

โปรแกรมอัจฉริยะสำหรับแปลภาษามือภาษาไทยและภาษาอังกฤษ Smart Application for Translating Thai and English Sign Language

สหโชค คุ่มวงษา¹, วียดา ยะไวทย์^{1,*}
Sahachok Khumwongsa¹, Wiyada Yawai^{1,*}

¹ หลักสูตรวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครราชสีมา นครราชสีมา 30000 ประเทศไทย

¹ Computer Science, Faculty of Science and Technology, Nakhon Ratchasima Rajabhat University, Nakhon Ratchasima 30000, Thailand

* Corresponding Author: Wiyada Yawai, wiyada.y@nrru.ac.th

Received:
5 March 2023
Revised:
20 May 2023
Accepted:
17 June 2023

คำสำคัญ:

แปลภาษามือ, มีเดียไปป์, ผู้พิการทางการได้ยิน, ภาษามือภาษาไทย, ภาษามือภาษาอังกฤษ

Keywords:

Translate Sign Language, MediaPipe, Hearing Impaired, Thai Sign Language, English Sign Language

บทคัดย่อ: ปัจจุบันผู้พิการทางการได้ยินสื่อสารด้วยการใช้ภาษามือ ทำให้อาจส่งผลต่อการสื่อสารกับบุคคลทั่วไป เนื่องจากบุคคลทั่วไปส่วนใหญ่ไม่มีความรู้ในการใช้ภาษามือ ดังนั้นในการสื่อสารอาจทำให้เกิดการสื่อสารผิดพลาดได้ งานวิจัยนี้ มีวัตถุประสงค์เพื่อพัฒนาโปรแกรมสำหรับแปลภาษามือภาษาไทยและภาษาอังกฤษ โดยสามารถแปลภาษามือภาษาไทย ภาษาอังกฤษ ตัวเลขอารบิก และคำภาษาอังกฤษ โดยใช้อัลกอริทึมมีเดียไปป์ในการตรวจจับมือ และหาจุดสำคัญบนมือที่มีจำนวน 21 จุด โดยงานวิจัยนี้สามารถตรวจจับมือสูงสุดได้จำนวน 2 มือ ทำให้สามารถนำจุดสำคัญทั้งสิ้น 42 จุดไปตัดสินใจเพื่อให้โปรแกรมจำแนก ภาษามือที่มีจำนวนทั้งสิ้น 68 สัญลักษณ์ ในการทดสอบนั้น ผู้วิจัยได้กำหนดให้มีผู้ทดลองทั้งสิ้นจำนวน 3 คนทำการทดลองให้โปรแกรมรู้จำภาษามือจำนวนสัญลักษณ์ละ 5 ครั้ง และกำหนดให้กล้องเว็บแคมจะต้องอยู่ห่างจากผู้ทดลองประมาณ 50 เซนติเมตร จากการทดลองพบว่าในการรู้จำภาษามือภาษาอังกฤษนั้นมีความถูกต้องมากกว่าภาษามือไทย เนื่องจากบางสัญลักษณ์ในภาษามือไทยนั้นจะต้องใช้การแสดงท่าทาง 2 ระดับ จึงทำให้เกิดความผิดพลาดในการรู้จำภาษามือ โดยงานวิจัยนี้มีผลการรู้จำภาษามือที่ 95.39%

Abstract: Currently, the sign language is used as a means of communication for individuals with hearing impairments. Since most people do not have knowledge to use sign language, as a result, miscommunication may occur. The aim of this project is to develop a program for translating Thai and English sign language. It can translate Thai sign language, English sign language, Arabic

numbers, and English words. MediaPipe algorithm is used for hand detection and find 21 Keypoints on each hand. This research is able to detect a maximum of two hands, allowing a total of 42 Keypoints to be used to make decisions for the classification of sign language, which encompasses a total of 68 symbols in the test. A total of 3 experimenters are assigned to experiment with the sign language recognition program 5 times per symbol and specified that the webcam must be approximately 50 centimeters away from the experimenter. As a result from the experiment, it was found that English sign language recognition was more accurate than Thai sign language. This is because some signs in Thai sign language require two levels of gestures. This causes errors in sign language recognition. This research has a sign language recognition result of 95.39%.

1. บทนำ

จากรายงานสถิติประชากรทางการทะเบียนพบว่าประเทศไทยมีประชากรจำนวน 66.2 ล้านคน (Department of Provincial Administration, 2022) และหากพิจารณาข้อมูลด้านคนพิการในประเทศไทยที่รายงานจากกระทรวงการพัฒนาสังคมและความมั่นคงของมนุษย์ กรมส่งเสริมและพัฒนาคุณภาพชีวิตพบว่า มีจำนวนคนพิการมากถึง 2,108,536 คน โดยแยกเป็นผู้พิการทางการได้ยินจำนวน 393,998 คน คิดเป็นร้อยละ 18.69% ของผู้พิการทั้งหมด (Department of Empowerment of Persons with Disabilities, 2022) จากรายงานข้างต้นจะเห็นว่าปัญหาด้านการสื่อสารนับว่าเป็นปัญหาใหญ่สำหรับผู้พิการและยังรวมถึงผู้พิการทางการได้ยินที่จะสื่อสารกับบุคคลทั่วไป โดยการสื่อสารสำหรับผู้พิการทางการได้ยินจะใช้ภาษามือในการสื่อสาร ในแต่ละภาษามือก็มีการใช้ท่าทางที่แตกต่างกัน (Chumchim & Maneerat, 2020) ดังนั้น จึงเป็นเรื่องยุ่งยากในการสื่อสารกับบุคคลทั่วไปไม่ว่าจะด้านเชื้อชาติและภาษาที่ใช้อาจส่งผลทำให้เกิดความเข้าใจคลาดเคลื่อน

ทั้งนี้ นักวิจัยได้พัฒนางานวิจัยการรู้จำภาษามือด้วยวิธีการที่แตกต่างกันเพื่อช่วยในการสื่อสารระหว่างผู้พิการและบุคคลทั่วไปทำให้สามารถสื่อสารกันได้ เช่น Wadhawan & Kumar (2020) ได้ใช้การเรียนรู้เชิงลึก (Deep Learning) เพื่อรู้จำภาษามือ

อินเดีย (Indian Sign Language) โดยมุ่งเน้นในการรู้จำภาษามือแบบคงที่ (Static Sign) โดยงานวิจัยดังกล่าวใช้วิธีโครงข่ายประสาทแบบคอนโวลูชัน (Convolutional Neural Network: CNN) โดยใช้เพื่อเรียนรู้จากรูปภาพสัญลักษณ์ภาษามือจำนวนทั้งสิ้น 35,000 รูปภาพจากภาษามือแบบคงที่จำนวนทั้งสิ้น 100 รูปแบบ (Static Sign)

