

การตรวจจับภาพความจริงเสริมที่มีการเพิ่มเครื่องประดับบนใบหน้า

ณัฐชา ยิ้มฉนวน

TNI

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

บัณฑิตศึกษา สาขาเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีไทย-ญี่ปุ่น

ปีการศึกษา 2568

DETECTION OF AUGMENTED REALITY IMAGES WITH FACIAL ACCESSORIES

Natchaya Yimtanom

TNI

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Information Technology

Graduate Studies

Thai-Nichi Institute of Technology

Academic Year 2025

หัวข้อวิทยานิพนธ์

การตรวจจับภาพความจริงเสริมที่มีการเพิ่มเครื่องประดับบน
ใบหน้า

โดย

ณัฐชา ยิ้มถนอม

สาขาวิชา

เทคโนโลยีสารสนเทศ

อาจารย์ที่ปรึกษาวิทยานิพนธ์

ดร.สะพรั่งสิทธิ์ มฤตุสาร

บัณฑิตศึกษา สถาบันเทคโนโลยีไทย-ญี่ปุ่น อนุมัติให้บัณฑิตวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่ง
ของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

.....รองอธิการบดีฝ่ายวิชาการ
(รองศาสตราจารย์ ดร.วรากร ศรีเชวงทรัพย์)
วันที่.....เดือน.....พ.ศ.....

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ
(รองศาสตราจารย์ ดร.ทิมพันธุ์ เจริญพงษ์)

.....กรรมการ
(ดร.ประมุข บุญเสียง)

.....กรรมการ
(ว่าที่ร้อยตรี ดร.พิชิตชัย คำอินทร์)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์
(ดร.สะพรั่งสิทธิ์ มฤตุสาร)

ณัฐชา ยัมณอม : การตรวจจับภาพความจริงเสริมที่มีการเพิ่มเครื่องประดับบนใบหน้า
อาจารย์ที่ปรึกษา : ดร.สพรั่งสิทธิ์ มฤตสาธร, 118 หน้า.

เทคโนโลยีปัญญาประดิษฐ์และเทคโนโลยีความจริงเสริมได้เข้ามามีบทบาทและผสมผสานเข้ากับชีวิตประจำวัน แม้เทคโนโลยีเหล่านี้จะมีประโยชน์ในการสร้างสรรค์สื่อ แต่ในขณะเดียวกันยังมีความเสี่ยงในด้านการปลอมแปลงข้อมูล เช่น การสร้างภาพที่บิดเบือนจากความเป็นจริง ความสมจริงที่เพิ่มมากขึ้นของเครื่องประดับเสมือนในรูปแบบ AR ทำให้การแยกแยะระหว่างเครื่องประดับจริงและสิ่งที่ถูกเสริมเติมแต่งด้วยระบบดิจิทัลทำได้ยากขึ้น ซึ่งส่งผลให้เกิดความสับสนและการตีความเนื้อหาดิจิทัลที่ผิดพลาดอย่างมีนัยสำคัญ งานวิจัยฉบับนี้นำเสนอวิธีการต่าง ๆ ในการจำแนกความแตกต่างระหว่างภาพเครื่องประดับจริงและภาพที่ถูกปรับแต่งด้วย AR โดยผู้วิจัยได้นำเสนอวิธีการแบบใช้กฎ เพื่อกำหนดกฎเกณฑ์และเงื่อนไขในการตัดสินใจ ผ่านการแปลงโมเดลสีของภาพเป็นแบบ HSV เพื่อวิเคราะห์ค่าความสว่างและคำนวณค่าของคุณลักษณะที่กำหนด เช่น ค่าส่วนเบี่ยงเบนเฉลี่ย หรือส่วนเบี่ยงเบนมาตรฐาน นอกจากนี้งานวิจัยยังได้นำโมเดล Deep Learning ที่มีอยู่มาใช้ในการเปรียบเทียบประสิทธิภาพของแต่ละกระบวนการด้วยตาราง Confusion Matrix โดยใช้ชุดข้อมูลที่สร้างขึ้นจำนวน 2,000 ภาพ ประกอบด้วยภาพเครื่องประดับจริงและภาพแบบ AR สำหรับเครื่องประดับทั้งสองประเภท ผลการวิจัยพบว่า แม้โมเดลแบบใช้กฎจะมีประสิทธิภาพโดยสามารถบรรลุค่าความแม่นยำร้อยละ 75 สำหรับทั้งแว่นตาและหน้ากากอนามัย ซึ่งยืนยันได้ว่าคุณลักษณะของ HSV ที่เลือกมานั้นมีความเข้มแข็งเพียงพอในการจำแนกเครื่องประดับ AR อย่างไรก็ตาม ประสิทธิภาพดังกล่าวกลับน้อยกว่าการใช้ Machine Learning โดยโมเดลที่เหมาะสมที่สุดจะแตกต่างกันไปตามประเภทของเครื่องประดับ ทั้งนี้ โมเดล Xception สามารถบรรลุค่าความแม่นยำสูงสุดสำหรับแว่นตาที่ร้อยละ 88 และโดดเด่นที่สุดสำหรับหน้ากากอนามัยด้วยความแม่นยำสูงสุดถึงร้อยละ 98 ผลลัพธ์นี้บ่งชี้ว่าโครงสร้าง Depthwise Separable Convolutions ของ Xception มีความเหมาะสมอย่างยิ่งในการตรวจจับร่องรอยการดัดแปลงที่เกิดขึ้นในภาพที่ถูกปรับแต่งด้วย AR เมื่อเปรียบเทียบกับโมเดลอื่นๆ

บัณฑิตศึกษา

สาขาวิชา เทคโนโลยีสารสนเทศ

ปีการศึกษา 2568

ลายมือชื่อนักศึกษา

ลายมือชื่ออาจารย์ที่ปรึกษา

NATCHAYA YIMTANOM : DETECTION OF AUGMENTED REALITY IMAGES WITH FACIAL ACCESSORIES. ADVISOR : DR.SAPRANGSIT MRUETUSATORN, 118 PP.

Technologies such as Artificial Intelligence and Augmented Reality have become deeply integrated into daily life. These technologies provide benefits in media creation, but also have risks regarding data forgery, such as deceptive imagery. The increasing realism of AR accessories makes it challenging to differentiate between digital enhancements and real accessories. This ambiguity leads to significant confusion and misinterpretation of digital content. This research presents various methods to distinguish between images of real accessories and those enhanced with AR. We propose a rule-based approach that identifies custom decision rules and criteria by converting images into the HSV color model to analyze luminance and calculate values based on defined features, such as Mean Absolute Deviation or Standard Deviation. Furthermore, we also use existing Machine Learning (Deep Learning) models to compare the performance of each methodology using Confusion Matrix. The custom dataset was made with 2,000 images, including real and AR versions of both types of accessory images. The results demonstrate that while the rule-based model is effective, achieving 75% accuracy for both glasses and facial masks, it confirms that the selected HSV features are sufficiently robust for distinguishing AR accessories. Nevertheless, this performance is surpassed by machine learning approaches, where the optimal model varies by accessory type. The Xception model achieved the highest accuracy for glasses at 88% and outperformed others for facial masks with a peak accuracy of 98%. The results suggest that Xception's depthwise separable convolutions are particularly well-suited for detecting the artifacts inherent in AR-modified images compared to other models.

Graduate Studies

Student's Signature.....

Field of Study Information Technology

Advisor's Signature.....

Academic Year 2025

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความช่วยเหลืออย่างยิ่ง จากอาจารย์ที่ปรึกษา ดร.สะพรั่งสิทธิ์ มฤตุสาธร์ ผู้ให้ความรู้ คำชี้แนะ และความอนุเคราะห์ช่วยเหลือทุกด้าน รวมถึงตรวจทานและให้คำแนะนำ ในการแก้ไขข้อบกพร่องต่างๆ ตลอดระยะเวลาในการทำวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอกราบขอบพระคุณ ไว้ ณ โอกาสนี้

ขอกราบขอบพระคุณประธานและคณะกรรมการสอบวิทยานิพนธ์ รองศาสตราจารย์ ดร.ทิมพันธุ์ เจริญพงศ์ ดร.ประมุข บุญเสียง และว่าที่ร้อยตรี ดร.พิชิตชัย คำอินทร์ ผู้ให้คำแนะนำข้อแก้ไขที่ควรปรับปรุงจนวิทยานิพนธ์ฉบับนี้สมบูรณ์

ขอกราบขอบพระคุณคณาจารย์ทุกท่านจากคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยี ไทย-ญี่ปุ่น ผู้ประสิทธิ์ประสาทวิชาความรู้ รวมทั้งบุคลากรคณะเทคโนโลยีสารสนเทศ ผู้อำนวยการความสะดวกและช่วยเหลือในทุกด้าน

ขอขอบพระคุณเพื่อนร่วมงานและบุคลากรในการไฟฟ้านครหลวง ที่ได้สละเวลาและให้ความร่วมมือในการเป็นกลุ่มตัวอย่างเพื่อจัดเก็บข้อมูลภาพถ่ายที่ใช้ภายในวิทยานิพนธ์ฉบับนี้

เหนือสิ่งอื่นใด ผู้วิจัยขอกราบขอบพระคุณบิดามารดาของผู้วิจัยที่ทำให้กำลังใจ และให้การสนับสนุน ในทุกๆ ด้านเสมอมา

สุดท้ายนี้ผู้วิจัยหวังเป็นอย่างยิ่งว่า วิทยานิพนธ์ฉบับนี้จะเป็นประโยชน์ต่อผู้ที่มีความสนใจ เพื่อนำไปต่อยอดองค์ความรู้ และประยุกต์ใช้งานให้เกิดคุณค่าต่อไป

ณัฐชา ยิ้มถนอม

TNI

THAI - NICHU INSTITUTE OF TECHNOLOGY

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ฌ
สารบัญรูป.....	ญ
บทที่	
1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 โจทย์วิจัย.....	3
1.3 วัตถุประสงค์ของการวิจัย.....	3
1.4 ขอบเขตของการวิจัย.....	3
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	4
2 แนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง.....	5
2.1 ทฤษฎี และองค์ความรู้.....	5
2.2 งานวิจัยที่เกี่ยวข้อง.....	23
3 กระบวนการวิจัย.....	39
3.1 การวิเคราะห์ปัญหาของงานวิจัย.....	40
3.2 การศึกษาข้อมูลที่เกี่ยวข้อง.....	40
3.3 การพัฒนาโมเดลด้วย Rule-based.....	44
3.4 การพัฒนาโมเดลด้วย Machine Learning.....	54
3.5 การประเมินและวัดผลประสิทธิภาพโมเดล.....	58

สารบัญ (ต่อ)

บทที่	หน้า
4 ผลการวิจัย.....	60
4.1 ผลการหากฎเกณฑ์และเงื่อนไขสำหรับโมเดล Rule-based	60
4.2 ผลการตรวจจับภาพด้วยโมเดล Rule-based	62
4.3 ผลการตรวจจับภาพด้วยโมเดล Machine Learning.....	67
4.4 ผลการเปรียบเทียบประสิทธิภาพของโมเดล Rule-based และ Machine Learning.....	73
5 บทสรุป อภิปราย และข้อเสนอแนะ.....	78
5.1 สรุปผลการวิจัย.....	78
5.2 อภิปรายผล.....	80
5.3 ปัญหาและข้อจำกัด	82
5.4 ข้อเสนอแนะ	82
บรรณานุกรม	84
ภาคผนวก	89
ภาคผนวก ก. เอกสารการนำเสนอภาพนิ่งสอบป้องกันวิทยานิพนธ์	90
ประวัติย่อผู้วิจัย.....	118

สารบัญตาราง

ตาราง	หน้า
2.1 ตาราง Confusion Matrix.....	21
2.2 สรุปงานวิจัยที่มีหัวข้อเกี่ยวกับ AR	29
2.3 สรุปงานวิจัยที่มีหัวข้อเกี่ยวกับ Image Detection หรือ Image Forensics	30
3.1 เงื่อนไขการแยกแยะภาพเครื่องประดับจริงและเครื่องประดับ AR.....	53
4.1 เงื่อนไขการแยกแยะภาพสวมแว่นตา	61
4.2 ผลลัพธ์การแยกแยะภาพแว่นตาดำด้วยเงื่อนไขการแยกแยะภาพสวมแว่นตา.....	61
4.3 เงื่อนไขการแยกแยะภาพสวมหน้ากาก.....	62
4.4 ผลลัพธ์การแยกแยะภาพหน้ากากด้วยเงื่อนไขการแยกแยะภาพสวมหน้ากาก	62
4.5 ผลลัพธ์การแยกแยะภาพแว่นตาชุดทดสอบ ด้วยเงื่อนไขการแยกแยะภาพสวมแว่นตา.....	63
4.6 ผลการคำนวณค่าจาก Confusion Matrix ภาพสวมแว่นตา	64
4.7 ผลการแยกแยะภาพแว่นตาของแต่ละคุณลักษณะตามประเภทภาพ	64
4.8 ผลลัพธ์การแยกแยะภาพหน้ากากชุดทดสอบ ด้วยเงื่อนไขการแยกแยะภาพสวม หน้ากาก	65
4.9 ผลการคำนวณค่าจาก Confusion Matrix ภาพสวมหน้ากาก.....	66
4.10 ผลการแยกแยะภาพหน้ากากของแต่ละคุณลักษณะตามประเภทภาพ	66
4.11 ผลการคำนวณค่าจาก Confusion Matrix โมเดล Xception (ภาพแว่นตา)	68
4.12 ผลการคำนวณค่าจาก Confusion Matrix โมเดล MobileNetV3 (ภาพแว่นตา).....	69
4.13 ผลการคำนวณค่าจาก Confusion Matrix โมเดล EfficientNetV2S (ภาพแว่นตา).....	70
4.14 ผลการคำนวณค่าจาก Confusion Matrix โมเดล Xception (ภาพหน้ากาก).....	71
4.15 ผลการคำนวณค่าจาก Confusion Matrix โมเดล MobileNetV3 (ภาพหน้ากาก).....	72
4.16 ผลการคำนวณค่าจาก Confusion Matrix โมเดล EfficientNetV2S (ภาพหน้ากาก)	73
4.17 นิยามค่าในการวัดประสิทธิภาพของโมเดล.....	74
4.18 ผลการเปรียบเทียบประสิทธิภาพของโมเดลในการตรวจจับภาพแว่นตา.....	75
4.19 ผลการเปรียบเทียบประสิทธิภาพของโมเดลในการตรวจจับภาพหน้ากาก	75

สารบัญรูป

รูป	หน้า
2.1 ภาพตัวอย่างการใช้เทคโนโลยี AR เป็นตัวช่วยในการฝึกอบรมผู้ใช้งาน (A) ภาพมุมมองของผู้ฝึกอบรม พร้อมกับอุปกรณ์เสมือนจริง (B) ห้องปฏิบัติการที่ติดตั้งกล้องจับการเคลื่อนไหว 24 ตัว และผู้เข้าร่วม 2 คนที่สวมอุปกรณ์ Immersive AR (C) ไม่ควบคุม แสดงด้วยตัวกระตุ้นสีเขียว.....	6
2.2 ภาพแสดงองค์ประกอบของ HSV Color Model	13
2.3 ประเภทการเรียนรู้ของ Machine Learning.....	19
3.1 แผนผังภาพรวมขั้นตอนดำเนินงานวิจัย.....	39
3.2 ตัวอย่างภาพใบหน้าสวมหน้ากาก (A) ภาพสวมหน้ากากจริง (B) ภาพสวมหน้ากาก AR.....	40
3.3 ตัวอย่างภาพใบหน้าสวมแว่นตา (A) ภาพสวมแว่นตาจริง (B) ภาพสวมแว่นตา AR.....	41
3.4 ภาพรวมข้อมูลภาพภายใน Dataset ที่ใช้ฝึกโมเดล Machine Learning.....	41
3.5 การแบ่งสัดส่วนข้อมูลภาพภายใน Dataset ตามขั้นตอนการพัฒนาโมเดล	43
3.6 แผนผังขั้นตอนการพัฒนาโมเดลด้วย Rule-based	44
3.7 แผนผังขั้นตอนการกำหนดกฎเกณฑ์และเงื่อนไขการตรวจจับภาพ (แว่นตา).....	46
3.8 (A) ภาพแสดงการเลือกบริเวณที่มีความสนใจ 3 จุดในภาพ (B) การกำหนด Mask สำหรับระบุพื้นที่ในการเก็บค่า Value ในภาพ	47
3.9 แผนผังขั้นตอนการกำหนดกฎเกณฑ์และเงื่อนไขการตรวจจับภาพ (หน้ากาก)	48
3.10 ตัวอย่างการเก็บค่าแสงสว่างทั้ง 3 จุดในแต่ละภาพ.....	48
3.11 ตัวอย่างการหาค่าที่ใช้เป็นเกณฑ์ของค่าส่วนเบี่ยงเบนเฉลี่ยสัมบูรณ์ (Mean AD).....	51
3.12 แผนผังขั้นตอนการพัฒนาโมเดลด้วย Machine Learning.....	55
3.13 แผนผังขั้นตอนการฝึกโมเดลด้วย Machine Learning.....	57
4.1 ผลลัพธ์การสร้าง Confusion Matrix ภาพสวมแว่นตา	63
4.2 ผลลัพธ์การสร้าง Confusion Matrix ภาพสวมหน้ากาก.....	65
4.3 ผลลัพธ์ Confusion Matrix ภาพสวมแว่นตาของโมเดล Xception	68
4.4 ผลลัพธ์ Confusion Matrix ภาพสวมแว่นตาของโมเดล MobileNetV3.....	69
4.5 ผลลัพธ์ Confusion Matrix ภาพสวมแว่นตาของโมเดล EfficientNetV2S.....	70
4.6 ผลลัพธ์ Confusion Matrix ภาพสวมหน้ากากของโมเดล Xception.....	71

สารบัญรูป (ต่อ)

รูป	หน้า
4.7	ผลลัพธ์ Confusion Matrix ภาพสวมหน้ากากของโมเดล MobileNetV3 72
4.8	ผลลัพธ์ Confusion Matrix ภาพสวมหน้ากากของโมเดล EfficientNetV2S 73



บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันนี้เป็นยุคสมัยที่มีการใช้เทคโนโลยีต่างๆ ในการใช้ชีวิตประจำวันมากขึ้นจนกลายเป็นเรื่องทั่วไป ซึ่งสร้างผลกระทบในหลายๆ ด้าน เช่น ด้านการค้นหาข้อมูลที่สามารถใช้อินเทอร์เน็ตช่วยให้ค้นหาข้อมูลได้รวดเร็ว แต่ในทางกลับกัน อาจหมายถึงความรวมถึงการนำเข้าสู่ข้อมูลเท็จที่ผู้คนสามารถค้นหาได้อย่างอิสระเช่นกัน โดยเฉพาะอย่างยิ่ง ในช่วงหลายปีที่ผ่านมา เทคโนโลยีหลายอย่างถูกพัฒนามากขึ้นหนึ่งในเทคโนโลยีที่มีคนให้ความสนใจเป็นจำนวนมากคือ เทคโนโลยี AI (Artificial Intelligence) หรือปัญญาประดิษฐ์ที่สามารถสร้างสรรค์ข้อมูลประเภทต่างๆ เช่น การใช้เทคโนโลยี Generative AI ในการช่วยสร้างข้อมูลประเภทข้อความ วิดีโอ หรือรูปภาพที่สามารถสร้างออกมาได้สมจริง ซึ่ง Generative AI นี้เป็นหัวข้อที่ถูกศึกษากันอย่างแพร่หลาย ยกตัวอย่างเช่น งานวิจัยของ Ana Elisa Sousa และ Paula Cardoso [1] ได้ทำการศึกษาเกี่ยวกับรูปแบบการนำเทคโนโลยีนี้ไปใช้งาน รวมถึงมุมมอง ประโยชน์ ความท้าทายของ Generative AI ที่มีต่อการนำไปใช้งานวิชาการของนักศึกษา เป็นต้น และด้วยความสามารถในการสร้างข้อมูลหลากหลายประเภทของ Generative AI นี้ อาจเป็นต้นเหตุที่ทำให้เกิดนำเข้าสู่ข้อมูลเท็จเข้าสู่ระบบอินเทอร์เน็ตได้ ซึ่งหากนำข้อมูลเหล่านี้เข้าสู่ระบบอินเทอร์เน็ตโดยไม่ผ่านการกลั่นกรอง อาจมีผู้ไม่หวังดีนำข้อมูลและเทคโนโลยีเหล่านี้ไปใช้เพื่อหาผลประโยชน์ส่วนตัว ทำให้เกิดปัญหาและความไม่สงบในสังคมได้หลากหลายรูปแบบ เช่น การปลอมแปลงตัวตนโดยใช้รูปบุคคลอื่นเพื่อหลอกเอาทรัพย์สิน หรือการสร้างข้อมูลข่าวปลอมที่ทำให้เกิดความเข้าใจผิด และการเลียนเสียงของบุคคลที่มีชื่อเสียงด้วย Generative AI เป็นต้น

นอกจากนี้ยังมีเทคโนโลยี AR (Augmented Reality) หรือเทคโนโลยีที่นำวัตถุ 3 มิติ มาจำลองและแสดงในโลกจริง โดยใช้เซนเซอร์ตรวจจับภาพหรือเสียง แล้วสร้างภาพ 3 มิติตามเงื่อนไขด้วยการประมวลผลจากซอฟต์แวร์ ซึ่งสามารถนำไปใช้งานได้อย่างหลากหลาย ยกตัวอย่างเช่น งานวิจัยของทรงธรรม สุขธานี และคณะ [2] ที่กล่าวถึงการใช้เทคโนโลยี AR ในการพัฒนาสื่อการสอนการเต้นโยนประกอบคำพากย์สำหรับนักเรียน หรือธุรกิจการขายเครื่องประดับของ The Banter by Piercing Pagoda [3] ที่ให้ลูกค้าสามารถอัปโหลดภาพใบหน้าของตัวเอง และเลือกทดลองใส่เครื่องประดับที่แสดงภาพด้วยเทคโนโลยี AR ได้โดยอิสระ เทคโนโลยี AR เป็นหนึ่งในตัวเลือกที่เพิ่มทางเลือกให้กับหลากหลายอุตสาหกรรม เช่น อุตสาหกรรมการผลิตที่สามารถใช้เทคโนโลยี AR เป็นตัวช่วยในการฝึกอบรมพนักงานได้ เป็นต้น

ในขณะเดียวกัน นอกจากข้อดีหลายอย่างในการใช้งานแล้ว ยังมีความเสี่ยงหลายอย่างที่อาจเกิดขึ้นหากนำเทคโนโลยีไปใช้ในทางที่ไม่ถูกต้อง เช่น เทคโนโลยี AI ที่สามารถสร้างภาพใบหน้าที่ไม่ได้อยู่จริง หรือทำการเลียนเสียงของบุคคลที่มีชื่อเสียง [4] และเทคโนโลยี AR ที่อาจทำให้เกิดความเข้าใจผิดได้

เช่น กรณีของธุรกิจขายเครื่องประดับที่กล่าวไปข้างต้น หากมีผู้ไม่ประสงค์ดีทำการสร้างภาพที่ตนสวมใส่เครื่องประดับที่ไม่ได้มีอยู่จริง และนำภาพนั้นไปหลอกลขายผู้อื่น จะทำให้เกิดความสับสนได้ ซึ่งวิธีการป้องกันเหตุการณ์เหล่านี้ในปัจจุบัน มีเครื่องมือที่สามารถตรวจจับภาพที่ถูกสร้างขึ้นจากเทคโนโลยี AI อยู่หลายเครื่องมือที่สามารถใช้เป็นตัวช่วยในการคัดกรองข้อมูลได้ ซึ่งแต่ละเครื่องมือสามารถเข้าถึงได้ง่ายผ่านระบบอินเทอร์เน็ต และถูกพัฒนาขึ้นจนมีความสามารถในการตรวจจับได้ค่อนข้างถูกต้องและแม่นยำในบางเครื่องมือ แต่สำหรับภาพที่ถูกปรับแต่งด้วยเทคโนโลยี AR นั้น ยังไม่มีเครื่องมือเฉพาะทางที่ใช้ในการตรวจจับแบบอัตโนมัติโดยเฉพาะ ซึ่งหากใช้เครื่องมือตรวจจับภาพ AI หรือ Large Language Models (LLMs) เช่น ChatGPT, Gemini หรือ Claude ในการตรวจจับ ผลลัพธ์การตรวจจับที่ได้อาจยังไม่มีความแม่นยำมากนัก ซึ่งจากการทดลองเบื้องต้น LLMs แต่ละแบบให้ผลลัพธ์ในการตรวจจับที่ต่างกัน แม้จะใช้รูปภาพเดียวกันและคำถามเดียวกัน ซึ่งทดลองโดยการส่งรูปภาพที่มีการใช้ AR ภาพเดียวกัน และใช้คำถามเดียวกันให้ ChatGPT, Gemini และ Claude ทำการตรวจจับภาพ ผลปรากฏว่า ChatGPT และ Claude ไม่สามารถตรวจจับภาพ AR และให้ผลลัพธ์ที่ถูกต้องได้ แต่ Gemini สามารถตรวจจับภาพ AR ได้ถูกต้อง แสดงถึงความสามารถในการตรวจจับภาพ AR ที่อาจยังไม่แม่นยำ และยังไม่มีความเสถียรในเครื่องมือที่นิยมใช้อย่าง LLMs ในปัจจุบัน

ผู้วิจัยได้ตระหนักถึงปัญหาในการตรวจสอบภาพที่ผ่านการปรับแต่งด้วยเทคโนโลยีเหล่านี้ ซึ่งโดยทั่วไปผู้คนมักจะใช้เครื่องมือสำหรับตรวจสอบภาพด้วย AI เพื่อตรวจสอบว่าภาพเหล่านั้นผ่านการปรับแต่งมาหรือไม่ แต่เนื่องจากผู้วิจัยพบว่า เครื่องมือตรวจสอบภาพด้วย AI ที่มีอยู่บนระบบอินเทอร์เน็ตในปัจจุบัน ยังไม่สามารถตรวจสอบภาพที่ผ่านการปรับแต่งด้วยเทคโนโลยี AR ได้มีประสิทธิภาพและแม่นยำมากเพียงพอ ผู้วิจัยจึงมีแนวคิดที่จะศึกษาทฤษฎีการตรวจจับภาพที่ถูกแต่งเติมด้วยเทคโนโลยี AR และพัฒนาเครื่องมือสำหรับการตรวจจับภาพประเภทนี้โดยเฉพาะ เพื่อตรวจสอบเบื้องต้นว่าภาพนั้นถูกแต่งเติมด้วยเทคโนโลยี AR หรือไม่ และใช้เป็นข้อมูลในการวิเคราะห์ความน่าเชื่อถือของภาพ และเนื่องจากเทคโนโลยี AR ในการใช้งานโดยบุคคลทั่วไป มักจะเป็นการใช้งานฟิลเตอร์ AR ผ่านแอปพลิเคชันถ่ายภาพต่างๆ เพื่อถ่ายภาพใบหน้าของตน ทำให้ผู้วิจัยมีความตั้งใจที่จะเริ่มศึกษาจากภาพใบหน้าบุคคล โดยเฉพาะภาพใบหน้าบุคคลที่มีการเพิ่มส่วนที่เป็นเครื่องประดับบนใบหน้า เป็นเป้าหมายสำคัญ เพื่อเป็นจุดเริ่มต้นในการศึกษาต่อไป

ในงานวิจัยนี้ ผู้วิจัยมุ่งเน้นศึกษาการตรวจจับเครื่องประดับบนใบหน้า 2 ประเภทหลัก คือ แว่นสายตา และหน้ากากอนามัย เนื่องจากเป็นอุปกรณ์ที่มีการสวมใส่จริงอย่างแพร่หลายในชีวิตประจำวัน โดยแว่นสายต้ามักถูกใช้งานทั้งในเชิงสุขภาพเพื่อแก้ไขทัศนวิสัยและในเชิงแฟชั่น ขณะที่หน้ากากอนามัยกลายเป็นอุปกรณ์พื้นฐานในด้านสุขอนามัย เพื่อป้องกันมลภาวะและลดการแพร่กระจายของเชื้อโรค นอกจากความสำคัญในโลกจริงแล้ว เครื่องประดับทั้งสองชนิดนี้ยังมีการใช้เป็นฟิลเตอร์ AR ที่ถูกนำมาใช้ในการตกแต่งภาพถ่ายผ่านแอปพลิเคชันโซเชียลมีเดีย ผู้วิจัยจึงเห็นว่าการเริ่มต้นศึกษาจากวัตถุที่มีความใกล้ชิดกับพฤติกรรมของผู้ใช้งานทั้งในโลกจริงและโลกดิจิทัล จะช่วยให้ผลการวิจัยสามารถนำไปประยุกต์ใช้เพื่อเพิ่มความน่าเชื่อถือในการตรวจสอบภาพบุคคลได้อย่างมีประสิทธิภาพ

ทิศทางของงานวิจัยนี้มุ่งเน้นไปที่การพัฒนาแนวทางในการตรวจจับภาพ โดยมีการเปรียบเทียบระหว่างการพัฒนาโมเดลตรวจจับภาพด้วยวิธี Rule-based ที่สามารถอธิบายเงื่อนไขและเหตุผลที่ใช้ในการทำนายผลลัพธ์ได้ ควบคู่ไปกับการพัฒนาโมเดลตรวจจับภาพด้วย Machine Learning (Deep Learning) ที่มีโครงสร้างซับซ้อนเพื่อขยายขีดความสามารถในการตรวจจับภาพ โดยงานวิจัยนี้จำกัดขอบเขตการศึกษาอยู่ที่เครื่องประดับบนใบหน้า เฉพาะแว่นตาและหน้ากาก เพื่อเป็นจุดเริ่มต้นในการพัฒนาโมเดลสำหรับตรวจจับภาพเครื่องประดับ AR บนใบหน้า

1.2 โจทย์วิจัย

ภาพใบหน้าที่ไม่สามารถแยกแยะได้อย่างแน่นอนว่าถูกแต่งเติมเครื่องประดับบนใบหน้าด้วยเทคโนโลยี AR หรือไม่ ทำให้เกิดความสับสนและเกิดความเข้าใจที่คลาดเคลื่อน

1.3 วัตถุประสงค์ของการวิจัย

เพื่อพัฒนาวิธีการตรวจจับภาพความจริงเสริม (Augmented Reality) ที่มีการเพิ่มเครื่องประดับในบริเวณใบหน้า ควบคู่กับการทดสอบด้วยโมเดล Machine Learning ที่ทำการเลือกไว้ และเปรียบเทียบประสิทธิภาพของแต่ละวิธีการตรวจจับ

1.4 ขอบเขตของการวิจัย

1.4.1 ประเภทข้อมูลภาพ

1.4.1.1 ภาพถ่ายใบหน้าบุคคล

1.4.1.2 ภาพสี

1.4.2 แหล่งที่มาของข้อมูลภาพ

1.4.2.1 ภาพถ่ายใบหน้าที่จริงจาก CelebA [5]

1.4.2.2 ภาพถ่ายใบหน้าที่จริงจาก Real-Time-Medical-Mask-Detection [6]

1.4.2.3 ภาพถ่ายจากกล้องโทรศัพท์มือถือ

1.4.3 ประเภทของเครื่องประดับบนใบหน้า

1.4.3.1 แว่นสายตา

1.4.3.2 หน้ากากอนามัย (ที่ไม่มีลวดลาย)

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1.5.1 ได้องค์ความรู้ใหม่เกี่ยวกับการพัฒนาวิธีการตรวจจับภาพใบหน้าที่มีลักษณะเฉพาะ เช่น มีการเพิ่มเติมเครื่องประดับ AR บนภาพ

1.5.2 ได้แนวทางการประยุกต์ใช้โมเดลการตรวจจับภาพกับการพัฒนาระบบด้านความปลอดภัย

1.5.3 ได้ข้อมูลเปรียบเทียบประสิทธิภาพของเทคนิคต่างๆ ในการตรวจจับใบหน้าที่มีเครื่องประดับ ปกติ และมีเครื่องประดับจากภาพ AR



บทที่ 2

แนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง

ในการวิจัยนี้ ผู้วิจัยได้ศึกษาแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้องจากบทความวิชาการ และงานวิจัยจากแหล่งที่น่าเชื่อถือ ตามหัวข้อดังต่อไปนี้

2.1 ทฤษฎี และองค์ความรู้

2.2 งานวิจัยที่เกี่ยวข้อง

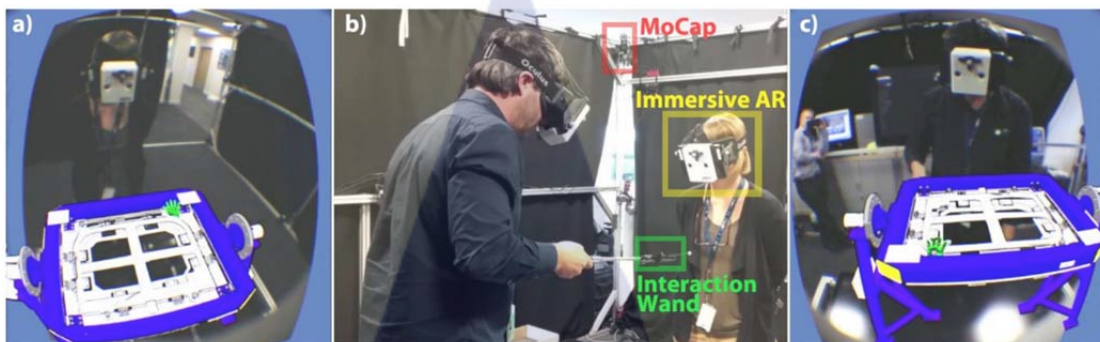
2.1 ทฤษฎี และองค์ความรู้

2.1.1 ความจริงเสริม

ความจริงเสริม (Augmented Reality: AR) คือ เทคโนโลยีที่ผสมผสานระหว่างข้อมูล ดิจิทัล เช่น ภาพ 3 มิติ หรือ วิดีโอ เข้ากับสภาพแวดล้อมในโลกแห่งความจริงที่มองเห็นผ่านอุปกรณ์ อิเล็กทรอนิกส์ประเภทต่างๆ เช่น สมาร์ทโฟน หรือแว่นตาสำหรับ AR โดยเฉพาะ เพื่อเสริมสร้างการรับรู้ ข้อมูลของผู้ใช้งานต่อข้อมูลที่ตอบสนองต่อการเปลี่ยนแปลงของสภาพแวดล้อมได้ โดยเทคโนโลยี AR จะ ใช้กล้องจากอุปกรณ์อิเล็กทรอนิกส์ หรือเซ็นเซอร์ เพื่อรับข้อมูลจากความเป็นจริงมายังอุปกรณ์ และแสดงผล ด้วยการแสดงข้อมูลซ้อนทับลงในสภาพแวดล้อมในชีวิตจริง

กระบวนการทำงานของเทคโนโลยี AR โดยเบื้องต้นนั้น เริ่มจากการใช้อุปกรณ์อิเล็กทรอนิกส์ ที่มีกล้อง หรืออุปกรณ์ตรวจจับการเคลื่อนไหว (Motion Detection) เพื่อทำการจับภาพวัตถุหรือสภาพแวดล้อมในโลกแห่งความจริงรอบๆ ตัวผู้ใช้งาน โดยเมื่อรับข้อมูลสภาพแวดล้อมโดยรอบแล้ว จะทำการ ใช้ซอฟต์แวร์ในการสร้างข้อมูลเป็นภาพจำลอง 3 มิติ หรือเป็นข้อมูลตัวอักษรต่างๆ เชื่อมต่อและเก็บข้อมูล ไว้บนระบบคลาวด์ (เทคโนโลยีที่ผู้ใช้สามารถจัดเก็บข้อมูลและซอฟต์แวร์ได้ผ่านอินเทอร์เน็ต) จากนั้นใช้ อุปกรณ์ AR ในการดาวน์โหลดข้อมูลจากระบบคลาวด์ เพื่อนำมาแสดงบนภาพวัตถุหรือสภาพแวดล้อมจริงที่จับภาพไว้ โดยที่ผู้ใช้งานสามารถโต้ตอบกับข้อมูลเหล่านั้นได้ทันทีด้วยการแตะหน้าจอ การพูด หรือ ใช้ท่าทางสื่อสารต่างๆ ซึ่งแตกต่างกันไปตามอุปกรณ์ AR ที่ใช้งาน

เทคโนโลยี AR นี้สามารถนำไปประยุกต์ใช้เพื่อให้เกิดประโยชน์ได้หลากหลายวิธี โดยมี งานวิจัยหลายงานที่นำเทคโนโลยี AR ไปใช้ในการพัฒนาด้านต่างๆ เช่น ใช้เป็นส่วนประกอบของระบบ ช่วยเหลืองานซ่อมวัตถุส่วนประกอบที่ซับซ้อน ซึ่งมีการนำเทคโนโลยี AR มาใช้แทนการใช้เอกสารกระดาษรูปแบบเดิม [7] หรือใช้เป็นตัวช่วยในการฝึกอบรมเกี่ยวกับการผลิตที่ซับซ้อน โดยให้ผู้รับการฝึกอบรมใช้ อุปกรณ์ AR ในการฝึก [8] ดังที่แสดงในรูปที่ 2.1



รูปที่ 2.1 ภาพตัวอย่างการใช้เทคโนโลยี AR เป็นตัวช่วยในการฝึกอบรมผู้ใช้งาน (A) ภาพมุมมองของผู้ฝึกอบรม พร้อมกับอุปกรณ์เสมือนจริง (B) ห้องปฏิบัติการที่ติดตั้งกล้องจับการเคลื่อนไหว 24 ตัว และผู้เข้าร่วม 2 คนที่สวมอุปกรณ์ Immersive AR (C) ไม่ควบคุม แสดงด้วยตัวกระตุ้นสีเขียว [8]

นอกจากนี้ ในปัจจุบันมีการนำเทคโนโลยี AR มาใช้ในภาคอุตสาหกรรมต่างๆ ยกตัวอย่างเช่น การออกแบบผลิตภัณฑ์ด้วยภาพจำลอง 3 มิติ เพื่อให้ผู้ใช้งานสามารถตรวจสอบรายละเอียดของวัตถุในทุกมุมได้ก่อนการผลิตจริง การควบคุมคุณภาพของสินค้าด้วยการใช้แว่นตา AR ในการตรวจสอบข้อมูลสินค้าและรับแจ้งเตือนเมื่อพบจุดบกพร่องในตัวสินค้า หรือการใช้เทคโนโลยี AR ในการเปลี่ยนสภาพแวดล้อมจริงให้เหมาะสมกับการฝึกอบรมด้านต่างๆ ที่จำเป็นต้องใช้อุปกรณ์หรือสถานที่ในการฝึก ซึ่งสามารถเพิ่มความปลอดภัยระหว่างการฝึก และเพิ่มประสิทธิภาพในการฝึกได้

ข้อดีของเทคโนโลยี Augmented Reality

1. ช่วยให้ผู้ใช้งานสามารถเข้าใจข้อมูลได้ง่ายขึ้นด้วยภาพ 3 มิติที่แสดงในโลกแห่งความจริง และความสามารถในการตอบโต้กับภาพได้ทันที
2. เพิ่มประสิทธิภาพในการทำงาน กรณีที่ใช้กับงานประเภทการออกแบบสินค้า หรืองานที่จำเป็นต้องเห็นภาพวัตถุจริงก่อนการผลิต
3. ลดต้นทุนและเพิ่มความปลอดภัยเมื่อใช้กับงานประเภทการฝึกอบรม หรือการฝึกฝนทักษะที่จำเป็นต้องใช้อุปกรณ์และสถานที่จริงในการฝึกฝน
4. สามารถเข้าถึงข้อมูลได้ง่ายหากเก็บข้อมูลไว้บนอินเทอร์เน็ต เมื่อใช้ร่วมกับอุปกรณ์ที่รองรับการแสดงผลข้อมูล AR
5. สามารถประยุกต์ใช้เทคโนโลยีได้กับหลายประเภทงาน เช่น อุตสาหกรรมการผลิต การศึกษา หรือการตลาด เป็นต้น

ข้อเสียของเทคโนโลยี Augmented Reality

1. จำเป็นต้องมีอุปกรณ์ที่รองรับการใช้เทคโนโลยี AR เช่น แว่นตา AR หรือสมาร์ทโฟน
2. จำเป็นต้องเชื่อมต่อกับอินเทอร์เน็ต ทำให้อาจมีข้อจำกัดในกรณีที่ใช้งานในสถานที่หรือบริเวณที่ไม่มีอินเทอร์เน็ตรองรับ หรือความเร็วของอินเทอร์เน็ตต่ำกว่าปกติ
3. ผู้ใช้งานต้องมีความรู้และความเข้าใจในเทคโนโลยีในระดับหนึ่ง จึงจะสามารถใช้เทคโนโลยี AR ร่วมกับอุปกรณ์ต่างๆ ได้อย่างถูกต้องและมีประสิทธิภาพ
4. อาจมีความเสี่ยงในด้านความปลอดภัยของข้อมูล เนื่องจากเทคโนโลยี AR จำเป็นต้องใช้กล้องเพื่อจับภาพ หากซอฟต์แวร์ที่ใช้งานมีความปลอดภัยต่ำ อาจถูกเจาะระบบเพื่อนำข้อมูลออกไปภายนอกได้

ในงานวิจัยนี้ ผู้วิจัยมีการนำเทคโนโลยี AR มาใช้งานในส่วนของการเพิ่มเครื่องประดับที่เป็น AR บนใบหน้า เพื่อทำการทดลองและพัฒนาโมเดลที่สามารถตรวจจับและแยกแยะภาพดังกล่าวกับภาพใบหน้าที่สวมเครื่องประดับตามปกติได้

2.1.2 การประมวลผลข้อมูลภาพดิจิทัล

การประมวลผลข้อมูลภาพดิจิทัล (Digital Image Processing) คือ กระบวนการที่ใช้คอมพิวเตอร์ในการวิเคราะห์ แก้ไข หรือปรับปรุงข้อมูลภาพเพื่อให้ได้ผลลัพธ์ที่ต้องการ โดยการประมวลผลข้อมูลภาพดิจิทัลนั้นจะเกี่ยวข้องกับการปรับปรุงข้อมูลภายในภาพ เช่น การลดสัญญาณรบกวนในภาพ (Noise Reduction) การปรับคอนทราสต์ของภาพ (Contrast Enhancement) หรือการปรับความคมชัดของภาพ (Image Sharpening) เป็นต้น ซึ่งการประมวลผลข้อมูลภาพดิจิทัลนี้ถือเป็นกระบวนการระดับต่ำ (Low-level Process) ที่เน้นเฉพาะการปรับปรุงคุณภาพของภาพเท่านั้น โดยนอกจากการประมวลผลข้อมูลภาพดิจิทัลแล้ว ยังมีกระบวนการระดับกลางอย่างการสกัดข้อมูลจากภาพ (Image Analysis) ที่เป็นกระบวนการสำหรับการสกัดข้อมูลบางอย่างออกมาจากภาพ เช่น การตรวจจับขอบของภาพ (Edge Detection) หรือการแบ่งส่วนของภาพ (Image Segmentation) เพื่อให้สามารถแบ่งแยกข้อมูลภายในภาพได้ง่ายขึ้น และมีกระบวนการระดับสูงอย่างคอมพิวเตอร์ วิทัศน์ (Computer Vision) ที่เป็นกระบวนการสำหรับให้คอมพิวเตอร์สามารถทำความเข้าใจภาพได้โดยอัตโนมัติ เช่น การจำแนกวัตถุ (Object Detection) หรือการจดจำใบหน้า (Face Recognition) เป็นต้น

ขั้นตอนของการประมวลผลข้อมูลภาพดิจิทัลนั้น สามารถแบ่งได้เป็น 10 ขั้นตอนหลักๆ ดังต่อไปนี้

1. การนำเข้าข้อมูลภาพ (Image Acquisition) คือ กระบวนการในการนำภาพเข้าสู่ระบบคอมพิวเตอร์ โดยการใช้อุปกรณ์ที่สามารถจับภาพได้ เช่น การใช้กล้องถ่ายภาพ การสแกนภาพเข้าสู่ระบบคอมพิวเตอร์ หรือการใช้กล้องประเภทพิเศษ เช่น กล้องความร้อน (Thermal Camera) เป็นต้น

2. Image Filtering & Emhancement คือ กระบวนการในการกรองข้อมูลภาพเพื่อปรับปรุงคุณลักษณะของภาพ เช่น การลดสัญญาณรบกวนในภาพ หรือการปรับภาพให้เบลอ การปรับภาพให้คมชัดขึ้น ตามวัตถุประสงค์ที่ต้องการ

3. Image Restoration คือ กระบวนการกู้ภาพที่มีจุดเสียหาย เช่น ใช้กระบวนการเพื่อแสดงจุดในภาพที่ไม่สามารถมองเห็นได้ด้วยตาเปล่าให้มองเห็นได้ รวมไปถึงการปรับภาพที่เบลอให้คมชัดขึ้น (Image Sharpening) เป็นต้น

4. Color Image Processing คือ กระบวนการประมวลผลสีในภาพ เพื่อปรับปรุงคุณภาพของสี เช่น การปรับความสว่าง การปรับคอนทราสต์ในภาพ เป็นต้น

5. Wavelets and Multiresolution Processing คือ กระบวนการประมวลผลภาพในระดับความละเอียดภาพ โดยสามารถปรับปรุงภาพในระดับของรายละเอียดความถี่ของภาพได้

6. Compression คือ กระบวนการบีบอัดภาพ เพื่อลดความซ้ำซ้อนของข้อมูลในส่วนที่มีความถี่ต่ำ (เช่น ส่วนที่เป็นพื้นหลังของภาพ)

7. Morphological Processing คือ กระบวนการในการปรับข้อมูลบางส่วนในภาพเพื่อให้เห็นรูปร่าง (Shape) ของส่วนประกอบในภาพได้ชัดเจนขึ้น เช่น การเพิ่มขนาดของตัวอักษรเพื่อเชื่อมเส้นอักษรให้ต่อกัน ทำให้ภาพมีความสมบูรณ์มากขึ้น เป็นต้น

8. Segmentation คือ กระบวนการแบ่งส่วนภาพ เพื่อแบ่งข้อมูลเฉพาะส่วนที่ต้องการออกมา เช่น การแบ่งขอบของภาพ หรือการแยกวัตถุออกจากพื้นหลัง

9. Representation & Description คือ กระบวนการแสดงข้อมูลภาพด้วยรหัส หรือรูปแบบข้อมูลที่สร้างขึ้นโดยเฉพาะ เช่น แผนภาพข้อมูล

10. Recognition and Interpretation คือ กระบวนการระบุความหมายของภาพ เป็นการให้คอมพิวเตอร์ทำการระบุความหมายของวัตถุที่อยู่ภายในภาพ ซึ่งนับเป็นกระบวนการระดับสูง หรือคอมพิวเตอร์วิทัศน์ด้วยเช่นกัน

การประมวลผลข้อมูลภาพดิจิทัลสามารถนำไปประยุกต์ใช้งานได้หลากหลายด้าน โดยเฉพาะในด้านที่จำเป็นต้องมีการวิเคราะห์ข้อมูลภาพ เช่น งานวิจัยของ W. Suklim et al. [9] ได้กล่าวถึงการใช้การประมวลผลข้อมูลภาพดิจิทัลมาประยุกต์ใช้กับกระบวนการผลิต เพื่อทำการทำนายน้ำหนักและคัดขนาดของปลาทุ่นำสานพันธุ์ทองแถบ (Skipjack) หรือ งานวิจัยของ Z. Li [10] ที่กล่าวถึงการใช้การประมวลผลข้อมูลภาพดิจิทัลมาประยุกต์ใช้กับการออกแบบกราฟิก เพื่อปรับปรุงคุณภาพภาพ และช่วยในการสร้างสรรค์งานออกแบบที่มีคุณภาพสูงขึ้น เป็นต้น จากตัวอย่างเหล่านี้ ทำให้สังเกตได้ว่าการประมวลผลข้อมูลภาพดิจิทัลมีส่วนร่วมในธุรกิจและการพัฒนาวงการต่างๆ เป็นอย่างมาก ซึ่งเป็นสิ่งที่ขาดไม่ได้ในยุคปัจจุบันที่มีข้อมูลภาพดิจิทัลอยู่รอบตัว

ในงานวิจัยนี้มีการใช้ทฤษฎีเกี่ยวกับ Digital Image Processing ในขั้นตอนการศึกษาข้อมูล เพื่อหาข้อสังเกตและทฤษฎีและนำไปใช้ในการกำหนดกฎและเงื่อนไขสำหรับการพัฒนาโมเดลแบบ Rule-based

2.1.3 Image Forensics

Image Forensics หรือ การพิสูจน์ภาพดิจิทัล คือ กระบวนการเกี่ยวกับการตรวจสอบและวิเคราะห์ความสมบูรณ์ของภาพดิจิทัล เพื่อระบุว่ามีการแก้ไขหรือดัดแปลงภาพหรือไม่ เป็นกระบวนการที่มีถูกใช้ในงานเกี่ยวกับด้านนิติวิทยาศาสตร์ดิจิทัล หรือนงานวิจัยต่างๆ ที่ต้องการตรวจสอบความสมบูรณ์หรือต้นกำเนิดของภาพดิจิทัล จากบทความในเว็บไซต์ Forensic Focus [11] ได้กล่าวถึงขั้นตอนในกระบวนการพิสูจน์ภาพดิจิทัลไว้ ดังต่อไปนี้

