



การระบุสายพันธุ์ข้าวจากลักษณะทางสัณฐานวิทยาของข้าวเปลือกด้วยวิธีต้นไม้ตัดสินใจร่วมกับตรรกะฟัซซี Rice Variety Identification Using Morphological Characteristics of Paddy Seed via Decision Tree and Fuzzy Logic

อรรถศาสตร์ นาคเทวัญ^{1*}, พัชรารณ นาคเทวัญ², นพดล มณีรัตน์³

Athasart Narkthewan^{1*}, Patcharaporn Narkthewan², Noppadol Maneerat³

¹หลักสูตรวิศวกรรมคอมพิวเตอร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง วิทยาเขตชุมพรเขตรอุดมศักดิ์, ชุมพร, 86160

²หลักสูตรรัตนตรัยอาหารและการจัดการ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง วิทยาเขตชุมพรเขตรอุดมศักดิ์, ชุมพร, 86160

³ภาควิชาวิศวกรรมการวัดและควบคุม คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, กรุงเทพมหานคร, 10520

¹Computer Engineering Program, King Mongkut's Institute of Technology Ladkrabang Prince of Chumphon Campus, Chumphon, 86160

²Food Innovation and Management Program King Mongkut's Institute of Technology Ladkrabang Prince of Chumphon Campus, Chumphon, 86160

³Department of Instrumentation and Control Engineering Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok, 10520

*Corresponding author: Tel: +66-8-1632-6779, E-mail: athasart.na@kmitl.ac.th

บทคัดย่อ

การรวบรวมข้าวสายพันธุ์ต่าง ๆ เป็นแนวทางในการอนุรักษ์เมล็ดพันธุ์ข้าว ซึ่งการอนุรักษ์พันธุ์ข้าวจะช่วยแก้ปัญหาการขาดแคลนเมล็ดพันธุ์ข้าวที่บริสุทธิ์และคุณภาพดี อีกทั้งยังช่วยในการรักษาพันธุ์ข้าวไม่ให้สูญหาย นอกจากนี้พันธุ์ข้าวที่คัดเลือกแล้วยังสามารถใช้เป็นเมล็ดพันธุ์ที่ดีในอนาคตได้อีกด้วย ในส่วนของการจำแนกสายพันธุ์ข้าวจะใช้ลักษณะทางสัณฐานวิทยาของพันธุ์ข้าวเป็นองค์ประกอบ ดังนั้นผู้เชี่ยวชาญและเครื่องมือเกี่ยวกับข้าวจึงมีบทบาทสำคัญในการจำแนกพันธุ์ข้าว การศึกษาในครั้งนี้มีวัตถุประสงค์เพื่อนำเสนอแนวคิดในการใช้คอมพิวเตอร์สร้างปัญญาประดิษฐ์เพื่อการจำแนกพันธุ์ข้าว โดยใช้วิธีตรรกะฟัซซีและต้นไม้ตัดสินใจเพื่อดำเนินการระบุสายพันธุ์ข้าว โดยการวิเคราะห์ข้อมูลทางสัณฐานวิทยาของเมล็ดข้าวเปลือกประกอบด้วยสองขั้นตอน ในขั้นแรกได้นำเทคนิคต้นไม้ตัดสินใจกรองข้อมูลคุณภาพของสัณฐานวิทยาของเมล็ดข้าวเปลือก ต่อจากนั้นข้อมูลเชิงปริมาณอื่น ๆ ของสัณฐานวิทยาของเมล็ดข้าวจะถูกประมวลผลด้วยเทคนิคฟัซซีลอจิกเพื่อจำแนกพันธุ์ข้าว ในการทดลองใช้พันธุ์ข้าวทั้งหมด 26 สายพันธุ์ และประเมินข้อมูลลักษณะทางสัณฐานวิทยาของเมล็ดข้าวทั้ง 12 แบบเพื่อจำแนกพันธุ์ข้าว ผลการศึกษาสรุปได้ว่าระบบคอมพิวเตอร์และเทคนิคปัญญาประดิษฐ์ในวิธีที่นำเสนอสามารถระบุสายพันธุ์ข้าวได้ตามวัตถุประสงค์ทั้ง 26 สายพันธุ์ และสายพันธุ์ข้าวที่มีค่าสัดส่วนของการจำแนกสายพันธุ์ที่ถูกต้องมากกว่า 70% จำนวน 18 สายพันธุ์

คำสำคัญ: การระบุสายพันธุ์ข้าว, สัณฐานวิทยา, ต้นไม้ตัดสินใจ, ตรรกะฟัซซี

Abstract

The collection of various rice varieties is a way of rice seed conservation. The conservation of rice varieties solves the problem of shortage of pure and good quality rice seeds as well as helps to conserve rice varieties from being lost. Moreover, the selected rice varieties can be used as good breeders for future breeding. The

Received: August 05, 2022

Revised: September 12, 2022

Accepted: September 12, 2022

Available online: December 20, 2022

morphological characteristics of rice varieties were used as the components for rice variety identification. Therefore, rice specialists and tools are an important role in rice variety identification. The aim of the study was to present the idea of using computers to build an artificial intelligence method for rice variety identification. The fuzzy logic and decision tree techniques were used to perform the proposed system for problem-solving rice variety identification. The morphological analysis of paddy seed is composed of two steps. In the first step, the decision tree technique was applied to filter the quality data of the paddy seed morphology. Subsequently, the other quantitative data of the paddy seed morphology were processed with fuzzy logic techniques to identify the rice varieties. A total of twenty-six rice varieties were used in the experiment as the proposed method and twelve characteristics of paddy seed morphological data were evaluated to identify rice varieties. The results summarized that the computer system and artificial intelligence technique in the proposed method successfully identified rice variety identification. The eighteen rice varieties of twenty-six rice varieties demonstrated a proportion of true prediction of more than 70% from a total prediction.

