

เครื่องช่วยอ่านฉลากยาด้วยปัญญาประดิษฐ์ผ่านการแปลงข้อความเป็นเสียง
AI-Powered Assistive Device for Reading Medicine Labels Using
Text-to-Speech Conversion

ขวัญแก้ว สมงาม พัฒนบุรี ตีที่
ณัฐธิดา ปะวะลี น้ำทิพย์ คำทอง¹

¹โรงเรียนวิทยาศาสตร์จุฬาราชวิทยาลัย สระแก้ว, ตำบลวัฒนานคร, อำเภอวัฒนานคร, จังหวัดสระแก้ว 27160

Email: 6700043@pcshssk.ac.th

บทคัดย่อ

เครื่องช่วยอ่านฉลากยาด้วยปัญญาประดิษฐ์ผ่านการแปลงข้อความเป็นเสียง มีวัตถุประสงค์เพื่อพัฒนาอุปกรณ์ต้นแบบสำหรับอ่านข้อความจากฉลากยา วิเคราะห์ข้อมูลสำคัญ และแปลงข้อมูลเป็นเสียงพูด เพื่อช่วยให้ผู้สูงอายุ ผู้ที่มีความบกพร่องทางการมองเห็น และผู้ที่มีข้อจำกัดในการเข้าถึงข้อมูลตัวอักษร สามารถรับรู้ข้อมูลบนฉลากยาได้อย่างถูกต้อง สะดวก และลดโอกาสการ用药ผิดพลาด ระบบต้นแบบประกอบด้วยบอร์ด Raspberry Pi 5 และกล้อง USB โดยประยุกต์ใช้ระบบรู้จำอักขระจากภาพแบบผสม (Hybrid OCR) สำหรับอ่านข้อความบนฉลากยา จากนั้นประมวลผลด้วยโมเดลภาษาขนาดใหญ่ (Large Language Model: LLM) เพื่อแก้ไข วิเคราะห์ และสรุปข้อมูลสำคัญ ก่อนแปลงข้อความเป็นเสียงด้วยระบบ Text-to-Speech และแสดงผลผ่านลำโพง ผลการทดสอบพบว่าการประมวลผลด้วย โมเดลภาษาขนาดใหญ่ (Large Language Model: LLM) ช่วยเพิ่มความถูกต้องของข้อมูลจากร้อยละ 77.85 เป็นร้อยละ 92.93 ระบบใช้เวลาในการประมวลผลเฉลี่ย 21.33 วินาที ต่อครั้ง และผู้ทดสอบจำนวน 5 คน มีความเข้าใจข้อมูลจากเสียงในระดับสูง โดยมีคะแนนเฉลี่ย 4.73 จาก 5 คะแนน หรือคิดเป็นร้อยละ 94.67 สรุปได้ว่าระบบสามารถทำงานได้ตามวัตถุประสงค์ ช่วยเพิ่มการเข้าถึงข้อมูลบนฉลากยา ลดความเสี่ยงจากการ用药ผิดพลาด และแสดงให้เห็นถึงศักยภาพของการประยุกต์ใช้เทคโนโลยีปัญญาประดิษฐ์เพื่อพัฒนาอุปกรณ์ช่วยเหลือด้านการใช้งานในอนาคต

คำสำคัญ: ปัญญาประดิษฐ์, ฉลากยา, การรู้จำอักขระจากภาพ, โมเดลภาษาขนาดใหญ่, การแปลงข้อความเป็นเสียง

1. บทนำ

ในปัจจุบัน โรงพยาบาล สถานพยาบาล และร้านขายยาส่วนใหญ่นิยมใช้ฉลากยาเพื่อแสดงข้อมูลสำคัญ เช่น ชื่อยา ขนาดยา วิธีการใช้ยา และข้อควรระวัง เพื่อให้ผู้ป่วยสามารถเข้าใจได้อย่างถูกต้อง และปลอดภัย อย่างไรก็ตาม ผู้สูงอายุ ผู้ที่มีความบกพร่องทางการมองเห็น และผู้ที่มีข้อจำกัดในการอ่านข้อความขนาดเล็ก ยังคงประสบปัญหาในการเข้าถึง

