

วิธีการจำแนกสัญญาณรบกวนในภาพเอกซเรย์ซีอาร์

The method for classification of noise in computed radiography image

บุญเสริม เนยสูงเนิน^{1,2*} บรรจง เชื้อนแก้ว³ เจษฎา ตันthanuch^{4,5}
Boonserm Nerysungnoen^{1,2*} Banjong Kheonkaew³ Jessada Tanthanuch^{4,5}

¹สาขาวิชาเทคโนโลยีสารสนเทศ สำนักวิชาเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีสุรนารี จ. นครราชสีมา

¹School of Information Technology, Institute of Social Technology, Suranaree University of Technology, Nakhon Ratchasima Province, Thailand

²กลุ่มงานรังสีวิทยา โรงพยาบาลมหาราชนครราชสีมา จ. นครราชสีมา

²Department of Radiology, Maharat Nakhon Ratchasima Hospital, Nakhon Ratchasima Province, Thailand

³ภาควิชารังสีวิทยา คณะแพทยศาสตร์ มหาวิทยาลัยขอนแก่น จ. ขอนแก่น

³Department of Radiology, Faculty of Medicine, Khon Kaen University, Khon Kaen Province, Thailand

⁴สาขาวิชาคณิตศาสตร์ สำนักวิชาวิทยาศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี จ. นครราชสีมา

⁴School of Mathematics, Institute of Science, Suranaree University of Technology, Nakhon Ratchasima Province, Thailand

⁵ศูนย์วิจัยความเป็นเลิศด้านชีวกลศาสตร์ทางการแพทย์ มหาวิทยาลัยเทคโนโลยีสุรนารี จ. นครราชสีมา

⁵Center of Excellent in Biomechanics Medicine, Suranaree University of Technology, Nakhon Ratchasima Province, Thailand

* ผู้รับผิดชอบบทความ (Email: booncartoonhm@hotmail.com)

* Corresponding author (Email: booncartoonhm@hotmail.com)

Received February 2017

Accepted as revised March 2017

Abstract

Introduction: Nowadays, conventional X-ray images are replaced by digital X-ray imaging, such as computed radiography (CR) images. Advantages of digital X-rays are, for example, the images can be easily viewed, duplicated and stored. The obtained images can be scaled, measured, compared and adjusted for brightness and contrast, thus enhance diagnosis performance. However, image noise is a key factor that reduces quality of X-ray images. It sometimes causes a deficiency in images leading to misdiagnosis.

Objectives: To classify the noise type in CR images by Naïve Bayes algorithm.

Materials and methods: Firstly, an original image was created for the study. Secondly, a number of instances were created. Each instance was constructed by overlaying various levels of known-noise over the original image. Here, the Gaussian, Poisson and impulse types of noise were considered. Also, the features, mean, SD, MSE, and PSNR were extracted from the modified CR images in this step. Thirdly, the most effective features were selected for classifying type of noise. Fourthly, models were built and evaluated. Finally, noise type in CR image was determined according to the model.

Results: The study was evaluated for classifying noise in CR image, which had 90% correctness.

Conclusion: The study showed that noise in CR system was Poisson noise with precision of 0.95, the recall of 0.73, and the F-measure of 1.13.

Journal of Associated Medical Sciences 2017; 50(2): 275-285. Doi: 10.14456/jams.2017.27

Keywords: Noise, computed radiography, Naïve Bayes

บทคัดย่อ

ที่มาและความสำคัญ: ปัจจุบันภาพเอกซเรย์ใช้ฟิล์มแบบดั้งเดิมส่วนใหญ่ถูกแทนที่ด้วยภาพดิจิทัล เช่น ภาพเอกซเรย์ซีอาร์ ซึ่ง มีข้อดีหลายอย่าง เช่น ภาพเอกซเรย์ดังกล่าวสามารถดูซ้ำง่าย สำเนาได้ ประหยัดพื้นที่ในการจัดเก็บ สามารถขยาย วัดเปรียบเทียบ ปรับความสว่างและความเปรียบต่างได้ง่าย ช่วยให้การวินิจฉัยมีประสิทธิภาพยิ่งขึ้น อย่างไรก็ตาม สัญญาณรบกวน เป็นปัจจัยสำคัญที่ลดคุณภาพของภาพเอกซเรย์ บางครั้งอาจทำให้การวินิจฉัยผิดพลาดได้

วัตถุประสงค์: เพื่อจำแนกชนิดของสัญญาณรบกวนในภาพเอกซเรย์ซีอาร์โดยใช้ขั้นตอนวิธีนาอิวเบส

วิธีการศึกษา: ขั้นตอนการศึกษาประกอบด้วย ขั้นตอนแรกสร้างภาพต้นฉบับสำหรับใช้ศึกษา ขั้นตอนที่สองสร้างกรณี ตัวอย่างด้วยการเติมสัญญาณรบกวนที่ทราบชนิด คือ เกาส์เซียน ปัวซอง และอิมพัลส์ในภาพต้นฉบับ และสกัดคุณลักษณะเฉพาะ คือ ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน ค่าคลาดเคลื่อนกำลังสองเฉลี่ย และค่าฟังก์ชันแนลทูนอยส์เรโซ ขั้นตอนที่สามเลือกคุณลักษณะ เฉพาะที่มีประสิทธิภาพสูงที่สุดในการจำแนกสัญญาณรบกวน ขั้นตอนที่สี่สร้างแบบจำลองและประเมินแบบจำลอง และ ขั้นตอนที่ห้าใช้แบบจำลองจำแนกสัญญาณรบกวนในภาพเอกซเรย์ซีอาร์

ผลการศึกษา: การจำแนกสัญญาณรบกวนในภาพเอกซเรย์ซีอาร์ที่ถูกประเมินมีความถูกต้องเท่ากับร้อยละ 90

สรุปผลการศึกษา: การศึกษาพบว่าสัญญาณรบกวนในภาพเอกซเรย์ซีอาร์ คือ สัญญาณรบกวนปัวซองซึ่งมีความเที่ยงเท่ากับ 0.95 ค่าคั่นคั้นเท่ากับ 0.73 และค่าเอนโทรปีเท่ากับ 1.13

