การเปรียบเทียบประสิทธิภาพของ LRT และ YOLO สำหรับการแบ่งส่วนพื้นที่น้ำท่วมจากภาพกล้องวงจรปิด: กรณีศึกษาจังหวัดน่าน ประเทศไทย

วรากร เลื่องลือวุฒิ[1, 2] พันศักดิ์ เทียนวิบูลย์[2*] กิตตากร วิริยะศาสตร์[1]
ชยยชญ์ เสรีจิตติมา[3] และ ปิยะรส มาลีเจริญ[3]

วันที่รับ 10 เมษายน 2568 วันที่แก้ไข 4 มิถุนายน 2568 วันตอบรับ 4 มิถุนายน 2568

บทคัดย่อ

งานวิจัยนี้นำเสนอแนวทางการจำแนกพื้นที่น้ำจากภาพกล้องโทรทัศน์วงจรปิด (Closed-Circuit Television หรือ CCTV) เพื่อการเฝ้าระวังน้ำท่วมในจังหวัดน่าน ประเทศไทย โดยใช้แบบจำลอง Likelihood Ratio Test (LRT) ภายใต้สมมติฐานการแจกแจงแบบ Gaussian ที่มีความแปรปรวนเท่ากันระหว่างคลาส (Equal-Variance Gaussian Distribution) โดยใช้ค่าความเข้ม (Intensity) ของภาพ Grayscale เฉพาะในบริเวณที่สนใจ (Region of Interest หรือ ROI) สำหรับการจำแนกจุดภาพว่าเป็น “น้ำ” หรือ “ไม่ใช่น้ำ” พร้อมเปรียบเทียบกับแบบจำลอง YOLOv11n-seg และ YOLOv11x-seg ซึ่งเป็นเทคนิคการเรียนรู้เชิงลึก (Deep Learning) ที่ได้รับความนิยม ผลการทดลองแสดงให้เห็นว่า แบบจำลอง LRT ให้ค่า F1 Score เท่ากับ 0.84 ซึ่งใกล้เคียงหรือดีกว่าแบบจำลอง YOLO ทั้งสองเวอร์ชันที่ได้รับการฝึกจากข้อมูลต่างโดเมน (Domain adaptation) และมีความโดดเด่นด้านความเร็วในการประมวลผล โดยสามารถทำงานได้สูงถึง 679.54 เฟรมต่อวินาที บน CPU จึงเหมาะสมอย่างยิ่งสำหรับระบบที่มีข้อจำกัดด้านทรัพยากรและต้องการประมวลผลแบบเรียลไทม์ แม้ว่า LRT จะให้ผลลัพธ์ที่ดีในบริบทของข้อมูลเฉพาะในพื้นที่ศึกษา แต่แบบจำลอง YOLO ยังคงแสดงให้เห็นถึงศักยภาพในการ Generalize ภายใต้สถานการณ์ Domain shift ได้อย่างมีประสิทธิภาพ จึงเหมาะสำหรับการใช้งานในกรณีที่ข้อมูลมีความหลากหลายหรือยังไม่แน่นอน งานวิจัยนี้จึงเสนอข้อพิจารณาเชิงเปรียบเทียบระหว่างแบบจำลองที่มีโครงสร้างเรียบง่ายกับแบบจำลองเชิงลึก เพื่อสนับสนุนการพัฒนาระบบเฝ้าระวังน้ำท่วมที่มีประสิทธิภาพและเหมาะสมกับบริบทการใช้งานที่แตกต่างกัน

คำสำคัญ: การแบ่งส่วนภาพ, การจำแนกน้ำ, กล้องวงจรปิด, Likelihood Ratio Test (LRT), ภาพระดับสีเทา, บริเวณที่สนใจ, ระบบเรียลไทม์ , การเรียนรู้เชิงลึก, You Only Live Once (YOLO)

[1] ส่วนงานวิศวกรรมการสื่อสารข้อมูลทางอิเล็กทรอนิกส์และเครือข่ายคอมพิวเตอร์, สถาบันเทคโนโลยีป้องกันประเทศ
[2] ภาควิชาวิศวกรรมไฟฟ้า, คณะวิศวกรรมศาสตร์, มหาวิทยาลัยเกษตรศาสตร์
[3] ส่วนงานวิศวกรรมระบบจำลองยุทธ์และเครื่องช่วยฝึกเสมือนจริง, สถาบันเทคโนโลยีป้องกันประเทศ
[*] ผู้แต่ง, อีเมล: phunsak.t@ku.ac.th

Performance Comparison of LRT and YOLO for Floodwater Segmentation from CCTV Images: A Case Study in Nan Province, Thailand

Warakorn Luangluewut[1, 2] Phunsak Thiennviboon[2*] Kittakorn Viriyasatr[1]
Chayayot Saerejittima[3] and Piyarose Maleecharoen[3]

Received 10 April 2025, Revised 4 June 2025, Accepted 4 June 2025

Abstract

This research article presents an approach for detecting water-covered areas in closed -circuit television (CCTV) images to support flood surveillance in Nan Province, Thailand. The proposed method employed a Likelihood Ratio Test (LRT) model under the assumption of equal-variance Gaussian distributions between classes. Pixel intensities from grayscale images within a predefined Region of Interest (ROI) were used as input features to classify each pixel as either "water" or "non-water." The LRT model was compared with two widely used deep learning-based segmentation models, YOLOv11n-seg and YOLOv11x-seg. Experimental results show that the LRT model achieved an F1 Score of 0.84. That was comparable to or better than the two YOLO variants trained using domain adaptation. Notably, LRT demonstrated a significant advantage in processing speed, reaching up to 679.54 frames per second (fps) on a CPU, making it highly suitable for real-time applications on resource-constrained systems. While the LRT model yielded strong performance within the specific context of the study area, the YOLO models exhibited better generalization capabilities under domain shift conditions. Therefore, YOLO models may be more appropriate for scenarios involving diverse or uncertain data characteristics. This research highlights a comparative perspective between simple -structured models and deep learning models, providing practical insights for the development of effective and context-appropriate flood surveillance systems.

Keywords: Image segmentation, Water classification, Closed-Circuit Television (CCTV), Likelihood Ratio Test (LRT), Grayscale images, Region of Interest (ROI), Real-time system, Deep Learning, You Only Live Once (YOLO)

---

[1] Data Communication Division, Defence Technology Institute

[2] Department of Electrical Engineering, Faculty of Engineering, Kasetsart University

[3] Simulation Systems and Virtual Training Division, Defence Technology Institute

[*] Corresponding author, E-mail: phunsak.t@ku.ac.th

1. บทนำ

ปัญหาน้ำท่วมเป็นหนึ่งในภัยพิบัติที่ส่งผลกระทบอย่างรุนแรงต่อความปลอดภัยของประชาชนและเศรษฐกิจของประเทศ โดยเฉพาะในเขตเมืองที่มีความหนาแน่นของสิ่งปลูกสร้างสูง การติดตามระดับน้ำอย่างต่อเนื่องและการวิเคราะห์พื้นที่น้ำท่วมจึงมีความสำคัญอย่างยิ่งต่อการแจ้งเตือนล่วงหน้าและการวางแผนบริหารจัดการภัยพิบัติอย่างมีประสิทธิภาพ