Keywords: Rice variety identification, Morphology, Decision tree, Fuzzy logic

1 บทนำ

ประเทศไทยเป็นประเทศเกษตรกรรมมีความหลากหลายทางพันธุศาสตร์ของพืช โดยเฉพาะข้าวซึ่งเป็นอาหารหลักของคนไทยและยังเป็นสินค้าส่งออกของประเทศไทยในอันดับต้นๆ จากภูมิประเทศที่แตกต่างกันของประเทศไทยในแต่ละภูมิภาคส่งผลให้ข้าวในแต่ละพื้นที่มีการปรับปรุงสายพันธุ์โดยธรรมชาติเพื่อให้มีความคงทนต่อสภาวะแวดล้อมนั้นๆ และจากการปรับตัวของพืชในลักษณะนี้จึงทำให้เกิดสายพันธุ์ข้าวที่มีลักษณะแตกต่างกัน ทั้งในด้านลักษณะทางสัณฐานวิทยา หรือแม้แต่วิทยาศาสตร์ของข้าว ตลอดจนคุณค่าทางโภชนาการของข้าวสายพันธุ์นั้นๆ ซึ่งได้มีผู้ที่สนใจศึกษาถึงสายพันธุ์ข้าวต่างๆ เหล่านี้ เพื่อเก็บรวบรวมสายพันธุ์ และการพัฒนาปรับปรุงพันธุ์กรรมข้าวอยู่มากมาย โดยเฉพาะได้มีการศึกษาถึงสายพันธุ์ข้าวพื้นเมืองของในแต่ละพื้นที่ เพื่อเก็บรวบรวมข้อมูลลักษณะทางสัณฐานวิทยาของสายพันธุ์ข้าวเหล่านั้นเพื่อใช้ประโยชน์ต่อไป

การจำแนกสายพันธุ์ข้าวเป็นการศึกษาในขั้นต้น สำหรับการเก็บรวบรวมสายพันธุ์ หรือการปรับปรุงพันธุ์กรรม ซึ่งจะต้องพิจารณาถึงองค์ประกอบต่างๆ มากมาย เช่น ลักษณะทางสัณฐานวิทยา กายวิภาคศาสตร์ คุณสมบัติทางเคมี ของข้าวชนิดนั้นที่กำลังศึกษา ซึ่งสิ่งต่างๆ เหล่านี้จำเป็นจะต้องอาศัยเครื่องมือ และผู้เชี่ยวชาญในการวิเคราะห์ โดยปกติแล้วในการจำแนกสายพันธุ์ข้าวจะพิจารณาจากลักษณะทางสัณฐานวิทยาและกายวิภาคศาสตร์เป็นขั้นต้น เช่น สีของข้าวเปลือก รูปร่างข้าวเปลือก ขนาดความกว้าง-ยาวข้าวเปลือก สีข้าวกล้อง รูปร่างข้าวกล้อง ขนาดความกว้าง-ยาวข้าวกล้อง เป็นต้นเพื่อให้สามารถระบุสายพันธุ์ได้

(ธีระ และคณะ, 2555) ซึ่งได้มีการศึกษาถึงวิธีการ และเครื่องมือในการจำแนกสายพันธุ์ข้าว ในลักษณะต่างๆ กันอย่างมากมาย

มีการนำเอาระบบปัญญาประดิษฐ์ (artificial intelligence) มาประยุกต์ใช้เพื่อแก้ปัญหาในการจำแนกวัตถุอยู่หลายลักษณะ โดยการนำเอาวิธีการของ machine learning เช่น ต้นไม้ตัดสินใจ (decision tree) ตรรกะฟัซซี (Fuzzy Logic), โครงข่ายประสาทเทียม (artificial neural network, ANN) มาใช้ในการแก้ปัญหา OuYang et al. (2010) ได้รายงานผลการศึกษารูปแบบแยกแยะความแตกต่างของเมล็ดพันธุ์ข้าวด้วยวิธีการเรียนรู้ของเครื่อง (machine learning) โดยการนำเอาข้อมูลทางสัณฐานวิทยาของเมล็ดข้าวบางประการมาใช้สำหรับการวิเคราะห์เช่นกัน และยังมีงานวิจัยที่ศึกษาถึงการนำเอาเทคนิคของตรรกะฟัซซีและต้นไม้ตัดสินใจมาออกแบบระบบผู้เชี่ยวชาญสำหรับการจัดเรียงถั่วพิตาชิโอ (Safavian and Landgrebe, 1991) ซึ่งแสดงให้เห็นว่าสามารถนำเอาเทคนิคตรรกะฟัซซีและต้นไม้ตัดสินใจมาใช้เพื่อประมวลผลข้อมูลเมล็ดธัญพืชได้ซึ่งวิธีการต่าง ๆ ที่กล่าวมามีข้อดีและข้อเสียที่แตกต่างกัน ความซับซ้อนของระบบก็แตกต่างกัน ความยุ่งยากในการออกแบบก็แตกต่างกัน ซึ่งในบทความนี้ได้หยิบเอาข้อดีของวิธีต้นไม้ตัดสินใจ และตรรกะฟัซซี ที่เป็นการตัดสินใจแบบมีการถ่วงน้ำหนักไม่ใช่การตัดสินใจแบบเด็ดขาด (ถูก หรือ ผิด) แต่จะมีค่าระดับความจริงของเงื่อนไขการตัดสินใจเป็นค่าความเป็นสมาชิกของฟัซซีเซต ซึ่งจะสามารถตัดสินใจในเงื่อนไขที่มีความไม่แน่นอนของข้อมูลได้ดีกว่าการตัดสินใจแบบเด็ดขาด โดยได้มีการนำเอาตรรกะฟัซซีมาประยุกต์ใช้ในการประเมินความสุกของแอปเปิ้ลโดยการวิเคราะห์ค่าสีของผลแอปเปิ้ลโดยใช้ตรรกะฟัซซี Mulyani and J. P. Susanto (2017)

ซึ่งจะมีความซับซ้อนน้อยกว่าวิธีการโครมาทอกราฟี โดยได้ออกแบบระบบด้วยวิธีตัดไม้ตัดลิ้นใจให้จำแนกกลุ่มของสายพันธุ์ข้าวในเบื้องต้น จากข้อมูลลักษณะทางสัณฐานวิทยาที่เป็นค่าข้อมูลเชิงคุณภาพ เช่น สีของข้าวเปลือก รูปร่างข้าวเปลือก สีข้าวกล้อง เป็นต้น และใช้ตรรกะฟัซซีซึ่งมีข้อดีในการตัดสินใจที่มีข้อมูลคลุมเครือ มาช่วยจำแนกสายพันธุ์ข้าวจากข้อมูลเชิงปริมาณของสายพันธุ์ข้าว เช่น ความยาวข้าวเปลือก ความกว้างข้าวเปลือก จำนวนขนที่กบาล่าง เป็นต้น เพื่อศึกษาวิธีการประมวลผลข้อมูลสำหรับการระบุสายพันธุ์ข้าว

