



การวิเคราะห์ข้อความตัวอักษรหน้ายาจากภาพอย่างอัตโนมัติ

โดย

นางสาวศิโรรัตน์ สุนทรสุข

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
วิทยาศาสตรมหาบัณฑิต (วิทยาการคอมพิวเตอร์)
สาขาวิชาวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์
ปีการศึกษา 2559
ลิขสิทธิ์ของมหาวิทยาลัยธรรมศาสตร์

การวิเคราะห์ข้อความตัวอักษรหน้ายาจากภาพอย่างอัตโนมัติ

โดย

นางสาวศิโรรัตน์ สุนทรสุข



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

วิทยาศาสตร์มหาบัณฑิต (วิทยาการคอมพิวเตอร์)

สาขาวิชาวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์

ปีการศึกษา 2559

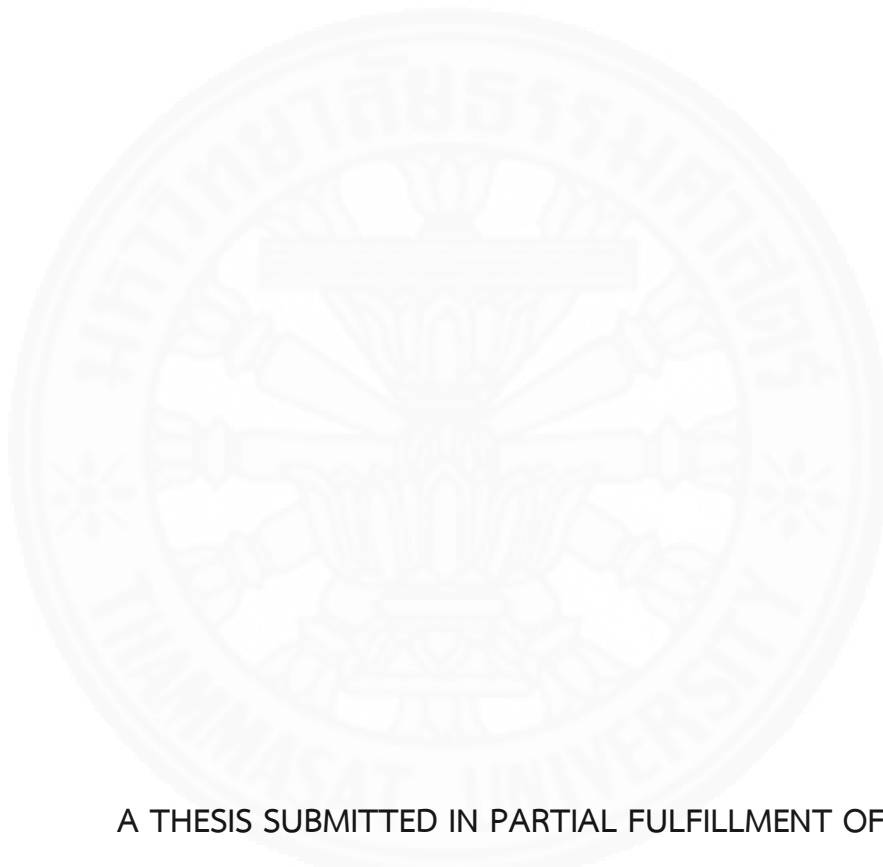
ลิขสิทธิ์ของมหาวิทยาลัยธรรมศาสตร์



Automatic Text Imprint Analysis from Pill Images

BY

Miss Sioratt Suntronsuk



A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE
(COMPUTER SCIENCE)

DEPARTMENT OF COMPUTER SCIENCE
FACULTY OF SCIENCE AND TECHNOLOGY
THAMMASAT UNIVERSITY

ACADEMIC YEAR 2016

COPYRIGHT OF THAMMASAT UNIVERSITY

มหาวิทยาลัยธรรมศาสตร์
คณะวิทยาศาสตร์และเทคโนโลยี

วิทยานิพนธ์

ของ

นางสาวศิริรัตน์ สุนทรสุข

เรื่อง

การวิเคราะห์ข้อความตัวอักษรหน้ายาจากภาพอย่างอัตโนมัติ
ได้รับการตรวจสอบและอนุมัติ ให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
วิทยาศาสตร์มหาบัณฑิต (วิทยาการคอมพิวเตอร์)

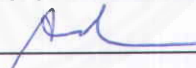
เมื่อ วันที่ 30 พฤศจิกายน พ.ศ. 2559

ประธานกรรมการสอบวิทยานิพนธ์



(ผู้ช่วยศาสตราจารย์ ดร. เสาวลักษณ์ วรรณภา)

กรรมการและอาจารย์ที่ปรึกษาวิทยานิพนธ์



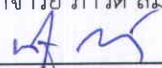
(อาจารย์ ดร. สุกัญญา รัตน์ททยานนท์)

กรรมการสอบวิทยานิพนธ์



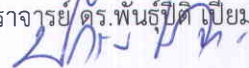
(รองศาสตราจารย์ ภาวดี สมภักดี)

กรรมการสอบวิทยานิพนธ์



(รองศาสตราจารย์ ดร. พันธุ์ปิติ เปี่ยมสง่า)

คณบดี



(รองศาสตราจารย์ ปกรณ์ เสริมสุข)

หัวข้อวิทยานิพนธ์	การวิเคราะห์ข้อความตัวอักษรหน้ายาจากภาพอย่างอัตโนมัติ
ชื่อผู้เขียน	นางสาวศิริโรรัตน์ สุนทรสุข
ชื่อปริญญา	วิทยาศาสตร์มหาบัณฑิต
สาขาวิชา/คณะ/มหาวิทยาลัย	สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์	อาจารย์ ดร. สุกัญญา รัตโนทยานนท์
ปีการศึกษา	2559

บทคัดย่อ

หนึ่งในปัญหาสำคัญของเภสัชกรคือการระบุข้อมูลเม็ดยาที่ไม่มีฉลาก ความยากของการระบุเม็ดยาเนื่องจากเม็ดยามีหลากหลายประเภทและมีลักษณะภายนอกคล้ายคลึงกัน รวมไปถึงความไม่ครบถ้วนของข้อมูลในฐานข้อมูลเม็ดยา สัญลักษณ์บนหน้ายามักแสดงคุณสมบัติที่เป็นประโยชน์ในการจำแนกเม็ดยา เช่น อักษรย่อผู้ผลิต สารประกอบสำคัญ เป็นต้น ดังนั้นการที่สามารถจำแนกตัวอักษรบนหน้ายาสามารถเป็นประโยชน์พัฒนาต่อยอดในการระบุเม็ดยา เช่น ใช้สำหรับเพิ่มข้อมูลในฐานข้อมูลด้วยรูปภาพ ใช้สร้างดัชนีสำหรับการสืบค้นข้อมูลเม็ดยาในฐานข้อมูลอัตโนมัติ เป็นต้น

เทคนิคการจำแนกสัญลักษณ์บนหน้ายาจากงานจำแนกเม็ดยาที่ผ่านมาให้ผลลัพธ์เป็นเวกเตอร์ทำให้ไม่สามารถทำงานร่วมกับฐานข้อมูลยาที่มีในปัจจุบันซึ่งเป็นฐานข้อมูลชนิดข้อความได้ ดังนั้นงานวิจัยนี้จึงนำเสนอวิธีการวิเคราะห์สัญลักษณ์บนหน้ายาที่ให้ผลลัพธ์อยู่ในรูปแบบข้อความตัวอักษร โดยงานวิจัยนี้ได้นำเสนอเทคนิคสำหรับหาตำแหน่งสัญลักษณ์หน้ายาบนภาพที่ได้แนวคิดจากงานวิจัย (T Kasar, J Kumar, & A G Ramakrishnan, 2007) และเสนอเทคนิค Noise elimination ที่ทำหน้าที่กำจัดกลุ่มพิกเซลปัญหาที่ปรากฏในภาพขาวดำที่สร้างจาก Otsu's thresholding ซึ่งกลุ่มพิกเซลปัญหาจะพบในภาพเม็ดยาที่มีเงาขีดบริเวณสัญลักษณ์บนหน้ายาแบบลึกและในภาพเม็ดยาหลากสี จากผลการทดลองพบว่าระบบวิเคราะห์สัญลักษณ์บนหน้ายาแบบใหม่นี้สามารถให้ความแม่นยำและความครบถ้วนเกิน 57 เปอร์เซ็นต์ โดยเฉพาะกับภาพเม็ดยาที่มีสัญลักษณ์แบบพิมพ์ที่ระบบสามารถให้ความแม่นยำและความครบถ้วนเกิน 77 เปอร์เซ็นต์

คำสำคัญ: การระบุข้อมูลเม็ดยา, เทคนิคการจำแนกสัญลักษณ์บนหน้ายา

Thesis Title	Automatic Text Imprint Analysis from Pill Images
Author	Miss Siroratt Suntronsuk
Degree	Master of Science
Major Field/Faculty/University	Computer Science Faculty of Science and Technology Thammasat University
Thesis Advisor	Dr.Sukanya Ratanotayanon
Academic Years	2016

ABSTRACT

Pill identification is a serious concern for pharmacists due to the sheer number of pills. Appearances of many pills are also similar. In addition, existing pill databases do not have complete pill information. As pill imprints often contain important information about the pills. It is beneficial to have a technique that can automatically extract texts on the imprints. The results can be extended to use for adding or searching pill information on existing pill databases using pill images.

Most of existing techniques for extracting pill imprints give results as vectors therefore the results cannot be used with existing pill databases which are text-based. Thus, this paper proposed new approach for automatically extracting text imprints as readable text. This paper proposes a technique for identifying imprint location based on (T Kasar et al., 2007) and proposes a noise elimination technique for removing problematic pixels in binary images obtained by OTSU's thresholding. This technique can deal with noises caused by shadow in engraved imprints and noises in multicolor pills. The experimental results shown that precision and recall of our proposed method can reach more than 57%. Especially, for printed imprint can sheer F-measure more than 0.77 but engraved imprint still must more improve further.

Keywords: Pill identification, extracting text imprints

กิตติกรรมประกาศ

ขอขอบพระคุณอาจารย์สุกัญญา รัตโนทยานนท์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ที่สละเวลาให้คำปรึกษา เสนอแนะแนวทาง สนับสนุน และช่วยเหลือ ใส่ใจและให้กำลังใจเสมอมาสำหรับความสำเร็จของวิทยานิพนธ์ฉบับนี้ต้องขอบพระคุณเป็นอย่างสูง

ขอขอบพระคุณคณะวิทยาศาสตร์และเทคโนโลยีสำหรับทุนบัณฑิตเรียนดีเพื่อศึกษาต่อระดับบัณฑิตศึกษาที่เป็นสิ่งค่าจุนช่วยเหลือภาระทางการเงินของครอบครัวส่งผลให้วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปด้วยดี

ขอขอบพระคุณคณาจารย์ประจำภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี และคณาจารย์ประจำคณะเภสัชศาสตร์ มหาวิทยาลัยธรรมศาสตร์ทุกท่านที่คอยประสิทธิ์ประสาทวิชาที่เป็นความรู้อันมีค่ายิ่งให้แก่ผู้วิจัยช่วยเพิ่มคุณค่าให้กับวิทยานิพนธ์ฉบับนี้มากขึ้น

ขอขอบพระคุณบิดาและมารดาที่คอยให้ความช่วยเหลือร่างกายและแรงใจเสมอมา

ท้ายที่สุดนี้ผู้วิจัยขอขอบพระคุณทุกท่านที่เกี่ยวข้องทั้งผู้ได้เอ่ยนามและไม่ได้เอ่ยนามก็ดีและขออวยพรให้ทุกท่านประสบความสำเร็จทั้งการงาน ครอบครัว และขอให้บุญบารมีเหล่านี้ส่งผลให้ไร้โรคภัยไข้เจ็บ มีแต่ความสุขความเจริญสถิตทุกท่านเทอญ

นางสาวศิโรรัตน์ สุนทรสุข

(4)

สารบัญ

หน้า

บทคัดย่อภาษาไทย	(1)
บทคัดย่อภาษาอังกฤษ	(2)
กิตติกรรมประกาศ	(3)
สารบัญตาราง	(7)
สารบัญภาพ	(8)
บทที่ 1 บทนำ	1
1.1. ที่มาและความสำคัญของปัญหา	1
1.2. ขั้นตอนการทำวิจัย	2
1.3. ขอบเขตของงานวิจัย	3
1.4. ประโยชน์ที่ได้รับ	4
บทที่ 2 วรรณกรรมและงานวิจัยที่เกี่ยวข้อง	5
2.1. ฐานข้อมูลยา	5
2.2. ทฤษฎีที่เกี่ยวข้อง	10

	(5)
เทคนิคการสร้างภาพขอบวัตถุ	10
เทคนิคการประมวลผลภาพโดยการเปลี่ยนแปลงลักษณะรูปร่างหรือโครงสร้างของภาพ	12
เทคนิคการแยกองค์ประกอบสำคัญของภาพ	13
หลักการพิจารณาพิกเซลวัตถุเดียวกันบนภาพขาวดำ	15
เทคนิคพื้นฐานสำหรับการลดพิกเซลรบกวน	17
เทคนิคพื้นฐานสำหรับการปรับขนาดภาพ	17
การรู้จำอักขระด้วยแสง	18
2.3. งานวิจัยที่เกี่ยวข้อง	19
งานวิจัยจำแนกเม็ดยาด้วยภาพ	19
งานวิจัยที่เกี่ยวข้องกับการวิเคราะห์ตัวอักษร	24
บทที่ 3 วิธีการวิจัย	28
3.1. การเตรียมภาพเม็ดยา	29
3.1.1 การกำจัดความหลากหลายของขนาดรูปเม็ดยา	29
3.1.2 การเพิ่มรายละเอียดสัญลักษณ์บนหน้ายา	30
3.2. การจำแนกพื้นที่บริเวณสัญลักษณ์บนหน้ายา	31
3.3. การจำแนกสัญลักษณ์หน้ายา	34
3.3.1 เทคนิคที่ประยุกต์ Otsu thresholding ร่วมกับ Noise elimination	35
3.3.2 เทคนิคที่ใช้ K-means clustering	37
3.4. การรู้จำตัวอักษรบนหน้ายา	39
3.5. กระบวนการสอนโปรแกรม Tesseract	39
3.5.1 การเตรียมภาพตัวอักษร	39
3.5.2 การติดป้ายข้อมูลให้ภาพตัวอักษร	40

	(6)
3.5.3 การสอนโปรแกรม Tesseract	41
3.5.4 การเพิ่มไฟล์ตั้งค่า	42
3.6. การทดลองเบื้องต้น	43
บทที่ 4 ผลการวิจัยและการอภิปรายผล	46
4.1. เกณฑ์การประเมินผล	47
4.2. การเลือกค่าพารามิเตอร์	48
4.3. ขั้นตอนการทดลอง	49
4.4. อภิปรายผลการทดลอง	50
บทที่ 5 บทสรุปและข้อเสนอแนะ	54
รายการอ้างอิง	57
ประวัติผู้เขียน	62

สารบัญตาราง

ตารางที่	หน้า
2.1	8
2.2	9
2.3	23
3.1	44
3.2	45
3.3	45
4.1	49
4.2	50
4.3	51
4.4	52

สารบัญภาพ

ภาพที่	หน้า
2.1 เมทริกซ์ค่าน้ำหนักของพิกเซลเพื่อนบ้านข้างเคียงของ Prewitt	11
2.2 เมทริกซ์ค่าน้ำหนักของพิกเซลเพื่อนบ้านข้างเคียงของ Sobel	11
2.3 ผลลัพธ์การทำงานของ Dilation	12
2.4 ผลลัพธ์การทำงานของ Erosion	13
2.5 พิกเซลเพื่อนบ้านข้างเคียง (สี่ฟ้า) สี่ทิศทางของ p	15
2.6 พิกเซลเพื่อนบ้านข้างเคียง (สี่ฟ้า) แปดทิศทางของ p	16
2.7 พิกเซลสีแดงแทนพิกเซลที่เป็นเพื่อนบ้านข้างเคียงสี่ทิศทางทั้งพิกเซล p และ q	16
2.8 ผลลัพธ์การจัดกลุ่มพิกเซลชั้น 1	20
2.9 ผลลัพธ์การคัดกรองกลุ่มพิกเซลชั้น 2	20
2.10 ตัวอย่าง EB ของตัวอักษรภาษาอังกฤษและตัวเลข	25
3.1 แผนภาพโดยรวมของกระบวนการวิจัย	28
3.2 ภาพเม็ดยาต้นฉบับ	29
3.3 ภาพเม็ดยาจากภาพที่ 3.2 ที่ผ่านขั้นตอนการเพิ่มรายละเอียดสัญลักษณ์บนหน้ายา	31
3.4 ภาพผลลัพธ์ (a) สมการ E (b) เทคนิค Dilation (c) เทคนิค Closing	32
3.5 ตัวอย่างพื้นที่จริงของสัญลักษณ์บนหน้ายาที่จำแนกได้ (กรอบสีขาว)	33
3.6 ภาพรวมของ (a) เทคนิคของ Otsu (b) เทคนิคของ Otsu ร่วมกับ Noise elimination	35
3.7 ภาพระดับเทา (a) จากขั้นตอนการจำแนกพื้นที่บริเวณสัญลักษณ์บนหน้ายา (b) เฉพาะพิกเซลรอบตัวอักษรบนหน้ายา (c) เฉพาะตัวอักษรหน้ายา (d) จากการกลับค่าสี ของภาพเฉพาะตัวอักษรหน้ายา	36
3.8 ภาพผลลัพธ์ของ (a) ขั้นตอนการจำแนกพื้นที่บริเวณสัญลักษณ์บนหน้ายา (b) เทคนิคของ Otsu (c) เทคนิค Noise elimination (d) เทคนิคร่วมระหว่าง Otsu และ Noise elimination	37

3.9 ตัวอย่างการสร้างภาพขาวดำด้วยเทคนิคที่ใช้ K-means clustering	38
3.10 ตัวอย่างภาพตัวอักษรผลลัพธ์จาก (ซ้าย) ขั้นตอนการจำแนกสัญลักษณ์บนหน้ายา (ขวา) หลังการเตรียมภาพตัวอักษร	40
3.11 ตัวอย่างค่า F-measure ของ 10 อันดับที่ดีที่สุดของแต่ละวิธีการ	44
4.1 ตัวอย่างไฟล์ข้อมูลเมื่อยาที่ได้จากฐานข้อมูล Drugs.com ที่มีข้อมูลรายการตัวอักษรหน้ายา (กรอบสีฟ้า)	47



บทที่ 1 บทนำ

1.1. ที่มาและความสำคัญของปัญหา

ปัญหาสำคัญสำหรับเภสัชกรคือไม่สามารถระบุข้อมูลเม็ดยาที่ปราศจากหีบห่อและคำบ่งชี้ได้เนื่องจากประเภทยาามีหลากหลายชนิดและมีลักษณะภายนอกคล้ายคลึงกัน การได้รับยาหลากหลายชนิดซึ่งอยู่ในชองยาเดียวกันที่ปราศจากฉลากสำหรับผู้ป่วยคลินิก รวมไปถึงความไม่ครบถ้วนของข้อมูลในฐานข้อมูลเม็ดยาเนื่องจากมีเม็ดยาที่จดทะเบียนใหม่ทุกวัน โดยเฉพาะเม็ดยาจากผู้ผลิตไทยเนื่องจากหน่วยงานที่รับผิดชอบฐานข้อมูลส่วนใหญ่เป็นชาวต่างชาติ ประกอบกับต้องใช้เวลาและแรงงานมากสำหรับการสร้างฐานข้อมูลเม็ดยาให้มีข้อมูลครบถ้วนอยู่เสมอ ปัญหานี้เห็นได้จากการที่เภสัชกรนิยมตั้งกลุ่มไลน์ในหมู่เภสัชกรร่วมรุ่นเพื่อสอบถามแทนการสืบค้นจากฐานข้อมูลที่ให้ข้อมูลที่น่าเชื่อถือกว่ามาก (จากคำปรึกษาของ อาจารย์ ดร. เภสัชกรหญิงกรกช กังวาลทัศน์) ระบบจำแนกค่าคุณลักษณะภายนอกของเม็ดยาด้วยภาพอัตโนมัติจึงถูกพัฒนาเพื่อสนับสนุนการสร้างฐานข้อมูลยาให้มีข้อมูลเม็ดยาที่ครบถ้วนมากขึ้น ประกอบกับสัญลักษณ์บนหน้ายามักเป็นคำสำคัญที่มีความหมาย (Chen, Yu, Kamata, & Yang, 2015; Lee, Park, & Jain, 2010) เช่น อักษรย่อผู้ผลิต ปริมาณสารสำคัญ รหัสครั้งการผลิต เป็นต้น ซึ่งเห็นได้จากงานวิจัยจำแนกเม็ดยาด้วยภาพที่ผ่านมา เช่น งานวิจัย (Yu, Chen, & Kamata, 2014) และงานวิจัย (Caban, Rosebrock, & Yoo, 2012) ต่างนิยมใช้สัญลักษณ์บนหน้ายาเป็นคุณลักษณะหนึ่งในการจำแนกเม็ดยารวมถึงการที่ทุกฐานข้อมูลยาในปัจจุบันใช้ตัวอักษรบนหน้ายาเป็นทางเลือกหนึ่งในการทำดัชนีสืบค้นอีกด้วย

จากการทบทวนเทคนิคการวิเคราะห์สัญลักษณ์บนหน้ายาที่ผ่านมาพบว่ามักให้ผลลัพธ์เป็นเวกเตอร์ทำให้ไม่สามารถทำงานร่วมกับฐานข้อมูลยาที่มีอยู่ได้เนื่องจากฐานข้อมูลยาในปัจจุบันเป็นฐานข้อมูลชนิดข้อความทั้งหมด

ดังนั้นงานวิจัยนี้จึงนำเสนอวิธีการวิเคราะห์สัญลักษณ์บนหน้ายาแบบใหม่ที่ให้ผลลัพธ์อยู่ในรูปแบบข้อความตัวอักษรได้อย่างอัตโนมัติเพื่อนำไปประยุกต์ใช้กับฐานข้อมูลยาในปัจจุบันได้ โดยขั้นตอนการทำงานเริ่มจากการเตรียมภาพให้เหลือเฉพาะพิทเชลพื้นที่บริเวณสัญลักษณ์บนหน้ายาซึ่งงานวิจัยนี้ได้เสนอเงื่อนไขการหาพื้นที่สัญลักษณ์บนหน้ายาที่ได้แนวคิดมาจากกฎการหาพื้นที่แต่ละตัวอักษรบนภาพถ่ายเอกสารที่นำเสนอโดย Kasar (T Kasar et al., 2007) ถัดมาเป็นการสร้างภาพขาวดำซึ่งงานวิจัยนี้ได้นำเสนอเทคนิค Noise elimination เพื่อพัฒนาเทคนิคการสร้างภาพขาวดำด้วย

Otsu's thresholding ให้เหมาะสมกับงานเม็ดยามากขึ้น โดยเทคนิค Noise elimination จะกำจัดกลุ่มพิกเซลปัญหาที่มักปรากฏในภาพขาวดำ ซึ่งกลุ่มพิกเซลปัญหาสามารถพบได้ใน 2 กรณีคือ กรณีที่เกิดจากเงาบริเวณสัญลักษณ์บนหน้ายาสำหรับตัวอักษรแบบลึก และกรณีที่เกิดจากจำนวนสีของเม็ดยามีมากกว่าหนึ่งสี และงานวิจัยนี้จะเปรียบเทียบประสิทธิภาพของเทคนิคการสร้างภาพขาวดำที่นำเสนอนี้กับเทคนิคที่ใช้ K-means clustering ซึ่งก่อนหน้านี้เทคนิค K-means clustering เคยถูกนำมาใช้ในการจำแนกตัวอักษรที่สลักลงบนโลหะในงานวิจัย (Mancas-Thillou & Mancas, 2007) ซึ่งมีลักษณะงานคล้ายคลึงกับการจำแนกตัวอักษรแบบลึกบนเม็ดยาที่มักมีลักษณะมันวาว หลังจากนั้นแปลงภาพตัวอักษรหน้ายาเป็นข้อความตัวอักษรด้วยระบบการรู้จำอักขระด้วยแสง

จากการทดลองพัฒนาระบบวิเคราะห์สัญลักษณ์บนหน้ายาที่นำเสนอบนภาษา Python โดยใช้เฟรมเวิร์ค Opencv ซึ่งทดลองกับภาพเม็ดยา 540 ภาพจากฐานข้อมูล Drugs.com ("Drugs.com Pill Identifier," n.d.) ที่มีความละเอียดภาพเม็ดยาระหว่าง 96 ถึง 150 DPI โดยประกอบไปด้วยเม็ดยาที่มีตัวอักษรพิมพ์จำนวน 180 ภาพและเม็ดยาที่มีตัวอักษรแบบสลักอีก 360 ภาพ ผลการทดลองพบว่าเทคนิคที่ประยุกต์ Otsu thresholding ร่วมกับ Noise elimination ที่ทำขั้นตอนเลือกค่าพารามิเตอร์และใช้กระบวนการสอนโปรแกรม OCR ร่วมด้วยสามารถให้ความแม่นยำและความครบถ้วนของระบบเกิน 57 เปอร์เซ็นต์ โดยเฉพาะกับสัญลักษณ์บนหน้ายาที่เป็นตัวพิมพ์ เทคนิคดังกล่าวสามารถให้ความแม่นยำและความครบถ้วนของระบบเกิน 77 เปอร์เซ็นต์ และให้ผลความแม่นยำและความครบถ้วนของระบบดีกว่าเทคนิคที่ใช้ K-means clustering ถึง 10 เท่าในขณะใช้เวลาประมวลผลเฉลี่ยต่อภาพน้อยกว่า ซึ่งสาเหตุหลักเกิดจากเทคนิคที่ใช้ K-means clustering ดั้งเดิมถูกพัฒนาเพื่อทำงานจำแนกพิกเซลตัวอักษรที่สลักบนวัตถุโลหะ ดังนั้นเมื่อนำมาจำแนกตัวอักษรบนหน้ายาที่มีทั้งตัวอักษรแบบพิมพ์และแบบลึกจึงทำให้ค่าความครบถ้วนและความแม่นยำลดลงอย่างเห็นได้ชัด แต่อย่างไรก็ดีเทคนิคสำหรับสัญลักษณ์หน้ายาแบบลึกที่นำเสนอได้ผลความแม่นยำและความครบถ้วนไม่ดิ่งจำเป็นต้องได้รับการพัฒนาต่อไปในอนาคต

1.2. ขั้นตอนการทำวิจัย

เพื่อให้งานวิจัยบรรลุได้ตามวัตถุประสงค์ ผู้วิจัยจะดำเนินการทำวิจัยตามลำดับขั้นตอน
ดังนี้

