



ระบบตรวจจับอุปกรณ์การเย็บแผลด้วยปัญญาประดิษฐ์

Surgical Suture Instrument Detection System

นาย อภินันท์ อายุยงค์

รหัสนักศึกษา 66112772

นางสาว ตรีฤทัย แเคยีหว่า

รหัสนักศึกษา 66120361

นางสาว ฟ้าใส ขวัญปาน

รหัสนักศึกษา 66126467

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

สาขาวิศวกรรมคอมพิวเตอร์และปัญญาประดิษฐ์

สำนักวิชาวิศวกรรมศาสตร์และเทคโนโลยี มหาวิทยาลัยวลัยลักษณ์

ปีการศึกษา 2568

คำนำ

โครงการฉบับนี้จัดทำขึ้นเพื่อนำเสนอแนวทางการประยุกต์ใช้เทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence) ในการยกระดับมาตรฐานการศึกษาทางการแพทย์ โดยเน้นไปที่การฝึกทักษะการเย็บแผล ซึ่งเป็นทักษะพื้นฐานที่สำคัญยิ่งสำหรับศัลยแพทย์และบุคลากรทางการแพทย์

ในปัจจุบัน การประเมินทักษะการผ่าตัดส่วนใหญ่ยังอาศัยดุลยพินิจของผู้เชี่ยวชาญเป็นหลัก คณะผู้จัดทำจึงได้พัฒนาระบบที่ใช้เทคนิคการประมวลผลภาพ (Computer Vision) ด้วยโมเดล Mask R-CNN ร่วมกับการวิเคราะห์ลำดับขั้นตอนด้วย Transformer เพื่อให้เกิดการประเมินผลที่เป็นกลาง แม่นยำ และสามารถให้คำแนะนำแก่ผู้ฝึกฝนได้ทันที

คณะผู้จัดทำหวังเป็นอย่างยิ่งว่า โครงการนี้จะเป็นก้าวสำคัญในการพัฒนาเครื่องมือสนับสนุนการเรียนการสอนทางการแพทย์ และเป็นประโยชน์ต่อวงการสาธารณสุขในอนาคต หากมีข้อผิดพลาดประการใด คณะผู้จัดทำขออภัยมา ณ ที่นี้

คณะผู้จัดทำ

บทที่ 1

บทนำ

1.1 ความสำคัญและที่มาของปัญหา

การฝึกฝนทักษะศัลยกรรมและการผ่าตัดในปัจจุบันยังคงพึ่งพารูปแบบการเรียนรู้แบบดั้งเดิมที่เรียกว่า Apprenticeship Model หรือระบบศิษย์มีครู ซึ่งเน้นให้ผู้เรียนเรียนรู้ผ่านการสังเกตการณ์และรับคำแนะนำจากอาจารย์แพทย์ผู้เชี่ยวชาญ แม้จะเป็นวิธีที่ทรงประสิทธิภาพในเชิงปฏิบัติ แต่ระบบนี้มีข้อจำกัดเชิงโครงสร้างหลายประการ โดยเฉพาะอย่างยิ่งในด้านความเป็นปรนัยของการวัดผล (Objectivity) เนื่องจากเกณฑ์การประเมินมักขึ้นอยู่กับดุลยพินิจส่วนบุคคล (Subjective Assessment) นำไปสู่ปัญหาความไม่สม่ำเสมอของมาตรฐาน ความล่าช้าในการให้ข้อมูลป้อนกลับ (Delayed Feedback) และปัญหาความเอนเอียง (Bias) ของผู้ประเมิน ปัจจัยเหล่านี้ส่งผลโดยตรงต่อการพัฒนาทักษะของผู้เรียนให้ได้ตามมาตรฐานในระยะเวลาที่กำหนดและอาจทำให้มาตรฐานการฝึกฝนมีความคลาดเคลื่อน

ประเด็นสำคัญที่สุดคือ ความท้าทายด้านความเป็นปรนัยของการวัดผล (Objectivity) เนื่องจากการให้คะแนนหรือการวิพากษ์ทักษะมักขึ้นอยู่กับดุลยพินิจส่วนบุคคลของผู้ประเมิน (Subjective Assessment) ส่งผลให้เกิดความไม่สม่ำเสมอของมาตรฐานการฝึกฝนในแต่ละบุคคล นำไปสู่ปัญหาความล่าช้าในการให้ข้อมูลป้อนกลับ (Delayed Feedback) ซึ่งมักจะเกิดขึ้นภายหลังจากการเสร็จสิ้นการปฏิบัติงานไปแล้ว รวมถึงปัญหาความเอนเอียง (Bias) ของผู้ประเมินที่อาจเกิดขึ้นโดยไม่ตั้งใจ ปัจจัยเหล่านี้ส่งผลกระทบโดยตรงต่อประสิทธิภาพในการพัฒนาทักษะของผู้เรียนให้ได้ตามมาตรฐานวิชาชีพภายในกรอบเวลาที่กำหนด และอาจทำให้ผลการฝึกฝนมีความคลาดเคลื่อนจากความเป็นจริง

ด้วยเหตุนี้ คณะผู้จัดทำจึงเล็งเห็นถึงความจำเป็นในการบูรณาการเทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence) เข้ามาเป็นเครื่องมือสนับสนุนการเรียนการสอนทางการแพทย์ยุคใหม่ โดยมุ่งเน้นการพัฒนาระบบที่ใช้เทคนิคการประมวลผลภาพ (Computer Vision) และการเรียนรู้เชิงลึก (Deep Learning) ด้วยแบบจำลอง Mask R-CNN สำหรับการตรวจจับและระบุตำแหน่งอุปกรณ์การ

เย็บแผลในลักษณะ Instance Segmentation ซึ่งมีขีดความสามารถในการจำแนกขอบเขตของอุปกรณ์เย็บแผลแต่ละชิ้นได้อย่างละเอียดแม่นยำถึงระดับพิกเซล แม้ในสภาพแวดล้อมการทำงานที่มีความซับซ้อน เช่น สถานะที่มีการบดบังของวัตถุ (Occlusion) จากมือของผู้ปฏิบัติงาน หรือสถานะที่มีการเคลื่อนไหวของอุปกรณ์อย่างรวดเร็วต่อเนื่อง

อย่างไรก็ตาม เนื่องจากการเย็บแผลเป็นทักษะวิชาชีพที่ประกอบด้วยลำดับขั้นตอนต่อเนื่องที่มีความละเอียดซับซ้อน (Surgical Workflow) การตรวจจับเพียงตำแหน่งเชิงพื้นที่ของวัตถุจึงไม่เพียงพอต่อการตัดสินใจระดับความเชี่ยวชาญ คณะผู้จัดทำจึงได้ผสานโมเดลวิเคราะห์ข้อมูลเชิงลำดับ (Sequential Analysis) อาทิ Transformer เข้ามาเพื่อประมวลผลความสัมพันธ์เชิงเวลา (Temporal Relationship) ของแต่ละการกระทำ เพื่อประเมินความถูกต้องของขั้นตอนการปฏิบัติงานตามหลักมาตรฐานสากล การพัฒนาระบบดังกล่าวจะช่วยสร้างบรรทัดฐานใหม่ในการประเมินทักษะที่สามารถตรวจสอบย้อนหลังได้ (Traceable Standards) ส่งเสริมการเรียนรู้ด้วยตนเองอย่างมีประสิทธิภาพ และลดภาระงานในการสังเกตการณ์ของอาจารย์แพทย์ในระยะยาวได้อย่างยั่งยืน

1.2 วัตถุประสงค์ของโครงการ

เพื่อให้การดำเนินงานเป็นไปตามเป้าหมายที่วางไว้ คณะผู้จัดทำได้กำหนดวัตถุประสงค์หลักของโครงการไว้ดังนี้

