

การพัฒนาเว็บแอปพลิเคชันประมาณค่าความขุ่นของน้ำ โดยใช้เทคโนโลยีการประมวลผลภาพ

Web Application Development for Water Turbidity Estimation using Image Processing Technology

สิรินยา องค์กรแห่ง
Sirinya Ongkhamhaeng
คณะวิทยาศาสตร์และเทคโนโลยี
มหาวิทยาลัยหอการค้าไทย
กรุงเทพมหานคร / ประเทศไทย
sry.ongkh@gmail.com

ฉัตรชวิน เจตะภัย
Chatchwin Jatapai
คณะวิทยาศาสตร์และเทคโนโลยี
มหาวิทยาลัยหอการค้าไทย
กรุงเทพมหานคร / ประเทศไทย
chatchwinj@gmail.com

ภูมินทร์ พานทอง
Pumin Panthong
คณะวิทยาศาสตร์และเทคโนโลยี
มหาวิทยาลัยหอการค้าไทย
กรุงเทพมหานคร / ประเทศไทย
puminpanthong@gmail.com

เอกรินทร์ วรุตบางกูร
Ekarin Varutbangkul
คณะวิทยาศาสตร์และเทคโนโลยี
มหาวิทยาลัยหอการค้าไทย
กรุงเทพมหานคร / ประเทศไทย
ekarin_var@utcc.ac.th

ปิยะวรรณ คอนาแฮน
Piyawan Conahan
คณะวิทยาศาสตร์และเทคโนโลยี
มหาวิทยาลัยหอการค้าไทย
กรุงเทพมหานคร / ประเทศไทย
piyawan.kas@gmail.com

บทคัดย่อ

น้ำมีความสำคัญต่อการดำรงชีวิตของมนุษย์ รวมถึงงานด้านเกษตรกรรมและอุตสาหกรรม ความขุ่นของน้ำถือเป็นลักษณะทางกายภาพรูปแบบหนึ่งที่ใช้เป็นดัชนีตรวจสอบคุณภาพน้ำ ซึ่งเป็นปัจจัยต่อทัศนคติในการเลือกอุปโภคบริโภคและผลกระทบต่อระบบกรองน้ำ ปัจจุบันการตรวจสอบความขุ่นของน้ำทำได้โดยใช้เครื่องวัดความขุ่น หรือการเก็บตัวอย่างน้ำไปทดสอบในห้องปฏิบัติการ ซึ่งอาจต้องใช้ระยะเวลาและค่าใช้จ่ายสูง การประยุกต์ใช้เทคโนโลยีการประมวลผลภาพ [1] ในการประมาณค่าความขุ่นของน้ำจึงเป็นทางเลือกหนึ่งในการตรวจสอบคุณภาพน้ำ งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของเทคนิคการเรียนรู้ของเครื่องเพื่อการประเมินความขุ่นของน้ำจากภาพ 3 เทคนิค คือ โครงข่ายประสาทเทียม, เพื่อนบ้านใกล้สุด และการวิเคราะห์การถดถอยแบบโพลิโนเมียล การเก็บข้อมูลภาพจากตัวอย่างน้ำสร้างจากสารแขวนลอย เช่น ครีมน้ำตาล ดิน และปุ๋ย ในปริมาณที่แตกต่างกัน จากการทดลองพบว่าวิธีการวิเคราะห์การถดถอยแบบโพลิโนเมียลให้ผลลัพธ์ค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด และนำมาใช้เป็นโมเดลประมาณค่าความขุ่นน้ำภายใต้เว็บแอปพลิเคชันที่พัฒนาขึ้น

คำสำคัญ: ความขุ่นของน้ำ, เทคโนโลยีการประมวลผลภาพ, การเรียนรู้ของเครื่อง, เว็บแอปพลิเคชัน

ABSTRACT

Water is essential to human life and is also important in agriculture and industry. One way to measure water quality and ecosystem health is to examine water turbidity. At present, it is measured using an instrument called a turbidimeter or turbidity sensor, or through an expensive and time-

consuming lab test. However, image processing for water turbidity estimation is an emerging technology that may prove to be both faster and cheaper than the alternatives. The objective of this research is to compare the efficacy of the following three machine learning techniques in estimating water turbidity using images: (1) Artificial Neural Network, (2) k-Nearest Neighbors and (3) Polynomial Regression Analysis. Water sample data was collected by creating a suspension using varying amounts of shower cream, soil and fertilizer. The experiment showed that the polynomial regression analysis yielded the least mean squared error and was used as a water turbidity estimation model under the developed web application.

Keywords: Water Turbidity, Image Processing, Machine Learning, Web Application

I. บทนำ

น้ำเป็นสารประกอบที่พบมากถึง 3 ใน 4 ส่วนของพื้นโลกที่มีความสำคัญมากทั้งด้านอุตสาหกรรม เกษตรกรรม การอุปโภคและบริโภคของมนุษย์ ลักษณะทางกายภาพของน้ำเป็นตัวชี้วัดหนึ่งที่ใช้วัดคุณภาพของน้ำที่สำคัญ คุณภาพของน้ำทางกายภาพ ได้แก่สิ่งเจือปนที่มีลักษณะที่สามารถรับรู้ได้ด้วยประสาทสัมผัส เช่น กลิ่น สี รส และความขุ่น งานวิจัยนี้มีความสนใจในการศึกษาเรื่องความขุ่นของน้ำ (Turbidity) เนื่องจากความขุ่นมีความสำคัญต่อทัศนคติในการเลือกอุปโภค บริโภคของผู้ใช้น้ำ มีผลกระทบต่อระบบการกรองน้ำ อุปกรณ์ชำระเร็วและประสิทธิภาพการฆ่าเชื้อโรคด้วยคลอรีนลดลง รวมถึงมีผลกระทบต่อสิ่งมีชีวิตในน้ำและสิ่งแวดล้อม