ขั้นตอนหลักในการทำงานของกระบวนการพิสูจน์ภาพดิจิทัล

1. การเก็บข้อมูล โดยสามารถเก็บข้อมูลได้หลากหลายวิธี ซึ่งขั้นตอนนี้มีความสำคัญเป็นอย่างมาก เนื่องจากการเก็บข้อมูลที่ผิดพลาดอาจทำให้การตรวจสอบและวิเคราะห์ข้อมูลภาพดิจิทัลในขั้นตอนต่อไป เกิดความผิดพลาดได้ง่าย การเก็บข้อมูลที่ถูกต้องสามารถทำได้หลากหลายวิธี เช่น การวิเคราะห์ Disk Image เพื่อค้นหาไฟล์ภาพที่สมบูรณ์ เป็นต้น
2. การถอดรหัสข้อมูล (Decoding) ซึ่งการดำเนินการอาจขึ้นอยู่กับข้อมูลที่ได้มาจากขั้นตอนก่อนหน้า เนื่องจากข้อมูลแต่ละแบบมีวิธีการเก็บหรือการเข้ารหัสที่ต่างกัน ขั้นตอนนี้มีไว้เพื่อให้ทำความเข้าใจว่าควรจัดการกับข้อมูลที่ได้มาอย่างไร ข้อมูลมีความเสียหายหรือไม่ หรือการค้นหาเครื่องมือที่สามารถเข้าถึงข้อมูลนั้นได้โดยสมบูรณ์ที่สุด เป็นต้น
3. การตรวจสอบข้อมูลที่ได้รับมา ในขั้นตอนนี้จะทำการตรวจสอบข้อมูลที่มี เพื่อค้นหาข้อมูลที่สามารถนำไปใช้งานหรือใช้ประโยชน์ต่อได้ เช่น ค้นหาภาพที่ต้องการจากฐานข้อมูลขนาดใหญ่ ซึ่งขั้นตอนนี้สามารถใช้การประมวลผลข้อมูลภาพดิจิทัลมาใช้งานด้วยได้ เช่น ใช้การจดจำใบหน้า (Face Recognition) เพื่อช่วยเหลือในการค้นหาข้อมูลที่ต้องการได้
4. การวิเคราะห์เพื่อระบุต้นทางของข้อมูล (Image Ballistics) หากสามารถระบุต้นทางของข้อมูลได้ จะทำให้การวิเคราะห์ภาพในขั้นต่อไปทำได้ชัดเจนมากขึ้น ยกตัวอย่างเช่น การระบุว่าข้อมูลภาพมีที่มาจากอุปกรณ์ประเภทใด เช่น กล้องดิจิทัล หรือการสร้างภาพขึ้นมาใหม่ด้วยปัญญาประดิษฐ์ (Artificial Intelligence: AI)
5. การตรวจสอบความสมบูรณ์ของข้อมูล ขั้นตอนนี้ทำให้สามารถรู้ได้ว่าข้อมูลที่มีผ่านการดัดแปลงมาหรือไม่ เพื่อเป็นตัวช่วยในการระบุความน่าเชื่อถือของข้อมูล โดยการดัดแปลงข้อมูลนั้นสามารถทำได้หลากหลายรูปแบบ เช่น การดัดแปลงข้อมูลที่บอกลักษณะของข้อมูลหลัก (Metadata)

การดัดแปลงตัวข้อมูลภาพโดยตรงด้วยการปรับขนาดภาพหรือตัดภาพ (Cropping) หรือแม้แต่การลบหรือเพิ่มวัตถุภายในภาพ รวมถึงการใช้วิธีการตรวจสอบ Error Level Analysis (ELA) เพื่อการตรวจจับการเปลี่ยนแปลงของภาพ โดยทำการบีบอัดภาพและเปรียบเทียบกับภาพต้นฉบับเพื่อหาร่องรอยความแตกต่าง และระดับความผิดพลาด (Error Level) ระหว่างภาพทั้งสองภาพ ซึ่งหากพบว่าส่วนใดส่วนหนึ่งของภาพมีระดับข้อผิดพลาดสูงกว่าปกติ จะทำให้รู้ได้ว่าภาพนั้นมีแนวโน้มว่าผ่านการแก้ไขหรือเพิ่มเติมข้อมูลบางอย่างเข้ามาในภาพ

6. การประเมินคุณภาพของข้อมูล สามารถประเมินโดยยึดจากจุดประสงค์ในการวิเคราะห์ว่ามีคุณภาพที่เพียงพอต่อเป้าหมายที่ต้องการหรือไม่ เช่น ต้องการระบุใบหน้า หากมีข้อมูลภาพใบหน้า ต้องทำการประเมินว่าภาพนั้นมีขนาดหรือจำนวนพิกเซลมากเพียงพอต่อการระบุใบหน้าหรือไม่ หากประเมินแล้วคุณภาพยังไม่เพียงพอกับความต้องการ สามารถใช้วิธีใดในการปรับปรุงภาพเพื่อให้ได้คุณภาพที่ต้องการได้หรือไม่ (เช่น ใช้การประมวลผลข้อมูลภาพดิจิทัลในการปรับปรุงคุณภาพของภาพให้ดีขึ้น เป็นต้น)

7. การปรับปรุงข้อมูล หลังจากขั้นตอนการประเมินคุณภาพ จะทำการปรับปรุงข้อมูลตามที่ได้ประเมินไว้ ซึ่งในขั้นตอนนี้สามารถใช้วิธีการได้หลากหลายในการปรับปรุงข้อมูลภาพ เช่น ใช้การประมวลผลข้อมูลภาพดิจิทัลในการปรับปรุงภาพ หรือการใช้โมเดลในการกู้คืนข้อมูลภาพโดยไม่ให้ส่งผลกับคุณภาพของภาพ เป็นต้น

8. การเปรียบเทียบผลลัพธ์ เมื่อทำการปรับปรุงข้อมูลจากขั้นตอนก่อนแล้ว จะต้องทำการเปรียบเทียบผลลัพธ์ระหว่างภาพก่อนและหลังการปรับปรุงเพื่อตรวจสอบว่าได้ผลลัพธ์ตรงตามความต้องการหรือไม่ หากยังไม่ได้ผลลัพธ์ที่ต้องการ สามารถย้อนกลับไปทำการปรับปรุงข้อมูลใหม่ได้

การพิสูจน์ภาพดิจิทัลเป็นกระบวนการหลักในการตรวจสอบข้อมูลภาพ ซึ่งอาจมีวัตถุประสงค์ได้หลากหลายรูปแบบ ในด้านนิติวิทยาศาสตร์ดิจิทัล การพิสูจน์ดิจิทัลสามารถใช้ในการเก็บรวบรวมหลักฐานของคดีได้ ซึ่งหากผ่านกระบวนการพิสูจน์ที่ถูกต้องแล้ว จะสามารถใช้ข้อมูลภาพนั้นเป็นหลักฐานในชั้นศาลได้ นอกจากด้านนิติวิทยาศาสตร์ดิจิทัลแล้ว ในด้านการศึกษายังสามารถใช้เป็นตัวช่วยในการค้นหาทฤษฎีหรือความเกี่ยวข้องของข้อมูลที่ต้องการได้ จากหนังสือ Image and Video Forensics [12] ในปัจจุบันมีการพัฒนาในเทคนิคของ Image Forensics หลายด้าน เช่น การใช้ Deep Learning ในการตรวจจับภาพที่สร้างโดย GAN Model การใช้ Deep Learning ตรวจจับการปลอมแปลงภาพภายในวิดีโอ หรือการใช้ Convolutional Neural Networks (CNNs) ในการจำแนกภาพที่มีสีจากภาพต้นฉบับ เพื่อตรวจสอบความถูกต้องของภาพที่ที่อาจถูกปรับแต่งหรือสร้างขึ้นใหม่

Image Forensics ถูกนำมาใช้ในงานวิจัยในส่วนของการปรับปรุงข้อมูลบางส่วนใน Dataset เช่นการตรวจสอบเบื้องต้นว่าข้อมูลภาพใน Dataset ที่ได้มา มีการดัดแปลงของข้อมูลมาก่อนหน้าหรือไม่ เป็นการตรวจสอบคุณภาพของข้อมูล โดยใช้วิธีการวิเคราะห์ทิศทางแสงและจุดสะท้อน

ภายในภาพ ร่วมกับการทำ Thresholding แยกระหว่างส่วนที่เป็นแสงสว่างและส่วนที่เป็นเงา เพื่อสังเกตทิศทางและความเปลี่ยนแปลงของแสงสว่างภายในภาพ ร่วมกับการตรวจสอบขอบของภาพ เพื่อสังเกตว่ามีส่วนใดในภาพที่เส้นขอบไม่ต่อเนื่องกันหรือไม่ เป็นข้อมูลเพิ่มเติมสำหรับการระบุว่ามี การเพิ่มเครื่องประดับลงในภาพที่เห็นหรือไม่ และใช้วิธีการตรวจสอบค่าความแปรปรวนของความต่อเนื่องของ Pixel (Local Variance) เพื่อระบุพื้นที่ที่มีความไม่ต่อเนื่องของ Pixel และอาจมีการปรับค่า Threshold เพื่อให้สามารถสังเกตความแตกต่างของความต่อเนื่องของ Pixel ได้

2.1.4 การตรวจจับใบหน้า

การตรวจจับใบหน้า (Face Recognition) คือเทคโนโลยีสำหรับจดจำและตรวจจับใบหน้าของมนุษย์ โดยทำการตรวจจับใบหน้าจากข้อมูลที่ได้รับผ่านกล้องถ่ายภาพ รูปภาพหรือวิดีโอ การตรวจจับใบหน้าเป็นเทคโนโลยีที่มีการใช้อย่างหลากหลายในยุคปัจจุบัน เช่น การยืนยันตัวตนด้วยการสแกนใบหน้าผ่านโทรศัพท์มือถือ การสแกนใบหน้าเพื่อเข้าอาคารหรือสถานที่ที่จำกัดการเข้าถึง เป็นต้น อ้างอิงจากบทความในเว็บไซต์ของศูนย์ศึกษายุทธศาสตร์และนานาชาติ (Center for Strategic and International Studies: CSIS) ที่กล่าวถึงความแม่นยำในระบบตรวจจับใบหน้า [13] ระบบตรวจจับใบหน้ามีการพัฒนาเป็นอย่างมาก จากในปี พ.ศ. 2557 อัตราความผิดพลาดของระบบตรวจจับใบหน้ามีถึง 4.1% แต่ได้ลดลงเหลือเพียง 0.08% ในปี พ.ศ. 2563 ซึ่งส่งผลให้ระบบจำนวนมากสามารถนำระบบตรวจจับใบหน้าไปพัฒนาและประยุกต์ใช้เพิ่มเติมในระบบของตนได้ และเมื่อระบบมีความแม่นยำสูงมากเท่าไร ความเสี่ยงจากการระบุข้อมูลผิดพลาดก็ลดลงเท่านั้น ทำให้สามารถนำไปใช้ประโยชน์ได้อีกหลากหลายวิธีเช่นกัน

ขั้นตอนในการตรวจจับใบหน้านั้นประกอบไปด้วยขั้นตอนหลักคือ การตรวจจับใบหน้า (Detection) การวิเคราะห์และระบุที่มา (Analysis and Attribution) และการจำแนกใบหน้า (Recognition) โดยมีรายละเอียดแต่ละขั้นตอนดังต่อไปนี้

1. การตรวจจับใบหน้า (Detection) เป็นขั้นตอนสำหรับการระบุและตรวจจับตำแหน่งของใบหน้ามนุษย์ในรูปภาพหรือวิดีโอ โดยสามารถใช้เทคนิคหลายอย่างในการระบุตำแหน่งและพื้นที่ของใบหน้า เช่น การใช้อัลกอริทึม Haar Cascade Classifier ของ OpenCV [14] ในการตรวจจับรูปแบบของใบหน้า หรือการใช้ Convolutional Neural Networks (CNNs) ในการตรวจจับใบหน้า เป็นต้น

2. การวิเคราะห์และระบุที่มา (Analysis and Attribution) หลังจากขั้นตอนก่อนหน้าที่ทำการระบุตำแหน่งของใบหน้าแล้ว จะทำการแยกแยะข้อมูลลักษณะเฉพาะบนใบหน้า เช่น รูปร่างของใบหน้า รูปร่างและระยะห่างระหว่างดวงตา จมูกและปาก เพื่อทำการประมวลผลข้อมูลสำหรับใช้ในขั้นตอนการระบุตัวตนของใบหน้า

3. การจำแนกใบหน้า (Recognition) หลังจากได้ข้อมูลลักษณะเฉพาะบนใบหน้าแล้ว ทำการเปรียบเทียบข้อมูลที่ได้กับฐานข้อมูลที่มี เพื่อยืนยันว่าบุคคลในรูปมีข้อมูลอยู่ในฐานข้อมูลหรือไม่ ซึ่งเหมาะกับการนำไปใช้งานในระบบที่เกี่ยวข้องกับการยืนยันตัวตน

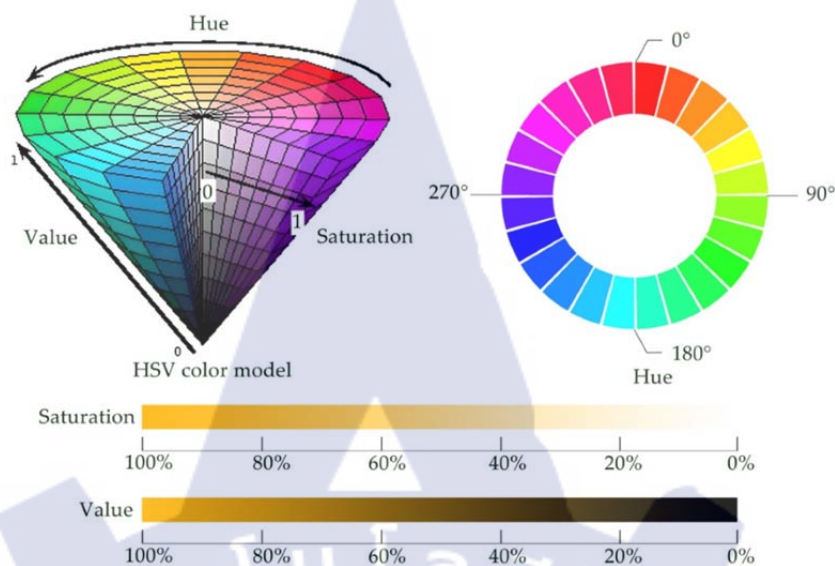
ระบบตรวจจับใบหน้าสามารถนำไปประยุกต์ใช้งานได้หลากหลายธุรกิจและอุตสาหกรรม ในด้านอาชญากรรม ระบบตรวจจับใบหน้าสามารถเป็นตัวช่วยในกรณีที่เจ้าหน้าที่ต้องการระบุตัวตนของ คนร้ายที่ปรากฏในภาพกล้องวงจรปิด หรือใช้ในด้านการทำธุรกรรม ปัจจุบันการทำธุรกรรมมักจะทำ ผ่านระบบอินเทอร์เน็ตที่ข้อมูลสามารถปลอมแปลงได้ง่าย การใช้ระบบตรวจจับใบหน้าในการยืนยัน ตัวตนก่อนการทำธุรกรรมถือเป็นตัวช่วยที่ดีมาก นอกจากการระบุตัวตนแล้ว การตรวจจับตำแหน่งของ ใบหน้ายังเป็นศาสตร์ที่สามารถนำไปต่อยอดในแง่ของงานวิจัยต่างๆ ได้ เนื่องจากวิธีการในการตรวจจับ ตำแหน่งของใบหน้านั้นสามารถทำได้หลายวิธี [15] ซึ่งยังมีการพัฒนาวิธีการหรือทฤษฎีใหม่ๆ เพื่อเพิ่ม ประสิทธิภาพในการตรวจจับตำแหน่งของใบหน้าให้แม่นยำมากยิ่งขึ้น และนำไปผสมผสานกับการตรวจจับ อื่นๆ ตามวัตถุประสงค์ของแต่ละระบบต่อไป

ซึ่งในงานวิจัยนี้ ผู้วิจัยเลือกใช้โมเดลที่มีชื่อว่า Multi-task Cascaded Convolutional Neural Network (MTCNN) ในการทำ Face Detection เพื่อระบุตำแหน่งของใบหน้าอยู่ส่วนไหน ของภาพ หลังจากนั้นจึงนำภาพที่ได้ไปทำการฝึกโมเดลในขั้นตอนต่อไป เนื่องจากโมเดล MTCNN สามารถ ระบุตำแหน่งของใบหน้า พร้อมกับการระบุตำแหน่งของ Facial Landmark บนใบหน้าอย่างดวงตา จมูก ปากได้

2.1.5 HSV Color Model

HSV Color Model คือโมเดลของสีประเภทหนึ่ง ซึ่งโมเดล HSV จะแสดงสีในรูปแบบ ที่ใกล้เคียงกับการรับรู้ของดวงตามนุษย์ โมเดลนี้จะแบ่งองค์ประกอบเป็น 3 ส่วน ได้แก่

1. Hue (ค่าสี) แสดงถึงประเภทของสีโดยวัดค่าเป็นองศา ตั้งแต่ 0° จนถึง 360° บน วงล้อของสี โดยแต่ละองศาบนวงล้อสีจะแสดงประเภทสีที่แตกต่างกัน เช่น ค่า 0° แสดงถึงสีแดง, 120° แสดงถึงสีเขียว และ 240° แสดงถึงสีน้ำเงิน เป็นต้น
2. Saturation (ค่าความอิ่มตัวสี) แสดงถึงความสดของสีนั้นๆ โดยวัดค่าระหว่าง 0 ถึง 100% ซึ่งยิ่งตัวเลขมีค่ามากเท่าไร แสดงถึงสีที่มีความสดมากขึ้นเท่านั้น กลับกันหากค่าเข้าใกล้ 0 มาก เท่าไร สีที่ปรากฏจะเข้าใกล้ความเป็นสีเทามากขึ้น
3. Value (ค่าความสว่าง) แสดงถึงความสว่างของสีนั้นๆ โดยวัดค่าระหว่าง 0 ถึง 100% ซึ่งยิ่งตัวเลขมีค่ามากเท่าไร แสดงถึงสีที่มีความสว่างมากขึ้นเท่านั้น กลับกันหากค่าเข้าใกล้ 0 มากเท่าไร สีที่ปรากฏจะมีลดลง



รูปที่ 2.2 ภาพแสดงองค์ประกอบของ HSV Color Model [16]

ความแตกต่างของ HSV Color Model เปรียบเทียบกับ RGB Color Model คือ RGB Model ใช้การผสมระหว่างค่าสีของ R (สีแดง), G (สีเขียว) และ B (สีน้ำเงิน) เพื่อแสดงออกมาเป็นสีหนึ่ง แต่ HSV Model จะใช้ค่า Hue ในการกำหนดสี และใช้ค่า Saturation, Value เพื่อกำหนดว่าสีนั้นมีความสดและความสว่างเท่าไร ซึ่งการแสดงสีด้วย HSV Model ถูกระบุโทนสีด้วยค่า Hue และระบุค่าความสว่างได้ด้วยค่า Value แบบแยกกัน แต่ถ้าเป็น RGB Model หากค่าความสว่างเปลี่ยนแปลงไป ทั้งค่า R, G และ B อาจถูกเปลี่ยนไปทั้งสามค่า ทำให้ยากในการระบุว่าการเปลี่ยนแปลงความสว่างแล้วสีนั้นยังเป็นสีเดิมอยู่หรือไม่ เนื่องจากมีความสัมพันธ์ระหว่างค่า R, G, B ในการแสดงสีอยู่ HSV Model จึงมีความเหมาะสมมากกว่าในกรณีที่ต้องการเก็บข้อมูลสีและข้อมูลความสว่างแยกออกจากกัน

T. Kusnandar et al. [17] ได้มีการอธิบายถึงวิธีการแปลงค่าจาก RGB ไปเป็นค่า HSV โดยในการหาค่า Value ที่เป็นค่าความสว่างจะใช้ค่าสีที่สูงที่สุดใน R, G, B เช่น หากค่า R, G, B เป็น (255, 200, 0) ค่า Value จะเท่ากับ 255 โดยมีสูตรคำนวณค่า Value ดังที่เห็นในสมการ 2.1 ต่อไปนี้

$$V = \max(r, g, b) \quad (2.1)$$

ในการหาค่า Saturation ที่เป็นค่าความสดของสี จะนำค่า Value ที่ได้ในขั้นตอนก่อนหน้ามาใช้ โดยใช้ร่วมกับค่าสีที่ต่ำที่สุดใน R, G, B เช่น หากค่า R, G, B เป็น (255, 200, 200) ค่า Saturation จะเท่ากับ 200 ยกเว้นในกรณีที่ค่า Value มีค่าเท่ากับ 0 ค่า Saturation จะมีค่าเท่ากับ 0 โดยอัตโนมัติ โดยมีสูตรคำนวณค่า Saturation ดังที่เห็นในสมการ 2.2 ต่อไปนี้

$$S = \begin{cases} 0, & \text{if } V = 0 \\ V - \frac{\min(r, g, b)}{V}, & \text{if } V > 0 \end{cases} \quad (2.2)$$

ในการหาค่า Hue ที่เป็นค่าสี จะนำค่า Value และค่า Saturation ที่ได้ในขั้นตอนก่อนหน้ามาใช้ โดยใช้ร่วมกับค่าสี R, G, B ทั้ง 3 ค่า ยกเว้นในกรณีที่ค่า Saturation มีค่าเท่ากับ 0 ค่า Hue จะมีค่าเท่ากับ 0 โดยอัตโนมัติ และสูตรการคำนวณจะเปลี่ยนไปตามกรณีที่ค่า V มีค่าเท่ากับค่าใดๆ ในค่า R, G, B หรือกรณีที่ค่า Hue ที่คำนวณออกมาได้ค่าน้อยกว่า 0 จะทำการบวก 360 เข้าไปเพื่อให้ค่าไม่ติดลบ และยังคงเป็นค่าองศาที่อยู่ในวงล้อของสี โดยมีสูตรคำนวณค่า Hue ดังที่เห็นในสมการ 2.3 ต่อไปนี้

$$H = \begin{cases} 0, & \text{if } V = 0 \\ 60^\circ \times \left[\frac{g - b}{s \times v} \right], & \text{if } V = r \\ 60^\circ \times \left[2 + \frac{b - r}{s \times v} \right], & \text{if } V = g \\ 60^\circ \times \left[4 + \frac{r - g}{s \times v} \right], & \text{if } V = b \\ H + 360^\circ, & \text{if } H < 0 \end{cases} \quad (2.3)$$

ซึ่งค่า HSV ที่ได้มาจากการแปลงค่า RGB นี้ สามารถนำไปใช้ในการคำนวณอื่นๆ ที่เกี่ยวข้องกับการตรวจจับภาพได้ เนื่องจากค่าเหล่านี้สามารถแยกแยะระหว่างข้อมูลที่เป็นสี และข้อมูลที่เป็นความสว่างออกจากกันได้อย่างชัดเจนต่างกับค่า RGB

2.1.6 การวัดการกระจายตัวของข้อมูล

การวัดการกระจายตัวของข้อมูล หมายถึง การวัดค่าทางสถิติที่เป็นระยะห่างเฉลี่ยระหว่างข้อมูลแต่ละค่ากับค่าเฉลี่ยของชุดข้อมูลนั้นๆ ซึ่งแสดงถึงลักษณะการกระจายตัวของข้อมูล ยิ่งค่าระยะห่างนี้มากเท่าไร หมายถึงมีการกระจายตัวของข้อมูลมาก (ระยะห่างของแต่ละค่าในข้อมูลมาก) โดยในงานวิจัยนี้จะพูดถึงวิธีการทั้งหมด 3 วิธีการ นั่นคือ Mean Absolute Deviation, Median Absolute Deviation และ Standard Deviation ซึ่งมีความแตกต่างกันในวิธีการหาค่าเฉลี่ย

1. Mean Absolute Deviation หรือ Mean AD หมายถึง ส่วนเบี่ยงเบนเฉลี่ยสัมบูรณ์ โดยค่า Mean AD จะใช้วิธีการหาค่าเฉลี่ย (Mean) ซึ่งเป็นการนำผลลัพธ์ทั้งหมดในข้อมูลมารวมกัน และหารด้วยจำนวนข้อมูลเหล่านั้นเพื่อให้ได้ค่าเฉลี่ยออกมา วิธีนี้มีข้อสังเกตอยู่ที่การนำข้อมูลทั้งหมดมาคำนวณรวมกัน ซึ่งหากมีข้อมูลที่มีค่าห่างจากข้อมูลอื่นมากๆ (Outliers) ข้อมูลนั้นจะถูกนำมาคำนวณด้วย ซึ่งอาจส่งผลต่อค่าเฉลี่ยที่เป็นผลลัพธ์ได้ ซึ่งหลักการของการหาค่าการกระจายตัวนี้คือการนำข้อมูลแต่ละค่าในชุดข้อมูลลบกับค่าเฉลี่ย แล้วใส่ค่าสัมบูรณ์ (Absolute) เพื่อไม่ให้ค่าติดลบ จากนั้นทำการวัดค่าเฉลี่ยของผลลัพธ์นั้น โดยการคำนวณ Mean Absolute Deviation มีสูตรคำนวณ ดังที่เห็นในสมการ 2.4 ต่อไปนี้

$$\text{Mean AD} = \frac{\sum |x_i - \bar{x}|}{n} \quad (2.4)$$

โดย n คือ จำนวนข้อมูลในชุดข้อมูล

x_i คือ ค่าของข้อมูลในตำแหน่งที่ i ของชุดข้อมูล

$\bar{x}$ คือ ค่าเฉลี่ย (Mean) ของชุดข้อมูลนั้น

2. Median Absolute Deviation หรือ Median AD หมายถึง ส่วนเบี่ยงเบนมัธยฐานสัมบูรณ์ โดยค่า Median AD จะใช้วิธีการหาค่ามัธยฐาน (Median) ซึ่งเป็นการนำผลลัพธ์ทั้งหมดในข้อมูลมาเรียงจากค่าที่น้อยที่สุดไปยังค่าที่มากที่สุด และทำการเลือกค่าที่อยู่ตรงกลางของชุดข้อมูลนั้นออกมา เรียกว่าค่ามัธยฐาน โดยในกรณีที่จำนวนข้อมูลเป็นเลขคี่ จะเลือกค่าที่อยู่ตรงกลางทันที หากเป็นกรณีที่จำนวนข้อมูลเป็นเลขคู่ จะนำค่าที่อยู่ตรงกลาง 2 ตำแหน่งมาทำการหาค่าเฉลี่ยด้วยการรวมกันและหารด้วย 2 ทำให้มีโอกาสได้ค่ามัธยฐานที่เป็นค่าที่ไม่ได้อยู่ในชุดข้อมูลเดิมได้ ข้อสังเกตของวิธีการนี้คือการนำข้อมูลทั้งหมดมาเรียงลำดับและเลือกค่าตรงกลาง วิธีนี้ทำให้ข้อมูลที่มีค่าห่างจากข้อมูลอื่นๆ มาก (Outliers) ไม่ถูกนำมาคำนวณ เนื่องจาก Outliers ที่เป็นค่าห่างจากค่าอื่นๆ มักจะอยู่ในตำแหน่งขอบของชุดข้อมูลที่ถูกเรียงค่าใหม่แล้ว ทำให้การเลือกค่าตรงกลางมีความเป็นไปได้น้อยที่จะเลือกค่าที่เป็น Outliers นอกจากนั้นวิธีการนี้ยังมีการใส่ค่าสัมบูรณ์ (Absolute) ที่ส่วนต่างระหว่างค่าในชุดข้อมูลกับค่ามัธยฐานเพื่อไม่ให้ค่าติดลบ จากนั้นทำการวัดค่ามัธยฐานของผลลัพธ์นั้น โดยการคำนวณ Median Absolute Deviation มีสูตรคำนวณ ดังที่เห็นในสมการ 2.5 ต่อไปนี้

$$\text{Median AD} = \text{median} (|x_i - \bar{x}|) \quad (2.5)$$

โดย x_i คือ ค่าของข้อมูลในตำแหน่งที่ i ของชุดข้อมูล
 $\bar{x}$ คือ ค่ามัธยฐาน (Median) ของชุดข้อมูลนั้น

3. Standard Deviation หรือ SD หมายถึง ส่วนเบี่ยงเบนมาตรฐาน โดยค่า Standard Deviation จะใช้วิธีการนำแต่ละค่าในชุดข้อมูล มาทำการลบกับค่าเฉลี่ย (Mean) ของข้อมูลทั้งหมด แล้วทำการยกกำลังสองค่านั้น จากนั้นรวมค่าความห่างของแต่ละค่าในชุดข้อมูลที่ยกกำลังสองแล้วมารวมกันหารด้วยจำนวนค่าในชุดข้อมูล -1 เพื่อลดความคลาดเคลื่อนในการประมาณค่าความแปรปรวนของจำนวนค่าในชุดข้อมูลนั้น และทำการหารากที่สองของค่านั้นเพื่อให้ได้ส่วนเบี่ยงเบนมาตรฐานของข้อมูลชุดนั้น ซึ่งมีความแตกต่างกับวิธีการของ Mean Absolute Deviation ที่ใช้ค่าเฉลี่ย (Mean) ในการคำนวณเหมือนกัน เนื่องจาก Mean AD จะใช้ค่าสัมบูรณ์ (Absolute) เพื่อไม่ให้ค่านั้นๆ ติดลบ แต่ SD จะใช้การยกกำลังเพื่อไม่ให้ค่านั้นๆ ติดลบแทน โดยการคำนวณ Standard Deviation มีสูตรคำนวณดังที่เห็นในสมการ 2.6 ต่อไปนี้

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}} \quad (2.6)$$

โดย n คือ จำนวนข้อมูลในชุดข้อมูล
 x คือ ค่าในชุดข้อมูล
 $\bar{x}$ คือ ค่าเฉลี่ย (Mean) ของชุดข้อมูลนั้น

ซึ่งทั้งสองวิธีการที่กล่าวไว้ข้างต้น เป็นวิธีการที่สามารถวัดการกระจายตัวของข้อมูลได้ โดยมีเงื่อนไขที่ต่างกันในการหาค่าเฉลี่ยของชุดข้อมูล ซึ่งมีจุดสังเกตที่แตกต่างกันในแง่ของการนำข้อมูลภายในชุดข้อมูลมาใช้ เช่น การใช้ค่าเฉลี่ยที่มีการนำค่า Outliers มาเป็นส่วนหนึ่งในการคำนวณค่าเฉลี่ย และการใช้ค่ามัธยฐานที่ไม่ใช้ส่วนที่เป็นค่า Outliers เป็นต้น

2.1.7 Euclidean Distance

Euclidean Distance คือ การวัดระยะทางระหว่างจุดสองจุดเป็นเส้นตรง โดยมีรากฐานมาจากทฤษฎีบทพีทาโกรัส สามารถคำนวณระยะทาง 2 มิติ, 3 มิติ หรือ n มิติ ได้ มีหลักการคือการหาส่วนต่างของระยะห่างในแต่ละมิติ จากนั้นทำการยกกำลังสองเพื่อทำให้ค่าเป็นค่าบวกเสมอ และทำให้ส่วนต่างของค่าถูกขยายให้ชัดเจนมากขึ้นเมื่อเทียบกับส่วนต่างของค่าที่น้อยกว่า จากนั้น

ทำการรวมค่าของทุกมิติเข้าด้วยกัน และทำการหารากที่สองของผลรวมนั้นเพื่อปรับให้กลับมาเป็นหน่วยปกติ (จากการยกกำลังสองในขั้นตอนก่อนหน้านี้) เพื่อให้ได้ระยะทางจริง โดยการคำนวณ Euclidean Distance มีสูตรในการคำนวณ ดังที่เห็นในสมการ 2.7 ต่อไปนี้ [18]

$$d = \sqrt{\sum_{i=1}^n (b_i - a_i)^2} \quad (2.7)$$

ในงานวิจัยนี้มีการใช้การคำนวณ Euclidean Distance ในขั้นตอนการวัดความต่างของ Pixel สีในภาพเพื่อทำการสร้าง Mask สำหรับใช้ในการกำหนดพื้นที่ในการเก็บข้อมูลแสงสว่างในแต่ละ Pixel ด้วยการเปรียบเทียบค่าสีใน HSV Model ซึ่งค่า H, S, V เปรียบเสมือนแกนพิกัดแบบ 3 มิติ ในสมการ 2.7 ข้างต้น จึงทำการแทนที่ค่า n ด้วย 3 การคำนวณความต่างของระยะทางจึงใช้สมการการคำนวณระยะทางแบบ 3 มิติ ดังที่เห็นในสมการ 2.8 ต่อไปนี้

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \quad (2.8)$$

ซึ่งในการใช้ Euclidean Distance ในการคำนวณความต่างของ Pixel สีใน HSV Model นั้น หากข้อมูลในแต่ละมิติมีช่วงความกว้างของค่าไม่เท่ากัน มิติที่มีค่าตัวเลขสูงกว่าอาจมีอิทธิพลต่อผลลัพธ์มากกว่ามิติอื่นๆ ได้ ทำให้มิติอื่นๆ มีผลต่อผลลัพธ์น้อยกว่าและให้ผลลัพธ์ที่ผิดได้ จึงจำเป็นต้องมีการปรับช่วงความกว้างของข้อมูลด้วยการทำ Normalization ข้อมูลให้อยู่ในช่วงความกว้างเท่ากัน เช่น ค่า Hue ที่เดิมมีช่วงความกว้างอยู่ที่ 0-360 องศา จำเป็นต้องทำการ Normalize ค่าเพื่อให้อยู่ในช่วงความกว้าง 0-1 เหมือนกับค่า Saturation และค่า Value ก่อนนำมาทำการคำนวณระยะทางด้วย Euclidean Distance

2.1.8 Rule-based System

Rule-based System คือ ระบบที่ใช้กฎเกณฑ์หรือเงื่อนไขที่กำหนดไว้ล่วงหน้าในการประมวลผลข้อมูล หากเกิดสถานการณ์ที่มีเงื่อนไขตรงตามเงื่อนไขที่ระบุไว้ ระบบจะตอบสนองด้วยการกระทำที่กำหนดไว้ โดยระบบนี้จะไม่พึ่งพาการเรียนรู้ข้อมูลใหม่ๆ แต่จะทำงานตามเงื่อนไขและสิ่งที่กำหนดไว้ ทำให้ต้องอาศัยความรู้และความเข้าใจของผู้ออกแบบระบบ เพื่อออกแบบเงื่อนไขที่ครอบคลุมสถานการณ์ต่างๆ ให้มากที่สุด และให้สามารถแสดงผลออกมาได้อย่างถูกต้อง โดย Rule-based System สามารถนำไปใช้งานได้หลากหลายรูปแบบ ตามวัตถุประสงค์ที่ถูกสร้างขึ้น และเนื่องจากเป็นระบบที่ต้องทำการออกแบบเงื่อนไขไว้ล่วงหน้า รวมถึงการกำหนดการแสดงผลต่อแต่ละเงื่อนไข ทำให้อาจมีความไม่แม่นยำ

ในกรณีที่สถานการณ์ไม่ตรงตามเงื่อนไขที่กำหนดไว้ หรือหากอยู่นอกเหนือขอบเขตที่ระบบถูกวางไว้ก็จะไม่สามารถแสดงผลออกมาได้อย่างถูกต้อง เป็นหนึ่งในข้อจำกัดของ Rule-based System

องค์ประกอบของ Rule-based System นั้นมีหลักๆ 3 ข้อ ข้อแรกคือกฎหรือคำสั่งที่ระบุสถานการณ์และการแสดงผลที่ถูกต้อง ซึ่งการออกแบบเงื่อนไขเหล่านี้จะต้องทำความเข้าใจกับสถานการณ์ที่อาจเกิดขึ้น และต้องระบุให้ครอบคลุมสถานการณ์ให้มากที่สุด เพื่อให้ระบบมีความสมบูรณ์ ข้อถัดมาคือตัวแปรและเงื่อนไข ตัวแปรในที่นี้หมายถึงข้อมูลที่ระบบจะนำไปใช้ตรวจสอบกับเงื่อนไข เพื่อเปรียบเทียบว่าข้อมูลตัวแปรที่ได้รับมานั้นตรงกับเงื่อนไขในข้อใดบ้าง และทำการแสดงผลตามเงื่อนไขที่ตรงกันองค์ประกอบข้อสุดท้ายคือผลลัพธ์ ซึ่งผลลัพธ์นั้นจะเกิดขึ้นเมื่อเงื่อนไขต่างๆ นั้นเป็นจริงเท่านั้น การออกแบบกฎและเงื่อนไขจึงต้องคำนึงถึงความเป็นไปได้ทั้งหมดที่อาจเกิดขึ้น เพื่อให้สามารถพัฒนา Rule-based System ที่มีความแม่นยำสูงได้

ข้อดีของ Rule-based

1. สามารถตรวจสอบได้ว่าผลลัพธ์ที่เกิดขึ้น เป็นผลมาจากเงื่อนไขข้อใดที่กำหนดไว้ ทำให้ง่ายต่อการแก้ไขและสามารถตรวจสอบย้อนหลังได้ง่าย
2. ไม่จำเป็นต้องใช้ข้อมูลจำนวนมากในการทดสอบ เนื่องจากการกำหนดเงื่อนไขให้ตรงตามที่ต้องการ ไม่ต้องพึ่งพาการป้อนข้อมูลซ้ำๆ แบบเดียวกับการฝึกโมเดลของ Machine Learning
3. สามารถทดสอบและประเมินผลลัพธ์ของระบบได้ง่าย เนื่องจากเห็นภาพรวมของเงื่อนไขและการตัดสินใจของระบบทั้งหมด

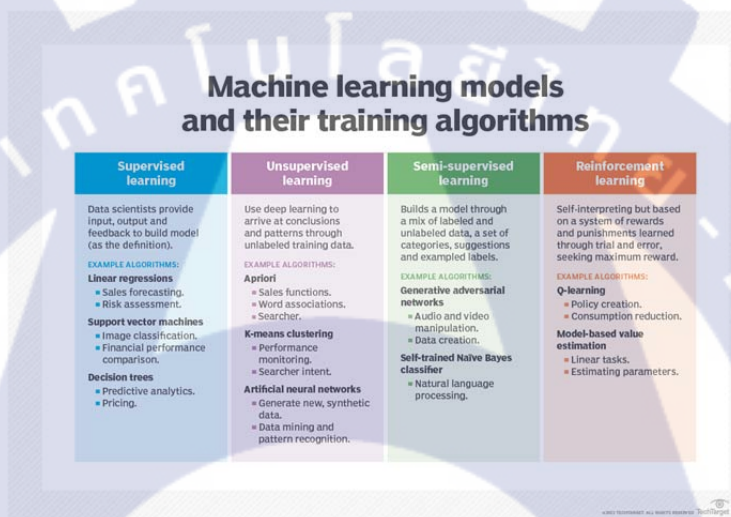
ข้อเสียของ Rule-based

1. มีโอกาสที่การทำงานของระบบจะผิดพลาด หากข้อมูลที่ได้รับไม่อยู่ในกฎหรือเงื่อนไขที่กำหนดไว้
2. ไม่เหมาะกับงานที่มีความซับซ้อนสูง หรืองานที่มีรูปแบบข้อมูลที่ไม่แน่นอน
3. จำเป็นต้องระบุกฎและเงื่อนไขให้ครอบคลุมทุกสถานการณ์ และไม่มีความขัดแย้งในตัวเอง หากสถานการณ์นั้นๆ มีข้อยกเว้นหรือกรณีพิเศษจำนวนมาก อาจเขียนเงื่อนไขได้ไม่ครอบคลุมทั้งหมด

ในงานวิจัยนี้ ผู้วิจัยมีการกำหนด Rule-based ในส่วนของการกำหนดเกณฑ์และเงื่อนไขสำหรับการพัฒนาโมเดลตรวจจับภาพ ซึ่งจำเป็นต้องผ่านการศึกษาทฤษฎีด้วย Digital Image Processing และ Image Forensics ดังที่ระบุไว้ก่อนหน้านี้

2.1.9 Machine Learning

ระบบการเรียนรู้ของเครื่อง (Machine Learning: ML) เป็นส่วนประกอบหนึ่งของปัญญาประดิษฐ์ (Artificial Intelligence: AI) หมายถึง การทำให้ระบบสามารถเรียนรู้ได้ด้วยตัวเอง โดยมีจุดประสงค์ในการพัฒนาระบบที่มีประสิทธิภาพและสามารถทำงานแทนมนุษย์ได้ โดย Machine Learning จะทำการตรวจสอบข้อมูลจำนวนมาก และทำการวิเคราะห์หรือตัดสินใจจากสิ่งที่เรียนรู้ได้จากข้อมูลเหล่านั้น โดยมนุษย์ไม่จำเป็นต้องระบุข้อมูลให้กับตัวระบบที่เรียนรู้ เพราะระบบ สามารถอนุมานตรรกะได้เองจากการวิเคราะห์ความสัมพันธ์ในข้อมูลที่ทำให้การเรียนรู้ทั้งหมด โดยสามารถแบ่งประเภทการเรียนรู้ของ Machine Learning ได้เป็น 4 ประเภทใหญ่ ดังที่แสดงในรูปที่ 2.3



รูปที่ 2.3 ประเภทการเรียนรู้ของ Machine Learning [19]

ประเภทการเรียนรู้ของ Machine Learning

1. Supervised learning หมายถึง การเรียนรู้แบบมีผู้สอน โดยมนุษย์จะเป็นผู้ป้อนข้อมูลให้คอมพิวเตอร์ พร้อมระบุว่าข้อมูลนั้นคืออะไร ระบบจะทำการจัดหมวดหมู่ตามข้อมูลที่ได้รับ (Classification) เพื่อให้สามารถจัดหมวดหมู่ของข้อมูลเองได้ในอนาคต เป็นประเภทการเรียนรู้ที่เหมาะสมกับระบบที่มีข้อมูลตัวอย่างที่ถูกระบุ Label ของข้อมูลไว้อยู่แล้ว (รู้คำตอบที่ถูกต้องของข้อมูลนั้นอยู่แล้ว) ยกตัวอย่างเช่น ระบบตรวจจับสแปมอีเมล (เนื่องจากเรียนรู้จากข้อมูลที่ระบุไว้ว่า ข้อมูลแบบใดถึงเข้าข่ายการเป็นสแปม) หรือระบบสำหรับการวินิจฉัยโรคจากภาพเอกซเรย์ (เนื่องจากเรียนรู้จากข้อมูลที่ระบุไว้ว่า ภาพแบบใดถึงเข้าข่ายการเป็นโรค) เป็นต้น

2. Unsupervised learning หมายถึง การเรียนรู้แบบไม่มีผู้สอน เป็นการเรียนรู้ที่ต่างกับ Supervised learning เนื่องจากข้อมูลที่ถูกป้อนเข้าระบบนั้นจะไม่มีการระบุว่าเป็นข้อมูลอะไร โดยระบบจะต้องทำการตรวจสอบข้อมูลทั้งหมด ค้นหารูปแบบและความสัมพันธ์ของข้อมูล และทำการจัดกลุ่มข้อมูลที่คล้ายกัน (Clustering) ด้วยตัวเอง เป็นประเภทการเรียนรู้ที่เหมาะสมกับระบบที่ต้องการหาความสัมพันธ์บางอย่างของข้อมูล หรือระบบที่ต้องการเข้าใจโครงสร้างของข้อมูลมากกว่าการระบุข้อมูลเหล่านั้นคืออะไร ยกตัวอย่างเช่น ระบบการจัดกลุ่มลูกค้า (ทำการแบ่งกลุ่มลูกค้าโดยวิเคราะห์จากพฤติกรรมของลูกค้า แม้ไม่มีการระบุลูกค้าคนใดอยู่ในกลุ่มไหน) หรือระบบสำหรับตรวจจับความผิดปกติบางอย่าง (ทำการมองหาข้อมูลที่มีจุดแตกต่างกับข้อมูลส่วนอื่น) ซึ่งในงานวิจัยนี้จะมีการใช้วิธีการเรียนรู้ที่เรียกว่า Transfer Learning ซึ่งเป็นการนำโมเดลที่ผ่านการเรียนรู้มาแล้วในระดับหนึ่ง มาทำการฝึกต่อด้วยชุดข้อมูล Dataset ที่มี โดยไม่ได้ระบุข้อมูลเหล่านั้นเป็นข้อมูลอะไร จึงนับว่าในงานวิจัยนี้ Machine Learning ใช้การเรียนรู้แบบ Unsupervised Learning

3. Semi-supervised Learning หมายถึง การเรียนรู้ที่แก้ไขข้อจำกัดของ Supervised Learning และ Unsupervised Learning โดยการเรียนรู้ประเภทนี้จะทำการเรียนรู้ด้วยข้อมูลที่ถูกระบุ Label ไว้เพียงส่วนหนึ่ง และค่อยๆ นำสิ่งที่เรียนรู้ได้จากข้อมูลนั้นไปใช้กับชุดข้อมูลที่ใหญ่กว่าในภายหลัง ทำให้สามารถเรียนรู้ได้โดยอาศัยข้อมูลที่มี Label เพียงส่วนเดียวผสมกับข้อมูลที่ไม่มี Label ได้ ทำให้ประหยัดเวลาในการเตรียมข้อมูลได้ การเรียนรู้ประเภทนี้เหมาะกับระบบที่มีข้อมูลจำนวนมาก แต่มีระบุ Label ไว้ส่วนหนึ่ง เช่น การจำแนกภาพ (Image Classification) ที่ระบุ Label ของข้อมูลภาพไว้เพียงส่วนเดียว (ในกรณีที่ไม่สามารถระบุ Label ของภาพทั้งหมดได้)

4. Reinforcement Learning หมายถึง การเรียนรู้แบบที่ใช้การโต้ตอบเป็นส่วนหนึ่งในการเรียนรู้ มักใช้กับระบบที่ผ่านการเรียนรู้มาแล้วในระดับหนึ่ง เพื่อทำการเพิ่มประสิทธิภาพของระบบนั้นให้ดีขึ้น โดยระบบจะทำการเรียนรู้ตามสถานการณ์ที่มนุษย์เป็นผู้ระบุให้ แล้วทำการปรับปรุงกระบวนการไปเรื่อยๆ ด้วยการลองผิดลองถูก เมื่อระบบทำการตัดสินใจ จะได้รับรางวัลหรือโทษตามความถูกต้องของการตัดสินใจนั้น เพื่อให้สามารถปรับปรุงระบบให้ดีขึ้นได้ต่อไป การเรียนรู้ประเภทนี้เหมาะกับระบบที่ต้องมีการตอบโต้กับสภาพแวดล้อม หรือมีการตัดสินใจแบบต่อเนื่อง เนื่องจากระบบจะสามารถลองผิดลองถูกและเรียนรู้จากผลลัพธ์ (รางวัลหรือโทษ) ที่ได้รับได้ เช่น ระบบรถยนต์ไร้คนขับ (จำเป็นต้องตัดสินใจในแต่ละขั้นตอน เช่น เลี้ยว หรือเบรก เพื่อให้สามารถขับได้อย่างปลอดภัย) เป็นต้น

การใช้ Machine Learning นั้นต้องเข้าใจในจุดประสงค์ของระบบที่ต้องการพัฒนา เพื่อให้สามารถตัดสินใจได้ว่าระบบนั้นต้องใช้การเรียนรู้แบบไหนจึงจะเหมาะสมและให้ผลลัพธ์ที่ดีที่สุด รวมถึงการวิเคราะห์ข้อมูลที่มีประกอบการตัดสินใจเพื่อให้สามารถเรียนรู้ได้อย่างมีประสิทธิภาพที่สุด โดยในงานวิจัยนี้ ผู้วิจัยมีการใช้ Machine Learning ในการพัฒนาโมเดลตรวจจับภาพ ซึ่งทำการเลือกมา

ทั้งหมด 3 โมเดล โดยทั้ง 3 โมเดลที่ทำการเลือกมา เป็นโมเดลที่ผู้วิจัยวิเคราะห์และตัดสินใจว่ามีความเหมาะสมกับระบบที่ต้องการพัฒนา โดยจะมีการลงรายละเอียดในส่วนของโมเดลในส่วนถัดไป

2.1.10 Confusion Matrix

Confusion Matrix คือ เครื่องมือที่ใช้ในการประเมินประสิทธิภาพของโมเดลสำหรับการแยกแยะประเภทของข้อมูล (Classification) หรือโมเดลที่สร้างขึ้นโดยการใช้ Machine Learning ซึ่งมีแนวคิดหลักเกี่ยวกับการวัดผลความถูกต้องระหว่างความจริง กับคำตอบที่ระบบทำการทำนาย (Prediction) โดยผลลัพธ์ที่ได้จะเป็นสัดส่วนระหว่างผลการทำนายที่ถูกต้อง (Positive) และผลการทำนายที่ผิด (Negative) ผลลัพธ์ของการทำนายมีได้ทั้งหมด 4 รูปแบบ คือ True Positive (TP), False Positive (FP), True Negative (TN) และ False Negative (FN) ซึ่งผลลัพธ์นี้เกิดจากการทำนายข้อมูลของโมเดล ว่าสามารถทำนายได้ถูกต้องตามข้อมูลจริง หรือทำนายผิดไปจากข้อมูลจริง สามารถแสดงในรูปแบบตารางได้ ดังที่แสดงในตารางที่ 2.1

