ภาพสี RGB ขนาด 640x640 พิกเซล (3 Channel) ในขณะที่วิธี YOLO + Canny ภาพป้อนเข้าถูกลดเหลือเพียง ภาพระดับสีเทาที่ถูกเสริมขอบ (enhanced grayscale image) ซึ่งเป็น 1 Channel การลดจำนวน Channel ของภาพป้อนเข้าในขั้นตอน Pre-processing นี้ช่วย ลดภาระการคำนวณ (computational load) ในเลเยอร์เริ่มต้นของโครงข่ายประสาทเทียม CNN ส่งผลให้เวลาในการอนุมานผล (inference time) ของโมเดล YOLO โดยรวมลดลงอย่างมีนัยสำคัญ (เร็วกว่า YOLO ปกติเกือบ 10 เท่า) ซึ่งเป็นการปรับปรุงประสิทธิภาพที่สอดคล้องกับข้อกำหนดด้าน Real-time ของงานวิจัย นอกจากนี้ Preprocessing ด้วย Canny ยังช่วยเน้นขอบเขตของใบหน้า ลดผลกระทบจากแสงที่ไม่สม่ำเสมอ และเพิ่มความแม่นยำควบคู่ไปกับความเร็วที่เพิ่มขึ้น

สรุปการตีความผลลัพธ์ของการผสมผสานเทคนิคนี้ ดังนี้ การผสมผสาน YOLO + Canny การผสมผสานนี้พิสูจน์แล้วว่า มีประโยชน์สูงสุด โดยให้ค่าความแม่นยำสูงสุดและใช้เวลาน้อยที่สุด ซึ่งแสดงให้เห็นถึงความเข้ากันได้ของ

เทคนิคการเน้นขอบเขต Canny กับสถาปัตยกรรมเชิงลึก YOLO ในการรับมือกับความแปรผันของแสง

การผสมผสาน Haar + Canny การ preprocessing ด้วย Canny ไม่ได้ช่วยปรับปรุงประสิทธิภาพของ Haar Cascade เท่าที่ควร และในบางกรณีกลับทำให้ประสิทธิภาพแย่ลง (accuracy) เนื่องจากการสร้างขอบภาพที่มากเกินไป อาจกลายเป็น noise ที่ทำให้ cascade classifier สับสน และไม่สามารถจับคู่คุณลักษณะได้แม่นยำ

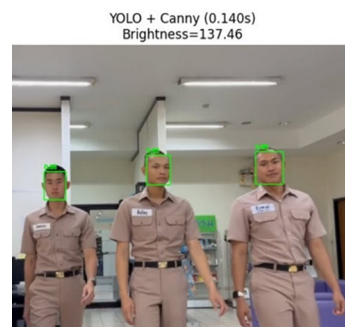
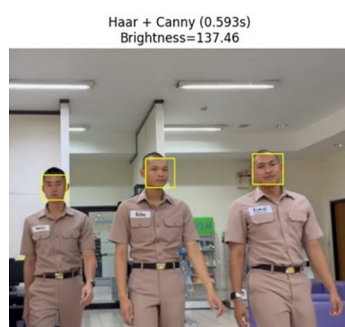
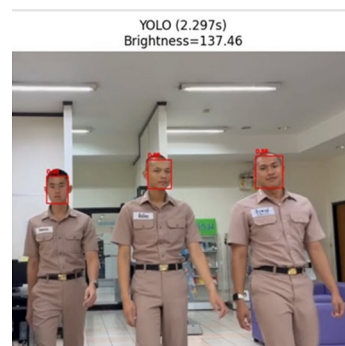
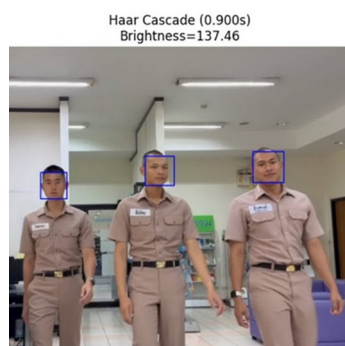
ภาพ 10 แสดงผลการตรวจจับใบหน้าในสภาวะแสงมาก ซึ่งภาพมีความสว่างสูงและอาจเกิดปัญหาโอเวอร์เอ็กซิโพสบางบริเวณ ส่งผลให้รายละเอียดของใบหน้าลดลง วิธี YOLO ร่วมกับ Canny Edge Detection สามารถลดผลกระทบจากความสว่างที่มากเกินไปได้ โดย Canny ช่วยเน้นโครงสร้างขอบเขตของใบหน้า ทำให้ YOLO ระบุตำแหน่งใบหน้าได้แม่นยำ แม้ในสภาวะที่มีความแปรปรวนของแสงสูง ขณะที่ Haar Cascade มีแนวโน้มเกิด False Positive จากขอบหรือเงาในภาพ

