

การประเมินความสามารถในการคำนวณท่าทางการเคลื่อนไหวของมนุษย์
โดยใช้ MediaPipe เพื่อการฟื้นฟูสมรรถภาพจากโรคหลอดเลือดสมองที่บ้าน
**Enhancing Stroke Rehabilitation at Home: Evaluating MediaPipe
for Human Pose Estimation**

นรภัทร ลาภชูรัตน์^{1*}, วรวัฒน์ เชิญสวัสดิ์² และ กิ่งกาญจน์ สุขคณาภิบาล²

Norapat Labchurat^{1*}, Worawat Choensawat², and Kingkarn Sookhanaphibarn²

¹คณะเทคโนโลยีสารสนเทศและนวัตกรรม สาขาวิชาเทคโนโลยีสารสนเทศและวิทยาการข้อมูล

*School of Information Technology and Innovation Department of Information Technology
and Data Science*

²ศูนย์นวัตกรรมพิเศษเฉพาะทาง มหาวิทยาลัยกรุงเทพ

Center of Specialty Innovation (CoSI), Bangkok University

Received: May 31, 2023; Revised: June 4, 2023; Accepted: June 25, 2023; Published: June 30, 2023

ABSTRACT – This research experiment evaluates the accuracy of MediaPipe, a computer vision technology, in estimating human pose compared to motion capture technology. The experiment utilizes poses derived from the Fugl-Meyer Assessment for upper extremity (FMA-UE). Webcams are installed at three positions: front 90°, left 40°, and right 140° of the participant, while participants wear motion capture suits. The experiment begins with simultaneous movements performed on both sides of the body, with data recorded from both systems simultaneously. We calculate angles of body joints and compare the accuracy percentage. The results indicate that MediaPipe's accuracy in estimating pose based on the FMA-UE was not significantly high. Due to limitations in estimating poses in the vertical plane from z-axis and the range resulted in wider angles than reality. However, the usage of normalization for 2D angle calculation presents the better accuracy. Therefore, MediaPipe can still be used for developing interactive applications and assist the doctor evaluate patient, but further research and development are needed for evaluating patients during FMA-UE exercises.

KEYWORDS: Computer vision, MediaPipe, Human pose estimation, Rehabilitation, Fugl-Meyer Assessment, Home-based rehabilitation, Upper limb rehabilitation

บทคัดย่อ -- งานวิจัยนี้เน้นการประเมินความแม่นยำของเทคโนโลยีคอมพิวเตอร์วิทัศน์ MediaPipe ในการประมาณค่าท่าทางของมนุษย์โดยเปรียบเทียบกับเทคโนโลยีโมชันแคปเจอร์ โดยจะใช้ท่าทางอ้างอิงมาจาก Fugl-Meyer Assessment for upper extremity (FMA-UE) ติดตั้งกล้องเว็บแคมสามจุดที่ด้านหน้า 90° ด้านซ้าย 40° และด้านขวา 140° ของผู้เข้าร่วมและผู้เข้าร่วมจะสวมใส่ชุดโมชันแคปเจอร์ตลอดการทดลอง เริ่มต้นการทดลองผู้เข้าร่วมจะทำท่าทางโดยการเคลื่อนไหวส่วนของร่างกายทั้งสองด้านพร้อมกัน และข้อมูลจากทั้งสองระบบถูกบันทึกพร้อมกัน ผู้วิจัยจะคำนวณองศาของข้อต่อในร่างกายและเปรียบเทียบร้อยละความแม่นยำ ผลการวิจัยพบว่าความแม่นยำของ MediaPipe ในการประมาณค่าท่าทาง

*Corresponding Author: narin@rmutl.ac.th

อ้างอิงจาก FMA-UE ไม่สูงมาก เนื่องจากมีข้อจำกัดในการประมาณค่าท่วงท่าในระนาบตั้งของแกนแนวลึกและมีช่วงขององศาที่ไต่กว้างเกินความเป็นจริง อย่างไรก็ตามการทำให้ข้อมูลขององศาจากการประมาณค่าท่วงท่าเป็นมาตรฐานเดียวกันกับ MoCap ในการคำนวณแบบสองมิติแสดงให้เห็นว่าการคำนวณมีความแม่นยำมากขึ้น ดังนั้น MediaPipe ยังคงสามารถใช้ในการพัฒนาแอปพลิเคชันเชิงตอบโต้และช่วยเหลือแพทย์ประเมินผู้ป่วยเบื้องต้นได้ แต่จำเป็นจะต้องวิจัยและพัฒนาต่อยอดจาก MediaPipe เพื่อใช้ในการประเมินผู้ป่วยขณะทำท่วงท่าจาก FMA-UE

คำสำคัญ: คอมพิวเตอร์วิทัศน์, MediaPipe, การประมาณค่าท่วงท่าของมนุษย์, การฟื้นฟูสมรรถภาพ, แบบประเมิน Fugl-Meyer, การฟื้นฟูสมรรถภาพที่บ้าน, การฟื้นฟูสมรรถภาพรยางค์บน

1. บทนำ

จากผลการสำรวจในปี 2562 โรคหลอดเลือดสมองเป็นสาเหตุสำคัญอันดับต้น ๆ ของโลกและของไทยที่ทำให้เกิดการเสียชีวิตและความพิการ [1] ผู้ป่วยที่รอดชีวิตจากโรคหลอดเลือดสมองส่วนใหญ่มักเกิดอาการอ่อนแรงครึ่งซีก หรือเกิดการอ่อนแรงของกล้ามเนื้อในด้านหนึ่งของร่างกายซึ่งทำให้ไม่สามารถช่วยเหลือตัวเองได้ จึงต้องได้รับการทำกายภาพบำบัด การออกกำลังกาย และการฟื้นฟูสมรรถภาพเพื่อให้สามารถกลับมาใช้ชีวิตประจำวันได้อย่างปกติที่สุด [2]

การประเมินและการติดตามผลการฟื้นฟูของผู้ป่วยแบบเดิม ผู้ป่วยจะต้องพบแพทย์ในโรงพยาบาลหรือคลินิก ฉะนั้นผู้ป่วยจำเป็นต้องเดินทางไปยังสถานที่ห่างไกลเพื่อขอรับการฟื้นฟู [3] จึงเป็นเหตุให้ผู้ป่วยขาดการนัดพบแพทย์และไม่สามารถได้รับการฟื้นฟูสมรรถภาพอย่างสม่ำเสมอ อีกทั้งสถานการณ์ระบาดของโรค โควิด 19 ที่กำลังเกิดขึ้นกับโลกในปัจจุบัน ทำให้มีการเข้ารับการฟื้นฟูสมรรถภาพที่โรงพยาบาลหรือคลินิกลดน้อยลงและทำได้ยากลำบาก [4] เนื่องจากความจำเป็นในการลดความเสี่ยงของการติดเชื้อไวรัสผ่านการสัมผัส หรือการลดความเสี่ยงระหว่างการเดินทาง