Camgz *et al.* (2020) พัฒนาการรู้จำภาษามือ (Sign Language Recognition) และการแปลภาษา (Translation) โดยใช้วิธีการที่เรียกว่าทรานส์ฟอร์มเมอร์ (Transformer) ซึ่งเป็นสถาปัตยกรรมที่ถูกใช้อย่างแพร่หลายในงานวิจัยด้านต่างๆ โดยเฉพาะทางด้านการประมวลผลภาษาธรรมชาติ (Natural Language Processing) และการรู้จำรูปภาพและวิดีโอ (Image and Video Recognition) โดยงานวิจัยดังกล่าวได้ออกแบบสถาปัตยกรรมชื่อ Single Layered Sign Language Transformer เพื่อใช้สำหรับเรียนรู้วิดีโอภาษามือและแปลงให้อยู่ในรูปแบบของข้อความตัวอักษร (Text) โดยการหาคำตอบที่ถูกต้องจากการแปลงวิดีโอภาษามือให้เป็นข้อความตัวอักษรใช้วิธีการ Connectionist Temporal Classification (CTC) และจากนั้นจึงนำตัวอักษรที่ได้ไปแปลเป็นประโยคภาษาพูด (Spoken Language Sentence) โดยใช้วิธีการ Sign2(Gloss+Text)

ทั้งนี้ กูเกิลได้พัฒนาซอฟต์แวร์มีเดียไปป์ (MediaPipe) (Grishchenko & Bazarevsky, 2022) ที่เป็นซอฟต์แวร์ที่พัฒนาด้วยเทคโนโลยีการเรียนรู้ของเครื่องในการประมวลผลเพื่อรองรับการทำงานทางด้านปัญญาประดิษฐ์ (Artificial Intelligence: AI) โดยหนึ่งในเทคนิคที่สำคัญของ MediaPipe คือการตรวจจับมือ (Hand Detection) และหาพื้นที่แลนด์มาร์กของมือ (Hand Landmark) เพื่อนำมาใช้สำหรับการรู้จำท่าทาง (Gesture) โดย Hand Landmark ที่ MediaPipe ได้พิจารณาว่ามีทั้งสิ้นจำนวน 21 จุด โดยสามารถตรวจจับจากทั้งสองมือในเวลาเดียวกัน ทำให้สามารถนำไปช่วยในการพัฒนาการตรวจจับท่าทางลักษณะของมือได้อย่างแม่นยำ

จากที่กล่าวมาข้างต้น ผู้วิจัยจึงมีแนวคิดที่จะพัฒนาต้นแบบโปรแกรมอัจฉริยะสำหรับแปลภาษามือแบบภาษาไทยและภาษาอังกฤษ โดยมีวัตถุประสงค์ของงานวิจัย เพื่อพัฒนาโปรแกรมอัจฉริยะสำหรับแปลภาษามือแบบภาษาไทยและภาษาอังกฤษโดยใช้เทคโนโลยี MediaPipe ในการตรวจจับมือและแปลภาษามือโดยสามารถแปลได้ทั้งตัวอักษรภาษาไทย ภาษาอังกฤษ ตัวเลข และคำในภาษาอังกฤษที่จำเป็น เช่น Stop, Left, Right, Forward และ Backward ทั้งนี้ ในการแสดงผลลัพธ์ ผู้วิจัยได้พัฒนาการแสดงผลลัพธ์ทำให้ต้นแบบโปรแกรมสามารถแสดงผลลัพธ์ได้ทั้งข้อความและเสียงพูด เพื่อให้สะดวกในการสื่อสารมากยิ่งขึ้น

ดังนั้น เห็นได้ว่างานวิจัยแปลภาษามือมีความสำคัญต่องานในหลากหลาย ด้าน เช่น 1) ด้านภาษาศาสตร์ เพื่อช่วยสร้างความเข้าใจในการสื่อสารด้วยภาษามือ ทั้งภาษาไทยและภาษาอังกฤษ 2) ด้านวิทยาการคอมพิวเตอร์และปัญญาประดิษฐ์ เพื่อช่วยในการพัฒนาและประยุกต์ใช้ MediaPipe สำหรับการแปลภาษามือ และช่วยให้ผู้วิจัยที่มีความสนใจงานทางด้านนี้มีความเข้าใจในวิธีการจำแนกข้อมูลภาพ และการเรียนรู้ของเครื่อง (Machine Learning) 3) ด้าน

การศึกษาและการเรียนรู้ เพื่อช่วยเสริมสร้างทักษะการสื่อสารของผู้ที่ใช้ภาษามือ ทั้งผู้ที่ใช้ภาษาไทยและภาษาอังกฤษ รวมถึงการสร้างความเข้าใจในการเรียนรู้ภาษามือเพื่อคนทั่วไป และ 4) ด้านการประยุกต์ใช้งานการแปลภาษามือ สามารถนำไปใช้ประยุกต์ใช้ในหลากหลาย สถานการณ์ เช่น การประยุกต์ใช้ได้ทั้งผู้พิการทางการได้ยินและบุคคลทั่วไปได้ใช้งานได้อย่างสะดวกรวดเร็ว ลดปัญหาการทำความเข้าใจคลาดเคลื่อนทั้งสองภาษา คือภาษาไทยและภาษาอังกฤษ ทั้งยังเพิ่มเติมกรณีที่ผู้พิการทางการได้ยินขอความช่วยเหลือกรณีเกิดเหตุฉุกเฉินได้

2. งานวิจัยและทฤษฎีที่เกี่ยวข้อง

2.1 งานวิจัยที่เกี่ยวข้อง

Jeekratok, Tongthaku, & Jampamoon (2019) ได้พัฒนาระบบภาษามือไทยในชีวิตประจำวัน โดยพัฒนาเป็นเว็บแอปพลิเคชัน (Web Application) โดยมีวัตถุประสงค์เพื่อใช้เป็นแหล่งรวบรวมข้อมูลภาษามือไทย เพื่อให้บุคคลที่มีความสนใจในภาษามือไทยสามารถเข้าไปศึกษาและเรียนรู้วิธีการทำภาษามือ โดยระบบภาษามือไทยนั้นได้ถูกประเมินประสิทธิภาพจากผู้เชี่ยวชาญ และศึกษาความพึงพอใจของนักศึกษา คณะวิทยาการจัดการในมหาวิทยาลัยราชภัฏอุดรธานี ผลวิจัยพบว่าระบบภาษามือไทยในชีวิตประจำวัน มีระดับคุณภาพประกอบที่ดีจากการประเมินโดยผู้เชี่ยวชาญโดยมีค่าเฉลี่ยเท่ากับ 4.24 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.55 และนักศึกษามีความพึงพอใจในการใช้ระบบนี้ในระดับมากโดยมีค่าเฉลี่ยเท่ากับ 4.24 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.66

Chumchim & Maneerat (2020) พัฒนาแอปพลิเคชันบนมือถือ (Mobile Application) เพื่อแปลงข้อความเสียง (Voice) ให้เป็นข้อความ (Text) และจากนั้นจึงพัฒนาโมเดลแบบสามมิติ (3D Model) จากโปรแกรม DAZ 3D โดยเป็นโมเดลที่สามารถ

เคลื่อนไหวได้ โดยมีวัตถุประสงค์คือนำข้อความนั้นไปค้นหากับพจนานุกรมที่ได้ออกแบบไว้คือ พจนานุกรมภาษาไทย-พจนานุกรมภาษามือไทย โดยนำคำศัพท์ภาษาไทยมาจัดเรียงใหม่ตามลักษณะของไวยากรณ์ที่ถูกออกแบบไว้ จากนั้นจึงนำผลลัพธ์ที่ได้ไปค้นหากจากพจนานุกรมภาษามือสามมิติ ซึ่งเป็นพจนานุกรมที่จัดเก็บข้อมูลความสัมพันธ์ของคำศัพท์ภาษามือ ข้อมูลตำแหน่ง และการเคลื่อนที่ของมือ โดยแอปพลิเคชันระบบการแปลภาษามือให้กับผู้พิการทางการได้ยินมีความถูกต้องในการแปลภาษามือ โดยพิจารณาจากความถูกต้องของไวยากรณ์ และภาษามือ โดยความถูกต้องของไวยากรณ์อยู่ที่ 91.40% และความถูกต้องของภาษามืออยู่ที่ 85.57%