Journal of Associated Medical Sciences 2560; 50(2): 275-285. Doi: 10.14456/jams.2017.27

คำสำคัญ: สัญญาณรบกวน คอมพิวเตอร์ เรดิโอกราฟีส์ นาอิวเบส

บทนำ

ขั้นตอนการประมวลผลภาพของภาพเอกซเรย์ซีอาร์ (computed radiography; CR) อาจมีสัญญาณรบกวนปะปน มากับภาพ และอาจลดคุณภาพของภาพเอกซเรย์ การวิเคราะห์ สัญญาณรบกวนเพื่อให้รู้ชนิดของสัญญาณรบกวนจะทำให้ สามารถใช้วิธีการลดสัญญาณรบกวนในภาพเอกซเรย์ได้ อย่างเหมาะสม เพื่อให้ได้ภาพเอกซเรย์ที่มีสัญญาณรบกวน น้อยที่สุด ทำให้สามารถวินิจฉัยได้อย่างถูกต้องและแม่นยำ เป็นประโยชน์ต่อการรักษาผู้ป่วยต่อไป

สัญญาณรบกวนแบบเกาส์เซียน (Gaussian) มีลักษณะ การแจกแจงสัญญาณรบกวนเป็นไปตามรูปแบบฟังก์ชัน ความหนาแน่นของความน่าจะเป็น (probability density function; PDF) ของเกาส์เซียน และรูปแบบของสัญญาณ รบกวนมีลักษณะคล้ายระฆังคว่ำ² เป็นหนึ่งในรูปแบบ การแจกแจงของสัญญาณรบกวนที่พบได้ทั่วไปมากที่สุด ในธรรมชาติ³ หากจำนวนกลุ่มตัวอย่างมีมากขึ้นจะทำให้การแจกแจง ใดๆ มีแนวโน้มเข้าสู่การแจกแจงปกติ ส่วนการแจกแจงปัวซอง (Poisson) มีรูปร่างขึ้นกับค่าเฉลี่ยเพียงค่าเดียว การแจกแจง

แบบนี้ยังมีประโยชน์ใช้กับจำนวนความสำเร็จหรือเหตุการณ์ ที่สนใจเกิดขึ้นในช่วงเวลาใดเวลาหนึ่ง เช่น การเกิดสัญญาณ รบกวนในภาพเอกซเรย์ และสัญญาณรบกวนอิมพัลส์ (impulse)² เป็นสัญญาณที่สามารถเกิดขึ้นได้จากหลายสาเหตุ เช่น จุดรับภาพ บางจุดในกล้องเสียหายทำให้ส่งสัญญาณภาพ ณ จุดนั้นมีด หรือสว่างกว่าที่ควรจะเป็น การบันทึกตำแหน่งหรือข้อมูลผิดพลาด ในหน่วยความจำ การส่งผ่านข้อมูลในช่องที่มีสัญญาณรบกวน ภาวะชั่วคราว (transient) ความผิดพลาดของการกำหนดเวลาใน การแปลงสัญญาณจากอนาล็อกเป็นดิจิทัล สัญญาณประเภทนี้ เป็นสัญญาณรบกวนที่ปรากฏกระจายอยู่บนจุดภาพของภาพ เป็นจุดๆ ทั่วไป โดยเปลี่ยนค่าความเข้มของจุดภาพที่ ตำแหน่งหนึ่งๆ ให้มีค่าความแตกต่างไปจากจุดภาพข้างเคียง โดยค่าความเข้มที่จำลองจะมีลักษณะความเข้มสีขาวและดำ

แบบจำลองของกระบวนการเติมสัญญาณรบกวน (equation 1) เมื่อ $\eta(x, y)$ คือ สัญญาณรบกวนที่เติมในภาพ ต้นฉบับ $f(x, y)$ คือ ภาพต้นฉบับ และ $g(x, y)$ คือ ภาพ ที่ถูกเติมสัญญาณรบกวน

$$f(x, y) \oplus \eta(x, y) = g(x, y) \dots\dots\dots (\text{Eq. 1})$$

การศึกษานี้เป็นการจำแนกประเภทของสัญญาณรบกวนในภาพเอกซเรย์ซี่อาร์ การจำแนกประเภทเป็นเทคนิคที่จำแนกกลุ่มข้อมูลด้วยคุณสมบัติต่างๆ โดยสำรวจรายการในฐานข้อมูลเพื่อจำแนกให้อยู่ในหมวดหมู่ที่ได้กำหนดไว้ก่อนล่วงหน้า ขั้นตอนการจำแนกประเภทข้อมูลมี 3 ขั้นตอน คือ ขั้นตอนแรก เป็นการสร้างแบบจำลองที่อาศัยการเรียนรู้จากการฝึกสอนให้คอมพิวเตอร์ได้เรียนรู้รูปแบบที่แน่นอนโดยใช้ข้อมูลฝึกสอน ขั้นตอนที่สอง คือขั้นตอนประเมินผลความถูกต้องการทดสอบความถูกต้องของแบบจำลองที่ถูกสร้างขึ้น โดยใช้ข้อมูลทดสอบ และขั้นตอนสุดท้าย เป็นการจำแนกประเภทข้อมูล เป็นการนำแบบจำลองที่ผ่านการตรวจสอบความถูกต้องมาใช้กับข้อมูลที่ไม่ทราบมาก่อน

