ระบบระบุพรรณพืชด้วยสัณฐานวิทยาของใบโดยใช้ซัพพอร์ตเวกเตอร์แมชชีน
Plant identification system by morphology of leaf using support vector machine

สุรเดช บุญลือ, ศิลา เต็มศิริฤกษ์กุล, อาภรณ์ จันทราภิรมย์

ภาควิชาวิทยาการคอมพิวเตอร์

คณะเทคโนโลยีสารสนเทศ วิทยาลัยนอร์ทกรุงเทพ

bosuradej@northbkk.ac.th, alisthonk@hotmail.com, arj_5104@hotmail.co,

บทคัดย่อ

การพัฒนาระบบระบุพรรณพืชด้วยสัณฐานวิทยาของใบโดยใช้ซัพพอร์ตเวกเตอร์แมชชีน มุ่งหวังสร้างเครื่องมือที่จะลดระยะเวลาในการระบุพรรณพืช และลดความผิดพลาดที่เกิดจากสายตาและการรู้จำของมนุษย์ ซึ่งพัฒนาด้วยโปรแกรม MATLAB, Peltarion และ Microsoft Visual Studio.NET 2005 จำนวนตัวอย่างพรรณพืชที่ใช้ในการพัฒนาระบบ 500 ตัวอย่าง จากพรรณพืช 5 ชนิด ได้รับการสนับสนุนจากสวนสมุนไพรสิรีรุกขชาติ มหาวิทยาลัยมหิดล ตัวอย่างทั้งหมดเก็บเป็นภาพถ่ายดิจิตอล และแปลงเป็นรูปแบบข้อมูลไบนารีด้วยเทคนิคการประมวลผลภาพดิจิตอล ใน 7 คุณลักษณะ ได้แก่ สีแดง สีเขียว สีน้ำเงิน รูปใบ เส้นใบ ความกว้างของใบ และความยาวของใบ นำไปผ่านตัวแบบที่ใช้เทคนิคซัพพอร์ตเวกเตอร์แมชชีน โดยมีค่าความแม่นตรงจากการทดสอบต้นแบบ 98.67% ซึ่งจากการทดสอบโดยผู้เชี่ยวชาญพบว่าความแม่นยำของการระบุพรรณพืชจากสุ่มตัวอย่าง 28 ตัวอย่าง ให้ความแม่นยำ96.43% นอกจากนั้นการประเมินประสิทธิภาพจากการใช้งานจริงของผู้เชียวชาญด้วยแบบสอบถาม ได้คะแนนเฉลี่ยอยู่ในระดับ 4.18 (SD = 0.60)

คำสำคัญ: ระบุพรรณพืช, ซัพพอร์ตเวกเตอร์แมชชีน

Abstract

Plants Identification Development by Leaf Morphology using Support Vector Machine aims to crate a tool assist to reduce the time it takes to identify various kinds of plants and also to reduce any errors from eye detection and human recognition. By doing so, MATLAB Programming, Microsoft Visual Studio.NET 2005 are introduced as the applications software to develop this system. Five hundred sample plants from five species are conducted and supported by Sirirukkhyachat Herbal Garden, Mahidol University. All such kinds of sample plants are stored in- form of digital image and then transform to binary data format by image-processing technique involve with seven characteristics of plants. For instance, red, green and blue, which are the pigments of photosynthesis, leaf shape, leaf venation, leaf width and leaf length. These characteristics are transition into the support vector machine model with the 98.67 percentages accuracy value of a prototype testing. Achieving expertise as a tester is a progressive to insist on plants identification consistence from twenty-eight sample plants with the accuracy of 96.43 percentages. In addition, this system is provided the experts perspective on the system evaluation in real work with 4.18 in SD 0.60 of an average satisfactory.

Keyword: plant classification, support vector machine

1. ความสำคัญของปัญหา

การจำแนกพรรณไม้นั้น อาศัยวิชาพฤกษศาสตร์สาขาอนุกรมวิธานพืช ซึ่งเป็นศาสตร์ในการศึกษาวิจัยเกี่ยวกับรูปพรรณสัณฐานของพืช ในการจัดหมวดหมู่เพื่อจำแนกชนิด [1] ซึ่งในการจำแนกตามหลักอนุกรมวิธานพืชนักพฤกษศาสตร์จะใช้ลักษณะเฉพาะของดอกในการจำแนกเป็นหลัก แต่ในการศึกษาพืชที่ได้รับการจำแนกชนิดแล้วเพื่อการนำมาใช้