Nadee & Yingkayun (2020) ได้การพัฒนาวิธีการแยกแยะรูปภาพภาษามือโดยอาศัยวิธีการปรับระนาบภาพโดยเน้นที่ภาษามืออเมริกัน (American Sign Language) ที่มีจำนวนทั้งสิ้น 24 สัญลักษณ์ งานวิจัยนี้เสนอวิธีการใหม่ในการแยกแยะรูปภาพภาษามือ โดยงานวิจัยได้ทำการประมวลผลภาพเบื้องต้นโดยการแปลงภาพสีขาวดำ ให้กลับเป็นภาพสีเทาที่มีค่าสีอยู่ในช่วงระหว่าง 0-255 และภาพสีนั้นจะใช้ค่าสีระดับ 0-255 ของค่าสีแบบ RGB จากนั้นเตรียมข้อมูลรูปภาพที่จะนำไปเรียนรู้ด้วยการเรียนรู้ของเครื่อง ได้แก่ K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Logistic Regression, Decision Tree และ Neural Network เพื่อหาค่าตัวถ่วงน้ำหนักสำหรับใช้พิจารณาการเคลื่อนที่ของชุดข้อมูลไปในทิศทางใด จากนั้นจึงนำค่าที่ได้จากการเรียนรู้ของเครื่องเป็นตัวพิจารณาเพื่อตัดสินใจที่จะปรับระดับระนาบภาพ หรือความเอียงของมือ (Rotation Adjustment) และนำรูปภาพมือที่ผ่านการปรับระดับความเอียงไปเรียนรู้ด้วยการเรียนรู้ของเครื่องอีกครั้ง เพื่อพิจารณาว่ามุมใดของมือที่จะทำให้การรู้จำสูงสุด โดยการทดลองพบว่าระบบที่ออกแบบสามารถรู้จำได้สูงสุดที่ 98.06%

Sriboonruang *et al.* (2022) ได้พัฒนาโปรแกรมแปลภาษามือเป็นข้อความ โดยใช้แพลตฟอร์ม (Platform) ของ MediaPipe เพื่อหา Hand Landmark จากวิดีโอจำนวนทั้งสิ้น 21 จุดจากมือ 1 ข้าง โดยค่าที่นำมาพิจารณาคือตำแหน่งของจุดพิกัดแกน x, y และ z ดังนั้น จึงได้จุดทั้งสิ้น 42 จุดจากมือทั้งสองข้าง จากนั้นจึงนำค่าทั้งหมดส่งเข้าไปยังโครงข่ายประสาทเทียมแบบประมวลผลลำดับ (Long Short-Term Memory: LSTM) เพื่อสร้างโมเดลที่ใช้สำหรับการรู้จำภาษามือ โดยสามารถรู้จำตัวเลข (1-9) และภาษามือจำนวน 20 คำ สำหรับการรู้จำตัวเลขมีความถูกต้อง 94%-99% และการรู้จำภาษามือจำนวน 20 คำมีความถูกต้อง 80.5% ทั้งนี้ ความถูกต้องในการรู้จำต่างกันเนื่องจากความซับซ้อนของท่าทาง

จากงานวิจัยที่กล่าวมา สังเกตว่าจำนวนการแปลข้อความมีจำนวนการแปลได้ไม่เกิน 30 คำ และภาษาในการแปลจะแปลเฉพาะภาษาใดภาษาหนึ่ง ดังนั้น ผู้วิจัยจึงได้พัฒนาโปรแกรมเพื่อช่วยเพิ่มประสิทธิภาพของการแปลภาษามือให้มีความหลากหลายมากขึ้น และช่วยลดปัญหาในการสื่อสารระหว่างผู้พิการทางการได้ยินและบุคคลทั่วไป โดยการใช้กล้องเว็บแคม (Webcam) เพื่อตรวจจับท่าทางของมือเพื่อแปลภาษามือแบบเรียลไทม์ จำนวนทั้งสิ้น 68 คำ โดยแบ่งเป็นภาษาไทย 31 คำ ภาษาอังกฤษ 28 คำ ตัวเลข 1 ถึง 9 ตลอดจนกระทั่งการแปลการขอความช่วยเหลือกรณีเกิดเหตุการณ์ฉุกเฉินอีกด้วย ทำให้สามารถขอความช่วยเหลือได้ทันท่วงทีป้องกันการเกิดเหตุร้ายขึ้น โดยโปรแกรมจะทำงานแสดงได้ทั้งข้อความและเสียง ดังมีรายละเอียดงานวิจัยในหัวข้อถัดไป

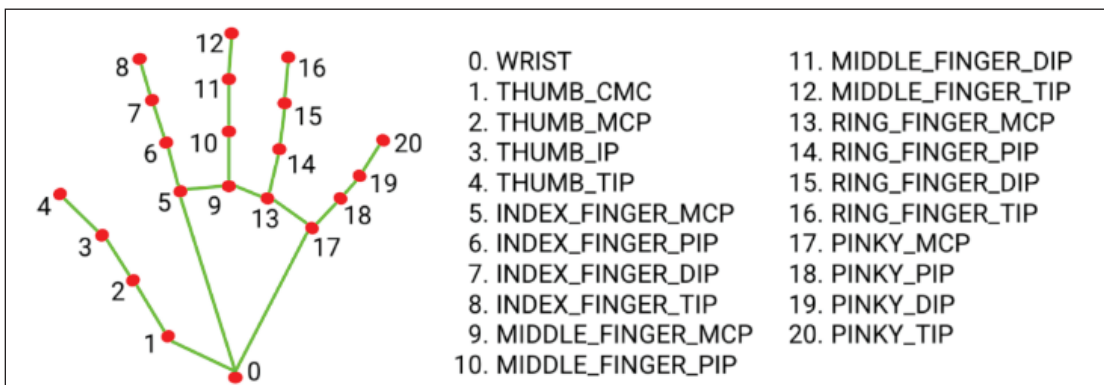
2.2 มีเดียไปป์ (MediaPipe)

MediaPipe เป็นเทคโนโลยีปัญญาประดิษฐ์ (Artificial intelligence: AI) ที่พัฒนาโดย Google ที่เป็นไปป์ไลน์ (Pipeline) ในการตรวจจับมือ ใบหน้า

และท่าทาง โดยมีความเร็วในการตรวจจับที่สูงจึง
เหมาะแก่การนำมาพัฒนาได้หลากหลายรูปแบบ
และสามารถนำไปใช้ได้ ในหลายแพลตฟอร์ม ในการ
ทำงานของ MediaPipe นั้นจะนำโมเดลของ มือ
ใบหน้า และท่าทางมารวมเข้าไว้ในตัวเดียวกัน จึง
สามารถแยกส่วนการทำงานของส่วนต่างๆ ของ
MediaPipe ได้ โดยในแต่ละส่วนของโมเดลนั้นจะ
มีจุดสำคัญ (Key Point) เป็นของตัวเอง โดยโมเดล
ท่าทางมีทั้งหมด 33 จุด มือข้างละ 21 จุด และใบหน้า
468 จุด ซึ่งเมื่อนำมารวมกันก็จะมีจุดสำคัญทั้งหมด
543 จุด (Grishchenko & Bazarevsky, 2022)

