

การตรวจจับอาการเป็นสัดตามเวลาจริงในโคนมโดยไม่ติดตั้งอุปกรณ์บนตัวโค
ด้วยเทคนิคการประมวลผลวิดีโอทัศนบนพื้นฐาน
โครงข่ายประสาทเทียมแบบสังวัตนาการ
NON-INVASIVE REAL-TIME ESTRUS DETECTION IN DAIRY COWS
WITH VIDEO PROCESSING TECHNIQUES BASED ON
CONVOLUTIONAL NEURAL NETWORKS

วัชร นิลเพชร¹ พยุง มีสัง² นพดล อ่ำดี¹ และอดิศักดิ์ แสงส่องฟ้า^{1*}

Watchara Ninphet¹, Phayung Meesad², Noppadol Amdee¹ and Adisak Sangsongfa^{1*}

บทคัดย่อ

การตรวจจับอาการเป็นสัด (estrus detection) ของโคนมเป็นปัจจัยสำคัญในการเพิ่มประสิทธิภาพการผสมพันธุ์ และการผลิตน้ำนม อย่างยั่งยืนในอุตสาหกรรมเลี้ยงโคนม การวิจัยนี้จึงมีวัตถุประสงค์เพื่อพัฒนาขั้นตอนวิธีและประเมินประสิทธิภาพของแบบจำลองการตรวจจับอาการเป็นสัดตามเวลาจริงในโคนมโดยไม่ติดตั้งอุปกรณ์บนตัวโคด้วยเทคนิคการประมวลผลวิดีโอทัศนบนพื้นฐานโครงข่ายประสาทเทียมแบบสังวัตนาการ โดยตรวจจับพฤติกรรมการเป็นสัดของโคนมที่มีระดับสายเลือดโฮลส์ไตน์ฟรီးเขียน 75% ขึ้นไป จำนวน 10 ตัวขึ้นไป 4 ท่าทาง ได้แก่ เดิน (walking) ดมหรือเลีย (flirting) ปีนทับ (mating) และปีนด้านศีรษะซึ่งเป็นการปีนเล่นหรือปีนผิวด้าน (climbing) โดยแต่ละท่าทางมีจำนวน 380 ภาพรวมทั้งหมด 1,520 ภาพ ผู้วิจัยทำการเพิ่มภาพ (image augmentation) ด้วยเทคนิคต่าง ๆ เป็นจำนวน 22,238 เท่าของภาพต้นฉบับ จนได้ชุดข้อมูลภาพรวมทั้งสิ้น 33,802 ภาพ ทำการสร้างแบบจำลองการตรวจจับ

¹คณะเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยราชภัฏหมู่บ้านจอมบึง อำเภอจอมบึง จังหวัดราชบุรี 70150

¹Faculty of Industrial Technology, Muban Chom Bueng Rajabhat University, Chom Bueng, Ratchaburi 70150

²คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ เขตบางซื่อ กรุงเทพฯ 10800

²Faculty of Information Technology and Digital Innovation, King Mongkut's University of Technology North Bangkok, Bang Sue District, Bangkok 10800

*corresponding author e-mail: adisaksan@mcru.ac.th

Received: 15 August 2024; Revised: 19 September 2024; Accepted: 17 October 2024

DOI: <https://doi.org/10.14456/lsej.2024.28>

ด้วยโครงข่ายประสาทเทียมแบบสังวัตนาการ ชนิด YOLOv7 จำนวนตัวอย่างในการเรียนรู้ที่ใช้ในแต่ละรอบการปรับปรุงแบบจำลอง (batch size) มีค่า 16 ข้อมูลเรียนรู้ผ่านเข้ามาในโครงข่าย (epochs) 30 รอบ ภาพถูกปรับขนาดให้มีความกว้างและความสูง (img-size) อยู่ที่ 640x640 พิกเซล พบว่า YOLOv7 สามารถตรวจจับอาการเป็นสัดได้อย่างมีประสิทธิภาพ โดยมีค่า F1-Score เฉลี่ย = 1.00 ค่า Precision = 1.00, ค่า Recall = 1.00, mAP.5 = 0.996 และค่า mAP.5:.95 = 0.845 แสดงให้เห็นว่าแบบจำลองนี้เหมาะสมที่จะนำไปใช้ในฟาร์มเกษตรกรผู้เลี้ยงโคนมทดแทนการตรวจจับอาการเป็นสัดของโคนมโดยใช้นักงานเลี้ยงสัตว์

คำสำคัญ: อาการเป็นสัด โคนม ประมวลผลวีดิทัศน์ โครงข่ายประสาทเทียมแบบสังวัตนาการ

Abstract

Estrus detection in dairy cows is a crucial factor in enhancing breeding efficiency and achieving sustainable milk production in the dairy farming industry. This research aims to develop an algorithm and evaluate the performance of a real-time estrus detection model for dairy cows without the use of wearable devices. The model utilizes video processing techniques based on evolutionary neural networks. Estrus behaviors in Holstein Friesian cows with at least 75% Holstein bloodline, with a sample size of at least 10 cows, were detected across four behaviors: walking, flirting (sniffing or licking), mating (mounting), and head mounting (playful or incorrect mounting). A total of 1,520 images, with 380 images per behavior, were used. The dataset was augmented using various image augmentation techniques, increasing the dataset size by 22,238 times, resulting in a total of 33,802 images. The estrus detection model was developed using the YOLOv7 evolutionary neural network. The batch size was set at 16, with 30 epochs, and images were resized to 640x640 pixels. The results indicated that YOLOv7 effectively detected estrus behaviors with an average F1-Score of 1.00, Precision of 1.00, Recall of 1.00, mAP@0.5 of 0.996, and mAP@0.5:.95 of 0.845. These results demonstrate that the model is suitable for practical application in dairy farming, providing an alternative to human labor for estrus detection in cows.

Keywords: estrus, dairy cows, video processing, convolutional neural networks

บทนำ

น้ำนมเป็นแหล่งโปรตีนคุณภาพสูง มีแคลเซียมและวิตามินดีช่วยเสริมสร้างกระดูก ฟัน และระบบภูมิคุ้มกัน ลดความเสี่ยงโรคกระดูกพรุน โรคเบาหวาน และโรคอ้วน การบริโภคน้ำนมสม่ำเสมอจึงจำเป็นต่อสุขภาพในทุกช่วงวัย โดยเฉพาะเด็กและผู้สูงอายุ (Pratelli et al., 2024; Marangoni et al., 2018) อุตสาหกรรมนมทั้งในประเทศและทั่วโลกเติบโตอย่างรวดเร็ว ความต้องการบริโภคน้ำนมในประเทศสำคัญทั่วโลกช่วงปี 2562-2566 เพิ่มขึ้นเฉลี่ย 1.32% ต่อปี โดยปี 2566 มีความต้องการ 195.091 ล้านตัน ผลผลิตน้ำนมดิบในประเทศผู้ผลิตหลักเพิ่มขึ้นเฉลี่ย 0.94% ต่อปี โดยปี 2566 มีผลผลิต 550.498 ล้านตัน จำนวนโคนมในประเทศผู้ผลิตหลักเพิ่มขึ้นเฉลี่ย 0.97% ต่อปี โดยปี 2566 มีโคนมรวม 141.457 ล้านตัว ในประเทศไทยความต้องการบริโภคนมเพิ่มขึ้นเฉลี่ย 0.44% ต่อปี ผลผลิตน้ำนมดิบใช้แปรรูปภายในประเทศเป็นหลัก จำนวนโคนมเพิ่มขึ้นเฉลี่ย 3.30% ต่อปี ในปี 2566 มีโคนม 774,461 ตัว แม่โครีดนม 277,215 ตัว และผลผลิตน้ำนมดิบ 1.141 ล้านตัน (Office of Agricultural Economics Research, 2023)

การเลี้ยงโคนมเพื่อผลิตน้ำนมดิบจำเป็นต้องพิจารณาพฤติกรรมกรรมการเป็นสัตว์ของโคนม ซึ่งมักเกิดในช่วงกลางคืน การสังเกตพฤติกรรมนี้ต้องใช้แรงงานคนที่มีความรู้เฉพาะทาง ทำให้มีต้นทุนสูงและอาจเกิดความผิดพลาดได้ง่าย วิธีการตรวจจำอาการเป็นสัตว์แทนการใช้แรงงานคนเฝ้าสังเกตโดยไม่ติดตั้งอุปกรณ์บนตัวโค ได้แก่ การวัดผลผลิตน้ำนมที่ลดลง (Senger, 1994) การตรวจวัดค่าฮอร์โมนโปรเจสเตอโรนที่ลดลงในช่วงเป็นสัตว์ (Royal et al., 2000) การวัดค่าการนำไฟฟ้าของน้ำนม (Aydin et al., 2008) และการตรวจจับพฤติกรรมกรรมการเคี้ยวเอื้องที่ลดลง (Reith & Hoy, 2012) เทคโนโลยีเหล่านี้ช่วยเพิ่มความแม่นยำในการจัดการฟาร์มโคนม แต่ควรพิจารณาความสะดวกในการใช้งาน ความแม่นยำ และความคุ้มค่า (Senger, 1994) นอกจากนี้ยังมีการพัฒนาเทคโนโลยีและเครื่องมือที่สามารถติดตั้งบนตัวโคเพื่อตรวจจำอาการเป็นสัตว์ โดยอุปกรณ์เหล่านี้ได้รับการออกแบบมาเพื่อบันทึกพฤติกรรมและลักษณะทางกายภาพของโคนมที่เกี่ยวข้องกับอาการเป็นสัตว์อย่างแม่นยำ ได้แก่ เซนเซอร์วัดความเร่งเพื่อตรวจจับการเคลื่อนไหว (Jónsson et al., 2011) อุปกรณ์ติดตามเสียงติดที่คอของโคเพื่อนำเสียงมาวิเคราะห์หาพฤติกรรมกรรมการเป็นสัตว์ของโค (Wang et al., 2023) อุปกรณ์ติดแผ่นหลังซึ่งใช้เซนเซอร์แรงดันและการเคลื่อนไหวเพื่อตรวจจับพฤติกรรมกรรมการป็นทับ (Benaissa et al., 2023) หรือใช้แผ่นสัติดแผ่นหลัง ซึ่งเมื่อโคถูกป็นทับแผ่นสันี้จะเปลี่ยนสีไปจากเดิม (Chowdhury et al., 2016) อุปกรณ์ติดหางเพื่อตรวจจับการยกหาง (Higaki et al., 2021) และเครื่องมือผูกคอสำหรับการตรวจจับการเคี้ยวเอื้องและการเคลื่อนไหวของศีรษะ (Roelofs et al., 2005; MacKay et al., 2012) การทาสีหางโค (Palmer et al., 2010) อุปกรณ์เหล่านี้ช่วยเพิ่มความแม่นยำในการตรวจจำอาการเป็นสัตว์เมื่อเทียบกับการสังเกตแบบดั้งเดิม อย่างไรก็ตามต้นทุนการติดตั้งและบำรุงรักษายังค่อนข้างสูง (Homer et al., 2013; Stangaferro et al., 2016; Royal et al., 2000)

