

ระบบวิเคราะห์ขนาดและคุณภาพมะม่วงน้ำดอกไม้โดยเทคนิคการประมวลผลภาพ

Analysis System for Size and Quality of Namdokmai Mangoes by Image Processing technique

วิไลพร กุลตั้งวัฒนา ภาณุกร ไกลสิทธิ์ วุฒิชัย โกสีย์สวัสดิ์ชัย และ กริช สมกันธา

สาขาวิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ

คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏอุดรธานี

บทคัดย่อ

งานวิจัยนี้นำเสนอระบบวิเคราะห์ขนาดและคุณภาพมะม่วงน้ำดอกไม้ โดยการประมวลผลภาพ พัฒนาขึ้นเพื่อช่วยเหลือเกษตรกรชาวสวนมะม่วง ช่วยเพิ่มประสิทธิภาพในการวิเคราะห์ขนาดและคุณภาพมะม่วงน้ำดอกไม้ โดยทำการวิเคราะห์ขนาดและคุณภาพ จากขนาด รูปร่าง และจุดดำบนผิวมะม่วง ในงานวิจัยที่ได้นำเสนอได้มีการสกัดคุณลักษณะเด่น 3 คุณลักษณะเด่น นั่นคือพื้นที่มะม่วง พื้นที่จุดดำ และคุณลักษณะรูปร่าง จากนั้นก็จะทำการจำแนกคุณภาพและขนาดมะม่วงด้วยอัลกอริทึมวิธีการเพื่อนบ้านที่ใกล้เคียงที่สุดจำนวน K ตัว (K-Nearest Neighbors Algorithm: K-NN) งานวิจัยนี้ผู้วิจัยได้ทำการเปรียบเทียบผลลัพธ์จากวิธีที่นำเสนอกับผู้เชี่ยวชาญ ซึ่งข้อมูลจากผู้เชี่ยวชาญคือข้อมูลความจริง อีกทั้งในการทดสอบประสิทธิภาพได้ใช้วิธีการแบ่งส่วนที่ละหนึ่งในการตรวจสอบ จากผลการทดลองแสดงให้เห็นว่าวิธีการที่นำเสนอมีความถูกต้องในการวิเคราะห์ใกล้เคียงกับผู้เชี่ยวชาญ ซึ่งจากงานวิจัยที่นำเสนอจะเป็นประโยชน์เป็นอย่างมากสำหรับเกษตรกรในการวัดคุณภาพของมะม่วงน้ำดอกไม้ได้อย่างมีประสิทธิภาพ

คำสำคัญ: วิธีการเพื่อนบ้านที่ใกล้เคียงที่สุด, การสกัดคุณลักษณะเด่น, มะม่วงน้ำดอกไม้

Abstract

This paper presents an efficient method to analysis of the size and quality of Namdokmai mangoes by image processing technique which was developed to help many cultivators. To analysis the size and quality of Namdokmai mangoes were done by feature extraction of size, shape, and speckle on the mango skin. In this paper, three important features i.e., size, shape, and speckle on the mango skin are taken as input in to the K-nearest neighbors. The K-nearest neighbors algorithm (K-NN) was used for analyzing the quality of Namdokmai mangoes. The results of skilled expert opinions and leave one out cross-validation were used be ground truths for testing the efficiency of the proposed method. From all experimental results show that the proposed technique is able to provide more accurate analysis results efficiently and visually close to the diagnosis by experts. The proposed method is very useful in assisting many cultivators to accomplish the task efficiently.

Keywords: K-Nearest Neighbors, Feature Extraction, Namdokmai mangoes.

1. บทนำ

มะม่วงมีถิ่นกำเนิดมาตั้งแต่สมัยโบราณจนถึงปัจจุบัน ในประเทศไทย จัดว่าเป็นผลไม้เศรษฐกิจที่สำคัญที่มีการปลูกกันอย่างแพร่หลาย นอกจากนี้ยังมีมะม่วงอีกมากมายหลากหลายขนาดและหลากหลายสายพันธุ์ เช่น มะม่วงแรด มะม่วงน้ำดอกไม้ มะม่วงเขียวเสวย เป็นต้น แต่ในปัจจุบันมะม่วงเป็นผลไม้ที่ต้องการของท้องตลาดเป็นอย่างมากและมีผลผลิตจำนวนมากต่อปี ยิ่งไปกว่านั้นยังสร้างรายได้ให้กับเกษตรกรชาวสวน ชาวไร่มะม่วง อย่างไรก็ตามเนื่องด้วยการวัดขนาดและคัดคุณภาพมะม่วงในแต่ละครั้งมีจำนวนมาก อาจเกิดการเหนื่อยล้าของ