1. **ทบทวนวรรณกรรม** ศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์สัญลักษณ์บนหน้ายาที่ให้ผลลัพธ์อยู่ในรูปแบบข้อความตัวอักษรบนหน้ายาได้อย่างอัตโนมัติ เช่น ฐานข้อมูลเม็ดยา การรู้จำอักขระด้วยแสง รวมถึงงานวิจัยการจำแนกเม็ดยาด้วยภาพ และงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์ตัวอักษรเพื่อสรุป หาจุดเด่นจุดด้อย และหาช่องทางการพัฒนาต่อยอด
2. **ศึกษาความเป็นไปได้** เรียนรู้การใช้เฟรมเวิร์คต่างๆ สำหรับการประมวลผลภาพและทดลองสร้างโปรแกรมเบื้องต้น เพื่อทดลองความเป็นไปได้ในการพัฒนาระบบวิเคราะห์ตัวอักษรบนเม็ดยาในงานวิจัยให้สำเร็จ
3. **พัฒนาระบบ** พัฒนาระบบให้สอดคล้องตาม 5 ขั้นตอนหลักคือ (1) การเตรียมภาพเม็ดยา (2) การจำแนกพื้นที่บริเวณสัญลักษณ์บนหน้ายา (3) การจำแนกสัญลักษณ์บนหน้ายา (4) การรู้จำตัวอักษรบนหน้ายา (5) กระบวนการสอนโปรแกรม Tesseract
4. **ประเมินผล** ประเมินสมรรถนะของระบบผ่านค่าความแม่นยำ (Precision) ค่าความครบถ้วน (Recall) ค่า F-measure ซึ่งสะท้อนความถูกต้องของระบบ และเวลาที่ใช้ในการประมวลผลเฉลี่ย รวมถึงการเปรียบเทียบกับงานวิจัยอื่นร่วมด้วย
5. **สรุปผล** สรุปผลที่ได้จากงานวิจัยและจัดทำวิทยานิพนธ์

1.3. ขอบเขตของงานวิจัย

การดำเนินงานวิจัยอยู่ภายใต้ขอบเขตที่กำหนดไว้เพื่อเป็นกรอบในการศึกษาและพัฒนาให้เป็นที่น่าพอใจที่สุด อีกทั้งเป็นการกำหนดข้อจำกัดของศึกษาและพัฒนา ทั้งนี้เพื่อให้เหมาะสมแก่เวลาและทรัพยากรในการทำวิจัย ขอบเขตของงานวิจัยฉบับนี้มีดังนี้

- a) รองรับเฉพาะภาพยาที่เป็นเม็ดยาเท่านั้น
- b) รองรับภาพเม็ดยาที่มีความละเอียดตั้งแต่ 70 DPI (dot per inch) และต้องมีพื้นที่เม็ดยาไม่น้อยกว่า 30 เปอร์เซ็นต์ของภาพ รวมถึงความสูงและความกว้างของภาพต้องยาวไม่น้อยกว่า 200*200 พิกเซล แต่ไม่เกิน 3000*3000 พิกเซลตามลำดับ
- c) รองรับตัวอักษรบนหน้ายาทั้งแบบตัวพิมพ์และแบบตัวลึก
- d) รองรับเฉพาะตัวอักษรบนหน้ายาที่มีแนวบรรทัดทางตรงแนวนอนและแนวตั้งเท่านั้น
- e) รองรับตัวอักษรบนหน้ายาที่ประกอบไปด้วยหลายบรรทัด
- f) รองรับเฉพาะเม็ดยาที่บแสงเท่านั้น
- g) รองรับเฉพาะเม็ดยาที่มีสีของสัญลักษณ์บนหน้ายาเพียงสีเดียวเท่านั้น
- h) รองรับทั้งสีสัญลักษณ์บนหน้ายาอ่อนกว่าสีเม็ดยาและในทางตรงกันข้าม

- i) รongรับภาพเม็ดยาที่ปรากฏเงาและการสะท้อนแสงแบบไม่คมชัดบนหน้ายา
- j) รongรับเฉพาะสัญลักษณ์บนหน้ายาที่เป็นตัวอักษรภาษาอังกฤษและตัวเลขอารบิกเท่านั้น
- k) ไม่รongรับตัวอักษรภาษาอังกฤษที่มีรูปแบบตัวอักษรเป็นลายมือเขียน
- l) รongรับยาที่มีมากกว่า 1 สีสำหรับเม็ดยาที่มีตัวอักษรบนหน้ายาแบบตัวพิมพ์ ดังนั้นสำหรับเม็ดยาดั้วลักรongรับเฉพาะเม็ดยาสีเดียวเท่านั้น
- m) รongรับเม็ดยาที่มีลักษณะภายนอกดังนี้
 - i. รูปรongรับเม็ดยาที่รongรับคือ ทรงกลม (round) ทรงรี (oval) และทรงแคปซูล (capsule) เนื่องจากเป็นรูปรongรับเม็ดยาที่พบมากที่สุด 3 อันดับ (Caban et al., 2012)
 - ii. สีเม็ดยาที่ทดลองคือ สีขาว สีเหลือง สีฟ้า และสีส้ม เป็นหลักเนื่องจากเป็นสีเม็ดยาที่พบมากที่สุด 4 อันดับ (Caban et al., 2012) โดยวิธีการที่นำเสนอนี้ควรจะทำงานกับสีอื่นได้ด้วย
- n) สีของพื้นหลังเม็ดยาต้องเป็นสีพื้นเท่านั้น กล่าวคือไม่รongรับภาพถ่ายเม็ดยาที่วางบนอุปกรณ์ที่มีลวดลายใดๆทั้งสิ้น
- o) สีของพื้นหลังและสีของเม็ดยาต้องมีโทนสีตรงข้ามกันเสมอ เช่น หากเม็ดยาสีเข้มสีพื้นหลังต้องสีอ่อน หากเม็ดยาสีอ่อนสีพื้นหลังต้องมีสีเข้ม เป็นต้น
- p) ไม่รongรับการหมุนของตัวอักษรบนหน้ายาและการเอียงของอุปกรณ์รับภาพ (Perspective distortion)

1.4. ประโยชน์ที่ได้รับ

หลังจากดำเนินงานวิจัยบรรลุตามวัตถุประสงค์ ผลงานวิจัยชิ้นนี้คาดว่าจะก่อให้เกิดประโยชน์ดังนี้

- a) เป็นเทคนิคซึ่งสามารถนำไปใช้ร่วมกับเทคนิคการวิเคราะห์ลักษณะภายนอกของเม็ดยาอื่นๆ เช่น เทคนิคการวิเคราะห์สีเม็ดยา เทคนิคการวิเคราะห์รูปรongรับเม็ดยา เป็นต้น เพื่อเพิ่มความแม่นยำให้กับระบบการจำแนกเม็ดยา
- b) เป็นเทคนิคที่นำไปใช้ในระบบจำแนกค่าคุณลักษณะภายนอกของเม็ดยาดั้วภาพ เพื่อเป็นประโยชน์สำหรับการสร้างฐานข้อมูลเม็ดยาในปัจจุบันซึ่งเป็นฐานข้อมูลชนิดข้อความให้มีข้อมูลเม็ดยาที่ครบถ้วนมากขึ้น
- c) เป็นเทคนิคสำหรับสร้างดัชนีจากสัญลักษณ์บนหน้ายาแบบอัตโนมัติเพื่อใช้ในการสืบค้น

บทที่ 2

วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะอธิบายเกี่ยวกับฐานข้อมูลยา ทฤษฎีและงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์สัญลักษณ์บนหน้ายาที่ให้ผลลัพธ์อยู่ในรูปแบบข้อความตัวอักษรได้อย่างอัตโนมัติ

2.1. ฐานข้อมูลยา

เนื่องจากงานวิจัยนี้นำเสนอวิธีการวิเคราะห์ข้อความตัวอักษรหน้ายาจากภาพซึ่งสามารถนำไปต่อยอดใช้ประโยชน์ร่วมกับฐานข้อมูลเม็ดยาที่มีอยู่ในปัจจุบันได้ ดังนั้นจึงศึกษารายละเอียดต่างๆของฐานข้อมูลยา เช่น ชนิดฐานข้อมูลที่ใช้ ข้อมูลยาที่เก็บ วิธีสืบค้น ความน่าเชื่อถือ เป็นต้น ซึ่งมีรายละเอียดดังต่อไปนี้

ฐานข้อมูล Pillbox (“Pillbox,” 2016) จุดเด่นอยู่ที่เป็นฐานข้อมูลยาที่มีความน่าเชื่อถือสูงเนื่องจากจัดทำโดย U.S. National Library of Medicine ที่เป็นฐานข้อมูลห้องสมุดยาที่ใหญ่ที่สุดในโลก แต่ในความเป็นจริงฐานข้อมูลดังกล่าวเก็บเฉพาะข้อมูลยาบริษัทต่างชาติเท่านั้น ไม่มีข้อมูลยาจากบริษัทไทย และสืบค้นด้วยวิธีรับข้อมูลจากคีย์บอร์ด เช่น ข้อมูลสัญลักษณ์หน้ายา รูปทรงเม็ดยา ชื่อยา ส่วนประกอบหลักของยา สี ชื่อผู้ผลิต ขนาด รหัสยา เป็นต้น โดยฐานข้อมูลจะเก็บข้อมูลต่างๆของยา เช่น รูปยา ชื่อทางการค้า ชื่อส่วนประกอบสำคัญ ปริมาณส่วนประกอบสำคัญ สัญลักษณ์บนหน้ายา ผู้ผลิต สี รูปทรง ขนาด

ฐานข้อมูล MIMS Thailand (“Image Advanced search-MIMS thailand,” 2016) เป็นอีกหนึ่งฐานข้อมูลที่มีความน่าเชื่อถือสูงเช่นเดียวกัน เนื่องจากเป็นฐานข้อมูลของ MIMS ซึ่งผลิตหนังสือเรียนสำหรับนักศึกษาเภสัชกรในไทย ซึ่งฐานข้อมูลนี้เก็บทั้งข้อมูลยาของบริษัทต่างชาติและยาจากบริษัทไทย และสืบค้นโดยให้ผู้ใช้กรอกข้อมูลเม็ดยาเช่น สี (เลือกได้สูงสุด 2 สี) รูปทรง รูปแบบ มีรอยบากหรือไม่ สัญลักษณ์บนหน้ายา ผู้ผลิต รหัสยาสำหรับสำนักงานอาหารและยาประจำประเทศไทย เป็นต้น ส่วนข้อมูลที่ฐานข้อมูลเก็บมีดังต่อไปนี้ ผู้ผลิต ผู้จัดจำหน่าย วิธีใช้ ส่วนประกอบสำคัญ ปริมาณที่ควรรับประทาน ยาที่ไม่ควรใช้ร่วมด้วย ระดับความอันตรายสำหรับผู้ตั้งครรภ์ หมวดยาประจำของสำนักงานอาหารและยาประจำประเทศไทย รูปยาประกอบ

ฐานข้อมูล Lexicomp (“Lexicomp,” 2016) เป็นฐานข้อมูลที่รองรับทั้งโหมดยออนไลน์และออฟไลน์ ซึ่งผลิตโดยบริษัท Wolters Kluwer ที่มีหุ้นส่วนเป็นองค์กรที่เกี่ยวข้องกับยามากมาย เช่น ‘American Pharmacists Association’ ‘ASHP’ ‘Canadian Pharmacists’ ‘The American Student Dental Association’ เป็นต้น ทำให้ฐานข้อมูล Lexicomp มีความน่าเชื่อถือสูง แต่ฐานข้อมูลมีเฉพาะข้อมูลยาจากบริษัทต่างประเทศเท่านั้น โดยเป็นฐานข้อมูลที่เก็บข้อมูลต่างๆของเม็ดยาตั้งแต่ปี 1980 ครอบคลุม รหัส NDC (รหัสของ FDA ประเทศสหรัฐอเมริกาโดยยาทุกชนิดที่อนุญาตให้จำหน่ายในประเทศสหรัฐอเมริกาต้องมีรหัสนี้) ชื่อทางการค้า ผู้ผลิต สี ลักษณะรอยบาก สัญลักษณ์บนหน้ายา ปริมาณสารประกอบหลัก วันที่อนุญาตเริ่มจัดจำหน่าย รูปทรง ซึ่งสืบค้นโดย สัญลักษณ์บนหน้ายา ลักษณะเม็ดยา รูปทรง สี (เลือกได้มากที่สุด 2 สี) แต่อย่างไรก็ดีฐานข้อมูลนี้สามารถเข้าถึงได้ก็ต่อเมื่อชำระค่าบริการก่อนเท่านั้น

ฐานข้อมูล Epocrates Rx Online (“epocrates,” 2016) เป็นฐานข้อมูลที่รองรับทั้งออนไลน์และออฟไลน์เช่นกัน แต่สร้างโดยบริษัท Athenahealth ซึ่งเป็นบริษัทขนาดใหญ่ที่ให้บริการระบบต่างๆที่เกี่ยวกับการดูแลสุขภาพของผู้คนโดยใช้ระบบ cloud และจัดทำแอปพลิเคชันด้านสุขภาพบนอุปกรณ์เคลื่อนที่อีกด้วย ทำให้ความน่าเชื่อถือของข้อมูลอยู่ในระดับสูงเทียบเท่ากับ MIMS แต่ปัญหาคือฐานข้อมูลนี้ไม่ได้เปิดเป็นข้อมูลสาธารณะกล่าวเข้าถึงได้ก็ต่อเมื่อชำระค่าบริการเสียก่อน อย่างไรก็ตามฐานข้อมูลนี้เก็บทั้งข้อมูลยาทั้งจากผู้ผลิตต่างประเทศและผู้ผลิตไทย และเป็นฐานข้อมูลที่สืบค้นโดยสัญลักษณ์บนหน้ายา รูปทรง สี (เลือกได้เพียงสีเดียว) และลักษณะรอยบากบนหน้ายา โดยข้อมูลทั้งหมดนี้ผู้ใช้ต้องกรอกข้อมูลอย่างน้อยหนึ่งอย่างผ่านเมนูสืบค้นของระบบ โดยข้อมูลที่ใช้ได้รับคือ วิธีการทานอย่างละเอียด ข้อควรระวัง ยาที่ไม่ควรใช้ร่วมด้วย กลุ่มบุคคลที่ควรระวังหลังจากใช้ยา ผู้ผลิต ราคา ยาอื่นๆที่สามารถใช้ทดแทนได้ และรูปประกอบ

ฐานข้อมูล YaAndYou (“YaAndYou,” 2016) เป็นฐานข้อมูลที่พัฒนาสำหรับอุปกรณ์เคลื่อนที่ที่จัดทำโดยมูลนิธิเพื่อการวิจัยและพัฒนาระบบยา (วพย.) จึงส่งผลให้ความน่าเชื่อถืออยู่ในระดับดี โดยจัดเก็บข้อมูลยาทั้งหมดที่อนุญาตจัดจำหน่ายได้ภายในประเทศไทย และสามารถสืบค้นโดยให้ผู้ใช้พิมพ์คีย์เวิร์ดที่เกี่ยวข้องกับยาลงบน Text field ซึ่งอาจเป็น ชื่อทางการค้า ชื่อส่วนประกอบสำคัญ เป็นต้น โดยข้อมูลยาที่ฐานข้อมูลนี้จัดเก็บประกอบไปด้วย ชื่อยาสามัญ ชื่อทางการค้า รูปแบบยา ข้อบ่งใช้ อาการไม่พึงประสงค์ แต่ไม่เก็บภาพเม็ดยา

ฐานข้อมูล WebMD (“WebMD,” 2016) เป็นฐานข้อมูลสาธารณะที่จัดทำโดยบริษัท WebMD เป็นบริษัทสัญชาติอเมริกาที่ให้บริการเกี่ยวกับข้อมูลยา ข่าวด้านสุขภาพ และคำแนะนำ ดังนั้นความน่าเชื่อถือของข้อมูลจึงอยู่ในระดับดี และฐานข้อมูลยังทำหน้าที่เก็บข้อมูลยาทุกชนิดที่อนุญาต

ให้จำหน่ายในประเทศสหรัฐอเมริกาได้ อนุญาตให้สืบค้นโดยกรอกข้อมูลสัญลักษณ์บนหน้ายา สี และ รูปทรง และตัวฐานข้อมูลเองจะให้ข้อมูลอย่างต่อเนื่อง รูปทรง ชื่อทางการค้า ปริมาณสารสำคัญ สี สัญลักษณ์บนหน้ายา วิธีใช้ ผลข้างเคียง ข้อควรระวัง

ฐานข้อมูล RxList (“RxList,” 2016) เป็นฐานข้อมูลสาธารณะและจัดทำโดย WebMD เช่นกันดังนั้นระดับความน่าเชื่อถือจึงอยู่ในระดับเดียวกันกับ WebMD และทำหน้าที่เก็บข้อมูลยา ทั้งหมดที่อนุญาตจำหน่ายในประเทศสหรัฐอเมริกา และสืบค้นโดยกรอกข้อมูลสัญลักษณ์บนหน้ายา สี รูปทรง ชื่อทางการค้า ชื่อสามัญ อย่างน้อยอย่างใดอย่างหนึ่ง สุดท้ายฐานข้อมูลจะให้ข้อมูลดังต่อไปนี้ ข้อควรระวัง วิธีใช้ ผลข้างเคียง โดยสรุปคือฐานข้อมูลนี้มีลักษณะคล้ายกับฐานข้อมูล WebMD มาก

ฐานข้อมูล Healthline (“Healthline,” 2016) จัดทำโดย Healthline Networks ที่เป็นเครือข่ายให้บริการด้านข้อมูลทางสุขภาพและเทคโนโลยีทางสุขภาพดังนั้นความน่าเชื่อถือของ ฐานข้อมูลจึงอยู่ในระดับดีซึ่งเก็บเฉพาะข้อมูลยาที่อนุญาตจำหน่ายในประเทศสหรัฐอเมริกาที่สืบค้น โดยกรอกชื่อยาหรือสัญลักษณ์หน้ายา หลังจากนั้นฐานข้อมูลจะให้ข้อมูลยาแก่ผู้ใช้นี้ รูปแบบยา รูปทรงยา ผลข้างเคียง ข้อบ่งใช้ รูปประกอบ ราคา

ฐานข้อมูล Drugs.com (“Drugs.com Pill Identifier,” n.d.) เป็นฐานข้อมูลออนไลน์ ที่สร้างจากการร่วมมือ 4 บริษัทเมดิคาลชั้นนำคือ ‘Wolters Kluwer Health’ ‘American Society of Health-System Pharmacists’ ‘Cerner Multum’ และสุดท้าย ‘Micromedex โดย Truven Health’ ทำให้ความน่าเชื่อถือของข้อมูลสูงมาก และเก็บข้อมูลยาทั้งหมดที่อนุญาตจัดจำหน่ายได้สำหรับ ประเทศสหรัฐอเมริกา และเป็นฐานข้อมูลเปิดที่อนุญาตให้เข้าถึงได้โดยไม่ต้องชำระค่าบริการ ส่วนการ สืบค้นข้อมูลยาทำได้โดยกรอกข้อมูลสัญลักษณ์บนหน้ายา รูปทรง สี ชื่อยา รหัส NDC และข้อมูลที่ ผู้ใช้ได้รับคือ รูปประกอบ สัญลักษณ์หน้ายา ปริมาณสารสำคัญ สี ขนาด (หน่วยมิลลิเมตร) รูปทรง จำหน่ายได้เฉพาะมีใบสั่งยาหรือไม่ คลาสของยา ระดับความรุนแรงสำหรับผู้ตั้งครุฑ ผู้ผลิต รหัส NDC

ฐานข้อมูล DailyMed (“DAILYMED,” 2016) เป็นฐานข้อมูลสาธารณะที่จัดทำโดย National Library of Medicine (NLM) ประเทศสหรัฐอเมริกาดังนั้นข้อมูลที่ได้รับจึงมีความ น่าเชื่อถือสูงมาก และเป็นฐานข้อมูลที่เก็บเฉพาะข้อมูลยาที่อนุญาตให้จำหน่ายในประเทศ สหรัฐอเมริกา และสืบค้นข้อมูลยาโดยให้ผู้ใช้กรอกข้อมูลชื่อยา รหัส NDC คลาสยา อย่างใดอย่างหนึ่ง จนสุดท้ายได้ข้อมูลอย่างต่อเนื่อง รหัส NDC ผู้ผลิต ผู้จำหน่าย หมู่ยา สถานการณ์จำหน่าย วิธีใช้ คำ เตือน เบอร์ดิตต่อสอบถาม ส่วนประกอบ ฤทธิ์ยา เป็นต้น

ฐานข้อมูลทะเบียนตำรับยา-องค์การอาหารและยาประจำประเทศไทย (“Pill information service, Thai food and drug administration,” 2016) จัดทำโดยสำนักงานคณะกรรมการอาหารและยาประจำประเทศไทยจึงฐานข้อมูลที่มีความน่าเชื่อถือสูงซึ่งเก็บข้อมูลยาทุกชนิดที่สามารถจัดจำหน่ายในประเทศไทยได้ และอนุญาตให้สืบค้นผ่านการกรอกข้อมูล ชื่อยาทางการค้า ชื่อสาระสำคัญ รหัสทะเบียนยา ชื่อผู้ผลิต ชื่อผู้แทนจำหน่าย และสิ่งที่ใช้ได้รับจากการสืบค้นฐานข้อมูลมีดังนี้ ชื่อทางการค้า เลขทะเบียนตามองค์การอาหารและยาประจำประเทศไทย ประเภทยา สถานการณ์จัดจำหน่ายในประเทศไทย รูปแบบยา ประเภทของยา ชนิดของการควบคุมทางกฎหมาย ข้อบ่งใช้ ส่วนประกอบสำคัญต่อหน่วย ชื่อผู้รับอนุญาตจำหน่าย ชื่อผู้ผลิต

จากการศึกษาฐานข้อมูลเมื่อยาที่น่าเชื่อถือในปัจจุบันข้างต้นสามารถสรุปเป็นตารางเปรียบเทียบคุณลักษณะได้ดังนี้ องค์กรที่รับผิดชอบ วิธีสืบค้น ประเภทข้อมูลที่เก็บ ประเภทฐานข้อมูล และมีสัญลักษณ์หน้ายาเป็นทางเลือกหนึ่งในการสืบค้นหรือไม่ดังตารางที่ 2.1 และตารางที่ 2.2

ตารางที่ 2.1 เปรียบเทียบคุณลักษณะฐานข้อมูลเมื่อยาที่ในปัจจุบัน

ชื่อฐานข้อมูล	องค์กรที่รับผิดชอบ	วิธีสืบค้น	ประเภทข้อมูลที่เก็บ	ประเภทฐานข้อมูล	มีสัญลักษณ์หน้ายาเป็นทางเลือกหนึ่งในการสืบค้นหรือไม่
(“Pillbox,” 2016)	หน่วยงาน U.S. National Library of Medicine	ผู้ใช้กรอกข้อมูลลักษณะภายนอกผ่านคีย์บอร์ด	เก็บเฉพาะข้อมูลยาบริษัทต่างชาติเท่านั้น ไม่มีข้อมูลยาจากบริษัทไทย	ฐานข้อมูลชนิดข้อความ	ใช่
(“Image Advanced search- MIMS thailand,” 2016)	หน่วยงาน MIMS	ผู้ใช้กรอกข้อมูลลักษณะภายนอกผ่านคีย์บอร์ด	เก็บทั้งข้อมูลยาของบริษัทต่างชาติและยาจากบริษัทไทย	ฐานข้อมูลชนิดข้อความ	ใช่
(“Lexicomp,” 2016)	บริษัท Wolters Kluwer	ผู้ใช้กรอกข้อมูลลักษณะภายนอกผ่านคีย์บอร์ด	เก็บเฉพาะข้อมูลยาบริษัทต่างชาติเท่านั้น ไม่มีข้อมูลยาจากบริษัทไทย	ฐานข้อมูลชนิดข้อความ	ใช่
(“epocrates,” 2016)	บริษัท athenahealth	ผู้ใช้กรอกข้อมูลลักษณะภายนอกผ่านคีย์บอร์ด	เก็บทั้งข้อมูลยาของบริษัทต่างชาติและยาจากบริษัทไทย	ฐานข้อมูลชนิดข้อความ	ใช่

ตารางที่ 2.2 เปรียบเทียบคุณลักษณะฐานข้อมูลเม็ดยาที่ในปัจจุบัน (ต่อ)

ชื่อฐานข้อมูล	องค์กรที่รับผิดชอบ	วิธีสืบค้น	ประเภทข้อมูลยาที่เก็บ	ประเภทฐานข้อมูล	มีสัญลักษณ์หน้าเป็นทางเลือกหนึ่งในการสืบค้นหรือไม่
("YaAndYou," 2016)	มูลนิธิเพื่อการวิจัยและพัฒนาาระบบยา (วพย.)	ผู้ใช้กรอกข้อมูลลักษณะภายนอกผ่านคีย์บอร์ด	เก็บเฉพาะข้อมูลยาที่อนุญาตให้จำหน่ายในประเทศไทยเท่านั้น	ฐานข้อมูลชนิดข้อความ	ไม่ใช่
("WebMD," 2016)	บริษัท WebMD	ผู้ใช้กรอกข้อมูลลักษณะภายนอกผ่านคีย์บอร์ด	เก็บเฉพาะข้อมูลยาที่อนุญาตให้จำหน่ายในประเทศสหรัฐอเมริกาเท่านั้น	ฐานข้อมูลชนิดข้อความ	ใช่
("RxList," 2016)	บริษัท WebMD	ผู้ใช้กรอกข้อมูลลักษณะภายนอกผ่านคีย์บอร์ด	เก็บเฉพาะข้อมูลยาที่อนุญาตให้จำหน่ายในประเทศสหรัฐอเมริกาเท่านั้น	ฐานข้อมูลชนิดข้อความ	ใช่
("Healthline," 2016)	เครือข่ายสุขภาพ Healthline	ผู้ใช้กรอกข้อมูลลักษณะภายนอกผ่านคีย์บอร์ด	เก็บเฉพาะข้อมูลยาที่อนุญาตให้จำหน่ายในประเทศสหรัฐอเมริกาเท่านั้น	ฐานข้อมูลชนิดข้อความ	ใช่
("Drugs.com Pill Identifier," n.d.)	ร่วมมือกันระหว่าง 1)บริษัท Wolters Kluwer Health 2) บริษัท American Society of Health-System Pharmacists 3)บริษัท Cerner Multum 4)บริษัท Micromedex	ผู้ใช้กรอกข้อมูลลักษณะภายนอกผ่านคีย์บอร์ด	เก็บเฉพาะข้อมูลยาที่อนุญาตให้จำหน่ายในประเทศสหรัฐอเมริกาเท่านั้น	ฐานข้อมูลชนิดข้อความ	ใช่
("DAILYMED," 2016)	หน่วยงาน National Library of Medicine (NLM)	ผู้ใช้กรอกข้อมูลลักษณะภายนอกผ่านคีย์บอร์ด	เก็บเฉพาะข้อมูลยาที่อนุญาตให้จำหน่ายในประเทศสหรัฐอเมริกาเท่านั้น	ฐานข้อมูลชนิดข้อความ	ไม่ใช่
("Pill information service, Thai food and drug administration," 2016)	องค์การอาหารและยาประจำประเทศไทย	ผู้ใช้กรอกข้อมูลลักษณะภายนอกผ่านคีย์บอร์ด	เก็บเฉพาะข้อมูลยาที่อนุญาตให้จำหน่ายในประเทศไทยเท่านั้น	ฐานข้อมูลชนิดข้อความ	ไม่ใช่

จากตารางที่ 2.1 และตารางที่ 2.2 พบว่างานที่เกี่ยวข้องกับเม็ดยาของหน่วยงานทางการแพทย์ที่น่าเชื่อถือมักปรากฏอยู่ในรูปฐานข้อมูลเม็ดยา ที่นิยมให้ผู้ใช้อกรอกข้อมูลสืบค้นผ่าน