ในปัจจุบัน หน่วยงานราชการในประเทศไทยได้ติดตั้งกล้องโทรทัศน์วงจรปิด (Closed-Circuit Television หรือ CCTV) ในหลายพื้นที่ ทำให้สามารถใช้ภาพจากกล้องเหล่านี้เป็นแหล่งข้อมูลต้นทุนต่ำสำหรับการตรวจจับพื้นที่น้ำ ควบคู่ไปกับการพัฒนาโปรแกรมวิเคราะห์ภาพที่สามารถจำแนกพื้นที่น้ำออกจากพื้นผิวอื่นได้อย่างรวดเร็วและแม่นยำ

งานวิจัยนี้จึงนำเสนอวิธีการจำแนกพื้นที่น้ำจากภาพที่ได้จากกล้อง CCTV ในจังหวัดน่าน ประเทศไทย โดยใช้ค่าความเข้มของภาพแบบ Grayscale ภายในบริเวณที่สนใจ (Region of Interest หรือ ROI) [1] ร่วมกับแบบจำลอง Likelihood Ratio Test (LRT) ซึ่งพัฒนาภายใต้สมมติฐานการแจกแจงแบบ Gaussian ที่มีความแปรปรวนเท่ากัน (Equal-Variance Gaussian Distribution) [2] เพื่อจำแนกจุดภาพ (Pixel) แต่ละตำแหน่งให้อยู่ในคลาส “น้ำ” หรือ “ไม่ใช่น้ำ” อย่างมีประสิทธิภาพ

กระบวนการวิเคราะห์ภาพโดยใช้ ROI ช่วยลดความซับซ้อนของข้อมูล ลดความต้องการของชุดข้อมูลขนาดใหญ่ และเพิ่มความแม่นยำในการจำแนก ทั้งยังสามารถประมวลผลได้อย่างรวดเร็วบนอุปกรณ์ที่มีทรัพยากรจำกัด เช่น ระบบที่ใช้เพียงหน่วยประมวลผลกลาง (CPU) เป็นต้น ซึ่งเหมาะสมสำหรับระบบเฝ้าระวังภัยพิบัติที่ต้องการการตอบสนองแบบเรียลไทม์ เพื่อประเมินประสิทธิภาพของแนวทางที่นำเสนอ คณะผู้วิจัยได้ทำการเปรียบเทียบผลกับแบบจำลอง YOLO (You Only Look Once) [3] - [5] ซึ่งเป็นแบบจำลอง Deep learning [6] ที่ได้รับความนิยมอย่างสูงในงานตรวจจับวัตถุ โดยเฉพาะเวอร์ชัน YOLOv11 ที่สามารถประมวลผลทั้งการตรวจจับและการแบ่งส่วนภาพ (segmentation) ได้ในขั้นตอนเดียว

2. ทฤษฎีที่เกี่ยวข้อง

2.1 Image Segmentation

Image Segmentation เป็นกระบวนการแบ่งภาพออกเป็นส่วนย่อย (Segments) โดยพิจารณาจากคุณลักษณะต่าง ๆ เช่น สี ความสว่าง พื้นผิว หรือบริบทของภาพ เป็นต้น เพื่อให้ง่ายต่อการวิเคราะห์และประมวลผลข้อมูลในขั้นตอนถัดไป โดยมีเป้าหมายเพื่อจัดกลุ่มจุดภาพที่มีลักษณะหรือบริบทที่เกี่ยวข้องกันให้อยู่ในกลุ่มเดียวกัน และแยกออกจากจุดภาพที่มีลักษณะหรือบริบทแตกต่างกัน [7]

2.2 ภาพ Grayscale

ภาพ Grayscale [8] คือภาพดิจิทัลที่แสดงเฉพาะระดับความเข้มของแสง (Intensity) โดยไม่มีข้อมูลเกี่ยวกับสี (Hue) หรือความอิ่มตัวของสี (Saturation) จุดภาพในภาพประเภทนี้จะมีค่าความสว่างอยู่ในช่วงตั้งแต่ 0 ถึง 255 การแปลงภาพสี (RGB) ให้เป็นภาพ Grayscale สามารถทำได้โดยใช้สมการแบบเชิงเส้น ดังสมการ (1)

$$Grayscale = (0.299 \times R) + (0.587 \times G) + (0.114 \times B) \quad (1)$$

โดยที่ R, G, และ B คือ ค่าความเข้มของจุดภาพในช่องสีแดง เขียว และน้ำเงิน ตามลำดับ ซึ่งค่าสัมประสิทธิ์ 0.299, 0.587 และ 0.114 มาจากการแปลงค่าความสว่าง (Luminance) ตามมาตรฐาน ITU-R BT.601 [9]

ในงานวิจัยนี้ การแปลงภาพจาก RGB เป็นภาพ Grayscale ช่วยให้สามารถพัฒนาแบบจำลองที่มีโครงสร้างเรียบง่าย และยังสามารถนำค่าความเข้ม (Intensity) ของจุดภาพในภาพไปใช้ในการจำแนกพื้นที่น้ำและพื้นที่ที่ไม่ใช่น้ำได้อย่างมีประสิทธิภาพ ดังแสดงในหัวข้อ 4.2

2.3 การจำแนกด้วยแบบจำลอง Likelihood Ratio Test (LRT)

ในทางสถิติ Likelihood Ratio Test (LRT) [2] เป็นวิธีการทดสอบสมมติฐาน (Hypothesis testing) ที่ใช้เปรียบเทียบความเหมาะสม (Goodness-of-Fit) ของสองแบบจำลองเชิงสถิติ โดยพิจารณาสัดส่วนของค่า likelihood ระหว่างแบบจำลองที่มีสมมติฐานสองประเภท วิธีการนี้สามารถประยุกต์ใช้ในงานจำแนกประเภท (Classification) โดยที่แต่ละประเภทหรือคลาส (Class) สามารถมองได้ว่าเป็นสมมติฐานหนึ่ง ดังนั้น LRT จึงสามารถถูกใช้เป็นแบบจำลองสำหรับการจำแนกประเภท (Classification model) ได้

ในกรณีของการจำแนกประเภทแบบทวิภาค (Binary classification) แบบจำลอง LRT จะประเมิน Likelihood หรือความน่าจะเป็นแบบมีเงื่อนไข (Conditional Probability Density Function) ของข้อมูลเข้า x ภายใต้แต่ละคลาส แล้วคำนวณอัตราส่วนของค่า Likelihood (Likelihood Ratio) ดังสมการ (2)

$$\Lambda(x) = \frac{p(x|C_1)}{p(x|C_0)} \underset{C_0}{\overset{C_1}{\gtrless}} \eta \tag{2}$$

โดยที่

x คือ ข้อมูลเข้า (Input) กำหนดให้มีค่าเป็นจำนวนจริง (Real number)

$\Lambda(x)$ คือ ค่า Likelihood ratio

$p(x|C_i)$ คือ ความน่าจะเป็นแบบมีเงื่อนไขของคลาส C_i สำหรับ $i \in \{0,1\}$

η คือ ค่า Threshold สำหรับการตัดสินใจ

หากสมมติว่า Likelihood ของแต่ละคลาสมีการแจกแจงแบบ Gaussian [10] – [11] ที่มีค่าความแปรปรวน (Variance) เท่ากัน (Equal-Variance Gaussian Distribution) จะได้ดังสมการ (3)

$$p(x|C_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu_i)^2}{2\sigma^2}} \tag{3}$$

