การศึกษาปัจจัยที่มีผลต่อการรู้จำตัวอักษรสำหรับเอกสารที่พิมพ์ด้วยเครื่องพิมพ์ดอทเมทริกซ์

A Study of Factors Effecting on Character Recognition of The Dot Matrix Printed Documents.

ปิยะนาถ เคว็จดำ

Piyanart Queddum

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญา
วิทยาศาสตรมหาบัณฑิต สาขาวิชาการจัดการเทคโนโลยีสารสนเทศ
มหาวิทยาลัยสงขลานครินทร์

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Management of Information Technology
Prince of Songkla University

2559

ลิขสิทธิ์ของมหาวิทยาลัยสงขลานครินทร์

การศึกษาปัจจัยที่มีผลต่อการรู้จำตัวอักษรสำหรับเอกสารที่พิมพ์ด้วยเครื่องพิมพ์ดอทเมทริกซ์

A Study of Factors Effecting on Character Recognition of The Dot Matrix Printed Documents.

ปิยะนาถ เคว็จดำ
Piyanart Queddum

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญา
วิทยาศาสตรมหาบัณฑิต สาขาวิชาการจัดการเทคโนโลยีสารสนเทศ
มหาวิทยาลัยสงขลานครินทร์
A Thesis Submitted in Partial Fulfillment of the Requirements for the
Degree of Master of Science in Management of Information Technology
Prince of Songkla University
2559

ลิขสิทธิ์ของมหาวิทยาลัยสงขลานครินทร์

ชื่อวิทยานิพนธ์ การศึกษาปัจจัยที่มีผลต่อการรู้จำตัวอักษรสำหรับเอกสารที่พิมพ์ด้วยเครื่องพิมพ์ดอทเมทริกซ์
ผู้เขียน ปิยะนาถ เคว็จดำ
สาขาวิชา การจัดการเทคโนโลยีสารสนเทศ

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

.....
(ดร.สมชัย หลิมศิโรรัตน์)

คณะกรรมการสอบ

.....ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.สุนทร วิทูสุรพจน์)

.....กรรมการ
(ดร.สมชัย หลิมศิโรรัตน์)

.....กรรมการ
(ดร.อนันท์ ชกสุริวงค์)

.....กรรมการ
(ดร.เดือนเพ็ญ กชกรจารุพงศ์)

บัณฑิตวิทยาลัย มหาวิทยาลัยสงขลานครินทร์ อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษา ตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาการจัดการเทคโนโลยีสารสนเทศ

.....
(รองศาสตราจารย์ ดร.ธีระพล ศรีชนะ)
คณบดีบัณฑิตวิทยาลัย

ขอรับรองว่า ผลงานวิจัยนี้มาจากการศึกษาวิจัยของนักศึกษาเอง และได้แสดงความขอบคุณบุคคลที่มีส่วนช่วยเหลือแล้ว

ลงชื่อ.....
(ดร.สมชัย หลิมศิโรรัตน์)
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

ลงชื่อ.....
(นางสาวปิยะนาถ เคว็จดำ)
นักศึกษา

ข้าพเจ้าขอรับรองว่า ผลงานวิจัยนี้ไม่เคยเป็นส่วนหนึ่งในการอนุมัติปริญญาในระดับใดมาก่อน และไม่ได้ถูกใช้ในการยื่นขออนุมัติปริญญาในขณะนี้

ลงชื่อ.....
(นางสาวปิยะนาถ เคว็จดำ)
นักศึกษา

ชื่อวิทยานิพนธ์ การศึกษาปัจจัยที่มีผลต่อการรู้จำตัวอักษรสำหรับเอกสารที่พิมพ์ด้วยเครื่องพิมพ์ดอทเมทริกซ์
ผู้เขียน ปิยะนาถ เคว็จดำ
สาขาวิชา การจัดการเทคโนโลยีสารสนเทศ
ปีการศึกษา 2558

บทคัดย่อ

เอกสารที่พิมพ์ด้วยเครื่องพิมพ์ดอทเมทริกซ์ ยังมีความสำคัญและใช้อย่างแพร่หลาย เนื่องจากประโยชน์ของจำนวนชั้นของกระดาษที่พิมพ์ได้ครั้งละหลายชั้น ทำให้ไม่เสียเวลา และประหยัด โดยเฉพาะสำหรับเอกสารที่เป็นแบบฟอร์ม เช่น ใบกำกับสินค้า ใบส่งของ ใบเสร็จรับเงิน เมื่อต้องการนำข้อมูลกลับมาใช้ในธุรกิจ ต้องใช้คนป้อนข้อมูลทำให้ควบคุมได้ยาก เนื่องจากความเหนื่อยล้า หรือการรู้จำอักขระด้วยแสงซึ่งเป็นระบบอัตโนมัติที่ดีกว่า แต่ยังมีข้อผิดพลาดเนื่องจากหลายปัจจัย วิทยานิพนธ์นี้ได้ศึกษาปัจจัยที่มีผลกระทบต่อการรู้จำตัวอักษรของกระดาษต่อเนื่องชนิดคาร์บอน และชนิดเคมี โดยศึกษาจากปัจจัยชนิดแบบอักษร ขนาดตัวอักษร ลักษณะตัวอักษร ความละเอียดในการพิมพ์ ความละเอียดและชนิดในการกราดภาพ เพื่อให้ได้มาซึ่งประสิทธิภาพ และประสิทธิผล นำไปสู่การออกแบบฟอร์มที่ดีที่สุด เหมาะสม มีประสิทธิผลที่สุด สำหรับธุรกิจ หลังจากเก็บสถิติการรู้จำมาแล้วพบว่า การรู้จำตัวอักษร ชนิดแบบอักษร Cordia New มีความแม่นยำสูงสุด ส่วนการรู้จำตัวเลข ชนิดแบบอักษร Angsana New มีความแม่นยำสูงสุด แต่ชนิดแบบอักษร TH SarabunPSK มีความผิดพลาดสูงสุดทั้งการรู้จำตัวอักษรและตัวเลข และควรเลือกใช้กระดาษชนิดเคมี

คำสำคัญ: การรู้จำ การรู้จำอักขระด้วยแสง ดอทเมทริกซ์

Thesis Title A Study of Factors Effecting on Character Recognition of The Dot Matrix Printed Documents
Auther Piyanart Queddum
Major Program Management Information Technology
Academic Year 2015

ABSTRACT

Documents printed with dot matrix printers also important and widely used. Because of the number of layers of paper to print multiple layers. This do not waste time and money, especially for documents that are forms such as invoices invoice receipt. When extract data back into business process, the input makes it hard to control due to fatigue or optical character recognition, which is automatically better but also has errors due to many factors. This thesis has studied the factors affecting on the character recognition of carbon and chemical continuous papers. Study from the font, font size, font style, printed resolution and scanned resolutions to obtain performance and effectiveness to best form suited for the most productive business. After recognition statistics report, the highest accuracy of font type on text recognition is the Cordia New, on numberic recognition is the Angsana New and the TH SarabunPSK has lowest recognition on both texts and numbers recognition. And should use chemicals continuous papers.

Keywords: Recognition, Optical Character Recognition, Dot Matrix

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สาเร็จลุล่วงได้ด้วยความกรุณาและคำชี้แนะอย่างดียิ่งจาก ดร.สมชัย หลิมศิโรรัตน์ อาจารย์ที่ปรึกษาหลัก ที่ได้ให้คำปรึกษาและแนะแนวทางในการดำเนินการวิจัย ตลอดจนช่วยตรวจสอบแก้ไขข้อบกพร่องต่างๆ ของวิทยานิพนธ์จนเสร็จสมบูรณ์ ด้วยความเอาใจใส่ และขอขอบพระคุณคณะกรรมการสอบทุกท่าน ได้แก่ ผู้ช่วยศาสตราจารย์ ดร.สุนทร วิทูสุรพจน์ ดร.อนันท์ ชกสุริวงค์ และดร.เดือนเพ็ญ กชกรจารุพงศ์ ที่สละเวลามาเป็นกรรมการสอบและกรุณาให้คำแนะนำตรวจทานเพื่อการแก้ไขวิทยานิพนธ์ ให้มีความสมบูรณ์ยิ่งขึ้น

สุดท้ายนี้ ขอขอบพระคุณ คุณพ่อโกวิทย์ คุณแม่สุกัลยา คุณปิยะวุฒิ เคว็จดำ พี่ชาย ผู้ที่คอยให้ความช่วยเหลือสนับสนุน ให้คำปรึกษา ให้กำลังใจด้วยความรัก ด้วยความห่วงใยต่อผู้วิจัยเสมอมา ตลอดจนผู้ที่ไม่ได้เอ่ยนามในที่นี้ทุกท่านที่ได้ให้ความช่วยเหลือ จนทำให้วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปด้วยดี ผู้วิจัยรู้สึกซาบซึ้ง และขอบพระคุณทุกท่านมา ณ ที่นี้ด้วย

ปิยะนาถ เคว็จดำ

สารบัญ

หน้า

บทคัดย่อ ..... **(5)**
ABSTRACT ..... **(6)**
กิตติกรรมประกาศ ..... **(7)**
สารบัญ ..... **(8)**
รายการตาราง ..... **(10)**
รายการภาพประกอบ ..... **(11)**

บทที่ 1 บทนำ **1**
1.1 ความเป็นมาและความสำคัญของปัญหา 1
1.2 วัตถุประสงค์ของงานวิจัย 4
1.3 ประโยชน์ที่คาดว่าจะได้รับ 4
1.4 ขอบเขตของการวิจัย 4

บทที่ 2 ทฤษฎีและหลักการ **5**
2.1 ความรู้พื้นฐานเกี่ยวกับการรู้จำอักขระด้วยแสง 5
2.2 ปัจจัยที่มีผลต่อการรู้จำอักขระด้วยแสง 6
2.3 การหมุนภาพ และการถอดข้อความ 14
2.4 การปรับปรุงภาพ และการขจัดสัญญาณรบกวน 14
2.5 การเลือกใช้โปรแกรมการรู้จำอักขระด้วยแสง 15
2.6 งานวิจัยที่เกี่ยวข้อง 16

บทที่ 3 วิธีการดำเนินงาน **18**
3.1 เตรียมเอกสารตัวอย่าง 18
3.2 การรู้จำ 20
3.3 วัดผลการรู้จำ 25
3.4 หาค่าความถูกต้องและค่าความผิดพลาดของการรู้จำแยกตามปัจจัย 30

บทที่ 4 ผลการดำเนินงาน **32**
4.1 ผลการรู้จำโดยรวม 32
4.2 ผลการรู้จำเฉพาะอักษรไทย 38
4.3 ผลการรู้จำเฉพาะอักษรอังกฤษ 44
4.4 ผลการรู้จำเฉพาะเลขไทย 50

สารบัญ (ต่อ)

หน้า

4.5 ผลการรู้จำเฉพาะเลขอารบิค 56
4.6 ปัจจัยที่ทำให้ผลการรู้จำดีขึ้นได้มากที่สุด 62

บทที่ 5 บทสรุปและข้อเสนอแนะ 66
5.1 สรุปผลการวิจัย 66
5.2 ประโยชน์ที่ได้จากงานวิจัย 67
5.3 ข้อเสนอแนะในการดำเนินการวิจัย 67

บรรณานุกรม 68
ภาคผนวก 70
ภาคผนวก ก คู่มือการใช้โปรแกรม ABBYY Fine Reader 11 71
ภาคผนวก ข คู่มือการใช้โปรแกรม Check Character 78
ภาคผนวก ค เปรียบเทียบผลการรู้จำ 84
ภาคผนวก ง ผลงานตีพิมพ์และเผยแพร่ 95
ประวัติผู้เขียน 102

รายการตาราง

หน้า

ตาราง 1-1 ผลการรู้จำอักขระด้วยแสงของภาพเอกสารที่พิมพ์ด้วยเครื่องพิมพ์เลเซอร์ 2
ตาราง 1-2 ผลการรู้จำอักขระด้วยแสงของภาพเอกสารที่พิมพ์ด้วยเครื่องพิมพ์ดอทเมทริกซ์ 3
ตาราง 2-1 ข้อดี/ข้อเสีย ของโปรแกรมการรู้จำอักขระด้วยแสง 15
ตาราง 4-1 ผลการรู้จำโดยรวม 32
ตาราง 4-2 ผลการรู้จำเฉพาะอักษรไทย 38
ตาราง 4-3 ตัวอย่างการรู้จำที่ผิดพลาดเฉพาะอักษรไทย 43
ตาราง 4-4 ผลการรู้จำเฉพาะอักษรอังกฤษ 44
ตาราง 4-5 ตัวอย่างการรู้จำที่ผิดพลาดเฉพาะอักษรอังกฤษ 49
ตาราง 4-6 ผลการรู้จำเฉพาะเลขไทย 50
ตาราง 4-7 ตัวอย่างการรู้จำที่ผิดพลาดเฉพาะเลขไทย 55
ตาราง 4-8 ผลการรู้จำเฉพาะเลขอารบิค 56
ตาราง 4-9 ตัวอย่างการรู้จำที่ผิดพลาดเฉพาะเลขอารบิค 61
ตาราง 4-10 เปรียบเทียบผลการรู้จำโดยรวมกับค่าเฉลี่ย 62
ตาราง ค-1 ผลการรู้จำเอกสารชนิดคาร์บอน ตามชนิดแบบอักษร 85
ตาราง ค-2 ผลการรู้จำเอกสารชนิดเคมี ตามชนิดแบบอักษร 86
ตาราง ค-3 ผลการรู้จำเอกสารชนิดคาร์บอน ตามขนาดตัวอักษร 87
ตาราง ค-4 ผลการรู้จำเอกสารชนิดเคมี ตามขนาดตัวอักษร 88
ตาราง ค-5 ผลการรู้จำเอกสารชนิดคาร์บอน ตามลักษณะตัวอักษร 89
ตาราง ค-6 ผลการรู้จำเอกสารชนิดเคมี ตามลักษณะตัวอักษร 90
ตาราง ค-7 ผลการรู้จำเอกสารชนิดคาร์บอน ตามความละเอียดในการพิมพ์ 91
ตาราง ค-8 ผลการรู้จำเอกสารชนิดเคมี ตามความละเอียดในการพิมพ์ 92
ตาราง ค-9 ผลการรู้จำเอกสารชนิดคาร์บอน ตามความละเอียดและชนิดในการกราดภาพ 93
ตาราง ค-10 ผลการรู้จำเอกสารชนิดเคมี ตามความละเอียดและชนิดในการกราดภาพ 94

รายการภาพประกอบ

หน้า

ภาพประกอบ 1-1 ตัวอย่างไฟล์ภาพเอกสารดอทเมทริกซ์ที่ไม่สมบูรณ์ 1
ภาพประกอบ 1-2 ปัญหาของการรู้จำอักขระด้วยแสง จากไฟล์ภาพที่ไม่สมบูรณ์ 2
ภาพประกอบ 2-1 กระบวนการทำงานของการรู้จำอักขระด้วยแสง 5
ภาพประกอบ 2-2 ตัวอย่างการรู้จำเอกสารพื้นหลังสี กราดภาพ 600 DPI GRAY 6
ภาพประกอบ 2-3 ตัวอย่างการรู้จำเอกสารพื้นหลังสี กราดภาพ 300 DPI COLOR 7
ภาพประกอบ 2-4 ตัวอย่างการรู้จำเอกสาร กราดภาพ 600 DPI COLOR 8
ภาพประกอบ 2-5 ตัวอย่างการรู้จำเอกสาร กราดภาพ 600 DPI GRAY 9
ภาพประกอบ 2-6 ตัวอย่างการรู้จำเอกสารสำเนา กราดภาพ 600 DPI COLOR 10
ภาพประกอบ 2-7 ตัวอักษรแบบมีเชิง 11
ภาพประกอบ 2-8 ตัวอักษรแบบไม่มีเชิง 11
ภาพประกอบ 2-9 ชนิดแบบอักษรที่ใช้ในการวิจัย 11
ภาพประกอบ 2-10 โปรแกรม ABBYY FINE READER 11 16
ภาพประกอบ 3-1 การออกแบบเอกสารตัวอย่าง 19
ภาพประกอบ 3-2 เอกสารตัวอย่าง 20
ภาพประกอบ 3-3 กระบวนการทำงานของโปรแกรม ABBYY FINE READER 11 21
ภาพประกอบ 3-4 ตัวอย่างการหมุนกระดาษผิด 22
ภาพประกอบ 3-5 ตัวอย่างการหมุนกระดาษผิดแล้วแปลงเป็นรูปภาพ 22
ภาพประกอบ 3-6 ตัวอย่างการหมุนกระดาษผิดแล้วแปลงเป็นรูปภาพ และตัวอักษร 23
ภาพประกอบ 3-7 ตัวอย่างการหมุนกระดาษผิดแล้วแปลงเป็นรูปภาพ ตาราง และตัวอักษร 24
ภาพประกอบ 3-8 ตัวอย่างการแบ่งหน้ากระดาษเป็น 2 หน้า แล้วแปลงเป็นรูปภาพ 25
ภาพประกอบ 3-9 กระบวนการทำงานของโปรแกรม CHECK CHARACTER 26
ภาพประกอบ 3-10 หน้าต่าง USER CHANGE INDEX ก่อนปรับปรุง 26
ภาพประกอบ 3-11 ตัวอย่างช่องว่างมาก 27
ภาพประกอบ 3-12 ตัวอย่างตัวอักษรซ้ำ 27
ภาพประกอบ 3-13 ตัวอย่างตัวอักษรขาดหาย 27
ภาพประกอบ 3-14 ตัวอย่างตัวอักษรเกิน 28
ภาพประกอบ 3-15 ตัวอย่างไม่มีตัวอักษร 28
ภาพประกอบ 3-16 ตัวอย่างตัวอักษรเกินและช่องว่างมาก 29
ภาพประกอบ 3-17 หน้าต่าง USER CHANGE INDEX หลังปรับปรุง 29
ภาพประกอบ 3-18 ตัวอย่างที่สามารถเลื่อนนับได้หลังปรับปรุง 30

รายการภาพประกอบ (ต่อ)

หน้า

ภาพประกอบ 3-19 การนำไฟล์ .csv มาหาค่าความถูกต้อง/ผิดพลาดแยกตามปัจจัย 30
ภาพประกอบ 4-1 ผลการรู้จำโดยรวม แยกตามชนิดแบบอักษร 33
ภาพประกอบ 4-2 ผลการรู้จำโดยรวม แยกตามขนาดตัวอักษร 34
ภาพประกอบ 4-3 ผลการรู้จำโดยรวม แยกตามลักษณะตัวอักษร 34
ภาพประกอบ 4-4 ผลการรู้จำโดยรวม แยกตามความละเอียดในการพิมพ์ 35
ภาพประกอบ 4-5 ผลการรู้จำโดยรวม แยกตามความละเอียดและชนิดในการกราดภาพ 36
ภาพประกอบ 4-6 ผลการรู้จำโดยรวมเฉลี่ย 36
ภาพประกอบ 4-7 ตัวอย่างภาพเอกสารสำเนาชนิดคาร์บอน 37
ภาพประกอบ 4-8 ผลการรู้จำเฉพาะอักษรไทย แยกตามชนิดแบบอักษร 39
ภาพประกอบ 4-9 ผลการรู้จำเฉพาะอักษรไทย แยกตามขนาดตัวอักษร 40
ภาพประกอบ 4-10 ผลการรู้จำเฉพาะอักษรไทย แยกตามลักษณะตัวอักษร 40
ภาพประกอบ 4-11 ผลการรู้จำเฉพาะอักษรไทย แยกตามความละเอียดในการพิมพ์ 41
ภาพประกอบ 4-12 ผลการรู้จำเฉพาะอักษรไทย แยกตามความละเอียดและชนิดในการกราดภาพ . 42
ภาพประกอบ 4-13 ผลการรู้จำเฉพาะอักษรไทยเฉลี่ย 42
ภาพประกอบ 4-14 ผลการรู้จำเฉพาะอักษรอังกฤษ แยกตามชนิดแบบอักษร 45
ภาพประกอบ 4-15 ผลการรู้จำเฉพาะอักษรอังกฤษ แยกตามขนาดตัวอักษร 46
ภาพประกอบ 4-16 ผลการรู้จำเฉพาะอักษรอังกฤษ แยกตามลักษณะตัวอักษร 46
ภาพประกอบ 4-17 ผลการรู้จำเฉพาะอักษรอังกฤษ แยกตามความละเอียดในการพิมพ์ 47
ภาพประกอบ 4-18 ผลการรู้จำเฉพาะอักษรอังกฤษ แยกตามความละเอียดและชนิดในการกราดภาพ. 48
ภาพประกอบ 4-19 ผลการรู้จำเฉพาะอักษรอังกฤษเฉลี่ย 48
ภาพประกอบ 4-20 ผลการรู้จำเฉพาะเลขไทย แยกตามชนิดแบบอักษร 51
ภาพประกอบ 4-21 ผลการรู้จำเฉพาะเลขไทย แยกตามขนาดตัวอักษร 52
ภาพประกอบ 4-22 ผลการรู้จำเฉพาะเลขไทย แยกตามลักษณะตัวอักษร 52
ภาพประกอบ 4-23 ผลการรู้จำเฉพาะเลขไทย แยกตามความละเอียดในการพิมพ์ 53
ภาพประกอบ 4-24 ผลการรู้จำเฉพาะเลขไทย แยกตามความละเอียดและชนิดในการกราดภาพ 54
ภาพประกอบ 4-25 ผลการรู้จำเฉพาะเลขไทยเฉลี่ย 54
ภาพประกอบ 4-26 ผลการรู้จำเฉพาะเลขอารบิค แยกตามชนิดแบบอักษร 57

รายการภาพประกอบ (ต่อ)

