

การทดสอบสมรรถนะการตรวจจับของโมเดลการตรวจจับจุดสังเกตบนใบหน้า

The Ability Test of Model the Facial Landmark Detecting

เพชร อินทนนท์* และ สุรศักดิ์ มั่งสิงห์

Petchara Inthanon*, & Surasak Mungsing

หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาเทคโนโลยีสารสนเทศ มหาวิทยาลัยศรีปทุม

Master of Science (Information Technology), Sripatum University

Submitted 23/5/2020 ; Revised 15/6/2020 ; Accepted 4/7/2020

บทคัดย่อ

ปัจจุบันการตรวจจับใบหน้าที่มีเทคนิคต่าง ๆ มากมาย ไม่ว่าจะเป็นการตรวจจับใบหน้าทั้งใบหน้า หรือ การตรวจจับใบหน้าแบบแยกเป็นส่วน ๆ ด้วยเทคโนโลยีปัญญาประดิษฐ์ แต่งานวิจัยชิ้นนี้มีการใช้โมเดลตรวจจับจุดสังเกตบนใบหน้า ซึ่งเป็นการตรวจจับใบหน้าส่วนต่าง ๆ จำนวน 68 จุด โดยแต่ละจุดบนใบหน้าที่ตรวจจับได้สามารถนำไปใช้คำนวณหาค่าต่าง ๆ ได้ งานวิจัยชิ้นนี้นำภาพถ่ายลักษณะอาการต่าง ๆ บนใบหน้า ซึ่งเป็น dataset จำนวน 1,522 ภาพ มาใช้ในการทดสอบการตรวจจับใบหน้าที่มีความผิดพลาดน้อยเพียงใด พบว่าโมเดลดังกล่าวมีความแม่นยำในการตรวจจับ 94.81% และเมื่อนำโมเดลดังกล่าวมาตรวจจับจุดสังเกตบริเวณตาและปาก โดยการเปรียบเทียบประสิทธิภาพระหว่างอุปกรณ์ 2 อย่าง ได้แก่ Nvidia Jetson Nano และ Raspberry Pi 4 + Intel® Neural Compute Stick 2 โดยใช้กล้องเว็บแคมในการตรวจจับจุดสังเกตบริเวณตาและปาก โดยจำลองสถานการณ์ 4 รูปแบบ พบว่าการตรวจจับจุดสังเกตบนใบหน้าของอุปกรณ์ทั้ง 2 ชนิดมีความแม่นยำเฉลี่ย 78.08% และ 80.59% ตามลำดับ

คำสำคัญ : จุดสังเกตบนใบหน้า ตรวจจับใบหน้า สมรรถนะ โมเดล

***ผู้ประสานงานหลัก (Corresponding Author)**

E-mail: petch.innon@gmail.com

Abstract

Recently, there are several techniques for face detection such as detecting the whole face or detecting specific parts of face by using artificial intelligence. For this project, facial landmarks model was conducted to detect 68 points of different face parts. Each facial point could be calculated to generate values. Moreover, various face motion photos were applied into 1,522 datasets to examine the accuracy of the facial landmarks model. The result showed that the accuracy of using model was very high which was 94.81%. Besides, the model was conducted to detect unusual long closed-eyes and opened-mouth by comparing 2 tools between Nvidia Jetson Nano and Raspberry Pi 4 + Intel® Neural Compute Stick 2 with webcam camera. There were 4 circumstances to test the detection of facial points and the average accuracy of each tool was 78.08% and 80.59% respectively.

Keyword : Facial landmark, Face detection, Competency, model

1. บทนำ

ในปัจจุบันมีการนำเทคโนโลยีปัญญาประดิษฐ์มาใช้อย่างมากมายในหลาย ๆ ด้านเพื่อช่วยในเรื่องต่าง ๆ หนึ่งในนั้นคือ การนำมาใช้ตรวจจับใบหน้า ไม่ว่าจะเป็นการเรียนรู้อัตโนมัติหรือการตรวจจับอาการต่าง ๆ บนใบหน้า ซึ่งมีหลากหลายโมเดลที่มีการนำมาใช้งาน งานวิจัยชิ้นนี้ทดสอบการตรวจจับใบหน้าที่กับภาพ dataset จำนวน 1,522 ภาพ เพื่อตรวจสอบว่ามีความแม่นยำในการตรวจจับมากน้อยเพียงใด เมื่อได้ค่าความแม่นยำข้างต้นแล้วจะนำไปทดสอบกับการตรวจจับใบหน้าแบบ real time โดยใช้โมเดลที่มีชื่อว่า facial landmark [1-3] ในการทดสอบประสิทธิภาพการตรวจจับตาและปาก