ความขุ่นของน้ำเกิดจากอินทรีย์และอนินทรีย์สารในน้ำ สารแขวนลอย เช่น ดิน โคลน หินทรายละเอียด สารละลายเกลือ โดยสิ่งเหล่านี้สามารถทำให้แสงเกิดการหักเหและอาจดูดซับปริมาณแสงที่ส่องผ่านเอาไว้ได้ หน่วยมาตรฐานที่ใช้ในการวัดค่าความขุ่นของน้ำคือ NTU (Nephelometric Turbidity Unit) เกณฑ์คุณภาพน้ำประปาของกรมอนามัย ปี พศ. 2553 [2] กำหนดให้ค่าความขุ่นมาตรฐานไม่เกิน 5 NTU หากเกินจากนี้ น้ำจะมีลักษณะขุ่นที่เห็นได้ชัดเจนขึ้น ช่วงความขุ่นของน้ำประปาจะอยู่ในช่วง 0 – 400 NTU น้ำเสียอาจจะมีความสูงถึง 1000 NTU ดังนั้นการตรวจวัดความขุ่นมีความสำคัญอย่างยิ่งต่อการปกป้องสิ่งแวดล้อมทั้งต่อสัตว์น้ำ สิ่งแวดล้อม ระบบนิเวศและน้ำดื่มที่สะอาดปลอดภัย

การตรวจวัดความขุ่นมีได้หลายวิธี เช่นการใช้เครื่องวัดความขุ่น-ใสในน้ำ หรือการตรวจสอบในห้องปฏิบัติการ ข้อดีของการตรวจสอบในห้องปฏิบัติการคือมีความแม่นยำที่สูง แต่ข้อเสียคือราคาแพง ต้องตรวจสอบโดยผู้เชี่ยวชาญเท่านั้น และใช้เวลานาน ซึ่งอาจจะไม่สะดวกสำหรับผู้ทั่วไปที่ต้องการตรวจวัดความขุ่น หรืออาจจะเป็นไปได้กับพื้นที่ตัวอย่างน้ำที่มีขนาดใหญ่ ด้วยความก้าวหน้าของเทคโนโลยีในปัจจุบัน และการใช้เทคโนโลยีการประมวลผลภาพ (Image Processing) [2] เข้ามาประยุกต์ในงานด้านต่างๆ เช่นการจดจำใบหน้า การระบุชนิดของต้นไม้ และการตรวจหาคลอโรฟิลล์และสารตกตะกอนในน้ำ จึงทำให้คณะผู้วิจัยเกิดแรงบันดาลใจในการประยุกต์ใช้เทคโนโลยีการประมวลผลภาพในการประมาณค่าความขุ่นของน้ำ เพื่อเป็นอีกทางเลือกหนึ่งในการตรวจวัดความขุ่นผ่านทางเว็บแอปพลิเคชัน

วัตถุประสงค์ของงานวิจัย คือ (1) การสร้างโมเดลการประมาณค่าความขุ่นของน้ำจากเทคโนโลยีการประมวลผลภาพ และ (2) การพัฒนาเว็บไซต์ประมาณค่าความขุ่นของน้ำจากโมเดลที่พัฒนาขึ้น

II. งานวิจัยที่เกี่ยวข้อง

Cheng En และคณะ [3] ได้พัฒนาระบบตรวจวัดและประเมินน้ำขุ่นอัตโนมัติโดยใช้เทคโนโลยีการประมวลผลภาพ ด้วยเหตุผลหลายประการในประเทศจีนที่ตัดสินใจลงทุนเงินกับสิ่งประดิษฐ์ด้านการประปาเป็นจำนวนมาก เพื่อแก้ไขปัญหาเรื่องระบบนิเวศทางน้ำ และการอุปโภคบริโภคของประชาชน ด้วยเหตุนี้จึงใช้การประมวลผลภาพดิจิทัล และเทคโนโลยีการจดจำรูปแบบบันทึกลักษณะของสารส้มในน้ำประปาจากภาพ โดยเปรียบเทียบผลกระทบของอัลกอริทึมหลายๆแบบ แล้วนำไปใช้กับ EmguCV ซึ่งเป็นไลบรารีที่ใช้ช่วยในการเขียนโปรแกรมเกี่ยวกับ computer vision แบบ real-time แล้วเลือกคุณสมบัติและจัดประเภทด้วย LibSVM ผลการวิจัยพบว่าระบบที่ออกแบบมีความเหมาะสมผล และเป็นไปได้

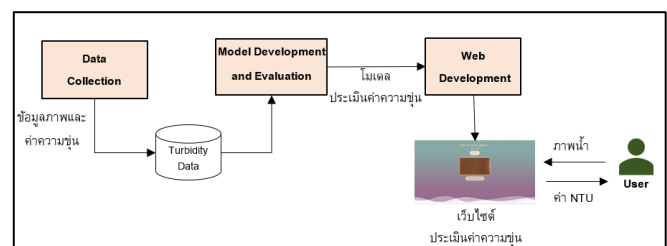
V. Karnawat และคณะ [4] ได้พัฒนาเทคนิคการตรวจสอบความขุ่นของตัวอย่างน้ำโดยใช้วิธีการประมวลผลภาพ การตรวจสอบความขุ่นโดยทั่วไปใช้เวลานาน มีข้อผิดพลาดเล็กน้อยๆ และทำซ้ำได้ลำบาก งานวิจัยนี้ได้แนะนำเทคนิคอันประกอบไปด้วยการเก็บภาพตัวอย่างน้ำด้วยกล้องความละเอียดสูง เตรียมภาพให้มีความขุ่นที่ต่างกันแล้วทำการแปลงภาพเป็นไบนารีเพื่อเกณฑ์เป็นค่า 0 แทนน้ำส่วนที่ใส และ 1 แทนน้ำส่วนที่มีความขุ่น และฐานข้อมูลที่มีข้อมูลของตัวอย่างน้ำที่มีค่าความขุ่นจากห้องปฏิบัติการสามารถตรวจสอบได้ทุกที่ทุกเวลา ผลลัพธ์ของการ

ประมวลผลภาพคือ เมื่อเปรียบเทียบค่า NTU จากเทคนิคการประมวลผลภาพ กับฐานข้อมูลทั้งหมด 3 ภาพ ผลปรากฏว่าค่อนข้างใกล้เคียงกับค่าในฐานข้อมูล ถือว่าแม่นยำภายในระยะเวลาที่รวดเร็ว

