

การพัฒนาวิธีการแยกแยะรูปภาพภาษามือ โดยอาศัยวิธีการปรับระนาบภาพ Development of Method for Categorization on Sign Language Picture using Image Plane Adjustment Technique

โชคมงคล นาดี้¹ และกฤษดา ยิงขยัน^{1*}

^{1,2} สาขาวิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา

128 ถนนห้วยแก้ว อำเภอเมือง จังหวัดเชียงใหม่ 50300

Chokemongkol Nadee¹ and Krisda Yingkayun^{2*}

^{1,2} Department of Electrical Engineering, Faculty of Engineering, Rajamangala University of Technology Lanna

128 Huay Kaew Road, Muang, Chiang Mai, Thailand, 50300

* ผู้รับผิดชอบบทความ: krisda@rmutl.ac.th เบอร์โทรศัพท์ 0-5392-1444 ต่อ 314

Received: 2 March 2020, Revised: 27 May 2020, Accepted: 31 May 2020

บทคัดย่อ

การแยกแยะรูปภาพโดยส่วนใหญ่อาศัยวิธีการตัดสินใจอย่างใดอย่างหนึ่งหรือใช้กระบวนการหลายอย่างร่วมกัน เช่นการปรับภาพก่อนเข้าร่วมกับวิธีการของแมทซ์ซึนเลนนิ่ง เพื่อกระบวนการดึงจุดเด่นหรือน้ำหนักของภาพให้ได้ค่าที่ดีที่สุดสำหรับการตัดสินใจ ในบทความวิจัยนี้ได้นำเสนอวิธีการแยกแยะรูปภาพภาษามือโดยวิธีการแบบหลายแมทซ์ซึนเลนนิ่ง การแยกแยะรูปภาพภาษามือที่ใช้ในงานวิจัยมีการเลือกพื้นหลังหรือสิ่งแวดล้อมในหลายรูปแบบเช่น พื้นหลังสีดำ พื้นหลังสีขาว พื้นหลังแบบมีลวดลาย โดยที่ทั้งรูปภาพแบบขาวดำและรูปภาพสี RGB โดยใช้รูปภาพของภาษามือจากฐานข้อมูลของ ASL จากการทดลองพบว่าวิธีการแบบหลายแมทซ์ซึนเลนนิ่งมีการแยกแยะรูปภาพภาษามือได้ถูกต้องถึง 98.06 เปอร์เซ็นต์

คำสำคัญ การแยกแยะรูปภาพ แมทซ์ซึนเลนนิ่ง ภาษามือ

Abstract

Most image classification is based on applying decision-making algorithms, in conjunction with image preprocessing, in order to extract distinctive features from the images and compute the most appropriate weights used in the decision making. This paper presents a method to classify sign language images by using multiple machine learning methods. The sign language images used in this work are obtained from the American Sign Language (ASL) database, which contains both black and white and color images (RGB). Experiments have shown that the multiple machine learning method is able to correctly classify sign language images with 98.06 percent accuracy.

Keyword: Classification, Machine Learning, Sign Language

1. บทนำ

ในปี ค.ศ. 2019 จำนวนประชากรของโลกมีประมาณจำนวน 7.7 พันล้านคน [1] และมีแนวโน้มเพิ่มมากขึ้น 8.5 พันล้านคนในปี ค.ศ. 2030 และคาดการณ์ว่าจะมีเพิ่มขึ้นถึง 10.9 พันล้านคนในปี ค.ศ. 2050 โดยจำนวนของผู้พิการทางการได้ยินอยู่ประมาณ 466 ล้านคน [2] และมีแนวโน้มเพิ่มขึ้นประมาณ 900 ล้านในปี ค.ศ. 2050 สำหรับประเทศไทยในปี พ.ศ. 2561 มีประชากรประมาณ 66 ล้านคน [3] ซึ่ง

เป็นจำนวนผู้พิการทางการได้ยินประมาณ 375,680 คน [4] ผู้พิการทางการได้ยินคือผู้ที่มีความบกพร่องในการได้ยินจึงไม่สามารถเลียนแบบเสียงที่ได้ยิน หรือแปลงเสียงพูดได้เหมือนคนปกติ ดังนั้นผู้พิการทางการได้ยินส่วนใหญ่ใช้ภาษามือในการติดต่อสื่อสารกันระหว่างผู้พิการทางการได้ยินกันเองและกับคนที่มีความสามารถในการแปลภาษามือ ซึ่งต้องมีการฝึกฝนเพราะภาษามือเป็นภาษาที่มีทั้งท่ามือ ตำแหน่งมือ การเคลื่อนไหวของมือ จึงทำให้มีความยากในการจดจำ ฉะนั้นจึง

ต้องมีระบบในการแปลภาษามือเพื่อช่วยให้สามารถสื่อสารกันได้ ภาษามือที่ผู้พิการทางการได้ยินในแต่ละประเทศใช้นั้นจะมีความแตกต่างกันเช่นเดียวกับภาษาพูด ซึ่งภาษามือของแต่ละประเทศก็ไม่เหมือนกัน แม้กระทั่งในประเทศเดียวกันแต่คนละภาคก็ยังไม่เหมือนกัน

ระบบการแปลภาษามือในอดีตนั้นมีหลายเทคนิคในการแปลภาษามือโดยส่วนใหญ่แบ่งออกเป็นสองกลุ่ม 1) กลุ่มภาพจริง 2) กลุ่มกราฟฟิค ซึ่งข้อดีและข้อเสียต่างกันเพราะถ้าใช้ในส่วนของภาพจริง อาจจะมีมุมมองของภาพที่มีการบิดเบือนของภาพสภาพแวดล้อมของแสงรวมถึงพื้นหลังทำให้การแปลภาษามือผิดพลาดได้ ซึ่งต้องมีการใช้เทคนิคในการปรับภาพก่อนเข้าเพื่อใช้ในการประมวลผล ส่วนการใช้กราฟฟิคส่วนใหญ่จะได้อารมณ์แบบท่าทาง การเคลื่อนไหวที่ใกล้เคียงกับภาษามือที่แท้จริง ทำให้การสื่อสารง่ายขึ้นและมีความถูกต้องแต่ยังมีความยากในส่วนของการควบคุมไวยากรณ์ ทำให้การแปลภาษามือที่ซับซ้อนอาจจะยังไม่ถูกต้องเท่าที่ควร

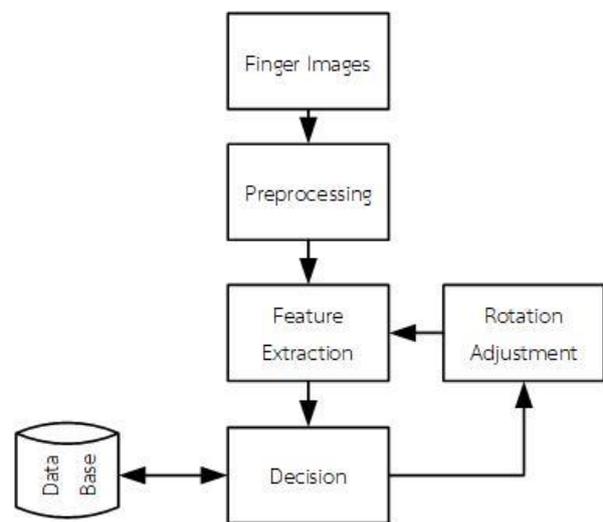
งานวิจัยนี้ได้นำเสนอการแยกแยะรูปภาพของภาษามือซึ่งเป็นกระบวนการดึงน้ำหนักของภาพโดยวิธีการของแบบหลายแมทซ์ซิง แล้วนำค่าน้ำหนักเหล่านั้นมาทำการตัดสินใจว่าตรงกับตัวอักษรใดในภาษามือ เมื่อเปรียบเทียบข้อดีข้อด้อยของระบบนี้ พบว่าข้อดีของระบบนี้คือความถูกต้องเพิ่มมากขึ้น เพราะใช้เทคนิคหลายแมทซ์ซิงในส่วนข้อดีคือความเร็วของระบบอาจจะช้าลงเพราะใช้การประมวลผลหลายอย่างในเวลาเดียวกัน