การใช้เทคโนโลยีคอมพิวเตอร์วิทัศน์ (computer vision) เพื่อตรวจจับพฤติกรรมกรรมการเป็นสัตว์ของโคนมได้รับการพัฒนาขึ้นอย่างมาก ได้แก่ การใช้โครงข่ายประสาทเทียมสังวัตนาการ (Convolutional

Neural Network; CNN) ที่มีการปรับพารามิเตอร์โดยขั้นตอนวิธีระบบภูมิคุ้มกันเทียม (Artificial Immune System Algorithm; AIS) ของ Nipheth et al. (2024) มีความถูกต้อง 0.984 แบบจำลอง Cow-YOLO ของ Li et al. (2024) ที่พัฒนาขึ้นจาก YOLOv5s ให้ค่าความแม่นยำ 0.997 และค่าดีกัล 0.995 แบบจำลองของ Arkan et al. (2023) ที่ใช้ CNN ให้ค่าความแม่นยำ 0.997 และค่าดีกัล 0.998 และยังใช้ YOLOv5 ซึ่งให้ค่าความแม่นยำ 0.998 และค่าดีกัล 0.998 ส่วน Higaki et al., (2021) ใช้เทคนิคแยกวัตถุที่เคลื่อนไหวยออกจากฉากหลังที่หนึ่งให้ค่าความแม่นยำ 0.54 การพัฒนาเหล่านี้ชี้ให้เห็นถึงศักยภาพของการประมวลผลภาพ แต่ยังมีข้อจำกัด เช่น จำนวนข้อมูลและการตั้งค่าพารามิเตอร์ที่ซับซ้อน ทั้งนี้ควรมีการพัฒนาวิธีการวิเคราะห์และเพิ่มจำนวนข้อมูลเพื่อเพิ่มความแม่นยำและการใช้งานในสภาพแวดล้อมจริง

You Only Look Once (YOLO) เป็นขั้นตอนวิธีการที่มีจุดเด่นคือประมวลผลได้ในครั้งเดียว (Single stage detector) และได้ผลลัพธ์ทันที มีความเร็วและความแม่นยำสูง ทำให้ได้รับความนิยมสำหรับการตรวจจับวัตถุโดยเฉพาะตามเวลาจริง พัฒนาขึ้นโดยประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึก (Deep Learning) จากความพิเศษของโครงข่ายประสาทเทียมสังวัตนาการ ทำให้สามารถแบ่งภาพออกเป็นส่วน ๆ พร้อมทำนายขอบเขตและความน่าจะเป็นของแต่ละคำตอบ (Morera et al., 2020) YOLOv7 (Wang et al., 2023) ซึ่งเปิดตัวในปี 2022 ได้รับการพัฒนาให้เหมาะกับการประมวลผลบน GPU สามารถลดพารามิเตอร์ได้ประมาณร้อยละ 40 ทำให้ประมวลผลได้รวดเร็ว ขนาดเล็ก และแม่นยำสูง จึงเป็นหนึ่งในสถาปัตยกรรมที่ได้รับความนิยมมากที่สุด

เพื่อให้การตรวจจับอาการเป็นสัตว์ของโคนมมีความแม่นยำและคุ้มค่า การนำเทคโนโลยีการเรียนรู้เชิงลึก โดยใช้โครงข่ายประสาทเทียมสังวัตนาการมาประยุกต์ใช้กับการประมวลผลภาพตามเวลาจริงสามารถเป็นทางออกที่มีประสิทธิภาพ งานวิจัยนี้จะมุ่งเน้นการใช้เทคโนโลยีดังกล่าวเพื่อทดแทนการใช้แรงงานคน ลดต้นทุน และเพิ่มผลผลิต ตอบสนองต่อความต้องการของอุตสาหกรรมเกษตรสมัยใหม่ ที่เน้นการใช้เทคโนโลยีเพื่อพัฒนาประสิทธิภาพการผลิตในอนาคต

วิธีดำเนินการวิจัย

วิธีดำเนินการวิจัยประกอบด้วยขั้นตอนสำคัญเริ่มตั้งแต่การเก็บรวบรวมและเตรียมข้อมูล การออกแบบและเรียนรู้แบบจำลอง ไปจนถึงการประเมินผลและการนำแบบจำลองไปใช้งานจริง ดังแสดงดังภาพที่ 1 (Figure 1)

1. การเก็บรวบรวมและเตรียมข้อมูล

1.1 เก็บข้อมูลวิดีโอทัศนพฤติกรรมการเป็นสัตว์ของโค โดยติดตั้งกล้องบันทึกวิดีโอที่โคนมที่มีระดับสายเลือดโฮลสไตน์ฟรีเซียนร้อยละ 75 ขึ้นไป ในลำดับท้องที่ 2 ขึ้นไป (โคนาง; โคที่เคยให้ลูกมาแล้ว) ในคอกพักโคที่มีการเลี้ยงดูปกติประจำวันตามวิธีการจัดการของเกษตรกร จำนวน 10 ตัว ขึ้นไป โดยมีพื้นที่อย่างน้อย 4 ตารางเมตร/ตัว จากฟาร์มโคนมมาตรฐานที่ได้รับการรับรองจากกรมปศุ

สัตว์ ติดตั้งกล้องวีดิทัศน์ IP camera จำนวน 4 ตัว ในมุมมองด้านหน้า ด้านข้าง (ความสูงระดับสายตามนุษย์ สูงจากพื้น 160 เซนติเมตร) ด้านบน (ติดตั้งบนหลังคาสูงจากพื้น 4 เมตร มุมมอง 90 องศากับหลังคา) และด้านบนมุมทแยงจากด้านหน้า (ติดตั้งบนหลังคาสูงจากพื้น 4 เมตร มุมมอง 45 องศากับหลังคา) ดังภาพที่ 2 (Figure 2) บันทึกวีดิทัศน์ที่ความละเอียด 1280 x 960 พิกเซล 25 เฟรมต่อวินาที เก็บข้อมูลเป็นระยะเวลา 3 เดือน จะได้วีดิทัศน์สำหรับตรวจจับอาการเป็นสัดของโคนมจากพฤติกรรมการปิ่นทับ

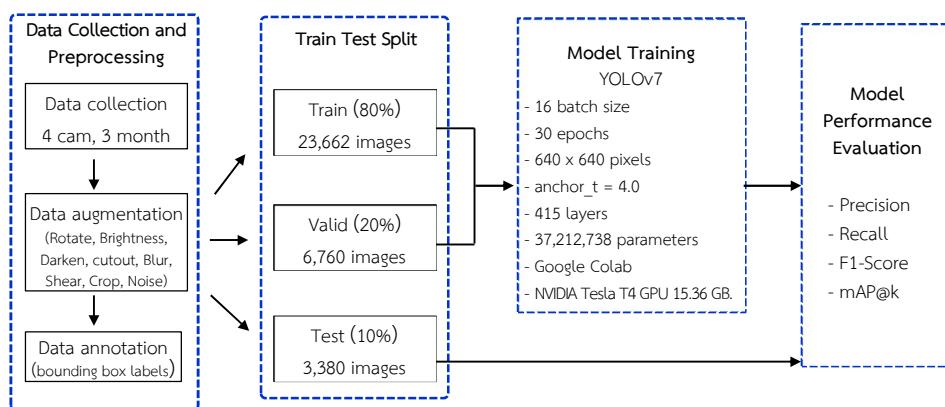


Figure 1 Framework of the proposed method.