หน้า

ภาพประกอบ 4-27 ผลการรู้จำเฉพาะเลขอารบิค แยกตามขนาดตัวอักษร 58
ภาพประกอบ 4-28 ผลการรู้จำเฉพาะเลขอารบิค แยกตามลักษณะตัวอักษร 58
ภาพประกอบ 4-29 ผลการรู้จำเฉพาะเลขอารบิค แยกตามความละเอียดในการพิมพ์ 59
ภาพประกอบ 4-30 ผลการรู้จำเฉพาะเลขอารบิค แยกตามความละเอียดและชนิดในการกราดภาพ... 60
ภาพประกอบ 4-31 ผลการรู้จำเฉพาะเลขอารบิคเฉลี่ย 60
ภาพประกอบ 4-32 ตัวอย่างการออกแบบฟอร์มที่ดี 63
ภาพประกอบ ก-1 หน้าต่างโปรแกรม ABBYY FINE READER 11 71
ภาพประกอบ ก-2 หน้าต่าง OPEN IMAGE 72
ภาพประกอบ ก-3 หน้าต่างการแปลงไฟล์รูปภาพเป็นตัวอักษร 72
ภาพประกอบ ก-4 หน้าต่างการบันทึกไฟล์ .TXT 73
ภาพประกอบ ก-5 เปิดไฟล์เอกสารที่ต้องการรู้จำ 73
ภาพประกอบ ก-6 การรู้จำและบันทึกไฟล์อัตโนมัติ 74
ภาพประกอบ ก-7 โปรแกรมเลือกการวิเคราะห์ประเภทให้อัตโนมัติ 74
ภาพประกอบ ก-8 การลบพื้นที่ที่ไม่ต้องการ 75
ภาพประกอบ ก-9 การเพิ่มพื้นที่ที่ต้องการ 75
ภาพประกอบ ก-10 การกำหนดพื้นที่การแปลงไฟล์ให้วิเคราะห์ตามที่ต้องการ 76
ภาพประกอบ ก-11 การกำหนดคุณสมบัติให้กับพื้นที่การแปลงไฟล์ 76
ภาพประกอบ ก-12 การบันทึกแบบฟอร์มเฉพาะ 77
ภาพประกอบ ข-1 หน้าต่างโปรแกรม CHECK CHARACTER 78
ภาพประกอบ ข-2 หน้าต่าง COUNT FORM 79
ภาพประกอบ ข-3 เมนู OPEN เปิดไฟล์ .TXT ที่ต้องการ 79
ภาพประกอบ ข-4 เปิดไฟล์ .TXT ที่ต้องการ 80
ภาพประกอบ ข-5 การนับคำถูก/ผิด 80
ภาพประกอบ ข-6 การเลื่อนนับตัวอักษรถูก/ผิด 81
ภาพประกอบ ข-7 การเปรียบเทียบตัวอักษร 81
ภาพประกอบ ข-8 การแสดงผลการนับตัวอักษรที่ถูก 82
ภาพประกอบ ข-9 ปุ่มบันทึกผลการนับตัวอักษรที่ถูก 82
ภาพประกอบ ข-10 บันทึกการวัดผลการรู้จำ 83
ภาพประกอบ ค-1 การตั้งชื่อไฟล์เอกสาร 84
ภาพประกอบ ง-1 เกียรติบัตรรับรองการนำเสนอผลงาน 95
ภาพประกอบ ง-2 หน้าปกหนังสือเรื่องเต็มประชุมวิชาการ 96
ภาพประกอบ ง-3 เนื้อหาบทความวิชาการหน้า 117 97

รายการภาพประกอบ (ต่อ)

หน้า

ภาพประกอบ ง-4 เนื้อหาบทความวิชาการหน้า 118 98
ภาพประกอบ ง-5 เนื้อหาบทความวิชาการหน้า 119 99
ภาพประกอบ ง-6 เนื้อหาบทความวิชาการหน้า 120 100
ภาพประกอบ ง-7 เนื้อหาบทความวิชาการหน้า 121 101

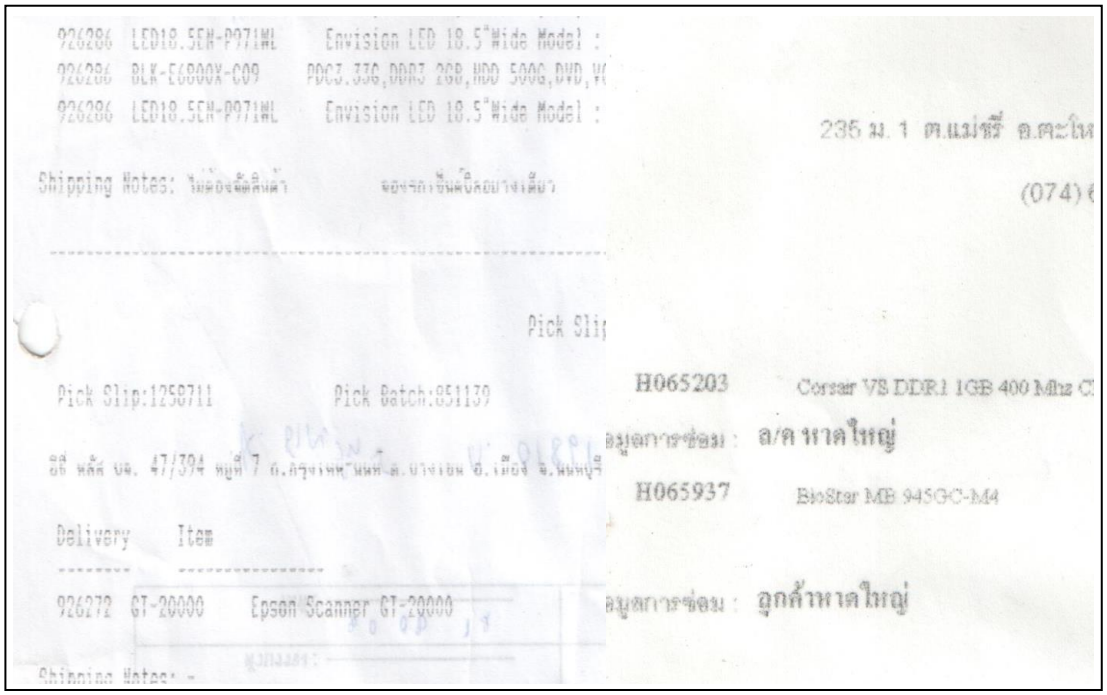
บทที่ 1

บทนำ

บทนี้จะกล่าวถึงความเป็นมาและความสำคัญของปัญหาซึ่งเป็นที่มาของวิทยานิพนธ์นี้ รวมถึงวัตถุประสงค์ของงานวิจัย ประโยชน์ที่คาดว่าจะได้รับ ขอบเขตของงานวิจัยซึ่งได้แก่ ขอบเขตด้านประชากรและกลุ่มตัวอย่าง และขอบเขตด้านเครื่องมือการวิจัย

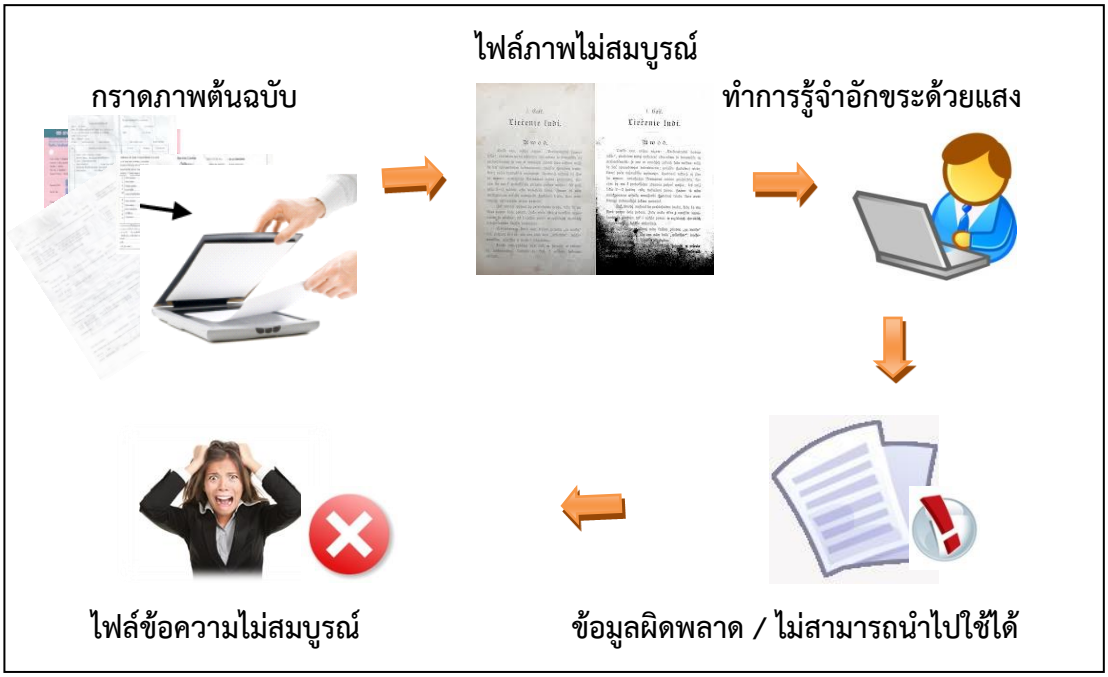
1.1 ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันการจัดการสารสนเทศมีความสำคัญมากสำหรับธุรกิจ ห้างร้าน หน่วยงานของรัฐ เอกชน และอื่นๆ แต่ในการณีที่ข้อมูลที่ได้รับมาจากหน่วยงานอื่นเป็นเอกสารกระดาษ จึงมีความจำเป็นต้องแปลงให้เป็นไฟล์เอกสาร โดยการกราดภาพแล้วแปลงให้เป็นไฟล์เอกสาร ด้วยวิธีการรู้จำอักขระด้วยแสง (OCR : Optical Character Recognition) เช่น ใบเสร็จรับเงิน ใบส่งของ ใบแสดงรายการสินค้าเข้า / สินค้าออก เอกสารต้นฉบับที่ได้มานั้นส่วนใหญ่จะเป็นกระดาษที่พิมพ์ด้วยเครื่องพิมพ์ดอทเมทริกซ์เนื่องจากสามารถพิมพ์ได้ครั้งละหลายแผ่น และมีราคาถูก แต่มักจะเกิดข้อผิดพลาดเมื่อนำไปรู้จำอักขระด้วยแสง เพราะจะมีลักษณะที่ไม่สมบูรณ์ เช่น ตัวอักษรติดกัน ตัวอักษรขาดเนื่องจากหัวเข็มหัก ไม่คมชัดเพราะผ้าหมึกจาง ภาพที่ได้จากการกราดภาพไม่ชัดเจนเนื่องจากการตั้งค่าไม่เหมาะสม คุณภาพของกระดาษ เป็นต้น ดังภาพประกอบ 1-1



ภาพประกอบ 1-1 ตัวอย่างไฟล์ภาพเอกสารดอทเมทริกซ์ที่ไม่สมบูรณ์

เมื่อนำข้อมูลกลับเข้าสู่กระบวนการธุรกิจ จะเกิดกรณีผิดพลาดต่างๆ ในแต่ละขั้นตอน ดังภาพประกอบ 1-2



ภาพประกอบ 1-2 ปัญหาของการรู้จำอักขระด้วยแสง จากไฟล์ภาพที่ไม่สมบูรณ์

ผู้วิจัยจึงได้มีการทำกรณีศึกษาเปรียบเทียบผลการรู้จำอักขระด้วยแสง ด้วยโปรแกรม ABBYY Fine Reader ระหว่างเอกสารที่พิมพ์ด้วยเครื่องพิมพ์เลเซอร์ เปรียบเทียบกับเอกสารที่พิมพ์ด้วยเครื่องพิมพ์ดอทเมทริกซ์ ดังตาราง 1-1 และตาราง 1-2

ตาราง 1-1 ผลการรู้จำอักขระด้วยแสงของภาพเอกสารที่พิมพ์ด้วยเครื่องพิมพ์เลเซอร์

ชนิดแบบอักษร	ขนาด	เอกสารเลเซอร์ กราดภาพ 600 dpi					
		ภาพเอกสาร		ค่าความถูกต้อง		ค่าความผิดพลาด	
		ตัวอักษร	จำนวน	จำนวน	%	จำนวน	%
Angsana New	8	a-z, 0-9	144	142	98.61	2	1.39
	11	a-z, 0-9	144	144	100.00	0	0.00
	14	a-z, 0-9	144	144	100.00	0	0.00
	16	a-z, 0-9	144	144	100.00	0	0.00
	18	a-z, 0-9	144	142	98.61	2	1.39
รวมทั้งสิ้น			**720**	**716**	**99.44**	**4**	**0.56**

ตาราง 1-2 ผลการรู้จำอักขระด้วยแสงของภาพเอกสารที่พิมพ์ด้วยเครื่องพิมพ์ดอทเมทริกซ์

ชนิดแบบอักษร	ขนาด	เอกสารดอทเมทริกซ์ กราดภาพ 600 dpi					
		ภาพเอกสาร		ค่าความถูกต้อง		ค่าความผิดพลาด	
		ตัวอักษร	จำนวน	จำนวน	%	จำนวน	%
Angsana New	8	a-z, 0-9	144	108	75.00	36	25.00
	11	a-z, 0-9	144	122	84.72	22	15.28
	14	a-z, 0-9	144	129	89.58	15	10.42
	16	a-z, 0-9	144	132	91.67	12	8.33
	18	a-z, 0-9	144	144	100.00	0	0.00
รวมทั้งสิ้น			**720**	**635**	**88.19**	**85**	**11.81**

จากผลการทดลองการรู้จำ เอกสารที่พิมพ์ด้วยเครื่องพิมพ์เลเซอร์จะมีความแม่นยำเฉลี่ย 99.44% ส่วนเอกสารดอทเมทริกซ์จะมีความแม่นยำเฉลี่ย 88.19% จะเห็นได้ว่าเอกสารที่พิมพ์ด้วยเครื่องพิมพ์เลเซอร์จะมีความแม่นยำมากกว่าเอกสารดอทเมทริกซ์ แต่ส่วนใหญ่เอกสารจะพิมพ์ด้วยเครื่องพิมพ์ดอทเมทริกซ์ เพราะในการพิมพ์เอกสารครั้งเดียวสามารถได้สำเนาหลายชั้นได้ ทำให้รวดเร็ว ประหยัดต้นทุน สะดวกในการทำงาน

จากการศึกษางานวิจัยเกี่ยวกับการรู้จำ จากแบบฟอร์มจะมีการทำมาบ้างแล้วแต่จะเป็นเอกสารที่พิมพ์ด้วยเครื่องพิมพ์เลเซอร์ เช่น งานวิจัยเกี่ยวกับการแก้ปัญหาการรู้จำ ที่ไม่แม่นยำซึ่งมีปัญหากับใบกำกับสินค้าที่มีแบบฟอร์มต่างกัน (T.A. Bayer และ H.U. Mogg-Schneider, ค.ศ. 1997) การออกแบบพัฒนาระบบรู้จำใบกำกับสินค้าภาษาจีน (Enrico Sorio, Alberto Bartoli, Giorgio Davanzo และ Eric, ค.ศ. 2012) ซึ่งเท่าที่สืบค้นงานวิจัยเมื่อปี พ.ศ. 2556 ไม่พบงานวิจัยที่ศึกษาจากแบบฟอร์มหลายๆแบบที่พิมพ์ด้วยเครื่องพิมพ์ดอทเมทริกซ์ซึ่งเป็นที่นิยมใช้มากกว่า

มีงานวิจัยที่เกี่ยวข้องกับการสกัดข้อมูลจากแบบฟอร์มที่แตกต่างกัน เช่น การสกัดบางรายการบนแบบฟอร์มตามโครงสร้างโมเดลและนำมาปรับปรุงด้วยฐานความรู้ที่สร้างไว้ (T.A. Bayer และ H.U. Mogg-Schneider, ค.ศ. 1997) วิธีการตรวจสอบอัตโนมัติและการแก้ไขข้อผิดพลาดของการรู้จำตัวอักษร โดยการใช้โดเมนฐานความรู้เกี่ยวกับชนิดของข้อมูล (Enrico Sorio, Alberto Bartoli, Giorgio Davanzo และ Eric, ค.ศ. 2012)

วิทยานิพนธ์นี้จะทำการศึกษาปัจจัยที่มีผลกระทบกับการรู้จำอักขระด้วยแสงเอกสารกระดาษที่พิมพ์ด้วยเครื่องพิมพ์ดอทเมทริกซ์ เช่น ชนิดตัวอักษร ขนาดตัวอักษร การตั้งค่าการกราดภาพ เป็นต้น เพื่อนำไปสู่แนวทางการออกแบบแบบฟอร์มที่ดีที่สุดให้สามารถรู้จำอักขระด้วยแสงได้ดีขึ้น และสกัดข้อมูลได้อย่างถูกต้องแม่นยำมากขึ้น

1.2 วัตถุประสงค์ของงานวิจัย

เพื่อศึกษาปัจจัยที่มีผลต่อความถูกต้องแม่นยำในการรู้จำอักขระด้วยแสง จากเอกสารดอทเมทริกซ์ และเสนอแนะแนวทางในการออกแบบฟอร์มที่ดี

1.3 ประโยชน์ที่คาดว่าจะได้รับ

1.3.1 เพื่อหลีกเลี่ยงปัจจัยที่มีผลกระทบต่อการรู้จำอักขระด้วยแสง

1.3.2 เพื่อเป็นแนวทางให้ธุรกิจสามารถเลือกใช้แบบฟอร์มที่ดีที่สุดเพื่อรองรับการรู้จำอักขระด้วยแสง

1.4 ขอบเขตของการวิจัย

1.4.1 ขอบเขตด้านประชากรและกลุ่มตัวอย่าง

ปัจจัยที่จะศึกษา ปัจจัยที่ผลกระทบกับการรู้จำอักขระด้วยแสง จากเอกสารดอทเมทริกซ์ เช่น ชนิดแบบอักษร ขนาดตัวอักษร ตัวหนา ตัวเอียง สีตัวอักษร สีเอกสาร จำนวนชั้นของแบบฟอร์ม ความละเอียดและชนิดในการกราดภาพ เป็นต้น

1.4.2 ขอบเขตด้านเครื่องมือการวิจัย

เครื่องมือที่ใช้ในการทดลอง มีดังนี้

1) อุปกรณ์ที่ใช้ในการวิจัย
   - เครื่องคอมพิวเตอร์แบบพกพา ตัวประมวลผล Intel Core i5 หน่วยความจำ 4 GB ฮาร์ดดิสก์ 500 GB
   - เครื่องพิมพ์ Epson LQ-310
   - เครื่องพิมพ์ HP Color Laser Jet Pro MFP M177fw
2) โปรแกรมและภาษาที่ใช้ในการวิจัย
   - ระบบปฏิบัติการ Windows
   - Microsoft Excel 2010
   - Microsoft Visual Studio C++ V.2012
   - Word Pad
   - ABBYY Fine Reader V.11

บทที่ 2

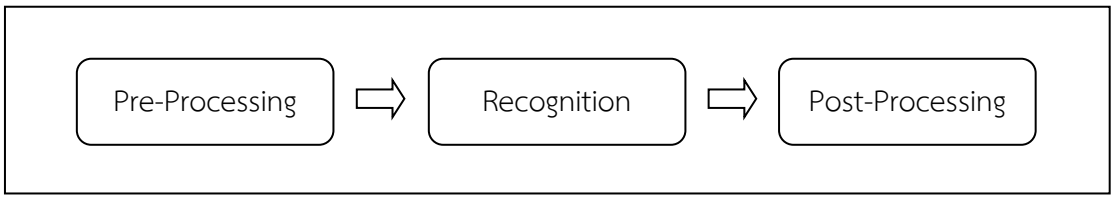
ทฤษฎีและหลักการ

บทนี้จะกล่าวถึงความรู้พื้นฐานของการรู้จำอักขระด้วยแสง ปัจจัยที่มีผลต่อการรู้จำอักขระด้วยแสง การหมุนภาพและการถอดข้อความ การปรับปรุงภาพและการขจัดสัญญาณรบกวน การเลือกใช้โปรแกรมรู้จำอักขระด้วยแสง และงานวิจัยที่เกี่ยวข้อง

2.1 ความรู้พื้นฐานเกี่ยวกับการรู้จำอักขระด้วยแสง

การรู้จำอักขระด้วยแสง หมายถึง การแปลงเอกสารที่อยู่ในรูปแบบของไฟล์ภาพให้เป็นเอกสารที่อยู่ในรูปแบบของไฟล์ข้อความโดยอัตโนมัติ ช่วยให้สะดวกต่อการจัดการงานด้านเอกสาร และการนำไฟล์ข้อความไปใช้

การรู้จำอักขระด้วยแสงมี 3 ขั้นตอน คือ การประมวลผลขั้นต้น (Pre-Processing), การรู้จำ (Recognition) และ การประมวลผลขั้นปลาย (Post-Processing) ตามลำดับ ดังภาพประกอบ 2-1 (วรพจน์ พีระวิทย์, พ.ศ. 2548)



ภาพประกอบ 2-1 กระบวนการทำงานของการรู้จำอักขระด้วยแสง

ขั้นตอนของการรู้จำอักขระด้วยแสง มี 3 ขั้นตอน (วรพจน์ พีระวิทย์, พ.ศ. 2548) คือ

2.1.1 การประมวลผลขั้นต้น

เป็นการจัดการกับไฟล์ภาพ ก่อนที่จะนำไปวิเคราะห์ข้อมูลตัวอักษร ให้มีความถูกต้องเพื่อการรู้จำมากที่สุด ซึ่งในขั้นตอนนี้จะเป็นการวิเคราะห์สิ่งแปลกปน สัญญาณรบกวน จุดดำ จุดขาว ที่ติดมาในภาพ เพื่อแยกแยะให้ได้ว่าไม่ใช่ตัวอักษร และไม่นำไปรู้จำ

2.1.2 การรู้จำ

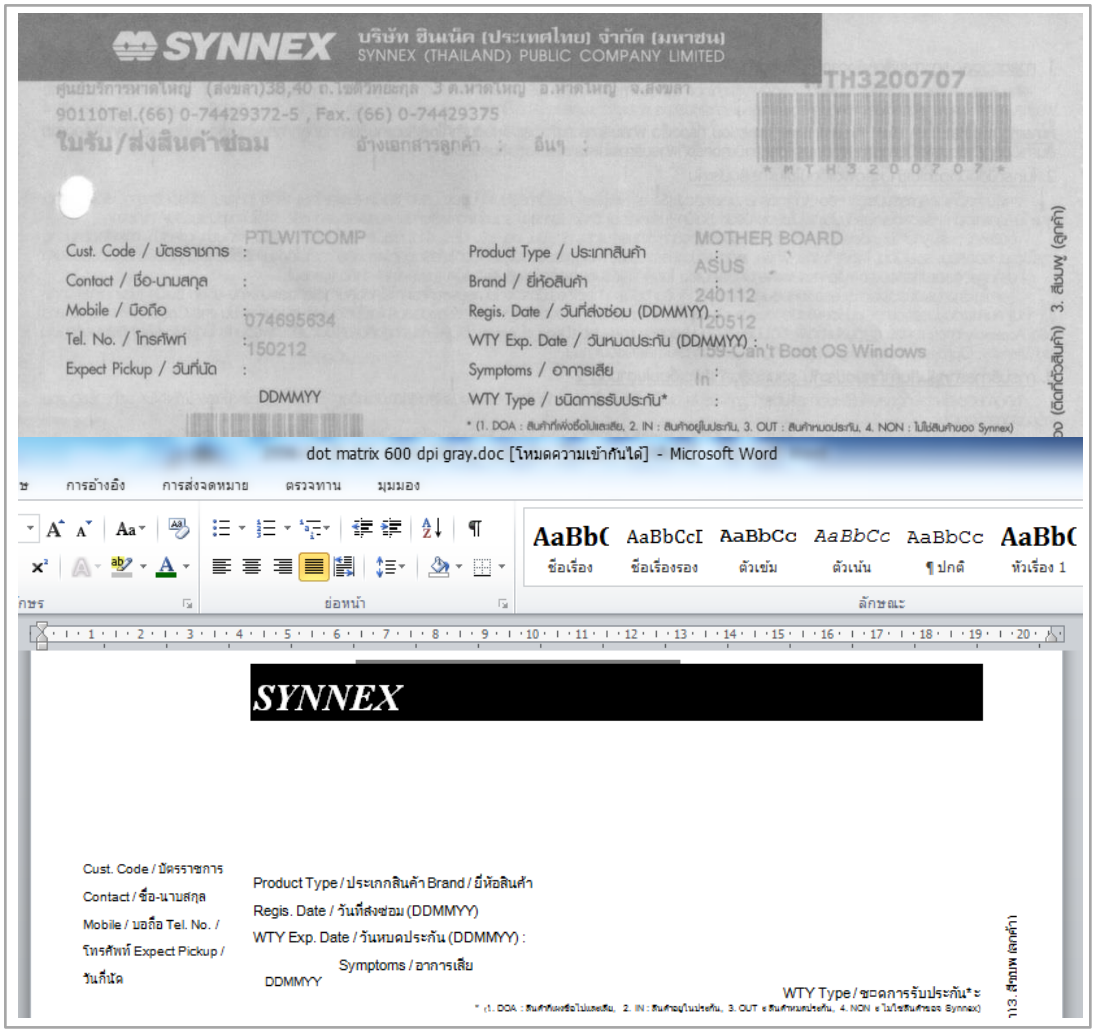
เป็นการวิเคราะห์ภาพ ที่ได้จากการประมวลผลขั้นต้นว่าเป็นตัวอักษรอะไร ทำให้เราได้มาซึ่งตัวอักษรหรือข้อความที่เราต้องการ ซึ่งมีวิธีในการรู้จำอยู่หลายวิธี เช่น การรู้จำโดยใช้หลักสถิติ การรู้จำโดยอาศัยโครงข่ายประสาทเทียม เป็นต้น