นอกจากนี้ การฟื้นฟูด้วยกระบวนการเดิม ๆ เป็นประจำอาจทำให้เกิดความเบื่อหน่ายแก่ผู้ป่วยได้ และการออกกำลังกายด้วยเทคโนโลยีความเป็นจริงเสมือน เทคโนโลยีจับการเคลื่อนไหวโมชันแคปเจอร์หรือการใช้หุ่นยนต์ฟื้นฟู ถึงแม้จะมีความแม่นยำสูง สามารถช่วยฟื้นฟู ประเมินสมรรถภาพ และลดความเบื่อหน่ายแก่ผู้ป่วยได้ แต่อุปกรณ์เหล่านี้ต้องติดตั้งโปรแกรมล่วงหน้าที่บ้านผู้ป่วย และมีต้นทุนที่สูง [3] ใน

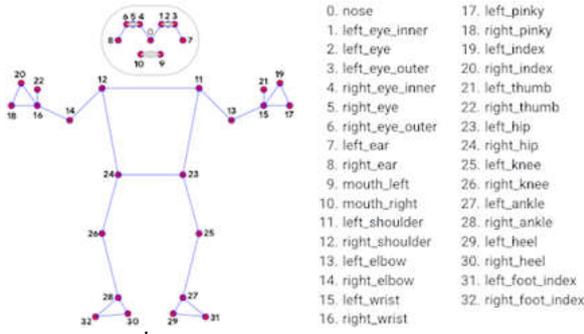
บทความนี้จึงเสนอการใช้อุปกรณ์เว็บแคมจับภาพทั่วไปเพื่อเป็นอีกทางเลือกหนึ่งสำหรับการฟื้นฟู

ทางผู้วิจัยจึงเล็งเห็นความสำคัญในการทำให้ผู้ป่วยสามารถเข้าถึงการฟื้นฟูสมรรถภาพได้อย่างง่ายที่บ้าน ส่งเสริมแรงจูงใจ และมีต้นทุนต่ำบนเว็บเบราว์เซอร์ด้วยเทคโนโลยีประมาณค่าท่วงท่าการเคลื่อนไหวของมนุษย์ (Pose estimation) ผ่านอุปกรณ์ที่สามารถส่งภาพไปในหน่วยประมวลผลกลาง (CPU) หรือหน่วยประมวลผลกราฟิกส์ (GPU) เช่น กล้องเว็บแคม กล้องไอแพดหรือกล้องที่สามารถจับค่าสี RGB เพื่อเปลี่ยนท่วงท่าออกมาเป็นอินพุตหรือคำสั่งสำหรับเกมออกกำลังกาย

อย่างไรก็ตาม คอมพิวเตอร์วิทัศน์ได้ใช้การประมาณค่าของท่วงท่าโดยการเรียนรู้ของเครื่องหรือ Machine Learning (ML) ผ่านภาพจากกล้องเว็บแคมซึ่งสามารถส่งข้อมูลได้เพียงรูปภาพเท่านั้น ดังนั้นมุมมองที่แตกต่างกันอาจจะทำให้ได้ความแม่นยำที่แตกต่างกัน และการออกแบบเพื่อใช้ท่วงท่าที่มาจาก FMA-UE จำเป็นจะต้องพิสูจน์ความแม่นยำในการประมาณค่าท่วงท่าที่มีการเคลื่อนไหวไปในทิศทางแนวแกนลึก (Z-Axis) ด้วยการคำนวณองศาแบบทั้งสองและสามมิติ และค้นหาวิธีการใช้งานคอมพิวเตอร์วิทัศน์ที่เหมาะสมก่อนจะนำมาใช้ในการพัฒนาเกมออกกำลังกาย หรือใช้เพื่อประเมินผู้ป่วยโรคหลอดเลือดสมองขณะออกกำลังกายได้จากที่บ้าน

การวิจัยนี้จึงใช้เครื่องมือคอมพิวเตอร์วิทัศน์ (Computer vision) ที่มีชื่อว่า MediaPipe ประมาณค่าท่วงท่าของมนุษย์ในขณะที่กำลังออกกำลังกาย และออกแบบท่วงท่าเพื่อการออกกำลังกายจาก FMA-UE ผู้วิจัยได้ทำการทดลองเปรียบเทียบความแม่นยำทั้งแบบสองมิติและสามมิติ ด้วยระบบจับการเคลื่อนไหวโมชันแคปเจอร์ เพื่อตรวจสอบความเป็นไปได้ใน

การสร้างแอปพลิเคชันให้กับผู้ป่วยโรคหลอดเลือดสมองออกกำลังกายเพื่อฟื้นฟูสมรรถนะได้จากที่บ้าน และตรวจสอบร้อยละความแม่นยำในการประเมินสมรรถภาพ เพื่อใช้เป็นเครื่องมือในการช่วยแพทย์สามารถเข้าถึง และให้การรักษาแก่ผู้ป่วยได้จากระยะไกลหรือที่เรียกว่า Telerehabilitation



รูปที่ 1. Blaze pose model (33 Landmarks)

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

เทคโนโลยีคอมพิวเตอร์วิทัศน์หรือ Computer vision เป็นการใช้อินเทอร์เน็ตเพื่อประมวลผลภาพ โดยใช้ปัญญาประดิษฐ์ (Artificial Intelligence) หรือการเรียนรู้ของเครื่องเพื่อให้คอมพิวเตอร์สามารถเข้าใจและวิเคราะห์ภาพเช่นเดียวกับมนุษย์ได้ผ่านอุปกรณ์ที่เป็นกล้อง RGB [5] ซึ่งปัจจุบันมีโปรแกรมและเครื่องมือมากมายที่สามารถนำไปสร้างแอปพลิเคชันเพื่อการประมวลผลภาพได้ เช่น OpenCV, MediaPipe, Openpose, TensorFlow.js และอื่น ๆ [6] [7] [8] โดยในการวิจัยนี้เราได้เลือกใช้ MediaPipe ซึ่งได้พิสูจน์แล้วใน [9] ว่ามีความแม่นยำสูงที่สุดจากเครื่องมือที่ยกตัวอย่างมา