ดังนั้นในงานวิจัยนี้จึงมีวัตถุประสงค์ในการศึกษาถึงการประยุกต์ใช้เทคนิคตัดไม้ตัดลิ้นใจและตรรกะฟัซซี เพื่อการระบุสายพันธุ์ข้าวเปลือก โดยได้ทำการศึกษาจากข้อมูลสัณฐานวิทยาของข้าวเปลือก จำนวน 26 สายพันธุ์ที่ได้มีการรวบรวมข้อมูลตามรายงานการศึกษาลักษณะทางสัณฐานวิทยาและกายวิภาคศาสตร์ของข้าวเปลือกข้าวต่างนาสวนในภาคตะวันออกเฉียงเหนือของประเทศไทย (ธีระ และคณะ, 2555) ซึ่งสามารถตรวจสอบย้อนกลับได้ถึงประสิทธิภาพของวิธีการที่ได้ศึกษา

2 อุปกรณ์และวิธีการ

2.1 ข้าวและการจำแนกสายพันธุ์

ข้าวมีผลแบบธัญพืช (caryopsis) คือ มีผนังผล (pericarp) เชื่อมกับเปลือกหุ้มเมล็ด (seed coat) รวมเรียกว่าเปลือกข้าวกล้อง (brown rice-coat) ซึ่ง ถูกหุ้มด้วยเปลือกข้าว (hull) อีกชั้นหนึ่ง ในทาง พฤกษศาสตร์เปลือกข้าวประกอบด้วย ด้วยเปลือกใหญ่ (lemma) และเปลือกเล็ก (palea) ทั้งผลและเปลือกข้าว รวมเรียกว่า ข้าวเปลือก (grain) ข้าวเปลือก และข้าวกล้องมีสีและขนาดแตกต่างกันไปในแต่ละพันธุ์ ข้าวเปลือกเป็นตัวอย่างเมล็ดพันธุ์ เพื่อการอนุรักษ์พันธุกรรมข้าว ซึ่งสถาบันวิจัยข้าวของประเทศไทย เรียกว่า เมล็ดเชื้อพันธุ์ข้าว (germplasm) (สุรีย์ และคณะ, 2539) ซึ่งสามารถแสดงลักษณะของข้าวเปลือกได้ดัง

Figure 1

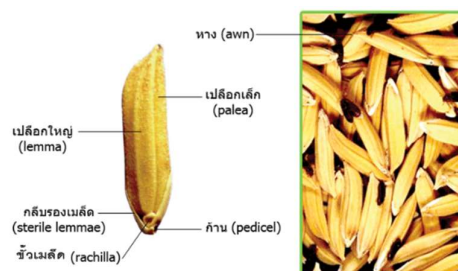


Figure 1 Whole grain rice (กองวิจัยและพัฒนาข้าว กรมการข้าว, 2559)

หากใช้ข้อมูลคุณสมบัติของเมล็ดข้าวดังกล่าวนี้มาใช้จำแนกสายพันธุ์ข้าวจะสามารถทำได้ตามหลักการทางคณิตศาสตร์ ซึ่งสอดคล้องกับ Appa et al. (1997) ที่รายงานถึงลักษณะสีและขนาดของข้าวเปลือก การมี หางข้าว การมีขนบนเปลือกข้าว ว่าสามารถนำมาใช้ในการแบ่งกลุ่มพันธุ์ปลูกของข้าวได้ และได้มีรายงานการศึกษาลักษณะทางสัณฐานวิทยาและกายวิภาคศาสตร์ของข้าวเปลือกข้าวต่างนาสวนในภาคตะวันออกเฉียงเหนือของประเทศไทย (ธีระ และคณะ 2555) ซึ่งแสดงตัวอย่างภาพเมล็ดข้าวดัง Figure 2 (ก)ข้าวเก่า1(LLR 028) (ข)ข้าวเก่า(LLR139) (ค)ข้าวเจ้าหอมด้าสุตะบุตร (LLR262) (ง)ข้าวเจ้าหอมนิล (LLR284) (จ)ข้าวเก่า (LLR077) (ฉ)ข้าวเก่า (LLR079) (ช)ข้าวเหนียวดำ (LLR006) (ซ)ข้าวเก่า (LLR026) (ฌ)ข้าวอีตาน้อย (LLR039)และได้แสดงผลของการเก็บข้อมูลในการศึกษาตาม Table 1 ได้แก่ ข้าวเหนียวดำ (LLR 005) ข้าวเหนียวดำ (LLR 006) ข้าวเหนียวดำ (LLR 015) ข้าวเหนียวดำ (LLR 019) ข้าวนาสวน (LLR 022) ข้าวเก่า (LLR 026) ข้าวเก่า (LLR 027) ข้าวเก่า 1 (LLR 028) ข้าวเก่า 2 (LLR 030) อีตาน้อย (LLR 039) ก้าเก่า (LLR 065) ก้า (LLR 069) ก้าฟ้าเหลือง 1 (LLR 073) ก้าฟ้าเหลือง 2 (LLR 075) ข้าวเก่า (LLR 077) ข้าวเก่า (LLR 078) ข้าวเก่า (LLR 079) ข้าวเก่า (LLR 080) ข้าวเก่า (LLR 081) ข้าวเก่า (LLR 139) ข้าวเก่า (LLR 145) ข้าวเก่า (LLR 182) เจ้าหอมด้าสุตะบุตร (LLR 262) เจ้าหอมนิล120 (LLR 284) ดำแม่แมวทนแล้ง (LLR 292) และเหนียวด้าอูบล (LLR 298) ซึ่งในบทความนี้จะใช้ตัวอย่างข้อมูลลักษณะทางสัณฐานวิทยาและกายวิภาคศาสตร์ของข้าวเปลือกข้าวต่างนาสวน ทั้ง 26 สายพันธุ์ที่ได้รายงานการศึกษาไว้ (ธีระ และคณะ, 2555) มาทำการประมวลผลเพื่อหาค่าความแม่นยำในการระบุสายพันธุ์ข้าวตามอัลกอริธึมที่ได้ออกแบบในงานวิจัยนี้ เนื่องจากเป็นข้อมูลที่ได้ทำการศึกษา และวิจัยแล้วมีความถูกต้องของข้อมูล สามารถตรวจสอบย้อนกลับได้ในภายหลัง และ