ข้อมูลบนฉลากยา ซึ่งอาจนำไปสู่ความคลาดเคลื่อนในการใช้ยาและส่งผลกระทบต่อสุขภาพ ทั้งนี้ องค์การอนามัยโลก (World Health Organization: WHO) ระบุว่าความคลาดเคลื่อนทางยา (Medication Error) เป็นหนึ่งในสาเหตุสำคัญของอันตรายที่สามารถป้องกันได้ และก่อให้เกิดความสูญเสียทางเศรษฐกิจทั่วโลก ประมาณ 42 พันล้านดอลลาร์สหรัฐต่อปี (World Health Organization [WHO], 2017)

แม้ว่าปัจจุบันจะมีเทคโนโลยีระบบรู้จำอักขระจากภาพ (Optical Character Recognition: OCR) ที่สามารถอ่านข้อความจากภาพได้ แต่ฉลากยามักประกอบด้วยข้อความภาษาไทย ภาษาอังกฤษ และข้อมูลที่เขียนด้วยลายมือ ทำให้การอ่านข้อความยังมีความคลาดเคลื่อนในบางกรณี งานวิจัยล่าสุดแสดงให้เห็นว่าการประยุกต์ใช้ระบบรู้จำอักขระจากภาพ (Optical Character Recognition: OCR) ร่วมกับโมเดลปัญญาประดิษฐ์เชิงกำเนิด (Generative AI) สามารถช่วยวิเคราะห์ แก้ไข และสรุปข้อมูลจากฉลากยาได้อย่างมีประสิทธิภาพ โดยเฉพาะสำหรับผู้สูงอายุ (Sathanarugsawait et al., 2025)

จากปัญหาดังกล่าว คณะผู้จัดทำจึงพัฒนาเครื่องช่วยอ่านฉลากยาด้วยปัญญาประดิษฐ์ผ่านการแปลงข้อความเป็นเสียง โดยใช้บอร์ด Raspberry Pi 5 ร่วมกับระบบรู้จำอักขระแบบผสม (Hybrid OCR) ซึ่งประกอบด้วย TrOCR และ EasyOCR เพื่ออ่านข้อความทั้งภาษาไทย ภาษาอังกฤษ และลายมือเขียน จากนั้นนำข้อมูลเข้าสู่โมเดลภาษาขนาดใหญ่ (Large Language Model: LLM) เพื่อวิเคราะห์ แก้ไขข้อความที่อ่านผิด และสรุปข้อมูลสำคัญของยา ก่อนแปลงผลลัพธ์เป็นเสียงด้วยระบบ Text-to-Speech เพื่อช่วยให้ผู้ใช้งานเข้าถึงข้อมูลบนฉลากยาได้สะดวก เข้าใจง่าย และลดความเสี่ยงจากการใช้ยาผิดพลาด

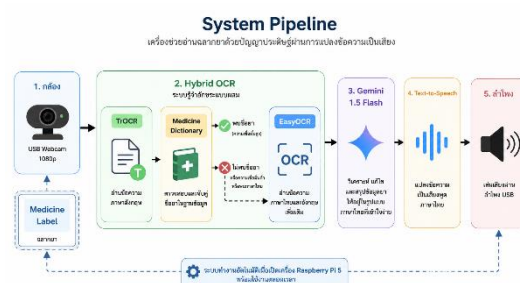
2. วิธีการดำเนินงาน

2.1 อุปกรณ์และชุดข้อมูล

ระบบต้นแบบพัฒนามาบนบอร์ด Raspberry Pi 5 ซึ่งทำหน้าที่เป็นหน่วยประมวลผลหลัก เชื่อมต่อกับกล้อง USB Webcam ความละเอียด 1080p สำหรับรับภาพฉลากยา และลำโพง USB สำหรับแสดงผลเสียงโดยการบันทึกภาพดำเนินการภายในกล่องควบคุมแสงเพื่อให้ได้สภาพแสงที่คงที่และลดผลกระทบจากแสงภายนอก สำหรับการพัฒนาแบบจำลอง ใช้ชุดข้อมูล IAM Handwriting Line Dataset (Teklia/IAM-line) จากแพลตฟอร์ม Hugging Face จำนวน 1,900 ภาพ