โดยที่ μ_i และ σ^2 คือ ค่าเฉลี่ย (Mean) ของ Likelihood ของคลาสที่ i $(i \in \{0,1\})$ และค่าความแปรปรวน (Variance) ของทั้งสองคลาสตามลำดับ โดยการแทนค่า (3) ในสมการ (2) แล้วจัดรูปจะได้กฎการตัดสินใจที่พึ่งพาเพียงค่า x ดังสมการ (4)

$$x \underset{C_0}{\overset{C_1}{\gtrless}} T = \frac{2\sigma^2 \ln(\eta) + \mu_1^2 - \mu_0^2}{2(\mu_1 - \mu_0)} \tag{4}$$

นั่นคือ

หาก $x > T$ ให้จำแนกว่าอยู่ในคลาส C_1

หาก $x < T$ ให้จำแนกว่าอยู่ในคลาส C_0

หาก $x = T$ การตัดสินสามารถทำได้ตามความเหมาะสม โดยในงานวิจัยนี้จะตัดสินให้ x อยู่ในคลาส C_1

ค่าคงที่ T นี้เรียกว่า Threshold ซึ่งสามารถกำหนดได้โดยใช้เกณฑ์หลายรูปแบบ ทั้งในเชิงทฤษฎีและเชิงประจักษ์ ในงานวิจัยนี้ คณะผู้วิจัยได้เลือกใช้แนวทางเชิงประจักษ์ โดยกำหนดค่า Threshold ที่ให้ค่า F1 Score สูงสุดบนชุดข้อมูลฝึก ซึ่งถือเป็นการเลือก Threshold โดยใช้วิธี Metric-based Optimization แทนการใช้เกณฑ์ทางทฤษฎี เช่น Bayes criterion หรือ Neyman–Pearson Criterion เป็นต้น ดังจะอธิบายเพิ่มเติมในหัวข้อ 3.3

2.4 การจำแนกพื้นที่น้ำด้วยแบบจำลอง Deep learning

งานวิจัยที่ผ่านมาได้มีการศึกษาการจำแนกพื้นที่น้ำและพื้นที่ที่ไม่ใช่น้ำด้วยวิธีการหลากหลาย เช่น การจำแนกพื้นที่แม่น้ำ [12] - [13] และพื้นที่น้ำท่วม [14] - [16] โดยส่วนใหญ่มักใช้โมเดล Deep learning ที่มีความซับซ้อนสูง ซึ่งแม้จะให้ผลลัพธ์ที่แม่นยำ แต่อาจไม่เหมาะสมกับการใช้งานแบบเรียลไทม์ภายใต้ข้อจำกัดด้านทรัพยากร เนื่องจากมีความเร็วในการประมวลผลที่ค่อนข้างต่ำ

ในงานวิจัยนี้ได้เลือกใช้เทคนิค YOLO Segmentation [17] เป็น Baseline สำหรับการเปรียบเทียบกับเทคนิคที่นำเสนอในงานวิจัยนี้ ซึ่งเป็นวิธีการที่เรียบง่ายกว่า

โดยใช้สำหรับการแบ่งส่วนภาพ (Segmentation) เพื่อจำแนกพื้นที่น้ำและพื้นที่ที่ไม่ใช่น้ำ โดยเน้นประสิทธิภาพและความเร็วที่เหมาะสมกับระบบที่มีข้อจำกัดด้านทรัพยากรและต้องการประมวลผลแบบเรียลไทม์

ทั้งนี้ มีงานวิจัยก่อนหน้าที่มีการประยุกต์ใช้ YOLO Segmentation สำหรับการแบ่งส่วนภาพเพื่อจำแนกพื้นที่น้ำท่วมเช่นกัน [15] - [16] โดยเวอร์ชันล่าสุดของ YOLO เช่น YOLOv8-seg, YOLOv10-mask, และ YOLOv11-seg ได้มีการเพิ่ม Segmentation head เข้าไปในโครงสร้างของโมเดล ทำให้สามารถแบ่งส่วนภาพได้โดยตรง และให้ผลลัพธ์ที่รวดเร็ว เหมาะสำหรับงานที่ต้องการการประมวลผลแบบเรียลไทม์

2.5 Domain adaptation

Domain adaptation [18] - [19] เป็นส่วนหนึ่งของกระบวนการเรียนรู้เชิงถ่ายโอน (Transfer learning) ที่มุ่งเน้นให้แบบจำลองสามารถทำงานได้อย่างมีประสิทธิภาพในโดเมนเป้าหมาย (Target domain) ซึ่งมีลักษณะของข้อมูลแตกต่างจากโดเมนต้นทาง (Source domain) ที่ใช้ในการฝึกแบบจำลอง

โดยทั่วไป แบบจำลองการเรียนรู้มักตั้งอยู่บนสมมติฐานที่ว่า ข้อมูลฝึก (Training data) และข้อมูลทดสอบ (Testing data) ต้องมาจากการแจกแจงข้อมูลเดียวกัน (Data distribution) อย่างไรก็ตาม ในสภาพแวดล้อมจริง ข้อมูลที่ใช้งานมักมีความแตกต่างหรือเบี่ยงเบนจากกัน เช่น ความแตกต่างของแสง เงา คุณภาพของกล้อง หรือสภาพแวดล้อมที่เปลี่ยนไป ทำให้แบบจำลองที่ถูกฝึกจากโดเมนหนึ่ง อาจไม่สามารถนำไปใช้งานได้ดีในอีกโดเมนหนึ่ง การปรับตัวระหว่างโดเมน (Domain adaptation) จึงเป็นเทคนิคที่ถูกพัฒนาขึ้นเพื่อลดช่องว่างระหว่างโดเมนต้นทางและโดเมนเป้าหมาย

ในบริบทของงานวิจัยนี้ การจำแนกพื้นที่น้ำบริเวณจุดเฝ้าระวังน้ำท่วมจากภาพกล้องวงจรปิด (CCTV) ประสบกับปัญหาข้อมูลจำกัดในโดเมนเป้าหมาย ดังนั้น จึงมีการใช้ข้อมูลจากอีกโดเมนหนึ่ง ได้แก่ พื้นที่แม่น้ำ เพื่อฝึกแบบจำลองเบื้องต้น ก่อนจะนำแบบจำลองดังกล่าวไปประเมินผลบนข้อมูลจากจุดเฝ้าระวังน้ำท่วมจริง

2.6 Confusion matrix

Confusion matrix [20] เป็นเครื่องมือพื้นฐานที่ใช้ในการประเมินประสิทธิภาพของโมเดลจำแนกประเภท (Classification model) โดยแสดงผลการทำนายที่ถูกและผิดของโมเดลในแต่ละคลาสในรูปของตาราง ซึ่งช่วยให้สามารถวิเคราะห์ข้อผิดพลาดและประสิทธิภาพของโมเดลได้อย่างชัดเจน ดังแสดงในตารางที่ 1

ตารางที่ 1 ตาราง Confusion Matrix

		Predicted Class	
		Positive (1)	**Negative (0)**
Actual Class	**Positive (1)**	*True Positive (TP)*	*False Negative (FN)*
	Negative	*False Positive (FP)*	*True Negative (TN)*

ตัวชี้วัดที่ใช้กันทั่วไป ได้แก่

ความถูกต้อง (Accuracy) คือ อัตราส่วนของการทำนายที่ถูกต้องต่อจำนวนตัวอย่างทั้งหมด โดยคำนวณได้จากสมการที่ (5)