1.2 จากนั้นนำวีดิทัศน์ที่ได้มาตัดส่วนเกินออก (trim) ให้เหลือเฉพาะส่วนที่สนใจ (ROI: Region of Interest) แล้วจึงทำการแยกเฟรมภาพแต่ละภาพออกจากวีดิทัศน์แล้วบันทึกเป็นไฟล์ภาพเดี่ยว ๆ ความละเอียด 300x300 dpi เพื่อนำติดฉลาก สร้างแบบจำลองการเรียนรู้ และประเมินประสิทธิภาพของแบบจำลองในขั้นตอนต่อไป

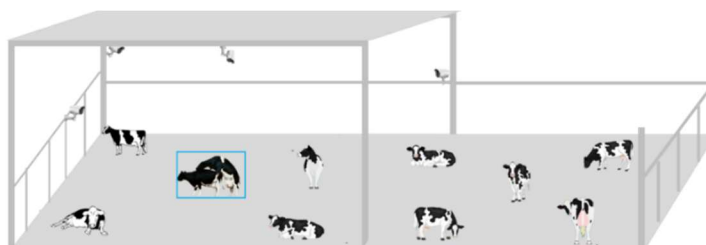


Figure 2 The installation site for video cameras to capture dairy cow estrus images.

1.3 ทำการติดฉลาก (label) เพื่อสร้างกรอบสี่เหลี่ยมล้อมรอบวัตถุ (bounding boxes) โดย bounding box แต่ละกล่องประกอบด้วยพารามิเตอร์ 5 ค่า ได้แก่ (x, y, w, h, confidence) ซึ่ง x และ y เป็นพิกัดจุดศูนย์กลางของ bounding box ส่วน w และ h เป็นความกว้างและความสูงของกล่อง ดังภาพที่ 3 (Figure 3)

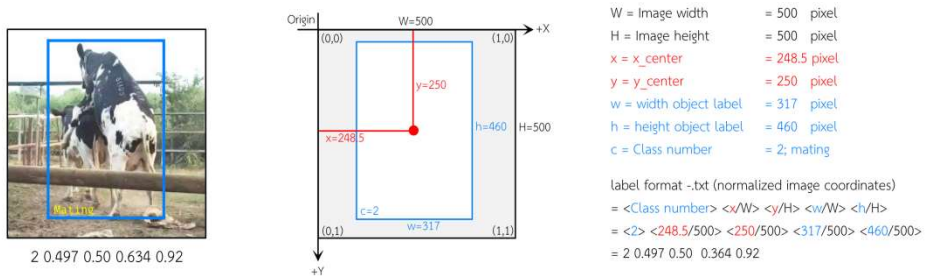


Figure 3 The general format of labels for creating bounding boxes in YOLO.

Table 1 Image data augmentation.

No.	Augmentation Techniques	Setting
1.	Rotate	Between -10 and +10 degree
2.	Brightness	Between (-20)% to 20%
3.	Darken	Between (-20)% to 20%
4.	1 count cutout	30%
5.	2 count cutout	20%
6.	Blur	1.5 px
7.	Horizontal Shear	Between -15 and +15 degree
8.	Vertical Shear	Between -15 and +15 degree
9.	Crop	Between 0% to 10%
10.	Noise	1.3% of pixel

1.4 แบ่งภาพพฤติกรรมการเป็นสัตว์ของโคตามลำดับเหตุการณ์จริง เริ่มตั้งแต่โคในฝูงจะเดินไปหาโคตัวที่เป็นสัตว์ จากนั้นจะดมหรือเลียโคตัวที่เป็นสัตว์ก่อน แล้วจึงขึ้นป็นทับจากด้านหลังตรงส่วนสะโพก โคตัวที่ยืนนิ่งเป็นเวลา 2-3 วินาที ให้โคตัวอื่นขึ้นทับคือโคที่มีอาการเป็นสัตว์ (standing heat) ในระยะเป็นสัตว์ (estrus) ผู้วิจัยจึงใช้พฤติกรรมเหล่านี้มาเป็นส่วนที่สนใจ ในการตรวจจับอาการเป็นสัตว์ของโคนม ทั้งนี้เพื่อให้แบบจำลองมีประสิทธิภาพมากยิ่งขึ้นสามารถตรวจจับการป็นทับที่แสดงถึงอาการเป็นสัตว์อย่างแท้จริง ไม่ได้ป็นเพื่อวัตถุประสงค์อื่น ผู้วิจัยจึงเพิ่มพฤติกรรมการป็นเล่นหรือป็นผิตด้าน เช่นป็นจากด้านศีรษะของโคมาอีก 1 ท่าทาง จึงแบ่งพฤติกรรมที่สนใจตรวจจับออกเป็น 4 ท่าทาง ได้แก่ เดิน (walking) โคจะก้าวขาไปข้างหน้าอย่างสม่ำเสมอ ดมหรือเลีย (flirting) โคจะยื่นศีรษะเข้าไปใกล้โคตัวอื่นแล้วใช้จมูกดมกลิ่นและอาจใช้ลิ้นเลียบริเวณที่ดม ป็นทับ (mating) โคจะยกขาหน้าขึ้นวางบนหลังของโคตัวที่อยู่ด้านหน้า ขาหลังจะยืนอยู่กับพื้นอย่างมั่นคง และป็นด้านศีรษะซึ่งเป็นการป็นเล่นหรือป็นผิตด้าน (climbing) โดยแต่ละท่าทางมีจำนวน 380 ภาพ รวมทั้งหมด 1,520 ภาพ ยิ่งไปกว่านี้ เพื่อเพิ่มจำนวนชุดข้อมูลภาพที่มีอยู่อย่างจำกัดให้มีจำนวนมากขึ้น ซึ่งจะส่งผลให้แบบจำลองทำการตรวจจับ

อย่างมีประสิทธิภาพมากยิ่งขึ้น ผู้วิจัยจึงทำการเพิ่มภาพ (image augmentation) ด้วย 10 เทคนิค โดยเลือกใช้เฉพาะเทคนิคที่ภาพมีโอกาสเกิดได้ในสถานการณ์จริงเท่านั้น ได้แก่ rotate ภาพจะถูกหมุนราวกับโคเดินไปมาได้ทุกทิศทาง brightness และ darken ภาพจะมีดลงและสว่างขึ้นราวกับกลางวันและกลางคืน 1-2 count cutout ภาพจะถูกบดบังบางส่วนราวกับมีเสาโรงเรือน ผงกั้นคอก หรือ ศีรษะของโคตัวอื่นมาบดบัง blur ภาพที่ได้มีความเนียนนุ่ม ความคมชัดและรายละเอียดลดลง ซึ่งอาจเกิดได้จากการมีฝุ่นเกาะติดเลนส์กล้อง หรือมีหมอกควันในโรงเรือน shear การเฉือนภาพทำให้วัตถุภายในภาพมีลักษณะเอียงหรือยืดออกในทิศทางแนวนอนหรือแนวตั้ง ตามแกนที่กำหนด crop การตัดขอบภาพออกบางส่วน ราวกับมีวัตถุมาบดบังขอบภาพ และ noise จะปรากฏเป็นจุดสีขาวดำ หรือจุดสีอื่นๆ กระจายอยู่ทั่วภาพ นอยส์เหล่านี้จะรบกวนความคมชัดและรายละเอียดของภาพ ซึ่งอาจเกิดจากความผิดพลาดทางอิเล็กทรอนิกส์ในเซนเซอร์ของกล้องหรือการมีฝุ่นผงรวมไปถึงแมลงบดบังกล้อง เป็นจำนวน 22,238 เท่าของภาพต้นฉบับ ตามที่แสดงดังตารางที่ 1 (Table 1) จนได้ชุดข้อมูลภาพรวมทั้งสิ้น 33,802 ภาพ จากนั้นผู้วิจัยทำการสร้างแบบจำลองการตรวจจับ โดยแบ่งชุดข้อมูลภาพแบบสุ่มออกเป็นสามส่วน ได้แก่ ชุดข้อมูลภาพสำหรับเรียนรู้ (training set) ชุดข้อมูลภาพสำหรับตรวจสอบความถูกต้อง (validation set) และชุดข้อมูลภาพสำหรับทดสอบแบบจำลอง (testing set) ในอัตราส่วนร้อยละ 70, 20 และ 10 ตามลำดับ ดังที่แสดงดังตารางที่ 2 (Table 2)

Table 2 Overview of splitting a dairy cow estrus image dataset.

No.	Purposes	Percentage	Initial Images Number	Augmentation Images Number
1.	Training	70	1,064	23,662
2.	Validation	20	304	6,760
3.	Testing	10	152	3,380
4.	Total	100	1,520	33,802

2. การออกแบบและเรียนรู้โดยโครงข่ายประสาทเทียมแบบสังวัตนาการด้วยแบบจำลอง YOLOv7

จากนั้นเข้าสู่การออกแบบและเรียนรู้โดยโครงข่ายประสาทเทียมแบบสังวัตนาการ มีขั้นตอนการทำงานหลัก ๆ เริ่มจากการทำคอนโวลูชันโดยใช้ฟิลเตอร์หรือเคอร์เนลสแกนข้อมูลภาพเพื่อดึงคุณลักษณะสำคัญเช่น ขอบ รูปทรง หรือพื้นผิว เลื่อนผ่านภาพทีละส่วน แล้วจึงลดขนาดของข้อมูลโดยการสุบพล เช่น ใช้ Max Pooling เพื่อลดจำนวนพารามิเตอร์และการคำนวณ แต่ยังคงรักษาคุณลักษณะที่สำคัญไว้ จากนั้นเป็นการเชื่อมต่อแบบสมบурณ์ เพื่อนำคุณลักษณะที่ได้มาประมวลผลผ่านชั้นของโครงข่ายประสาทเทียมแบบทั่วไป เพื่อตัดสินใจตรวจจับรูปภาพ สุดท้ายเป็นการตรวจจับรูปภาพใช้ฟังก์ชัน Softmax ซึ่ง YOLOv7 ถูกออกแบบด้วยสถาปัตยกรรมโครงข่ายประสาทเทียมแบบสังวัตนาการที่ทันสมัยและมีประสิทธิภาพ โดยประกอบด้วยสามส่วนหลัก ได้แก่ Backbone, Neck, และ Head ดังตารางที่ 3 (Table 3) กระบวนการทำงานเริ่มต้นที่ Backbone ซึ่งใช้ Efficient Layer Aggregation Network