การจำแนกประเภทข้อมูลด้วยวิธีนาอีฟเบส์ (Naïve Bayes)⁴ เป็นวิธีที่ได้รับความนิยมวิธีหนึ่ง เนื่องจากการสร้างแบบจำลองง่ายและไม่ซับซ้อน อาศัยทฤษฎีความน่าจะเป็น โดยใช้การเรียนรู้แบบเบส์ (Bayesian learning) ซึ่งใช้ทฤษฎีความน่าจะเป็นที่มีพื้นฐานจากทฤษฎีของเบส์ (Bayes theorem) เข้ามาช่วยในการเรียนรู้ จุดมุ่งหมายเพื่อต้องการสร้างแบบจำลองที่อยู่ในรูปของความน่าจะเป็นซึ่งเป็นค่าที่บันทึกได้จากการสังเกต จากนั้นนำแบบจำลองมาหาว่าสมมติฐานใดถูกต้องที่สุดโดยใช้ความน่าจะเป็นเข้ามาช่วย

$$P(A | B) = \frac{P(B | A)P(A)}{P(B)} \dots\dots\dots (\text{Eq. 2})$$

จาก equation 2 กำหนดให้ A และ B เป็นเหตุการณ์ใดๆ $P(A | B)$ (posterior probability) คือ ความน่าจะเป็นของ A เมื่อรู้ค่า B หรือความน่าจะเป็นที่จะเกิดเหตุการณ์ A โดยมีเงื่อนไขว่าเหตุการณ์ B ได้เกิดขึ้นแล้ว สามารถคำนวณได้จากผลคูณของ $P(A | B)$ (likelihood) คือ ความน่าจะเป็นของ B เมื่อรู้ A กับ $P(A)$ (prior probability) คือ ความน่าจะเป็นของ A หากด้วย $P(A)$ คือ ความน่าจะเป็นของ B

แนวคิดของการวิจัยในครั้งนี้ เลือกใช้ภาพที่รู้ชนิดของสัญญาณรบกวนซึ่งมีลักษณะเฉพาะที่ได้สกัดจากภาพ แล้วนำไปฝึกสอนให้คอมพิวเตอร์จดจำรูปแบบ เพื่อให้คอมพิวเตอร์สามารถทำนายสัญญาณรบกวนจากภาพที่ไม่ทราบชนิดของสัญญาณรบกวนโดยใช้ขั้นตอนวิธีของนาอีฟเบส์สำหรับจำแนกสัญญาณรบกวน ในขั้นตอนดังกล่าว คอมพิวเตอร์จะคำนวณความน่าจะเป็นของลักษณะเฉพาะที่ได้ป้อนเข้าไป โดยคำนวณความน่าจะเป็นสูงสุดแล้วทำนายผลลัพธ์ออกมา (Figure 1)

วัตถุประสงค์ของการศึกษา คือ เพื่อจำแนกชนิดของสัญญาณรบกวนในภาพเอกซเรย์ซี่อาร์ ด้วยขั้นตอนวิธีนาอีฟเบส์ โดยใช้หลักการของ machine learning ซึ่งเป็นสาขาหนึ่งของปัญญาประดิษฐ์ (artificial intelligence; AI) หมายถึงการทำให้คอมพิวเตอร์มีความสามารถเรียนรู้จำแบบ เกี่ยวข้องกับการศึกษาและสร้างขั้นตอนวิธีที่สามารถเรียนรู้ข้อมูล และทำนายข้อมูลได้ โดยงานวิจัยนี้ใช้แบบจำลองที่ได้เพื่อทำนายหรือจำแนกชนิดสัญญาณรบกวนที่อยู่ในภาพเอกซเรย์ซี่อาร์

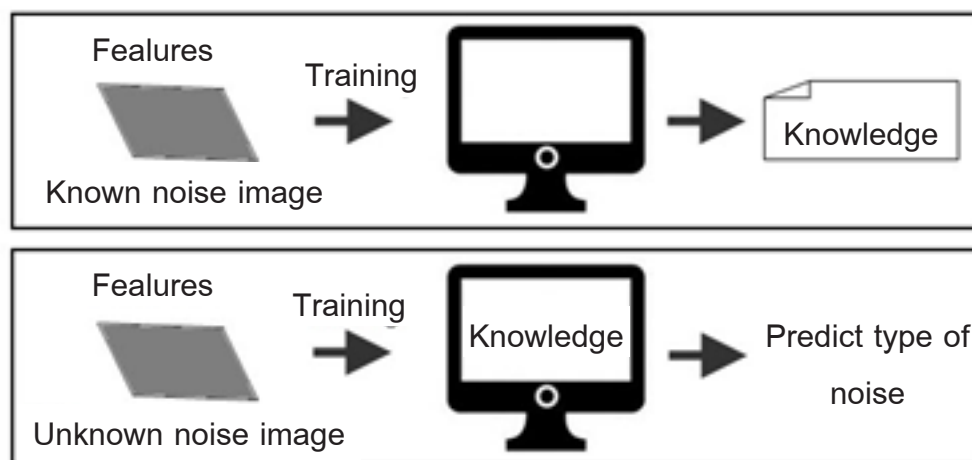


Figure 1 The concept of research.

วิธีการศึกษา

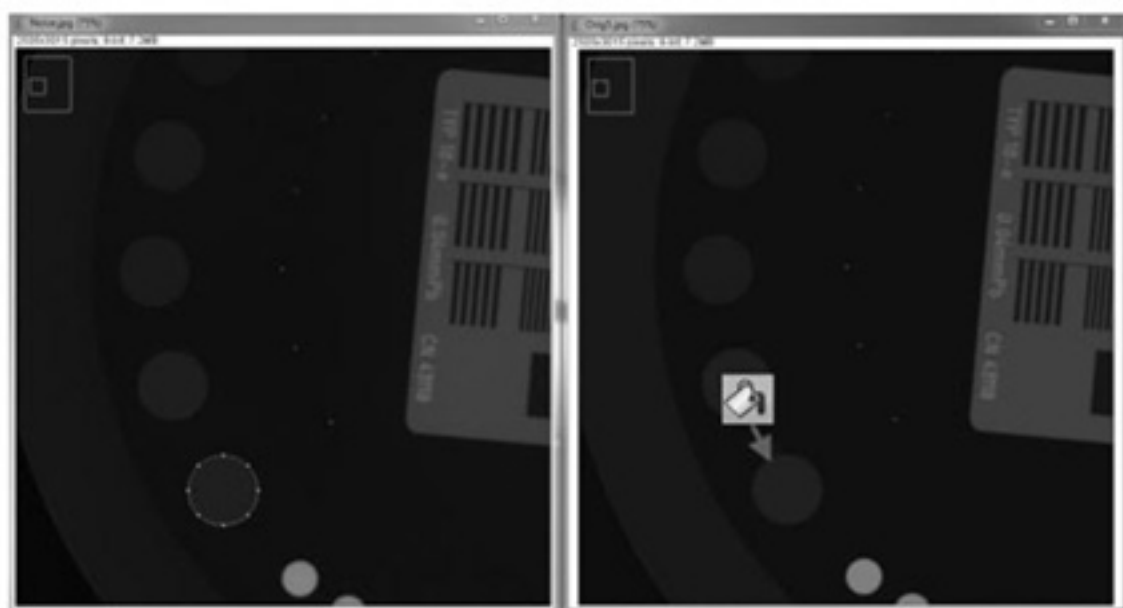
1. สร้างภาพต้นฉบับ (Original image)