2.1.3 การประมวลผลขั้นปลาย

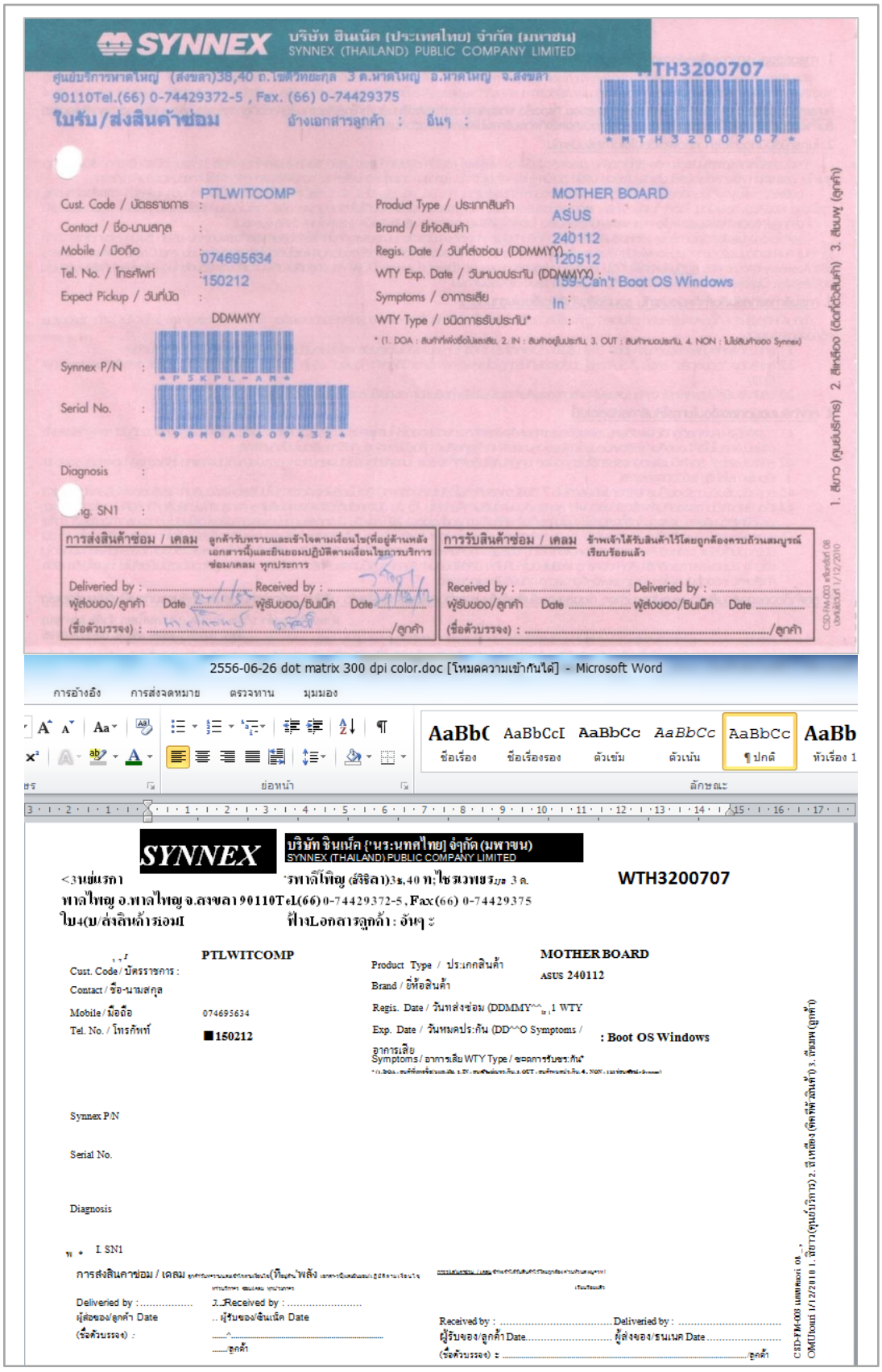
เป็นขั้นตอนของการนำรูปตัวอักษรที่ได้จากการรู้จำไปเข้ารหัสแสดงผลลัพธ์เป็นตัวอักษรออกมา ซึ่งในการรู้จำอาจจะไม่ถูกต้องทั้งหมด การประมวลผลหรือการตรวจสอบก็จะเพิ่มประสิทธิภาพในการทำงานของการรู้จำให้ทำงานได้ดีขึ้น ในการตรวจสอบนี้จะทำการแก้ไขความถูกต้องของภาษา ตัวสะกด ไวยากรณ์ การนำพจนานุกรมมาช่วยในการตรวจสอบคำศัพท์ หาคำผิดและแก้ไขให้ถูกต้อง หรือแสดงให้เห็นตัวอักษรที่ผิด หรือแก้ไขให้อัตโนมัติ หากในขั้นตอนของการรู้จำได้ข้อความไวยากรณ์ที่ผิดพลาด โปรแกรมในส่วนนี้ก็จะทำการตรวจสอบไวยากรณ์ที่ถูกต้อง

2.2 ปัจจัยที่มีผลต่อการรู้จำอักขระด้วยแสง

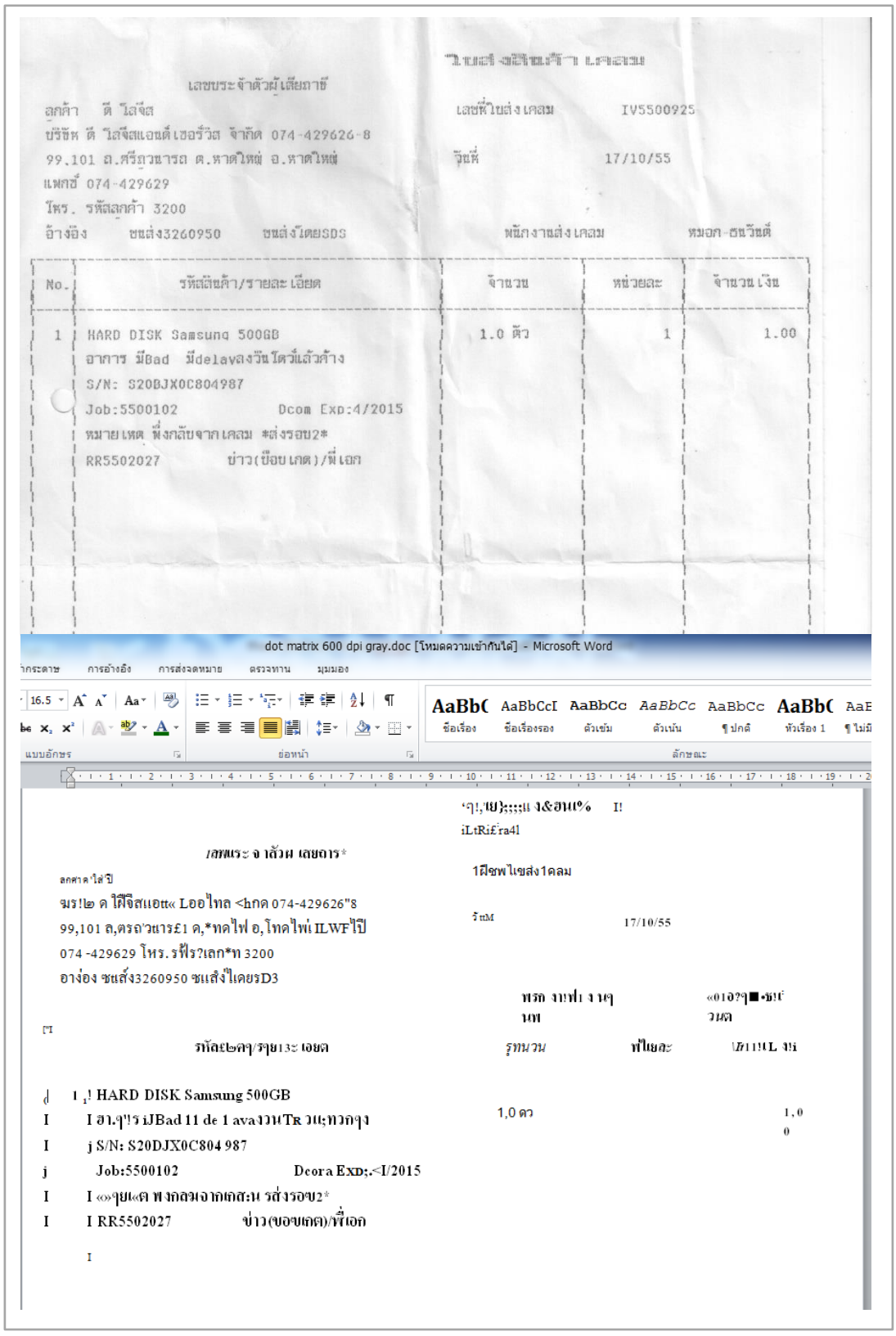
ผู้วิจัยได้ทำกรณีศึกษาผลการรู้จำอักขระด้วยแสงจากเอกสารดอทเมทริกซ์ ด้วยการนำเอกสารกระดาษมากราดภาพ แล้วรู้จำด้วยโปรแกรม ABBYY Fine Reader 11 แปลงเป็นไฟล์เอกสาร ตามตัวอย่างดังภาพประกอบ 2-2 ถึงภาพประกอบ 2-6



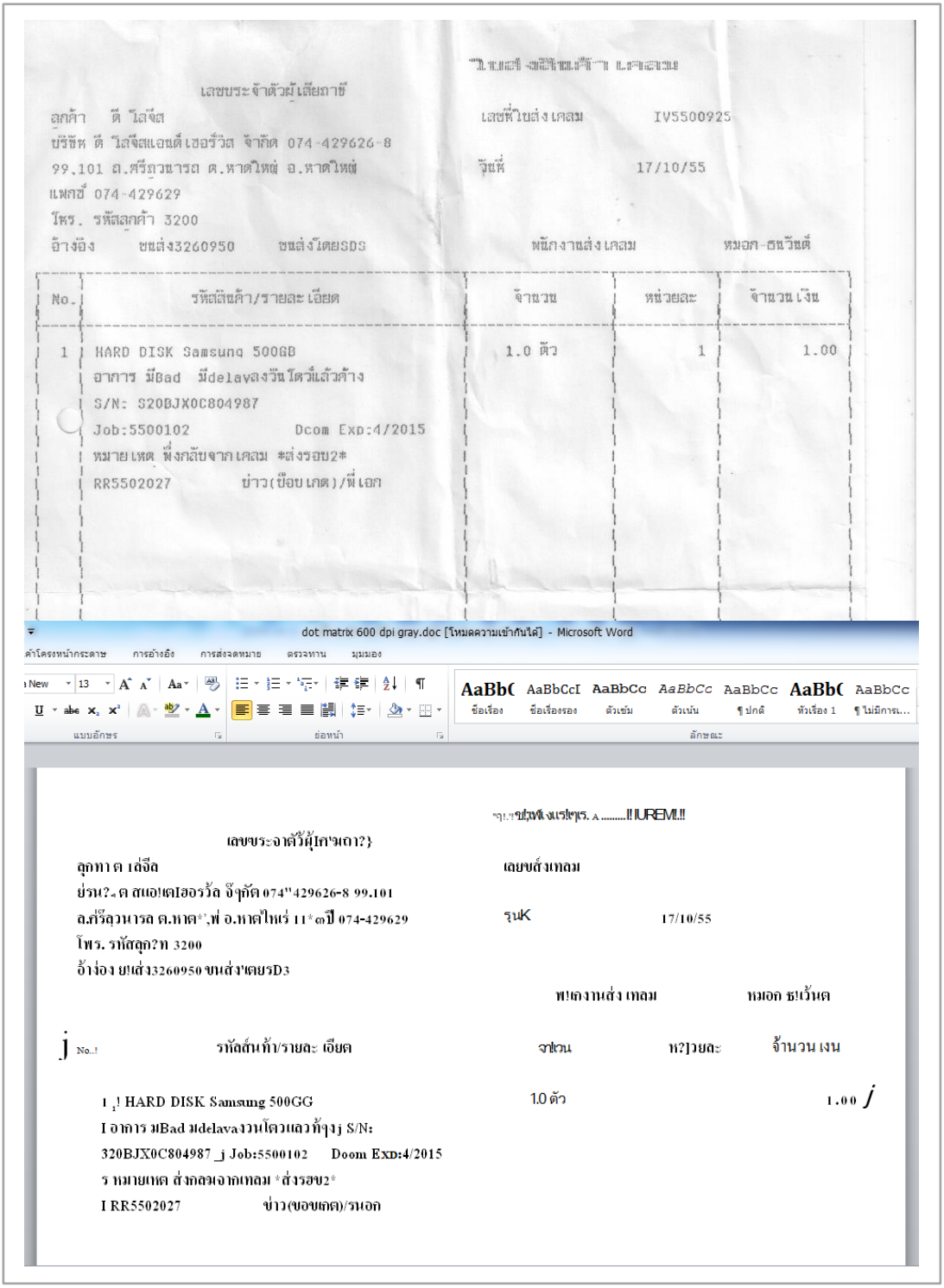
ภาพประกอบ 2-2 ตัวอย่างการรู้จำเอกสารพื้นหลังสี กราดภาพ 600 dpi Gray



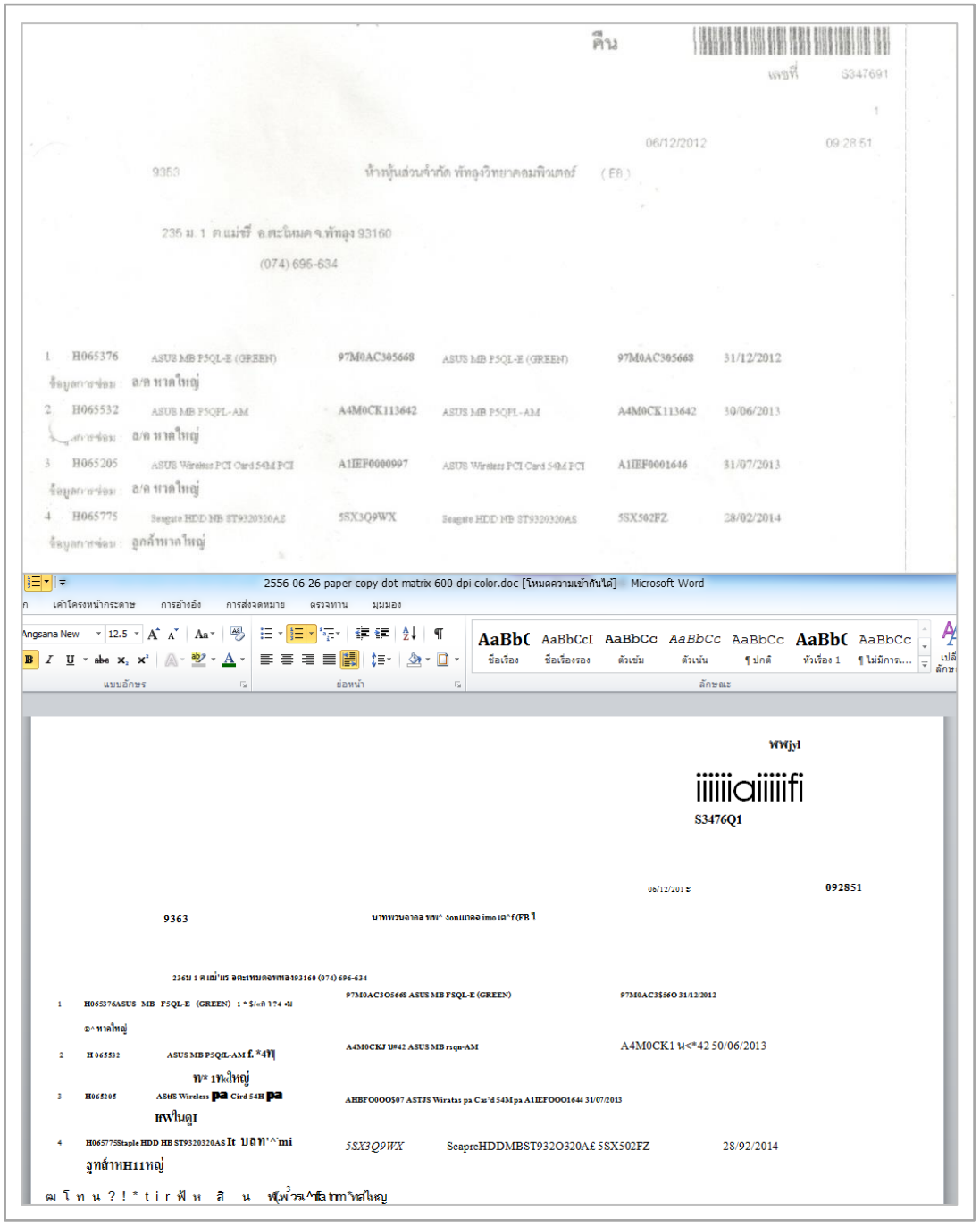
ภาพประกอบ 2-3 ตัวอย่างการรู้จำเอกสารพื้นหลังสี กราดภาพ 300 dpi Color



ภาพประกอบ 2-4 ตัวอย่างการรู้จำเอกสาร กราดภาพ 600 dpi Color



ภาพประกอบ 2-5 ตัวอย่างการรู้จำเอกสาร กราดภาพ 600 dpi Gray



ภาพประกอบ 2-6 ตัวอย่างการรู้จำเอกสารสำเนา กราดภาพ 600 dpi Color

จากภาพประกอบ 2-2 ถึงภาพประกอบ 2-6 แสดงให้เห็นว่าผลการรู้จำจะถูกต้องหรือผิดพลาดขึ้นอยู่กับปัจจัยต่างๆ ที่มีผลต่อการรู้จำ ได้แก่

2.2.1 ชนิดแบบอักษร

ชนิดแบบอักษร (Font) สามารถแบ่งได้ 2 แบบ คือ แบ่งตามเชิงตัวอักษร และแบ่งตามความกว้างของตัวอักษร

1) แบ่งตามเชิงตัวอักษร สามารถแบ่งเป็น 2 ประเภท คือ แบบมีเชิง (Serif) จะมีขีดเล็กๆ ที่ปลายตัวอักษร และแบบไม่มีเชิง (Sans Serif) จะไม่มีขีดเล็กๆ ที่ปลายตัวอักษร ดัง ภาพประกอบ 2-7 และภาพประกอบ 2-8

ภาพประกอบ 2-7 ตัวอักษรแบบมีเชิง

ภาพประกอบ 2-8 ตัวอักษรแบบไม่มีเชิง

2) แบ่งตามความกว้างของตัวอักษร สามารถแบ่งเป็น 2 ประเภท คือ ชนิดแบบอักษรที่มีความกว้างไม่เท่ากัน (Variable Width Font) มีทั้งตัวที่ผอมและตัวที่อ้วน อักษรอังกฤษ เช่น A I M O อักษรไทย เช่น ว ร ข ฒ ณ ญ เช่นชนิดแบบอักษร TH SarabanPSK, Angsana New, Browallia New หรือชนิดแบบอักษรที่ลงท้ายด้วย UPC เช่น AngsanaUPC, BrowalliaUPC และชนิดแบบอักษรที่มีความกว้างเท่ากันทุกตัวอักษร (Fixed Width Font) เช่น ชนิดแบบอักษร Cordia new DB, ThaiTextFixed, LotusSuttabun, Courier Thai

ชนิดแบบอักษรมีความสำคัญมากในการรู้จำอักขระด้วยแสง เช่น TH SarabanPSK เส้นตัวอักษรจะบาง และเส้นตัวอักษรส่วนปลายจะสั้น เช่น ช ซ ศ ส ฬ ฮ เมื่อพิมพ์ด้วยเครื่องพิมพ์ดอทเมทริกซ์ทำให้ตัวอักษรไม่คมชัดทำให้การรู้จำผิดพลาดได้

ชนิดแบบอักษรที่ผู้วิจัยเลือกใช้ในการวิจัย ได้แก่ Angsana New, Cordia New และ TH SarabunPSK ดังภาพประกอบ 2-9

ภาพประกอบ 2-9 ชนิดแบบอักษรที่ใช้ในการวิจัย

2.2.2 ขนาดตัวอักษร

จากกรณีศึกษาที่ผู้วิจัยได้ทดลองนำเอกสารที่มีตัวอักษรขนาด 8 11 14 16 18 มากราดภาพแล้วทำการรู้จำผลที่ได้คือ ตัวอักษรขนาด 8 ให้ค่าความถูกต้องเท่ากับ 75.00% ขนาด 11 ให้ค่าความถูกต้องเท่ากับ 84.72% ขนาด 14 ให้ค่าความถูกต้องเท่ากับ 89.58% ขนาด 16 ให้ค่าความถูกต้องเท่ากับ 91.67% และขนาด 18 ให้ค่าความถูกต้องเท่ากับ 100.00% แสดงให้เห็น

ว่าตัวอักษรที่มีขนาดเล็กจะมีความแม่นยำต่ำกว่าตัวอักษรที่มีขนาดใหญ่ ดังที่ได้ทดลองเปรียบเทียบไว้ในตาราง 1-2 งานวิจัยนี้จึงเลือกตัวอักษรขนาด 14 16 และ 18 มาทำการทดลอง

2.2.3 ความเข้มของผ้าหมึกและแรงกดของหัวเข็ม

เนื่องจากเครื่องพิมพ์ดอทเมทริกซ์ใช้ผ้าหมึกและแรงกดของหัวเข็มในการพิมพ์ หากผ้าหมึกใหม่สีของหมึกจะเข้มมากและแรงกดของหัวเข็มตัวอักษรที่ได้ก็จะมีลักษณะขอบตัวอักษรที่เบลอล้นออกมา และในทางกลับกันถ้าผ้าหมึกผ่านการพิมพ์มามากหรือหมึกใกล้หมดตัวอักษรที่ได้ก็จะบางหรือขาดเนื่องจากหัวเข็มหักหรือเสื่อมคุณภาพ ดังภาพประกอบ 2-4 งานวิจัยนี้เลือกใช้เครื่องพิมพ์ดอทเมทริกซ์เครื่องใหม่ทั้งผ้าหมึกและหัวเข็มมีคุณภาพดีมาทำการทดลอง