(ELAN) ในการสกัดฟีเจอร์สำคัญจากภาพที่ป้อนเข้า ELAN นี้พัฒนาจากสถาปัตยกรรม CSPDarknet ของ YOLOv4 ทำให้การรวมฟีเจอร์จาก convolutional layers หลายชั้นเป็นไปได้อย่างหลากหลายและแม่นยำ หลังจาก Backbone ทำการสกัดฟีเจอร์แล้ว ข้อมูลจะถูกส่งต่อไปยัง Neck ซึ่งทำหน้าที่เชื่อมต่อฟีเจอร์เหล่านั้นเข้ากับ Head ของแบบจำลอง Neck ใน YOLOv7 ใช้โครงสร้าง Feature Pyramid Networks (FPN) หรือ Path Aggregation Network (PANet) เพื่อเพิ่มประสิทธิภาพในการตรวจจับวัตถุทั้งขนาดเล็กและขนาดใหญ่ที่อยู่ในระดับต่าง ๆ ของภาพ YOLOv7 ยังได้ปรับปรุง PANet เพื่อเพิ่มประสิทธิภาพในการรวมฟีเจอร์จากหลายระดับของ Backbone

Table 3 The YOLOv7 Architecture.

	Type	Filter	Size	Output
Input	Layers	128	52 × 52	128 × 128
Backbone; ELAN	Conv2D 32 × 3 × 3	32	416 × 416	128 × 128
	Residual Block 1 × 64	64	208 × 208	-
	Residual Block 2 × 128	128	104 × 104	64 × 64
	Residual Block 8 × 256	256	52 × 52	-
	Residual Block 8 × 512	512	26 × 26	64 × 64
	Residual Block 4 × 1024	1024	13 × 13	32 × 32
Neck; FPN	Conv2D 5 L	384	52 × 52	16 × 16
	Concat	384	52 × 52	-
	Conv2D + Up2D	128	52 × 52	8 × 8
	Conv2D 5 L	1024	13 × 13	-
Head; Loss Function	Conv2D 3 × 3 + Conv2D 1 × 1	255	52 × 52	8 × 8
	Conv2D 3 × 3 + Conv2D 1 × 1	255	26 × 26	-
	Conv2D 3 × 3 + Conv2D 1 × 1	255	13 × 13	-
	Avg Pool	-	Global	-
	Connected	-	1000	-
	Softmax	-	-	-

เมื่อข้อมูลถูกส่งมาถึง Head แบบจำลองจะเริ่มทำการทำนาย bounding boxes, objectness scores และ class probabilities ของวัตถุในภาพ การทำนาย bounding box สำหรับแต่ละเซลล์ YOLO จะทำนาย bounding box หลายกล่อง โดยใช้กลไก anchor-based detection ซึ่งทำนาย anchor boxes หลายขนาดและรูปร่างเพื่อให้ครอบคลุมวัตถุในภาพได้อย่างเหมาะสม โครงสร้างนี้ถูกปรับปรุงให้การทำนายทำได้เร็วและแม่นยำยิ่งขึ้น การตรวจจับวัตถุแบบใช้แอ่งเคอร์ ระบบจะสร้างกรอบแอ่งเคอร์ขึ้นมาจำนวนมากหลาย ๆ ขนาดและอัตราส่วน ในแต่ละตำแหน่งบนภาพเพื่อให้ครอบคลุมวัตถุที่มีขนาดและรูปร่างต่างกัน และพยายามทำนายว่ามีวัตถุอยู่ในแอ่งเคอร์เหล่านั้นหรือไม่ การตรวจจับวัตถุแบบใช้แอ่งเคอร์มีขั้นตอน เริ่มจากการสร้างแอ่งเคอร์ ระบบจะสร้างกรอบแอ่งเคอร์ขนาดต่าง ๆ ในทุกจุดของฟีเจอร์แมพที่ได้จากการผ่านโครงข่ายคอนโวลูชัน โดยแอ่งเคอร์เหล่านี้มีขนาดและอัตราส่วนต่าง ๆ เพื่อจับวัตถุในทุกขนาดและรูปร่างในภาพ จากนั้นจึงจับคู่กับกราวด์ทรู ระบบจะจับคู่แอ่งเคอร์กับกรอบ

วัตถุจริง (ground truth boxes) โดยใช้ Intersection over Union (IoU) ซึ่งเป็นค่าที่บ่งบอกถึงการทับซ้อนระหว่างแอ่งเคอร์กับกรอบวัตถุจริง แอ่งเคอร์ที่มี IoU สูงสุดจะถูกพิจารณาว่าเป็นแอ่งเคอร์ที่ดีที่สุดสำหรับการตรวจจับวัตถุนั้น ค่า IoU มีค่าตั้งแต่ 0 ถึง 1 ซึ่งค่าใกล้เคียง 1 หมายถึง bounding box ที่ทำนายมีความแม่นยำสูง ดังแสดงดังภาพที่ 4 (Figure 4) และสมการ (1)

$$IOU = \frac{\text{Area of the intersection between } B_1 \text{ and } B_2}{\text{Area of Union between } B_1 \text{ and } B_2} \quad (1)$$

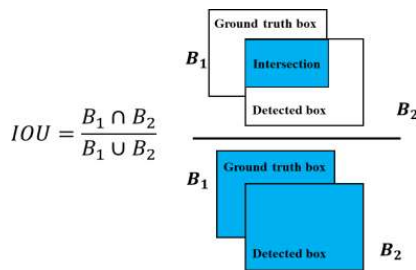


Figure 4 Intersection over Union; IoU.

จากนั้นระบบจะทำนายว่ามีวัตถุอยู่หรือไม่ และทำนายชนิดของวัตถุที่พบ พร้อมทั้งให้ค่า confidence score ซึ่งบ่งบอกถึงความมั่นใจในผลการทำนาย ยิ่งค่าความมั่นใจสูงก็ยิ่งหมายความว่าผลการทำนายมีความแม่นยำสูง ระบบจะเรียนรู้ค่า offset เพื่อปรับตำแหน่งและขนาดของแอ่งเคอร์ที่จับคู่กับกรอบวัตถุจริงให้เหมาะสมที่สุด เพื่อให้แอ่งเคอร์ครอบคลุมวัตถุได้อย่างแม่นยำ เมื่อได้ผลการทำนายทั้งหมดแล้ว ระบบจะใช้เทคนิค Non-Maximum Suppression (NMS) ดังแสดงดังภาพที่ 5 (Figure 5) เพื่อลดการซ้ำซ้อนของแอ่งเคอร์ที่ทับกันโดยรักษาเฉพาะแอ่งเคอร์ที่มี confidence score สูงสุดในแต่ละกลุ่มที่ทับซ้อนกันไว้ วิธีนี้ช่วยลดการทำนายที่ไม่จำเป็นและทำให้ผลลัพธ์มีความชัดเจนมากขึ้น

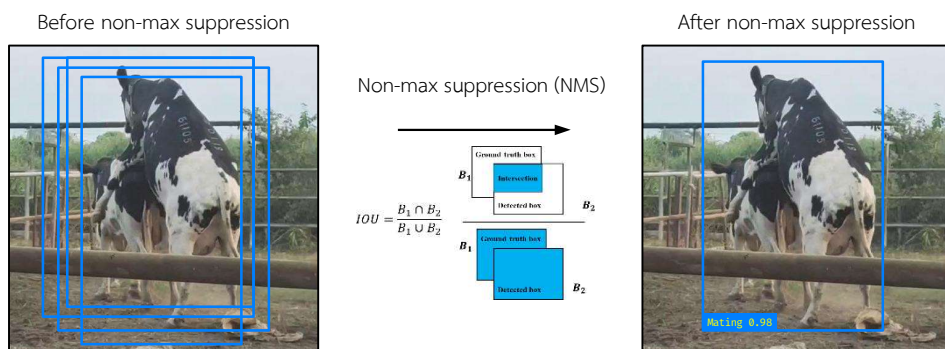


Figure 5 Non-max suppression (NMS).

Table 4 YOLOv7 Hyperparameter Configuration for Training.