ในงานวิจัยนี้ใช้ MediaPipe Hands ซึ่ง
เป็นการทำงานแบบการเรียนรู้ของเครื่อง (Machine
Learning) ประมวลผลโดยการนำโมเดลมือมารวม

เข้าใจด้วยกัน (MediaPipe, 2023) แต่ในการทำวิจัยนี้จะใช้เพียงส่วนของมือแต่ละข้างเท่านั้น สำหรับการเขียนโปรแกรมด้วยภาษาไพธอน (Python) เพื่อเรียกใช้ MediaPipe Hands นั้นจะต้อง Import ตัว Library ของ MediaPipe เสียก่อน โปรแกรมที่เขียนจะสามารถเรียกใช้ฟังก์ชันทั้งหมดของ MediaPipe ได้ ตัวของ MediaPipe Hands จะตรวจหาจุดสนใจบนฝ่ามือในรูปภาพหรือวิดีโอ จากนั้น MediaPipe จะทำเครื่องหมายไว้ที่จุดต่างๆ ของมือ มีทั้งหมด 21 จุด หากโปรแกรมตรวจพบมือทั้ง 2 ข้างก็จะทำเครื่องหมายบนมือทั้งสอง ดังนั้น การเขียนโปรแกรมจะต้องทำการเทียบจุดบนมือทั้ง 21 จุดเพื่อทำการแปลให้เป็นภาษามือ โดยตำแหน่งของจุดบนฝ่ามือที่ตรวจจับด้วย MediaPipe Hands แสดงดังภาพประกอบ 1



ภาพประกอบ 1 ตำแหน่งของจุดทั้งสิ้น 21 จุดบนฝ่ามือที่ตรวจจับด้วย MediaPipe Hand (MediaPipe, 2023)



ภาพประกอบ 2 ต้นแบบภาษามือไทย (Prajan, 2022)

2.3 ต้นแบบภาษามือ

งานวิจัยฉบับนี้ได้พิจารณาเลือกใช้ภาษามือภาษาไทย ภาษาอังกฤษ ตัวเลขอารบิก และภาษามือคำภาษาอังกฤษ โดยรายละเอียดของภาษามือแสดงดังต่อไปนี้

2.3.1 ต้นแบบภาษามือภาษาไทย

ต้นแบบภาษามือภาษาไทย เป็นการแสดงแบบสะกดนิ้วมือไทย ซึ่งประกอบด้วยตัวอักษร

ภาษาไทย (ก-ฮ) และสระภาษาไทย (Prajan, 2022) โดยภาพประกอบ 2 แสดงตัวอย่างต้นแบบภาษามือไทย

2.3.2 ต้นแบบภาษามือภาษาอังกฤษ

ต้นแบบภาษามือภาษาอังกฤษ มีทั้งหมด 26 ตัวอักษร ประกอบด้วยตัวอักษร a-z (Disabled World, 2022) ตัวอย่างภาษามืออังกฤษทั้ง 26 รูปแบบแสดงดังภาพประกอบ 3



ภาพประกอบ 3 ต้นแบบภาษามืออังกฤษ (Disabled World, 2022)

1 หนึ่ง	2 สอง	3 สาม	4 สี่	5 ห้า	6 หก	7 เจ็ด	8 แปด	9 เก้า	10 สิบ
สิบเอ็ด	สิบสอง	สิบสาม	สิบสี่	สิบห้า	สิบหก	สิบเจ็ด	สิบแปด	สิบเก้า	
ยี่สิบ	สามสิบ	สี่สิบ	ห้าสิบ	หกสิบ	เจ็ดสิบ	แปดสิบ	เก้าสิบ		
หนึ่งร้อย	หนึ่งพัน	หนึ่งหมื่น	หนึ่งแสน	หนึ่งล้าน	สิบล้าน	หนึ่งร้อยล้าน			
100	1,000	10,000	100,000	1,000,000	10,000,000	100,000,000			

ภาพประกอบ 4 ต้นแบบภาษามือตัวเลขอารบิก (Daily News Online, 2021)

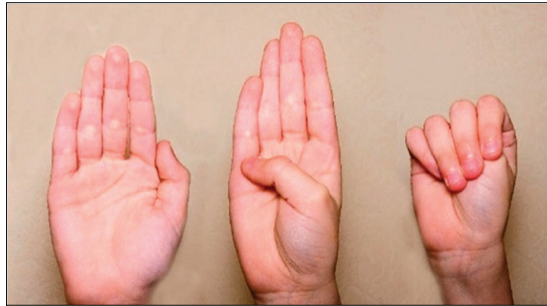
2.3.3 ต้นแบบภาษามือตัวเลขอารบิก

ต้นแบบภาษามือตัวเลขอารบิก เป็นการสื่อสารตั้งแต่เลข 1 จนกระทั่งหนึ่งร้อยล้าน (Daily News Online, 2021) ตัวอย่างภาษามือเพื่อสะกดตัวเลขแสดงดังภาพประกอบ 4

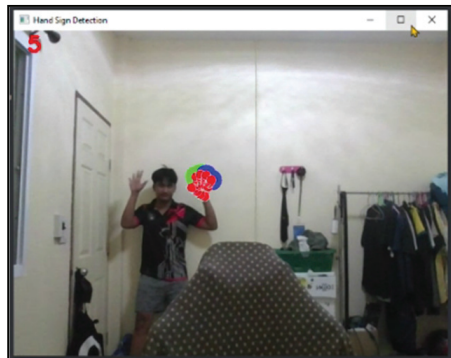
2.3.4 ต้นแบบภาษามือคำภาษาอังกฤษ

ต้นแบบภาษามือคำภาษาอังกฤษ ที่ถูกนำ

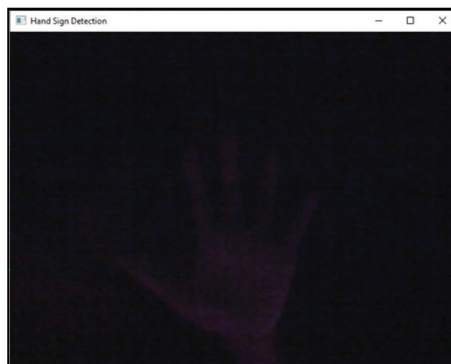
มาใช้กันอย่างแพร่หลายได้แก่ การขอความช่วยเหลือ โดยสากลการขอความช่วยเหลือจะต้องแสดงท่าทาง 3 ระดับ (Agency, 2021) ตัวอย่างการขอความช่วยเหลือด้วยภาษามือแสดงดังภาพประกอบ 5 โดยระดับที่ 1 ยกมือเหยียดนิ้วทุกนิ้วตั้งตรงและชิดกัน ระดับที่ 2 พับหัวนิ้วโป้งเข้ากับฝ่ามือ และระดับที่ 3 พับนิ้วมือทั้ง 4 นิ้ว คือ นิ้วชี้ นิ้วกลาง นิ้วนาง และนิ้วก้อย เข้ากับฝ่ามือ