ถ่ายภาพเอกซเรย์ โดยวาง TOR CDR Phantom (Figure 2)⁵ (high-contrast sensitivity: 17 circular details, diameter 0.5 mm, low-contrast sensitivity: 17 circular details, diameter 11 mm, grey scale: 10 discs, diameter 5.6 mm, and resolution test patterns) ไว้ที่จุดกึ่งกลางของแผ่นรับสัญญาณภาพ (imaging Plate: Fuji, 200 mm x 251 mm) แล้วเอกซเรย์ด้วยเครื่องเอกซเรย์ Toshiba KXO-50R (54 kVp, 3.2 mAs, no filtration, SID=100 cm) หลังจากนั้น แปลงสัญญาณ imaging plate เป็นภาพเอกซเรย์ที่เครื่อง FCR PROPECT โดยขั้นตอนการประมวลผลภาพไม่ผ่านซอฟต์แวร์การลด

สัญญาณรบกวนของบริษัท หลังจากนั้นนำภาพเอกซเรย์ที่ได้ไปสร้างเป็นภาพต้นฉบับโดยใช้ซอฟต์แวร์ (ImageJ 1.50i; <https://imagej.nih.gov/ij/download.html>) ตกแต่งและปรับภาพหลังการประมวลผลให้เป็นภาพที่มีสัญญาณรบกวนน้อยที่สุด (พิจารณาจากค่าเบี่ยงเบนมาตรฐานของภาพที่น้อยที่สุด) โดยใช้เครื่องมือการเลือก (selection tool) เพื่อเลือกพื้นที่ที่ต้องการ หลังจากนั้นคำนวณค่าเฉลี่ยเพื่อใช้เป็นตัวแทนระดับสีของพื้นที่ที่เลือก หลังจากนั้นใช้เครื่องมือ (flood fill tool) เทสีให้บริเวณที่เลือกมีค่าระดับสีเท่ากับค่าเฉลี่ยของพื้นที่นั้น ทำแบบนี้จนครบทุกบริเวณของภาพ (Figure 3) และสมมติให้ภาพต้นฉบับเป็นภาพที่ไม่มีสัญญาณรบกวน



Figure 2 TOR CDR Phantom.



Measured average in the area

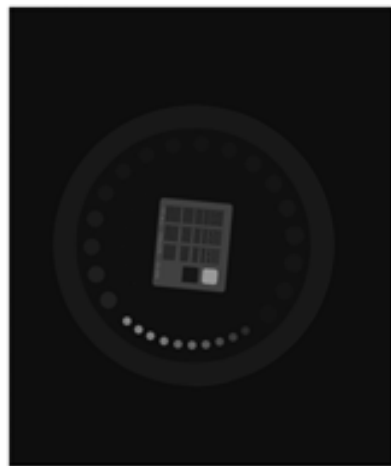
Used Flood Fill Tool

Figure 3 Image retouching.

Figure 3 เป็นตัวอย่างการตกแต่งภาพของ low-contrast sensitivity วงที่ 1 โดยสมมติตามหลักการวัตถุที่มีเนื้อเดียว เมื่อนำไปถ่ายภาพเอกซเรย์ควรมีค่าความเข้มแสงเท่ากัน ในการศึกษาครั้งนี้ผู้วิจัยตกแต่งภาพ high-contrast sensitivity ทั้ง 17 วง low-contrast sensitivity ทั้ง 17 วง grey scale ทั้ง 10 วง และพื้นหลังของภาพ

2. สร้างกรณีตัวอย่าง (Instance) และสกัดคุณลักษณะเฉพาะ (Feature extraction)

สร้างกรณีตัวอย่างด้วยการเติมสัญญาณรบกวนที่ทราบชนิด คือ เกาส์เซียน ปัวซอง และอิมพัลส์ในภาพต้นฉบับโดยเติมสัญญาณรบกวนปริมาณน้อยๆ ไปจนถึงปริมาณมากๆ เติมสัญญาณรบกวนเกาส์เซียน โดยตั้งค่าความแปรปรวน (variance)



Original Image &
Known-noise

ระหว่าง 0.000001-0.000251 ได้กรณีตัวอย่าง จำนวน 100 ภาพ แล้วเติมสัญญาณรบกวนปัวซอง โดยตั้งค่า scale ระหว่าง 10000000-245000000 ได้กรณีตัวอย่าง จำนวน 100 ภาพ และเติมสัญญาณรบกวนอิมพัลส์ โดยตั้งค่า noise density ระหว่าง 0.0001-0.0060 ได้กรณีตัวอย่าง จำนวน 100 ภาพ (ใช้คำสั่งสำหรับสร้างภาพกรณีตัวอย่างโดย MATLAB R2012a⁶)

เมื่อได้ภาพกรณีตัวอย่างทั้งหมด 300 ภาพ สกัดคุณลักษณะเฉพาะ (Figure 4) ของภาพเอกซเรย์ที่ได้ โดยคำนวณค่าจากภาพ ซึ่งคำนวณจากค่าความเข้มแสงในภาพ ได้แก่ ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (mean square error; MSE) และค่าพิกัดสัญญาณต่อสัญญาณรบกวน (peak signal to noise ratio; PSNR)

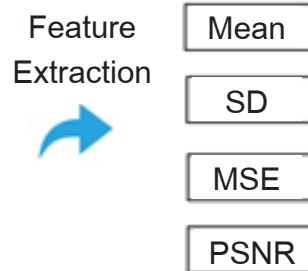


Figure 4 Feature extraction.