ประโยชน์ ส่วนใหญ่ใช้หลักการระบุจากลักษณะเฉพาะของดอกและใบก็จริง พบว่าพืชบางชนิดไม่สามารถรอให้ออกดอกหรือติดผลได้ บางครั้งจึงระบุชนิดโดยอาศัยลักษณะเฉพาะของใบเพียงอย่างเดียว ซึ่งจำเป็นต้องอาศัยประสบการณ์ที่สั่งสมเป็นระยะเวลานาน ปัจจุบันมีผู้เชี่ยวชาญที่สามารถระบุชนิดพืชด้วยวิธีการดังกล่าวได้ แต่มีจำนวนน้อย เพราะต้องอาศัยความสามารถพิเศษ และระยะเวลาในการฝึกฝน [2] การนำเทคโนโลยีคอมพิวเตอร์เข้ามาใช้ช่วยเหลือในการจำแนก จึงช่วยย่นระยะเวลา และมีความแม่นตรงในการระบุชนิดพืชเพื่อนำมาใช้ประโยชน์

จากการทบทวนงานวิจัย พบว่างานวิจัยทางการรู้จำรูปลักษณะของใบส่วนใหญ่ จะใช้เทคนิคกระบวนการที่หลากหลาย อาทิเช่น Miao และคณะ ได้นำเสนอทฤษฎีการจำแนกใบของกุหลาบโดยใช้คุณลักษณะของใบ ในขณะที่ Gu และคณะ ก็พยายามที่จะศึกษาการรู้จำรูปแบบของใบพืชโดยใช้การแบ่งส่วนโครงสร้างใบด้วย wavelet transforms และ Gaussian interpolation [12] ส่วน Wang และคณะ ใช้เทคนิค moving median center (MMC) ในการจำแนกไฮเปอร์สเฟียร [13] ซึ่งเทคนิคกระบวนการนี้ Du และคณะก็นำไปประยุกต์ใช้เช่นกัน [9]

งานวิจัยที่กล่าวถึงนี้ ส่วนใหญ่จะใช้เทคนิคการจำแนกแบบ k-nearest neighbor (k-NN) ในการจำแนก [8] [12] [10] ขณะที่มีบางงานวิจัยนำเทคนิคโครงข่ายประสาทเทียมมาประยุกต์ใช้ อาทิเช่น Saitoh และคณะ ได้นำเทคนิคนี้ มาประยุกต์ใช้กับการจำแนกดอกไม้ป่าด้วยรูปร่างของใบ [15] ส่วน Heymans และคณะ ก็นำเสนอการประยุกต์ใช้โครงข่ายประสาทเทียมในการจำแนกพรรณพืช [4]

งานวิจัยที่นำเสนอนี้ยังมีจุดที่ต้องการปรับปรุงพัฒนาให้ดีขึ้น ด้วยว่ามีบางงานวิจัยที่ใช้วิธีจำแนกพรรณพืชแบบเฉพาะเจาะจงสายพันธุ์เท่านั้น [14] [10] [4] อีกทั้งยังมีปัญหาอยู่ที่ความแม่นตรงของการจำแนกรูปแบบ ซึ่งจากการสำรวจพบว่าสูงสุดอยู่ที่ 90.31% [11] โดยใช้เทคนิคโครงข่ายประสาทเทียม ดังนั้นทีมคณะผู้วิจัยจึงได้เสนอวิธีซัพพอร์ตเว็คเตอร์แมชชีน เพื่อใช้ในการระบุชนิดพืชด้วยสัณฐานวิทยาของใบ ด้วยว่าใบของพืชแต่ละชนิดมีความหลากหลายทางกายภาพ มีรูปใบที่แตกต่างกัน ทำให้ต้องใช้ความละเอียดอย่างสูงในการจำแนก และการอนุมานคำตอบ ซึ่ง ซัพพอร์ตเวกเตอร์แมชชีนสามารถกระทำได้อย่างมีประสิทธิภาพสูงกว่า

การพัฒนาระบบระบุพรรณพืชด้วยสัณฐานวิทยาของใบโดยใช้ซัพพอร์ตเวกเตอร์แมชชีน คณะผู้วิจัยมุ่งหวังให้เป็นเครื่องมือพื้นฐานสำหรับการรู้จักพืชในระยะสั้นและเป็นเครื่องมือหนึ่งที่สนับสนุนการใช้พืชให้ถูกชนิด ซึ่งเป็นปัจจัยที่สำคัญที่สุดในการนำพืชมาใช้ประโยชน์ และลดความเสี่ยงในการเกิดอันตราย อีกทั้งหากมีการเก็บข้อมูลได้อย่างครบถ้วนตามตัวอย่างที่มีอยู่จริงเครื่องมือนี้น่าจะสามารถสนับสนุนข้อมูลสำหรับการวิเคราะห์และพัฒนาองค์ความรู้ใหม่เกี่ยวกับพืชได้ในอนาคต

2. วัตถุประสงค์

พัฒนาระบบระบุพรรณพืชด้วยสัณฐานวิทยาของใบโดยใช้ซัพพอร์ตเวกเตอร์แมชชีน เพื่อจำแนกพรรณพืชในระยะเวลาสั้นและมีความถูกต้องแม่นตรงสูง