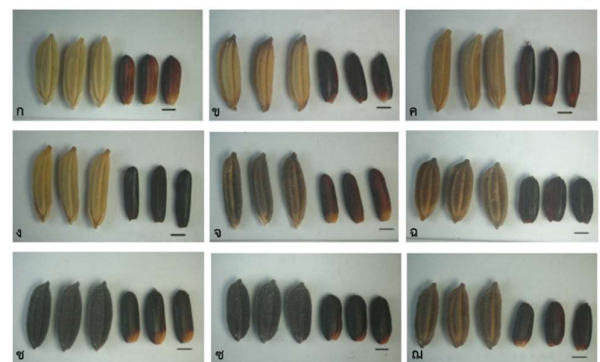


Figure 2 Rice seed varieties (ธีระ และคณะ, 2555)

2.2 การสร้างแผนผังต้นไม้ตัดสินใจ

ต้นไม้ตัดสินใจเป็นเทคนิคหนึ่งของระบบปัญญาประดิษฐ์ที่ถูกนำมาแก้ปัญหาที่ซับซ้อนในการตัดสินใจได้อย่างเป็นระบบโดยการพิจารณาค่า information gain ของข้อมูลเพื่อประกอบการตัดสินใจ และจะวางโครงสร้างของการตัดสินใจในรูปแบบโครงสร้างของต้นไม้ (Safavian and Landgrebe, 1991) ในการสร้างระบบต้นไม้ตัดสินใจสำหรับบทความนี้ได้พิจารณาข้อมูลใน Table 1 ซึ่งทำการเลือกเอาข้อมูล สีและรูปร่างของข้าวเปลือก รูปร่างและการกระจายของสีในข้าวกล้อง และลักษณะความหนาของเปลือกข้าวกล้อง ซึ่งกองวิจัยและพัฒนาข้าว กรมการข้าว. (2559) ได้รายงานว่ามีสีของข้าวเปลือกจะแยกออกเป็น สีขาว สีฟาง สีม่วง และสีดำ สีของข้าวกล้องส่วนใหญ่มีสีขาว บางพันธุ์มีสีแดงน้ำตาลหรือสีม่วงจนเกือบดำ และธีระ และคณะ (2555) ได้รายงานผลการศึกษารูปร่างของข้าวเปลือกและข้าวกล้อง โดยใช้อัตราส่วนความยาวต่อความกว้าง (L/W) สามารถแบ่งรูปร่างของข้าวเปลือกออกเป็น 3 แบบ ได้แก่ รูปร่างป้อม ($L/W < 3.1$) รูปร่างเรียว ($L/W = 3.3-4.2$) และรูปร่างเรียวยาว ($L/W > 4.2$) รูปร่างของข้าวกล้อง แบ่งเป็น 4 แบบ ได้แก่ รูปร่างเรียวยาว (slender, $L/W > 3.0$) รูปร่างเรียว (medium, $L/W = 2.0-3.0$) รูปร่างป้อม (bold, $L/W = 1.0-2.0$) และรูปร่างกลม (round,

$L/W < 1.1$) และลักษณะความยาวของข้าวกล้อง แบ่งเป็น 4 แบบ ได้แก่ ข้าวกล้องยาวมาก (extra-long, ยาวมากกว่า 7.60 มม.) ข้าวกล้องยาว (long, ยาว 6.6-7.50 มม.) ข้าวกล้องสั้น (medium, ยาว 5.51-6.60 มม.) และข้าวกล้องสั้นมาก (short, ยาวน้อยกว่า 5.51 มม. ซึ่งถูกนำมาใช้มาสร้างระบบต้นไม้ตัดสินใจ โดยใช้ค่า information gain เป็นตัวตัดสินใจเลือก attribute เพื่อมาเป็น node ซึ่งค่า information gain สามารถหาได้จากค่า entropy ของข้อมูล (Dengchao et al., 2013) ดังสมการที่ (1) ซึ่งเป็นสมการตามทฤษฎีของการหาค่า entropy และสามารถสร้างแผนผังโครงสร้างของต้นไม้ตัดสินใจได้ดัง Figure 3 ซึ่งสามารถใช้ต้นไม้ตัดสินใจแยกกลุ่มของสายพันธุ์ข้าวออกเป็น 10 กลุ่ม และมีบางกลุ่มสามารถระบุสายพันธุ์ข้าวได้เมื่อเสร็จสิ้นกระบวนการต้นไม้ตัดสินใจ แต่ในกลุ่มที่มีสายพันธุ์ข้าวมากกว่า 1 สายพันธุ์ จะต้องนำเข้าสู่กระบวนการของตรรกะฟัซซีต่อไป

$$entropy(p_1, p_2, \dots, p_n) = -p_1 \log_2 p_1 - p_2 \log_2 p_2 \dots - p_n \log_2 p_n \quad (1)$$

เมื่อ $p_1 + p_2 + \dots + p_n = 1$

Table 1 The morphology of twenty-six paddy varieties

Number	Code	Color	Shape	Length (L) (mm.)	Width (W) (mm.)	Texture of Lemma
1	LLR 005	Black	Bold	8.50-8.99	3.44-3.66	Smooth
2	LLR 006	Black	Bold	9.96-10.60	3.40-3.72	Smooth
3	LLR 015	Black	Bold	8.28-8.99	3.24-3.57	Smooth
4	LLR 019	Black	Bold	8.72-9.14	3.28-3.66	Rough
5	LLR 022	Black	Medium	10.00-10.90	2.76-3.05	Smooth
6	LLR 026	Black	Bold	8.59-8.94	3.25-3.48	Smooth
7	LLR 027	Brown	Bold	8.41-8.97	3.14-3.36	Rough
8	LLR 028	Brown	Bold	9.00-9.89	3.17-3.35	Rough
9	LLR 030	Black	Bold	8.57-9.31	3.40-3.60	Smooth
10	LLR 039	Black	Bold	8.73-9.46	3.24-3.51	Rough
11	LLR 065	Black	Bold	9.04-9.56	3.09-3.49	Rough
12	LLR 069	Black	Bold	9.03-9.45	3.20-3.60	Rough
13	LLR 073	Black	Bold	8.99-9.36	3.44-3.67	Rough
14	LLR 075	Black	Bold	9.00-9.38	3.47-3.67	Rough
15	LLR 077	Black	Medium	9.30-10.60	2.66-2.96	Rough
16	LLR 078	Black	Bold	8.48-8.87	3.59-3.61	Rough
17	LLR 079	Black	Bold	8.13-8.76	3.15-3.75	Rough
18	LLR 080	Black	Bold	8.65-9.12	3.29-3.59	Rough
19	LLR 081	Black	Bold	8.98-9.38	3.41-3.62	Rough
20	LLR 139	Brown	Bold	8.57-9.05	3.01-3.28	Smooth
21	LLR 145	Black	Bold	8.71-8.88	3.11-3.30	Rough
22	LLR 182	Black	Bold	8.65-9.12	3.29-3.59	Rough
23	LLR 262	Brown	Slender	10.30-11.30	2.41-2.75	Rough
24	LLR 284	Brown	Medium	10.20-10.80	2.60-2.84	Rough
25	LLR 292	Black	Bold	8.66-9.17	3.09-3.29	Rough
26	LLR 298	Black	Bold	8.56-9.05	3.42-3.68	Rough