แบ่งเป็นชุดข้อมูลสำหรับฝึก 1,500 ภาพ และชุดข้อมูลสำหรับประเมินผล 400 ภาพ โดยแบ่งข้อมูลด้วยวิธี Stratified Sampling ตามจำนวนอักขระของแต่ละภาพ เพื่อให้ชุดข้อมูลมีการกระจายของความยาวข้อความอย่างสมดุล นอกจากนี้ยังใช้ตัวอย่างฉลากยาจริงที่มีทั้งข้อความภาษาไทย ภาษาอังกฤษ และลายมือเขียน จากบรรจุภัณฑ์ยาเม็ด ยาน้ำ และยาทาภายนอก สำหรับทดสอบการทำงานของระบบในสภาพแวดล้อมจริง ภาพในชุดข้อมูลสำหรับฝึกได้รับการทำ Data Augmentation ได้แก่ การปรับความสว่าง การหมุนภาพ และการเพิ่ม Gaussian Blur เพื่อเพิ่มความทนทานของแบบจำลองต่อความแปรผันของภาพจากการใช้งานจริง

2.2 สถาปัตยกรรมการประมวลผลของระบบ



ภาพที่ 1 System Pipeline

การทำงานของระบบแบ่งออกเป็น 5 ขั้นตอนหลัก ดังแสดงในภาพที่ 1

2.2.1 การรับภาพฉลากยา

กล้อง USB Webcam ทำหน้าที่บันทึกภาพฉลากยาภายในกล่องควบคุมแสง และส่งภาพเข้าสู่ Raspberry Pi 5 เพื่อเตรียมสำหรับการประมวลผล

2.2.2 การอ่านข้อความด้วยระบบรู้จำอักขระแบบผสม Hybrid OCR

ภาพฉลากยาถูกแบ่งออกเป็นแต่ละบรรทัดข้อความก่อนส่งเข้าสู่แบบจำลอง TrOCR เพื่ออ่านข้อความภาษาอังกฤษ จากนั้นระบบตรวจสอบชื่อยาด้วยฐานข้อมูลยา (Medicine Dictionary) หากตรวจ

พบข้อมูลที่ตรงกับฐานข้อมูล ระบบจะนำข้อมูลเข้าสู่ขั้นตอนถัดไปทันที แต่หากค่าความเชื่อมั่นต่ำหรือพบข้อความภาษาไทย ระบบจะเรียกใช้ EasyOCR เพื่ออ่านข้อความเพิ่มเติมและรวมผลลัพธ์ก่อนดำเนินการต่อ

2.2.3 การวิเคราะห์ข้อมูลด้วยโมเดลภาษาขนาดใหญ่

ข้อความที่ได้จากระบบรู้จำอักขระจากภาพ (Optical Character Recognition: OCR) จะถูกส่งเข้าสู่โมเดล Google Gemini 1.5 Flash ผ่าน OpenRouter API เพื่อแก้ไขคำที่อ่านผิด วิเคราะห์บริบท และสรุปข้อมูลสำคัญของฉลากยาให้อยู่ในรูปแบบภาษาไทยที่เข้าใจง่าย โดยประกอบด้วยชื่อยาสรรพคุณ วิธีใช้ และข้อควรระวัง