3. ทฤษฎีที่เกี่ยวข้อง

ทฤษฎีที่นำมาใช้ในงานวิจัยนี้ เป็นองค์ความรู้ทางด้านอนุกรมวิธานพืช ซึ่งเป็นพื้นความรู้ในการจำแนกพืชพันธุ์ใหม่ๆ โดยจะนำมาประยุกต์ใช้ร่วมกับเทคนิคการประมวลผลภาพและเทคนิคซัพพอร์ตเว็คเตอร์แมชชีน เพื่อพัฒนาแอพลิเคชั่นสำหรับระบุพรรณพืชโดยใช้ลักษณะทางสัณฐานวิทยาของใบพืชได้อย่างถูกต้อง

3.1 ทฤษฎีทางด้านพฤกษศาสตร์

3.1.1 อนุกรมวิธานพืช

หลักอนุกรมวิธานพืชเป็นหลักที่ใช้พิสูจน์พันธุ์พืชซึ่งมีวิธีกระบวนการพิสูจน์หลายวิธีด้วยกัน

1) การระบุ (Identification) คือ การพิสูจน์ชนิดพืชโดยใช้วิธีการเปรียบเทียบการตัดสินใจ การใช้ความจำอันแม่นตรงและประสบการณ์อย่างหนึ่งอย่างใดหรือหลายๆอยู่รวมกัน ซึ่งในขั้นตอนแรกจะใช้วิธีการจำเหมือนกับการจำหน้าบุคคลต่างๆ ซึ่งต้องจำลักษณะของพืชแต่ละชนิด แต่ละสกุล หรือแต่ละวงศ์ให้ได้แต่อย่างไรก็ตามเพื่อความมั่นใจ จำเป็นต้องทำการยืนยันโดยการนำพรรณไม้ชนิดนั้นไปเปรียบเทียบกับตัวอย่างที่มีชื่อแล้วในหอพรรณไม้ต่อไป สำหรับในประเทศไทย มีหอพรรณไม้อยู่ที่กรมป่าไม้

2) การจำแนก (Classification) คือ การจำแนกพรรณพืชขึ้นเป็นหมวดหมู่ต่างๆ การจำแนกอย่างง่ายๆก็ได้แก่การจัดพืชเป็นพวกไม้ต้นไม้พุ่มไม้เถาและไม้ล้มลุกหรือจัดเป็นพืชใบเลี้ยงคู่กับใบเลี้ยงเดี่ยวกล่าวคือจะจัดหมวดหมู่พืชที่มีลักษณะคล้ายคลึงกันไว้ด้วยกัน ส่วนการจำแนกพืชในขั้นสูงก็จะจำแนกออกตามลำดับชั้นต่างๆจากระดับกว้างสุดมาหาแคบสุด

3) การใช้รูปวิธาน เป็นวิธีการหนึ่งในการพิสูจน์พันธุ์ไม้ โดยการคัดลักษณะต่างๆ ที่ ไม่ปรากฏในพันธุ์ไม้ที่ต้องการจะพิสูจน์คงเหลือลักษณะที่ปรากฏอยู่ แล้วพิจารณาในขั้นต่อไปซึ่งจะมีขอบเขตที่แคบลง อย่างนี้ไปเรื่อยๆจนกว่าจะพบคำตอบ

รูปวิธานที่ใช้กันมากในปัจจุบันนี้เรียกว่า “Dichotomous key” คือการนำเอาลักษณะที่แตกต่างกันหรือตรงกันข้ามมาเข้าคู่กัน แล้วพิจารณาเปรียบเทียบในคู่นั้นๆว่าจะเข้าลักษณะไหนตัดลักษณะที่ไม่ใช่ออกไปแล้วไปพิจารณา ในส่วนที่แคบลงไปซึ่งจะมีให้เปรียบเทียบเป็นคู่ๆ อีก ทำเช่นนี้ไปเรื่อยๆ จนเหลือแต่ลักษณะที่ปรากฏอยู่ในพรรณไม้ที่กำลังตรวจสอบ ก็จะได้ผลลัพธ์ในขั้นสุดท้าย คือชนิดพันธุ์ไม้นั้นๆ

อย่างไรก็ตามการวิเคราะห์พันธุ์ไม้ในระดับนี้ ผู้ที่จะทำการวิเคราะห์จะต้องมีประสบการณ์และพื้นฐานความรู้ในด้านอนุกรมวิธาน (Plant taxonomy) และวิชาการด้านอื่นๆ เช่น รุกขวิทยา(Dendrology) และ พฤกษศาสตร์ (Botany) มาพอสมควร ในส่วนของผู้สนใจและประชาชนทั่วไปควรเริ่มต้นจากการจดลักษณะของพันธุ์ไม้แต่ละชนิดให้ได้เสียก่อน โดยเฉพาะกับชนิดพันธุ์ไม้ที่สำคัญๆและมีคุณค่าทางเศรษฐกิจหรือแม้แต่พันธุ์ไม้ที่พบเห็นได้ทั่วไปในชีวิตประจำวัน เช่นไม้ผลชนิดต่างๆ [1]