No.	Hyperparameters	Description
1	batch size 16	The batch size used per training iteration is 16.
2	img-size 640	Images are resized to 640x640 pixels.
3	Epochs = 30	Training data passes through the network for 30 epochs.
4	lr0 = 0.01	The initial learning rate is set to 0.01, ensuring a moderate training speed.
5	lrf = 0.1	The final learning rate is reduced to 10% of the initial rate (0.1) at the end of training.
6	momentum = 0.937	A momentum value of 0.937 maintains 93.7% of the previous weight update direction for stability.
7	weight_decay = 0.0005	The weight decay rate of 0.0005 prevents overfitting while preserving model flexibility.
8	warmup_epochs = 3.0	The learning rate is warmed up over the first three epochs.
9	warmup_momentum = 0.8	A momentum of 0.8 is applied during the warmup phase for smoother initial weight changes.
10	warmup_bias_lr = 0.1	The learning rate for bias during warmup is 0.1.
11	box = 0.05	The weight for box loss is set to 0.05, prioritizing accurate object localization.
12	cls = 0.3	Class loss is weighted at 0.3, emphasizing correct classification.
13	cls_pw = 1.0	The weight for class loss adjustment is 1.0, indicating no further adjustments.
14	obj = 0.7	Objectness loss, determining the presence of an object within a bounding box, has a weight of 1.0.
15	obj_pw = 1.0	The weight for objectness loss adjustment is 1.0.
16	iou_t = 0.2	The IoU threshold is 0.2, considering predictions with over 20% overlap as correct.
17	anchor_t = 4.0	The anchor-matching threshold is set at 4.0, selecting anchors up to four times the ground truth size.
18	fl_gamma = 0.0	The Focal Loss parameter is 0.0, indicating it is not used.
19	hsv_h = 0.015, hsv_s = 0.7, hsv_v = 0.4	Parameters for hue, saturation, and value are adjusted during data augmentation by ± 0.015 , $\pm 70\%$, and $\pm 40\%$, respectively.
20	degrees = 0.0, translate = 0.2, scale = 0.9, shear = 0.0, perspective = 0.0	Parameters for data augmentation include rotation (degrees), translation, scaling, shearing, and perspective transformation.
21	flipud = 0.0, fliplr = 0.5	Vertical flip probability is 0.0, while horizontal flip probability is 0.5.
22	mosaic = 1.0	Mosaic augmentation is applied to 100% of the training data.
23	mixup = 0.15	Mixup augmentation is applied to 15% of the training data.
24	loss_ota = 1	The OTA (Optimal Transport Assignment) Loss function is enabled.
Model Summary: 415 layers, 37212738 parameters, 37212738 gradients, 105.2 GFLOPS		

ระหว่างการฝึกแบบจำลอง กระบวนการทั้งหมดจะเกิดขึ้นซ้ำหลายรอบ โดยแบบจำลองเรียนรู้จากข้อมูลที่ป้อนเข้ามา ผู้วิจัยใช้ไฟล์ yolov7_training.pt ขนาด 72.1 MB ซึ่งเป็น pre-trained weights จาก COCO dataset และทำการ fine-tuning ด้วยชุดข้อมูลพฤติกรรมการณ์การเป็นสัตว์ของโคนมที่สร้างขึ้น

เพื่อให้แบบจำลองปรับตัวกับข้อมูลใหม่ได้อย่างมีประสิทธิภาพ ดังตารางที่ 4 (Table 4) การฝึกถูกตั้งค่าให้มี 30 รอบการเรียนรู้ (epochs) โดยในแต่ละรอบแบบจำลองจะทดสอบและปรับปรุงผลการทำนายด้วย validation set เพื่อลดโอกาสเกิด overfitting และเพิ่มความแม่นยำ

การตั้งค่าไฮเปอร์พารามิเตอร์ เช่น ขนาดชุดข้อมูล (batch size) 16 และขนาดภาพ 640x640 พิกเซล ดำเนินการบนแพลตฟอร์ม Google Colab โดยใช้ GPU รุ่น NVIDIA Tesla T4 ซึ่งมีหน่วยความจำ 15.36 GB และรองรับการประมวลผล FP32, FP16, INT8, และ INT4 เหมาะสำหรับการประมวลผลเชิงลึก

หลังจากฝึกเสร็จ แบบจำลองจะถูกนำไปใช้ในกระบวนการทำนาย (inference) โดยใช้โครงสร้างเดียวกับระหว่างการฝึก แบบจำลองจะรับข้อมูลใหม่ ประมวลผลผ่าน Backbone, Neck, และ Head เพื่อทำนาย bounding boxes, objectness scores, และ class probabilities จากนั้นจะแสดงผล bounding boxes พร้อม class labels และ confidence scores ตามเวลาจริง

3. การประเมินประสิทธิภาพของแบบจำลอง

เครื่องมือที่ใช้ในการประเมินประสิทธิภาพของแบบจำลอง มีดังนี้

3.1 Precision หรือ ความแม่นยำ หมายถึงความสามารถของแบบจำลองในการทำนายผลลัพธ์ที่เป็นกลุ่มที่สนใจ (Positive Class) ได้อย่างถูกต้อง เมื่อแบบจำลองทำนายผลว่าเป็นกลุ่มที่สนใจ สูตรการคำนวณมีดังนี้ (สมการ 2)

$$Precision = \frac{\text{Number of True Positive}}{\text{Number the model predicts to be Positive predictions}} \quad (2)$$

3.2 Recall หรือ การดึงกลับ หมายถึงความสามารถของแบบจำลองในการระบุคำตอบที่เป็นกลุ่มที่สนใจ (Positive Class) ได้ครบถ้วน โดยวัดจากจำนวนคำตอบที่ถูกต้องทั้งหมดในกลุ่มที่สนใจ สูตรการคำนวณมีดังนี้ (สมการ 3)

$$Recall = \frac{\text{Number of True Positive}}{\text{Number of real Positive}} \quad (3)$$

3.3 F1-Score หรือคะแนน F1 หมายถึงการวัดค่าเฉลี่ยที่เน้นความสมดุลระหว่างความแม่นยำ (Precision) และการดึงกลับ (Recall) ซึ่งจะให้ผลลัพธ์ที่ดีเมื่อทั้งความแม่นยำและการดึงกลับมีค่าสูงและใกล้เคียงกัน การคำนวณคะแนน F1 นี้ จะให้ค่าอยู่ในช่วง 0-1 โดยค่า 1 หมายถึงแบบจำลองมีความแม่นยำและการดึงกลับที่ดีเยี่ยม สูตรการคำนวณมีดังนี้ (สมการ 4)

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

3.4 Mean Average Precision at K (mAP@k) หรือค่าเฉลี่ยความแม่นยำเฉลี่ยที่อันดับ K คือการคำนวณค่า Average Precision (AP) ของแต่ละชุดข้อมูล โดยตัดค่าเทรชโฮลด์ (Threshold) ที่อันดับใด ๆ ที่เราสนใจ (Rank K) แล้วคำนวณค่าเฉลี่ยของ AP@k สำหรับทุกชุดข้อมูลที่ได้มา เพื่อให้ได้ค่าของ mAP@k สูตรการคำนวณมีดังนี้ (สมการ 5)

$$mAP@k = \frac{1}{N} \sum_{i=1}^N (AP@_i) \quad (5)$$

โดยที่

- $AP@k$ หมายถึงค่า Average Precision ที่อันดับ K สำหรับชุดข้อมูลที่ i
- N หมายถึงจำนวนชุดข้อมูลทั้งหมด
- $\sum_{i=1}^N (AP@_i)$ คือการรวมค่าของ $AP@k$ จากชุดข้อมูลที่ i ถึงชุดข้อมูลที่ N แล้วหารด้วยจำนวนชุดข้อมูลทั้งหมด เพื่อหาค่าเฉลี่ยของความแม่นยำที่อันดับ K

$mAP@.5$ หมายถึง ค่า mean Average Precision เมื่อ IoU (Intersection over Union) เท่ากับ 0.5 เป็นตัวชี้วัดที่นิยมใช้ในการวัดประสิทธิภาพของแบบจำลอง ส่วน $mAP@.5:.95$ หมายถึงค่า mean Average Precision โดยเฉลี่ยในช่วง IoU ตั้งแต่ 0.5 ถึง 0.95 เป็นตัวชี้วัดที่มีความเข้มงวดมากขึ้น

ผลการวิจัย

การพัฒนาขั้นตอนวิธีและประเมินประสิทธิภาพของแบบจำลองการตรวจจับอาการเป็นสัดตามเวลาจริงในโคนมโดยไม่ติดตั้งอุปกรณ์บนตัวโคด้วยเทคนิคการประมวลผลวีดิทัศน์บนพื้นฐานโครงข่ายประสาทเทียมแบบสังวัตนาการ ใช้เวลาประมวลผลทั้งหมดสำหรับ 30 รอบการฝึกอยู่ที่ 10.924 ชั่วโมง ได้ผลลัพธ์แบบจำลองที่ดีที่สุด (best model) ขนาด 74.8MB ได้ผลลัพธ์ที่เหมาะสมและมีประสิทธิภาพ มีค่า F1-Score เฉลี่ย = 1.00 ค่า Precision = 1.00, ค่า Recall = 1.00, $mAP.5 = 0.996$ และค่า $mAP.5:.95 = 0.845$ ผู้วิจัยได้นำเสนอผลการประมวลผลทั้งหมดตามลำดับ ดังตารางที่ 5 (Table 5) และภาพที่ 6-10 (Figure 6-10)

Table 5 Traing Result.

Model	Time (hr)	Epoch	Precision	Recall	F1-score	mAP.5	mAP.5:.95
YOLOv7	10.924	30	1.00	1.00	1.00	0.996	0.845

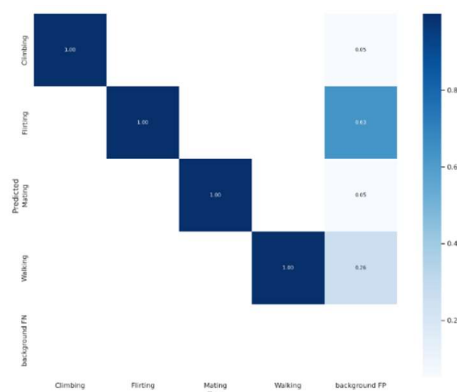


Figure 6 YOLOv7 confusion matrix and accuracy for cow estrus detection.