2.2.4 จำนวนชั้นของเอกสาร

เอกสารดอทเมทริกซ์จะมีการตั้งค่าในการพิมพ์เอกสารตามจำนวนชั้นของเอกสารแต่ละชุดได้ เช่นแบบฟอร์มเอกสารที่มี 5 สำเนาหรือชั้น ชั้นแรกๆ ย่อมมีแรงกดของหัวพิมพ์ที่มากกว่า ตัวอักษรก็คมชัดกว่า ขึ้นอยู่กับสีและความเข้มของกระดาษคาร์บอน ในแต่ละชั้นด้วยดังภาพประกอบ 2-2 และภาพประกอบ 2-6 เช่นคาร์บอนสีน้ำเงิน สีดำ งานวิจัยนี้เลือกใช้กระดาษต่อเนื่องชนิดคาร์บอน 2 ชั้น หมึกสีดำ และกระดาษต่อเนื่องชนิดเคมี 3 ชั้น หมึกสีดำและสีน้ำเงินมาทำการทดลอง

2.2.5 ลักษณะตัวอักษร

การพิมพ์ตัวอักษรตัวปกติ ตัวหนา และตัวเอียง ด้วยเครื่องพิมพ์ดอทเมทริกซ์สำหรับอักษรตัวหนาที่อยู่ใกล้กันขอบตัวอักษรที่เบลอล้นออกมาทำให้ตัวอักษรติดกันได้ ส่วนตัวเอียงจะทำให้ตัวอักษรขาดไม่เป็นเส้นเดียวกัน ขอบตัวอักษรไม่เรียบ หรือตัวอักษรที่บางของชนิดแบบอักษร TH SarabanPSK ดังภาพประกอบ 2-9 ทำให้ตัวอักษรที่ได้ไม่คมชัด งานวิจัยนี้จึงเลือกลักษณะตัวอักษรตัวปกติ ตัวหนา และตัวเอียง มาทำการทดลอง

2.2.6 สีตัวอักษร และสีพื้นเอกสาร

ถ้าตัวอักษรเป็นสีดำหรือสีน้ำเงิน และสีพื้นเอกสารเป็นสีขาว การกราดภาพเอกสารจะมีความชัดเจนเมื่อนำไปรู้จำอักขระด้วยแสง ก็สามารถถอดตัวอักษรออกมาได้ง่าย แต่ถ้าสีตัวอักษรเป็นสีดำหรือสีเข้มและสีพื้นของเอกสารก็เข้ม เมื่อกราดภาพก็จะได้ไฟล์ภาพที่เข้มทั้งสีตัวอักษรและสีพื้นทำให้การรู้จำผิดพลาด ดังภาพประกอบ 2-2 และภาพประกอบ 2-3 งานวิจัยนี้จึงเลือกกระดาษต่อเนื่องชนิดคาร์บอน 2 ชั้น สีตัวอักษรสีดำ กระดาษสีขาว และกระดาษต่อเนื่องชนิดเคมี 3 ชั้น สีตัวอักษรสีดำและสีน้ำเงิน กระดาษสีขาว สีชมพู และสีน้ำเงิน มาทำการทดลอง

2.2.7 ความละเอียดในการพิมพ์เอกสาร

ความละเอียดในการพิมพ์เอกสารมีความสำคัญ เพราะคุณภาพของตัวอักษรขึ้นอยู่กับการตั้งค่าในการพิมพ์ ความละเอียดในการพิมพ์เอกสารของเครื่องพิมพ์ดอทเมทริกซ์จะแบ่งเป็น 3 ระดับ คือ ความละเอียด 120 x 180 Dpi, 180 x 180 Dpi และ 360 x 180 Dpi การตั้งค่าความละเอียดในการพิมพ์ต่ำทำให้คุณภาพการพิมพ์ต่ำแต่จะพิมพ์ได้รวดเร็วและประหยัด ทำให้ผ้าหมึกสามารถพิมพ์ได้จำนวนมากขึ้น หัวพิมพ์ก็จะสึกหรอน้อย ช่วยยืดอายุการใช้งานของเครื่องพิมพ์ ในทางกลับกันการพิมพ์ที่ความละเอียดสูงขึ้นก็จะช่วยให้คุณภาพการพิมพ์สูงขึ้นด้วยแต่การพิมพ์จะช้ากว่าและสิ้นเปลืองมากกว่า ทำให้ผ้าหมึกสามารถพิมพ์ได้จำนวนน้อยลง หัวพิมพ์ก็จะสึกหรอมากขึ้น อายุการใช้งานของเครื่องพิมพ์ก็จะน้อยลง

การพิมพ์ที่ความละเอียดสูงเมื่อนำเอกสารไปกราดภาพช่วยให้คุณภาพในการกราดภาพเอกสารสามารถกราดภาพได้คมชัดถูกต้องกว่าการพิมพ์ที่ความละเอียดต่ำกว่า

งานวิจัยนี้จึงเลือกความละเอียดในการพิมพ์ 120 x 180 Dpi, 180 x 180 Dpi และ 360 x 180 Dpi มาทำการทดลอง

2.2.8 การกราดภาพเอกสาร

การกราดภาพ คือการนำเอกสารกระดาษแปลงเป็นไฟล์รูปภาพโดยเครื่องกราดภาพ ในวิทยานิพนธ์นี้เอกสารกระดาษที่นำมากราดภาพทั้งเอกสารสี และเอกสารขาวดำ กราดภาพเป็นไฟล์รูปภาพชนิด Color (สี) และชนิด Gray (สีเทา) ที่ความละเอียด 300 dpi และ 600 dpi เท่านั้น

ความละเอียดในการกราดภาพเอกสาร จะมีผลอย่างมากกับการนำไฟล์ภาพเอกสารไปรู้จำอักขระด้วยแสง ถ้ากราดภาพที่ความละเอียดต่ำเกินไปทำให้ภาพแตก รายละเอียดของจุดสีบนภาพจะได้มาน้อย ถ้ากราดภาพที่ความละเอียดมากเกินไปก็จะทำให้ใช้เวลานาน

การตั้งค่าการกราดภาพชนิด Color หรือ Gray มีความสำคัญมาก การกราดภาพชนิด Color กรณีที่เอกสารกระดาษมีทั้งสีและขาวดำหรือตัวอักษรสีจางทำให้ได้รายละเอียดของจุดสีบนภาพมากกว่าการกราดภาพชนิด Gray ดังภาพประกอบ 2-4 และภาพประกอบ 2-5

2.2.9 ลักษณะ คุณภาพของเอกสารกระดาษ

ลักษณะ คุณภาพของเอกสารกระดาษ มีผลต่อการรู้จำ เอกสารที่มีลักษณะคุณภาพที่ดี ไม่มีสัญญาณรบกวน ไม่มีจุดดำ จุดขาว ทำให้ผลการรู้จำดีกว่าเอกสารที่มีสัญญาณรบกวน เอกสารที่มีสัญญาณรบกวน กระดาษเสื่อมคุณภาพหรือเก่า รอยพับ รอยขีด ดังภาพประกอบ 2-5 เมื่อกราดภาพทำให้มีสัญญาณรบกวน ทำให้ผลการรู้จำผิดพลาดได้ งานวิจัยนี้จึงเลือกกระดาษที่มีคุณภาพดีสำหรับพิมพ์เอกสารมาทำการทดลอง

2.2.10 คุณภาพ อายุการใช้งานของเครื่องพิมพ์

คุณภาพ อายุการใช้งานของเครื่องพิมพ์มีผลต่อการรู้จำ เครื่องพิมพ์ที่ใหม่ คุณภาพดี อายุการใช้งานน้อย คุณภาพการพิมพ์จะดีกว่า เครื่องพิมพ์ที่หัวเข็มสึกหรอ หรืออายุการใช้งานนานทำให้คุณภาพการพิมพ์ต่ำลง ทำให้ตัวอักษรไม่ชัดเจน แรงกดไม่สม่ำเสมอ ดังภาพประกอบ 2-4 ทำให้ผลการรู้จำผิดพลาดได้ งานวิจัยนี้เลือกใช้เครื่องพิมพ์ใหม่สำหรับพิมพ์เอกสารมาทำการทดลอง

2.3 การหมุนภาพ และการถอดข้อความ

เมื่อรู้ว่าเป็นแบบฟอร์มชนิดใดแล้ว เราสามารถหมุน/เลื่อน/ย่อ-ขยายแบบฟอร์มให้ตรงกับต้นแบบ จากนั้น จึงตัดเอาภาพข้อความเฉพาะตรงตำแหน่งของข้อมูลในแบบฟอร์มนั้น ๆ ได้ ซึ่งส่วนของภาพที่ได้จะนำไปผ่านกระบวนการรู้จำตัวอักษรต่อไป

2.4 การปรับปรุงภาพ และการขจัดสัญญาณรบกวน

การปรับปรุงภาพจะปรับปรุงโดยการปรับค่าสีของภาพ และการขจัดสัญญาณรบกวนออกจากภาพก่อนทำการรู้จำ เพื่อให้ได้ผลการรู้จำที่ถูกต้องมากขึ้น การปรับปรุงภาพมีหลายวิธี เช่น

2.4.1 การปรับปรุงรูปร่างตัวอักษร

การปรับปรุงรูปร่างตัวอักษร ได้แก่ การวัดช่องว่างระยะห่างของตัวอักษร (C. H. Chen และ J. L. DeCurtins, ค.ศ. 1992) การพิจารณาจากรูปร่างและคำศัพท์ตัวอักษรจากหนังสือที่มีความหลากหลายทั้งเนื้อหาและแบบอักษร (Lee Dar-Shyang และ Smith Ray, ค.ศ. 2012) การปรับปรุงภาพเอกสารดอทเมทริกซ์ที่ตัวอักษรขาดกลายเป็นชิ้นๆ ไม่ต่อกัน ตัวอักษรติดกัน (Berrin A. Yanikoglu, ค.ศ. 2000) โดยการวิเคราะห์ความสัมพันธ์โดยอัตโนมัติ (Auto-Correlation Analysis) การประมาณการอย่างมีประสิทธิภาพที่น่าเชื่อถือ การแปลงและการวิเคราะห์ระยะห่าง ช่องว่างในแนวตั้ง การวิเคราะห์องค์ประกอบหลัก (Component Analysis)

การวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis) อาศัยหลักความสัมพันธ์เชิงเส้นระหว่างตัวแปรที่ใช้เป็นข้อมูลองค์ประกอบหลักตัวแปรคือ การผสมเชิงเส้นตรง (Linear Combination) ของตัวแปรที่อธิบายการผันแปรของข้อมูลได้มากที่สุด จากนั้นหาการผสมที่สองที่สามารถอธิบาย การผันแปรได้มากที่สุดเป็นอันดับที่สอง โดยที่ไม่สัมพันธ์กับการผสมแรก ทำเช่นนี้เรื่อยไปจนได้องค์ประกอบหลัก (หรือปัจจัย) ที่สามารถอธิบายการผันแปรของทุกตัวแปรได้ครบถ้วน ซึ่งองค์ประกอบหลักจะอธิบายการผันแปรได้ น้อยลงตามลำดับและทุกองค์ประกอบไม่สัมพันธ์กัน

2.4.2 การขจัดปัญหาสัญญาณรบกวน

การขจัดปัญหาสัญญาณรบกวน ได้แก่ สิ่งแปลกปลอมเป็นเส้นบางๆ ระหว่างตัวอักษร หรือสัญลักษณ์ (Iuliu Konya, Stefan Eickeler และ Christoph Seibert, ค.ศ. 2011) โดยการตรวจสอบการเชื่อมต่อของตัวอักษร

2.5 การเลือกใช้โปรแกรมการรู้จำอักขระด้วยแสง

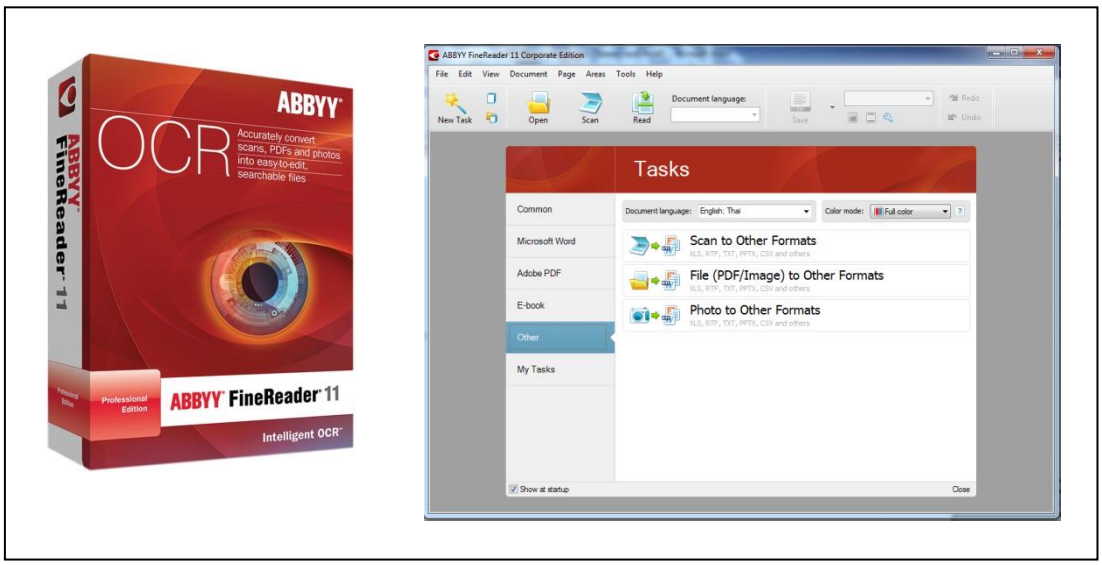
การเลือกใช้โปรแกรมการรู้จำอักขระด้วยแสง มีความสำคัญมาก เพราะมีผลโดยตรงต่อการวิจัย ควรใช้โปรแกรมที่มีความแม่นยำสูงช่วยให้การทำงานมีประสิทธิภาพ สามารถนำข้อมูลไปใช้ได้ดีที่สุดในการรู้จำ โปรแกรมรู้จำเท่าที่ได้ศึกษามีตัวอย่าง ดังตาราง 2-1

ตาราง 2-1 ข้อดี/ข้อเสีย ของโปรแกรมการรู้จำอักขระด้วยแสง

โปรแกรมรู้จำ	ข้อดี	ข้อเสีย
ABBYY Fine Reader	มีความสามารถครบถ้วน มีประสิทธิภาพดี และสามารถรู้จำการวางผังหน้ากระดาษ	ต้องซื้อซอฟต์แวร์ ไม่สามารถใช้รวมกับโปรแกรมประยุกต์อื่น
GOCR	ไม่ต้องสอน มีสามารถในการเพิ่มชนิดแบบอักษรได้ และสามารถเปิดแฟ้มข้อมูลได้หลายรูปแบบ	ไม่สามารถใช้กับลักษณะตัวอักษรเอียง ไม่รู้จำการวางผังหน้ากระดาษ ตัวอักษรต้องแยกจากกันโดยเด็ดขาด
Tesseract	มีความถูกต้อง แม่นยำดี สามารถนำไปใช้รวมกับโปรแกรมประยุกต์อื่นได้	ไม่รู้จำการวางผังหน้ากระดาษ นำมาใช้ยาก

งานวิจัยนี้เลือกใช้ ABBYY Fine Reader ดังภาพประกอบ 2-10 เนื่องจากมีงานวิจัยได้ทดสอบความแม่นยำในการรู้จำ โดยใช้โปรแกรม ABBYY Fine Reader พบว่ามีความถูกต้อง 97 – 99% ใช้งานได้ง่าย มีประสิทธิภาพ (Tanushyam Chattopdhyyay, Priyanka Sinha and Provot Biswas, ค.ศ. 2011) ใช้เทคโนโลยีการรู้จำ Integrity, Purposefulness and Adaptability (IPA) เป็นเทคโนโลยีที่อยู่เบื้องหลัง หลักการทั้ง 3 อย่างนี้ทำให้สามารถมีการรู้จำคล้ายมนุษย์ (ABBYY, ค.ศ. 2015) และนอกจากนี้มีงานวิจัยที่ได้ศึกษาการรู้จำชนิดแบบอักษร (Font) แบบเก่าๆ โดยเลือกใช้ ABBYY Fine Reader ในการศึกษาปัจจัยที่มีผลกระทบต่อการรู้จำเช่นกัน (Iuliu Konya, Stefan Eickeler and Christoph Seibert, ค.ศ. 2011)

งานวิจัยนี้ไม่นำโปรแกรม Tasseract และ GOCR มาใช้เนื่องจากไม่ใช่เครื่องมือสำเร็จรูป ต้องพัฒนาอีกทำให้ไม่สะดวกในการใช้ ดังตาราง 2-1



ภาพประกอบ 2-10 โปรแกรม ABBYY Fine Reader 11

2.6 งานวิจัยที่เกี่ยวข้อง

2.6.1 งานวิจัยที่เกี่ยวข้องกับการสกัดข้อมูลจากแบบฟอร์ม

T.A. Bayer และ H.U. Mogg-Schneider ทำวิจัยเมื่อปี ค.ศ. 1997 ได้นำเสนอระบบอัตโนมัติในการสกัดรายการจากใบกำกับสินค้าที่ไม่จำกัดรูปแบบของฟอร์มจากใบกำกับสินค้า 473 ใบ โดยเลือกดึงตัวอักษรออกมาบางรายการ เช่น รายการ เลขที่ วันที่ ผู้ส่งของ จำนวนเงิน โดยอาศัยโครงสร้างตัวแบบ

MING Delie, LIU Jian, TIAN Jinwen ทำวิจัยเมื่อปี ค.ศ. 2002 ได้มีการออกแบบพัฒนาระบบรู้จำใบกำกับสินค้าภาษาจีน โดยการตรวจสอบหาชนิดของแบบฟอร์ม โดยวิธีการเคลื่อนที่ทั้งบล็อกตรงๆ ในส่วนของแนวตั้ง เพื่อหาตำแหน่งของเขตข้อมูลในฟอร์ม การกรองบรรทัดให้ฟอร์ม ฯลฯ คุณสมบัติของใบกำกับสินค้าแต่ละชนิดจะถูกสร้างขึ้นโดยวิธีจักรกลเรียนรู้แบบกึ่งอัตโนมัติบนพื้นฐานของการรู้จำชนิดของใบกำกับสินค้า โดยแบบฟอร์มจริงจะถูกจัดวางใหม่ให้ตรงกับแบบฟอร์มเปล่าตามคุณลักษณะของชนิดของแบบฟอร์มนั้นๆ

Enrico Sorio, Alberto Bartoli, Giorgio Davanzo และ Eric Medvet ทำวิจัยเมื่อปี ค.ศ. 2012 นักวิจัยได้นำเสนอวิธีการตรวจสอบอัตโนมัติและการแก้ไขข้อผิดพลาดของการรู้จำตัวอักษร โดยการใช้ฐานความรู้เกี่ยวกับชนิดของข้อมูล การดึงข้อมูลมาจากใบกำกับสินค้ามาเกือบ 800 ค่า จากใบกำกับสินค้า 100 ใบ ซึ่งในแต่ละใบจะดึงมา 8 ตำแหน่ง เช่น วันที่ เลขที่ใบกำกับสินค้า เลขที่ประจำตัวผู้เสียภาษี มูลค่าสินค้า ภาษีมูลค่าเพิ่ม จำนวนเงินรวม

งานวิจัยข้างต้นได้มีการศึกษาเกี่ยวกับแบบฟอร์มใบกำกับสินค้าเหมือนกับวิทยานิพนธ์นี้ แต่แตกต่างที่ใบกำกับสินค้าที่นำมาวิจัยนั้นไม่ได้พิมพ์จากเครื่องพิมพ์ดอทเมทริกซ์

2.6.2 งานวิจัยที่เกี่ยวข้องกับการปรับปรุงปัจจัยสภาพแวดล้อม

C. H. Chen และ J. L. DeCurtins ทำวิจัยเมื่อปี ค.ศ. 1992 นักวิจัยได้นำเสนอแนวคิดใหม่ในการวัดช่องว่างของตัวอักษรในแนวตั้งลงมาแล้วดูว่าจุดตัดอยู่ตำแหน่งไหน ดูค่าจากจุดตัดมาเป็นข้อมูล แล้วนำไปวิเคราะห์กับค่าของจุดตัดของตัวอักษรแต่ละตัว ใช้วิธีที่กำบังคุณลักษณะ (Mask Feature) และเส้นแสดงรูปร่าง (Contour) แล้วนำไปนับคะแนน (Mask Feature) ว่าเป็นตัวอักษรอะไร แล้วนำไปรู้จำ

Berrin A. Yanikoglu ทำวิจัยเมื่อปี ค.ศ. 2000 นักวิจัยได้เสนอระบบการรู้จำเพื่อเอาชนะกับความยากลำบากของกระบวนการตัดตัวอักษรข้อความดอทเมทริกซ์ โดยเฉพาะอย่างยิ่งเมื่อตัวอักษรขาดกลายเป็นชิ้นๆ ไม่ต่อกัน ระบบการรู้จำข้อความดอทเมทริกซ์จะค้นหาระยะห่างส่วนใหญ่ของข้อความ มาตัดสินว่าข้อความนี้เขียนด้วยชนิดแบบอักษร แบบระยะห่างคงที่หรือเป็นชนิดแบบอักษรแบบมีสัดส่วน ข้อความแบบระยะห่างคงที่จะถูกตัดด้วยระยะห่างซึ่งสามารถตัดได้สำเร็จทั้งตัวอักษรที่ติดกันและตัวอักษรขาด

งานวิจัยข้างต้นจะเป็นการศึกษาปัจจัยที่มีผลกระทบต่อการรู้จำอักขระด้วยแสง ได้แก่ ช่องว่าง ระยะห่างตัวอักษร ชนิดแบบอักษร ซึ่งวิทยานิพนธ์นี้ก็จะศึกษาปัจจัยต่าง ๆ ที่มีผลกระทบต่อการรู้จำอักขระด้วยแสง เช่นกัน

Iuliu Konya, Stefan Eickeler และ Christoph Seibert ทำวิจัยเมื่อปี ค.ศ. 2011 จากปัญหาสิ่งแปลกปลอมที่เกิดขึ้นจากการพิมพ์หนังสือพิมพ์ในโบราณ ที่พิมพ์ด้วยแม่พิมพ์แบบโลหะที่ใช้กันอย่างแพร่หลาย ในช่วงปลายศตวรรษที่ 19 และต้นศตวรรษที่ 20 สิ่งแปลกปลอมจะปรากฏเป็นเส้นบาง ๆ เส้นเดียวระหว่างตัวอักษร หรือสัญลักษณ์ เชื่อมต่อกับตัวอักษรที่อยู่ใกล้เคียงเป็นส่วนใหญ่ ทำให้การรู้จำผิดพลาด นักวิจัยได้เสนอวิธีการใหม่ ที่ใช้กับชนิดแบบตัวอักษรแบบเก่า ๆ โดยการตรวจสอบการเชื่อมต่อของตัวอักษรทิศทางเดียว (Directional Single-Connected Chains) สำหรับเอกสารที่ตกแต่งซับซ้อน สิ่งแปลกปลอมที่มีความกว้าง ความสูง เส้นที่บางกว่าตัวอักษร เช่น ชนิดแบบตัวอักษร Fraktur โดยเฉพาะการตรวจสอบชุดข้อมูลที่ประกอบด้วย งานหนังสือพิมพ์เก่า โดยใช้ ABBYY Fine Reader และ Tesseract ของ Google