ตารางที่ 2.1 ตาราง Confusion Matrix

	ผลลัพธ์เป็นจริง	ผลลัพธ์ไม่เป็นจริง
ทำนายว่าเป็นจริง	True Positive (TP)	False Positive (FP)
ทำนายว่าไม่เป็นจริง	False Negative (FN)	True Negative (TN)

ผลลัพธ์ของการทำนายทั้ง 4 รูปแบบ

1. True Positive (TP) หมายถึง ผลการทำนายตรงกับข้อมูลจริง กรณีที่ข้อมูลเป็นจริง และทำนายว่าเป็นจริง
2. True Negative (TN) หมายถึง ผลการทำนายตรงกับข้อมูลจริง กรณีที่ข้อมูลไม่เป็นจริง และทำนายว่าไม่เป็นจริง
3. False Positive (FP) หมายถึง ผลการทำนายไม่ตรงกับข้อมูลจริง กรณีที่ข้อมูลไม่เป็นจริง แต่ทำนายว่าเป็นจริง
4. False Negative (FN) หมายถึง ผลการทำนายไม่ตรงกับข้อมูลจริง กรณีที่ข้อมูลเป็นจริง แต่ทำนายว่าไม่เป็นจริง

จากผลลัพธ์การทำนายทั้ง 4 รูปแบบนี้ สามารถนำไปใช้ในการคำนวณเพื่อหาค่าตัวชี้วัดได้ ดังต่อไปนี้

1. Accuracy หรือ ค่าความถูกต้อง เป็นอัตราส่วนที่โมเดลทำนายได้ถูกต้องตรงกับข้อมูลจริง เทียบกับจำนวนข้อมูลทั้งหมดที่ให้โมเดลทำการทำนาย มีสูตรคำนวณดังที่เห็นในสมการ 2.9 ต่อไปนี้

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (2.9)$$

โดยนำค่า TP และ TN (เฉพาะคำตอบที่ทำนายได้ถูกต้องตรงตามข้อมูลจริง) มาหารกับจำนวนคำตอบทั้งหมดที่ทำการทำนาย เพื่อหาค่า Accuracy

2. Precision หรือ ค่าความแม่นยำ เป็นอัตราส่วนที่เปรียบเทียบระหว่างค่าที่โมเดลทำนายว่าเป็นจริง และข้อมูลเป็นจริง (TP) กับค่าที่โมเดลทำนายว่าข้อมูลเป็นจริง นับรวมทั้งข้อมูลที่เป็นจริงและไม่เป็นจริง (TP และ FP) มีสูตรคำนวณดังที่เห็นในสมการ 2.10 ต่อไปนี้

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2.10)$$

3. Recall หรือ ค่าความอ่อนไหว คืออัตราส่วนที่โมเดลทำนายว่าเป็นจริง และข้อมูลเป็นจริง (TP) กับค่าที่เป็นจริงทั้งหมด นับรวมทั้งที่โมเดลทำนายว่าเป็นจริง และที่โมเดลทำนายว่าไม่เป็นจริง (TP และ FN) มีสูตรคำนวณดังที่เห็นในสมการ 2.11 ต่อไปนี้

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2.11)$$

4. F1-Score หรือ ค่าวัดประสิทธิภาพโดยรวม คือการนำค่าความแม่นยำ (Precision) และค่าความอ่อนไหว (Recall) มาทำการหาค่าเฉลี่ยแบบ Harmonic Mean (ค่าเฉลี่ยสำหรับอัตราส่วน) เพื่อให้ได้ค่าเฉลี่ยที่สามารถบ่งบอกถึงประสิทธิภาพของโมเดลได้ มีสูตรคำนวณดังที่เห็นในสมการ 2.12 ต่อไปนี้

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (2.12)$$

โดยการคำนวณแต่ละค่านั้น สามารถนำไปใช้ในการวัดผลโมเดลได้หลากหลายรูปแบบตามความเหมาะสมของโมเดลนั้นๆ ซึ่งนอกเหนือจากค่าชี้วัด 4 ค่าที่กล่าวถึงข้างต้น อาจมีวิธีการอื่นเพิ่มเติมในการคำนวณความแม่นยำหรือความคลาดเคลื่อนของการทำนาย ขึ้นอยู่กับรูปแบบโมเดลที่ต้องการวัดผลว่ามีข้อจำกัดหรือเงื่อนไขพิเศษใดเพิ่มเติมหรือไม่

ในงานวิจัยนี้ ผู้วิจัยจะเน้นไปที่การคำนวณหาค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่าความอ่อนไหว (Recall) และค่าวัดประสิทธิภาพโดยรวม (F1-Score) เพื่อวัดประสิทธิภาพของโมเดล โดยมุ่งเน้นไปที่การวัดสัดส่วนความถูกต้องของผลลัพธ์ที่โมเดลทำนายออกมา และนำผลลัพธ์ที่ได้ของระหว่างแต่ละโมเดลมาทำการเปรียบเทียบกัน เพื่อหาความแตกต่างของแต่ละโมเดล

2.2 งานวิจัยที่เกี่ยวข้อง

2.2.1 งานวิจัยที่มีหัวข้อเกี่ยวกับ AR

W. Wei et al. [20] ได้ศึกษาเกี่ยวกับ การวิเคราะห์ความสมมาตรและความสวยงามของใบหน้าด้วยเทคโนโลยีความจริงเสริม โดยมีวัตถุประสงค์เพื่อสร้างระบบที่ช่วยวัดความสมมาตร และวัดความสวยงามของใบหน้า โดยใช้เทคโนโลยี AR ผู้วิจัยใช้ Google Face API ในการตรวจจับจุดสังเกตบนใบหน้า (Landmarks) และทิศทางศีรษะ (Euler Angles) ซึ่งสามารถตรวจจับได้ 12 จุด แต่พบว่าการตรวจจับที่หยาบเกินไป จึงเลือกใช้งานเพียง 10 จุดหลัก จากนั้นทำการคำนวณพีเจอรเพื่อประเมินความสมมาตรจากจุด landmarks เมื่อพัฒนาพีเจอรเสร็จสิ้น ทำการทดสอบทำนายความดึงดูดของใบหน้าบุคคลจริง ด้วย 4 ชุดข้อมูลมาตรฐาน ได้แก่ SCUT-FBP, SCUT-FBP5500, FACES, และ Chicago Face Database (CFD) จากนั้นทำการวัดผลแบบ Regression ด้วยวิธี 5-Fold Cross-Validation และประเมินผลด้วย MAE (Mean Absolute Error) และ RMSE (Root Mean Squared Error) ผลการศึกษาปรากฏว่า ประสิทธิภาพของพีเจอรที่เสนอในงานวิจัยนี้ สามารถเทียบได้กับพีเจอรที่พัฒนาขึ้นด้วยชุดข้อมูลที่มีจุดสังเกตบนใบหน้าเยอะและซับซ้อนกว่า แสดงให้เห็นว่าใช้ทรัพยากรน้อยกว่าแต่สามารถให้ผลลัพธ์ที่ใกล้เคียงกันได้ รวมถึงสามารถนำไปพัฒนาเป็นแอปพลิเคชันบน Android OS ได้สำเร็จ และแอปพลิเคชันสามารถวัดภาพใบหน้าได้แบบ Real time โดยแอปพลิเคชันนี้สามารถนำไปใช้ในการช่วยเหลือนักแสดงในด้านการแสดงความสมมาตรของใบหน้าระหว่างการวางแผนเพื่อทำการผ่าตัดได้

D. Kang and L. Ma [21] ได้ศึกษาเกี่ยวกับ การติดตามการเคลื่อนไหวของดวงตาแบบเรียลไทม์สำหรับผู้สวมแว่นตาธรรมดาและแว่นตากันแดด เพื่อใช้ในระบบแสดงผล 3D Head-Up Displays (HUD) ในเทคโนโลยีความเป็นจริงเสริม (AR) โดยมีวัตถุประสงค์เพื่อพัฒนาเทคนิคการติดตามการเคลื่อนไหวของดวงตาที่มีความแม่นยำและเสถียร โดยไม่ขึ้นอยู่กับแว่นตาธรรมดาหรือแว่นตากันแดด ซึ่งเป็นความท้าทายในการใช้งาน AR HUD ในสภาพแวดล้อมจริง ผู้วิจัยใช้ AdaBoost Classifier

ร่วมกับ Local Binary Pattern (LBP) Feature ในการตรวจจับและจำแนกใบหน้าว่ามีการสวมแว่นกันแดดหรือไม่ โดยสร้างฐานข้อมูลภาพสำหรับฝึกด้วยภาพใบหน้าปกติ 50,000 ภาพ และภาพใบหน้าสวมแว่นกันแดด 30,000 ภาพ จากนั้นแยกเป็นสองเหตุการณ์ สำหรับภาพใบหน้าปกติ จะใช้วิธี Supervised Descent Method (SDM) พร้อมด้วย Scale-Invariant Feature Transform (SIFT) เพื่อติดตามรูปร่างใบหน้า และระบุตำแหน่งศูนย์กลางม่านตา กรณีเป็นภาพใบหน้าสวมแว่นกันแดด จะใช้การติดตามบริเวณใบหน้าที่ไม่ถูกบัง (จมูก ปาก ขอบใบหน้า) และใช้ Supervised Regression-based Pupil Position Estimation เพื่อประเมินตำแหน่งม่านตาที่อยู่หลังแว่น สุดท้ายทำการวัดประสิทธิภาพของระบบด้วยค่าความคลาดเคลื่อนเฉลี่ย (Mean Error)

ผลการศึกษาปรากฏว่า ประสิทธิภาพของระบบที่นำเสนอในงานวิจัยนี้ มีค่าความคลาดเคลื่อนเฉลี่ย (Mean Error) สำหรับภาพใบหน้าปกติอยู่ที่ 2 มิลลิเมตร และสำหรับภาพใบหน้าสวมแว่นกันแดดจะอยู่ที่ 10 มิลลิเมตร ซึ่งผลลัพธ์นี้สามารถนำไปใช้งานร่วมกับ AR 3D HUD ได้อย่างเหมาะสม

P. Riccio et al. [22] ได้ทำการพัฒนา OpenFilter: เฟรมเวิร์คสำหรับเปิดโอกาสในการเข้าถึงฟิลเตอร์ AR บนโซเซียลมีเดียสำหรับงานวิจัยอย่างเท่าเทียม โดยมีจุดประสงค์เพื่อพัฒนาชุดข้อมูลสาธารณะที่มีข้อมูลฟิลเตอร์ AR เพื่อเปิดโอกาสในการเข้าถึงข้อมูลสำหรับงานวิจัยในด้านต่างๆ ผู้วิจัยได้ใช้โปรแกรม Android Emulator บนคอมพิวเตอร์ โดยติดตั้งแอปพลิเคชันโซเซียลมีเดียที่ต้องการโดยใน Emulator นี้ สามารถเข้าถึงฟิลเตอร์ AR ที่โซเซียลมีเดียนั้นเปิดให้ใช้งาน แต่ฟิลเตอร์ส่วนใหญ่นั้นสามารถใช้งานได้เฉพาะภาพที่ถ่ายจากกล้อง ณ ขณะนั้นเท่านั้น จึงใช้เว็บแคมเสมือน ฉายภาพจากชุดข้อมูลที่มีไปยังกล้องของแอปพลิเคชันใน Emulator ทำให้สามารถใช้ฟิลเตอร์ AR กับภาพเหล่านั้นได้ จากนั้นใช้ระบบ Auto-Clicker เพื่อนำภาพจากชุดข้อมูลมาฉายผ่านกล้อง ใช้ฟิลเตอร์กับภาพนั้นๆ และบันทึกภาพที่ผ่านการใส่ฟิลเตอร์ไว้บนเครื่องคอมพิวเตอร์ ผลการศึกษาปรากฏว่า ด้วยวิธีการที่งานวิจัยนี้นำเสนอ ทำให้ได้ชุดข้อมูลตัวอย่างใหม่ 2 ชุด คือ FairBeauty (ดัดแปลงจากชุดข้อมูล FairFace โดยใส่ฟิลเตอร์ยอดนิยม 8 รูปแบบจากโซเซียลมีเดีย) และ B-LFW (ดัดแปลงจากชุดข้อมูล LFW โดยใส่ฟิลเตอร์จำนวนหลายฟิลเตอร์บนภาพ เพื่อศึกษาว่าฟิลเตอร์มีผลต่อระบบ Face Recognition หรือไม่)

2.1.2 งานวิจัยที่มีหัวข้อเกี่ยวกับ Image Detection หรือ Image Forensics

T. Kusnandar et al. [17] ได้ทำการศึกษาเทคนิคการเลือกสีที่มีประสิทธิภาพยิ่งขึ้นใน HSV Color Model โดยการประยุกต์ใช้ฟังก์ชันขั้นบันไดของเฮวีไซด์ (Heaviside Step Function) เพื่อเพิ่มประสิทธิภาพในการคัดแยกสีและความสว่างภาพเป้าหมาย ผลการศึกษาพบว่า การแยกช่องค่าสี (Hue) ออกจากค่าความเข้มของแสง (Intensity) และความอิ่มตัวสี (Saturation) ช่วยให้ค่าสีมีความเสถียรมากขึ้นและเพิ่มประสิทธิภาพในการประมวลผล งานวิจัยดังกล่าวจึงสนับสนุนแนวทางของงานวิจัยนี้ในการใช้แบบจำลองสี HSV เพื่อวิเคราะห์ค่าความสว่าง (Luminance) แยกจากข้อมูลสี ซึ่งช่วยให้

สามารถสกัดข้อมูลความสว่างได้โดยไม่ส่งผลกระทบต่อองค์ประกอบอื่นๆ ส่งผลให้การตรวจจับอุปกรณ์เสริมในระบบ AR มีความแม่นยำและเสถียรภายใต้สภาวะแสงที่ไม่คงที่

J. Yoneyama et al. [23] ได้ทำการศึกษาในหัวข้อ การสนทนาด้วยความจริงเสริม: ฟิลเตอร์ AR บนใบหน้าสำหรับเพิ่มความผ่อนคลายในการสนทนาแบบตัวต่อตัว โดยมีจุดประสงค์เพื่อเสนอแนวทางในการลดความวิตกกังวลของผู้มีสภาวะวิตกกังวลทางสังคมที่รู้สึกกังวลเมื่อต้องสนทนาแบบตัวต่อตัว ผู้วิจัยได้ใช้โปรแกรม Unity พัฒนาแอปพลิเคชันสำหรับใช้งานกับอุปกรณ์แว่นตา AR HoloLens2 โดยพัฒนาแอปพลิเคชันให้สามารถซ้อนภาพฟิลเตอร์ 2 แบบลงบนใบหน้าของคู่สนทนา แบ่งเป็นประเภทอวาตาร์สไตล์อนิเมะ (Anime-style avatar) ภาพใบหน้าแบบยิ้มแย้ม ในการตรวจจับใบหน้า ใช้ Dlib Face Landmark Detector ในการตรวจจับจุด Landmark บนใบหน้าทั้งหมด 17 จุด และทำการชิงค์การเคลื่อนไหวของตาและปากให้เคลื่อนไหวพร้อมกับใบหน้าของคู่สนทนาแบบเรียลไทม์ และทำการวัดผลผ่านแบบสอบถาม (แบบสอบถามวัดความกังวล หรือ STAI) และอัตราการเต้นของหัวใจ (วัดด้วย Apple Watch) ผลการศึกษาปรากฏว่า ด้วยวิธีการที่งานวิจัยนี้นำเสนอ การใช้ฟิลเตอร์ที่เป็นภาพอวาตาร์สไตล์อนิเมะ ช่วยลดความวิตกกังวล และเพิ่มความรู้สึกสบายใจในการสนทนาให้กับผู้เข้าทดสอบ โดยเฉพาะผู้ที่มีภาวะวิตกกังวลทางสังคมสูง ส่วนฟิลเตอร์ที่เป็นภาพใบหน้าที่ยิ้มแย้มนั้นได้ผลน้อยกว่า เนื่องจากผู้ทดสอบยังรับรู้ว่าภาพใบหน้าที่ยิ้มยังเคลื่อนไหวได้ไม่เหมือนจริง ทำให้ส่งผลด้านบวกน้อยกว่าฟิลเตอร์ที่เป็นภาพอวาตาร์สไตล์อนิเมะ

L. Nataraj et al. [24] ได้ทำการศึกษา การตรวจจับการปลอมแปลงภาพแบบองค์รวม โดยใช้เมตริกซ์การเกิดร่วมของพิกเซล โดยมีจุดประสงค์เพื่อพัฒนาวิธีการตรวจจับการปลอมแปลงภาพแบบองค์รวม (Holistic) โดยใช้เทคนิคการวิเคราะห์เมตริกซ์การเกิดร่วมของพิกเซล (Pixel Co-occurrence Matrices: PCM) ร่วมกับ Deep Learning เพื่อเพิ่มความแม่นยำในการตรวจจับภาพ ผู้วิจัยเริ่มจากการเตรียมชุดข้อมูลภาพจำนวนมากที่ประกอบไปด้วยภาพปกติและภาพที่ถูกปลอมแปลงในรูปแบบต่างๆ เช่น การตัดแปะ การลบภาพบางส่วน หรือการปรับแต่งอื่นๆ และทำการคำนวณเมตริกซ์ PCM ในแนวตั้งและแนวนอนของแต่ละช่องสี (RGB) เพื่อจับลักษณะและสถิติของพิกเซลที่บ่งบอกถึงความผิดปกติที่เกิดจากการปลอมแปลงข้อมูล จากนั้นใช้เมตริกซ์ PCM ที่ได้มา ไปใช้ในการฝึกโมเดล CNN เพื่อให้สามารถแยกแยะได้ระหว่างภาพปกติและภาพที่ผ่านการปลอมแปลง โดยใช้ภาพจำนวนกว่า 86,000 ภาพในการฝึกโมเดล และใช้ภาพที่เป็นชุดข้อมูลใหม่อีก 19,740 ภาพในการทดสอบ จากนั้นวัดความแม่นยำของโมเดลด้วย Area Under the Curve หรือ AUC ผลการศึกษาปรากฏว่า ระบบที่งานวิจัยนี้นำเสนอได้ผลความแม่นยำจาก AUC สูงกว่า 0.99 ในชุดข้อมูลฝึก และได้ประมาณ 0.81 ในชุดข้อมูลทดสอบ ซึ่งเป็นผลงานที่ดีที่สุด ณ ขณะนั้น

C. L. Li et al. [25] ได้ทำการศึกษาในหัวข้อ CutPaste: การเรียนรู้แบบ Self-Supervised สำหรับการตรวจจับและระบุตำแหน่งความผิดปกติ โดยมีวัตถุประสงค์เพื่อสร้างโมเดลตรวจจับความผิดปกติในภาพ โดยใช้เฉพาะข้อมูลภาพแบบปกติ (ไม่มีความผิดปกติในภาพ) และสามารถบอกได้ว่าจุดที่ผิดปกติอยู่ที่ส่วนใดในภาพ โดยใช้วิธี Self-Supervised Learning บนข้อมูลภาพปกติเพื่อเรียนรู้ฟีเจอร์ที่สามารถแยกแยะระหว่างภาพปกติ และภาพที่ไม่ปกติได้ดี ด้วยการใช้เทคนิค CutPaste ซึ่งทำการตัดส่วนหนึ่งของภาพ เอามาวางลงในตำแหน่งสุ่มในภาพเดียวกัน เพื่อสร้างภาพที่มีความผิดปกติด้วยตัวเอง จากนั้นทำการฝึกโมเดลให้สามารถแยก (Classification) ระหว่างภาพปกติ กับภาพที่ผ่านการใช้ Cut Paste หลังจากโมเดลได้เรียนรู้ฟีเจอร์การแยกภาพแล้ว ผู้วิจัยได้ใช้ข้อมูลภาพปกติในการสร้างโมเดล One-Class Classifier (ฝึกให้โมเดลรับรู้ข้อมูลเพียงคลาสเดียว เพื่อให้โมเดลเข้าใจว่าสิ่งใดคือความปกติ เมื่อเจอสิ่งที่ไม่เหมือนจากที่เรียนรู้ จะตีความว่าเป็นความผิดปกติ) บนฟีเจอร์ที่ทำการเรียนรู้มาแล้ว เพื่อใช้ในการตรวจสอบความผิดปกติในภาพใหม่ๆ นอกเหนือจากภาพที่ใช้ในการฝึกโมเดลก่อนหน้านี้ ผลการศึกษาปรากฏว่า เมื่อทดสอบด้วยชุดข้อมูล MVTEC Anomaly Detection ได้ผลความแม่นยำจาก AUC สูงกว่าวิธีก่อนหน้านี้ โดยมีความแม่นยำเฉลี่ยอยู่ที่ประมาณ 96.6% แม้ใช้เฉพาะรูปภาพปกติในการฝึกโมเดล และยังสามารถบอกตำแหน่งของจุดที่ผิดปกติได้ แม้ไม่ได้ระบุตำแหน่งให้ล่วงหน้า

จากงานวิจัย J. Yu et al. [26] ผู้วิจัยได้ทำการศึกษาในหัวข้อ FastFlow: การตรวจจับและระบุตำแหน่งความผิดปกติแบบไม่ต้องการกำกับ ด้วยการแปลงเชิงปกติสองมิติ (2D Normalizing Flows) โดยมีวัตถุประสงค์เพื่อพัฒนาวิธีที่สามารถตรวจจับและระบุตำแหน่งของความผิดปกติในภาพได้ โดยไม่ต้องใช้ข้อมูลที่มีการระบุความผิดปกติล่วงหน้า (Unsupervised Learning) โดยใช้เทคนิค 2D Normalizing Flows เพื่อเรียนรู้การกระจายของลักษณะภาพปกติอย่างแม่นยำ และนำมาใช้ในการระบุว่าภาพใหม่ๆ ผิดปกติหรือไม่ พร้อมกับชี้ตำแหน่งของสิ่งผิดปกตินั้น ผลการศึกษาปรากฏว่า เมื่อทดสอบด้วยชุดข้อมูล MVTEC Anomaly Detection ได้ผลความแม่นยำจาก AUC อยู่ที่ 99.4% สำหรับการตรวจจับความผิดปกติ และ 97.5% สำหรับการระบุตำแหน่งของความผิดปกติ)

D. Nimma et al. [27] ได้ทำการศึกษาในหัวข้อ นิติวิทยาศาสตร์ภาพขั้นสูง : การตรวจจับและสร้างภาพที่ถูกปรับแต่งด้วย Deep Learning เพื่อพัฒนาโมเดลที่สามารถแยกแยะภาพที่แท้จริงและภาพที่ถูกปรับแต่งด้วยความแม่นยำสูง ผู้วิจัยได้ใช้ Error Level Analysis (ELA) วิเคราะห์ความแตกต่างของระดับการบีบอัดภายในภาพ และใช้ Wavelet Denoising เพื่อลดสัญญาณรบกวนภายในภาพ จากนั้นส่งภาพที่เตรียมไว้เข้า Convolutional Neural Networks (CNNs) เพื่อให้เรียนรู้ความแตกต่างของภาพปกติและภาพที่ถูกแก้ไข และทำให้โมเดลสามารถจำแนกแยะระหว่างภาพปกติและภาพที่ถูกแก้ไขได้ จากนั้นใช้ Generative Adversarial Network (GAN) ระบุบริเวณที่ถูกแก้ไขในภาพ และฟื้นฟูให้ใกล้เคียงกับภาพจริง ผลการศึกษาปรากฏว่า โมเดล CNN สามารถจำแนกแยะระหว่างภาพปกติและภาพที่ถูกแก้ไขได้ และ Heatmap ที่ได้จากโมเดลสามารถระบุตำแหน่งที่ถูกแก้ไขได้อย่างแม่นยำ นอกจากนี้ GAN ยัง

สามารถฟื้นฟูภาพได้ใกล้เคียงกับภาพจริง สรุปได้ว่าวิธีการที่งานวิจัยนี้นำเสนอ มีความยืดหยุ่น สามารถนำไปใช้งานเพื่อพิสูจน์ความถูกต้องของภาพดิจิทัลได้อย่างมีประสิทธิภาพ

C. Shi et al. [28] ได้ทำการศึกษาในหัวข้อ การทบทวนเทคนิคการพิสูจน์ภาพด้วย Deep Learning โดยมีวัตถุประสงค์เพื่อทบทวนเทคนิคการพิสูจน์ภาพ (Image Forensics) ที่ใช้เทคโนโลยี Deep Learning ในช่วงหลายปีที่ผ่านมา และเปรียบเทียบกับเทคนิคการพิสูจน์ภาพแบบดั้งเดิม ผู้วิจัยได้ทำการศึกษาเกี่ยวกับเทคนิคของ Image Forensics โดยแบ่งเป็น 2 ประเภทคือ Passive Forensics และ Active Forensics โดยในส่วนของ Passive Forensics จะเกี่ยวกับการตรวจจับการปลอมแปลงภาพ ซึ่งมุ่งเน้นไปที่การตรวจจับและระบุพื้นที่ที่ถูกดัดแปลงในภาพโดยการวิเคราะห์ร่องรอยที่เกิดจากการแก้ไขภาพ และในส่วนของ Active Forensics จะเกี่ยวกับการฝังลายน้ำดิจิทัล (Digital Watermarking) ซึ่งมีการแนะนำโมเดลลายน้ำที่ทำงานแบบ end-to-end (สามารถฝังลายน้ำลงในภาพดิจิทัลและดึงลายน้ำออกจากภาพนั้นได้โดยตรงผ่าน Machine Learning หรือ Deep Learning) และจำแนกประเภทการโจมตีตามลักษณะของการโจมตีที่สามารถแยกแยะได้ เช่น การบีบอัดภาพ (JPEG Attacks) หรือการโจมตีจากการถ่ายภาพหน้าจอ (Screen-Shooting) ผลการศึกษาปรากฏว่า โมเดลที่ใช้ Deep Learning มีความสามารถในการตรวจจับและแก้ไขการปลอมแปลงได้ดีกว่าเทคนิคเก่าหลายเท่า และในส่วนของลายน้ำ เทคนิคที่ใช้การฝังลายน้ำสามารถป้องกันและปรับปรุงความทนทานต่อการโจมตีแบบต่างๆ ได้ดีขึ้น (เช่น การโจมตีจากการบีบอัดภาพ หรือการโจมตีจากการถ่ายภาพหน้าจอ) แต่ยังคงต้องพัฒนาเพิ่มเติมเพื่อรับมือกับการโจมตีที่ซับซ้อนขึ้นในอนาคต

M. Dang and T. N. Nguyen [29] ได้ทำการศึกษาในหัวข้อ การทบทวนเกี่ยวกับการหัวข้อ การสร้างและการตรวจจับการปรับแต่งใบหน้าดิจิทัล โดยมีวัตถุประสงค์เพื่อวิเคราะห์องค์รวมเกี่ยวกับวิธีการที่ใช้ในการสร้างภาพใบหน้าที่ถูกดัดแปลง โดยเน้นเทคโนโลยี Deepfake และเทคนิคใหม่ในการตรวจจับภาพปลอม ผู้วิจัยได้ทำการศึกษาเกี่ยวกับเทคนิคการสร้างใบหน้าที่ถูกดัดแปลงทั้งหมด 4 ประเภทคือ Attribute Manipulation (การเปลี่ยนลักษณะบางอย่างบนใบหน้าโดยไม่เปลี่ยนภาพบุคคล), การเลียนแบบการแสดงออกทางใบหน้า, การสลับใบหน้า และการสร้างภาพใบหน้าใหม่ทั้งหมด โดยตรวจสอบฐานข้อมูลที่ใช้บ่อย วิธีการสร้างและการตรวจจับใบหน้าที่ถูกดัดแปลง รวมถึงวิธีประเมินประสิทธิภาพของโมเดลหรือเทคนิคที่ใช้ ผลการศึกษาปรากฏว่า วิธีการตรวจจับที่มีการใช้บ่อยในหัวข้อเกี่ยวกับการตรวจจับใบหน้าคือ CNN, RNN, Vision Transformer และ Attention-Based Models มีการระบุถึงฐานข้อมูลที่มีการใช้บ่อยอย่าง FF++, Celeb-DF และ DeeperForensics จากนั้นทำการวิเคราะห์จุดแข็งและจุดอ่อนของแต่ละแนวทางการตรวจจับ

S. Polat et al. [30] ได้ทำการศึกษาในหัวข้อ วิธีการทำ Rule-based Classification โดยใช้คุณสมบัติรูปทรงของเส้นสเปกตรัม โดยมีวัตถุประสงค์เพื่อพัฒนาวิธีการจำแนกข้อมูลโดยใช้คุณสมบัติรูปทรงของเส้นสเปกตรัม (Spectral Curves) ซึ่งเป็นข้อมูลที่ได้จากการถ่ายภาพด้วยเทคนิค Hyperspectral

Imaging (HSI) เพื่อเพิ่มประสิทธิภาพในการจำแนกวัตถุหรือพืชในสภาพแวดล้อมที่มีความซับซ้อน ผลการศึกษาปรากฏว่า ระบบ Rule-based Classification ที่พัฒนาขึ้นสามารถจำแนกวัตถุจากสเปกตรัมได้อย่างแม่นยำ วัดผลด้วยการคำนวณสัดส่วนของตัวอย่างที่จำแนกถูกต้องทั้งหมด โดยมีความแม่นยำรวม (Overall Accuracy) ของชุดข้อมูลพลาสติกที่ 96.94% และชุดข้อมูลขยะจริงที่ 98.42%

P. Duszejko et al. [31] ได้ทำการศึกษางานวิจัยเกี่ยวกับวิธีการพิสูจน์หลักฐานภาพดิจิทัล (Digital Image Forensics) ทั้งในรูปแบบเชิงรุก (Active) และเชิงรับ (Passive) โดยมุ่งเน้นไปที่วิวัฒนาการของการตรวจจับด้วย Deep Learning ซึ่งผลการศึกษาพบว่า โมเดล Deep Learning มีประสิทธิภาพเหนือกว่าแนวทางทางสถิติแบบดั้งเดิม โดยสามารถระบุรูปแบบที่ซับซ้อนและความไม่สอดคล้องของระดับพิกเซลเพียงเล็กน้อยที่บ่งบอกถึงการปลอมแปลงภาพได้อย่างเป็นอิสระ งานวิจัยชิ้นนี้มีความเกี่ยวข้องกับงานวิจัยของเรา เนื่องจากการยืนยันถึงความจำเป็นในการใช้แบบจำลองขั้นสูง เช่น Xception เพื่อตรวจจับร่องรอยการดัดแปลงของ AR ที่มีความละเอียดสูง ซึ่งมักจะเกินขีดความสามารถในการตรวจจับของแบบจำลองที่ใช้ Rule-based models

S. Yadav and A. Kumar [32] ได้ทำการเปรียบเทียบวิธีการพิสูจน์หลักฐานแบบดั้งเดิมกับเทคนิคการเรียนรู้เชิงลึก (Deep Learning) ในการตรวจจับการดัดแปลงภาพ ผลการศึกษาปรากฏว่า กฎเกณฑ์ทางพิสูจน์หลักฐานแบบดั้งเดิมจะให้ความสามารถในการตีความที่สูง (อธิบายที่มาที่ได้ชัดเจน) แต่บ่อยครั้งมักถูกแซงหน้าโดยแบบจำลองการเรียนรู้เชิงลึก เช่น CNN (Convolutional Neural Networks) ซึ่งสามารถระบุความผิดปกติในระดับพิกเซลที่ซับซ้อนเกินกว่าที่กฎเกณฑ์แบบ Manual จะตรวจจับได้ ข้อมูลนี้สนับสนุนการนำการเรียนรู้แบบ Transfer Learning มาประยุกต์ใช้ในงานวิจัยของเรา เพื่อจัดการกับความซับซ้อนของร่องรอยการดัดแปลงที่ถูกสร้างขึ้นโดย AR

จากการศึกษาวิจัยที่เกี่ยวข้อง สามารถสรุปออกมาเป็นตารางโดยแบ่งเป็น 2 ตารางแยกตามประเภทของงานวิจัย ได้ดังนี้

ตารางที่ 2.2 สรุปงานวิจัยที่มีหัวข้อเกี่ยวกับ AR

ลำดับ	ชื่องานวิจัย	วัตถุประสงค์	วิธีการดำเนินงาน	ผลลัพธ์ที่ได้
1	Assessing Facial Symmetry and Attractiveness Using Augmented Reality [20]	เพื่อสร้างระบบที่ช่วยวัดความสมมาตร และวัดความสวยงามของใบหน้า โดยใช้เทคโนโลยี AR	1. ใช้ Google Face API จับจุดสังเกตบนใบหน้า 2. คำนวณพีเจอร์เพื่อประเมินความสมมาตรจากจุด landmarks 3. ทดสอบด้วยชุดข้อมูล 4 ชุด 4. วัดผลแบบ regression ด้วยวิธี 5-fold Cross-validation	ประสิทธิภาพของพีเจอร์เทียบได้กับพีเจอร์ที่พัฒนาขึ้นด้วยชุดข้อมูลที่มีจุดสังเกตบนใบหน้าเยอะและซับซ้อนกว่า และสามารถนำไปพัฒนาเป็นแอปพลิเคชันบน Android OS ได้สำเร็จ
2	Real-Time Eye Tracking for Bare and Sunglasses-Wearing Faces for Augmented Reality 3D Head-Up Displays [21]	พัฒนาระบบการติดตามการเคลื่อนไหวของดวงตาที่มีความแม่นยำและเสถียร	1. ใช้ AdaBoost classifier และ Local Binary Pattern (LBP) feature ในการตรวจจับใบหน้า 2. สร้างฐานข้อมูลภาพสำหรับฝึกโดยแยกเป็นใบหน้าปกติและใบหน้าสวมแว่นตา 3. กรณิภาพใบหน้า ใช้ Supervised Descent Method (SDM) และ Scale-Invariant Feature Transform (SIFT) 4. กรณิภาพแว่นกันแดด ใช้ Supervised Regression-based Pupil Position Estimation	ประสิทธิภาพของระบบที่นำเสนอในงานวิจัยนี้ มีค่าความคลาดเคลื่อนเฉลี่ย (Mean Error) สำหรับภาพใบหน้าปกติอยู่ที่ 2 มิลลิเมตร และสำหรับภาพใบหน้าสวมแว่นกันแดดจะอยู่ที่ 10 มิลลิเมตร

ตารางที่ 2.2 สรุปงานวิจัยที่มีหัวข้อเกี่ยวกับ AR (ต่อ)

ลำดับ	ชื่องานวิจัย	วัตถุประสงค์	วิธีการดำเนินงาน	ผลลัพธ์ที่ได้
3	OpenFilter: A Framework to Democratize Research Access to Social Media AR Filters [22]	พัฒนาชุดข้อมูลสาธารณะที่มีข้อมูลฟิลเตอร์ AR	1. ใช้ Android Emulator ติดตั้งโซเซียลมีเดีย 2. ใช้เว็บแคมเสมือนเพื่อให้สามารถใช้ฟิลเตอร์ AR บน Emulator ได้ 3. ใช้ Auto-Clicker นำภาพจากชุดข้อมูลมาฉายผ่านกล้อง	ได้ชุดข้อมูลตัวอย่างใหม่ 2 ชุด คือ FairBeauty (ดัดแปลงจากชุดข้อมูล FairFace) และ B-LFW (ดัดแปลงจากชุดข้อมูล LFW)

ตารางที่ 2.3 สรุปงานวิจัยที่มีหัวข้อเกี่ยวกับ Image Detection หรือ Image Forensics

ลำดับ	ชื่องานวิจัย	วัตถุประสงค์	วิธีการดำเนินงาน	ผลลัพธ์ที่ได้
1	Enhancing Color Selection in HSV Color Space [17]	เพื่อศึกษาวิธีการเพิ่มประสิทธิภาพในการคัดแยกสีและความสว่างภาพเป้าหมาย	1. แปลงภาพสีจากภาพ RGB เป็นภาพ HSV 2. ใช้สูตรคำนวณ กำหนดช่วงค่าสี (Hue) ของภาพส่วนที่ต้องการ เพื่อแยกสีที่ต้องการออกจากภาพ 3. เปรียบเทียบความแม่นยำและเวลาที่ใช้ในการประมวลผล ระหว่างวิธีที่นำเสนอกับวิธีแบบดั้งเดิม	วิธีที่งานวิจัยนี้เสนอ สามารถลดเวลาที่ใช้คำนวณเพื่อแยกสีได้ลดลง และแยกสีออกจากพื้นหลังในภาพได้อย่างชัดเจน

ตารางที่ 2.3 สรุปงานวิจัยที่มีหัวข้อเกี่ยวกับ Image Detection หรือ Image Forensics (ต่อ)

ลำดับ	ชื่องานวิจัย	วัตถุประสงค์	วิธีการดำเนินงาน	ผลลัพธ์ที่ได้
2	Augmented Conversations: AR Face Filters for Facilitating Comfortable Inperson Interactions [23]	เพื่อเสนอแนวทางในการลดความวิตกกังวลของผู้มีสภาวะวิตกกังวลทางสังคมที่รู้สึกกังวลเมื่อต้องสนทนาแบบตัวต่อตัว	1. ใช้ Unity พัฒนาแอปพลิเคชัน 2. ใช้ Dlib Face Landmark Detector ในการตรวจจับจุด Landmark บนใบหน้า 3. Sync การเคลื่อนไหวของฟิลเตอร์กับใบหน้า	การใช้ฟิลเตอร์ที่เป็นภาพอวตารสไตล์อนิเมะ ช่วยลดความวิตกกังวล และเพิ่มความรู้สึกสบายใจในการสนทนาให้กับผู้เข้าทดสอบ โดยเฉพาะผู้ที่มีภาวะวิตกกังวลทางสังคมสูง ส่วนฟิลเตอร์ที่เป็นภาพใบหน้าที่มึนนั้นได้ผลน้อยกว่า
3	Holistic Image Manipulation Detection using Pixel Cooccurrence Matrices [24]	พัฒนาวิธีการตรวจจับการปลอมแปลงภาพแบบองค์รวม (Holistic) โดยใช้เทคนิคการวิเคราะห์เมทริกซ์การเกิดร่วมของพิกเซล (Pixel Co-occurrence Matrices: PCM) ร่วมกับ Deep Learning	1. เตรียมชุดข้อมูลภาพ 2. คำนวณเมทริกซ์ PCM ในแนวดิ่งและแนวนอนของแต่ละช่องสี (RGB) 3. นำค่าเมทริกซ์ PCM ไปใช้ในการฝึกโมเดล CNN	ระบบได้ผลความแม่นยำจาก AUC สูงกว่า 0.99 ในชุดข้อมูลฝึก และได้ประมาณ 0.81 ในชุดข้อมูลทดสอบ

ตารางที่ 2.3 สรุปงานวิจัยที่มีหัวข้อเกี่ยวกับ Image Detection หรือ Image Forensics (ต่อ)

ลำดับ	ชื่องานวิจัย	วัตถุประสงค์	วิธีการดำเนินงาน	ผลลัพธ์ที่ได้
4	CutPaste: Self-Supervised Learning for Anomaly Detection and Localization [25]	เพื่อสร้างโมเดลตรวจจับความผิดปกติในภาพ โดยใช้เฉพาะข้อมูลภาพแบบปกติ (ไม่มี ความผิดปกติในภาพ) และสามารถบอกได้ว่าจุดที่ผิดปกติ อยู่ในส่วนใดในภาพ	1. ใช้วิธี Self-Supervised Learning เพื่อให้โมเดลเรียนรู้ค่าปกติ 2. ใช้เทคนิค CutPaste สร้างภาพที่มีความผิดปกติด้วยตัวเอง 3. ฟังก์ชันฟิเจอร์ให้สามารถแยกแยะระหว่างภาพปกติ กับภาพที่ผ่านการใช้ CutPaste ได้ 4. ใช้ข้อมูลภาพปกติในการสร้างโมเดล One-Class Classifier บนฟิเจอร์ที่ฝึกมาแล้ว	เมื่อทดสอบด้วยชุดข้อมูล MVTec Anomaly Detection ได้ผลความแม่นยำ จาก AUC สูงกว่าวิธีก่อนหน้านี้ โดยมีความแม่นยำเฉลี่ยอยู่ที่ประมาณ 96.6% แม้ใช้เฉพาะรูปภาพปกติในการฝึกโมเดล และยังสามารถบอกตำแหน่งของจุดที่ผิดปกติได้ แม้ไม่ได้ระบุตำแหน่งให้ล่วงหน้า
5	FastFlow: Unsupervised Anomaly Detection and Localization via 2D Normalizing Flows [26]	พัฒนาวิธีที่สามารถตรวจจับ และระบุตำแหน่งของความผิดปกติในภาพได้โดย ไม่ต้องใช้ข้อมูลที่มีการระบุความผิดปกติ ล่วงหน้า (Unsupervised Learning)	1. ใช้วิธีตรวจจับ FastFlow เพื่อดึง Feature map 2. ใช้โมเดล 2D normalizing flow เรียนรู้การกระจายความน่าจะเป็นของ ฟิเจอร์ปกติ 3. ใช้ heatmap แสดงบริเวณผิดปกติ ในภาพ	เมื่อทดสอบด้วยชุดข้อมูล MVTec Anomaly Detection ได้ผลความแม่นยำ จาก AUC อยู่ที่ 99.4% สำหรับการ ตรวจจับความผิดปกติ และ 97.5% สำหรับการระบุตำแหน่งของความผิดปกติ

ตารางที่ 2.3 สรุปงานวิจัยที่มีหัวข้อเกี่ยวกับ Image Detection หรือ Image Forensics (ต่อ)

ลำดับ	ชื่องานวิจัย	วัตถุประสงค์	วิธีการดำเนินงาน	ผลลัพธ์ที่ได้
6	Advanced Image Forensics: Detecting and Reconstructing Manipulated Images with Deep Learning [27]	พัฒนาโมเดลที่สามารถแยกแยะภาพจริงและภาพที่ถูกปรับแต่งด้วยความแม่นยำสูง	1. ใช้ Error Level Analysis (ELA) วิเคราะห์ความแตกต่างของระดับการบีบอัดภายในภาพ 2. ใช้ Wavelet Denoising เพื่อลดสัญญาณรบกวนภายในภาพ 3. ส่งภาพเข้า Convolutional Neural Networks (CNNs) เพื่อให้เรียนรู้ความแตกต่างของภาพปกติและภาพที่ถูกแก้ไข 4. ใช้ Generative Adversarial Network (GAN) ระบุบริเวณที่ถูกแก้ไขในภาพ และฟื้นฟูภาพ	โมเดล CNN สามารถจำแนกระหว่างภาพปกติและภาพที่ถูกแก้ไขได้ และ Heatmap ที่ได้จากโมเดลสามารถระบุตำแหน่งที่ถูกแก้ไขได้อย่างแม่นยำ นอกจากนี้ GAN ยังสามารถฟื้นฟูภาพได้ใกล้เคียงกับภาพจริง

ตารางที่ 2.3 สรุปงานวิจัยที่มีหัวข้อเกี่ยวกับ Image Detection หรือ Image Forensics (ต่อ)

ลำดับ	ชื่องานวิจัย	วัตถุประสงค์	วิธีการดำเนินงาน	ผลลัพธ์ที่ได้
7	Review of Image Forensic Techniques Based on Deep Learning [28]	ทบทวนเทคนิคการพิสูจน์ภาพ (Image Forensics) ที่ใช้เทคโนโลยี Deep Learning	1. พูดยถึงเทคนิคการปลอมแปลงภาพหลายรูปแบบ เช่น การปลอมแปลงแบบคัดลอกและวาง (Image Copy-Move Forgery Detection: CMFD) หรือ การปลอมแบบ Splicing (คัดลอกพื้นที่จากภาพหนึ่งแล้วนำไปวางลงในอีกภาพหนึ่ง) 2. อธิบายวิธีการตรวจจับภาพปลอมแปลงแต่ละชนิด โดยแบ่งเป็นหัวข้อใหญ่ๆ 2 หัวข้อ ได้แก่ การตรวจจับด้วยวิธีแบบดั้งเดิม (manual feature-based) และการตรวจจับโดยใช้ Deep Learning (DL-based) ซึ่งสำหรับแต่ละการปลอมแปลงภาพจะมีรายละเอียดที่แตกต่างกันออกไป (เช่น ใช้การตรวจจับ Noise ของภาพสำหรับการตรวจจับแบบดั้งเดิม หรือการเรียนรู้ฟีเจอร์ของภาพโดยอัตโนมัติของ Deep Learning)	โมเดลที่ใช้ Deep Learning มีความสามารถในการตรวจจับและแก้ไขการปลอมแปลงได้ดีกว่าเทคนิคเก่า

ตารางที่ 2.3 สรุปงานวิจัยที่มีหัวข้อเกี่ยวกับ Image Detection หรือ Image Forensics (ต่อ)

ลำดับ	ชื่องานวิจัย	วัตถุประสงค์	วิธีการดำเนินงาน	ผลลัพธ์ที่ได้
8	Digital Face Manipulation Creation and Detection: A Systematic Review [29]	วิเคราะห์องค์รวมเกี่ยวกับวิธีการที่ใช้ในการสร้างภาพใบหน้าที่ถูกดัดแปลง โดยเน้นเทคโนโลยี Deepfake และเทคนิคใหม่ในการตรวจจับภาพปลอม	1. มีการพูดถึงเทคนิคการปลอมแปลงภาพใบหน้า เช่น การสังเคราะห์ใบหน้า (Face Synthesis), การสลับใบหน้า (Face Swapping), การแก้ไขคุณลักษณะของใบหน้า (Face Attribute Editing) และการจำลองการแสดงออกทางใบหน้า (Facial Re-enactment) รวมถึง Dataset ที่นิยมใช้ในแต่ละเทคนิค 2. มีการเจาะลึกในแต่ละเทคนิคการปลอมแปลงใบหน้า โดยอธิบายหลักการทำงานและวิธีการตรวจจับภาพในแต่ละประเภท โดยส่วนมากจะพูดถึงการใช้โมเดล Deep Learning ในการตรวจจับภาพ และมีการระบุถึงโมเดลที่มีการพูดถึงบ่อยในการตรวจจับภาพแต่ละประเภท	วิธีการตรวจจับที่มีการใช้บ่อยในหัวข้อเกี่ยวกับการตรวจจับใบหน้าคือ CNN, RNN, Vision Transformer และ Attention-Based Models มีการระบุถึงฐานข้อมูลที่มีการใช้บ่อยอย่าง FF++, Celeb-DF และ DeeperForensics จากนั้นทำการวิเคราะห์จุดแข็งและจุดอ่อนของแต่ละแนวทางการตรวจจับ

ตารางที่ 2.3 สรุปงานวิจัยที่มีหัวข้อเกี่ยวกับ Image Detection หรือ Image Forensics (ต่อ)

ลำดับ	ชื่องานวิจัย	วัตถุประสงค์	วิธีการดำเนินงาน	ผลลัพธ์ที่ได้
9	A New Rule-Based Classification Method Using Shape-Based Properties of Spectral Curves [30]	พัฒนาวิธีการจำแนกข้อมูลโดยใช้คุณสมบัติรูปทรงของเส้นสเปกตรัม โดยใช้ Rule-based	1. ถ่ายภาพวัตถุพลาสติกและพืชด้วยกล้อง Hyperspectral 2. สร้างกฎเกณฑ์ขึ้นจากการวิเคราะห์คุณสมบัติของข้อมูล โดยวิเคราะห์ลักษณะโค้งของเส้นสเปกตรัมแต่ละช่วง และวิเคราะห์ความเข้มของสเปกตรัม 3. ตรวจสอบคุณสมบัติของสเปกตรัมที่ได้จากภาพ นำค่าที่ตรวจสอบได้ไปเทียบกับเงื่อนไขในกฎเพื่อจำแนกวัตถุ 4. ทดสอบด้วยชุดข้อมูลประเภทพลาสติก และชุดข้อมูลประเภทพืช	ระบบ Rule-based Classification ที่พัฒนาขึ้นสามารถจำแนกวัตถุจากสเปกตรัมได้อย่างแม่นยำ โดยมีความแม่นยำของชุดข้อมูลพลาสติกที่ 96.94% และชุดข้อมูลพืชจริงที่ 98.42%

ตารางที่ 2.3 สรุปงานวิจัยที่มีหัวข้อเกี่ยวกับ Image Detection หรือ Image Forensics (ต่อ)