1. เพื่อวิจัยและพัฒนาระบบตรวจจับและแบ่งส่วนวัตถุ (Instance Segmentation) สำหรับคัดแยกและระบุตำแหน่งอุปกรณ์ทางการแพทย์ที่เกี่ยวข้องกับการเย็บแผล เช่น ปลายคีมจับเข็ม เข็มเย็บแผล และคีมคีบ (Forceps) ได้อย่างแม่นยำระดับพิกเซล โดยใช้แบบจำลอง Mask R-CNN
2. เพื่อพัฒนาระบบวิเคราะห์ลำดับขั้นตอนอัตโนมัติ (Sequential Modeling) โดยการประยุกต์ใช้โมเดล Transformer ในการประมวลผลข้อมูลความสัมพันธ์เชิงเวลา เพื่อตรวจสอบความถูกต้องของลำดับขั้นตอนในการเย็บแผลตามหลักจริยธรรมและมาตรฐานทางการแพทย์
3. เพื่อสร้างเครื่องมือประเมินผลทักษะศัลยกรรมที่เป็นกลาง (Objective Evaluation) ที่สามารถประมวลผลและให้ข้อมูลป้อนกลับ (Feedback) แก่ผู้เรียนได้ในทันที เพื่อลดความ

เอนเอียงที่เกิดจากอคติของตัวบุคคล และเสริมสร้างประสิทธิภาพในการฝึกฝนทักษะด้วยตนเอง

4. เพื่อประยุกต์ใช้ฐานข้อมูลวิดีโอของขั้นตอนการเย็บแผล ทั้งรูปแบบการปฏิบัติที่ “ถูกต้อง” และรูปแบบที่ “ควรปรับปรุง” ในการสร้างแนวทางการฝึกทักษะตามรูปแบบ Apprenticeship Model ที่มีความแม่นยำและมีประสิทธิภาพ
5. เพื่อพัฒนาระบบที่เอื้ออำนวยให้ทั้งผู้เรียนและอาจารย์แพทย์สามารถตรวจสอบขั้นตอนการเย็บแผลแบบอัตโนมัติ นำไปสู่การปรับปรุงและแก้ไขจุดบกพร่องในการฝึกปฏิบัติ
6. ใช้ฐานข้อมูลวิดีโอของขั้นตอนการเย็บแผลทั้งแบบ “ถูกต้อง” และ “คลิป์ที่ต้องปรับปรุง” เพื่อสร้างแนวทางฝึกทักษะแบบ Apprenticeship Model ที่แม่นยำและมีประสิทธิภาพ
7. รวมผลจากทั้ง 2 โมเดล (ตรวจจับอุปกรณ์ + วิเคราะห์ลำดับ) เพื่อสรุปการประเมิน และสร้างฐานข้อมูลสำหรับฝึกทักษะแบบ Apprenticeship Model

1.3 ขอบเขตของงานวิจัย

คณะผู้จัดทำได้กำหนดขอบเขตของการดำเนินงานวิจัยเพื่อให้ครอบคลุมเป้าหมาย สอดคล้องกับสถานการณ์การใช้งานจริงในช่วงการนำไปใช้งาน (Deployment) ดังนี้

1. ผู้เกี่ยวข้องและสภาพแวดล้อมการใช้งาน

- ในระยะการนำระบบไปใช้งานจริง ระบบมุ่งเน้นกลุ่มผู้ใช้งานหลัก 1 กลุ่ม ได้แก่ ผู้ใช้งานทั่วไป (User) หรือผู้รับบริการ (Trainee/Customer) เช่น นักศึกษาแพทย์ หรือนักศึกษาสัตวแพทย์
- ผู้ใช้งานจะทำหน้าที่ปฏิสัมพันธ์กับระบบเพื่อเริ่มกระบวนการฝึกปฏิบัติ โดยระบบจะทำการประมวลผลข้อมูลจากกล้องหรือไฟล์วิดีโอที่นำเข้า เพื่อวิเคราะห์และส่งคืนผลการประเมินในรูปแบบคะแนนหรือร้อยละความถูกต้อง พร้อมข้อมูลข้อเสนอแนะ (Feedback) เพื่อใช้เป็นแนวทางพัฒนาทักษะ

2. ชุดข้อมูลและคุณสมบัติของข้อมูลนำเข้า (Dataset)

- ระบบรองรับข้อมูลนำเข้าประเภทไฟล์วิดีโอ (Video Clips) ที่บันทึกขั้นตอนการฝึกปฏิบัติการเย็บแผลทางการแพทย์ (Surgical Suturing)
- วิดีโอที่ใช้ต้องมีความละเอียดและค่าความสว่างที่เพียงพอต่อการจำแนกวัตถุ โดยมีข้อกำหนดเรื่องมุมมองภาพที่ต้องเป็นมุมมองจากด้านบน (Top-down View) หรือมุมมองที่เห็นตำแหน่งอุปกรณ์และเป้าหมายการเย็บได้อย่างชัดเจน

3. ขอบเขตด้านการตรวจจับและแบ่งส่วนวัตถุ (Instance Segmentation)

- ระบบมุ่งเน้นการตรวจจับและระบุขอบเขตพิกเซลของอุปกรณ์ศัลยกรรมหลัก 3 ชนิด ได้แก่
 1. คีมจับเข็ม (Needle Holder)
 2. เข็มเย็บแผล (Surgical Needle)
 3. คีมคีบ (Forceps)
- ใช้สถาปัตยกรรมแบบจำลอง Mask R-CNN เป็นโครงสร้างหลักในการประมวลผลภาพ

4. ขอบเขตด้านการวิเคราะห์ลำดับขั้นตอน (Sequential Analysis)

- ระบบจะทำการวิเคราะห์เหตุการณ์สำคัญ (Key Events) ในกระบวนการเย็บแผล ซึ่งประกอบด้วย 3 ระยะหลัก ได้แก่ การจับเข็ม (Needle Loading), การแทงเข็ม (Needle Insertion) และการถอนเข็ม (Needle Exit)
- ใช้แบบจำลอง Transformer ในการวิเคราะห์ความสัมพันธ์เชิงเวลาจากข้อมูลพิกัดและคุณลักษณะที่ส่งต่อมาจากส่วนการตรวจจับวัตถุ

5. ขอบเขตด้านการวัดผลและเกณฑ์การประเมิน (Evaluation & Metrics)

- ประเมินความแม่นยำของโมเดล Mask R-CNN ผ่านค่าทางสถิติ mean Average Precision (mAP)
- ประเมินความถูกต้องเชิงลำดับขั้นตอนด้วยการเปรียบเทียบเชิงลึกกับข้อมูลมาตรฐานของผู้เชี่ยวชาญ (Ground Truth)

- ระบบสามารถประมวลผลและแสดงผลคะแนนความแม่นยำ (Accuracy Score) ของการปฏิบัติงานผ่านอินเทอร์เฟซผู้ใช้งาน

1.4 ผลที่คาดว่าจะได้รับเมื่อเสร็จสิ้นโครงการ

เมื่องานวิจัยชิ้นนี้เสร็จสมบูรณ์ คณะผู้จัดทำคาดหวังผลลัพธ์ที่จะเกิดขึ้นในมิติต่างๆ ดังนี้