3.1.2 สัณฐานวิทยาของใบพืช

ใบพืชประกอบด้วย 3 ส่วนด้วยกัน ได้แก่ แผ่นใบ ก้านใบ และหูใบ ในงานวิจัยนี้ จะมุ่งเน้นการวิเคราะห์โครงสร้างของแผ่นใบ ซึ่งจะมีองค์ประกอบในการวิเคราะห์ดังต่อไปนี้

รูปใบ จะมีลักษณะแตกต่างกันไปได้หลายรูปแบบ เช่น รูปลิ่มแคบ รูปเข็มรูปใบหอก รูปหอกกลับ รูปขอบขนาน

ปลายใบ เป็นส่วนที่นับจากบริเวณกลางใบไปถึงส่วนบนสุดของใบ มีได้หลายรูปแบบเช่นปลายตัด ปลายมน เว้าบุ๋ม เว้าตื้น

โคนใบ ส่วนที่ลงมาจากบริเวณกลางใบ จนถึงก้านใบ มีได้หลายรูปแบบเช่น โคนรูปลิ่ม โคนสอบเรียว โคนเฉียง

ขอบใบคือบริเวณโดยรอบของใบ ลักษณะต่างกันไปเช่น ขอบใบเรียบ ขอบใบเป็นคลื่น

3.2 เทคนิคหรือเทคโนโลยีที่ใช้ในการพัฒนา

3.2.1 การประมวลผลภาพ (Image processing)

การนำภาพมาประมวลผลหรือคิดคำนวณด้วยคอมพิวเตอร์เพื่อให้ได้ข้อมูลที่เราต้องการทั้งในเชิงคุณภาพและปริมาณ โดยมีขั้นตอนต่าง ๆ ที่สำคัญ คือ การทำให้ภาพมีความคมชัดมากขึ้น การกำจัดสัญญาณรบกวนออกจากภาพ การแบ่งส่วนของวัตถุที่เราสนใจออกมาจากภาพ เพื่อนำภาพวัตถุที่ได้ไปวิเคราะห์หาข้อมูลเชิงปริมาณ เช่น ขนาด รูปร่าง และทิศทางการเคลื่อนของวัตถุในภาพ จากนั้นเราสามารถนำข้อมูลเชิงปริมาณเหล่านี้ไปวิเคราะห์ และสร้างเป็นระบบต่อไป

3.2.2 ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machines)

วิธีการของซัพพอร์ตเวกเตอร์แมชชีน [5] [6] [7] จัดเป็นเทคนิคที่ใช้ในการแก้ปัญหาทางด้านการรู้จำรูปแบบข้อมูล โดยอาศัยหลักการของการหาสัมประสิทธิ์ของสมการเพื่อสร้างเส้นแบ่งแยกกลุ่มข้อมูลที่ถูกป้อนเข้าสู่กระบวนการสอนให้ระบบเรียนรู้ โดยเน้นไปยังเส้นแบ่งแยกแยะกลุ่มข้อมูลได้ดีที่สุด (Optimal Separating Hyperplane) สำหรับรากฐานเดิมของซัพพอร์ตเวกเตอร์แมชชีน ถูกนำมาใช้กับข้อมูลที่เป็นเชิงเส้น แต่ในความเป็นจริงแล้วข้อมูลที่นำมาใช้ในระบบการสอนให้ระบบเรียนรู้ส่วนใหญ่มักเป็นข้อมูลแบบไม่เป็นเชิงเส้น ซึ่งสามารถแก้ปัญหาดังกล่าวด้วยเคอร์เนล ที่เป็นการเปลี่ยนแปลงมิติของข้อมูลให้สูงขึ้น เพื่อช่วยในการเรียงตัวของข้อมูลเสียใหม่ที่เรียกว่า “พื้นที่มิติสูง” (Higher Dimensional Space) เคอร์เนลที่นิยมใช้มีอยู่ 3 ชนิด [6] ด้วยกันคือ โพลิโนเมียล (Polynomial), เรเดียลเบสิสฟังก์ชัน (Radial Basis Function-RBF) และ ซิกมอยด์ (Sigmoid) โดยงานวิจัยนี้จะใช้เคอร์เนลซิกมอยด์เป็นหลัก