3. การเลือกคุณลักษณะเฉพาะ (Feature selection)

เป็นขั้นตอนคัดเลือกคุณลักษณะเฉพาะที่มีประสิทธิภาพสูงสุด โดยทดสอบคุณลักษณะเฉพาะทั้ง 4 เพื่อให้ทราบว่า คุณลักษณะเฉพาะใดมีอิทธิพลต่อการจำแนกสัญญาณรบกวนที่แท้จริง การศึกษาได้แบ่งคุณลักษณะเฉพาะที่เป็นไปได้สำหรับการทดสอบดังนี้ คือ {mean}, {SD}, {MSE}, {PSNR}, {mean, SD}, {mean, MSE}, {mean, PSNR}, {SD, MSE}, {SD, PSNR}, {MSE, PSNR}, {mean, SD, MSE}, {mean, SD, PSNR}, {mean, MSE, PSNR}, {SD, MSE, PSNR} และ {mean, SD, MSE, PSNR} จากนั้นทดสอบคร่าวๆ ด้วยขั้นตอนวิธีของนาอีฟเบส์ โดย validation technique คือ k -folds cross validation คือ การแบ่งข้อมูลออกเป็น k ส่วนเท่ากัน ข้อมูล 1 ส่วนใช้เป็นตัวอย่างทดสอบ ประสิทธิภาพของแบบจำลอง ข้อมูลในส่วนที่เหลือใช้ในการ

สร้างแบบจำลอง ทำวนไปเช่นนี้จนครบจำนวน k ครั้ง ซึ่งกำหนดให้ $k=8$ ทำการทดสอบเพื่อเปรียบเทียบว่า ชุดของคุณลักษณะเฉพาะชุดใดทำนายผลลัพธ์ได้ถูกต้องมากที่สุด และจะใช้ชุดของคุณลักษณะเฉพาะที่ได้นี้สำหรับสร้างแบบจำลองต่อไป

4. การสร้างแบบจำลองและการประเมินผล

เป็นขั้นตอนสร้างแบบจำลองสำหรับจำแนกสัญญาณรบกวน โดยสร้างกรณีตัวอย่างขึ้นมาใหม่จากภาพต้นฉบับจำนวน 900 ภาพ ด้วยเงื่อนไขตามขั้นตอนที่ 2 แต่เพิ่มจำนวนของภาพที่เติมสัญญาณรบกวนเกาส์เซียน ปัวซอง และอิมพัลส์อย่างละ 300 ภาพ โดยสกัดคุณลักษณะเฉพาะตามชุดที่ได้จากขั้นตอนที่ 3 จากนั้นสร้างแบบจำลองเพื่อจำแนกสัญญาณรบกวนในภาพเอกซเรย์ซีอาร์ ด้วยขั้นตอนวิธีนาอีฟเบส์

(WEKA Version 3.7.4) แบบจำลองของ dataset ที่ประกอบด้วย training data และ testing data ใช้เทคนิค *k*-folds cross validation โดยแปรผันค่า *k* (*k* = 2, 3, 4, ..., 18) เพื่อเปรียบเทียบว่าแบบจำลองใดสามารถจำแนกสัญญาณรบกวนได้ถูกต้องมากที่สุด

การประเมินผล (evaluation) สำหรับประมาณค่าประสิทธิภาพของแบบจำลองกระทำโดยประเมินค่าจากการใช้แบบจำลองด้วยค่าความเที่ยง (precision) คือ ความสามารถในการจัดข้อมูลที่ไม่เกี่ยวข้องออกไป precision เป็น

อัตราส่วนของจำนวนข้อมูลที่เกี่ยวข้องและถูกดึงออกกับจำนวนข้อมูลที่ถูกดึงออกมาทั้งหมด ค่าค้นคืน (recall) คือ การวัดความสามารถของระบบในการดึงข้อมูลที่เกี่ยวข้องทั้งหมดออกมา recall เป็นอัตราส่วนของจำนวนข้อมูลที่เกี่ยวข้องและถูกดึงออกมากับจำนวนข้อมูลที่เกี่ยวข้องทั้งหมด และค่าเอฟเมเชอร์ (F-measure) คือ ค่าที่บ่งบอกว่าข้อมูลโน้มเอียงไปทางค่า precision หรือ recall ซึ่งสามารถคำนวณได้จาก equation 3-5 และตารางที่ 1

Table 1 Confusion matrix for evaluating the categorical output.

	Data retrieved (+)	Data not retrieved (-)
Relevant data (+)	TP	FP
Irrelevant data (-)	FN	TN

Denote: TP = True Positive, FP = False Positive, FN = False Negative, TN = True Negative

$$\text{Precision} = \frac{TP}{(TP + FN)} \dots\dots\dots (\text{Eq. 3})$$

$$\text{Recall} = \frac{TP}{(TP + FP)} \dots\dots\dots (\text{Eq. 4})$$

$$\text{F - Measure} = \frac{2\text{Precision}}{(\text{Precision} + \text{Recall})}, \text{ where } \beta = 1 \dots\dots\dots (\text{Eq. 5})$$

5. การใช้งานแบบจำลอง (Model usage)

เป็นขั้นตอนใช้งานแบบจำลองสำหรับข้อมูลที่ไม่ทราบค่า การศึกษาครั้งนี้ต้องการทราบชนิดของสัญญาณรบกวนที่อยู่ในภาพเอกซเรย์ซี่อาร์ ภาพ unseen data ได้จากการถ่ายภาพเอกซเรย์ TOR CDR phantom (ใช้เทคนิค 54 kVp, 3.2 mAs, no filtration, SID = 100 cm) โดยขั้นตอนการประมวลผลภาพ

ด้วย CR reader ไม่ให้ผ่านซอฟต์แวร์การลดสัญญาณรบกวนของบริษัท จึงได้เป็นภาพเอกซเรย์ซี่อาร์ที่มีสัญญาณรบกวนผสมอยู่ หลังจากนั้นนำภาพไปจำแนกประเภทด้วยแบบจำลองที่ได้จากการศึกษาในครั้งนี้ เพื่อระบุชนิดสัญญาณรบกวน (Figure 5)