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

ความแม่นยำ (Precision) คือ ค่าความถูกต้องเฉพาะในกรณีที่โมเดลทำนายว่าเป็นบวก (Predicted positive) โดยคำนวณได้จากสมการที่ (6)

$$Precision = \frac{TP}{TP+FP} \quad (6)$$

Recall คือ ค่าความถูกต้องเฉพาะในกรณีเหตุการณ์ที่เป็นบวกจริง (Actual Positive) โดยคำนวณได้จากสมการที่ (7)

$$Recall = \frac{TP}{TP+FN} \quad (7)$$

F1 Score คือ ค่าประสิทธิภาพที่เป็นค่าเฉลี่ยฮาร์มอนิก (Harmonic Mean) ของ Precision และ Recall โดยเป็นการรวมทั้งสองค่าดังกล่าวเข้าด้วยกันเป็นค่าประสิทธิภาพเพียงค่าเดียว ซึ่งเหมาะสำหรับกรณีที่ต้องการพิจารณาทั้งความแม่นยำในการทำนายและความสามารถในการตรวจจับเหตุการณ์ที่เกิดขึ้นจริง โดยคำนวณได้จากสมการที่ (8)

$$F1\ Score = \frac{2}{\left(\frac{1}{Recall}\right)+\left(\frac{1}{Precision}\right)} \quad (8)$$

IoU (Intersection over Union) มักใช้ในงานตรวจจับวัตถุ (Object detection) หรือการแบ่งส่วนภาพ (Image segmentation) เพื่อวัดความแม่นยำเชิงพื้นที่ โดยนิยามเป็นอัตราส่วนของพื้นที่ที่เป็นบวกที่ทำนายได้ถูกต้อง (True positive หรือ Intersection) กับพื้นที่รวมทั้งหมด (Union) ของพื้นที่ที่เป็นบวก (Actual Positive) กับพื้นที่ที่โมเดลทำนายว่าเป็นบวก (Predicted Positive) ดังสมการที่ (9)

$$IoU = \frac{TP}{TP+FP+FN} \quad (9)$$

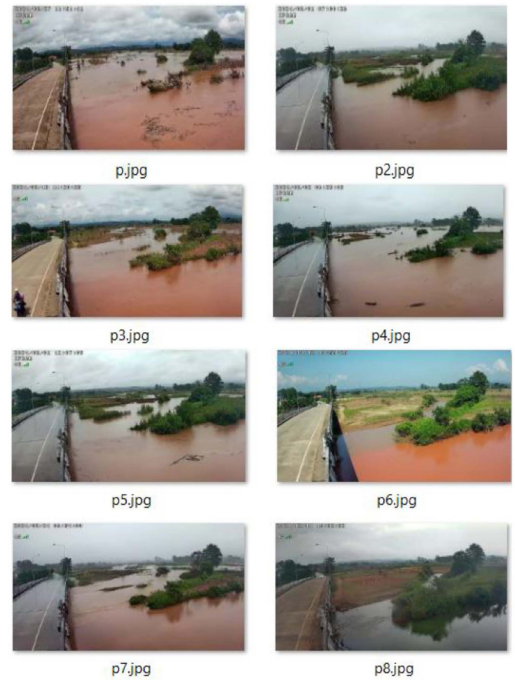
3. วิธีการดำเนินการ

3.1 การเตรียมชุดข้อมูลฝึกและชุดข้อมูลทดสอบจากภาพกล้องวงจรปิด

ข้อมูลภาพที่ใช้ในการฝึก (สำหรับแบบจำลอง LRT) และข้อมูลภาพที่ใช้ในการทดสอบ (สำหรับแบบจำลอง LRT และ YOLO Segmentation) ในงานวิจัยนี้ได้มาจากกล้องวงจรปิด (CCTV) ซึ่งติดตั้ง ณ จุดเฝ้าระวังน้ำท่วมในจังหวัดน่าน โดยรวมมีจำนวนทั้งสิ้น 8 ภาพ โดยแต่ละภาพมีขนาด 640x360 ซึ่งมีภาพเป็นจำนวนที่จำกัด เนื่องจากพื้นที่ที่ศึกษาเป็นพื้นที่ใหม่ที่ยังมีข้อมูลสะสมอยู่น้อย ลักษณะเฉพาะของข้อมูลชุดนี้ คือ สีของน้ำในภาพปรากฏเป็นสีแดง ซึ่งแตกต่างจากลักษณะของน้ำทั่วไปอย่างชัดเจน เนื่องจากมีตะกอนดินปนอยู่ในน้ำอย่างเข้มข้น จึงทำให้ภาพจากพื้นที่นี้มีความแตกต่างจากโดเมนอื่นที่เคยใช้ในงานวิจัยที่ผ่านมา

ภาพทั้ง 8 นี้ถูกบันทึกในช่วงเวลาต่าง ๆ ที่มีระดับน้ำท่วมแตกต่างกัน เพื่อแสดงให้เห็นความหลากหลายของภาพระดับน้ำในพื้นที่ศึกษาดังแสดงในรูปที่ 1 คณะผู้วิจัยได้สุ่มเลือกภาพ p3.jpg เพื่อนำไปสร้างชุดข้อมูลฝึกสำหรับแบบจำลอง LRT ในขณะที่ภาพที่เหลืออีก 7 ภาพ ถูกนำไปสร้างชุดข้อมูลทดสอบสำหรับประเมินแบบจำลอง LRT และ YOLO Segmentation บน CPU ของ Google Colab

สำหรับกระบวนการเตรียมฉลาก (Label) ซึ่งเป็นกระบวนการสำคัญในการออกแบบและพัฒนาแบบจำลอง คณะผู้วิจัยได้ทำการจำแนกจุดภาพในภาพออกเป็นสองประเภท ได้แก่ พื้นที่น้ำ และ พื้นที่ที่ไม่ใช่น้ำ โดยกำหนดให้จุดภาพที่เป็นน้ำมีสีขาว (ค่า 1) และจุดภาพที่ไม่ใช่น้ำมีสีดำ (ค่า 0) โดยมีตัวอย่างแสดงในรูปที่ 2



รูปที่ 1 ข้อมูลภาพจากกล้องวงจรปิดทั้งหมดที่ใช้ในการทดลอง

Image

(a)

Label

(b)

รูปที่ 2 ตัวอย่างข้อมูลภาพและฉลาก (Label)
(a) ข้อมูลภาพจริง, (b) ฉลากของข้อมูลภาพ (Label)

3.2 การเตรียมชุดข้อมูลฝึกสำหรับแบบจำลอง YOLOv11 Segmentation

ในงานวิจัยนี้ คณะผู้วิจัยเลือกใช้แบบจำลอง YOLOv11 Segmentation [21] เพราะเป็นหนึ่งในเวอร์ชันของ YOLO ที่มีความเร็วในการประมวลผลสูงที่สุดในปัจจุบัน [4] เนื่องจากการฝึกแบบจำลอง YOLO จำเป็นต้องใช้ข้อมูลจำนวนมาก คณะผู้วิจัยจึงเลือกใช้ชุดข้อมูล River Water Segmentation Computer Vision Project [22] ในการฝึกโมเดล และนำโมเดลที่ได้มาทดสอบกับชุดข้อมูลทดสอบในรูปแบบ Domain Adaptation โดยชุดข้อมูลที่ใช้ในการฝึกถูกแบ่งออกเป็นข้อมูลสำหรับการฝึก (Training) จำนวน 241 ภาพ และข้อมูลสำหรับการตรวจสอบ (Validation) จำนวน 80 ภาพ โดยขนาดภาพแต่ละภาพมีขนาด 640x640 ตัวอย่างของข้อมูลที่ใช้แสดงใน รูปที่ 3