2. ทฤษฎีและวิธีการดำเนินการวิจัย

งานวิจัยที่เกี่ยวข้องกับการพัฒนาวิธีการแยกแยะรูปภาพภาษามือ สามารถแบ่งออกเป็น 2 กลุ่ม คือ 1) ระบบการอ่านท่าทางจากเซ็นเซอร์ โดยเครื่องตรวจจับจะถูกติดตั้งที่ข้อต่อและตำแหน่งที่สำคัญของมือและนิ้วมือ เซ็นเซอร์จะทำการรวบรวมข้อมูลตำแหน่งและข้อมูลจะถูกประมวลผลเพื่อตีความสัญญาณ วิธีการเหล่านี้อาจให้ตำแหน่งที่ต้องการ แต่อาจมีข้อจำกัดบางประการ ตามลักษณะของผู้ใช้งาน [5-7] และ 2) ระบบการอ่านท่าทางจากกล้องหรือการมองเห็น ซึ่งแบ่งเป็น 2 กลุ่ม ได้แก่ 1) เป็นกลุ่มที่อาศัยมือเปล่าในการวิเคราะห์ [8-10] โดยแยกออกเป็นกลุ่มที่ใช้หน้ามือและหลังมือ 2) เป็นกลุ่มที่อาศัยการสวมถุงมือ [11-12] ระบบการแปลภาษามืออัตโนมัติส่วนใหญ่ใช้การแยกแยะรูปแบบของท่าทางโดยอาศัยวิธีการแมทซ์ซิงในหลาย ๆ เทคนิค เช่น Canonical correlation analysis (CCA), Convolutional neural network (CNN) และอีกหลายเทคนิค เป็นต้น

วิธีการปรับปรุงระบบอัตโนมัติในการแปลภาษามือ ในงานวิจัยนี้ได้คำนึงถึงปัญหาที่เกิดขึ้นจากการวิจัยที่ผ่านมาสรุป

ปัญหาได้ 3 รูปแบบหลัก คือ 1) ผลกระทบและการเปลี่ยนแปลงของแสงต่อภาพมือ 2) ผลกระทบและการเปลี่ยนแปลงท่าทางและการหมุนของภาพมือ 3) ผลกระทบการเปลี่ยนแปลงของภาพมือที่เกิดจากการสูญเสียจากการวิเคราะห์ภาพโดยใช้เทคนิคหลายแมทซ์ซิง ผลของการวิเคราะห์จากระบบจะได้เป็นค่าน้ำหนักของตัวอักษรและตัวเลข จากลักษณะของน้ำหนักของภาพแต่ละภาพในชุดของตัวอักษรเดียวกัน ลักษณะของน้ำหนักจะคล้ายกัน เพื่อบ่งชี้ว่าข้อมูลลักษณะที่ถอดออกมาจากภาพตรงกับข้อมูลลักษณะภาพมือในฐานข้อมูล จากปัญหาข้างต้นงานวิจัยนี้จึงได้รวมวิธีการวิเคราะห์ลักษณะโดยรวมของภาพมือ และได้เพิ่มวิธีการตรวจวัดลักษณะของการเอียงของภาพมือเพื่อแก้ส่วนของปัญหาที่เกิดขึ้นในกรณีที่มีการเปลี่ยนแปลงลักษณะของภาพมือที่มีแนวเอียงเกิดขึ้น และวิธีการวิเคราะห์ส่วนประกอบหลัก ซึ่งการวิเคราะห์ทั้งสองจะทำการวิเคราะห์องค์ประกอบที่สัมพันธ์กันคือ ภาพมือจะถูกวิเคราะห์ลักษณะการเอียงของภาพมือ ถ้ามีการเอียงของภาพมือแล้วจะส่งภาพที่ได้นั้นเข้าสู่ระบบปรับภาพเพื่อหาระนาบที่ถูกต้องในการปรับภาพ จากนั้นเมื่อผ่านระบบในการปรับภาพแล้ว ภาพที่ถูกปรับแล้วนั้นจะถูกทำการวิเคราะห์โดยวิธีการวิเคราะห์ส่วนขององค์ประกอบหลัก จากโครงสร้างของระบบการแปลภาษามือ อธิบายรายละเอียดดังต่อไปนี้ ดังรูปที่ 1



รูปที่ 1 แผนภาพระบบแปลภาษามือ

2.1 ทฤษฎีพื้นฐานที่เกี่ยวข้อง

การประมวลผลภาพกับรูปร่างและโครงสร้างของภาพ (Morphological image processing) การประมวลผลภาพกับรูปร่างและโครงสร้างของภาพ คือ กระบวนการประมวลผลภาพโดยการเปลี่ยนแปลงลักษณะรูปร่างหรือโครงสร้างของภาพ โอเปอเรชันพื้นฐานโดยทั่วไปได้แก่ การขยายภาพ



(Dilation) การย่อภาพ (Erosion) และ การหาโครงสร้างหลักของภาพ (Skeleton) โดยการขยายภาพ คือ การขยายภาพโดยมีสัดส่วนเท่ากันทั่วทั้งภาพ (Uniform) การย่อภาพคือการย่อภาพ ส่วนการทำ การหาโครงสร้างของภาพเป็นกระบวนการหาโครงสร้างหลักของวัตถุ โดยจะกล่าวรายละเอียดต่อไปตามหัวข้อ นอกจากนี้ยังมีวิธีการใช้โอเพอเรชันอื่นๆ นี้ได้แก่การ Opening และ Closing เป็นต้น

2.1.1 แมทซึนลินนิ่ง

ลักษณะการเรียนรู้โดยวิธีการของแมทซึนลินนิ่งสามารถแบ่งออกได้เป็นสามกลุ่มดังนี้

- 1) การเรียนรู้แบบมีผู้สอน (Supervised learning)
- 2) การเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning)
- 3) การเรียนรู้แบบการเสริมแรงใจ (Reinforcement learning)

มนุษย์มีความสามารถในการเรียนรู้จำสิ่งต่าง ๆ จำแนกได้เป็นหลายลักษณะการเรียนรู้ เช่นการเรียนรู้โดยมีผู้สอน การเรียนรู้จากประสบการณ์ ด้วยตนเอง เป็นต้น ผลลัพธ์จากการเรียนรู้ก็คือ พฤติกรรมที่เปลี่ยนไป ความรู้ที่ได้ถูกนำไปประกอบการตัดสินใจในการปฏิบัติต่อปัญหาต่าง ๆ ที่พบ การเรียนรู้เป็นสิ่งที่มนุษย์รับรู้ถึงสิ่งต่าง ๆ ที่ เกิดขึ้นรอบ ๆ ตัว โดยผ่านกระบวนการของการรับความรู้ ซึ่งสภาวะแวดล้อมต่างเป็นตัวกระตุ้นให้มนุษย์มีการตอบสนองทางพฤติกรรมต่าง ๆ ความรู้ของมนุษย์ถูกจัดเก็บไว้เป็นความจำและสามารถเรียกกลับมาใช้ใหม่ได้ นักจิตวิทยาได้ศึกษาถึงการจดจำของมนุษย์ และพบว่า มนุษย์มีการเรียนรู้เป็นแบบไม่มีความสัมพันธ์ (Non-associative learning) และการเรียนรู้แบบมีความสัมพันธ์ (Associative learning) และความจำของมนุษย์ แบ่งออกได้เป็น ความจำระยะสั้น (Short-term memory: STM) และ ความจำระยะยาว (Long-term memory: LTM) สำหรับโครงข่ายประสาทเทียมซึ่งเป็นการจำลองสมองมนุษย์ จะต้องมีการเรียนรู้เช่นกัน การเรียนรู้ของโครงข่ายประสาทเทียมจำแนกออกเป็น 3 กลุ่ม ได้แก่ 1) การเรียนรู้แบบมีผู้สอน (Supervised learning) 2) การเรียนรู้ แบบไม่มีผู้สอน (Unsupervised learning) และ 3) การเรียนรู้แบบการเสริมแรงใจ (Reinforcement learning)