คีย์บอร์ดที่ละเมิดยาและเป็นฐานข้อมูลชนิดข้อความที่เก็บและใช้สัญลักษณ์บนหน้ายาเป็นทางเลือกหนึ่งสำหรับการสืบค้น นอกจากนี้ฐานข้อมูลส่วนใหญ่ยังมีข้อมูลเมตาดาต้าที่ไม่ครบถ้วนเนื่องจากการขึ้นทะเบียนเมตาดาต้าใหม่ในทุกๆวันโดยเฉพาะข้อมูลเมตาดาต้าจากผู้ผลิตไทยเนื่องจากองค์กรส่วนใหญ่เป็นชาวต่างชาติ ประกอบกับการใช้เวลาสำหรับการสร้างฐานข้อมูลเมตาดาต้าให้มีข้อมูลเมตาดาต้าที่ครบถ้วนในแต่ละวัน ส่งผลให้เป็นปัญหาสำคัญสำหรับเภสัชกรไทยกรณีต้องการสืบค้นข้อมูลเมตาดาต้าที่ผลิตจากผู้ผลิตไทยที่ปราศจากหีบห่อหรือคำบ่งชี้

2.2. ทฤษฎีที่เกี่ยวข้อง

ทฤษฎีที่เกี่ยวข้องกับงานวิจัยเรื่องการวิเคราะห์สัญลักษณ์บนหน้ายาที่ให้ผลลัพธ์อยู่ในรูปแบบข้อความตัวอักษรบนหน้ายาได้อย่างอัตโนมัติเพื่อเป็นประโยชน์สำหรับการสร้างฐานข้อมูลเมตาดาต้าด้วยภาพมีทั้งหมด 7 เรื่องได้แก่ เทคนิคการสร้างภาพขอบวัตถุ เทคนิคการประมวลผลภาพโดยการเปลี่ยนแปลงลักษณะรูปร่างหรือโครงสร้างของภาพ เทคนิคการแยกองค์ประกอบสำคัญของภาพ หลักการพิจารณาพิกเซลวัตถุเดียวกันบนภาพขาวดำ เทคนิคพื้นฐานสำหรับการลดพิกเซลขอบทวน เทคนิคพื้นฐานสำหรับการปรับขนาดภาพ และการรู้จำอักขระด้วยแสง

เทคนิคการสร้างภาพขอบวัตถุ

เป็นเทคนิคการแปลงภาพระดับเทาเป็นภาพขาวดำที่เหลือเฉพาะพิกเซลขอบวัตถุซึ่งมีอัลกอริธึมที่นิยมเช่น เทคนิคของ Canny เทคนิคของ Sobel เทคนิคของ Prewitt เป็นต้น โดยสร้างภาพขอบวัตถุจากค่าความต่างของค่าสีที่พิกเซลตำแหน่งที่ $(i-1)$ ลบด้วยพิกเซลตำแหน่งที่ i ทำให้ขอบวัตถุคือจุดยอดของค่าความต่างของค่าสีซึ่งคือจุดที่มีค่าสีของพิกเซลเปลี่ยนไป ดังนั้นการหาขอบวัตถุในทิศแนวตั้ง (G_x)และทิศขวาง (G_y) ที่พิกัด $(1,1)$ จะคำนวณจาก

$$G_x = f(1,1)-f(2,1) \text{ หรือ } f(0,1)-f(1,1)$$

$$G_y = f(1,1)-f(1,2) \text{ หรือ } f(1,2)-f(1,1)$$

โดยฟังก์ชัน f แทนค่าสีของแต่ละพิกเซลในภาพ

-1	-1	-1	-1	0	1
0	0	0	-1	0	1
1	1	1	-1	0	1

ภาพที่ 2.1 เมทริกซ์ค่าน้ำหนักของฟิกเซลเพื่อนบ้านข้างเคียงของ Prewitt

จึงเกิดเมทริกซ์ค่าน้ำหนักของฟิกเซลเพื่อนบ้านข้างเคียงที่นำเสนอโดย Prewitt (Gonzalez & Woods, 2008) ดังภาพที่ 2.1 ต่อมา Sobel พัฒนาต่อยอดโดยเพิ่มค่าน้ำหนักของเพื่อนบ้านข้างเคียงในเมริกซ์ ณ แถวและคอลัมน์ตำแหน่งกลางตามทิศทางการหาขอบจนเกิดเป็นเมทริกซ์ค่าน้ำหนักของฟิกเซลเพื่อนบ้านข้างเคียงของ Sobel (Gonzalez & Woods, 2008) ดังภาพที่ 2.2

-1	-2	-1	-1	0	1
0	0	0	-2	0	2
1	2	1	-1	0	1

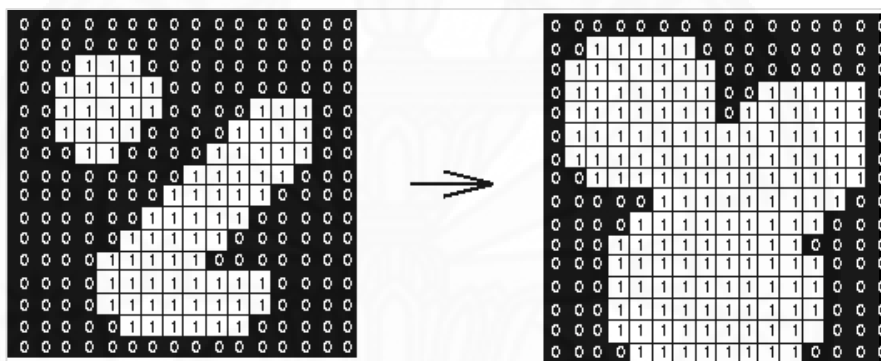
ภาพที่ 2.2 เมทริกซ์ค่าน้ำหนักของฟิกเซลเพื่อนบ้านข้างเคียงของ Sobel

ต่อมา Canny นำเสนอเมทริกซ์ค่าน้ำหนักของฟิกเซลเพื่อนบ้านข้างเคียงที่สามารถช่วยเบลอฟิกเซลรบกวน ณ ตำแหน่ง (1,1) ได้โดยนำหลักการของ Gaussian smoothing (Gonzalez & Woods, 2008) มาประยุกต์ ในขณะที่เทคนิคของ Prewitt และ Sobel จะพิจารณาว่าเป็นฟิกเซลขอบวัตถุทันที

เมื่อพิจารณาเทคนิคการหาภาพขอบวัตถุทั้ง 3 ทั้งในแง่ความแม่นยำ ความไว (Sensitivity) และความเจาะเพาะ (Specificity) รวมถึงเวลาที่ใช้ในการประมวลผลพบว่าเทคนิคที่เสนอโดย Canny สามารถให้ผลลัพธ์การประเมินดีที่สุดในทุกๆด้านเมื่อเปรียบเทียบกับเทคนิคการหาภาพขอบวัตถุอื่นที่นิยมเช่น เทคนิคที่นำเสนอโดย Sobel และเทคนิคที่นำเสนอโดย Prewitt (E.Annasaro et al, 2014)

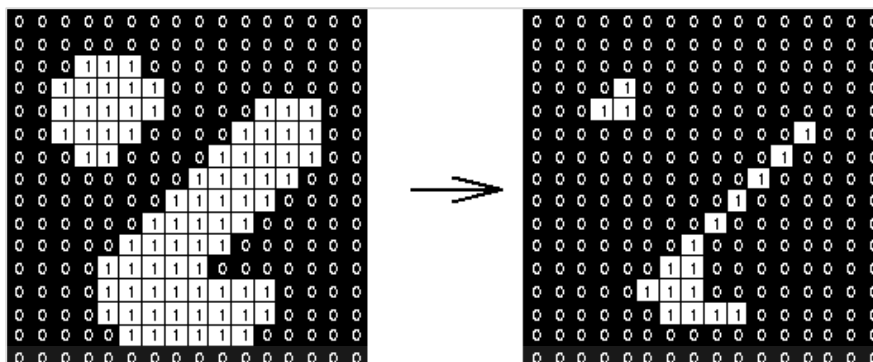
เทคนิคการประมวลผลภาพโดยการเปลี่ยนแปลงลักษณะรูปร่างหรือโครงสร้างของภาพ

เป็นเทคนิคที่ดัดแปลงมาจากทฤษฎีของเซตทางคณิตศาสตร์สามารถนำไปใช้ประโยชน์ได้หลากหลายหรือเป็นที่รู้จักกันในชื่อเทคนิค Morphology เช่น ใช้ถมรูฟิกเซลในวัตถุ (Fill object) ใช้หาตัวแทนคุณลักษณะ (Descriptor) ใช้เป็นเทคนิคเบลอฟิกเซลรบกวน (Smoothing filter) ใช้จำแนกฟิกเซลที่น่าสนใจ (Segmentation) หาฟิกเซลขอบวัตถุ เป็นต้น โดยอาศัยการทำงานกับภาพขนาดเล็กเรียกว่า Structuring element มักเป็นภาพสมมาตร โดย Morphology จะประกอบไปด้วย 2 การทำงานหลักคือ การทำงาน Dilation การทำงาน Erosion



ภาพที่ 2.3 ผลลัพธ์การทำงานของ Dilation

การทำงาน Dilation คือการขยายภาพโดยมีสัดส่วนเท่ากันทั่วทั้งภาพโดยเป็นการทำงานบนภาพขาวดำ (Binary image) การขยายภาพขึ้นอยู่กับ Structuring element หลักการคือ สแกนให้ทั่วทั้งภาพตามลำดับบนลงล่างซ้ายไปขวา และพิจารณาว่า ณ ตำแหน่งฟิกเซลที่ทับกันนั้นในภาพต้นฉบับมีตำแหน่งใดมีค่าฟิกเซลเท่ากับ 1 หรือไม่ ถ้าใช่ค่าฟิกเซลในภาพผลลัพธ์ ณ ตำแหน่งจุดศูนย์กลางของ Structuring element จะมีค่าเท่ากับ 1 ดังภาพที่ 2.3



ภาพที่ 2.4 ผลลัพธ์การทำงานของ Erosion

สำหรับการ Erosion ซึ่งเป็นการย่อภาพดังนั้นจะให้ผลลัพธ์ตรงข้ามกับการทำงาน Dilation หลักการคือค่าพิกเซลในภาพผลลัพธ์ ณ ตำแหน่งจุดศูนย์กลางของ Structuring element จะมีค่าเท่ากับ 1 ก็ต่อเมื่อทุกตำแหน่งพิกเซลที่ทับกันในภาพต้นฉบับต้องมีค่าเท่ากับ 1 ดังภาพที่ 2.4

นอกจากนี้ยังมีอีก 2 การทำงานที่เกิดจากการทำงานต่อเนื่องกันระหว่างการทำงาน Dilation ตามด้วยการทำงาน Erosion เช่น การทำงาน Closing และการทำงาน Opening

เทคนิคการแยกองค์ประกอบสำคัญของภาพ

เทคนิคการแยกองค์ประกอบสำคัญของภาพหรือ Image segmentation เป็นขั้นตอนหนึ่งที่สำคัญสำหรับการวิเคราะห์ภาพเพื่อหากลุ่มพิกเซลที่ประกอบกันเป็นวัตถุที่สนใจในภาพซึ่งสามารถจำแนกกลุ่มวิธีการได้ 5 กลุ่มดังนี้

กลุ่มวิธีการ Threshold based segmentation หรือ Histogram thresholding เป็นเทคนิคที่พยายามแบ่งฮิสโตแกรมสีของพิกเซลภาพเป็นส่วนๆ และเป็นกลุ่มเทคนิคที่นิยมใช้บ่อยสำหรับการจำแนกองค์ประกอบสำคัญในภาพระดับเทาซึ่งมีหลักการทำงานดังฟังก์ชันด้านล่าง

$$g(v) = \begin{cases} 0, & v < t \\ 1, & v \geq t \end{cases}$$

ซึ่ง v แทนค่าพิกเซลระดับเทา และ t คือค่า threshold ที่หาได้ ดังนั้นจึงใช้เป็นเทคนิคการแปลงภาพระดับเทาเป็นภาพขาวดำที่มีค่าพิกเซล 2 ค่าคือ 0 และ 1 เทคนิคประเภทนี้จะทำงานได้ดีกับกรณีที่มีลักษณะฮิสโตแกรมสีระหว่างกลุ่มพิกเซลที่สนใจและกลุ่มพิกเซลพื้นหลังแยกกันอย่างเห็นได้ชัด และในกรณีที่มีวัตถุที่สนใจมากกว่าหนึ่งที่มีกลุ่มฮิสโตแกรมสีแยกจากกันชัดเจนเทคนิคนี้สามารถจำแนกกลุ่มวัตถุที่สนใจแต่ละกลุ่มได้โดยใช้เทคนิค Labelling algorithm บนภาพผลลัพธ์การทำ Histogram

thresholding นอกจากนี้กลุ่มวิธีการนี้ยังมีกลุ่มวิธีการย่อยที่สำคัญมากมายเช่น กลุ่ม Minimum variance within segments คือกลุ่มเทคนิคที่พยายามหาค่า threshold ที่ทำให้ค่าความต่างของค่าพิกเซลระดับเทาภายในกลุ่มเดียวกันต่างกันอย่างน้อยที่สุดในขณะที่ให้ค่าความต่างระดับเทาระหว่างกลุ่มพิกเซลมีค่ามากที่สุด เช่น เทคนิค Otsu's thresholding นอกจากนี้ยังมีกลุ่ม Local thresholding ที่พยายามแบ่งภาพเป็นภาพย่อยหลายๆภาพและทำ Histogram thresholding บนภาพย่อยๆเพื่อให้ได้ภาพขาวดำผลลัพธ์จากการประกอบภาพย่อยกลับเป็นภาพตั้งต้นซึ่งเหมาะกับภาพที่ต้องการการจำแนกวัตถุโดยละเอียดที่มีกลุ่มสีพิกเซลมากกว่า 2 กลุ่ม และกลุ่มสุดท้ายที่น่าสนใจคือกลุ่ม Multispectral thresholding ซึ่งเป็นกลุ่มที่ใช้ค่าสีของพิกเซลให้เป็นประโยชน์ในการจำแนกวัตถุเช่น มีภาพสี RGB ซึ่งประกอบไปด้วย 3 channel สีให้พิจารณาว่าเป็นภาพระดับเทา 3 ภาพประกอบกัน ดังนั้นจะทำการ threshold โดยทำทีละภาพ channel สี และภาพขาวดำผลลัพธ์ได้จากการรวมภาพโดยใช้ Boolean operation ซึ่งจะให้ผลการจำแนกวัตถุในภาพละเอียดกว่าการทำงานบนภาพระดับเทาปกติ

กลุ่มวิธีการ Edge based segmentation เทคนิคประเภทนี้จะเริ่มจากการหาเส้น (edge) ทั้งหมดที่ปรากฏในภาพ และให้เส้นเหล่านั้นแทนขอบวัตถุที่สนใจพร้อมถมพิกเซลภายในขอบวัตถุเพื่อจำแนกวัตถุที่สนใจในภาพ โดยมักทำงานร่วมกับเทคนิคจำพวก Edge linking หรือเทคนิค Watershed segmentation เพื่อหาขอบวัตถุเชื่อมเป็นเส้นเดียวกัน ดังนั้นจึงไม่เหมาะกับภาพที่ไม่สามารถเชื่อมขอบวัตถุเป็นเส้นเดียวกันได้

กลุ่มวิธีการ Region based segmentation เป็นเทคนิคที่มีหลักการตรงกันข้ามกับกลุ่มวิธีการ Edge based segmentation โดยเริ่มจากการหาพิกเซลใดๆภายในวัตถุหลังจากนั้นขยายกลุ่มพิกเซลรอบๆ เรียกว่า growing จนกว่าจะพบพิกเซลที่เป็นขอบวัตถุ ซึ่งกลุ่มวิธีการนี้มีการทำงานพื้นฐาน 2 การทำงานคือ การแยก (splitting) และการรวม (merging) หลักการทำงานของการทำงานรวมเริ่มจากการหาพิกเซลตั้งแต่และพยายามขยายพิกเซลตั้งต้นไปยังเพื่อนบ้านข้างเคียงที่มีลักษณะบางอย่างเหมือนกันและขยายกลุ่มพิกเซลจนกระทั่งไม่สามารถขยายต่อได้ ส่วนการทำงานของการทำงานแยกเริ่มจากมองทั้งภาพเป็นพิกเซลกลุ่มเดียวกันหากพบว่ามีความสว่างต่างกันจะแตกภาพออกเป็นชิ้นส่วนย่อยและแตกภาพซ้ำๆจนกว่าทุกส่วนย่อยของภาพประกอบไปด้วยพิกเซลที่มีความสว่างใกล้เคียงกัน

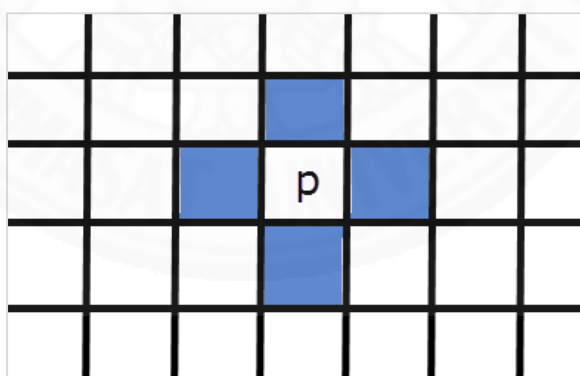
กลุ่มวิธีการ Clustering techniques เป็นกลุ่มเทคนิคการจัดกลุ่มพิกเซลที่มีลักษณะเหมือนกันไม่ทางใดก็ทางหนึ่ง (agglomerative segmentation) โดยมีจุดมุ่งหมายเพื่อจำแนกวัตถุที่สนใจภายในภาพ โดยส่วนใหญ่นิยมทำงานกับภาพ 2 ถึง 3 มิติ หลักการจัดกลุ่มจะพิจารณาจากค่าระยะห่างซึ่งมักใช้อัลกอริธึม Euclidean distance โดยให้ค่าความต่างระหว่าง 2 พิกเซลใดๆแทน

ด้วยระยะห่างระหว่างค่าสีของ 2 พิกเซล โดยกลุ่มวิธีการนี้แบ่งได้เป็น 2 วิธีการพื้นฐาน หนึ่งคือกลุ่มวิธีการ Agglomerative clustering ซึ่งเริ่มต้นจะพิจารณาให้แต่ละพิกเซลเป็น 1 กลุ่มพิกเซลย่อย หลังจากนั้นตัดสินใจว่ากลุ่มพิกเซลใดควรรวมเป็นกลุ่มเดียวกันโดยดูจากกลุ่มสองกลุ่มที่มีค่าระยะห่างน้อยที่สุด และรวมกลุ่มพิกเซลไปเรื่อยๆจนกว่าจะได้จำนวนกลุ่มพิกเซลที่ต้องการ สองคือกลุ่มวิธีการ Partitional clustering เช่นเทคนิค K-means clustering ซึ่งนิยมจัดพิกเซลลงกลุ่มที่เหมาะสมโดยพิจารณาจากระยะห่างที่น้อยที่สุดระหว่างพิกเซลนั้นและพิกเซลตัวแทนของแต่ละกลุ่ม

กลุ่มวิธีการ Matching เหมาะกับงานที่ทราบลักษณะของวัตถุที่ต้องการจำแนกในภาพ หรือ template โดยจะพยายามหาตำแหน่งวัตถุที่สนใจโดยหา template ของวัตถุที่สนใจไปยังตำแหน่งต่างๆบนภาพจนกว่าจะเจอ

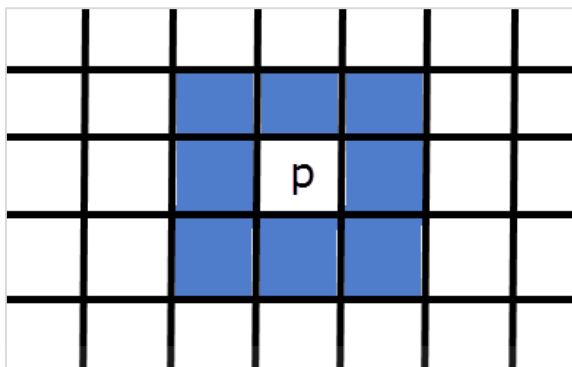
หลักการพิจารณาพิกเซลวัตถุเดียวกันบนภาพขาวดำ

หลักการพิจารณาพิกเซลวัตถุเดียวกันบนภาพขาวดำซึ่งประกอบไปด้วยค่าพิกเซลที่เป็นไปได้เพียง 2 ค่าคือ 0 และ 1 โดยให้ 1 แทนพิกเซลวัตถุที่สนใจ หรือเรียกว่า pixel connectivity คือการพิจารณาคู่ของพิกเซลใดๆที่เกี่ยวข้องกันหรือไม่ ณ ที่นี้คือเป็นพิกเซลที่อยู่ในวัตถุเดียวกันหรือไม่ (กำหนดให้แทนเป็นพิกเซล p และพิกเซล q) ซึ่งมีหลักการพิจารณา 3 แบบดังนี้



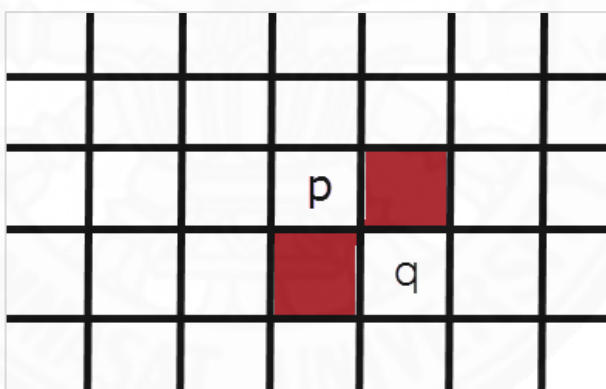
ภาพที่ 2.5 พิกเซลเพื่อนบ้านข้างเคียง (สีฟ้า) สี่ทิศทางของ p

หลักการ 4-connectivity จะตัดสินพิกเซลที่มีค่าเท่ากับ 1 สองพิกเซลใดๆที่เกี่ยวข้องกันหรือไม่จะพิจารณาจากพิกเซล q เป็นหนึ่งในเพื่อนบ้านข้างเคียงแบบสี่ทิศทางของพิกเซล p หรือไม่ ดังภาพที่ 2.5



ภาพที่ 2.6 พิกเซลเพื่อนบ้านข้างเคียง (สี่ฟ้า) แปรทิศทางของ p

หลักการ 8-connectivity จะตัดสินพิกเซลที่มีค่าเท่ากับ 1 สองพิกเซลใดๆว่าเกี่ยวข้องกันหรือไม่จะพิจารณาจากพิกเซล q เป็นหนึ่งในเพื่อนบ้านข้างเคียงแบบแปรทิศทางของพิกเซล p หรือไม่ดังภาพที่ 2.6



ภาพที่ 2.7 พิกเซลสีแดงแทนพิกเซลที่เป็นเพื่อนบ้านข้างเคียงสี่ทิศทางทั้งพิกเซล p และ q

หลักการ m -connectivity จะตัดสินพิกเซลที่มีค่าเท่ากับ 1 สองพิกเซลใดๆว่าเกี่ยวข้องกันก็ต่อเมื่อเป็นจริงจากหนึ่งในสองเงื่อนไขดังนี้

- พิกเซล q ต้องเป็นหนึ่งในเพื่อนบ้านข้างเคียงแบบสี่ทิศทางของพิกเซล p
- พิกเซล q เป็นหนึ่งในเพื่อนบ้านข้างเคียงแบบตามแนวทะแยงของพิกเซล p (เป็นเพื่อนบ้านข้างเคียงแบบแปรทิศทางและต้องไม่เป็นเพื่อนบ้านข้างเคียงแบบสี่ทิศทาง) และต้องไม่พบว่ามีพิกเซลที่เป็นทั้งเพื่อนบ้านข้างเคียงสี่ทิศทางของทั้ง p และ q ที่มีค่าเท่ากับ 1 ดังภาพที่ 2.7

ดังนั้นการหาวัตถุที่ได้จากหลักการ 8-connectivity จะพบวัตถุได้ง่ายที่สุด รองลงมาคือ m-connectivity และหลักการที่เกิดวัตถุในภาพยากที่สุดคือ 4-connectivity

เทคนิคพื้นฐานสำหรับการลดพิกเซลรบกวน

เทคนิคพื้นฐานสำหรับการลดพิกเซลรบกวนแบ่งได้เป็น 3 เทคนิคดังนี้

เทคนิค Mean filter ซึ่งจะจางพิกเซลรบกวนด้วยการแทนค่าสีในภาพผลลัพธ์ด้วยค่าเฉลี่ยของสีพิกเซลที่คำนวณจากพิกเซลเพื่อนบ้านข้างเคียงด้วยวิธี convolution ข้อดีของเทคนิคนี้คือพิกเซลสำคัญของภาพตั้งต้นเช่น พิกเซลขอบวัตถุจางไปด้วยส่งผลให้ภาพผลลัพธ์เบลอทั้งภาพและเนื่องจากใช้มีการทำงาน convolution ร่วมด้วยทำให้เกิดการประมวลผลมาก

เทคนิค Median filter เป็นเทคนิคที่จางพิกเซลรบกวนด้วยการหาตำแหน่งค่ากลาง (median) ของค่าสีพิกเซลเพื่อนบ้านใกล้เคียง ภาพผลลัพธ์ที่ได้พิกเซลรบกวนจะหายไปเลยโดยไม่ใช้การทำงาน convolution ซึ่งจะเหมาะกับการลดพิกเซลรบกวนประเภท Salt and pepper noise แต่เนื่องจากวิธีการหาค่าตำแหน่งกลาง (median) จำเป็นต้องมีการเรียงลำดับข้อมูลดังนั้นจึงก่อให้เกิดการประมวลผลมาก

สุดท้ายเทคนิค Gaussian filter จะจางพิกเซลรบกวนโดยทำ convolution กับหน้ากาก(mask) ซึ่งเป็นภาพขนาดเล็กที่มีค่าพิกเซลศูนย์กลางมากที่สุดและมีค่าพิกเซลบนหน้ากากลดหลั่นลงเรื่อยๆจากพิกเซลศูนย์กลาง ซึ่งเหมาะสำหรับการจางพิกเซลรบกวนประเภท Gaussian noise ซึ่งเป็นพิกเซลรบกวนประเภทที่พบมากที่สุดที่สุดในธรรมชาติ

เทคนิคพื้นฐานสำหรับการปรับขนาดภาพ

เทคนิคการปรับขนาดภาพ (Interpolation) คือ วิธีการสำหรับหาตัวแทนให้กับค่าพิกเซลในภาพผลลัพธ์จากการขยายหรือการย่อภาพ ซึ่งมีเทคนิคพื้นฐานดังต่อไปนี้

เทคนิค Nearest-neighbor interpolation เป็นเทคนิคที่ง่ายที่สุดสำหรับการปรับขนาดภาพ เนื่องจากค่าพิกเซลในภาพผลลัพธ์ได้จากการพิจารณาค่าพิกเซลในภาพตั้งต้นที่มีตำแหน่งใกล้เคียงที่สุด หรือรู้จักกันในนามเทคนิค Point shift algorithm หรือเทคนิค Pixel replication ซึ่งภาพผลลัพธ์จะได้จากการทำงาน convolution ที่มีค่าสมาชิกบน Interpolation kernel ที่คำนวณจากฟังก์ชันด้านล่าง

$$h(x) = \begin{cases} 1, & 0 \leq |x| < 0.5 \\ 0, & 0.5 \leq |x| \end{cases}$$

โดย x แทนระยะห่างระหว่างตำแหน่งพิกเซลในภาพผลลัพธ์และพิกเซลเพื่อนบ้านข้างเคียง ณ ตำแหน่งที่ kernel ทาบ ดังนั้นภาพผลลัพธ์ที่ได้จะสามารถคงรายละเอียดของภาพตั้งต้นแต่ไม่เกิดอาการไม่เรียบบริเวณพิกเซลขอบวัตถุ (jagged edge) ซึ่งเป็นลักษณะเฉพาะของภาพผลลัพธ์ที่ได้จากการปรับขนาดด้วยเทคนิค Nearest-neighbor