- ด้านประสิทธิภาพการเรียนรู้ ผู้เรียนจะได้รับระบบประเมินผลอัตโนมัติที่สามารถให้ข้อมูลป้อนกลับได้ทันที ช่วยให้ระบุข้อบกพร่องในลำดับขั้นตอนได้ด้วยตนเอง นำไปสู่การพัฒนาทักษะที่รวดเร็วและแม่นยำ
- ด้านมาตรฐานการประเมิน เกิดเกณฑ์การวัดผลที่เป็นกลางและเป็นปรนัย (Objective Measurement) โดยใช้ปัญญาประดิษฐ์เป็นมาตรฐานตัดสิน ช่วยลดความเอนเอียงและสร้างบรรทัดฐานการประเมินที่โปร่งใสและตรวจสอบได้ (Traceability)
- ด้านการบริหารจัดการทรัพยากร ช่วยลดภาระงานในการเฝ้าสังเกตการณ์ของอาจารย์แพทย์ในขั้นตอนพื้นฐาน ส่งผลให้บุคลากรผู้สอนสามารถใช้เวลาไปกับการสอนในกรณีที่มีความซับซ้อนสูงหรือเคสที่มีความยากพิเศษได้มากขึ้น
- ด้านนวัตกรรมทางการแพทย์ คณะผู้จัดทำจะได้รับองค์ความรู้และต้นแบบระบบวิเคราะห์วิดีโออัจฉริยะที่ผสมผสานระหว่าง Instance Segmentation และ Sequential Modeling ซึ่งเป็นพื้นฐานสำคัญในการต่อยอดไปสู่ระบบประเมินทักษะการผ่าตัดรูปแบบอื่นๆ ในอนาคต

อุปกรณ์ที่ใช้ในการพัฒนางาน

การพัฒนา ระบบ ผู้พัฒนาได้เลือกใช้เครื่องมือที่หลากหลาย ทั้งด้านซอฟต์แวร์และฮาร์ดแวร์ เพื่อรองรับการพัฒนาโมเดล AI สำหรับการตรวจจับวัตถุ (Instance Segmentation) การวิเคราะห์ลำดับการเคลื่อนไหวด้วย Transformer รวมถึงการประเมินประสิทธิภาพของผู้ฝึกฝน โดยมีรายละเอียดดังนี้

1.7.1 ด้านซอฟต์แวร์ (Software)

ระบบปฏิบัติการและฐานข้อมูลระบบปฏิบัติการ ใช้ระบบปฏิบัติการ Windows 10 หรือ macOS เป็นสภาพแวดล้อมหลักในการพัฒนาโค้ดและทดสอบระบบ

- ฐานข้อมูล (Database): MongoDB 3.2.11 ถูกเลือกใช้เพื่อจัดเก็บข้อมูลสำคัญของระบบแบบ NoSQL ซึ่งรวมถึง ข้อมูลภาพ/วิดีโอ ที่ใช้ในการฝึกสอน, ค่าการประเมิน (Evaluation Metrics) ของโมเดล, และ ผลการฝึกซ้อม (Log) ของผู้ใช้งานแต่ละคน

เครื่องมือและเทคโนโลยีสำหรับ AI Core การทำ Data Annotation และ Preprocessing

- Roboflow (Premium License) เป็นเครื่องมือสำคัญที่ใช้สำหรับ ทำ Data Annotation อย่างมีประสิทธิภาพ, การเตรียมข้อมูลสำหรับ Instance Segmentation, และการ Deploy โมเดล

- LabelMe / CVAT ใช้ช่วยในการ ทำ Label ภาพ (Bounding Box/Segmentation Mask) และ ตรวจสอบคุณภาพข้อมูล ที่จะนำไปฝึกสอนโมเดล

โมเดล Deep Learning

- Mask R-CNN โมเดลหลักที่ใช้สำหรับการ ตรวจจับและแบ่งส่วนวัตถุ (Instance Segmentation) เพื่อระบุตำแหน่งที่แน่นอนและขอบเขตของอุปกรณ์ทางการแพทย์ (เข็มเย็บแผล, คีม) และวัตถุที่เกี่ยวข้อง (บาดแผล, มือ) ในแต่ละเฟรม

- TensorFlow เฟรมเวิร์ก Deep Learning ระดับขั้นสูงที่ใช้ในการสร้าง, ฝึกสอน (Training), และปรับปรุงประสิทธิภาพของโมเดล AI

- Transformer-based Models (Multimodal Transformer): โมเดลขั้นสูงที่ถูกนำมาใช้สำหรับการ วิเคราะห์ลำดับและขั้นตอน ของท่าทางการเย็บแผล (Surgical Gestures) โดยการประมวลผลข้อมูลลำดับที่ได้จากการตรวจจับวัตถุ

ไลบรารีสำหรับงานเฉพาะ

- OpenCV ไลบรารีสำหรับงาน Computer Vision ใช้ในการจัดการ, ประมวลผล, และ วิเคราะห์ข้อมูลภาพและวิดีโอที่ใช้ในการฝึกฝน

- Pandas / NumPy / Matplotlib ไลบรารี Python มาตรฐานสำหรับการจัดการ ข้อมูลเชิงตัวเลข, การ วิเคราะห์เชิงสถิติ, และการ แสดงผลข้อมูล (เช่น กราฟประสิทธิภาพโมเดล และตารางสรุปผล

4. ซอฟต์แวร์สำหรับจัดทำเอกสารและการนำเสนอ

- Microsoft Office (Word, Excel, PowerPoint) ใช้สำหรับจัดทำรายงาน, ตารางสรุปผล, และเอกสารนำเสนอโครงการอย่างเป็นทางการ

- Google Docs และ Google Drive ใช้สำหรับจัดทำเอกสารและการ ทำงานร่วมกัน (Collaboration) ในทีมแบบเรียลไทม์

ด้านฮาร์ดแวร์ (Hardware)

การพัฒนาระบบ ผู้พัฒนาได้ใช้เครื่องคอมพิวเตอร์สำหรับพัฒนา ซึ่งคอมพิวเตอร์ทั้ง 3 เครื่องของผู้พัฒนา มีรายละเอียดคุณสมบัติทางฮาร์ดแวร์ได้ดังนี้

- acer รุ่น - Nitro AN515-46

CPU: AMD Ryzen 7 6800H with Radeon Graphics (3.20 GHz)

GPU: NVIDIA GeForce RTX 3050 Tiptop GPU

RAM: 16 GB

OS: Windows 10/11 64-bit

- Lenovo รุ่น - DESKTOP-9KEB509

CPU: Intel Core i5-11320H 11th Gen (3.2 GHz)

GPU: 4 GB (Multiple GPUs installed)

RAM: 16 GB

OS: Windows 10/11 64-bit

- acer รุ่น - Nitro 5

CPU: 12th Gen Intel Core i7-12700H (2.70 GHz)

GPU: NVIDIA GeForce RTX 3050 Laptop GPU

RAM: 8 GB

OS: Windows 64-bit

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในการพัฒนาโครงงานระบบตรวจจับอุปกรณ์และวิเคราะห์ลำดับขั้นตอนการเย็บแผลอัตโนมัติ คณะผู้จัดทำได้ศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้องเพื่อเป็นแนวทางในการดำเนินงาน ดังนี้

2.1 เทคโนโลยีการแบ่งส่วนภาพแบบ Instance Segmentation และ Mask R-CNN การตรวจจับวัตถุในงานศัลยกรรมมีความซับซ้อนสูง เนื่องจากอุปกรณ์ทางการแพทย์มีลักษณะเรียวยาวและมักมีการวางซ้อนทับหรือบดบังกันเองในขณะปฏิบัติงาน คณะผู้จัดทำจึงเลือกใช้ Mask R-CNN ซึ่งเป็นสถาปัตยกรรมที่พัฒนาต่อยอดมาจาก Faster R-CNN เพื่อเพิ่มประสิทธิภาพในการจำแนกวัตถุ