ในปี 2561 ประเทศไทยมีสถิติการเกิดอุบัติเหตุเป็นอันดับที่ 9 ของโลก [4] มีคนไทยเสียชีวิตจากการเกิดอุบัติเหตุทางถนนปีละประมาณ 22,491 ราย คิดเป็น 32.7 คน ต่อประชากร 1 แสนคน เฉลี่ยแล้วคนไทยเสียชีวิตจากอุบัติเหตุ ชั่วโมงละ 3 คน และพบว่าการเกิดอุบัติเหตุกับรถโดยสารสาธารณะในปีงบประมาณ 2562 นั้น “อาการหลับใน” เป็นสาเหตุลำดับที่ 3 ของการเกิดอุบัติเหตุ รองจากการขับรดัดหน้าและขับรดตามหลัง ในระยะกระชั้นชิด ซึ่งเกิดขึ้นถึง 43 ครั้ง จากจำนวนอุบัติเหตุทั้งหมด 414 ครั้ง คิดเป็น 10.39 % [5] ซึ่งหากนำข้อมูลจากงานวิจัยชิ้นนี้ไปพัฒนาต่อยอด จะสามารถนำไปควบคุมหรือบังคับใช้เป็นกฎหมายที่สามารถลดอุบัติเหตุและความเสียหายที่เกิดจากอาการผิดปกติของผู้ขับขี่รถโดยสารสาธารณะได้อีกด้วย

งานวิจัยก่อนหน้านี้นำ facial landmark มาใช้ตรวจจับอาการง่วงนอนของผู้ขับขี่รถยนต์ โดยการตรวจจับส่วนตา เพื่อตรวจจับอาการหลับตานานเกิน 2 วินาที เมื่อตรวจพบก็จะมีการแจ้งเตือน ซึ่งหากพบว่าผู้ขับขี่มีอาการง่วงนอนจะมีการแจ้งเตือนผ่านหน้าจอและแจ้งเตือนไปยังอุปกรณ์สั่งการภายในรถยนต์ [6] ผู้วิจัยจึงนำสมการจากงานวิจัยชิ้นนี้มาพัฒนาให้มีความเหมาะสมกับการใช้งานในส่วนการตรวจจับการกระพริบตา นอกจากนี้ยังมีงานวิจัยที่นำ facial landmark มาใช้เพื่อตอบบนโมเดลใบหน้าในรูปแบบ 3 มิติ โดยใช้ Convolutional Neural Network (CNN) หรือ โครงข่ายประสาทแบบคอนโวลูชัน เป็นโครงข่ายประสาทเทียมหนึ่งในกลุ่ม Bio-Inspired ในการตรวจจับจุดสำคัญต่าง ๆ บนใบหน้า [7]

จากงานวิจัยข้างต้น มีการนำโมเดล facial landmark มาพัฒนาและประยุกต์ใช้ในด้านต่าง ๆ โดยประสิทธิภาพในการตรวจจับอาจขึ้นอยู่กับอุปกรณ์ที่ใช้ในการตรวจจับหน้า ซึ่งจะมีผลต่อการประมวลผลของ facial landmark ในงานวิจัยนี้ผู้วิจัยเลือกเปรียบเทียบโดยใช้อุปกรณ์ 2 อย่าง ได้แก่ Nvidia Jetson Nano และ Raspberry Pi 4 + Intel® Neural Compute Stick 2 เนื่องจากอุปกรณ์ทั้งสองอย่างนั้นมีราคาไม่สูง ทำให้บุคคลทั่วไปหรือนักศึกษาสามารถเข้าถึงและนำผลการทดลองในงานวิจัยชิ้นนี้ไปใช้ให้เหมาะสมกับงานและต่อยอดได้ โดยอุปกรณ์ 2 อย่างนี้ใช้เพื่อทดสอบว่าอุปกรณ์ใดมีประสิทธิภาพในการตรวจจับตาและปากโดยใช้โมเดล facial landmark มากที่สุด เพื่อนำไปสู่การพัฒนาต่อไปในอนาคต

2. วัตถุประสงค์

1. เพื่อทดสอบประสิทธิภาพ facial landmark ในการตรวจจับใบหน้าจาก dataset
2. เพื่อเปรียบเทียบประสิทธิภาพในการตรวจจับจุดบริเวณตาและปาก ระหว่าง Nvidia Jetson Nano และ Raspberry Pi 4 + Intel® Neural Compute Stick 2 โดยใช้กล้องเว็บแคม

