

การพัฒนาระบบคัดแยกระดับความสูงของฟักทองด้วยการผสานข้อมูลภาพและเสียง โดยใช้เทคนิคการเรียนรู้ของเครื่อง

ฐานันดร มุลสาร¹, วิกรม วรรณสุทธิ์² และ ดนุ วิโรจน์อุไรเรือง^{1,*}

¹ สาขาเทคโนโลยีวิศวกรรม คณะเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยราชภัฏอุบลราชธานี

² สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏสุรินทร์

*ผู้ประสานงานบทความต้นฉบับ: tanandonr.mg65@ubru.ac.th โทรศัพท์: 099-1569463

(รับบทความ: 20 เมษายน 2568; แก้ไขบทความ: 17 กรกฎาคม 2568; ตอรับบทความ: 15 สิงหาคม 2568)

บทคัดย่อ

ฟักทองเป็นพืชเศรษฐกิจที่มีความต้องการสูง แต่การประเมินระดับความสุกยังคงเป็นปัญหา เนื่องจากเกษตรกรส่วนใหญ่ขาดความชำนาญ และกระบวนการคัดแยกแบบดั้งเดิมใช้เวลานาน งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบจำแนกระดับความสุกของฟักทองพันธุ์อำพัน 342 โดยใช้ค่าความหวาน (Brix) จากตัวอย่าง 30 ตัวอย่าง แบ่งเป็น 3 ระดับ ได้แก่ ดิบ สุก และสุกเกินไป โดยเก็บข้อมูลภาพและเสียงมาใช้ฝึกโมเดลโครงข่ายประสาทเทียมแบบคอนโวลูชัน (CNN) ผลการทดลองพบว่า โมเดลที่ใช้ภาพเพียงอย่างเดียวให้ค่า F1-Score เท่ากับ 0.84 และ Sensitivity เท่ากับ 0.85 ขณะที่โมเดลที่ใช้เสียงเพียงอย่างเดียวให้ค่า F1-Score เท่ากับ 0.87 และ Sensitivity เท่ากับ 0.85 สำหรับโมเดลที่รวมข้อมูลภาพและ Spectrogram ให้ผลลัพธ์ดีที่สุด โดยมีค่า F1-Score เท่ากับ 0.93 และ Sensitivity เท่ากับ 0.92 ชุดข้อมูลที่ใช้ในการทดลองประกอบด้วยภาพจำนวน 900 ภาพ (สำหรับทดสอบ 150 ภาพ) ไฟล์เสียง 900 ไฟล์ (สำหรับทดสอบ 150 ไฟล์) และข้อมูลแบบรวมภาพ-สเปกโตรแกรม 1,800 รายการ (สำหรับทดสอบ 300 รายการ) ข้อจำกัดของงานวิจัยนี้คือขนาดของชุดข้อมูลที่ยังไม่ใหญ่มาก และความไวของข้อมูลเสียงต่อสภาพแวดล้อม อย่างไรก็ตาม ระบบที่พัฒนาขึ้นมีศักยภาพในการประยุกต์ใช้จริงภาคสนาม เพื่อเพิ่มความแม่นยำและลดความสูญเสียจากการคัดแยกแบบใช้แรงงานคน

คำสำคัญ: ฟักทองอำพัน 342 โครงข่ายประสาทเทียม CNN การจำแนกภาพและเสียง Spectrogram เทคนิคของเครื่อง

การอ้างอิงบทความ: ฐานันดร มุลสาร, วิกรม วรรณสุทธิ์ และ ดนุ วิโรจน์อุไรเรือง, "การพัฒนาระบบคัดแยกระดับความสุกของฟักทองด้วยการผสานข้อมูลภาพและเสียงโดยใช้เทคนิคการเรียนรู้ของเครื่อง," วารสารวิศวกรรมและเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยกาฬสินธุ์, ปีที่ 3, ฉบับที่ 4, หน้า 59-75, 2568.

Development of Pumpkin Ripeness Classification System by Integrating Image and Sound Data Using Machine Learning Techniques

Tanandonr Moolsanr¹, Wikrom Wonnasuthi² and Danu Wiroteurairuang^{1,*}

¹ Engineering Technology Program, Faculty of Industrial Technology, Ubon Ratchathani Rajabhat University

² Computer Science Program, Faculty of Science and Technology, Surindra Rajabhat University

* Corresponding Author: tanandonr.mg65@ubru.ac.th, Tel: 099-1569463

(Received: April 20, 2025; Revised: July 17, 2025; Accepted: August 15, 2025)

Abstract

Pumpkin is a high-demand economic crop, but determining its ripeness remains a challenge due to limited expertise among farmers and the labor-intensive manual sorting process. This study aimed to develop a classification model for assessing the ripeness levels of "Ampan 342" pumpkins using Brix values from 30 samples categorized into three levels: unripe, ripe, and overripe. Image and sound data were collected and used to train Convolutional Neural Network (CNN) models. The image-based model achieved an F1-score of 0.84 and a sensitivity of 0.85, while the sound-based model yielded an F1-score of 0.87 and a sensitivity of 0.85. The combined image and spectrogram-based model achieved the best performance, with an F1-score of 0.93 and a sensitivity of 0.92. The dataset comprised 900 images (150 for testing), 900 audio recordings (150 for testing), and 1,800 combined image-spectrogram samples (300 for testing). Limitations of the study include the relatively small dataset and the susceptibility of audio data to environmental noise. Nevertheless, the proposed approach shows promise for real-world applications as an automated system for evaluating pumpkin ripeness in the field, enhancing accuracy and reducing post-harvest losses.

Keywords: AMPHAN 342 Pumpkin, Convolutional Neural Network, Image and Sound Classification, Spectrogram, Machine Learning

Please cite this article as: T. Moolsanr, W. Wonnasuthi and D. Wiroteurairuang, "Development of Pumpkin Ripeness Classification System by Integrating Image and Sound Data Using Machine Learning Techniques," *The Journal of Engineering and Industrial Technology, Kalasin University*, vol. 3, no. 4, pp. 59-75, 2025.

บทความวิจัย (Research Article)

1. บทนำ

ฟักทอง (Pumpkin) เป็นพืชเศรษฐกิจที่ได้รับ ความนิยมสูงในประเทศไทย โดยในปี 2562 มีพื้นที่ เก็บเกี่ยวรวมประมาณ 23,384 ไร่ และให้ผลผลิตรวม กว่า 51,520 ตัน [1] แม้จะมีปริมาณการผลิตจำนวนมาก แต่การประเมินความสุกของฟักทองยังคงอาศัย แรงงานคนเป็นหลัก ซึ่งมีข้อจำกัดด้านความแม่นยำ และความรวดเร็ว การคัดแยกที่ไม่ถูกต้องอาจนำไปสู่ การสูญเสียทางเศรษฐกิจ ทั้งในกระบวนการขนส่ง และการจัดจำหน่าย ในระดับสากล มีงานวิจัยที่พัฒนา เทคโนโลยีเพื่อประเมินความสุกของผลผลิตทาง การเกษตรโดยใช้เทคนิคด้านปัญญาประดิษฐ์ เช่น งานของ Smith et al. (2021) [2] ที่ใช้ข้อมูลภาพ และเสียงในการประเมินความสุกของมะม่วงด้วยความ แม่นยำสูง แนวทางดังกล่าวสามารถประยุกต์ใช้กับ ฟักทองได้เช่นกัน งานวิจัยนี้จึงมุ่งพัฒนาระบบจำแนก ระดับความสุกของฟักทองพันธุ์อำพัน 342 โดยใช้ภาพ และเสียงร่วมกับเทคนิคการเรียนรู้เชิงลึก เพื่อเพิ่ม ประสิทธิภาพของการคัดแยก ลดการสูญเสีย และ รองรับการใช้งานในภาคสนามจริง

หนึ่งในปัญหาหลักของการผลิตฟักทองคือการคัด แยกระดับความสุก เนื่องจากระดับความสุกมีผล โดยตรงต่อรสชาติ เนื้อสัมผัส และความต้องการของ ตลาด อย่างไรก็ตาม ปัจจุบันการคัดแยกยังคงพึ่งพา ทักษะส่วนบุคคล การสังเกตด้วยสายตา หรือการเคาะ เพื่อฟังเสียง ซึ่งมีแนวโน้มเกิดความผิดพลาดสูง ไม่สม่ำเสมอ และต้องใช้แรงงานจำนวนมาก โดยเฉพาะในฤดูเก็บเกี่ยวที่มีผลผลิตจำนวนมาก เพื่อแก้ปัญหาดังกล่าว งานวิจัยนี้จึงมีวัตถุประสงค์ เพื่อพัฒนา ระบบจำแนกระดับความสุกของฟักทอง พันธุ์อำพัน 342 โดยใช้เทคโนโลยีปัญญาประดิษฐ์ ระดับสากล ได้แก่ โครงข่ายประสาทเทียมแบบคอน โวลูชัน (Convolutional Neural Network: CNN) ซึ่งเป็นเทคนิคด้านการเรียนรู้เชิงลึก (Deep Learning) ที่ได้รับความนิยมในการประมวลผลภาพ