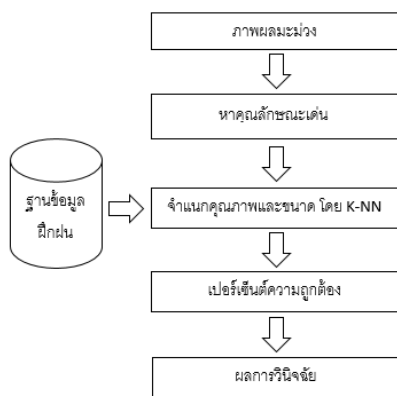
ชาวไร่มะม่วงและจะทำให้การวัดขนาดและคัดคุณภาพ Angsana New # 14 ผิดพลาด ซึ่งจะมีผลทำให้มะม่วงที่จะส่งออกไปยังผู้บริโภคทั้งในและต่างประเทศมีราคาที่ไม่สูง ทำให้ผู้บริโภคไม่มั่นใจมะม่วงของประเทศไทย ซึ่งมะม่วงบางส่วนอาจเกิดรอยแตกหรือการเกิดรอยขีดข่วนขึ้น แต่ถ้ามีการตรวจพบก่อนและทำการแยกคุณภาพของมะม่วง ก็จะทำให้มะม่วงได้ราคาที่ดี

ดังนั้นในงานวิจัยนี้จึงมีแนวคิดที่จะใช้หลักการประมวลผลภาพ (Image Processing) [1] มาใช้ในงานวิจัยคือการสกัดคุณลักษณะและเทคนิคทางปัญญาประดิษฐ์ซึ่งใช้อัลกอริทึมเพื่อนบ้านที่ใกล้เคียงที่สุด

จำนวน K ตัว (K-Nearest Neighbors) [5],[6] มาช่วยในการวิเคราะห์ขนาดและคัลคุณภาพมะม่วง เพื่อให้เกิดความถูกต้องแม่นยำมากขึ้น เพื่อให้การตรวจสอบของเกษตรกรมีความรวดเร็วมากยิ่งขึ้น และได้ระบบได้ผู้เชี่ยวชาญเรื่องมะม่วงที่มีความน่าเชื่อถือ ที่สามารถทำงานแทนมนุษย์ได้อย่างมีประสิทธิภาพ ซึ่งส่งผลทำให้ได้สินค้าเกษตรที่มีคุณภาพสำหรับการส่งออกที่ดี

2. วิธีการดำเนินการวิจัย

การที่จะวินิจฉัยภาพจะต้องมีการหาคุณลักษณะเด่นก่อนและนำคุณลักษณะเด่นที่ได้ไปจำแนกโดยอัลกอริทึม ซึ่งต้องอาศัยการฝึกฝนข้อมูลที่มีปริมาณมากๆ จะทำให้มีประสิทธิภาพสูง ซึ่งกระบวนการงานวิจัยแสดงดังรูปที่ 1



รูปที่ 1 กระบวนการวิเคราะห์ขนาดและคุณภาพมะม่วง

2.1. ความรู้เกี่ยวกับมะม่วงเกี่ยวกับจุดดำหนิ

การหาจุดดำหนิแบ่ง ได้ 3 ชั้นคุณภาพ

2.1.1 ชั้นพิเศษ (Extra class)

มะม่วงในชั้นนี้ต้องมีคุณภาพดีที่สุด ไม่มีความผิดปกติด้านรูปทรง ไม่มีตำหนิที่ผิว ในกรณีที่มี ความผิดปกติหรือตำหนิ ต้องมองเห็นได้ไม่ชัดเจน และไม่มีผลกระทบต่อรูปลักษณะทั่วไป คุณภาพของเนื้อมะม่วง คุณภาพระหว่างการเก็บรักษา และการจัดเรียงเสนอในภาชนะบรรจุ

2.1.2 ชั้นหนึ่ง (Class I)