จากภาพที่ 6 (Figure 6) Confusion Matrix ผลการทำนายของแบบจำลองสำหรับแต่ละประเภทพฤติกรรม (Climbing, Flirting, Mating, Walking) โดยภาพรวมมีความแม่นยำสูงสำหรับแต่ละประเภทพฤติกรรม (Precision = 1.00) แต่มี False Positives (FP) และ False Negatives (FN) เล็กน้อยในบางประเภท ตัวอย่างเช่น ในการทำนายพฤติกรรม Flirting มี FP อยู่ที่ 0.63 ซึ่งหมายความว่ามีการตรวจจับที่ผิดพลาดบ้างในภาพพื้นหลัง (background) แสดงให้เห็นว่าแบบจำลองมีประสิทธิภาพที่ดีในการตรวจจับ แต่ควรระวังการตรวจจับที่ผิดพลาดในบางกรณี

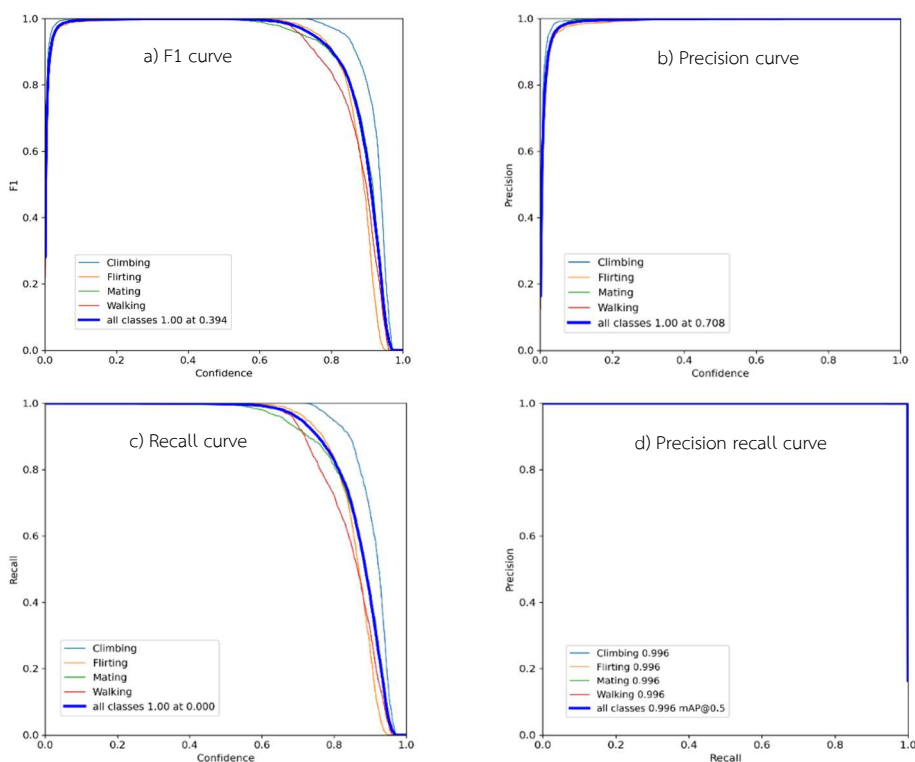


Figure 7 Model performance metrics: a) F1 curve, b) precision curve, c) recall curve, d) precision-recall curve for estrus detection in cows across 4 YOLOv7 classes.

จากภาพที่ 7a (Figure 7a) กราฟ F1 Curve แสดงความสัมพันธ์ระหว่างค่า F1-score กับค่า Confidence ในการทำนายของแบบจำลองสำหรับแต่ละคลาส สังเกตได้ว่า F1-score ของทุกคลาสอยู่ในระดับสูงมาก (ใกล้ 1) ที่ช่วงค่า confidence ระหว่าง 0.2 ถึง 0.8 ซึ่งแสดงให้เห็นว่าแบบจำลองมีความแม่นยำสูงและสามารถตรวจจับได้อย่างดีในหลายระดับความมั่นใจ อย่างไรก็ตาม ที่ค่า confidence สูงขึ้น (ใกล้ 1) ค่า F1-score มีแนวโน้มลดลงเล็กน้อย เนื่องจากแบบจำลองอาจทำนายผิดพลาดในบางกรณี จากภาพที่ 7b (Figure 7b) กราฟ P Curve แสดงความสัมพันธ์ระหว่างค่า Precision และค่า Confidence สำหรับแต่ละคลาสของแบบจำลอง จากกราฟนี้เราสามารถเห็นได้ว่า ค่า Precision

สำหรับทุกคลาสอยู่ในระดับสูงมาก (ใกล้ 1) เมื่อค่า Confidence อยู่ในช่วงตั้งแต่ประมาณ 0.2 ขึ้นไป ซึ่งแสดงให้เห็นว่าแบบจำลองมีความแม่นยำในการตรวจจับสูง โดยเฉพาะเมื่อค่า Confidence สูงกว่า 0.7 ค่า Precision จะอยู่ที่ 1.00 ในทุกคลาส จากภาพที่ 7c (Figure 7c) จากกราฟ R Curve แสดงความสัมพันธ์ระหว่างค่า Recall และ Confidence สำหรับแต่ละคลาส สามารถสังเกตได้ว่า ค่า Recall อยู่ในระดับสูงมาก (ใกล้ 1) ในช่วงค่า Confidence ที่ต่ำ (ประมาณ 0.0 ถึง 0.6) ซึ่งหมายความว่าแบบจำลองสามารถตรวจจับวัตถุได้เกือบทั้งหมดเมื่อค่า Confidence อยู่ในช่วงนี้ อย่างไรก็ตาม เมื่อค่า Confidence เพิ่มขึ้น ค่า Recall จะลดลงเล็กน้อย แสดงให้เห็นว่ามีการพลาดการตรวจจับในบางกรณี เมื่อค่า Confidence สูงขึ้น จากภาพที่ 7d (Figure 7d) จากกราฟ PR Curve แสดงความสัมพันธ์ระหว่างค่า Precision และ Recall สำหรับแต่ละคลาส สามารถสรุปได้ว่าค่า Precision และ Recall ของทุกคลาสอยู่ในระดับสูงมาก (ใกล้ 1) ซึ่งหมายความว่าแบบจำลองของคุณมีความแม่นยำและความสามารถในการตรวจจับวัตถุได้ดีมาก นอกจากนี้ ค่า mAP.5 อยู่ที่ 0.996 ซึ่งบ่งบอกถึงประสิทธิภาพที่ยอดเยี่ยมของแบบจำลองในการจำแนกวัตถุในทุกคลาส

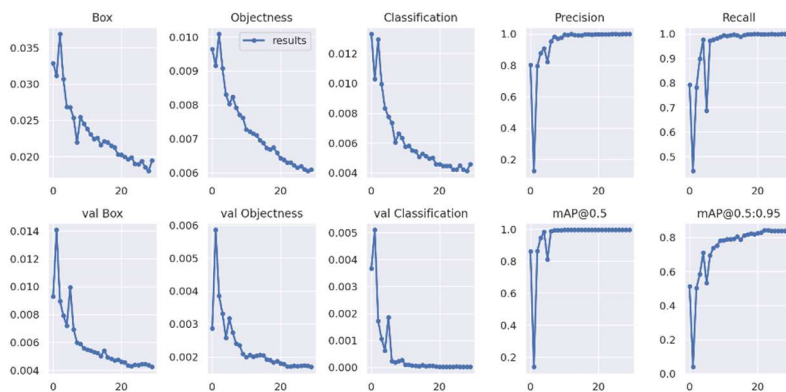


Figure 8 Performance of the improved model.

จากภาพที่ 8 (Figure 8) ค่าของ Box, Objectness, Classification Losses พบว่าค่า Loss ลดลงอย่างต่อเนื่อง ซึ่งบ่งบอกว่าแบบจำลองเรียนรู้ได้ดีขึ้นเมื่อ epoch เพิ่มขึ้น ส่วนค่าของ Validation Losses พบว่าค่า Loss สำหรับชุด validation ก็ลดลงตามไปด้วย แสดงว่าแบบจำลองทำงานได้ดีทั้งในชุดฝึกและชุดทดสอบ จะเห็นได้ว่าเส้นกราฟค่าการเรียนรู้และค่าที่ทำนายได้มีเส้นเกือบจะเป็นเส้นเดียวกัน ซึ่งบ่งชี้ว่าการเตรียมข้อมูล การกำหนดรอบการทำงาน และการกำหนดค่าพารามิเตอร์ต่าง ๆ มีความสอดคล้องกัน เส้นโค้งการสูญเสียแสดงให้เห็นว่าแบบจำลองที่ถูกฝึกมีการทำงานดีในทั้งชุดข้อมูลฝึกและชุดข้อมูลทดสอบ สามารถเรียนรู้รูปแบบพื้นฐานจากข้อมูลโดยไม่เกิดการยึดติดกับข้อมูลฝึกมากเกินไป (Good Fit) ซึ่งแสดงให้เห็นว่าการเรียนรู้และการทำนายสามารถทำได้อย่างมีประสิทธิภาพ นอกจากนี้ ค่าของ Precision, Recall, mAP.5, mAP.5:0.95 และภาพที่ 9 (Figure 9) ตัวอย่างชุดข้อมูลภาพ

สำหรับเรียนรู้ และภาพที่ 10 (Figure 10) ตัวอย่างผลการตรวจจับ เหล่านี้เพิ่มขึ้นและคงที่ในระดับสูง แสดงถึงประสิทธิภาพการตรวจจับที่ดีของแบบจำลอง นอกจากนี้แบบจำลองจะมีการปรับปรุงค่าต่าง ๆ ไปในทิศทางที่ดีขึ้นเรื่อย ๆ ตลอดการฝึก เช่น ค่า loss ที่ลดลงและค่า mAP ที่สูงขึ้น ซึ่งเป็นสัญญาณว่า การฝึกแบบจำลองมีประสิทธิภาพ



Figure 9 An example of YOLOv7 training for Climbing, Flirting, Mating, and Walking.