- หลักการทำงาน Mask R-CNN ไม่เพียงแต่สร้างกรอบล้อมรอบวัตถุ (Bounding Box) เท่านั้น แต่ยังสามารถสร้าง Mask หรือการระบุพิกเซลที่ประกอบเป็นวัตถุนั้น ๆ ได้อย่างแม่นยำในระดับพิกเซล
- ส่วนประกอบสำคัญ (Heads) ส่วนท้ายของโมเดลจะแยกการทำงานออกเป็น 3 ทางหลัก ได้แก่ การจำแนกชนิดวัตถุ (Classification), การปรับกรอบล้อมรอบ (Bounding Box Regression) และการสร้างหน้ากากพิกเซล (Mask Prediction)
- ประสิทธิภาพในการระบุตำแหน่ง โครงงานเลือกใช้โมเดลนี้เนื่องจากมีความแม่นยำสูงในการระบุตำแหน่งวัตถุที่มีความซับซ้อน แม้ในสภาวะที่มีการวางทับซ้อนกัน โดยสามารถระบุได้ทั้งกรอบและรูปร่างพิกเซลของอุปกรณ์
- การวิเคราะห์เชิงลำดับ (Sequential Modeling) นอกเหนือจากการตรวจจับภาพนิ่ง ระบบยังใช้หลักการวิเคราะห์ข้อมูลแบบอนุกรมเวลา (Time-series) เพื่อทำความเข้าใจลำดับท่าทางการเย็บแผลว่าถูกต้องตามขั้นตอนมาตรฐานหรือไม่

2.2 การวิเคราะห์ลำดับข้อมูลเชิงเวลาด้วย Transformer ในการรับรู้ลำดับขั้นตอนการเย็บแผล (Surgical Phase Recognition) ระบบไม่สามารถตัดสินใจได้จากภาพนิ่งเพียงเฟรมเดียว แต่ต้องอาศัยการวิเคราะห์พฤติกรรมที่เกิดขึ้นต่อเนื่องกัน คณะผู้จัดทำจึงเลือกใช้สถาปัตยกรรม Transformer ซึ่งมีกลไกสำคัญดังนี้

- Self-Attention Mechanism เป็นหัวใจหลักที่ช่วยให้โมเดลสามารถคำนวณน้ำหนักหรือความสำคัญของเหตุการณ์ในแต่ละช่วงเวลาได้ ทำให้ระบบเข้าใจความสัมพันธ์ระหว่างพฤติกรรมปัจจุบันกับสถานะที่เกิดขึ้นในอดีต เช่น ข้อมูลจาก 50 เฟรมก่อนหน้า
- Positional Encoding เนื่องจาก Transformer ประมวลผลข้อมูลแบบขนาน (Parallel) จึงจำเป็นต้องมีกลไกนี้เพื่อระบุ "ลำดับเวลา" (Time Steps) เพื่อให้โมเดลทราบลำดับเหตุการณ์ก่อนหลัง ซึ่งสำคัญต่อการประเมินว่าผู้เรียนปฏิบัติข้ามขั้นตอนหรือไม่
- Temporal Dependency Transformer สามารถจับความสัมพันธ์ระยะไกล (Long-range Dependencies) ได้ดีกว่าโมเดลแบบ RNN หรือ LSTM จึงมีความเหมาะสมอย่างยิ่งสำหรับการวิเคราะห์วิดีโอการผ่าตัดที่มีความยาวและมีความซับซ้อนสูง

2.3 งานวิจัยที่เกี่ยวข้อง

คณะผู้จัดทำได้รวบรวมและวิเคราะห์งานวิจัยที่เกี่ยวข้องเพื่อเป็นแนวทางในการพัฒนาระบบ ดังนี้

1. งานวิจัยด้านการตรวจจับเครื่องมือ (Surgical Instrument Detection)

Laina et al. ทดลองใช้ Fully Convolutional Networks (FCNs) เพื่อแยกส่วนเครื่องมือแพทย์ โดยพบว่าการใช้ข้อจำกัดทางพื้นที่ (Spatial Constraints) ช่วยลดการตรวจจับที่ผิดพลาดได้ เช่น เข็ม มักอยู่ใกล้ปลายคีมจับเสมอ

YOLO-Suturing เน้นความเร็วและความแม่นยำด้วยโมเดลตระกูล YOLO (v5/v8) โดยพบว่าการทำ Data Augmentation ด้วยการปรับ "แสงสะท้อน" (Specular Reflection) บนโลหะ ช่วยเพิ่มความแม่นยำขึ้นร้อยละ 20 ในสภาพแสงที่ต่างกัน

2. งานวิจัยด้านการจำแนกท่าทาง (Surgical Gesture Recognition)

The JIGSAWS Dataset Paper รายงานสำคัญที่สร้างชุดข้อมูลสำหรับการเย็บแผล โดยใช้เทคนิค HMMs หรือ TCNs เพื่อวิเคราะห์ลำดับการเคลื่อนที่ โครงการนี้สามารถอ้างอิงเกณฑ์การแบ่งท่าทาง

(Surgical Vocabulary) เช่น การจับเข็ม (Reaching for needle) หรือการจัดวางเข็ม (Positioning needle) ได้จากงานวิจัยนี้

3. งานวิจัยด้านการประเมินทักษะอัตโนมัติ (Automated Skill Assessment) Zia et al.นำเสนอการวัดผลด้วยค่า Kinematics ผ่านวิดีโอ ตัวชี้วัดสำคัญประกอบด้วย ระยะทางรวม (Path Length), ความราบรื่นของการเคลื่อนที่ (Motion Fluidity/Jerk) และเวลาที่ใช้ในแต่ละขั้นตอน (Completion Time)

4. งานวิจัยด้านการประมวลผลวิดีโอและการให้คำแนะนำ (Video Analysis & XAI)

Islam et al. (2023) ใช้สถาปัตยกรรมแบบผสมผสานในการวิเคราะห์วิดีโอผ่าตัด โดยใช้ Attention Mechanism ใน Transformer เพื่อโฟกัสเฉพาะเฟรมสำคัญ (Key Frames) และตัดข้อมูลรบกวน (Noise) ออกไป ส่งผลให้ความแม่นยำเพิ่มขึ้นร้อยละ 15 เมื่อเทียบกับวิธีเดิม

Explainable AI (XAI) ศึกษาการใช้ Attention Maps เพื่อแสดงให้เห็นจุดที่ถูกหักคะแนนผ่าน Visual Feedback (เช่น วงกลมสีแดง) ซึ่งช่วยให้ผู้ฝึกพัฒนาทักษะได้รวดเร็วกว่าการได้รับเพียงคะแนนรวมถึงร้อยละ 30

5. การให้ข้อมูลป้อนกลับแบบเรียลไทม์ (Real-time Feedback Systems)

Lin et al. (2025) พัฒนาระบบที่เน้นความเร็วในการประมวลผล (Inference Time) โดยกำหนดความล่าช้าไม่เกิน 200 มิลลิวินาที มีการปรับปรุง Backbone ของ Mask R-CNN ให้เล็กลงแต่ยังคงความแม่นยำ ซึ่งเป็นแนวทางสำคัญสำหรับการพัฒนาแอปพลิเคชันเพื่อการศึกษา

บทที่ 3

การวิเคราะห์และออกแบบระบบ

ในการพัฒนาระบบตรวจจับอุปกรณ์การเย็บแผล คณะผู้จัดทำได้ดำเนินการวิเคราะห์ความต้องการและออกแบบสถาปัตยกรรม เพื่อให้เห็นโครงสร้างการทำงานและปฏิสัมพันธ์ระหว่างผู้ใช้กับระบบอย่างชัดเจน โดยมีรายละเอียดดังนี้