และเสียง ระบบที่พัฒนาขึ้นใช้ข้อมูลจากภาพสีของ ฟักทองและเสียงที่เกิดจากการเคาะผิวผล โดยนำมา ผ่านกระบวนการวิเคราะห์และฝึกโมเดลด้วย CNN เพื่อจำแนกระดับความสุก ได้แก่ ดิบ สุก และสุก เกินไป โดยเฉพาะการใช้เสียงร่วมกับภาพในรูปแบบ Spectrogram ช่วยเพิ่มความแม่นยำในการจำแนก อย่างมีนัยสำคัญ

ข้อได้เปรียบของเทคโนโลยีนี้ คือความสามารถใน การวิเคราะห์ข้อมูลจำนวนมากได้อย่างแม่นยำ สม่ำเสมอ และไม่ขึ้นกับบุคคล ช่วยลดต้นทุนแรงงาน เพิ่มความเร็ว และสามารถนำไปประยุกต์ใช้ใน ภาคสนามได้จริงผ่านอุปกรณ์กล้องและไมโครโฟน ตามมาตรฐานทั่วไป

2. ทบทวนวรรณกรรมที่เกี่ยวข้อง

กระบวนการประมวลผลภาพด้วยคอมพิวเตอร์ (Image Processing) คือ กระบวนการที่ทำให้ได้ภาพ ที่มีความละเอียดคมชัด เพื่อให้สามารถนำไปใช้งาน หรือการแปลงภาพให้อยู่ในรูปแบบที่เหมาะสมสำหรับ การวิเคราะห์ต่อไป [3]

2.1 สัญญาณรบกวนภาพแบบเกาส์เซียน (Gaussian Noise)

เป็นสัญญาณรบกวนประเภทหนึ่งทางด้านสถิติ ที่ค่าฟังก์ชันความหนาแน่นของความน่าจะเป็นมีค่า เท่ากับค่าการกระจายตัวแบบปกติ ซึ่งค่าของสัญญาณ รบกวนจะมีค่าอยู่ในการกระจายตัวแบบเกาส์เซียน โดยสมการของสัญญาณรบกวนภาพแบบเกาส์เซียน [4] สามารถแสดงได้ในสมการที่ (1)

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

บทความวิจัย (Research Article)

โดยที่

$G(x,y)$ คือ ค่าของฟิลเตอร์เกาส์เซียนที่ตำแหน่ง (x,y)
 σ คือ ส่วนเบี่ยงเบนมาตรฐานที่กำหนด ความกว้างของฟิลเตอร์

โดยที่

CDF_{min} คือ ค่า CDF ที่น้อยที่สุดในภาพ
 $M \times N$ คือ จำนวนพิกเซลในภาพ
 L คือ ค่าสูงสุดของสี (255 สำหรับ Grayscale)

2.2 Histogram Equalization

การปรับปรุงความเปรียบต่าง (Contrast) ของภาพโดยการกระจายความถี่ของค่าความสว่าง (Intensity) ให้กระจายอย่างสม่ำเสมอมากขึ้นในช่วงของความสว่างทั้งหมดที่มีอยู่ในภาพ เทคนิคนี้ช่วยให้รายละเอียดในภาพที่มีความเปรียบต่างต่ำถูกขยายให้ชัดเจนขึ้น โดยเฉพาะในภาพที่มีการกระจุกตัวของค่าความสว่างในบางช่วงแคบ ๆ และการปรับรูปภาพให้มีการกระจายตัวของความเข้มแสงอย่างสม่ำเสมอซึ่งทำให้กราฟความถี่สะสมเพิ่มขึ้นเป็นเส้นตรง [5]

คำนวณ Histogram $H(r_k)$ สำหรับทุก ๆ ค่าสี $H(r_k)$ ในช่วงค่าสีทั้งหมด (0-255 สำหรับ Grayscale)
 $H(r_k)$ = จำนวนพิกเซลที่มีค่าสี r_k ในภาพ คำนวณฟังก์ชัน CDF (Cumulative Distribution Function) สร้าง CDF $CDF(r_k)$ โดยบวกความถี่ของ Histogram จากซ้ายไปขวา

$$CDF(r_k) = \sum_{i=0}^k H(r_i) \quad (2)$$

ปรับปรุงค่าสีในภาพ ปรับปรุงค่าสีของแต่ละพิกเซลในภาพ $I(x,y)$ โดยใช้ค่า CDF

$$I_{\text{equalized}}(x,y) = \left\lfloor \frac{CDF(I(x,y)) - CDF_{min}}{M \times N - CDF_{min}} \times L - 1 \right\rfloor \quad (3)$$

สร้างภาพที่ปรับปรุงแล้วใช้ค่า $I_{\text{equalized}}(x,y)$ เพื่อสร้างภาพใหม่ที่มีการกระจายความสว่างที่ปรับปรุงแล้ว

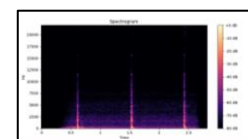
2.3 การกำจัดเสียงรบกวน (Eliminate Noise)

การลดหรือขจัดเสียงที่ไม่ต้องการที่มีอยู่ในสัญญาณเสียง เช่น เสียงลม เสียงจากเครื่องจักร หรือเสียงรบกวนจากสภาพแวดล้อมต่าง ๆ โดยมักจะใช้เทคนิคต่าง ๆ เช่น การกรอง (Filtering) หรือการใช้เทคโนโลยีการประมวลผลสัญญาณ (Signal Processing) เช่น Noise Reduction หรือ Noise Cancellation เพื่อแยกแยะเสียงรบกวนออกจากเสียงที่ต้องการ

2.4 การเสริมประสิทธิภาพเสียง (Enhancements)

การปรับแต่งหรือปรับปรุงคุณภาพของเสียงเพื่อให้ฟังได้ชัดเจนขึ้น เช่น การเพิ่มความคมชัดของเสียง การปรับความสมดุลของเสียง หรือการเพิ่มรายละเอียดในบางส่วนของสัญญาณเสียง

2.5 สเปกโตรแกรม (Spectrogram)



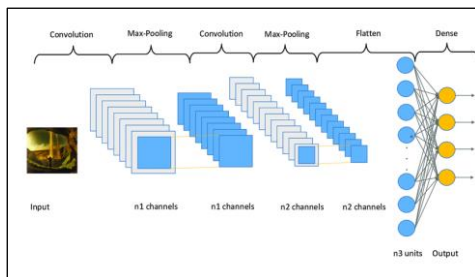
รูปที่ 1 สเปกโตรแกรม (Spectrogram)

บทความวิจัย (Research Article)

จากรูปที่ 1 แสดงสเปกโตรแกรมของการบันทึกเสียงเป็นการแสดงผลข้อมูลเสียงในรูปแบบกราฟที่แสดงการเปลี่ยนแปลงของความถี่ (Frequency) และแอมพลิจูด (Amplitude) ของเสียงตลอดช่วงเวลา โดยใช้ในการวิเคราะห์รายละเอียดของเสียงที่บันทึกไว้ [6]

2.6 Convolutional Neural Network (CNN)

เป็นโมเดลประสาทเทียมที่มีโครงสร้างที่ออกแบบมาเพื่อการประมวลผลภาพและข้อมูลที่มีลักษณะเป็นตาราง (Grid-Like Structure) โดยมีช่วงเวลาหรือสถานที่ที่ข้อมูลมีความสำคัญเช่นกัน ภายในภาพโมเดล CNN นี้มีความสามารถในการเรียนรู้และจดจำลักษณะที่เป็นลักษณะที่สำคัญของข้อมูลด้วยตนเอง [7]



รูปที่ 2 Convolutional Neural Network (CNN)

จากรูปที่ 2 คือ ขั้นตอนการทำงานของโมเดลโครงข่ายประสาทเทียมแบบคอนโวลูชัน และสมการในแต่ละเลเยอร์

ขั้นที่ 1 Input Layer : รับภาพอินพุต

ขั้นที่ 2 ชั้น Convolutional Layer ในเลเยอร์คอนโวลูชันฟิลเตอร์ (หรือเคอร์เนล) ถูกใช้งานเพื่อตรวจจับคุณลักษณะต่าง ๆ ของภาพสมการคอนโวลูชันระหว่างอินพุต / และฟิลเตอร์ K จะเป็น

$$(I * K)(i, j) = \sum_m \sum_n I(i+m, j+n) K(m, n) \quad (4)$$

โดยที่

I คือ อินพุต (เช่น ภาพ)