ภาพประกอบ 5 ต้นแบบขอความช่วยเหลือแบบภาษามือ (Agency, 2021)



ภาพประกอบ 6 การตรวจจับมือในระยะประมาณ 350 เซนติเมตร



ภาพประกอบ 7 แสดงให้เห็นถึงภาพที่ถ่ายในพื้นที่แสงน้อย

3. วิธีการดำเนินงานวิจัย

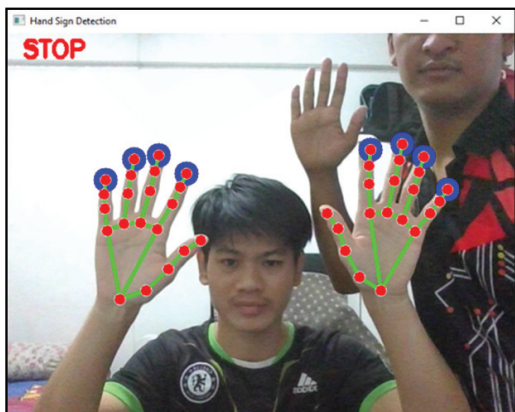
3.1 เครื่องมือที่ใช้พัฒนา

งานวิจัยฉบับนี้เลือกพัฒนาโปรแกรมอัจฉริยะสำหรับแปลภาษามือแบบภาษาไทยและภาษาอังกฤษบนเครื่องคอมพิวเตอร์โน้ตบุ๊ก (Notebook Computer) ที่ติดตั้งระบบปฏิบัติการ Window 10 Home 64-bit และเลือกใช้ภาษาไพธอน (Python) ในการพัฒนาโปรแกรม โดยเฟรมเวิร์กหลักที่ใช้ในการหาจุดสำคัญบนมือคือ MediaPipe

3.2 ขอบเขตของงานวิจัย

3.2.1 ด้านความสามารถของโปรแกรม

- 1) สามารถตรวจจับจุดบนฝ่ามือในพื้นที่ที่มีแสงเพียงพอ
- 2) สามารถตรวจจับจุดบนฝ่ามือได้ไกลถึงประมาณ 350 เซนติเมตร ในกรณีที่เจอจุดฝ่ามือแล้วเท่านั้น ตัวอย่างการตรวจจับมือในระยะ 350 เซนติเมตรแสดงดังภาพประกอบ 6
- 3) สามารถตรวจจับจุดบนฝ่ามือได้แม้จะเปลี่ยนผู้ทดลองหรือฉากหลัง



ภาพประกอบ 8 การตรวจจับจุดบนฝ่ามือได้เพียง 2 ข้าง

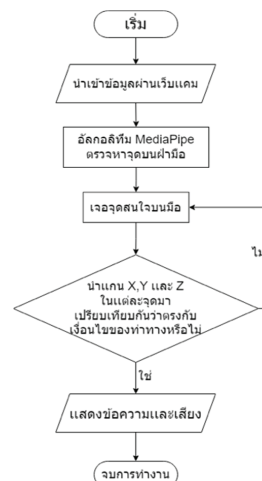
3.2.2 ด้านข้อจำกัดของโปรแกรม

- 1) ไม่สามารถตรวจจับได้ในพื้นที่ที่มีแสงน้อยได้ ตัวอย่างภาพที่มีแสงน้อยแสดงดังภาพประกอบ 7
- 2) สามารถตรวจจับจุดบนฝ่ามือได้จากมือจำนวน 2 ข้างเท่านั้น ตัวอย่างภาพที่มีจำนวนมือมากกว่า 2 ข้าง แสดงดังภาพประกอบ 8

3.3 ขั้นตอนการดำเนินงาน

3.3.1 การวิเคราะห์และออกแบบโปรแกรมสำหรับแปลภาษามือ

การทำงานของโปรแกรมสำหรับแปลภาษามือแบบภาษาไทยและภาษาอังกฤษแสดงดังภาพประกอบ 9 โดยมีกระบวนการดังนี้ นำเข้าข้อมูลผ่านเว็บแคม จากนั้นอัลกอริทึมของ MediaPipe จะทำการตรวจหาจุดสนใจบนฝ่ามือ เมื่อเจอจุดสนใจบนฝ่ามือแล้วจะทำการเปรียบเทียบจุดแต่ละจุดว่าตรงตามเงื่อนไขที่โปรแกรมตั้งไว้หรือไม่ ถ้าหากจุดสนใจนั้นตรงตามเงื่อนไขของโปรแกรมจะแสดงข้อความของภาษามือที่ตั้งไว้ แต่ถ้าหากไม่ตรงกัน โปรแกรมจะกลับไปหาจุดสนใจใหม่และเช็คเงื่อนไขใหม่อีกจนกว่าจะเข้าเงื่อนไข



ภาพประกอบ 9 กระบวนการทำงานของโปรแกรมสำหรับแปลภาษามือแบบภาษาไทยและภาษาอังกฤษ

3.3.2 การพัฒนาโปรแกรมสำหรับแปลภาษามือ

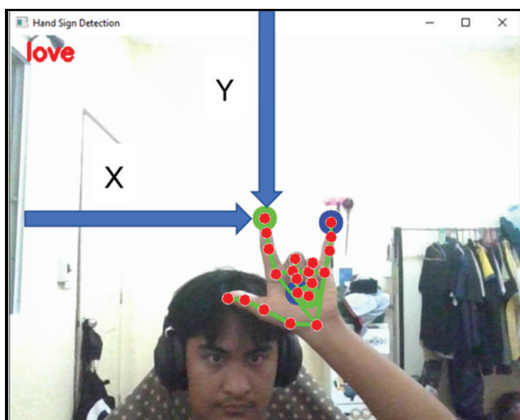
การพัฒนาโปรแกรมสำหรับแปลภาษามือใช้โปรแกรม Visual Studio Code ในการเขียนโปรแกรมภาษา Python ในการเขียนโปรแกรมจำเป็นต้องติดตั้ง Library ที่มีชื่อว่า OpenCV-python และ MediaPipe ผ่านคำสั่ง pip install ดังนั้น ก่อนเขียนโปรแกรมจึงต้องนำเข้าทั้ง Library ทั้ง 2 ตัวเสียก่อน ดังนั้นในการเขียนโปรแกรมต้องดำเนินการคือ

1) นำเข้า Library ทั้ง 2 library (OpenCV-python และ MediaPipe) เพื่อใช้ในการเรียกใช้งานคำสั่งต่างๆ ในการทำงานของโปรแกรม

2) เขียนโปรแกรมเพื่อตรวจหาจุดบนฝ่ามือทั้ง 2 ข้าง และตรวจสอบตำแหน่งและท่าทางของนิ้วมือ จากนั้นจึงนำตำแหน่งของนิ้วมือไปตรวจสอบตามเงื่อนไขที่กำหนดไว้

3) ผู้วิจัยทำการบันทึกเสียงเพื่อใช้แสดงผลของการรู้จำจากการแปลภาษามือแบบภาษาไทยและภาษาอังกฤษ เพื่อให้บุคคลทั่วไปสามารถเห็นภาพ อ่านข้อความ และฟังเสียงที่แปลออกมาได้