ภาพ 11 แสดงผลการตรวจจับใบหน้าในสภาวะแสงปานกลาง ซึ่งเป็นสภาวะที่เหมาะสมที่สุด เนื่องจากมีความสว่างและความแปรปรวนที่สมดุล ผลการทดลองพบว่า YOLO + Canny ให้ประสิทธิภาพสูงสุด สามารถตรวจจับใบหน้าได้ครบถ้วนและแม่นยำ การใช้ Canny Edge Detection ในขั้นตอน preprocessing ช่วยเสริมโครงสร้างของใบหน้าให้เด่นชัด ส่งผลให้ YOLO สร้าง feature ได้อย่างมีประสิทธิภาพ ขณะที่ Haar Cascade ยังมีข้อจำกัดด้านความเสถียรเมื่อเทียบกับโมเดลเชิงลึก

ภาพ 12 แสดงผลการตรวจจับใบหน้าในสภาวะแสงน้อย ซึ่งเป็นกรณีที่ท้าทาย เนื่องจากภาพมีความสว่างต่ำและมีสัญญาณรบกวนสูง วิธี YOLO + Canny Edge Detection สามารถรักษาประสิทธิภาพได้ดีกว่าวิธีอื่น โดย Canny ช่วยดึงข้อมูลเชิงโครงสร้างที่คงทนต่อการเปลี่ยนแปลงของแสง ทำให้ YOLO ตรวจจับใบหน้าได้แม่นยำมากขึ้น ในขณะที่ Haar Cascade และ Haar + Canny มีอัตราการตรวจจับพลาดสูงจากข้อจำกัดในการแยกแยะใบหน้ากับพื้นหลังที่มีด



ภาพ 10 ตรวจจับใบหน้าสภาวะแสงมาก



ภาพ 11 ตรวจจับใบหน้าสภาวะปานกลาง



ภาพ 12 ตรวจจับใบหน้าสภาวะแสงน้อย

สรุปผลการวิจัย

งานวิจัยนี้ได้ทำการศึกษาและเปรียบเทียบประสิทธิภาพการตรวจจับใบหน้าภายใต้สภาวะแสงที่ต่างกัน (แสงมาก แสงปานกลาง และแสงน้อย) โดยใช้วิธี ได้แก่ Haar Cascade Haar Cascade ร่วมกับ Canny Edge Detection YOLO และ YOLO ร่วมกับ Canny Edge Detection พร้อมประเมินผลด้วยตัวชี้วัดมาตรฐาน ได้แก่ Accuracy Precision Recall F1-Score และเวลาในการประมวลผล

ผลการทดลองพบว่า YOLO + Canny ให้ผลลัพธ์ที่ดีที่สุดในทุกสภาวะแสง ทั้งในด้านความแม่นยำ (สูงสุด 92.0% ในสภาวะแสงปานกลาง) และยังคงมีความทนทานในสภาวะแสงน้อย ขณะเดียวกันยังมีเวลาในการประมวลผลสั้นที่สุด (<0.2 วินาที/ภาพ) ทำให้มีความเหมาะสมอย่างยิ่งสำหรับการใช้งานที่ต้องการความเร็วและความถูกต้องสูง ส่วน YOLO ปกติ ให้ผลใกล้เคียงกันในด้านความแม่นยำ แต่ใช้เวลามากกว่าอย่างมีนัยสำคัญ (~2 วินาที/ภาพ)