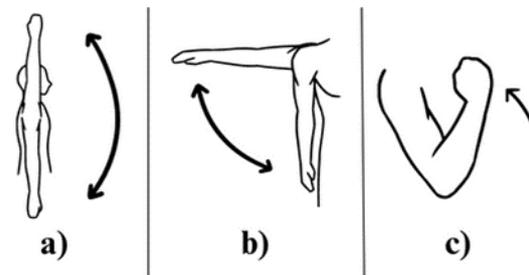
ในวรรณกรรมที่ผ่านมาได้แสดงให้เห็นถึงการใช้เทคโนโลยีคอมพิวเตอร์วิทัศน์ และเทคนิคการประมาณค่าท่วงท่าของมนุษย์ในการพัฒนาแอปพลิเคชันเพื่อเสริมสร้างแรงจูงใจและการฟื้นฟูสมรรถภาพแก่ผู้ป่วยที่มีความพิการหรือเกมออกกำลังกายส่งเสริมสุขภาพ [10] และการตรวจสอบความถูกต้องในขณะที่ท่วงท่าจากการเล่นโยคะ [9] นอกจากนี้การประมาณค่าท่วงท่ายังมีการใช้เทคนิคที่เกี่ยวข้องกับคอมพิวเตอร์วิทัศน์อย่างอื่นอีกเช่นการอ่านค่าสีจากพิกเซล (Pixel) ของรูปภาพในขณะที่กำลังสวมใส่อุปกรณ์เสริมอย่างเช่นถุงมือสี [11] หรือการถือสิ่งของเพื่อการใช้เทคนิคตรวจจับวัตถุ (Object detection) [12] แทนที่การตรวจจับพิกัดของร่างกายมนุษย์โดยตรง เทคนิค

เหล่านี้เพิ่มความแม่นยำในการโต้ตอบ และผลลัพธ์ของการกระทำที่จะเกิดขึ้นภายในแอปพลิเคชัน

2.1 MediaPipe และการประมาณค่าท่วงท่า

MediaPipe เป็นเครื่องมือที่ผู้วิจัยเลือกใช้ในการวิจัยนี้ และเป็นเครื่องมือที่ถูกพัฒนาโดย Google ซึ่งมีเป้าหมายเพื่อให้นักพัฒนาสามารถสร้างและพัฒนาแอปพลิเคชันคอมพิวเตอร์วิทัศน์ได้ง่ายขึ้น อย่างเช่นเรื่องของการประมาณค่าท่วงท่าของมนุษย์ (Pose estimation) [9] ซึ่งเป็นกระบวนการหาตำแหน่งและทิศทางของส่วนต่างๆ ของร่างกายในภาพหรือวิดีโอ

การประมาณค่าท่วงท่าเป็นกระบวนการที่ใช้ในการตรวจจับการเคลื่อนไหวและระบุตำแหน่งส่วนต่างๆ ของร่างกาย โดยส่วนใหญ่ใช้เทคนิคการเรียนรู้ของเพื่อสร้างแบบจำลองที่สามารถระบุโครงสร้างกระดูกของมนุษย์ ซึ่ง MediaPipe มีการใช้การประมวลผลร่วมกันของ CPU และ GPU ผ่านเครื่องมือ OpenGL จึงทำให้มีการประมวลผลที่รวดเร็ว [13] และผลลัพธ์ของการประมาณค่าท่วงท่าจะได้เป็นพิกัด Landmarks จำนวน 33 จุด [14] ดังรูปที่ 1 นอกจากนี้ยังมีโมเดลอื่น ๆ เช่น การตรวจจับใบหน้า มือ และสิ่งของ เป็นต้น



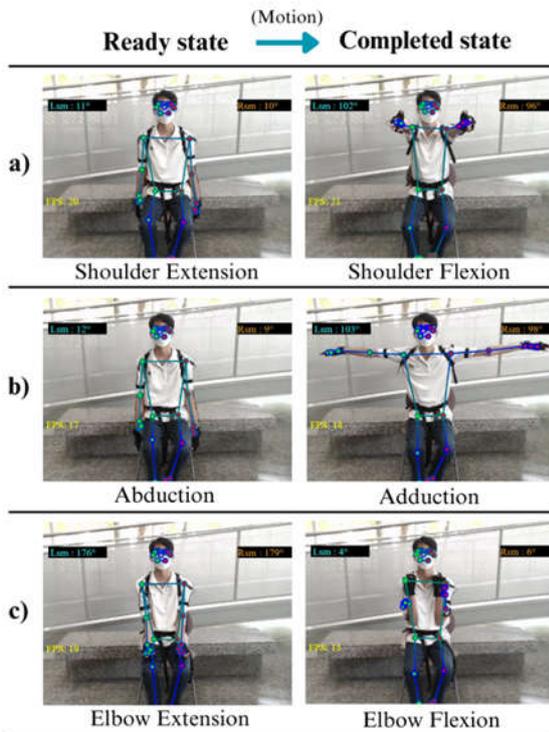
รูปที่ 2. ท่วงท่าที่ใช้ในการวิจัยจาก FMA-UE

2.2 Motion Capture (MoCap)

การวิจัยนี้ผู้วิจัยใช้ระบบจับการเคลื่อนไหวโมชันแคปเจอร์ประเภท Inertial Measurement Unit (IMU) [15] ซึ่งใช้เซ็นเซอร์ในการจับและบันทึกการเคลื่อนไหวของร่างกาย ซึ่งแต่ละเซ็นเซอร์ IMU ประกอบด้วยเซ็นเซอร์จับความเร่ง (accelerometer) และเซ็นเซอร์วัดความเร็วมุม (gyroscope) เพื่อวัดและบันทึกการเคลื่อนไหวของส่วนต่างๆ ของร่างกาย [16] และสามารถนำผลลัพธ์จากเซ็นเซอร์มาสร้างโมเดลที่แสดงการเคลื่อนไหวของร่างกายในรูปแบบสามมิติ ผ่านโปรแกรม

สำหรับโมชันแคปเจอร์ ด้วยเฟรมเรทมาตรฐานที่ 120 Fps (Frame per second)

ปัจจุบันมีการใช้งาน MoCap เพื่อเปรียบเทียบคุณภาพของเทคโนโลยีที่มีการประมาณค่าท่วงท่าของมนุษย์ [17] และใช้เพื่อการฟื้นฟูสมรรถภาพและประเมินผู้ป่วยโรคหลอดเลือดสมองเพื่อหาระดับของความคิด หรือประเมินความบกพร่องของร่างกายในการเคลื่อนไหวซึ่งสามารถประเมินได้อย่างละเอียด [18] แต่ข้อเสียคือราคาที่สูง ส่วนในการวิจัยนี้ผู้วิจัยได้ใช้ฟังก์ชันของ MoCap จำนวนมุมและเป็นข้อมูลที่ที่จะทำนาย (Predict) เพื่อหาความแม่นยำการประมาณค่าท่วงท่าของ MediaPipe