ลำดับ	ชื่องานวิจัย	วัตถุประสงค์	วิธีการดำเนินงาน	ผลลัพธ์ที่ได้
10	Detection of Manipulations in Digital Images: A Review of Passive and Active Methods [31]	ทบทวนและวิเคราะห์วิธีการตรวจสอบความถูกต้องของภาพดิจิทัล และเปรียบเทียบประสิทธิภาพเทคนิคการตรวจจ็บบรูปแบบต่างๆ โดยเน้นไปที่การประยุกต์ใช้ Deep Learning	- รวบรวมงานวิจัยที่เกี่ยวข้องกับการตรวจจ็บบการปลอมแปลงภาพ (Image Manipulation Detection) และการระบุตำแหน่งที่มีการแก้ไข (Localization) - แบ่งกลุ่มวิธีการตรวจสอบออกเป็น 2 แนวทางหลัก คือ Passive Methods (เน้นการวิเคราะห์คุณสมบัติทางสถิติของภาพและร่องรอยความไม่สอดคล้องระดับพิกเซลที่เกิดขึ้นจากการตัดต่อ) และ Active Methods (เน้นการใช้ลายน้ำดิจิทัลหรือข้อมูลที่ฝังซ่อนไว้ในโครงสร้างภาพเพื่อยืนยันแหล่งที่มา) - ประเมินความสามารถของโมเดล Deep Learning ในสถานการณ์ต่างๆ - สรุปวิวัฒนาการของโมเดลตรวจจ็บบตั้งแต่แนวทางสถิติดั้งเดิมไปจนถึงโมเดล Deep Learning	โมเดล Deep Learning มีความสามารถเหนือกว่าวิธีดั้งเดิมในการระบุรูปแบบการปลอมแปลงที่ซับซ้อนและตรวจจ็บบร่องรอยพิกเซลได้โดยอัตโนมัติ

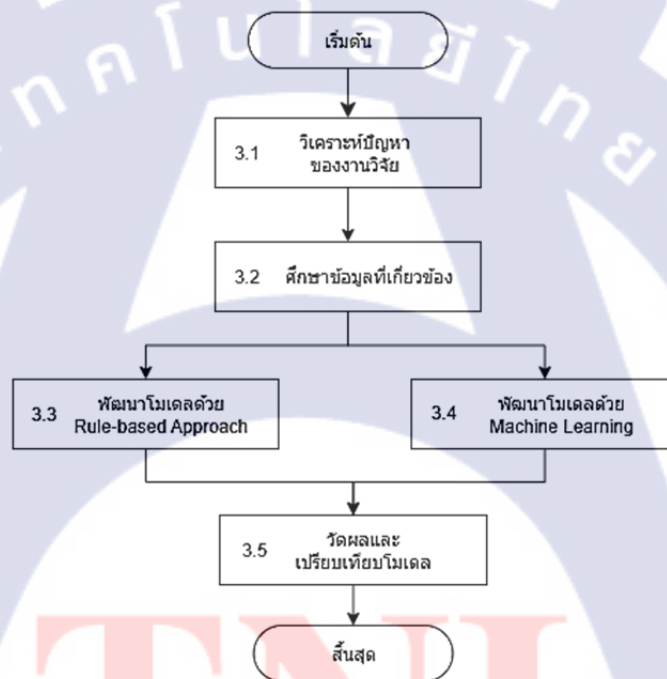
ตารางที่ 2.3 สรุปงานวิจัยที่มีหัวข้อเกี่ยวกับ Image Detection หรือ Image Forensics (ต่อ)

ลำดับ	ชื่องานวิจัย	วัตถุประสงค์	วิธีการดำเนินงาน	ผลลัพธ์ที่ได้
11	Combining Deep Learning and Forensic Analysis for Robust Fake Image Detection [32]	นำเสนอวิธีการแบบผสมผสานในการตรวจจับภาพปลอม โดยการผสมผสานจุดแข็งของโครงข่ายประสาทเทียมแบบคอนโวลูชัน (CNN) และการวิเคราะห์ทางฟิสิกส์พื้นฐานภาพถ่ายดิจิทัลแบบดั้งเดิม (Conventional Image Forensics)	- เตรียมข้อมูลภาพ - ฝึกพีเจอาร์โดยใช้ CNN เพื่อให้เรียนรู้ภาพในระดับพิกเซล - ใช้ Error Level Analysis (ELA), Photo-response non-uniformity noise (PRNU) และ Metadata & Frequency Domain วิเคราะห์คุณลักษณะของภาพ - นำผลการทำนายจาก CNN และการวิเคราะห์ทางฟิสิกส์พื้นฐานมารวมกันโดยการกำหนดค่าน้ำหนักความสำคัญของแต่ละผลลัพธ์	วิธีการแบบผสมผสานที่งานวิจัยนำเสนอ ทำคะแนนความแม่นยำได้สูงถึง 95.8% และมีค่า AUC อยู่ที่ 0.971 ซึ่งสูงกว่าการใช้ CNN เพียงอย่างเดียว (91.3%) หรือการใช้วิธีทางฟิสิกส์พื้นฐานเพียงอย่างเดียว (83.6%)

บทที่ 3

กระบวนการวิจัย

งานวิจัยนี้เกี่ยวข้องกับการเปรียบเทียบประสิทธิภาพของโมเดลที่ถูกพัฒนาขึ้นเพื่อการตรวจจับภาพความจริงเสริม (Augmented Reality) ที่มีการเพิ่มเครื่องประดับในบริเวณใบหน้ามีวัตถุประสงค์เพื่อพัฒนาโมเดลที่สามารถตรวจจับภาพที่มีการปรับแต่งด้วย AR ได้ และทดลองเปรียบเทียบประสิทธิภาพของวิธีการพัฒนาโมเดล 2 รูปแบบ ระหว่างการพัฒนาโมเดลโดยใช้วิธี Rule-based Approach และการพัฒนาโมเดลด้วย Machine Learning โดยมีภาพรวมขั้นตอนการดำเนินงาน ดังที่แสดงในรูปที่ 3.1



รูปที่ 3.1 แผนผังภาพรวมขั้นตอนดำเนินงานวิจัย

การดำเนินงานวิจัยแบ่งเป็น 5 ขั้นตอนใหญ่ โดยเริ่มต้นจากการวิเคราะห์ปัญหาของงานวิจัย จากนั้นทำการศึกษาข้อมูลที่เกี่ยวข้อง และเริ่มทำการพัฒนาโมเดลด้วยวิธีการทั้ง 2 รูปแบบ คือการพัฒนาโมเดลด้วย Rule-based Approach และการพัฒนาโมเดลด้วย Machine Learning จากนั้นจึงทำการประเมินและวัดผลประสิทธิภาพของโมเดลที่ถูกพัฒนาด้วยวิธีการทั้ง 2 รูปแบบ และนำผลลัพธ์มาเปรียบเทียบระหว่าง 2 วิธีการ

3.1 การวิเคราะห์ปัญหาของงานวิจัย

ผู้วิจัยได้ทำการศึกษาและวิเคราะห์ปัญหาของงานวิจัยเพื่อทำความเข้าใจเกี่ยวกับปัญหาที่เกิดขึ้น จากข้อมูลที่ได้พบว่าปัจจุบันยังไม่มีเครื่องมือที่ทำการตรวจจับภาพ AR โดยเฉพาะที่สามารถใช้งานได้ทั่วไป ผู้วิจัยจึงมีความสนใจและต้องการทดลองพัฒนาเครื่องมือสำหรับการตรวจจับภาพ AR โดยเฉพาะ โดยทำการกำหนดเงื่อนไขเบื้องต้นเกี่ยวกับภาพใบหน้ามนุษย์และเครื่องประดับบนใบหน้า เนื่องจากเทคโนโลยี AR ที่ใกล้ตัวและมีการใช้งานโดยทั่วไปในชีวิตประจำวันส่วนมากจะเกี่ยวกับการใช้ฟิลเตอร์ AR ในการถ่ายภาพใบหน้าสำหรับเผยแพร่บนโซเชียล เมื่อวิเคราะห์จนได้ข้อสรุปเกี่ยวกับขอบเขตของงานวิจัยแล้ว จากนั้นจึงทำการศึกษาข้อมูลเบื้องต้นเพิ่มเติมเกี่ยวกับวิธีการดำเนินงานในขั้นตอนต่อไป

3.2 การศึกษาข้อมูลที่เกี่ยวข้อง

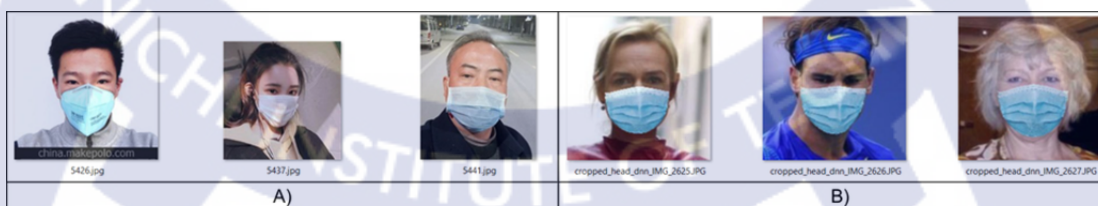
ในขั้นตอนนี้ ผู้วิจัยได้ทำการศึกษาข้อมูลเบื้องต้นที่เกี่ยวข้องกับงานวิจัย รวมถึงการค้นคว้างานวิจัยในอดีต และศึกษาเพิ่มเติมเกี่ยวกับทฤษฎีและวิธีการที่มีส่วนเกี่ยวข้องกับหัวข้องานวิจัย

3.2.1 วิธีการเตรียม Dataset

Dataset สำหรับงานวิจัยนี้ ประกอบไปด้วยข้อมูลภาพใบหน้าในสถานการณ์ต่างๆ ซึ่งส่งผลต่อการตรวจสอบภาพ โดยมีประเภทของข้อมูลดังต่อไปนี้

1. ภาพใบหน้าที่สวมเครื่องประดับ

กำหนดให้ภายใน Dataset มีภาพใบหน้าที่สวมเครื่องประดับตามปกติ เพื่อให้โมเดลสามารถเรียนรู้และแยกแยะความเป็นธรรมชาติของภาพที่มีเครื่องประดับได้ โดยเครื่องประดับที่จะทำการมุ่งเน้นในงานวิจัยนี้คือ หน้ากากอนามัย และแว่นตา ในส่วนนี้จะนำภาพจาก Dataset Celeb A [5] และ Dataset Real-Time-Medical-Mask-Detection [6] ซึ่งประกอบไปด้วยภาพใบหน้าคนที่สวมหน้ากากอนามัย และภาพใบหน้าคนที่สวมแว่นตารวมเข้าด้วยกัน เป็นชุดข้อมูลภาพที่ประกอบไปด้วยภาพใบหน้าคนที่สวมเครื่องประดับ 2 ชนิด จำนวน 1,000 ภาพ โดยมีภาพตัวอย่างข้อมูลภายใน Dataset ดังที่แสดงในรูปที่ 3.2 และรูปที่ 3.3 ตามลำดับ



รูปที่ 3.2 ตัวอย่างภาพใบหน้าที่สวมหน้ากาก (A) ภาพสวมหน้ากากจริง (B) ภาพสวมหน้ากาก AR



รูปที่ 3.3 ตัวอย่างภาพใบหน้าสวมแว่นตา (A) ภาพสวมแว่นตาจริง (B) ภาพสวมแว่นตา AR

2. ภาพใบหน้าที่เพิ่มเครื่องประดับด้วยภาพ AR

นอกเหนือจากภาพใบหน้าที่สวมเครื่องประดับตามปกติแล้ว กำหนดให้ภายใน Dataset มีภาพใบหน้าที่เพิ่มเครื่องประดับด้วยภาพ AR เพื่อให้โมเดลสามารถเรียนรู้ความแตกต่างระหว่างภาพที่สวมเครื่องประดับตามปกติกับภาพที่ถูกเพิ่มด้วย AR ได้ ในส่วนนี้จะนำภาพจาก Dataset CelebA [5] และ Dataset Real-Time-Medical-Mask-Detection [6] มาทำการสุ่มภาพและเพิ่มเครื่องประดับ AR ลงไปบนภาพโดยอัตโนมัติด้วยฟิลเตอร์ของแอปพลิเคชัน Snapchat ซึ่ง

ทำการเพิ่มหน้ากากอนามัยและแว่นตา โดยรูปภาพหนึ่งรูป จะเพิ่มเครื่องประดับเพียง 1 ชนิดเท่านั้น

ในส่วนของการละเอียดข้อมูลภาพ Dataset นี้จะประกอบไปด้วย ภาพใบหน้าที่สวมเครื่องประดับทั้งหมด 1,000 ภาพ (ภาพสวมหน้ากากอนามัย 500 ภาพ และภาพสวมแว่นตา 500 ภาพ) และภาพใบหน้าที่เพิ่มเครื่องประดับด้วยภาพ AR 1,000 ภาพ (ภาพสวมหน้ากากอนามัย 500 ภาพ และภาพสวมแว่นตา 500 ภาพ) รวมทั้งหมดเป็น 2,000 ภาพ ซึ่งมีการแบ่งสัดส่วน Dataset นี้ ดังที่แสดงในรูปที่ 3.4



รูปที่ 3.4 ภาพรวมข้อมูลภาพภายใน Dataset ที่ใช้ฝึกโมเดล Machine Learning

Dataset ที่เตรียมขึ้นนี้ จะถูกนำไปใช้งานในการฝึกและทดสอบโมเดลสำหรับตรวจจับภาพ เพื่อทำการแบ่งกลุ่ม (Classification) ระหว่างภาพที่สวมเครื่องประดับตามธรรมชาติ และภาพที่ถูกเพิ่มเครื่องประดับด้วย AR โดยโมเดลที่พัฒนาด้วย Machine Learning จะใช้ Dataset ส่วนนี้ในการฝึกโมเดล แม้ว่าชุดข้อมูลจำนวน 2,000 ภาพจะมีความสมดุลในเชิงโครงสร้าง แต่ปริมาณดังกล่าวยังอาจไม่เพียงพอสำหรับการสร้างการเรียนรู้เชิงลึกของโมเดลที่มีความซับซ้อนสูง นอกจากนี้ การเลือกใช้ข้อมูลจากฟิลเตอร์ของแอปพลิเคชัน Snapchat เพียงอย่างเดียว ส่งผลให้ผลลัพธ์ที่ได้ถูกจำกัดอยู่ภายใต้กลไกการแสดงผลเฉพาะเจาะจงเพียงรูปแบบเดียว ซึ่งอาจส่งผลกระทบต่อประสิทธิภาพของโมเดลเมื่อนำไปประยุกต์ใช้กับแพลตฟอร์มอื่น เช่น TikTok หรือ Instagram ดังนั้น การตรวจสอบความถูกต้องในอนาคตด้วยชุดข้อมูลที่มีขนาดใหญ่และมีความหลากหลายมากขึ้น จึงมีความจำเป็นอย่างยิ่ง เพื่อรับประกันความน่าเชื่อถือและความแม่นยำเมื่อนำไปใช้งานในสถานการณ์จริง

ในส่วนของ Dataset ที่ใช้กับโมเดลที่พัฒนาด้วย Rule-Based จะใช้ข้อมูลที่เป็นส่วนหนึ่งของ Dataset ที่ใช้กับ Machine Learning (Dataset ที่มี 2,000 ภาพ) โดยจะรวมภาพถ่ายจากมือถือเข้าไปใน Dataset และทำการกำหนดเงื่อนไขภาพที่ใช้ โดยภาพจะต้องประกอบไปด้วยภาพที่สวมเครื่องประดับ คือ แว่นตาและหน้ากากอนามัย โดยแบ่งสัดส่วนระหว่างภาพที่สวมเครื่องประดับจริงและเครื่องประดับ AR เท่าๆ กัน (ทั้งภาพจาก Dataset และภาพถ่ายจากมือถือ)

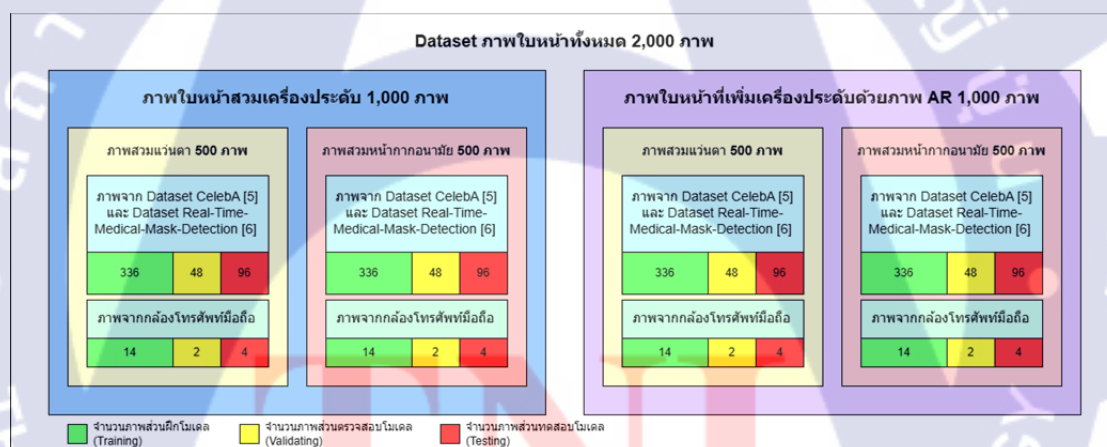
ใน Dataset ส่วนที่ใช้กับโมเดลที่พัฒนาด้วย Rule-Based นั้น จะนำข้อมูลจาก Dataset CelebA [5] และ Dataset Real-Time-Medical-Mask-Detection [6] มาประกอบกับภาพถ่ายจากมือถือ แบ่งข้อมูลภาพเป็น 2 ส่วน ได้แก่

1. ภาพส่วนที่ใช้ในการค้นหากฎเกณฑ์การตรวจจับภาพ 120 ภาพ แบ่งเป็นภาพสวมหน้ากาก 60 ภาพ (ภาพสวมหน้ากากจริง 30 ภาพ และภาพสวมหน้ากาก AR 30 ภาพ) และภาพสวมแว่นตา 60 ภาพ (ภาพสวมแว่นตาจริง 30 ภาพ และภาพสวมแว่นตา AR 30 ภาพ)
2. ภาพส่วนที่ใช้ในการทดสอบกฎเกณฑ์การตรวจจับภาพ 400 ภาพ แบ่งเป็นภาพสวมหน้ากาก 200 ภาพ (ภาพสวมหน้ากากจริง 100 ภาพ และภาพสวมหน้ากาก AR 100 ภาพ) และภาพสวมแว่นตา 200 ภาพ (ภาพสวมแว่นตาจริง 100 ภาพ และภาพสวมแว่นตา AR 100 ภาพ)

จากหนังสือทางสถิติ Statistical Inference [33] มีการกล่าวถึงทฤษฎี Central Limit Theorem (CLT) ที่กล่าวถึงสถิติเกี่ยวกับจำนวนตัวอย่างที่เหมาะสมเพื่อให้การวิเคราะห์เชิงสถิติสามารถเชื่อถือได้ ทฤษฎีดังกล่าวสนับสนุนว่าการใช้ข้อมูลที่มีจำนวนมากพอ (ไม่ได้ระบุจำนวนตัวอย่างที่แน่นอน แต่ในทางปฏิบัติมักใช้เกณฑ์ที่จำนวนตัวอย่างเท่ากับหรือมากกว่า 30 ตัวอย่าง) จะทำให้การแจกแจงของค่าเฉลี่ยตัวอย่างใกล้เคียงการแจกแจงปกติ ซึ่งช่วยให้ผลการประเมินมีความน่าเชื่อถือมากขึ้น และลดความเสี่ยงจากความผันผวนของข้อมูลตัวอย่างที่มีขนาดเล็ก ผู้วิจัยจึงกำหนดจำนวน Dataset ที่ใช้ในการค้นหากฎเกณฑ์การตรวจจับภาพไว้ที่ 120 ภาพ และกำหนดจำนวนภาพที่ใช้ในการทดสอบ

กฎเกณฑ์ไว้ที่ 400 ภาพ ซึ่งแบ่งย่อยออกมาแล้วจะได้ประเภทภาพละ 100 ภาพ ทำให้สามารถวัดผลการทดสอบโดยเทียบเป็นเปอร์เซ็นต์ความถูกต้องได้โดยตรง ซึ่งภาพทั้ง 2 ส่วนรวม 520 ภาพนี้จะเป็นส่วนหนึ่งของ Dataset ทั้งหมด 2,000 ภาพที่ใช้ในการฝึกโมเดลด้วย Machine Learning ด้วยเช่นกัน

ในส่วนของการฝึกโมเดล การตรวจสอบโมเดล และการทดสอบโมเดล จะทำการแบ่งข้อมูลภายใน Dataset อยู่ที่สัดส่วน 70 เปอร์เซ็นต์ ต่อ 10 เปอร์เซ็นต์ ต่อ 20 เปอร์เซ็นต์ ตามลำดับในทุกๆ ส่วนของ Dataset โดยทำการแบ่งจากจำนวนภาพทั้งหมด 2,000 ภาพ ซึ่งเมื่อแบ่งสัดส่วนแยกตามแหล่งที่มาของภาพแล้ว ในหนึ่งส่วนเครื่องประดับที่มี 500 ภาพ ภาพส่วนที่มาจาก Dataset CelebA [5] และ Dataset Real-Time-Medical-Mask-Detection [6] จะถูกแบ่งเป็นส่วนฝึกโมเดล 336 ภาพ ส่วนตรวจสอบโมเดล 48 ภาพ และส่วนทดสอบโมเดล 96 ภาพ ตามลำดับ รวม 480 ภาพ และภาพส่วนที่มาจากกล้องโทรศัพท์มือถือ จะถูกแบ่งเป็นส่วนฝึกโมเดล 14 ภาพ ส่วนตรวจสอบโมเดล 2 ภาพ และส่วนทดสอบโมเดล 4 ภาพ ตามลำดับ รวม 20 ภาพ โดยสัดส่วนนี้ถูกแบ่งเท่ากันในทุกๆ เครื่องประดับทั้งภาพส่วนที่เป็นใบหน้าสวมเครื่องประดับจริง และภาพส่วนที่เพิ่มเครื่องประดับด้วยภาพ AR ดังที่แสดงในรูปที่ 3.5



รูปที่ 3.5 การแบ่งสัดส่วนข้อมูลภาพภายใน Dataset ตามขั้นตอนการพัฒนาโมเดล

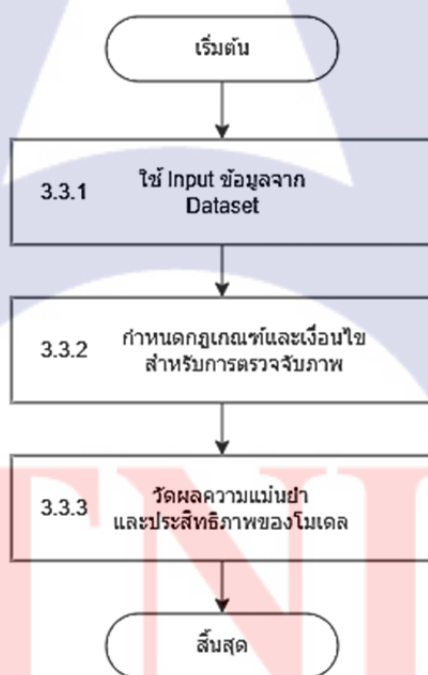
3.2.2 วิธีการตรวจจับตำแหน่งใบหน้าในภาพ

การตรวจจับตำแหน่งใบหน้าในภาพ เป็นขั้นตอนแรกและขั้นตอนสำคัญในการระบุพื้นที่ข้อมูลที่ต้องการตรวจสอบข้อมูลในขั้นตอนถัดไป เนื่องจากในงานวิจัยนี้จะมุ่งเน้นไปที่การตรวจจับเครื่องประดับบนใบหน้า การระบุตำแหน่งหรือพื้นที่ของใบหน้าจึงเป็นสิ่งสำคัญ ซึ่งโมเดลที่ผู้วิจัยเลือกใช้ในการตรวจจับตำแหน่งใบหน้าในภาพ คือ Multi-task Cascaded Convolutional Neural Network (MTCNN) ซึ่งเป็นโมเดลที่สามารถตรวจจับตำแหน่งใบหน้าในภาพได้ด้วยการระบุตำแหน่งของ Landmark

บนใบหน้า เช่น ตำแหน่งของดวงตา จมูกและปาก และเป็นโมเดลที่มีความแม่นยำ มีการระบุตำแหน่ง Landmark บนใบหน้า ซึ่งเหมาะกับการใช้งานที่เกี่ยวข้องกับเทคโนโลยี AR ที่อาจมีการแต่งเติมวัตถุบางอย่างลงไปตามตำแหน่งของใบหน้า เช่น การเพิ่มแว่นตาลงบนใบหน้า เมื่อรู้ Landmark ตำแหน่งของดวงตาแล้วจะทำให้การเพิ่มวัตถุสามารถทำได้แม่นยำมากขึ้น เป็นต้น

3.3 การพัฒนาโมเดลด้วย Rule-based

การพัฒนาโมเดลด้วยวิธีการแบบ Rule-based คือการสร้างโมเดลด้วยการกำหนดกฎเกณฑ์หรือเงื่อนไขให้กับโมเดลโดยตรง เพื่อให้โมเดลคำนวณและให้ผลลัพธ์ตามเงื่อนไขที่กำหนดไว้ จุดเด่นของโมเดลนี้คือสามารถอธิบายได้ว่าการตรวจจับภาพแต่ละประเภท มีความแตกต่างของข้อมูลอย่างไรบ้าง เนื่องจากผู้วิจัยต้องทำการกำหนดกฎเกณฑ์และเงื่อนไขต่างๆ ด้วยตัวเอง โดยมีขั้นตอนทั้งหมด 3 ขั้นตอน ดังที่แสดงในรูปที่ 3.6



รูปที่ 3.6 แผนผังขั้นตอนการพัฒนาโมเดลด้วย Rule-based

3.3.1 ใช้ Input ข้อมูลจาก Dataset

ใช้ภาพจาก Dataset ที่เตรียมไว้ในขั้นตอนก่อนหน้านี้ โดยแบ่งภาพที่ใช้ออกเป็น 2 ประเภทหลัก

1. ภาพส่วนสำหรับใช้ในการค้นหากฎเกณฑ์ในการตรวจจับภาพ 120 ภาพ

ใช้ภาพใบหน้าสวมเครื่องประดับจริง 60 ภาพ (แบ่งเป็นภาพสวมแว่นตา 30 ภาพ และภาพสวมหน้ากาก 30 ภาพ) และภาพใบหน้าสวมเครื่องประดับ AR อีก 60 ภาพ (แบ่งเป็นภาพสวมแว่นตา 30 ภาพ และภาพสวมหน้ากาก 30 ภาพ)

2. ภาพส่วนสำหรับใช้ในการทดสอบเกณฑ์ในการตรวจจับภาพ 400 ภาพ

ใช้ภาพใบหน้าสวมเครื่องประดับจริง 200 ภาพ (แบ่งเป็นภาพสวมแว่นตา 100 ภาพ และภาพสวมหน้ากาก 100 ภาพ) และภาพใบหน้าสวมเครื่องประดับ AR อีก 200 ภาพ (แบ่งเป็นภาพสวมแว่นตา 100 ภาพ และภาพสวมหน้ากาก 100 ภาพ) โดยใน 400 ภาพนี้ จะไม่มีภาพในส่วนที่ใช้สำหรับการค้นหากฎเกณฑ์รวมอยู่ด้วย

ในส่วนของข้อมูล Dataset นี้ อาจเป็นหนึ่งในจุดด้อยของโมเดล Rule-based เมื่อเปรียบเทียบกับโมเดล Machine Learning ที่จะทำการพัฒนาในขั้นตอนถัดไปของงานวิจัย เนื่องจากจำนวนข้อมูลที่ใช้ในการพัฒนาค้นหากฎเกณฑ์ของโมเดล Rule-based นั้นมีจำนวนน้อยกว่า เพราะต้องทำการวิเคราะห์ข้อมูลและหารูปแบบของข้อมูลเองทั้งหมด

3.3.2 กำหนดกฎเกณฑ์และเงื่อนไขสำหรับการตรวจจับภาพ

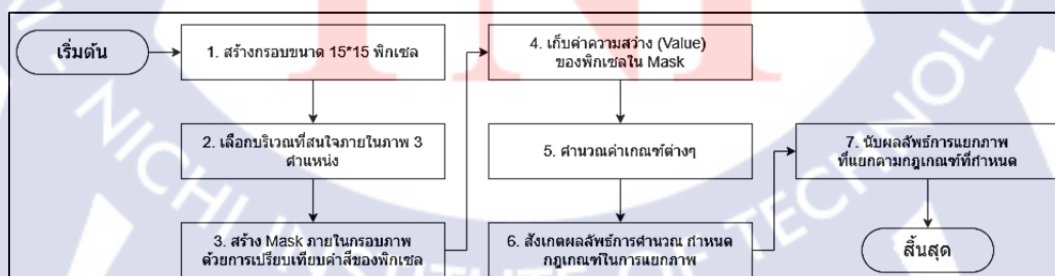
ทำการศึกษาทฤษฎีการตรวจจับภาพด้วยภาพจากภายใน Dataset ที่รวบรวมไว้ เพื่อให้สามารถกำหนดกฎและเงื่อนไขที่จะกำหนดให้โมเดลทำการตรวจจับภาพได้ ซึ่งกรอบบริเวณวัตถุที่เป็นจุดสนใจภายในภาพของงานวิจัยนี้คือ ภาพส่วนที่เป็นแว่นตา และภาพส่วนที่เป็นหน้ากาก ในงานวิจัยนี้ ได้ทำการทดลองค้นหาปัจจัยที่สามารถใช้ในการตรวจจับความแตกต่างของภาพระหว่างภาพถ่ายจริงและภาพถ่ายที่เพิ่มเครื่องประดับ AR เข้าไป ซึ่งในการทดลองแรกเริ่ม ผู้วิจัยได้ทำการศึกษาเกี่ยวกับการวัดค่าแสงสว่างโดยการปรับภาพจาก RGB Color Model ให้เป็นภาพ Grayscale และทำการเลือกเฉพาะส่วนที่เป็นรอยต่อระหว่างขอบหน้ากากกับส่วนที่เป็นใบหน้า เพื่อทำการสังเกตความเปลี่ยนแปลงของค่าความสว่างในส่วนที่เป็นรอยต่อระหว่างภาพถ่ายจริงและภาพถ่ายที่เพิ่มเครื่องประดับ AR แต่เมื่อทำการเปรียบเทียบและวิเคราะห์ข้อมูลแล้ว พบว่ารูปแบบของข้อมูลไม่ชัดเจนมากพอที่จะนำมากำหนดเป็นเกณฑ์สำหรับการตรวจจับได้ (เช่น รูปแบบความเปลี่ยนแปลงของค่าความสว่างของภาพทั้งสองประเภท มีความใกล้เคียงกันมากจนไม่สามารถระบุได้ว่าลักษณะข้อมูลแบบใดคือลักษณะการเปลี่ยนแปลงของภาพประเภทภาพถ่ายจริง หรือภาพถ่ายเพิ่มเครื่องประดับ AR เป็นต้น หลังจากนั้น ผู้วิจัยได้ทำการทดลองเกี่ยวกับการแปลงภาพจาก RGB Color Model ให้เป็น HSV Color Model และทำการหีบข้อมูลส่วนที่เป็นค่า Hue (ค่าสี) จากนั้นทำการเลือกเฉพาะส่วนที่เป็นรอยต่อระหว่างขอบหน้ากากกับส่วนที่เป็นใบหน้า เพื่อสังเกตการเปลี่ยนแปลงของค่า Hue ในส่วนที่เป็นรอยต่อระหว่างภาพทั้งสองประเภท ผลลัพธ์ที่ได้คือ การสังเกตความเปลี่ยนแปลงของค่า Hue สามารถวิเคราะห์ความแตกต่างของรูปแบบการเปลี่ยนแปลง

ของข้อมูลทั้งสองประเภทได้ในระดับหนึ่ง แต่ความแม่นยำไม่สูงมากพอที่จะนำมากำหนดเป็นเงื่อนไขในการตรวจจับได้ และวิธีการนี้เมื่อนำไปทดลองกับภาพแว่นตา ผลลัพธ์คือไม่สามารถแยกแยะลักษณะความแตกต่างของภาพทั้งสองประเภทได้ เนื่องจากภาพเครื่องประดับแต่ละประเภทมีลักษณะที่ต่างกัน และอาจใช้ปัจจัยที่แตกต่างกันในการตรวจจับ ด้วยความตั้งใจในการกำหนดเกณฑ์และเงื่อนไขที่สามารถใช้ร่วมกันได้ระหว่างภาพเครื่องประดับทั้งสองประเภท ความตั้งใจในการค้นหาปัจจัยที่สามารถใช้ตรวจจับได้ชัดเจนที่สุด จึงทำการเปลี่ยนวิธีการและค้นหาวิธีกำหนดเงื่อนไขต่อไป

หลังจากทำการทดลองหลายวิธีการตามที่อธิบายไปข้างต้น ผลสรุปคือ วิธีการที่ใช้ในงานวิจัยนี้ จะใช้วิธีการแปลงภาพถ่ายที่เป็น RGB Color Model ให้เป็นภาพ HSV Color Model และทำการเก็บค่าแสงสว่างของสี (Value) เนื่องจากการใช้ HSV Color Model มีการเก็บค่าแสงสว่างเป็นมิติแยกออกมาจากค่าสี (Hue) และค่าความสดของสี (Saturation) ทำให้สามารถนำค่าแสงสว่างนี้ออกมาใช้และวัดผลอื่นๆ ได้โดยตรงและไม่ส่งผลกับมิติอื่นๆ ต่างกับการใช้ RGB Color Model ที่การเปลี่ยนแปลงของแสงสว่างมีผลกับค่าทั้ง 3 มิติ ทำให้การวัดค่าแสงสว่างทำได้ยากกว่า หลังจากเลือกกรอบบริเวณที่เป็นจุดสนใจแล้ว จะทำการเก็บค่าแสงสว่างในส่วนภาพนั้นๆ ทั้งหมด 3 บริเวณในภาพ และนำค่าแสงสว่างในบริเวณนั้นๆ มาทำการคำนวณค่าที่จะใช้เป็นเกณฑ์ในการแยกระหว่างภาพใบหน้าสวมเครื่องประดับจริง และภาพใบหน้าที่ไม่สวมเครื่องประดับด้วยภาพ AR โดยขั้นตอนทั้งหมดนี้จะทำการศึกษาด้วยโปรแกรม MATLAB เพื่อหาความแตกต่างและข้อสังเกตระหว่างข้อมูลภาพใน Dataset แต่ละประเภท (สวมเครื่องประดับและเพิ่มเครื่องประดับด้วย AR) โดยขั้นตอนในการเก็บข้อมูลค่าแสงสว่างในภาพเพื่อหาเงื่อนไขสำหรับตรวจจับภาพ จะทำแยกกันระหว่างภาพเครื่องประดับประเภทแว่นตา และภาพเครื่องประดับประเภทหน้ากาก

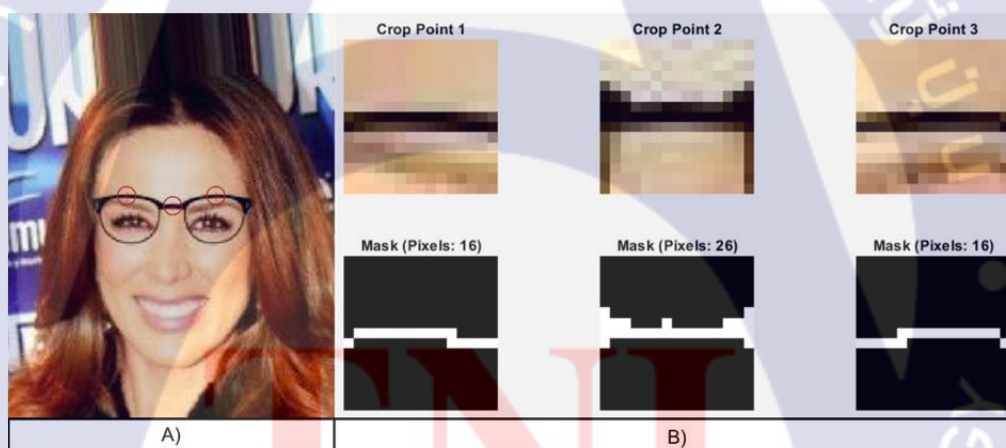
1. ภาพที่สวมเครื่องประดับประเภทแว่นตา

ในส่วนของภาพที่สวมเครื่องประดับประเภทแว่นตา มีขั้นตอนในการกำหนดกฎเกณฑ์และเงื่อนไขสำหรับการตรวจจับภาพ ดังรูปที่ 3.7 ต่อไปนี้



รูปที่ 3.7 แผนผังขั้นตอนการกำหนดกฎเกณฑ์และเงื่อนไขการตรวจจับภาพ (แว่นตา)

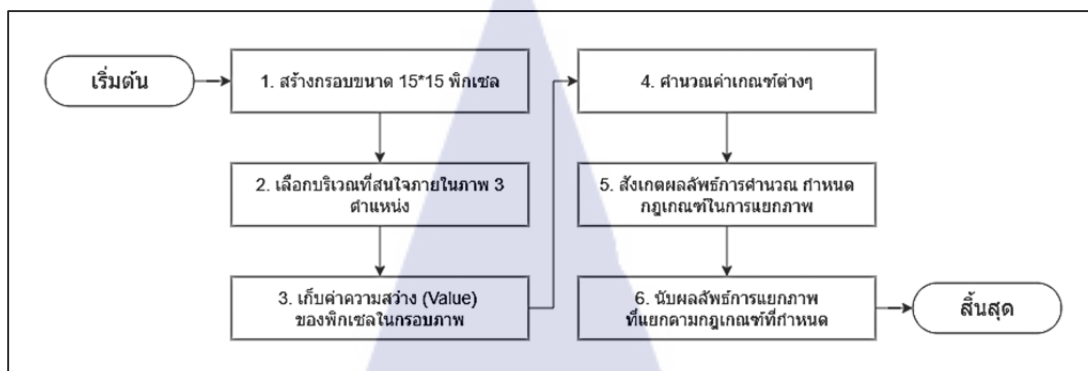
โดยเริ่มทำการสร้างกรอบ (Window) ขนาด 15×15 พิกเซล เพื่อเลือกบริเวณที่มีความสนใจ (ROI: Region of Interest) ภายในภาพ จากนั้นเลือกบริเวณที่มีความสนใจทั้งหมด 3 จุด โดยในบริเวณนั้นๆ ต้องมีส่วนที่เป็นกรอบของแว่นตาภายในภาพ (จุดที่ 1 เลือกกรอบแว่นตาส่วนบนซ้าย จุดที่ 2 เลือกกรอบแว่นตาส่วนบนขวา และจุดที่ 3 เลือกกรอบแว่นตาส่วนบนขวา) จากนั้นเมื่อเลือกจุด 3 จุดได้แล้ว ทำการสร้าง Mask เพื่อระบุพื้นที่ในการเก็บค่า Value เนื่องจากการเลือกบริเวณที่มีความสนใจด้วยกรอบนี้จะทำให้ติดภาพส่วนอื่นๆ นอกจากกรอบแว่นตามาด้วย (เช่น ผิวหนัง เป็นต้น) จึงใช้วิธีการแปลงภาพจาก RGB Color Model ให้เป็น HSV Color Model และทำการเปรียบเทียบสีระหว่างพิกเซลที่กดเลือก (จุดกึ่งกลางของกรอบขนาด 15×15) กับพิกเซลอื่นๆ ในกรอบนั้นๆ ด้วยการใช้ Euclidean Distance [18] โดยกำหนดไว้ว่าหากมีความต่างของสีระหว่างพิกเซลที่จุดกึ่งกลางของภาพกับพิกเซลที่ตรวจสอบปัจจุบันเกิน 15% จะทำการตัดพิกเซลนั้นออกจาก Mask และไม่เก็บค่าความสว่างของพิกเซลนั้น เนื่องจากหากความต่างของสีมากเกินไป แสดงว่าจุดนั้นๆ ไม่ใช่จุดที่เป็นกรอบแว่นตา (ที่เป็นสีของพิกเซลที่จุดกึ่งกลางของภาพ) ทำให้ไม่นับพิกเซลนั้นมารวมในการเก็บค่าและคำนวณค่าแสงสว่าง เพราะต้องการเก็บค่าเฉพาะส่วนที่เป็นตัววัตถุที่สงสัยเท่านั้น ดังตัวอย่างในรูปที่ 3.8 ต่อไปนี้



รูปที่ 3.8 (A) ภาพแสดงการเลือกบริเวณที่มีความสนใจ 3 จุดในภาพ (B) การกำหนด Mask สำหรับระบุพื้นที่ในการเก็บค่า Value ในภาพ

2. ภาพที่สวมเครื่องประดับประเภทหน้ากาก

ในส่วนของภาพที่สวมเครื่องประดับประเภทหน้ากาก จะมีความแตกต่างกับภาพที่สวมเครื่องประดับประเภทแว่นตาเล็กน้อยในส่วนของการเก็บค่าความสว่างของพิกเซล โดยมีขั้นตอนในการกำหนดกฎเกณฑ์และเงื่อนไขสำหรับการตรวจจับภาพ ดังรูปที่ 3.9 ต่อไปนี้



รูปที่ 3.9 แผนผังขั้นตอนการกำหนดกฎเกณฑ์และเงื่อนไขการตรวจจับภาพ (หน้ากาก)

เริ่มแรกทำการสร้างกรอบขนาด 15×15 พิกเซล เพื่อเลือกบริเวณที่มีความสนใจ (ROI: Region of Interest) ภายในภาพ จากนั้นเลือกบริเวณที่มีความสนใจทั้งหมด 3 จุด โดยในบริเวณนั้นๆ ต้องเป็นส่วนที่เป็นผิวหน้ากากเท่านั้น (ไม่มีส่วนผิวหนังหรือวัตถุอื่นๆ ภายในกรอบ) ด้วยเหตุนี้จึงไม่จำเป็นต้องคำนวณเปรียบเทียบความแตกต่างของสีในแต่ละพิกเซลเพื่อสร้าง Mask สำหรับระบุพื้นที่การเก็บค่าความสว่าง เนื่องจากทุกพิกเซลภายในกรอบภาพเป็นบริเวณที่มีความสนใจ หรือก็คือส่วนที่เป็นหน้ากากทั้งหมดอยู่แล้วนั่นเอง

หลังจากสามารถระบุพื้นที่ที่ต้องการเก็บค่าแสงสว่างภายในกรอบได้แล้ว ทำการเก็บค่า Value ที่เป็นค่าแสงสว่างเพื่อนำไปทำการคำนวณหากฎเกณฑ์ โดยการสำรวจและวิเคราะห์ข้อมูลค่าแสงสว่างที่เก็บได้จากภาพ (หากเป็นภาพแว่นตา จะเก็บเฉพาะค่าแสงสว่างในตำแหน่งพิกเซลที่ถูกระบุไว้ใน Mask แต่หากเป็นภาพหน้ากาก จะเก็บค่าทุกพิกเซลที่อยู่ในกรอบภาพ) โดยจำนวนพิกเซลที่เก็บค่าแสงสว่างในแต่ละจุดของภาพมีโอกาสที่จะไม่เท่ากัน เนื่องจากแต่ละจุดจะมี Mask ที่ระบุตำแหน่งพิกเซลที่ต้องทำการเก็บค่าไม่เหมือนกัน ผลในการเก็บค่าแสงสว่างจะแยกระหว่างแต่ละภาพ และแต่ละจุดในภาพ ดังที่แสดงในรูปที่ 3.10 ต่อไปนี้

FileName	PointNumber	Value_Pixels_1	Value_Pixels_2	Value_Pixels_3	Value_Pixels_4	Value_Pixels_5	Value_Pixels_6
000618.jpg	1	0.074509804	0.137254902	0.105882353	0.133333333	0.070588235	0.11372549
000618.jpg	2	0.066666667	0.105882353	0.070588235	0.070588235	0.125490196	0.078431373
000618.jpg	3	0.258823529	0.2	0.250980392	0.17254902	0.231372549	0.231372549
000619.jpg	1	0.11372549	0.109803922	0.101960784	0.117647059	0.125490196	0.160784314
000619.jpg	2	0.278431373	0.258823529	0.270588235	0.282352941	0.341176471	0.294117647
000619.jpg	3	0.152941176	0.141176471	0.180392157	0.08627451	0.08627451	0.074509804

รูปที่ 3.10 ตัวอย่างการเก็บค่าแสงสว่างทั้ง 3 จุดในแต่ละภาพ

เมื่อรวบรวมค่าแสงสว่างทั้งหมดแล้ว ทำการคำนวณค่าต่างๆ โดยจะใช้การคำนวณค่าที่สามารถแยกระหว่างความแตกต่างระหว่างภาพใบหน้าสวมเครื่องประดับจริงและภาพใบหน้าที่ไม่สวมเครื่องประดับด้วยภาพ AR ด้วยค่าแสงสว่างในภาพที่ทำการเก็บผลลัพธ์จากขั้นตอนก่อนหน้า ซึ่งการกำหนดค่าที่จะใช้ในการแบ่งแยกภาพ จะกำหนดตามสมมติฐานที่มีในการแยกแยะภาพ โดยมีทั้งหมด 5 ค่าที่จะใช้ในการกำหนดเกณฑ์

1. ค่าเฉลี่ยของค่ารวมแสงสว่างรายกรอบบริเวณที่สนใจ (Mean Box Intensity)

เป็นการหาค่ารวมแสงสว่างรายกรอบบริเวณที่สนใจก่อน จากนั้นนำค่ารวมแสงสว่างที่ได้มาทำการเฉลี่ย (หารด้วยจำนวนประชากรหรือ 3 เนื่องจากเลือกทั้งหมด 3 บริเวณภายในหนึ่งภาพ) เพื่อให้ได้ค่าเฉลี่ยของค่ารวมแสงสว่างของทั้งสามกรอบภาพ ซึ่งทำการเก็บรวบรวมค่านี้โดยมีสมมติฐานแยกกันระหว่างภาพแว่นตาและหน้ากากว่า ภาพแว่นตาที่เป็นเครื่องประดับจริงอาจมีความสว่างมากกว่าภาพแว่นตา AR เนื่องจากภาพเครื่องประดับจริงอาจมีแสงสะท้อนหรือแสงสว่างภายในบริเวณวัตถุมากกว่าภาพแว่นตา AR ส่วนภาพหน้ากากที่เป็นเครื่องประดับจริงอาจมีความสว่างน้อยกว่าภาพหน้ากาก AR เนื่องจากภาพหน้ากากจริงอาจมีส่วนที่เป็นเงา หรือรอยยับของหน้ากากมากกว่าภาพหน้ากาก AR

2. ค่าเฉลี่ยแสงสว่างของพิกเซลทั้งหมด (Mean Intensity)

หาค่าเฉลี่ยแสงสว่างของพิกเซลทั้งหมดในกรอบบริเวณที่สนใจ โดยนำค่าความสว่างของแต่ละบริเวณที่สนใจมารวมกันทั้งหมด และหารด้วยจำนวนพิกเซลของทั้งสามกรอบบริเวณที่สนใจ เพื่อหาค่าเฉลี่ยแสงสว่างของพิกเซลทั้งหมด เนื่องจากต้องการหาค่าเฉลี่ยแสงสว่างของบริเวณที่สนใจทั้งภาพโดยไม่สนใจการแบ่งค่าเฉลี่ยตามกรอบเหมือนค่าเฉลี่ยแสงสว่างของค่าเฉลี่ยรายกรอบบริเวณที่สนใจ (ที่มีการเฉลี่ยรายกรอบก่อนนำมาเฉลี่ยอีกครั้ง) โดยมีสมมติฐานเหมือนกับค่าเฉลี่ยแสงสว่างของค่าเฉลี่ยรายกรอบบริเวณที่สนใจ ซึ่งแยกกันระหว่างภาพแว่นตาและหน้ากากว่า ภาพแว่นตาที่เป็นเครื่องประดับจริงอาจมีความสว่างมากกว่าภาพแว่นตา AR เนื่องจากภาพเครื่องประดับจริงอาจมีแสงสะท้อนหรือแสงสว่างภายในบริเวณวัตถุมากกว่าภาพแว่นตา AR ส่วนภาพหน้ากากที่เป็นเครื่องประดับจริงอาจมีความสว่างน้อยกว่าภาพหน้ากาก AR เนื่องจากภาพหน้ากากจริงอาจมีส่วนที่เป็นเงา หรือรอยยับของหน้ากากมากกว่าภาพหน้ากาก AR

3. ค่าส่วนเบี่ยงเบนเฉลี่ยสัมบูรณ์ (Mean AD: Mean Absolute Deviation)

หาค่า Mean AD โดยการนำค่าเฉลี่ยแสงสว่างของแต่ละกรอบบริเวณมาเข้าสมการคำนวณ โดยใช้สมการที่ 2.4 ซึ่งใช้ค่าเฉลี่ย (Mean) ของชุดข้อมูลในการคำนวณ เพื่อวิเคราะห์การกระจายตัวของค่าแสงสว่างของพิกเซลในภาพนั้นๆ ซึ่งสะท้อนถึงความสม่ำเสมอของวัตถุในภาพ โดยมีสมมติฐานว่า หากค่า Mean AD ต่ำ หมายถึงค่าความสว่างในแต่ละจุดพิกเซลเท่ากันหรือมีความใกล้เคียงกันมาก ทำให้มีโอกาสมากกว่าที่จะเป็นภาพ AR ที่ตัววัตถุจะมีสีที่เรียบเนียนกว่าวัตถุที่เป็นภาพถ่ายจริง และหากค่า Mean AD สูง จะหมายถึงค่าความสว่างในแต่ละจุดพิกเซลมีความห่างกัน มีโอกาสที่จะเป็นวัตถุจริงที่สี

ของวัตถุจะได้รับผลกระทบจากสภาพแวดล้อมต่างๆ เช่น แสงสว่างภายในภาพ ทำให้มีสีที่ไม่เรียบเนียนเท่าวัตถุที่เป็น AR