3. วิธีดำเนินการวิจัย

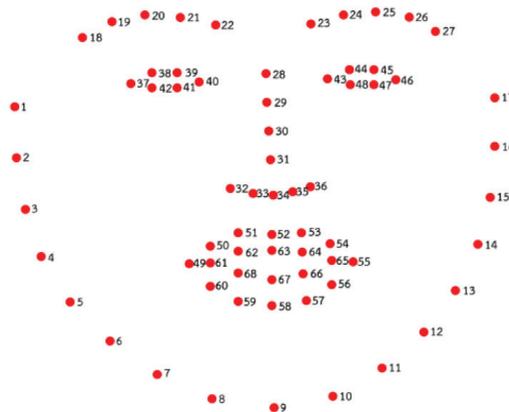
งานวิจัยชิ้นนี้ มีการทดสอบ 2 รูปแบบ คือ การนำ dataset จำนวน 1,522 ภาพ มาใช้ในการทดสอบว่าโมเดล facial landmark นั้นมีความแม่นยำมากน้อยเพียงใด และการนำโมเดล facial landmark มาประมวลผลโดยอุปกรณ์ 2 อย่าง ได้แก่ Nvidia Jetson Nano และ Raspberry Pi 4 + Intel® Neural Compute Stick 2 ซึ่งใช้วิธีตรวจจับใบหน้าแบบ real time โดยใช้กล้องเว็บแคมนำข้อมูลภาพใบหน้าเข้าสู่โปรแกรม

แล้วใช้โมเดล facial landmark ตรวจจับจุดสังเกตต่าง ๆ บนใบหน้า โดยจำลองสถานการณ์การตรวจจับใบหน้า 4 รูปแบบ ได้แก่ รูปแบบที่ 1 คือ ใบหน้าปกติ ไม่มีสิ่งใดปกปิด รูปแบบที่ 2 คือ ใบหน้าที่มีการสวมแว่นสายตา รูปแบบที่ 3 คือ ใบหน้าที่มีการสวมหมวก และรูปแบบที่ 4 คือ ใบหน้าที่มีการสวมทั้งแว่นสายตาและหมวก ดังภาพที่ 1



ภาพที่ 1 รูปแบบการตรวจจับใบหน้า

งานวิจัยนี้การทดสอบการใช้งานโมเดล facial landmark จะกำหนดพิกัดสำคัญบนใบหน้าที่แสดงในภาพที่ 2



ภาพที่ 2 พิกัดสำคัญบนใบหน้า จำนวน 68 จุด [8]

การทดสอบแบ่งออกเป็น 2 รูปแบบ ได้แก่

3.1 การทดสอบความแม่นยำของโมเดล facial landmark

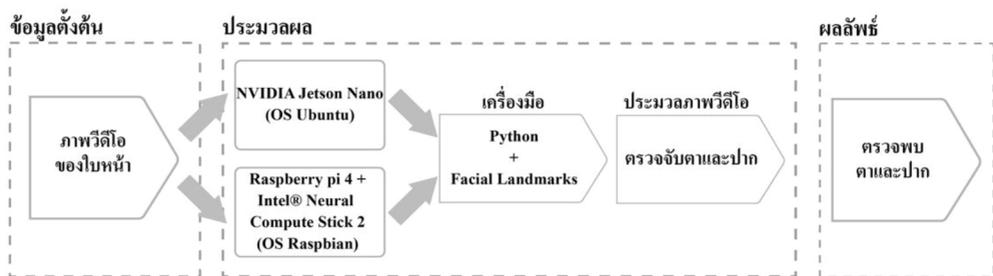
การทดสอบนี้จะทดสอบความแม่นยำของโมเดล facial landmark โดยใช้งานชุดข้อมูลภาพใบหน้า ที่ชื่อว่า Bio ID Dataset ซึ่งมีภาพใบหน้าของคน 21 คน จำนวนภาพใบหน้า 1,522 ภาพ ขนาดของภาพ 384 x 286 พิกเซล และเป็นภาพระดับสีเทา (grey level) ในการทดสอบความแม่นยำของการตรวจจับจุดต่าง ๆ บนใบหน้า จะแบ่งการทดสอบเป็น 3 รูปแบบ ได้แก่

- 1) ทดสอบความแม่นยำในการตรวจจับทุกจุดบนใบหน้า
- 2) ทดสอบความแม่นยำในการตรวจจับจุดบริเวณตา
- 3) ทดสอบความแม่นยำในการตรวจจับจุดบริเวณปาก