รูปที่ 3. ท่วงท่าที่ใช้ในการวิจัยจาก FMA-UE เมื่อทำการประมาณค่าท่วงท่าใน MediaPipe

3. วิธีดำเนินการวิจัย

ในส่วนนี้ผู้วิจัยจะอธิบายขั้นตอนในการดำเนินการวิจัย วิธีเก็บข้อมูลจากผู้เข้าร่วมการทดลอง และขั้นตอนการดำเนินการกับข้อมูลที่ได้นำมาเพื่อเตรียมนำไปหาผลลัพธ์ของการทดลองต่อไป

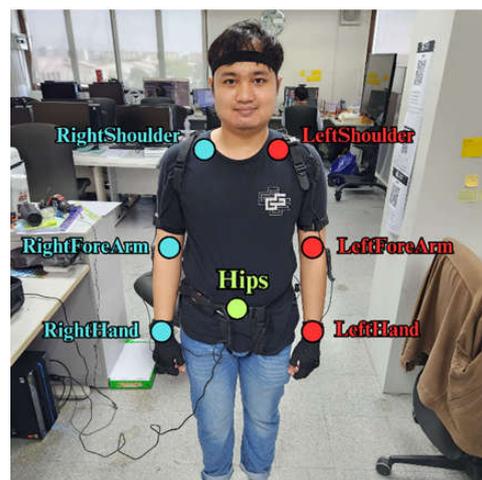
3.1 การออกแบบท่วงท่า

การออกแบบท่วงท่าที่ใช้ในการดำเนินการวิจัยได้อ้างอิงการเคลื่อนไหวจาก Fugl-Meyer Assessment-Upper Extremity หรือ FMA-UE ซึ่งเป็นเครื่องมือทางการแพทย์ที่ใช้ในการประเมินความสามารถของผู้ป่วยอัมพาตครึ่งซีกส่วนรยางค์บน [19] ทางผู้วิจัยเลือกใช้ท่วงท่าเหล่านี้เพื่อหาความเป็นไปได้ในการประเมินสมรรถนะผู้ป่วยระหว่างการฟื้นฟูสมรรถภาพด้วยเกมออกกำลังกายและการกระทำท่วงท่าที่มาจาก FMA-UE

ท่วงท่าที่ใช้ในการวิจัยนี้มีทั้งหมดสามท่วงท่าดังรูปที่ 2 และการเคลื่อนไหวดังรูปที่ 3 ได้แก่ a) การงอหรือเคลื่อนไหวไหล่ขึ้น (Shoulder Flexion), b) การเคลื่อนไหวของไหล่ไปทิศทางด้านข้างหรือออกจากลำตัว (Abduction) และ c) การงอข้อศอก (Elbow Flexion) โดยภาพทางด้านซ้ายจะเป็นท่าเตรียมพร้อม และภาพด้านขวาจะเป็นท่าที่ต้องทำให้สำเร็จ

3.2 การคำนวณองศาของตำแหน่งในร่างกาย

การเลือกท่วงท่าในการวิจัยนั้นสำคัญต่อตำแหน่งของร่างกายที่ใช้ในการคำนวณผลลัพธ์ของข้อต่อของ MoCap หรือผลลัพธ์ของ Landmark ของ MediaPipe ทั้งนี้เพื่อให้มั่นใจว่าองศาที่ได้จากการเคลื่อนไหวนั้นมาจากตำแหน่งของร่างกายจุดเดียวกัน หรือใกล้เคียงกันมากที่สุด การกำหนดตำแหน่งของร่างกายสามจุดเพื่อเอาไว้ใช้คำนวณองศาจึงจำเป็นอย่างมากที่จะทำให้ผลลัพธ์การทดลองออกมาถูกต้องมากที่สุด



รูปที่ 4. ตำแหน่งข้อต่อของ MoCap ที่ใช้ในการทดลอง

3.2.1 การเลือกข้อต่อของ MoCap

ข้อต่อ MoCap ที่ผู้วิจัยเลือกใช้แสดงดังรูปที่ 4 เพื่อคำนวณหาองศาของไหล่ในท่วงท่าอไหล่ขึ้น และกางไหล่ออกจากลำตัวในแต่ละด้าน (แทนคำว่า [Side] ด้วยคำว่า Left ด้านซ้ายหรือ Right ด้านขวา) ได้แก่ Hips, [Side]Shoulder [Side]ForeArm ส่วนข้อต่อที่ใช้เพื่อคำนวณองศาของข้อศอกแต่ละด้านคือ [Side]Shoulder, [Side]ForeArm และ [Side]Hand ตามลำดับ

3.2.2 การเลือก Landmarks ของ MediaPipe

หมายเลข Landmark ของ MediaPipe ดังรูปที่ 1 เพื่อใช้พิกัดมาคำนวณหาองศาของท่วงท่าอไหล่ขึ้นและกางไหล่ออกจากลำตัว เฉพาะในส่วนด้านขวาของร่างกายได้แก่ [24, 12, 14] ส่วนด้านซ้ายของร่างกายได้แก่ [23, 11, 13] ส่วนหมายเลขที่เอามาคำนวณองศาของข้อศอกเฉพาะด้านขวาของร่างกายได้แก่ [12, 14, 16] ส่วนด้านซ้ายของร่างกายได้แก่ [11, 13, 15]



รูปที่ 5. องศาการติดตั้งกล้อง

3.3 ขั้นตอนและวิธีการทดลอง

การติดตั้งกล้องเพื่อส่งภาพให้ MediaPipe ผู้วิจัยได้แบ่งตำแหน่งให้คอมพิวเตอร์แต่ละเครื่องทำมุมกับผู้เข้าร่วมการทดลองดังรูปที่ 5 ได้แก่ มุมที่ 90° เน้นจับทางด้านหน้า เนื่องจากเป็นมุมที่สามารถติดตั้งกล้องเว็บแคมบนมอนิเตอร์ได้ และเป็นมุมปกติที่ผู้คนจะติดตั้งกล้องกัน, มุมที่ 40° และมุมที่ 140° เป็นมุมที่อยู่ด้านซ้ายและขวาตามลำดับ ซึ่งเป็นมุมเฉียงที่สามารถมองเห็น