4. ค่าส่วนเบี่ยงเบนมาตรฐานสัมบูรณ์ (Median AD: Median Absolute Deviation)

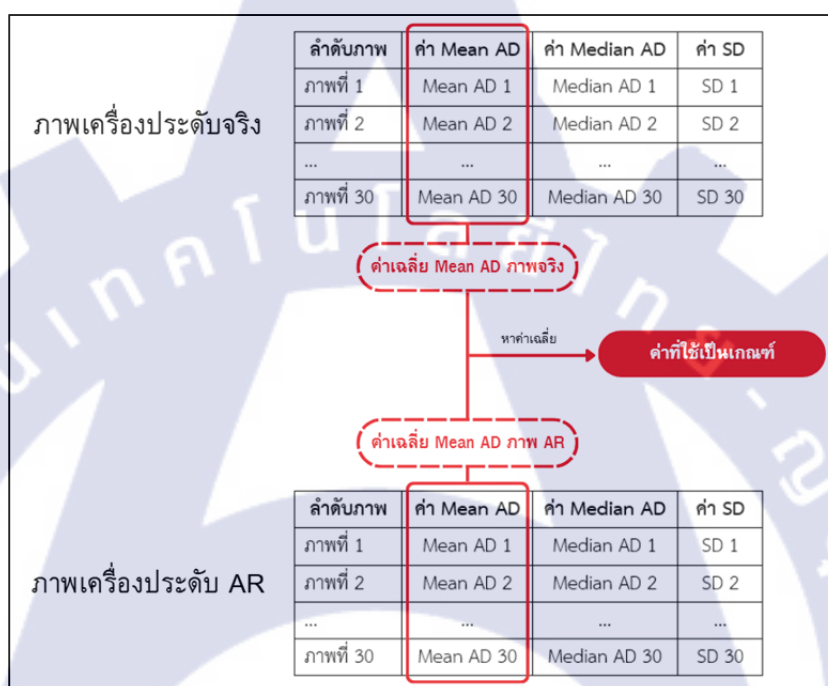
หาค่า Median AD โดยการนำค่าเฉลี่ยแสงสว่างของแต่ละกรอบบริเวณมาเข้าสมการคำนวณ โดยใช้สมการที่ 2.5 ซึ่งใช้ค่ามัธยฐาน (Median) ของชุดข้อมูลในการคำนวณ ทำให้มีความเป็นไปได้ในการตัดค่าที่เป็น Outliers ออกไปซึ่งต่างกับการหาค่า Mean Absolute Deviation เพื่อวิเคราะห์การกระจายตัวของค่าแสงสว่างของพิกเซลในภาพนั้นๆ ซึ่งสะท้อนถึงความสม่ำเสมอของวัตถุในภาพ โดยมีสมมติฐานว่า หากค่า Median AD ต่ำ หมายถึงค่าความสว่างในแต่ละจุดพิกเซลเท่ากัน หรือมีความใกล้เคียงกันมาก ทำให้มีโอกาสมากกว่าที่จะเป็นภาพ AR ที่ตัววัตถุจะมีสีที่เรียบเนียนกว่าวัตถุที่เป็นภาพถ่ายจริง และหากค่า Median AD สูง จะหมายถึงค่าความสว่างในแต่ละจุดพิกเซลมีความห่างกัน มีโอกาสที่จะเป็นวัตถุจริงที่สีของวัตถุจะได้รับผลกระทบจากสภาพแวดล้อมต่างๆ เช่น แสงสว่างภายในภาพ ทำให้มีสีที่ไม่เรียบเนียนเท่าวัตถุที่เป็น AR

5. ค่าส่วนเบี่ยงเบนมาตรฐาน (SD: Standard Deviation)

หาค่า SD โดยการนำค่าเฉลี่ยแสงสว่างของแต่ละกรอบบริเวณมาเข้าสมการคำนวณ โดยใช้สมการที่ 2.6 ซึ่งใช้ค่าเฉลี่ย (Mean) ของชุดข้อมูล และใช้วิธีการยกกำลังสองในการคำนวณซึ่งแตกต่างกับการหาค่า Mean AD และ Median AD ที่ใช้ค่าสัมบูรณ์ในการคำนวณ เพื่อวิเคราะห์การกระจายตัวของค่าแสงสว่างของพิกเซลในภาพนั้นๆ ซึ่งสะท้อนถึงความสม่ำเสมอของวัตถุในภาพ และด้วยสูตรคำนวณของการหาค่า SD ที่มีการใช้การยกกำลัง ทำให้ค่า SD มีความเปลี่ยนแปลงต่อค่าที่เป็น Outliers สูงกว่า Mean AD และ Median AD ทำให้สามารถแยกความแตกต่างของการกระจายตัวของค่าได้ชัดเจนกว่า Mean AD และ Median AD โดยมีสมมติฐานว่า หากค่า SD ต่ำ หมายถึงค่าความสว่างในแต่ละจุดพิกเซลมีความใกล้เคียงกับค่าเฉลี่ย และมีความสม่ำเสมอสูง อาจตีความได้ว่าเป็นวัตถุจากภาพ AR ที่มีสีและความสว่างเรียบเนียนกว่าวัตถุจริง และหากค่า SD สูง หมายถึงค่าความสว่างในแต่ละจุดพิกเซลห่างจากค่าเฉลี่ย และมีความสม่ำเสมอต่ำ อาจตีความได้ว่าเป็นวัตถุที่เป็นภาพถ่ายจริงๆ ซึ่งมีแสงแต่ละจุดแตกต่างกันจากการถ่ายภาพ

เมื่อทำการคำนวณครบทั้ง 5 ค่าแล้ว ทำซ้ำจนครบจำนวนภาพที่ใช้ในการทดลองเพื่อหาเงื่อนไข คือ 60 ภาพ (เครื่องประดับจริง 30 ภาพ และเครื่องประดับ AR 30 ภาพ) ต่อหนึ่งประเภทเครื่องประดับ (หากรวมทั้งสองประเภทเครื่องประดับ จะใช้ภาพรวม 120 ภาพ ได้แก่ ภาพหน้ากาก 60 ภาพ และภาพแว่นตา 60 ภาพ) จากนั้นทำการสังเกตรูปแบบของข้อมูลที่ได้เพื่อทำการหาเกณฑ์ที่จะใช้ในการแยกภาพ ยกตัวอย่างเช่น หากเป็นภาพแว่นตา ถ้ามีค่า Mean AD น้อยกว่าค่าที่ระบุไว้ จะตีความว่าเป็นภาพเครื่องประดับ AR เป็นต้น ซึ่งในส่วนของคุณค่าเฉลี่ยของค่ารวมแสงสว่างรายกรอบบริเวณที่สนใจ (Mean Box Intensity) และค่าเฉลี่ยแสงสว่างของพิกเซลทั้งหมด (Mean Intensity) จะสามารถกำหนดเกณฑ์ใน

การแยกแยะได้หลังจากทำการสังเกตจากรูปแบบของข้อมูลเครื่องประดับที่เป็นภาพจริงและเครื่องประดับที่เป็นภาพ AR ทำให้เห็นเส้นแบ่งของข้อมูลที่เหมาะสมและสามารถใช้เป็นเกณฑ์ในการแยกแยะได้ และในส่วนของการหาเกณฑ์ของค่าส่วนเบี่ยงเบน (ทั้งส่วนเบี่ยงเบนเฉลี่ยสัมบูรณ์ ส่วนเบี่ยงเบนมาตรฐานสัมบูรณ์ และส่วนเบี่ยงเบนมาตรฐาน) จะใช้วิธีการหาค่าเฉลี่ยของแต่ละค่าเพื่อหาค่ากลางที่จะใช้เป็นเกณฑ์ ดังตัวอย่างรูปที่ 3.11 ต่อไปนี้



รูปที่ 3.11 ตัวอย่างการหาค่าที่ใช้เป็นเกณฑ์ของค่าส่วนเบี่ยงเบนเฉลี่ยสัมบูรณ์ (Mean AD)

ในการกำหนดเงื่อนไขเพื่อแยกแยะระหว่างภาพถ่ายที่สวมเครื่องประดับจริงและเครื่องประดับ AR ผู้วิจัยได้ทำการศึกษาเพิ่มเติมเกี่ยวกับความแตกต่างของภาพถ่ายและภาพที่ถูกแต่งเติมโดย AR จากงานวิจัยของ [34] H. Yao et al. [34] มีการกล่าวถึงลักษณะของภาพถ่าย ที่มีความสัมพันธ์ระหว่างค่า Noise กับค่าความสว่างภายในภาพ ซึ่งภาพถ่ายจริงจะมีส่วนที่เป็น Noise หลักๆ 2 ประเภท ได้แก่ Photo-response non-uniformity noise (PRNU) และ Random Noise ซึ่ง PRNU นั้นเกิดจากความคลาดเคลื่อนของเซนเซอร์ภายในกล้อง ซึ่งมีรูปแบบที่คงที่ และสามารถระบุเอกลักษณ์ของกล้องได้ ส่วนภาพที่ถูกแต่งเติมโดย AR นั้น เปรียบได้กับการนำภาพอื่นมาใส่เพิ่มลงไปบนภาพถ่ายเดิม ทำให้รูปแบบของ Noise ที่เดิมในภาพถ่ายนั้นมีอยู่ หายไปเนื่องจากโดนภาพอื่นมาใส่ทับตำแหน่งเดิม หรือทำให้มีรูปแบบของ Noise ที่ไม่ตรงกันกับ Noise ส่วนที่เหลือภายในภาพซึ่งมีผลกับความสม่ำเสมอของ

พิกเซลภายในภาพ เป็นจุดที่สามารถแยกแยะความแตกต่างระหว่างภาพถ่ายจริงและภาพถ่ายที่ถูกแต่งเติมด้วย AR ได้ด้วยเช่นกัน

เพื่อให้สามารถแยกแยะระหว่างภาพที่สวมเครื่องประดับจริงและเครื่องประดับ AR ได้ จึงกำหนดเกณฑ์การตัดสินใจจากค่าที่คำนวณออกมาได้ โดยใช้ทั้งหมด 5 เงื่อนไข ซึ่งแตกต่างกันระหว่างภาพเครื่องประดับที่เป็นแว่นตา และภาพเครื่องประดับที่เป็นหน้ากาก ดังนี้

1. เงื่อนไขของค่าเฉลี่ยของค่ารวมแสงสว่างรายกรอบบริเวณที่สนใจ (Mean Box Intensity)

กรณีแว่นตา หากภาพนั้นๆ มีค่าเฉลี่ยของค่ารวมแสงสว่างรายกรอบบริเวณที่สนใจ น้อยกว่าค่าที่เป็นเกณฑ์ จะถูกจัดกลุ่มเป็นภาพเครื่องประดับ AR เนื่องจากมีสมมติฐานว่าภาพแว่นตาที่เป็น AR จะมีความสว่างน้อยกว่าภาพแว่นตาจริง เนื่องจากภาพแว่นตา AR จะมีแสงสะท้อนน้อยกว่าภาพแว่นตาจริง

กรณีหน้ากาก หากภาพนั้นๆ มีค่าเฉลี่ยของค่ารวมแสงสว่างรายกรอบบริเวณที่สนใจ มากกว่าค่าที่เป็นเกณฑ์ จะถูกจัดกลุ่มเป็นภาพเครื่องประดับ AR เนื่องจากมีสมมติฐานว่าภาพหน้ากากที่เป็น AR จะมีความสว่างมากกว่าภาพหน้ากากจริง เนื่องจากหน้ากากจริงจะมีส่วนที่เป็นรอยยับหรือเงาที่เกิดตามแหล่งกำเนิดแสงขณะถ่ายภาพอยู่ด้วย

2. เงื่อนไขของค่าเฉลี่ยแสงสว่างของพิกเซลทั้งหมด (Mean Intensity)

กรณีแว่นตา หากภาพนั้นๆ มีค่าเฉลี่ยแสงสว่างของพิกเซลทั้งหมด น้อยกว่าค่าที่เป็นเกณฑ์ จะถูกจัดกลุ่มเป็นภาพเครื่องประดับ AR เนื่องจากมีสมมติฐานว่าภาพแว่นตาที่เป็น AR จะมีความสว่างน้อยกว่าภาพแว่นตาจริง เนื่องจากภาพแว่นตา AR จะมีแสงสะท้อนน้อยกว่าภาพแว่นตาจริง

กรณีหน้ากาก หากภาพนั้นๆ มีค่าเฉลี่ยแสงสว่างของพิกเซลทั้งหมด มากกว่าค่าที่เป็นเกณฑ์ จะถูกจัดกลุ่มเป็นภาพเครื่องประดับ AR เนื่องจากมีสมมติฐานว่าภาพหน้ากากที่เป็น AR จะมีความสว่างมากกว่าภาพหน้ากากจริง เนื่องจากหน้ากากจริงจะมีส่วนที่เป็นรอยยับหรือเงาที่เกิดตามแหล่งกำเนิดแสงขณะถ่ายภาพอยู่ด้วย

3. เงื่อนไขของค่าส่วนเบี่ยงเบนเฉลี่ยสัมบูรณ์ (Mean AD)

กรณีแว่นตา หากภาพนั้นๆ มีค่าส่วนเบี่ยงเบนเฉลี่ยสัมบูรณ์ น้อยกว่าค่าที่เป็นเกณฑ์ จะถูกจัดกลุ่มเป็นภาพเครื่องประดับ AR เนื่องจากมีสมมติฐานว่าภาพแว่นตาที่เป็น AR จะมีความสม่ำเสมอของพิกเซล หรือค่าส่วนเบี่ยงเบนน้อยกว่าภาพแว่นตาจริง เนื่องจากภาพแว่นตา AR อาจมีสีหรือลักษณะของวัตถุที่เรียบเนียนกว่าแว่นตาจริง

กรณีหน้ากาก หากภาพนั้นๆ มีค่าส่วนเบี่ยงเบนเฉลี่ยสัมบูรณ์ น้อยกว่าค่าที่เป็นเกณฑ์ จะถูกจัดกลุ่มเป็นภาพเครื่องประดับ AR เนื่องจากมีสมมติฐานว่าภาพหน้ากากที่เป็น AR จะมี

ความสม่ำเสมอของพิกเซล หรือค่าส่วนเบี่ยงเบนน้อยกว่าภาพหน้ากากจริง เนื่องจากภาพหน้ากาก AR อาจมีสีหรือลักษณะของวัตถุที่เรียบเนียนกว่าแว่นตาจริง

4. เงื่อนไขของค่าส่วนเบี่ยงเบนมัธยฐานสัมบูรณ์ (Median AD)

กรณีแว่นตา หากภาพนั้นๆ มีค่าส่วนเบี่ยงเบนมัธยฐานสัมบูรณ์ น้อยกว่าค่าที่เป็นเกณฑ์ จะถูกจัดกลุ่มเป็นภาพเครื่องประดับ AR เนื่องจากมีสมมติฐานว่าภาพแว่นตาที่เป็น AR จะมีความสม่ำเสมอของพิกเซล หรือค่าส่วนเบี่ยงเบนน้อยกว่าภาพแว่นตาจริง เนื่องจากภาพแว่นตา AR อาจมีสีหรือลักษณะของวัตถุที่เรียบเนียนกว่าแว่นตาจริง

กรณีหน้ากาก หากภาพนั้นๆ มีค่าส่วนเบี่ยงเบนมัธยฐานสัมบูรณ์ น้อยกว่าค่าที่เป็นเกณฑ์ จะถูกจัดกลุ่มเป็นภาพเครื่องประดับ AR เนื่องจากมีสมมติฐานว่าภาพหน้ากากที่เป็น AR จะมีความสม่ำเสมอของพิกเซล หรือค่าส่วนเบี่ยงเบนน้อยกว่าภาพหน้ากากจริง เนื่องจากภาพหน้ากาก AR อาจมีสีหรือลักษณะของวัตถุที่เรียบเนียนกว่าแว่นตาจริง

5. เงื่อนไขของค่าส่วนเบี่ยงเบนมาตรฐาน (SD)

กรณีแว่นตา หากภาพนั้นๆ มีค่าส่วนเบี่ยงเบนมาตรฐาน น้อยกว่าค่าที่เป็นเกณฑ์ จะถูกจัดกลุ่มเป็นภาพเครื่องประดับ AR เนื่องจากมีสมมติฐานว่าภาพแว่นตาที่เป็น AR จะมีความสม่ำเสมอของพิกเซล หรือค่าส่วนเบี่ยงเบนน้อยกว่าภาพแว่นตาจริง เนื่องจากภาพแว่นตา AR อาจมีสีหรือลักษณะของวัตถุที่เรียบเนียนกว่าแว่นตาจริง

กรณีหน้ากาก หากภาพนั้นๆ มีค่าส่วนเบี่ยงเบนมาตรฐาน น้อยกว่าค่าที่เป็นเกณฑ์ จะถูกจัดกลุ่มเป็นภาพเครื่องประดับ AR เนื่องจากมีสมมติฐานว่าภาพหน้ากากที่เป็น AR จะมีความสม่ำเสมอของพิกเซล หรือค่าส่วนเบี่ยงเบนน้อยกว่าภาพหน้ากากจริง เนื่องจากภาพหน้ากาก AR อาจมีสีหรือลักษณะของวัตถุที่เรียบเนียนกว่าแว่นตาจริง สามารถกำหนดเงื่อนไขของการแยกแยะภาพเครื่องประดับจริงและเครื่องประดับ AR ระหว่างภาพแว่นตาและภาพหน้ากากได้ ตามตารางที่ 3.1

ตารางที่ 3.1 เงื่อนไขการแยกแยะภาพเครื่องประดับจริงและเครื่องประดับ AR

คุณลักษณะ	สัญลักษณ์ เงื่อนไข แว่นตา	สัญลักษณ์ เงื่อนไข หน้ากาก	เงื่อนไขภาพ แว่นตา	เงื่อนไขภาพ หน้ากาก
ค่าเฉลี่ยของค่ารวมแสงสว่าง รายกรอบบริเวณที่สนใจ	T_1	T'_1	$V \leq T_1$	$V > T'_1$
ค่าเฉลี่ยแสงสว่างของพิกเซล ทั้งหมด	T_2	T'_2	$V \leq T_2$	$V > T'_2$

ตารางที่ 3.1 เงื่อนไขการแยกแยะภาพเครื่องประดับจริงและเครื่องประดับ AR (ต่อ)

คุณลักษณะ	สัญลักษณ์ เงื่อนไข แว่นตา	สัญลักษณ์ เงื่อนไข หน้ากาก	เงื่อนไขภาพ แว่นตา	เงื่อนไขภาพ หน้ากาก
ค่าส่วนเบี่ยงเบนเฉลี่ยสัมบูรณ์ (Mean AD)	T_3	T'_3	$V \leq T_3$	$V > T'_3$
ค่าส่วนเบี่ยงเบนมัธยฐาน สัมบูรณ์ (Median AD)	T_4	T'_4	$V \leq T_4$	$V > T'_4$
ค่าส่วนเบี่ยงเบนมาตรฐาน (SD)	T_5	T'_5	$V \leq T_5$	$V > T'_5$

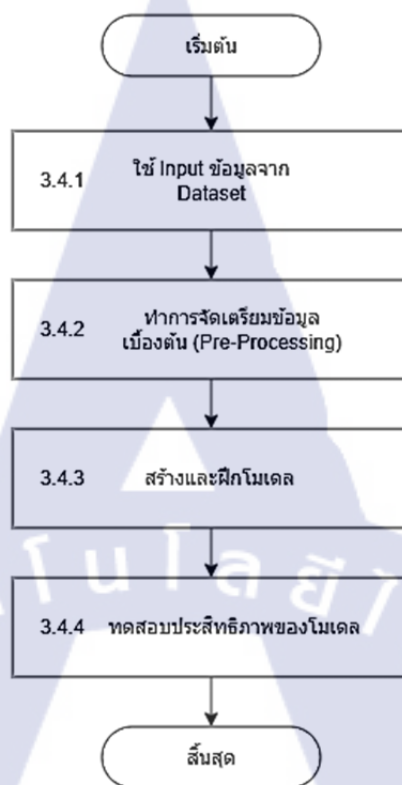
เมื่อกำหนดเงื่อนไขได้แล้ว ทำการวัดผลด้วยการทดลองการแยกภาพโดยใช้วิธีนับเสียงข้างมาก (Voting) ในการแยก เนื่องจากเงื่อนไขที่ระบุไว้มีทั้งหมด 5 ข้อ (ตามค่าที่ได้จากการคำนวณ) หากภาพนั้นๆ มีผลลัพธ์จากการแยกแยะภาพด้วยแต่ละเงื่อนไขมากกว่า 3 ใน 5 จะถูกนับว่าเป็นภาพประเภทนั้นๆ ยกตัวอย่างเช่น หากภาพหนึ่งใช้เงื่อนไขวัด 5 ข้อ ใน 3 ข้อนั้นตีความออกมาเป็นภาพเครื่องประดับ AR และอีก 2 ข้อตีความว่าเป็นภาพเครื่องประดับจริง ภาพนั้นจะถูกนับว่าเป็นภาพเครื่องประดับ AR เนื่องจากเป็นเสียงข้างมากจากการตรวจสอบทุกเงื่อนไข เมื่อได้เงื่อนไขครบทุกข้อ ตามค่าที่คำนวณได้แล้ว นำเงื่อนไขเหล่านี้ไปใช้ในขั้นตอนการวัดผลความแม่นยำต่อไป

3.3.3 วัดผลความแม่นยำและประสิทธิภาพของโมเดล

หลังจากทำการกำหนดเงื่อนไขแล้ว ทำการทดสอบประสิทธิภาพของโมเดลด้วยข้อมูลภาพใน Dataset ส่วนสำหรับการทดสอบ (Testing) และทำการประเมินผลประสิทธิภาพของโมเดลที่พัฒนาขึ้น ด้วย Confusion Matrix และจัดเก็บผลลัพธ์ไว้สำหรับนำไปประเมินผลในขั้นตอนสุดท้าย

3.4 การพัฒนาโมเดลด้วย Machine Learning

การพัฒนาโมเดลด้วย Machine Learning คือการสร้างโมเดลที่สามารถเรียนรู้จากข้อมูลที่มอบให้โมเดลได้โดยตรง ซึ่งโมเดลจะสามารถเรียนรู้ผ่านข้อมูลนั้นๆ และสามารถจำแนกหรือแบ่งกลุ่มข้อมูลได้ โดยที่มนุษย์ไม่ต้องกำหนดกฎหรือเงื่อนไขให้โดยตรงแบบการใช้ Rule-based โดยมีขั้นตอนทั้งหมด 4 ขั้นตอน ดังที่แสดงในรูปที่ 3.12



รูปที่ 3.12 แผนผังขั้นตอนการพัฒนาโมเดลด้วย Machine Learning

3.4.1 ใช้ Input ข้อมูลจาก Dataset

ใช้ภาพจาก Dataset ที่เตรียมไว้ในขั้นตอนก่อนหน้า โดยแบ่งภาพที่ใช้เป็น 3 ส่วน

1. ภาพสำหรับการฝึกโมเดล (Training) 1,400 ภาพ

ใช้ภาพใบหน้าสวมเครื่องประดับจริง 700 ภาพ (แบ่งเป็นภาพสวมแว่นตา 350 ภาพ และภาพสวมหน้ากาก 350 ภาพ) และภาพใบหน้าสวมเครื่องประดับ AR อีก 700 ภาพ (แบ่งเป็นภาพสวมแว่นตา 350 ภาพ และภาพสวมหน้ากาก 350 ภาพ)

2. ภาพสำหรับการตรวจสอบโมเดล (Validation) 200 ภาพ

ใช้ภาพใบหน้าสวมเครื่องประดับจริง 100 ภาพ (แบ่งเป็นภาพสวมแว่นตา 50 ภาพ และภาพสวมหน้ากาก 50 ภาพ) และภาพใบหน้าสวมเครื่องประดับ AR อีก 100 ภาพ (แบ่งเป็นภาพสวมแว่นตา 50 ภาพ และภาพสวมหน้ากาก 50 ภาพ)

3. ภาพสำหรับการทดสอบโมเดล (Testing) 400 ภาพ

ใช้ภาพใบหน้าสวมเครื่องประดับจริง 200 ภาพ (แบ่งเป็นภาพสวมแว่นตา 10 ภาพ และภาพสวมหน้ากาก 100 ภาพ) และภาพใบหน้าสวมเครื่องประดับ AR อีก 200 ภาพ (แบ่งเป็นภาพ

สวมแว่นตา 100 ภาพ และภาพสวมหน้ากาก 100 ภาพ) โดยภาพในส่วนนี้จะป็นภาพกลุ่มเดียวกับภาพที่ใช้ในการทดสอบเกณฑ์ในการตรวจจับภาพของวิธี Rule-based

3.4.2 ทำการเตรียมข้อมูลเบื้องต้น (Pre-Processing)

หลังจากได้ข้อมูลภาพจาก Dataset แล้ว ทำการเตรียมข้อมูลภาพนั้นให้พร้อมต่อการนำไปใช้ในการฝึกโมเดล เนื่องจากงานวิจัยนี้นำภาพจาก Dataset ที่ต่างกัน 2 ชุดมาใช้งาน จึงทำการเตรียมข้อมูลโดยการปรับขนาดภาพให้เท่ากัน โดยไม่ทำให้เกิดความเปลี่ยนแปลงของมิติภาพ (ไม่เกิดการบีบภาพ หรือการปรับภาพที่ทำให้ภาพบิดเบี้ยว) เพื่อให้ขนาดของภาพใน Dataset เท่ากันทุกภาพ และการปรับภาพให้ไม่เกิดความเปลี่ยนแปลงของมิติภาพ ทำให้ข้อมูลภาพที่มียังคงอยู่ครบถ้วน โดยใช้โมเดล MTCNN ในการตรวจจับตำแหน่งใบหน้า และทำการตัดภาพเฉพาะส่วนที่เป็นใบหน้าออกมาเพื่อจำกัดบริเวณของข้อมูลให้มุ่งเน้นเฉพาะส่วนที่เป็นใบหน้า (ตั้งแต่ช่วงศีรษะถึงคาง) เพื่อให้ข้อมูลสามารถส่งเข้าโมเดลเพื่อทำการฝึกได้ง่าย

3.4.3 สร้างและฝึกโมเดล

การสร้างและฝึกโมเดล เริ่มจากการเลือกประเภทของโมเดลที่จะนำมาใช้ในการพัฒนาในงานวิจัยนี้จะใช้วิธีการที่เรียกว่า Transfer Learning หรือการใช้โมเดลที่มีการเรียนรู้มาอยู่ก่อนแล้ว (Pre-trained model) มาทำการฝึกต่อในด้านที่ต้องการมุ่งเน้น ในแง่ของงานวิจัยนี้คือการตรวจจับภาพใบหน้ามนุษย์ วิธีการนี้จะช่วยประหยัดเวลาและทรัพยากรที่ใช้ในการพัฒนาโมเดลมากกว่าการเริ่มพัฒนาโมเดลใหม่ตั้งแต่เริ่มแรก ซึ่งในงานวิจัยนี้จะเลือกใช้โมเดลทั้งหมด 3 โมเดลที่เลือกมาใช้ในการทดสอบ

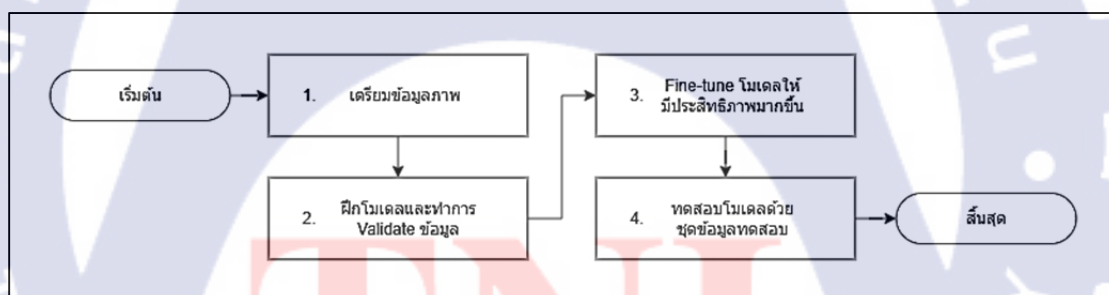
1. XceptionNet [35] เป็นโมเดลที่ผ่านการฝึกมาแล้วในระดับหนึ่ง (มีการ Pre-trained ด้วย ImageNet [36] ที่มีจำนวนประมาณ 1.2 ล้านภาพ) ทำให้โมเดลมีการเรียนรู้ Feature พื้นฐานของภาพมาแล้ว มีจำนวนพารามิเตอร์อยู่ที่ 22.8 ล้านพารามิเตอร์ ซึ่งเปรียบเทียบกับโมเดล Inception V3 [37] ที่มีจำนวนพารามิเตอร์ใกล้เคียงกันที่ 23.6 ล้านพารามิเตอร์แล้ว โมเดล Xception มีประสิทธิภาพที่สูงกว่า และมีการแยก Spatial Correlations (ความสัมพันธ์เชิงพื้นที่) และ Cross-Channel Correlations (ความสัมพันธ์ระหว่างช่องสัญญาณ) ออกจากกัน ซึ่งส่งผลต่อการสกัดฟีเจอร์ที่มีความละเอียดและซับซ้อนได้ดีกว่าการใช้ Convolution แบบปกติที่เรียนรู้ทั้งสองอย่างพร้อมกัน

2. MobileNetV3 [38] เป็นโมเดลที่มีจุดเด่นในด้านความสามารถในการใช้งานบนโทรศัพท์มือถือ หรืออุปกรณ์ที่มีทรัพยากรจำกัดโดยไม่ลดความสามารถของโมเดล โดย MobileNetV3 ถูกปรับแต่งให้โทรศัพท์มือถือสามารถใช้งานได้ ผ่านการใช้ Hardware Aware Network Architecture Search (NAS) ที่ถูกเสริมด้วยอัลกอริทึม NetAdapt [39] โดยโมเดล MobileNetV3 จะถูกแยกเป็น 2 โมเดลย่อยตามความเหมาะสมของทรัพยากรในการใช้งาน ได้แก่โมเดล MobileNetV3-Large ที่เหมาะ

กับอุปกรณ์ที่มีทรัพยากรสูงในระดับหนึ่ง และโมเดล MobileNetV3-Small ที่เหมาะกับอุปกรณ์ที่มีทรัพยากรน้อยกว่า โดยภายในงานวิจัยนี้จะใช้งานโมเดล MobileNetV3-Small เนื่องจากขนาดของโมเดลที่เล็กกว่าและความเร็วในการฝึกโมเดลที่มากกว่า

3. EfficientNetV2 [40] โมเดลที่มีการใช้แนวทางของ Compound Scaling ที่สามารถสร้างสมดุลระหว่างความลึก (จำนวน Layer ของโมเดล) ความกว้าง (จำนวน Channels ภายในแต่ละ Layer) และความละเอียด (ขนาดของภาพที่ใช้ในโมเดล) เพื่อให้ได้ประสิทธิภาพสูงที่สุด โดยโมเดล EfficientNetV2 ถูกแบ่งเป็น 3 โมเดลย่อยตามการปรับขนาด (Scaling) ของโมเดล ได้แก่โมเดล EfficientNetV2-S, EfficientNetV2-M และ EfficientNetV2-L ที่ขนาดของโมเดลเรียงจากเล็กที่สุดไปมากที่สุด ซึ่งผู้วิจัยได้เลือกใช้โมเดล EfficientNetV2-S ในงานวิจัยนี้ เนื่องจากขนาดของโมเดลที่เล็ก จากงานวิจัยของ T. Shahriar [41] โมเดล EfficientNetV2 มีค่าความแม่นยำโดยรวมสูงที่สุดในโมเดลทั้งหมด 5 โมเดลที่ผู้วิจัยเลือกใช้ในการทดสอบเปรียบเทียบประสิทธิภาพของโมเดล

ทำการฝึกโมเดลที่เลือกมา โดยแยกการฝึกระหว่างโมเดลตรวจจับภาพแว่นตาและโมเดลตรวจจับภาพหน้ากาก เนื่องจากต้องการวัดผลการตรวจจับแยกกันเพื่อให้สามารถเปรียบเทียบผลลัพธ์การตรวจจับกับโมเดล Rule-based ที่ทำการตรวจจับแบบแยกเครื่องประดับได้ มีขั้นตอนในการฝึกโมเดลก่อนนำโมเดลไปทดสอบ ตามแผนผังในรูปที่ 3.13 ดังต่อไปนี้



รูปที่ 3.13 แผนผังขั้นตอนการฝึกโมเดลด้วย Machine Learning

เริ่มจากการเตรียมข้อมูลภาพ โดยแบ่งเป็นข้อมูลส่วนการฝึกโมเดล (Training) ส่วนการตรวจสอบโมเดล (Validation) และส่วนทดสอบโมเดล (Testing) จากนั้นใช้โมเดลที่เลือกมาทำการฝึกด้วยภาพส่วนการฝึกโมเดล โดยกำหนดไว้ที่ 30 Epochs ต่อการฝึก 1 ครั้ง โดยในแต่ละ Epochs จะใช้ข้อมูลส่วนการตรวจสอบโมเดล (Validation) ในการตรวจสอบควบคู่ไปด้วย เพื่อทำการดูว่าเกิดการ Overfit ของข้อมูลระหว่างการฝึกโมเดลหรือไม่ และกำหนดให้โมเดลทำการสังเกตค่า Validation Loss ในแต่ละ Epochs หากค่า Validation Loss ไม่ดีขึ้นเลยเกิน 5 Epochs จะทำการหยุดการฝึกโมเดล

ทันที และเมื่อโมเดลฝึกสำเร็จแล้ว จะทำการบันทึกไฟล์โมเดลที่ฝึกไว้ ก่อนนำไปทำการ Fine-Tune ในขั้นตอนต่อไป

ในการ Fine-Tune โมเดลที่ฝึกมาแล้วในระดับหนึ่ง จะทำการเปิด Layer ส่วนหลักของโมเดลเพิ่มเติม โดยจะเปิดเฉพาะ Layer ในส่วนท้ายของโมเดล เนื่องจาก Layer ส่วนท้ายของโมเดลจะเกี่ยวกับการมองรูปร่างของสิ่งของ จึงต้องการเปิด Layer ส่วนนี้เพื่อให้โมเดลสามารถเรียนรู้ตาม Dataset ที่เรากำหนดไว้ได้ นอกจากนั้นยังปรับค่าอัตราการเรียนรู้ของโมเดล (Learning Rate) ให้น้อยลง เนื่องจากหากไม่ปรับค่าอัตราการเรียนรู้ของโมเดลให้ลดลง ค่า Loss อาจเพิ่มขึ้นหรือลดลงได้อย่างรวดเร็วในขณะที่ฝึก และอาจส่งผลให้ค่า Accuracy ลดลง รวมถึงเป็นการป้องกันไม่ให้ความรู้เดิมที่โมเดลได้ทำการฝึกและเรียนรู้ไปแล้วหายไป เนื่องจากหากใช้อัตราการเรียนรู้ของโมเดลที่สูงเกินไป ความรู้ที่โมเดลเรียนรู้ไปแล้วจาก Dataset ที่เราเตรียมไว้ อาจโดนความรู้เดิมของโมเดลเขียนทับไปได้ หลังจากปรับการตั้งค่าของโมเดลแล้ว นำโมเดลที่เราบันทึกไว้ก่อนหน้านี้มาทำการฝึกเพิ่มเติมด้วยการตั้งค่าใหม่ที่ตั้งไว้ โดยในครั้งนี้จะกำหนดให้ฝึกโมเดลเพิ่มอีก 20 Epochs และใช้เงื่อนไขเดิมคือการสังเกตค่า Validation Loss เพื่อดูว่า จะให้โมเดลหยุดการฝึกก่อนครบ 20 Epochs หรือไม่ และทำการบันทึกโมเดล Fine-Tune ที่มีค่า Validation Loss ต่ำที่สุดไว้

เมื่อทำการ Fine-Tune เสร็จสิ้น นำโมเดลที่บันทึกไว้ล่าสุดมาทำการทดสอบกับข้อมูลชุดทดสอบ (Testing) และทำการบันทึกผลเพื่อนำไปเปรียบเทียบข้อมูลในขั้นตอนวัดประสิทธิภาพ

3.4.4 ทดสอบประสิทธิภาพของโมเดล

ทำการทดสอบประสิทธิภาพของโมเดลด้วยข้อมูลภาพใน Dataset ส่วนสำหรับการทดสอบ (Testing) และทำการประเมินผลประสิทธิภาพของโมเดลที่พัฒนาขึ้น ด้วย Confusion Matrix และจัดเก็บผลลัพธ์ไว้สำหรับนำไปประเมินผลในขั้นตอนสุดท้าย ซึ่งในขั้นตอนนี้จะทำงานใกล้เคียงกับขั้นตอนวัดผลความแม่นยำและประสิทธิภาพของโมเดลที่พัฒนาด้วย Rule-based

3.5 การประเมินและวัดผลประสิทธิภาพโมเดล

ขั้นตอนสุดท้ายในการดำเนินงานวิจัยนี้คือการประเมินและวัดผลประสิทธิภาพโมเดลทั้ง 2 โมเดลที่ทำการพัฒนามาในขั้นตอนก่อนหน้านี้ ในการทดสอบโมเดลทั้ง 2 โมเดล ใช้อุปกรณ์เครื่องเดียวกันในการทดสอบ คือ เครื่องคอมพิวเตอร์โน้ตบุ๊ก รุ่น ASUS TUF Gaming F15 RAM 24GB และในการฝึกโมเดล ผู้วิจัยได้ใช้งาน Google Colab ในการฝึก โดยใช้ Python เวอร์ชัน 3.0 ขึ้นไป และใช้ตัวเร่งฮาร์ดแวร์ T4 GPU เนื่องจากเป็นตัวเร่งฮาร์ดแวร์ที่ไม่เสียเงินในการใช้งาน ซึ่งหลังจากการฝึกและทดสอบโมเดล จะทำการเปรียบเทียบค่าผลลัพธ์ที่ได้จาก Confusion Matrix โดยนำค่าที่ได้จากการวัดผลประสิทธิภาพ

ของแต่ละโมเดลมาทำการเปรียบเทียบกัน เพื่อดูและเปรียบเทียบประสิทธิภาพของโมเดลในแต่ละแง่มุม เช่น เปรียบเทียบสัดส่วนที่โมเดลสามารถทำนายได้ถูกต้อง ตรงตามความจริง เป็นต้น

โดยในส่วนของ การประเมินและวัดผลประสิทธิภาพโมเดลนี้ จะใช้ตัวชี้วัดทั้งหมด 4 อย่าง คือ ค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่าความอ่อนไหว (Recall) และค่าวัดประสิทธิภาพโดยรวม (F1-Score) สุดท้ายนี้ ทำการสรุปผลลัพธ์ทั้งหมดในรูปแบบตารางเพื่อให้ง่ายต่อการเปรียบเทียบ และทำความเข้าใจ



บทที่ 4

ผลการวิจัย

จากการศึกษาทฤษฎีและข้อมูลที่เกี่ยวข้องกับงานวิจัยนี้ ผู้วิจัยได้ทำการดำเนินงานวิจัยเพื่อศึกษาประสิทธิภาพของโมเดลที่พัฒนาด้วยวิธีการแบบ Rule-based และโมเดลที่พัฒนาด้วยการใช้ Machine Learning โดยมีผลการวิจัยแบ่งเป็น 4 ขั้นตอน ดังต่อไปนี้

- 4.1 ผลการหากฎเกณฑ์และเงื่อนไขสำหรับโมเดล Rule-based
 - 4.1.1 ผลการหาเงื่อนไขของภาพเครื่องประดับประเภทแว่นตา
 - 4.1.2 ผลการหาเงื่อนไขของภาพเครื่องประดับประเภทหน้ากาก
- 4.2 ผลการตรวจจับภาพด้วยโมเดล Rule-based
 - 4.2.1 ผลการตรวจจับภาพเครื่องประดับประเภทแว่นตา
 - 4.2.2 ผลการตรวจจับภาพเครื่องประดับประเภทหน้ากาก
- 4.3 ผลการตรวจจับภาพด้วยโมเดล Machine Learning
 - 4.3.1 ผลการตรวจจับเครื่องประดับประเภทแว่นตา
 - 4.3.2 ผลการตรวจจับภาพเครื่องประดับประเภทหน้ากาก
- 4.4 ผลการเปรียบเทียบประสิทธิภาพของโมเดล Rule-based และ Machine Learning

4.1 ผลการหากฎเกณฑ์และเงื่อนไขสำหรับโมเดล Rule-based

4.1.1 ผลการหาเงื่อนไขของภาพเครื่องประดับประเภทแว่นตา

ผู้วิจัยได้ทำการเก็บค่าแสงสว่างของพิกเซลที่อยู่ใน Mask ทั้ง 3 จุดในภาพชุดการฝึกโมเดลที่เป็นภาพแว่นตาจำนวน 60 ภาพ (ภาพเครื่องประดับจริง 30 ภาพ และเครื่องประดับ AR 30 ภาพ) จากนั้นทำการสังเกตและวิเคราะห์รูปแบบของข้อมูลที่ได้เพื่อทำการหาค่าที่สามารถกำหนดเป็นเกณฑ์ที่จะใช้ในการแยกแยะ โดยอ้างอิงคุณลักษณะจากตารางที่ 3.1 ได้ผลลัพธ์เงื่อนไขการแยกแยะภาพสวมแว่นตา ตามตารางที่ 4.1 ต่อไปนี้

ตารางที่ 4.1 เงื่อนไขการแยกแยะภาพสวมแว่นตา

คุณลักษณะ	สัญลักษณ์	ค่าในการแยกแยะ
ค่าเฉลี่ยของค่ารวมแสงสว่างรายกรอบบริเวณที่สนใจ	T_1	5.000
ค่าเฉลี่ยแสงสว่างของพิกเซลทั้งหมด	T_2	0.200
ค่าส่วนเบี่ยงเบนเฉลี่ยสมบูรณ์ (Mean AD)	T_3	0.070
ค่าส่วนเบี่ยงเบนมัธยฐานสมบูรณ์ (Median AD)	T_4	0.051
ค่าส่วนเบี่ยงเบนมาตรฐาน (SD)	T_5	0.084

จากตารางที่ 4.1 ผลลัพธ์เงื่อนไขที่ได้นี้มาจากการที่ผู้วิจัยได้ทำการทดลองเก็บค่าแสงสว่างของข้อมูลภาพชุดการฝึกโมเดล (Training) และทำการคำนวณค่าในแต่ละคุณลักษณะ จากนั้นทำการวิเคราะห์และปรับตัวเลขเกณฑ์ของค่าเฉลี่ยของค่ารวมแสงสว่างรายกรอบบริเวณที่สนใจ และค่าเฉลี่ยแสงสว่างของพิกเซลทั้งหมดที่ต้องทำการวิเคราะห์จากรูปแบบของข้อมูล เพื่อให้ได้ตัวเลขเงื่อนไขที่สามารถแยกแยะภาพได้ดีที่สุด ซึ่งจากเงื่อนไขที่ได้ เมื่อนำไปทดสอบกับภาพสำหรับการฝึกโมเดลทั้งหมด 60 ภาพแล้วได้ผลลัพธ์ ดังนี้

ตารางที่ 4.2 ผลลัพธ์การแยกแยะภาพแว่นตาด้วยเงื่อนไขการแยกแยะภาพสวมแว่นตา

ประเภทเครื่องประดับ	จำนวนผลการแยกแยะที่ถูกต้อง (ภาพ)	คิดเป็นอัตราส่วน (%)
ภาพแว่นตาจริง	19 จาก 30	63
ภาพแว่นตา AR	25 จาก 30	83

จากตารางที่ 4.2 ผลการแยกแยะภาพที่ถูกต้องของภาพแว่นตาจริงอยู่ที่ 63% และผลการแยกแยะภาพที่ถูกต้องของภาพแว่นตา AR อยู่ที่ 83% ซึ่งเป็นตัวเลขที่ดีที่สุดในการทดลอง

4.1.2 ผลการหาเงื่อนไขของภาพเครื่องประดับประเภทหน้ากาก

ผู้วิจัยได้ทำการเก็บค่าแสงสว่างของพิกเซลที่อยู่ใน Mask ทั้ง 3 จุดในภาพชุดการฝึกโมเดลที่เป็นภาพหน้ากากจำนวน 60 ภาพ (ภาพเครื่องประดับจริง 30 ภาพ และเครื่องประดับ AR 30 ภาพ) จากนั้นทำการสังเกตและวิเคราะห์รูปแบบของข้อมูลที่ได้เพื่อทำการหาค่าที่สามารถกำหนดเป็นเกณฑ์ที่จะใช้ในการแยกแยะ โดยอ้างอิงคุณลักษณะจากตารางที่ 3.1 ได้ผลลัพธ์เงื่อนไขการแยกแยะภาพสวมแว่นตา ตามตารางที่ 4.3 ต่อไปนี้

ตารางที่ 4.3 เงื่อนไขการแยกแยะภาพสวมหน้ากาก

คุณลักษณะ	สัญลักษณ์	ค่าในการแยกแยะ
ค่าเฉลี่ยของค่ารวมแสงสว่างรายกรอบบริเวณที่สนใจ	T'_1	190.000
ค่าเฉลี่ยแสงสว่างของพิกเซลทั้งหมด	T'_2	0.850
ค่าส่วนเบี่ยงเบนเฉลี่ยสมบูรณ์ (Mean AD)	T'_3	0.070
ค่าส่วนเบี่ยงเบนมัธยฐานสมบูรณ์ (Median AD)	T'_4	0.030
ค่าส่วนเบี่ยงเบนมาตรฐาน (SD)	T'_5	0.056

จากตารางที่ 4.3 ผลลัพธ์เงื่อนไขที่ได้นี้มาจากการที่ผู้วิจัยได้ทำการทดลองเก็บค่าแสงสว่างของข้อมูลภาพชุดการฝึกโมเดล (Training) และทำการคำนวณค่าในแต่ละคุณลักษณะ จากนั้นทำการวิเคราะห์และปรับตัวเลขเกณฑ์ของค่าเฉลี่ยของค่ารวมแสงสว่างรายกรอบบริเวณที่สนใจ และค่าเฉลี่ยแสงสว่างของพิกเซลทั้งหมดที่ต้องทำการวิเคราะห์จากรูปแบบของข้อมูล เพื่อให้ได้ตัวเลขเงื่อนไขที่สามารถแยกแยะภาพได้ดีที่สุด ซึ่งจากเงื่อนไขที่ได้ เมื่อนำไปทดสอบกับภาพสำหรับการฝึกโมเดลทั้งหมด 60 ภาพแล้วได้ผลลัพธ์ ดังนี้

ตารางที่ 4.4 ผลลัพธ์การแยกแยะภาพหน้ากากด้วยเงื่อนไขการแยกแยะภาพสวมหน้ากาก

ประเภทเครื่องประดับ	จำนวนผลการแยกแยะที่ถูกต้อง (ภาพ)	คิดเป็นอัตราส่วน (%)
ภาพหน้ากากจริง	17 จาก 30	56
ภาพหน้ากาก AR	26 จาก 30	87

จากตารางที่ 4.4 ผลการแยกแยะภาพที่ถูกต้องของภาพหน้ากากจริงอยู่ที่ 56% และผลการแยกแยะภาพที่ถูกต้องของภาพหน้ากาก AR อยู่ที่ 87% ซึ่งเป็นตัวเลขที่ดีที่สุดในการทดลอง

4.2 ผลการตรวจจับภาพด้วยโมเดล Rule-based

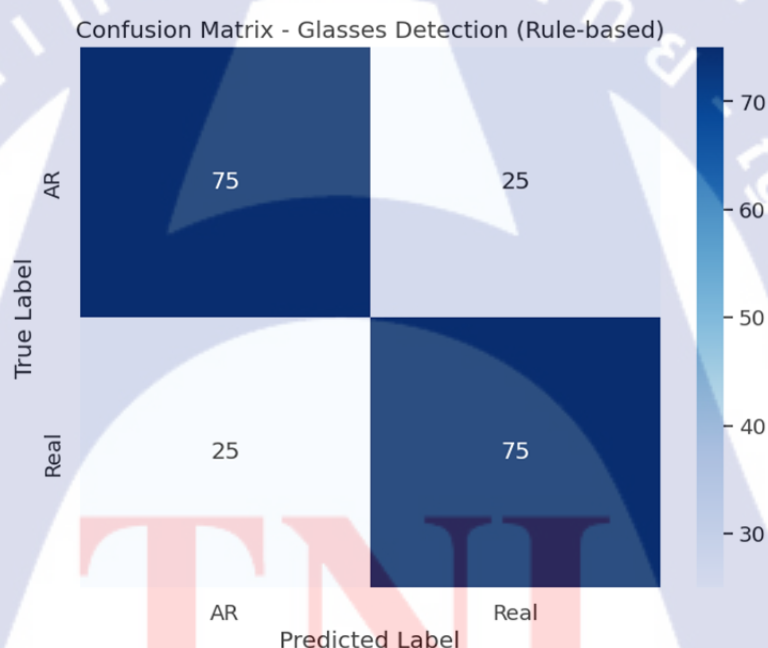
4.2.1 ผลการตรวจจับภาพเครื่องประดับประเภทแว่นตา