ลักษณะการทำงานของ MediaPipe แสดงได้ดังภาพประกอบ 10 โดยโปรแกรมจะทำการครอบ



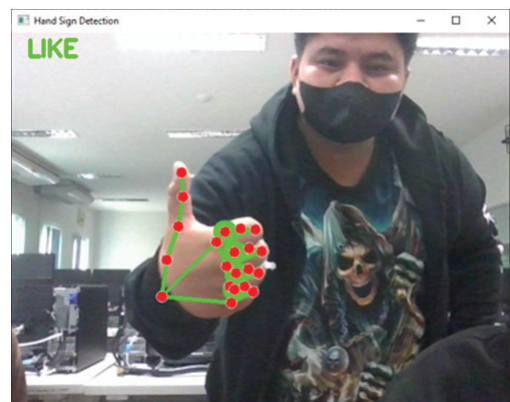
ภาพประกอบ 10 การกำหนดค่า X และ Y ของจุดสนใจบนมือ

(Crop) ตัดภาพในส่วนของมือและทำการตรวจหาจับจุดสนใจ (Region of Interest : ROI) ทั้งหมด 21 จุดบนฝ่ามือในแต่ละจุดบนฝ่ามือนั้นจะมีค่า X, Y และ Z โดย X จะแทนค่าจากขอบจอด้านซ้ายไปหาจุดสนใจ โดย Y จะแทนค่าจากขอบจอด้านบนลงมาหาจุดสนใจ และ Z จะแทนค่าความลึกของจุดสนใจ

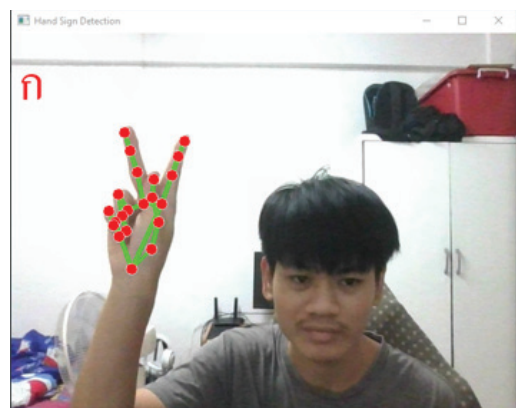
4. ผลการทดลองและอภิปรายผล

4.1 ข้อมูลที่ใช้ทดลอง

ข้อมูลลักษณะท่าทางของภาษามือในแต่ละคำนั้นมีความยากง่ายไม่เหมือนกันบางคำใช้ท่าทางเพียงท่าทางเดียว บางคำอาจใช้ท่าทาง 2-3 ท่าหรือ



ภาพประกอบ 11 ภาษามือคำว่า Like



ภาพประกอบ 12 ผลลัพธ์การแสดงภาษามือภาษาไทยของตัวอักษร ก

มากกว่านั้นในการแปลภาษาเพียง 1 คำ ในงานวิจัยฉบับนี้ใช้ทำทางในการพัฒนาและทดสอบทั้งหมด 68 สัญลักษณ์ ประกอบด้วย

4.1.1 ภาษามือภาษาไทย 31 ตัวอักษร คือ ก, ข, ค, ช, ฉ, ญ, ฎ, ฐ, ท, ฒ, ณ, ต, ถ, ท, น, บ, ป, ผ, ฝ, พ, ฟ, ภ, ม, ย, ร, ล, ว, ห, ฬ, อ และ ฮ

4.1.2 ภาษามือภาษาอังกฤษ 20 ตัวอักษร คือ a, b, c, d, f, h, i, k, L, m, n, o, p, q, r, t, v, w, x และ y

4.1.3 ภาษามือคำภาษาอังกฤษ 8 คำ คือ stop, left, right, forward, backward, like, dislike และ love

4.1.4 ภาษามือตัวเลขอารบิก 9 ตัวเลข คือ 1, 2, 3, 4, 5, 6, 7, 8 และ 9

4.2 วิธีการทดลอง

4.2.1 ทดลองการตรวจจับมือโดยใช้มือทั้งสองข้างของผู้วิจัยและมือของบุคคลอื่นๆ

4.2.2 ทดลองการตรวจจับมือโดยเปลี่ยนฉากหลัง (Background)

4.2.3 ทดลองการตรวจจับมือโดยเปลี่ยนความสว่างของแสง

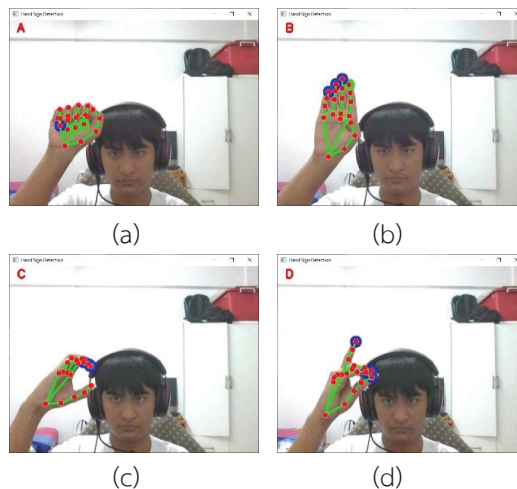
4.2.4 ทดลองการตรวจจับมือโดยเปลี่ยนระยะห่างระหว่างกล้องเว็บแคมและบุคคล

4.3 ผลการทดลอง

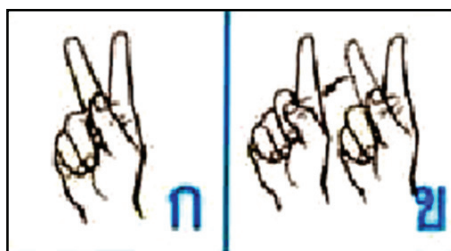
ภาพประกอบ 11-13 แสดงให้เห็นถึงการทดลองการตรวจจับมือ การค้นหาจุดสนใจบนมือ และการแปลภาษามือ (ภาษาไทย ภาษาอังกฤษ ตัวเลขอารบิก และคำภาษาอังกฤษ) จากการทดลองพบว่า ระยะห่างระหว่างบุคคลและกล้องเว็บแคมที่เหมาะสมที่สุดอยู่ที่ระยะ 50 เซนติเมตร ดังนั้น ในการวัดประสิทธิภาพจึงกำหนดไว้ที่ระยะดังกล่าว

ตัวอย่างของการทดลองการแปลภาษามือ คำภาษาอังกฤษ แสดงดังภาพประกอบ 11 ซึ่งแสดงถึงการแปลคำภาษาอังกฤษคือ Like

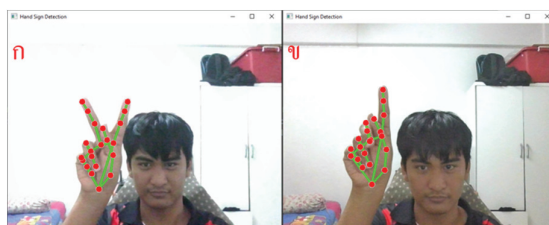
ตัวอย่างของการทดลองการแปลภาษามือ ภาษาไทย แสดงดังภาพประกอบ 12 โดยแสดงให้เห็นถึงการตรวจจับมือ การหาจุดสนใจบนมือ และการแปลภาษามือภาษาไทยของตัวอักษร ก และ ภาพประกอบ 13 แสดงให้เห็นถึงการทดลองการแปลภาษามือภาษาอังกฤษของตัวอักษร a-d