YEQI LIU และคณะ [5] ได้สรุปวิธีการตรวจสอบความขุ่นน้ำโดยใช้ภาพจากนักวิจัยหลายๆ กลุ่ม การศึกษาตัวอย่างน้ำแบ่งออกเป็น 4 ประเภทจากแหล่งที่มาที่ต่างกัน ได้แก่ ภาพตัวอย่างน้ำทั่วไป (Sample image) ภาพใต้น้ำ (Under water image) ภาพบนผิวน้ำ (Water surface image) และภาพที่ใช้แสงที่มองไม่เห็น (Invisible light image) นักวิจัยมีการใช้ปัจจัยหลากหลายเช่น คุณภาพของรูป การสกัดตัวแปรของรูป การเลือกแบบจำลอง และอื่นๆ ดังนั้นผู้วิจัยจึงทำการเปรียบเทียบและกำหนดขอบเขตเพื่อค้นหาการสกัดตัวแปร และวิธีการตรวจสอบความขุ่น ผลการวิจัยพบว่า ค่า NTU อยู่ในช่วง 1 – 500 NTU และแบบจำลองความสัมพันธ์แบบพหุนาม (Polynomial relation) มีความคลาดเคลื่อนต่ำที่สุด คือ 2.8%

III. วิธีดำเนินงานวิจัย

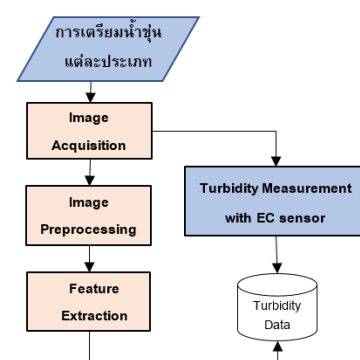
งานวิจัยนี้มีขั้นตอนดำเนินงานวิจัยทั้งหมด 3 ขั้นตอนหลัก คือ (1) การเก็บรวบรวมข้อมูล (Data Collection), (2) การพัฒนาและการประเมินโมเดล (Model Development and Evaluation) และ (3) การพัฒนาเว็บแอปพลิเคชัน (Web Development) แสดงดังรูปที่ 1. รายละเอียดของแต่ละขั้นตอนจะอธิบายในหัวข้อต่อไป



รูปที่ 1. ภาพรวมของวิธีการดำเนินงานวิจัย

III.1 การเก็บรวบรวมข้อมูล (Data Collection)

วัตถุประสงค์ของขั้นตอนนี้ คือ การสร้างข้อมูลตัวอย่างน้ำขุ่นและน้ำใส พร้อมทั้งค่าความขุ่นในหน่วย NTU เพื่อใช้ในการพัฒนาโมเดลประเมินความขุ่นน้ำด้วยการเรียนรู้แบบมีผู้สอน (Supervised Learning) ขั้นตอนการเก็บรวบรวมข้อมูลแสดงดังรูปที่ 2.



รูปที่ 2. ขั้นตอนการเก็บรวบรวมข้อมูล

- การเตรียมน้ำขุ่น คณะผู้วิจัยได้ออกแบบการสร้างตัวอย่างน้ำที่มีความขุ่นต่างกัน 4 ประเภท คือ (1) น้ำใส, (2) น้ำผสมครีมอาบน้ำ, (3) น้ำผสมดิน และ (4) น้ำผสมปุ๋ย ซึ่งแต่ละประเภทมีการใช้ปริมาตรของส่วนผสมที่ต่างกันต่อน้ำปริมาตร 7 ลิตร เพื่อให้ได้ความขุ่นของน้ำที่แตกต่างกัน 16 แบบ ดังแสดงในตารางที่ I

ตารางที่ I. แสดงสัดส่วนของการสร้างตัวอย่างน้ำทั้ง 16 แบบ

แบบที่	ประเภทน้ำ	ปริมาตรของส่วนผสมต่อน้ำ 7 ลิตร
1	น้ำเปล่า	-
2	น้ำเปล่าผสมครีมอาบน้ำ	4 มิลลิลิตร
3		8 มิลลิลิตร
4		12 มิลลิลิตร
5		16 มิลลิลิตร
6		20 มิลลิลิตร
7	น้ำเปล่าผสมดิน	20 กรัม
8		40 กรัม
9		60 กรัม
10		80 กรัม
11		100 กรัม
12	น้ำเปล่าผสมปุ๋ย	15 กรัม
13		30 กรัม
14		45 กรัม
15		60 กรัม
16		75 กรัม

- Image Acquisition เป็นการถ่ายภาพตัวอย่างน้ำที่สร้างขึ้นทั้ง 16 แบบ โดยใช้อุปกรณ์เป็นตู้เลี้ยงปลาขนาดกว้าง 17 เซนติเมตร ยาว 26 เซนติเมตร และสูง 19 เซนติเมตร บรรจุน้ำได้ 7 ลิตร อุปกรณ์ที่ใช้ในการถ่ายภาพ คือ กล้องของ Smartphone รุ่น Samsung S10 โดยมีระยะห่างระหว่างตู้เลี้ยงปลากับกล้องระหว่าง 20-30 เซนติเมตร ใช้แสงธรรมชาติในการถ่ายภาพ และกำหนดขนาดภาพเท่ากับ 1920x1080 พิกเซล โดยเริ่มต้นจากการทดลองถ่ายภาพจากมุมมองด้านข้าง และมุมมองด้านบน (ตัวอย่างแสดงดังรูปที่ III) จำนวนแบบละ 10 ภาพ เพื่อนำมาทำการประเมินประสิทธิภาพเบื้องต้นด้วยเทคนิคการเรียนรู้ทั้ง 3 เทคนิค และพบว่าภาพถ่ายจากมุมมองด้านบนให้ผลลัพธ์ค่าความคลาดเคลื่อนของการประมาณค่าที่น้อยกว่า (ดีกว่า) ภาพถ่ายจากมุมมองด้านข้าง ซึ่งเป็นจริงทั้ง 3 เทคนิค ดังแสดงในตารางที่ 2