เทคนิค Bilinear interpolation หรือคือการคำนวณค่าในฟังก์ชันเส้นตรงบนทิศทางสองมิติซึ่งค่าความแม่นยำและการประมวลผลจะขึ้นกับขนาด Kernel โดยตรงดังนั้นการออกแบบและการวิเคราะห์หา kernel ที่เหมาะสมจึงเป็นหัวใจหลักของเทคนิคนี้ ภาพผลลัพธ์ของเทคนิคบริเวณพิกเซลขอบจะเรียบเนียนมากกว่าการปรับขนาดภาพด้วยเทคนิค Nearest-neighbor

เทคนิค Bicubic interpolation เป็นอัลกอริธึมสำหรับการปรับขนาดภาพที่ใช้ค่ายกกำลังสาม ซึ่งเป็นการประมาณค่าที่ได้จากฟังก์ชัน Sinc interpolation ขั้นสูงสุด โดยภาพผลลัพธ์ที่ได้จะมีลักษณะคล้ายกับภาพผลลัพธ์จากเทคนิค Lanczos interpolation คือ บริเวณพิกเซลขอบวัตถุในภาพผลลัพธ์จะเรียบเนียนที่สุดเมื่อเปรียบเทียบกับ การปรับขนาดภาพด้วยเทคนิค Nearest-neighbor และเทคนิค Bilinear

การรู้จำอักขระด้วยแสง

การรู้จำอักขระด้วยแสงหรือ Optical character recognition หรือ OCR (Mantas, 1986) เป็นกระบวนการการเปลี่ยนตัวอักขระจากภาพไม่ว่าจะเป็นอักขระพิมพ์หรืออักขระลายมือเป็นข้อความตัวอักษรที่คอมพิวเตอร์อ่านได้ ซึ่งใช้ในหน่วยงานประเภทที่ต้องจำแนกข้อความตัวอักษรจากแบบฟอร์มในปริมาณมาก ๆ ต่อวันเช่น งานธนาคาร เป็นต้น นอกจากนี้นิยมใช้สำหรับจำแนกข้อมูลจากใบเสร็จต่าง รวมถึงข้อมูลนามบัตรด้วย อย่างไรก็ตามระบบรู้จำส่วนใหญ่ต้องมีกระบวนการสอนร่วมด้วยสำหรับแต่ละอักขระเพื่อเพิ่มความแม่นยำ

เดิมทีเทคนิคการรู้จำอักขระด้วยแสงพัฒนาขึ้นเพื่อใช้ในอุปกรณ์ช่วยอ่านสำหรับผู้บกพร่องทางสายตา ซึ่งในปัจจุบันนำไปใช้อย่างแพร่หลายในงานวิจัยเช่น งานรู้จำเลขทะเบียนรถยนต์ (Automatic number plate recognition) งานวิจัยเกี่ยวกับการพัฒนาแคปต์ชา (CAPCHA) เพื่อแยกแยะมนุษย์และคอมพิวเตอร์ (โปรแกรมอัตโนมัติ) ซึ่ง OCR ช่วยสำหรับจำลองความสามารถของโปรแกรมอัตโนมัติเพื่อทดสอบประสิทธิภาพของงานวิจัย

อัลกอริธึมของ OCR แบ่งได้เป็น 2 ประเภทได้แก่ ประเภท Pattern recognition และ ประเภท Feature extraction อัลกอริธึมประเภท Pattern recognition หรือ Image correlation จะรู้จำสัญลักษณ์บนภาพพิกเซลต่อพิกเซลโดยความแม่นยำขึ้นอยู่กับคุณภาพภาพอักขระสีบคั่นและภาพอักขระเทียบของระบบ อัลกอริธึมประเภทนี้ทำงานได้ดีกับอักขระลายมือ สำหรับอัลกอริธึมประเภท Feature extraction จะพยายามจำแนกค่าคุณลักษณะต่างๆของอักขระในภาพให้อยู่ในรูปแบบเวกเตอร์ซึ่งเป็นเทคนิคที่นิยมนำมาใช้ในโปรแกรม OCR ปัจจุบัน เช่น โปรแกรม Tesseract เป็นต้น โดยการเปรียบเทียบความคล้ายของเวกเตอร์อักขระนิยมใช้เทคนิคประเภท Nearest neighbor classifiers

2.3. งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับการวิเคราะห์สัญลักษณ์บนหน้ายาที่ให้ผลลัพธ์อยู่ในรูปแบบข้อความตัวอักษรบนหน้ายาได้อย่างอัตโนมัติได้แก่ งานวิจัยจำแนกเม็ดยาด้วยภาพ และงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์ตัวอักษร

งานวิจัยจำแนกเม็ดยาด้วยภาพ

งานวิจัยจำแนกเม็ดยาด้วยภาพนำเสนอเทคนิคการวิเคราะห์คุณลักษณะภายนอกของเม็ดยาเพื่อสร้างระบบการจำแนกเม็ดยาอัตโนมัติซึ่งมีรายละเอียดดังนี้

งานวิจัย (Yu et al., 2014) นำเสนอการจำแนกเม็ดยาด้วยลักษณะภายนอกเม็ดยาดังนี้ รูปทรง สี สัญลักษณ์บนหน้ายาจากภาพ โดยใช้รูปทรงและสีเม็ดยาเพื่อคัดกรองเม็ดยาค่าตอบครั้งแรก และใช้สัญลักษณ์หน้ายาเพื่อคัดกรองเม็ดยาค่าตอบครั้งสุดท้าย เทคนิคที่ใช้วิเคราะห์สัญลักษณ์บนหน้ายา คือ Two-step sampling distance sets (TSDS)(Grigorescu & Petkov, 2003) ซึ่งให้ผลลัพธ์อยู่ในรูปเวกเตอร์ขนาดใหญ่ก่อให้เกิดการประมวลผลมาก

เทคนิค TSDS เป็นการพัฒนาต่อยอดจากเทคนิค Distance set (DS) ที่เสนอในงานวิจัย (Grigorescu & Petkov, 2003) ซึ่งเป็นเทคนิคสำหรับอธิบายรูปทรงในภาพสองมิติที่มีค่าอธิบายคุณลักษณะผลลัพธ์ขึ้นกับค่าของพิกเซลเพื่อนบ้านข้างเคียง ดังนั้นสำหรับเทคนิค TSDS จึงนำเทคนิค DS มาพัฒนาต่อยอดเพื่อใช้สำหรับอธิบายคุณลักษณะของสัญลักษณ์บนหน้ายาที่ประกอบไปด้วยตัวอักษร สัญลักษณ์ต่างๆ โดยหลักการทำงานของ TSDS เริ่มจากหา DS ของระยะห่างระหว่างพิกัด

ใดๆไปยังพิกเซลเพื่อนบ้านข้างเคียง k พิกเซลบริเวณหน้ายา ดังนั้นผลลัพธ์ที่ได้คือเซตของระยะห่างที่สามารถบอกลักษณะการจัดเรียงพิกเซลที่อยู่รอบๆพิกเซลนั้นได้ ส่งผลให้เซตของเซตระยะห่างสามารถใช้อธิบายภาพรวมของสัญลักษณ์บนหน้ายาได้เช่นกัน ลักษณะเด่นของพิกเซลสัญลักษณ์บนหน้ายาคือซับซ้อนและไม่มีรูปแบบตายตัว ดังนั้นเป็นไปได้ที่จะถูกพิจารณาเป็นพิกเซลรบกวน จึงเกิดวิธี Two step sampling strategy เพราะจะมีขั้นตอนการหาพิกเซลที่เป็นสัญลักษณ์บนหน้ายาจริงถึง 2 ขั้นตอนก่อนนำไปหาค่าอธิบายคุณลักษณะผลลัพธ์ ขั้นที่1 เริ่มจากดึงพิกเซลที่คาดว่าเป็นสัญลักษณ์บนหน้ายาทั้งหมดออกมาจากภาพต้นฉบับก่อนในช่วงเตรียมภาพด้วยการสร้างภาพขาวดำของสัญลักษณ์บนหน้ายาต่อมาสุ่มหาพิกเซลตัวแทนของพิกเซลขอบของแต่ละวัตถุบนหน้ายาด้วยวิธีการสุ่มความถี่เดียว และจัดกลุ่มพิกเซลตัวแทนที่อยู่บนขอบวัตถุเดียวกันเป็น 1 พื้นที่ย่อยดังภาพที่ 2.8



ภาพที่ 2.8 ผลลัพธ์การจัดกลุ่มพิกเซลขั้น 1

ต่อมาเป็นการคัดกรองพิกเซลสัญลักษณ์บนหน้ายาขั้นที่ 2 ดังภาพที่ 2.9 เพื่อกำจัดกลุ่มพิกเซลรบกวนซึ่งมักเป็นพื้นที่ขนาดเล็ก กล่าวคือให้พิจารณากลุ่มพิกเซลที่มีสมาชิกน้อยกว่าเกณฑ์เป็นกลุ่มพิกเซลรบกวน



ภาพที่ 2.9 ผลลัพธ์การคัดกรองกลุ่มพิกเซลขั้น 2

หลังจากนั้นเป็นขั้นตอนการคำนวณหาเวกเตอร์ตัวแทนคุณลักษณะของสัญลักษณ์บนหน้ายา ดังนั้นจะได้ค่าตัวแทนคุณลักษณะที่ไม่ขึ้นกับทิศทางการหมุนและตำแหน่งของเม็ดยาในภาพ ต่อมาในปี 2015 งานวิจัย (Chen et al., 2015) และงานวิจัย (Chen & Kamata, 2013) นำเสนอวิธีการสร้างภาพขาวดำที่ต่อยอดจากเทคนิค Stroke width transform (SWT)(Epshtein, Ofek, & Wexler, 2010) ซึ่งเป็นเทคนิคสร้างภาพขาวดำที่จำแนกเฉพาะตัวอักษรบนภาพจากภาพระดับเทาโดยมีแนวคิดที่ว่าเส้นทุกเส้นที่ปรากฏในภาพจะเป็นตัวอักษรก็ต่อเมื่อตลอดทั้งเส้นมีความหนาที่ใกล้เคียงกัน ความหนาของเส้นพิจารณาจากทิศทางค่าความต่างของแต่ละพิกเซลบนภาพขอบวัตถุ ดังนั้น Chen จึงเสนอให้ประยุกต์ใช้เทคนิค Switch function ในขั้นตอนการเตรียมภาพก่อนเข้าเทคนิค SWT โดยจะใช้สมการ Bounded function ซึ่งเป็นสมการหนึ่งมิติ เพื่อกำจัดพิกเซลหน้ายาที่ไม่เกี่ยวข้องกับสัญลักษณ์

บนหน้ายา และเพิ่มขึ้นตอน Accumulated gradient เพื่อปรับปรุงค่าพิกเซลในภาพขอบวัตถุที่ SWT ใช้ไม่ให้โดดเด่นมาจากพิกเซลเพื่อนบ้านใกล้เคียง และนำเทคนิคที่ต่อยอดจาก SWT ข้างต้นมาหาค่าตัวแทนคุณลักษณะสัญลักษณ์บนหน้ายาด้วยเทคนิค TSDS ซึ่งผลการทดลองพบว่าสามารถเพิ่มความแม่นยำของระบบได้ถึง 4 เปอร์เซ็นต์โดยประมาณ

งานวิจัย (Caban et al., 2012) เสนอวิธีการจำแนกเม็ดยาด้วยภาพที่ใช้คุณลักษณะร่วมกันของรูปทรง สี และสัญลักษณ์บนหน้ายา จุดเด่นอยู่ที่วิธีการวิเคราะห์สัญลักษณ์บนหน้ายาและรูปทรงได้จากการต่อยอดเทคนิค Shape distribution (Osada, Funkhouser, Chazelle, & Dobkin, 2002) ซึ่งเป็นเทคนิคสำหรับหาค่าตัวแทนลักษณะรูปทรงของรูปทรง 3 มิติด้วยหลักการหาค่าความสัมพันธ์ระหว่าง 2 พิกเซลตัวแทนที่ให้ผลลัพธ์อยู่ในรูปเวกเตอร์ ดังนั้นตัวแทนลักษณะสัญลักษณ์บนหน้ายาจึงเป็นเวกเตอร์ของคุณลักษณะของข้อมูลเวกเตอร์ผลลัพธ์จากเทคนิค Shape distribution เช่น ค่าระยะห่างที่น้อยที่สุดจากเวกเตอร์ของ Osada ค่าระยะห่างที่มากที่สุดจากเวกเตอร์ของ Osada ค่าระยะห่างเฉลี่ย ค่าเบี่ยงเบนมาตรฐานของข้อมูลระยะห่างจากเวกเตอร์ของ Osada เป็นต้น ส่งผลให้ตัวแทนคุณลักษณะผลลัพธ์ของสัญลักษณ์บนหน้ายาไม่ขึ้นกับองศาการหมุน ตำแหน่ง และขนาดของเม็ดยาในภาพ

งานวิจัย (Lee, Park, Jain, & Lee, 2012) เป็นงานวิจัยที่ต่อยอดมาจากงาน (Lee, Park, & Jain, 2010) ซึ่ง (Lee et al., 2012) นำเสนอระบบการจำแนกเม็ดยาด้วยภาพโดยใช้รูปทรง สี สัญลักษณ์บนหน้ายา ที่ประยุกต์ใช้เทคนิค Scale-invariant keypoints (SIFT)(Lowe, 2004) หลังการปรับปรุงร่วมกับเทคนิค Multi-scale Local Binary Pattern (MLBP)(Ojala, T., Pietikainen, M., Maenpaa, T., 2002) เพื่อหาตัวแทนคุณลักษณะของสัญลักษณ์บนหน้ายาส่งผลให้เกิดภาระการคำนวณค่อนข้างสูง โดยขั้นตอนการหาค่าตัวแทนสัญลักษณ์บนหน้ายาจะทำงานบนภาพขอบวัตถุซึ่งเลือกใช้เทคนิคการหาภาพขอบวัตถุที่นำเสนอโดย Sobel ที่มีรายละเอียดในหัวข้อ 2.2 เนื่องจากเทคนิค SIFT และ MLBP จะทำงานได้ดีกว่ากับภาพขอบวัตถุเมื่อเทียบกับการทำงานภาพขาวดำและภาพสี ข้อดีของ SIFT คือตัวแทนคุณลักษณะที่ได้จะไม่ขึ้นกับตำแหน่งของเม็ดยาในภาพ องศาการหมุนของเม็ดยา รวมถึงขนาดของเม็ดยาในภาพ ส่วนเทคนิค MLBP จะนำมาประยุกต์ที่หลังเพื่อจำแนกลักษณะพื้นผิวเม็ดยาบริเวณที่มีสัญลักษณ์บนหน้ายา สุดท้ายเวกเตอร์ตัวแทนเม็ดยาจะได้รับการจัดเรียงค่าคุณลักษณะต่างๆทั้งหมด

การปรับปรุงเทคนิค SIFT มีจุดประสงค์เพื่อลดขั้นตอนการทำงานที่ก่อให้เกิดการประมวลผลสูง ยกตัวอย่างเช่น ขั้นตอนการหาพิกเซลตัวแทนหรือ keypoint จะใช้หน้ากากเป็นตัวระบุตำแหน่งพิกเซลตัวแทนแทนวิธีการคำนวณดั้งเดิม ส่งผลให้ตำแหน่งและจำนวนจุดของพิกเซลตัวแทนคงที่เสมอ และสำหรับขั้นตอนการสร้างตัวแทนอธิบายคุณลักษณะของสัญลักษณ์บนหน้ายาจะใช้หน้ากากที่มีขนาด $64 * 64$ พิกเซลทาบบนพิกเซลตัวแทนเพื่อหาค่าทิศทางและขนาดของภาพขอบวัตถุ ดังนั้นจะได้เวกเตอร์ผลลัพธ์ที่มีสมาชิก 3712 ค่าเสมอ นอกจากนี้มีการลดขั้นตอนบางขั้นตอนจาก SIFT ดั้งเดิมซึ่งแบ่งได้เป็น 5 ขั้นตอนเหลือ 4 ขั้นตอนเนื่องจากตัดขั้นตอนการเกลี่ยค่าน้ำหนักของพิกเซลขอบวัตถุทิ้งเนื่องจากสังเกตเห็นว่ามีขั้นตอนการเกลี่ยหรือจางพิกเซลรบกวนอยู่แล้ว

งานวิจัย (E.Annasaro et al, 2014) นำเสนอระบบจำแนกเม็ดยาด้วยภาพที่ใช้รูปทรงสี่เหลี่ยมและตัวอักษรบนหน้ายา เทคนิคการวิเคราะห์สัญลักษณ์บนหน้ายาใช้เทคนิค Alphabet distance sets ซึ่งหาเซตของระยะห่างระหว่างพิกเซลจุดกึ่งกลางตัวอักษรไปยังพิกเซลขอบตัวอักษรด้วยวิธีของยูคิลิดีเยน (Gonzalez & Woods, 2008) ซึ่งผลลัพธ์จะอยู่ในรูปแบบเวกเตอร์ที่ประกอบไปด้วยเซตย่อยของแต่ละตัวอักษร จุดเด่นของเวกเตอร์คือทำให้ตัวแทนคุณลักษณะไม่ขึ้นกับองศาการหมุนของเม็ดยาในภาพ

นอกจากนี้ยังมีงานวิจัยจำแนกเม็ดยาด้วยภาพที่ไม่ใช้ประโยชน์ของสัญลักษณ์บนหน้ายา เช่น งานวิจัย (Cunha, Adão, & Trigueiros, 2014) งานวิจัย (Andreas Hartl, 2010) งานวิจัย (Hartl, Arth, & Schmalstieg, 2011) งานวิจัย (Hartl et al., 2011) และงานวิจัย (Kim & Chun, 2011) ซึ่งพบว่าคุณลักษณะภายนอกเม็ดยาของงานวิจัยเหล่านี้นิยมใช้คือ เทคนิคการจำแนกรูปทรงและเทคนิคการจำแนกสีเม็ดยา

จากงานวิจัยการจำแนกเม็ดยาด้วยภาพที่ทบทวนข้างต้นนั้นสามารถเขียนสรุปเป็นตารางเปรียบเทียบคุณลักษณะที่เกี่ยวข้องได้ดังนี้ เทคนิคการการจำแนกรูปทรง เทคนิคการการจำแนกขนาด เทคนิคการการจำแนกสี เทคนิคการการจำแนกสัญลักษณ์บนหน้ายา และผลลัพธ์การวิเคราะห์สัญลักษณ์บนหน้ายาดังตารางที่ 2.3

ตารางที่ 2.3 เปรียบเทียบคุณลักษณะงานวิจัยการจำแนกเม็ดยาด้วยภาพ

งานจำแนกเม็ดยา	เทคนิคการจำแนกรูปทรง	เทคนิคการจำแนกขนาด	เทคนิคการจำแนกสี	เทคนิคการจำแนกสัญลักษณ์บนหน้ายา	ผลลัพธ์การวิเคราะห์สัญลักษณ์บนหน้ายา
Yu et al., 2014	Shape signature	-	HSV	TSDS	vector
Chen et al., 2015	Shape signature	-	HSV	TSDS	vector
Caban et al., 2012	Shape signature	-	HSV	Shape distribution (extended version)	vector
Lee et al., 2012	HU's seven moments invariant	-	RGB	SIFT	vector
E.Annasaro et al., 2014	Shape signature	-	RGB	Alphabet distance sets	vector
Cunha et al., 2014	HU's seven moments invariant	Bounding box	HSV	-	-
Andreas Hartl, 2010	Modified pairwise geometric histogram	Convex hull	RGB	-	-
(Hartl et al., 2011)	Modified pairwise geometric histogram	Pill projection	RGB	-	-
(Kim & Chun, 2011)	Shape signature	-	RGB และ HSV	-	-

จากตารางที่ 2.3 สามารถจำแนกเทคนิคการจำแนกสัญลักษณ์บนหน้ายาได้เป็น 2 กลุ่ม ดังนี้

เทคนิคจำพวก DS ซึ่งค่าตัวแทนคุณลักษณะคำนวณจากค่าระยะห่างระหว่าง 2 พิกเซลที่มีความสัมพันธ์กันหลายๆคู่ ตัวอย่างเทคนิคประเภทนี้ เช่น เทคนิค Two-step sampling distance sets เทคนิค Shape distribution และเทคนิค Alphabet distance sets

แต่สำหรับเทคนิคจำพวก Keypoint ค่าตัวแทนคุณลักษณะได้จากการคำนวณทางคณิตศาสตร์เพื่อดึงค่าคุณลักษณะต่างๆของพิกเซลตัวแทน (keypoint) ออกมา เช่น ค่าองศาการหมุน (Orientation) ค่าขนาด (Scale) เป็นต้น ตัวอย่างของเทคนิคประเภทนี้คือ SIFT

ความแม่นยำของเทคนิคจากทั้งสองกลุ่มข้างต้นจะขึ้นอยู่กับฟิสิกส์ตัวแทน ซึ่งสำหรับงานเมดิคาเทคนิคประเภท DS จะทำงานได้ดีกว่าเทคนิคประเภท keypoint เนื่องจากลักษณะเด่นของฟิสิกส์สัญลักษณ์บนหน้ายาคือ ชับซ้อนและไม่มีรูปแบบตายตัว ดังนั้นการใช้ฟิสิกส์กลุ่มเล็กๆเพื่อเป็นตัวแทนตัวอักษรหน้ายาจึงมีความเสี่ยงสูงที่จะไม่สะท้อนลักษณะสัญลักษณ์บนหน้ายา และยิ่งแย่ไปกว่าหากได้ฟิสิกส์ครบถ้วนเป็นฟิสิกส์ตัวแทน ในขณะที่เทคนิคประเภท DS ที่เก็บข้อมูลลักษณะหน้ายาในเชิงค่าความสัมพันธ์ของกลุ่มฟิสิกส์ตัวแทนจึงได้เปรียบกว่าแม้ได้ฟิสิกส์ตัวแทนส่วนหนึ่งเป็นฟิสิกส์ครบถ้วน อย่างไรก็ตามแต่ละเทคนิคจำเป็นต้องสร้างฐานข้อมูลภาพเมดิคาจำเพาะเพื่อเก็บค่าเวกเตอร์ของตนทำให้ไม่สามารถใช้ฐานข้อมูลยาร่วมกันได้ นอกจากนี้ทุกเทคนิคการจำแนกสัญลักษณ์บนหน้ายาข้างต้นให้ผลลัพธ์เป็นเวกเตอร์เสมอดังนั้นค่าตัวแทนคุณลักษณะที่ได้จึงไม่ขึ้นกับทิศทางการหมุนของเมดิคา ขนาดเมดิคา และตำแหน่งของเมดิคาบนภาพ แต่ส่งผลให้ไม่สามารถทำงานร่วมกับฐานข้อมูลยาที่มีอยู่ได้เนื่องจากฐานข้อมูลยาในปัจจุบันเป็นฐานข้อมูลชนิดข้อความทั้งหมด ดังนั้นงานวิจัยนี้จึงนำเสนอวิธีการวิเคราะห์สัญลักษณ์บนหน้ายาที่ให้ผลลัพธ์อยู่ในรูปแบบข้อความตัวอักษรอย่างอัตโนมัติเพื่อนำไปประยุกต์ใช้กับฐานข้อมูลยาที่มีในปัจจุบัน

งานวิจัยที่เกี่ยวข้องกับการวิเคราะห์ตัวอักษร

สำหรับงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์ตัวอักษรมีรายละเอียดดังต่อไปนี้

งานวิจัย (T Kasar et al., 2007) นำเสนอเทคนิคการจำแนกตัวอักษรบนภาพถ่ายเอกสารโดยภาพขาวดำผลลัพธ์ตัวอักษรเป็นสีดำในขณะที่ส่วนประกอบอื่นในภาพเป็นสีขาว จุดเด่นคือสามารถรองรับภาพที่มีตัวอักษรหลากหลายสีบนพื้นหลังหลากหลายสีได้ รวมถึงรองรับตัวอักษรได้หลากหลายขนาด

เริ่มจากหาภาพ Edge map E จากการรวมภาพ Edge map ของแต่ละช่องสี RGB จากการทำ Canny edge detection (Gonzalez & Woods, 2008) ด้วยคำสั่ง OR ดังสมการด้านล่าง

$$E = ER \vee EG \vee EB$$

โดย ER และ EG และ EB คือ ภาพ Edge detection ของแต่ละช่องสี RGB

ขั้นตอนต่อไปคือทำ Labeling เพื่อหาวัตถุทั้งหมดที่ปรากฏในภาพโดยอิงทฤษฎีเพื่อนบ้านข้างเคียง 8 ทิศทาง หลังจากนั้นให้สร้าง Bounding box ครอบแต่ละวัตถุในภาพและจะเรียกแต่ละ Bounding box ว่าเป็น Edge-box หรือ EB ดังภาพที่ 2.10

เงื่อนไขการพิจารณา EB ของตัวอักษร มี 3 ข้อดังนี้

- a) สัดส่วนของ Bounding box ต้องไม่เกิน 0.1 ถึง 10 ถึงจะพิจารณาว่าเป็นตัวอักษร
 b) ขนาดของ EB ต้องมากกว่า 15 พิกเซลแต่ต้องน้อยกว่า 1/5 ของขนาดของภาพ



ภาพที่ 2.10 ตัวอย่าง EB ของตัวอักษรภาษาอังกฤษและตัวเลข

- c) จะเห็นได้ว่าไม่มีตัวอักษรใดที่มี EB อื่นอยู่ภายในเกิน 2 EB ดังนั้นหาก EB ใดมี EB อื่นอยู่ภายในเกิน 2 จะถูกพิจารณาว่าไม่เป็นตัวอักษร ในทางตรงกันข้ามจะถูกพิจารณาว่าเป็นตัวอักษร ดังเงื่อนไขด้านล่าง

```

if ( $N_{int} < 3$ )
  {Reject  $EB_{int}$ , Accept  $EB_{out}$ }
else
  {Reject  $EB_{out}$ , Accept  $EB_{int}$ }

```

โดย EB_{int} คือ EB ที่อยู่ภายใน EB อื่น และ EB_{out} คือ EB ที่ครอบ EB อื่นอยู่

ดังนั้น EB ที่ผ่านเงื่อนไขทั้งสามจะถูกพิจารณาว่าเป็น EB ของตัวอักษร หรือ Text-like elements

งานวิจัย (Gastal & Oliveira, 2011) เสนอเทคนิค Edge-aware smoothing ที่สามารถกำจัดพิกเซลรบกวนพร้อมคงความคมชัดของขอบวัตถุในภาพได้เทียบเท่ากับเทคนิค Edge-preserving smoothing filter ยอดนิยมนั้น เช่น Anisotropic diffusion (Perona & Malik, 1990) และเทคนิค Bilateral filter (Tomasi & Manduchi, 1998) แต่ให้ประสิทธิภาพดีกว่าเนื่องจากเป็นการประยุกต์ใช้ Space-invariant filter บน High dimensional space ดังนั้นจึงใช้ทรัพยากรและมีภาระในการประมวลผลน้อยกว่าเมื่อเทียบกับการทำ High dimensional filter บนโดเมนภาพปกติ (Spatial domain) ของเทคนิค Edge-preserving smoothing filter ข้างต้นส่งผลให้เทคนิคของ Gastal ประมวลผลได้เร็วกว่าจึงสามารถนำไปใช้กับงานประเภท Real time ได้ ดังนั้นจึงเป็นเทคนิค