3.1 ภาพรวมของการวิเคราะห์ระบบ

3.1.1 ขอบเขตงาน (Scope of Work) โครงการนี้มุ่งเน้นการพัฒนาระบบอัตโนมัติเพื่อสนับสนุนการฝึกทักษะศัลยกรรมพื้นฐาน โดยใช้เทคโนโลยี Computer Vision ในการระบุอุปกรณ์ทางการแพทย์ 3 ชนิดหลัก (คีมจับเข็ม, เข็มเย็บแผล, คีมคีบ) จากวิดีโอการฝึกปฏิบัติ และนำข้อมูลพิกัดมาวิเคราะห์ลำดับขั้นตอน (Sequential Analysis) เพื่อให้คะแนนความถูกต้องตามมาตรฐานทางการแพทย์

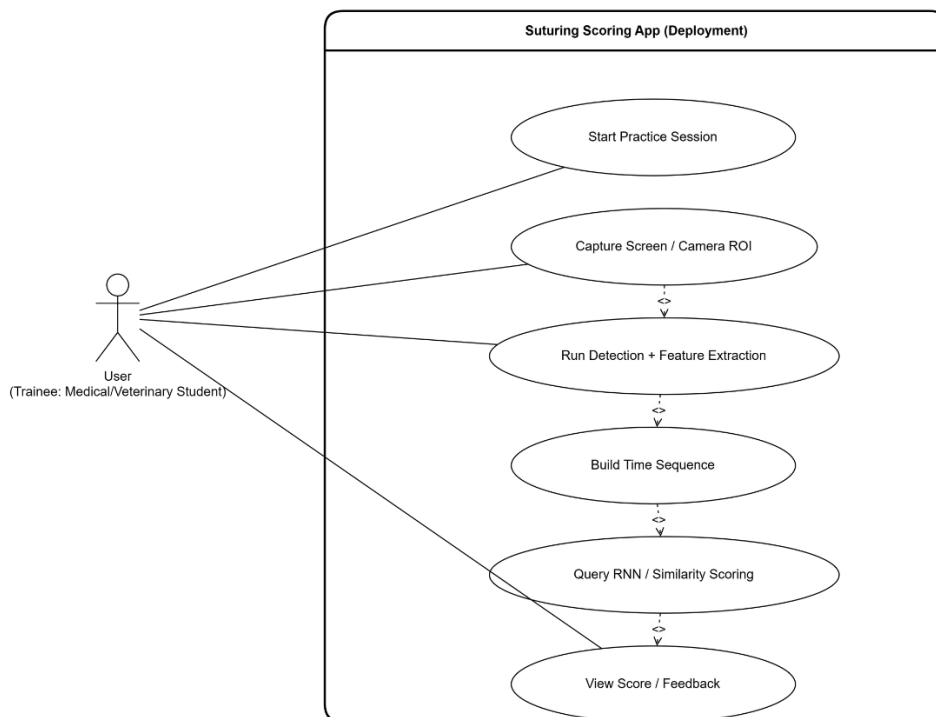
3.1.2 กลุ่มผู้ใช้งานระบบและหน้าที่ (System Users and Roles) จากการวิเคราะห์ระบบสามารถจำแนกผู้เกี่ยวข้องออกเป็น 1 กลุ่ม ได้แก่

User กลุ่มผู้ฝึกฝนทักษะ (เช่น นักศึกษาแพทย์/สัตวแพทย์) ทำหน้าที่ใช้งานระบบผ่านการประมวลผลวิดีโอหรือกล้องสดเพื่อรับผลการประเมิน

3.1.3 วิธีการเก็บข้อมูล (Data Collection) คณะผู้จัดทำดำเนินการรวบรวมข้อมูลผ่านการบันทึกวิดีโอการฝึกปฏิบัติเย็บแผลจริง (Suturing VDO Clip) และนำมาผ่านกระบวนการสุ่มเฟรมภาพ (Frame Sampling) จำนวนประมาณ 50 ภาพต่อคลิป เพื่อใช้เป็นตัวแทนข้อมูลในการวิเคราะห์และจัดทำชุดข้อมูลสอน

3.2 การวิเคราะห์ความต้องการด้วยแผนภาพยูสเคส (Use Case Analysis)

จากการวิเคราะห์หน้าที่การทำงานของระบบ (Functional Requirements) สามารถจำแนกผู้เกี่ยวข้องและฟังก์ชันการทำงานหลักผ่านแผนภาพ Use Case Diagram ได้ดังนี้



จากการวิเคราะห์โครงสร้างหน้าที่การทำงานของระบบ สามารถจำแนกผู้ใช้งานและฟังก์ชันการทำงานได้ดังภาพประกอบ Use Case โดยมีรายละเอียดดังนี้

1. ผู้เกี่ยวข้องกับระบบ (Actors) ระบบมีผู้เกี่ยวข้องหลักจำนวน 1 กลุ่ม ได้แก่

- User ผู้ฝึกเย็บแผล (เช่น นักศึกษาแพทย์/สัตวแพทย์) เป็นผู้ใช้งานระบบเพื่อเริ่มการฝึก โดยระบบจะรับภาพจากกล้องในมุมมองที่สอดคล้องกับข้อมูลที่ใช้ฝึกโมเดล จากนั้นคำนวณคะแนน/เปอร์เซ็นต์ความเหมือน พร้อมแสดงข้อเสนอแนะ (Feedback) เพื่อช่วยชี้แนะและปรับปรุงทักษะระหว่างการฝึกปฏิบัติ

2. รายละเอียดของยูสเคส (Use Cases) ฟังก์ชันการทำงานของระบบประกอบด้วย

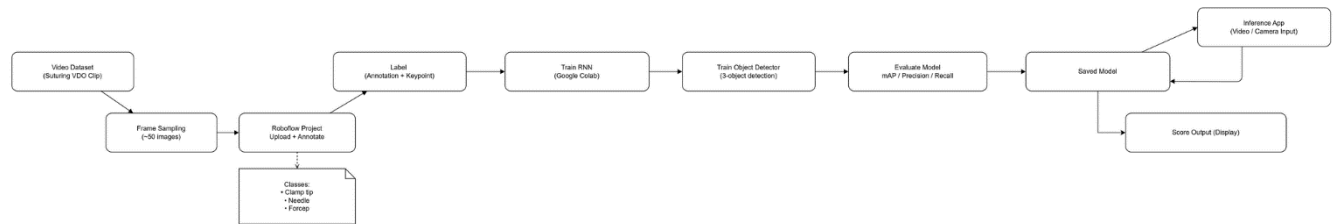
- Start Practice Session ผู้ใช้งานเริ่มต้นการฝึกเย็บแผลในระบบ เพื่อให้ระบบเข้าสู่สถานะพร้อมรับสัญญาณจากกล้อง