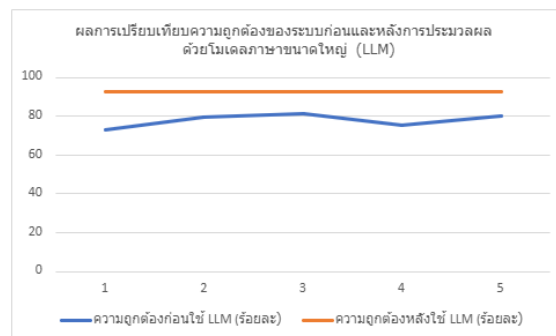
2.2.4 การแปลงข้อความเป็นเสียง

ข้อความสรุปถูกส่งเข้าสู่ระบบ Text-to-Speech เพื่อสังเคราะห์เสียงภาษาไทยและแสดงผลผ่านลำโพง USB โดยระบบถูกตั้งค่าให้เริ่มทำงานโดยอัตโนมัติเมื่อเปิดเครื่อง Raspberry Pi 5

3. สรุปและอภิปรายผล

3.1 การประเมินประสิทธิภาพของระบบ

เพื่อประเมินความถูกต้องของระบบ ได้ใช้ชุดข้อมูลภาพฉลากยาสังเคราะห์จำนวน 10,000 ภาพต่อการทดสอบ 1 ครั้ง ซึ่งสร้างขึ้นจากภาพฉลากยาจริงจำนวน 10 ภาพ และทำการทดสอบทั้งหมด 5 ครั้ง โดยเปรียบเทียบผลการทำงานของระบบก่อนและหลังการใช้โมเดลภาษาขนาดใหญ่ (Large Language Model: LLM) การประเมินใช้วิธีเปรียบเทียบผลลัพธ์ของระบบกับข้อมูลอ้างอิงจากฉลากยาจริง โดยพิจารณาความถูกต้องจากการระบุข้อมูลสำคัญ ได้แก่ ชื่อยา ความแรงหรือขนาดยา วิธีใช้ และข้อควรระวัง



กราฟที่ 3.1 ผลการเปรียบเทียบความถูกต้องของระบบก่อนและหลังการประมวลผลด้วยโมเดลภาษาขนาดใหญ่ (Large Language Model: LLM)

3.2 การประเมินระยะเวลาการทำงานของระบบ

ดำเนินการโดยใช้ภาพฉลากยาจริงจำนวน 5 ภาพ และทดสอบภาพละ 3 ครั้ง เพื่อหาค่าเฉลี่ยของระยะเวลาที่ใช้ในการประมวลผล โดยเริ่มจับเวลาตั้งแต่ระบบรับภาพจากกล้อง USB จนถึงการประมวลผลข้อความด้วยระบบรู้จำอักขระจากภาพ (Optical Character Recognition: OCR) การวิเคราะห์ข้อมูลด้วยโมเดลภาษาขนาดใหญ่ (Large Language Model: LLM) และการแปลงข้อความเป็นเสียง (Text-to-Speech) พร้อมแสดงผลเสียงผ่านลำโพง

ภาพฉลากยา	ครั้งที่ 1	ครั้งที่ 2	ครั้งที่ 3	ค่าเฉลี่ย
1	20	23	24	22.33
2	19	21	22	20.67
3	23	20	19	20.67
4	21	22	21	21.33
5	23	22	20	21.67
เฉลี่ยรวม	21.20	21.60	21.20	21.33

ตารางที่ 3.2 แสดงผลการทดสอบเวลาการทำงานของระบบตั้งแต่รับภาพจนแปลงเป็นเสียง

3.3 การประเมินความเข้าใจข้อมูลจากเสียงที่ระบบสร้างขึ้น

ความเข้าใจข้อมูลจากเสียงที่ระบบสร้างขึ้น ประเมินโดยผู้ทดสอบจำนวน 5 คน ซึ่งรับฟังเสียงที่แปลงจากข้อความบนฉลากยาจำนวน 3 ภาพที่มี