แขนทั้งสองด้านของร่างกายได้อย่างชัดเจน อีกทั้งภาพจากกล้องเว็บแคมเป็นแบบสองมิติ และทำการประมาณค่าท่วงท่าด้วย MediaPipe เป็นพิกัดสามมิติ ผู้วิจัยจึงตั้งสมมติฐานว่ามุมกล้องส่งผลต่อการประมาณค่าท่วงท่าในแต่ละมิติหรือไม่ การเบี่ยงมุมกล้องดังกล่าวจะสามารถแสดงความแม่นยำที่แตกต่างกันของการประมาณค่าในแต่ละมิติของแต่ละมุมกล้องได้

นอกจากนี้ผู้เข้าร่วมการทดลองจะต้องสวมใส่ชุดเซนเซอร์ IMU ของ Motion capture ตลอดการทดลอง และเมื่อการทดลองเริ่มขึ้น ผู้วิจัยจะทำการบันทึกการเคลื่อนไหวของ MoCap และบันทึกการประมาณค่าท่วงท่าของ MediaPipe ทั้งหมดพร้อมเรียงกันแบบเรียลไทม์

การบันทึกข้อมูลจะแบ่งออกไปในการทดลองแต่ละท่วงท่าดังรูปที่ 3 เริ่มต้นผู้เข้าร่วมการทดลองจะอยู่ในท่าเตรียมพร้อมเพื่อเข้าสู่สถานะ Ready state และทำการเคลื่อนไหวร่างกายส่วนบนทั้งสองด้านของร่างกายพร้อมกัน (Synchronous) เพื่อทำช่วงของการเคลื่อนไหว (Range of motion) ของท่วงท่าให้สำเร็จ และเข้าสู่สถานะต่อไปคือ Completed state หลังจากนั้นจะทำการเคลื่อนไหวกลับมายังท่าเตรียมพร้อมอีกครั้ง และยังคงกระทำท่วงท่าเดิมวนซ้ำมากที่สุดที่ยังทำให้มีข้อมูลเพื่อนำมาเปรียบเทียบมากขึ้นและได้ผลลัพธ์ที่ชัดเจนมากขึ้น ซึ่งในการวิจัยนี้จึงจะทำการวนซ้ำทั้งหมด 30 วินาที ซึ่งจะได้ประมาณ 7 รอบ หลังจากทำครบก็จะสิ้นสุดการทดลองของท่วงท่า นั้น ๆ และทำท่วงท่าที่เหลือต่อจนครบทั้งหมดสามท่วงท่า จึงจะเปลี่ยนเป็นผู้เข้าร่วมการทดลองคนถัดไป

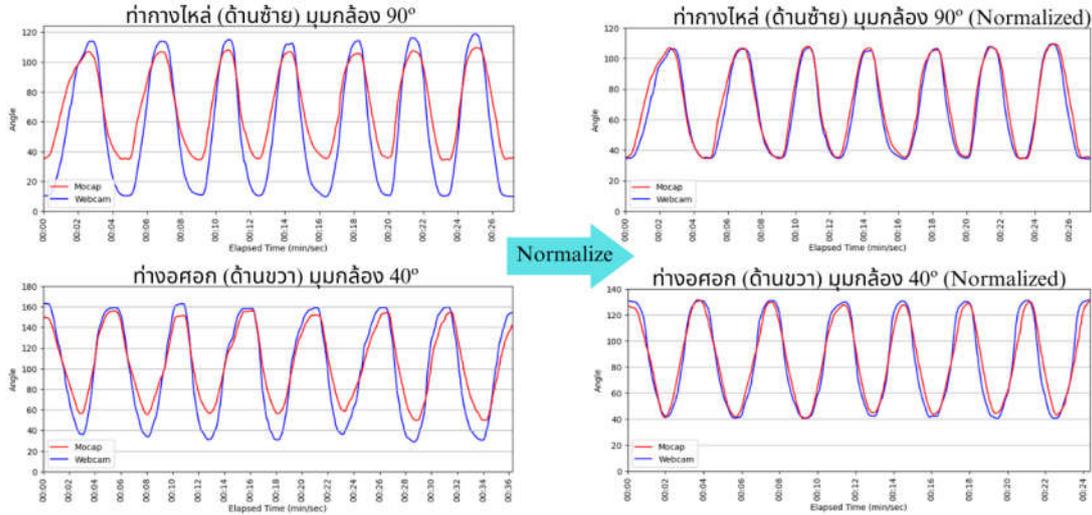
3.4 การประมวลผลการเคลื่อนไหว

พิกัด Landmark ที่ได้จาก MediaPipe จะถูกคำนวณองศาออกมาดังข้อ 3.2.2 และผลลัพธ์ที่ได้อาจจะมีความแปรปรวนเนื่องจากสีเสื้อผ้าของผู้เข้าร่วมการทดลอง สีของพื้นหลัง และสิ่งของที่อยู่ในพื้นหลัง อาจจะทำให้กระทบกับการประมาณค่าท่วงท่าของเทคโนโลยีคอมพิวเตอร์วิทัศน์ ผู้วิจัยจึงต้องทำการลดสัญญาณข้อมูลหรือ Smooth ข้อมูลขององศาด้วยหลักการ Moving Average (MV) [20] ตามขนาดหน้าต่าง

ขนาดหน้าต่าง (Window size) หรือช่วงที่เคลื่อน ถ้า FPS สูงสามารถกำหนดให้มีขนาดใหญ่ได้ แต่กลับกันถ้า FPS ต่ำ ช่วงที่เคลื่อนก็ควรมีขนาดเล็ก เพราะฉะนั้นการกำหนดขนาดมากเกินไปจะส่งผลกระทบต่อ delay ของคำสั่งหรืออินพุตในระบบจากการทดลองในการวิจัยนี้ ซึ่งเฟรมเรทของ MediaPipe เคลื่อนใน

การวิจัยอยู่ที่ 25 FPS และจากการทดลองปรับขนาดต่าง ๆ ผู้วิจัยได้สังเกตว่าขนาดที่ดีที่สุดของการวิจัยนี้คือ 5 หรือ 1 ใน 5 ของ

เฟรมเรททั้งหมด และ Window size นี้ได้ให้รูปแบบของกราฟที่ไม่แตกต่างกับข้อมูลดิบมากนัก และมีผลลัพธ์ใกล้เคียงกับ MoCap มากที่สุด