2.1.2 SVM (Support vector machine)

SVM เป็นอัลกอริทึมในการตัดแยกที่มีการนำมาใช้กันอย่างกว้างขวางในด้านการประมวลผลเป็นภาพดิจิทัล หลักการของ SVM คือการให้อินพุตที่ใช้ฝึกเป็นเวกเตอร์ในสเปซ N มิติ เช่นถ้าในกรณีของ 2 มิติ และ 3 มิติ จะเป็นจุดที่อยู่ในระนาบ xy และสเปซ xyz ตามลำดับ จากนั้นทำการสร้างไฮเปอร์เพลน (Hyperplane) ที่จะแยกกลุ่มของเวกเตอร์

อินพุตออกเป็นประเภทต่างๆ ในกรณีที่ เป็น 2 มิติ และ 3 มิติ ไฮเปอร์เพลน คือเส้นตรงและระนาบตามลำดับ ข้อเด่นของ SVM จะทำการเก็บแมพ (Map) เวกเตอร์ ในสเปซอินพุตให้เข้าสู่ Feature Space โดยใช้ฟังก์ชันหรือเรียกว่าเคอร์เนล (Kernel) ชนิดต่าง ๆ เช่น โพลีโนเมียล (Polynomial) เรเดียล (Radial) เป็นต้น ใน Feature Space ดังกล่าวเวกเตอร์อินพุตสามารถแยกประเภทได้โดยไฮเปอร์เพลน [5]

2.1.3 ต้นไม้ตัดสินใจ (Decision tree)

ต้นไม้ตัดสินใจเป็นเทคนิคที่ช่วยในการตัดสินใจโดยใช้วิธีการโครงสร้างของกราฟที่มีรูปร่างเหมือนกับต้นไม้ ทำให้สามารถแปลความหมายเข้าใจง่ายมากขึ้นเพื่อใช้ในการแยกแยะทั้งแบบสองทาง (Binary) และปัญหาแบบหลายคลาส (Multiple classes) [13]

2.1.4 การวิเคราะห์การถดถอยโลจิสติก

การวิเคราะห์การถดถอยโลจิสติกแบ่งออกเป็น 2 ประเภท ได้แก่ (1) การวิเคราะห์การถดถอยโลจิสติกทวิ ใช้กับตัวแปรเกณฑ์ที่แบ่งออกเป็น 2 กลุ่มย่อย เช่น 1) กลุ่มที่ปรากฏเหตุการณ์ที่สนใจ มีค่าเป็น 1 กับกลุ่มที่ไม่ปรากฏเหตุการณ์ที่สนใจ มีค่าเป็น 0 และ 2) การวิเคราะห์การถดถอยโลจิสติกพหุกลุ่ม ใช้กับตัวแปรเกณฑ์ที่มีหลายกลุ่มย่อยหรือมีมากกว่า 2 กลุ่มย่อย เช่น ขนาดสถานศึกษา แบ่งเป็นกลุ่มขนาดใหญ่/กลาง/เล็ก หรือโรงพยาบาลมีมาตรฐานการให้บริการสูง/ปานกลาง/ต่ำการวิเคราะห์การถดถอยโลจิสติก ความสัมพันธ์ระหว่างตัวแปรทำนายกับตัวแปรเกณฑ์จึงไม่เป็นความสัมพันธ์เชิงเส้นในการวิเคราะห์ จะต้องมีการปรับให้ความสัมพันธ์อยู่ในรูปเชิงเส้น ในรูปของ odds และในการเขียนโมเดลโลจิสติก จะต้องเขียนให้อยู่ในรูป log ของ odds เรียกว่า logit [14]

2.1.5 โครงข่ายประสาทเทียม

โครงข่ายประสาทเทียม (Artificial Neuron Network : ANN) เป็นศาสตร์แขนงหนึ่งของทางด้านปัญญาประดิษฐ์ (Artificial Intelligence : AI) มีรูปแบบโครงสร้างและการทำงานของการทำงานที่ประมวลผลเหมือนกับสมองของสิ่งมีชีวิตซึ่งมีปรับเปลี่ยนตัวเองต่อการตอบสนองของอินพุตตามกฎของการเรียนรู้ (Learning rule) หลังจากที่โครงข่ายได้เรียนรู้สิ่งที่ต้องการแล้ว โครงข่ายนั้นจะสามารถทำงานที่กำหนดไว้ได้ โครงข่ายประสาทเทียมได้ถูกพัฒนาคิดค้นจากการทำงานของสมองมนุษย์โดยสมองมนุษย์ประกอบไปด้วยหน่วยประมวลผลเรียกว่า นิวรอน (เซลล์ประสาท หรือ Neuron) จำนวน นิวรอนในสมองมนุษย์มีอยู่ประมาณและมีการเชื่อมต่อกันอย่างมากมาย สมองมนุษย์จึงสามารถกล่าวได้ว่าเป็น

คอมพิวเตอร์ที่มีการปรับตัวเอง (Adaptive) ไม่เป็นเชิงเส้น (Nonlinear) และทำงานแบบขนาน (Parallel) ในการดูแลจัดการการทำงานร่วมกันของนิวรอนในสมอง การคำนวณเชิงนิวรอลเป็นการคำนวณที่เลียนแบบมาจากการทำงานของสมองมนุษย์นั่นเอง [15]

2.1.6 KNN (K-Nearest neighbor)

เป็นวิธีการในการจัดแบ่งคลาส โดยเทคนิคแบบนี้จะตัดสินใจว่า คลาสใดสามารถแทนเงื่อนไขหรือมีกรณีใหม่ ๆ อย่างไรบ้าง โดยตรวจสอบจำนวนบางจำนวนในขั้นตอน KNN ของกรณีหรือเงื่อนไขที่เหมือนกันหรือใกล้เคียงกันมากที่สุด โดยจะหาผลรวม (Summary or count up) ของจำนวนเงื่อนไข หรือกรณีต่าง ๆ สำหรับแต่ละคลาส และกำหนดเงื่อนไขใหม่ ๆ ให้คลาสที่เหมือนกันกับคลาสที่ใกล้เคียงกันมากที่สุด

2.2 วิธีการดำเนินการวิจัย

ในการทดลองแบ่งออกเป็นขั้นตอนดังนี้

ขั้นตอนที่ 1 รับภาพอินพุต

ขั้นตอนที่ 2 ทำการปรับภาพก่อนเข้า เนื่องจากการรับภาพมาวิเคราะห์มีอยู่ 2 กลุ่มคือกลุ่มภาพขาวดำ และกลุ่มภาพสีแบบ RGB จึงต้องทำการแยกระดับของสี ถ้าในอยู่ในรูปแบบของภาพขาวดำต้องทำเป็นภาพเทา 256 ระดับ ส่วนภาพสีเลือกระดับแม่สี RGB 256 ระดับเพื่อใช้ในการแยกข้อมูล