นอกจากนี้ เพื่อประเมินผลลัพธ์ที่ได้จากโมเดลที่มีความซับซ้อนต่างกัน คณะผู้วิจัยได้ทดลองฝึกแบบจำลอง 2 ขนาด [23] ได้แก่ YOLOv11n-seg (โมเดลขนาดเล็กและมีความซับซ้อนต่ำ) และ YOLOv11x-seg (โมเดลขนาดใหญ่และมีความซับซ้อนสูง) โดยที่โมเดลที่มีความซับซ้อนสูงอาจมีความเสี่ยงที่จะเกิดปัญหา Overfitting ได้ [24] เนื่องจากข้อมูลที่ใช้มีจำนวนจำกัด

รูปที่ 3 ตัวอย่างภาพจากชุดข้อมูล River Water Segmentation Computer Vision Project ที่ใช้สำหรับฝึกแบบจำลอง YOLO11n-seg และ YOLO11x-seg

3.3 กระบวนการดำเนินการและการฝึกแบบจำลอง LRT

ตามที่ได้อธิบายไว้ในหัวข้อ 2.3 แบบจำลอง LRT ที่ใช้ในงานวิจัยนี้มีโครงสร้างเรียบง่าย ภายใต้สมมติฐานว่า Likelihood ของข้อมูลเข้ามีการแจกแจงแบบ Gaussian และมีความแปรปรวนเท่ากันสำหรับทุกคลาส (Equal-Variance Gaussian Distribution)

ในการทดลองนี้ ข้อมูลเข้าของแบบจำลอง LRT คือค่าความเข้ม (Intensity) ของจุดภาพในภาพ Grayscale ภายในบริเวณที่สนใจ (ROI) ซึ่งได้จากภาพสี (RGB) ของกล้อง CCTV โดยผ่านการ Normalize ให้อยู่ในช่วง 0.0 - 1.0 แบบจำลองจะเปรียบเทียบค่าความเข้มแต่ละจุดภาพกับค่า Threshold เพื่อจำแนกเป็น “น้ำ” หรือ “ไม่ใช่น้ำ” และรวมผลลัพธ์เพื่อสร้าง Segment ของภาพที่แบ่งพื้นที่ออกเป็นสองประเภทดังกล่าว

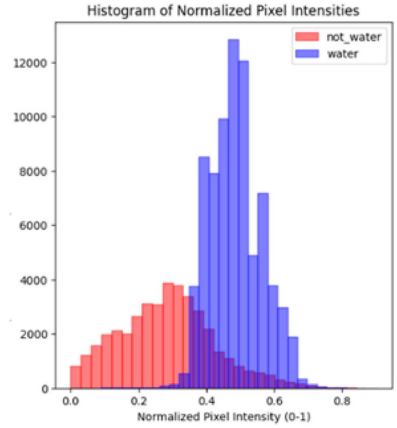
จากการวิเคราะห์ฮิสโตแกรม (Histogram) ของค่าความเข้มของจุดภาพที่ได้จากภาพ p3.jpg (ดูรูปที่ 1) ดังแสดงในรูปที่ 4 พบว่า รูปแบบการแจกแจงของจุดภาพทั้งในพื้นที่น้ำและไม่ใช่น้ำมีลักษณะใกล้เคียงกับการแจกแจงแบบ Gaussian และแม้ว่าค่า Variance ของทั้งสองคลาสอาจไม่เท่ากันโดยสมบูรณ์ แต่ด้วยจำนวนข้อมูลที่จำกัดและเพื่อให้ได้โมเดลที่เรียบง่ายและประมวลผลได้รวดเร็ว คณะผู้วิจัยจึงเลือกใช้สมมติฐาน Equal-Variance ซึ่งสามารถทดสอบความเหมาะสมผ่านผลการประเมินกับชุดข้อมูลทดสอบ

ขั้นตอนการดำเนินการของแบบจำลอง LRT มีดังนี้

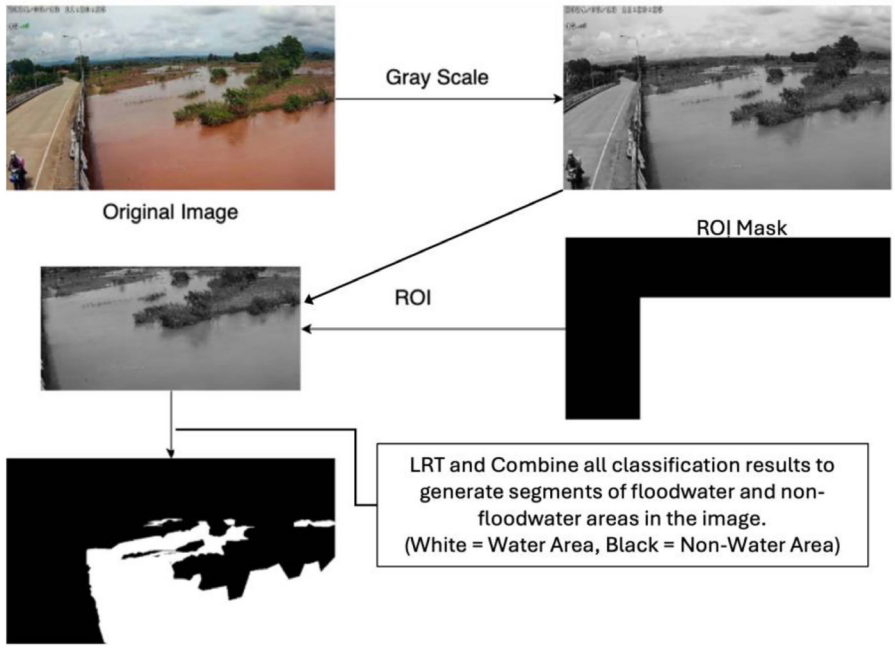
1. แปลงภาพ RGB เป็นภาพ Grayscale และ Normalize ค่าความเข้มของจุดภาพให้อยู่ในช่วง 0.0 - 1.0
2. ระบุบริเวณที่สนใจ (ROI) โดยใช้ ROI Mask ที่กำหนดไว้ล่วงหน้าจากขอบฟ้าและขอบสะพาน โดยจุดภาพสีขาว ใน ROI Mask หมายถึงบริเวณที่สนใจ
3. เปรียบเทียบค่าความเข้มของจุดภาพใน ROI กับค่า Threshold
   - หาก $\geq$ Threshold $\rightarrow$ จำแนกว่าเป็น “น้ำ”
   - หาก $<$ Threshold $\rightarrow$ จำแนกว่าเป็น “ไม่ใช่น้ำ”
4. รวมผลการจำแนกของจุดภาพทั้งหมดเพื่อสร้างแผนภาพแบ่งส่วน (Segmented Image)