รูปที่ 6. ตัวอย่างของข้อมูลจากการคำนวณองศาแบบสองมิติเมื่อนำมาทำการ Normalize

เฟรมเรทการบันทึกข้อมูลต่อวินาทีของ MoCap อยู่ที่ 120 fps เพราะฉะนั้นการบันทึกข้อมูลเฟรมเรทของระบบทั้งสองจะไม่เท่ากัน ทางผู้วิจัยจึงต้องจัดเรียงให้ข้อมูลของทั้งสองระบบมีจำนวนเฟรมที่เท่ากัน โดยวิธีทำคือจับเฟรมที่มีระยะเวลา (Elapsed time) ใกล้เคียงมากที่สุดเมื่อเทียบกับ MediaPipe มาสร้างเป็นชุดข้อมูลใหม่สำหรับ MoCap เพื่อที่จะสามารถเปรียบเทียบความแตกต่าง และใช้เครื่องมือสถิติคำนวณผลลัพธ์ออกมาได้เฟรมต่อเฟรม

4. การทดลองและผลการทดลอง

ในการทดลองครั้งนี้มีผู้เข้าร่วมทั้งหมด 5 คน ประกอบด้วยเพศชาย 3 คนและหญิง 2 คน อายุของผู้เข้าร่วมอยู่ในช่วง 20-23 ปี

4.1 ชุดข้อมูล

ผู้วิจัยแบ่งชุดผลการทดลองออกเป็นหัวข้อตามตัวแปรที่ใช้ในการสรุปผลได้แก่ ท่า่งท่า มิติในการคำนวณ ด้านของร่างกาย และ มุมกล้อง ฉะนั้นการจับคู่ของข้อมูลทั้งสองระบบเพื่อใช้ในการหาค่าผลลัพธ์สถิติจะกระทำโดยตัวแปรที่กำหนดไว้เช่นเดียวกัน การแบ่งชุดข้อมูลลักษณะนี้ทำให้สามารถจำลองสถานการณ์ต่าง ๆ เพื่อให้ได้ผลการทดลองที่มีความละเอียด

4.2 การทดลอง

การหาความแม่นยำของการประมาณค่าท่า่งท่าโดย MediaPipe ผู้วิจัยได้ใช้ผลลัพธ์ของทั้งสองระบบตามที่ได้อธิบายไป ข้อ 3.2 และใช้สูตรคำนวณหาค่าความผิดพลาดเฉลี่ยสัมบูรณ์ หรือ Mean Absolute Error (MAE) ดังสมการที่ (1) จำนวนช่วงของการเคลื่อนไหวโดยใช้ค่าองศาที่ได้จาก MoCap ดังสมการที่ (2) และหาร้อยละความแม่นยำ (Accuracy) ในแต่ละชุดการทดลองดังสมการที่ (3)

$$MAE = \frac{1}{n} \sum_{i=1}^n |Me_i - Mo_i| \quad (1)$$

$$ROM = Mo_{max} - Mo_{min} \quad (2)$$

$$Accuracy = 1 - (MAE/ROM) \quad (3)$$

โดยกำหนดให้ Mo และ Me เป็นองศาที่ได้มาจาก MoCap และ MediaPipe ตามลำดับ i เป็นดัชนีของข้อมูลทั้งสอง n เป็นจำนวนเฟรมที่ใช้ในการเปรียบเทียบ และ ROM เป็น Range of motion หรือช่วงของการเคลื่อนไหว

นอกจากการวัดความแม่นยำขององศาปกติที่ได้จากการประมาณค่าท่า่งท่า ผู้วิจัยยังได้นำข้อมูลทั้งหมดมาทำ Normalization หรือทำให้ข้อมูลขององศาจากการประมาณค่า

ทั้งทำเป็นมาตรฐานเดียวกันกับ MoCap ดังตารางที่ 1 และรูปที่ 6 ตัวอย่างของข้อมูลเมื่อนำองศาทำการ Normalize เพื่อตรวจสอบความแม่นยำเพิ่มเติมว่าหลังจากทำให้องศาเป็นมาตรฐานเดียวกัน ข้อมูลจะมีความแม่นยำมากขึ้นหรือไม่

4.3 ผลการทดลอง

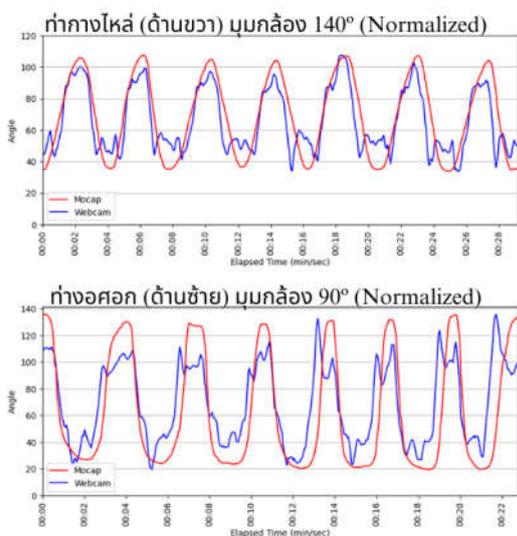
4.3.1 ผลลัพธ์จากการคำนวณสามมิติ

จากผลการทดลองทำให้ผู้วิจัยพบว่าการคำนวณทั้งทำฟิสิกส์สามมิติของ MediaPipe มีความแปรปรวนเป็นอย่างมาก

ตารางที่ 1. แสดงสถิติต่าง ๆ โดยเฉลี่ยของท่วงท่าที่ใช้ในการทดลองจากผู้เข้าร่วมการทดลองทั้งหมดห้าคน โดยแบ่งเป็นด้านซ้ายและด้านขวา แบ่งตามการคำนวณองศาแบบปกติและองศาแบบ Normalize และแบ่งตามมุมกล้องทั้งสาม

สถิติ	ท่วงท่า	ด้านของร่างกายในแต่ละมุมของเว็บแคม องศาปกติและหลังจาก Normalize											
		ซ้าย			ขวา			ซ้าย (Normalized)			ขวา (Normalized)		
		40°	90°	140°	40°	90°	140°	40°	90°	140°	40°	90°	140°
Mean	งอไหล่ขึ้น	31	32	22	23	31	28	8	10	4	4	11	6
Absolute	กางไหล่	26	18	30	30	18	29	8	4	9	8	5	11
Error	งอข้อศอก	15	46	21	20	47	13	6	16	7	8	16	7
Accuracy	งอไหล่ขึ้น	42	40	59	60	46	53	86	82	92	93	81	89
Percentage	กางไหล่	60	73	54	56	73	58	88	93	87	89	93	84
	งอข้อศอก	84	51	77	81	54	87	94	83	92	92	84	93