ลักษณะแตกต่างกัน และประเมินระดับความเข้าใจด้วยแบบประเมินมาตราส่วนประมาณค่า (Rating Scale) 5 ระดับ โดยพิจารณาจากข้อมูลสำคัญที่ระบบอ่านออกมา ได้แก่ ชื่อยา สรรพคุณ วิธีใช้ ขนาดรับประทาน และข้อควรระวัง

ภาพ ผู้ทดสอบ	A	B	C	ค่าเฉลี่ย
1	5	5	5	5.00
2	5	4	5	4.67
3	4	5	4	4.33
4	5	5	5	5.00
5	4	5	5	4.67
ค่าเฉลี่ย	4.60	4.80	4.80	4.73

ตารางที่ 3.3 แสดงผลการประเมินความเข้าใจข้อมูลจากการรับฟังเสียงที่ระบบสร้างขึ้น

สรุปผล

โครงการเครื่องช่วยอ่านฉลากยาด้วยปัญญาประดิษฐ์ผ่านการแปลงข้อความเสียงมีวัตถุประสงค์เพื่อพัฒนาระบบที่สามารถอ่านข้อความบนฉลากยา วิเคราะห์ข้อมูลสำคัญ และแปลงข้อความเป็นเสียงพูด เพื่อช่วยอำนวยความสะดวกในการเข้าถึงข้อมูลบนฉลากยา โดยเฉพาะผู้สูงอายุ ผู้ที่มีความบกพร่องทางการมองเห็น และผู้ที่มีความยากลำบากในการอ่านข้อความขนาดเล็กบนฉลากยา

การพัฒนาระบบดำเนินการโดยใช้บอร์ด Raspberry Pi 5 ร่วมกับกล้อง USB สำหรับรับภาพฉลากยา จากนั้นประมวลผลภาพด้วยไลบรารี OpenCV และอ่านข้อความด้วยระบบรู้จำอักขระจากภาพแบบผสม (Hybrid OCR) ซึ่งประกอบด้วย TrOCR และ EasyOCR พร้อมทั้งตรวจสอบชื่อยาด้วยฐานข้อมูลยาก่อนนำข้อมูลเข้าสู่กระบวนการวิเคราะห์และสรุปสาระสำคัญด้วยโมเดลภาษาขนาดใหญ่ (Large Language Model: LLM) และแปลงข้อความเสียงด้วยระบบ Text-to-Speech เพื่อให้ผู้ใช้งานสามารถรับฟังข้อมูลผ่านลำโพงได้อย่างสะดวก

ผลการทดสอบประสิทธิภาพของระบบพบว่าการประยุกต์ใช้โมเดลภาษาขนาดใหญ่ (Large Language Model: LLM) ช่วยเพิ่มความถูกต้องในการอ่านและวิเคราะห์ข้อมูลบนฉลากยา โดยค่าความถูกต้องเฉลี่ยของระบบเพิ่มขึ้นจากร้อยละ 77.85 เป็นร้อยละ 92.93 ขณะที่ระบบใช้ระยะเวลาในการประมวลผลตั้งแต่รับภาพจนแปลงข้อความเสียงเฉลี่ย 21.33 วินาทีต่อครั้ง นอกจากนี้ ผลการประเมินความเข้าใจข้อมูลจากเสียงที่ระบบสร้างขึ้นมีคะแนนเฉลี่ย 4.73 จากคะแนนเต็ม 5 หรือคิดเป็นร้อยละ 94.67 แสดงให้เห็นว่าผู้ใช้งานสามารถเข้าใจข้อมูลที่ระบบนำเสนอได้อย่างถูกต้องและชัดเจน