3.2 การทดสอบประสิทธิภาพ

การทดสอบประสิทธิภาพของการตรวจจับบริเวณตาและปากโดยใช้ฟังก์ชัน facial landmark ในการทดสอบประสิทธิภาพของอุปกรณ์ประมวลผล โดยเปรียบเทียบประสิทธิภาพของอุปกรณ์ในการตรวจจับใบหน้าระหว่าง Nvidia Jetson Nano และ Raspberry Pi 4 + Intel® Neural Compute Stick 2

กระบวนการตรวจจับใบหน้าประกอบด้วยขั้นตอนต่าง ๆ ได้แก่ ส่วนข้อมูลตั้งต้น (input) มีการรับข้อมูลภาพวิดีโอของใบหน้าตามเวลาจริง (real time) ส่วนประมวลผล (process) จะใช้งานอุปกรณ์ 2 อย่าง ได้แก่ 1) Nvidia Jetson Nano ซึ่งใช้ระบบปฏิบัติการ Ubuntu และ Raspberry Pi 4 และ 2) Intel® Neural Compute Stick 2 ซึ่งใช้ระบบปฏิบัติการ Raspbian โดยทั้งสองอุปกรณ์จะใช้เครื่องมือ (tool) เป็นภาษา Python เพื่อเรียกใช้โมเดล facial landmark สำหรับตรวจจับบริเวณตาและปาก และส่วนผลลัพธ์ (output) ที่จะแสดงผลการตรวจจับบริเวณตาและปาก โดยในภาพจะแสดงเส้นรอบ ๆ ตาและปากที่จับภาพได้ โดยกระบวนการตรวจจับภาพแสดงดังภาพที่ 3



ภาพที่ 3 กระบวนการตรวจจับภาพ

อุปกรณ์ที่ใช้ในการวิจัยนี้มีคุณสมบัติต่าง ๆ ดังตารางที่ 1

ตารางที่ 1 คุณสมบัติของอุปกรณ์ที่ใช้ในการทดลอง

| ฮาร์ดแวร์ | Raspberry Pi 4 + Intel® Neural Compute Stick 2 | NVIDIA Jetson Nano |
|-----------|---|---|
| CPU | Quad-core ARM Cortex-A72 64-bit @ 1.5 Ghz | Quad-Core ARM Cortex-A57 64-bit @ 1.42 Ghz |
| GPU | Broadcom VideoCore VI (32-bit) | NVIDIA Maxwell w/ 128 CUDA cores @ 921 Mhz |
| Memory | 4 GB LPDDR4** | 4 GB LPDDR4 |
| VPU | 700 MHz | - |

จากตารางที่ 1 อุปกรณ์ที่ใช้ในการทดลองอุปกรณ์ทั้ง 2 อย่าง มีคุณสมบัติที่แตกต่างกันไม่มาก ยกเว้น GPU ที่มีการใช้งาน GPU ต่างยี่ห้อกัน ซึ่งอาจทำให้การประมวลผลภาพมีประสิทธิภาพไม่เท่ากัน และในส่วนของ Raspberry Pi ยังมี NCS2 ซึ่งเป็นระบบ Vision Processing Unit (VPU) สำหรับงานประมวลผลปัญญาประดิษฐ์โดยเฉพาะด้วย ทำให้ผู้วิจัยเลือกมาใช้ทดสอบการประมวลผลภาพที่เป็นปัญญาประดิษฐ์ในงานวิจัยชิ้นนี้

การตรวจจับส่วนต่างๆ ของใบหน้าใน 4 รูปแบบ (ภาพที่ 1) จะทดสอบความแม่นยำของโมเดลทั้ง 3 รูปแบบ โดยใช้โมเดล facial landmark โดยนำข้อมูลจากผลการทดสอบมาใช้ในการคำนวณประสิทธิภาพในการตรวจจับบริเวณตาและบริเวณปาก จากสูตรดังนี้

$$\text{ร้อยละความถูกต้องแม่นยำ} = \frac{A}{B} \times 100 \quad (1)$$

$$\text{จำนวนเฟรมทั้งหมดจากการทดสอบ} = C \times D \quad (2)$$

$$\text{อัตราความถูกต้อง (\%)} = \frac{E}{F} \times 100 \quad (3)$$

โดยที่

- A คือ จำนวนภาพที่มีความถูกต้องแม่นยำ (ภาพ)
- B คือ จำนวนภาพทั้งหมด (1522 ภาพ)
- C คือ ระยะเวลาวิดีโอ (วินาที)
- D คือ จำนวนเฟรมใน 1 วินาที (25 FPS)
- E คือ จำนวนเฟรมที่ตรวจจับได้ (เฟรม)
- F คือ จำนวนเฟรมทั้งหมดจากการทดสอบ (เฟรม)