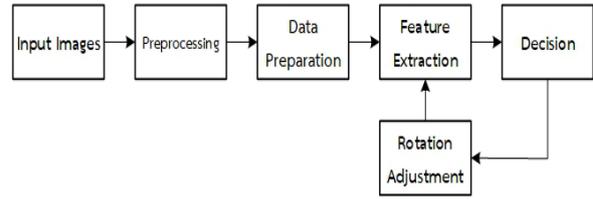
ขั้นตอนที่ 3 การเตรียมข้อมูล (Data preparation) เมื่อได้ชุดภาพภาษามือแล้ว ทำการกำหนด Labels ของชุดภาพภาษามือของแต่ละตัวอักษรโดยกำหนดชุดข้อมูลในการเรียนรู้อยู่ที่ 100 ภาพต่อตัวอักษร และชุดของการทดสอบอีกอย่างละ 72 ภาพตามสภาพแวดล้อมต่าง ๆ เช่น 1.พื้นหลังสีดำ 2.พื้นหลังสีขาว 3.พื้นหลังแบบมีลวดลาย 4. ภาพสีแบบ RGB

ขั้นตอนที่ 4 การหาค่าตัวถ่วงน้ำหนักโดยวิธีการของแมทซินเลนนิ่งแต่ละแบบนั้นจะต้องหาตัวถ่วงน้ำหนักเพื่อหาการเคลื่อนที่ของชุดข้อมูลว่าไปในทิศทางใด ถ้าค่า cost function ออกมามีค่าน้อยเกินไปก็ต้องกลับไปเข้าสู่วิธีการเรียนรู้หลายรอบ เพื่อปรับค่าให้ได้ค่าตัวถ่วงน้ำหนักที่ดีที่สุดในการนำไปใช้งาน

ขั้นตอนที่ 5 การตัดสินใจ (Decision) ขั้นตอนนี้จะเป็นการเปรียบเทียบข้อมูลที่นำมาทดสอบกับชุดข้อมูลที่ทำกรเรียนรู้แล้ว จึงได้ผลลัพธ์จากแมทซินเลนนิ่งหลาย ๆ แบบส่วนที่มีการผิดพลาดจากการตัดสินใจ

ขั้นตอนที่ 6 การปรับระนาบของภาพของมือใหม่ ในกรณีที่มีความผิดพลาดอยู่จึงนำภาพที่ผิดพลาดมาหาจุด

ซ้ายขวาของขอบมือเพื่อหามุมในการปรับระนาบของภาพใหม่ จากนั้นนำภาพที่ปรับมุมใหม่เข้าสู่ขั้นตอนที่ 4 เพื่อทำการตัดสินใจอีกครั้งขั้นตอนการทำงานของระบบแสดงดังรูปที่ 2



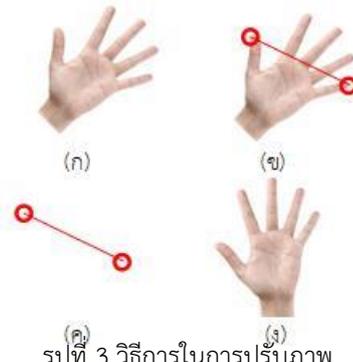
รูปที่ 2 ขั้นตอนการทำงานของระบบแปลภาษามือ

2.2.1 การรับภาพอินพุต (Input images)

ในการรับภาพที่ใช้เป็นอินพุตจะทำการอ่านไฟล์ภาพเข้ามาด้วยโปรแกรมสำเร็จรูปโดยไฟล์ภาพที่อ่านเข้ามาเป็นภาพสีขนาด 105x122 จุดภาพ โดยมีตัวอักษรภาษาอังกฤษทั้งหมด 24 ตัวโดยเลือกรูปภาพตัวอักษรมาอย่างละ 100 ภาพสำหรับการเรียนรู้โดยแบ่งเป็นกลุ่มที่มีสีพื้นหลังแบบสีเดียว และกลุ่มที่มีพื้นหลังมีลวดลายโดยจะส่งผ่านเป็นภาพอินพุตให้กับวิธีการวิเคราะห์หาจุดเด่นและกระบวนการเทียบน้ำหนักจากระบบแมทซินเลนนิ่ง

2.2.2 การวิเคราะห์ส่วนเอียงของรูป

กระบวนการในการนำภาพมือเพื่อใช้ในการทดลอง อาจจะมีการเปลี่ยนแปลงของท่าทางของมือได้ ทำให้ข้อมูลที่ใช้วิธีการวิเคราะห์แบบจับคู่แม่แบบ และกระบวนการเทียบน้ำหนักจากระบบแมทซินเลนนิ่งแล้วนั้นอาจเกิดความผิดพลาดได้ จึงต้องมีส่วนของการวิเคราะห์การเอียงโดยวิเคราะห์ด้วยวิธีการของราดอนทรานฟอร์ม แล้วนำไปหาจุดอ้างอิงจากรอบของรูปมือเพื่อหาระนาบของมุมที่ถูกต้องเพื่อใช้ในการตัดสินใจและการปรับมุมให้ถูกต้องเพื่อนำภาพที่ได้ไปวิเคราะห์ได้อย่างถูกต้องดังแสดงแนวความคิดการปรับแก้ดังรูปที่ 3



รูปที่ 3 วิธีการในการปรับภาพ

2.2.3 การวิเคราะห์ภาพ



ในขั้นตอนของการวิเคราะห์ส่วนประกอบหลักนั้น เป็นขั้นตอนต่อจากการวิเคราะห์การเอียงของภาพมือ โดยเมื่อการปรับภาพให้อยู่ในระนาบตรงแล้วนั้น ภาพที่ได้จะถูกส่งผ่านมาวิเคราะห์หาค่าน้ำหนักของตัวอักษร ซึ่งสาเหตุที่ใช้การวิเคราะห์การหาน้ำหนักของตัวอักษรนี้ เนื่องจากเป็นวิธีการที่เร็วและสามารถลดขนาดของข้อมูลลงได้มาก โดยที่ยังสามารถอธิบายลักษณะทั้งหมดของภาพตัวอักษรที่คล้ายกันได้ อย่างชัดเจนโดยที่ภาพอินพุตจะถูกนำมาหาความแตกต่างจากภาพเฉลี่ยซึ่งถือว่าเป็นภาพที่แสดงลักษณะของแต่ละตัวอักษรแล้ว นำมาค่าที่หาได้นำไปทำการแยกแยะต่อไป

2.2.4 การแยกแยะและการตัดสินใจ

ในขั้นตอนนี้การตัดสินใจเพื่อพิจารณาว่าเวกเตอร์น้ำหนักที่ใช้แทนภาพตัวอักษร จะอยู่ในขอบเขตการตัดสินใจของการแยกแยะ โดยใช้การวัดจากน้ำหนักโดยแบ่งเป็นช่วงของการยอมรับโดยที่ค่าตัดสินใจน้อยกว่า ค่าตัดสินใจที่เกิดจากการประมาณค่าโดยใช้เส้นโค้งของความผิดพลาดที่ตำแหน่งความผิดพลาดร้อยละ 1 จากกราฟ False Acceptance Error และช่วงที่ไม่แน่ใจอยู่ระหว่างค่าตัดสินใจแรก และค่าตัดสินใจที่ False Acceptance Error มีค่าเท่ากับ False Rejection Error ซึ่งข้อมูลที่อยู่ในช่วงของความไม่แน่ใจนี้ จะนำไปทำการแยกแยะโดยการจับคู่แบบรูป (Template matching) อีกครั้ง โดยที่แบบรูปที่นำมาใช้ ได้มาจากการคำนวณเวกเตอร์น้ำหนักของแต่ละตัวอักษรจากชุดภาพที่นำมาใช้สร้างฐานข้อมูล เมื่อผ่านกระบวนการจับคู่แบบรูปเพื่อแยกแยะตัวอักษรอีกครั้งแล้ว จะทำให้การตัดสินใจมีความถูกต้องเพิ่มมากยิ่งขึ้น