กิตติกรรมประกาศ

คณะผู้จัดทำขอขอบพระคุณอาจารย์ที่ปรึกษาโครงการที่ได้ให้คำปรึกษา ข้อเสนอแนะ และแนวทางในการดำเนินงานตลอดระยะเวลาการจัดทำโครงการรวมทั้งขอขอบพระคุณคณะครูและโรงเรียนที่ให้การสนับสนุนด้านสถานที่และอุปกรณ์ที่จำเป็นต่อการดำเนินโครงการ ขอขอบคุณผู้ทดสอบระบบทุกท่านที่ให้ความร่วมมือในการทดลองใช้งานและประเมินผล พร้อมทั้งให้ข้อเสนอแนะอันเป็นประโยชน์ต่อการปรับปรุงและพัฒนาระบบให้มีประสิทธิภาพยิ่งขึ้น คณะผู้จัดทำขอขอบพระคุณทุกท่านที่มีส่วนสนับสนุนและให้กำลังใจจนโครงการเรื่อง "เครื่องช่วยอ่านฉลากยาด้วยปัญญาประดิษฐ์ผ่านการแปลงข้อความเสียง" สำเร็จลุล่วงด้วยดี

เอกสารอ้างอิง

- สำนักงานคณะกรรมการอาหารและยา. (2567). *การใช้ยาอย่างสมเหตุผล*. กรุงเทพมหานคร: กระทรวงสาธารณสุข. สืบค้นเมื่อ 18 พฤศจิกายน 2568, จาก <https://www.fda.moph.go.th>
- สำนักงานหลักประกันสุขภาพแห่งชาติ. (2567). *การใช้ยาอย่างปลอดภัย*. กรุงเทพมหานคร: สำนักงาน



- หลักประกันสุขภาพแห่งชาติ. สืบค้นเมื่อ 10
ธันวาคม 2568, จาก <https://www.nhso.go.th>
- Google. (2024). *Gemini API Documentation*.
Mountain View, California: Google.
Retrieved January 8, 2026, from
<https://ai.google.dev>
- Hugging Face. (2024). *Teklia/IAM-line Dataset*.
New York: Hugging Face. Retrieved January
15, 2026, from
<https://huggingface.co/datasets/Teklia/IAM-line>
- Intel Corporation. (2024). *OpenCV Documentation*. Santa Clara, California:
Intel Corporation. Retrieved February 2,
2026, from <https://docs.opencv.org>
- Jaided AI. (2024). *EasyOCR Documentation*.
Seoul: Jaided AI. Retrieved February 6,
2026, from <https://www.jaided.ai/easyocr>
- Microsoft. (2024). *TrOCR Documentation*.
Redmond, Washington: Microsoft.
Retrieved February 18, 2026, from
[https://huggingface.co/docs/transformers/
model_doc/trocr](https://huggingface.co/docs/transformers/model_doc/trocr)
- OpenRouter. (2024). *OpenRouter Documentation*. Retrieved March 1, 2026,
from <https://openrouter.ai/docs>
- Raspberry Pi Foundation. (2024). *Raspberry Pi Documentation*. Cambridge: Raspberry Pi
Foundation. Retrieved March 5, 2026,
from <https://www.raspberrypi.com>
- OpenRouter.(2024).
OpenRouterDocumentation. Retrieved
July 14, 2026, from <https://openrouter.ai>
- Raspberry Pi Foundation. (2024). *Raspberry Pi Documentation*. Cambridge: Raspberry Pi
Foundation. Retrieved July 14, 2026, from
<https://www.raspberrypi.com/documentation/>
- <https://www.raspberrypi.com/documentation/>
on
- Sathanarugsawait, B., Ketsiriluck, P.,
Chokchuenchob, S., &
Kaewrattanapat, N. (2025). *Application
of Generative Artificial Intelligence
Models for Accurate Prescription
Label Identification and Information
Retrieval for the Elderly in Northern
East of Thailand*. Retrieved July 15,
2026, from
[https://doi.org/10.3390/jimaging11010
011](https://doi.org/10.3390/jimaging11010011)
- World Health Organization. (2017). *Medication
Without Harm: WHO Global Patient
Safety Challenge*. Geneva: World
Health Organization. Retrieved July
15, 2026, from
<https://www.who.int/initiatives>