4. ผลการวิจัย

4.1 การทดสอบความแม่นยำของโมเดล facial landmark

ความแม่นยำของการตรวจสอบใบหน้าของโมเดล facial landmark ระหว่าง Nvidia Jetson Nano และ Raspberry Pi 4 + Intel® Neural Compute Stick 2 โดยการทดสอบความแม่นยำของการตรวจสอบทั้ง 3 รูปแบบ มีผลการทดสอบดังตารางที่ 2

ตารางที่ 2 ความแม่นยำของการตรวจสอบใบหน้าของโมเดล facial landmark

| รายการทดสอบ | จำนวนภาพที่มีความถูกต้องแม่นยำ (ภาพ) | ร้อยละ |
|--------------------------|--------------------------------------|--------|
| การตรวจจับทุกจุดบนใบหน้า | 1,443 | 94.81 |
| การตรวจจับจุดบริเวณดวงตา | 1,501 | 98.02 |
| การตรวจจับจุดบริเวณปาก | 1,465 | 96.25 |

จากตารางที่ 2 พบว่า ความแม่นยำในการตรวจจับจุดบริเวณดวงตาส่งผลดีที่สุด รองลงมาคือ การตรวจจับจุดบริเวณปาก และการตรวจจับทุกจุดบนใบหน้า ตามลำดับ

4.2 การทดสอบประสิทธิภาพ

การทดสอบประสิทธิภาพของอุปกรณ์ 2 อย่าง ในการตรวจจับจุดบริเวณตาและบริเวณปาก ตามรูปแบบที่ 1 – 4 ได้ผลการทดสอบดังตารางที่ 3 และตารางที่ 4

ตารางที่ 3 ความแม่นยำของการตรวจจับจุดบริเวณตาและปากของอุปกรณ์ Nvidia Jetson Nano

| รูปแบบ | ระยะเวลาวิดีโอ (วินาที) | เฟรมทั้งหมด (เฟรม) | เฟรมที่ตรวจจับได้ (เฟรม) | อัตราความ ถูกต้อง (%) |
|-------------|----------------------------|-----------------------|-----------------------------|--------------------------|
| รูปแบบที่ 1 | 59.01 | 1475.25 | 1229 | 83.31 |
| รูปแบบที่ 2 | 58.82 | 1470.50 | 1184 | 80.52 |
| รูปแบบที่ 3 | 59.71 | 1492.75 | 1096 | 73.42 |
| รูปแบบที่ 4 | 57.33 | 1433.25 | 1076 | 75.07 |

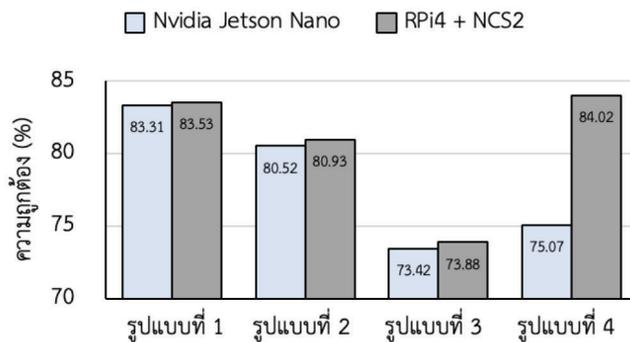
จากตารางที่ 3 พบว่า รูปแบบที่ 1 มีอัตราความถูกต้องในการตรวจจับจุดบริเวณตาและปากมากที่สุด รองลงมาคือ รูปแบบที่ 2 รูปแบบที่ 4 และรูปแบบที่ 3 ตามลำดับ

ตารางที่ 4 ความแม่นยำของการตรวจจับจุดบริเวณตาและปากของอุปกรณ์ Raspberry Pi 4 + Intel® Neural Compute Stick 2

| รูปแบบ | ระยะเวลาวิดีโอ (วินาที) | เฟรมทั้งหมด (เฟรม) | เฟรมที่ตรวจจับได้ (เฟรม) | อัตราความ ถูกต้อง (%) |
|-------------|----------------------------|-----------------------|-----------------------------|--------------------------|
| รูปแบบที่ 1 | 60.58 | 1514.50 | 1265 | 83.53 % |
| รูปแบบที่ 2 | 61.19 | 1529.75 | 1184 | 80.93 % |
| รูปแบบที่ 3 | 60.54 | 1513.50 | 1096 | 73.88 % |
| รูปแบบที่ 4 | 57.89 | 1447.25 | 1216 | 84.02 % |