Lee Dar-Shyang และ Smith Ray ทำวิจัยเมื่อปี ค.ศ. 2012 นักวิจัยได้เสนอการแก้ปัญหาการรู้จำตัวอักษรจากหนังสือที่มีความหลากหลายทั้งเนื้อหาและแบบอักษร ด้วยระบบที่รวมการแก้ไขแบบคู่ขนานเพื่อให้ผลของการรู้จำถูกต้องมากขึ้น โดยใช้ภาพเฉพาะของเอกสาร และแบบของภาษาแต่ละแบบ โดยพิจารณาจากรูปร่างและคำศัพท์ เพื่อระบุความไม่แน่นอนเป็นสมมุติฐานของการแก้ไขแต่อยู่บนฐานการตรวจสอบที่มีประสิทธิภาพ โดยการใช้โปรแกรม Tasseract ซึ่งเป็น Open Source ของ Google เป็นผลลัพธ์ของข้อมูลหนังสือที่กราดภาพจำนวนมาก ทำให้อัตราความผิดพลาดลดลง 25%

งานวิจัยข้างต้นนอกจากจะศึกษาปัจจัยที่มีผลกระทบต่อการรู้จำอักขระด้วยแสงและใช้โปรแกรม ABBYY Fine Reader ในการรู้จำอักขระด้วยแสงเหมือนกับวิทยานิพนธ์นี้ แต่มีความแตกต่างที่นำโปรแกรม Tasseract มาช่วยในการปรับปรุงความผิดพลาดเพื่อเสริมการรู้จำอีกทาง

บทที่ 3

วิธีการดำเนินงาน

บทนี้จะกล่าวถึงวิธีการดำเนินงานตั้งแต่การเตรียมเอกสารตัวอย่าง ได้แก่ รูปแบบของเอกสารตัวอย่าง ชนิดของกระดาษ และการกราดภาพเอกสาร ต่อด้วยการทำการรู้จำ วัดผลการรู้จำ จนถึงการหาค่าความถูกต้องและค่าความผิดพลาดของการรู้จำ

3.1 เตรียมเอกสารตัวอย่าง

3.1.1 รูปแบบของเอกสารตัวอย่าง

งานวิจัยนี้ได้จัดพิมพ์เอกสารตัวอย่างที่พิมพ์ด้วยเครื่องพิมพ์ดอทเมทริกซ์ใช้กระดาษต่อเนื่องชนิดคาร์บอน 2 ชั้น พื้นกระดาษเป็นสีขาว ตัวอักษรสีดำ และกระดาษต่อเนื่องชนิดเคมี 3 ชั้น พื้นกระดาษสีขาว สีชมพู และสีเหลือง ตัวอักษรสีดำ และสีน้ำเงิน ตามปัจจัยที่สามารถควบคุมได้ดังนี้

1) ชนิดแบบอักษร มี 3 ปัจจัย ได้แก่ Angsana New (นิยมใช้ทั่วไป), Cordia New (นิยมใช้ทั่วไป) และ TH SarabunPSK (ใช้ในราชการ)

2) ขนาดตัวอักษร มี 3 ปัจจัย ได้แก่ ขนาด 14 (ถ้าเล็กกว่าอักษรไทยจะใช้ไม่ได้), 16 และ 18 (ถ้าใหญ่กว่าก็ใหญ่เกินไป)

3) ลักษณะตัวอักษร มี 3 ปัจจัย ได้แก่ ตัวปกติ ตัวหนา และตัวเอียง

4) ความละเอียดในการพิมพ์ มี 3 ปัจจัย ได้แก่ 120 x 180 dpi, 180 x 180 dpi และ 360 x 180 dpi (เครื่องพิมพ์ที่ใช้พิมพ์มี 3 ความละเอียด)

5) ความละเอียดและชนิดในการกราดภาพ มี 4 ปัจจัย ได้แก่ ความละเอียด 300 dpi (เพียงพอที่จะใช้ในการรู้จำอักขระด้วยแสง ไม่เร็วไม่ช้าเกินไป) มี 2 แบบ คือ ชนิด Color และ Gray ความละเอียด 600 dpi มี 2 แบบ คือ ชนิด Color และ Gray บันทึกเป็นไฟล์รูปภาพแบบ PNG (Portable Network Graphic)

6) ตัวอักษรที่ใช้ทดสอบการรู้จำ มีทั้งอักษรไทย อักษรอังกฤษตัวพิมพ์ใหญ่และตัวพิมพ์เล็ก เลขไทย และเลขอารบิค

จากปัจจัยชนิดแบบอักษร (3) ขนาดตัวอักษร (3) ลักษณะตัวอักษร (3) ความละเอียดในการพิมพ์ (3) เมื่อนำมาพิมพ์ลงบนกระดาษต่อเนื่องชนิดคาร์บอน 2 ชั้น และชนิดเคมี 3 ชั้น (5) ตามที่ได้กำหนดไว้ ทำให้สามารถคำนวณเป็นจำนวนรูปแบบได้ทั้งหมด 3 x 3 x 3 x 3 x 5 = 405 หน้า

จากปัจจัยความละเอียดและชนิดในการกราดภาพ (4) ทั้งชนิด Color และชนิด Gray จะได้ไฟล์รูปภาพทั้งหมด 405 x 4 = 1,620 ไฟล์

ในการออกแบบการวิจัยในตอนแรกสร้างเอกสารตัวอย่าง ได้แก่ อักษรไทย 1 ชุด อักษรอังกฤษ 1 ชุด และตัวเลข 1 ชุด ซึ่งได้ทำตามขั้นตอน จนเสร็จสิ้นตามกระบวนการ 1 ชุด คืออักษรอังกฤษ แต่เกิดข้อผิดพลาดที่อักษรอังกฤษ ผู้วิจัยได้ออกแบบเอกสารตัวอย่าง a ถึง z แบบพบกันทุกตัวอักษรไว้หน้าเดียวกันแต่มีเฉพาะตัวพิมพ์เล็ก ได้ทำขั้นตอนตั้งแต่พิมพ์เอกสาร การกราดภาพ การรู้จำ การนับตัวอักษรถูก/ผิด รายงานผล จนครบทุกตัวอย่างการวิจัย แต่การวิจัยที่ทำไปแล้วนั้นรายงานผลการนับตัวอักษรถูก/ผิดออกมา มีเฉพาะตัวพิมพ์เล็ก ซึ่งในการทำงานจริงต้องการตัวพิมพ์ใหญ่ด้วย ดังนั้นก็ต้องสร้างเอกสารตัวอย่างตัวพิมพ์ใหญ่อีกชุด และต้องทำตามขั้นตอนทั้งหมดอีกรอบซึ่งทำให้เสียเวลาเป็นอย่างมาก อีกทั้งเอกสารตัวอย่างการวิจัยอักษรไทย และตัวเลขที่ได้ออกแบบไว้ก็ยังไม่ได้ดำเนินการตามขั้นตอน

ดังนั้นผู้วิจัยจึงได้ออกแบบตัวอย่างใหม่ให้มีอักษรอังกฤษทั้งตัวพิมพ์เล็ก ตัวพิมพ์ใหญ่ อักษรไทย เลขไทย เลขอารบิค อยู่ในหน้าเดียวกันและดำเนินการวิจัยใหม่ทั้งหมด ดังภาพประกอบ 3-1 การออกแบบเอกสารตัวอย่าง

abcdefghijklmnopqrstuvwxyzABCDEFGHIJKLMNOPQRSTUVWXYZ1234567890
๑๒๓๔๕๖๗๘๙๐กขฃคฅฆงจฉชซฌญฎฏฐฑฒณดตถทธนบปผฝพฟภมยรฤลฦวศษสหฬอฮ
abcdefghijklmnopqrstuvwxyzABCDEFGHIJKLMNOPQRSTUVWXYZ1234567890
๑๒๓๔๕๖๗๘๙๐กขฃคฅฆงจฉชซฌญฎฏฐฑฒณดตถทธนบปผฝพฟภมยรฤลฦวศษสหฬอฮ

ภาพประกอบ 3-1 การออกแบบเอกสารตัวอย่าง

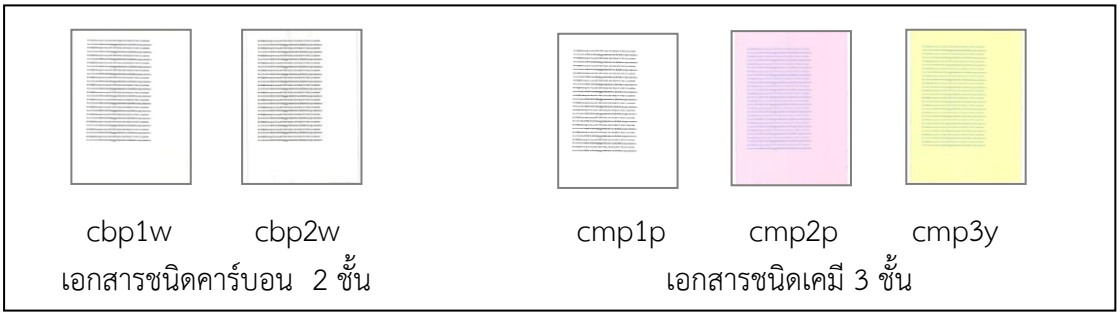
3.1.2 ชนิดของกระดาษ

ชนิดของกระดาษที่ใช้ในงานวิจัยนี้คือ กระดาษต่อเนื่อง ชนิดคาร์บอน และชนิดเคมี เป็นชนิดกระดาษที่ใช้กันแพร่หลาย สะดวกประหยัด สำหรับการพิมพ์เอกสารหลายชุดในครั้งเดียว มีคุณลักษณะดังนี้

1) กระดาษต่อเนื่องชนิดคาร์บอน 2 ชั้น ยี่ห้อ TKS ขนาด 9x11 นิ้ว กระดาษทั้ง 2 ชั้น สีขาว หนา 50 แกรม

2) กระดาษต่อเนื่องชนิดเคมี 3 ชั้น ยี่ห้อ GNR ขนาด 9x11 นิ้ว ชั้นที่ 1 สีขาว ชั้นที่ 2 สีชมพู ชั้นที่ 3 สีเหลือง กระดาษชั้นที่ 1 และชั้นที่ 2 หนา 50 แกรม ชั้นที่ 3 หนา 54 แกรม

เอกสารดอทเมทริกซ์ทั้ง 5 หน้า ดังภาพประกอบ 3-2 ประกอบด้วย cbp1w คือชนิดคาร์บอน หน้า 1 กระดาษสีขาว ตัวอักษรสีดำ cbp2w คือชนิดคาร์บอน หน้า 2 (สำเนา) กระดาษสีขาว ตัวอักษรสีดำ cmp1w คือชนิดเคมี หน้า 1 กระดาษสีขาว ตัวอักษรสีดำ cmp2p คือชนิดเคมี หน้า 2 (สำเนา) กระดาษสีชมพู ตัวอักษรสีน้ำเงิน cmp3y คือชนิดเคมี หน้า 3 (สำเนา) กระดาษสีเหลือง ตัวอักษรสีน้ำเงิน



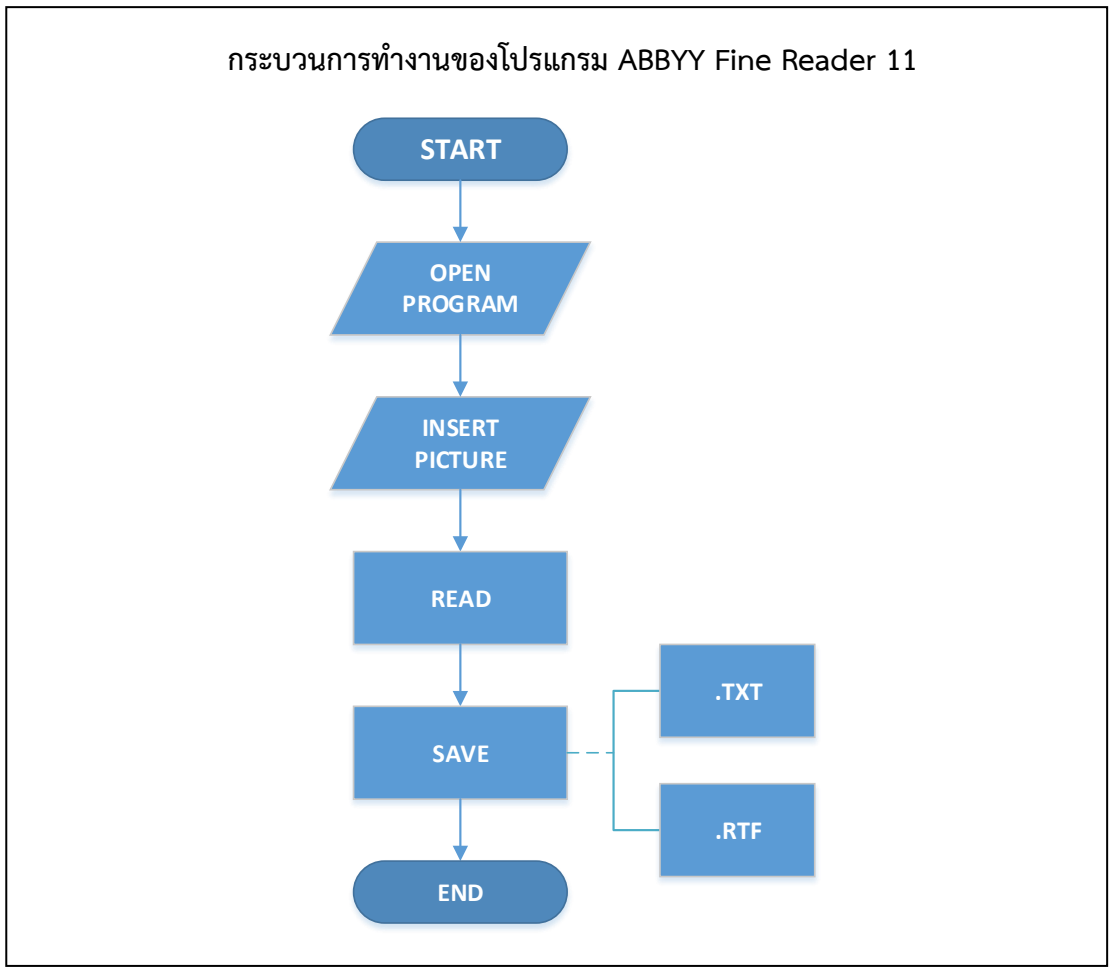
ภาพประกอบ 3-2 เอกสารตัวอย่าง

การกราดภาพ ต้องเลือกการกราดภาพเป็นชนิดภาพ .png เพราะให้ค่าความละเอียดที่เหมาะสำหรับการนำไปใช้มากที่สุด ภาพที่ได้ไม่เล็กไม่ใหญ่เกินไป ในการตั้งค่าการกราดภาพก็ต้องระมัดระวังไม่ให้ตั้งค่าผิด การวางกระดาษ การใส่กระดาษก็สำคัญ ต้องทำด้วยความละเอียด รอบคอบ อีกทั้งเอกสารที่พิมพ์มีทั้งตัวปกติ ตัวหนา และตัวเอียง ขนาด 14, 16 และ 18 ชนิดตัวอักษร Angsana New, Cordia New และ TH SarabunPSK ความละเอียดในการพิมพ์เอกสาร 120x180 dpi 180x180 dpi และ 360x180 dpi เอกสารแต่ละแผ่นจะแตกต่างกันที่รูปแบบที่เราตั้งไว้ ควรแยกเอกสารอย่างละเอียดรอบคอบ การบันทึกก็ต้องระมัดระวังไม่ให้ผิดพลาด

ขั้นตอนการรวบรวมเอกสารตัวอย่างมีความสำคัญมากเพราะหมายถึงเวลาในการทำงาน ขั้นตอนในการทำงานทุกขั้นตอนจะใช้เวลาและความละเอียดรอบคอบ จากการออกแบบในตอนแรกได้ทำการออกแบบเอกสารตัวอย่างโดยแยกเอกสารอักษรไทย อักษรอังกฤษ และตัวเลข เป็นวิธีการที่ทำให้ตัวอย่างเพิ่มมากขึ้น ยุ่งยากหลายขั้นตอนมากขึ้น ทั้งการพิมพ์ การกราดภาพ การรู้จำ การหาค่าความถูกต้อง และที่สำคัญเอกสารที่เราพบมักจะรวมกันทั้งหมด ทั้งอักษรไทย อักษรอังกฤษ ตัวเลข ดังนั้นการรวมไว้หน้าเดียวกันทำให้ใกล้เคียงกับความเป็นจริงมากกว่าและสามารถช่วยลดขั้นตอนในการทำงานได้

3.2 การรู้จำ

ทำการรู้จำโดยการนำไฟล์ภาพที่ได้จากการกราดภาพ มาทำการรู้จำด้วยโปรแกรม ABBYY Fine Reader 11 ระบบปฏิบัติการ Windows 7 บันทึกผลลัพธ์เป็นไฟล์ข้อความ .txt และไฟล์เอกสาร .rtf การทำงานของโปรแกรม ABBYY Fine Reader 11 เมื่อผู้ใช้แทรกไฟล์ภาพที่ต้องการรู้จำ โปรแกรมก็จะวิเคราะห์ภาพและแปลงไฟล์อัตโนมัติ แล้วแสดงหน้าต่างให้บันทึกไฟล์ .txt และ .rtf ตามต้องการ ดังภาพประกอบ 3-3

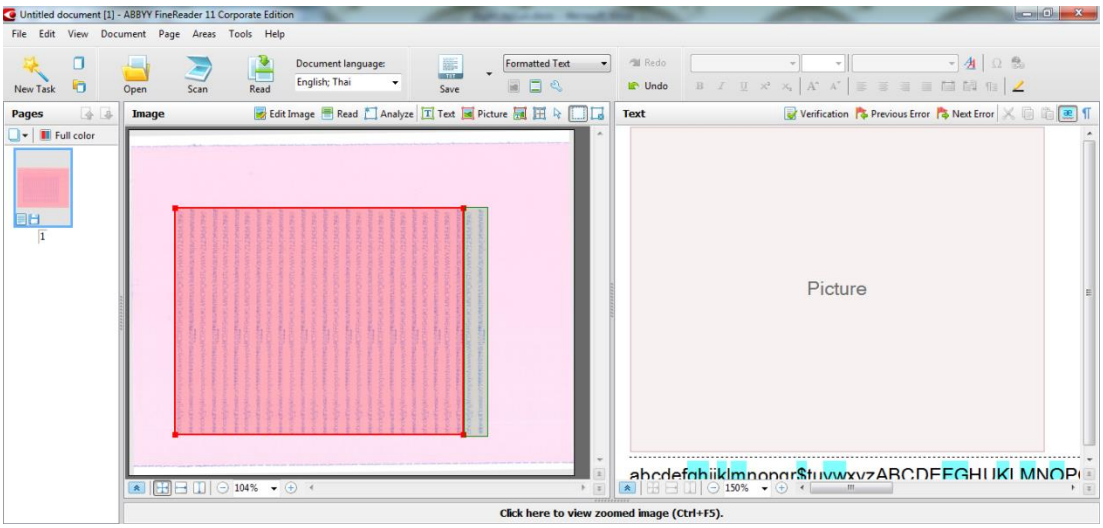


ภาพประกอบ 3-3 กระบวนการทำงานของโปรแกรม ABBYY Fine Reader 11

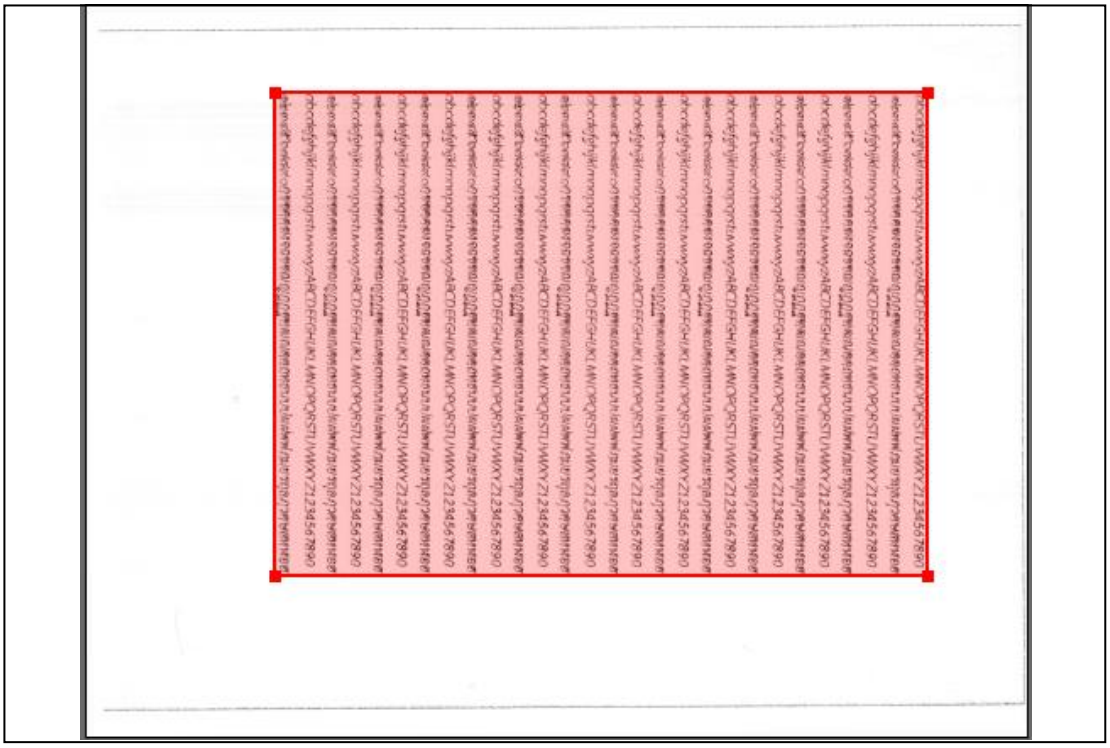
3.2.1 ปัญหาการรู้จำด้วยโปรแกรม ABBYY Fine Reader 11