ในทางกลับกัน Haar Cascade แม้จะมีความเร็วในระดับปานกลาง แต่ความแม่นยำลดลงชัดเจน โดยเฉพาะในสภาวะแสงน้อย (Accuracy เพียง 66.6%) และการเพิ่ม Canny Edge Detection เข้ามาไม่ได้ช่วยปรับปรุงผลการตรวจจับ กลับเพิ่ม noise และลดประสิทธิภาพลงในบางกรณี

โดยสรุป การผสมผสาน YOLO และ Canny Edge Detection แสดงให้เห็นถึงความสามารถในการตรวจจับใบหน้าได้อย่างมีประสิทธิภาพและทนทานต่อการเปลี่ยนแปลงของแสง จึงเป็นวิธีที่มีศักยภาพสูงในการนำไปประยุกต์ใช้ในงานจริง ไม่ว่าจะเป็นระบบตรวจจับความปลอดภัย การเฝ้าระวัง หรือการยืนยันตัวตนที่ต้องการทั้งความแม่นยำและความรวดเร็ว

ผลการศึกษานี้ชี้ให้เห็นว่าความเหนือกว่าของวิธี YOLO ร่วมกับ Canny Edge Detection เกิดจากความสอดคล้องเชิงโครงสร้างระหว่างข้อมูลขอบภาพซึ่งมีความคงทนต่อการเปลี่ยนแปลงของแสง และสถาปัตยกรรมการเรียนรู้เชิงลึกของ YOLO ซึ่งช่วยลดความแปรปรวนจากความเข้มของแสงและข้อมูลพื้นหลังที่ไม่จำเป็น ส่งผลให้โมเดลสามารถมุ่งเน้นคุณลักษณะเชิงเรขาคณิตของใบหน้าได้อย่างมีประสิทธิภาพ ทั้งยังช่วยลดภาระการคำนวณและเวลาในการประมวลผล ในขณะที่ผลลัพธ์ของ Haar Cascade และ Haar + Canny สะท้อนข้อจำกัดของวิธีการตรวจจับแบบ feature-based ที่อ่อนไหวต่อ noise และสภาวะแสง โดยเฉพาะเมื่อข้อมูลขอบภาพไม่สอดคล้องกับกลไก cascade classifier โดยสรุปงานวิจัยนี้ตอกย้ำว่า การออกแบบระบบตรวจจับใบหน้าที่มีประสิทธิภาพสูงควรพิจารณาความสอดคล้องของทั้งขั้นตอน preprocessing โครงสร้างข้อมูล และโมเดลตรวจจับร่วมกัน เพื่อให้เกิดสมดุลที่เหมาะสมระหว่างความแม่นยำ ความทนทานต่อสภาวะแสง และความเร็วในการประมวลผลสำหรับการใช้งานจริง

กิตติกรรมประกาศ

คณะผู้วิจัยขอแสดงความขอบคุณอย่างสูงต่อ กอง
วิชาวิศวกรรมศาสตร์ ฝ่ายศึกษา โรงเรียนนายเรือ กองทัพ

เรือ ที่ได้กรุณาให้ความอนุเคราะห์ในการเอื้อเฟื้อทั้งอุปกรณ์
และสถานที่สำหรับการดำเนินการวิจัย ทำให้งานวิจัยเรื่องนี้
สำเร็จลุล่วงไปได้ด้วยดีตามวัตถุประสงค์ที่ตั้งไว้ทุกประการ