มะม่วงในชั้นนี้ต้องมีคุณภาพดี อาจมีความผิดปกติหรือตำหนิได้เล็กน้อย ทั้งนี้ความผิดปกติหรือตำหนิดังกล่าวจะต้องไม่มีผลกระทบต่อรูปลักษณะทั่วไป ซึ่งอาจมีความผิดปกติเล็กน้อยด้านรูปทรง หรือคราบหรือรอยดำที่ติดจากยางของมะม่วง เป็นต้น

2.1.3 ชั้นสอง (Class II)

มะม่วงในชั้นนี้รวมมะม่วงที่มีคุณภาพไม่เข้าชั้นที่สูงกว่า แต่มีคุณภาพตามข้อกำหนดขั้นต่ำที่กำหนดใน มะม่วงในชั้นนี้มีความผิดปกติ

หรือตำหนิได้ทั้งนี้ความผิดปกติหรือตำหนิจะต้องไม่มีผลกระทบต่อรูปลักษณะทั่วไป คุณภาพของเนื้อมะม่วง คุณภาพระหว่างการเก็บรักษา และการจัดเรียงเสนอในภาชนะ ซึ่งอาจมีความผิดปกติด้านรูปทรง มีตำหนิที่ผิวเกิดจากการเสียดสีหรือแตกผา ครอบหรือรอยต่างๆ และมีจุดสำน้ำตาลประปราย แต่ต้องไม่มีรอยแผลเป็น

2.2 การแทนคุณลักษณะเด่น

2.2.1 Area detection

การแทนคุณลักษณะบางอย่างของวัตถุ ซึ่งอาจจะเป็นการเปลี่ยนแปลงค่าของวัตถุอยู่ในรูปของตัวเลขเพื่อที่จะนำไปเป็นค่าบ่งบอกหรือนำไปประยุกต์ในการทำงาน สำหรับการให้ Area คือคุณลักษณะพื้นฐานของวัตถุชนิดหนึ่งที่ยอมรับกันมากนั่นคือพื้นที่หรือบริเวณหรือขนาดของวัตถุนั้นเอง [2] หรือจำนวนของ Pixel ที่เราสนใจของวัตถุซึ่งสามารถคำนวณได้จากสมการ ดังนี้

$$A = \sum_{i=1}^n \sum_{j=1}^m B[i, j] \quad (1)$$

โดย B คือ ภาพขาวดำ

2.2.2 Compactness

เป็นลักษณะของรูปทรงวัตถุ [2] ที่เข้าใกล้ทรงกลม ซึ่งเราสามารถนำเอาค่านี้ไปวัดลักษณะความกลมของวัตถุได้ ว่าวัตถุใดมีลักษณะกลมหรือไม่กลม ซึ่งจะได้ค่าออกมาเป็นตัวเลขที่บ่งบอกลักษณะของวัตถุ ซึ่งสามารถคำนวณได้จาก

$$C = P^2 / A \quad (2)$$

โดย P คือ Perimeter ส่วน A คือ Area

2.2.3 Perimeter จาก Edge detection operator

ในการหาขอบของวัตถุในภาพ (edge) ซึ่งมี 2 ชนิดคือ Horizontal edges และ Vertical edges [3]-[4](เป็นขอบตามแนวแกน x และ แกน y) และ ใช้ Gradient operator มาช่วยในการหา โดยหลักการหาขอบแต่ละแบบจะต่างกันที่ filter ที่ใช้ในการคำนวณ ในงานวิจัยนี้ใช้หลักการของ sobel operator จะมี filters mask เป็น

$$H_x^s = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \text{ and } H_y^s = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (3)$$

2.2.3 Speckle Detection

ในการหาจุดดำหนิจะถูกทำโดยใช้เทคนิคเทรชโฮลด์ดัง [1] ดังสมการที่ 4

$$f_{thr}(x, y) = \begin{cases} 1, & f_f(x, y) < Threshold \\ 0, & f_f(x, y) \geq Threshold \end{cases} \quad (4)$$

โดยค่าที่มากกว่า threshold จะทำให้กลายเป็น “1” และ น้อยกว่า threshold จะทำให้กลายเป็น “0” ซึ่งจะทำให้ได้จุดหนีบริเวณที่สนใจ จากนั้นทำการนับจำนวนจุดดำหนึ่ที่เกิดขึ้นพื้นผิวของมะม่วง ซึ่งถ้ามีจุดดำหนึ่มากแสดงว่ามะม่วงไม่มีคุณภาพ ซึ่งในงานวิจัยได้ใช้ค่า threshold เท่ากับ 98