*ข้อมูลจากรายงานการศึกษาลักษณะทางสัณฐานวิทยาและกายวิภาคศาสตร์ของข้าวเปลือกข้าวดำนาสวนในภาคตะวันออกเฉียงเหนือของประเทศไทย (ธีระ และคณะ, 2555)

Table 2 The morphology of twenty-six brown rice varieties

Number	Code	Shape	Distribution of Black	Length (L) (mm.)	Width (W) (mm.)	Peel thickness
1	LLR 005	Bold	Irregular	5.92-6.50	2.57-2.92	15.0-17.5
2	LLR 006	Bold	Irregular	7.06-7.41	2.63-2.78	15.0-17.5
3	LLR 015	Bold	Irregular	5.59-6.17	2.60-2.79	22.5-25.0
4	LLR 019	Bold	Regular	5.96-6.12	2.38-2.93	10.0-15.0
5	LLR 022	Medium	Irregular	6.72-7.40	2.15-2.44	22.5-25.0
6	LLR 026	Bold	Irregular	6.21-6.54	2.54-2.79	22.5-25.0
7	LLR 027	Bold	Irregular	5.95-6.16	2.51-2.68	22.5-25.0
8	LLR 028	Bold	Irregular	6.60-6.97	2.32-2.73	10.0-17.5
9	LLR 030	Bold	Irregular	6.11-6.55	2.66-2.80	22.5-25.0
10	LLR 039	Bold	Irregular	6.18-6.64	2.57-2.79	22.5-25.0
11	LLR 065	Bold	Regular	6.28-6.55	2.48-2.77	15.0-17.5
12	LLR 069	Bold	Regular	6.32-6.55	2.48-2.89	12.5-17.5
13	LLR 073	Bold	Irregular	5.97-6.45	2.50-2.80	15.0-17.5
14	LLR 075	Bold	Regular	6.01-6.49	2.68-2.87	15.0-17.5
15	LLR 077	Medium	Irregular	7.17-7.55	2.30-2.45	10.0-15.0
16	LLR 078	Bold	Regular	5.89-6.28	2.70-3.02	22.5-25.0
17	LLR 079	Bold	Regular	5.58-6.00	2.58-2.89	22.5-25.0
18	LLR 080	Bold	Regular	6.11-6.46	2.62-2.83	22.5-27.5
19	LLR 081	Bold	Regular	6.03-6.50	2.62-2.80	12.5-15.0
20	LLR 139	Bold	Irregular	5.69-6.19	2.32-2.53	10.0-15.0
21	LLR 145	Bold	Irregular	5.93-6.43	2.14-2.67	12.5-15.0
22	LLR 182	Bold	Regular	6.11-6.46	2.62-2.83	12.5-17.5
23	LLR 262	Medium	Regular	6.98-7.61	1.89-2.22	10.0-17.5
24	LLR 284	Medium	Regular	7.16-7.74	2.02-2.13	12.5-17.5
25	LLR 292	Bold	Irregular	6.04-6.30	2.40-2.62	22.5-27.5
26	LLR 298	Bold	Irregular	5.84-6.24	2.63-2.83	22.5-27.5

*ข้อมูลจากรายงานการศึกษาลักษณะทางสัณฐานวิทยาและกายวิภาคศาสตร์ของข้าวเปลือกข้าวดำนาสวนในภาคตะวันออกเฉียงเหนือของประเทศไทย (ธีระ และคณะ, 2555)

2.3 การออกแบบตรรกะพีซี

เมื่อนำข้อมูลข้าวจาก Table 1 ผ่านกระบวนการต้นไม้ตัดสินใจ จะเหลือกลุ่มของสายพันธุ์ข้าวที่ต้องวิเคราะห์เพียง 6 กลุ่ม ส่วนอีก 4 กลุ่ม ได้แก่ สายพันธุ์ ข้าวเก่า (LLR 139) ข้าวเก่า (LLR 027) ข้าวเจ้าหอมดำสุตะบุตร (LLR 262) และข้าวเจ้าหอมนิล 120 (LLR 284) สามารถระบุได้โดยไม่ต้องผ่านกระบวนการตรรกะพีซี เนื่องจากต้นไม้ตัดสินใจสามารถระบุ

สายพันธุ์ข้าวของทั้ง 4 กลุ่ม ได้แก่ Group 2, Group 6, Group 9 และ Group 10 ได้เรียบร้อยแล้วดัง Figure 3 ที่สร้างมาจากการประมวลผลข้อมูลในงานวิจัยนี้ตามหลักการของการสร้างโครงสร้างต้นไม้ตัดสินใจ โดยในขั้นตอนการออกแบบตรรกะพีซี จะเริ่มต้นด้วยการคำนวณหาฟังก์ชันการตัดสินใจในแต่ละกลุ่ม โดยการนำข้อมูลข้าวที่เป็นข้อมูลเชิงปริมาณในแต่ละชนิดมาสุ่มสร้างข้อมูลใหม่ แต่ยังคงมีลักษณะสัณฐานในแต่ละสายพันธุ์เดิม จากนั้นนำข้อมูลในแต่ละกลุ่มมาวิเคราะห์ฮิสโตแกรมของ