มุมมองภาพถ่ายด้านข้าง



มุมมองภาพถ่ายด้านบน

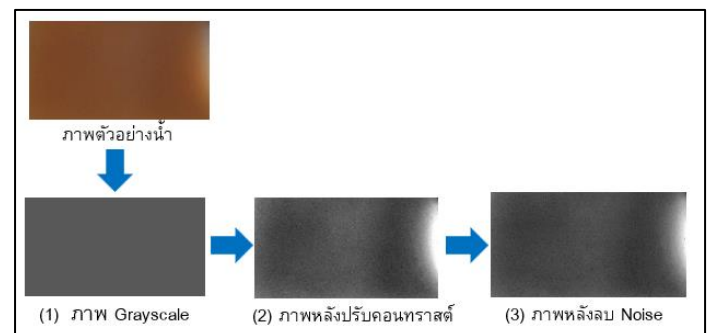
รูปที่ III. ตัวอย่างมุมมองการถ่ายภาพทั้ง 2 มุมมอง

ตารางที่ II. แสดงประสิทธิภาพด้วยเทคนิคการเรียนรู้ทั้ง 3 เทคนิคบนภาพถ่ายมุมมองด้านบนและมุมมองด้านข้าง

แหล่งที่มาของภาพ	ค่า Root Mean Square Error (RMSE) ของชุดข้อมูล		
	โมเดลโครงข่ายประสาทเทียม	โมเดลเพื่อนบ้านใกล้เคียง (K=3)	โมเดลการวิเคราะห์การถดถอยโพลีโนเมียลพหุคูณ
ภาพด้านบน	13.7827	9.4747	9.5376
ภาพด้านข้าง	17.2390	10.3327	10.8570

จากนั้นคณะผู้วิจัยดำเนินการเก็บข้อมูลภาพถ่ายมุมมองด้านบนจากตัวอย่างน้ำทั้ง 16 แบบ แบบละ 600 ภาพ ดังนั้นจำนวนรูปภาพที่นำมาใช้ในการสร้าง Turbidity Data มีทั้งหมด 9,600 ภาพ และนำไปใช้ในสกัดข้อมูลภาพ และเก็บข้อมูลความขุ่นของน้ำต่อไปด้วยเซ็นเซอร์

- Image Preprocessing เป็นขั้นตอนการเตรียมภาพก่อนนำไปสกัดคุณลักษณะเฉพาะของภาพ การเตรียมภาพประกอบด้วย 3 ขั้นตอนตามลำดับ คือ (1) การเปลี่ยนภาพต้นฉบับให้เป็น Grayscale, (2) การปรับ Contrast ด้วยไลบรารี Pillow ของภาษา Python ในการปรับแต่งความเข้มของภาพ, (3) การลบ Noise ของภาพด้วยตัวกรองสัญญาณรบกวนแบบ Minimum Filter ของไลบรารี Pillow ตัวอย่างการทำ Image Preprocessing แสดงดังรูปที่ IV.



รูปที่ IV. ขั้นตอนการทำ Image Preprocessing

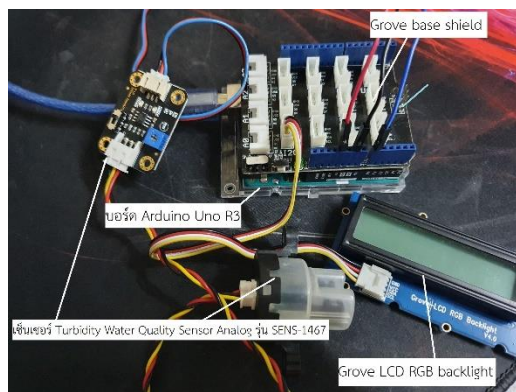
- Feature Extraction เป็นขั้นตอนสกัดคุณลักษณะของภาพให้อยู่ในรูปแบบที่นำไปใช้งานกับการเรียนรู้ของเครื่อง (Machine Learning) คุณลักษณะที่สกัดจากภาพ ประกอบด้วย 12 ตัวแปรที่มาจากค่าแนะนำตัวแปรจากวิธีการของภาพถ่ายตัวอย่างน้ำของงานวิจัย [5] และทำการเพิ่มเติมตัวแปรต่างๆ ดังนี้ ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน ของตัวแปรต่าง ๆ (grayscale, red, green, blue, luminance) ค่าสูงสุด และค่าต่ำสุดของ grayscale คำอธิบายของแต่ละตัวแปรแสดงดังตารางที่ III.

ตารางที่ III. แสดงตัวแปรจากการสกัดตัวแปรของภาพและคำอธิบาย

ชื่อตัวแปร	คำอธิบาย
max_gry	ค่าสูงสุดของค่า grayscale
min_gry	ค่าต่ำสุดของค่า grayscale
mean_gry	ค่าเฉลี่ยของค่า grayscale

ชื่อตัวแปร	คำอธิบาย
std_gry	ค่าเบี่ยงเบนมาตรฐานของค่า grayscale
mean_lu	ค่าเฉลี่ยของค่าแสง
std_lu	ค่าเบี่ยงเบนมาตรฐานของค่าแสง
mean_r	ค่าเฉลี่ยของค่าสีแดง
std_r	ค่าเบี่ยงเบนมาตรฐานของค่าสีแดง
mean_g	ค่าเฉลี่ยของค่าสีเขียว
std_g	ค่าเบี่ยงเบนมาตรฐานของค่าสีเขียว
mean_b	ค่าเฉลี่ยของค่าสีน้ำเงิน
std_b	ค่าเบี่ยงเบนมาตรฐานของค่าสีน้ำเงิน

- Turbidity Measurement เป็นการวัดค่าความขุ่นของน้ำในหน่วย NTU สำหรับภาพตัวอย่างแต่ละภาพ อุปกรณ์ที่ใช้ คือ Analog Turbidity Sensor Module เซ็นเซอร์วัดความขุ่นในน้ำ โดยวัดค่าแสงที่หักเหในน้ำ และให้ผลลัพธ์ออกมาเป็นค่า Analog และ Digital รูปที่ V. แสดงภาพอุปกรณ์ที่ใช้ในการวัดค่าความขุ่นน้ำ