หลังจากได้กฎเกณฑ์สำหรับการแยกแยะภาพเครื่องประดับประเภทแว่นตาแล้ว ผู้วิจัยได้ทำการทดลองใช้กฎเกณฑ์นั้นกับข้อมูลชุดทดสอบ (Testing) ทั้งหมด 200 ภาพ (ภาพเครื่องประดับจริง 100 ภาพ และภาพเครื่องประดับ AR 100 ภาพ) ได้ผลลัพธ์การตรวจจับภาพ ดังนี้

ตารางที่ 4.5 ผลลัพธ์การแยกแยะภาพแว่นตาชุดทดสอบ ด้วยเงื่อนไขการแยกแยะภาพสวมแว่นตา

ประเภทเครื่องประดับ	จำนวนผลการแยกแยะที่ถูกต้อง (ภาพ)	คิดเป็นอัตราส่วน (%)
ภาพแว่นตาจริง	75 จาก 100	75
ภาพแว่นตา AR	75 จาก 100	75

จากตารางที่ 4.5 ผลการแยกแยะภาพที่ถูกต้องของภาพแว่นตาจริงอยู่ที่ 75% และผลการแยกแยะภาพที่ถูกต้องของภาพแว่นตา AR อยู่ที่ 75% เช่นเดียวกัน เมื่อนำผลลัพธ์นี้ไปสร้าง Confusion Matrix เพื่อคำนวณค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่าความอ่อนไหว (Recall) และค่า F1-Score โดยมีผลลัพธ์การสร้าง Confusion Matrix ของการแยกแยะภาพแว่นตา ดังรูปที่ 4.1 ต่อไปนี้



รูปที่ 4.1 ผลลัพธ์การสร้าง Confusion Matrix ภาพสวมแว่นตา

จากรูปที่ 4.1 พบว่า ค่า True Positive (TP) อยู่ที่ 75, ค่า True Negative (TN) อยู่ที่ 75, ค่า False Positive (FP) อยู่ที่ 25 และ ค่า False Negative (FN) อยู่ที่ 25 ซึ่งเมื่อนำค่าเหล่านี้มาคำนวณเพิ่มเติม จะได้ผลลัพธ์ดังตารางที่ 4.6 ต่อไปนี้

ตารางที่ 4.6 ผลการคำนวณค่าจาก Confusion Matrix ภาพสวมแว่นตา

ค่าในการวัดผล	ผลลัพธ์ที่ได้ (ภาพสวมแว่นตา)
Accuracy	0.75
Precision	0.75
Recall	0.75
F1-Score	0.75

นอกจากนั้นยังมีการเก็บข้อมูลการแยกแยะภาพของแต่ละคุณลักษณะ โดยมีรายละเอียดจำนวนผลการแยกแยะภาพของแต่ละคุณลักษณะ ดังนี้

ตารางที่ 4.7 ผลการแยกแยะภาพแว่นตาของแต่ละคุณลักษณะตามประเภทภาพ

คุณลักษณะ	แยกแยะถูกว่าเป็นภาพจริง	แยกแยะถูกว่าเป็นภาพ AR	รวมแยกแยะถูกทั้งหมด	คิดเป็นอัตราส่วน
ค่าเฉลี่ยของค่ารวมแสงสว่างรายกรอบบริเวณที่สนใจ	86/100	83/100	169/200	84.5%
ค่าเฉลี่ยแสงสว่างของฟิกเซลทั้งหมด	89/100	80/100	169/200	84.5%
ค่าส่วนเบี่ยงเบนเฉลี่ยสมบูรณ์ (Mean AD)	65/100	69/100	134/200	67%
ค่าส่วนเบี่ยงเบนมัธยฐานสมบูรณ์ (Median AD)	70/100	72/100	142/200	71%
ค่าส่วนเบี่ยงเบนมาตรฐาน (SD)	69/100	68/100	137/200	68.5%

จากตารางที่ 4.7 ผลการแยกแยะภาพแว่นตาของแต่ละคุณลักษณะตามประเภทสามารถสังเกตได้ว่า คุณลักษณะที่สามารถแยกแยะภาพแว่นตาจริงได้ถูกต้องมากที่สุดคือ ค่าเฉลี่ยแสงสว่างของฟิกเซลทั้งหมด ซึ่งสามารถแยกแยะได้ถูกว่าเป็นภาพจริงที่ 89 จาก 100 ภาพ หรือคิดเป็นอัตราส่วนที่ 89% และคุณลักษณะที่สามารถแยกแยะภาพแว่นตา AR ได้ถูกต้องมากที่สุดคือ ค่าเฉลี่ยของค่ารวมแสงสว่างรายกรอบบริเวณที่สนใจ ซึ่งสามารถแยกแยะได้ถูกว่าเป็นภาพ AR ที่ 83 จาก 100 ภาพ หรือคิดเป็นอัตราส่วนที่ 83% หากรวมสถิติของแต่ละคุณลักษณะโดยไม่แบ่งแยกตามประเภทของภาพแล้ว ค่าเฉลี่ยของค่ารวมแสงสว่างรายกรอบบริเวณที่สนใจ และค่าเฉลี่ยแสงสว่างของฟิกเซลรวมแล้วทำการแยกแยะภาพได้ถูกต้องมากที่สุด เท่ากันที่ 169 จาก 200 ภาพ หรือคิดเป็นอัตราส่วนที่ 84.5% ซึ่งจากผลลัพธ์นี้

สังเกตได้ว่าคุณลักษณะที่สามารถแยกแยะได้ถูกต้องน้อยที่สุดโดยรวมแล้วคือ ค่าส่วนเบี่ยงเบนเฉลี่ยสัมบูรณ์ (Mean AD) ซึ่งสันนิษฐานได้ว่าอาจเป็นเพราะการหาค่าเฉลี่ยโดยการรวมค่าที่เป็น Outliers อาจทำให้ผลลัพธ์ไม่ตรงกับสมมติฐานและเกณฑ์ที่ตั้งไว้ ส่งผลให้แยกแยะได้ถูกต้องน้อยลง

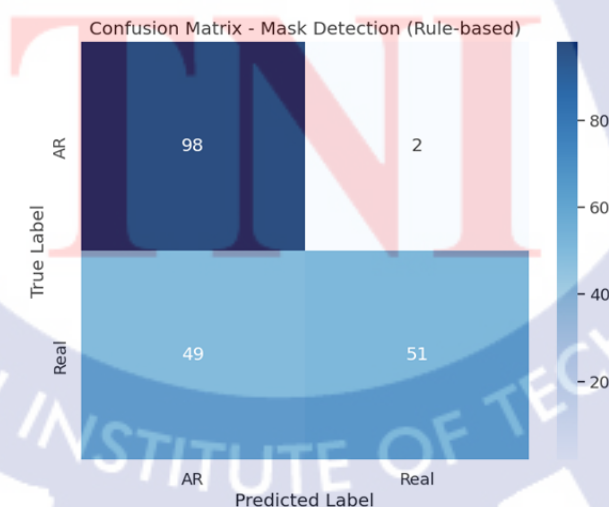
4.2.2 ผลการตรวจจับภาพเครื่องประดับประเภทหน้ากาก

หลังจากได้กฎเกณฑ์สำหรับการแยกแยะภาพเครื่องประดับประเภทหน้ากากแล้ว ผู้วิจัยได้ทำการทดลองใช้กฎเกณฑ์นั้นกับข้อมูลชุดทดสอบ (Testing) ทั้งหมด 200 ภาพ (ภาพเครื่องประดับจริง 100 ภาพ และภาพเครื่องประดับ AR 100 ภาพ) ได้ผลลัพธ์การตรวจจับภาพ ดังนี้

ตารางที่ 4.8 ผลลัพธ์การแยกแยะภาพหน้ากากชุดทดสอบ ด้วยเงื่อนไขการแยกแยะภาพสวมหน้ากาก

ประเภทเครื่องประดับ	จำนวนผลการแยกแยะที่ถูกต้อง (ภาพ)	คิดเป็นอัตราส่วน (%)
ภาพหน้ากากจริง	51 จาก 100	51
ภาพหน้ากาก AR	98 จาก 100	98

จากตารางที่ 4.8 ผลการแยกแยะภาพที่ถูกต้องของภาพหน้ากากจริงอยู่ที่ 51% และผลการแยกแยะภาพที่ถูกต้องของภาพหน้ากาก AR อยู่ที่ 98% ซึ่งมีความห่างกันค่อนข้างมาก เมื่อนำผลลัพธ์นี้ไปสร้าง Confusion Matrix เพื่าคำนวณค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่าความอ่อนไหว (Recall) และค่าวัดประสิทธิภาพโดยรวม (F1-Score) โดยมีผลลัพธ์การสร้าง Confusion Matrix ของการแยกแยะภาพหน้ากาก ดังรูปที่ 4.2 ต่อไปนี้



รูปที่ 4.2 ผลลัพธ์การสร้าง Confusion Matrix ภาพสวมหน้ากาก

จากรูปที่ 4.2 พบว่า ค่า True Positive (TP) อยู่ที่ 98, ค่า True Negative (TN) อยู่ที่ 51, ค่า False Positive (FP) อยู่ที่ 49 และค่า False Negative (FN) อยู่ที่ 2 ซึ่งเมื่อนำค่าเหล่านี้มาคำนวณเพิ่มเติม จะได้ผลลัพธ์ดังตารางที่ 4.9 ต่อไปนี้

ตารางที่ 4.9 ผลการคำนวณค่าจาก Confusion Matrix ภาพสวมน้ำกาก

ค่าในการวัดผล	ผลลัพธ์ที่ได้ (ภาพสวมน้ำกาก)
Accuracy	0.75
Precision	0.67
Recall	0.98
F1-Score	0.80

นอกจากนั้นยังมีการเก็บข้อมูลการแยกแยะภาพของแต่ละคุณลักษณะ โดยมีรายละเอียดจำนวนผลการแยกแยะภาพของแต่ละคุณลักษณะ ดังนี้

ตารางที่ 4.10 ผลการแยกแยะภาพน้ำกากของแต่ละคุณลักษณะตามประเภทภาพ

คุณลักษณะ	แยกแยะถูกว่าเป็นภาพจริง	แยกแยะถูกว่าเป็นภาพ AR	รวมแยกแยะถูกทั้งหมด	คิดเป็นอัตราส่วน
ค่าเฉลี่ยของค่ารวมแสงสว่าง รายการรอบบริเวณที่สนใจ	47/100	99/100	146/200	73%
ค่าเฉลี่ยแสงสว่างของพิกเซลทั้งหมด	47/100	97/100	144/200	72%
ค่าส่วนเบี่ยงเบนเฉลี่ยสมบูรณ์ (Mean AD)	53/100	97/100	150/200	75%
ค่าส่วนเบี่ยงเบนมัธยฐานสมบูรณ์ (Median AD)	52/100	94/100	146/200	73%
ค่าส่วนเบี่ยงเบนมาตรฐาน (SD)	54/100	88/100	142/200	71%

จากตารางที่ 4.10 ผลการแยกแยะภาพน้ำกากของแต่ละคุณลักษณะตามประเภทสามารถสังเกตได้ว่า คุณลักษณะที่สามารถแยกแยะภาพน้ำกากจริงได้ถูกต้องมากที่สุดคือ ค่าส่วนเบี่ยงเบนมาตรฐาน (SD) ซึ่งสามารถแยกแยะได้ถูกว่าเป็นภาพจริงที่ 54 จาก 100 ภาพ หรือคิดเป็นอัตราส่วนที่

54% และคุณลักษณะที่สามารถแยกแยะภาพหน้ากาก AR ได้ถูกต้องมากที่สุดคือ ค่าเฉลี่ยของค่ารวมแสงสว่างรายกรอบบริเวณที่สนใจ ซึ่งสามารถแยกแยะได้ถูกกว่าเป็นภาพ AR ที่ 99 จาก 100 ภาพ หรือคิดเป็นอัตราส่วนที่ 99% หากรวมสถิติของแต่ละคุณลักษณะโดยไม่แบ่งแยกตามประเภทของภาพแล้ว ค่าเฉลี่ยของค่ารวมแสงสว่างรายกรอบบริเวณที่สนใจ และค่าเฉลี่ยแสงสว่างของพิกเซลรวมแล้วทำการแยกแยะภาพได้ถูกต้องมากที่สุด ที่ 150 จาก 200 ภาพ หรือคิดเป็นอัตราส่วนที่ 75% ซึ่งจากผลลัพธ์นี้สังเกตได้ว่า เงื่อนไขที่ใช้อยู่ในปัจจุบันสามารถใช้ในการแยกแยะภาพเครื่องประดับที่เป็นภาพ AR ได้ดีมาก โดยแต่ละคุณลักษณะสามารถแยกแยะได้ถูกต้องเฉลี่ย 95% แต่ในส่วนของการแยกแยะภาพเครื่องประดับที่เป็นภาพถ่ายจริงนั้น ยังไม่สามารถแยกแยะได้ดี โดยแต่ละคุณลักษณะสามารถแยกแยะได้ถูกต้องเฉลี่ยเพียง 51% เท่านั้น จึงเป็นข้อสังเกตว่าเงื่อนไขที่ใช้อยู่ในปัจจุบันยังไม่สามารถแยกแยะระหว่างภาพเครื่องประดับที่เป็นภาพจริง และภาพเครื่องประดับที่เป็นภาพ AR ได้ดีมากพอ และจำเป็นต้องมีการปรับปรุงเงื่อนไขต่อไป

4.3 ผลการตรวจจับภาพด้วยโมเดล Machine Learning

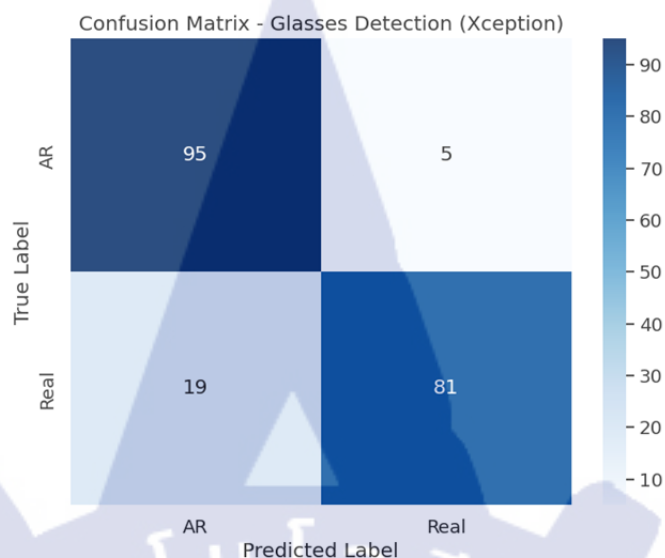
ในส่วนของการทดลองตรวจจับภาพด้วยโมเดล Machine Learning ผู้วิจัยได้เลือกโมเดลที่ต้องการใช้ในการทดสอบไว้ 3 โมเดล ได้แก่ โมเดล Xception, โมเดล MobileNetV3 และโมเดล EfficientNetV2 ซึ่งทั้ง 3 โมเดลนี้จะใช้ข้อมูลแบบเดียวกันทั้งหมดในส่วนการฝึกโมเดล (Training) และใช้ข้อมูลส่วนเดียวกันการตรวจสอบโมเดล (Validation) สุดท้ายจึงใช้ข้อมูลส่วนการทดสอบโมเดล (Testing) ในการทดสอบประสิทธิภาพของโมเดล โดยข้อมูลส่วนการทดสอบนี้ จะใช้ข้อมูลภาพแบบเดียวกันกับข้อมูลส่วนการทดสอบที่ใช้ในการทดสอบประสิทธิภาพของโมเดล Rule-based เพื่อให้สามารถเปรียบเทียบการตรวจจับภาพได้ในเงื่อนไขที่คล้ายกันมากที่สุด

ในการตรวจจับภาพด้วยโมเดล Machine Learning จะทำการวัดผลแยกกันระหว่างการตรวจจับภาพแว่นตาและการตรวจจับภาพหน้ากาก โดยใช้โมเดลแต่ละประเภททำการฝึกทีละประเภทเครื่องประดับ ในการตรวจจับภาพด้วยโมเดล และผู้วิจัยได้ทำการทดลองตรวจจับโมเดลด้วย Google Colab โดยใช้ Run-time ประเภท T4 GPU เพื่อให้สามารถฝึกโมเดลและทดสอบตรวจจับภาพด้วยโมเดลต่างๆ ได้รวดเร็วยิ่งขึ้น

4.3.1 ผลการตรวจจับภาพเครื่องประดับประเภทแว่นตา

1. โมเดล Xception

จากการใช้โมเดล Xception ในการฝึกโมเดลและทำการทดสอบการตรวจจับภาพเครื่องประดับประเภทแว่นตา มีผลลัพธ์การตรวจจับภาพ ดังนี้



รูปที่ 4.3 ผลลัพธ์ Confusion Matrix ภาพสวมแว่นตาของโมเดล Xception

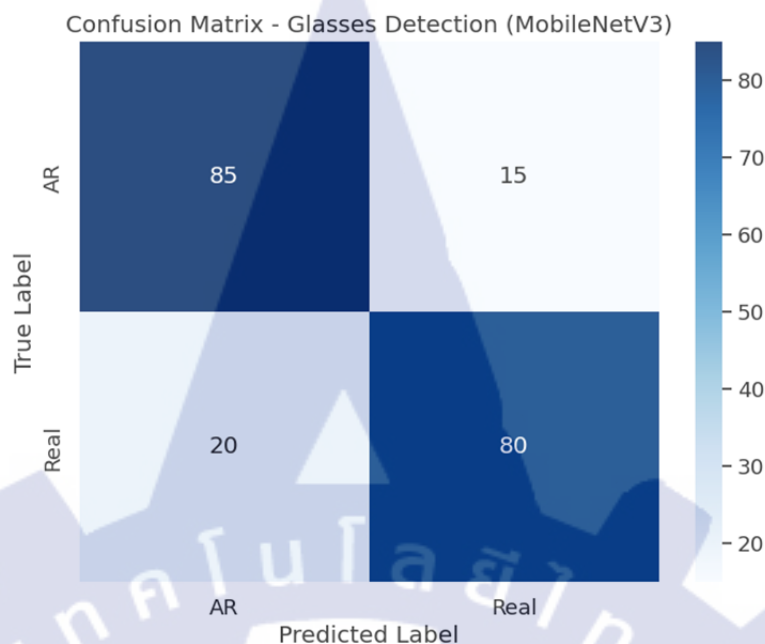
จากรูปที่ 4.3 พบว่า โมเดล Xception ที่ทำการฝึกโมเดลด้วยภาพแว่นตา ได้ผลค่า True Positive (TP) อยู่ที่ 95, ค่า True Negative (TN) อยู่ที่ 81, ค่า False Positive (FP) อยู่ที่ 19 และค่า False Negative (FN) อยู่ที่ 5 ซึ่งเมื่อนำค่าเหล่านี้มาคำนวณเพิ่มเติม จะได้ผลลัพธ์ดังตารางที่ 4.11 ต่อไปนี้

ตารางที่ 4.11 ผลการคำนวณค่าจาก Confusion Matrix โมเดล Xception (ภาพแว่นตา)

ค่าในการวัดผล	ผลลัพธ์ที่ได้ (โมเดล Xception)
Accuracy	0.88
Precision	0.83
Recall	0.95
F1-Score	0.88

2. โมเดล MobileNetV3

จากการใช้โมเดล MobileNetV3 ในการฝึกโมเดลและทำการทดสอบการตรวจจับภาพเครื่องประดับประเภทแว่นตา มีผลลัพธ์การตรวจจับภาพ ดังนี้



รูปที่ 4.4 ผลลัพธ์ Confusion Matrix ภาพสวมแว่นตาของโมเดล MobileNetV3

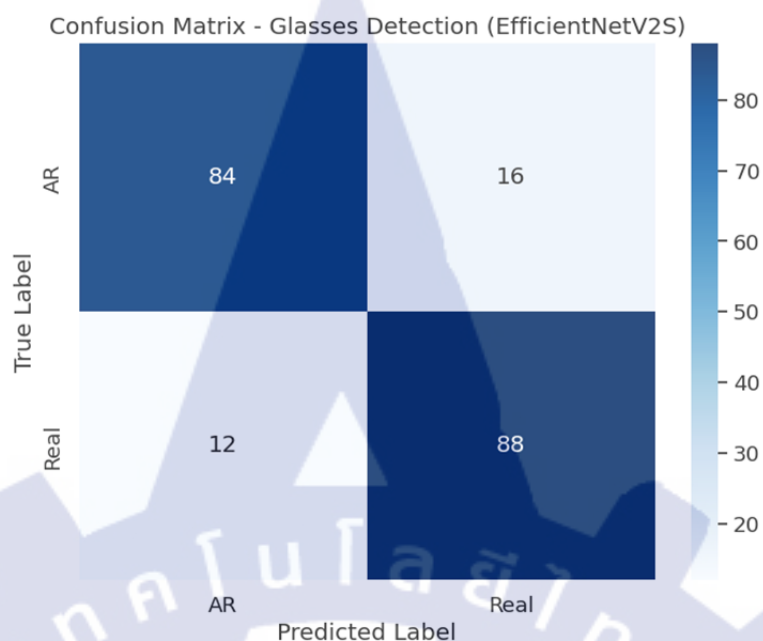
จากรูปที่ 4.4 พบว่า โมเดล MobileNetV3 ที่ทำการฝึกโมเดลด้วยภาพแว่นตา ได้ผลค่า True Positive (TP) อยู่ที่ 85, ค่า True Negative (TN) อยู่ที่ 80, ค่า False Positive (FP) อยู่ที่ 20 และค่า False Negative (FN) อยู่ที่ 15 ซึ่งเมื่อนำค่าเหล่านี้มาคำนวณเพิ่มเติม จะได้ผลลัพธ์ดังตารางที่ 4.12 ต่อไปนี้

ตารางที่ 4.12 ผลการคำนวณค่าจาก Confusion Matrix โมเดล MobileNetV3 (ภาพแว่นตา)

ค่าในการวัดผล	ผลลัพธ์ที่ได้ (โมเดล MobileNetV3)
Accuracy	0.82
Precision	0.80
Recall	0.85
F1-Score	0.83

3. โมเดล EfficientNetV2S

จากการใช้โมเดล EfficientNetV2S ในการฝึกโมเดลและทำการทดสอบการตรวจจับภาพเครื่องประดับประเภทแว่นตา มีผลลัพธ์การตรวจจับภาพ ดังนี้



รูปที่ 4.5 ผลลัพธ์ Confusion Matrix ภาพสวมแว่นตาของโมเดล EfficientNetV2S

จากรูปที่ 4.5 พบว่า โมเดล EfficientNetV2S ที่ทำการฝึกโมเดลด้วยภาพแว่นตา ได้ผลค่า True Positive (TP) อยู่ที่ 84, ค่า True Negative (TN) อยู่ที่ 88, ค่า False Positive (FP) อยู่ที่ 12 และค่า False Negative (FN) อยู่ที่ 16 ซึ่งเมื่อนำค่าเหล่านี้มาคำนวณเพิ่มเติม จะได้ผลลัพธ์ดังตารางที่ 4.13 ต่อไปนี้

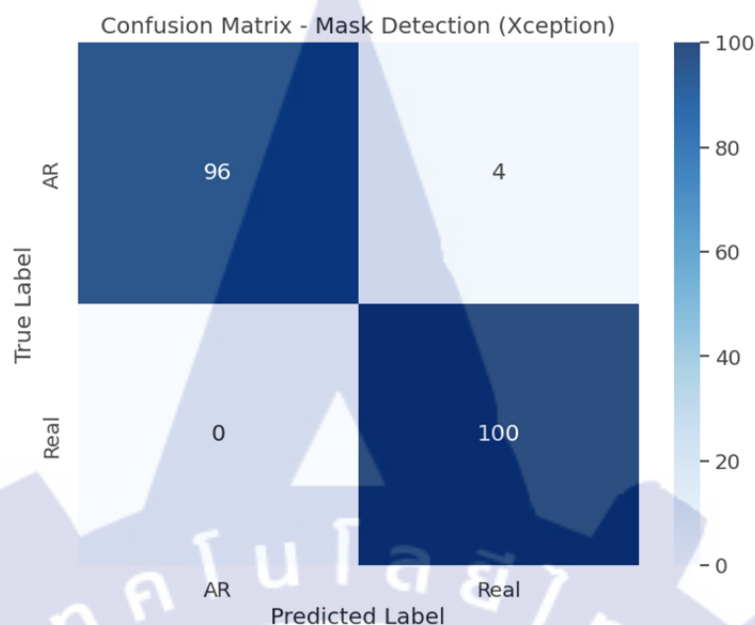
ตารางที่ 4.13 ผลการคำนวณค่าจาก Confusion Matrix โมเดล EfficientNetV2S (ภาพแว่นตา)

ค่าในการวัดผล	ผลลัพธ์ที่ได้ (โมเดล EfficientNetV2S)
Accuracy	0.86
Precision	0.87
Recall	0.84
F1-Score	0.86

4.3.2 ผลการตรวจจับภาพเครื่องประดับประเภทหน้ากาก

1. โมเดล Xception

จากการใช้โมเดล Xception ในการฝึกโมเดลและทำการทดสอบการตรวจจับภาพเครื่องประดับประเภทหน้ากาก มีผลลัพธ์การตรวจจับภาพ ดังนี้



รูปที่ 4.6 ผลลัพธ์ Confusion Matrix ภาพสวมหน้ากากของโมเดล Xception

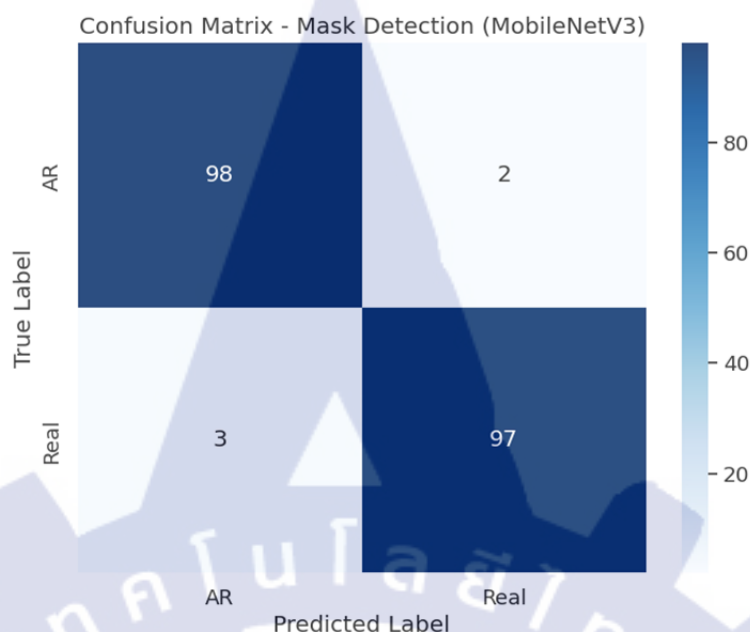
จากรูปที่ 4.6 พบว่า โมเดล Xception ที่ทำการฝึกโมเดลด้วยภาพหน้ากาก ได้ผลค่า True Positive (TP) อยู่ที่ 96, ค่า True Negative (TN) อยู่ที่ 100, ค่า False Positive (FP) อยู่ที่ 0 และค่า False Negative (FN) อยู่ที่ 4 ซึ่งเมื่อนำค่าเหล่านี้มาคำนวณเพิ่มเติม จะได้ผลลัพธ์ดังตารางที่ 4.14 ต่อไปนี้

ตารางที่ 4.14 ผลการคำนวณค่าจาก Confusion Matrix โมเดล Xception (ภาพหน้ากาก)

ค่าในการวัดผล	ผลลัพธ์ที่ได้ (โมเดล Xception)
Accuracy	0.98
Precision	1
Recall	0.96
F1-Score	0.98

2. โมเดล MobileNetV3

จากการใช้โมเดล MobileNetV3 ในการฝึกโมเดลและทำการทดสอบการตรวจจับภาพเครื่องประดับประเภทหน้ากาก มีผลลัพธ์การตรวจจับภาพ ดังนี้



รูปที่ 4.7 ผลลัพธ์ Confusion Matrix ภาพสวมหน้ากากของโมเดล MobileNetV3

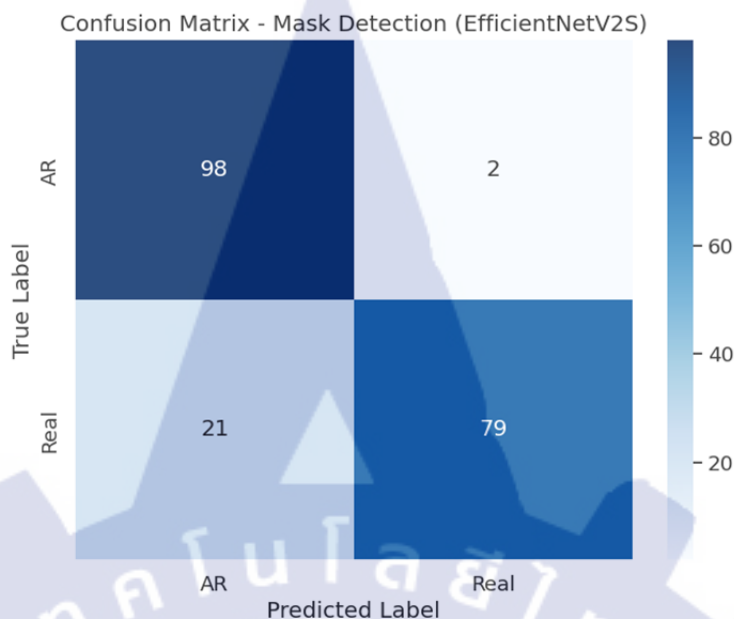
จากรูปที่ 4.7 พบว่า โมเดล MobileNetV3 ที่ทำการฝึกโมเดลด้วยภาพหน้ากาก ได้ผลค่า True Positive (TP) อยู่ที่ 98, ค่า True Negative (TN) อยู่ที่ 97, ค่า False Positive (FP) อยู่ที่ 3 และค่า False Negative (FN) อยู่ที่ 2 ซึ่งเมื่อนำค่าเหล่านี้มาคำนวณเพิ่มเติม จะได้ผลลัพธ์ดังตารางที่ 4.15 ต่อไปนี้

ตารางที่ 4.15 ผลการคำนวณค่าจาก Confusion Matrix โมเดล MobileNetV3 (ภาพหน้ากาก)

ค่าในการวัดผล	ผลลัพธ์ที่ได้ (โมเดล MobileNetV3)
Accuracy	0.97
Precision	0.97
Recall	0.98
F1-Score	0.97

3. โมเดล EfficientNetV2S

จากการใช้โมเดล EfficientNetV2S ในการฝึกโมเดลและทำการทดสอบการตรวจจับภาพเครื่องประดับประเภทหน้ากาก มีผลลัพธ์การตรวจจับภาพ ดังนี้



รูปที่ 4.8 ผลลัพธ์ Confusion Matrix ภาพสวมหน้ากากของโมเดล EfficientNetV2S

จากรูปที่ 4.8 พบว่า โมเดล EfficientNetV2S ที่ทำการฝึกโมเดลด้วยภาพหน้ากาก ได้ผลค่า True Positive (TP) อยู่ที่ 98, ค่า True Negative (TN) อยู่ที่ 79, ค่า False Positive (FP) อยู่ที่ 21 และค่า False Negative (FN) อยู่ที่ 2 ซึ่งเมื่อนำค่าเหล่านี้มาคำนวณเพิ่มเติม จะได้ผลลัพธ์ดังตารางที่ 4.16 ต่อไปนี้

ตารางที่ 4.16 ผลการคำนวณค่าจาก Confusion Matrix โมเดล EfficientNetV2S (ภาพหน้ากาก)

ค่าในการวัดผล	ผลลัพธ์ที่ได้ (โมเดล EfficientNetV2S)
Accuracy	0.88
Precision	0.82
Recall	0.98
F1-Score	0.89

4.4 ผลการเปรียบเทียบประสิทธิภาพของโมเดล Rule-based และ Machine Learning

ในการตรวจสอบประสิทธิภาพของโมเดล ทำการทดสอบด้วยข้อมูลภาพใน Dataset ส่วนสำหรับการทดสอบ (Testing) และทำการประเมินผลประสิทธิภาพของโมเดลที่พัฒนาขึ้น ว่าสามารถตรวจจับและแบ่งแยกระหว่างภาพสวมเครื่องประดับและภาพที่เพิ่มเครื่องประดับด้วย AR ได้ถูกต้องตามวัตถุประสงค์ที่ต้องการหรือไม่ โดยทำการกำหนดนิยามค่าในการวัดประสิทธิภาพของโมเดลดังต่อไปนี้

1. ค่าความถูกต้อง (Accuracy) คือ อัตราส่วนข้อมูลที่โมเดลทำนายถูกต้อง
 2. ค่าความแม่นยำ (Precision) คือ อัตราส่วนข้อมูลที่โมเดลทำนายว่าเป็น Positive และถูกต้องจริง
 3. ค่าความอ่อนไหว (Recall) คือ อัตราส่วนข้อมูลที่โมเดลทำนายว่าเป็น Positive เทียบกับจำนวนข้อมูล Positive ทั้งหมด
 4. ค่าวัดประสิทธิภาพโดยรวม (F1-Score) คือ ค่าที่เกิดขึ้นจากการคำนวณค่าเฉลี่ยของค่า Precision และค่า Recall
- สามารถแสดงนิยามในรูปแบบตารางได้ ดังตารางที่ 4.17 ต่อไปนี้

ตารางที่ 4.17 นิยามค่าในการวัดประสิทธิภาพของโมเดล

ลำดับ	ค่าในการวัดผล	นิยาม	คำอธิบาย
1	ค่าความถูกต้อง (Accuracy)	อัตราส่วนข้อมูลที่โมเดลทำนายถูกต้อง	จากข้อมูลทั้งหมด โมเดลทำนายถูกต้องเป็นจำนวนเท่าไร
2	ค่าความแม่นยำ (Precision)	อัตราส่วนข้อมูลที่โมเดลทำนายว่าเป็น Positive และถูกต้องจริง	เมื่อโมเดลทำนายว่าเป็นภาพสวมเครื่องประดับจริง มีความถูกต้องกี่เปอร์เซ็นต์
3	ค่าความอ่อนไหว (Recall)	อัตราส่วนข้อมูลที่เป็น Positive ทั้งหมด เทียบกับจำนวนที่โมเดลทำนายถูกต้อง	ในจำนวนข้อมูลที่เป็นภาพสวมเครื่องประดับจริงทั้งหมด โมเดลสามารถทำนายถูกกี่เปอร์เซ็นต์
4	ค่าวัดประสิทธิภาพโดยรวม (F1-Score)	ค่าที่เกิดขึ้นจากการคำนวณค่าเฉลี่ยของค่า Precision และค่า Recall	ค่าเฉลี่ยระหว่างค่า Precision และค่า Recall

จากผลลัพธ์การตรวจจับภาพเครื่องประดับจริงและเครื่องประดับ AR ของโมเดล Rule-based และโมเดล Machine Learning แล้ว แต่ละโมเดลมีประสิทธิภาพแตกต่างกันไปทั้งการตรวจจับภาพเครื่องประดับที่เป็นภาพจริงและภาพ AR รวมถึงความแตกต่างระหว่างประสิทธิภาพในการตรวจจับภาพเครื่องประดับที่เป็นแว่นตาและหน้ากาก ในการเปรียบเทียบประสิทธิภาพของแต่ละประเภทโมเดลสามารถแสดงผลการเปรียบเทียบโมเดลโดยแบ่งตามประเภทภาพเครื่องประดับได้ ดังตารางต่อไปนี้

ตารางที่ 4.18 ผลการเปรียบเทียบประสิทธิภาพของโมเดลในการตรวจจับภาพแว่นตา

ค่าในการวัดผล	Rule-based Model	Machine Learning (Deep Learning Model)		
		Xception	MobileNetV3	EfficientNetV2S
Accuracy	0.75	0.88	0.82	0.86
Precision	0.75	0.83	0.80	0.87
Recall	0.75	0.95	0.85	0.84
F1-Score	0.75	0.88	0.83	0.86

จากข้อมูลในตารางที่ 4.18 สามารถพิจารณาได้ว่า ในการตรวจจับภาพแว่นตา โมเดลที่ได้ค่าความถูกต้อง (Accuracy) สูงที่สุดคือ โมเดล Xception โดยมีค่าความถูกต้อง (Accuracy) อยู่ที่ 88% ค่าความแม่นยำ (Precision) อยู่ที่ 83% ค่าความอ่อนไหว (Recall) อยู่ที่ 95% และค่าวัดประสิทธิภาพโดยรวม (F1-Score) อยู่ที่ 88% โดยโมเดลที่มีค่าความถูกต้องสูงเป็นลำดับถัดมาคือ โมเดล EfficientNetV2S ที่มีค่าความถูกต้องที่ 86% ลำดับถัดมาคือโมเดล MobileNetV3 ที่มีค่าความถูกต้องอยู่ที่ 82% และโมเดล Rule-based ที่มีค่าความถูกต้องน้อยที่สุดอยู่ที่ 75%

ซึ่งเมื่อนำตัวเลขเหล่านี้มาทำการวิเคราะห์แล้ว สามารถทำความเข้าใจได้ว่า จากข้อมูลทั้งหมด โมเดล Xception ทำนายได้ถูกต้องมากที่สุด ในโมเดลทั้งหมด เป็นจำนวน 88% และมีค่าความแม่นยำ (Precision) ที่ 0.83 หมายถึง เมื่อโมเดล Xception ทำนายว่าเป็นภาพที่เห็นเป็นภาพเครื่องประดับ AR จะมีความถูกต้องอยู่ที่ 83%

ในส่วนของค่าความอ่อนไหว (Recall) โมเดลที่มีค่านี้สูงที่สุดยังคงเป็นโมเดล Xception หมายถึง ในจำนวนข้อมูลที่เป็นภาพสวมเครื่องประดับ AR ทั้งหมด โมเดล Xception สามารถทำนายได้ถูกต้องมากที่สุด ที่ 95% และในส่วน of ค่าวัดประสิทธิภาพโดยรวม (F1-Score) หมายถึง ค่าเฉลี่ยระหว่างค่าความแม่นยำ (Precision) และค่าความอ่อนไหว (Recall) โมเดล Xception มีค่านี้สูงที่สุดเช่นกัน ซึ่งมีค่าเท่ากับ 88%

ตารางที่ 4.19 ผลการเปรียบเทียบประสิทธิภาพของโมเดลในการตรวจจับภาพหน้ากา

ค่าในการวัดผล	Rule-based Model	Machine Learning (Deep Learning Model)		
		Xception	MobileNetV3	EfficientNetV2S
Accuracy	0.75	0.98	0.97	0.88
Precision	0.67	1	0.97	0.82
Recall	0.98	0.96	0.98	0.98
F1-Score	0.8	0.98	0.97	0.89

จากข้อมูลในตารางที่ 4.19 สามารถพิจารณาได้ว่า ในการตรวจจับภาพหน้ากาก โมเดลที่ได้ค่าความถูกต้อง (Accuracy) สูงที่สุดคือ โมเดล Xception โดยมีค่าความถูกต้อง (Accuracy) อยู่ที่ 98% ค่าความแม่นยำ (Precision) อยู่ที่ 100% ค่าความอ่อนไหว (Recall) อยู่ที่ 96% และค่าวัดประสิทธิภาพโดยรวม (F1-Score) อยู่ที่ 98% โดยโมเดลที่มีค่าความถูกต้องสูงเป็นลำดับถัดมาคือ โมเดล MobileNetV3 ที่มีค่าความถูกต้องที่ 97% ลำดับถัดมาคือโมเดล EfficientNetV2S ที่มีค่าความถูกต้องอยู่ที่ 88% และโมเดล Rule-based ที่มีค่าความถูกต้องน้อยที่สุดอยู่ที่ 75%

ซึ่งเมื่อนำตัวเลขเหล่านี้มาทำการวิเคราะห์แล้ว สามารถทำความเข้าใจได้ว่า ค่าความถูกต้อง (Accuracy) ของโมเดลนั้น หมายถึงจากข้อมูลทั้งหมด โมเดล Xception ทำนายได้ถูกต้องมากที่สุด โมเดลทั้งหมด เป็นจำนวน 98% และมีค่าความแม่นยำ (Precision) ที่ 1 หมายถึง เมื่อโมเดล Xception ทำนายว่าเป็นภาพที่เห็นเป็นภาพเครื่องประดับ AR จะมีความถูกต้องอยู่ที่ 100%

ในส่วนของค่าความอ่อนไหว (Recall) โมเดลที่มีค่านี้สูงที่สุดไม่ใช่โมเดล Xception เหมือนกับสองค่าก่อนหน้า แต่เป็นโมเดล Rule-based, โมเดล MobileNetV3 และโมเดล EfficientNetV2S ที่มีค่านี้สูงที่สุดเท่ากัน หมายถึง ในจำนวนข้อมูลที่เป็นภาพสวมเครื่องประดับ AR ทั้งหมด โมเดล 3 โมเดลนี้สามารถทำนายได้ถูกต้องมากที่สุดเท่ากัน ที่ 98% และในส่วน of ค่าวัดประสิทธิภาพโดยรวม (F1-Score) หมายถึง ค่าเฉลี่ยระหว่างค่าความแม่นยำ (Precision) และค่าความอ่อนไหว (Recall) โมเดล Xception มีค่านี้สูงที่สุด ซึ่งมีค่าเท่ากับ 98%

เมื่อเปรียบเทียบระหว่าง 2 ประเภทเครื่องประดับ ได้มีข้อสังเกตเกี่ยวกับโมเดลที่มีค่าความถูกต้องสูงที่สุดในแต่ละประเภทเครื่องประดับ โดยทั้งภาพแว่นตาและภาพหน้ากาก โมเดล Xception มีค่าความถูกต้องสูงที่สุด เนื่องจากในการฝึกโมเดลทุกโมเดลได้ใช้ข้อมูลภาพแบบเดียวกันทั้งชุดข้อมูลฝึกโมเดล (Training) ข้อมูลตรวจสอบโมเดล (Validation) และข้อมูลทดสอบโมเดล (Testing) ซึ่งมีข้อสันนิษฐานว่าแต่ละโมเดลอาจมีการเรียนรู้และสังเกตรูปแบบของข้อมูลแต่ละประเภทเครื่องประดับแตกต่างกัน ซึ่งการที่โมเดล Xception มีค่าความถูกต้องสูงที่สุดนั้น คาดว่ามีปัจจัยหลักมาจากโครงสร้าง Depthwise Separable Convolutions ซึ่งสามารถแยกการประมวลผลความสัมพันธ์เชิงพื้นที่ (Spatial Correlations) และความสัมพันธ์ระหว่างช่องสัญญาณ (Cross-channel Correlations) ได้อย่างมีประสิทธิภาพ ส่งผลให้โมเดล Xception สามารถตรวจจับร่องรอยการดัดแปลงที่มีความละเอียดสูง นอกจากนี้ การที่ความแม่นยำในการตรวจจับหน้ากากอนามัยสูงกว่าการตรวจจับแว่นตา แสดงให้เห็นว่าพื้นที่ผิวของเครื่องประดับที่มีขนาดใหญ่กว่า ช่วยให้กระบวนการสกัดคุณลักษณะ หรือ Feature ของภาพ สามารถสร้างแผนผังของคุณลักษณะดังกล่าวที่เป็นตัวแทนของข้อมูลได้ดีกว่า ในส่วนของโมเดล Rule-based ที่ทำการพัฒนากฎเกณฑ์และเงื่อนไขขึ้นเองนั้น มีค่าความถูกต้องในระดับหนึ่ง แต่ยังไม่มากเท่ากับโมเดล Machine Learning ที่ทำการเรียนรู้รูปแบบของภาพเองโดยอัตโนมัติ ผลการวิเคราะห์ชี้ให้เห็นถึงความเอนเอียงในการตรวจจับหน้ากากอนามัยแบบ AR โดยมีค่าความอ่อนไหว (Recall) สูงถึง 0.98 แต่ค่าความแม่นยำ (Precision)

ต่ำกว่า อยู่ที่ 0.67 สะท้อนว่าโมเดลมีการจำแนกหน้ากากจริงผิดว่าเป็นหน้ากาก AR เนื่องจากการซ้อนทับกันของความสว่างภายใต้สภาพที่มีแสงจ้า ในทางกลับกัน เครื่องประดับประเภทแว่นตาแสดงค่าที่สมมาตรกันในทุกตัววัดที่ 0.75 ซึ่งบ่งชี้ว่าข้อจำกัดอยู่ที่ขั้นตอนการตัดแบ่งส่วนพื้นที่ที่สนใจ (ROI) และการขาดกฎเกณฑ์ที่จำเป็นต่อการแยกแยะโครงสร้างของกรอบแว่นที่ค่อนข้างบาง ออกจากภาพซ้อนทับทางดิจิทัล อีกหนึ่งปัจจัยที่อาจส่งผลให้ประสิทธิภาพของโมเดล Rule-based ไม่ดีเทียบเท่าโมเดล Machine Learning อาจเป็นเพราะจำนวนข้อมูลภาพที่ใช้ในการฝึกฝนโมเดล Rule-based มีไม่มากเท่าจำนวนข้อมูลภาพที่ใช้ในการฝึกฝนโมเดล Machine Learning ทำให้กฎเกณฑ์ที่กำหนดอาจยังไม่ครอบคลุมลักษณะข้อมูลของภาพเครื่องประดับแต่ละประเภทได้ ซึ่งจำเป็นต้องทำการศึกษาเพิ่มเติมและพัฒนาวิธีการในการหารูปแบบของข้อมูลและการกำหนดกฎเกณฑ์ต่อไป



บทที่ 5

บทสรุป อภิปราย และข้อเสนอแนะ

งานวิจัยนี้ในหัวข้อ การตรวจจับภาพความจริงเสริมที่มีการเพิ่มเครื่องประดับบนใบหน้า มีวัตถุประสงค์หลักเพื่อการพัฒนาวิธีการตรวจจับภาพความจริงเสริม (Augmented Reality) ที่มีการเพิ่มเครื่องประดับในบริเวณใบหน้า ควบคู่กับการทดลองด้วยวิธีการที่มีอยู่ในปัจจุบัน และเปรียบเทียบประสิทธิภาพของแต่ละวิธีการตรวจจับ โดยในบทนี้จะทำการสรุปผลการวิจัย และอภิปรายเกี่ยวกับปัญหาและข้อจำกัด และนำเสนอข้อเสนอแนะสำหรับการพัฒนางานวิจัยต่อไปในอนาคต

5.1 สรุปผลการวิจัย

จากการเปรียบเทียบประสิทธิภาพของโมเดลที่ใช้ในการตรวจจับภาพความจริงเสริมที่มีการเพิ่มเครื่องประดับบนใบหน้า โดยเปรียบเทียบระหว่างโมเดล Rule-based ที่ทำการพัฒนาและกำหนดเกณฑ์ในการตรวจจับภาพขึ้นเอง กับโมเดล Machine Learning ที่มีอยู่แล้วทั้งหมด 3 โมเดล ได้แก่ โมเดล Xception, โมเดล MobileNetV3 และโมเดล EfficientNetV2S สามารถสรุปผลการวิจัยได้ ดังนี้

5.1.1 จากการศึกษาทฤษฎีต่างๆ ที่เกี่ยวข้องกับการตรวจจับภาพ เพื่อกำหนดกฎเกณฑ์และเงื่อนไขที่สามารถนำมาใช้ในการตรวจจับระหว่างภาพเครื่องประดับจริงและเครื่องประดับ AR นั้น ได้กับการสังเกตรูปแบบค่าแสงสว่างของพิกเซลในบริเวณที่เป็นจุดสนใจ (ยกตัวอย่างเช่น การคำนวณค่าเฉลี่ยแสงสว่างภายในภาพ การคำนวณค่าส่วนเบี่ยงเบนด้วยวิธีการต่างๆ ทั้งการคำนวณส่วนเบี่ยงเบนมาตรฐาน หรือส่วนเบี่ยงเบนเฉลี่ยสัมบูรณ์ เป็นต้น) สามารถแบ่งแยกระหว่างรูปแบบข้อมูลของเครื่องประดับที่เป็นภาพถ่ายจริง และเครื่องประดับที่เป็นภาพ AR ที่แต่งเติมเข้ามาในภาพได้ในระดับหนึ่ง

5.1.2 ในการทดลองค้นหากฎเกณฑ์และเงื่อนไขที่นำมากำหนดในการตรวจจับภาพเครื่องประดับของโมเดล Rule-based สามารถกำหนดกฎเกณฑ์และเงื่อนไขที่ใช้ในการตรวจจับภาพได้ โดยเงื่อนไขของเครื่องประดับแต่ละชนิดจะใช้เงื่อนไขแบบเดียวกัน แต่มีความแตกต่างกันในส่วน of ค่าที่เป็นตัวกำหนดเพื่อแบ่งแยกว่าเป็นภาพเครื่องประดับจริงหรือเครื่องประดับ AR ซึ่งผลการทดลองแสดงให้เห็นว่า เกณฑ์ที่กำหนดไว้เมื่อทดสอบกับภาพส่วนของการฝึกโมเดลจำนวน 120 ภาพ (แบ่งเป็นภาพแว่นตา 60 ภาพ ซึ่งประกอบไปด้วยภาพแว่นตาจริง 30 ภาพ และภาพแว่นตา AR 30 ภาพ กับภาพหน้ากาก 60 ภาพ ซึ่งประกอบไปด้วยภาพหน้ากากจริง 30 ภาพ และภาพหน้ากาก AR 30 ภาพ) สามารถแบ่งแยกได้ในระดับหนึ่ง โดยเงื่อนไขของภาพแว่นตาสามารถแบ่งแยกภาพแว่นตาจริงได้ถูกต้องที่ 63% และแบ่งแยกภาพแว่นตา AR ได้ถูกต้องที่ 83% ส่วนเงื่อนไขของภาพหน้ากากนั้น สามารถแบ่งแยกภาพหน้ากากจริงได้ถูกต้องที่ 56% และแบ่งแยกภาพหน้ากาก AR ได้ถูกต้องที่ 87%