3. ผลการวิจัยและอภิปราย

การทดลองในงานวิจัยนี้ได้ใช้รูปภาพภาษามือโดยแบ่งออกเป็น 24 ตัวอักษรในภาษาอังกฤษเว้นตัว j กับตัว z เพราะมีความใกล้เคียงกันกับตัวอักษรอื่น โดยใช้องค์ประกอบด้วยภาพระดับ สีแดง สีเขียว สีน้ำเงิน 256 ระดับ และภาพที่ใช้สำหรับการสอนให้ระบบการแปลภาษามือที่ใช้สำหรับแมชชีนเลนนิ่งโดยแบ่งออกเป็น ตัวอักษรภาษาอังกฤษ 24 ตัว มีขนาดภาพ 106x119 จุดภาพ โดยให้แต่ละตัวอักษรมี 100 ภาพ ต่อหนึ่งตัวอักษร การทดลองเริ่มด้วยการนำภาพอินพุตที่เป็นภาพ มาทำการสอนโดยวิธีการแมชชีนเลนนิ่ง หลายแบบ โดยใช้วิธีการ 1. KNN 2. SVM 3. Neural Network 4. การวิเคราะห์การถดถอยเชิงเส้น 5. การวิเคราะห์การถดถอยโลจิสติก จากนั้นนำภาพมาคำนวณหาเวกเตอร์น้ำหนักด้วยกระบวนการของแมชชีนเลนนิ่งในสภาพแวดล้อมแบบต่าง ๆ จากนั้นทำการทดสอบในสภาพแวดล้อมต่าง ๆ กันเช่นการเปลี่ยนแปลงของแสงสว่าง ทำทาง ขนาด ภาพที่มีการเอียง และภาพที่มีการบดบังเป็นบางส่วนซึ่งเปรียบการสูญหายของ

ภาพภาษามือ โดยนำส่วนที่ตัดสินใจผิดพลาดกลับปรับระนาบที่ถูกต้องใหม่แล้วกลับไปเข้าสู่กระบวนการของแมชชีนเลนนิ่งใหม่ บนความเปลี่ยนแปลงของข้อมูลที่แตกต่างกันคือ ภาพที่ถูกกำหนดให้เห็นลักษณะของภาพภาษามือชัดเจนจากชุดภาพที่ใช้สอน และชุดภาพภาษามือที่ใช้ทดสอบ ภาพที่มีการบดบัง โดยรายละเอียดเงื่อนไขของระบบแยกแยะที่ใช้ทดสอบแสดงดังต่อไปนี้

จากการออกแบบการทดลอง ได้กำหนดรูปภาพของมือจากฐานข้อมูลของ ASL จำนวน 24 ตัวอักษร โดยจัดทำทางในแบบต่าง ๆ ดังแสดงในรูปที่ 4 เพื่อใช้ในการเรียนรู้ ส่วนรูปที่ 5 เป็นภาพของมือที่นำมาทดสอบโดยมีการเปลี่ยนแปลงของพื้นหลังหรือสภาพแวดล้อมแบบต่าง ๆ เช่น รูปที่ 5 (ก) เป็นภาพมือที่มีพื้นหลังหลากหลายเช่น พื้นหลังสีขาว พื้นหลังสีดำ พื้นหลังที่มีลวดลาย รูปที่ 5 (ข) เป็นรูปภาพภาษามือที่มีการดึงเฉพาะมือออกมาเป็นแบบขาวดำในพื้นหลังสีขาว ส่วนรูปที่ 5 (ค) เป็นรูปภาพภาษามือที่มีการดึงเฉพาะมือออกมาในรูปแบบภาพสีระดับเทา ในพื้นหลังสีขาว และรูปที่ 5 (ง) เป็นรูปของภาษามือที่มีการบดบังของภาพหรือการสูญเสียข้อมูล

ตารางที่ 1 แสดงรายละเอียดของระบบที่ใช้ทดสอบ

Micro-computer	Processor	Intel corei7
	Memory	12GB
Database Input	Finger images	106x119
	ASL database	



รูปที่ 4 ตัวอย่างภาพภาษามือที่ระบบใช้นำมาใช้ในการเรียนรู้



ก) รูปภาพภาษามือที่มีพื้นหลังแบบต่าง ๆ [16]



ข) รูปภาพภาษามือที่มีติ่งเฉพาะส่วนของมือแบบขาว-ดำ [16]



ค) รูปภาพภาษามือที่มีติ่งเฉพาะส่วนของมือแบบระดับเทา [16]



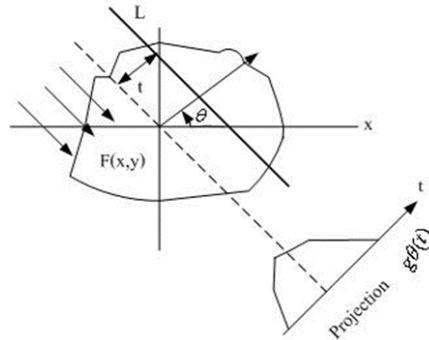
ง) รูปภาพภาษามือที่มีการบังหรือสูญหาย [16]

รูปที่ 5 ตัวอย่างภาพภาษามือที่นำมาใช้ทดสอบ

3.1 การปรับขนาดของภาพ

3.1.1 การวิเคราะห์ส่วนของระนาบภาพโดยอาศัยวิธีการของราดอนทรานฟอร์ม

การหาส่วนของระนาบของภาพวิธีการแรกนั้น อาศัยหลักการของราดอนทรานฟอร์ม เพราะวิธีการของราดอนทรานฟอร์มนั้นมีความดีในการหาการเอียงของภาพได้รวดเร็ว โดยสามารถอธิบายได้ดังรูปที่ 6



รูปที่ 6 ภาพหลักการของราดอนทรานฟอร์ม

จากปัญหาเบื้องต้นของการวางตัวของภาพใน 1 มิติ มุมของภาพที่ได้นั้นสามารถที่จะสร้างใหม่ให้อยู่ในรูปของ 2 มิติได้ สิ่งที่ต้องพิจารณานั้นต้องมองจากลักษณะการวางตัวของภาพ โดยอาศัยหลักการจากสมการของราดอน ดังแสดงไว้ในสมการที่ 1

$$\mathcal{R}\{I\}(d, \phi) = \int_{-\infty}^{\infty} I(d \cos \phi - s \sin \phi + s \cos \phi) ds \quad (1)$$

d = ระยะในแกน x จากจุด Origin

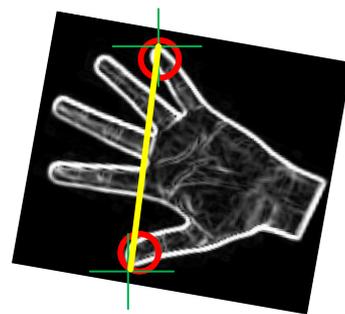
s = ระยะในแกน y จากจุด Origin

ϕ = มุม

I = ขนาดของกรอบหาได้จากตัวแปร d และ s

3.1.2 การหามุมของ (Reference point)

การหาขอบล่างสุดและขอบบนสุดของภาพมือเพื่อหาระนาบโดยวิธีการหาขอบแบบ Candy Edge จากนั้นจะหาค่ายอดที่สูงที่สุด และยอดที่ต่ำสุดของภาพดังรูปที่ 7