ข้อมูลในแต่ละ Feature เพื่อหาค่าเทรโดโลสำหรับกำหนดตำแหน่งการตัดสินใจสำหรับตระกูลพืชซีของแต่ละกลุ่ม โดยค่าตำแหน่งเทรโดโลจะเป็นจุดที่แบ่งข้อมูลฮิสโตรแกรมออกเป็นสองชุดโดยใช้หลักการทางสถิติ Figure 4 แสดงถึงฮิสโตรแกรมของลักษณะเด่นทั้ง 7 กลุ่มข้อมูลที่เป็นข้อมูลเชิงปริมาณของสายพันธุ์ข้าว ได้แก่ Feature 1 (ความยาวข้าวเปลือก), Feature 2 (ความกว้างข้าวเปลือก), Feature 3 (ความยาวข้าวกล้อง), Feature 4 (ความกว้างข้าวกล้อง), Feature 5 (ความหนาเปลือกข้าวกล้อง), Feature 6 (จำนวนขนที่กาบล่าง) และ Feature 7 (ความยาวขนขนาดใหญ่ของกาบล่าง) แต่ละกลุ่มข้อมูล และนำมาสร้างฟังก์ชันการตัดสินใจของตระกูลพืชซี เพื่อกำหนดค่าของข้อมูลสมาชิกพืชซีในเซตแบบฉบับ (classic set) โดยใช้สมการที่ (2) และการกำหนดค่าความเป็นสมาชิกของพืชซีเซต ดังสมการที่ (3) เมื่อ A คือพืชซีเซต x คือสมาชิกในเซต และ μ_A คือค่าความเป็นสมาชิกในเซต โดยได้กำหนดเงื่อนไขการตัดสินใจการเป็นสมาชิกของพืชซีเซตไว้ที่ 70%

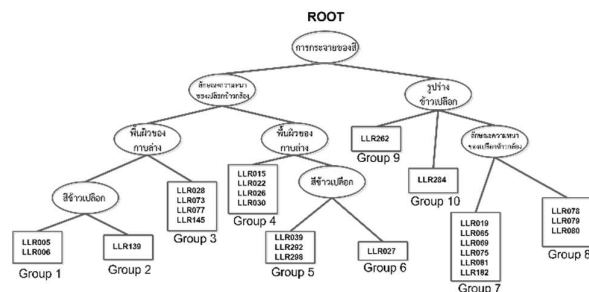


Figure 3 Decision tree construction for rice seed varieties classification.

$$\mu_A(x): X \rightarrow [0,1] \quad (2)$$

$$A = \{(x \cdot \mu_A(x)) | x \in X\} \quad (3)$$

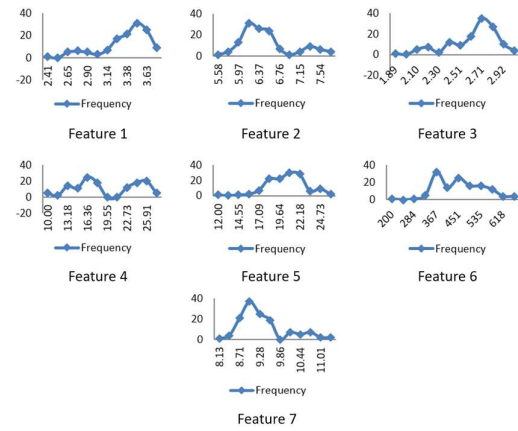


Figure 4 The 7 Features morphology data of the rice seed varieties.

Table 3 แสดงค่าลักษณะของกลุ่มต่าง ๆ ที่ได้จากการวิเคราะห์ฮิสโตรแกรมใน Figure 4 เพื่อแทนในสมการที่ (2) และนำไปสร้างเป็นกฎในการตัดสินใจของตระกูลพืชซี สำหรับสายพันธุ์ข้าวทั้ง 26 สายพันธุ์ ได้ทั้งหมดจำนวน 26 กฎตามจำนวนสายพันธุ์ข้าวที่นำมาทดสอบในงานวิจัยนี้ เพื่อจำแนกสายพันธุ์ข้าวทั้ง 26 สายพันธุ์

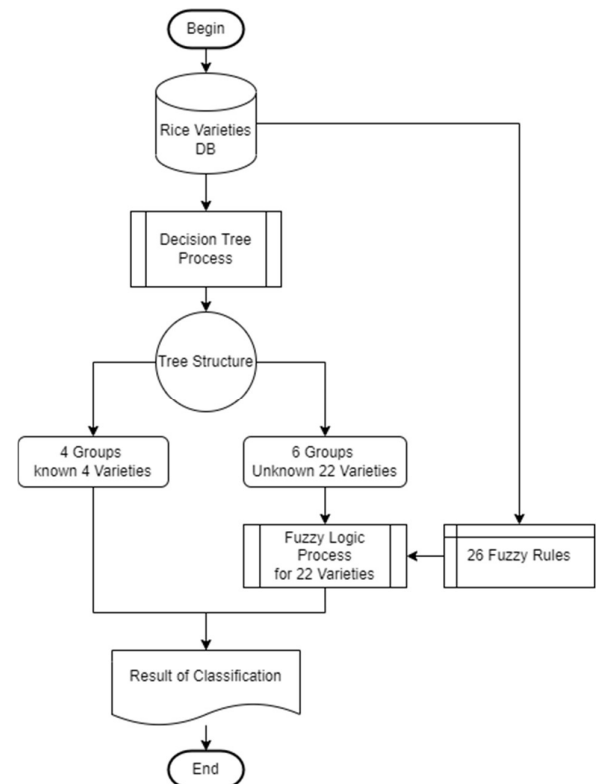


Figure 5 Block diagram of rice varieties classification

2.4 ขั้นตอนการวิเคราะห์ข้อมูลเพื่อจำแนกสายพันธุ์ข้าว

เริ่มต้นกระบวนการวิเคราะห์โดยการดึงข้อมูลข้าวที่ต้องการวิเคราะห์เข้าสู่ระบบตามผังงานดัง Figure 5 โดยนำข้อมูลข้าวมา