4. วิธีการดำเนินการ

4.1 การเตรียมข้อมูลกลุ่มพืชตัวอย่างที่ใช้ในงานวิจัย

4.1.1 การคัดเลือกกลุ่มพืชตัวอย่าง

พืชตัวอย่างทั้ง 5 ชนิด ซึ่งคัดสรรพืชที่พบเจอได้ง่ายในประเทศไทย ซึ่งมีลักษณะใบมีความแหมือนและแตกต่างกัน โดยได้รับคำแนะนำ จากรองศาสตราจารย์รุ่งระวี เต็มศิริฤกษ์กุล นักวิชาการทางด้านพฤกษศาสตร์จากคณะเภสัชศาสตร์ มหาวิทยาลัยมหิดล พืชทั้ง 5 ชนิด ซึ่งได้แก่ กระดังงาไทย ชบา ตำลึง จินจีเหมาเยี่ย และ โหระพา

4.1.2 การเก็บตัวอย่าง

การเก็บตัวอย่างพืชทั้ง 5 ชนิด ใช้ใบโตเต็มที่ ที่มีความสมบูรณ์ตั้งแต่ปลายใบ จรดโคนใบ ไม่มีรอยขาดหรือทะลุ หรือความเสียหายใดๆ นำเข้าแฟ้มข้อมูลภาพดิจิตอล โดยกล้องถ่ายภาพดิจิตอล ความละเอียด 6 เม็กกะพิกเซลโดยเก็บภาพภายในกล่องเก็บภาพที่มีพื้นสีขาว หรือกระดาษสีขาว เก็บภาพ 2 ด้าน ได้แก่ หน้าใบ และ หลังใบ ทั้งหมด 5 ชนิด ชนิดละ 100 ใบ รวมทั้งสิ้น 500 ภาพ ดังภาพตัวอย่างที่ 1

ภาพที่ 1 ตัวอย่างการเก็บข้อมูลพืช

4.2 การประมวลผลภาพ

แฟ้มข้อมูลภาพที่ได้มาจะต้องผ่านการกรองภาพเพื่อตัดสัญญาณรบกวนที่มาจากสิ่งแวดล้อมในขณะถ่ายภาพ โดยใช้วิธี Convolution ซึ่งเป็นคือการกระทำกันระหว่าง Template ที่เป็น Matrix ขนาด n x m กับ ภาพ (image) ดังแสดงในสมการที่ 3-1 [2] ซึ่งจะสามารถลดทอนคุณสมบัติที่ไม่ต้องการลง และเน้นคุณสมบัติที่ต้องการให้เด่นชัดขึ้น เมื่อเสร็จสิ้นการกรองข้อมูลภาพ ค่าสีของภาพจะถูกนำไปสร้าง Histogram ในระบบสีแบบ RGB เพื่อใช้ในการเปรียบเทียบสีของใบพืช

$$I'(x,y) = T * I = \sum_{i=0}^{n-1}\sum_{j=0}^{m-1} T(i,j)_{\mathrm{g}} I(x-i, y-j) \quad (3\text{-}1)$$

เมื่อ $I'(x,y)$ คือ ภาพผลลัพธ์ที่ผ่านการกรองภาพ

T คือ Template ที่ใช้ในการกรองภาพ

I คือ ภาพที่ต้องการกรอง

ในลำดับต่อไปเป็นการแปลงภาพ จากระดับสี RGB สู่ระดับ Gray Scale ซึ่งจะใช้ค่าความเข้มของสีทั้ง 3 ได้แก่ สีแดง เขียว น้ำเงิน มาทำการหาค่าความเข้มในระดับ Gray Scale ดังสมการที่ 3-2 ในแต่ละจุดของภาพ

$$f'(x,y) = \begin{pmatrix} 1 \,|\, if\ f(x,y) \geq T \\ 0 \,|\, if\ f(x,y) < T \end{pmatrix} \quad (3\text{-}2)$$

เมื่อ $f'(x,y)$ คือ ค่าที่ได้จากการทำ Threshold

$f(x,y)$ คือ ค่าความเข้มในระดับ Grayscale

เมื่อทำ Threshold จนได้ความละเอียดที่เหมาะสมกับการวิเคราะห์ จึงทำการหาขอบภาพโดยใช้ทฤษฎีการหาขอบภาพตามแบบของโซเบล ซึ่งจะเป็นการหาขอบภาพในลักษณะที่ไม่เป็นเชิงเส้นโดยมีลักษณะ Template เป็น Matrix ขนาดปรับเปลี่ยนตามความเหมาะสม กระทำกับรูปที่ต้องการหาขอบดังสมการที่ 3-4 กับ 3-5

$$C_{gray}(x,y) = \left\{ \frac{C_{red}(x,y) + C_{green}(x,y) + C_{blue}(x,y)}{3} \right\} \quad (3\text{-}3)$$

เมื่อ $C_{gray}(x,y)$ คือ ความเข้มในระดับสีเทา ณ จุด (x,y)