การทำการรู้จำหรือการแปลงไฟล์ภาพ .png 1,620 ไฟล์ เป็นไฟล์ข้อความ .txt พบปัญหา คือ

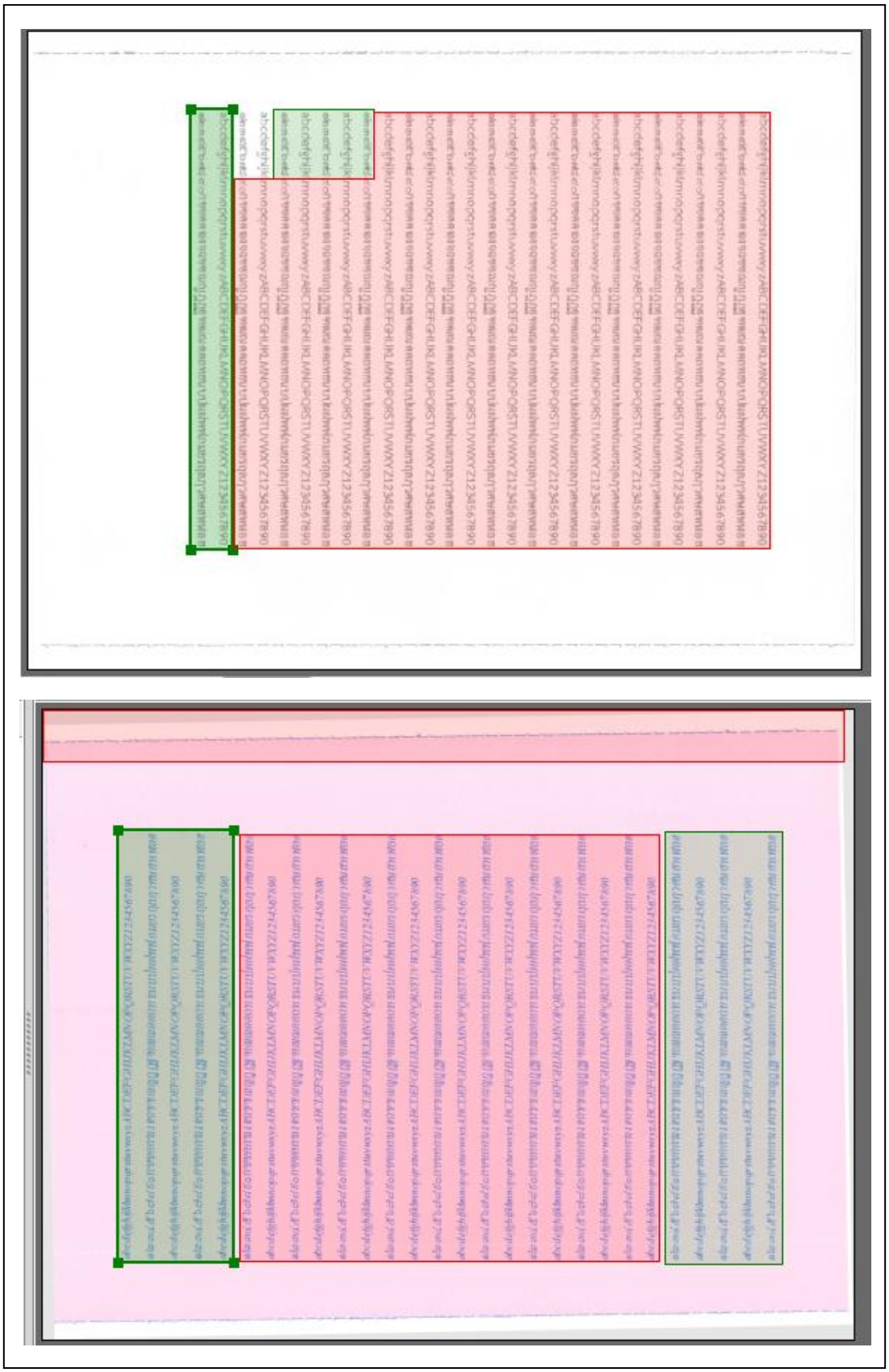
1) โปรแกรมวิเคราะห์การหมุนกระดาษผิด ไฟล์ภาพแสดงเป็นแนวนอน แล้ววิเคราะห์ผิดพลาด ดังภาพประกอบ 3-4 ถึงภาพประกอบ 3-7



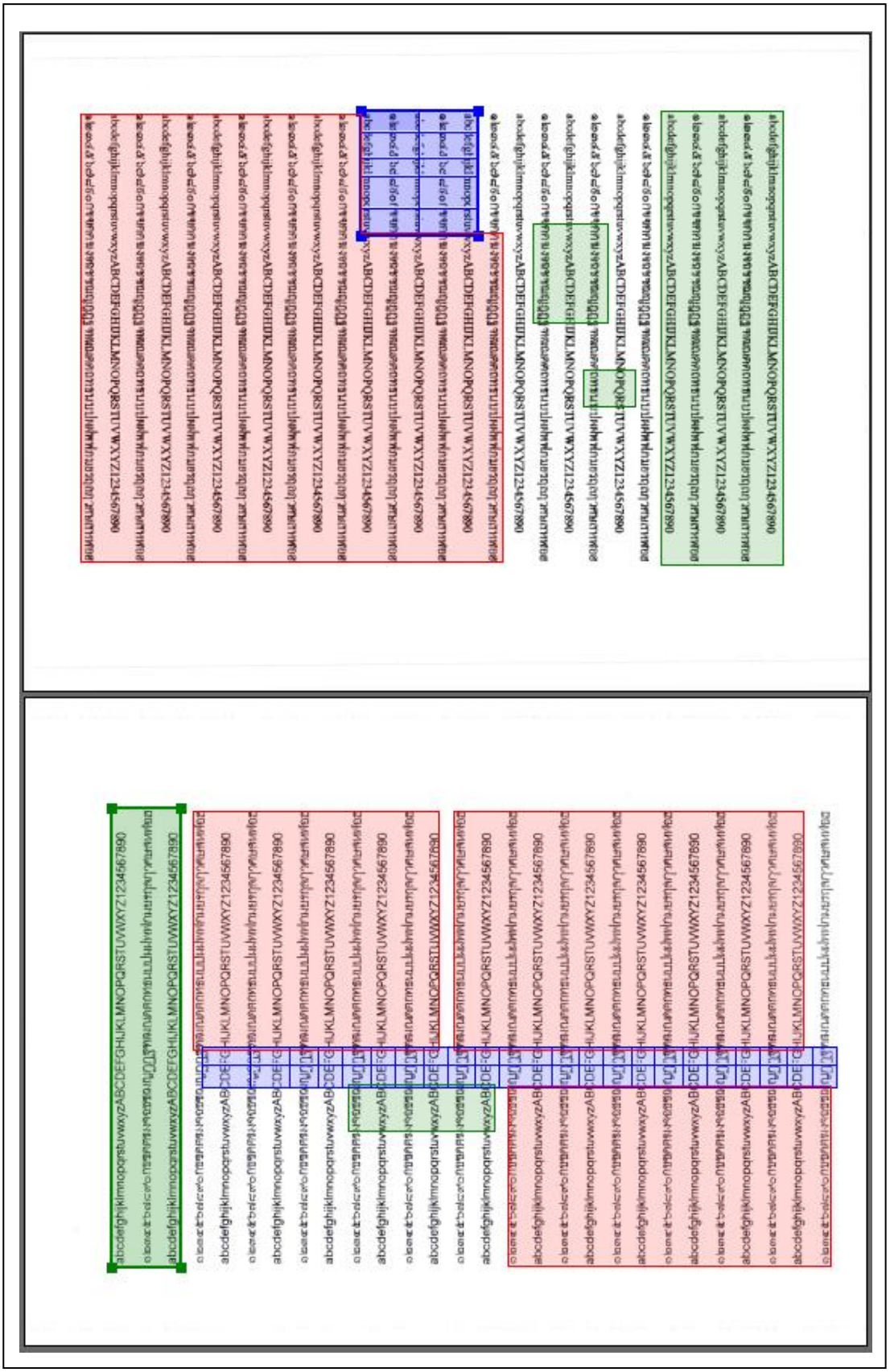
ภาพประกอบ 3-4 ตัวอย่างการหมุนกระดาษผิด



ภาพประกอบ 3-5 ตัวอย่างการหมุนกระดาษผิดแล้วแปลงเป็นรูปภาพ



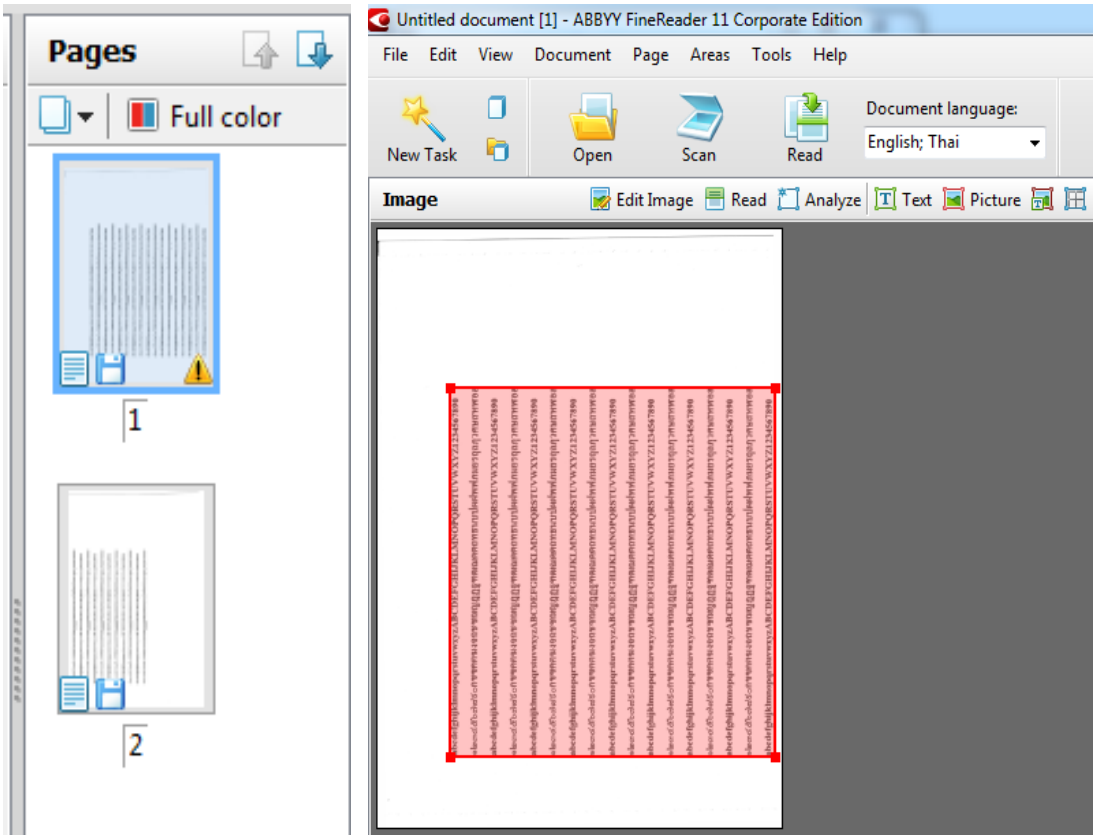
ภาพประกอบ 3-6 ตัวอย่างการหมุนกระดาษผิดแล้วแปลงเป็นรูปภาพ และตัวอักษร



ภาพประกอบ 3-7 ตัวอย่างการหมุนกระดาษผิดแล้วแปลงเป็นรูปภาพ ตาราง และตัวอักษร

โปรแกรมวิเคราะห์การหมุนของกระดาษผิด แสดงไฟล์ภาพเอกสารจากแนวตั้งเป็นแนวนอน วิเคราะห์ไฟล์ภาพเอกสารเป็นรูปภาพ ตาราง และข้อความ จากเอกสารจริงที่เป็นข้อความอย่างเดียว ดังนั้น ต้องแก้ไขโดยการหมุนภาพใหม่ก่อนการแปลงไฟล์ ถ้าไม่ได้อีกก็นำเอกสารมากราดภาพใหม่

2) ไฟล์ภาพเอกสารแบ่งเป็น 2 หน้า ดังภาพประกอบ 3-8



ภาพประกอบ 3-8 ตัวอย่างการแบ่งหน้ากระดาษเป็น 2 หน้า แล้วแปลงเป็นรูปภาพ

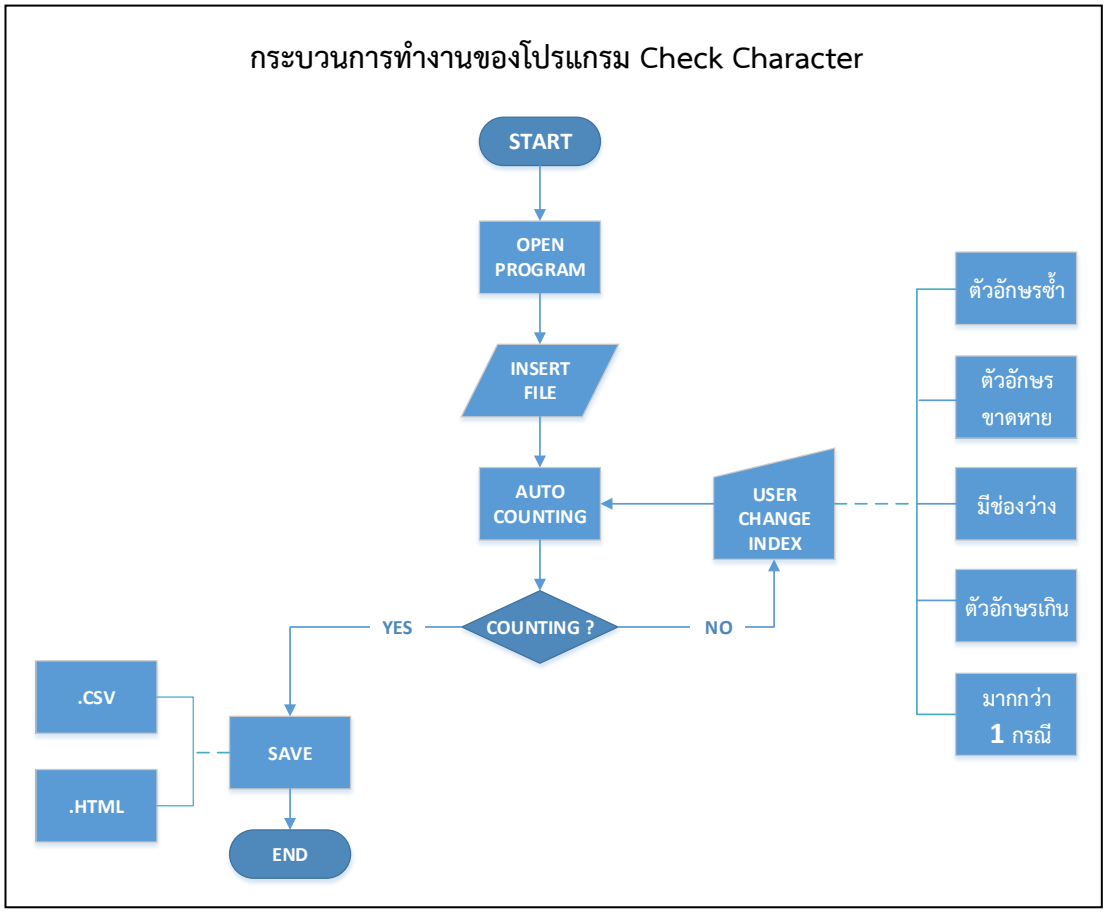
บางไฟล์ภาพตัดมาเป็น 2 หน้า ต้องนำเอกสารมากราดภาพใหม่ และทำการรู้จำใหม่

3.3 วัดผลการรู้จำ

วัดผลการรู้จำโดยการนำไฟล์ .txt 1,620 ไฟล์ ที่ได้จากการรู้จำมานับตัวอักษรถูก/ผิด ด้วยโปรแกรม Check Character ที่พัฒนาขึ้นด้วยโปรแกรม Microsoft Visual Studio 2012 บันทึกเป็นไฟล์ .csv และ .html

โปรแกรม Check Character ที่พัฒนาขึ้นเพื่อนับจำนวนตัวอักษรถูก/ผิด ของการรู้จำตัวอักษรดอทเมทริกซ์แยกตามปัจจัย เพื่อให้ทราบว่าตัวอักษรตัวไหนถูก ตัวไหนผิด เป็นจำนวนกี่

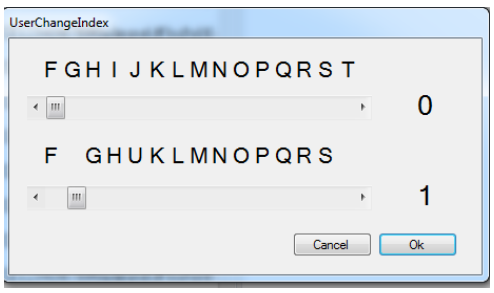
ตัวอักษรเพื่อใช้ในการวิเคราะห์ผลการรู้จำ โปรแกรม Check Character มีกระบวนการทำงานดังภาพประกอบ 3-19



ภาพประกอบ 3-9 กระบวนการทำงานของโปรแกรม Check Character

3.3.1 ปัญหาการสร้างโปรแกรม Check Character

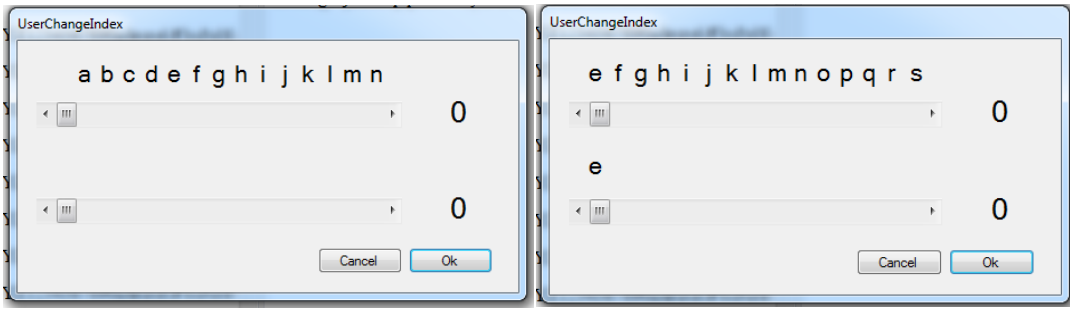
การสร้างโปรแกรม Check Character เพื่อช่วยในการนับคำถูก/ผิด ในครั้งแรกหน้าต่าง User Change Index สามารถเลื่อนเปรียบเทียบได้เพียงครั้งละ 15 ตัวอักษร ดังภาพประกอบ 3-10



ภาพประกอบ 3-10 หน้าต่าง User Change Index ก่อนปรับปรุง

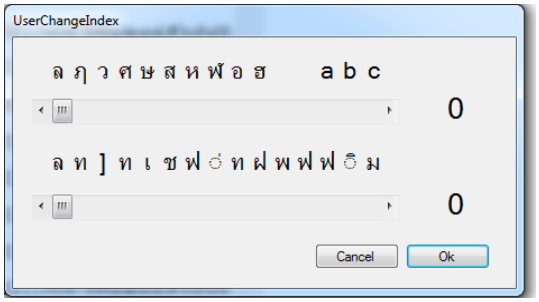
ทำให้ไฟล์ .txt บางไฟล์ที่มีช่องว่างมาก ตัวอักษรซ้ำ ตัวอักษรขาดหาย ตัวอักษรเกิน ไม่มีตัวอักษร ความผิดพลาดมากกว่า 15 ตัวอักษร หรือหายไปทั้งบรรทัด ทำให้ไม่สามารถเลื่อนนับได้ ไม่ได้ผลการทดลองจากไฟล์นั้น เมื่อไปสู่ขั้นตอนการหาค่าความถูกต้องและความผิดพลาดของการรู้จำ ไม่สามารถเปรียบเทียบผลการทดลองได้ ทำให้ตัวอย่างการทดลองไม่เท่ากัน ผิดพลาดลักษณะนี้ 74 ไฟล์ แยกเป็นกรณีดังนี้