K คือ ฟิลเตอร์

(i, j) คือ พิกัดในผลลัพธ์ของคอนโวลูชัน

ขั้นที่ 3 Activation Function ฟังก์ชันเปิดใช้ (Activation Function) ถูกนำมาใช้เพื่อเพิ่มไม่เชิงเส้นให้กับโมเดล ซึ่งช่วยให้สามารถเรียนรู้จากข้อมูลที่ซับซ้อนขึ้นได้ ฟังก์ชันที่นิยมใช้คือ ReLU (Rectified Linear Unit)

$$f(x) = \max(0, x) \quad (5)$$

ขั้นที่ 4 Pooling Layer เลเยอร์พูลลิงมักจะใช้ Max Pooling เพื่อลดขนาดของข้อมูล

$$P(i, j) = \max_{(m, n)} \epsilon_{pool_region}(i+m, j+n) \quad (6)$$

ขั้นที่ 5 Fully Connected Layer ในเลเยอร์ Fully Connected แต่ละนิวรอนจะเชื่อมต่อกับทุกนิวรอนในเลเยอร์ก่อนหน้า ผลลัพธ์คำนวณโดยใช้

$$y = Wx + b \quad (7)$$

โดยที่

y คือ ผลลัพธ์

W คือ เวกเตอร์น้ำหนัก

x คือ เวกเตอร์อินพุต

b คือ ไบแอส

บทความวิจัย (Research Article)

ขั้นตอนที่ 6 Loss Function เพื่อฝึกฝนโมเดล CNN ต้องมีฟังก์ชันการสูญเสีย (Loss Function) เพื่อวัดความแตกต่างระหว่างการคาดการณ์ของโมเดล และค่าจริง (Ground Truth) ตัวอย่างหนึ่งคือ Cross-Entropy Loss สำหรับปัญหาการจำแนก

$$L = -\sum y_i \log(p_i) \quad (8)$$

โดยที่

Y_i คือ ค่าจริง

P_i คือ ค่าที่คาดการณ์

ขั้นตอนที่ 7 การปรับน้ำหนักในระหว่างการฝึกฝนโมเดลใช้วิธี Backpropagation และการลดค่า Gradient (Gradient Descent)

2.7 ความสัมพันธ์ของค่า Brix กับระดับความสุกของผลไม้

ค่า Brix เป็นการวัดความเข้มข้นของสารละลายน้ำตาลในของเหลว ซึ่งมักใช้เพื่อวัดความหวานของผลไม้ โดยค่า Brix หนึ่งหน่วยเท่ากับ 1 กรัมของซูโครสในน้ำ 100 กรัม ดังนั้น ค่า Brix จึงสามารถแสดงถึงปริมาณน้ำตาลในผลไม้ได้

ผลไม้ไม่สุก ในระยะแรกการพัฒนาของผลไม้ ปริมาณน้ำตาลในผลไม้จะน้อยและค่า Brix จะต่ำซึ่งหมายความว่าผลไม้ยังไม่สุกและมีรสชาติที่ไม่หวานมาก

ผลไม้สุกเต็มที่ เมื่อผลไม้พัฒนาและเข้าสู่ระยะที่สุกเต็มที่ ปริมาณน้ำตาลในผลไม้จะเพิ่มขึ้นอย่างมาก ทำให้ค่า Brix สูงขึ้น ซึ่งหมายความว่าผลไม้มีความหวานและมีรสชาติที่ดีที่สุด

ผลไม้สุกเกินไป หากผลไม้สุกเกินไปปริมาณน้ำตาลอาจลดลงเนื่องจากการเปลี่ยนแปลงทางเคมี

ตารางที่ 1 ค่า Brix ของฟักทอง

ระดับการสุก	ค่า Brix	ระดับความหวาน
ฟักทองดิบ	5-8	หวานน้อย
ฟักทองสุก	9-14	หวานปานกลาง
ฟักทองสุกเกินไป	15 ขึ้นไป	หวานมาก

จากตารางที่ 1 สรุปว่า ค่า Brix สามารถใช้บ่งชี้ระดับความสุกของฟักทองได้ โดยค่า Brix ยิ่งสูง ฟักทองยิ่งสุกและมีความหวานมากขึ้น [8]

2.8 เปรียบเทียบและการพัฒนาต่อยอดจากงานวิจัยที่ผ่านมา

จากการทบทวนวรรณกรรมพบว่า มีงานวิจัยหลายฉบับที่นำเทคโนโลยี การประมวลผลภาพ และเสียง มาใช้ในการจำแนกระดับความสุกของผลไม้ เช่น มะม่วง มะเขือเทศ และแตงโม โดยใช้ CNN หรือ Machine Learning ในการเรียนรู้ลักษณะเฉพาะ (Feature) ของข้อมูล

Lee et al. (2020) ใช้ข้อมูลเสียงจากการเคาะแตงโมแปลงเป็น Spectrogram แล้ววิเคราะห์ด้วย CNN ซึ่งได้ความแม่นยำประมาณ 88% [9]

Chen et al. (2019) ใช้ภาพ RGB ของมะเขือเทศกับโมเดล CNN ในการแยกความสุก ได้ F1-Score ประมาณ 0.86 [10]

งานวิจัยก่อนหน้านี้มักใช้ข้อมูลเพียงประเภทเดียว (ภาพหรือเสียง) และมักเน้นไปที่ผลไม้บางชนิด

บทความวิจัย (Research Article)

ตารางที่ 2 ตารางเปรียบเทียบงานวิจัยที่เกี่ยวข้อง

ลำดับ	งานวิจัย (อ้างอิง)	เทคโนโลยีที่ใช้	ข้อมูลนำเข้า	จุดเด่น	ข้อจำกัด
1	[8] Smith et al.	Deep Learning	ภาพฟักทอง	แนวทางอัตโนมัติลดแรงงานคน	ไม่ระบุโมเดลหรือผลลัพธ์เชิงปริมาณ
2	[9] Wang et al.	CNN (Convolutional)	ภาพ RGB	ไม่ทำลายผลผลิต จัดระดับความสุกได้	ยังไม่ทดสอบข้อมูลจริงภาคสนาม
3	[10] Chen et al.	CNN	ภาพฟักทอง	ใช้เทคนิค AI ในการจัดการเกษตร	ยังไม่ครอบคลุมข้อมูลหลายแหล่ง
4	งานวิจัยนี้	CNN + Multimodal	ภาพ + เสียง	ความแม่นยำสูงสุด (F1 = 0.93) ใช้ข้อมูลหลายมิติ	ข้อมูลจำนวนน้อย ไวต่อเสียงรบกวน

จากตารางที่ 2 พบว่าเทคโนโลยี Deep Learning โดยเฉพาะ Convolutional Neural Network (CNN) ถูกนำมาใช้ในการจำแนกระดับความสุกของฟักทอง โดยอาศัยข้อมูลภาพเป็นหลัก เสนอแนวทางอัตโนมัติที่ช่วยลดภาระแรงงานมนุษย์ แต่ยังมีขาดรายละเอียดเชิงเทคนิคและผลลัพธ์เชิงปริมาณ ขณะที่ Wang et al. พัฒนาโมเดล CNN จากภาพ RGB ซึ่งสามารถจัดระดับความสุกได้โดยไม่ทำลายผลผลิต แต่ยังไม่ผ่านการทดสอบในสภาพแวดล้อมจริง ส่วนงานของ Chen et al. ชี้ให้เห็นถึงศักยภาพของ AI ในการจัดการเกษตรกรรม แต่ยังมีข้อจำกัดในด้านความหลากหลายของแหล่งข้อมูล

การคัดแยกคุณภาพของทุเรียนโดยใช้เทคโนโลยีการประมวลผลภาพ (Image Processing) ร่วมกับโครงข่ายประสาทเทียมแบบคอนโวลูชัน (CNN) เพื่อจำแนกคุณภาพของทุเรียนอัตโนมัติ ใช้เทคนิคการเรียนรู้เชิงลึก (Deep Learning) โดยเฉพาะ CNN ในการวิเคราะห์ภาพถ่ายของทุเรียน เพื่อจำแนกคุณภาพ เช่น ความสุก หรือสภาพของทุเรียน [11]

การพัฒนาแอปพลิเคชันเพื่อจำแนกความสุกของทุเรียนจากเสียงเคาะโดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network: CNN) การใช้เทคโนโลยีการวิเคราะห์เสียง (Acoustic