แม้จะ Smooth และ Normalize ข้อมูลแล้วดังตัวอย่างกราฟที่แสดงในรูปที่ 7 เนื่องจากการใช้ ML ประมาณค่าท่วงท่าในแนวแกนลึก และเป็นขีดจำกัดของอุปกรณ์เว็บแคมที่จับได้เพียงภาพสองมิติ จึงทำให้การประมาณค่าท่วงท่าในระนาบตัดแนวตั้ง (Sagittal Plane) ของคอมพิวเตอร์วิทัศน์จึงไม่มีประสิทธิภาพเท่ากับอุปกรณ์หรือฮาร์ดแวร์ที่มีการติดตั้งเซนเซอร์ตรวจจับความลึก



รูปที่ 7. ตัวอย่างของข้อมูลองศาสามมิติ

แต่ละมุมกล้อง ถ้าหากวิเคราะห์ด้วยสีพื้นหลังจากตารางดังกล่าวว่ามุมกล้องใดสามารถประมาณค่าท่วงท่าในแต่ละด้านได้ร้อยละความแม่นยำมากที่สุด ก็จะพบได้ว่าจากตารางที่ 1 จะมี Pattern หรือรูปแบบชัดเจนว่าผลลัพธ์ความแม่นยำที่ได้จากการคำนวณองศาสองมิติและหลังจาก Normalization ผลลัพธ์ในแต่ละมุมกล้องมีความสัมพันธ์อย่างมีนัยสำคัญ ซึ่งผู้วิจัยจะอธิบายต่อไปในย่อหน้าถัดจากนี้

องศาของท่วงท่ากางไหล่จะมีความแม่นยำมากที่สุดเมื่อมุมกล้องหันเข้าหาผู้เข้าร่วมการทดลอง 90° เนื่องจากรวมและการเคลื่อนไหวของท่วงท่าขนานกับการมองเห็นของกล้องพอดี กลับกันท่วงท่าที่ใช้การยืดส่วนของร่างกายออกไปด้านหน้าอย่างท่วงท่างอข้อศอกและงอไหล่ขึ้น กลับมีความแม่นยำต่อการคำนวณองศาในมุมด้านข้างหรือมุมเฉียงมากกว่ามุมกล้องที่อยู่ด้านหน้า เพราะส่วนของร่างกายได้ยืดออกไปทางด้านหน้าและขนานกับการจับภาพของมุมกล้องด้านข้าง อย่างไรก็ตามท่วงท่างอไหล่ขึ้นกลับได้ความแม่นยำของด้านตรงข้ามมากกว่าด้านของร่างกายด้านเดียวกับกล้อง นั่นเป็นเพราะผลลัพธ์ของการคำนวณองศาสองมิติในส่วนนั้นจะได้แคบกว่าความเป็นจริงจึงทำให้ความแม่นยำลดลง

นอกจากนี้การทำ Normalization กับข้อมูลสองมิติได้ให้ผลลัพธ์การประมวลค่าท่วงท่าที่แม่นยำสูงขึ้นของทุกท่วงท่าจนความแม่นยำเกินร้อยละ 90 ดังตารางที่ 1 อีกทั้งลำดับความแม่นยำยังคงเป็นลำดับเดียวกับข้อมูลสองมิติจึงแสดงให้เห็นถึงความสัมพันธ์ของข้อมูลที่ไม่เปลี่ยนแปลง

5. อภิปรายและสรุปผลการวิจัย

งานวิจัยนี้เน้นการทดลองความแม่นยำของเทคโนโลยีประมวลค่าท่วงท่าเมื่อเปรียบเทียบกับเทคโนโลยีจับการเคลื่อนไหวของร่างกายด้วยท่วงท่าที่ออกแบบมาจาก FMA-UE เพื่อตรวจสอบความเป็นไปได้ในการใช้เทคโนโลยีประมวลค่าท่วงท่าและ MediaPipe สร้างสื่อเชิงตอบโต้ แอปพลิเคชันหรือเกมออกกำลังกายเพื่อใช้ในการฟื้นฟูสมรรถภาพ จูงใจให้ผู้ป่วยออกกำลังกายมากขึ้น การประเมินสมรรถนะ และเปิดโอกาสให้ผู้ป่วยโรคหลอดเลือดสมองได้เข้ารับการรักษาและฟื้นฟูสมรรถภาพทางไกลจากที่บ้าน

จากผลการทดลองสรุปได้ว่าเทคโนโลยีการประมวลค่าท่วงท่าโดยโปรแกรม MediaPipe การประมวลค่าและคำนวณองศาโดยใช้พิกัดแนวลึกหรือ Z-Axis มีความแปรปรวนมากเกินไป ไม่เหมาะกับการนำมาใช้สำหรับการประเมินผู้ป่วย และไม่เหมาะกับการใช้เป็นอินพุตเพื่อออกคำสั่งภายในเกมออกกำลังกาย เพราะหากส่วนของร่างกายใด ๆ โคนบบังแม้เพียงเล็กน้อยก็อาจจะทำให้การประมวลค่าในแนวแกนลึกผิดพลาดได้ ส่วนการคำนวณแบบสองมิติ ร่างกายที่ขนานกับกล้องจะได้ผลลัพธ์ที่แม่นยำที่สุด แต่โดยรวมแล้วยังคงมีร้อยละความแม่นยำที่ต่ำ เนื่องจากช่วงของการเคลื่อนไหวที่จับได้กว้างหรือแคบเกินจริง อย่างไรก็ตาม การใช้ Normalization เพื่อให้ข้อมูลองศาเป็นมาตรฐานเดียวกันกับ MoCap พบว่ามีความแม่นยำที่เพิ่มสูงขึ้นเป็นอย่างมาก เพราะ Normalization ช่วยปรับมาตรฐานช่วงขององศาให้เป็นมาตรฐานที่ถูกต้อง และทำให้สามารถใช้ MediaPipe เพื่อประเมินหรือช่วยแพทย์ประเมินผู้ป่วยในขณะที่ทำกายภาพบำบัดด้วยท่วงท่าที่ออกแบบมาจาก FMA-UE

การคำนวณองศาแบบสองมิติผนวกกับการ Normalization จึงเหมาะสมเป็นอย่างมากในการพัฒนาแอปพลิเคชันหรือเกมออกกำลังกาย เพื่อการฟื้นฟูสมรรถภาพจากที่บ้านหรือการเข้ารับการรักษาทางไกล และลดจำนวนผู้ป่วยที่ขาดการเข้ารับการรักษา แต่วิธีการประเมินผู้ป่วยเบื้องต้นในระหว่างใช้งานเกม