ขั้นตอนดังกล่าวได้รับการออกแบบมาโดยเฉพาะเพื่อให้เหมาะสมกับบริบทของข้อมูลในงานวิจัยนี้ (ดูภาพรวมในรูปที่ 5)

แม้แบบจำลอง LRT จะมีโครงสร้างที่ไม่ซับซ้อน แต่สามารถถือเป็นรูปแบบหนึ่งของการเรียนรู้ของเครื่อง (Machine Learning) [25] โดยขั้นตอนการกำหนด Threshold ที่เหมาะสมจัดเป็นส่วนของกระบวนการฝึกโมเดล (Training)



รูปที่ 4 ฮิสโตแกรมของค่าความเข้มของจุดภาพในภาพ Grayscale ที่ถูก Normalize ให้มีค่าในช่วง 0.0 - 1.0 สำหรับพื้นที่น้ำและพื้นที่ที่ไม่ใช่น้ำภายในบริเวณที่สนใจ (ROI) ของภาพ p3.jpg



รูปที่ 5 กระบวนการดำเนินการสำหรับแบบจำลอง LRT

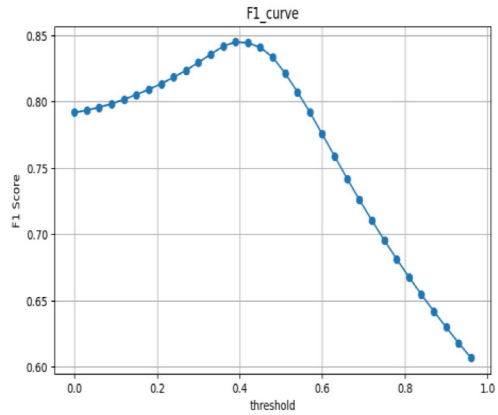
ในการฝึกแบบจำลอง LRT ครั้งนี้ ใช้ภาพ p3.jpg จากกล้องวงจรปิดเป็นข้อมูลฝึก โดยใช้ค่าความเข้มของจุดภาพภายใน ROI และเปรียบเทียบกับ Label ที่กำหนดไว้ล่วงหน้า จากนั้นดำเนินการค้นหา Threshold ที่ให้ค่า F1 Score สูงสุดผ่านกระบวนการ Grid Search โดยมีรายละเอียดดังนี้

1. สร้างชุดค่า Threshold ตั้งแต่ 0.00 ถึง 0.96 โดยเพิ่มทีละ 0.03 ได้แก่ {0.00, 0.03, 0.06, ..., 0.96}
2. สำหรับแต่ละ Threshold ที่กำหนด ดำเนินการจำแนกจุดภาพใน ROI ของภาพ p3.jpg
3. คำนวณ F1 Score และเลือก Threshold ที่ให้ค่าดังกล่าวสูงสุด

4. ผลการทดลอง

4.1 ผลการฝึกแบบจำลอง (Training results)

สำหรับแบบจำลอง LRT ผลการฝึกแสดงในรูปที่ 6 โดยแสดงค่า F1 Score ที่ได้จากการเปรียบเทียบค่าความเข้มของจุดภาพในภาพ p3.jpg กับค่า Threshold ที่แตกต่างกัน (ตามขั้นตอนที่อธิบายไว้ในหัวข้อ 3.3) จากผลการทดลองพบว่า ค่า Threshold ที่ให้ค่า F1 Score สูงสุดคือ 0.39 โดยให้ค่า F1 Score เท่ากับ 84.9% ซึ่งค่า Threshold ดังกล่าวจะถูกนำไปใช้ในการทดสอบแบบจำลองกับภาพอื่น ๆ ในขั้นตอนถัดไป ส่วนการฝึกแบบจำลอง YOLOv11n-seg และ YOLOv11x-seg นั้น ใช้การประเมินผลกับชุดข้อมูล

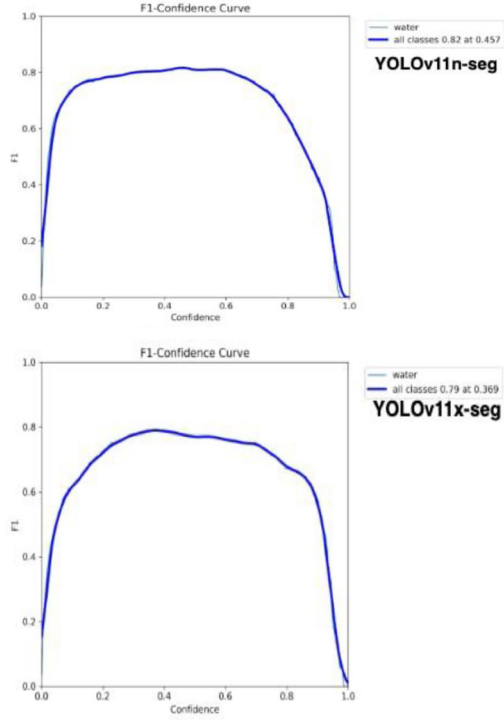


รูปที่ 6 F1 Score ที่ค่า Threshold ต่าง ๆ จากกระบวนการฝึกแบบจำลอง LRT

Validation โดยทดสอบกับค่า Confidence threshold ที่หลากหลายตั้งแต่ 0.0 ถึง 1.0 เพื่อหาค่าที่ให้ผลลัพธ์ F1 Score ที่ดีที่สุด ดังแสดงในรูปที่ 7 โดยมีรายละเอียดดังนี้

- YOLOv11n-seg ให้ค่า F1 Score สูงสุดเท่ากับ 82% เมื่อใช้ค่า Confidence Threshold เท่ากับ 0.457
- YOLOv11x-seg ให้ค่า F1 Score สูงสุดเท่ากับ 79% เมื่อใช้ค่า Confidence Threshold เท่ากับ 0.369

ค่าความเชื่อมั่น (Confidence threshold) ที่ให้ผลลัพธ์สูงสุดเหล่านี้จะถูกนำไปใช้ในการทดสอบแบบจำลองกับชุดข้อมูลทดสอบในขั้นตอนต่อไป



รูปที่ 7 ค่า F1 Score ที่ระดับ Confidence ต่าง ๆ จากการฝึก YOLOv11n-seg และ YOLOv11x-seg

4.2 ผลการทดสอบแบบจำลอง (Testing results)

การประเมินประสิทธิภาพของแบบจำลอง LRT, YOLOv11n-seg และ YOLOv11x-seg ดำเนินการกับชุดข้อมูลทดสอบจำนวน 7 ภาพ จากกล้องวงจรปิด (ดูรูปที่ 1 โดยไม่รวมภาพ p3.jpg) ซึ่งไม่เคยถูกใช้ในการฝึกแบบจำลองหรือกำหนดพารามิเตอร์ใดมาก่อน ทั้งนี้ ค่าคงที่ Threshold สำหรับแบบจำลอง LRT และค่า Confidence threshold ของแบบจำลอง YOLO ทั้งสองเวอร์ชัน ได้มาจากกระบวนการฝึกที่ระบุไว้ในหัวข้อ 4.1

ตัวชี้วัดที่ใช้ในการประเมินผล ได้แก่ Intersection over Union (IoU), Precision, Recall, Accuracy, F1 Score และความเร็วในการประมวลผล (Frames per Second หรือ fps) โดยทำการประมวลผลบน CPU ของ Google Colab และใช้ค่าเฉลี่ยจากผลลัพธ์ของแต่ละภาพเพื่อประเมินภาพรวมของแต่ละโมเดล ซึ่งสรุปไว้ในตารางที่ 2