หนึ่งที่เหมาะสมสำหรับใช้ในขั้นตอนการเตรียมภาพของงานวิจัยที่เกี่ยวกับการประมวลผลภาพ อย่างไรก็ตามก็ ดีเวลาที่ใช้ในการประมวลผลไม่ขึ้นอยู่กับค่าพารามิเตอร์แต่ขึ้นต่อขนาดภาพในลักษณะความสัมพันธ์เชิงเส้น นอกจากนี้ยังสามารถประยุกต์ใช้เทคนิค Edge-aware smoothing สำหรับขั้นตอนการเพิ่มความชัดของภาพ (Detail enhancement) ได้อีกด้วย

พารามิเตอร์ที่สำคัญสำหรับเทคนิค Edge-aware smoothing และเทคนิค Detail enhancement มี 2 ค่าคือค่าขนาดของพิกเซลเพื่อนบ้านข้างเคียงที่จะมาคำนวณเพื่อให้ได้ค่าพิกเซลผลลัพธ์ของภาพหลังการปรับปรุงซึ่งแทนด้วย σ_s ซึ่งมีค่าที่เป็นไปได้ตั้งแต่ 0 ถึง 200 และค่าพารามิเตอร์ที่แทนช่วงของค่าความต่างของสีของพิกเซลเพื่อนบ้านข้างเคียงกับสีพิกเซลศูนย์กลางที่อนุญาตเพื่อนำมาคำนวณค่าสีพิกเซลของภาพผลลัพธ์การปรับปรุงซึ่งแทนด้วย σ_r ซึ่งมีค่าที่เป็นไปได้ตั้งแต่ 0 ถึง 1

งานวิจัย (Mancas-Thillou & Mancas, 2007) ได้นำเสนอเทคนิคการจำแนกตัวอักษรพิมพ์ลึกลับวัตถุโลหะภาพจากกล้องดิจิทัลซึ่งหากพิจารณาแล้วคุณภาพของภาพที่ใช้ในงานวิจัยคล้ายกับภาพเม็ดยามากโดยเฉพาะภาพเม็ดยาที่มีตัวอักษรแบบตัวลึกลับเนื่องจากเม็ดยามักมีริ้วมันง่ายต่อการสะท้อนแสงคล้ายผิวโลหะ

ขั้นตอนการสร้างภาพขาวดำภาพที่เริ่มจากการแก้ปัญหาความไม่ชัดของภาพและพิกเซลที่เกิดจากการสะท้อนแสงด้วยเทคนิค Contrast Enhancement (Mancas, Mancas-Thillou, Gosselin, & Macq, 2006) ซึ่งมีแนวคิดมาจากพฤติกรรมการรับรู้ภาพของเรตินาของมนุษย์โดย H_{gangON} และ $H_{gangOFF}$ แทนเซลล์ ON และเซลล์ OFF ของเซลล์ประสาท Ganglion ส่วน H_{amac} แทนเซลล์ Retina amacrine จนได้ภาพ ($I_{enhanced}$) ซึ่งภาพผลลัพธ์จะทนต่อพิกเซลรบกวนมากกว่าภาพที่ได้จากเทคนิค Enhancement ทั่วไป หลังจากนั้นเพื่อกำจัดรายละเอียดพื้นผิวของวัตถุโลหะจึงทำ Median filter จนได้เป็นภาพ $I_{enhanced}^m$

การสร้างภาพขาวดำเริ่มจากการลดระดับสีในภาพจาก 24-บิต RGB เป็น 12-บิต RGB เพื่อให้เหมาะสมกับการรับรู้ของมนุษย์ หลังจากนั้นใช้เทคนิค K-means clustering เพื่อจำแนกพิกเซลในภาพเป็นสามกลุ่มคือ กลุ่มพื้นหลัง กลุ่มตัวอักษร และกลุ่มพิกเซลรบกวน การจำแนกกลุ่มพิกเซลพื้นหลังทำโดยนับจำนวนความถี่ที่พบมากที่สุดของแต่ละกลุ่มพิกเซล ณ บริเวณขอบภาพ สำหรับการจำแนกกลุ่มพิกเซลตัวอักษรและกลุ่มพิกเซลรบกวนทำโดยพิจารณาค่า M ของแต่ละกลุ่มซึ่งคำนวณจากสมการด้านล่าง โดย M จะสะท้อนค่าเบี่ยงเบนของขนาดของพื้นที่ของแต่ละ

Connected component ภายในกลุ่มพิกเซลจากขนาดพื้นที่เฉลี่ย ดังนั้นกลุ่มพิกเซลที่มีค่า M น้อยที่สุดคือกลุ่มพิกเซลตัวอักษร

$$M = \sum_i^N \left| area_i - \left(\frac{1}{N} * \sum_i^N area_i \right) \right|$$

โดย N แทนจำนวน Connected component ของแต่ละกลุ่มพิกเซล

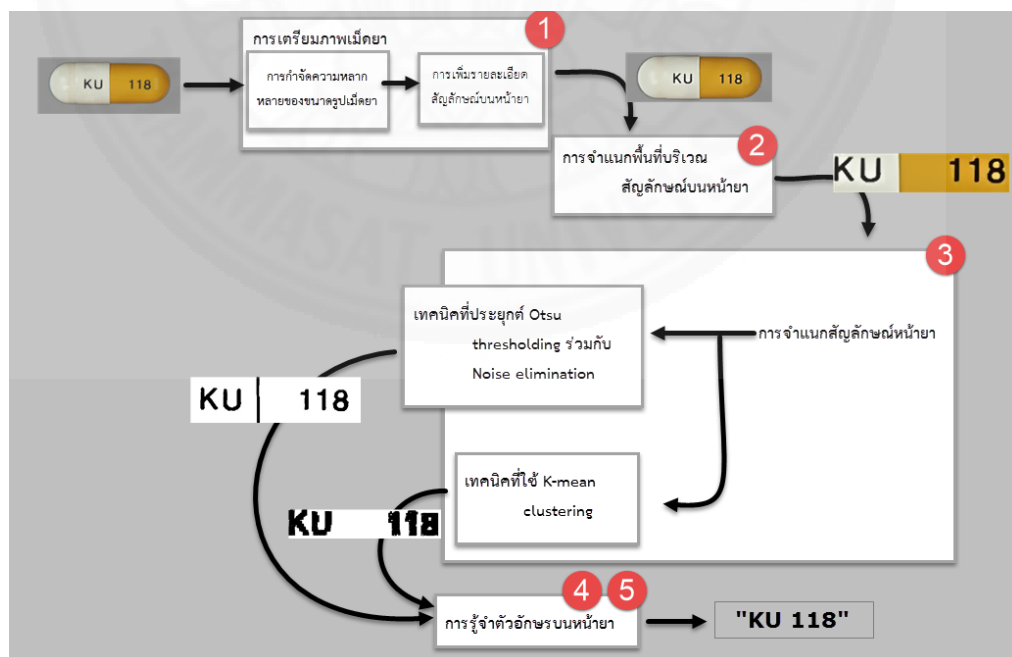
และ $area_i$ แทนขนาดพื้นที่ของ Connected component ที่ i

งานวิจัย (Wang, Kamata, & Ma, 2013) นำเสนอวิธีการกำจัดพิกเซลที่เกิดจากการสะท้อนแสงของภาพเม็ดยาในซองพลาสติกใสด้วยการสร้างอุปกรณ์รับภาพแบบใหม่ที่ประกอบไปด้วยอุปกรณ์รับภาพ 3 ตำแหน่งเพื่อใช้เป็นเทคนิคหนึ่งในขั้นตอนการเตรียมภาพเม็ดยาก่อนนำไปวิเคราะห์ด้วยเทคนิคการหาค่าตัวแทนคุณลักษณะภายนอกเม็ดยาจากภาพต่อไป

บทที่ 3 วิธีการวิจัย

จุดประสงค์หลักของงานวิจัยนี้คือการสร้างระบบวิเคราะห์สัญลักษณ์บนหน้ายาที่ให้ผลลัพธ์อยู่ในรูปแบบข้อความตัวอักษรได้อย่างอัตโนมัติที่ทนต่อภาพเม็ดยาหลากหลายขนาด การกระทบของแสงและเงาบนหน้ายา เม็ดยาที่มีสีมากกว่าหนึ่งสี รวมถึงความหลากหลายของรูปร่าง (ทรงกลม ทรงรี และแคปซูล) สี (ขาว เหลือง ส้ม ฟ้า) รูปแบบตัวอักษรภาษาอังกฤษและตัวเลขอารบิก เป็นต้น

โดยกระบวนการที่ใช้ในงานวิจัยนี้แบ่งออกเป็น 5 ขั้นตอนหลักดังภาพที่ 3.1 ดังนี้ เริ่มจากขั้นตอนการเตรียมภาพทำหน้าที่ลดความหลากหลายของขนาดภาพด้วยการย่อขยายภาพให้มีขนาดเท่าๆกันและปรับปรุงคุณภาพภาพด้วยการเพิ่มรายละเอียดตัวอักษรหน้ายา ขั้นตอนถัดมาคือหาภาพเฉพาะพื้นที่สัญลักษณ์บนหน้ายาด้วยเทคนิคที่ประยุกต์จากงาน (T Kasar et al., 2007) และนำมาสร้างภาพขาวดำด้วยเทคนิคที่ประยุกต์ Otsu thresholding ร่วมกับ Noise elimination หรือเทคนิคที่ใช้ K-means clustering ต่อมาขั้นตอนที่สี่แปลงภาพตัวอักษรหน้ายาเป็นข้อความตัวอักษรด้วยเทคนิค OCR ซึ่งต้องทำงานร่วมกับขั้นตอนที่ห้าคือกระบวนการสอนโปรแกรม OCR



ภาพที่ 3.1 แผนภาพโดยรวมของกระบวนการวิจัย

3.1. การเตรียมภาพเม็ดยา

จุดประสงค์หลักของขั้นตอนนี้คือการเตรียมภาพเพื่อกำจัดปัญหาที่เกิดขึ้นจากสภาพแวดล้อมของภาพที่หลากหลาย เช่น ขนาดภาพ แสงและเงา พิกเซลรบกวน ความคมชัดของภาพเนื่องจากอุปกรณ์รับภาพ เป็นต้น ขั้นตอนนี้ประกอบด้วยกระบวนการ 2 ขั้นตอนคือ การกำจัดความหลากหลายของขนาดรูปเม็ดยา การเพิ่มรายละเอียดสัญลักษณ์บนหน้ายา ซึ่งมีรายละเอียดดังต่อไปนี้

3.1.1 การกำจัดความหลากหลายของขนาดรูปเม็ดยา



ภาพที่ 3.2 ภาพเม็ดยาด้านฉบับ

เพื่อขจัดปัญหาภาพที่รับมามีขนาดแตกต่างกันดังภาพที่ 3.2 โดยการปรับขนาดรูปเริ่มจากการระบุตำแหน่งเม็ดยาในภาพด้วยเทคนิค K-means clustering โดยกำหนดค่า K เท่ากับ 3 เพื่อจำแนกเม็ดยาออกเป็น 3 กลุ่มเนื่องจากจำนวนสีของเม็ดยามากสุดที่เป็นไปได้คือ 2 สี การตัดสินใจว่ากลุ่มใดเป็นกลุ่มพิกเซลพื้นหลังทำโดยหากกลุ่มสีพิกเซลที่พบมากที่สุดบริเวณขอบภาพ เมื่อได้กลุ่มสีพื้นหลังแล้วจะถือว่ากลุ่มสีอีกสองกลุ่มที่เหลือเป็นกลุ่มสีพิกเซลเม็ดยา หลังจากนั้นตัดภาพให้เหลือเฉพาะพื้นที่เม็ดยาโดยใช้สีที่เหลือน้อยที่สุดที่สามารถครอบคลุมกลุ่มสีพิกเซลเม็ดยาและปรับขนาดภาพให้มีความสูงเท่ากับขนาดความสูงเฉลี่ยของเม็ดยาในภาพต้นฉบับของชุดข้อมูลภาพทดลอง

ซึ่ง ณ ที่นี้คือ 300 พิกเซลเพื่อให้ภาพที่ผ่านการปรับขนาดมีจำนวนน้อยที่สุด ส่วนความกว้างของภาพให้ปรับขนาดโดยคงสัดส่วนดั้งเดิมของเมตาดาต้าฉบับซึ่งใช้เทคนิค Bicubic interpolation เนื่องจากเมื่อเปรียบเทียบเทคนิคการยืดหดภาพอื่นๆ เช่น เทคนิค Nearest-Neighbor interpolation เทคนิค Bilinear interpolation และเทคนิค Lanczos interpolation พบว่าเทคนิค Bicubic interpolation ให้ภาพผลลัพธ์ที่มีขอบวัตถุเรียบที่สุดใกล้เคียงกับเทคนิคของ Lanczos เนื่องจากมีการทำ Anti-aliasing แต่เทคนิค Lanczos interpolation มีการพิจารณาเพื่อนบ้านข้างเคียงขนาดใหญ่กว่าจึงใช้เวลาในการประมวลผลมากกว่า (Reddy & Reddy, 2013)

3.1.2 การเพิ่มรายละเอียดสัญลักษณ์บนหน้ายา

จากการทดลองเบื้องต้นกับเทคนิคประเภท Edge detection เช่น เทคนิคของ Canny เทคนิคของ Sobel รวมถึงเทคนิคของ Prewitt พบว่าภาพขอบวัตถุที่ได้ไม่สามารถให้ขอบของตัวอักษรบนหน้ายาได้ดีเท่าที่ควรเนื่องจากขอบของตัวอักษรบนหน้ายาไม่ชัดเจนโดยเฉพาะตัวอักษรแบบลิก ดังนั้นจึงต้องเพิ่มความคมชัดของตัวอักษรในภาพซึ่งเลือกใช้เทคนิค Detail enhancement ที่เสนอโดย Gastal (Gastal & Oliveira, 2011) ก่อนการหาภาพขอบวัตถุ เทคนิคของ Gastal สามารถเพิ่มความคมชัดของภาพโดยที่ไม่เพิ่มพิกเซลรบกวน ดังนั้นจึงเกิดพิกเซลรบกวนน้อยกว่าเมื่อเปรียบเทียบกับเทคนิค Histogram equalization แต่การเพิ่มความคมชัดของภาพส่งผลให้เพิ่มพิกเซลรบกวนอย่างหลีกเลี่ยงไม่ได้ดังนั้นจึงใช้เทคนิค Edge-aware smoothing ของ Gastal ร่วมด้วยเพื่อเกลี่ยพิกเซลรบกวน เนื่องจากเมื่อเปรียบเทียบกับเทคนิคการเบลอพิกเซลรบกวนพื้นฐาน เช่น เทคนิค Mean filter เทคนิค Gaussian filter และเทียบกับเทคนิค Edge-preserving smoothing filter ยอดนิยม (Perona & Malik, 1990; Tomasi & Manduchi, 1998) พบว่าเทคนิคที่เสนอโดย Gastal สามารถให้ผลลัพธ์ได้ดีเทียบเท่า Edge-preserving smoothing filter ข้างต้นแต่ใช้เวลาในการประมวลผลน้อยกว่ามาก (Gastal & Oliveira, 2011) โดยภาพผลลัพธ์จากการเพิ่มรายละเอียดสัญลักษณ์บนหน้ายาจะมีลักษณะดังภาพที่ 3.3



ภาพที่ 3.3 ภาพเม็ดยาจากภาพที่ 3.2 ที่ผ่านขั้นตอนการเพิ่มรายละเอียดสัญลักษณ์บนหน้ายา

3.2. การจำแนกพื้นที่บริเวณสัญลักษณ์บนหน้ายา

เป็นขั้นตอนหาตำแหน่งที่แท้จริงของสัญลักษณ์บนหน้ายาเพื่อให้ได้ภาพเฉพาะส่วนที่เป็นพื้นที่สัญลักษณ์บนหน้ายา เพื่อลดปัญหาที่เกิดจากอิทธิพลของค่าสีพิกเซลที่ไม่เกี่ยวข้องสำหรับขั้นตอนการสร้างภาพขาวดำโดยเฉพาะกับเทคนิคจำพวก Global thresholding

เริ่มจากการหาขอบสัญลักษณ์บนหน้ายาโดยใช้เทคนิค Canny edge detection ซึ่งอธิบายไว้ในหัวข้อ 2.3 กับภาพเม็ดยาที่ละช่องสี RGB หลังจากนั้นนำภาพผลลัพธ์การทำ Edge detection ของแต่ละช่องสีมารวมกันเป็นภาพผลลัพธ์ดังสมการ

$$E = E_P \vee E_R \vee E_B$$

เนื่องจากขอบตัวอักษรที่ได้เป็นเส้นบางดังภาพที่ 3.4 จึงใช้เทคนิค Morphology dilation เพื่อเพิ่มขนาดเส้นขอบตัวอักษร หลังจากเพิ่มขนาดขอบวัตถุพบว่ามีรูภายในขอบวัตถุซึ่ง

ก่อให้เกิดเส้นรอบวงของรูที่ไม่พึงประสงค์ส่งผลให้ผลลัพธ์การหา Character contour ในขั้นตอนต่อไปแม่นยำน้อยลง ดังนั้นเพื่อกลบรูเหล่านี้เลือกใช้เทคนิค Morphology closing ที่มี Kernel ทรงสี่เหลี่ยมจัตุรัสขนาดเล็ที่สุดที่เป็นไปได้ซึ่งก็คือขนาด 3×3 พิกเซลเพื่อก่อให้เกิดการประมวลผลน้อยที่สุด



ภาพที่ 3.4 ภาพผลลัพธ์ (a) สมการ E (b) เทคนิค Dilation (c) เทคนิค Closing

หลังจากนั้นหา Bounding box ที่ครอบคลุมเส้นรอบรูปของแต่ละพื้นที่พิกเซล เพื่อนำมาพิจารณาว่าเป็น Character contour หรือไม่ซึ่งพิจารณาตามเงื่อนไข 5 ข้อดังต่อไปนี้

1. Bounding box ที่มีความกว้างไม่เกิน 6.67 เปอร์เซ็นต์ของความสูงของภาพและความสูงไม่เกิน 6.67 เปอร์เซ็นต์ของความสูงของภาพ จะพิจารณาให้เป็นพิกเซลรบกวน
2. Bounding box ที่มีความกว้างใหญ่เกินครึ่งหนึ่งของความกว้างของภาพหรือมีความสูงมากเกินครึ่งหนึ่งของความสูงของภาพจะพิจารณาให้เป็นขอบเม็ดยา
3. สัดส่วนของ Bounding box สามารถนำมาเป็นส่วนหนึ่งของการพิจารณาหากมีสัดส่วนความกว้างเกิน 3 เท่าของความสูงของ Bounding box จะถูกพิจารณาว่าไม่เป็น Character contour เนื่องจากไม่พบตัวอักษรใดที่มีความกว้างเกินสามเท่าของความสูง
4. ขณะเดียวกัน Bounding box ใดที่อยู่ภายใน Bounding box อื่นจะพิจารณาว่าไม่เป็น Character contour เช่นกัน
5. จากการทดลองกับภาพจาก Training set พบว่า Character contour ที่ใหญ่ที่สุดที่ผ่านทุกเงื่อนไขข้างต้นมักเป็น Contour ของ Character เสมอ ดังนั้นเพื่อกำจัดพิกเซลรบกวนที่แอบแฝงมาทำโดยหาก Bounding box ของ Character contour ใดมีพื้นที่น้อยกว่า 1 ใน

4 ของ Bounding box ของ Character contour ที่ใหญ่ที่สุดจะพิจารณาว่าเป็นพิกเซลรวมกันทันที

เงื่อนไข 5 ข้อข้างต้นนี้ได้แนวคิดมาจากกฎการหาพื้นที่ของแต่ละตัวอักษรบนภาพถ่ายเอกสารในงานวิจัย (Kasar et al., 2007) โดยเฉพาะเงื่อนไขข้อ 1 กับข้อ 4 แต่จากการทดลองเบื้องต้นพบว่ามักพบรูภายในตัวอักษรเสมอแม้ผ่านขั้นตอนการทำงาน Morphology closing จึงปรับปรุงเงื่อนไขข้อที่ 3 ของ Kasar กลายเป็นเงื่อนไขข้อที่ 4 ที่นำเสนอ แต่เงื่อนไขข้อที่ 2 ที่นำเสนอ มีจุดประสงค์เพื่อกำจัดขอบเม็ดยาทิ้งเพื่อให้เหลือเฉพาะกลุ่มพิกเซลตัวอักษรเท่านั้น นอกจากนี้เงื่อนไขข้อ 3 และ 5 ที่นำเสนอ นั้นมีจุดประสงค์เพื่อกำจัดกลุ่มพิกเซลรวมกันที่มักพบในงานเม็ดยาจากการทดลองเบื้องต้น

เมื่อได้ Character contour ทั้งหมดถัดไปคือการหาพื้นที่จริงของสัญลักษณ์บนหน้ายา ด้วย Bounding box ที่เล็กที่สุดที่สามารถครอบคลุมทุกๆ Character contour ในภาพดังภาพที่ 3.5 และใช้ Bounding box ดังกล่าวตัดภาพสีของเม็ดยาที่ได้จากขั้นตอน 3.1.2 เพื่อให้ได้ภาพเฉพาะพื้นที่ของสัญลักษณ์บนหน้ายา



ภาพที่ 3.5 ตัวอย่างพื้นที่จริงของสัญลักษณ์บนหน้ายาที่จำแนกได้ (กรอบสีขาว)

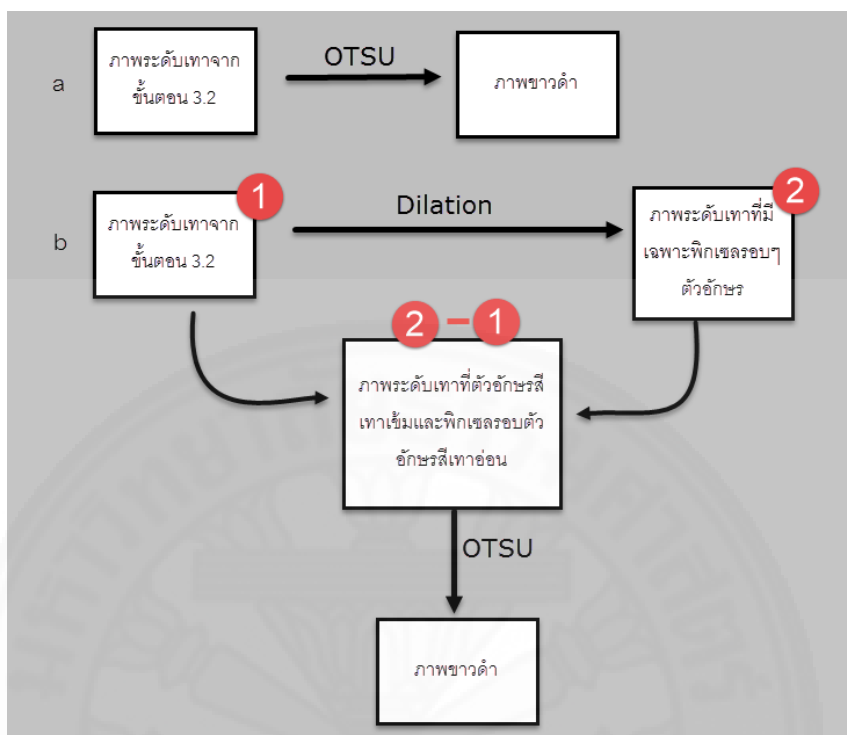
3.3. การจำแนกสัญลักษณ์หน้ายา

เป็นขั้นตอนเพื่อจำแนกสัญลักษณ์บนหน้ายากับภาพสีที่มีเฉพาะพื้นที่ตัวอักษรบนหน้ายาที่ได้จากขั้นตอน 3.2 ออกจากส่วนอื่นของภาพซึ่งจะถือว่าเป็นพื้นหลังดังนั้นภาพผลลัพธ์ที่ได้พื้นหลังจะกำจัดทิ้งเป็นสีขาวในขณะที่ตัวอักษรบนหน้ายาเป็นสีดำ สำหรับขั้นตอนนี้จะนำเสนอ 2 เทคนิคหลักคือ

เทคนิคที่ประยุกต์ Otsu thresholding ร่วมกับ Noise elimination เทคนิค Otsu thresholding เป็นเทคนิค Histogram thresholding ประเภท Minimum variance within segments ที่นิยมซึ่งเหมาะกับงานเมดิคยาเมื่อเทียบกับเทคนิคประเภท Local based method เนื่องจากการหาค่า threshold เพื่อจำแนกตัวอักษรจะทำบนภาพเฉพาะพื้นที่ตัวอักษรบนหน้ายา ดังนั้นหากทำกับทีละส่วนย่อยของภาพจะก่อให้เกิดพิกเซลรบกวนจำนวนมากโดยเฉพาะเมดิคยาที่มีตัวอักษรหน้ายาแบบเล็ก

เทคนิคที่สองคือเทคนิคที่ใช้ K-means clustering ซึ่งถูกนำไปใช้เพื่อจำแนกตัวอักษรแบบเล็กที่สลักลงบนโลหะ (Mancas-Thillou & Mancas, 2007) ซึ่งมีลักษณะมันวาวคล้ายการจำแนกตัวอักษรแบบเล็กผิวเมดิคยา หลักการของเทคนิค K-means clustering คือจะจัดกลุ่มสีพิกเซลเป็น K กลุ่มย่อยจึงเหมาะสมกับการจำแนกพิกเซลตัวอักษรบนหน้ายาเนื่องจากตัวอักษรบนหน้ายามักมีสีเดียวกันแม้อยู่บนเมดิคยาหลากสี

3.3.1 เทคนิคที่ประยุกต์ Otsu thresholding ร่วมกับ Noise elimination



ภาพที่ 3.6 ภาพรวมของ (a) เทคนิคของ Otsu (b) เทคนิคของ Otsu ร่วมกับ Noise elimination

เมื่อรับภาพสีที่มีเฉพาะพื้นที่บริเวณหน้ายาจากขั้นตอน 3.2 หากเป็นเทคนิค Otsu thresholding จะนำมาแปลงเป็นภาพระดับเทา (Grayscale image) เพื่อหาค่า threshold ที่เหมาะสมสำหรับการแบ่งกลุ่มสีพิกเซล 2 กลุ่ม หลังจากนั้นจะสร้างภาพขาวดำ (Binary image) จากค่า threshold ซึ่งได้ภาพผลลัพธ์ดังภาพที่ 3.8 และมีขั้นตอนการทำงานดังภาพที่ 3.6 (a)