5.1.3 ฟังก์ชันการตรวจจับภาพด้วยโมเดล Machine Learning ด้วยโมเดลทั้งหมด 3 โมเดลที่ได้ทำการเลือกไว้ ได้แก่ โมเดล Xception, โมเดล MobileNetV3 และโมเดล EfficientNetV2S ด้วยข้อมูลทั้งหมด 2,000 ภาพ ที่ทำการแบ่งส่วนไว้ล่วงหน้า ได้แก่ข้อมูลส่วนทดสอบ (Training) 1,400 ภาพ ซึ่งประกอบไปด้วยภาพเครื่องบินระดับแนวหน้า 700 ภาพ (ภาพแนวหน้าจริง 350 ภาพ และภาพแนวหน้า AR 350 ภาพ) และภาพเครื่องบินระดับหน้ากัก 700 ภาพ (ภาพหน้ากักจริง 350 ภาพ และภาพหน้ากัก AR 350 ภาพ) ข้อมูลส่วนตรวจสอบโมเดล (Validation) 200 ภาพ ซึ่งประกอบไปด้วยภาพเครื่องบินระดับแนวหน้า 100 ภาพ (ภาพแนวหน้าจริง 50 ภาพ และภาพแนวหน้า AR 50 ภาพ) และภาพเครื่องบินระดับหน้ากัก 100 ภาพ (ภาพหน้ากักจริง 50 ภาพ และภาพหน้ากัก AR 50 ภาพ) และข้อมูลส่วนสุดท้ายที่ยังไม่ใช้ในขั้นตอนการฝึกโมเดล นั่นคือข้อมูลส่วนทดสอบโมเดล 400 ภาพ โดยทำการฝึกโมเดลเบื้องต้นก่อน และทำการ Fine-tune โมเดลด้วยการปรับค่า Parameter ในการฝึก และนำโมเดลที่ผ่านการฝึกเบื้องต้นแล้วมาฝึกซ้ำอีกครั้งเพื่อเพิ่มประสิทธิภาพในการตรวจจับภาพ

5.1.4 ในการตรวจจับภาพด้วยโมเดล Rule-based โดยใช้เงื่อนไขที่กำหนดขึ้นเองกับภาพชุดข้อมูลทดสอบจำนวน 400 ภาพ (แบ่งเป็นภาพแนวหน้า 200 ภาพ ซึ่งประกอบไปด้วยภาพแนวหน้าจริง 100 ภาพ และภาพแนวหน้า AR 100 ภาพ กับภาพหน้ากัก 200 ภาพ ซึ่งประกอบไปด้วยภาพหน้ากักจริง 100 ภาพ และภาพหน้ากัก AR 100 ภาพ) ผลลัพธ์การตรวจจับภาพแนวหน้า สามารถแบ่งแยกแนวหน้าจริงได้ถูกต้องที่ 75% และแบ่งแยกภาพแนวหน้า AR ได้ถูกต้องที่ 75% ส่วนผลลัพธ์การตรวจจับภาพหน้ากัก นั้น สามารถแบ่งแยกหน้ากักจริงได้ถูกต้องที่ 51% และแบ่งแยกภาพหน้ากัก AR ได้ถูกต้องที่ 98%

5.1.5 ในการตรวจจับภาพด้วยโมเดล Machine Learning ด้วยโมเดลทั้งหมด 3 โมเดลที่ได้ทำการเลือกไว้ ทดสอบการตรวจจับภาพด้วยข้อมูลภาพส่วนทดสอบ (Testing) จำนวน 400 ภาพ ซึ่งประกอบไปด้วยภาพเครื่องบินระดับแนวหน้า 200 ภาพ (ภาพแนวหน้าจริง 100 ภาพ และภาพแนวหน้า AR 100 ภาพ) และภาพเครื่องบินระดับหน้ากัก 200 ภาพ (ภาพหน้ากักจริง 100 ภาพ และภาพหน้ากัก AR 100 ภาพ) ผลลัพธ์ในการตรวจจับภาพด้วยโมเดลทั้ง 3 โมเดลนั้น ได้ผลลัพธ์ว่า ในการตรวจจับภาพแนวหน้า โมเดล Xception มีค่าความถูกต้องสูงที่สุดที่ 88% ส่วนในการตรวจจับภาพหน้ากัก โมเดล Xception มีค่าความถูกต้องสูงที่สุดที่ 98%

5.1.6 จากการเปรียบเทียบผลลัพธ์การตรวจจับภาพของโมเดลทั้ง 4 โมเดล (โมเดล Rule-based ที่ทำการกำหนดเงื่อนไขขึ้นเอง และโมเดล Machine Learning 3 โมเดล) แสดงให้เห็นว่า โดยภาพรวมโมเดล Machine Learning มีค่าความถูกต้องสูงกว่าโมเดล Rule-based ที่กำหนดเงื่อนไขขึ้นเองทั้งในส่วนของการตรวจจับภาพเครื่องบินระดับแนวหน้าและภาพเครื่องบินระดับหน้ากัก จึงสันนิษฐานได้ว่า นอกจากเงื่อนไขที่ใช้ในการตรวจจับภาพในโมเดล Rule-based ปัจจุบัน ยังมีเงื่อนไขอื่นๆ หรือรูปแบบข้อมูลอื่นที่สามารถแบ่งแยกระหว่างภาพเครื่องบินระดับจริงและภาพเครื่องบินระดับ AR ได้แม่นยำมากกว่านี้ ซึ่งต้องทำการพัฒนาต่อไปในอนาคต

5.1.7 จากผลลัพธ์ของงานวิจัย สามารถวิเคราะห์ได้ถึงสาเหตุที่โมเดล Rule-based มีประสิทธิภาพโดยรวมน้อยกว่าโมเดล Machine Learning เนื่องจากการพัฒนาโมเดล Rule-based ผู้พัฒนาต้องทำการวิเคราะห์ปัจจัยที่ใช้ในการแบ่งแยกและตรวจจับภาพแต่ละประเภท ซึ่งปัจจัยที่สามารถใช้ในการตรวจจับภาพนั้นมีมากจนการพัฒนาโมเดลด้วยตัวผู้พัฒนาเองอาจทำได้ไม่ครอบคลุมในทุกปัจจัย ในขณะเดียวกัน โมเดล Machine Learning มีการพัฒนาให้สามารถวิเคราะห์รูปแบบข้อมูลได้ด้วยตนเอง รวมถึงการวิเคราะห์ปัจจัยที่ใช้ในการตรวจจับภาพ ซึ่งโมเดล Machine Learning ที่สามารถวิเคราะห์ข้อมูลได้ด้วยตนเองนั้น สามารถมองได้หลากหลายปัจจัยมากกว่าโมเดล Rule-based ที่จำเป็นต้องใช้ความสามารถของผู้พัฒนาในการวิเคราะห์ปัจจัยแต่ละข้อ ซึ่งเป็นสาเหตุที่ทำให้โมเดล Machine Learning ที่สามารถวิเคราะห์ปัจจัยได้ครบถ้วนมากกว่า มีประสิทธิภาพในการตรวจจับภาพมากกว่าโมเดล Rule-based

5.2 อภิปรายผล

จากการศึกษาวิธีการต่างๆ ในการพัฒนาโมเดลที่สามารถตรวจจับภาพความจริงเสริมที่มีการเพิ่มเครื่องประดับบนใบหน้าได้ ผู้วิจัยได้ใช้วิธีการในการพัฒนาโมเดลทั้งหมด 2 รูปแบบ ได้แก่ การพัฒนาโมเดลตรวจจับภาพด้วยวิธี Rule-based ที่สามารถอธิบายเงื่อนไขและเหตุผลที่ใช้ในการทำนายผลลัพธ์ได้ และการพัฒนาโมเดลตรวจจับภาพด้วยการใช้ Machine Learning หรือการนำโมเดล Deep Learning มาใช้ในการพัฒนาเพิ่มเติมเพื่อให้โมเดลสามารถตรวจจับภาพในเงื่อนไขที่กำหนดในหัวข้อวิจัยนี้ได้ ในส่วนของการพัฒนาโมเดลตรวจจับภาพด้วยวิธี Rule-based ผู้วิจัยได้ใช้วิธีการดึงข้อมูลที่เป็นค่าแสงสว่างออกมาจากวัตถุเครื่องประดับภายในภาพเพื่อทำการคำนวณค่าต่างๆ และนำมากำหนดเงื่อนไขในการตรวจจับภาพ โดยโมเดล Rule-based ที่ผู้วิจัยพัฒนาขึ้นนั้น ในส่วนของภาพเครื่องประดับที่เป็นแว่นตานั้น สามารถตรวจจับได้โดยมีค่าความถูกต้อง (Accuracy) อยู่ที่ 75% และในส่วนของภาพเครื่องประดับที่เป็นหน้ากาก สามารถทำนายได้โดยมีค่าความถูกต้อง (Accuracy) อยู่ที่ 75% เช่นเดียวกัน ถึงแม้ทั้งสองเครื่องประดับจะมีค่าความถูกต้องเท่ากัน แต่ในส่วนของค่าอื่นๆ มีความแตกต่างกัน โดยการทำนายภาพแว่นตามีค่าความแม่นยำ (Precision) ที่ 75% ซึ่งมากกว่าภาพหน้ากากที่มีค่า 67% ข้อสังเกตนี้ชี้ให้เห็นว่า โมเดลมีความสามารถในการระบุภาพแว่นตา AR ได้ถูกต้องมากกว่าภาพหน้ากาก AR อย่างไรก็ตาม การที่ภาพหน้ากากมีค่า Precision ต่ำกว่า อาจมีสาเหตุมาจากการที่ภาพหน้ากากมีหลากหลายสี หรือมีค่าความสว่างที่ใกล้เคียงกันกับภาพหน้ากากจริง ซึ่งส่งผลให้โมเดลเกิดความสับสนในการจำแนกภาพได้ ในส่วนของค่าความอ่อนไหว (Recall) โมเดลสามารถตรวจจับภาพแว่นตา AR จากข้อมูลทั้งหมดที่มีได้ครอบคลุมถึง 75% และภาพหน้ากาก AR ที่ 98% ค่าความอ่อนไหวที่สูงนี้ชี้ให้เห็นว่าโอกาสที่ภาพหน้ากาก AR จะหลุดรอดจากการตรวจจับมีเพียง 2% เท่านั้น ซึ่งสอดคล้องกับวัตถุประสงค์ของงานวิจัยที่ต้องการตรวจจับภาพใบหน้าที่มีการเพิ่มเครื่องประดับ AR และในส่วนของค่าวัดประสิทธิภาพโดยรวม

(F1-Score) ภาพหน้ากากมีค่าวัดประสิทธิภาพโดยรวมอยู่ที่ 80% ซึ่งมากกว่าภาพแว่นตา AR ทั้งหมด 5% แสดงให้เห็นว่าสำหรับโมเดล Rule-based นี้ การตรวจจับภาพหน้ากากมีความสมดุระหว่างจำนวนการระบุภาพให้ถูกต้อง กับจำนวนภาพที่หลุดรอดจากการตรวจจับมากกว่าการตรวจจับภาพแว่นตา

นอกจากผลลัพธ์การตรวจจับภาพของโมเดล Rule-based แล้ว งานวิจัยนี้ยังแสดงถึงผลลัพธ์การตรวจจับภาพของโมเดล Machine Learning ซึ่งมีการเปรียบเทียบระหว่างโมเดล Deep Learning ที่ทำการเลือกมาทั้งหมด 3 โมเดล ได้แก่ โมเดล Xception [36], MobileNetV3 [39] และโมเดล EfficientNetV2S [41] จากนั้นนำผลลัพธ์การตรวจจับจากทั้ง 3 โมเดลมาเปรียบเทียบกันเพื่อสังเกตและวิเคราะห์ประสิทธิภาพของแต่ละโมเดล จากการเปรียบเทียบพบว่า ในการตรวจจับภาพแว่นตา AR โมเดล ที่มีค่าความถูกต้อง (Accuracy) สูงที่สุดได้แก่โมเดล Xception ซึ่งมีค่าอยู่ที่ 88% นอกจากนั้นโมเดล Xception ยังเป็นโมเดลที่มีค่าความอ่อนไหว (Recall) และค่าวัดประสิทธิภาพโดยรวม (F1-Score) สูงที่สุดในโมเดลตรวจจับภาพทั้งหมด แสดงให้เห็นว่าโมเดล Xception สามารถตรวจจับภาพแว่นตาได้ถูกต้องทั้งภาพที่เป็นภาพแว่นตาจริงและภาพแว่นตา AR และผลลัพธ์ดังกล่าวนี้แสดงให้เห็นว่า โมเดล Xception มีความเหมาะสมในการตรวจจับภาพแว่นตาสูงที่สุด ในส่วนของภาพหน้ากานั้น โมเดล Xception ยังคงเป็นโมเดลที่มีค่าความถูกต้อง (Accuracy) สูงที่สุดที่ 98% รวมถึงมีค่าความแม่นยำ (Precision) สูงถึง 100% และมีค่าวัดประสิทธิภาพโดยรวม (F1-Score) สูงที่สุดเช่นกัน ผลลัพธ์นี้แสดงให้เห็นว่าโมเดล Xception มีความสามารถในการตรวจจับภาพหน้ากากได้ถูกต้องเหนือกว่าโมเดลอื่นๆ

เมื่อสรุปจากผลลัพธ์ของทั้งการตรวจจับภาพแว่นตาและภาพหน้ากาก แสดงให้เห็นว่าโมเดล Xception มีความสามารถในการตรวจจับภาพที่มีการเพิ่มเครื่องประดับ AR โดยรวมสูงกว่าโมเดลอื่นๆ ที่นำมาเปรียบเทียบ ทั้งโมเดล Rule-based และโมเดล Machine Learning ส่งผลให้ในงานวิจัยนี้ โมเดล Xception เป็นโมเดลที่มีความเหมาะสมที่สุดในการตรวจจับภาพเครื่องประดับ AR ตามข้อมูลทำงานวิจัยนี้กำหนดไว้ ผลลัพธ์นี้มีความสอดคล้องกับข้อค้นพบล่าสุดในงานวิจัยของ G. Casella and R. L. Berger [33] ซึ่งได้พิสูจน์ให้เห็นว่า โมเดล Deep Learning มีประสิทธิภาพที่เหนือกว่าการวิเคราะห์แบบดั้งเดิมในด้านการตรวจจับภาพปลอม ความสอดคล้องดังกล่าวช่วยเน้นข้อสรุปของงานวิจัยนี้ที่ว่า กระบวนการเรียนรู้คุณลักษณะแบบอัตโนมัติมีความยืดหยุ่นและเสถียรมากกว่าการใช้เกณฑ์การตัดสินใจจากกฎที่กำหนดไว้ล่วงหน้า โดยเฉพาะในการระบุร่องรอยของภาพดิจิทัลที่มีความซับซ้อนซึ่งพบได้ในเครื่องประดับ AR ซึ่งจากผลลัพธ์การเปรียบเทียบประสิทธิภาพของโมเดลเหล่านี้ สามารถนำไปต่อยอดเกี่ยวกับการใช้งานในด้านระบบรักษาความปลอดภัย หรือการยืนยันตัวตนผ่านใบหน้า (Face Recognition) เพื่อใช้ในการตรวจจับภาพใบหน้าที่อาจมีการเพิ่มส่วนประกอบที่เป็นภาพ AR เพื่อปลอมแปลงตัวตนได้ โดยโมเดล Xception ที่มีประสิทธิภาพการตรวจจับโดยรวมสูงที่สุดในโมเดลทั้งหมดที่ใช้ในงานวิจัยนี้ อาจเป็นจุดเริ่มต้นที่ดีในการพัฒนาระบบตรวจสอบภาพใบหน้า โดยหากมีการฝึกและใช้งานโมเดลอย่างเหมาะสมด้วยข้อมูลประเภทอื่นๆ เช่น เครื่องประดับประเภทอื่นๆ บนใบหน้า หรือองค์ประกอบที่เป็นภาพ AR อื่นๆ

เพิ่มเติม อาจทำให้โมเดลสามารถแยกแยะระหว่างภาพใบหน้าจริงและภาพใบหน้าที่ถูกเพิ่มองค์ประกอบเสมือนได้อย่างมีประสิทธิภาพมากขึ้น ซึ่งสามารถนำไปต่อยอดในการพัฒนาระบบรักษาความปลอดภัยได้

5.3 ปัญหาและข้อจำกัด

5.3.1 ในการกำหนดเงื่อนไขของโมเดล Rule-based ในแง่ของการพิสูจน์ภาพดิจิทัล (Image Forensics) มีปัจจัยและข้อมูลอีกหลากหลายมิติที่สามารถนำมาใช้ในการกำหนดกฎเกณฑ์และเงื่อนไขในการตรวจจับความแตกต่างระหว่างภาพเครื่องประดับจริงและภาพเครื่องประดับ AR ซึ่งมีความเป็นไปได้หลากหลายรูปแบบ ด้วยข้อจำกัดด้านเวลาและกำลัง งานวิจัยนี้จึงได้ทำการเลือกมาเพียงวิธีการเดียวคือการวิเคราะห์รูปแบบข้อมูลค่าความสว่างของพิกเซลใน HSV Color Model

5.3.2 ในการทดลองเก็บค่าแสงสว่างของพิกเซลในขั้นตอนการฝึกโมเดล Rule-based มีข้อจำกัดในด้านจำนวนข้อมูลที่ใช้ในการเก็บค่าแสงสว่างเพื่อนำมาวิเคราะห์รูปแบบของข้อมูลในแต่ละประเภทเครื่องประดับ ซึ่งแตกต่างกับการฝึกโมเดล Machine Learning ที่สามารถใช้งานข้อมูลได้มากกว่าหลายเท่าตัว ทำให้ผลลัพธ์ในการตรวจจับภาพด้วยโมเดล Rule-based อาจไม่ได้ผลลัพธ์ที่ดีเท่าที่ควรจะเป็น

5.4 ข้อเสนอแนะ

ในการพัฒนาโมเดลสำหรับการตรวจจับภาพด้วยเงื่อนไขใดๆ โดยเฉพาะการพัฒนาโมเดลที่กำหนดเงื่อนไขเองอย่างโมเดล Rule-based ควรพิจารณาข้อมูลภาพในมิติอื่นๆ เพิ่มเติม เช่น การวิเคราะห์ข้อมูลในมิติความถี่ (Frequency Domain) ซึ่งแตกต่างกับการวิเคราะห์ข้อมูลในมิติเชิงพื้นที่ (Spatial Domain) แบบเดียวกันกับในงานวิจัยนี้ เพื่อพัฒนาความแม่นยำในการตรวจจับของโมเดล Rule-based และสามารถนำไปใช้เป็นหนึ่งในวิธีการตรวจจับภาพที่แม่นยำได้ในอนาคต รวมถึงการพัฒนาโมเดลด้วย Machine Learning ที่สามารถปรับปรุงส่วนการตั้งค่าที่ใช้ในการฝึกโมเดลให้เพิ่มการตรวจจับคุณลักษณะ (Feature) เพิ่มเติม โดยสามารถเพิ่มการตรวจจับให้สอดคล้องกับการตรวจจับด้วยคุณลักษณะแบบเดียวกันกับโมเดล Rule-based เพื่อเพิ่มประสิทธิภาพในการตรวจจับได้มากขึ้น รวมถึงการตั้งค่าอื่นๆ ที่มีส่วนช่วยในการฝึกโมเดลให้มีประสิทธิภาพมากขึ้น เช่น การฝึกโมเดลโดยไม่กำหนดจำนวน Epochs ซึ่งส่งผลให้โมเดลทำการฝึกฝนไปเรื่อยๆ จนกว่าค่า Validate_loss ที่แสดงถึงการเรียนรู้ของโมเดลจะหยุดนิ่งไปด้วยตัวเอง โมเดลที่ผ่านการฝึกด้วยวิธีการนี้จะได้ผลลัพธ์ออกมาเป็นโมเดลที่ทำการเรียนรู้จากข้อมูลที่มีจนถึงที่สุด โดยไม่มีการกำหนดจำนวนรอบในการเรียนรู้มาเป็นข้อจำกัดในการฝึกฝนโมเดล

นอกจากนี้ ในงานวิจัยในอนาคต การปรับปรุงข้อมูลที่ใช้ในการฝึกโมเดลอาจช่วยเพิ่มประสิทธิภาพในการตรวจจับของโมเดลได้ดียิ่งขึ้น ไม่ว่าจะเป็นในแง่ของปริมาณข้อมูลที่เพิ่มขึ้น หรือรูปแบบเครื่องประดับภายในภาพที่อาจทำให้สามารถวิเคราะห์และมองเห็นรูปแบบของข้อมูลที่ทำให้กำหนดกฎเกณฑ์ที่ตรวจจับได้ดียิ่งขึ้น



บรรณานุกรม

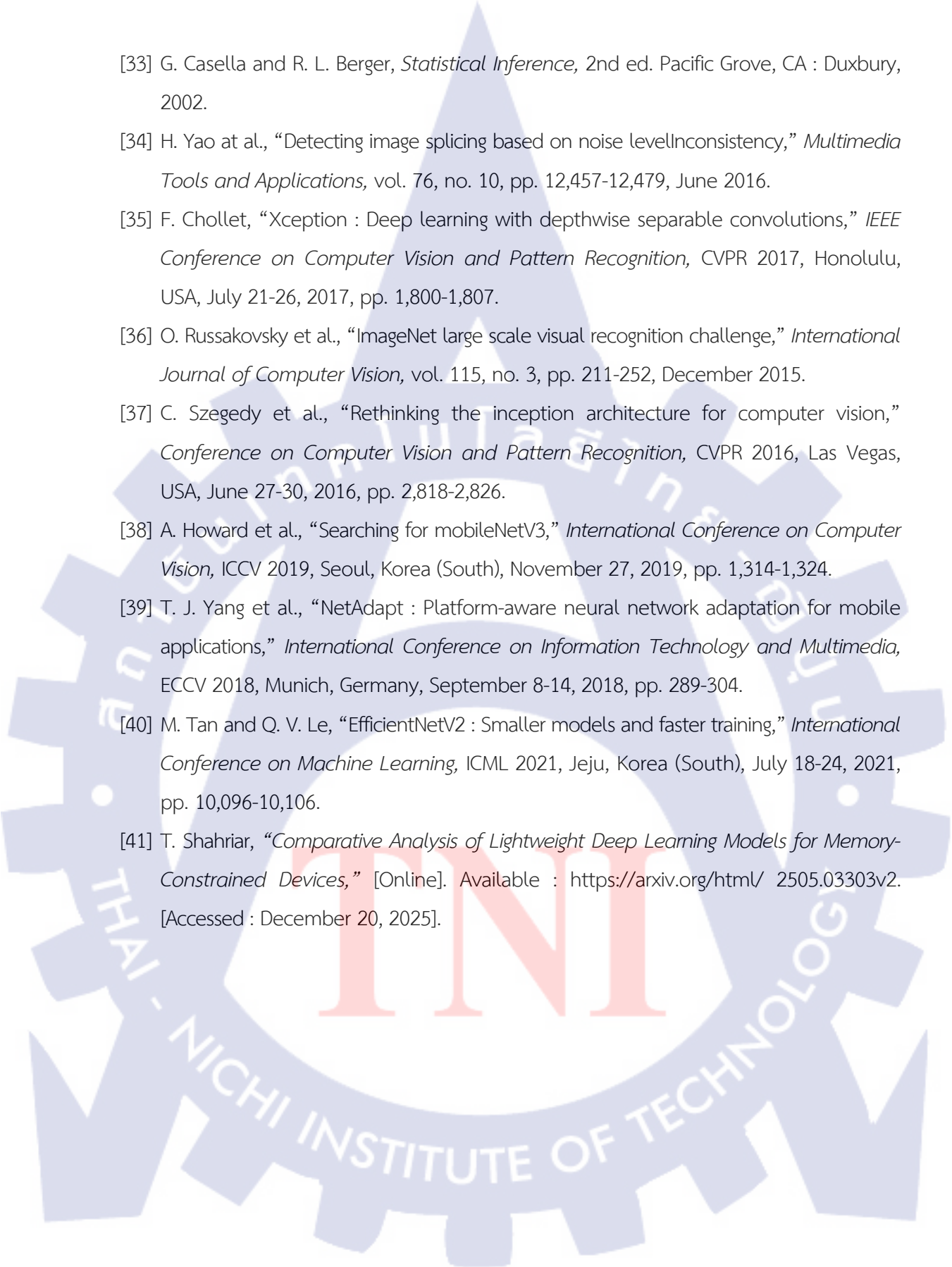
TNI

บรรณานุกรม

- [1] A. E. Sousa and P. Cardoso, "Use of generative AI by higher education students," *Electronics* 2025, vol. 14, no. 7, pp. 1-20, March 2025.
- [2] S. Sukyanee et al., "Development of instructional media of thai classical masked dance (Khon) conform to chant by using augmented technology (AR) for grade 6th students," *Journal of Technical Education Development*, vol. 35, no. 126, pp. 45-52, July 2023.
- [3] TJM, "Customers Pay \$40 for Signet's Digital-Only Jewelry," [Online]. Available : <https://www.thejewelrymagazine.com/customers-pay-40-for-signets-digital-only-jewelry>. [Accessed : August 20, 2025].
- [4] Thai PBS, "จับตาอาชญากรรมออนไลน์ปี 67 ใช้ AI สร้างคลิป-เลียนเสียงหลอกเหยื่อ," [Online]. Available : <https://www.thaipbs.or.th/news/content/335055>. [Accessed : 20 สิงหาคม 2568].
- [5] Z. Liu et al., "Deep learning face attributes in the wild in proceedings," *International Conference on Computer Vision, ICCV 2015*, Santiago, Chile, December 7-13, 2015, pp. 3,730-3,738.
- [6] TheSSJ2612, "Real-Time-Medical-Mask-Detection," [Online]. Available : <https://github.com/TheSSJ2612/Real-Time-Medical-Mask-Detection>. [Accessed : December 20, 2025].
- [7] L. Eversberg and J. Lambrecht, "Evaluating digital work instructions with augmented reality versus paper-based documents for manual, object-specific repair tasks in a case study with experienced workers," *The International Journal of Advanced Manufacturing Technology*, vol. 127, no. 3-4, pp. 1,859-1,871, June 2023.
- [8] M. G. Franco et al., "Immersive mixed reality for manufacturing training," *Frontiers in Robotics and AI*, vol. 4, no. 12, pp. 1-8, April 2017.
- [9] W. Suklim et al., "Development of image processing for tuna sizing classification," *International Conference on Information and Communication Technology Convergence, ICIT 2022*, Bangkok, Thailand, January 30 - February 2, 2022, pp. 811-818.

- [10] Z. Li, "Application research of digital image technology in graphic design," *Journal of Visual Communication and Image Representation*, vol. 65, no. 3, pp. 102,689, December 2019.
- [11] Forensic Focus, "The Complete Workflow of Forensic Image and Video Analysis," [Online]. Available : <https://www.forensicfocus.com/articles/the-complete-workflow-of-forensic-image-and-video-analysis>. [Accessed : October 12, 2025].
- [12] I. Amerini et al., *Image and Video Forensics*, Basel, Switzerland : MDPI, 2022.
- [13] CSIS, "How Accurate are Facial Recognition Systems and Why Does It Matter?," [Online]. Available : <https://www.csis.org/blogs/strategic-technologies-blog/how-accurate-are-facial-recognition-systems-and-why-does-it>. [Accessed : September 6, 2025].
- [14] OpenCV Tutorial, "Cascade Classifier," [Online]. Available : https://docs.opencv.org/3.4/db/d28/tutorial_cascade_classifier.html. [Accessed : September 8, 2025].
- [15] M. K. Hasan et al., "Human face detection techniques : A comprehensive review and future research directions," *Electronics 2021*, vol. 10, no. 19, pp. 2,354, September 2021.
- [16] D. Jin et al., "A study on the color and glossiness of polypropylene (PP) films based on the visual perception of elderly users," *Coatings*, vol. 16, no. 1, pp. 68, January 2026.
- [17] T. Kusnandar et al., "Enhancing color selection in HSV color space," *Ingénierie des Systèmes d'Information*, vol. 29, no. 4, pp. 1,483-1,491, May 2024.
- [18] V. Chugani, "Understanding Euclidean Distance : From Theory to Practice," [Online]. Available : <https://www.datacamp.com/tutorial/euclidean-distance>. [Accessed : February 6, 2026].
- [19] G. Lawton, "The Different Types of Machine Learning Explained," [Online]. Available : <https://www.techtarget.com/searchenterpriseai/tip/Types-of-learning-in-machine-learning-explained>. [Accessed : September 9, 2025].
- [20] W. Wei et al., "Assessing facial symmetry and attractiveness using augmented reality," *Pattern Analysis and Applications*, vol. 25, no. 6, pp. 635-651, August 2022.
- [21] D. Kang and L. Ma, "Real-time eye tracking for bare and sunglasses-wearing faces for augmented reality 3D head-up displays," *IEEE Access*, vol. 9, no. 1, pp. 125,508-125,522, September 2021.

- [22] P. Riccio et al., "OpenFilter : A framework to democratize research access to social media AR filters," *Proceedings of the 2022 Northeast Decision Science Institute Conference, NEDSI 2022*, New Orleans, USA, September 28 - December 9, 2022, pp. 1-13.
- [23] J. Yoneyama et al., "Augmented conversations : AR face filters for facilitating comfortable in-person interactions," *Journal on Multimodal User Interfaces*, vol. 19, no. 4, pp. 57-74, December 2024.
- [24] L. Nataraj et al., "Holistic image manipulation detection using pixel co-occurrence matrices," *Journal of Imaging*, vol. 33, no. 4, pp. 277-1-277-7, April 2021.
- [25] C. L. Li et al., "CutPaste : Self-supervised learning for anomaly detection and localization," *Conference on Computer Vision and Pattern Recognition, CVPR 2021*, Balaclava, Mauritius, June 19-25, 2021, pp. 9,659-9,669.
- [26] J. Yu et al., "FastFlow : Unsupervised Anomaly Detection and Localization via 2D Normalizing Flows," [Online]. Available : <https://arxiv.org/abs/2111.07677>. [Accessed : December 20, 2025].
- [27] D. Nimma et al., "Advanced image forensics : Detecting and reconstructing manipulated images with deep learning," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 4, pp. 283-290, May 2024.
- [28] C. Shi et al., "Review of image forensic techniques based on deep learning," *Mathematics 2023*, vol. 11, no. 14, pp. 3,134, July 2023.
- [29] M. Dang and T. N. Nguyen, "Digital face manipulation creation and detection : A systematic review," *Electronics 2023*, vol. 12, no. 16, pp. 3,407, August 2023.
- [30] S. Polat et al., "A new rule-based classification method using shape-based properties of spectral curves," *Journal of Spectroscopy*, vol. 20, no. 6, pp. 1-17, February 2022.
- [31] P. Duszejko et al., "Detection of manipulations in digital images : A review of passive and active methods utilizing deep learning," *Applied Sciences*, vol. 15, no. 2, pp. 881, January 2025.
- [32] S. Yadav and A. Kumar, "Combining deep learning and forensic analysis for robust fake image detection," *International Journal on Science and Technology*, vol. 16, no. 2, pp. 1-6, June 2025.

- 
- [33] G. Casella and R. L. Berger, *Statistical Inference*, 2nd ed. Pacific Grove, CA : Duxbury, 2002.
- [34] H. Yao et al., "Detecting image splicing based on noise level inconsistency," *Multimedia Tools and Applications*, vol. 76, no. 10, pp. 12,457-12,479, June 2016.
- [35] F. Chollet, "Xception : Deep learning with depthwise separable convolutions," *IEEE Conference on Computer Vision and Pattern Recognition*, CVPR 2017, Honolulu, USA, July 21-26, 2017, pp. 1,800-1,807.
- [36] O. Russakovsky et al., "ImageNet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211-252, December 2015.
- [37] C. Szegedy et al., "Rethinking the inception architecture for computer vision," *Conference on Computer Vision and Pattern Recognition*, CVPR 2016, Las Vegas, USA, June 27-30, 2016, pp. 2,818-2,826.
- [38] A. Howard et al., "Searching for mobileNetV3," *International Conference on Computer Vision*, ICCV 2019, Seoul, Korea (South), November 27, 2019, pp. 1,314-1,324.
- [39] T. J. Yang et al., "NetAdapt : Platform-aware neural network adaptation for mobile applications," *International Conference on Information Technology and Multimedia*, ECCV 2018, Munich, Germany, September 8-14, 2018, pp. 289-304.
- [40] M. Tan and Q. V. Le, "EfficientNetV2 : Smaller models and faster training," *International Conference on Machine Learning*, ICML 2021, Jeju, Korea (South), July 18-24, 2021, pp. 10,096-10,106.
- [41] T. Shahriar, "Comparative Analysis of Lightweight Deep Learning Models for Memory-Constrained Devices," [Online]. Available : <https://arxiv.org/html/2505.03303v2>. [Accessed : December 20, 2025].

ภาคผนวก

TNI

ภาคผนวก ก.

เอกสารการนำเสนอภาพนิ่งสอบป้องกันวิทยานิพนธ์

TNI

เอกสารการนำเสนอภาพนิ่งสอบป้องกันวิทยานิพนธ์

Detection of Augmented Reality Images with Facial Accessories

การตรวจจับภาพความจริงเสริมที่มีการเพิ่มเครื่องประดับบนใบหน้า

Natchaya Yimtanom

7 March, 2026

Agenda Overview

- 01 บทนำ
- 02 แนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้อง
- 03 กระบวนการวิจัย
- 04 ผลการวิจัย
- 05 บทสรุปและข้อเสนอแนะ
- 06 Q&A Session

01 บทนำ

ที่มาและความสำคัญ
ของงานวิจัย

วัตถุประสงค์และ
ประโยชน์ของงานวิจัย

ขอบเขตของงานวิจัย

ที่มาและความสำคัญของงานวิจัย

- ปัจจุบันเทคโนโลยี AI ถูกนำมาใช้ในชีวิตประจำวันอย่างแพร่หลาย แม้จะมีประโยชน์ต่อการเรียนรู้และสร้างสรรค์ แต่อาจก่อให้เกิดข้อมูลเท็จ การปลอมแปลงตัวตน หรือข่าวปลอมได้
- เทคโนโลยี AR (Augmented Reality) ถูกนำไปใช้ในหลายด้าน แต่การใช้ AR โดยขาดการควบคุม อาจก่อให้เกิดความเข้าใจผิด เช่น การสร้างภาพปลอมเพื่อหลอกลวงผู้อื่น



3

01 บทนำ

ที่มาและความสำคัญ
ของงานวิจัย

วัตถุประสงค์และ
ประโยชน์ของงานวิจัย

ขอบเขตของงานวิจัย

ที่มาและความสำคัญของงานวิจัย

- ปัจจุบันมีเครื่องมือหลายชนิดที่สามารถตรวจจับภาพปลอมที่สร้างด้วย AI ได้ค่อนข้างแม่นยำ
- แต่ภาพที่ผ่านการปรับแต่งด้วยเทคโนโลยี AR ยังไม่มีเครื่องมือเฉพาะทางที่สามารถตรวจจับได้อย่างแม่นยำ

พัฒนาวิธีการสำหรับตรวจจับภาพใบหน้าที่ถูกแต่งเติมด้วยฟิลเตอร์หรือภาพ AR ควบคู่กับการทดลองตรวจจับภาพด้วยโมเดล Machine Learning



4

01 บทนำ

ที่มาและความสำคัญ
ของงานวิจัย

วัตถุประสงค์และ
ประโยชน์ของงานวิจัย

ขอบเขตของงานวิจัย

วัตถุประสงค์และประโยชน์ของงานวิจัย

“

เพื่อพัฒนาวิธีการตรวจจับภาพความจริงเสริม
(Augmented Reality) ที่มีการเพิ่มเครื่องประดับในบริเวณใบหน้า
ควบคู่กับการทดลองด้วยโมเดล Machine Learning
และเปรียบเทียบประสิทธิภาพของแต่ละวิธีการตรวจจับ

”

โดยมีความคาดหวังว่า

- ได้องค์ความรู้ใหม่เกี่ยวกับการพัฒนาวิธีการตรวจจับภาพใบหน้าที่มีลักษณะเฉพาะ
- ได้แนวทางการประยุกต์ใช้โมเดลการตรวจจับภาพ
- ได้ข้อมูลเปรียบเทียบประสิทธิภาพของเทคนิคต่างๆ ในการตรวจจับใบหน้า



5

01 บทนำ

ที่มาและความสำคัญ
ของงานวิจัย

วัตถุประสงค์และ
ประโยชน์ของงานวิจัย

ขอบเขตของงานวิจัย

ขอบเขตของงานวิจัย

ประเภทข้อมูลภาพ



ภาพใบหน้าบุคคล



ภาพสี

ประเภทเครื่องประดับ



แว่นสายตา



หน้ากากอนามัย
(ที่ไม่มีลวดลาย)

6

01 บทนำ

ที่มาและความสำคัญ
ของงานวิจัย

วัตถุประสงค์และ
ประโยชน์ของงานวิจัย

ขอบเขตของงานวิจัย

ขอบเขตของงานวิจัย

แหล่งที่มาของภาพ



ภาพจากกล้องโทรศัพท์มือถือ



ภาพจาก Dataset ที่มีอยู่แล้ว
CelebA , Real-Time-Medical-Mask-Detection

7

02 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ทฤษฎีที่เกี่ยวข้อง

Augmented Reality	1	6	การวัดการกระจายตัวข้อมูล
Digital Image Processing	2	7	Euclidean Distance
Image Forensics	3	8	Rule-based System
Face Recognition	4	9	Machine Learning
HSV Color Model	5	10	Confusion Matrix

8

02 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ทฤษฎีที่เกี่ยวข้อง

1 2 3 4 5 6 7 8 9 10

1 ความจริงเสริม (Augmented Reality: AR)

- มีการใช้ AR เพิ่มภาพเครื่องประดับลงบนใบหน้าเพื่อนำไปใช้ในการตรวจจับภาพ



2 การประมวลผลข้อมูลภาพดิจิทัล (Digital Image Processing)

- มีการใช้ Digital Image Processing ในหลายขั้นตอนในงานวิจัย

- ปรับขนาดภาพก่อนนำไปเทรนโมเดล
- แปลงภาพจาก RGB ไปเป็น HSV
- คำนวณค่าทางสถิติจากค่าแสงสว่าง (Value)

9

02 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ทฤษฎีที่เกี่ยวข้อง

1 2 3 4 5 6 7 8 9 10

3 การพิสูจน์ภาพดิจิทัล (Image Forensics)

- ใช้ Image Forensics ในขั้นตอนค้นหาเงื่อนไขการตรวจจับภาพ

- ตรวจสอบค่าแสงสว่าง (Value) ในแต่ละพิกเซลเพื่อสังเกตรูปแบบของภาพจริงและภาพที่ผ่านการดัดแปลงโดยเพิ่มเครื่องประดับ AR

4 การตรวจจับใบหน้า (Face Recognition)

- มีการใช้ Face Detection ในขั้นตอนการพัฒนาโมเดล

- ใช้ขั้นตอน **Pre-processing** เตรียมภาพเพื่อนำไปใช้ในการเทรนโมเดล

10

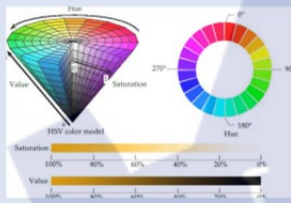
02 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ทฤษฎีที่เกี่ยวข้อง

12345678910

5 HSV Color Model

แปลงภาพเป็น HSV เพื่อวัดค่าแสงสว่าง (Value) และนำไปคำนวณค่าอื่นๆ ที่ใช้ในการตรวจจับภาพ



6 การวัดการกระจายตัวข้อมูล

คำนวณการกระจายตัวข้อมูลเพื่อสังเกตลักษณะการกระจายตัวข้อมูลของภาพแต่ละประเภท

- Mean Absolute Deviation (Mean AD)
- Median Absolute Deviation (Median AD)
- Standard Deviation (SD)

11

02 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ทฤษฎีที่เกี่ยวข้อง

12345678910

7 Euclidean Distance

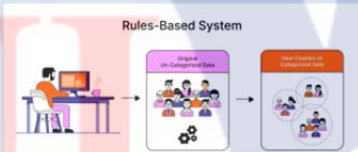
การวัดระยะทางระหว่างจุดสองจุดเป็นเส้นตรง มีรากฐานมาจากทฤษฎีบทพีทาโกรัส สามารถคำนวณระยะทาง 2 มิติ, 3 มิติ หรือ n มิติ ได้

ใช้ในการคำนวณความต่างของพิกเซลสีใน HSV Model เพื่อสร้าง mask สำหรับกำหนด Region of Interest (ROI)

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$$

8 Rule-based System

กำหนดเงื่อนไขในการตรวจจับภาพ และกำหนดค่า Threshold เพื่อตีความว่าเป็นภาพประเภทใด



12

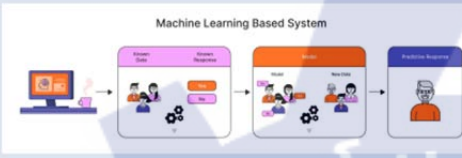
02 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ทฤษฎีที่เกี่ยวข้อง

1
2
3
4
5
6
7
8
9
10

9
Machine Learning

มีการใช้ Machine Learning ในการพัฒนาโมเดลสำหรับตรวจจับภาพ



10
Confusion Matrix

ใช้ Confusion Matrix แสดงค่าการทำนายกับค่าจริงที่เกิดขึ้น และใช้เป็นส่วนประกอบในการประเมินผลประสิทธิภาพโมเดล

	ผลลัพธ์เป็นจริง	ผลลัพธ์ไม่เป็นจริง
ทำนายว่าเป็นจริง	True Positive (TP)	False Positive (FP)
ทำนายว่าไม่เป็นจริง	False Negative (FN)	True Negative (TN)

13

02 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง งานวิจัยที่เกี่ยวข้อง		
ชื่องานวิจัย	วัตถุประสงค์ของงานวิจัย	จุดที่เหมือนกับการงานวิจัยนี้
1 A New Rule-Based Classification Method Using Shape-Based Properties of Spectral Curves [1]	พัฒนาระบบ Rule-based Classification โดยใช้คุณสมบัติรูปทรงของเส้นสเปกตรัม	การใช้ Rule-based ในการพัฒนาระบบสำหรับการทำ Image Classification
2 Advanced Image Forensics: Detecting and reconstructing Manipulated Images with Deep Learning [2]	พัฒนาโมเดลที่สามารถแยกแยะภาพจริงและภาพที่ถูกปรับแต่งด้วยความแม่นยำสูงได้	ใช้กระบวนการวิเคราะห์เพื่อหาความผิดปกติภายในภาพ และมีการใช้ AI ในการสร้างโมเดลจำแนกภาพ
3 Holistic Image Manipulation Detection using Pixel Cooccurrence Matrices [3]	พัฒนาวิธีการตรวจจับการปลอมแปลงภาพโดยใช้เทคนิคการวิเคราะห์เหตุการณ์การเกิดร่วมของ Pixel ร่วมกับ Deep Learning	ใช้กระบวนการวิเคราะห์เพื่อหาความผิดปกติภายในภาพ และมีการใช้ AI ในการสร้างโมเดลจำแนกภาพ
4 FastFlow: Unsupervised Anomaly Detection and Localization via 2D Normalizing Flows [4]	พัฒนาวิธีที่สามารถตรวจจับและระบุตำแหน่งของความผิดปกติในภาพ โดยไม่ใช้ข้อมูลที่มีการระบุความผิดปกติล่วงหน้า	ใช้กระบวนการวิเคราะห์เพื่อหารูปแบบของภาพเพื่อกำหนดมาตรฐานในการตรวจสอบภาพผิดปกติ

02 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

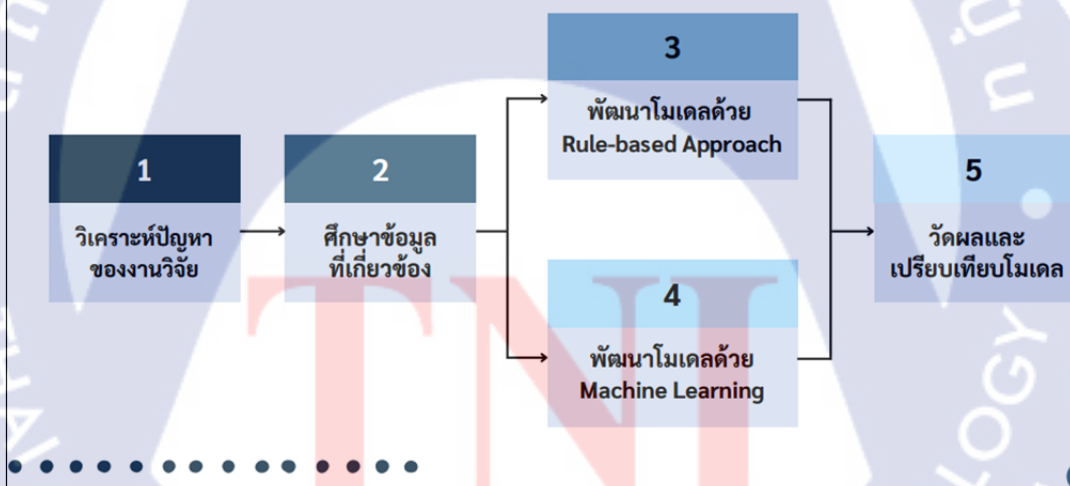
เอกสารอ้างอิงของงานวิจัยที่เกี่ยวข้อง

- 1 [1] S. Polat et al., "A New Rule-Based Classification Method Using Shape-Based Properties of Spectral Curves", Journal of Spectroscopy, vol. 2022, ISSN 2314-4920 (Print), 2314-4939 (Online), 2020.
- 2 [2] D. Nimma et al., "Advanced Image Forensics: Detecting and reconstructing Manipulated Images with Deep Learning", International Journal of Intelligent Systems and Applications in Engineering (IJISAE), vol. 12, no. 4, 2024.
- 3 [3] L. Nataraj et al., "Holistic Image Manipulation Detection using Pixel Cooccurrence Matrices", IS&T International Symposium on Electronic Imaging (MWSF – Media Watermarking, Security, and Forensics 2021), vol. 33, no. 4, pp. 277-1 – 277-7, 2021
- 4 [4] J. Yu et al., "FastFlow: Unsupervised Anomaly Detection and Localization via 2D Normalizing Flows", arXiv preprint arXiv:2111.07677, 2021. [Online]. Available: <https://arxiv.org/abs/2111.07677>

15

03 วิธีดำเนินงานวิจัย

ภาพรวมขั้นตอนงานวิจัย



16

03 วิธีดำเนินงานวิจัย

- 1 วิเคราะห์ปัญหาของงานวิจัย
- 2 ศึกษาข้อมูลที่เกี่ยวข้อง
- 3 พัฒนาโมเดลด้วย Rule-based Approach
- 4 พัฒนาโมเดลด้วย Machine Learning
- 5 วัดผลและเปรียบเทียบโมเดล

วิเคราะห์ปัญหาของงานวิจัย

- ผู้วิจัยศึกษาและวิเคราะห์ปัญหาเพื่อทำความเข้าใจประเด็นหลักของงาน พบว่า **ยังไม่มีเครื่องมือสำหรับตรวจจับภาพ AR โดยเฉพาะ** ที่สามารถใช้งานได้ทั่วไป และมีความแม่นยำ
- ผู้วิจัยจึงสนใจและต้องการทดลองพัฒนาเครื่องมือสำหรับตรวจจับภาพ AR โดยกำหนดเงื่อนไขเบื้องต้นคือ ภาพใบหน้ามนุษย์ และเครื่องประดับบนใบหน้า
- รวมถึงการทดสอบตรวจจับภาพด้วยโมเดล Machine Learning ที่มีอยู่แล้ว เพื่อทดสอบประสิทธิภาพของโมเดลที่มีอยู่แล้ว



17

03 วิธีดำเนินงานวิจัย

- 1 วิเคราะห์ปัญหาของงานวิจัย
- 2 ศึกษาข้อมูลที่เกี่ยวข้อง
- 3 พัฒนาโมเดลด้วย Rule-based Approach
- 4 พัฒนาโมเดลด้วย Machine Learning
- 5 วัดผลและเปรียบเทียบโมเดล

ศึกษาข้อมูลที่เกี่ยวข้อง

การจัดเตรียม Dataset

Dataset ประกอบไปด้วยภาพใบหน้าทั้งหมด 2,000 ภาพ



18

03 วิธีดำเนินงานวิจัย

- 1 วิเคราะห์ปัญหาของงานวิจัย
- 2 **ศึกษาข้อมูลที่เกี่ยวข้อง**
- 3 พัฒนาโมเดลด้วย Rule-based Approach
- 4 พัฒนาโมเดลด้วย Machine Learning
- 5 วัดผลและเปรียบเทียบโมเดล