2.3 การจำแนก (Classification)

การจำแนกลักษณะเด่นที่ได้มาเพื่อที่จะทราบว่ลักษณะเด่นแต่ละตัวอยู่ Class ไหน ซึ่งในงานวิจัยนี้ได้ใช้เทคนิคการจำแนกวิธีเพื่อนบ้านที่ใกล้เคียงที่สุดจำนวน K ตัว (K-Nearest Neighbors Algorithm) ซึ่งเป็นเทคนิคที่เหมาะสมกับปัญหาแบบ Non parametric estimation สำหรับการจำแนกข้อมูลที่ไม่เป็นรูปร่างที่ดี (ข้อมูลกระจาย) วิธีการนี้จะไม่กำหนด blockตายตัวแต่จะใช้วิธีขยาย block ไปจนกว่าจะเจอข้อมูลตามจำนวนที่ k ต้องการ จากนั้นก็จะทำการจำแนกข้อมูลว่าอยู่ใกล้กับข้อมูลจุดไหนมากที่สุด [5]-[6] ในการจำแนกจะได้สมการดังนี้

$$P(W_i/X)=k_i/k \quad (5)$$

k_i = จำนวนข้อมูลของ Class ที่อยู่ใน bin

K = จำนวนข้อมูลทั้งหมดทุก Class ที่อยู่ใน bin

จากวิธีการเพื่อนบ้านที่ใกล้เคียงที่สุดในการวัดระยะทางระหว่างข้อมูลฝึกกับข้อมูลทดสอบจะใช้ Euclidean Distance คือ การวัดระยะทางความใกล้เคียงของข้อมูลที่เราจะจำแนก

$$D(\text{PointTest}, \text{PointTrain}) =$$

$$\sqrt{(F1\text{Test} - F1\text{Train})^2 + (F2\text{Test} - F2\text{Train})^2 + \dots + (Fn\text{Test} - Fn\text{Train})^2} \quad (6)$$

2.4 การหาเปอร์เซ็นต์ความถูกต้อง Leave One Out cross-validation

ในงานวิจัยนี้เพื่อวัดประสิทธิภาพของวิธีการที่นำเสนอจะใช้วิธีการ Leave One Out cross-validation [7] ซึ่งวิธีการหาเปอร์เซ็นต์ความถูกต้องของข้อมูล เช่น มีข้อมูลทั้งหมด 10 ข้อมูล จะเลือกมา 1 ข้อมูลและจะใช้ข้อมูลที่เหลือ 9 ข้อมูลเป็นตัว Train และจะใช้ 1 ข้อมูลที่ดึงออกมาเพื่อใช้ในการ Test และจะทำไปเรื่อยๆ จำนวน 10 ครั้ง

3. ผลการทดลอง

ในการทดลองได้แบ่งผลการทดลองออกเป็น 2 ส่วนคือ ส่วนของการวัดประสิทธิภาพของวิธีการที่นำเสนอ และส่วนของระบบสารสนเทศในการออกแบบระบบ

3.1 การวัดประสิทธิภาพ

การตรวจสอบความถูกต้องของข้อมูลทั้งหมดที่มีในระบบด้วยวิธีการที่ได้มาซึ่งเปอร์เซ็นต์ความถูกต้องโดยการใช้วิธี Leave One Out cross-validation โดยการทดลองกับรูปมะม่วงทั้ง 197 รูป แบ่งเป็น 3 เกรด 5 ขนาด (คุณภาพ 3 กลุ่ม และขนาด 5 กลุ่ม) ซึ่งในการทดลองจะใช้ 3-NN ในการทดสอบประสิทธิภาพ ซึ่งจากผลการทดลองสรุปได้ดังตารางที่ 1 และตารางที่ 2