Analysis) เพื่อวัดเสียงเคาะทุเรียนและประเมินระดับความสุกของทุเรียน [12]

3. วิธีดำเนินการวิจัย

3.1 ศึกษาเอกสาร และงานวิจัยที่เกี่ยวข้อง

เพื่อใช้ในการกำหนดวัตถุประสงค์และขอบเขตของการวิจัย

3.2 การเตรียมข้อมูล

3.2.1 ข้อมูลภาพ

ใช้ชุดข้อมูลรูปภาพ (Data set) ฟักทอง โดยทำการแบ่งข้อมูลสำหรับใช้ในการเรียนรู้ (Training Set) ตามค่า Brix จำนวน 900 ภาพ ใช้สำหรับชุดทดสอบความถูกต้อง (Validation Set) จำนวน 150 ภาพ และชุดทำนาย (Prediction) จำนวน 100 ภาพ โดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันในการสร้างโมเดล และทำการทดสอบการทำงาน การทดสอบโดยได้แบ่งประเภทของข้อมูลออกเป็น 3 คลาส ได้แก่ ฟักทองดิบ ฟักทองสุก ฟักทองสุกเกินไป

บทความวิจัย (Research Article)

3.2.2 ข้อมูลเสียง

ใช้ชุดข้อมูลเสียง (Data set) ฟักทอง โดยทำการแบ่งข้อมูลสำหรับใช้ในการเรียนรู้ (Training Set) ตามค่า Brix จำนวน 900 เสียง ใช้สำหรับชุดทดสอบความถูกต้อง (Validation Set) จำนวน 150 เสียง และชุดทำนาย (Prediction) จำนวน 100 เสียง โดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันในการสร้างโมเดล และทำการทดสอบการทำงานการทดสอบโดยได้แบ่ง ประเภทของข้อมูลออกเป็น 3 คลาส ได้แก่ ฟักทองดิบ ฟักทองสุก ฟักทองสุกเกินไป และทำการแปลงเสียงฟักทองแต่ละคลาส เป็น Spectrogram

3.2.3 ข้อมูลภาพและเสียง

ใช้ชุดข้อมูลภาพและเสียง โดยเสียงฟักทองแปลงเป็น Spectrogram ก่อนจากนั้นนำข้อมูลภาพและ Spectrogram (Data Set) ฟักทอง โดยทำการแบ่งข้อมูลสำหรับใช้ในการเรียนรู้ (Training Set) ตามค่า Brix จำนวนภาพและ Spectrogram อย่างละ 900 ภาพ/Spectrogram ใช้สำหรับชุดทดสอบความถูกต้อง (Validation Set) จำนวนภาพและ Spectrogram 150 ภาพ/Spectrogram และชุดทำนาย (Prediction) จำนวนภาพและ Spectrogram 100 ภาพ/Spectrogram โดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันในการสร้างโมเดล

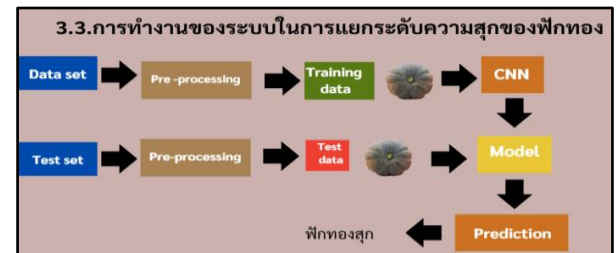
การทดสอบโดยได้แบ่งประเภทของข้อมูลออกเป็น 3 คลาส ได้แก่ ฟักทองดิบ ฟักทองสุกและ ฟักทองสุกเกินไปดังรูปที่ 3



รูปที่ 3 การแบ่งฟักทองออกเป็น 3 คลาส ตามค่า Brix

3.3 ออกแบบแนวคิดระบบการทำงานการแยกระดับความสุกของฟักทอง

3.3.1 การทำงานของระบบการแยกระดับความสุกของฟักทองด้วยภาพ



รูปที่ 4 การทำงานของระบบการแยกระดับความสุกของฟักทองด้วยภาพ

จากรูปที่ 4 บล็อกที่ 1 คือ Data Set เป็นส่วนของชุดข้อมูลฟักทอง ที่ได้จัดเตรียมข้อมูลรูปภาพฟักทอง จากการวัดค่า Brix โดยมีการเปรียบเทียบข้อมูลค่า Brix ของฟักทอง พร้อมชุดข้อมูลการเรียนรู้และข้อมูลชุดทดสอบความถูกต้องเพื่อประเมินประสิทธิภาพโมเดล

บล็อกที่ 2 คือ Pre-Processing เป็นกระบวนการที่ใช้สำหรับการกำจัดและลดสัญญาณรบกวน จากภาพ จากนั้นจึงนำไปให้โมเดลตรวจจับวัตถุที่ถูกฝึกสอนแล้ว ทำการวิเคราะห์

บล็อกที่ 3 คือ CNN เป็นอัลกอริทึมหลักที่ใช้สำหรับสร้างโมเดลเพื่อจำแนกประเภทการแยกระดับความสุก

บล็อกที่ 4 Model เป็นโมเดลที่ถูกสร้างด้วยอัลกอริทึม จากกระบวนการเรียนรู้ของข้อมูลและทำการปรับค่าน้ำหนัก (Weight) ให้เหมาะสมกับชุดข้อมูลฟักทองในการนำไปใช้งาน

บล็อกที่ 5 Test Data เป็นชุดข้อมูลรูปภาพฟักทองที่ใช้ในการทดสอบระบบและค่า Brix เพื่อหา

บทความวิจัย (Research Article)

ประสิทธิภาพความแม่นยำในการจำแนกประเภทของ
ฟักทอง โดยเป็นชุดข้อมูลที่ไม่ได้ผ่านกระบวนการ
เรียนรู้มาก่อน

บล็อกที่ 6 Training Data เป็นส่วนการนำชุด
ข้อมูลรูปภาพฟักทองและค่า Brix มาเรียนรู้และฝึกฝน

บล็อกที่ 7 Prediction เป็นกระบวนการที่ได้จาก
การวิเคราะห์และการทำนายผลการแยกแยะโดยการ

โมเดล Convolutional Neural Network (CNN)
ถูกฝึกด้วยพารามิเตอร์ต่อไปนี้ โดยใช้ TensorFlow

จำนวน Epoch: 25 รอบการฝึก

Batch size: 32 ตัวอย่างต่อรอบการฝึก

Optimizer: Adam

Learning Rate: 0.001

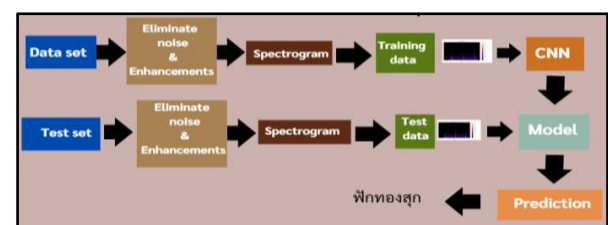
Loss Function: Categorical Cross Entropy

ตารางที่ 3 แสดงผลการทำงานของโมเดล CNN ภาพฟักทอง

Layer	Output Shape	Param #	คำอธิบาย
Conv2D	(126, 126, 32)	896	เลเยอร์คอนโวลูชัน 32 ฟิลเตอร์ ขนาด 3x3 ใช้ฟังก์ชัน ReLU
MaxPooling2D	(63, 63, 32)	0	เลเยอร์ Max Pooling ลดขนาดภาพโดยใช้ฟิลเตอร์ 2x2
Conv2D	(61, 61, 64)	18,496	เลเยอร์คอนโวลูชัน 64 ฟิลเตอร์ ขนาด 3x3 ใช้ฟังก์ชัน ReLU
MaxPooling2D	(30, 30, 64)	0	เลเยอร์ Max Pooling ลดขนาดภาพโดยใช้ฟิลเตอร์ 2x2
Conv2D	(28, 28, 128)	73,856	เลเยอร์คอนโวลูชัน 128 ฟิลเตอร์ ขนาด 3x3 ใช้ฟังก์ชัน ReLU
MaxPooling2D	(14, 14, 128)	0	เลเยอร์ Max Pooling ลดขนาดภาพโดยใช้ฟิลเตอร์ 2x2
Flatten	(25088)	0	แปลงข้อมูลจาก 3D เป็น 1D
Dense	(128)	3,221,120	เลเยอร์ Fully connected 128 หน่วย ใช้ฟังก์ชัน ReLU
Dense (Output)	(3)	387	เลเยอร์เอาต์พุต 3 หน่วย ใช้ฟังก์ชัน SoftMax สำหรับการจำแนกคลาส