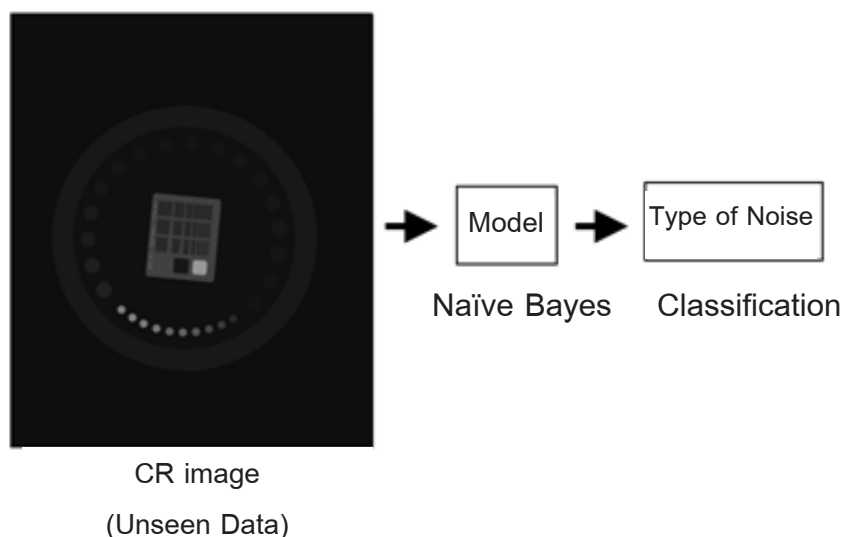


Figure 5 Model usage.

ผลการศึกษา

ภาพเอกซเรย์ซีอาร์ของ TOR CDR phantom เป็นภาพที่ไม่ถูกลดสัญญาณรบกวนจากซอฟต์แวร์ของบริษัทในขั้นตอนการประมวลผล ยังมีสัญญาณรบกวนผสมอยู่ และพบมีค่า SD=13.675 และภาพต้นฉบับที่ได้ทำตกแต่งภาพให้มีสัญญาณรบกวนลดลง มีค่า SD = 12.892 (Figure 6)

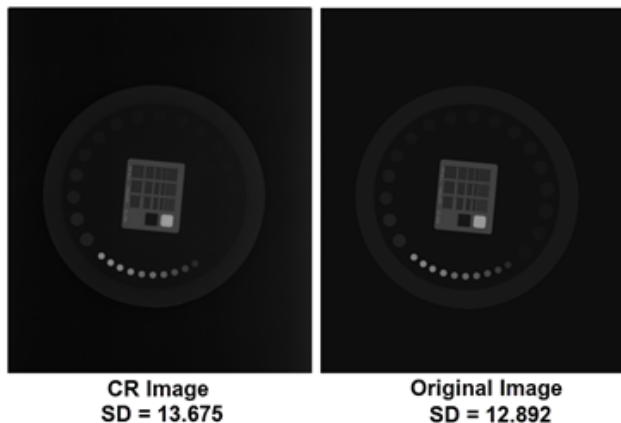


Figure 6 CR and original images.

รายละเอียดของคุณลักษณะเฉพาะที่ได้จากการสกัดจากภาพเอกซเรย์ที่ได้เติมสัญญาณรบกวนที่ทราบชนิดจำนวน 300 ภาพในขั้นตอนการคัดเลือกคุณลักษณะเฉพาะแสดงในตารางที่ 2 และผลการคัดเลือกคุณลักษณะเฉพาะที่มีประสิทธิภาพสูงสุดที่มีอิทธิพลกับการจำแนกสัญญาณรบกวนที่แท้จริงแสดงในตารางที่ 3

Table 2 Detail of features in feature extraction step.

Feature	Maximum	Minimum	Mean	SD
Mean	18.463	17.766	17.870	0.103
SD	18.12	12.89	13.433	0.856
MSE	8.128	0.011	1.647	2.137
PSNR	67.747	39.031	50.704	7.339
Class = Gaussian (100), Poisson (100), Impulse (100)				

Table 3 Performance estimation of feature selection.

Subset	Performance estimation (%)
{Mean}	75.00
{SD}	73.67
{MSE}	67.33
{PSNR}	63.67
{Mean, SD}	82.33
{Mean, MSE}	86.00
{Mean, PSNR}	88.67
{SD, MSE}	81.67
{SD, PSNR}	83.00
{MSE, PSNR}	68.00
{Mean, SD, MSE}	84.00
{Mean, SD, PSNR}	86.33
{Mean, MSE, PSNR}	83.67
{SD, MSE, PSNR}	77.67
{Mean, SD, MSE, PSNR}	86.33

จากตารางที่ 3 จะเห็นได้ว่าชุดของคุณลักษณะเฉพาะที่มีประสิทธิภาพสูงสุด คือ {mean, PSNR} ซึ่งสามารถจำแนกสัญญาณรบกวนได้ถูกต้องร้อยละ 88.67 และถูกใช้เพื่อสร้างแบบจำลองต่อไป

รายละเอียดของคุณลักษณะเฉพาะที่ได้จากการสกัดภาพเอกซเรย์ซึ่งถูกเติมสัญญาณรบกวนที่ทราบชนิดจำนวน 900 ภาพ เพื่อใช้สำหรับการสร้างแบบจำลองแสดงในตารางที่ 4 ผลการสร้างแบบจำลองสำหรับจำแนกสัญญาณรบกวนในภาพเอกซเรย์ซีอาร์ด้วยขั้นตอนวิธีนาอิวเบสได้ผลการประเมินดังแสดงในตารางที่ 5

Table 4 Detail of features in modeling step.

Feature	Maximum	Minimum	Mean	SD
Mean	18.482	17.662	17.869	0.106
PSNR	67.747	38.991	50.624	7.280

Class = Gaussian (300 instances), Poisson (300 instances), Impulse (300 instances)

Table 5 Modeling evaluation.