รูปที่ V: อุปกรณ์วัดค่าความขุ่นของภาพตัวอย่างน้ำ

ผลลัพธ์ค่าความขุ่นที่ได้จะถูกเพิ่มเข้าไปใน Turbidity Data ในส่วนของ Output เพื่อใช้เป็นข้อมูลฝึกฝนสำหรับการพัฒนาโมเดลประเมินค่าความขุ่นของน้ำ ตารางที่ IV. แสดงรายละเอียดช่วงค่าความขุ่นของน้ำในหน่วย NTU ที่มีส่วนผสมเหมือนกัน วัดในเวลาที่แตกต่างกัน ค่าความขุ่นเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของค่าความขุ่น

ตารางที่ IV. แสดงรายละเอียดค่าความขุ่นของน้ำทั้ง 16 แบบ

แบบที่	ประเภทน้ำ	ช่วงค่าความขุ่น	ค่าความขุ่นเฉลี่ย	ค่าเบี่ยงเบนมาตรฐาน
1	น้ำเปล่า	0-1	0.03	0.17
2	น้ำเปล่าผสมครีมอาบ น้ำ 4 มิลลิลิตร	3-4	3.22	0.41
3	น้ำเปล่าผสมครีมอาบ น้ำ 8 มิลลิลิตร	8-10	8.76	0.63
4	น้ำเปล่าผสมครีมอาบ น้ำ 12 มิลลิลิตร	14-16	14.53	0.68
5	น้ำเปล่าผสมครีมอาบ น้ำ 16 มิลลิลิตร	20-23	20.97	0.97
6	น้ำเปล่าผสมครีมอาบ น้ำ 20 มิลลิลิตร	24-25	24.48	0.73

แบบที่	ประเภทน้ำ	ช่วงค่าความขุ่น	ค่าความขุ่นเฉลี่ย	ค่าเบี่ยงเบนมาตรฐาน
7	น้ำเปล่าผสมดิน 20 กรัม	24-26	34.97	1.25
8	น้ำเปล่าผสมดิน 40 กรัม	54-59	56.36	1.35
9	น้ำเปล่าผสมดิน 60 กรัม	67-70	67.87	0.80
10	น้ำเปล่าผสมดิน 80 กรัม	76-79	77.37	1.16
11	น้ำเปล่าผสมดิน 100 กรัม	85-87	85.79	0.84
12	น้ำเปล่าผสมปุ๋ย 15 กรัม	24-26	24.97	0.72
13	น้ำเปล่าผสมปุ๋ย 30 กรัม	45-47	45.89	0.79
14	น้ำเปล่าผสมปุ๋ย 45 กรัม	63-66	64.38	1.28
15	น้ำเปล่าผสมปุ๋ย 60 กรัม	69-72	70.22	1.22
16	น้ำเปล่าผสมปุ๋ย 75 กรัม	77-82	79.93	1.67

III. II การพัฒนาและการประเมินโมเดล

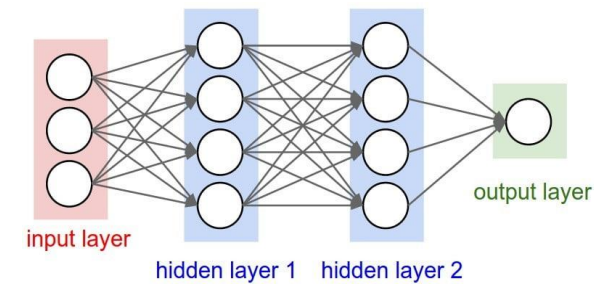
(Model Development and Evaluation)

การวิจัยครั้งนี้ได้แบ่งข้อมูล Turbidity Data ที่ได้จากขั้นตอนการเก็บรวบรวมข้อมูลออกเป็น 2 ส่วนแบบสุ่ม โดยส่วนที่ 1 ร้อยละ 67 ของข้อมูลทั้งหมดเป็นชุดข้อมูลฝึกฝนสำหรับการสร้างโมเดล และส่วนที่ 2 ร้อยละ 33 ของข้อมูลทั้งหมดเป็นชุดข้อมูลทดสอบสำหรับการทดสอบและประเมินสมรรถนะของโมเดล เครื่องมือที่ใช้ในการวิจัยคือ ไลบรารีของ Scikit-Learn สำหรับการเรียนรู้ของเครื่องด้วยภาษา Python คณะผู้วิจัยเลือกใช้เทคนิคในการเรียนรู้โมเดลสำหรับการประเมินค่าความขุ่นของน้ำจำนวน 3 เทคนิค คือ การวิเคราะห์การถดถอยแบบพหุนาม (Polynomial Regression), โครงข่ายประสาทเทียม (Artificial Neural Network) และ เพื่อนบ้านใกล้สุด (k-Nearest Neighbors) เนื่องจาก 3 เทคนิคดังกล่าวเป็นเทคนิคที่นิยมใช้สำหรับการประมาณค่า โดยการวิเคราะห์การถดถอย และเพื่อนบ้านใกล้สุด เป็นเทคนิคที่เข้าใจได้ง่ายไม่ซับซ้อนและใช้เวลาในการประมวลผลไม่มากนัก ส่วนโครงข่ายประสาทเทียมแม้เป็นเทคนิคที่ซับซ้อนและใช้เวลาในการประมวลผลนานกว่าแต่โมเดลที่ได้มักมีสมรรถนะที่ดีและนิยมประยุกต์ใช้ในงานหลากหลายด้าน

การเตรียมข้อมูล (Data Preprocessing) ก่อนการนำข้อมูลไปสร้างโมเดลผู้วิจัยมีการเตรียมข้อมูลโดยการทำ Normalization เพื่อแปลงข้อมูลให้ทุกตัวแปรมีค่าอยู่ในช่วงสเกลเดียวกันคือ 0 ถึง 1 ซึ่งเป็นวิธีการเตรียมข้อมูลที่เหมาะสมสำหรับเทคนิคโครงข่ายประสาทเทียม เทคนิคเพื่อนบ้านใกล้สุด และเทคนิคการวิเคราะห์การถดถอยแบบพหุนาม