ภาพประกอบ 13 ผลลัพธ์การแสดงผลภาษามือภาษาอังกฤษของ (a) ตัวอักษร A (b) ตัวอักษร B (c) ตัวอักษร C และ (d) ตัวอักษร D

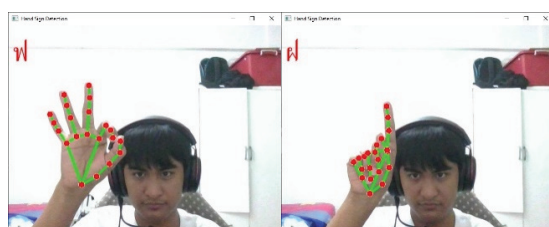


(a)

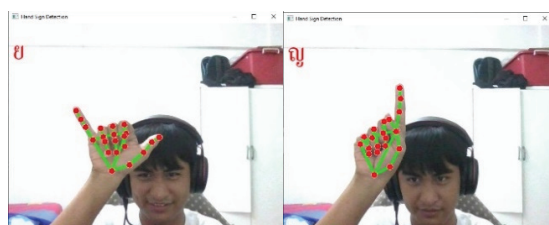


(b)

ภาพประกอบ 14 การเปลี่ยนท่าแบบ 2 ระดับ (a) เปลี่ยนจากตัวอักษร ก เป็นตัวอักษร ข และ (b) แสดงผลลัพธ์การแปลภาษามือภาษาไทยจากตัวอักษร ก เป็นตัวอักษร ข

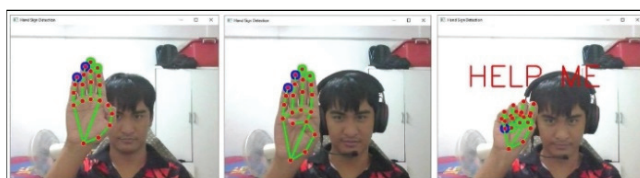


(a)



(b)

ภาพประกอบ 15 ผลลัพธ์การแปลภาษามือภาษาไทยแบบแบบ 2 ระดับ (a) เปลี่ยนจากตัวอักษร ฟ เป็นตัวอักษร ฝ และ (b) เปลี่ยนจากตัวอักษร ย เป็นตัวอักษร ญ



ภาพประกอบ 16 ผลลัพธ์การแปลคำภาษาอังกฤษ โดยเป็นการตรวจจับภาษามือแบบ 3 ระดับ ซึ่งหมายถึง สัญญาณขอความช่วยเหลือ

ในส่วนของการแปลภาษามือแบบ 2 ระดับ นั้นจะมีรูปแบบของคำในการแปลคำให้เป็นความหมายที่ 2 แสดงดังภาพประกอบ 14(a) นั้นหมายถึง จะต้องแปลตัวอักษรที่ 1 และต่อด้วยตัวอักษรที่ 2 จึงจะสามารถระบุความหมายของตัวอักษรได้ โดยการทดลองการแปลภาษามือแบบ 2 ระดับจากตัวอักษร ก เป็นตัวอักษร ข แสดงดังภาพประกอบ 14(b)

ภาพประกอบ 15(a) แสดงให้เห็นถึงการแปลภาษาเมื่อภาษาไทยจากตัวอักษร ฟ เป็นตัวอักษร ฝ และภาพประกอบ 15(b) แปลจากตัวอักษร ย เป็นตัวอักษร ญ และภาพประกอบ 16 เป็นการทดลองการแปลภาษาเมื่อแบบ 3 ระดับ โดยเป็นการแปลคำภาษาอังกฤษ ซึ่งแปลว่า Help Me หรือ สัญญาณขอความช่วยเหลือ

4.3 ผลการทดลอง

ในงานวิจัยฉบับนี้ได้ทดสอบความถูกต้องของการแปลภาษามือโดยกำหนดให้มีผู้ทดสอบโปรแกรมจำนวน 3 คน และทำการทดสอบสัญลักษณ์ละ 5 ครั้ง จากการทดลองพบว่าการแปลภาษามือภาษาไทยจะมีความผิดพลาดมากกว่าภาษามือภาษาอังกฤษ เนื่องจากภาษาไทยจะมีรูปแบบของคำเป็นแบบ 2 ระดับ เช่น ตัวอักษร ฐ, ฑ, ป, ผ, ฝ, ภ และ ท สำหรับภาษามือภาษาอังกฤษนั้นจะมีแค่ระดับเดียวจึงทำให้การแปลภาษามือทำได้ถูกต้องมากกว่า ผลการทดลองแปลภาษามือภาษาไทยแสดงดังตาราง 1 ผลการทดลองแปลภาษามือภาษาอังกฤษแสดงดังตาราง 2 ผลการทดลองแปลตัวเลขอารบิกแสดงดังตาราง 3 และผลการทดลองแปลคำภาษาอังกฤษแสดงดังตาราง 4

จากตาราง 1-4 สัญลักษณ์ / หมายถึงการ
แปลภาษาถูกต้อง และ x หมายถึงการแปลภาษา
ผิดพลาด

ตาราง 1 ผลการทดลองการแปลภาษามือภาษาไทย

[illegible]

[illegible]

[illegible][illegible]

ตาราง 4 ผลการทดลองการแปลภาษามือคำภาษาอังกฤษ

คำ	คนที่ 1					คนที่ 2					คนที่ 3				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Backward	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Forward	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Left	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Love	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Right	/	x	x	/	/	/	/	/	/	/	/	/	/	/	/
Stop	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Like	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Dislike	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/

ตาราง 2 สรุปผลการทดสอบโปรแกรมของทั้ง 3 คน

คนที่ 1	คนที่ 2	คนที่ 3
95.58%	93.82%	96.76%
เฉลี่ยทั้งหมด	95.39%	

ตาราง 2 แสดงให้เห็นว่าเมื่อทดสอบโปรแกรมจากผู้ทดสอบโปรแกรมทั้ง 3 คนพบว่าโปรแกรมมีความถูกต้องในการแปลภาษามือเฉลี่ยที่ 95.39%

5. สรุปผลการดำเนินการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการดำเนินการวิจัย

ระบบอัจฉริยะสำหรับแปลภาษามือให้เป็นภาษาไทยนี้สามารถตรวจจับจุดต่างๆ บนฝ่ามือได้ทั้ง 2 ข้างแบบ real-time แต่จะสามารถระบุเป็นคำต่างๆ ได้แค่ครั้งละ 1 ข้างเท่านั้นนอกจากนี้ยังสามารถตรวจจับจุดบนฝ่ามือของคนอื่นได้ดี และสามารถตรวจจับจุดบนฝ่ามือได้ในระยะประมาณ 350 เซนติเมตรและเปลี่ยนฉากหลังก็ยังสามารถตรวจจับจุดบนฝ่ามือได้ แต่ในพื้นที่ที่มีแสงน้อยจะไม่สามารถตรวจจับจุดบน

ฝ่ามือได้ โดยระยะที่เหมาะสมที่สุดที่จะตรวจจับและแปลภาษามือคือระยะ 50 เซนติเมตร จากผลการทดลองแต่ละตัวอักษรมีความแม่นยำถูกต้องโดยเฉลี่ยที่ 95.39% แต่จะต้องวางมือในตำแหน่งที่ถูกต้องตามเงื่อนไขที่ได้ตั้งไว้เท่านั้น