จากตารางที่ 4 อัตราความถูกต้องในการตรวจจับจุดบริเวณตาและปากของรูปแบบที่ 4 มีค่ามากที่สุด รองลงมาคือ รูปแบบที่ 1 รูปแบบที่ 2 และรูปแบบที่ 3 ตามลำดับ

ผลการทดสอบจากตารางที่ 3 และ 4 ในการตรวจจับจุดบริเวณตาและปากตามรูปแบบที่ 1 – 4 เป็นการแสดงผลของความแม่นยำของการตรวจจับจุดบริเวณตาและปากของอุปกรณ์ 2 อย่าง ทดสอบโดยใช้กล้องเว็บแคม ความเร็ว 25 เฟรมต่อวินาที ตรวจจับส่วนต่าง ๆ ของใบหน้าใน 4 รูปแบบที่ต่างกัน ดังภาพที่ 2 ซึ่งจากการทดสอบ พบว่า การตรวจจับจุดบริเวณตาและปากโดยใช้ Raspberry Pi 4 + Intel® Neural Compute Stick 2 มีความแม่นยำมากกว่าในทุก ๆ รูปแบบ ดังภาพที่ 4



ภาพที่ 4 เปรียบเทียบร้อยละความแม่นยำของการตรวจจับจุดบริเวณตาและปากด้วยอุปกรณ์ 2 อย่าง

5. อภิปรายผลการวิจัย

จากการทดสอบความแม่นยำของโมเดล facial landmark และการทดสอบประสิทธิภาพของอุปกรณ์สรุปผลการทดสอบได้ดังนี้

5.1 การทดสอบความแม่นยำของโมเดล facial landmark

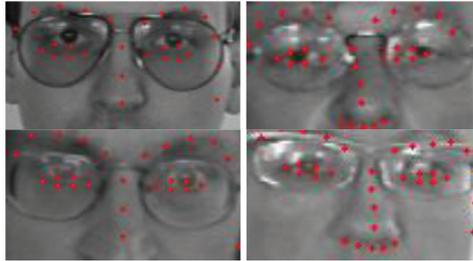
ในงานวิจัยนี้ เมื่อนำ dataset ภาพใบหน้าคน จำนวน 1,522 ภาพ มาทดสอบด้วยโมเดล facial landmark มีอัตราความแม่นยำในการตรวจจับใบหน้า 94.81% เมื่อเปรียบเทียบกับงานวิจัยของ Khunthi et al. (2019) พบว่าการทดสอบด้วยโมเดล facial landmark มีความแม่นยำสูงกว่าวิธีการ Haar-Cascade หรือการตรวจจับและรู้จำวัตถุต่าง ๆ จากลักษณะเฉพาะที่มีอยู่ในรูปนั้น ๆ [9] เช่น รู้จำว่าลักษณะแบบไหนคือใบหน้า ลักษณะไหนคือดวงตา ซึ่งมีความแม่นยำในการตรวจจับใบหน้า 93.79% และความแม่นยำสูงกว่าวิธีการ Faced ซึ่งมี 2 ขั้นตอน คือ การนำ Fully Convolutional Neural Network (FCNN) ในการพัฒนาบน YOLO และการใช้มาตรฐานของ CNN (Convolutions + Fully Connected Layer) โดยใช้กรอบสี่เหลี่ยมในการจับหน้าและทำนายตำแหน่งที่จะตรวจจับ ซึ่งมีความแม่นยำในการตรวจจับใบหน้า 88.70% นั้น โดยทั้ง 2 วิธีเป็นการตรวจจับใบหน้าเช่นเดียวกับ facial landmark ทำให้โมเดล facial landmark สามารถนำไปประยุกต์ใช้งานได้จริง และมีความผิดพลาดในระดับที่ยอมรับได้ จึงนำมาสู่การทดสอบการตรวจจับจุดบริเวณตาและปาก โดยมีผลการทดสอบดังนี้

1) ความแม่นยำของการตรวจจับทุกจุดบนใบหน้า มีความแม่นยำ 94.81% โดยมีตัวอย่างภาพที่ผิดพลาด เนื่องจากในบางมุมของภาพใบหน้าทีนำมาใช้ในการทดสอบนั้นมีสีที่กลมกลืนกัน ทำให้โมเดลไม่สามารถแยกแยะได้ว่าส่วนใดคือส่วนของใบหน้า จึงทำให้การตรวจจับใบหน้ามีความผิดพลาด ดังภาพที่ 5