จากตารางที่ 3 โมเดลนี้เป็นโครงข่ายประสาท
เทียมแบบ CNN ที่ประกอบด้วยเลเยอร์คอนโวลูชัน 3
ชั้นสลับกับ MaxPooling เพื่อลดขนาดภาพ ตามด้วย
การแปลงข้อมูลเป็นเวกเตอร์ 1 มิติ (Flatten) และ
เลเยอร์ Dense 2 ชั้น โดยใช้ ReLU และ SoftMax
ตามลำดับ เหมาะสำหรับการจำแนกภาพออกเป็น 3
คลาส มีจำนวนพารามิเตอร์รวมประมาณ 3.3 ล้านค่า

3.3.2 การทำงานของระบบการแยกระดับความสุก ของฟักทองด้วยเสียง



รูปที่ 5 การทำงานของระบบการแยกระดับความสุก ของฟักทองด้วยเสียง

บทความวิจัย (Research Article)

จากรูปที่ 5 บล็อกที่ 1 คือ Data Set เป็นส่วนของชุดข้อมูลฟักทอง ที่ผู้วิจัยได้จัดเตรียมข้อมูลเสียงฟักทอง จากการวัดค่า Brix โดยมีการเปรียบเทียบข้อมูลค่า Brix ของฟักทอง พร้อมชุดข้อมูลการเรียนรู้และข้อมูลชุดทดสอบความถูกต้อง

บล็อกที่ 2 คือ Eliminate Noise & Enhancements กระบวนการในการกำจัดสัญญาณรบกวน (Noise) และปรับปรุงคุณภาพ (Enhancements) ของข้อมูลหรือสัญญาณ เพื่อให้ได้ผลลัพธ์ที่ชัดเจนและมีประสิทธิภาพ

บล็อกที่ 3 คือ Spectrogram การแสดงผลสัญญาณเสียงหรือสัญญาณคลื่นในรูปแบบของภาพ

บล็อกที่ 4 Training Data เป็นส่วนการนำชุดข้อมูล Spectrogram ตามค่า Brix มาเรียนรู้

บล็อกที่ 5 Model เป็นโมเดลที่ถูกสร้างด้วยอัลกอริทึม จากกระบวนการเรียนรู้ของข้อมูลและทำการปรับค่าน้ำหนัก (Weight) ให้เหมาะสมกับชุดข้อมูล Spectrogram ฟักทองในการนำไปใช้งาน

บล็อกที่ 6 Test set เป็นชุดข้อมูลเสียงฟักทองที่ใช้ในการทดสอบระบบและค่า Brix เพื่อหาประสิทธิภาพความแม่นยำในการจำแนกประเภทของฟักทอง โดยเป็นชุดข้อมูลที่ไม่ได้ผ่านกระบวนการเรียนรู้มาก่อน

บล็อกที่ 7 Test Data เป็นส่วนการนำชุดข้อมูล Spectrogram ตามค่า Brix มาทดสอบ

บล็อกที่ 8 Prediction เป็นกระบวนการที่ได้จากการวิเคราะห์และการทำนายผลการแยกแยะโดยการตรวจจ็ับวัตถุในภาพเพื่อทำนายวัตถุที่อยู่ในกรอบ

ตารางที่ 4 แสดงผลการทำงานของโมเดล CNN เสียงฟักทอง

Layer	Output Shape	Param #	คำอธิบาย
Conv2D	(126, 126, 32)	896	เลเยอร์คอนโวลูชัน 32 ฟิลเตอร์ ขนาด 3x3 ใช้ฟังก์ชัน ReLU
MaxPooling2D	(63, 63, 32)	0	เลเยอร์ Max Pooling ลดขนาดภาพโดยใช้ฟิลเตอร์ 2x2
Conv2D	(61, 61, 64)	18,496	เลเยอร์คอนโวลูชัน 64 ฟิลเตอร์ ขนาด 3x3 ใช้ฟังก์ชัน ReLU
MaxPooling2D	(30, 30, 64)	0	เลเยอร์ Max Pooling ลดขนาดภาพโดยใช้ฟิลเตอร์ 2x2
Conv2D	(28, 28, 128)	73,856	เลเยอร์คอนโวลูชัน 128 ฟิลเตอร์ ขนาด 3x3 ใช้ฟังก์ชัน ReLU
MaxPooling2D	(14, 14, 128)	0	เลเยอร์ Max Pooling ลดขนาดภาพโดยใช้ฟิลเตอร์ 2x2
Flatten	(25088)	0	แปลงข้อมูลจาก 3D เป็น 1D
Dense	(256)	6,436,096	การเชื่อมต่อแบบ Fully Connected มี 256 หน่วย
Dense (Output)	(3)	771	เลเยอร์เอาต์พุต 3 หน่วย ใช้ฟังก์ชัน SoftMax สำหรับการจำแนกคลาส

จากตารางที่ 4 โมเดลนี้เป็นโครงข่ายประสาทเทียมแบบ CNN ที่ประกอบด้วยเลเยอร์คอนโวลูชัน 3 ชั้น ร่วมกับ MaxPooling เพื่อลดขนาดภาพ

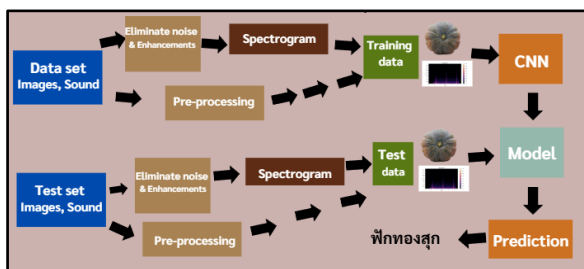
ตามด้วยเลเยอร์ Flatten สำหรับแปลงข้อมูลเป็นเวกเตอร์ 1 มิติ และเลเยอร์ Dense 2 ชั้น โดยมีหน่วย 256 หน่วยในชั้นแรก และ 3 หน่วยในชั้นเอาต์พุต

บทความวิจัย (Research Article)

ใช้ฟังก์ชัน SoftMax สำหรับการจำแนกคลาส รวมจำนวนพารามิเตอร์ประมาณ 6.5 ล้านค่า เหมาะสำหรับงานจำแนกภาพหลายคลาส

3.3.3 การทำงานของระบบการแยกระดับความสูงของฟักทองด้วยภาพและเสียงเพื่อประเมินประสิทธิภาพโมเดล

จากรูปที่ 6 บล็อกที่ 1 คือ Data Set Images, Sound เป็นส่วนของชุดข้อมูลฟักทอง ที่ได้จัดเตรียมข้อมูลเสียงและภาพฟักทอง จากการวัดค่า Brix โดยมีการเปรียบเทียบข้อมูลค่า Brix ของฟักทอง พร้อมชุดข้อมูลการเรียนรู้และข้อมูลชุดทดสอบความถูกต้องเพื่อประเมินประสิทธิภาพโมเดล



รูปที่ 6 การทำงานของระบบการแยกระดับความสูงของฟักทองด้วยภาพและเสียง

บล็อกที่ 2 คือ Eliminate Noise & Enhancements & Pre-Processing กระบวนการในการกำจัดสัญญาณรบกวน (Noise) และ ปรับปรุงคุณภาพ (Enhancements) ของข้อมูลหรือสัญญาณเพื่อให้ได้ผลลัพธ์ที่ชัดเจน

บล็อกที่ 3 คือ Spectrogram การแสดงผลสัญญาณเสียงหรือสัญญาณคลื่นในรูปแบบของภาพ

บล็อกที่ 4 Training data เป็นส่วนการนำชุดข้อมูล Spectrogram และภาพตามค่า Brix มาเรียนรู้

บล็อกที่ 5 Model เป็นโมเดลที่ถูกสร้างด้วยอัลกอริทึม จากกระบวนการเรียนรู้ของข้อมูลและทำการปรับค่าน้ำหนัก (Weight) ให้เหมาะสมกับชุดข้อมูล Spectrogram และภาพฟักทองในการนำไปใช้งาน

บล็อกที่ 6 Test set เป็นชุดข้อมูลเสียงฟักทองที่ใช้ในการทดสอบระบบและค่า Brix เพื่อหาประสิทธิภาพความแม่นยำในการจำแนกประเภทของฟักทอง โดยเป็นชุดข้อมูลที่ไม่ได้ผ่านกระบวนการเรียนรู้มาก่อน

บล็อกที่ 7 Test Data เป็นส่วนการนำชุดข้อมูล Spectrogramและภาพตามค่า Brix มาทดสอบ

บล็อกที่ 8 Prediction เป็นกระบวนการที่ได้จากการวิเคราะห์และการทำนายผลการแยกแยะโดยการตรวจจับวัตถุในภาพเพื่อทำนายวัตถุที่อยู่ในกรอบพร้อมระบุค่าความมั่นใจ (Confidence) ด้วย