ผ่านกระบวนการแรก คือ ระบบต้นไม้ตัดสินใจ เพื่อสร้างโครงสร้างต้นไม้ โดยใช้ลักษณะพื้นฐานของข้าว แบ่งออกเป็น 10 กลุ่ม สามารถระบุสายพันธุ์ได้ 4 กลุ่ม รวม 4 สายพันธุ์ และไม่สามารถระบุสายพันธุ์ได้ 6 กลุ่ม รวม 22 สายพันธุ์ จากนั้นนำข้อมูลข้าวทั้ง 22 สายพันธุ์ ที่ไม่สามารถระบุสายพันธุ์ได้ผ่านเข้าสู่กระบวนการตรรกะฟัซซี ซึ่งได้ออกแบบกฎในการตัดสินใจเอาไว้ รวมทั้งสิ้น 26 กฎ แต่ใช้เพียง 22 กฎของสายพันธุ์ที่ไม่สามารถระบุสายพันธุ์ได้ เมื่อผ่านกระบวนการประมวลผลต่าง ๆ ที่ได้กำหนดขึ้นจะสามารถระบุสายพันธุ์ข้าวได้ครบทั้ง 26 สายพันธุ์

ในขั้นตอนการทดลองระบุสายพันธุ์ข้าวได้นำเอาข้อมูลการศึกษาลักษณะทางสัณฐานวิทยาและกายวิภาคศาสตร์ของข้าวเปลือกข้าวดำนาสวนในภาคตะวันออกเฉียงเหนือของประเทศไทยซึ่งได้เก็บรวบรวมข้อมูลของข้าวถึง 26 สายพันธุ์ มาทำการเพิ่มจำนวนของข้าวในแต่ละสายพันธุ์ ด้วยวิธีการสุ่มค่าข้อมูลในส่วนที่เป็นข้อมูลเชิงปริมาณ โดยให้มีค่าอยู่ในช่วงของข้อมูลพื้นฐานของข้าวในแต่ละสายพันธุ์นั้น ๆ โดยเพิ่มจำนวนเป็นสายพันธุ์ละ 100 ข้อมูล และในการทดลองระบุสายพันธุ์กำหนดให้สุ่มข้อมูลสายพันธุ์ที่ถูกต้องจำนวน 5 ชุด และสุ่มข้อมูลสายพันธุ์อื่น ๆ อีกสายพันธุ์ละ 5 ชุด รวมทั้งหมด 130 ชุดข้อมูล

3 ผลและวิจารณ์

ผลการทดลองพบว่าระบบที่ออกแบบสามารถทำงานได้ตามที่ได้ออกแบบ ดังผลการทดลองใน Table 4 และมีค่าสัดส่วนของการจำแนกสายพันธุ์ที่ถูกต้องจากการประมวลผลว่าเป็นสายพันธุ์นั้นทั้งหมด (positive predicted value, PPV) จากการทดลองพบว่าสายพันธุ์ข้าวที่มีค่าสัดส่วนของการจำแนกสายพันธุ์ตามเงื่อนไขการเป็นสมาชิกของฟัซซีเซต มีอยู่ 18 สายพันธุ์ จาก

ทั้งหมด 26 สายพันธุ์ คิดเป็น 69% ของสายพันธุ์ที่ได้ศึกษาสำหรับระบบนี้ จะพบว่าข้าวที่มักจะระบุเป็นหลายชนิดนั้นจะมีลักษณะพื้นฐานของข้าวที่ใกล้เคียงกันมาก ซึ่งจากการทดลองแสดงให้เห็นว่าสายพันธุ์ที่มีความผิดพลาด หรือสัดส่วนของการจำแนกถูกต้องต่ำจะเป็นพันธุ์ข้าวในกลุ่ม 7 ได้แก่ ข้าวเหนียวดำ (LLR 019) ก่ำแกดำ (LLR 065) ก่ำ (LLR 069) ก่ำฟ้าเหลือง 2 (LLR 075) และ ข้าวก่ำ (LLR 182) เนื่องจากมีค่า Feature สำหรับการกำหนดค่าความเป็นสมาชิกของฟัซซีเซตที่ใกล้เคียงกัน

ในกระบวนการออกแบบโครงสร้างต้นไม้ตัดสินใจในการศึกษานี้ได้นำเอาข้อมูลลักษณะพื้นฐานของข้าวที่มีค่าเฉพาะเพียงค่าเดียวสำหรับแต่ละสายพันธุ์ ซึ่งเหมาะกับการสร้างต้นไม้ตัดสินใจ ได้แก่ สีและรูปร่างของข้าวเปลือก รูปร่างและการกระจายของสีในข้าวกล้อง และลักษณะความหนาของเปลือกข้าวกล้อง ซึ่งทำให้ไม่สามารถระบุสายพันธุ์ข้าวได้ครบทั้ง 26 สายพันธุ์โดยวิธีการนี้ ซึ่งข้อมูลลักษณะทางสัณฐานของข้าวในกลุ่มต่าง ๆ ที่แบ่งได้จากต้นไม้ตัดสินใจ จะมีค่าซ้อนทับอยู่ในช่วงเดียวกันของแต่ละลักษณะ ไม่สามารถระบุสายพันธุ์ได้ จึงนำเอาตรรกะฟัซซีที่มีความสามารถในการตัดสินใจในสิ่งที่มีความคลุมเครือมาช่วยในการตัดสินใจระบุสายพันธุ์ในขั้นตอนสุดท้าย

การออกแบบนี้แสดงให้เห็นว่าระบบที่ออกแบบสามารถใช้เทคนิคต้นไม้ตัดสินใจจัดกลุ่มสายพันธุ์ข้าวได้ และในขั้นตอนการประมวลผลด้วยตรรกะฟัซซี จะประมวลผลเฉพาะ 22 กฎของสายพันธุ์ข้าวจำนวน 6 กลุ่ม ที่ไม่สามารถระบุสายพันธุ์ได้โดยต้นไม้ตัดสินใจเท่านั้น ซึ่งจะช่วยลดระยะเวลาในการประมวลผลลงไปได้ ซึ่งหากเป็นระบบตรรกะฟัซซีปกติจะต้องประมวลผลในทุก ๆ กฎที่มีอยู่ในระบบ

Table 3 Data for the fuzzy set from histogram data analysis.

	Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	Feature 6	Feature 7
Group 1	9.51	3.56	6.72	2.73	16.25	21.75	407.50
Group 2*	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Group 3	9.66	3.17	6.74	2.47	13.75	20.50	425.00
Group 4	9.59	3.18	6.49	2.47	23.75	19.50	445.00
Group 5	8.94	3.37	6.21	2.64	25.01	20.33	475.00
Group 6*	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Group 7	8.95	3.47	6.35	2.74	15.00	20.00	460.00
Group 8	8.63	3.45	6.02	2.80	25.00	22.50	420.00
Group 9*	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Group 10*	1.00	1.00	1.00	1.00	1.00	1.00	1.00

* The group of rice varieties can be classified with the decision tree

Table 4 Result of the experiment

Number	Code of Rice Varieties	Number of Correct data	Number of Not Correct data	Number of identified	PPV
1	LLR 005	5	125	5	1.00
2	LLR 006	5	125	5	1.00
3	LLR 015	5	125	12	0.42
4	LLR 019	5	125	15	0.33
5	LLR 022	5	125	5	1.00
6	LLR 026	5	125	7	0.71
7	LLR 027	5	125	5	1.00
8	LLR 028	5	125	10	0.50
9	LLR 030	5	125	6	0.83
10	LLR 039	5	125	4	1.00
11	LLR 065	5	125	17	0.29
12	LLR 069	5	125	19	0.26
13	LLR 073	5	125	5	1.00
14	LLR 075	5	125	8	0.63
15	LLR 077	5	125	5	1.00
16	LLR 078	5	125	5	1.00
17	LLR 079	5	125	5	1.00
18	LLR 080	5	125	9	0.56
19	LLR 081	5	125	5	1.00
20	LLR 139	5	125	5	1.00
21	LLR 145	5	125	5	1.00
22	LLR 182	5	125	30	0.17
23	LLR 262	5	125	5	1.00
24	LLR 284	5	125	5	1.00
25	LLR 292	5	130	6	0.83
26	LLR 298	5	130	7	0.71

4 สรุป

จากการออกแบบระบบโดยใช้วิธีการของต้นไม้ตัดสินใจ และตรรกะฟัซซีร่วมกันประมวลผลข้อมูลเพื่อจำแนกสายพันธุ์ข้าวพบว่าสามารถจำแนกสายพันธุ์ข้าวได้ตามสมมุติฐานที่ตั้งไว้ โดยมีค่าสัดส่วนของการจำแนกสายพันธุ์ที่ถูกต้องต่อการจำแนกทั้งหมดตามเงื่อนไขการเป็นสมาชิกของฟัซซีเซตที่มีค่ามากกว่า 70% มีอยู่จำนวน 18 สายพันธุ์ จากทั้งหมด 26 สายพันธุ์ ซึ่งนับว่าเป็นผลการจำแนกที่ดี ส่วนสายพันธุ์ข้าวที่มักจะระบุได้เป็น

หลายชนิดนั้นจะมีลักษณะสัณฐานของข้าวที่ใกล้เคียงกันมาก ในทางวิทยาศาสตร์การเกษตรนั้น เมื่อเกิดความไม่มั่นใจในลักษณะสายพันธุ์จะต้องนำเข้าสู่กระบวนการทางเคมีวิเคราะห์เพื่อข้อมูลเชิงเคมีประกอบการวิเคราะห์ต่อไป แต่ในมุมมองของการประมวลผลทางคณิตศาสตร์อาจจะต้องมีการปรับปรุงกฎของตรรกะฟัซซีให้เหมาะสมมากขึ้น หรืออาจจะนำเอาข้อมูลอย่างอื่นมาช่วยประกอบการประมวลผล หรือแม้กระทั่งการนำเอา

โครงการขยายประสาทเทียมที่มีความซับซ้อนมากขึ้นมาใช้ในการวิเคราะห์จำแนกสายพันธุ์ข้าว

5 เอกสารอ้างอิง

กองวิจัยและพัฒนาข้าว กรมการข้าว. 2559. คุณภาพข้าว.

แหล่งข้อมูล: <http://webold.ricethailand.go.th/rkb3/title-index.php-file=content.php&id=6-2.htm>. เข้าถึงเมื่อ 9 ตุลาคม 2562.

ธีระ ธรรมวงศา, จิรวัฒน์ สนธิชน, อมรรัตน์ มีสวาสดี, และ ปิยะรัตน์ อัฐรัตน์ 2555. ลักษณะทางสัณฐานวิทยาและกายวิภาคศาสตร์ของข้าวเปลือกข้าวดำนาสวน ในภาคตะวันออกเฉียงเหนือของประเทศไทย, วารสารวิทยาศาสตร์ มข. ฉบับที่ 40. หน้า 1138-1148.

สุรีย์ ศรีวันพนิยกุล, สรรเสริญ ศรีสุนทร, ศิริชัย สมบูรณ์พงษ์, เล็กจันทร์เกษม, สมบัติ รุจาคม และ อรพิน วัฒนเสถียร. 2539. การรวบรวม อนุรักษ์ ประเมิน ลักษณะและจัดหมวดหมู่ข้าวพื้นเมืองในเขตศูนย์วิจัย ข้าวพิษณุโลก. ศูนย์วิจัยข้าวพิษณุโลก สถาบันวิจัย ข้าว กรมวิชาการเกษตร กระทรวงเกษตรและ สหกรณ์. 215 หน้า.

Appa, R.S., Bounphanousoy, C., Kanyavong, K., Phetpaseuth, V., Sengthong, B., Schiller, J.M., Thirasack, S., and Jackson, M.T. 1997. Collection and Classification of Rice Germplasm from the Lao PDR. Part 2: Northern, Southern and Central Regions, LaoIRRI Project. (pp. 41). Vientiane: National Agricultural Research Center, Lao PDR.

Dengchao, H., Wenning, H., Wenyan, G., Gang C., and Dawei, J. 2013. A Decision Tree Algorithm Based on Dispersion Measure of Attribute Information. 2013 10th Web Information System and Application Conference. pp. 84-89.

Mahmoud, O. 2011. Design of an expert system for sorting pistachio nuts through decision tree and fuzzy logic classifier. Expert Systems with Applications. 38(4), 4339-4347.

Mulyani, E. D. S., and Susanto, J. P. 2017. Classification of maturity level of fuji apple fruit with fuzzy logic method. 2017 5th International Conference on Cyber and IT Service Management (CITSM), pp. 1-4.

OuYang, A.G., Gao, R. J., Liu, Y. D., Sun, X. D., Pan, Y. Y., and Xiao-ling Dong. 2010. An automatic method for identifying different variety of rice seeds using machine vision technology. 2010 Sixth International Conference on Natural Computation. pp. 84-88.

Safavian, S. R., and Landgrebe, D. 1991. A survey of decision tree classifier methodology. IEEE Transactions on Systems, Man, and Cybernetics. 21(3), 660-674.