5.2 อภิปรายผล

ระบบไม่สามารถตรวจจับจุดบนฝ่ามือในที่ที่มีแสงน้อยได้ และท่าทางของคำต่างๆ นั้นไม่สามารถย้ายมือได้เนื่องจากจุดบนฝ่ามือที่ใช้ MediaPipe Hands นั้นจะใช้จุด X, Y และ Z เดียวกัน และเงื่อนไขในแต่ละคำนั้นก็ถูกกำหนดให้อยู่ในมือข้างใดข้างหนึ่ง และในพื้นที่ที่มีแสงน้อยจะไม่สามารถตรวจจับจุดบนฝ่ามือและแปลเป็นคำได้

เนื่องจากในการกำหนดคำแต่ละคำนั้นจะต้องใช้เงื่อนไขในการกำหนดหลายๆ จุดพร้อมกัน เพื่อ

ให้เกิดเป็นคำขึ้นมา บางครั้งอาจจะมีคำบางคำหรือจุดบางจุดนั้นเกิดการซ้ำกันของจุดทำให้ต้องมาเชื่อกว่าจุดตรงไหนซ้ำกันกับตรงไหนจึงทำให้เกิดความสับสนในจุดบนฝ่า ในการตรวจจับจุดบนฝ่ามือนั้นในบางครั้งก็ไม่เจอจึงต้องพลิกมือหรือหมุนมก้องในการตรวจจับจุดบนฝ่ามือให้เจอมือ

5.3 ข้อเสนอแนะ

ในส่วนของการแปลค่านั้นคำบางคำมีความแม่นยำที่คลาดเคลื่อนเนื่องจากการกำหนดจุดบนฝ่ามือนั้นไม่ละเอียดเพียงพอ ควรกำหนดจุดบนฝ่ามือให้มีความแม่นยำมากกว่านี้ นอกจากนี้ยังมีท่าทางบางท่าทางที่ต้องใช้ 2-3 ท่าทางในการแปลความหมายให้เป็น 1 คำ แต่ระบบยังทำท่าทางได้แค่ท่าทางเดียวเท่านั้น รวมทั้งท่าทางที่เป็นประโยชน์ด้วย

สำหรับงานวิจัยในอนาคต ผู้วิจัยจะใช้ MediaPipe Hands ในการตรวจจับบริเวณที่เป็นมือทั้ง 2 ข้าง และจะตัดเฉพาะ (Crop) บริเวณมือทั้งสองข้างไปสร้างเป็นชุดข้อมูลภาษามือภาษาไทย (Thai Hand Sign Dataset) และจะนำข้อมูลไปเรียนรู้ (Train) ด้วยการเรียนรู้เชิงลึก (Deep Learning) เช่น โครงข่ายประสาทแบบคอนโวลูชัน (Convolutional Neural Network) (Saisangchan, Chamchong, & Suwannasa, 2022; Lapthanachai et al., 2023)

เอกสารอ้างอิง

Agency, A. (2021). *Viral hand sign leads Spanish authorities to detain domestic abuser*. Retrieved 2 March 2023. Retrieved from <https://www.dailysabah.com/world/europe/viral-hand-sign-leads-spanish-authorities-to-detain-domestic-abuser>

Camgz, N. C., Koller, O., Hadfield, S., & Bowden, R. (2020). Sign language transformers: Joint end-to-end sign language recognition and translation. In *the Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, pp. 10023-10033. IEEE. <https://doi.org/10.1109/CVPR42600.2020.01004>

Chumchim, P. & Maneerat, P. (2020). Application development of sign language translation systems for the hearing-impaired people. *PKRU SciTech Journal*, 4(1), 22–32. <https://ph01.tci-thaijo.org/index.php/pkrusciotech/article/view/240583> [in Thai]

Daily News Online. (2021). *September 23, International day of sign languages [23 ก.ย. วันภาษามือโลก]*. Retrieved 21 February 2023. Retrieved from <https://www.dailynews.co.th/articles/289869/> [in Thai]

Department of Empowerment of Persons with Disabilities. (2022). *Report on the situation of persons with disabilities in Thailand [รายงานข้อมูลสถานการณ์ด้านคนพิการในประเทศไทย]*. Retrieved 2 May 2023. Retrieved from https://dep.go.th/images/uploads/files/situation_sep65.pdf [in Thai]

Department of Provincial Administration. (2022). *Population statistics from the civil registration (Monthly) [สถิติประชากรทางการทะเบียนราษฎร (รายเดือน)]*. Retrieved 1 May 2023. Retrieved from <https://stat.bora.dopa.go.th/stat/statnew/statMONTH/statmonth/#/view> [in Thai]

- Disabled World. (2022). *Deaf communication: Sign language and assistive hearing devices*. Retrieved 2 May 2023. Retrieved from <https://www.disabled-world.com/disability/types/hearing/communication>
- Grishchenko, I. & Bazarevsky, V. (2022). *MediaPipe holistic - Simultaneous face, hand and pose prediction, on device*. Retrieved 25 February 2023. Retrieved from <https://blog.research.google/2020/12/mediapipe-holistic-simultaneous-face.html?m=1>
- Jeekratok, K., Tongthaku, P., & Jampamoon, A. (2019). Thai sign language system in daily Life. *Udon Thani Rajabhat University Journal of Science and Technology*, 7(1), 101-113. <https://ph01.tci-thaijo.org/index.php/scudru/article/view/181628> [in Thai]
- Lapthanachai, N., Chomthong, A., Waijanya, S., & Promrit, N. (2023). Classification of nail abnormalities using convolutional neural network. *Journal of Applied Informatics and Technology*, 5(1), 18–35. <https://doi.org/10.14456/jait.2023.2> [in Thai]
- MediaPipe. (2023). *Hand landmarks detection guide*. Retrieved 20 February 2023. Retrieved from https://developers.google.com/mediapipe/solutions/vision/hand_landmarker
- Nadee, C. & Yingkayun, K. (2020). Development of method for categorization on sign language picture using image plane adjustment technique. *RMUTL Engineering*, 5(1), 25–34. <https://doi.org/10.14456/rmutlengj.2020.4> [in Thai]
- Prajan, S. (2022). *Thai fingerspelling in sign language [แบบสะกดนิ้วมือไทยในภาษามือ]*. Retrieved 20 February 2023. Retrieved from <https://www.gotoknow.org/posts/604277> [in Thai]
- Saisangchan, U., Chamchong, R., & Suwannasa, A. (2022). Analysis of lime leaf disease using deep learning. *Journal of Applied Informatics and Technology*, 4(1), 71-86. <https://doi.org/10.14456/jait.2022.6> [in Thai]
- Sriboonruang, P., Tananchai, P., Khaggathog, K., Vijitkunsawat, W., & Anunvrapong, P. (2022). Software sign language translator to text and speech by using landmarks technique of MediaPipe. *The Journal of Industrial Technology: Suan Sunandha Rajabhat University*, 10(2), 66–76. <https://ph01.tci-thaijo.org/index.php/fit-ssru/article/view/249506> [in Thai]
- Wadhawan, A. & Kumar, P. (2020). Deep learning-based sign language recognition system for static signs. *Neural Computing and Applications*, 32, 7957–7968. <https://doi.org/10.1007/s00521-019-04691-y>