$C_{red}(x,y)$ คือ ความเข้มของสีแดง ณ จุด (x,y)

$C_{green}(x,y)$ คือ ความเข้มของสีเขียว ณ จุด (x,y)

$C_{blue}(x,y)$ คือ ความเข้มของสีน้ำเงิน ณ จุด (x,y)

ทำ Threshold เพื่อแปลงภาพสู่ระดับ Binary Image ด้วยสมการที่ 3-3 ในการแปลง Threshold ตามค่าที่ถูกกำหนดโดยผู้ใช้ ความละเอียด จะขึ้นอยู่กับการกำหนดค่า Threshold ที่เหมาะสมด้วยเช่นกัน

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * I \quad (3\text{-}4)$$

$$G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * I \quad (3\text{-}5)$$

เมื่อ G_x คือ ค่า Gradient ตามแนวแกน x

G_y คือ ค่า Gradient ตามแนวแกน y

และมีสมการในการคำนวณเพื่อหาขนาดของขอบภาพดังสมการที่ 3-6 และ การหาทิศทางของขอบภาพดังสมการที่ 3-7

$$M_G = \sqrt{G_x^2 + G_y^2} \quad (3\text{-}6)$$

$$\Theta = \arctan(\frac{G_y}{G_x}) \qquad (3\text{-}7)$$

เมื่อ M_G คือ ขนาดของขอบภาพ

Θ คือ ทิศทางของขอบภาพ

เมื่อเสร็จสิ้นกระบวนการวิเคราะห์ตามขั้นตอนดังกล่าวอาจมีการทำการแบ่งส่วนของรูปภาพในการวิเคราะห์เพื่อแบ่งจุดในการวิเคราะห์ตามลักษณะทางสัณฐานวิทยาของใบพืช อีกครั้งหนึ่ง และนำข้อมูลที่ได้จัดทำ Dataset เพื่อใช้ในการเรียนรู้ของเครื่องต่อไป

4.3 การสร้างตัวแบบเรียนรู้

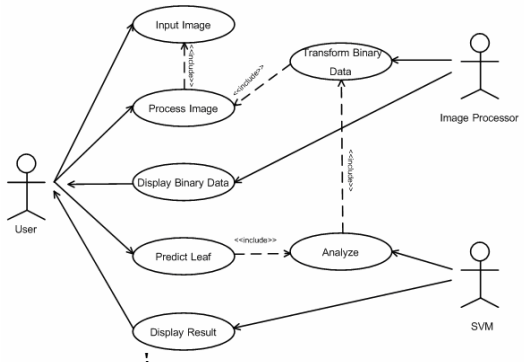
เป็นส่วนการสร้างตัวแบบเรียนรู้โดยใช้ซัพพอร์ตเวกเตอร์แมชชีนเพื่อเรียนรู้ข้อมูลที่ป้อนเข้าไป การฝึกสอนให้จดจำข้อมูลทั้งหมดที่ป้อนเข้าไป โดยชุดข้อมูลที่ใช้สอนได้แก่ ค่าข้อมูลที่ได้จากการประมวลผลภาพ ประกอบไปด้วย สีแดง สีเขียว สีน้ำเงิน รูปใบ เส้นใบ ความกว้างของใบ และความยาวของใบ ซึ่งชุดข้อมูลทั้งหมดจะถูกเก็บไว้ในไฟล์ข้อมูล

ซัพพอร์ตเวกเตอร์แมชชีนสามารถเรียกไฟล์ข้อมูลทั้งหมดนี้ขึ้นมาทำการฝึกสอน ในที่นี้จะกำหนดรอบในการฝึกสอนทั้งหมด 30,000 รอบ ทำให้ซัพพอร์ตเวกเตอร์แมชชีนเกิดความรู้ และประสบการณ์ ซึ่งสามารถนำไปใช้ในการวิเคราะห์ตีความข้อมูลที่อยู่ในลักษณะคล้ายกัน โดยได้ค่าความแม่นยำ 98.67 %

4.4 ขั้นตอนการพัฒนาแอพพลิเคชั่น

4.4.1 การออกแบบระบบ

ใช้การวิเคราะห์และออกแบบระบบเชิงวัตถุ และใช้โปรแกรม Microsoft Visio 2003 เป็นเครื่องมือช่วยพัฒนา โดยภาพรวมของระบบสามารถแสดงได้ดังภาพที่ 2 ข้างล่างต่อไปนี้



ภาพที่ 2 Use Case Diagram

4.4.2 การพัฒนาซอฟต์แวร์ระบบงาน

ส่วนการประมวลผลภาพ จะใช้โปรแกรม MATLAB สร้างฟังก์ชันประมวลผลภาพ แล้วนำเข้าสู่ตัวโปรแกรมที่พัฒนาด้วย Microsoft Visual Studio .NET 2005 ซึ่งจะมีส่วนคอมโพเน็นต์ของซัพพอร์ทเวกเตอร์แมชชีนที่สร้างขึ้นด้วยโปรแกรม Peltarion