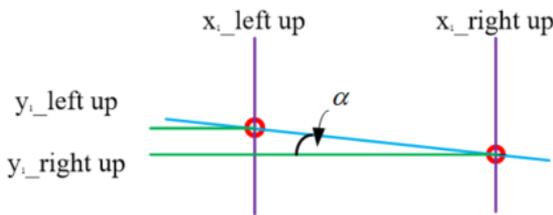


รูปที่ 7 ลักษณะการหาขอบบนและขอบล่างของภาพมือ



เป็นการนำเอาจุดอ้างอิงนำมาใช้งานโดยจะกำหนดได้ดังนี้ $x1_left\ up, y1_left\ up$ คือจุดที่อยู่ตรงขอบด้านบนหรือขอบด้านล่างซ้าย และ $x1_right\ up, y1_right\ up$ คือจุดที่อยู่ตรงขอบด้านบนหรือขอบด้านล่างทางด้านขวา ดังระยะในรูปที่ 7 เพื่อนำมาหาค่ามุม และนำค่าที่ได้ไปทำการปรับมุมต่อไปโดยหาได้จากสมการที่ 2

$$\text{Angle} = \arctan\left(\frac{x1_left\ up - x1_right\ up}{abs(y1_left\ up) - y1_right\ up}\right) \quad (2)$$

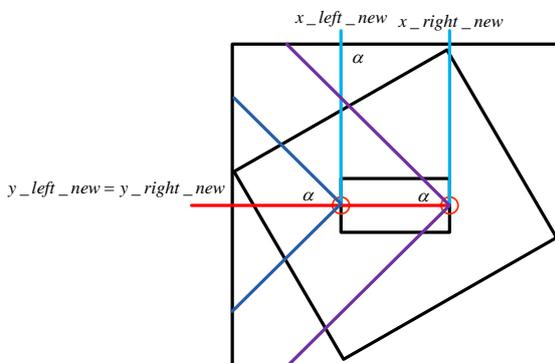


รูปที่ 8 การหาส่วนของจุดอ้างอิงที่ใช้ในการปรับภาพ
จากรูปที่ 8 เป็นภาพของการหาจุดอ้างอิง $x1_left\ up, y1_left\ up$ และ $x1_right\ up, y1_right\ up$ เมื่อได้ 2 จุดข้างต้นแล้วใช้วิธีการจากสมการที่ 2 เพื่อหามุมอัลฟา เพื่อใช้ในการปรับระนาบในวิธีการต่อไป

3.1.3 ผลที่ได้จากมุม (Result degree)

เมื่อได้มุมจากสมการที่ 2 โดยการหาค่าเฉลี่ยของมุมแล้วนำมุนั้นมาทำการหาค่าขององศาที่ถูกต้องเพราะค่าของมุมที่ได้จากสมการแรกนั้นจะได้มุมที่เป็นเรเดียนต้องนำกลับมาทำเป็นมุมที่เป็นองศาโดยใช้สมการที่ 3

$$\text{Angle}(\text{degree}) = \text{Angle}(\text{radian}) * (180/\text{Pi}) \quad (3)$$



รูปที่ 9 ภาพการหมุนกลับมาอยู่ในมุมที่ถูกต้อง

จากรูปที่ 9 แสดงการหมุนกลับของภาพมาอยู่ในระนาบมุมที่ถูกต้องโดยอาศัยวิธีการของราดอนทรานฟอร์มมาอาศัยการ

แบ่งส่วนของภาพออกเป็นส่วนย่อย ซึ่งแบ่งโดยใช้จุด $x1_left\ up, y1_left\ up$ และ $x1_right\ up, y1_right\ up$ เพื่อใช้อ้างอิงการแบ่งส่วนของรูปมือ โดยเมื่อได้มุมอัลฟาแล้วซึ่งเป็นมุมเรเดียนจากสมการที่ 2 แล้วต้องหาค่ามุมที่เป็นองศา โดยเปลี่ยนได้จากสมการที่ 3 จากนั้นเมื่อได้ค่ามุมมองศาที่ถูกต้องแล้วจึงนำมุมที่ได้ป้อนค่าให้กับโปรแกรมเพื่อนำมุมไปปรับระนาบของรูปภาพ

3.1.4 การวิเคราะห์ขั้นตอนวิธี k-Nearest neighbors

กำหนดให้ชุดข้อมูลประกอบไปด้วยกลุ่มของรูปภาพภาษามือจำนวน 24 ตัวอักษร (A-Y ยกเว้นตัว J และ Z) โดยเมื่อต้องการทราบประเภทของข้อมูลที่น่ามาทดสอบทำได้ดังนี้
ขั้นที่ 1 กำหนดค่า k ซึ่งหมายถึงจำนวนตัวอย่างที่ใกล้เคียงที่สุด k ตัว โดยจะมีตั้งแต่ 1 จนถึงจำนวนตัวอย่างทั้งหมดในปริภูมิ

ขั้นที่ 2 กำหนดวิธีการวัดความใกล้เคียงหรือความเหมือน (Similarity) ซึ่งโดยส่วนใหญ่จะวัดจากค่าผกผันของระยะทาง (Distance) ระหว่างตัวอย่าง ระยะที่นิยมมากที่สุดคือ ระยะแบบยูคลิด (Euclidean distance) ระยะทางแบบยูคลิดจากจุด $P = (p1, p2, \dots, pN)$ ไปยังจุด $q = (q1, q2, \dots, qN)$ ในปริภูมิ N มิติ หาได้จากสมการที่ 4

$$d(p, q) = \sqrt{\sum_{i=1}^N (p_i - q_i)^2} \quad (4)$$

ขั้นที่ 3 นับจำนวนตัวอย่างที่ใกล้เคียงที่สุด k ตัวโดยหากจำนวนตัวอย่างประเภทใดมีมากที่สุด ใน k ตัวให้กำหนดประเภทของตัวอย่างที่ต้องการทราบประเภทเป็นประเภทเดียวกับตัวอย่างประเภทที่มีมากที่สุดนั้น ซึ่งความแตกต่างกันไปตามค่า k ที่กำหนด

3.1.5 SVM (Support vector machine)

ข้อดีของการใช้วิธีการแบบ SVM คือสามารถรองรับข้อมูลที่เป็นเชิงเส้นและไม่เป็นเชิงเส้นได้ โดยอาศัยฟังก์ชันของเคอร์เนล (Kernel) ในการลดความซับซ้อนของข้อมูลโดยการส่ง (Mapping) ไปยังปริภูมิพิเศษที่มีความเป็นเชิงเส้น

3.1.6 ขั้นตอนการถดถอยโลจิสติก

การถดถอยโลจิสติก (Logistic regression) เป็นเทคนิคการวิเคราะห์ทางสถิติสำหรับการพยากรณ์ความน่าจะเป็นของการเกิดเหตุการณ์ที่สนใจจากชุดตัวแปรอิสระ หรือตัวแปรตามที่มีการวัดแบบกลุ่ม (Categorical dependent variables) ส่วนใหญ่ใช้โมเดลทางคณิตศาสตร์ดังสมการที่ 5

$$g(x) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k \quad (5)$$

g = ความน่าจะเป็นหรือ โอกาสที่จะเกิดเหตุการณ์

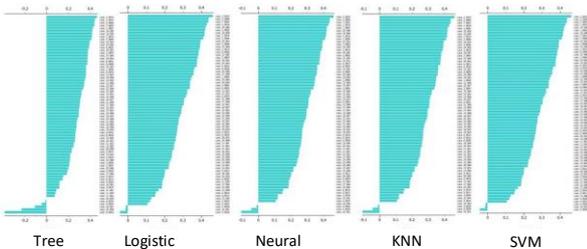
β = ส่วนตัดแกน Y (intercept)