ตารางที่ 1 ความถูกต้องของขนาด

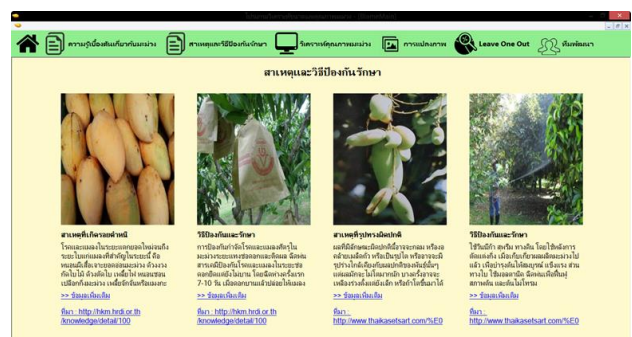
ประเภทของขนาด (Class)	ค่าพิทเชลพื้นที่	เปอร์เซ็นต์ความถูกต้อง
ขนาดใหญ่	302329	99 %
ขนาดใหญ่อ	224186	86 %
ขนาดกลาง	178055	100 %
ขนาดเล็ก	147779	98 %
ขนาดเล็กมาก	103544	87 %

ตารางที่ 2 ความถูกต้องของคุณภาพ

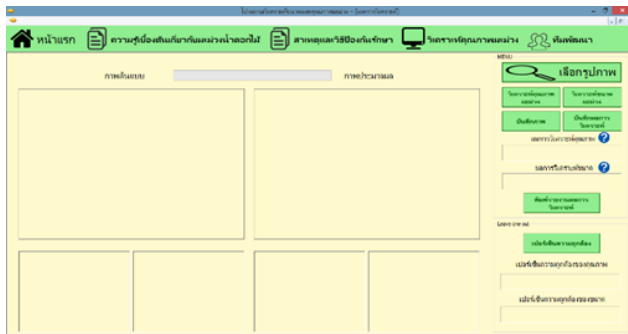
ประเภทของคุณภาพ (Class)	ค่าพิทเชลจุดดำหนึ่	เปอร์เซ็นต์ความถูกต้อง
ชั้นพิเศษ	105	97 %
ชั้น I	44	94 %
ชั้น II	16	93 %

3.2 ระบบสารสนเทศในการออกแบบระบบ

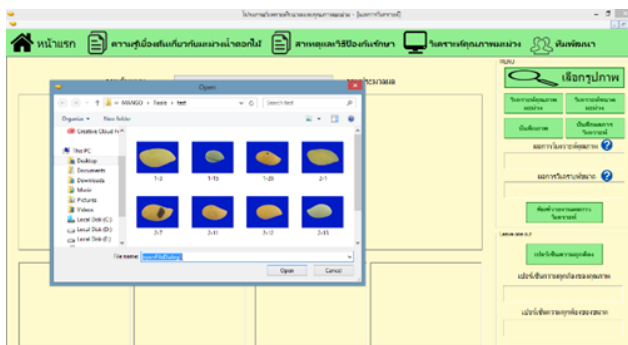
ในงานวิจัยได้มีการออกแบบระบบสารสนเทศเพื่อให้ใช้งานได้ง่ายเพื่อการตรวจสอบที่มีประสิทธิภาพดังแสดงในตัวอย่างรูปของโปรแกรม และสรุปผลจากผู้ใช้งานระบบได้ดังตารางที่ 3



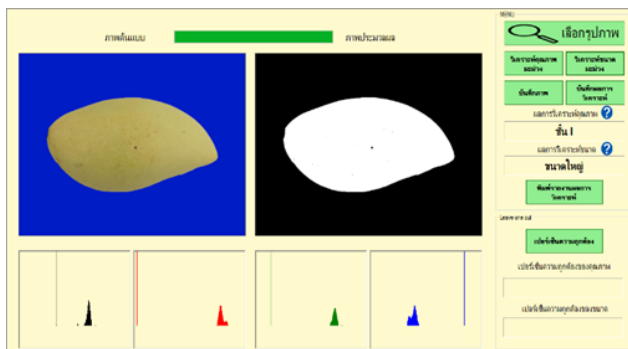
รูปที่ 2 รายละเอียด “สาเหตุและวิธีป้องกันรักษา”