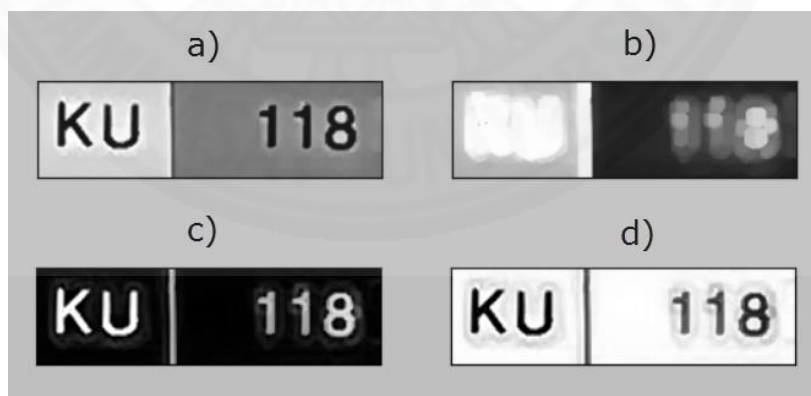
จากการทดลองเบื้องต้นพบว่าเม็ดยาที่มีสัญลักษณ์บนหน้ายาแบบตัวพิมพ์ที่มีสีเม็ดยามากกว่าหนึ่งสีกล่าวคือมีกลุ่มสีของพิกเซลมากกว่า 2 กลุ่มที่แตกต่างกันซึ่งเกินขอบเขตความสามารถของเทคนิค Otsu thresholding จึงทำให้ภาพขาวดำที่ได้มีเพียงสีของเม็ดยาสีเดียวเท่านั้นที่สามารถจำแนกตัวอักษรได้ดังภาพที่ 3.8 ดังนั้นจำเป็นต้องกำจัดสีระดับเทาของพิกเซลรอบๆตัวอักษรทิ้งก่อนสร้างภาพขาวดำด้วยค่า threshold จากเทคนิคของ Otsu โดยใช้เทคนิค Noise elimination

เทคนิค Noise elimination จะหาภาพระดับเทาที่มีเฉพาะพิกเซลรอบตัวอักษรบนหน้ายาด้วยเทคนิค Morphology dilation ที่ใช้ Kernel ทรงกลมขนาดเท่ากับสองเปอร์เซ็นต์ของขนาดความสูงของภาพเม็ดยา ซึ่งทำเทคนิค dilation กับภาพเทาของภาพสีจาก

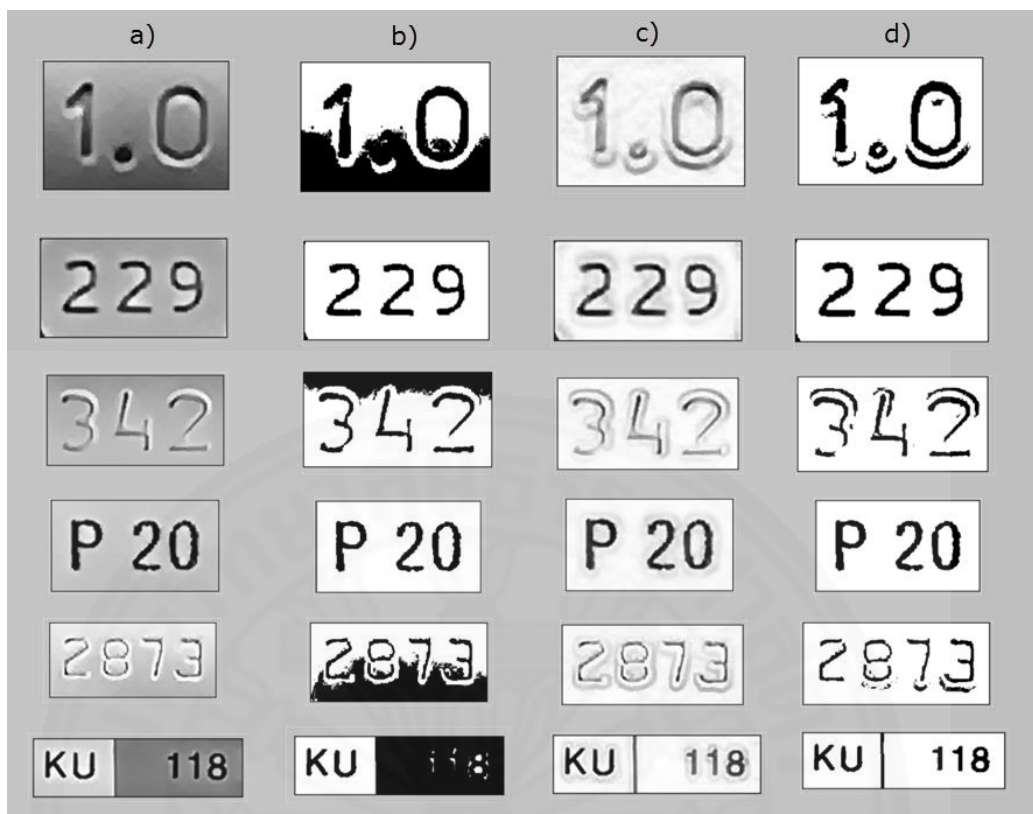
ขั้นตอน 3.2 และนำภาพระดับเทาที่มีเฉพาะพิกเซลรอบตัวอักษรบนหน้าyalobด้วยภาพเทาของภาพสี จากขั้นตอน 3.2 เพื่อให้ได้ภาพระดับเทาที่จำแนกตัวอักษรบนหน้าyalobดังภาพที่ 3.7 (c) หลังจากนั้น กลับค่าสีพิกเซลให้สีตัวอักษรมีสีเทาเข้มโดยมีสีพิกเซลรอบๆตัวอักษรสีเทาอ่อนดังภาพที่ 3.7 (d) แล้ว จึงนำไปสร้างเป็นภาพขาวดำผลลัพธ์ดังภาพที่ 3.8 ด้วยค่า threshold จากเทคนิค Otsu ซึ่งมีขั้นตอน การทำงานดังภาพที่ 3.6 (b)

สำหรับเฟรมเวิร์ค Opencv ค่าสีของพิกเซลในภาพระดับเทาเป็นไปได้ ตั้งแต่ 0 ซึ่งแทนสีเทาเข้มถึง 255 ซึ่งแทนสีเทาอ่อน ดังนั้นการลบพิกเซลระดับเทาด้วยพิกเซลระดับ เทามี 4 กรณีที่เป็นไปได้คือ พิกเซลสีเทาเข้มลบด้วยพิกเซลสีเทาอ่อน (0-255) พิกเซลสีเทาเข้มลบ ด้วยพิกเซลสีเทาเข้ม (0-0) พิกเซลสีเทาอ่อนลบด้วยพิกเซลสีเทาอ่อน (255-255) พิกเซลสีอ่อนเข้มลบ ด้วยพิกเซลสีเทาเข้ม (255-0) ซึ่งจะได้สีพิกเซลในภาพผลลัพธ์ดังนี้ สีเทาเข้ม สีเทาเข้ม สีเทาเข้ม สีเทา อ่อน ตามลำดับ

นอกจากนี้สำหรับสัญลักษณ์บนหน้าyalobแบบตัวลึกลับที่มีเงาตกกระทบบริเวณ หน้าyalobมักเจอปัญหาในกลุ่มพิกเซลเงาปรากฏในภาพผลลัพธ์เนื่องจากสีพิกเซลบริเวณเงาใกล้เคียงกับสี พิกเซลตัวอักษรส่งผลให้เทคนิค Otsu พิจารณาพิกเซลเงาอยู่ในกลุ่มสีเดียวกันกับสีของตัวอักษร ซึ่ง ปัญหาที่เกิดขึ้นนี้เทคนิค Noise elimination สามารถกำจัดสีระดับเทาของพิกเซลเงาที่ก่อก่อนส่งเข้า เทคนิคของ Otsu ได้เช่นกัน



ภาพที่ 3.7 ภาพระดับเทา (a) จากขั้นตอนการจำแนกพื้นที่บริเวณสัญลักษณ์บนหน้าyalob (b) เฉพาะ พิกเซลรอบตัวอักษรบนหน้าyalob (c) เฉพาะตัวอักษรหน้าyalob (d) จากการกลับค่าสีของภาพเฉพาะ ตัวอักษรหน้าyalob



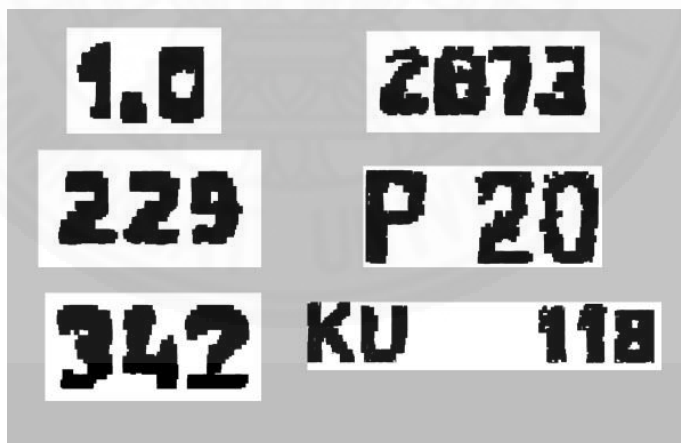
ภาพที่ 3.8 ภาพผลลัพธ์ของ (a) ขั้นตอนการจำแนกพื้นที่บริเวณสัญลักษณ์บนหน้ายา (b) เทคนิคของ Otsu (c) เทคนิค Noise elimination (d) เทคนิคร่วมระหว่าง Otsu และ Noise elimination

3.3.2 เทคนิคที่ใช้ K-means clustering

สาเหตุที่เลือกเทคนิค K-means clustering เป็นเทคนิคหนึ่งที่น่าเสนอสำหรับการจำแนกสัญลักษณ์บนหน้ายาเนื่องจาก สามารถทำงานได้ดีกับภาพตัวอักษรสลักลึกลงบนวัตถุโลหะมันวาวจากงานวิจัย (Mancas-Thillou & Mancas, 2007) ซึ่งสอดคล้องกับลักษณะงานเม็ดยาที่มีสัญลักษณ์บนหน้ายาแบบลึกซึ่งเป็นส่วนที่ยากที่สุดของงานนี้ เพื่อเตรียมภาพให้เหมาะสมกับเทคนิค K-means clustering จะใช้เทคนิค Contrast enhancement (Mancas et al., 2006) หรือ *I_{enhanced}* ที่มีรายละเอียดในหัวข้อ 2.3 กับภาพระดับเทา 3 ภาพที่ได้จากการแยกช่องสี RGB จากภาพสีที่ได้จากขั้นตอน 3.2 หลังจากนั้นลดระดับสีของภาพด้วยการแปลงภาพสี 24-บิต RGB ที่ผ่านการทำเทคนิค *I_{enhanced}* แล้วเป็น 12-บิต RGB เพื่อลดภาระการประมวลผลในขั้นตอนการสร้างภาพขาวดำโดยใช้เทคนิค K-means clustering โดยที่ค่า K เท่ากับ 16 เพื่อลดค่าสีของแต่ละช่องสี RGB ให้เหลือเพียงช่องสีละ 16 ค่าสีที่เป็นไปได้ (0-15) ท้ายที่สุดสร้างภาพขาวดำผลลัพธ์จากภาพสี 12-บิต RGB ที่ได้จากเทคนิค 16-mean clustering ด้วยเทคนิค K-means clustering ที่กำหนดให้

ค่า K เท่ากับ 3 เพื่อจัดกลุ่มพิกเซลของภาพสี RGB เป็น 3 กลุ่มเนื่องจากภาพเฉพาะพื้นที่บริเวณตัวอักษรบนหน้ายามักประกอบไปด้วยกลุ่มพิกเซลสามสีหลักคือ พิกเซลสีเม้ตยา พิกเซลสีตัวอักษรบนหน้ายา และสีของพิกเซลรบกวน ประกอบกับพิกเซลตัวอักษรบนหน้ายามักสีและขนาดหรือรูปแบบตัวอักษรที่เหมือนกันดังนั้นจึงใช้เทคนิค K-means clustering เพื่อจัดกลุ่มพิกเซลที่มีสีใกล้เคียงกัน เพื่อหากกลุ่มพิกเซลตัวอักษรบนหน้ายาจากกลุ่มพิกเซลที่มีขนาดของ connected component ที่ใกล้เคียงกันมากที่สุดหรือกลุ่มสีที่มีค่า M น้อยสุดนั่นเองโดยค่า M มีรายละเอียดตามหัวข้อ 2.3

การสร้างภาพขาวดำเริ่มจากแยกแยะกลุ่มสีพิกเซลสีเม้ตยาออกก่อนพิจารณาจากการนับความถี่ของกลุ่มพิกเซลที่เกิดขึ้นมากที่สุดบริเวณขอบภาพ หลังจากนั้นการตัดสินใจระหว่างกลุ่มสีพิกเซลตัวอักษรและกลุ่มสีพิกเซลรบกวนพิจารณาจากขนาดของแต่ละ connected component หรือจำนวนพิกเซลของแต่ละ connected component ซึ่งสะท้อนขนาดของตัวอักษร โดย connected component ที่ได้สร้างจากหลักการพิจารณาพิกเซลวัตถุเดียวกันแบบ 8-connectivity ดังนั้นหากเป็นกลุ่มพิกเซลตัวอักษรหน้ายาจริงทุกๆ connected component ที่เป็นสมาชิกของกลุ่มพิกเซลนั้นควรมีขนาดใกล้เคียงกันซึ่งสะท้อนผ่านค่า M หลังจากนั้นสร้างภาพขาวดำผลลัพธ์จากกลุ่มพิกเซลตัวอักษรและขยายพิกเซลตัวอักษรด้วยเทคนิค Morphology dilation ด้วย kernel สีเหลี่ยมจัตุรัสที่มีขนาดเท่ากับ 5 เพลอร์เซ็นต์ของด้านของรูปที่สั้นที่สุดระหว่างด้านสูงและด้านกว้าง และภาพขาวดำผลลัพธ์มีลักษณะดังภาพที่ 3.9



ภาพที่ 3.9 ตัวอย่างการสร้างภาพขาวดำด้วยเทคนิคที่ใช้ K-means clustering

3.4. การรู้จำตัวอักษรบนหน้ายา

จุดประสงค์ของขั้นตอนนี้คือการแปลงภาพเป็นตัวอักษรด้วยโปรแกรม Optical Character Recognition (OCR) ในงานวิจัยนี้ผู้วิจัยเลือกใช้โปรแกรม Tesseract (Smith, 2007) ซึ่งเป็น Open-source ที่นิยมที่สุดในปัจจุบัน โดยภาพที่รับเข้ามาคือภาพผลลัพธ์ที่ได้จากขั้นตอน 3.3 ซึ่งคือภาพขาวดำที่มีตัวอักษรสีดำและมีพื้นหลังสีขาว

สาเหตุที่เลือก OCR เพื่ออ่านตัวอักษรบนหน้ายาเนื่องจากรองรับตัวอักษรได้หลายรูปแบบ (Font และ Layout) และหลายภาษารวมถึงทนต่อ Perspective distortion

เนื่องจากโดยพื้นฐานแล้ว OCR ถูกพัฒนาเพื่ออ่านตัวอักษรบนเอกสารเป็นหลักซึ่งมีอุปกรณ์รับภาพที่มีประสิทธิภาพที่อยู่ในสภาพแวดล้อมควบคุม เช่น เครื่องสแกนเอกสาร รวมถึงการมีรูปแบบ (Layout) ของตัวอักษรที่ตายตัวทำให้เมื่อนำมาประยุกต์กับงานเม็ดยาซึ่งตัวอักษรมีเส้นขอบที่ไม่คมชัด โดยเฉพาะตัวอักษรตัวเล็ก รวมถึงการปรากฏของฟิกเชลรบกวนเพียงเล็กน้อยสามารถทำให้ความแม่นยำของ OCR ต่ำลงได้ดังนั้นจึงต้องมีการสอนโปรแกรม OCR ร่วมด้วย

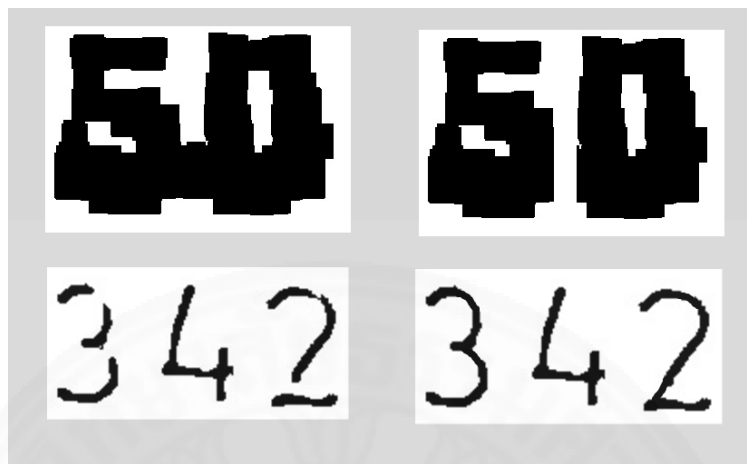
3.5. กระบวนการสอนโปรแกรม Tesseract

เพื่อสอนโปรแกรม Tesseract ให้รู้จักฟอนต์ของแต่ละตัวอักษรหน้ายาในภาพขาวดำจากขั้นตอน 3.3 จำเป็นต้องสอน 4 ฟอนต์คือ ฟอนต์ของภาพขาวดำจากเทคนิค Otsu ซึ่งมีลักษณะผอมสูง และฟอนต์ของภาพขาวดำจากเทคนิค K-means clustering ที่มีลักษณะกลมสั้น ทั้งนี้ทั้งสองฟอนต์ข้างต้นประกอบไปด้วยอีก 2 ฟอนต์ย่อยคือฟอนต์สำหรับตัวพิมพ์และตัวเล็ก และเนื่องจากตัวอักษรบนหน้ายาส่วนใหญ่เป็นตัวเลขดังนั้นจึงเน้นสอนที่ภาพตัวเลขมากกว่าตัวอักษรโดยใช้ 30 ภาพโดยเฉลี่ยสำหรับแต่ละตัวเลขของแต่ละฟอนต์ตั้งแต่เลข 0 ถึง 9 กระบวนการสอนโปรแกรม Tesseract จะเป็นการสอนให้รู้จำแต่ละภาพของแต่ละตัวอักษรที่ละฟอนต์ตัวอักษรซึ่งสามารถแบ่งได้เป็น 4 ขั้นตอนดังนี้

3.5.1 การเตรียมภาพตัวอักษร

ขั้นตอนสำหรับเตรียมภาพเพื่อใช้สอนโปรแกรม Tesseract เตรียมโดยนำภาพขาวดำที่ได้จากขั้นตอน 3.3 มาทำให้สมบูรณ์มากขึ้นดังภาพที่ 3.10 โดยภาพที่ตัวอักษรขาดจะ

เชื่อมให้ต่อกันรวมถึงตัวอักษรหลายตัวที่เชื่อมกันอยู่จะทำให้แยกจากกันทีละตัวมากขึ้นแต่ยังคงลักษณะเฉพาะของตัวอักษรดั้งเดิมอยู่



ภาพที่ 3.10 ตัวอย่างภาพตัวอักษรผลลัพธ์จาก (ซ้าย) ขั้นตอนการจำแนกสัญลักษณ์บนหน้ายา (ขวา) หลังการเตรียมภาพตัวอักษร

หลังจากนั้นสร้างไฟล์ pdf เพื่อเก็บภาพตัวอักษรทั้งหมดที่ใช้ในการสอนเพื่อสร้างไฟล์ tiff ที่มีความละเอียดอย่างน้อย 300*300 พิกเซลเพื่อใช้สำหรับการสอนด้วยคำสั่งของโปรแกรม GhostScript9 ด้านล่าง

```
Gswin64.exe -o [lang][font-name].tif -sDEVICE=tifffg4 -r300x300  
-g2550x3300 [lang][font-name].pdf
```

3.5.2 การติดป้ายข้อมูลให้ภาพตัวอักษร

ขั้นตอนการติดป้ายข้อมูลของแต่ละภาพตัวอักษรเป็นขั้นตอนทำมือทั้งหมด โดยข้อมูลที่ต้องติดคือ ข้อมูลที่บอกตำแหน่งของแต่ละภาพตัวอักษร และข้อมูลที่ระบุว่าแต่ละภาพตัวอักษรเป็นตัวอักษรอะไรซึ่งข้อมูลเหล่านี้จะเก็บในไฟล์นามสกุล box อย่างไรก็ตามก็ยังสามารถใช้โปรแกรม bbTesseract ของโปรแกรม Tesseract ด้านล่างเพื่อสร้างไฟล์ที่เก็บข้อมูลการติดป้ายตั้งต้นได้ (box file) แต่ความถูกต้องของข้อมูลมักไม่ดี ผลลัพธ์ที่ได้จากขั้นตอนนี้คือไฟล์ box ที่เก็บข้อมูลภาพตัวอักษรที่อยู่ภายในไฟล์ tiff

```
tesseract.exe [lang][font-name].tif [lang][font-name]  
batch.no chop makebox
```

3.5.3 การสอนโปรแกรม Tesseract

สำหรับการสอนโปรแกรม Tesseract มี 8 ไฟล์ที่จำเป็นดังนี้

tessdata/[lang][fontName].freq-dawg

tessdata/[lang][fontName].word-dawg

tessdata/[lang][fontName].user-words

tessdata/[lang][fontName].inttemp

tessdata/[lang][fontName].normproto

tessdata/[lang][fontName].pffmtable

tessdata/[lang][fontName].unicarset

tessdata/[lang][fontName].DangAmbigs

โดย [lang] แทน ตัวอักษรย่อของภาษาที่ต้องการสอนตามมาตรฐาน ISO 639-3 standard language code เพื่อให้ได้ไฟล์ผลลัพธ์การสอนคือไฟล์นามสกุล traineddata เพื่อให้โปรแกรม Tesseract ใช้งานโดยนำไปวางภายในโฟลเดอร์ย่อยชื่อ tessdata

เริ่มจากการสร้างไฟล์นามสกุล tr ซึ่งมีหน้าที่เก็บข้อมูล Character feature ของแต่ละภาพตัวอักษรด้วยคำสั่งโปรแกรม Tesseract ดังนี้

```
tesseract [lang][font-name].tif junk nobatch box.train
```

เพื่อจัดกลุ่ม Character shape feature จะใช้โปรแกรม mftraining และ cntraining โดยการรัน mftraining จะให้ไฟล์ที่สำคัญ 3 ไฟล์คือไฟล์นามสกุล inttemp ไฟล์นามสกุล pffmtable และไฟล์นามสกุล Microfeat ส่วนการรัน cntraining จะให้ไฟล์นามสกุล normproto เท่านั้น คำสั่งที่ใช้มีดังนี้

```
mftraining [lang][font-name]_1.tr [lang][font-name]_2.tr ...
```

```
cntraining fontfile_1.tr fontfile_2.tr ...
```

ต่อมาเป็นการสร้างไฟล์ข้อมูลชื่อ unicarset ที่ทำหน้าที่เก็บคุณลักษณะของแต่ละภาพตัวอักษร เช่น isAlpha isDigit isUpper และ isLower โดยใช้โปรแกรม unicarset_extractor ด้วยคำสั่งด้านล่าง

```
unicarset_extractor fontfile_1.box fontfile_2.box ...
```

ถัดมาเป็นขั้นตอนการสร้างไฟล์ดิกชันนารีที่โปรแกรม Tesseract รองรับไฟล์ dictionary เพียง 3 ไฟล์ดังนี้ ไฟล์ freq-dawg เป็นไฟล์ประเภท Directed acyclic word graph (DAWG) ซึ่งมีหน้าที่เก็บคำศัพท์ที่ภาษาของพอนต์ที่มักพบบ่อยๆ หนึ่งบรรทัดต่อหนึ่งคำ ต่อมาคือไฟล์ word-dawg เป็นไฟล์ DAWG เช่นกันซึ่งมีหน้าที่เก็บคำศัพท์ทั้งหมดที่เป็นไปได้ของภาษา ไฟล์ dictionary สุดท้ายชื่อ user-words อยู่ในรูปแบบไฟล์ UTF-8 ซึ่งมักเป็นไฟล์ว่างเสมอ อย่างไรก็ตามก็มีคำสั่งที่ทำหน้าที่แปลงไฟล์ UTF-8 เป็นไฟล์ประเภท DAWG เพื่ออำนวยความสะดวกดังกล่าว

Wordlist2dawg frequent_words_list freq-dawg

Wordlist2dawg words_list word-dawg

สำหรับไฟล์ DangAmbigs ที่เป็นหนึ่งในไฟล์ตั้งต้นเพื่อสร้างไฟล์ผลลัพธ์ traineddata ที่ทำหน้าที่เก็บข้อมูลตัวอักษรหรือคำที่มักเกิดความกำกวม โดยไฟล์นี้ผู้ใช้ต้องเป็นผู้ใส่ข้อมูลเท่านั้น

ก่อนการสร้างไฟล์ traineddata จะต้องมั่นใจว่าทุกๆ Language files เช่น normproto Microfeat inttemp และ pffmtable ล้วนมีชื่อเดียวกันคือ [lang][font-name] หลังจากจึงสร้างไฟล์ traineddata ด้วยคำสั่งด้านล่างหลังจากนั้นให้วางไฟล์ traineddata ไว้ที่โฟลเดอร์ย่อย tessdata ก็เป็นอันสำเร็จ

combine_tessdata lang.