1) กรณีช่องว่างมาก ดังภาพประกอบ 3-11



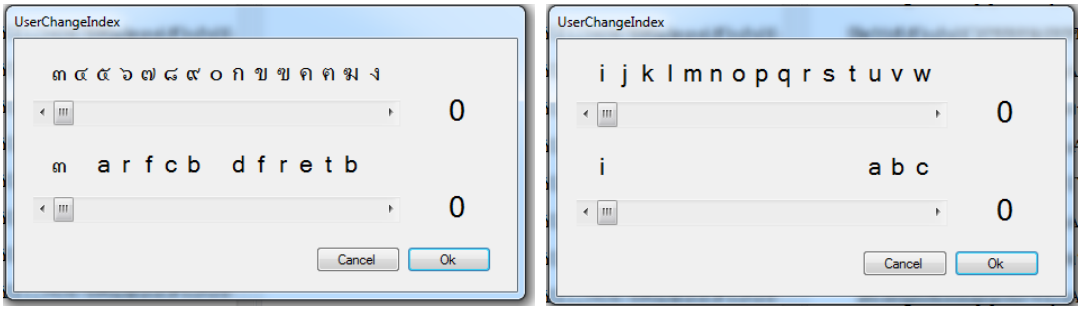
ภาพประกอบ 3-11 ตัวอย่างช่องว่างมาก

2) กรณีตัวอักษรซ้ำ ดังภาพประกอบ 3-12



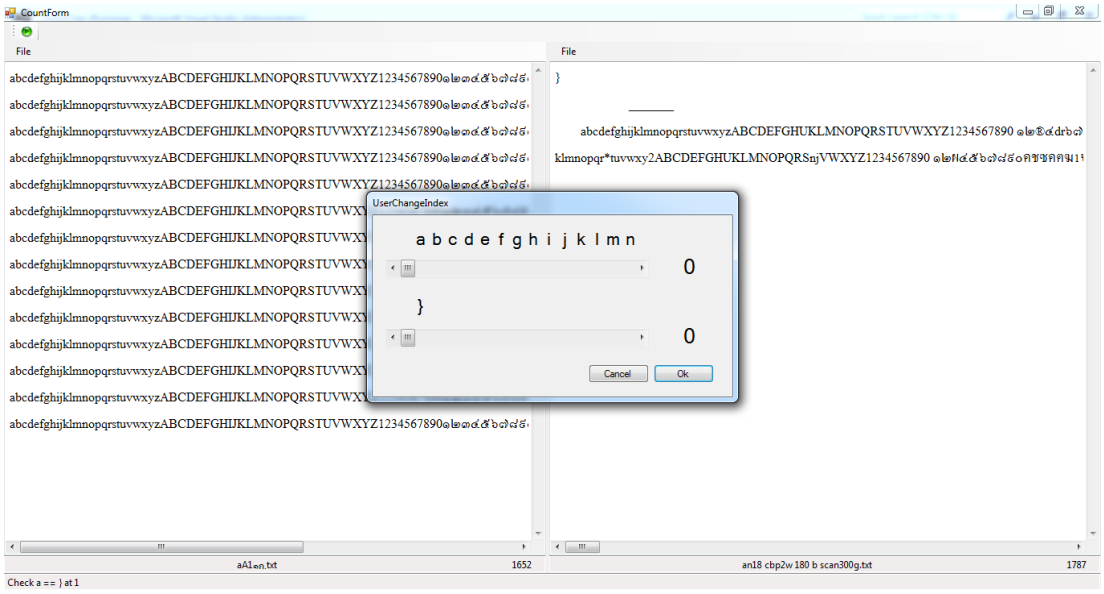
ภาพประกอบ 3-12 ตัวอย่างตัวอักษรซ้ำ

3) กรณีตัวอักษรขาดหาย ดังภาพประกอบ 3-13



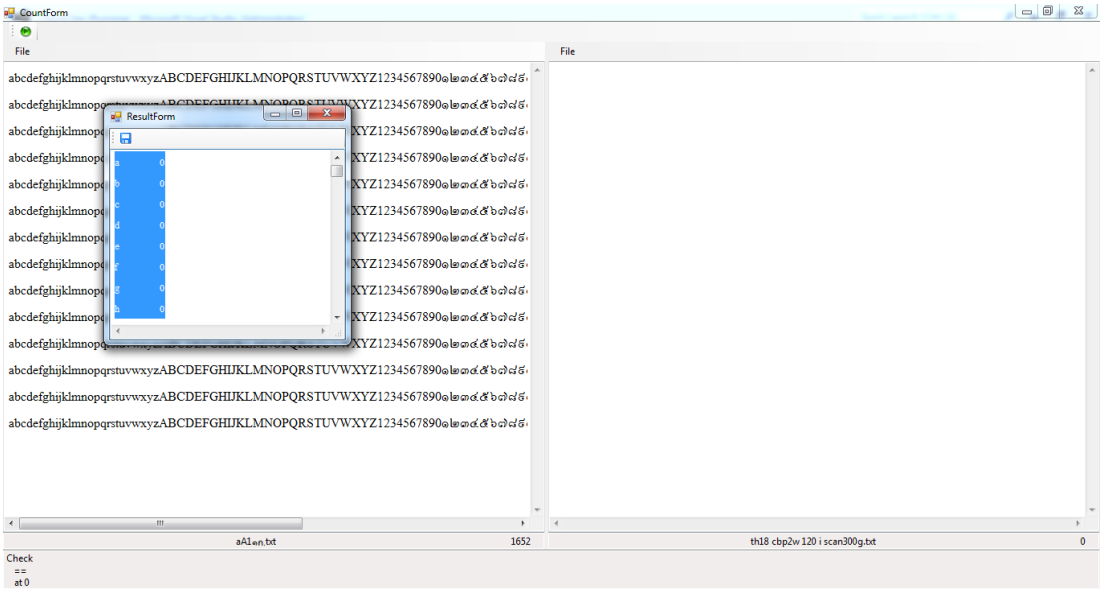
ภาพประกอบ 3-13 ตัวอย่างตัวอักษรขาดหาย

4) กรณีตัวอักษรเกิน ดังภาพประกอบ 3-14



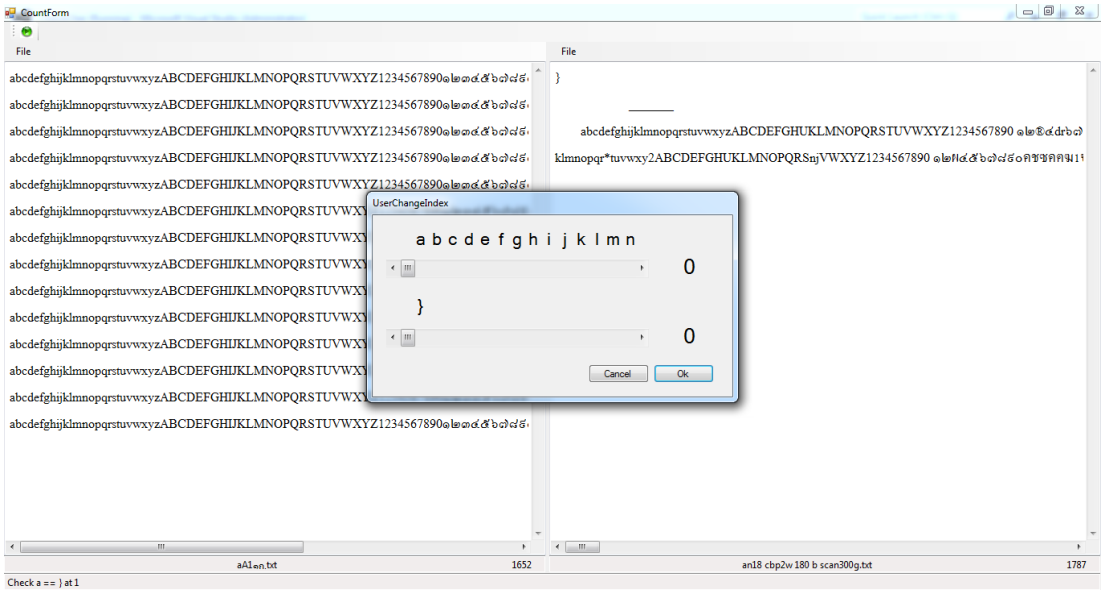
ภาพประกอบ 3-14 ตัวอย่างตัวอักษรเกิน

5) กรณีไม่มีตัวอักษร ดังภาพประกอบ 3-15



ภาพประกอบ 3-15 ตัวอย่างไม่มีตัวอักษร

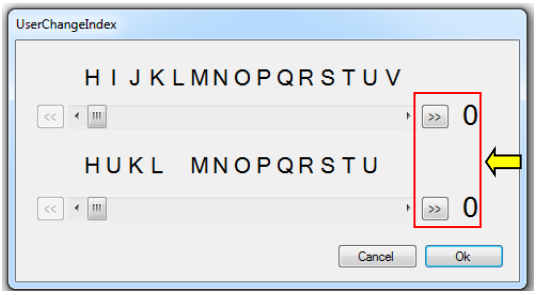
6) กรณีมากกว่า 1 แบบ ดังภาพประกอบ 3-16



ภาพประกอบ 3-16 ตัวอย่างตัวอักษรเกินและช่องว่างมาก

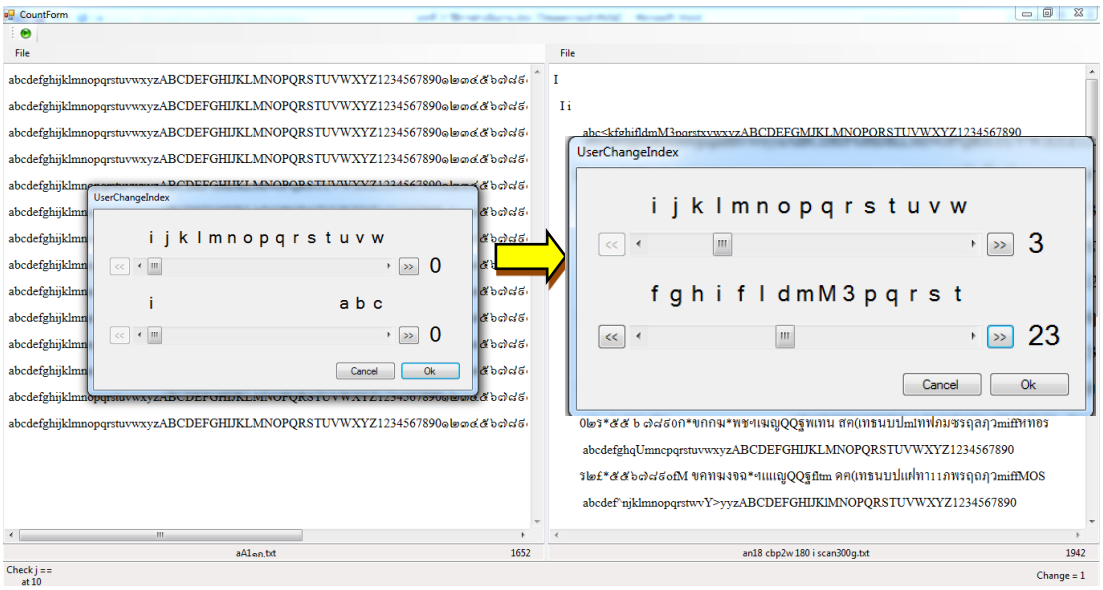
3.3.2 การปรับปรุงโปรแกรม Check Character

จากปัญหาดังกล่าวข้างต้น จึงต้องทำการปรับปรุงโปรแกรม Check Character ดังภาพประกอบ 3-17



ภาพประกอบ 3-17 หน้าต่าง User Change Index หลังปรับปรุง

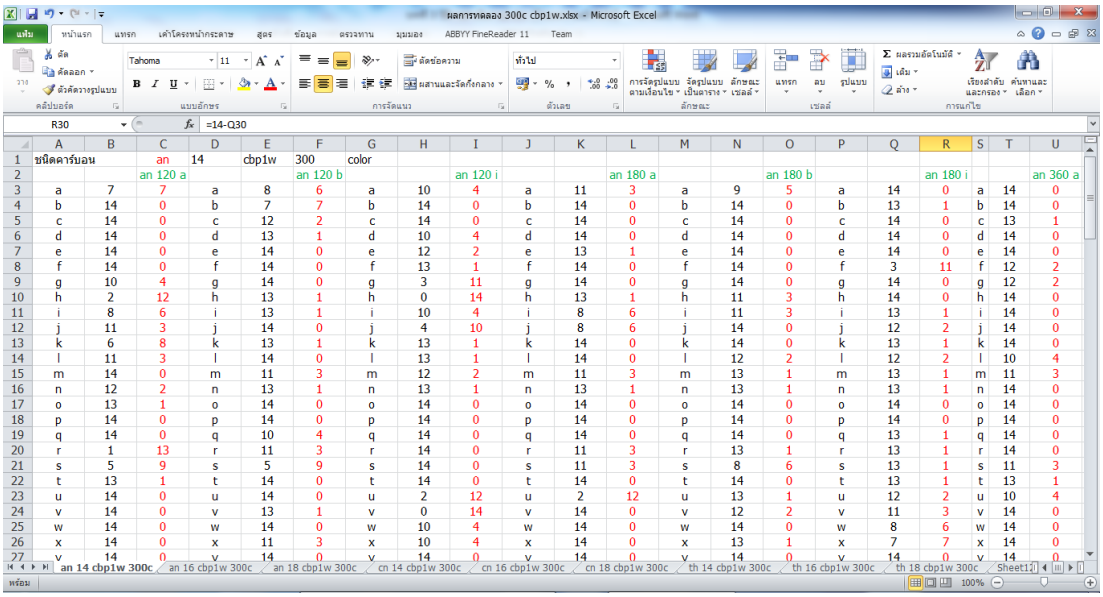
ปรับปรุงโปรแกรมให้สามารถเลื่อนนับได้ไม่จำกัด และนำไฟล์ที่ไม่สามารถนับได้มานับใหม่ทั้งหมด บางไฟล์กราดภาพภาพใหม่ ทำการรู้จำใหม่ นับใหม่ จนได้ไฟล์ครบทุกไฟล์สามารถนำไปสรุปผลการทดลองได้ ดังภาพประกอบ 3-18



ภาพประกอบ 3-18 ตัวอย่างที่สามารถเลื่อนนับได้หลังปรับปรุง

3.4 หาค่าความถูกต้องและค่าความผิดพลาดของการรู้จำแยกตามปัจจัย

หาค่าความถูกต้องและค่าความผิดพลาด โดยการนำไฟล์ .csv 1,620 ไฟล์ที่ได้จากการวัดผลการรู้จำ ด้วยการนับตัวอักษรถูก/ผิด ประมวลผล วิเคราะห์ และเปรียบเทียบ ด้วยโปรแกรม Microsoft Excel ดังภาพประกอบ 3-19



ภาพประกอบ 3-19 การนำไฟล์ .csv มาหาค่าความถูกต้อง/ผิดพลาดแยกตามปัจจัย

การประมวลผล เปรียบเทียบ วิเคราะห์ปัจจัยที่มีผลต่อการรู้จำเอกสารดอทเมทริกซ์ ชนิดคาร์บอน และชนิดเคมี โดยแบ่งเป็นอักษรอังกฤษ อักษรไทย เลขอารบิค และเลขไทย โดยการวิเคราะห์หาค่าความถูกต้อง/ความผิดพลาดตามปัจจัยต่างๆ ดังต่อไปนี้

ชนิดแบบอักษร จะหาค่าจากไฟล์ .csv ที่ได้จากการหาค่าความถูกต้องของการรู้จำจากชนิดแบบอักษร Angsana New, Cordia New และ TH SarabunPSK ทุกขนาดตัวอักษร ทุกลักษณะตัวอักษร ทุกความละเอียดในการพิมพ์ และทุกความละเอียดและชนิดในการกราดภาพ

ขนาดตัวอักษร จะหาค่าจากไฟล์ .csv ที่ได้จากการหาค่าความถูกต้องของการรู้จำจากขนาดตัวอักษร 14 16 และ 18 ทุกชนิดแบบอักษร ทุกลักษณะตัวอักษร ทุกความละเอียดในการพิมพ์ และทุกความละเอียดและชนิดในการกราดภาพ

ลักษณะตัวอักษร จะหาค่าจากไฟล์ .csv ที่ได้จากการหาค่าความถูกต้องของการรู้จำจากลักษณะตัวอักษรตัวปกติ ตัวหนา และตัวเอียง ทุกชนิดแบบอักษร ทุกขนาดตัวอักษร ทุกความละเอียดในการพิมพ์ และทุกความละเอียดและชนิดในการกราดภาพ

ความละเอียดในการพิมพ์ จะหาค่าจากไฟล์ .csv ที่ได้จากการหาค่าความถูกต้องของการรู้จำ จากความละเอียดในการพิมพ์ที่ 120 x 180 dpi, 180 x180 dpi และ 360 x 180 dpi ทุกชนิดแบบอักษร ทุกลักษณะตัวอักษร ทุกขนาดตัวอักษร และทุกความละเอียดและชนิดในการกราดภาพ

ความละเอียดและชนิดในการกราดภาพ จะหาค่าจากไฟล์ .csv ที่ได้จากการหาค่าความถูกต้องของการการรู้จำ จากความละเอียดในกราดภาพ 300 และ 600 dpi ทั้งชนิด Color และ Gray ทุกชนิดแบบอักษร ทุกลักษณะตัวอักษร ทุกขนาดตัวอักษร ทุกความละเอียดในการพิมพ์

บทที่ 4

ผลการดำเนินงาน

บทนี้จะกล่าวถึง ผลการรู้จำโดยรวม ผลการรู้จำเฉพาะอักษรไทย ผลการรู้จำเฉพาะอักษรอังกฤษ ผลการรู้จำเฉพาะเลขไทย และผลการรู้จำเฉพาะเลขอารบิค แยกตามปัจจัยชนิดแบบอักษร ขนาดตัวอักษร ลักษณะตัวอักษร ความละเอียดในการพิมพ์ และความละเอียดและชนิดในการกราดภาพ

4.1 ผลการรู้จำโดยรวม

จากการทดลอง ประมวลผล เปรียบเทียบ วิเคราะห์ปัจจัยที่มีผลต่อการรู้จำเอกสารดอทเมทริกซ์ ชนิดคาร์บอน และชนิดเคมี ของเอกสารทั้งหมด (อักษรไทย อักษรอังกฤษ เลขไทย และเลขอารบิค) ได้ผลการทดลอง ดังตาราง 4-1

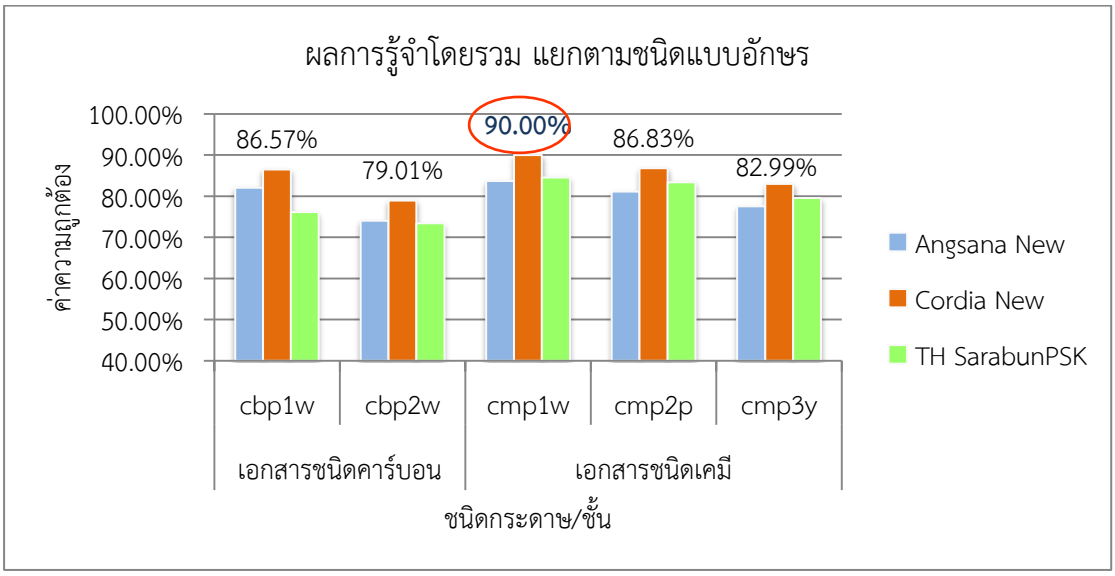
ตาราง 4-1 ผลการรู้จำโดยรวม

ปัจจัย	ชนิดแบบฟอร์ม	ค่าความถูกต้อง (%)									
		เอกสารชนิดคาร์บอน				เอกสารชนิดเคมี					
		cbp1w		cbp2w		cmp1w		cmp2p		cmp3y	
		$\bar{x}$	S.D.	$\bar{x}$	S.D.	$\bar{x}$	S.D.	$\bar{x}$	S.D.	$\bar{x}$	S.D.
ชนิดแบบอักษร	Angsana New	82.11	1.97	74.08	2.56	83.72	1.93	81.12	2.81	77.57	1.45
	Cordia New	86.57	1.30	79.01	1.53	90.00	2.75	86.83	0.74	82.99	1.70
	TH SarabunPSK	76.14	13.41	73.41	2.45	84.61	3.34	83.43	2.26	79.63	1.28
ขนาดตัวอักษร	14	78.93	2.19	68.50	1.57	81.69	1.63	79.48	1.35	74.05	1.87
	16	86.08	1.79	78.04	2.32	87.61	1.90	86.15	0.99	82.38	1.41
	18	79.80	13.91	79.97	2.50	86.80	3.48	85.75	3.49	83.75	1.18
ลักษณะตัวอักษร	ตัวปกติ	82.98	4.39	76.85	1.49	85.91	3.40	86.45	0.44	81.56	1.17
	ตัวหนา	83.64	4.42	75.77	3.00	87.58	1.13	84.28	2.38	81.13	1.11
	ตัวเอียง	78.19	3.73	75.05	5.18	82.61	2.48	80.65	3.37	77.49	2.05
ความละเอียดในการพิมพ์	120 x 180 dpi	77.56	3.95	72.59	1.20	81.40	3.22	78.99	3.90	75.66	2.08
	180 x 180 dpi	80.39	3.93	72.59	3.25	84.91	1.68	83.80	1.21	78.82	1.43
	360 x 180 dpi	86.23	2.90	78.84	2.50	87.01	3.58	86.30	1.55	83.99	1.55
ความละเอียดและชนิดในการกราดภาพ	300 dpi Color	82.31	4.57	73.13	5.56	83.23	2.71	83.50	3.91	79.34	5.76
	300 dpi Gray	82.71	5.03	74.36	5.94	83.48	3.02	81.22	3.61	78.39	5.47
	600 dpi Color	85.57	4.14	77.02	6.79	87.23	3.78	85.32	3.82	81.36	4.66
	600 dpi Gray	75.83	14.95	77.50	6.33	87.54	3.67	85.13	4.40	81.15	5.13
ค่าเฉลี่ย		**81.57**	**5.41**	75.42	**3.39**	85.33	**2.73**	**83.65**	**2.51**	**79.95**	**2.46**

หมายเหตุ ตัวเลขสีน้ำเงิน หมายถึง ค่ามากที่สุด ตัวเลขสีแดง หมายถึง ค่าน้อยที่สุด

4.1.1 ผลการรู้จำโดยรวม แยกตามชนิดแบบอักษร

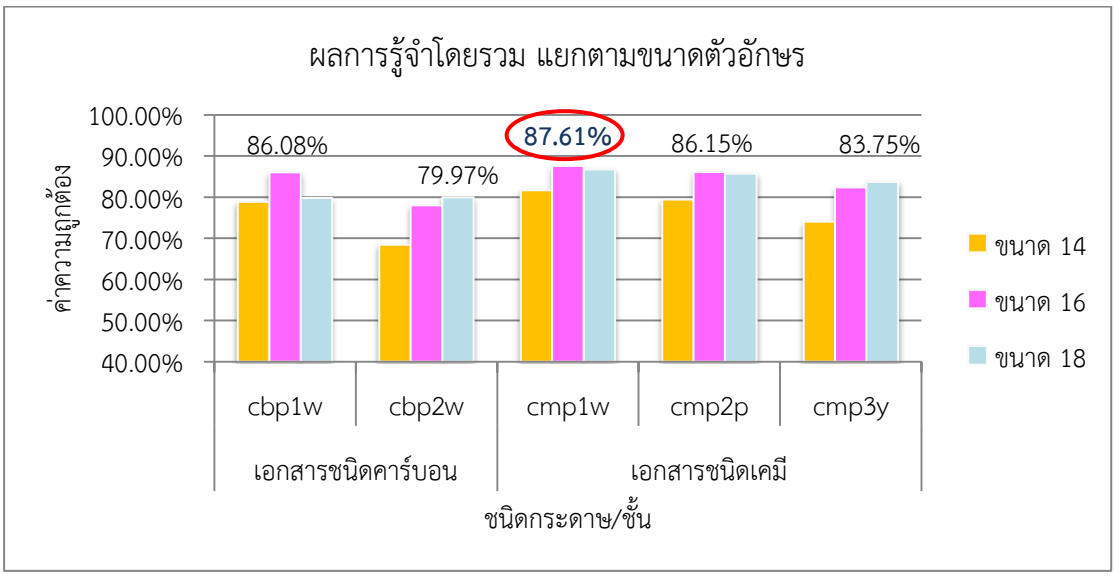
จากตาราง 4-1 ในส่วนของปัจจัยชนิดแบบอักษร สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-1 จากการดำเนินงาน ชนิดแบบอักษร Cordia New ให้ค่าความถูกต้องสูงสุด ทั้งเอกสารชนิดคาร์บอน cbp1w เท่ากับ 86.57% และเอกสารชนิดเคมี cmp1w เท่ากับ 90.00% และเอกสารสำเนาชนิดเคมีทั้ง cmp2p และ cmp3y ให้ค่าความถูกต้องสูงกว่าเอกสารสำเนาชนิดคาร์บอน cbp2w เท่ากับ 7.82% และ 3.98% ตามลำดับ