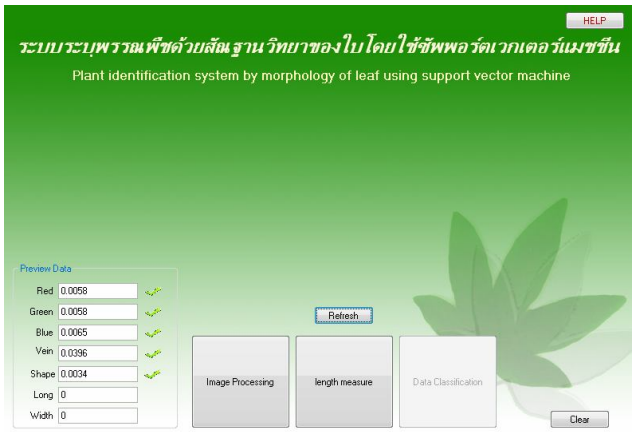
4.4.3 การทดสอบระบบงาน

จัดทำแบบประเมินประสิทธิภาพและความพึงพอใจของผู้ใช้ และนำไปให้ผู้เชี่ยวชาญซึ่งเป็นนักวิชาการทางด้านพฤกษศาสตร์ จากคณะเภสัชศาสตร์ มหาวิทยาลัยมหิดล

5. ผลของการดำเนินงาน

5.1 ส่วนของระบบงาน

จากการพัฒนาระบบงาน ผลที่ได้จะเป็นโปรแกรมที่สามารถนำข้อมูลจากภาพถ่ายรูปใบไม้มาประมวลผลภาพเพื่อให้ได้ข้อมูลดิจิตอล ต่อจากนั้นข้อมูลที่ได้จะส่งเข้าสู่ส่วนเรียนรู้แบบซัพพอร์ทเวกเตอร์แมชชีน ซึ่งจะตัดสินใจจากเงื่อนไขที่ได้จากการเรียนรู้แล้ว มาใช้ในการระบุพรรณพืชต่อไป สามารถแสดงหน้าจอโปรแกรมได้ดังภาพที่ 3 ต่อไปนี้



ภาพที่ 3 แสดงการโปรแกรมหลังจากประมวลผลภาพ

5.2 ส่วนการเรียนรู้โดยใช้ซัพพอร์ทเวกเตอร์แมชชีน

5.2.1 การทดสอบความแม่นตรงจากตัวแบบ

จากการทดสอบกับตัวแบบจากกลุ่มข้อมูลที่ใช้ จำนวน 500 ตัวอย่าง โดยจะทำการแบ่งข้อมูลออกเป็นสองส่วน คือ ส่วนของการสอน 80 เปอร์เซ็นต์ และข้อมูลสำหรับการทดสอบคำตอบ 20 เปอร์เซ็นต์ ปรากฏว่าได้ค่าความแม่นยำ98.67 %

5.2.2 การทดสอบความแม่นตรงจากผู้เชียวชาญ

จากการทดสอบโดยผู้เชี่ยวชาญพบว่าความแม่นตรงของการระบุพรรณพืชจากสุ่มตัวอย่าง 28 ตัวอย่าง ให้ความแม่นยำ 96.43%

5.3 การประเมินหาประสิทธิภาพของระบบโดยผู้เชี่ยวชาญ

เมื่อได้นำระบบไปประเมินโดยนักวิชาการทางด้านพฤกษศาสตร์ จากคณะเภสัชศาสตร์ มหาวิทยาลัยมหิดล สามารถสรุปผลการประเมินแต่ละด้าน ได้ดังตารางที่ 1 ข้างล่างต่อไปนี้

ตารางที่ 1 การประเมินประสิทธิภาพโดยผู้เชี่ยวชาญ

รายการประเมิน	$\bar{x}$	SD	ระดับ
1. ผลการประเมินด้านความสามารถในการทำงาน	3.80	0.75	ดี
2. ผลการประเมินด้านความต้องการของผู้ใช้	4.33	0.39	ดี
3. ผลการประเมินด้านการใช้งานของโปรแกรม	4.13	0.69	ดี
4. ผลการประเมินด้านผลลัพธ์ที่ได้จากโปรแกรม	4.30	0.94	ดี
5. ผลการประเมินด้านความปลอดภัย	4.34	0.24	ดี
สรุปประเมินประสิทธิภาพโดยผู้เชี่ยวชาญ	4.18	0.60	ดี