- โครงข่ายประสาทเทียม (ใช้คลาส MLPRegressor ของไลบรารี neural_network)

โครงข่ายประสาทเทียม [6] เป็นเทคนิคที่เกิดจากการพยายามจำลองให้เครื่องคอมพิวเตอร์ทำงานคล้ายกับสมองของสิ่งมีชีวิตโดยการจำลองเป็นโครงข่ายของหน่วยประมวลผลจำนวนมากที่มีการเชื่อมโยงกันเหมือนกับโครงข่ายของเซลล์ประสาทในสมอง (neurons) โดยหน่วยประมวลผลเหล่านี้มีการเชื่อมโยงกันเป็นโครงข่ายหลายระดับชั้น (multilayer network) ซึ่งแต่ละระดับชั้นสามารถแบ่งออกได้เป็น 3 ประเภท คือ ระดับชั้นนำเข้า (input layer) 1 ชั้น ระดับชั้นที่ถูกซ่อน (hidden layers) อย่างน้อยหนึ่งชั้นขึ้นไปหรืออาจมีได้หลายชั้น และระดับชั้นส่งออก (output layer) 1 ชั้น ดังแสดงในรูปที่ VI.



รูปที่ VI. ตัวอย่างโครงข่ายประสาทเทียมซึ่งมีระดับชั้นที่ถูกซ่อน 2 ชั้น

การเชื่อมโยงในโครงข่ายประสาทเทียมทุกเส้นการเชื่อมโยงจะมีค่าน้ำหนักกำกับ ซึ่งใช้ในการคำนวณเพื่อส่งค่าไปยังระดับชั้นถัดไป ซึ่งโมเดลที่ได้ของโครงข่ายประสาทเทียมคือค่าน้ำหนักที่ถูกปรับให้เหมาะสมที่สุดทั้งโครงข่ายเพื่อการประมาณค่าผลลัพธ์ที่มีความคลาดเคลื่อนต่ำสุดด้วยอัลกอริทึมการแพร่ย้อนกลับ (Backpropagation algorithm) โดยมีการกำหนดพารามิเตอร์ hidden_layer_sizes = (512, 256, 128, 56, 32) hidden layer เป็น 5 ชั้น แต่ละชั้นมี nodes ดังนี้ ชั้นที่ 1 มี 512 nodes ชั้นที่ 2 มี 256 nodes ชั้นที่ 3 มี 128 nodes ชั้นที่ 4 มี 56 nodes และชั้นที่ 5 มี 32 nodes ใช้ฟังก์ชันกระตุ้น (activation function) แบบ relu กำหนดจำนวนรอบการฝึกฝน 200 รอบ

- เพื่อนบ้านใกล้สุด (ใช้คลาส KNeighborsRegressor ของไลบรารี neighbors)

เพื่อนบ้านใกล้สุด [7] เป็นเทคนิคการเรียนรู้ของเครื่องที่ง่ายที่สุดเทคนิคหนึ่ง ไม่มีการสร้างโมเดลขึ้นมาจากชุดข้อมูลฝึกฝนแต่ในการประมาณค่าข้อมูลแต่ละตัวอย่างจะต้องเปรียบเทียบข้อมูลแต่ละตัวอย่างนั้นกับชุดข้อมูลฝึกฝนทั้งหมดโดยตรงเพื่อหาเพื่อนบ้านหรือตัวอย่างในชุดข้อมูลฝึกฝนที่ใกล้เคียงหรือคล้ายคลึงกับตัวอย่างที่ต้องการประมาณค่าที่สุดจำนวน k ตัวอย่าง แล้วประมาณค่าตัวแปรเป้าหมายของตัวอย่างที่ต้องการประมาณค่าโดยใช้ค่าเฉลี่ยของตัวแปรเดียวกันของเพื่อนบ้าน k ตัวอย่างนั้น โดยมีการกำหนดค่าพารามิเตอร์ที่สำคัญดังนี้ กำหนดค่า k หรือพารามิเตอร์ n_neighbors จากการทดลองด้วยค่า 1, 3, 5, 7, 9 และ 11 ค่าที่เลือกไม่ควรเกินรากที่สองของจำนวนตัวอย่างในชุดข้อมูลฝึกฝนซึ่งเท่ากับ 97 แต่จากการทดลองพบว่าตั้งแต่ k เท่ากับ 11 ขึ้นไป ค่าความคลาดเคลื่อนยิ่งมีค่ามากขึ้นจึงไม่ได้นำมาเปรียบเทียบ จากการทดลองพบว่าค่า k ที่เหมาะสมที่สุดคือ 3 และจากการทดสอบความคลาดเคลื่อนของชุดข้อมูลฝึกฝนและชุดข้อมูลทดสอบแตกต่างกันไม่มากนักซึ่งแสดงว่าไม่เกิดปัญหาการเข้ารูปมากเกินไป (overfitting) นอกจากนี้มีการกำหนดพารามิเตอร์ p ซึ่งเป็นเลขยกกำลังของตัววัดระยะห่างของ Minkowski เป็น 2 ซึ่งเรียกชื่อระยะห่างเมื่อเลขยกกำลังเป็น 2 นี้ว่าระยะห่างยูคลิด (Euclidean Distance)

- การวิเคราะห์การถดถอยแบบโพลิโนเมียล (ใช้คลาส PolynomialFeatures ของไลบรารี preprocessing ในการหาพจน์ทั้งหมดตามค่าดีกรีที่กำหนด แล้วหาค่าสัมประสิทธิ์ถดถอยของแต่ละพจน์ เพื่อสร้างสมการถดถอยด้วยคลาส PolynomialLinearRegression ของไลบรารี linear_model)

การวิเคราะห์การถดถอยแบบโพลิโนเมียล [8] เป็นเทคนิคการประมาณค่าทางสถิติที่นิยมใช้เมื่อความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม

ไม่เป็นเส้นตรง โมเดลที่ได้ในการเรียนรู้จากข้อมูลฝึกฝนคือ สมการโพลิโนเมียลสำหรับการประมาณค่าตัวแปรตาม โดยในการกำหนดพารามิเตอร์ต่าง ๆ ของคลาส PolynomialFeatures ในไลบรารี preprocessing และคลาส PolynomialLinearRegression ในไลบรารี linear_model ใช้ค่าเริ่มต้น (default) ทั้งหมดยกเว้นค่าดีกรีของโพลิโนเมียล โดยมีการทดสอบดีกรีของโพลิโนเมียลเป็น 1, 2, และ 3 พบว่าเมื่อดีกรีของโพลิโนเมียลเท่ากับ 3 จะให้ค่าความคลาดเคลื่อนจากการประมาณค่าต่ำที่สุด ตัวอย่างของสมการถดถอยแบบโพลิโนเมียลเมื่อมีจำนวนตัวแปรอิสระ 2 ตัว และมีดีกรีของโพลิโนเมียลเท่ากับ 2 เป็นดังแสดงในสมการที่ 1

$$\hat{y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + \beta_4 X_1^2 + \beta_5 X_2^2 \quad (1)$$

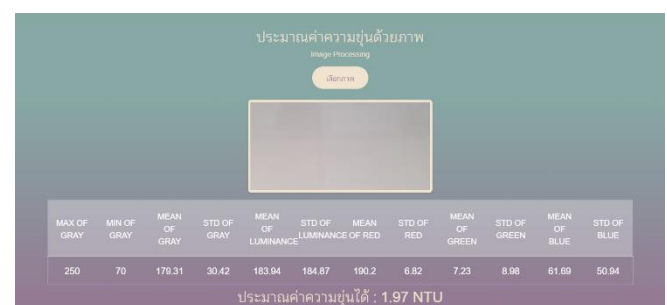
โดย \hat{y} คือตัวแปรตาม, β_i คือสัมประสิทธิ์ถดถอย และ X_i คือตัวแปรอิสระ จากสมการที่ 1 จะเห็นว่าแต่ละพจน์จะเป็นชุดค่าผสม (combinations) ที่เป็นไปได้ทั้งหมดของตัวแปรอิสระ 2 ตัวเมื่อมีค่าผลรวมของเลขยกกำลังไม่เกินดีกรีของโพลิโนเมียล แต่เนื่องจากในงานวิจัยนี้มีตัวแปรอิสระ 12 ตัว เมื่อกำหนดดีกรีของโพลิโนเมียลเป็น 3 จะมีจำนวนพจน์ในสมการเป็นจำนวนมากและไม่เหมาะในการแสดงทั้งสมการในบทความนี้ จึงขอยกเว้นการแสดงสมการดังกล่าว

การประเมินประสิทธิภาพของโมเดลใช้ตัววัดความคลาดเคลื่อนของการประมาณค่าคือค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) โดยมีสูตรการคำนวณตามสมการที่ 2 โมเดลที่มีสมรรถนะในการประมาณค่าสูงจะมีความคลาดเคลื่อน RMSE ต่ำ และโมเดลที่มีค่า RMSE ต่ำที่สุดจะถูกเลือกเพื่อใช้ในการพัฒนาเว็บแอปพลิเคชันเพื่อนำไปใช้งานจริงในหัวข้อต่อไป

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Error)^2} \quad (2)$$

III.III การพัฒนาเว็บแอปพลิเคชัน (Web Development)

คณะผู้วิจัยได้พัฒนาเว็บแอปพลิเคชันสำหรับการประเมินค่าความขุ่นของน้ำจากภาพ โดยใช้โมเดลประเมินค่าความขุ่นที่สร้างขึ้นมา เครื่องมือที่ใช้ในการพัฒนาเว็บไซต์ คือ การใช้ Flask เป็นเว็บเฟรมเวิร์ค และใช้ภาษา HTML, CSS, และ JavaScript ในการพัฒนา ผู้ใช้สามารถ Upload รูปภาพตัวอย่างน้ำ จากนั้นระบบจะประมวลผลภาพและแสดงผลค่าความขุ่นน้ำในหน่วย NTU กลับไปให้ผู้ใช้งาน ตัวอย่างการประมาณค่าความขุ่นด้วยภาพแสดงดังรูปที่ VII. (ตัวอย่างน้ำที่มีความขุ่นน้อย) และรูปที่ VIII. (ตัวอย่างน้ำที่มีความขุ่นเยอะ)



รูปที่ VII. ตัวอย่างการประมาณค่าความขุ่นของน้ำที่มีความขุ่นน้อย



รูปที่ VIII. ตัวอย่างการประมาณค่าความขุ่นของน้ำที่มีความขุ่นเยอะ

IV. ผลการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อการสร้างโมเดลประเมินค่าความขุ่นของน้ำจากการประมวลผลภาพ คณะผู้วิจัยได้ออกแบบงานวิจัยให้มีการฝึกฝนโมเดลโดยใช้เทคนิคการเรียนรู้ทั้งหมด 3 เทคนิคด้วยพารามิเตอร์ที่แตกต่างกันบนข้อมูลชุดเดียวกัน จากนั้นทำการเปรียบเทียบประสิทธิภาพของเทคนิคการเรียนรู้ของเครื่องทั้ง 3 เทคนิค เพื่อเลือกโมเดลที่ให้ค่า RMSE จากการทดสอบโมเดลด้วยชุดข้อมูลทดสอบต่ำที่สุด

จากข้อมูลตัวอย่างภาพน้ำ 9,600 ภาพที่ประกอบด้วยน้ำใส และน้ำขุ่นที่มีค่าความขุ่นแตกต่างกัน โดยในชุดข้อมูลมีค่า NTU เฉลี่ย คือ 42.4818 ค่า NTU สูงสุด คือ 82 และค่า NTU ต่ำสุด คือ 0 ภาพตัวอย่างน้ำทั้งหมดผ่านขั้นตอนการเตรียมภาพและสกัดข้อมูลให้พร้อมสำหรับการฝึกฝนและทดสอบโมเดลทั้ง 3 เทคนิค โดยการแบ่งข้อมูลแบบสุ่มออกเป็นชุดข้อมูลฝึกฝนร้อยละ 67 และชุดข้อมูลทดสอบร้อยละ 33 เพื่อทำการทดสอบเทคนิควิธีการเรียนรู้ที่มีประสิทธิภาพที่สุด ผลการเปรียบเทียบประสิทธิภาพโมเดลด้วยค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ยแสดงดังตารางที่ IV.