- Capture Camera ROI ระบบรับภาพจาก “กล้อง” และกำหนดพื้นที่สนใจ (ROI: Region of Interest) เพื่อโฟกัสบริเวณที่เกี่ยวข้องกับการเฝ้าพล โดยมีเป้าหมายให้มุมมองการถ่ายภาพสอดคล้องกับข้อมูลที่ใช้ในการฝึกสอนโมเดล มุ่งเน้นลดสิ่งรบกวน และเพิ่มความเสถียรของผลการประเมิน
- Run Detection + Feature Extraction ระบบประมวลผลภาพแต่ละเฟรมด้วยโมเดลตรวจจับวัตถุ (Object Detector) เพื่อระบุองค์ประกอบสำคัญที่เกี่ยวข้องกับการเฝ้าพล (เช่น ปลายคีม, เข็มเย็บ, คีมคีบ) จากนั้นสกัดคุณลักษณะ (Feature) ที่แทนการเคลื่อนไหวในแต่ละช่วงเวลา เช่น พิกัดตำแหน่ง (x, y) ของวัตถุ/จุดสำคัญในแต่ละเฟรม
- Build Time Sequence ระบบนำค่า Feature จากหลายเฟรมมาเรียงเป็นข้อมูลแบบลำดับเวลา (Time Sequence) เพื่อสะท้อนรูปแบบการเคลื่อนไหวอย่างต่อเนื่อง ซึ่งเหมาะสำหรับการประเมินทักษะที่เป็น “กระบวนการ” มากกว่าการพิจารณาจากภาพนิ่งเพียงเฟรมเดียว
- Query RNN / Similarity Scoring ระบบส่ง Time Sequence เข้าโมเดล RNN เพื่อคำนวณค่าความเหมือน (Similarity) เมื่อเทียบกับรูปแบบอ้างอิงที่โมเดลเรียนรู้จากข้อมูลตัวอย่างที่ถูกต้อง (จากอาจารย์) โดยสามารถสรุปผลเป็นเปอร์เซ็นต์ตามเกณฑ์ที่กำหนด
- View Score / Feedback ระบบแสดงคะแนน/เปอร์เซ็นต์และข้อเสนอแนะ (Feedback) ให้ผู้ใช้งาน เพื่อบอกระดับความถูกต้องของการปฏิบัตินั้น และช่วยชี้แนะแนวทางปรับปรุงระหว่างการฝึก

3.3 การออกแบบขั้นตอนการทำงานของระบบ (System Flowchart)

คณะผู้จัดทำได้ออกแบบลำดับการทำงานเชิงอัลกอริทึม (Algorithm Flow) ตั้งแต่เริ่มต้นการรับข้อมูลจนถึงการแสดงผลลัพธ์คะแนน ดังภาพประกอบ

flowchart



ขั้นตอนการทำงานของระบบ (System Flowchart) กระบวนการทำงานของระบบตั้งแต่นำเข้าข้อมูลจนถึงการแสดงผลลัพธ์ มีลำดับขั้นตอนตาม Flowchart ดังนี้

1. ขั้นตอนการเตรียมข้อมูล (Data Preparation)

เริ่มจากการนำเข้า Video Dataset (Suturing VDO Clip) ซึ่งเป็นคลิปวิดีโอสาธิตการเย็บแผลเข้าสู่กระบวนการ Frame Sampling เพื่อสุ่มภาพหนึ่งออกมาประมาณ 50 ภาพ นำข้อมูลเข้าสู่ Roboflow Project เพื่อทำการ Upload และ Annotate ข้อมูล โดยมีการกำหนดคลาส (Classes) ของวัตถุที่ต้องการตรวจจับ 3 ประเภท ได้แก่

1. Clamp tip (ปลายคีมจับเข็ม)
2. Needle (เข็มเย็บแผล)
3. Forcep (คีมคีบ)

2. ขั้นตอนการฝึกสอนและประเมินโมเดล (Training & Evaluation)

- ทำการ Label (Annotation + Keypoint) เพื่อระบุตำแหน่งวัตถุและจุดพิกัดสำคัญ
- ดำเนินการ Train RNN โดยใช้ Google Colab เพื่อเรียนรู้ลักษณะข้อมูลแบบอนุกรมเวลา
- ดำเนินการ Train Object Detector เพื่อสร้างโมเดลสำหรับตรวจจับวัตถุทั้ง 3 ประเภท (3-object detection)
- ทำการ Evaluate Model เพื่อวัดประสิทธิภาพของโมเดล โดยพิจารณาจากค่าสถิติสำคัญ ได้แก่ mAP (mean Average Precision), Precision และ Recall
- เมื่อได้โมเดลที่มีประสิทธิภาพตามเกณฑ์ จะทำการบันทึกเป็น Saved Model

3. ขั้นตอนการนำไปใช้งาน (Deployment)

- นำ Saved Model ไปใช้งานร่วมกับ Inference App
- ระบบรองรับข้อมูลนำเข้า (Input) ทั้งแบบไฟล์วิดีโอและภาพจากกล้อง (Video / Camera Input)
- สุดท้ายระบบจะทำการประมวลผลและแสดงผลลัพธ์ออกมาเป็นคะแนนผ่านหน้าจอแสดงผล (Score Output / Display)

3.4 การออกแบบฐานข้อมูลของระบบ (ER Diagram)

การออกแบบฐานข้อมูลในโครงการนี้มีวัตถุประสงค์เพื่อรองรับการจัดเก็บข้อมูลสำหรับการฝึกสอนโมเดล (Training Dataset) โดยขอบเขตข้อมูลที่จัดเก็บในระยะปัจจุบันมุ่งเน้นที่ “วิดีโอการเย็บแผลที่ถูกต้องของอาจารย์” ซึ่งใช้เป็นข้อมูลอ้างอิงหลักในการสร้างชุดข้อมูลและกระบวนการฝึกโมเดล ดังนั้นโครงสร้างฐานข้อมูลจึงออกแบบให้เรียบง่ายและเน้นการเก็บไฟล์วิดีโอพร้อมข้อมูลกำกับ (metadata) ที่จำเป็นต่อการประมวลผล

Training Dataset (Current)

VIDEO
video_id (PK)
file_name
file_path / storage_url
format (e.g., mp4)
fps
duration_sec
resolution (e.g., 1920x1080)
recorded_at (optional)
recorded_at (optional)
camera_view/setup (optional)
note (optional)

เอนทิตีและแอตทริบิวต์ (Entities & Attributes)

ฐานข้อมูลประกอบด้วยเอนทิตีหลัก ได้แก่

1) VIDEO

ใช้จัดเก็บข้อมูลวิดีโอที่ใช้สำหรับฝึกสอนโมเดล โดยเป็นวิดีโอสาธิตการเย็บแผลที่ถูกต้องจากอาจารย์ (Expert/Instructor Video) ซึ่งระบบจัดเก็บทั้งตำแหน่งไฟล์และคุณสมบัติพื้นฐานของวิดีโอ (metadata) เพื่อรองรับขั้นตอนการสุ่มเฟรม (Frame Sampling) และการเตรียมข้อมูลก่อนนำไปทำ Annotation และฝึกสอนโมเดล (Training) โดยมีแอตทริบิวต์สำคัญที่ควรจัดเก็บ เช่น

- video_id (PK) รหัสวิดีโอ (คีย์หลัก)
- file_name ชื่อไฟล์วิดีโอ
- file_path ที่อยู่ไฟล์/ตำแหน่งจัดเก็บ
- format ประเภทไฟล์ (เช่น mp4)
- fps ค่าเฟรมต่อวินาที
- duration_sec ความยาววิดีโอ (วินาที)
- resolution ความละเอียดวิดีโอ
- recorded_at วันที่บันทึกวิดีโอ
- note หมายเหตุ เช่น มุมกล้อง

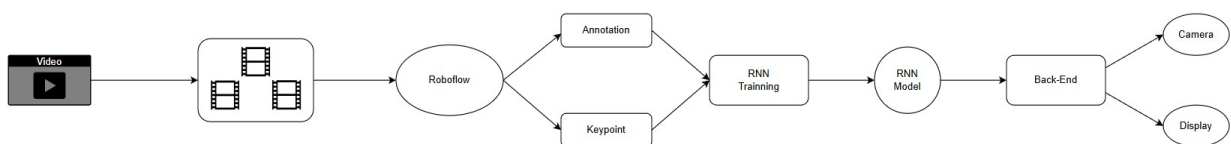
หมายเหตุ: ในระยะปัจจุบันเป็นการพัฒนาต้นแบบ (Prototype) ระบบใช้วิดีโอสำหรับฝึกสอนเพียง 1 ไฟล์ ดังนั้นตาราง VIDEO จึงมีข้อมูลเพียง 1 รายการ