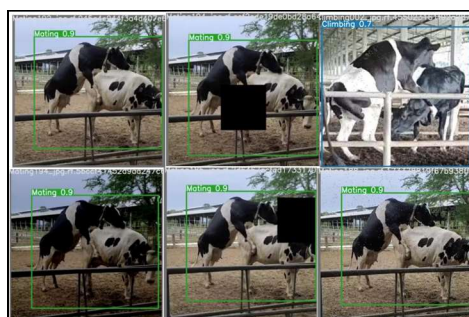


Figure 10 An example of YOLOv7 prediction for Climbing, Flirting, Mating, and Walking.

อภิปรายผล

ในการวิจัยนี้ ผู้วิจัยใช้ปัญญาประดิษฐ์เพื่อปรับปรุงการจัดการฟาร์มโคนมและเปลี่ยนให้เป็นฟาร์มอัจฉริยะ นักวิจัยใช้การเรียนรู้เชิงลึกเพื่อตรวจจับอาการเป็นสัดในโค การเป็นสัดในโคเป็นช่วงสั้น ๆ และไม่อาจคาดเดาได้ ทำให้เกษตรกรติดตามดูได้ยาก โดยเฉพาะในเวลากลางคืน นักวิจัยใช้กล้องวงจรปิดเพื่อตรวจจับและติดตามพฤติกรรมของโคโดยใช้การเรียนรู้เชิงลึกประเภท YOLO การศึกษาพบว่า YOLOv7 สามารถตรวจจับอาการเป็นสัดได้อย่างมีประสิทธิภาพ โดยเห็นได้จากค่าที่คำนวณและทำนายในตารางที่ 5 (Table 5) ตัวอย่างเช่น ค่า F1-Score = 1.00 โดยมีค่า Precision = 1.00, Recall = 1.00, mAP.5 = 0.996 และ mAP.5:95 = 0.845 โดยรวมแบบจำลองแสดงการเรียนรู้ที่ดีและมีความแม่นยำสูง แสดงให้เห็นว่าการทำนายสามารถตรวจจับได้อย่างแม่นยำด้วยข้อผิดพลาดที่น้อยมาก

Table 6 Comparison of the proposed non-invasive real-time estrus detection method in dairy cows with related research on model performance.

Author (Year)	Methods	Precision	Recall	F1-core	mAP.5:.95	Accuracy
Arago et al. (2020)	3 Pan-tilt-zoom camera, top view 1 class (Estrus)	NA	NA	NA	NA	0.5 (TP) 0.5 (FP)
Ankan et al. (2023)	pan-tilt camera and drone camera - 2 class (Estrus, Non-estrus) - total 1,638 images	CNN: 0.97 VGG-19: 0.99 YOLOv5: 0.98	CNN: 0.98 VGG-19: 0.98 YOLOv5: 0.98	CNN: 0.97 VGG-19: 0.97 YOLOv5: 0.97	NA	NA
Wang et al. (2023)	Thermal infrared temperatures of cows image - 2 class (Estrus, Non-estrus) - total 400 dataset	NA	LOGISTIC: 0.89 SVM: 0.89	NA	NA	NA
Aradhya et al. (2023)	- CCTV camera, Top-view - 2 class (Estrus, Non-estrus) - total 2,200 images	NA	XgBoost: 0.99	NA	NA	XgBoost: 0.99
Higaki et al. (2021)	- 3 webcams (front, top, and rear views of a cow) - ten estrous cycles in six cows - tie-stalled cows - background subtraction technique	0.54	NA	0.61	NA	NA
Ninphet et al. (2024)	- 4 CCTV camera (front, top, and 2 Diagonal view from top views) - 4 class (walkinf, firting, mating climbing) - total 1,520 images	YOLOv5: 0.987	YOLOv5: 1.00	YOLOv5: 0.993	YOLOv5: 0.887	NA
Proposed YOLOv7 model	- 4 CCTV camera (front, top, and 2 Diagonal view from top views) - 4 class (walkinf, firting, mating, climbing) - total 33,802 images	YOLOv7: 1.00	YOLOv7: 1.00	YOLOv7: 1.00	YOLOv7: 0.845	NA

เมื่อเปรียบเทียบประสิทธิภาพของแบบจำลองที่ผู้วิจัยพัฒนาขึ้นกับการตรวจจับอาการเป็นสัดตามเวลาจริงในโคนมโดยไม่ติดตั้งอุปกรณ์บนตัวโคด้วยเทคนิคการประมวลผลวีดิทัศน์ที่มีก่อนหน้านี้พบว่าแบบจำลองที่ผู้วิจัยพัฒนาขึ้นนี้มีประสิทธิภาพสูงกว่ามาก ๆ หรืออาจใกล้เคียงกันในบางค่าประสิทธิภาพ ดังภาพที่ 6 (Table 6) ทั้งนี้ประสิทธิภาพของแบบจำลองที่แตกต่างกันนั้น เป็นผลมาจากการบันทึกและเตรียมภาพ การแบ่งคลาส จำนวนภาพ เทคนิคที่ใช้ แบบจำลองที่ใช้ และการปรับไฮเปอร์พารามิเตอร์ เป็นต้น

จากการวิจัยนี้ แนะนำให้นำเทคโนโลยีปัญญาประดิษฐ์มาใช้ในการจัดการฟาร์มสมัยใหม่ หากเกษตรกรเลือกใช้เทคโนโลยีที่เหมาะสมในราคาที่สามารถจ่ายได้ จะสามารถลดต้นทุนการผลิตและเพิ่มผลผลิตได้อย่างมีประสิทธิภาพ ทำให้เป็นการลงทุนที่คุ้มค่า ในการวิจัยอนาคต สามารถนำวิธีการตรวจจับที่ใช้ในงานวิจัยนี้มาผสมผสานกับการตรวจสอบพฤติกรรมอื่น ๆ ของโค เช่น การเจ็บป่วยและการคลอดลูก ที่เคยพึ่งพาการสังเกตจากมนุษย์ นอกจากนี้ การพัฒนาวิธีการตรวจจับได้ก่อให้เกิดหลาย

เวอร์ชันของการเรียนรู้เชิงลึกด้วย YOLO นักวิจัยในอนาคตจึงควรใช้เวอร์ชันล่าสุดของ การเรียนรู้เชิงลึกแบบ YOLO ที่จะทำให้การเตรียมข้อมูลง่ายขึ้น ซึ่งจะปรับปรุงประสิทธิภาพของนักวิจัย ทำให้มีเวลาน้อยลงในการเตรียมข้อมูลและมีเวลามากขึ้นในการทำงานจริง

การวิจัยนี้มุ่งเน้นพัฒนาการตรวจจับอาการเป็นสัตว์ตามเวลาจริงในโคนมโดยไม่ติดตั้งอุปกรณ์บนตัวโคด้วยเทคนิคการประมวลผลวิดีโอที่สนับสนุนพื้นฐานโครงข่ายประสาทเทียมแบบสังวัตนาการ ซึ่งเป็นวิธีที่ไม่จำเป็นต้องติดตั้งอุปกรณ์ผู้ติดบนตัวโค มีข้อได้เปรียบที่สำคัญในหลายด้าน ได้แก่ การลดความเครียดของโคนม การใช้อุปกรณ์ที่ผู้ติดกับตัวโคสามารถสร้างความเครียดให้กับโคนมได้ในบางกรณี เนื่องจากโคนมอาจรู้สึกไม่สบายตัวหรือมีปฏิกิริยาเชิงลบต่ออุปกรณ์ที่ติดตั้งบนร่างกาย โดยเฉพาะอย่างยิ่งในช่วงเริ่มต้นที่ต้องการให้โคนมคุ้นเคยกับอุปกรณ์ และช่วยลดความเสี่ยงในการเกิดความเครียดที่อาจนำไปสู่ปัญหาสุขภาพอื่น ๆ (Homer et al., 2013) การลดภาระในการบำรุงรักษาอุปกรณ์ที่ผู้ติดกับตัวโคมักต้องการการบำรุงรักษา และการตรวจสอบอย่างสม่ำเสมอเพื่อให้มั่นใจว่าอุปกรณ์ทำงานได้อย่างถูกต้อง ตัวอย่างเช่น เซนเซอร์ที่ติดตั้งบนขาโคนมอาจเกิดปัญหาการเสียหายหรือการเคลื่อนย้ายตำแหน่ง ซึ่งอาจทำให้ข้อมูลที่ได้รับไม่แม่นยำ (Saint-Dizier & Chastant-Maillard, 2018) การลดต้นทุน การใช้อุปกรณ์ผู้ติดกับตัวโคอาจมีค่าใช้จ่ายสูงทั้งในด้านการจัดซื้อและการบำรุงรักษา (Royal et al., 2000)

สรุปผลการวิจัย

การพัฒนาขั้นตอนวิธีและประเมินประสิทธิภาพของแบบจำลองการตรวจจับอาการเป็นสัตว์ตามเวลาจริงในโคนมโดยไม่ติดตั้งอุปกรณ์บนตัวโคด้วยเทคนิคการประมวลผลวิดีโอที่สนับสนุนพื้นฐานโครงข่ายประสาทเทียมแบบสังวัตนาการ