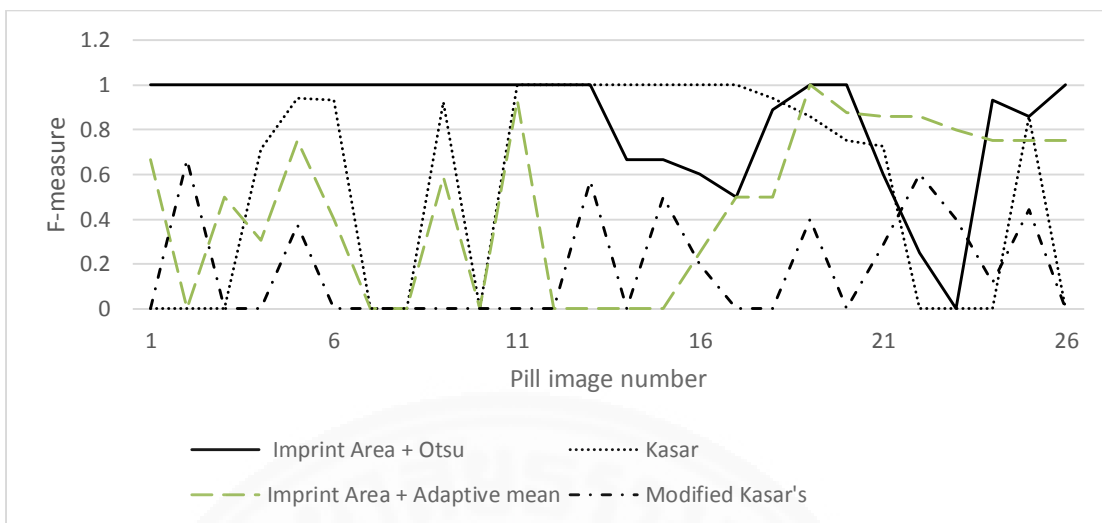
3.5.4 การเพิ่มไฟล์ตั้งค่า

เพื่อเพิ่มความแม่นยำให้กับโปรแกรม Tesseract มากขึ้นจะยกเลิกการใช้ไฟล์ดิกชันนารีที่ให้มากับโปรแกรม Tesseract ของภาษาอังกฤษเนื่องจากสัญลักษณ์บนหน้ายาส่วนใหญ่ไม่ใช่คำที่มีความหมายที่ปรากฏในภาษาอังกฤษแต่เป็นเพียงอักษรย่อที่ผลิตจากผู้ผลิตยาเท่านั้นด้วยการตั้งค่าตัวแปร load_system_dawg และ load_freq ให้มีค่าเท่ากับ F (false) ลงในไฟล์ตั้งค่าของโปรแกรม Tesseract พร้อมกับกำหนดตัวอักษรที่เป็นไปได้ของสัญลักษณ์บนหน้ายาผ่านตัวแปร tessedit_char_whitelist และกำหนดชุดของคำที่มักพบบนสัญลักษณ์บนหน้ายาผ่านไฟล์ dictionary ที่ชื่อ user-words

3.6. การทดลองเบื้องต้น

จุดประสงค์ของการทดลองเบื้องต้นคือการทดสอบประสิทธิภาพขั้นตอนการหาพื้นที่บริเวณสัญลักษณ์บนหน้ายาที่เสนอในงานวิจัยนี้ซึ่งประยุกต์จากกฎการสร้าง Text area ที่นำเสนอในงาน (T Kasar et al., 2007) เพื่อให้สอดคล้องกับงานวิเคราะห์ตัวอักษรบนหน้ายามากขึ้น และต้องการเปรียบเทียบประสิทธิภาพเทคนิคการจำแนกสัญลักษณ์บนหน้ายาระหว่างเทคนิค Otsu's thresholding เทคนิค Local adaptive thresholding และเทคนิคที่นำเสนอในงานวิจัย (T Kasar et al., 2007) การทดลองเบื้องต้นพัฒนาบนภาษา Python โดยใช้เฟรมเวิร์ค Opencv ซึ่งเป็นเฟรมเวิร์คที่มีชุดคำสั่งสำหรับการประมวลผลภาพ โดยทำทดลองกับภาพเม็ดยาจากเว็บไซต์ฐานข้อมูลออนไลน์ ("Drugs.com Pill Identifier," 2016) และจากกิจกรรม Pill Image Recognition Challenge ของ U.S. National Library of Medicine (NLM) โดยแบ่งเป็น Training set 30 ภาพ และ Test set จำนวน 423 ภาพซึ่งภาพเม็ดยาทั้งหมดสัญลักษณ์บนหน้ายาจะประกอบไปด้วยตัวอักษรภาษาอังกฤษ ตัวเลขอารบิกและสัญลักษณ์ (symbol) จากผู้ผลิต

เกณฑ์การพิจารณาประกอบไปด้วย 3 ส่วนหลักคือค่าความแม่นยำ (Precision) ค่าความครบถ้วน (Recall) และค่า F-measure ซึ่งสะท้อนถึงความถูกต้อง (accuracy) ของระบบโดยมีหลักการคำนวณตามหัวข้อ 4.1 เมื่อพิจารณาจากภาพเม็ดยา 10 อันดับที่ให้ค่า F-measure สูงสุดของแต่ละวิธีการตามภาพที่ 3.11 พบว่าวิธีการจำแนกพื้นที่บริเวณสัญลักษณ์บนหน้ายาที่นำเสนอในงานวิจัยนี้เมื่อทำงานร่วมกับเทคนิค Otsu's thresholding สามารถให้ค่า F-measure ได้อย่างโดดเด่นมากเมื่อเทียบกับทำงานร่วมกับเทคนิค Adaptive thresholding หรือเมื่อเปรียบเทียบผลกับอัลกอริธึมที่เสนอในงาน (T Kasar et al., 2007)



ภาพที่ 3.11 ตัวอย่างค่า F-measure ของ 10 อันดับที่ดีที่สุดของแต่ละวิธีการ

นอกจากนี้วิธีการหาพื้นที่สัญลักษณ์ของเม็ดยาในภาพที่นำเสนอเมื่อเทียบกับวิธีดั้งเดิมที่เสนอโดย Kasar พบว่าวิธีที่นำเสนอสามารถให้ค่าความถูกต้องได้ถึง 62.41 เปอร์เซ็นต์ซึ่งดีกว่าวิธีดั้งเดิมกว่า 2 เท่าตัวจากตารางที่ 3.1 แต่อย่างไรก็ดีค่าความแม่นยำ ค่าความครบถ้วน และค่า F-measure เฉลี่ยของวิธีการที่นำเสนอจากตารางที่ 3.2 พบว่ายังอยู่ในเกณฑ์ต่ำมากแม้ว่าจะได้ค่าทั้งสามจะมีค่าสูงต่างจากวิธีอื่นๆเกือบเท่าตัว และเมื่อพิจารณา

ตารางที่ 3.3 ร่วมด้วยจะแสดงให้เห็นว่าสาเหตุที่ทำให้ค่าทั้งสามของวิธีที่นำเสนอต่ำนั้นคือ การทำงานกับเม็ดยาที่มีตัวอักษรบนหน้ายาแบบตัวลึกลง ดังนั้นหากสามารถพัฒนาเทคนิคการจำแนกตัวอักษรบนหน้ายาสำหรับตัวอักษรแบบลึกลงที่มีประสิทธิภาพได้อาจส่งผลให้ค่าความแม่นยำ ค่าความครบถ้วน และค่า F-measure สูงกว่านี้มาก

ตารางที่ 3.1 ผลการจำแนกพื้นที่บริเวณสัญลักษณ์บนหน้ายา

	วิธีที่ทำงานวิจัยนี้นำเสนอ	วิธีที่นำเสนอโดย Kasar
ความแม่นยำของการจำแนกพื้นที่บริเวณสัญลักษณ์บนหน้ายา (%)	62.41	29.31

ตารางที่ 3.2 ค่าเฉลี่ยของค่าความแม่นยำ ค่าความครบถ้วนและค่า F-measure จากผลลัพธ์โปรแกรม OCR

	ตัวอักษรหน้ายาแบบพิมพ์			ตัวอักษรหน้าแบบลึก		
	เทคนิคที่นำเสนอนในหัวข้อ3.2 +Otsu	เทคนิคที่นำเสนอในหัวข้อ3.2 + Adaptive mean	เทคนิคการสร้างภาพขาวดำของKasar	เทคนิคที่นำเสนอนในหัวข้อ3.2 +Otsu	เทคนิคที่นำเสนอนในหัวข้อ3.2 + Adaptive mean	เทคนิคการสร้างภาพขาวดำของKasar
Precision	0.710	0.291	0.405	0.111	0.038	0.059
Recall	0.610	0.282	0.419	0.113	0.040	0.067
F-measure	0.656	0.286	0.412	0.112	0.039	0.062

ตารางที่ 3.3 ค่าความแม่นยำ ค่าความครบถ้วนและค่า F-measure เฉลี่ยของแต่ละระบบ

	เทคนิคที่นำเสนอนในหัวข้อ3.2 +Otsu	เทคนิคที่นำเสนอนในหัวข้อ3.2 + Adaptive mean	เทคนิคการสร้างภาพขาวดำของKasar
Precision	0.258	0.100	0.144
Recall	0.237	0.097	0.152
F-measure	0.247	0.099	0.148

บทที่ 4

ผลการวิจัยและการอภิปรายผล

ในการทดลองนี้มีจุดประสงค์เพื่อประเมินประสิทธิภาพวิธีการวิเคราะห์สัญลักษณ์บนหน้ายาที่ให้ผลลัพธ์อยู่ในรูปแบบข้อความตัวอักษรได้อย่างอัตโนมัติที่นำเสนอในบทที่ 3 รวมถึงวิเคราะห์ผลการเปรียบเทียบความสามารถของเทคนิคการสร้างภาพขาวดำระหว่างเทคนิคที่ประยุกต์ Otsu thresholding ร่วมกับ Noise elimination และเทคนิคที่ใช้ K-means clustering

ระบบพัฒนามาตรฐานภาษา Python โดยใช้เฟรมเวิร์ค Opencv (“OpenCV 3.1.0,” n.d.) ซึ่งเป็นเฟรมเวิร์คที่มีชุดคำสั่งสำหรับการประมวลผลภาพ ภาพเม็ดยาที่ใช้ในการทดลองแบ่งเป็น 2 ชุดคือ ชุดที่ใช้สำหรับการทดลอง (Test set) และชุดที่ใช้สำหรับการเลือกค่าพารามิเตอร์ (Training set) โดยภาพเม็ดยาทั้ง 2 ชุดสัญลักษณ์บนหน้ายาจะประกอบจากตัวอักษรภาษาอังกฤษและตัวเลขอารบิกเท่านั้น สำหรับชุดภาพเม็ดยาที่ใช้ในการทดลองได้จากฐานข้อมูลเม็ดยาออนไลน์ (“Drugs.com Pill Identifier,” 2016) ที่เกิดจากการร่วมมือกันระหว่าง 4 บริษัททางการแพทย์ชั้นนำดังรายละเอียดในหัวข้อ 2.1 โดยชุดภาพทดลองประกอบไปด้วยเม็ดยาที่มีสัญลักษณ์บนหน้ายาแบบลึกลับจำนวน 360 ภาพ และเม็ดยาที่มีสัญลักษณ์บนหน้ายาแบบพิมพ์จำนวน 180 ภาพ รวมเม็ดยาทั้งหมด 540 ภาพที่มีความละเอียดภาพ 96 ถึง 150 DPI (Dots Per Inch)

เม็ดยาที่มีสัญลักษณ์บนหน้ายาแบบลึกลับประกอบไปด้วยเม็ดยา 4 สีซึ่งเป็นสีเม็ดยาที่พบบากสุด (Caban et al., 2012) คือ สีขาว สีเหลือง สีฟ้า และสีส้ม โดยแต่ละสีจำแนกได้อีก 3 ประเภทตามรูปร่างเม็ดยาที่พบบากที่สุดคือ ทรงกลม (round) ทรงรี (oval) และทรงแคปซูล (capsule) จำนวนกลุ่มละ 30 ภาพ ดังนั้นแต่ละกลุ่มสีเม็ดยาจะมีชุดภาพที่ใช้ในการทดลองจำนวน 90 ภาพ

สำหรับเม็ดยาที่มีสัญลักษณ์บนหน้ายาแบบพิมพ์แบ่งได้เป็น 2 กลุ่มหลักคือเม็ดยาที่มีหนึ่งสี และเม็ดยาที่มีมากกว่าหนึ่งสี โดยเม็ดยาที่มีเพียงหนึ่งสีแบ่งตามกลุ่มสีของเม็ดยาได้ดังนี้ สีขาว สีเหลือง สีส้ม และสีฟ้า และแต่ละกลุ่มสีเม็ดยาจะประกอบไปด้วยชุดภาพที่ใช้ในการทดลองจำนวน 30 ภาพ ส่วนเม็ดยาที่มีสีมากกว่าหนึ่งจะมีชุดภาพที่ใช้ในการทดลองจำนวน 60 ภาพ

ส่วนชุดภาพเม็ดยาที่ใช้สำหรับการเลือกค่าพารามิเตอร์ได้จากภาพเม็ดยาในกิจกรรม Pill Image Recognition Challenge ของ U.S. National Library of Medicine (NLM) จำนวน 40 ภาพซึ่งแบ่งเป็นเม็ดยาที่มีตัวอักษรแบบลึกลับ 20 ภาพและตัวอักษรแบบพิมพ์ 20 ภาพ โดยตัวอักษรแบบลึกลับประกอบไปด้วยเม็ดยาสีขาว สีส้ม สีเหลือง และสีฟ้าที่มีรูปทรงกลม ทรงรี และแคปซูล ส่วน

เมื่อยาที่มีตัวอักษรแบบพิมพ์ประกอบไปด้วยเมื่อยาที่มีมากกว่าหนึ่งสีและเมื่อยาสีเดียวคือ สีขาว สีเหลือง สีฟ้า และสีส้มที่มีรูปทรงกลม ทรงรี และแคปซูล

4.1. เกณฑ์การประเมินผล

เกณฑ์ที่ใช้สำหรับประเมินประสิทธิภาพของวิธีการที่นำเสนอในบทที่ 3 ประกอบไปด้วย 4 ส่วนหลักคือ ความแม่นยำ (Precision) ความครบถ้วน (Recall) และ F-measure และเวลาที่ใช้โดยเฉลี่ยสำหรับเมื่อยา 1 ภาพ ซึ่งมีรูปแบบสมการดังนี้

$$precision = \frac{|{\text{relevant items}} \cap {\text{retrieved items}}|}{|{\text{retrived items}}|}$$

$$recall = \frac{|{\text{relevant items}} \cap {\text{retrieved items}}|}{|{\text{relevant items}}|}$$

$$F - Measure = \frac{2 \times precision \times recall}{precision + recall}$$

$$\text{เวลาที่ใช้โดยเฉลี่ย} = \frac{\sum \text{เวลาที่ใช้ตั้งแต่รับภาพจนถึงได้ผลลัพธ์จากโปรแกรม OCR}}{\text{จำนวนงานที่เสร็จ}}$$

เกณฑ์การประเมินผลความแม่นยำและความครบถ้วนจะประเมินในระดับการเปรียบเทียบตัวอักษรกับตัวอักษรของแต่ละประโยคสัญลักษณ์บนหน้ายา ซึ่งเปรียบเทียบระหว่างประโยคผลลัพธ์จากระบบและประโยคที่ต้องการจากฐานข้อมูลจากไฟล์ข้อมูลเมื่อยาที่ถูกเตรียมไว้ก่อนแล้วดังภาพที่ 4.1 คอลัมน์ที่ 2

1	Ace00010.jpg;M361;engrave;Blue;Capsule-shape;
2	Ace00018.jpg;WATSON540;engrave;Blue;Elliptical/Oval;
3	Ace00095.jpg;WATSON;engrave;Blue;Capsule-shape;
4	Ace00100.JPG;44235;engrave;Blue;Capsule-shape;
5	Acy00065.jpg;400;engrave;Blue;Capsule-shape;
6	Alp00059.JPG;605;engrave;Blue;Elliptical/Oval;
7	Alp01214.JPG;1.0;engrave;Blue;Elliptical/Oval;

ภาพที่ 4.1 ตัวอย่างไฟล์ข้อมูลเมื่อยาที่ได้จากฐานข้อมูล Drugs.com ที่มีข้อมูลรายการตัวอักษรหน้ายา (กรอบสีฟ้า)

โดย Relevant items คือ เซตของตัวหนังสือแต่ละตัวบนข้อมูลหน้ายาจริงซึ่งได้จากฐานข้อมูล ยกตัวอย่างเมื่อยาที่ 1 จากภาพที่ 4.1 Relevant items คือ 'M361' ส่วน Retrieved

items คือ เซตของตัวหนังสือแต่ละตัวบนข้อมูลหน้ายาที่ได้จากระบบที่นำเสนอในบทที่ 3 ซึ่งอาจไม่ตรงกับ Relevant items ของภาพเมื่อดูเดียวกัน ส่งผลให้ค่าความแม่นยำและค่าความครบถ้วนของระบบคำนวณจากการเปรียบเทียบระหว่างเซตตัวหนังสือดังนั้นผลการวัดจึงไม่สะท้อนความถูกต้องของลำดับตัวอักษรในประโยค

สำหรับค่า f-measure จะเป็นค่าที่สะท้อนความแม่นยำและความครบถ้วนของระบบ ส่วน เวลาที่ใช้โดยเฉลี่ยได้จากสัดส่วนของเวลาที่ใช้ตั้งแต่ระบบอ่านภาพแรกของชุดภาพทดสอบเข้าสู่ระบบจนได้ประโยคสัญลักษณ์หน้ายาของภาพสุดท้ายหารด้วยจำนวนภาพทั้งหมดในชุดภาพทดสอบ

4.2. การเลือกค่าพารามิเตอร์

ก่อนการประเมินผลประสิทธิภาพวิธีการวิเคราะห์สัญลักษณ์บนหน้ายาผู้วิจัยได้ทำการทดลองย่อยโดยใช้ภาพจาก Test set จำนวน 40 ภาพเพื่อเลือกค่าพารามิเตอร์ที่ใช้สำหรับเทคนิค Detail enhancement และเทคนิค Edge-aware smoothing ที่นำเสนอโดย Gastal ที่ใช้ในขั้นตอน 3.1.2 ด้วยการเปลี่ยนแปลงค่าพารามิเตอร์ที่สำคัญ 2 ค่าคือค่า σ_s ซึ่งมีค่าที่เป็นไปได้ตั้งแต่ 0 ถึง 200 โดยเพิ่มค่าไปที่ละ 50 หน่วย ส่วนค่าพารามิเตอร์ σ_r มีค่าที่เป็นไปได้ตั้งแต่ค่า 0 ถึง 1 ดังนั้นจึงเพิ่มค่าไปที่ละ 0.25 หน่วย หลังจากนั้นเมื่อทราบช่วงค่าพารามิเตอร์ที่เหมาะสมจะทำการเปลี่ยนแปลงค่าพารามิเตอร์อีกรอบเป็น 5 ค่าย่อยคล้ายกับรอบแรก โดยจะตัดสินค่าพารามิเตอร์ที่เหมาะสมดูจากความสมบูรณ์ของตัวอักษรภายในภาพผลลัพธ์ที่ได้จากขั้นตอน 3.2 โดยพิจารณาจาก 2 ข้อคือ ภาพผลลัพธ์ต้องมีตัวอักษรหน้ายาครบถ้วน และต้องไม่มีฟิกเซลพื้นหลังปรากฏในภาพผลลัพธ์เกิน 10 เปอร์เซ็นต์ของฟิกเซลทั้งหมด

จากการเลือกค่าพารามิเตอร์พบว่าค่าพารามิเตอร์ที่เหมาะสมที่สุดสำหรับเทคนิค Detail enhancement คือค่า σ_s เท่ากับ 10 และค่า σ_r เท่ากับ 0.15 ส่วนเทคนิค Edge-aware smoothing ค่าที่เหมาะสมคือค่า σ_s เป็น 35 และค่า σ_r เป็น 0.4 ตามลำดับซึ่งจะถูกนำไปใช้ในการทดลองในหัวข้อ 4.3 ต่อไป

4.3. ขั้นตอนการทดลอง

ผู้วิจัยได้ทำการทดลองเพื่อประเมินประสิทธิภาพวิธีการจำแนกตัวอักษรบนหน้ายาที่นำเสนอในบทที่ 3 โดยใช้ภาพเม็ดยาจาก Test set จำนวน 540 ภาพดังที่อธิบายในข้างต้น โดยระบบพัฒนาบนภาษา Python เวอร์ชัน 2.7 ร่วมกับเฟรมเวิร์ค Opencv (“OpenCV 3.1.0,” n.d.) ซึ่งการทำงานเริ่มจากการเตรียมภาพด้วยเทคนิค Gatal และใช้เทคนิคการจำแนกพื้นที่บริเวณสัญลักษณ์บนหน้ายาที่นำเสนอในหัวข้อ 3.2 และจำแนกสัญลักษณ์บนหน้ายาด้วยการสร้างภาพขาวดำด้วยเทคนิคที่ประยุกต์ Otsu thresholding ร่วมกับ Noise elimination หรือเทคนิคที่ใช้ K-means clustering ขั้นตอนสุดท้ายแปลงภาพขาวดำของตัวอักษรให้อยู่ในรูปแบบรายการตัวอักษรด้วยโปรแกรม OCR ดังรายละเอียดจากบทที่ 3

จุดประสงค์ของการทดลองแบ่งได้เป็น 2 จุดประสงค์ดังนี้ จุดประสงค์แรกต้องการประเมินค่าพารามิเตอร์ที่ได้มาจากหัวข้อ 4.1 ว่าสามารถส่งผลให้ประสิทธิภาพของระบบดีขึ้นจริงเกณฑ์การพิจารณาดูจากค่าความแม่นยำของการจำแนกพื้นที่สัญลักษณ์บนหน้ายาเมื่อเปรียบเทียบกับก่อนและหลังเลือกค่าพารามิเตอร์ซึ่งพิจารณาจากภาพผลลัพธ์ต้องมีตัวอักษรหน้ายาครบถ้วน และต้องไม่มีพิกเซลพื้นหลังปรากฏในภาพผลลัพธ์เกิน 10 เปอร์เซ็นต์ของพิกเซลทั้งหมด เกณฑ์ถัดมาคือความแม่นยำและความครบถ้วนของระบบ

ตารางที่ 4.1 ผลการจำแนกพื้นที่บริเวณสัญลักษณ์บนหน้ายาที่นำเสนอในงานวิจัยนี้

	ก่อนการเลือกค่าพารามิเตอร์	หลังการเลือกค่าพารามิเตอร์
ความแม่นยำของการจำแนกพื้นที่บริเวณสัญลักษณ์บนหน้ายา (%)	79.44	89.814

จากการทดลองพบว่า การเลือกค่าพารามิเตอร์ส่งผลให้ความแม่นยำของการจำแนกพื้นที่บริเวณสัญลักษณ์บนหน้ายาดีขึ้นถึง 10 เปอร์เซ็นต์จากตารางที่ 4.1 เมื่อเทียบกับความแม่นยำก่อน

กระบวนการเลือกค่าพารามิเตอร์ซึ่งใช้ค่าพารามิเตอร์ที่เป็นค่าแนะนำในงานวิจัย (Gastal & Oliveira, 2011)

จุดประสงค์ที่สองคือต้องการประเมินประสิทธิภาพวิธีการวิเคราะห์สัญลักษณ์บนหน้ายา ที่นำเสนอในบทที่ 3 รวมถึงวิเคราะห์ผลการเปรียบเทียบความแม่นยำและความครบถ้วนของเทคนิค การสร้างภาพขาวดำระหว่างเทคนิคที่ประยุกต์ Otsu thresholding ร่วมกับ Noise elimination และเทคนิคที่ใช้ K-means clustering

4.4. อภิปรายผลการทดลอง

จากการทดลองสร้างระบบวิเคราะห์สัญลักษณ์บนหน้ายาที่ให้ผลลัพธ์อยู่ในรูปแบบ ข้อความตัวอักษรจากภาพที่มีขั้นตอนการทำงานตามรายละเอียดในบทที่ 3 โดยพัฒนาบนภาษา Python และใช้เฟรมเวิร์ค Opencv (“OpenCV 3.1.0,” n.d.) ผลการทดลองสามารถสรุปผลได้ ดังต่อไปนี้

ตารางที่ 4.2 ค่าเฉลี่ยของค่าความแม่นยำ ค่าความครบถ้วนและค่า F-measure จากผลลัพธ์โปรแกรม OCR

	เทคนิค K-means clustering ที่ทำ ขั้นตอนในหัวข้อ4.2		เทคนิค Otsu ที่ไม่ ประยุกต์เทคนิค Noise elimination และไม่มี การสอนโปรแกรม OCR และไม่ทำขั้นตอนใน หัวข้อ4.2	เทคนิค Otsu ที่ทำขั้นตอนในหัวข้อ4.2			
	ไม่สอนโปรแกรม OCR	สอนโปรแกรม OCR		ไม่ประยุกต์เทคนิค Noise elimination		ประยุกต์เทคนิค Noise elimination	
				ไม่สอน โปรแกรม OCR	สอน โปรแกรม OCR	ไม่สอน โปรแกรม OCR	สอน โปรแกรม OCR
Precision	0.036	0.055	0.465	0.491	0.544	0.537	0.585
Recall	0.045	0.061	0.464	0.514	0.541	0.562	0.588
F-measure	0.04	0.058	0.467	0.502	0.542	0.548	0.586
เวลาที่ใช้โดย เฉลี่ยต่อภาพ (Millisecond)	15.883	16.922	9.254	9.672	11.731	11.470	13.278

การเปรียบเทียบค่า F-measure ของเทคนิคการจำแนกสัญลักษณ์บนหน้ายา 2 เทคนิค ที่เสนอในหัวข้อ 3.3 พบว่าเทคนิคที่ประยุกต์ Otsu thresholding ร่วมกับ Noise elimination

สามารถให้ค่า F-measure ต่ำกว่ามากและใช้เวลาในการประมวลผลน้อยกว่าดังตารางที่ 4.2 โดยเฉพาะเมื่อดูค่าที่มีสัญลักษณ์หน้ายาแบบลิกเทคนิคที่ใช้ K-means clustering ให้ค่า F-measure ที่ต่ำที่สุดจากตารางที่ 4.3 ซึ่งสาเหตุหลักมาจากเมื่อดูค่าและตัวอักษรบนหน้ายามีสีที่ใกล้เคียงกันมาก ส่งผลให้เกิดการปนกันระหว่างกลุ่มพิกเซลสีเม็ดยาและกลุ่มพิกเซลสีตัวอักษรเมื่อจัดกลุ่มพิกเซลตามค่าสีทำให้ภาพขาวดำผลลัพธ์มีตัวอักษรไม่ครบถ้วนและเกิดพิกเซลรบกวน

นอกจากนี้สาเหตุหลักที่ทำให้ค่า F-measure ของเทคนิคที่ใช้ K-means clustering มีค่าเฉลี่ยน้อยกว่าเทคนิคที่ใช้ Otsu's thresholding ถึง 10 เท่าเนื่องจากเทคนิคที่ใช้ K-means clustering ดั้งเดิมถูกพัฒนาขึ้นเพื่อจำแนกตัวอักษรพิมพ์ลิกบนวัตถุโลหะ เช่น จำแนกหมายเลขบนอาวุธปืน ดังนั้นเมื่อนำมาทำงานกับภาพเม็ดยาที่มีทั้งตัวอักษรแบบพิมพ์และตัวอักษรแบบลิกจึงได้ค่า F-measure ที่ไม่ดีนัก

ตารางที่ 4.3

ค่าความแม่นยำ ค่าความครบถ้วน และค่า F-measure เฉลี่ยของการสร้างภาพขาวดำด้วยเทคนิคของ K-means clustering

	ตัวอักษรหน้ายาแบบลิก		ตัวอักษรหน้ายาแบบพิมพ์	
	ไม่สอนโปรแกรม OCR	สอนโปรแกรม OCR	ไม่สอนโปรแกรม OCR	สอนโปรแกรม OCR
Precision	0.018	0.035	0.053	0.075
Recall	0.024	0.039	0.066	0.082
F-measure	0.021	0.037	0.059	0.078

ตารางที่ 4.4

ค่าความแม่นยำ ค่าความครบถ้วน และค่า F-measure เฉลี่ยของการสร้างภาพขาวดำด้วยเทคนิค Otsu ที่ทำขั้นตอนการเลือกค่าพารามิเตอร์

	ตัวอักษรหน้ายาแบบลิก					ตัวอักษรหน้ายาแบบพิมพ์				
	เทคนิค Otsu ที่ไม่ ประยุกต์เทคนิค Noise elimination และไม่มีการสอน โปรแกรม OCR และ ไม่ทำขั้นตอนในหัวข้อ 4.2	ไม่ประยุกต์เทคนิค Noise elimination		ประยุกต์เทคนิค Noise elimination		เทคนิค Otsu ที่ไม่ ประยุกต์เทคนิค Noise elimination และไม่มีการสอน โปรแกรม OCR และ ไม่ทำขั้นตอนในหัวข้อ 4.2	ไม่ประยุกต์เทคนิค Noise elimination		ประยุกต์เทคนิค Noise elimination	
		ไม่สอน โปรแกรม OCR	สอน โปรแกรม OCR	ไม่สอน โปรแกรม OCR	สอน โปรแกรม OCR		ไม่สอน โปรแกรม OCR	สอน โปรแกรม OCR	ไม่สอน โปรแกรม OCR	สอน โปรแกรม OCR
Precision	0.280	0.308	0.390	0.314	0.388	0.659	0.674	0.698	0.759	0.781
Recall	0.304	0.345	0.399	0.360	0.410	0.624	0.683	0.683	0.764	0.765
F- measure	0.292	0.325	0.394	0.335	0.399	0.641	0.678	0.690	0.761	0.773