ตารางที่ 5 แสดงผลการทำงานของโมเดล CNN ภาพกับเสียงฟักทอง

Layer	Output Shape	Param #	คำอธิบาย
Conv2D (32)	(126, 126, 32)	1,792	เลเยอร์ Convolution แรกใช้ 32 ฟิลเตอร์ ขนาด 3x3, ขนาด input (128, 128, 6)
MaxPooling2D	(63, 63, 32)	0	การลดขนาดภาพ (Down Sampling) ด้วย Max Pooling ขนาด (2, 2)
Conv2D	(61, 61, 32)	9,248	เลเยอร์ Convolution ที่สอง ใช้ 32 ฟิลเตอร์ขนาด 3x3

ฐานันดร มูลสาร และคณะ, การพัฒนาระบบคัดแยกระดับความสูงของฟักทองด้วยการผสานข้อมูลภาพและเสียงโดยใช้เทคนิคการเรียนรู้ของเครื่อง

บทความวิจัย (Research Article)

ตารางที่ 5 แสดงผลการทำงานของโมเดล CNN ภาพกับเสียงฟักทอง (ต่อ)

Layer	Output Shape	Param #	คำอธิบาย
MaxPooling2D	(30, 30, 32)	0	Max Pooling ขนาด (2, 2)
Conv2D	(28, 28, 64)	18,496	เลเยอร์ Convolution ที่สาม ใช้ 64 ฟิลเตอร์ขนาด 3x3
MaxPooling2D	(14, 14, 64)	0	Max Pooling ขนาด (2, 2)
Conv2D	(12, 12, 64)	36,928	เลเยอร์ Convolution ที่สี่ ใช้ 64 ฟิลเตอร์ขนาด 3x3
MaxPooling2D	(6, 6, 64)	0	Max Pooling ขนาด (2, 2)
Dropout (0.2)	(6, 6, 64)	0	ใช้ Dropout เพื่อป้องกันการ Overfitting
Conv2D	(4, 4, 128)	73,856	เลเยอร์ Convolution ที่ห้า ใช้ 128 ฟิลเตอร์ขนาด 3x3
MaxPooling2D	(2, 2, 128)	0	Max Pooling ขนาด (2, 2)
Dropout (0.2)	(2, 2, 128)	0	ใช้ Dropout เพื่อป้องกันการ Overfitting
Conv2D	(0, 0, 128)	147,584	เลเยอร์ Convolution ที่หก ใช้ 128 ฟิลเตอร์ขนาด 3x3
MaxPooling2D	(0, 0, 128)	0	Max Pooling ขนาด (2, 2)
Dropout (0.2)	(0, 0, 128)	0	ใช้ Dropout เพื่อป้องกันการ overfitting
Flatten	(0, 0)	0	แปลงข้อมูลที่มีหลายมิติให้เป็นเวกเตอร์ 1 มิติ
Dense	(256)	32,768	เลเยอร์ Dense ที่มี 256 หน่วย ใช้ฟังก์ชัน ReLU
Dense	(3)	771	เลเยอร์เอาต์พุต 3 หน่วย ใช้ฟังก์ชัน SoftMax สำหรับการจำแนกคลาส 3 คลาส

จากตารางที่ 5 งานวิจัยเลือกใช้ CNN เนื่องจากเหมาะสมกับการประมวลผลทั้ง ภาพฟักทอง และเสียงที่แปลงเป็น Spectrogram โดย CNN สามารถเรียนรู้ลักษณะสำคัญของข้อมูลได้อัตโนมัติโดยไม่ต้องกำหนดคุณลักษณะ (Feature) ล่วงหน้า โครงสร้างของโมเดลถูกออกแบบให้เหมาะกับชุดข้อมูลขนาดเล็ก ลดความซับซ้อนเพื่อป้องกัน Overfitting และสามารถประยุกต์ใช้กับข้อมูลภาพและเสียงได้อย่างมีประสิทธิภาพในการจำแนกระดับความสุกของฟักทอง

3.4 เก็บรวบรวมข้อมูล

ได้ศึกษาข้อมูลฟักทองอำพัน 342 ที่ตำบลนากระแซง อำเภอดงเจริญ จังหวัดอุบลราชธานี จำนวน 30 ลูก โดยถ่ายฟักทองมูมด้านบน หมุนภาพของฟักทองครั้งละ 20 องศา พร้อมด้วยการบันทึกเสียงฟักทองโดยใช้ตะเกียบเหล็กป้อนมากระแทบผลฟักทองด้านบน การถ่ายภาพและบันทึกเสียงฟักทองตามค่า Brix ผลลัพธ์ค่า Brix ของฟักทอง ผู้วิจัยได้นำฟักทองไปตรวจหาค่า Brix ที่ห้องปฏิบัติการกลาง คณะเกษตรศาสตร์ มหาวิทยาลัยราชภัฏอุบลราชธานี ในการตรวจค่า Brix ของฟักทอง 1 ลูกจะตรวจหาค่า Brix อยู่ 3 ครั้ง โดยการนำเนื้อฟักทองมาบดให้ละเอียด

บทความวิจัย (Research Article)

แล้วใส่บีกเกอร์ให้มีปริมาตรเท่ากัน คั้นน้ำของฟักทอง แล้วมาตรวจด้วย Brix Refractometer 0-32% Brix รุ่น RHB-32ATC S ก็จะได้อ่านค่า Brix ของ ฟักทอง ดังตารางที่ 6

ตารางที่ 6 ตัวอย่างค่า Brix ของฟักทองที่ใช้การเทรน

ตัวอย่างที่	ชื่ออย่าง	ค่า Brix	ภาพฟักทอง
1	T1/1	10.0	
	T1/2	10.4	
	T1/3	11.0	
	เฉลี่ย	10.4666	
2	T2/1	7.0	
	T2/2	7.2	
	T2/3	7.4	
	เฉลี่ย	7.2000	
3	T3/1	17.0	
	T3/2	17.0	
	T3/3	16.8	
	เฉลี่ย	16.9333	

จากตารางที่ 6 แสดงให้เห็นค่า Brix ของฟักทอง นำรูปฟักทองมาเทรนข้อมูล โดยแบ่งการเทรนข้อมูลตามค่า Brix ของฟักทอง ภาพที่ 1 แสดงฟักทองสุก ภาพที่ 2 แสดงฟักทองดิบและภาพที่ 3 แสดงฟักทองสุกเกินไป

3.5 เครื่องมือการวิจัย

1) Brix Refractometer 0-32% Brix รุ่น RHB - 32ATC S เครื่องวัดความหวานของผลไม้

2) ชุดอุปกรณ์และเครื่องมือ ซึ่งประกอบด้วยฐานรองรับซึ่งเป็นแท่นโต๊ะแม่เหล็กที่มีส่วนต่าง ๆ

ประกอบด้วยแกนหลัก ที่ยึดจับโทรศัพท์มือถือและอุปกรณ์ปรับความละเอียดของระยะที่ต้องการทดลอง ตะเกียบเหล็กที่ใช้ในการทดลอง

3) นำฟักทองแต่ละคลาสไปตรวจความหวาน ในห้องปฏิบัติการกลาง คณะเกษตรศาสตร์ มหาวิทยาลัยราชภัฏอุบลราชธานี เพื่อตรวจสอบความ หวานของฟักทองแต่ละลูก โดยฟักทอง 1 ลูก จะทำการตรวจทั้งหมด 3 ครั้ง ผลที่ได้จะมีค่าเป็น Brix



รูปที่ 7 การตรวจความหวานจากห้องปฏิบัติการกลาง คณะเกษตรศาสตร์ มหาวิทยาลัยราชภัฏอุบลราชธานี

จากรูปที่ 7 แสดงการตรวจความหวานจากห้องปฏิบัติการกลาง คณะเกษตรศาสตร์ มหาวิทยาลัยราชภัฏอุบลราชธานี

3.6. กลุ่มตัวอย่าง

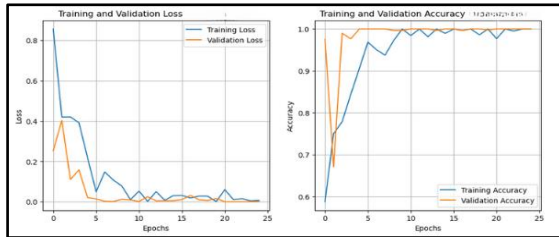
กลุ่มตัวอย่าง คือ ฟักทองอำพัน 342 จำนวน 30 ลูก โดยแบ่งฟักทองออกเป็น 3 คลาส ตามค่า Brix ได้แก่ฟักทองดิบ ฟักทองสุก และฟักทองสุกเกินไป