ตารางที่ IV. แสดงผลการวัดสมรรถนะของโมเดลต่าง ๆ ด้วยการวัดค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย

โมเดล	ค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ยของชุดข้อมูลฝึกฝน		ค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ยของชุดข้อมูลทดสอบ
โครงข่ายประสาทเทียม	5.0851		5.2482
เพื่อนบ้านใกล้เคียง	ค่า k		
	1	0.0	6.3646
	3	3.8220	6.0030
	5	4.5950	6.0397
	7	4.9895	6.1075
	9	5.3151	6.2840
	11	5.5465	6.3911
การวิเคราะห์การถดถอยแบบพอลิโนเมียล	3.3058		3.5534

จากตารางที่ IV. พบว่าโมเดลการประมาณค่าความขุ่นของน้ำที่มีค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ยของชุดข้อมูลทดสอบต่ำที่สุดคือการวิเคราะห์การถดถอยแบบพอลิโนเมียลด้วยค่าพารามิเตอร์ที่กำหนด

โดยมีค่า RMSE = 3.5534 และนำโมเดลนี้ไปใช้ในการพัฒนาเว็บแอปพลิเคชัน ซึ่งผลจากการทดสอบด้วยภาพใหม่ที่ไม่ได้นำมาใช้ในการฝึกฝนโมเดลพบว่าสามารถประมาณค่าความขุ่นของน้ำได้ใกล้เคียงความเป็นจริง

V. สรุปและข้อเสนอแนะ

ความขุ่นของน้ำถือเป็นตัวบ่งชี้ที่สำคัญในการวัดคุณภาพน้ำ งานวิจัยนี้เสนอทางเลือกในการวัดความขุ่นของน้ำด้วยเทคโนโลยีการประมวลผลภาพผ่านการใช้งานบนเว็บไซต์ ขั้นตอนดำเนินการวิจัยประกอบด้วย การเก็บรวบรวมข้อมูลตัวอย่างน้ำที่มีความขุ่นแบบต่างๆ ทั้งหมด 16 แบบ จำนวน 9600 ภาพตัวอย่างพร้อมข้อมูลความขุ่นในหน่วย NTU จากเซ็นเซอร์วัดความขุ่นเพื่อใช้เป็นข้อมูลผู้ฝึกสอนโมเดล การพัฒนาโมเดลฯ ใช้เทคนิคการเรียนรู้ 3 เทคนิค คือ โครงข่ายประสาทเทียม, เพื่อนบ้านใกล้เคียง, และการวิเคราะห์การถดถอยแบบพอลิโนเมียล จากการทดลองพบว่า การวิเคราะห์การถดถอยแบบพอลิโนเมียลให้ผลลัพธ์ค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ยน้อยที่สุด และนำมาใช้เป็นโมเดลประมาณค่าความขุ่นน้ำภายใต้เว็บแอปพลิเคชันที่พัฒนาขึ้น แต่อย่างไรก็ตามงานวิจัยนี้มีข้อจำกัดในเรื่องของภาพน้ำที่นำมาใช้ในการทดสอบโมเดลฯ เป็นภาพน้ำอย่างเดียวก่อนที่ไม่มีองค์ประกอบอื่นๆ ในภาพ โดยเป็นภาพถ่ายจากมุมมองด้านบนและใช้แสงธรรมชาติ งานวิจัยในอนาคตควรจะศึกษาเพิ่มเติม เกี่ยวกับ Texture ของสารละลายที่ปนอยู่ในน้ำชนิดต่างๆ รวมถึงการเพิ่มตัวอย่างน้ำขุ่นที่มีความหลากหลายมากขึ้น เพื่อให้ผลการประเมินค่าความขุ่นของน้ำมีความถูกต้องแม่นยำยิ่งขึ้น

เอกสารอ้างอิง

- [1] Admin, "หลักการพื้นฐานในการประมวลผลภาพดิจิทัล (Fundamental Steps in Digital Image Processing).", 2561. [ออนไลน์]. แหล่งที่มา: <http://www.imageprocessinglab.org/image-theory/หลักการพื้นฐานในการประมวลผลภาพ/>. [เข้าถึงเมื่อ: 28 กันยายน 2564].
- [2] ศูนย์ห้องปฏิบัติการ กรมอนามัย., "คุณภาพน้ำทางกายภาพและผลกระทบต่อสุขภาพ," 2553. [ออนไลน์]. แหล่งที่มา: https://rldcold.anamai.moph.go.th/images/FileDownloads/Chemical_New.pdf. [เข้าถึงเมื่อ: 1 มกราคม 2565].
- [3] Cheng En., "Automatic Detection and Assessment System of Water Turbidity based on Image Processing.", TELKOMNIKA, pp.1506 – 1513, Mar. 2013.
- [4] Vaibhav Karnawat and Dr.S.L. Patil, "Turbidity detection using Image Processing," in 2016 International Conference on Computing, IEEE Communication and Automation (ICCCA), pp. 1-4, Apr. 2016.
- [5] YEQI LIU, YINGYI CHEN and XIAOMIN FANG, "A Review of Turbidity Detection Based on Computer Vision," IEEE Access, pp. 3-17, Sep – Oct, 2018.
- [6] John J. Hopfield, "Artificial neural networks," IEEE Circuits and Devices Magazine., vol 4 (5), pp. 3-10, Sep. 1988.
- [7] Cover, Thomas M. "Estimation by the Nearest Neighbor Rule." IEEE Transactions on Information Theory, vol 14, pp. 50–55. Jan. 1968.
- [8] Ralph A Bradley and Sushil S Srivastava, "Correlation in polynomial regression.", The American Statistician, vol 33, no 1, pp. 11–14, Feb. 1979.