6. บทสรุป

งานวิจัยนี้ได้แนะนำถึงการนำเทคนิคซัพพอร์ตเวกเตอร์แมชชีนนำมาประยุกต์ใช้กับการจำแนกพรรณพืชด้วยสัณฐานวิทยาของใบพืช ซึ่งมีการนำภาพถ่ายรูปใบจำนวน 500 ตัวอย่าง จากพืชจำนวน 5 ชนิด มาผ่านกระบวนการประมวลผลภาพแล้วส่งผ่านให้ตัวแบบเรียนรู้ซัพพอร์ตเวกเตอร์แมชชีนจำแนกพรรณพืช ซึ่งพบว่าระบบมีค่าความแม่นตรงจากการทดสอบต้นแบบ 98.67% แต่จากการทดสอบโดยผู้เชี่ยวชาญพบว่าความแม่นตรงของการระบุพรรณพืชจากสุ่มตัวอย่าง 28 ตัวอย่าง ให้ความแม่นยำ96.43% นอกจากนั้นการประเมินประสิทธิภาพจากการใช้งานจริงของผู้เชียวชาญด้วยแบบสอบถาม ได้คะแนนเฉลี่ยอยู่ในระดับ 4.18 (SD = 0.60)

7. กิติกรรมประกาศ

ขอขอบพระคุณ วิทยาลัยนอร์ทกรุงเทพ ที่ให้ทุนสนับสนุนงานวิจัยครั้งนี้ รวมทั้งรองศาสตราจารย์รุ่งระวี เต็มศิริฤกษ์กุล และคณะ จากสวนสมุนไพรสิรีรุกขชาติ คณะเภสัชศาสตร์ มหาวิทยาลัยมหิดล ที่อำนวยความสะดวกด้านการให้คำปรึกษาอนุเคราะห์การเก็บข้อมูลและประเมินผลระบบงาน

8.บรรณานุกรม

[1] ก่องกานดา ชยามฤต. “คู่มือจำแนกพรรณไม้” กรุงเทพ: กรมป่าไม้, 2541.

[2] มนตรี กาญจนะเดชะ. “เอกสารประกอบการสอนวิชาประมวลผลภาพ” คณะวิศวกรรมศาสตร์ มหาวิทยาลัยสงขลานครินทร์, 2544.

[3] รุ่งระวี เต็มศิริฤกษ์กุล. “พรรณไม้มีพิษ” กรุงเทพฯ: คณะเภสัช-ศาสตร์ มหาวิทยาลัยมหิดล, 2537.

[4] B. C. Heymans, J. P. Onema, and J. O. Kuti, “A neural network for opuntia leaf-form recognition,” in *Proceedings of IEEE International Joint Conference on Neural Networks*, 1991.

[5] Chapelle O. Chapelle, P. Haffner and V. Vapnik. “SVMs for Histogram-Based Image Classification.” Submitted to IEEE Transactions on Neural Networks. 1999.

[6] Chih-Chung Chang and Chih-Jen Lin. “LIBSVM.” A Library for Support Vector Machines. September, 2006.

[7] E. Mayoraz and E. Alpaydm. “Support vector machines for multi-class classification.” in: IDIAP Research Report. May, 1998.

[8] J.-X. Du, X.-F. Wang, and G.-J. Zhang, “Leaf shape based plant species recognition,” *Applied Mathematics and Computation*, vol. 185, 2007.

[9] J. Du, D. Huang, X. Wang, and X. Gu, “Shape recognition based on radial basis probabilistic neural network and application to plant species identification,” in *Proceedings of 2005 International Symposium of Neural Networks*, ser. LNCS 3497. Springer, 2005.

[10] F. Gouveia, V. Filipe, M. Reis, C. Couto, and J. Bulas-Cruz, “Biometry: the characterisation of chestnut-tree leaves using computer vision,” in *Proceedings of IEEE International Symposium on Industrial Electronics*, Guimar˜aes, Portugal, 1997.

[11] Stephen Gang Wu and et al. “A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network.” Institute of Applied Chemistry, Chinese Academy of Science, P. R. China. 2007.

[12] X. Gu, J.-X. Du, and X.-F. Wang, “Leaf recognition based 2on the combination of wavelet transform and gaussian interpolation,” in *Proceedings of International Conference on Intelligent Computing 2005*, ser. LNCS 3644. Springer, 2005.

[13] X.-F. Wang, J.-X. Du, and G.-J. Zhang, “Recognition of leaf images based on shape features using a hypersphere classifier,” in *Proceedings of International Conference on Intelligent Computing 2005*, ser. LNCS 3644. Springer, 2005.

[14] Z. Miao, M.-H. Gandelin, and B. Yuan, “An oopr-based rose variety recognition system,” *Engineering Applications of Artificial Intelligence*, vol. 19, 2006.

[15] T. K. Takeshi Saitoh, “Automatic recognition of wild flowers,” in *Proceedings of 15th International Conference on Pattern Recognition (ICPR’00)*, vol. 2, 2000.