4. ผลการวิจัย

4.1 ผลการเทรนข้อมูลภาพฟักทอง

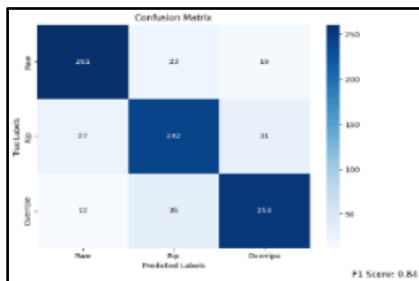
ข้อมูลฟักทองประกอบด้วยชุดข้อมูลการฝึกชุดข้อมูลการทดสอบเข้าสู่ระบบการฝึกฝน ทั้งหมดครบ 25 รอบเต็ม เพื่อปรับปรุงค่าเวท (Weights) และอคติ (Biases) ของโมเดลในแต่ละรอบตามค่าความสูญเสีย (Loss) ที่คำนวณได้ ดังรูปที่ 8

บทความวิจัย (Research Article)



รูปที่ 8 กราฟแสดงการเทรนข้อมูลภาพฟักทอง

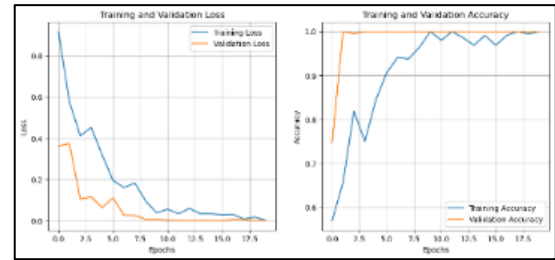
กราฟแสดงค่าความแม่นยำของโมเดลชุดฝึกและชุดทดสอบความถูกต้องของฟักทอง ในจำนวน 25 รอบแสดงค่าว่าโมเดลมีค่าความถูกต้อง ในการเทรนและทดสอบอยู่ที่ 100.00 เปอร์เซนต์ ตั้งแต่รอบที่ 4 เป็นต้นไป ประสิทธิภาพโมเดลของภาพฟักทอง ค่า F1 Score มีค่าเป็น 0.84 ซึ่งเป็นค่าที่ใช้ในการวัด แสดงได้ดังรูปที่ 9



รูปที่ 9 Confusion Matrix ภาพฟักทอง

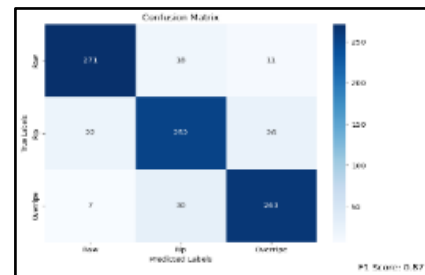
4.2 ผลการเทรนข้อมูลเสียงฟักทอง

ในรูปที่ 10 ค่า Loss ของข้อมูล Spectrogram ฟักทองและข้อมูล Validation Loss ข้อมูลจำนวน 20 รอบ ค่า Loss ของทั้ง Training และ Validation ลดลงเรื่อย ๆ เมื่อ Epochs เพิ่มขึ้น Training Loss และ Validation Loss ใกล้เคียงกันมากในช่วงหลัง แสดงว่าไม่มี Overfitting



รูปที่ 10 กราฟแสดงการเทรนข้อมูลเสียงฟักทอง

กราฟแสดงค่าความแม่นยำของโมเดลชุดฝึกและชุดทดสอบความถูกต้องของเสียงฟักทอง ใน จำนวน 20 รอบ แสดงค่าว่าโมเดลมีค่าความถูกต้องตั้งแต่รอบที่ 1 และการเทรนอยู่ที่ 100.00 เปอร์เซนต์ ในรอบที่ 8 แล้วก็ผันผวนขึ้นลง ประสิทธิภาพโมเดลของภาพฟักทอง ค่า F1 Score มีค่าเป็น 0.87 ซึ่งเป็นค่าที่ใช้ในการวัด แสดงได้ดังรูปที่ 11

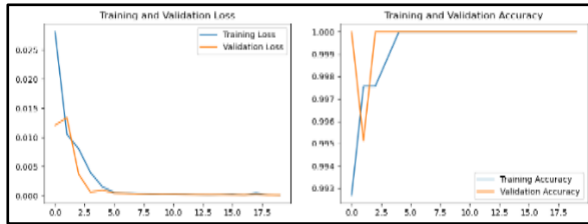


รูปที่ 11 Confusion Matrix เสียงฟักทอง

4.3 ผลการเทรนข้อมูลภาพและ Spectrogram

จากรูปที่ 12 แสดงค่า Loss ของข้อมูลชุดเทรนภาพและ Spectrogram ฟักทองและข้อมูลการตรวจสอบความถูกต้องของการเทรนข้อมูลจำนวน 20 รอบ แสดงให้เห็นว่าข้อมูลการเทรนและข้อมูลการตรวจสอบความถูกต้องมีค่าใกล้เคียง 0 ตั้งแต่รอบที่ 5

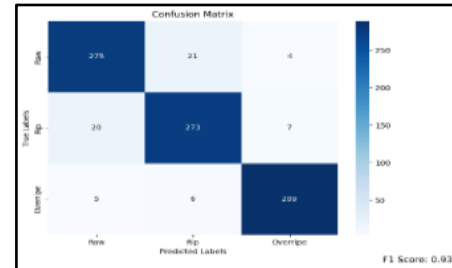
บทความวิจัย (Research Article)



รูปที่ 12 กราฟแสดงการเทรนข้อมูลภาพกับเสียง
ฟัททอง

กราฟแสดงค่าความแม่นยำของโมเดลชุดฝึกและชุดทดสอบความถูกต้องของภาพและ Spectrogram ฟัททอง จำนวน 20 รอบ แสดงค่าว่าโมเดลมีความถูกต้องตั้งแต่รอบที่ 2 และการเทรนอยู่ที่ 100.00 เปอร์เซนต์ ในรอบที่ 5

ประสิทธิภาพโมเดลของภาพฟัททอง ค่า F1 Score มีค่าเป็น 0.93 ซึ่งเป็นค่าที่ใช้ในการวัดประสิทธิภาพของโมเดลในการจำแนกข้อมูล (Classification Model) แสดงได้ดังรูปที่ 13



รูปที่ 13 Confusion Matrix ภาพและ
Spectrogram ฟัททอง

ตารางที่ 7 สรุป Precision, Recall และ F1-Score ของโมเดลจำแนกระดับความสูงของฟัททอง

ข้อมูลที่ใช้	ระดับความสูง	Precision	Recall	F1-score	เฉลี่ยรวม
ภาพ (Image)	ดิบ	0.85	0.82	0.83	
	สูง	0.83	0.84	0.84	
	สูงเกินไป	0.80	0.85	0.82	
	เฉลี่ยรวม	0.83	0.84	0.84	✓
เสียง (Audio)	ดิบ	0.88	0.85	0.86	
	สูง	0.86	0.87	0.87	
	สูงเกินไป	0.85	0.88	0.86	
	เฉลี่ยรวม	0.86	0.87	0.87	✓
ภาพ + Spectrogram	ดิบ	0.94	0.91	0.92	
	สูง	0.93	0.94	0.94	
	สูงเกินไป	0.92	0.93	0.93	
	เฉลี่ยรวม	0.93	0.93	0.93	✓

บทความวิจัย (Research Article)

จากตารางที่ 7 การใช้ข้อมูลหลายมิติ (ภาพร่วมกับเสียง) ช่วยเพิ่มความแม่นยำในการจำแนกระดับความสุกของฟักทองได้ดีที่สุด