ศึกษาข้อมูลที่เกี่ยวข้อง

การจัดเตรียม Dataset

รายละเอียดการแบ่งบทบาทข้อมูลภาพใน Dataset

บทบาทข้อมูล	จำนวนภาพแว่นตา (ภาพจริง / ภาพ AR)	จำนวนภาพหน้ากาก (ภาพจริง / ภาพ AR)	รวมทั้งหมด
Training	350 / 350	350 / 350	1,400
Validating	50 / 50	50 / 50	200
Testing	100 / 100	100 / 100	400
รวมทั้งหมด	1,000	1,000	2,000

* ข้อมูลส่วน Training : โมเดล Rule-based ใช้เป็นกลุ่มตัวอย่าง 120 ภาพ
โมเดล Machine Learning ใช้ทั้งหมด 1,400 ภาพ
* ข้อมูลส่วน Testing จะใช้เหมือนกัน 400 ภาพ ทั้งโมเดล Rule-based และโมเดล Machine Learning

19

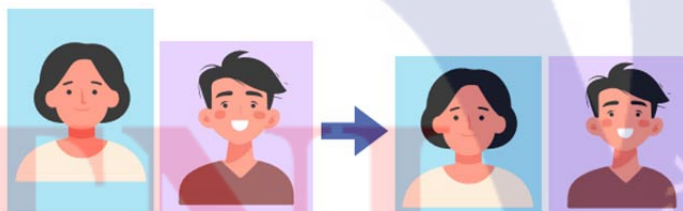
03 วิธีดำเนินงานวิจัย

- 1 วิเคราะห์ปัญหาของงานวิจัย
- 2 **ศึกษาข้อมูลที่เกี่ยวข้อง**
- 3 พัฒนาโมเดลด้วย Rule-based Approach
- 4 พัฒนาโมเดลด้วย Machine Learning
- 5 วัดผลและเปรียบเทียบโมเดล

ศึกษาข้อมูลที่เกี่ยวข้อง

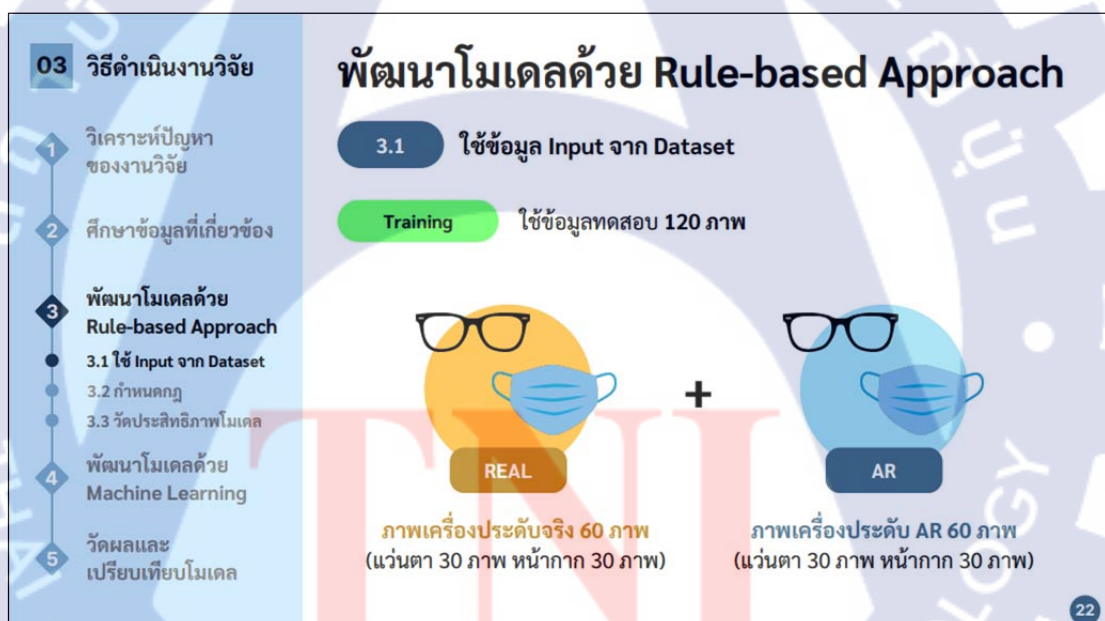
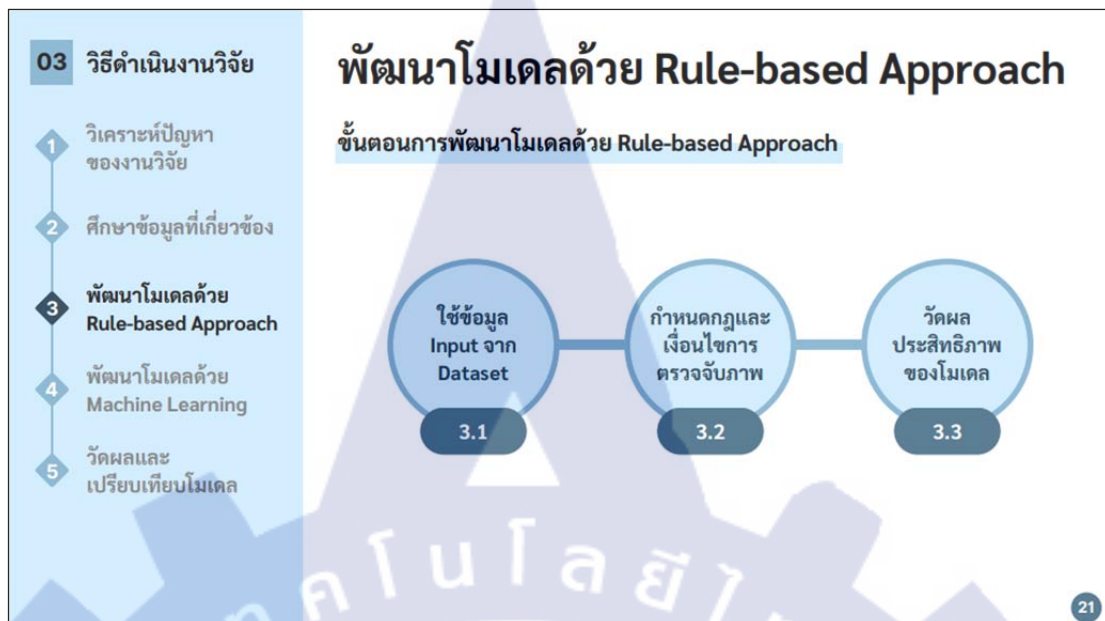
วิธีการตรวจจับตำแหน่งใบหน้าในภาพ

ใช้โมเดล Multi-task Cascaded Convolutional Neural Network (MTCNN)
ตรวจจับตำแหน่งใบหน้าในภาพ และ Crop ภาพใน Dataset ให้มีขนาดเท่ากัน



- ใช้ Multi-task Cascaded Convolutional Neural Network (MTCNN) ในการตรวจจับใบหน้า

20



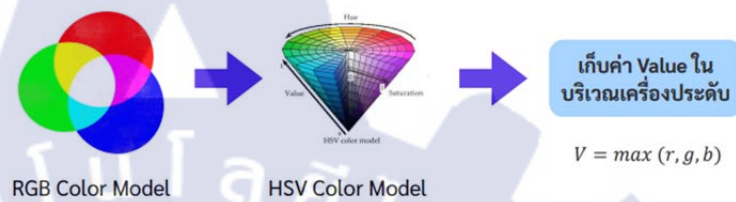
03 วิธีดำเนินงานวิจัย

- 1 วิเคราะห์ปัญหาของงานวิจัย
- 2 ศึกษาข้อมูลที่เกี่ยวข้อง
- 3 พัฒนาโมเดลด้วย Rule-based Approach
 - 3.1 ใช้ Input จาก Dataset
 - 3.2 กำหนดกฎ
 - 3.3 วัดประสิทธิภาพโมเดล
- 4 พัฒนาโมเดลด้วย Machine Learning
- 5 วัดผลและเปรียบเทียบโมเดล

พัฒนาโมเดลด้วย Rule-based Approach

3.2 กำหนดกฎและเงื่อนไขการตรวจจับภาพ

- แปลงภาพถ่ายที่เป็น RGB Color Model ให้เป็นภาพ HSV Color Model และทำการเก็บค่าแสงสว่างของสี (Value)
- เก็บค่าแสงสว่างในส่วนที่เป็นวัตถุเครื่องประดับ และนำค่าแสงสว่างนั้นมาคำนวณค่าที่สามารถใช้แยกแยะระหว่างเครื่องประดับจริงและ AR ได้



23

03 วิธีดำเนินงานวิจัย

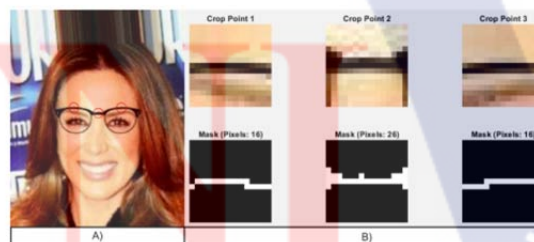
- 1 วิเคราะห์ปัญหาของงานวิจัย
- 2 ศึกษาข้อมูลที่เกี่ยวข้อง
- 3 พัฒนาโมเดลด้วย Rule-based Approach
 - 3.1 ใช้ Input จาก Dataset
 - 3.2 กำหนดกฎ
 - 3.3 วัดประสิทธิภาพโมเดล
- 4 พัฒนาโมเดลด้วย Machine Learning
- 5 วัดผลและเปรียบเทียบโมเดล

พัฒนาโมเดลด้วย Rule-based Approach

3.2 กำหนดกฎและเงื่อนไขการตรวจจับภาพ

ภาพแว่นตา

- สร้างกรอบ 15 * 15 pixel จากการเลือกจุดตรงกลางกรอบ ทั้งหมด 3 จุดในภาพ
- ในแต่ละกรอบ ใช้ Euclidean Distance เทียบระหว่างค่าสีของแต่ละ pixel กับสีของ pixel จุดกลาง
- สร้าง mask (ส่วนสีขาว) เพื่อกำหนดว่าในกรอบนั้นจะเก็บค่า Value ใน pixel ไດบ้าง
- เก็บค่า value จากแต่ละ pixel ใน mask



24

03 วิธีดำเนินงานวิจัย

- 1 วิเคราะห์ปัญหาของงานวิจัย
- 2 ศึกษาข้อมูลที่เกี่ยวข้อง
- 3 **พัฒนาโมเดลด้วย Rule-based Approach**
 - 3.1 ใช้ Input จาก Dataset
 - 3.2 **กำหนดกฎ**
 - 3.3 วัดประสิทธิภาพโมเดล
- 4 พัฒนาโมเดลด้วย Machine Learning
- 5 วัดผลและเปรียบเทียบโมเดล

พัฒนาโมเดลด้วย Rule-based Approach

3.2

กำหนดกฎและเงื่อนไขการตรวจจับภาพ

ภาพหน้ากาก

1. สร้างกรอบ 15 * 15 pixel จากการเลือกจุดตรงกลางกรอบ ทั้งหมด 3 จุดในภาพ
2. ในกรณีของภาพหน้ากาก กรอบ 15*15 จะมีเฉพาะข้อมูลส่วนที่เป็นหน้ากากทั้งหมด จึง**ไม่มีการใช้ Euclidean Distance** เทียบค่าสีเหมือนภาพแว่นตา
3. เก็บค่า value จากแต่ละ pixel ในกรอบ

เลือก 3 จุดในภาพ → ตีกรอบ 15*15 pixel → เก็บค่า Value ภายในกรอบภาพ

25

03 วิธีดำเนินงานวิจัย

- 1 วิเคราะห์ปัญหาของงานวิจัย
- 2 ศึกษาข้อมูลที่เกี่ยวข้อง
- 3 **พัฒนาโมเดลด้วย Rule-based Approach**
 - 3.1 ใช้ Input จาก Dataset
 - 3.2 **กำหนดกฎ**
 - 3.3 วัดประสิทธิภาพโมเดล
- 4 พัฒนาโมเดลด้วย Machine Learning
- 5 วัดผลและเปรียบเทียบโมเดล

พัฒนาโมเดลด้วย Rule-based Approach

3.2

กำหนดกฎและเงื่อนไขการตรวจจับภาพ

1

Mean Box Intensity

- หาค่ารวมแสงสว่างรายกรอบ ROI
- นำค่าแสงสว่างมาเฉลี่ย (หารด้วย 3)
- ได้ค่าเฉลี่ยของค่ารวมแสงสว่างของทั้ง 3 กรอบภาพ

Sum + Sum + Sum → หารค่าเฉลี่ย

สมมติฐาน

- ▶ **ภาพแว่นตา :** ภาพแว่นตาจริงมีความสว่าง**มากกว่า**ภาพแว่นตา AR (ภาพจริงอาจมีแสงสะท้อนหรือแสงสว่าง ภายในบริเวณวัตถุมากกว่าภาพ AR)
- ▶ **ภาพหน้ากาก :** ภาพหน้ากากจริงมีความสว่าง**น้อยกว่า**ภาพหน้ากาก AR (ภาพจริงอาจมีส่วนที่เป็นเงา หรือรอยยับของหน้ากากมากกว่าภาพ AR)

26

03 วิธีดำเนินงานวิจัย

- 1 วิเคราะห์ปัญหาของงานวิจัย
- 2 ศึกษาข้อมูลที่เกี่ยวข้อง
- 3 **พัฒนาโมเดลด้วย Rule-based Approach**
 - 3.1 ใช้ Input จาก Dataset
 - 3.2 **กำหนดกฎ**
 - 3.3 วัดประสิทธิภาพโมเดล
- 4 พัฒนาโมเดลด้วย Machine Learning
- 5 วัดผลและเปรียบเทียบโมเดล

พัฒนาโมเดลด้วย Rule-based Approach

3.2 กำหนดกฎและเงื่อนไขการตรวจจับภาพ

2 Mean Intensity

- นำค่าความสว่างของ ROI มารวมกันทั้งหมด
- หารด้วยจำนวนพิกเซลของทั้ง 3 กรอบ เพื่อหาค่าเฉลี่ยแสงสว่างของพิกเซลทั้งหมด

สมมติฐาน

- ภาพแว่นตา :** ภาพแว่นตาจริงมีความสว่าง**มากกว่า**ภาพแว่นตา AR (ภาพจริงอาจมีแสงสะท้อนหรือแสงสว่าง ภายในบริเวณวัตถุมากกว่าภาพ AR)
- ภาพหน้ากาก :** ภาพหน้ากากจริงมีความสว่าง**น้อยกว่า**ภาพหน้ากาก AR (ภาพจริงอาจมีส่วนที่เป็นเงา หรือรอยยับของหน้ากากมากกว่าภาพ AR)

Sum

หาค่าเฉลี่ย

27

03 วิธีดำเนินงานวิจัย

- 1 วิเคราะห์ปัญหาของงานวิจัย
- 2 ศึกษาข้อมูลที่เกี่ยวข้อง
- 3 **พัฒนาโมเดลด้วย Rule-based Approach**
 - 3.1 ใช้ Input จาก Dataset
 - 3.2 **กำหนดกฎ**
 - 3.3 วัดประสิทธิภาพโมเดล
- 4 พัฒนาโมเดลด้วย Machine Learning
- 5 วัดผลและเปรียบเทียบโมเดล

พัฒนาโมเดลด้วย Rule-based Approach

3.2 กำหนดกฎและเงื่อนไขการตรวจจับภาพ

3 Mean Absolute Deviation (Mean AD)
คำนวณส่วนเบี่ยงเบนเฉลี่ยสัมบูรณ์

4 Median Absolute Deviation (Median AD)
คำนวณส่วนเบี่ยงเบนมัธยฐานสัมบูรณ์

5 Standard Deviation (SD)
คำนวณส่วนเบี่ยงเบนมาตรฐาน

สมมติฐาน

- ค่า Deviation ต่ำ :** ค่าความสว่างในแต่ละพิกเซลมีความ**ใกล้เคียงกัน** มีโอกาสเป็นภาพ **AR**
- ค่า Deviation สูง :** ค่าความสว่างในแต่ละพิกเซลมีความ**ห่างกัน** มีโอกาสเป็น**วัตถุจริง**

$$Mean AD = \frac{\sum |x_i - \bar{x}|}{n}$$

$$Median AD = median(|x_i - \bar{x}|)$$

$$S = \sqrt{\frac{\sum (X - \bar{x})^2}{n-1}}$$

28

03 วิธีดำเนินงานวิจัย

- 1 วิเคราะห์ปัญหาของงานวิจัย
- 2 ศึกษาข้อมูลที่เกี่ยวข้อง
- 3 **พัฒนาโมเดลด้วย Rule-based Approach**
 - 3.1 ใช้ Input จาก Dataset
 - 3.2 กำหนดกฎ
 - 3.3 วัดประสิทธิภาพโมเดล
- 4 พัฒนาโมเดลด้วย Machine Learning
- 5 วัดผลและเปรียบเทียบโมเดล

พัฒนาโมเดลด้วย Rule-based Approach

3.2 กำหนดกฎและเงื่อนไขการตรวจจับภาพ

เงื่อนไขการแยกแยะภาพเครื่องประดับจริงและเครื่องประดับ AR

ค่าที่ใช้ในการกำหนดเกณฑ์	สัญลักษณ์เงื่อนไขแว่นตา	สัญลักษณ์เงื่อนไขหน้ากาก	เงื่อนไขของภาพ AR (แว่นตา)	เงื่อนไขของภาพ AR (หน้ากาก)
Mean Box Intensity	T_1	T'_1	$V \leq T_1$	$V > T'_1$
Mean Intensity	T_2	T'_2	$V \leq T_2$	$V > T'_2$
Mean AD	T_3	T'_3	$V \leq T_3$	$V \leq T'_3$
Median AD	T_4	T'_4	$V \leq T_4$	$V \leq T'_4$
Standard Deviation	T_5	T'_5	$V \leq T_5$	$V \leq T'_5$

V = Value (ค่าแสงสว่าง)

29

03 วิธีดำเนินงานวิจัย

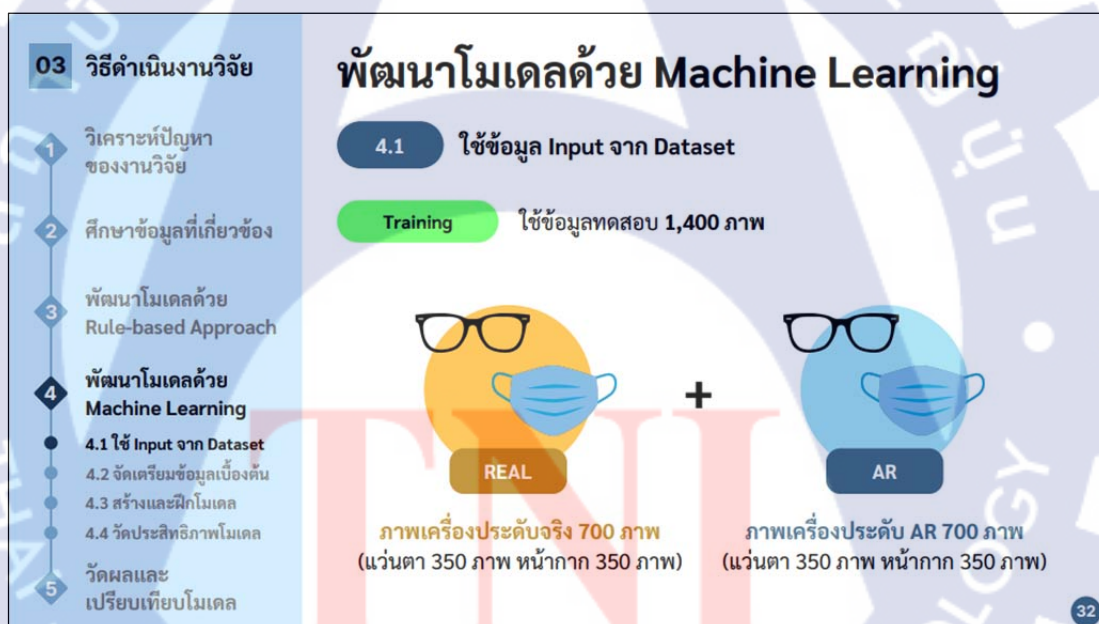
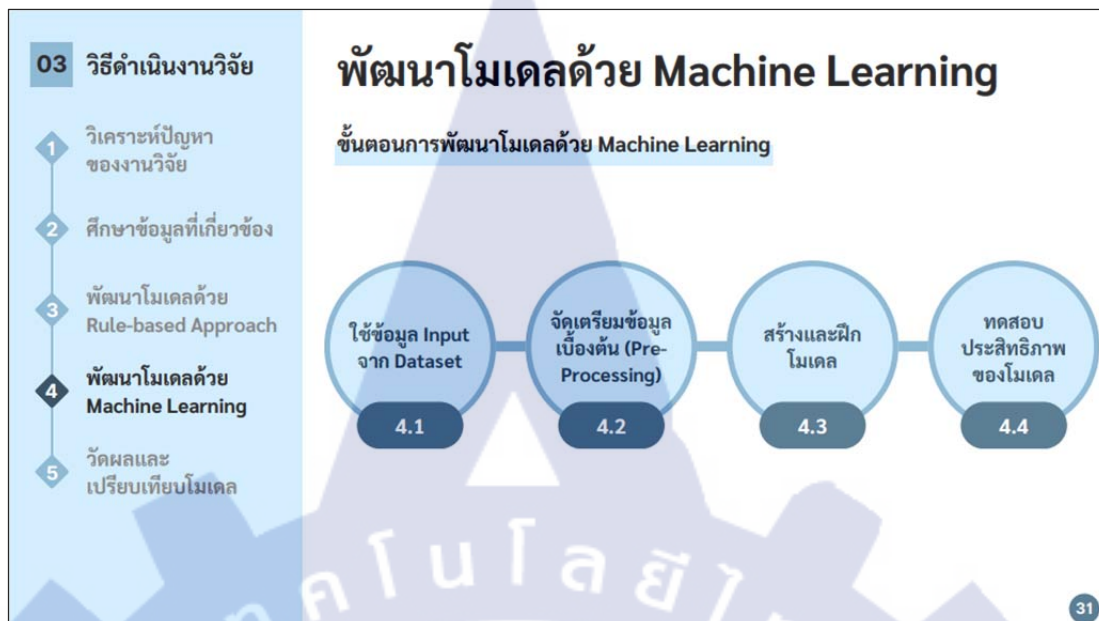
- 1 วิเคราะห์ปัญหาของงานวิจัย
- 2 ศึกษาข้อมูลที่เกี่ยวข้อง
- 3 **พัฒนาโมเดลด้วย Rule-based Approach**
 - 3.1 ใช้ Input จาก Dataset
 - 3.2 กำหนดกฎ
 - 3.3 วัดประสิทธิภาพโมเดล
- 4 พัฒนาโมเดลด้วย Machine Learning
- 5 วัดผลและเปรียบเทียบโมเดล

พัฒนาโมเดลด้วย Rule-based Approach

3.3 วัดผลประสิทธิภาพของโมเดล

- วัดผลโมเดลแยกกันระหว่างภาพแว่นตาและภาพหน้ากากด้วยข้อมูลชุด Testing
- ใช้ Confusion Matrix แสดงค่าการทำนายกับค่าจริงที่เกิดขึ้น และประเมินผลประสิทธิภาพโมเดล
- จัดเก็บผลลัพธ์สำหรับเปรียบเทียบในขั้นตอน **วัดผลและเปรียบเทียบโมเดล**

30



03 วิธีดำเนินงานวิจัย

- 1 วิเคราะห์ปัญหาของงานวิจัย
- 2 ศึกษาข้อมูลที่เกี่ยวข้อง
- 3 พัฒนาโมเดลด้วย Rule-based Approach
- 4 พัฒนาโมเดลด้วย Machine Learning**
 - 4.1 ใช้ Input จาก Dataset
 - 4.2 จัดเตรียมข้อมูลเบื้องต้น**
 - 4.3 สร้างและฝึกโมเดล
 - 4.4 วัดประสิทธิภาพโมเดล
- 5 วัดผลและเปรียบเทียบโมเดล

พัฒนาโมเดลด้วย Machine Learning

4.2

จัดเตรียมข้อมูลเบื้องต้น (Pre-Processing)

- ใช้โมเดล MTCNN **ตรวจจับตำแหน่งใบหน้าในภาพ** และ Crop ภาพใน Dataset ให้มีขนาดเท่ากัน
- แบ่งส่วนภาพใน Dataset ให้พร้อมสำหรับแต่ละขั้นตอนการพัฒนาโมเดล

Training	Validating	Testing
1,400 ภาพ	200 ภาพ	400 ภาพ

↑
 ใช้ข้อมูลทดสอบส่วนเดียวกันกับโมเดล Rule-based

33

03 วิธีดำเนินงานวิจัย

- 1 วิเคราะห์ปัญหาของงานวิจัย
- 2 ศึกษาข้อมูลที่เกี่ยวข้อง
- 3 พัฒนาโมเดลด้วย Rule-based Approach
- 4 พัฒนาโมเดลด้วย Machine Learning**
 - 4.1 ใช้ Input จาก Dataset
 - 4.2 จัดเตรียมข้อมูลเบื้องต้น
 - 4.3 สร้างและฝึกโมเดล**
 - 4.4 วัดประสิทธิภาพโมเดล
- 5 วัดผลและเปรียบเทียบโมเดล

พัฒนาโมเดลด้วย Machine Learning

4.3

สร้างและฝึกโมเดล

เลือกโมเดลที่ใช้ในการฝึก

Xception	สามารถเรียนรู้รายละเอียดเชิงโครงสร้างของภาพได้แม่นยำและมีประสิทธิภาพ ✓
MobileNetV3	มีความสามารถในการใช้งานบนอุปกรณ์ที่มีทรัพยากรน้อย หรือมีข้อจำกัดได้ ✓
EfficientNetV2S	มีความสมดุลในแง่จำนวนพารามิเตอร์ที่ใช้งาน และความแม่นยำในการตรวจจับ ✓

34

03 วิธีดำเนินงานวิจัย

- 1 วิเคราะห์ปัญหาของงานวิจัย
- 2 ศึกษาข้อมูลที่เกี่ยวข้อง
- 3 พัฒนาโมเดลด้วย Rule-based Approach
- 4 **พัฒนาโมเดลด้วย Machine Learning**
 - 4.1 ใช้ Input จาก Dataset
 - 4.2 จัดเตรียมข้อมูลเบื้องต้น
 - 4.3 **สร้างและฝึกโมเดล**
 - 4.4 วัดประสิทธิภาพโมเดล
- 5 วัดผลและเปรียบเทียบโมเดล

พัฒนาโมเดลด้วย Machine Learning

4.3 สร้างและฝึกโมเดล

การฝึกโมเดล

```

graph LR
    START([START]) --> IT[Initial Training]
    IT --> FT[Fine-tuning]
    FT --> END([END])
          
```

ทำ Data Augmentation

- เช่น หมุนภาพ กลับด้านภาพ ซูมภาพ
- จำนวนภาพที่ใช้เทรนจริงอยู่ที่ประมาณ 4,900 ภาพ

ฝึกโมเดล

- กำหนดไว้ที่ 30 Epochs
- ตั้งค่า Early Stopping เพื่อป้องกันการ Overfit ของข้อมูล
- บันทึกโมเดลที่ฝึกสำเร็จ

35

03 วิธีดำเนินงานวิจัย

- 1 วิเคราะห์ปัญหาของงานวิจัย
- 2 ศึกษาข้อมูลที่เกี่ยวข้อง
- 3 พัฒนาโมเดลด้วย Rule-based Approach
- 4 **พัฒนาโมเดลด้วย Machine Learning**
 - 4.1 ใช้ Input จาก Dataset
 - 4.2 จัดเตรียมข้อมูลเบื้องต้น
 - 4.3 **สร้างและฝึกโมเดล**
 - 4.4 วัดประสิทธิภาพโมเดล
- 5 วัดผลและเปรียบเทียบโมเดล

พัฒนาโมเดลด้วย Machine Learning

4.3 สร้างและฝึกโมเดล

การฝึกโมเดล

```

graph LR
    START([START]) --> IT[Initial Training]
    IT --> FT[Fine-tuning]
    FT --> END([END])
          
```

เปิด Layer ของโมเดลเพิ่มเติม

- เปิด Layer ส่วนท้ายของโมเดล (ที่ทำการเรียนรู้เรื่องรูปร่างของสิ่งของ) เพื่อให้โมเดลเรียนรู้ตาม Dataset ที่กำหนดไว้

ปรับอัตราการเรียนรู้ของโมเดลให้ลดลง

- ป้องกันไม่ให้ความรู้เดิมที่โมเดลฝึกและเรียนรู้ไปแล้ว หายไป
- ฝึกเพิ่มอีก 20 Epochs

36

03 วิธีดำเนินงานวิจัย

- 1 วิเคราะห์ปัญหาของงานวิจัย
- 2 ศึกษาข้อมูลที่เกี่ยวข้อง
- 3 พัฒนาโมเดลด้วย Rule-based Approach
- 4 พัฒนาโมเดลด้วย Machine Learning
 - 4.1 ใช้ Input จาก Dataset
 - 4.2 จัดเตรียมข้อมูลเบื้องต้น
 - 4.3 สร้างและฝึกโมเดล
 - 4.4 วัดประสิทธิภาพโมเดล
- 5 วัดผลและเปรียบเทียบโมเดล

พัฒนาโมเดลด้วย Machine Learning

4.4 วัดผลประสิทธิภาพของโมเดล

- วัดผลโมเดลแยกกันระหว่างภาพแว่นตาและภาพหน้ากากด้วยข้อมูลชุด Testing
- ใช้ Confusion Matrix แสดงค่าการทำนายกับค่าจริงที่เกิดขึ้น และประเมินผลประสิทธิภาพโมเดล
- จัดเก็บผลลัพธ์สำหรับเปรียบเทียบในขั้นตอน **วัดผลและเปรียบเทียบโมเดล**



37

03 วิธีดำเนินงานวิจัย

- 1 วิเคราะห์ปัญหาของงานวิจัย
- 2 ศึกษาข้อมูลที่เกี่ยวข้อง
- 3 พัฒนาโมเดลด้วย Rule-based Approach
- 4 พัฒนาโมเดลด้วย Machine Learning
- 5 วัดผลและเปรียบเทียบโมเดล

วัดผลและเปรียบเทียบโมเดล

เปรียบเทียบค่าผลลัพธ์ที่ได้จาก Confusion Matrix

- นำค่าที่ได้จากการวัดผลประสิทธิภาพของแต่ละโมเดลมาคำนวณค่าตัวชี้วัด

1 ค่าความถูกต้อง (Accuracy)

$$\frac{TP + TN}{TP + TN + FP + FN}$$

2 ค่าความแม่นยำ (Precision)

$$\frac{TP}{TP + FP}$$

3 ค่าความอ่อนไหว (Recall)

$$\frac{TP}{TP + FN}$$

4 ค่าวัดประสิทธิภาพโดยรวม (F1-Score)

$$2 * \frac{Precision * Recall}{Precision + Recall}$$

38

04 ผลการวิจัย

ผลการวิจัย

- 1 ผลการหากฎเกณฑ์และเงื่อนไขสำหรับโมเดล Rule-based
- 2 ผลการตรวจจับภาพด้วยโมเดล Rule-based
- 3 ผลการตรวจจับภาพด้วยโมเดล Machine Learning
- 4 ผลการเปรียบเทียบประสิทธิภาพของโมเดล Rule-based และ Machine Learning

39

04 ผลการวิจัย

ผลการหากฎเกณฑ์และเงื่อนไขสำหรับโมเดล Rule-based

วิธีการหาค่าคุณลักษณะแต่ละค่า

- 1 ผลการหากฎเกณฑ์และเงื่อนไขสำหรับโมเดล Rule-based
- 2 ผลการตรวจจับภาพด้วยโมเดล Rule-based
- 3 ผลการตรวจจับภาพด้วยโมเดล Machine Learning
- 4 ผลการเปรียบเทียบประสิทธิภาพของโมเดล Rule-based และ Machine Learning

- 1 Mean Box Intensity
- 2 Mean Intensity

- สังเกตรูปแบบข้อมูลระหว่างภาพ AR และภาพจริง
- กำหนดตัวเลข Threshold ที่ทำให้ผลการแยกแยะออกมาดีที่สุด

- 3 Mean Absolute Deviation (Mean AD)
- 4 Median Absolute Deviation (Median AD)
- 5 Standard Deviation (SD)

- หาค่าเฉลี่ยของค่า Deviation รวมทุกภาพ
- นำค่าเฉลี่ยของภาพจริงและภาพ AR มาหาค่าเฉลี่ยอีกครั้งเพื่อให้ได้ค่า Threshold

ลำดับภาพ	ค่า Mean AD	ค่า Median AD	ค่า SD
ภาพที่ 1	Mean AD 1	Median AD 1	SD 1
ภาพที่ 2	Mean AD 2	Median AD 2	SD 2
...	...	...	...
ภาพที่ 30	Mean AD 30	Median AD 30	SD 30

ค่าเฉลี่ย Mean AD ภาพจริง	ค่าเฉลี่ย Mean AD ภาพ AR
ค่าเฉลี่ย	ค่าเฉลี่ย

ลำดับภาพ	ค่า Mean AD	ค่า Median AD	ค่า SD
ภาพที่ 1	Mean AD 1	Median AD 1	SD 1
ภาพที่ 2	Mean AD 2	Median AD 2	SD 2
...	...	...	...
ภาพที่ 30	Mean AD 30	Median AD 30	SD 30

40

04 ผลการวิจัย

1 ผลการหากฎเกณฑ์และเงื่อนไขสำหรับโมเดล Rule-based

- 1.1 เงื่อนไขภาพสวมแว่นตา
1.2 เงื่อนไขภาพสวมหน้ากาก

2 ผลการตรวจจับภาพด้วยโมเดล Rule-based

3 ผลการตรวจจับภาพด้วยโมเดล Machine Learning

4 ผลการเปรียบเทียบประสิทธิภาพของโมเดล Rule-based และ Machine Learning

ผลการหากฎเกณฑ์และเงื่อนไขสำหรับโมเดล Rule-based



ภาพแว่นตา

เงื่อนไขการแยกแยะภาพสวมแว่นตา

คุณลักษณะ	สัญลักษณ์	ค่าในการแยกแยะ	ภาพจริง	ภาพ AR
ค่าเฉลี่ยของความสว่างของกรอบบริเวณที่สนใจ	T_1	5.000	1.4 - 95.2	0.4 - 7.7
ค่าเฉลี่ยแสงสว่างของปีกเซลล์ทั้งหมด	T_2	0.200	0.15 - 0.90	0.03 - 0.41
ค่าส่วนเบี่ยงเบนเฉลี่ยสัมบูรณ์ (Mean AD)	T_3	0.070	0.02 - 0.20	0.01 - 0.22
ค่าส่วนเบี่ยงเบนมัธยฐานสัมบูรณ์ (Median AD)	T_4	0.051	0.02 - 0.13	0.007 - 0.176
ค่าส่วนเบี่ยงเบนมาตรฐาน (SD)	T_5	0.084	0.03 - 0.24	0.01 - 0.24

ผลลัพธ์การแยกแยะภาพสวมแว่นตาดูด้วยเงื่อนไขข้างต้น กับภาพชุดทดสอบ

ประเภทเครื่องประดับ	จำนวนผลการแยกแยะที่ถูกต้อง (ภาพ)	คิดเป็นอัตราส่วน (%)
ภาพแว่นตาจริง	19 จาก 30	63
ภาพแว่นตา AR	25 จาก 30	83

41

04 ผลการวิจัย

1 ผลการหากฎเกณฑ์และเงื่อนไขสำหรับโมเดล Rule-based

- 1.1 เงื่อนไขภาพสวมแว่นตา
1.2 เงื่อนไขภาพสวมหน้ากาก

2 ผลการตรวจจับภาพด้วยโมเดล Rule-based

3 ผลการตรวจจับภาพด้วยโมเดล Machine Learning

4 ผลการเปรียบเทียบประสิทธิภาพของโมเดล Rule-based และ Machine Learning

ผลการหากฎเกณฑ์และเงื่อนไขสำหรับโมเดล Rule-based



ภาพหน้ากาก

เงื่อนไขการแยกแยะภาพสวมหน้ากาก

คุณลักษณะ	สัญลักษณ์	ค่าในการแยกแยะ	ภาพจริง	ภาพ AR
ค่าเฉลี่ยของความสว่างของกรอบบริเวณที่สนใจ	T'_1	190.000	30.6 - 224.8	187.7 - 222.3
ค่าเฉลี่ยแสงสว่างของปีกเซลล์ทั้งหมด	T'_2	0.850	0.13 - 0.99	0.83 - 0.98
ค่าส่วนเบี่ยงเบนเฉลี่ยสัมบูรณ์ (Mean AD)	T'_3	0.070	0.001 - 0.17	0.009 - 0.052
ค่าส่วนเบี่ยงเบนมัธยฐานสัมบูรณ์ (Median AD)	T'_4	0.030	0.01 - 0.18	0.007 - 0.027
ค่าส่วนเบี่ยงเบนมาตรฐาน (SD)	T'_5	0.056	0.003 - 0.18	0.01 - 0.06

ผลลัพธ์การแยกแยะภาพหน้ากากด้วยเงื่อนไขข้างต้น กับภาพชุดทดสอบ

ประเภทเครื่องประดับ	จำนวนผลการแยกแยะที่ถูกต้อง (ภาพ)	คิดเป็นอัตราส่วน (%)
ภาพหน้ากากจริง	17 จาก 30	56
ภาพหน้ากาก AR	26 จาก 30	87

42

04 ผลการวิจัย

1 ผลการหาคุณลักษณะ
เงื่อนไขสำหรับโมเดล
Rule-based

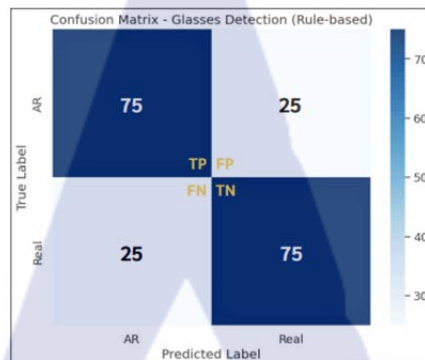
2 ผลการตรวจจับภาพ
ด้วยโมเดล Rule-based

- 2.1 ตรวจจับภาพสวมแว่นตา
- 2.2 ตรวจจับภาพสวมหน้ากาก

3 ผลการตรวจจับภาพ
ด้วยโมเดล Machine
Learning

4 ผลการเปรียบเทียบ
ประสิทธิภาพของโมเดล
Rule-based และ
Machine Learning

ผลการตรวจจับภาพด้วยโมเดล Rule-based



ภาพแว่นตา

ค่าในการวัดผล	ผลลัพธ์
Accuracy	0.75
Precision	0.75
Recall	0.75
F1-Score	0.74

▶ **ตรวจจับภาพ AR :** โมเดลสามารถตรวจจับภาพ AR ได้ถูกต้อง 75 จาก 100 ภาพ (75%)

▶ **ตรวจจับภาพจริง :** โมเดลสามารถตรวจจับภาพจริงได้ถูกต้อง 75 จาก 100 ภาพ (75%)

43

04 ผลการวิจัย

1 ผลการหาคุณลักษณะ
เงื่อนไขสำหรับโมเดล
Rule-based

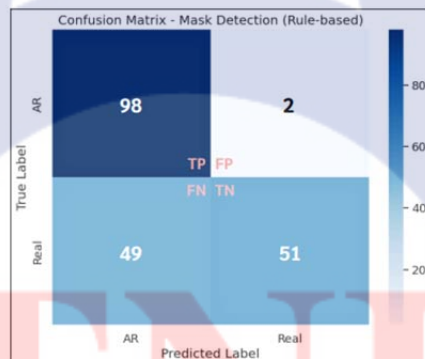
2 ผลการตรวจจับภาพ
ด้วยโมเดล Rule-based

- 2.1 ตรวจจับภาพสวมแว่นตา
- 2.2 ตรวจจับภาพสวมหน้ากาก

3 ผลการตรวจจับภาพ
ด้วยโมเดล Machine
Learning

4 ผลการเปรียบเทียบ
ประสิทธิภาพของโมเดล
Rule-based และ
Machine Learning

ผลการตรวจจับภาพด้วยโมเดล Rule-based



ภาพหน้ากาก

ค่าในการวัดผล	ผลลัพธ์
Accuracy	0.75
Precision	0.67
Recall	0.98
F1-Score	0.80

▶ **ตรวจจับภาพ AR :** โมเดลสามารถตรวจจับภาพ AR ได้ถูกต้อง 98 จาก 100 ภาพ (98%)

▶ **ตรวจจับภาพจริง :** โมเดลสามารถตรวจจับภาพจริงได้ถูกต้อง 51 จาก 100 ภาพ (51%)

44

04 ผลการวิจัย

1 ผลการหาจุดเด่นและ
เงื่อนไขสำหรับโมเดล
Rule-based

2 ผลการตรวจจับภาพ
ด้วยโมเดล Rule-based

3 ผลการตรวจจับภาพ
ด้วยโมเดล Machine
Learning

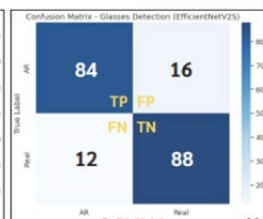
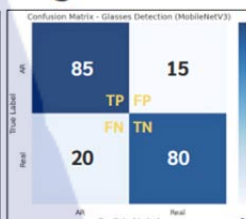
3.1 ตรวจจับภาพสวมแว่นตา

3.2 ตรวจจับภาพสวมหน้ากาก

4 ผลการเปรียบเทียบ
ประสิทธิภาพของโมเดล
Rule-based และ
Machine Learning

ผลการตรวจจับภาพด้วยโมเดล Machine Learning

ภาพแว่นตา



Xception	
ค่าในการวัดผล	ผลลัพธ์
Accuracy	0.88
Precision	0.83
Recall	0.95
F1-Score	0.88

MobileNetV3	
ค่าในการวัดผล	ผลลัพธ์
Accuracy	0.82
Precision	0.80
Recall	0.85
F1-Score	0.83

EfficientNetV2S	
ค่าในการวัดผล	ผลลัพธ์
Accuracy	0.92
Precision	0.87
Recall	0.84
F1-Score	0.86

45

04 ผลการวิจัย

1 ผลการหาจุดเด่นและ
เงื่อนไขสำหรับโมเดล
Rule-based

2 ผลการตรวจจับภาพ
ด้วยโมเดล Rule-based

3 ผลการตรวจจับภาพ
ด้วยโมเดล Machine
Learning

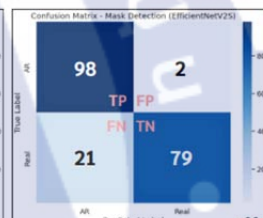
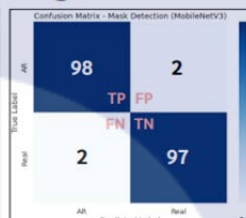
3.1 ตรวจจับภาพสวมแว่นตา

3.2 ตรวจจับภาพสวมหน้ากาก

4 ผลการเปรียบเทียบ
ประสิทธิภาพของโมเดล
Rule-based และ
Machine Learning

ผลการตรวจจับภาพด้วยโมเดล Machine Learning

ภาพหน้ากาก



Xception	
ค่าในการวัดผล	ผลลัพธ์
Accuracy	0.98
Precision	1.00
Recall	0.96
F1-Score	0.98

MobileNetV3	
ค่าในการวัดผล	ผลลัพธ์
Accuracy	0.97
Precision	0.97
Recall	0.98
F1-Score	0.97

EfficientNetV2S	
ค่าในการวัดผล	ผลลัพธ์
Accuracy	0.88
Precision	0.82
Recall	0.98
F1-Score	0.89

46

04 ผลการวิจัย

1 ผลการหากฎเกณฑ์และเงื่อนไขสำหรับโมเดล Rule-based

2 ผลการตรวจจับภาพด้วยโมเดล Rule-based

3 ผลการตรวจจับภาพด้วยโมเดล Machine Learning

4 ผลการเปรียบเทียบประสิทธิภาพของโมเดล Rule-based และ Machine Learning

4.1 ภาพส่วนแว่นตา

4.2 ภาพส่วนหน้ากาก

ผลการเปรียบเทียบประสิทธิภาพของโมเดล Rule-based และ Machine Learning

ภาพส่วนตา

ค่าในการวัดผล	Rule-based Model	Machine Learning (Deep Learning Model)		
		Xception	MobileNetV3	EfficientNetV2S
Accuracy	0.75	0.88	0.82	0.92
Precision	0.75	0.83	0.80	0.875
Recall	0.75	0.95	0.85	0.84
F1-Score	0.74	0.88	0.83	0.86

เรียงลำดับตามค่า Accuracy

1. EfficientNetV2S
2. Xception
3. MobileNetV3
4. Rule-based

เรียงลำดับตามค่าอื่นๆ

- EfficientNetV2S มีค่า Precision สูงที่สุด
- Xception มีค่า Recall สูงที่สุด
- Xception มีค่า F1-Score สูงที่สุด

47

04 ผลการวิจัย

1 ผลการหากฎเกณฑ์และเงื่อนไขสำหรับโมเดล Rule-based

2 ผลการตรวจจับภาพด้วยโมเดล Rule-based

3 ผลการตรวจจับภาพด้วยโมเดล Machine Learning

4 ผลการเปรียบเทียบประสิทธิภาพของโมเดล Rule-based และ Machine Learning

4.1 ภาพส่วนแว่นตา

4.2 ภาพส่วนหน้ากาก

ผลการเปรียบเทียบประสิทธิภาพของโมเดล Rule-based และ Machine Learning

ภาพหน้ากาก

ค่าในการวัดผล	Rule-based Model	Machine Learning (Deep Learning Model)		
		Xception	MobileNetV3	EfficientNetV2S
Accuracy	0.75	0.98	0.97	0.88
Precision	0.67	1	0.97	0.82
Recall	0.98	0.96	0.98	0.98
F1-Score	0.8	0.98	0.97	0.89

เรียงลำดับตามค่า Accuracy

1. Xception
2. MobileNetV3
3. EfficientNetV2S
4. Rule-based

เรียงลำดับตามค่าอื่นๆ

- Xception มีค่า Precision สูงที่สุด
- Rule-based, MobileNetV3 และ EfficientNetV2 มีค่า Recall สูงที่สุดเท่ากัน
- Xception มีค่า F1-Score สูงที่สุด

48



05 บทสรุปและข้อเสนอแนะ

สรุปผลการวิจัย

1 สรุปผลการวิจัย

2 ปัญหาและข้อจำกัด

3 ข้อเสนอแนะ

ผลการศึกษากฎเกณฑ์ต่างๆ ที่เกี่ยวข้องกับการตรวจจับภาพ

- พบวิธีการวัดค่าแสงสว่างของพิกเซลจาก HSV Color Model
- การคำนวณค่าที่เกี่ยวข้องกับการสังเกตรูปแบบค่าแสงสว่าง

ผลการหากฎเกณฑ์และเงื่อนไขสำหรับโมเดล Rule-based

- กำหนดกฎเกณฑ์ที่ใช้ในการตรวจจับภาพได้
- ใช้กฎเกณฑ์แบบเดียวกัน แตกต่างกันในส่วนค่า Threshold

ผลการเปรียบเทียบผลลัพธ์การตรวจจับภาพของโมเดล

- โมเดล Machine Learning มีค่า Accuracy สูงกว่าโมเดล Rule-based

Rule-based Machine Learning

50

05 บทสรุปและข้อเสนอแนะ

- 1 สรุปผลการวิจัย
- 2 ปัญหาและข้อจำกัด
- 3 ข้อเสนอแนะ

ปัญหาและข้อจำกัด



ปัจจัยที่ใช้ในการแยกแยะภาพ

- โมเดล Rule-based ใช้ปัจจัยในการตรวจจับความแตกต่างน้อยกว่า Machine Learning ที่สามารถใช้หลายปัจจัยในการตรวจจับภาพ



จำนวนข้อมูลของโมเดล Rule-based

- โมเดล Rule-based ใช้จำนวนข้อมูลในการ Train น้อยกว่าโมเดล Machine Learning หลายเท่า

51

05 บทสรุปและข้อเสนอแนะ

- 1 สรุปผลการวิจัย
- 2 ปัญหาและข้อจำกัด
- 3 ข้อเสนอแนะ

ข้อเสนอแนะ

ด้านการค้นหาปัจจัยในการตรวจจับภาพ

- โมเดล Rule-based ควรพิจารณาข้อมูลภาพในมิติอื่นๆ เพิ่มเติม เช่น การวิเคราะห์ข้อมูลในมิติความถี่ (Frequency Domain) เพื่อพัฒนาความแม่นยำในการตรวจจับของโมเดล

ด้านการปรับปรุงข้อมูลที่ใช้ในการฝึกโมเดล

- การปรับปรุงจำนวนข้อมูล และคุณภาพข้อมูลใน Dataset อาจช่วยเพิ่มประสิทธิภาพในการตรวจจับของโมเดล



52



Q&A SESSION

**THANK YOU
FOR YOUR ATTENTION**

TNI