บทที่ 4

ผลการดำเนินงานเบื้องต้น

ในการพัฒนาระบบตรวจจับอุปกรณ์การเฝ้าและประเมินผล คณะผู้จัดทำได้แบ่งระยะการดำเนินงานออกเป็น 2 ส่วนหลัก เพื่อให้การบริหารจัดการวงจรชีวิตของระบบปัญญาประดิษฐ์ (AI Lifecycle) เป็นไปอย่างมีประสิทธิภาพ ดังรายละเอียดต่อไปนี้

4.1 โครงการที่ 1 การจัดการชุดข้อมูลและเตรียมความพร้อมโมเดล (Dataset & Setup)



ในระยะแรกมุ่งเน้นไปที่การสร้างฐานข้อมูลที่มีคุณภาพสูง เพื่อเป็นรากฐานในการฝึกสอนโมเดลตรวจจับวัตถุและวิเคราะห์ท่าทาง

4.1.1 การเตรียมชุดข้อมูล (Data Preparation) การเตรียมข้อมูลที่มีประสิทธิภาพส่งผลโดยตรงต่อความแม่นยำของโมเดล โดยมีขั้นตอนดำเนินการดังนี้

- การนำเข้าวิดีโอ (Video Import) ทำการรวบรวมและนำเข้าไฟล์วิดีโอคลิปต้นฉบับ (Dataset) ที่บันทึกขั้นตอนการเฝ้าและการแพทย์เข้าสู่ระบบประมวลผล
- การแยกเฟรมภาพ (Frame Extraction) ดำเนินการแตกไฟล์วิดีโอออกเป็นภาพนิ่ง (Frames) อย่างต่อเนื่อง เพื่อสร้างชุดข้อมูลสำหรับฝึกสอน (Training Data)
- การสุ่มตัวอย่างข้อมูล (Sampling Strategy) เพื่อลดความซ้ำซ้อนของข้อมูลและประหยัดทรัพยากรในการประมวลผล คณะผู้จัดทำได้เลือกใช้วิธีการสุ่มตัวอย่างภาพ เช่น การตัดภาพในทุกๆ 5 วินาที ทำให้ได้ภาพที่เป็นตัวแทนของแต่ละขั้นตอนการเฝ้าอย่างเหมาะสม
- การบริหารจัดการผ่าน Roboflow นำชุดภาพที่ผ่านการคัดเลือกเข้าสู่แพลตฟอร์ม Roboflow เพื่อดำเนินการบริหารจัดการชุดข้อมูล (Dataset Management) อย่างเป็นระบบ

4.1.2 การกำหนดสัญลักษณ์และตรวจจับจุดสำคัญ (Annotation & Keypoint Detection) เมื่อได้เฟรมภาพที่ต้องการแล้ว ระบบจะเข้าสู่ขั้นตอนการระบุข้อมูลอัตลักษณ์ (Data Labeling) ดังนี้

1. Annotation (การระบุวัตถุ) ทำการสร้างกรอบล้อมรอบ (Bounding Box) บนอุปกรณ์คล้ายกรรมหลักทั้ง 3 ชนิด รวมถึงตัวผู้ปฏิบัติงานในแต่ละเฟรม เพื่อให้ระบบระบุตำแหน่งเชิงพื้นที่ของวัตถุได้อย่างถูกต้อง

2. Keypoint Detector (การตรวจจับจุดพิกัดสำคัญ) นอกเหนือจากตำแหน่งวัตถุ ระบบจะทำการกำหนดจุดพิกัดสำคัญ (Keypoints) บนร่างกายผู้ฝึกและตัวอุปกรณ์ เช่น **บาดแผลจำลอง**

3. การวิเคราะห์เชิงลึก ข้อมูล Keypoints นี้ช่วยให้ระบบสามารถวิเคราะห์การเคลื่อนไหว (Motion Analysis) ได้ละเอียดกว่าการตรวจจับเพียงตำแหน่งยืนเพียงอย่างเดียว

การใช้โปรแกรม vs code ในการการแยกเฟรมภาพหรือการ sampling frame

```
import cv2
import os
video_path = "vdo.mp4"
output_dir = "sampled_frames"
os.makedirs(output_dir, exist_ok=True)
cap = cv2.VideoCapture(video_path)
if not cap.isOpened():
    raise IOError("ไม่สามารถเปิดวิดีโอได้")
total_frames = int(cap.get(cv2.CAP_PROP_FRAME_COUNT))
fps = cap.get(cv2.CAP_PROP_FPS)
duration = total_frames / fps
print(f"Total frames: {total_frames}")
print(f"FPS: {fps}")
print(f"Duration (sec): {duration:.2f}")
num_samples = 50
```

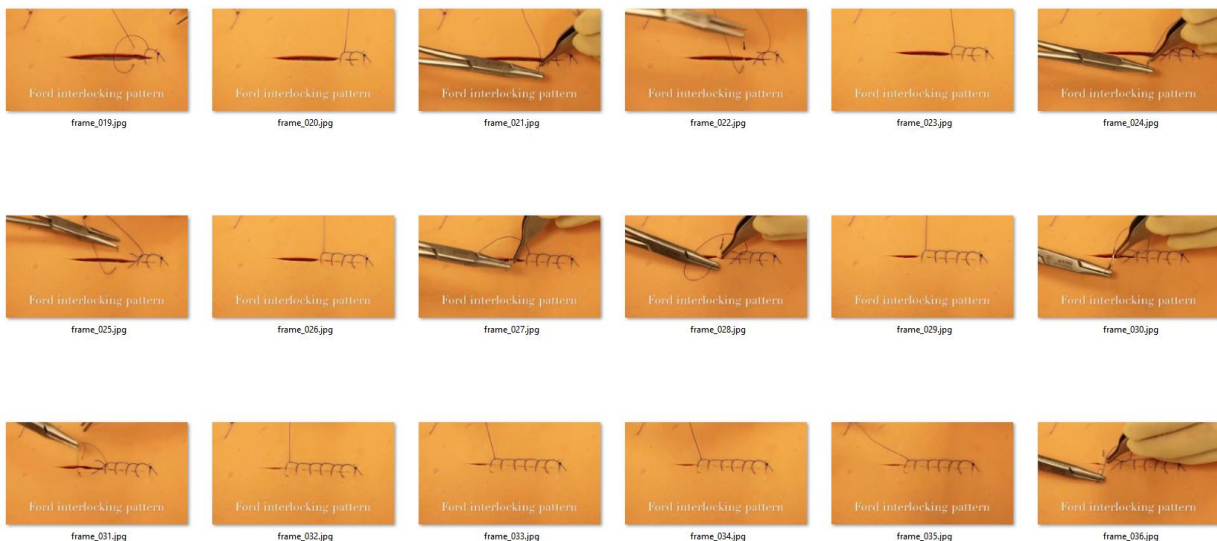
```

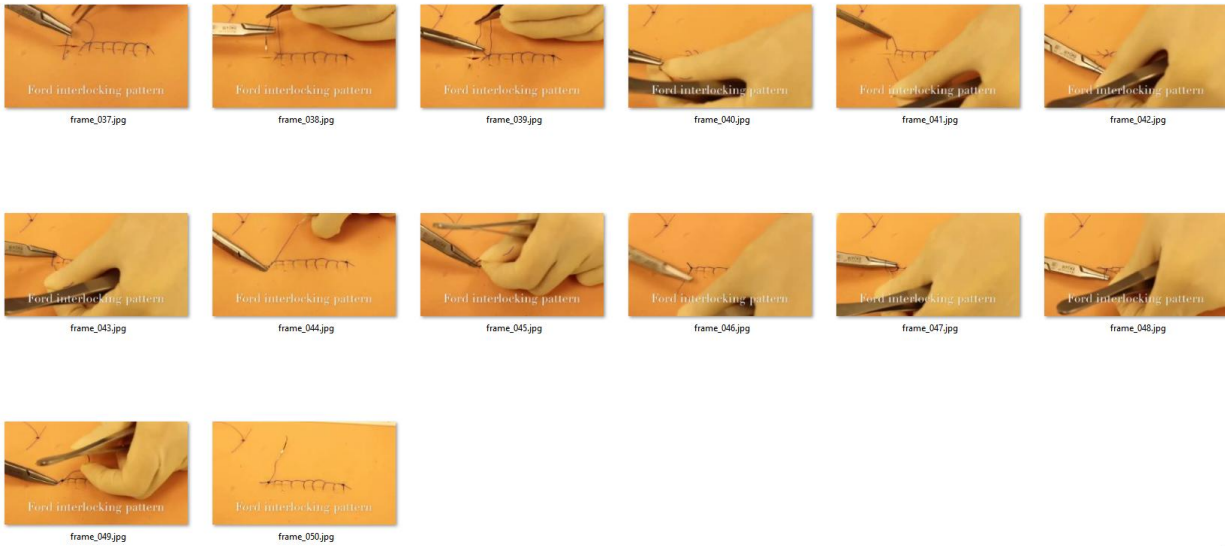
frame_interval = total_frames // num_samples
print(f"Frame interval: {frame_interval}")

saved_count = 0
for i in range(num_samples):
    frame_id = i * frame_interval
    cap.set(cv2.CAP_PROP_POS_FRAMES, frame_id)
    ret, frame = cap.read()
    if not ret:
        continue
    frame_filename = os.path.join(
        output_dir, f"frame_{i+1:03d}.jpg"
    )
    cv2.imwrite(frame_filename, frame)
    saved_count += 1
cap.release()
print(f"Saved {saved_count} frames successfully!")