<i>k</i> -fold cross validation	Correctly classified instance (%)
2	89.11
3	88.89
4	88.33
5	89.33
6	88.89
7	89.56
8	89.22
9	89.89
10	89.56
11	89.89
12	89.78
13	89.67
14	90.00
15	89.44
16	89.89
17	89.11
18	88.89

ข้อมูลจากตารางที่ 5 จะเห็นได้ว่า แบบจำลองสำหรับจำแนกสัญญาณรบกวนในภาพเอกซเรย์ซี่อาร์ไคเทคนิค 14-folds cross validation สามารถจำแนกสัญญาณรบกวนได้ถูกต้องมากที่สุด คือ ร้อยละ 90.00

จากการศึกษาพบว่า การประเมินผลของแบบจำลองสามารถจำแนกได้ถูกต้องว่าเป็นสัญญาณรบกวนเกาส์เซียน 292 จาก 300 กรณีตัวอย่าง มีความผิดพลาดโดยจำแนก

เป็นปัวซอง 8 กรณีตัวอย่าง ภาพที่มีสัญญาณรบกวนแบบปัวซองสามารถจำแนกได้ถูกต้อง 219 จาก 300 กรณีตัวอย่าง มีความผิดพลาดโดยจำแนกเป็นเกาส์เซียน 78 กรณีตัวอย่างและอิมพัลส์ 3 กรณีตัวอย่าง และภาพที่มีสัญญาณรบกวนแบบอิมพัลส์สามารถจำแนกได้ถูกต้อง 296 จาก 300 กรณีตัวอย่าง มีความผิดพลาดโดยจำแนกเป็นปัวซอง 4 กรณีตัวอย่าง (Table 6)

Table 6 Cross-validated confusion matrix for classifier.

Retrieved or Classified as	(Gaussian)	(Poisson)	(Impulse)	Test set
Gaussian	292	8	0	300
Poisson	78	219	3	300
Impulse	0	4	296	300
	370	231	299	900

และพบว่า การประมาณค่าประสิทธิภาพของแบบจำลองสำหรับสัญญาณรบกวนเกาส์เซียนมีค่าความเที่ยงเท่ากับ 0.79 ค่าค้นคืนเท่ากับ 0.97 และค่าเอฟเมเชอร์เท่ากับ 0.90 สำหรับสัญญาณรบกวนแบบปัวซองมีค่าความเที่ยงเท่ากับ 0.95 ค่าค้นคืนเท่ากับ 0.73 และค่าเอฟเมเชอร์เท่ากับ 1.13 และสัญญาณรบกวนแบบอิมพัลส์มีค่าความเที่ยงเท่ากับ 0.99 ค่าค้นคืนเท่ากับ 0.99 และค่าเอฟเมเชอร์เท่ากับ 1.00 (Table 7)

ภาพเอกซเรย์ซีอาร์ หรือ unseen data เป็นภาพเอกซเรย์ซีอาร์ที่ไม่ถูกลดสัญญาณรบกวนจากซอฟต์แวร์ของบริษัทในขั้นตอนการประมวลผลภาพ มีค่า mean = 17.947 และ PSNR = 30.091 นำไปทดสอบกับแบบจำลองที่ได้สามารถจำแนกได้ว่าเป็นสัญญาณรบกวนแบบปัวซอง

Table 7 Detailed accuracy by class.

	Precision	Recall	F-measure
Gaussian	0.79	0.97	0.90
Poisson	0.95	0.73	1.13
Impulse	0.99	0.99	1.00

วิจารณ์ผลการศึกษา

ภาพเอกซเรย์ซีอาร์ที่ได้จากการประมวลโดย CR reader ในการศึกษาครั้งนี้เป็นภาพสเกลสีเทา (gray scale) 10 bit มีระดับความเข้มแสง 1024 ระดับ (0–1023) แต่ใช้ตารางค้นหา (look-up table) ปรับให้เป็นภาพ 8 bit มีระดับความเข้มแสง 256 เพื่อให้สอดคล้องกับความสามารถในการแสดงผลของจอภาพคอมพิวเตอร์ทั่วไป

ภาพต้นฉบับได้รับการตกแต่งให้มีสัญญาณรบกวนลดลงโดยเปรียบเทียบค่า SD กับภาพซีอาร์ที่ไม่ถูกลดสัญญาณรบกวนจากซอฟต์แวร์ของบริษัทในขั้นตอนการประมวลผล โดยภาพต้นฉบับที่ได้มีค่า SD น้อยกว่า และสมมติให้เป็นภาพที่มีสัญญาณรบกวนในภาพน้อยที่สุด¹

สำหรับการสร้างกรณีตัวอย่างในขั้นตอนการคัดเลือกคุณลักษณะเฉพาะ ผู้วิจัยเติมสัญญาณรบกวนที่ทราบชนิดโดยกำหนดค่า variance สำหรับเติมสัญญาณรบกวนเกาส์เซียน ค่า scale สำหรับเติมสัญญาณรบกวนปัวซอง และ ค่า noise density สำหรับเติมสัญญาณรบกวนอิมพัลส์เป็นการสุ่ม

กลุ่มตัวอย่างตามสะดวก เพราะง่ายต่อการสร้างข้อมูล แต่ปริมาณสัญญาณรบกวนแต่ละชนิดที่สร้างขึ้นมามีปริมาณน้อยมากถึงปริมาณมากจนสามารถสังเกตเห็นได้ชัดเจน ซึ่งจะครอบคลุมสัญญาณรบกวนที่จะทำการศึกษา⁶ คุณลักษณะเฉพาะที่สกัดจากการศึกษานี้เป็นค่าที่คำนวณได้จากภาพ เนื่องจากยังไม่ทราบว่าคุณลักษณะเฉพาะทั้งหมดที่สกัดได้จะมีอิทธิพลที่แท้จริงต่อการสร้างแบบจำลองหรือไม่ จึงจำเป็นต้องมีขั้นตอนการคัดเลือกคุณลักษณะเฉพาะ

การคัดเลือกคุณลักษณะเฉพาะที่มีประสิทธิภาพสูงสุดเหมือนกับการเลือกปัจจัยเพื่อเป็นค่านำเข้าที่ส่งผลต่อค่าผลลัพธ์ที่แท้จริง เนื่องจากบางปัจจัยอาจไม่มีความสำคัญหรือไม่ส่งผลต่อค่าผลลัพธ์จึงต้องตัดทิ้งไป คุณลักษณะเฉพาะที่ถูกคัดเลือกในการศึกษานี้ได้รับการทดสอบเพื่อให้ได้คุณลักษณะเฉพาะที่มีอิทธิพลสำหรับจำแนกสัญญาณรบกวนที่แท้จริง