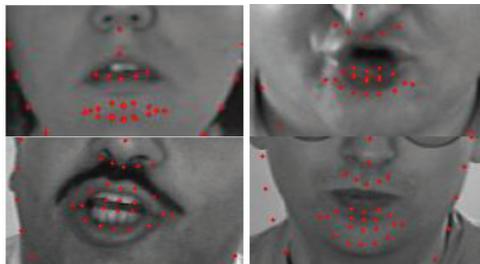
ภาพที่ 5 ภาพตัวอย่างการตรวจจับจุดบนใบหน้าที่มีความผิดพลาด

2) ความแม่นยำของการตรวจจับจุดบริเวณตา มีความแม่นยำ 98.02% โดยพบว่าตัวอย่างภาพที่มีความผิดพลาดของการตรวจจับจุดบริเวณตาทุกภาพนั้นเป็นภาพที่มีการสวมแว่นสายตา ซึ่งมีการหักเหของแสงหรือมีแสงสะท้อน ทำให้การตรวจจับจุดบริเวณตาที่มีความผิดพลาด ดังภาพที่ 6



ภาพที่ 6 ภาพตัวอย่างการตรวจจับจุดบริเวณตาที่มีความผิดพลาด

3) ความแม่นยำของการตรวจจับจุดบริเวณปาก มีความแม่นยำ 96.25% โดยพบว่าตัวอย่างภาพที่มีความผิดพลาดในการตรวจจับจุดบริเวณปากนั้นเป็นภาพที่มีรูปปากในลักษณะที่ผิดปกติ โดยเฉพาะการห่อปากซึ่งโมเดลไม่สามารถระบุได้ว่าส่วนใดคือริมฝีปาก จึงทำให้การตรวจจับจุดบริเวณปากมีความผิดพลาด ดังภาพที่ 7



ภาพที่ 7 ภาพตัวอย่างการตรวจจับจุดบริเวณปากที่มีความผิดพลาด

5.2 การทดสอบประสิทธิภาพของอุปกรณ์

จากผลการทดสอบจะเห็นว่า เมื่อเปรียบเทียบความแม่นยำของการตรวจจับจุดบริเวณตาและปาก ระหว่าง 2 อุปกรณ์ พบว่า รูปแบบที่ 4 ซึ่งเป็นรูปแบบที่มีการสวมแว่นตาและสวมหมวกนั้นมีความแม่นยำในการตรวจจับบริเวณตาและปากต่างกันมากที่สุด โดย Raspberry Pi 4 + Intel® Neural Compute Stick 2 นั้น มีความแม่นยำในการตรวจจับสูงกว่า Nvidia Jetson Nano ถึง 8.95% เนื่องจาก Raspberry Pi 4 + Intel® Neural Compute Stick 2 นั้น มี VPU ซึ่งผลิตมาเพื่อการประมวลผลด้านปัญญาประดิษฐ์โดยเฉพาะ ทำให้สามารถแยกแยะส่วนโครงสร้างใบหน้าได้แม่นยำกว่า Nvidia Jetson Nano ที่ไม่มี VPU และในการตรวจจับจุดบริเวณตา ไม่ได้มีการระบุว่ามีจุดบริเวณตาและปากที่ตรงตามที่เราสวมแว่นตาแล้วจะสามารถตรวจจับจุดบริเวณตาได้ดีขึ้น ซึ่งอาจจะผิดเพี้ยนไม่ตรงบ้าง ประกอบกับการสวมหมวกเป็นการตัดแสงตกกระทบจากมุมอื่น ๆ ทำให้การตรวจจับมีความแม่นยำสูงที่สุดนั่นเอง โดยผลการทดสอบที่ได้นี้ ผู้ทำการวิจัยได้ทำการทดสอบซ้ำ ๆ หลายครั้ง และพบว่าผลการทดลองมีความแม่นยำเช่นเดิม