β_k = ค่าคงที่ หรือ ความชัน (slope) ของเส้นตรง แสดงถึงอัตรา การเปลี่ยนแปลงของ Y เมื่อ X_k เปลี่ยนไป 1 หน่วย

X = ค่าตัวแปรอิสระ

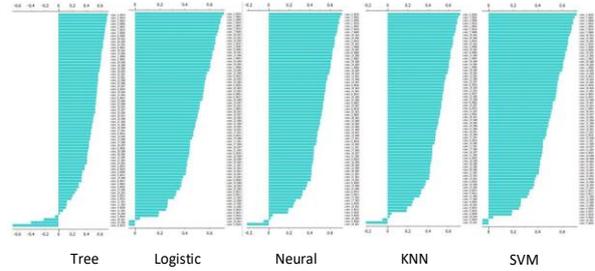
3.17 ขั้นตอนการโครงข่ายประสาทเทียม

เมื่อผ่านขั้นตอนการตัดตัวเลขและตัวอักษร ส่วนรู้ของระบบการแปลภาษาหรือการระบุตัวอักษรของภาษามือ ซึ่งขั้นตอนการระบุตัวอักษร โดยที่เครือข่ายประสาทที่สร้างนี้เป็นแบบ ReLu (Rectified Linear Unit) ซึ่งเป็นฟังก์ชันเส้นตรงที่ถูกปรับแก้ เพื่อนำการใช้งานเครือข่ายประสาทให้สามารถจดจำตัวอักษรภาษาอังกฤษ A-Y ยกเว้นตัว (I,Z) ได้นั้นจะต้องทำการ สอน(train) ให้กับเครือข่ายประสาทก่อนนำไปใช้งาน โดยได้เข้าภาพตัวอักษรอย่างละ 100 ภาพแล้วนำภาพทั้งหมดมาหาค่า Central Moment และ Sum ค่าออกได้จากจุด Central Moment ตามแนวรัศมี ซึ่งจะได้ค่าเฉพาะออกมาแต่ละภาพทั้งหมดจำนวน 24 ค่า จากนั้นก็นำค่า 24 ค่าของแต่ ละรูปภาพ (ภาพ A-Y อย่างละ 100 ภาพ) นี้ ส่งเข้าไป สอน (Train) ให้กับเครือข่ายประสาทได้จดจำเพื่อใช้เป็น ตัวอ้างอิงในการเปรียบเทียบภาพภาษามือที่จะรับเข้ามาว่าเป็นตัวอักษรใด



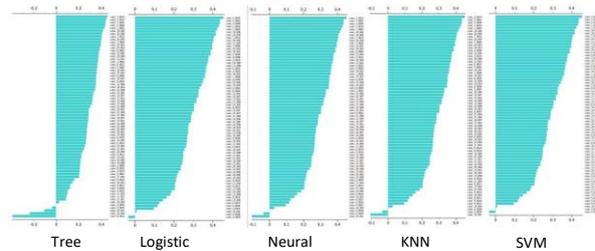
รูปที่ 10 ผลการทดลองโดยใช้ระยะทางแบบยุคลิด

จากรูปที่ 10 แสดงผลการทดลองของการใช้แมทซิ่นเลนนิ่งแบบหลายแมทซิ่นเลนนิ่ง ใช้รูปแบบระยะทางแบบยุคลิด โดยผลการทดลองสามารถแยกแยะภาพภาษามือโดยได้ผลลัพธ์ จากการใช้วิธีการของทรีผลลัพธ์เท่ากับ 93.05% วิธีการของ KNN ผลลัพธ์เท่ากับ 94.44% วิธีการของโครงข่ายประสาท ผลลัพธ์เท่ากับ 94.44 % วิธีการของ SVM ผลลัพธ์เท่ากับ 95.53% และ วิธีการของ Logistic ผลลัพธ์เท่ากับ 95.83% โดยรวมประมาณ 94.72%



รูปที่ 11 ผลการทดลองโดยใช้วิธีแบบโคไซน์

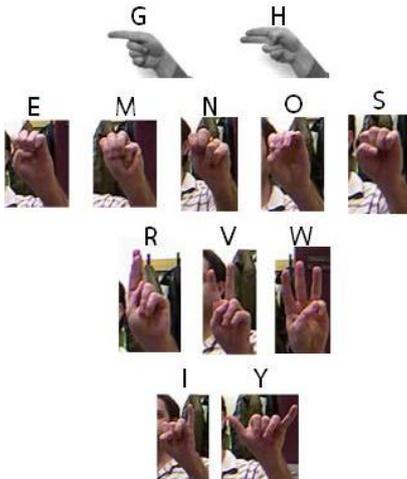
จากรูปที่ 11 แสดงผลการทดลองของการใช้แมทซิ่นเลนนิ่งแบบหลายแมทซิ่นเลนนิ่งใช้วิธีแบบโคไซน์ โดยผลการทดลองสามารถแยกแยะภาพภาษามือโดยได้ผลลัพธ์ จากการใช้วิธีการของทรี ผลลัพธ์เท่ากับ 94.44% วิธีการของ KNN ผลลัพธ์เท่ากับ 97.22% วิธีการของโครงข่ายประสาท ผลลัพธ์เท่ากับ 97.22 % วิธีการของ SVM ผลลัพธ์เท่ากับ 97.22% และ วิธีการของ Logistic ผลลัพธ์เท่ากับ 97.22% โดยรวมประมาณ 96.66%



รูปที่ 12 ผลการทดลองโดยใช้วิธีการของแมนฮัตตัน

จากรูปที่ 12 แสดงผลการทดลองของการใช้แมทซิ่นเลนนิ่งแบบหลายแมทซิ่นเลนนิ่ง ใช้วิธีการของแมนฮัตตัน โดยผลการทดลองสามารถแยกแยะภาพภาษามือโดยได้ผลลัพธ์ จากการใช้วิธีการของทรีผลลัพธ์เท่ากับ 94.44% วิธีการของ KNN ผลลัพธ์เท่ากับ 95.83% วิธีการของโครงข่ายประสาท ผลลัพธ์เท่ากับ 95.83 % วิธีการของ SVM ผลลัพธ์เท่ากับ 97.22% และ วิธีการของ Logistic ผลลัพธ์เท่ากับ 97.22% โดยรวมประมาณ 96.11%

จากการทดลองผลลัพธ์ที่ได้จากการแยกแยะทำให้สามารถทราบในส่วนของความถูกต้อง โดยรวมของระบบประมาณ 95.83% โดยผลของการผิดพลาดอยู่ประมาณ 4.17% จากการวิเคราะห์ทำให้ทราบว่าความผิดพลาดของระบบเกิดจากสามสาเหตุหลัก ๆ คือ 1.ความคล้ายกันของภาพภาษามือดังรูปที่ 13



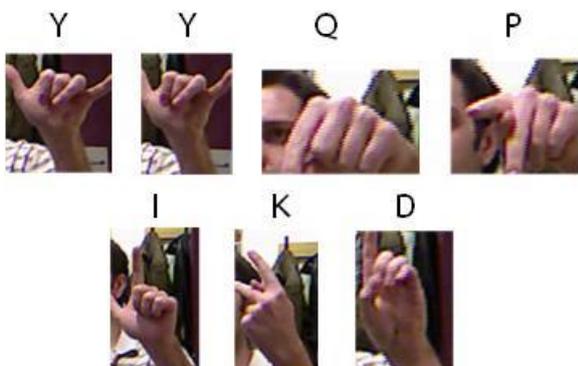
รูปที่ 13 สัญลักษณ์ของภาษามือที่มีความคล้ายกัน