ออกกำลังกายผ่านเว็บไซต์ จำเป็นจะต้องให้ข้อมูลระหว่างการถ่ายภาพบำบัดในรูปแบบของวิดีโอ และแสดงกราฟองศาแต่ละส่วนของร่างกายเป็นข้อมูลประกอบ เพื่อให้แพทย์หรือผู้ดูแลทำการวินิจฉัยได้ด้วยตนเองต่อไป

6. ข้อเสนอแนะ

การทดลองของงานวิจัยนี้คือการหาร้อยละความแม่นยำของ MediaPipe เมื่อเปรียบเทียบกับเทคโนโลยีตรวจจับการเคลื่อนไหวของร่างกาย การทดลองปัจจุบันผู้วิจัยใช้ Moving Average ซึ่งอาจจะมีประสิทธิภาพไม่พอ ถ้าหากต้องการเพิ่มแนวโน้มที่จะนำเทคโนโลยีดังกล่าวเพื่อการประเมินผู้ป่วยที่กระทำท่วงท่าที่มีแนวแกนลึก อาจจะต้องมีการเพิ่มเทคนิคการคาดคะเนหรือการกรองข้อมูล (Filter) เพื่อเพิ่มความแม่นยำของการประมวลค่าท่วงท่า และการออกแบบท่วงท่าที่ใช้ในการประเมินหรือการออกแบบการประเมินผู้ป่วยอาจจะต้องเป็นท่วงท่าที่เน้นการเคลื่อนไหวระนาบในแนวตั้งและไม่มีการเคลื่อนไหวในแนวลึกเกี่ยวข้อง หรือใช้เทคนิคอื่นอย่างการตรวจจับสิ่งของร่วมเพื่อเพิ่มความแม่นยำ

เอกสารอ้างอิง

- [1] สมศักดิ์ เทียมเก่า, “อุบัติเหตุโรครหลอดเลือดสมองประเทศไทย,” *วารสารประสาทวิทยาแห่งประเทศไทย*, vol. 37, p. 54–60, 2021.
- [2] กองโรคไม่ติดต่อหรือสำนักสื่อสารความเสี่ยงฯ กรมควบคุมโรค, “วันอัมพาตโลก 2565 เน้นสร้างความตระหนักรู้เกี่ยวกับสัญญาณเตือนโรคหลอดเลือดสมองให้กับประชาชน,” [ออนไลน์], 2565. เข้าถึงได้: <https://pr.moph.go.th/?url=pr/detail/all/02/180623>. [เข้าถึงเมื่อ 27 เมษายน 2566].
- [3] ตรีนุช อมรภิญโญเกียรติ, “บทบาทของการใช้หุ่นยนต์ฟื้นฟูผู้ป่วยโรคหลอดเลือดสมองในโรงพยาบาล ดากสิน,” *วารสารสาธารณสุขและวิทยาศาสตร์สุขภาพ*, vol. 5, p. 160–171, 2022.
- [4] ปิยธิดา ชูรักษ์, “กรอบแนวคิดการกายภาพบำบัดผู้ป่วยโรคหลอดเลือดสมองในชุมชน,” *Thaksin University Online Journal*, vol. 1, p. 65, 2022.

- [5] J. Adolf *et al.*, “Automatic telerehabilitation system in a home environment using computer vision.,” in *pHealth*, pp. 142–148, 2020.
- [6] G. Bradski and A. Kaehler, *Learning OpenCV: Computer vision with the OpenCV library*. “O’Reilly Media, Inc.”, 2008.
- [7] N. Nakano *et al.*, “Evaluation of 3D markerless motion capture accuracy using OpenPose with multiple video cameras,” *Frontiers in sports and active living*, vol. 2, art. 50, 2020.
- [8] M. Nour, M. Gardoni, J. Renaud and S. Gauthier, “Real-time detection and motivation of eating activity in elderly people with dementia using pose estimation with TensorFlow and OpenCV,” *Adv. Soc. Sci. Res. J*, vol. 8, p. 28–34, 2021.
- [9] D. M. Kishore, S. Bindu and N. K. Manjunath, “Estimation of yoga postures using machine learning techniques,” *International Journal of Yoga*, vol. 15, no. 2, p. 137 - 143, 2022.
- [10] C. El-Habr, X. Garcia, P. Paliyawan and R. Thawonmas, “Runner: A 2D platform game for physical health promotion,” *SoftwareX*, vol. 10, p. 100329, 2019.
- [11] J. W. Burke *et al.*, “Vision based games for upper-limb stroke rehabilitation,” in *2008 International Machine Vision and Image Processing Conference*, pp. 159–164, IEEE, 2008.
- [12] L. E. Sucar, R. Luis, R. Leder, J. Hernández and I. Sánchez, “Gesture therapy: A vision-based system for upper extremity stroke rehabilitation,” in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, 2010.
- [13] C. Lugaresi *et al.*, “Mediapipe: A framework for building perception pipelines,” *arXiv preprint arXiv:1906.08172*, 2019.
- [14] A. Latreche *et al.*, “Reliability and validity analysis of MediaPipe-based measurement system for some human rehabilitation motions,” *Measurement*, vol. 214, p. 112826, 2023.
- [15] R. Sers *et al.*, “Validity of the Perception Neuron inertial motion capture system for upper body motion analysis,” *Measurement*, vol. 149, p. 107024, 2020.
- [16] P. Palani, S. Panigrahi, S. A. Jammi and A. Thondiyath, “Real-time Joint Angle Estimation using Mediapipe Framework and Inertial Sensors,” in *2022 IEEE 22nd International Conference on Bioinformatics and Bioengineering (BIBE)*, 2022.
- [17] A. Fernández-Baena, A. Susin and X. Lligadas, “Biomechanical validation of upper-body and lower-body joint movements of kinect motion capture data for rehabilitation treatments,” in *2012 fourth international conference on intelligent networking and collaborative systems*, 2012.
- [18] M. Yahya *et al.*, “Motion capture sensing techniques used in human upper limb motion: A review,” *Sensor Review*, vol. 39, p. 504–511, 2019.
- [19] A. R. Fugl-Meyer *et al.*, “A method for evaluation of physical performance,” *Scand. J. Rehabil. Med*, vol. 7, p. 13–31, 1975.
- [20] I. Hen *et al.*, “The dynamics of spatial behavior: how can robust smoothing techniques help?,” *Journal of neuroscience methods*, vol. 133, p. 161–172, 2004.