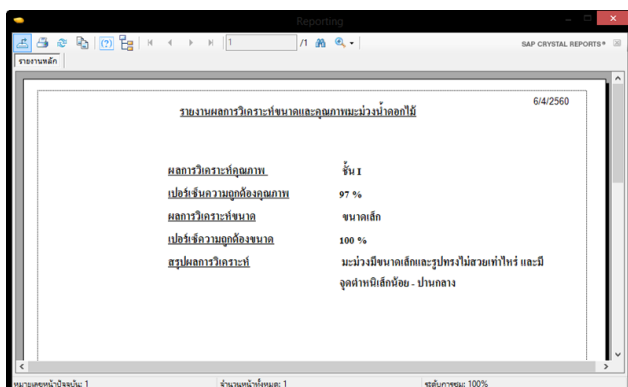
รูปที่ 3 หน้าจอหลักของการวิเคราะห์ขนาดและคุณภาพ



รูปที่ 4 การเลือกภาพมะม่วง



รูปที่ 5 หน้าจอการวิเคราะห์ขนาดและคุณภาพและผลลัพธ์



รูปที่ 6 ข้อมูลรายงานผลการวิเคราะห์ขนาดและคุณภาพ

ตารางที่ 3 สรุปผลการประเมินประสิทธิภาพของระบบจากผู้ใช้งานระบบ

รายการประเมิน	$\bar{X}$	S.D	ผล
1. ข้อมูลที่นำมาใช้ในโปรแกรม	3.59	0.06	ปานกลาง
2. ความง่ายต่อการใช้งานของระบบ	4.55	0.08	ดี
3. ผลของการวิเคราะห์ข้อมูลมีความถูกต้อง	4.75	0.09	ดี
4. ระยะเวลาในการประมวลผลการวิเคราะห์	3.56	0.06	ปานกลาง
5. ความสามารถในการบันทึกไฟล์	4.59	0.08	ดี
6. ความถูกต้องของกราฟระดับสี	3.57	0.06	ปานกลาง
7. ลิสต์ของโปรแกรมมีความน่าเชื่อถือ	4.50	0.08	ดี
8. เปอร์เซ็นต์ความถูกต้องมีความน่าเชื่อถือ	4.57	0.07	ดี
ผลรวมเฉลี่ย	4.21	0.07	ดี

4. สรุป

ระบบวิเคราะห์ขนาดและคุณภาพมะม่วงน้ำดอกไม้ ได้ใช้ 3 คุณลักษณะเด่นคือ ขนาด รูปร่าง และจุดดำหนิ เพื่อเข้าสู่กระบวนการเพื่อนบ้านที่ใกล้เคียงที่สุดจำนวน K ตัว ซึ่งสามารถแยกขนาดของมะม่วงได้เป็น 5 ขนาด ได้แก่ ขนาดใหญ่มาก ขนาดใหญ่ ขนาดกลาง ขนาดเล็ก และขนาดเล็กมาก อีกทั้งยังสามารถคัดแยกคุณภาพของมะม่วงได้เป็น 3 คุณภาพ ได้แก่ ชั้นพิเศษ ชั้น I และ ชั้น II ซึ่งจากผลการทดลองแสดงให้เห็นว่าวิธีการที่นำเสนอสามารถทำการวัดขนาดและคุณภาพได้อย่างถูกต้อง ซึ่งจะประโยชน์ต่อเกษตรกรในการคัดมะม่วงที่มีคุณภาพ

5. กิตติกรรมประกาศ

ขอขอบคุณสถาบันวิจัยและพัฒนา มหาวิทยาลัยราชภัฏอุดรธานี ผู้ให้ทุนสนับสนุนการทำวิจัย

เอกสารอ้างอิง

- [1] R.C. Gonzalez and R.E. Woods, Digital Image Processing. Reading, MA: Addison Wesley, reprint, 1992.
- [2] J.M.S. Prewitt, "Object enhancement and extraction", Picture Processing and Psychopictorics, pp.75-149, 1970.
- [3] G.S. Robinson, "Edge detection by compass gradient masks", Compute, Graph, Image Process, Vol.6, pp. 492-501, 1977.
- [4] E. Argyle, "Techniques for edge detection", in Proc. IEEE, pp.285-287, 1970.
- [5] B.V. Dasarasthy, Nearest Neighbor Pattern Classification Techniques, IEEE Computer Society Press, 1991.
- [6] G. Shakhnarovich, T. Darrel and P. Indyk, Nearest-Neighbor Methods in Learning and Vision, MIT Press, 2006.
- [7] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection", Proceedings of the 14th International Joint Conference on Artificial Intelligence, 1995, pp.1137-1143.