สาเหตุหลักอย่าง 2 ความผิดพลาดเกิดขึ้นจากการเคลื่อนตัวของภาพมือนั่งตัวอย่างรูปที่ 14



รูปที่ 14 ภาพของภาษามือที่มีการเคลื่อนตัว

สาเหตุหลักอย่าง 3 ความผิดพลาดเกิดขึ้นจากการสูญเสียของภาพมือนั่งตัวอย่างรูปที่ 15



รูปที่ 15 ภาพของภาษามือที่มีการสูญเสีย

จากการทดลองใช้ ตัวอักษร 24 ตัวอักษรสีขาวดำ อย่างละ 100 ภาพต่อตัวและตัวอักษรสี 24 ตัว อย่างละ 100 ภาพต่อตัวอักษรรวมทั้งหมด รวมทั้งหมด 4,800 ภาพในการทดลอง

เพื่อใช้ในการเรียนรู้ โดยแบ่งเป็นชุดของภาพขาวดำและสี เพื่อให้รวดเร็วในการแยกแยะ จากการทดลองโดยดึงผลระยะทางแบบยุคลิดได้ 94.72% การทดลองแบบใช้วิธีแบบโคไซน์ได้ 96.66% และ การทดลองแบบได้ 96.11% ระบบโดยรวมที่ใช้วิธีการแยกแยะได้ถูกต้องโดยรวม 95.83% จึงนำมาวิเคราะห์หาสาเหตุการผิดพลาดของระบบจากที่ได้กล่าวมานั้นกระบวนการเคลื่อนไหว สามารถปรับแก้โดยวิธีการหาระนาบใหม่แล้วนำภาพกลับไปสู่กระบวนการตัดสินใจอีกครั้ง ส่วนสาเหตุการมีความคล้ายกันของรูปภาพนั้น ใช้วิธีการเพิ่มรูปแบบความคล้ายกันของภาพมือให้มากขึ้น เพื่อหาความค่าน้ำหนักในการตัดสินใจใหม่ ส่วนสาเหตุสูญเสียหายของภาพนั้นยังต้องหาวิธีการต่อไปในการเพิ่มความถูกต้อง ในการทดลองสามารถเพิ่มความถูกต้องเพิ่มขึ้นได้อีก 2.23% จากการนำภาพมาทำหาระนาบใหม่แล้วทำการเข้าสู่กระบวนการดึงลักษณะหรือน้ำหนักของภาพ และเข้าสู่กระบวนการตัดสินใจอีกครั้ง จึงทำให้ระบบโดยรวมมีความถูกต้องอยู่ที่ 98.06 %

4. บทสรุป

งานวิจัยฉบับนี้ได้นำเสนอการเพิ่มประสิทธิภาพในระบบการแปลภาษามือ โดยการใช้ภาพภาษามือภาษาอังกฤษด้วยการเพิ่มส่วนของการปรับการเอียงของภาพภาษามือ ร่วมกับการวิเคราะห์โดยแมทซึนเลนนิ่งแบบหลายแมทซึนเลนนิ่ง โดยระบบที่นำเสนอนี้ได้ตั้งสมมติฐานที่ว่า ภาพภาษามือเป็นลักษณะข้อมูลคุณลักษณะภาพที่มีเป็นเอกลักษณ์ทางตัวอักษร ซึ่งระบบที่ได้นำเสนอประกอบกับการปรับการเอียงของภาพภาษามือและการวิเคราะห์ภาพโดยใช้แมทซึนเลนนิ่ง ซึ่งถ้า นำเอาข้อมูลที่ผิดพลาดกลับมาหมุนภาพใหม่ แล้วส่งข้อมูลไปวิเคราะห์โดยวิธีการของแมทซึนเลนนิ่งอีกครั้ง โดยหากมีการเปลี่ยนแปลงลักษณะท่าทางไปมากกว่าขอบเขตที่สามารถแยกแยะได้ จะทำให้ระบบสับสนได้ ยิ่งถ้าใช้การแบบวัดระยะทางด้วยน้ำหนักตัวอักษรด้วยแล้ว จะทำให้ผลการแยกแยะคลาดเคลื่อนเนื่องจากลักษณะข้อมูล ที่คล้ายคลึงกัน จึงได้ใช้วิธีการปรับภาพเอียงของภาพภาษามือเข้ามาช่วยปรับภาพภาษามือของข้อมูลให้มีลักษณะของระนาบให้ถูกต้องเพื่อใช้สำหรับการวิเคราะห์โดยรวม โดยสรุปการแยกแยะสามารถทำได้อยู่ที่ 98.06% ของความถูกต้องโดยการรวมแมทซึนเลนนิ่งแบบหลายวิธีร่วมกับการปรับระนาบภาพใหม่เพื่อใช้ในการตัดสินใจอีกครั้ง ทำให้การแยกแยะตัวอักษรภาพภาษามือได้ประสิทธิภาพดีขึ้น

5. กิตติกรรมประกาศ

ผู้วิจัยขอขอบคุณมหาวิทยาลัยเทคโนโลยีราชมงคลล้านนาที่ได้ให้สถานที่และทุนในการทำงานวิจัย

6. เอกสารอ้างอิง

- [1] United Nations, Department of Economic and Social Affairs, Population Division (2019). World Population Prospects 2019: Highlights (ST/ESA/SER.A/423).
- [2] United Nations, Multi-country assessment of national capacity to provide hearing care, WHO Report Hearing Care: (NLM classification: WV 270)
- [3] Demographic Statistics Report Population and Housing, National Statistical Office. Available from: <http://statbbi.nso.go.th/>
- [4] Report of the situation of the disabled in Thailand. Available from: <http://www.dep.go.th>
- [5] Gafka, Jakub, et al. "Inertial motion sensing glove for sign language gesture acquisition and recognition," *IEEE Sensors Journal* 16.16 (2016): pp.6310-6316.
- [6] Javed, Shahrukh. "Wireless Glove for Hand Gesture Acknowledgment: Sign Language to Discourse Change Framework in Territorial Dialect," *ICRA 2018* (2018).
- [7] Saleh, Neven, et al. "Smart glove-based gestures recognition system for Arabic sign language," *Innovative Trends in Computer Engineering (ITCE)*, International Conference on (2020).
- [8] Rokade, Rajeshree S., and Dharmpal D. Doye. "Spelled sign word recognition using key frame." *IET Image Processing* , 9.5 (2014): pp 381-388.
- [9] Kshirsagar, Ketki P. "Key Frame Selection for One-Two Hand Gesture Recognition with HMM." *International Journal of Advanced Computer Research* 5.19 (2015): pp. 192.
- [10] Hosain, Al Amin, et al. "FineHand: Learning Hand Shapes for American Sign Language Recognition." arXiv preprint arXiv:2003.08753 (2020).
- [11] Tubaiz, Noor, Tamer Shanableh, and Khaled Assaleh. "Glove-based continuous Arabic sign language recognition in user-dependent mode." *IEEE Transactions on Human-Machine Systems*, 45.4 (2015): pp.526-533.
- [12] Jiang, Shuo, et al. "Stretchable E-Skin Patch for Gesture Recognition on the Back of the Hand." *IEEE Transactions on Industrial Electronics* (2019).
- [13] Sanguansat, Parinya.(2019). Artificial Intelligence with Machine Learning, Nonthaburi : IDC Premier.
- [14] Miroslav Kubat. (2015). An Introduction to Machine Learning. Coral Gables: Springer.
- [15] Srikaew, Athit. (2009). Computational intelligence. Teaching documents Department of Electrical Engineering Suranaree University of Technology.
- [16] Life Print. American Sign Language. Available from:<https://www.lifeprint.com/asl101/fingerspelling/fingerspelling.htm/> [Accessed 6th April 2020].