ผลการทดสอบสามารถสรุปได้ดังนี้

- แบบจำลอง LRT ซึ่งมีโครงสร้างเรียบง่าย ให้ค่า F1 Score สูงถึง 0.84 ซึ่งสูงกว่า YOLOv11x-seg (0.80) และใกล้เคียงกับ YOLOv11n-seg (0.83) ทั้งนี้เป็นผลมาจากการออกแบบโมเดลให้เหมาะสมกับลักษณะเฉพาะของข้อมูลจากกล้อง CCTV ในพื้นที่ศึกษา
- แม้ว่าแบบจำลอง YOLO ทั้งสองเวอร์ชันจะไม่เคยฝึกกับภาพจากกล้องวงจรปิดที่ใช้ในการทดสอบ แต่ก็สามารถรักษาค่า F1 Score ใกล้เคียงกับค่าสูงสุดที่ได้จากชุด Validation ขณะฝึก (YOLOv11n-seg = 0.82 และ YOLOv11x-seg = 0.79) สะท้อนถึงความสามารถในการ Generalize ภายใต้สถานการณ์ Domain shift ได้ดี
- แบบจำลอง LRT แสดงความได้เปรียบอย่างชัดเจนในด้านความเร็ว โดยให้ค่าเฉลี่ย 679.54 fps บน CPU ซึ่งสูงกว่า YOLOv11n-seg (2.45 fps) และ YOLOv11x-seg (0.13 fps) หลายร้อยเท่า

ตารางที่ 2 การเปรียบเทียบประสิทธิภาพของแบบจำลอง LRT, YOLOv11n-seg และ YOLOv11x-seg

Model	IOU	Precision	recall	Accuracy	F1 Score	Frame per second (fps)
LRT (LRT with equal-variance Gaussian assumption)	0.73	0.83	0.85	0.90	0.84	679.54
YOLOv11n-seg	0.70	0.74	0.93	0.88	0.83	2.45
YOLOv11x-seg	0.67	0.73	0.89	0.87	0.80	0.13

- แบบจำลอง YOLOv11n-seg ทำงานได้เร็วกว่า YOLOv11x-seg เนื่องจากมีโครงสร้างที่ซับซ้อนน้อยกว่า และยังให้ค่า F1 Score ที่สูงกว่า ซึ่งอาจเป็นเพราะโครงสร้างที่ซับซ้อนน้อยนั้นเพียงพอต่องานนี้ หรือ YOLOv11x-seg ยังไม่มีข้อมูลฝึกที่มากเพียงพอทำให้เกิด Overfitting

ข้อสรุปจากผลการทดสอบมีดังนี้

- แบบจำลอง LRT เหมาะสำหรับการประยุกต์ใช้ในบริบทเฉพาะของงานวิจัยนี้ โดยเฉพาะในกรณีที่มีข้อจำกัดด้านปริมาณข้อมูลและทรัพยากรประมวลผล อีกทั้งสามารถรองรับการทำงานแบบเรียลไทม์บนอุปกรณ์ขนาดเล็กและต้นทุนต่ำได้อย่างมีประสิทธิภาพ
- ขณะที่แบบจำลอง YOLO ทั้งสองเวอร์ชันแสดงศักยภาพในการ Generalize ได้ดีในสถานการณ์ที่มีความไม่แน่นอนของข้อมูลหรือยังไม่ทราบลักษณะเฉพาะของภาพ จึงเหมาะสำหรับงานที่ต้องการความยืดหยุ่นสูง แม้จะมีข้อจำกัดในด้านความเร็วและทรัพยากรที่ใช้ในการประมวลผลมากกว่า

5. สรุปและวิเคราะห์ผลการทดลอง

จากผลการทดลองในงานวิจัยนี้ พบว่าแบบจำลอง LRT ซึ่งพัฒนาขึ้นจากแนวคิด Likelihood Ratio Test ภายใต้สมมติฐานว่าแต่ละคลาสมีการแจกแจงแบบ Gaussian ที่มีความแปรปรวนเท่ากัน (Equal-Variance Gaussian Distribution) สามารถจำแนกพื้นที่น้ำจากภาพกล้องโทรทัศน์วงจรปิดได้อย่างมีประสิทธิภาพ โดยให้ค่า F1 Score เฉลี่ยสูงถึง 0.84 ซึ่งใกล้เคียงหรือดีกว่าแบบจำลอง YOLOv11n-seg และ YOLOv11x-seg ที่เป็นโมเดล Deep Learning และได้รับการฝึกจากข้อมูลต่างโดเมน (Domain Adaptation)

ข้อได้เปรียบสำคัญของแบบจำลอง LRT คือ โครงสร้างที่เรียบง่าย ใช้พารามิเตอร์เพียงหนึ่งค่า (Threshold) และอาศัยคุณลักษณะพื้นฐานที่สังเกตได้ชัดเจน เช่น ความแตกต่างของค่าสีในภาพ Grayscale ส่งผลให้สามารถประมวลผลได้ด้วยความเร็วสูงมาก โดยให้ค่าเฉลี่ย fps ถึง 679.54 บน CPU ซึ่งเหมาะสำหรับระบบแจ้งเตือนแบบเรียลไทม์ โดยเฉพาะในบริบทที่มีข้อจำกัดด้านทรัพยากรคอมพิวเตอร์และข้อมูลฝึก

ในทางตรงกันข้าม แบบจำลอง YOLOv11n-seg และ YOLOv11x-seg แม้จะมีโครงสร้างที่ซับซ้อนและต้องการข้อมูลฝึกจำนวนมาก แต่แสดงให้เห็นถึงความสามารถในการ Generalize ได้ดี กล่าวคือ แม้ไม่ได้ฝึกจากภาพกล้องวงจรปิดชุดเดียวกันกับที่ใช้ทดสอบ แต่ก็สามารถให้ค่า F1 Score ใกล้เคียงกับผลลัพธ์จากชุด Validation ได้ ซึ่งสะท้อนถึงความเหมาะสมของโมเดล Deep Learning สำหรับสถานการณ์ที่ข้อมูลมีความหลากหลายหรือยังไม่ทราบลักษณะเฉพาะของโดเมน

อย่างไรก็ตาม ข้อจำกัดของแบบจำลอง LRT คือความสามารถในการจำแนกที่จำเพาะต่อบริบทของชุดข้อมูล โดยเฉพาะของข้อมูลที่ศึกษาในงานวิจัยนี้ หากนำไปใช้กับข้อมูลที่มีลักษณะต่างไปจากที่ใช้พัฒนาแบบจำลอง อาจส่งผลให้ความแม่นยำลดลง ในขณะที่แบบจำลอง YOLO สามารถรับมือกับความหลากหลายของข้อมูลและสภาพแวดล้อมที่เปลี่ยนแปลงได้ดีกว่า

โดยสรุป แบบจำลอง LRT เหมาะสำหรับบริบทที่ลักษณะของข้อมูลมีความจำเพาะเจาะจงของงานวิจัยนี้ และระบบมีข้อจำกัดด้านทรัพยากร โดยให้ประสิทธิภาพดีเยี่ยมภายใต้ข้อจำกัดเหล่านั้น ในขณะที่แบบจำลอง YOLO เหมาะกับการใช้งานที่ต้องการความยืดหยุ่นและความสามารถในการปรับตัวต่อข้อมูลใหม่ในบริบทที่ไม่แน่นอนหรือเปลี่ยนแปลงได้บ่อย