เทคนิคที่ประยุกต์ Otsu thresholding ร่วมกับ Noise elimination จากตารางที่ 4.4 พบว่าทำงานกับตัวอักษรบนหน้ายาแบบพิมพ์ได้ดีกว่าตัวอักษรบนหน้ายาแบบลิกมากเกือบ 2 เท่า เนื่องจากเม็ดยาที่มีตัวอักษรแบบพิมพ์มีสีเม็ดยาและสีตัวอักษรแตกต่างกันชัดเจนทำให้จำแนกสัญลักษณ์บนหน้ายาได้ครบถ้วนและเกิดฟิสิกเซลรบกวนน้อย นอกจากนี้เทคนิค Noise elimination ก็เป็นส่วนสำคัญที่ทำให้ค่า F-measure ของตัวอักษรบนหน้ายาแบบพิมพ์สูงขึ้น เนื่องจากตัวอักษรหน้ายาแบบพิมพ์ที่มีเม็ดยามากกว่าหนึ่งสีล้วนประสบปัญหาที่มีเพียงตัวอักษรบนเม็ดยาสีใดสีหนึ่งเท่านั้นที่ค่า threshold ของเทคนิค Otsu จำแนกได้ดังภาพที่ 3.8 (b) แฉกที่หกซึ่งเทคนิค Noise elimination เข้ามาปรับปรุงภาพระดับเทา ก่อนสร้างภาพขาวดำผลลัพธ์ส่งผลให้ภาพขาวดำผลลัพธ์ จำแนกตัวอักษรได้ครบถ้วนทั้งเม็ดยา แต่สำหรับตัวอักษรแบบลิกเทคนิค Noise elimination ไม่ได้มีส่วนช่วยเท่าไรนักเนื่องจากภาพที่เทคนิค Noise elimination ช่วยคือภาพที่เกิดเงาตกกระทบ บริเวณหน้ายาซึ่งมีจำนวนภาพน้อยกว่า 15 เปอร์เซ็นต์จาก Test set ยิ่งไปกว่านั้นภาพเม็ดยาอีก 85 เปอร์เซ็นต์ที่เหลือเทคนิค Noise elimination มีแนวโน้มพิจารณาบางส่วนของฟิสิกเซลตัวอักษรเป็นเงาตกกระทบส่งผลให้ภาพขาวดำมีตัวอักษรไม่ครบถ้วน สิ่งที่มีส่วนช่วยเพิ่มค่า F-measure ของตัวอักษรบนหน้ายาแบบลิกสำหรับเทคนิคที่ประยุกต์ Otsu thresholding กลับเป็นกระบวนการสอนโปรแกรม OCR จากตารางที่ 4.2 เนื่องจากตัวอักษรในภาพขาวดำที่ได้จากสัญลักษณ์บนหน้ายาแบบลิกมีลักษณะเส้นตัวอักษรไม่คมชัดเกิดเป็นฟอนต์ใหม่ดังภาพที่ 3.8 (d) ในขณะที่ภาพขาวดำของตัวอักษรแบบพิมพ์มีลักษณะเส้นและรูปทรงตัวอักษรคมชัดคล้ายฟอนต์พื้นฐานภาษาอังกฤษ ดังนั้นกระบวนการสอนโปรแกรม OCR จึงไม่มีอิทธิพลกับค่า F-measure ที่ได้จากตัวอักษรบนหน้ายาแบบ

พิมพ์ นอกจากนี้ขั้นตอนในหัวข้อ 4.2 มีส่วนช่วยทำให้ค่า F-measure ของทั้งระบบดีขึ้น 0.04 โดยประมาณจากตารางที่ 4.2

ดังนั้นจากผลการทดลองข้างต้นสรุปได้ว่าวิธีการวิเคราะห์ตัวอักษรบนหน้ายาจากภาพด้วยเทคนิคที่ประยุกต์ Otsu's thresholding ร่วมกับ Noise elimination ที่ทำขั้นตอนเลือกค่าพารามิเตอร์และใช้กระบวนการสอนโปรแกรม OCR ร่วมด้วยเป็นวิธีที่ดีที่สุดทั้งสำหรับเมื่อยามีสัญลักษณ์หน้ายาแบบลึกและแบบพิมพ์ และสามารถให้ค่า F-measure โดยเฉลี่ยทั้งระบบที่ดีที่สุดคือ 0.586 จากตารางที่ 4.2 โดยเฉพาะกับสัญลักษณ์หน้ายาแบบพิมพ์ที่ระบบสามารถให้ค่า F-measure ได้เกิน 0.77 จากตารางที่ 4.4 แต่สำหรับสัญลักษณ์บนหน้ายาแบบลึกยังได้ค่า F-measure ไม่ดีนัก จำเป็นต้องได้รับการพัฒนาต่อไป

เมื่อพิจารณาวิธีที่ดีที่สุดที่มีการสอนโปรแกรม Tesseract ร่วมด้วยโดยเน้นไปที่การสอนรู้จำตัวเลขอารบิก จากผลการทดลองพบว่าอัตราการรู้จำผิดพลาดของตัวอักษรเปรียบเทียบกับตัวเลขมีเปอร์เซ็นต์ความผิดพลาดใกล้เคียงกันมากคือ ผิดพลาดไป 40.9 (ตัวอักษรภาษาอังกฤษ) และ 38.3 (ตัวเลขอารบิก) เปอร์เซ็นต์ตามลำดับ เนื่องจากตั้งเดิมโปรแกรม Tesseract ได้รับการสอนให้รู้จำพอนต์ตัวอักษรภาษาอังกฤษพื้นฐานอยู่แล้วจากทางผู้พัฒนา Tesseract อย่างไรก็ตามสำหรับการสอนโปรแกรม Tesseract โดยเฉพาะพอนต์ตัวเลขจากภาพขาวดำของ Otsu's thresholding สามารถเพิ่มค่า F-measure เฉลี่ยทั้งระบบได้ 0.04 โดยประมาณจากตารางที่ 4.2

บทที่ 5

บทสรุปและข้อเสนอแนะ

งานวิจัยนี้นำเสนอวิธีการสร้างระบบวิเคราะห์สัญลักษณ์บนหน้ายาแบบใหม่ที่ให้ผลลัพธ์อยู่ในรูปแบบข้อความตัวอักษรได้อย่างอัตโนมัติซึ่งประกอบไปด้วย 5 ขั้นตอนหลัก ขั้นตอนแรกการเตรียมภาพเพื่อลดความหลากหลายของขนาดภาพพร้อมเพิ่มรายละเอียดขอบของตัวอักษรหน้ายาเพื่อให้งานหาพื้นที่เฉพาะตัวอักษรหน้ายาจากภาพขอบวัตถุทำได้แม่นยำมากขึ้น หลังจากนั้นในขั้นตอนการหาพื้นที่ตัวอักษรหน้ายาโดยใช้เงื่อนไข 5 ข้อที่นำเสนอในงานวิจัยนี้ซึ่งเป็นผลลัพธ์การปรับปรุงจากกฎ 3 ข้อสำหรับการหาพื้นที่ตัวอักษรบนภาพถ่ายเอกสารจากงาน (T Kasar et al., 2007) เพื่อให้สามารถทำงานได้แม่นยำมากขึ้นกับงานเม็ดยา เนื่องจากกฎ 3 ข้อดั้งเดิมเกณฑ์การหาพื้นที่ตัวอักษรมักขึ้นกับขนาดภาพ แต่สำหรับงานเม็ดยาที่ไม่ได้กำหนดระยะยกกล้องดังนั้นจึงไม่สามารถใช้กฎของ Kasar ได้ จึงต้องหาวัตถุอ้างอิงอื่นเป็นเกณฑ์การพิจารณาแทนสำหรับการจำแนกพื้นที่ตัวอักษรและพื้นที่ฟิกเชลรบกวน ซึ่งในงานวิจัยนี้เลือกใช้ขอบเม็ดยาเป็นวัตถุอ้างอิงโดยการตัดภาพในขั้นตอนก่อนเข้าเงื่อนไข 5 ข้อให้พอดีครอบคลุมฟิกเชลเม็ดยา นอกจากนี้งานของ Kasar ไม่จำเป็นต้องจัดการกับขอบเม็ดยาเพราะทำงานบนภาพถ่ายเอกสาร ดังนั้นเงื่อนไข 5 ข้อที่นำเสนอจำเป็นต้องเพิ่มเงื่อนไขบางข้อเพื่อจำแนกกลุ่มฟิกเชลขอบเม็ดยาออกจากกลุ่มฟิกเชลตัวอักษร รวมไปถึงจากการทดลองเบื้องต้นกับภาพเม็ดยาในหัวข้อ 3.6 พบว่า กฎ 3 ข้อของ Kasar ไม่สามารถแยกแยะกลุ่มฟิกเชลตัวอักษรและกลุ่มฟิกเชลรบกวนได้ส่งผลให้ความแม่นยำของการจำแนกพื้นที่บริเวณสัญลักษณ์บนหน้ายาเท่ากับ 29.31 เปอร์เซ็นต์เมื่อเทียบกับเงื่อนไข 5 ข้อที่นำเสนอที่สามารถให้ค่าความแม่นยำถึง 62.41 เปอร์เซ็นต์จากการทดลองเบื้องต้น หลังจากนั้นเป็นขั้นตอนการสร้างภาพขาวดำซึ่งนำเสนอ 2 เทคนิคคือ เทคนิคที่ประยุกต์ Otsu thresholding ร่วมกับ Noise elimination และเทคนิคที่ใช้ K-means clustering และเมื่อได้ภาพขาวดำของตัวอักษรหน้ายาจะใช้โปรแกรม Tesseract ที่ผ่านกระบวนการสอนแล้วแปลงให้เป็นข้อความตัวอักษรผลลัพธ์

สำหรับเทคนิค Noise elimination ที่นำเสนอเป็นวิธีการกำจัดกลุ่มฟิกเชลปัญหาที่มักปรากฏบนภาพขาวดำที่สร้างจากเทคนิค Otsu thresholding บนภาพเม็ดยาที่ทำให้โปรแกรม Tesseract ไม่สามารถแปลงภาพตัวอักษรหน้ายาได้ ซึ่งสาเหตุมาจาก 2 กรณีคือ กรณีเม็ดยามีสีมากกว่าหนึ่งสี และกรณีเกิดเงาบนหน้ายาซึ่งส่งผลให้ฟิกเชลบริเวณเงามีสีใกล้เคียงกับสัญลักษณ์บนหน้ายาดังนั้นกรณีนี้จึงเกิดกับเฉพาะตัวอักษรหน้ายาแบบลึกลงเท่านั้น ซึ่งส่งผลอย่างมีนัยสำคัญกับ

ตัวอักษรแบบพิมพ์เนื่องจากมีภาพเม็ดยาเกือบครึ่งที่มีสีเม็ดยามากกว่าหนึ่งสีโดยสามารถเพิ่มค่า F-measure ของระบบได้ถึง 0.08

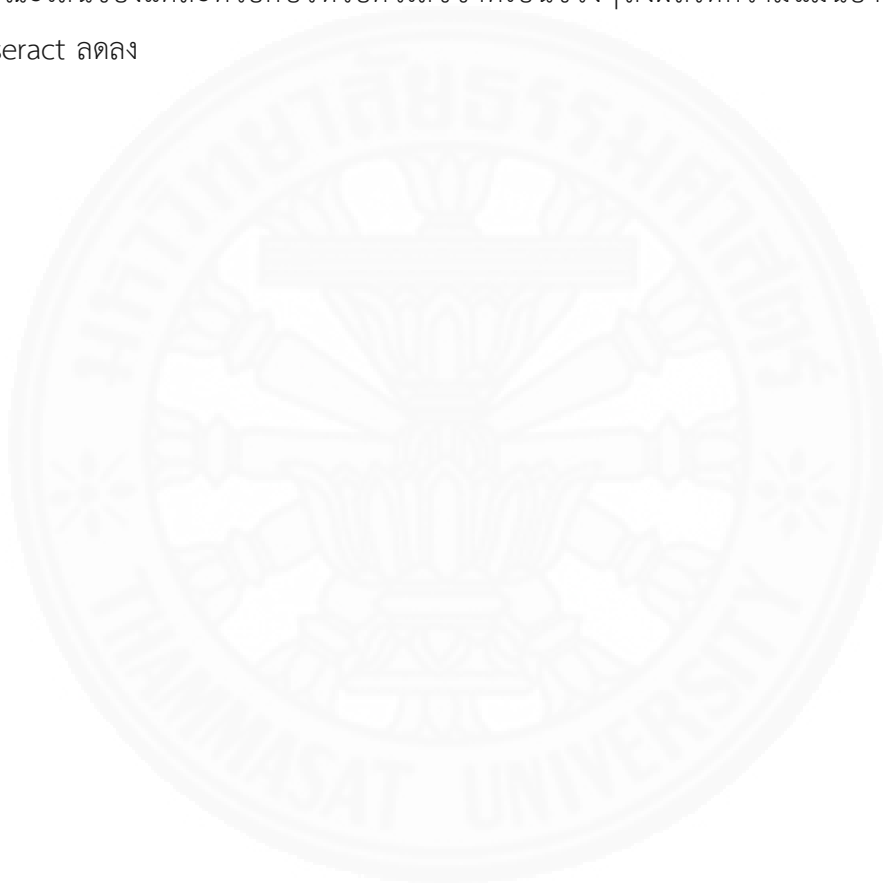
จากการทดลองพบว่าวิธีการที่นำเสนอสามารถสร้างระบบวิเคราะห์สัญลักษณ์บนหน้ายาแบบใหม่ที่ให้ผลลัพธ์ในรูปแบบตัวอักษรที่สามารถให้ความแม่นยำและความครบถ้วนของระบบเกิน 57 เปอร์เซ็นต์ โดยเฉพาะสัญลักษณ์หน้ายาแบบพิมพ์ที่สามารถให้ความแม่นยำและความครบถ้วนของระบบเกิน 77 เปอร์เซ็นต์แต่เทคนิคสำหรับสัญลักษณ์หน้ายาแบบลึกลงยังต้องได้รับการพัฒนาต่อไป โดยเฉพาะขั้นตอนการสร้างภาพขาวดำควรได้รับการพัฒนาเนื่องจากมีบางภาพที่ได้ฝึกเซลล์ตัวอักษรหน้ายาไม่สมบูรณ์ทำให้ความแม่นยำของโปรแกรม Tesseract ไม่ดีนัก นอกจากนี้เมื่อพิจารณาผลการทดลองของวิธีที่ดีที่สุดที่มีการสอนเน้นให้รู้จักพอนต์ตัวเลขอารบิกพบว่าเปอร์เซ็นต์ความผิดพลาดของการรู้จำตัวอักษรหรือตัวเลขของระบบมีค่าใกล้เคียงกันมากคือ 40.9 และ 38.3 เปอร์เซ็นต์ตามลำดับ แม้การสอนโปรแกรม Tesseract จะเน้นการสอนไปที่ตัวเลขก็ตาม เนื่องจากดั้งเดิมโปรแกรม Tesseract ได้รับการสอนให้รู้จักพอนต์พื้นฐานของตัวอักษรภาษาอังกฤษก่อนหน้าแล้วจากผู้พัฒนาโปรแกรม Tesseract แต่อย่างไรก็ดีการสอนที่เน้นไปที่ตัวเลขอารบิกสามารถเพิ่มความแม่นยำและค่าความครบถ้วนของระบบได้ถึง 0.04 โดยประมาณ

ผลการเปรียบเทียบเทคนิคการสร้างภาพขาวดำระหว่างเทคนิคที่ประยุกต์ Otsu's thresholding ร่วมกับ Noise elimination และเทคนิคที่ใช้ K-means clustering พบว่าเทคนิคที่ประยุกต์ Otsu's thresholding ร่วมกับ Noise elimination สามารถให้ค่าความแม่นยำและค่าความครบถ้วนของระบบมากกว่าถึง 10 เท่า ซึ่งสาเหตุหลักมาจากการที่เทคนิคที่ใช้ K-means clustering ถูกพัฒนาเพื่อทำงานกับภาพตัวอักษรสลักลึกลงบนแผ่นวัตถุโลหะจึงส่งผลให้การทำงานกับชุดภาพเม็ดยาที่มีทั้งรูปแบบตัวอักษรแบบพิมพ์และตัวอักษรแบบลึกลงให้ผลลัพธ์ไม่ดีนักรวมไปถึงลักษณะพื้นผิวโลหะที่มันวาวกว่าผิวเม็ดยามากส่งผลให้ลักษณะแสงและเงาที่เกิดขึ้นในภาพต่างกัน ประกอบกับงานวิจัยที่นำเสนอเทคนิคที่ใช้ K-means clustering พัฒนาระบบรู้จำตัวอักษรโดยเฉพาะเพื่อใช้ในขั้นตอนการรู้จำตัวอักษรบนหน้ายา ดังนั้นระบบรู้จำตัวอักษรนี้อาจถูกพัฒนาขึ้นเพื่อให้ทำงานได้ดีกับภาพที่ได้จากเทคนิคที่ใช้ K-means clustering มากกว่าการรู้จำตัวอักษรที่ใช้โปรแกรมรู้จำอักษรด้วยแสง เช่น Tesseract ทำให้ผลลัพธ์ที่ได้ในงานวิจัยนี้ไม่ดีเท่าที่รายงานในผลงานวิจัยต้นฉบับ

สำหรับงานวิจัยนี้การวัดผลค่าความแม่นยำและความครบถ้วนจะวัดในระดับการเปรียบเทียบตัวอักษรต่อตัวอักษรสำหรับแต่ละประโยคบนเม็ดยา แต่ถ้าต้องการนำไปใช้จริงสำหรับการต่อยอด เช่น นำไปใช้ร่วมกับเทคนิคการวิเคราะห์ลักษณะภายนอกของเม็ดยาอื่นๆ เพื่อเพิ่มความแม่นยำในการจำแนกเม็ดยาด้วยภาพ หรือนำไปใช้ต่อยอดเพื่อเป็นประโยชน์สำหรับฐานข้อมูลเม็ดยา อาจจะต้องคำนวณค่าความแม่นยำและความครบถ้วนในระดับประโยคของแต่ละเม็ดยาแทน ซึ่งจะส่งผลให้ค่า

ความแม่นยำและครบถ้วนค่าใหม่ลดลงเนื่องจากประโยชน์ผลลัพธ์จากระบบส่วนใหญ่ถูกเพียงบางส่วน
ของประโยคเท่านั้น

สำหรับแนวทางการพัฒนาได้แก่ การพัฒนาเทคนิค Noise elimination เพื่อให้สามารถลบ
กลุ่มพิกเซลปัญหาที่หมดจคมมากขึ้นสำหรับกรณีที่เกิดจากเงาบนหน้ายา รวมถึงการนำเทคนิคจำพวก
Broken character restoration มาใช้หลังขั้นตอนการสร้างภาพขาวดำเพื่อเชื่อมเส้นของแต่ละ
ตัวอักษรหรือตัวเลขก่อนเข้าโปรแกรม Tesseract เนื่องจากพบตัวอักษรหรือตัวเลขจำนวนมากที่มี
ลักษณะเส้นของแต่ละตัวอักษรหรือตัวเลขขาดเป็นช่วงๆ ส่งผลให้ความแม่นยำของโปรแกรม
Tesseract ลดลง



รายการอ้างอิง

- Andreas Hartl. (2010). Computer-Vision based Pharmaceutical Pill Recognition on Mobile Phones. In *Proceedings of the 14th Central European Seminar on Computer Graphics* (pp. 51–58). Budmerice, Slovakia.
- Caban, J. J., Rosebrock, A., & Yoo, T. S. (2012). Automatic identification of prescription drugs using shape distribution models (pp. 1005–1008). IEEE.
<https://doi.org/10.1109/ICIP.2012.6467032>
- Canny, J. (1986). A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-8(6)*, 679–698.
<https://doi.org/10.1109/TPAMI.1986.4767851>
- Chen, Z., & Kamata, S. (2013). A new accurate pill recognition system using imprint information. In B. Vuksanovic, J. Zhou, & A. Verikas (Eds.) (p. 906711).
<https://doi.org/10.1117/12.2051168>
- Chen, Z., Yu, J., Kamata, S., & Yang, J. (2015). Accurate system for automatic pill recognition using imprint information. *IET Image Processing, 9(12)*, 1039–1047.
<https://doi.org/10.1049/iet-ipr.2014.1007>
- Cunha, A., Adão, T., & Trigueiros, P. (2014). HelpmePills: A Mobile Pill Recognition Tool for Elderly Persons. *Procedia Technology, 16*, 1523–1532.
<https://doi.org/10.1016/j.protcy.2014.10.174>
- DAILYMED-National library of medicine. (2016). Retrieved November 7, 2016, from <http://dailymed.nlm.nih.gov/dailymed/index.cfm>
- Drug.com Pill Identifier. (2016). Retrieved November 7, 2016, from http://www.drugs.com/pill_identification.html

- E. Annasaro et al. (2014). Color and Shape Feature Extraction and Matching in Pill Identification Systems. *IJCSIT*, 5(2), 1011–1015.
- Epocrates-an Athenahealth service. (2016). Retrieved November 7, 2016, from <http://www.epocrates.com/>
- Epshtein, B., Ofek, E., & Wexler, Y. (2010). Detecting text in natural scenes with stroke width transform (pp. 2963–2970). IEEE. <https://doi.org/10.1109/CVPR.2010.5540041>
- Gastal, E. S. L., & Oliveira, M. M. (2011). Domain transform for edge-aware image and video processing. *ACM Transactions on Graphics*, 30(4), 1. <https://doi.org/10.1145/2010324.1964964>
- Gonzalez, R. C., & Woods, R. E. (2008). *Digital image processing* (3rd ed). Upper Saddle River, N.J: Prentice Hall.
- Grigorescu, C., & Petkov, N. (2003). Distance sets for shape filters and shape recognition. *IEEE Transactions on Image Processing*, 12(10), 1274–1286. <https://doi.org/10.1109/TIP.2003.816010>
- Hartl, A., Arth, C., & Schmalstieg, D. (2011). Instant Medical Pill Recognition on Mobile Phones. ACTAPRESS. <https://doi.org/10.2316/P.2011.740-017>
- Healthline pill identifier. (2016). Retrieved November 7, 2016, from <http://www.healthline.com/pill-identifier>
- Image Advanced search-MIMS thailand. (2016). Retrieved November 7, 2016, from <https://www.mims.com/Thailand/image/AdvancedSearch>
- Kim, D., & Chun, J. (2011). Drug Image Retrieval by Shape and Color Similarity of the Medication (pp. 387–390). IEEE. <https://doi.org/10.1109/CNSI.2011.40>

- Lee, Y.-B., Park, U., & Jain, A. K. (2010). PILL-ID: Matching and Retrieval of Drug Pill Imprint Images (pp. 2632–2635). IEEE. <https://doi.org/10.1109/ICPR.2010.645>
- Lee, Y.-B., Park, U., Jain, A. K., & Lee, S.-W. (2012). Pill-ID: Matching and retrieval of drug pill images. *Pattern Recognition Letters*, 33(7), 904–910. <https://doi.org/10.1016/j.patrec.2011.08.022>
- Lexicomp - drug information. (2016). Retrieved November 7, 2016, from <https://itunes.apple.com/th/app/lexicomp/id313401238?mt=8>
- Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2), 91–110. <https://doi.org/10.1023/B:VISI.0000029664.99615.94>
- Mancas, M., Mancas-Thillou, C., Gosselin, B., & Macq, B. (2006). A Rarity-Based Visual Attention Map - Application to Texture Description (pp. 445–448). IEEE. <https://doi.org/10.1109/ICIP.2006.312489>
- Mancas-Thillou, C., & Mancas, M. (2007). Comparison between Pen-scanner and Digital Camera Acquisition for Engraved Character Recognition (pp. 130–137). Presented at the International workshop on camera-based document analysis and recognition (CBDAR 2007), Curitiba, Brazil.
- Mantas, J. (1986). An overview of character recognition methodologies. *Pattern Recognition*, 19(6), 425–430. [https://doi.org/10.1016/0031-3203\(86\)90040-3](https://doi.org/10.1016/0031-3203(86)90040-3)
- Ojala, T., Pietikainen, M., Maenpaa, T. (2002). *Multiresolution gray-scale and rotation invariant texture classification with local binary patterns*. IEEE Trans. Pattern Anal. Machine Intell.
- OpenCV 3.1.0. (n.d.). Retrieved November 7, 2016, from <http://code.opencv.org/projects/opencv>

Osada, R., Funkhouser, T., Chazelle, B., & Dobkin, D. (2002). Shape distributions. *ACM Transactions on Graphics*, 21(4), 807–832.

<https://doi.org/10.1145/571647.571648>

Perona, P., & Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7), 629–639. <https://doi.org/10.1109/34.56205>

Pill information service, Thai food and drug administration. (2016). Retrieved November 7, 2016, from

<http://fdaolap.fda.moph.go.th/logistics/drgdrug/Dserch.asp>

“Pillbox” National Library of Medicine, National Institutes of Health, United States. (2016). Retrieved November 14, 2016, from <http://pillbox.nlm.nih.gov/>

Reddy, K. S., & Reddy, D. K. R. L. (2013). Enlargement of Image Based Upon Interpolation Techniques. *IJASCSE*, 2(12), 4631–4640.

RxList pill identification tool. (2016). Retrieved November 7, 2016, from <http://www.rxlist.com/pill-identification-tool/article.htm>

Smith, R. (2007). An Overview of the Tesseract OCR Engine (pp. 629–633). IEEE. <https://doi.org/10.1109/ICDAR.2007.4376991>

T Kasar, J Kumar, & A G Ramakrishnan. (2007). Font and background color independent text binarization. <https://doi.org/10.13140/RG.2.1.1349.4880>

Tomasi, C., & Manduchi, R. (1998). Bilateral filtering for gray and color images (pp. 839–846). Narosa Publishing House. <https://doi.org/10.1109/ICCV.1998.710815>

Wang, C., Kamata, S., & Ma, L. (2013). A fast multi-view based specular removal approach for pill extraction (pp. 4126–4130). IEEE. <https://doi.org/10.1109/ICIP.2013.6738850>

WebMD pill identification tool. (2016). Retrieved November 7, 2016, from
<http://www.webmd.com/pill-identification/>.

YaAndYou. (2016). Retrieved November 7, 2016, from
<https://itunes.apple.com/th/app/yaandyou/id459400481?mt=8>

Yu, J., Chen, Z., & Kamata, S. (2014). Pill Recognition Using Imprint Information by
Two-Step Sampling Distance Sets (pp. 3156–3161). IEEE.
<https://doi.org/10.1109/ICPR.2014.544>



ประวัติผู้เขียน

ชื่อ	นางสาวศิริรัตน์ สุนทรสุข
วันเดือนปีเกิด	28 สิงหาคม พ.ศ. 2535
วุฒิการศึกษา	ปีการศึกษา 2553: วิทยาศาสตร์บัณฑิต (วิทยาการคอมพิวเตอร์) มหาวิทยาลัยธรรมศาสตร์
ตำแหน่ง	นักพัฒนาซอฟต์แวร์ บริษัท ดีเอสที เวิลด์วายด์ เซอร์วิส (ประเทศไทย) จำกัด
ทุนการศึกษา	ปีงบประมาณ 2557: ทุนบัณฑิตเรียนดีเพื่อศึกษาต่อระดับ บัณฑิตศึกษา คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์

ผลงานทางวิชาการ

Suntronsuk, S., & Ratanotayanon, S. (2016). Pill Image Binarization for Detecting Text Imprints (Vol. 13). Presented at the International Joint Conference on Computer Science and Software Engineering, Thailand: IEEE.

Suntronsuk, S., & Ratanotayanon, S. (2017). Automatic Text Imprint Analysis from Pill Images (Vol. 9). Presented at the International Conference on Knowledge and Smart Technology, Thailand: IEEE.

ประสบการณ์ทำงาน	2557-2558: นักพัฒนาซอฟต์แวร์(JAVA,C,COBOL) บริษัท ดีเอสที เวิลด์วายด์ เซอร์วิส (ประเทศไทย) จำกัด
-----------------	---