```

ผลลัพธ์ที่ได้





4.2 โครงการที่ 2 การออกแบบและประมวลผลการประเมิน (Modeling & Scoring)

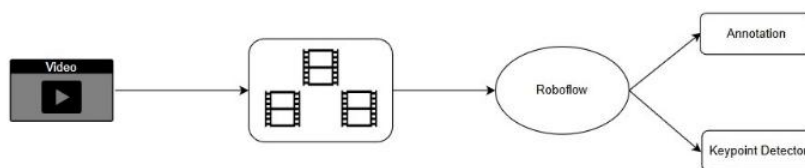
ในระยะที่สองมุ่งเน้นการพัฒนาสถาปัตยกรรมของโมเดลเพื่อให้สามารถวิเคราะห์ข้อมูลที่มีความต่อเนื่องเชิงเวลาและสรุปผลออกมาเป็นคะแนน

4.2.1 การออกแบบและเทรนโมเดล (Model Architecture & Training) เนื่องจากข้อมูลนำเข้าเป็นรูปแบบวิดีโอที่มีการเคลื่อนไหวต่อเนื่อง (Sequence) สถาปัตยกรรมโมเดลจึงถูกออกแบบให้รองรับข้อมูลประเภทนี้

- การเลือกใช้ RNN / LSTM เลือกใช้สถาปัตยกรรมแบบ Recurrent Neural Network (RNN) หรือ Long Short-Term Memory (LSTM) เนื่องจากมีความโดดเด่นในการจดจำและเรียนรู้ลำดับเหตุการณ์ (Time Series)
- การประมวลผลเชิงเวลา (Temporal Data) โมเดลชนิดนี้เหมาะสมกับการวิเคราะห์ท่าทางที่มีการเปลี่ยนแปลงไปตามเวลา ซึ่งมีประสิทธิภาพสูงกว่าการใช้ CNN ทั่วไปที่วิเคราะห์ได้เพียงภาพนิ่ง
- เป้าหมายเพื่อฝึกสอนให้โมเดลจดจำ "จังหวะ" และ "แบบแผนที่ถูกต้อง" ของการเย็บแผล (Pattern Recognition) ตามมาตรฐานทางการแพทย์

4.2.2 กระบวนการประมวลผลและแสดงผล (Inference & Scoring System) เมื่องานพัฒนาโมเดลเสร็จสมบูรณ์ ระบบจะถูกเชื่อมต่อเข้ากับส่วนงานหลังบ้าน (Backend) เพื่อใช้งานในสถานการณ์จริงดังนี้

1. การรับภาพสด (Camera Input) ระบบรองรับการรับสัญญาณภาพสดจากกล้องที่กำลังบันทึกพฤติกรรมของผู้ใช้งานในขณะที่ฝึกปฏิบัติ
2. การคำนวณดัชนีความคล้ายคลึง (Similarity Index): ระบบจะนำข้อมูลพิกัดและลำดับท่าทางจากการตรวจจับสด ไปเปรียบเทียบกับแบบจำลองต้นแบบ (Gold Standard) เพื่อหาดัชนีความคล้ายคลึง
3. การแสดงผลการประเมิน (Scoring Output) ระบบจะคำนวณความถูกต้องของท่าทางและลำดับขั้นตอนออกมาเป็นคะแนนเชิงปริมาณ และแสดงผลผ่านหน้าจอ Score Display เพื่อให้ผู้เรียนทราบคุณภาพของการปฏิบัติงาน (Quality Evaluation)



4.2 การกำหนดสัญลักษณ์และตรวจจับจุดสำคัญ (Annotation & Keypoint Detection)

เมื่อได้เฟรมภาพแล้ว ขั้นตอนถัดไปคือการระบุสิ่งที่สนใจในภาพ เพื่อนำไปสอนโมเดล โดยแบ่งออกเป็น 2 ส่วนหลัก ตามแผนภาพ

1. Annotation (การระบุวัตถุ) ทำการตีขอบล้อมรอบวัตถุหรือบุคคล (Bounding Box) ในแต่ละเฟรม เพื่อให้ระบบทราบตำแหน่งของบุคคลในภาพ
2. Keypoint Detector (การระบุจุดพิกัดสำคัญ) ทำการกำหนดจุดพิกัดสำคัญบนร่างกาย (Keypoints) เช่น ข้อศอก หัวไหล่ เข่า หรือจุดที่ต้องการโฟกัสเป็นพิเศษ เพื่อใช้ในการวิเคราะห์การเคลื่อนไหวที่ละเอียดกว่าเพียงแค่ตำแหน่งยืน

4.3 สรุปและข้อเสนอแนะในการดำเนินงาน

จากการดำเนินงานทั้ง 2 ระยะ สามารถสรุปประเด็นสำคัญได้ดังนี้

Project 1 เน้นความสำคัญไปที่การสร้างชุดข้อมูลคุณภาพสูง (High-Quality Dataset) การทำ Labeling ที่แม่นยำระดับจุดพิกเซล (Keypoint) และการเทรนให้โมเดลพื้นฐานจดจำอุปกรณ์และท่าทางเบื้องต้นได้

Project 2 เน้นการประยุกต์ใช้โมเดลในระดับสูงเพื่อคำนวณคะแนน (Scoring) โดยใช้ Similarity Index เป็นตัวชี้วัดหลักเพื่อบอกความแตกต่างระหว่างท่าทางของผู้ฝึกกับต้นแบบที่ถูกต้อง

อ้างอิง

Islam, M., et al. (2023). "Surgical Workflow Recognition Using Transformer-Based Spatio-Temporal Modeling." *International Journal of Computer Assisted Radiology and Surgery*.

Wang, J., et al. (2024). "Robust Instance Segmentation of Surgical Tools under Occlusion using Augmented Mask R-CNN." *IEEE Robotics and Automation Letters*.

Lin, H., et al. (2025). "Real-time AI-based Feedback System for Surgical Skill Training: Architecture and Evaluation." *Journal of Medical Systems*.