1. ขั้นตอนการพัฒนา โดยตรวจจับพฤติกรรมการเป็นสัตว์ของโค 4 ท่าทาง ได้แก่ เดิน (Walking) ดมหรือเลีย (Flirting) ปีนทับ (Mating) และปีนด้านศีรษะซึ่งเป็นการปีนเล่นหรือปีนผิวด้าน (Climbing) โดยแต่ละท่าทางมีจำนวน 380 ภาพ รวมทั้งหมด 1,520 ภาพ ผู้วิจัยทำการเพิ่มภาพ (Image Augmentation) ด้วยเทคนิคต่าง ๆ เป็นจำนวน 22,238 เท่าของภาพต้นฉบับ จนได้ชุดข้อมูลภาพรวมทั้งสิ้น 33,802 ภาพ ทำการสร้างแบบจำลองการตรวจจับ สร้างแบบจำลองการตรวจจับด้วยโครงข่ายประสาทเทียมแบบสังวัตนาการ ชนิด YOLOv7 จำนวนตัวอย่างในการเรียนรู้ที่ใช้ในแต่ละรอบการปรับปรุงแบบจำลอง (batch size) มีค่า 16 ข้อมูลเรียนรู้ผ่านเข้ามาในโครงข่าย (epochs) 30 รอบ ภาพถูกปรับขนาดให้มีความกว้างและความสูง (img-size) อยู่ที่ 640x640 พิกเซล

2. แบบจำลองที่พัฒนาขึ้นสามารถตรวจจับอาการเป็นสัตว์ของโคได้อย่างมีประสิทธิภาพ โดยมีค่า F1-Score เฉลี่ย = 1.00 ค่า Precision = 1.00, ค่า Recall = 1.00, mAP.5 = 0.996 และ mAP.5:95 = 0.845

แสดงให้เห็นว่าแบบจำลองนี้เหมาะสมที่จะนำไปใช้การฟาร์มเกษตรกรผู้เลี้ยงโคนมทดแทนการตรวจจับอาการเป็นสัตว์ของโคนมโดยใช้พนักงานเลี้ยงสัตว์

กิตติกรรมประกาศ

ผู้วิจัยขอขอบพระคุณมหาวิทยาลัยราชภัฏหมู่บ้านจอมบึงที่ให้ทุนการศึกษาในระดับบัณฑิตศึกษาและสนับสนุนการวิจัยในครั้งนี้ ขอขอบพระคุณชัยตระกราด์ฟาร์ม อำเภोजอมบึง และอุดมแต่ร์ฟาร์ม อำเภอบ้านโป่ง จังหวัดราชบุรี ที่ให้ความอนุเคราะห์เก็บข้อมูลวิจัยภาพวิดีโอทัศนอาการเป็นสัตว์ของโคนมที่ใช้ในการวิจัยครั้งนี้

เอกสารอ้างอิง

- Aradhya AS, Patil AS, Rao KP, Monika M, Shireesha V. COWOZE: An Automated Estrus Detection System. *International Journal of Creative Research Thoughts* 2023;11(4):492-496.
- Arago NM, Alvarez CI, Mabale AG, Legista CG, Repiso NE, Robles RRA, Amado TM. Automated estrus detection for dairy cattle through neural networks and bounding box corner analysis. *International Journal of Advanced Computer Science and Applications* 2020;11(9):303-311.
- Ankan İ, Ayav T, Seçkin AÇ, Soygazi F. Estrus Detection and dairy cow identification with cascade deep learning for augmented reality-ready livestock farming. *Sensors*. 2023;23(24):9795.
- Aydin M, Risvanli A, Timurkan H, Kaygusuzoglu E. Lack of correlation between the electrical conductivity of milk and the blood progesterone levels in cows. *Journal of the South African Veterinary Association* 2008;79(3):153-154.
- Benaissa S, Tuytens FAM, Plets D, Martens L, Vandaele L, Joseph W, Sonck B. Improved cattle behaviour monitoring by combining Ultra-Wideband location and accelerometer data. *Animal* 2023;17(4):100730.
- Chowdhury S, Verma B, Roberts J, Corbet N, Swain D. Deep learning based computer vision technique for automatic heat detection in cows. In: 2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA);2016:1-6.
- Higaki S, Horiata K, Suzuki C, Sakurai R, Suda T, Yoshioka K. Estrus detection using background image subtraction technique in tie-stalled cows. *Animals* 2021;11(6):1795.
- Higaki S, Okada H, Suzuki C, Sakurai R, Suda T, Yoshioka K. Estrus detection in tie-stall housed cows through supervised machine learning using a multimodal tail-attached device. *Computers and Electronics in Agriculture* 2021;191:106513.
- Homer EM, Gao Y, Meng X, Dodson A, Webb R, Garnsworthy PC. Technical note: A novel approach to the detection of estrus in dairy cows using ultra-wideband technology. *Journal of Dairy Science* 2013;96(10):6529-6534.
- Jónsson R, Blanke M, Poulsen NK, Caponetti F, Højsgaard S. Oestrus detection in dairy cows from activity and lying data using on-line individual models. *Computers and Electronics in Agriculture*. 2011;76(1):6-15.
- Li D, Wang JH, Zhang Z, Dai BS, Zhao KX, Shen WZ, et al. Cow-YOLO: Automatic cow mounting detection based on non-local CSPDarknet53 and multiscale neck. *International Journal of Agricultural and Biological Engineering* 2024;17(3):193-202.

- MacKay JRD, Deag JM, Haskell MJ. Establishing the extent of behavioural reactions in dairy cattle to a leg mounted activity monitor. *Applied Animal Behaviour Science* 2012;139(1-2):35-41.
- Marangoni F, Pellegrino L, Verduci E, Ghiselli A, Bernabei R, Calvani R, Cetin I, Giampietro M, et al. Cow's Milk Consumption and Health: A Health Professional's Guide. *Journal of the American College of Nutrition* 2018;38(3):197-208.
- Morera Á, Sánchez Á, Moreno AB, Sappa ÁD, Vélez JF. SSD vs. YOLO for Detection of Outdoor Urban Advertising Panels under Multiple Variabilities. *Sensors* 2020;20(16):4587.
- Ninphet W, Amdee N, Sangsongfa A. Adaptive deep learning for image-based estrus prediction and detection in Dairy Cows. *PriMera Scientific Engineering* 2024;5(2):12-37.
- Ninphet W, Amm-Dee N, Sangsongfa A. Prediction cows estrus images using convolutional neural network with optimized parameters by the artificial immune system algorithm. In: *Proceedings of the 20th International Conference on Computing and Information Technology (IC2IT 2024)*. Springer Link. 2024;973:105-120.
- Office of Agricultural Economics Research. Situation of important agricultural products and trends in 2024. Dairy. Office of Agricultural Economics. Published Dec, 2023. Available at: <https://www.oae.go.th/assets/portals/1/files/journal/2566/trend2567.pdf>. Accessed August 20, 2024.
- Palmer MA, Olmos G, Boyle LA, Mee JF. Estrus detection and estrus characteristics in housed and pastured Holstein–Friesian cows. *Theriogenology* 2010;74(2):255-264.
- Pratelli G, Tamburini B, Badami GD, Lo Pizzo M, De Blasio A, Carlisi D, Di Liberto D. Cow's Milk: A Benefit for Human Health? Omics Tools and Precision Nutrition for Lactose Intolerance Management. *Nutrients* 2024;16(2):320.
- Reith S, Hoy S. Relationship between daily rumination time and estrus of dairy cows. *Journal of Dairy Science* 2012;95(11):6416-6420.
- Roelofs JB, van Eerdenburg FJCM, Soede NM, Kemp B. Pedometer readings for estrous detection and as predictor for time of ovulation in dairy cattle. *Theriogenology* 2005;64(8):1690-1703.
- Royal MD, Darwash AO, Flint APF, Webb R, Woolliams JA, Lamming GE. Declining fertility in dairy cattle: changes in traditional and endocrine parameters of fertility. *Animal Science*. 2000;70(3):487-501.
- Saint-Dizier M, Chastant-Maillard S. Potential of connected devices to optimize cattle reproduction. *Theriogenology* 2018;112:53-62.
- Senger PL. The estrus detection problem: New concepts, technologies, and possibilities. *Journal of Dairy Science* 1994;77(9):2745-2753.
- Stangaferro ML, Wijma R, Caixeta LS, Al-Abri MA, Giordano JO. Use of rumination and activity monitoring for the identification of dairy cows with health disorders: Part I. Metabolic and digestive disorders. *Journal of Dairy Science* 2016;99(9):7395-7410.

- Wang CY, Bochkovskiy A, Liao HY. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2023. p. 7464-7474. Available from: https://openaccess.thecvf.com/content/CVPR2023/papers/Wang_YOLOv7_Trainable_Bag-of-Freebies_Sets_New_State-of-the-Art_for_Real-Time_Object_Detectors_CVPR_2023_paper.pdf
- Wang J, Chen H, Wang J, Zhao K, Li X, Liu B, Zhou Y. Identification of oestrus cows based on vocalisation characteristics and machine learning technique using a dual-channel-equipped acoustic tag. *Animal* 2023;17(6):100811.
- Wang Z, Wang S, Wang C, Zhang Y, Zong Z, Wang H, Su L, et al. A Non-Contact Cow Estrus Monitoring Method Based on the Thermal Infrared Images of Cows. *Agriculture* 2023;13(2):385.