6. กิตติกรรมประกาศ

บทความวิจัยฉบับนี้เป็นส่วนหนึ่งของผลผลิตจากโครงการการต่อยอดระบบจำลองสถานการณ์น้ำท่วมด้วยข้อมูลที่ทันสมัยและระบบเตือนภัย ซึ่งได้รับการสนับสนุนทุนอุดหนุนจากสำนักงานการวิจัยแห่งชาติ (วช.) ประจำปีงบประมาณ พ.ศ. 2568 ด้านแผนงานวิจัยและนวัตกรรม สิ่งประดิษฐ์ เพื่อแก้ปัญหาและผลกระทบจากภัยพิบัติทางธรรมชาติและการเปลี่ยนแปลงสภาพภูมิอากาศ ตามสัญญาเลขที่ N81A680725 และคณะผู้วิจัยขอขอบคุณ สถาบันเทคโนโลยีป้องกันประเทศ มหาวิทยาลัยเกษตรศาสตร์ มหาวิทยาลัยเชียงใหม่ และมหาวิทยาลัยพะเยา ในฐานะสถาบันต้นสังกัดของนักวิจัยร่วมในโครงการนี้

7. เอกสารอ้างอิง

[1] R. Chityala, K. R. Hoffmann, D. R. Bednarek, and S. Rudin, "Region of Interest (ROI) Computed Tomography," vol. 5368, no. 2, *SPIE,* pp. 534 - 541, 2004.

[2] S. M. Kay, *Fundamentals of Statistical Signal Processing, Vol. II: Detection Theory.* Upper Saddle River, NJ, USA: Prentice Hall PTR, 1998.

[3] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *2016 IEEE Conf. Comput. Vision Pattern Recognit. (CVPR),* Jun. 2016, pp. 779 - 788, doi: 10.1109/CVPR.2016.91.

[4] L. He, Y. Zhou, L. Liu, and J. Ma, "Research and Application of YOLOv11-Based Object Segmentation in Intelligent Recognition at Construction Sites," *Buildings,* vol. 14, no. 12, p. 3777, 2024, doi: 10.3390/buildings14123777.

[5] J. Terven, D. - M. Cordova-Esparza, and J. - A. Romero-González, "A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS," *Mach. Learn. Knowl. Extr.*, vol. 5, no. 4, pp. 1680 - 1716, 2023, doi: 10.3390/make5040083.

[6] D. A. Roberts, S. Yaida, and B. Hanin, *The Principles of Deep Learning Theory.* NY, USA: Cambridge University Press, 2022.

[7] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image Segmentation Using Deep Learning: A Survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 7, pp. 3523 - 3542, 2021.

[8] C. Kanan and G. W. Cottrell, "Color-to-Grayscale: Does the Method Matter in Image Recognition?," *PloS ONE,* vol. 7, no. 1, p. e29740, 2012.

[9] *Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide-Screen 16:9 Aspect Ratios,* Recommendation ITU-R BT.601-7, Mar. 2011. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.601-7-201103-I!!PDF-E.pdf

[10] L. A. Curtiss, P. C. Redfern, and K. Raghavachari, "Gaussian 4 Theory," *J. Chem. Phys.*, vol. 126, no. 8, p. 084108, Mar. 2007, doi: 10.1063/1.2436888.

[11] G. Shafer, "Conditional Probability," *Int. Stat.* Rev., vol. 53, no. 3, pp. 261-275, 1985, doi: 10.2307/1402890.

[12] N. A. Muhadi, A. F. Abdullah, S. K. Bejo, M. R. Mahadi, and A. Mijic, "Deep Learning Semantic Segmentation for Water Level Estimation Using Surveillance Camera," *Appl. Sci.*, vol. 11, no. 20, p. 9691, 2021.

[13] L. Lopez-Fuentes, C. Rossi, and H. Skinnemoen, "River Segmentation for Flood Monitoring," in *2017 IEEE Int. Conf. Big Data (Big Data),* Boston, MA, USA, 2017, pp. 3746 - 3749, doi: 10.1109/BigData.2017.8258373.

[14] Y. Wang *et al.*, "Urban Flood Extent Segmentation and Evaluation from Real-World Surveillance Camera Images Using Deep Convolutional Neural Network," *Environ. Model. Softw.*, vol. 173, p. 105939, 2024.

[15] J. Wan *et al.*, "Automatic Segmentation of Urban Flood Extent in Video Image with DSS-YOLOv8n," *J. Hydrol.*, vol. 655, p. 132974, 2025.

[16] C. Sazara, M. Cetin, and K. M. Iftekharuddin, "Detecting Floodwater on Roadways from Image Data with Handcrafted Features and Deep Transfer Learning," in *2019 IEEE Intell. Transp. Syst. Conf. (ITSC),* Auckland, New Zealand, 2019, pp. 804 - 809.

[17] M. G. Ragab *et al.*, "A Comprehensive Systematic Review of YOLO for Medical Object Detection (2018 to 2023)," *IEEE Access,* vol. 12, pp. 57815 - 57836, 2024.

[18] W. Luangluewut, P. Thiennviboon, K. Viriyasatr, P. Maleecharoen, and T. Kasetkasem, "Investigating Domain Adaptation Feasibility for Drone Detection: A CPU-Based YOLO Approach Using RGB-Infrared Images," in *2025 Int. Conf. Artif. Intell. Inf. Commun. (ICAIIC),* Fukuoka, Japan, 2025, pp. 0926 - 0931, doi: 10.1109/ICAIIC64266.2025.10920826.

[19] P. Akiva, M. Purri, K. Dana, B. Tellman, and T. Anderson, "H2O-Net: Self-Supervised Flood Segmentation via Adversarial Domain Adaptation and Label Refinement," *2021 IEEE Winter Conf. Appl. Comput. Vision (WACV),* 2021, pp. 111 - 122, doi: 10.1109/WACV48630.2021.00016.

[20] S. Visa, B. Ramsay, A. Ralescu, and E. van der Knaap, "Confusion Matrix-based Feature Selection," *CEUR Workshop Proc.*, vol. 710, pp. 120 - 127, 2011.

[21] Ultralytics YOLO Docs. "Instance Segmentation." DOC.ULTRALYTICS.com. https://docs.ultralytics.com/tasks/segment (accessed Mar. 23, 2025).

[22] A. Hisyam, "River Water Segmentation Dataset." 2024. Distributed by Roboflow.

[23] M. Tan, R. Pang, and Q. V. Le, "Efficient Det: Scalable and Efficient Object Detection," in *2020 IEEE/CVF Conf. Comput. Vision Pattern Recognit. (CVPR),* Seattle, WA, USA, 2020, pp. 10778 - 10787, doi: 10.1109/CVPR42600.2020.01079.

[24] A. Safonova, G. Ghazaryan, S. Stiller, M. Main-Knorn, C. Nendel, and M. Ryo, "Ten Deep Learning Techniques to Address Small Data Problems with Remote Sensing," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 125, p. 103569, 2023, doi: 10.1016/j.jag.2023.103569.

[25] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification,* 2nd ed., NY, USA: Wiley-Interscience Publication, 2001.