การสร้างแบบจำลองสำหรับจำแนกสัญญาณรบกวนในภาพเอกซเรย์ซีอาร์ด้วยขั้นตอนวิธีน้ำอูฟเวเบสโดยใช้เทคนิค k-folds cross validation ในการศึกษาพบว่า ใช้เทคนิค

14-folds cross validation ทำให้แบบจำลองมีความถูกต้องมากที่สุด⁷ สามารถใช้สำหรับจำแนกสัญญาณรบกวนในภาพเอกซเรย์ซีอาร์ทีซึ่งครอบคลุมสัญญาณรบกวน 3 แบบ คือ เกาส์เซียน ปัวซอง และ อิมพัลส์

ภาพเอกซเรย์ซีอาร์ที หรือ unseen data เป็นภาพเอกซเรย์ซีอาร์ทีที่มีสัญญาณรบกวนปนอยู่ในภาพ เนื่องจากไม่ถูกลดสัญญาณรบกวนจากซอฟต์แวร์ของบริษัทในขั้นตอนการประมวลผลภาพ เมื่อใช้คุณลักษณะเฉพาะของภาพ คือ mean และ PSNR จำแนกสัญญาณรบกวนจากแบบจำลองพบว่าสามารถจำแนกได้ว่าเป็นสัญญาณรบกวนชนิดปัวซองสอดคล้องกับการศึกษาของ Bushberg และคณะ¹ และ Papoulis⁸ ที่กล่าวว่าสัญญาณรบกวนที่เกิดขึ้นในภาพเอกซเรย์ส่วนมากเป็นสัญญาณรบกวนที่มีการแจกแจงปัวซอง ในทำนองเดียวกัน หากสมมติว่าในภาพเอกซเรย์ซีอาร์ทีมีสัญญาณรบกวน เกาส์เซียน หรือ อิมพัลส์ แบบจำลองที่ได้ก็สามารถจำแนกได้เช่นเดียวกัน

การศึกษาในครั้งนี้ศึกษาด้วย Phantom ที่ใช้เป็นตัวแทนของผู้ป่วยและสมมติให้ภาพต้นฉบับเป็นภาพที่ไม่มีสัญญาณรบกวนโดยการปรับภาพให้มีสัญญาณรบกวนน้อยที่สุดจาก post processing ซึ่งไม่สามารถทำกับภาพเอกซเรย์ผู้ป่วยจริงได้ และรูปแบบของการศึกษาดังกล่าวสามารถใช้เป็นต้นแบบเพื่อประยุกต์ใช้งานสำหรับจำแนกสัญญาณรบกวนชนิดต่างๆ ที่อยู่ในภาพสเกลสีเทาโดยสามารถเพิ่มชนิดของสัญญาณรบกวนที่เป็นไปได้ชนิดอื่นๆ เช่น สัญญาณรบกวนชนิด gamma, uniform, rayleigh หรือ exponential หรือใช้ขั้นตอนวิธีอื่นๆ เพื่อจำแนกสัญญาณรบกวน เช่น artificial neural network, Bayesian network หรือ decision tree

สรุปผลการศึกษา

จากการศึกษาได้แบบจำลองสำหรับจำแนกสัญญาณรบกวนในภาพเอกซเรย์ซีอาร์ทีด้วยขั้นตอนวิธีนาอีฟเบส์ มีความถูกต้องร้อยละ 90.00 เมื่อนำแบบจำลองที่ได้ไปจำแนกสัญญาณรบกวนในภาพเอกซเรย์ซีอาร์ทีพบว่าสัญญาณรบกวนที่มีอยู่ในภาพเอกซเรย์ซีอาร์ที คือ สัญญาณรบกวนชนิดปัวซอง ดังนั้นเมื่อทราบวิธีการแยกชนิดสัญญาณรบกวนในภาพเอกซเรย์ซีอาร์ทีแล้ว จะสามารถนำไปประยุกต์ขั้นตอนวิธีที่เหมาะสมสำหรับการลดสัญญาณรบกวนต่อไป

กิตติกรรมประกาศ

คณะผู้วิจัยขอขอบพระคุณโรงพยาบาลมหาราชนครราชสีมา สาขาเทคโนโลยีสารสนเทศ สำนักวิชาเทคโนโลยีสังคม และสาขาคณิตศาสตร์ สำนักวิชาวิทยาศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี ที่สนับสนุน และเป็นสถานที่สำหรับทำวิจัย ขอขอบพระคุณภาควิชารังสีวิทยา คณะแพทยศาสตร์ มหาวิทยาลัยขอนแก่น ที่สนับสนุนเครื่องมือสำหรับทำวิจัย และขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.ธรา อังสกุล และ ผู้ช่วยศาสตราจารย์ ดร.จิตติมนต์ อังสกุล สำหรับคำแนะนำและสอนเรื่องปัญญาประดิษฐ์

เอกสารอ้างอิง

1. Bushberg JT, Seibert JA, Leidholdt EM, Boone JM. The essential physics of medical imaging. 2nd Baltimore: Lippincott Williams & Wilkins; 2002.
2. Gonzalez RC, Woods RE. Digital image processing. 2nd Massachusetts: Addison-Wesley; 1992.
3. Goodman JW. Statistical optics. New York: John Wiley & Sons; 2000.
4. Bayes T, Price R. An essay towards solving a Problem in the Doctrine of Chances. By the late Rev. Mr. Bayes, F.R.S. Communicated by Mr. Price in a Letter to John Canton A.M.F.R.S., Philosophical Transactions of the Royal Society of London 1763; 53: 370-418.
5. Leads Test Objects. TOR CDR Radiography Phantom [cited 2016 November 20]. Available from: <http://www.leadstestobjects.com>.
6. MathWorks. MATLAB R2012a [cited 2016 November 20]. Available from: <http://www.mathworks.com>.
7. The University of Waikato. WEKA (Waikato Environment for Knowledge Analysis) [cited 2017 February 27]. Available from: <http://www.cs.waikato.ac.nz/ml/weka>.
8. Papoulis A. Probability, random variables, and stochastic processes. 3rd New York: McGraw-Hill; 1991.