5. อภิปรายผล

ผลการทดลองแสดงให้เห็นว่าโมเดลที่ผสานข้อมูลภาพและเสียงของฟักทองสามารถจำแนกระดับความสุกได้อย่างแม่นยำ โดยมีค่าความถูกต้องเฉลี่ย 93.00 เปอร์เซ็นต์ ซึ่งสูงกว่าโมเดลที่ใช้ข้อมูลภาพอย่างเดียว (84.00 เปอร์เซ็นต์) และเสียงอย่างเดียว (87.00 เปอร์เซ็นต์) สะท้อนให้เห็นว่าเทคนิคการเรียนรู้แบบหลายมิติ (Multimodal Learning) สามารถช่วยเพิ่มประสิทธิภาพของโมเดลได้อย่างมีนัยสำคัญ สาเหตุที่โมเดลผสานข้อมูลให้ผลลัพธ์ดีกว่า มาจากการที่ข้อมูลภาพให้รายละเอียดภายนอก เช่น สีและพื้นผิว ในขณะที่ข้อมูลเสียงสะท้อนคุณสมบัติภายใน เช่น ความแน่นและความกลวงของเนื้อฟักทอง การรวมข้อมูลจากทั้งสองแหล่งจึงช่วยให้โมเดลเข้าใจลักษณะของฟักทองในมิติต่าง ๆ ได้ครอบคลุมยิ่งขึ้น แนวคิดนี้สอดคล้องกับงานวิจัยของเวียงชัย คาระมาตย์ [12] ที่ใช้เสียงเคาะในการจำแนกระดับความสุกของทุเรียน โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน (CNN) อย่างไรก็ตามงานวิจัยนี้ยังมีข้อจำกัดเชิงเทคนิคหลายประการ ได้แก่ ความหลากหลายของข้อมูลที่ยังจำกัดอยู่ในฟักทองพันธุ์อำพัน 342 เพียงชนิดเดียว ซึ่งอาจส่งผลกระทบต่อความสามารถในการนำโมเดลไปใช้งานกับฟักทองสายพันธุ์อื่นที่มีค่า Brix แตกต่างกัน รวมถึงความเสี่ยงของปัญหา Overfitting จากจำนวนข้อมูลฝึกที่ยังน้อยเมื่อเทียบกับขนาดของโมเดล นอกจากนี้ การเก็บข้อมูลเสียงในภาคสนามอาจมีปัญหาเรื่องเสียงรบกวนหรือความไม่สม่ำเสมอของวิธีการเคาะ จึงควรมีการพัฒนาเครื่องมือเก็บเสียงที่มีมาตรฐานมากขึ้น

6. สรุป

งานวิจัยนี้ได้นำเสนอแนวทางการจำแนกระดับความสุกของฟักทองโดยใช้เทคนิคการเรียนรู้ของเครื่อง พบว่าโมเดลที่ใช้การผสานข้อมูลภาพและเสียงให้ผลลัพธ์ที่ดีที่สุด โดยมีค่าความถูกต้องเฉลี่ย 93.00 เปอร์เซ็นต์ ซึ่งสูงกว่าโมเดลที่ใช้ข้อมูลภาพหรือเสียงเพียงอย่างเดียวอย่างชัดเจน ผลการวิจัยชี้ให้เห็นถึงศักยภาพของการใช้เทคโนโลยีปัญญาประดิษฐ์ โดยเฉพาะโครงข่ายประสาทเทียมแบบคอนโวลูชัน (CNN) ในการวิเคราะห์ลักษณะของผลผลิตทางการเกษตรได้อย่างมีประสิทธิภาพ ทั้งนี้แนวทางดังกล่าวสามารถนำไปประยุกต์ใช้ในการคัดเกรดผลผลิตฟักทองในภาคสนาม หรือพัฒนาเป็นอุปกรณ์ช่วยตรวจสอบความสุกที่สะดวกรวดเร็ว ลดภาระแรงงาน และเพิ่มคุณภาพการผลิต ในอนาคต ควรมีการขยายการทดลองให้ครอบคลุมฟักทองหลายสายพันธุ์ และพัฒนาอุปกรณ์ต้นแบบสำหรับการตรวจสอบความสุกที่สามารถนำไปใช้จริงในภาคเกษตรกรรมได้อย่างยั่งยืน

7. ข้อเสนอแนะ

งานวิจัยนี้ยังพบข้อจำกัด ได้แก่ การเพิ่มจำนวนของข้อมูลในการเทรนข้อมูล การหาภาพที่มีความแตกต่างกันและการหาค่า Brix ให้มากกว่านี้ จะใช้เวลาและค่าใช้จ่ายค่อนข้างสูง อย่างไรก็ตามคณะผู้วิจัยจะใช้ข้อมูลอื่น ๆ เพิ่มเติมเพื่อช่วยในการจำแนกระดับความสุกของฟักทองร่วมด้วย เช่น น้ำหนัก ขนาด ความแข็งแรงของฟักทอง เป็นต้น

8. เอกสารอ้างอิง

- [1] ระบบสารสนเทศการผลิตทางด้านเกษตร (รต.) กรมส่งเสริมการเกษตร, "รายงานสถานการณ์การปลูกฟักทอง ปี 2562," [ออนไลน์].

บทความวิจัย (Research Article)

- Available: <https://production.doae.go.th>.
[เข้าถึงเมื่อ: 16 ก.ค. 2568].
- [2] A. Amna, M. W. Akram, G. Li, M. Z. Akram, M. Faheem, M. M. Omar, and M. G. Hassan, "Machine vision-based automatic fruit quality detection and grading," *Front. Agr. Sci. Eng.*, vol. 12, no. 2, pp. 274–287, 2025, doi: 10.15302/J-FASE-2023532.
- [3] เอกรัตน์ สุขสุคนธ์, ชลทิศ สันติวงศ์งาม, วัชรพล คำแขก, และ ศศิชา ปานสุวรรณ, "การวิเคราะห์การสุกของผลไม้ด้วยเทคโนโลยีการประมวลผล," *วารสารวิทยาศาสตร์และเทคโนโลยีสารสนเทศ*, ปีที่ 12, ฉบับที่ 1, หน้า 61-66, 2565.
- [4] ศุภกร ชูประภาวรรณ, "การลดเสียงรบกวนของภาพสำหรับเสียงรบกวนแบบเกาส์เซียนโดยใช้การเรียนรู้เชิงลึกและพีเออร์ชอบ," *วิทยานิพนธ์, จุฬาลงกรณ์มหาวิทยาลัย*, 2561. [ออนไลน์]. Available: Chula ETD, No. 3276.
- [5] บงการ หอมนาน และวาทีต เบญจพลกุล, "การจำแนกภาพแบบฟังก์ชันเชิงเส้นการปรับเทียบฮิสโตแกรมผูกพันบนพื้นฐานทฤษฎีสารสนเทศและขั้นตอนวิธีการจัดกลุ่มค่าเฉลี่ยกลุ่มแบบฟัซซีในการส่งสารสนเทศ," *วารสารสุทธิปริทัศน์*, ปีที่ 27, ฉบับที่ 4, หน้า 105-115, ต.ค. 2566 – ม.ค. 2567.
- [6] "File:Spectrogram," Wikimedia Commons, CC BY-SA 3.0. [ออนไลน์]. Available: <https://commons.wikimedia.org/w/index.php?curid=202335>. [Accessed: 12 Mar. 2024].
- [7] J. A. Benitez-Andrades, "How to cite electronic documents," 2020. [ออนไลน์]. Available: https://miro.medium.com/v2/resize:fit:720/format:webp/1*7kRxfXK4J-wCR2l8G1Jtuw.png. [Accessed: Feb. 2020].
- [8] จานุลักษณ์ ขนบดี, สุภาวดี ตั้งธีระวัฒน์, และ ศิริพร มุลาลินน์, "เรื่องสายพันธุ์ฟักทองคุณภาพสูง," ใน *เอกสารประชุมทางวิชาการของมหาวิทยาลัยเกษตรศาสตร์ ครั้งที่ 41: สาขาพืชสาขาส่งเสริมและนิเทศศาสตร์เกษตร*, กรุงเทพฯ, 2546, หน้า 64-69.
- [9] N. A. Ayunda, E. Haryatmi, and T. A. Riyadi, "Classification of Tomato Ripeness Based on Convolutional Neural Network Methods," *J. Inf. Syst. Informatics*, vol. 5, no. 4, 2023, doi: 10.51519/journalisi.v5i4.613.
- [10] K.-E. Ko, S.-H. Yang, and I.-H. Jang, "Real-Time Tomato Ripeness Classification System based on Deep Learning Model for Object Detection," *J. Inst. Control, Robot. Syst.*, vol. 24, no. 11, pp. 999–1004, Nov. 2018, doi: 10.5302/J.ICROS.2018.18.0166.
- [11] มัณฑิณี ถิ่นรัมย์ และ เยวรัตน์ จำปา, "การคัดแยกระดับความสุกของทุเรียนด้วยการเรียนรู้ของเครื่อง," *วารสารวิจัยวิทยาศาสตร์และเทคโนโลยี*, ปีที่ 26, ฉบับที่ 4, หน้า 105-115, 2564.
- [12] เวียงชัย คาระมาตย์, "การพัฒนาแอปพลิเคชันสำหรับการจำแนกความสุกของทุเรียนจากเสียงเคาะโดยใช้เครือข่ายประสาทเทียมแบบคอนโวลูชัน," *Chulalongkorn University Theses and Dissertations (Chula ETD)*, No. 9605, 2563.