6. สรุปผลการวิจัย

ในการตรวจจับใบหน้าจาก dataset โดยใช้ Bio ID Dataset ของคน 21 คน และมีจำนวนภาพใบหน้า 1,522 ภาพ จากการทดสอบประสิทธิภาพของโมเดล facial landmark พบว่า การตรวจจับจุดสังเกตบนใบหน้า มีความแม่นยำ 94.81% เมื่อตรวจจับบริเวณส่วนตา มีความแม่นยำถึง 98.02% และเมื่อตรวจจับบริเวณ ส่วนปาก มีความแม่นยำถึง 96.25% ซึ่งถือว่ามีความแม่นยำสูง ทำให้สามารถนำไปพัฒนาต่อยอดในการตรวจจับ อากัปกิริยาที่ผิดปกติต่อยอดได้ เช่น การตรวจจับอาการง่วง โดยวิธีการตรวจการหลับตานานผิดปกติหรือ เรียกว่า อาการหลับใน หรืออาการอ้าปากนานกว่าปกติหรือเรียกว่าอาการหาวได้

จากการเปรียบเทียบประสิทธิภาพในการตรวจจับจุดบริเวณตาและปาก ระหว่าง Nvidia Jetson Nano และ Raspberry Pi 4 + Intel® Neural Compute Stick 2 โดยใช้กล้องเว็บแคม ทดสอบรูปแบบการตรวจจับ จุดบริเวณตาและปากในรูปแบบต่าง ๆ 4 รูปแบบ พบว่า Raspberry Pi 4 + Intel® Neural Compute Stick 2 มีประสิทธิภาพในการตรวจจับบริเวณตาและปากมากกว่า Nvidia Jetson Nano

ผู้วิจัยได้ปฏิบัติงานภายในกรมการขนส่งทางบก จึงได้คิดค้นงานวิจัยชิ้นนี้ขึ้นเพื่อเป็นต้นแบบในการพัฒนา และต่อยอดในการตรวจจับอาการผิดปกติต่าง ๆ ของผู้ขับขี่รถโดยสารสาธารณะซึ่งมีผลกระทบต่อผู้โดยสาร จำนวนมาก โดยสามารถพัฒนาให้ตรวจจับอาการหลับตานานผิดปกติ (อาการหลับใน) หรืออาการอ้าปากนาน ผิดปกติ (อาการหาว) ซึ่งสามารถนำผลการวิจัยนี้ไปใช้เป็นต้นแบบในการพัฒนาและควบคุมพฤติกรรมของผู้ขับขี่ รถโดยสารสาธารณะได้

7. เอกสารอ้างอิง

- [1] Soukupová, T., & Cech, J. (2016). *Eye blink detection using facial landmarks*. (Research Reports). Slovenia: Czech Technical University in Prague.
- [2] Taufique, A. M. N., Savakis, A., & Leckenby, J. (2019). Automatic quantification of facial asymmetry using facial landmarks (pp. 1-5). In *2019 IEEE Western New York Image and Signal Processing Workshop (WNYISPW)*. Rochester, New York.
- [3] Wei, W., Tian, C., & Zhang, Y. (2015). A two-stage facial landmark localization method (pp. 157-160). In *2015 International Conference on Orange Technologies (ICOT)*. Hong Kong, China.
- [4] พรพิมล ดินะน้อย และ ณิชมน ทองพัฒน์. (2562). มุ่งเป้าปลอดภัย ‘จยย.’ ลดตายติดอันดับโลก. [ออนไลน์], สืบค้นจาก <https://tdri.or.th/2019/02/motorbikes-key-to-solving-road-deaths/> (30 มีนาคม 2563).
- [5] กลุ่มสถิติการขนส่ง กองแผนงาน กรมการขนส่งทางบก. (2562). *รายงานสรุปผลอุบัติเหตุรถโดยสารสาธารณะ ประจำปีงบประมาณ พ.ศ. 2562*. [ออนไลน์], สืบค้นจาก <https://web.dlt.go.th/statistics/> (25 มีนาคม 2563).
- [6] Wong, J. Y., & Lau, P. Y. (2019). Real-time driver alert system using raspberry pi. *ECTI Transactions on Electrical Engineering, Electronics, and Communications*, 17(2), 193-203.
- [7] Terada, T., Chen, Y. W., & Kimura, R. (2018). 3D facial landmark detection using deep convolutional neural networks (pp. 390-393). In *14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*. Huangshan, China.
- [8] Huber, P. (2017). *Real-time 3D morphable shape model fitting to monocular in-the-wild videos* (PhD Thesis). UK: University of Surrey.
- [9] Khunthi, S., Saichua, P., & Surinta, O. (2019). Effective face verification systems based on the histogram of oriented gradients and deep learning techniques (pp. 1-6). In *14th International Joint Symposium on Artificial Intelligence and Natural Language Processing (ISAI-NLP)*. Chiang Mai, Thailand.