References

- Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934. <https://doi.org/10.48550/arXiv.2004.10934>
- Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6), 679–698. <https://doi.org/10.1109/TPAMI.1986.4767851>
- Hong, M., Cheng, S., Huang, H., Fan, H., & Liu, S. (2024). You only look around: Learning illumination-invariant feature for low-light object detection. arXiv:2410.18398. <https://doi.org/10.48550/arXiv.2410.18398>
- Jobson, D. J., Rahman, Z., & Woodell, G. A. (1997). A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transactions on Image Processing*, 6, 965-976. <https://doi.org/10.1109/83.597272>
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90. <https://doi.org/10.1145/3065386>
- Li, Y., & Deng, W. (2020). *Deep face recognition: A survey. Neurocomputing*, 429, 215–244. <https://doi.org/10.1016/j.neucom.2020.01.085>
- Lienhart, R., & Maydt, J. (2002). An extended set of Haar-like features for rapid object detection. In B. Mercer & L. Cotton (Eds.), *Proceedings of the 2002 International Conference on Image Processing (ICIP)* (pp. 900–903). IEEE. <https://doi.org/10.1109/ICIP.2002.1038171>
- Liu, Y., Li, S., Zhou, L., Liu, H., & Li, Z. (2025). Dark-YOLO: A low-light object detection algorithm integrating multiple attention mechanisms. *Applied Sciences*, 15(9), 5170. <https://doi.org/10.3390/app15095170>
- Maini, R., & Aggarwal, H. (2009). Study and comparison of various image edge detection techniques. *International Journal of Image Processing*, 3(1), 1–11. <https://www.cscjournals.org/manuscript/Journals/IJIP/Volume3/Issue1/IJIP-15.pdf>
- Mittal, P. (2024). A comprehensive survey of deep learning-based lightweight object detection models for edge devices. *Artificial Intelligence Review*, 57, 242. <https://doi.org/10.1007/s10462-024-10877-1>

- Pizer, S. M., Amburn, E. P., Austin, J. D., Cromartie, R., Geselowitz, A., Greer, T., Romeny, B. H., Zimmerman, J. B., & Zuiderveld, K. (1987). Adaptive histogram equalization and its variations. *Computer Vision, Graphics, and Image Processing*, 39(3), 355–368.
[https://doi.org/10.1016/S0734-189X\(87\)80186-X](https://doi.org/10.1016/S0734-189X(87)80186-X)
- Powers, D. M. W. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. arXiv:2010.16061. <https://doi.org/10.48550/arXiv.2010.16061>
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 779–788). IEEE.
- Sun, R., Lei, T., Chen, Q., Wang, Z., Du, X., Zhao, W., & Nandi, A. K. (2022). Survey of image edge detection. *Frontiers in Signal Processing*, 2, 826967. <https://doi.org/10.3389/frsip.2022.826967>
- Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (Vol. 1, pp. I–I). IEEE.
- Yisihak, H. M., & Li, L. (2024). Advanced face detection with YOLOv8: Implementation and integration into AI modules. *Open Access Library Journal*, 11(11), e12474.
<https://doi.org/10.4236/oalib.1112474>
- Zareiforush, H., Minaei, S., Alizadeh, M. R., & Banakar, A. (2016). An intelligent system design for classification of milled rice grains using metaheuristic methods. *Measurement*, 90, 350–361.
<https://doi.org/10.1016/j.measurement.2016.04.091>
- Zhang, C., & Zhang, Z. (2010). *A survey of recent advances in face detection* [Technical Report MSR-TR-2010-66]. Microsoft Research.
- Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint face detection and alignment using multi-task cascaded convolutional networks. *IEEE Signal Processing Letters*, 23(10), 1499-1503.
<https://doi.org/10.1109/lsp.2016.2603342>
- Zhou, Z., Cao, Z., & Pi, Y. (2017). A robust face detection method based on skin color and edge detection. *Multimedia Tools and Applications*, 76(7), 9711–9726.
<https://doi.org/10.1007/s11042-016-3507-3>