ภาพประกอบ 4-1 ผลการรู้จำโดยรวม แยกตามชนิดแบบอักษร

4.1.2 ผลการรู้จำโดยรวม แยกตามขนาดตัวอักษร

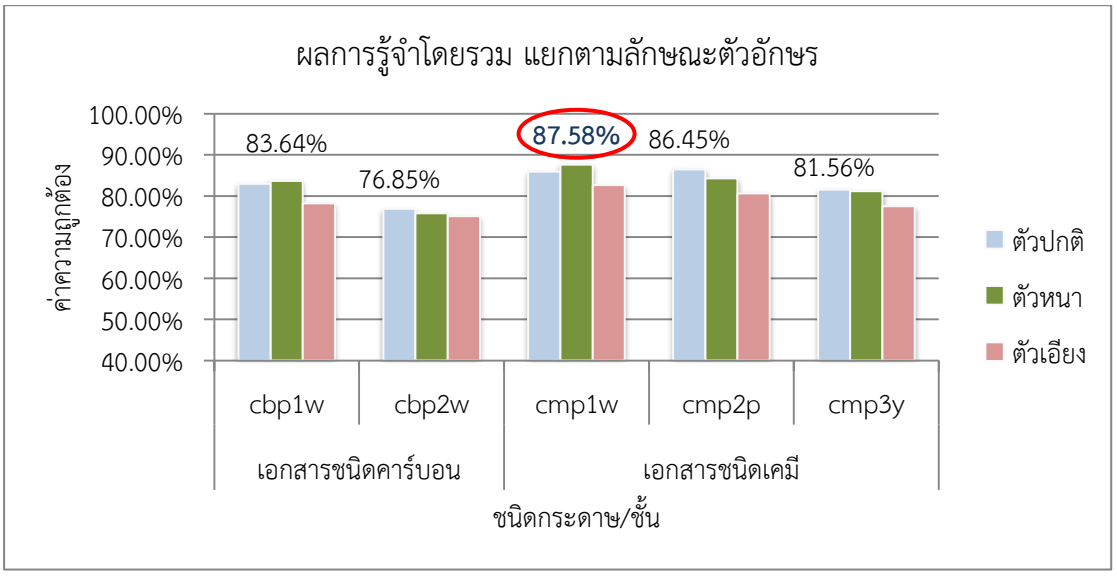
จากตาราง 4-1 ในส่วนของปัจจัยขนาดตัวอักษร สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-2 จากการดำเนินงาน เอกสารชนิดเคมี cmp1w ขนาดตัวอักษร 16 ให้ค่าความถูกต้องสูงสุด เท่ากับ 87.61% และเอกสารชนิดคาร์บอน cmp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 86.08% ซึ่งใกล้เคียงกัน และเอกสารสำเนาชนิดเคมีทั้ง cmp2p และ cmp3y ให้ค่าความถูกต้องสูงกว่าเอกสารสำเนาชนิดคาร์บอน cbp2w เท่ากับ 6.18% และ 3.78% ตามลำดับ



ภาพประกอบ 4-2 ผลการรู้จำโดยรวม แยกตามขนาดตัวอักษร

4.1.3 ผลการรู้จำโดยรวม แยกตามลักษณะตัวอักษร

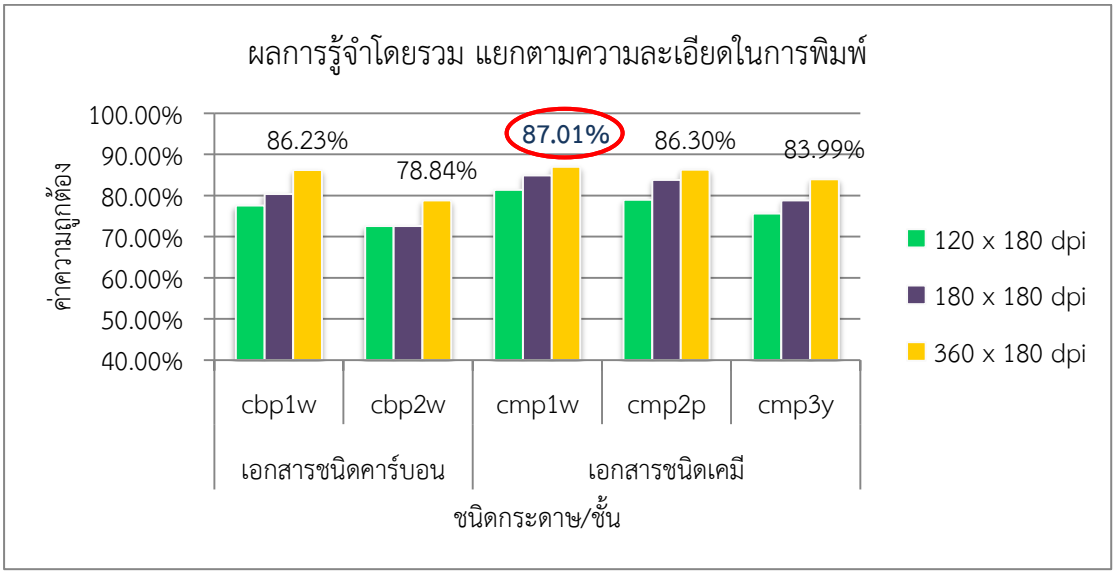
จากตาราง 4-1 ในส่วนของปัจจัยลักษณะตัวอักษร สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-3 จากการดำเนินงาน ลักษณะตัวอักษร ตัวหนา ให้ค่าความถูกต้องสูงสุด ทั้งเอกสารชนิดคาร์บอน cbp1w เท่ากับ 83.64% และเอกสารชนิดเคมี cmp1w เท่ากับ 87.58% และเอกสารสำเนาชนิดเคมีทั้ง cmp2p และ cmp3y ให้ค่าความถูกต้องสูงกว่าเอกสารสำเนาชนิดคาร์บอน cbp2w เท่ากับ 9.60% และ 4.71% ตามลำดับ



ภาพประกอบ 4-3 ผลการรู้จำโดยรวม แยกตามลักษณะตัวอักษร

4.1.4 ผลการรู้จำโดยรวม แยกตามความละเอียดในการพิมพ์

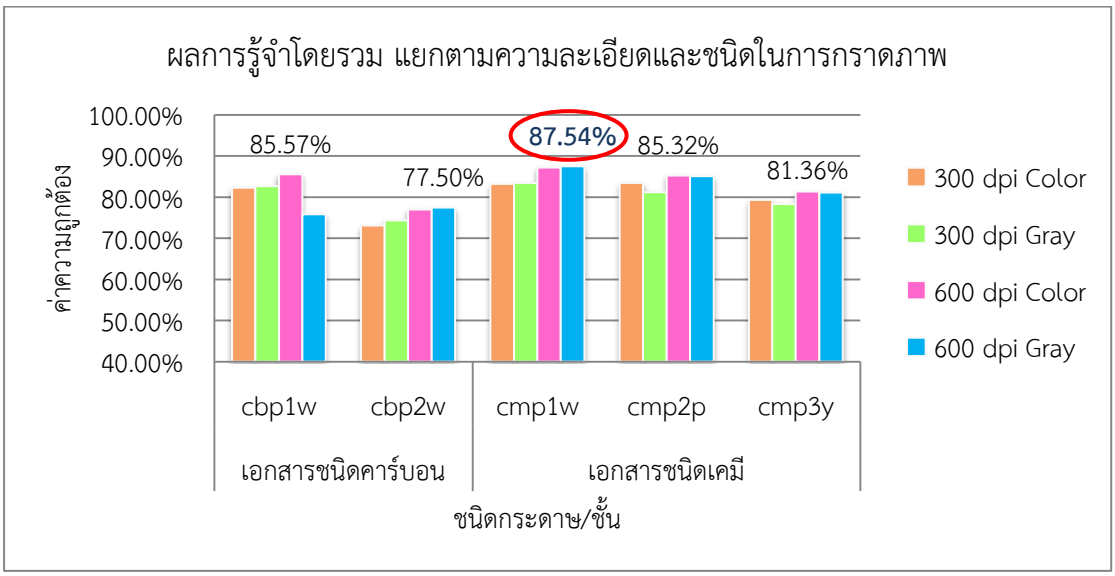
จากตาราง 4-1 ในส่วนของปัจจัยความละเอียดในการพิมพ์ สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-4 จากการดำเนินงาน เอกสารชนิดเคมี cmp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 87.01% ที่การพิมพ์ที่ความละเอียด 360 x 180 dpi เอกสารชนิดคาร์บอน cbp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 86.23% ซึ่งได้ผลใกล้เคียงกัน และเอกสารชนิดเคมีทั้ง cmp2p และ cmp3y ให้ค่าความถูกต้องสูงกว่าเอกสารสำเนาชนิดคาร์บอน cbp2w เท่ากับ 7.46% และ 5.15% ตามลำดับ



ภาพประกอบ 4-4 ผลการรู้จำโดยรวม แยกตามความละเอียดในการพิมพ์

4.1.5 ผลการรู้จำโดยรวม แยกตามความละเอียดและชนิดในการกราดภาพ

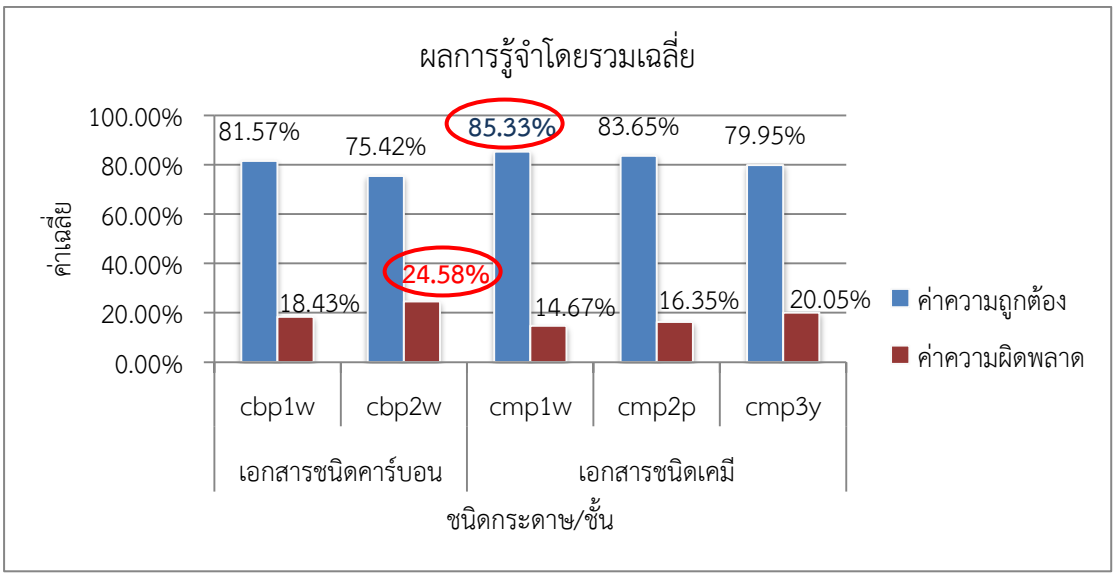
จากตาราง 4-1 ในส่วนของปัจจัยความละเอียดและชนิดในการกราดภาพ สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-5 จากการดำเนินงาน ความละเอียดและชนิดในการกราดภาพของเอกสารชนิดเคมี cmp1w ที่กราดภาพที่ความละเอียด 600 dpi ชนิด Gray ให้ค่าความถูกต้องสูงสุด เท่ากับ 87.54% และเอกสารชนิดคาร์บอน cbp1w ที่กราดภาพที่ความละเอียด 600 dpi ชนิด Color ให้ค่าความถูกต้องสูงสุดเท่ากับ 85.57% ซึ่งใกล้เคียงกัน และเอกสารชนิดเคมีทั้ง cmp2p และ cmp3y ให้ค่าความถูกต้องสูงกว่าเอกสารสำเนาชนิดคาร์บอน cbp2w เท่ากับ 7.82% และ 3.86% ตามลำดับ



ภาพประกอบ 4-5 ผลการรู้จำโดยรวม แยกตามความละเอียดและชนิดในการกราดภาพ

4.1.6 ผลการรู้จำโดยรวมเฉลี่ย

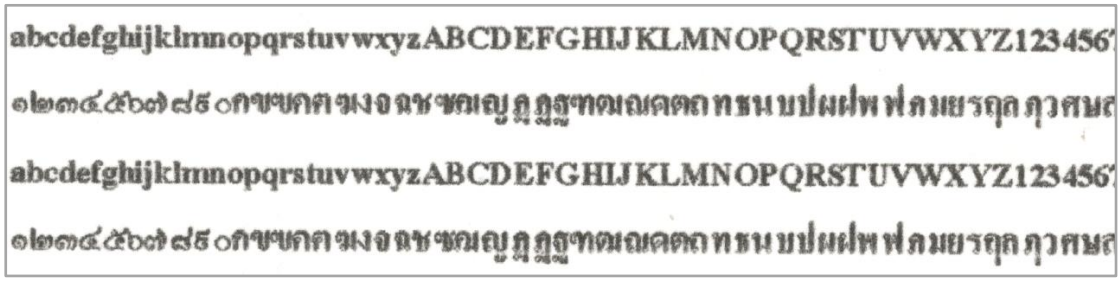
จากตาราง 4-1 สามารถสรุปเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-6 จากผลการทดลองได้ผลลัพธ์ค่าเฉลี่ยโดยรวม เอกสารชนิดเคมี cmp1w ให้ค่าความถูกต้องสูงสุดทุกปัจจัย เท่ากับ 85.33% และเอกสารชนิดคาร์บอน cbp1w ให้ค่าความถูกต้องเฉลี่ย เท่ากับ 81.57% และเอกสารสำเนาชนิดคาร์บอน cbp2w ให้ค่าความถูกต้องเฉลี่ยต่ำสุด เท่ากับ 75.42% ซึ่งยังต่ำกว่าเอกสารสำเนาชนิดเคมีทั้ง cmp2p และ cmp3y เท่ากับ 8.23% และ 4.53% ตามลำดับ



ภาพประกอบ 4-6 ผลการรู้จำโดยรวมเฉลี่ย

สรุปได้ว่า เอกสารที่พิมพ์ด้วยกระดาษชนิดเคมีสามารถรู้จำอักขระด้วยแสงดีกว่าเอกสารที่พิมพ์ด้วยกระดาษชนิดคาร์บอนทั้งเอกสารหน้าแรก (cmp1w) เอกสารสำเนาที่ 2 (cmp2p) และเอกสารสำเนาที่ 3 (cmp3y) สามารถรู้จำอักขระด้วยแสงดีกว่าเอกสารสำเนาที่ 2 (cbp2w) ที่มีค่าความผิดพลาดสูงสุด เท่ากับ 24.58% ดังนั้นเพื่อประสิทธิภาพของการรู้จำอักขระด้วยแสงควรเลือกใช้กระดาษชนิดเคมีเพราะมีคุณภาพดีกว่าชนิดคาร์บอน และจำนวนชั้นของสำเนามีผลการรู้จำแทบจะไม่แตกต่างกัน

เอกสารสำเนาที่ 2 ชนิดคาร์บอนจะมีสิ่งแปลกปลอมที่เกิดจากหมึกเลอะบนกระดาษหมึกกระจาย ดังภาพประกอบ 4-7 ทำให้การรู้จำผิดพลาดมากกว่าเอกสารสำเนาชนิดเคมี



ภาพประกอบ 4-7 ตัวอย่างภาพเอกสารสำเนาชนิดคาร์บอน

4.2 ผลการรู้จำเฉพาะอักษรไทย

จากการทดลอง ประมวลผล เปรียบเทียบ วิเคราะห์ปัจจัยที่มีผลต่อการรู้จำเอกสารดอทเมทริกซ์ ชนิดคาร์บอน และชนิดเคมี ของเอกสารอักษรไทย ได้ผลการทดลอง ดังตาราง 4-2

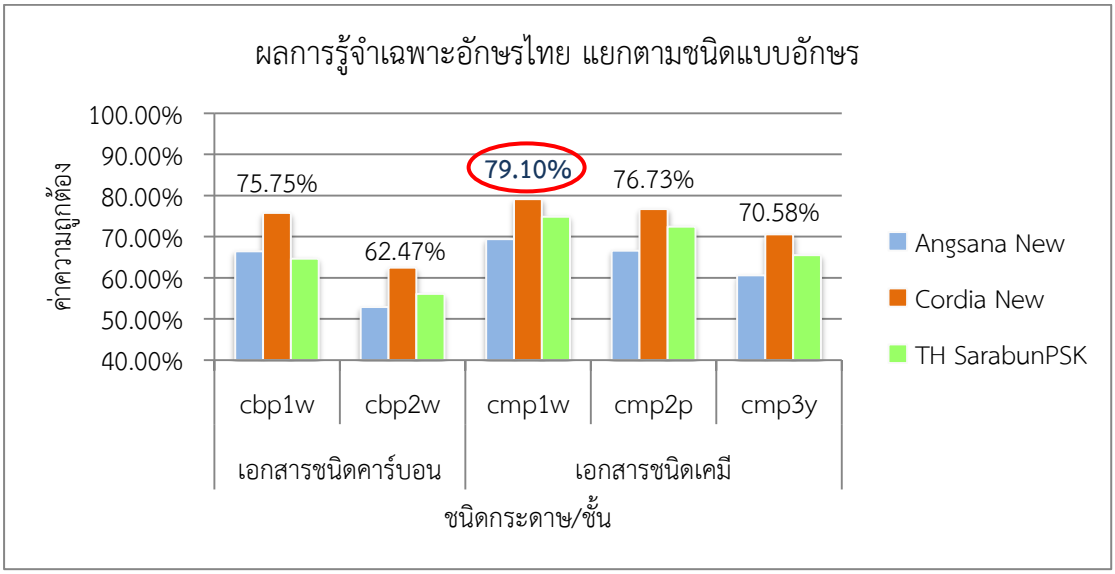
ตาราง 4-2 ผลการรู้จำเฉพาะอักษรไทย

ปัจจัย	ชนิดแบบฟอร์ม	ค่าความถูกต้อง (%)									
		เอกสารชนิดคาร์บอน				เอกสารชนิดเคมี					
		cbp1w		cbp2w		cmp1w		cmp2p		cmp3y	
		$\bar{x}$	S.D.	$\bar{x}$	S.D.	$\bar{x}$	S.D.	$\bar{x}$	S.D.	$\bar{x}$	S.D.
ชนิดแบบอักษร	Angsana New	66.47	3.78	52.93	4.76	69.42	3.75	66.62	3.22	60.64	2.18
	Cordia New	75.75	2.95	62.47	2.78	79.10	3.43	76.73	1.42	70.58	3.66
	TH SarabunPSK	64.67	11.04	56.10	3.22	74.86	5.91	72.40	2.42	65.49	1.81
ขนาดตัวอักษร	14	61.97	3.67	44.94	2.47	67.04	2.42	63.96	1.78	55.70	2.18
	16	74.38	3.87	60.44	4.35	77.44	4.02	75.12	1.85	68.88	2.93
	18	70.55	11.79	66.12	3.78	78.02	5.23	76.67	3.41	72.14	2.47
ลักษณะตัวอักษร	ตัวปกติ	72.95	3.65	63.07	4.57	77.55	4.74	77.80	1.25	69.94	1.58
	ตัวหนา	71.71	4.02	56.08	4.10	76.96	2.39	71.69	3.79	65.93	2.45
	ตัวเอียง	62.23	4.10	57.54	8.76	67.98	4.45	66.26	3.95	60.85	3.30
ความละเอียดในการพิมพ์	120 x 180 dpi	62.74	3.61	53.27	2.23	68.18	4.06	65.45	2.76	59.22	3.39
	180 x 180 dpi	66.84	3.75	52.89	3.74	73.03	3.16	71.50	1.70	63.21	2.18
	360 x 180 dpi	75.42	2.26	60.39	4.42	76.55	4.96	74.46	3.31	71.25	3.20
ความละเอียดและชนิดในการกราดภาพ	300 dpi Color	67.66	8.14	53.25	10.01	70.21	4.34	70.27	6.99	63.70	9.32
	300 dpi Gray	68.72	8.55	55.20	10.50	71.53	5.82	69.65	5.87	63.33	8.00
	600 dpi Color	73.98	7.79	59.50	11.89	77.11	7.70	73.99	6.87	67.80	8.55
	600 dpi Gray	65.51	12.43	60.71	11.49	77.80	6.94	73.74	8.10	67.44	9.10
ค่าเฉลี่ย		**68.85**	**5.96**	57.18	**5.82**	73.93	**4.58**	**71.64**	**3.67**	**65.38**	**4.14**

หมายเหตุ ตัวเลขสีน้ำเงิน หมายถึง ค่ามากที่สุด ตัวเลขสีแดง หมายถึง ค่าน้อยที่สุด

4.2.1 ผลการรู้จำเฉพาะอักษรไทย แยกตามชนิดแบบอักษร

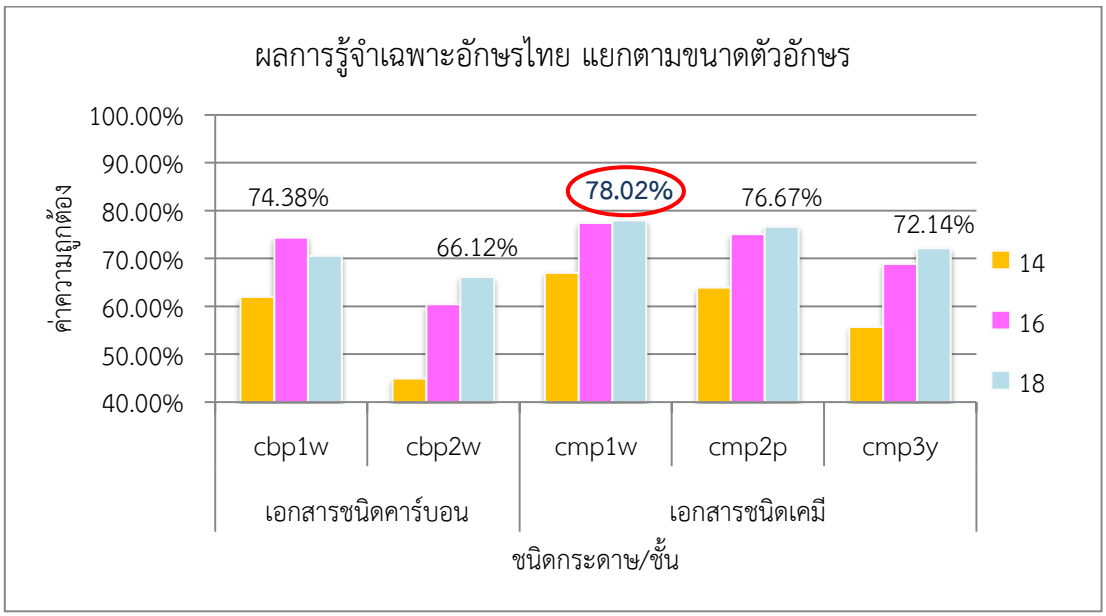
จากตาราง 4-2 ในส่วนของปัจจัยชนิดแบบอักษร สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-8 จากการดำเนินงาน ชนิดแบบอักษร Cordia New ให้ค่าความถูกต้องสูงสุด ทั้งเอกสารชนิดคาร์บอน cbp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 75.75% และเอกสารชนิดเคมี cmp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 79.10% และเอกสารสำเนาชนิดเคมีทั้ง cmp2p และ cmp3y ให้ค่าความถูกต้องสูงกว่าเอกสารสำเนาชนิดคาร์บอน เท่ากับ 14.26% และ 8.12% ตามลำดับ



ภาพประกอบ 4-8 ผลการรู้จำเฉพาะอักษรไทย แยกตามชนิดแบบอักษร

4.2.2 ผลการรู้จำเฉพาะอักษรไทย แยกตามขนาดตัวอักษร

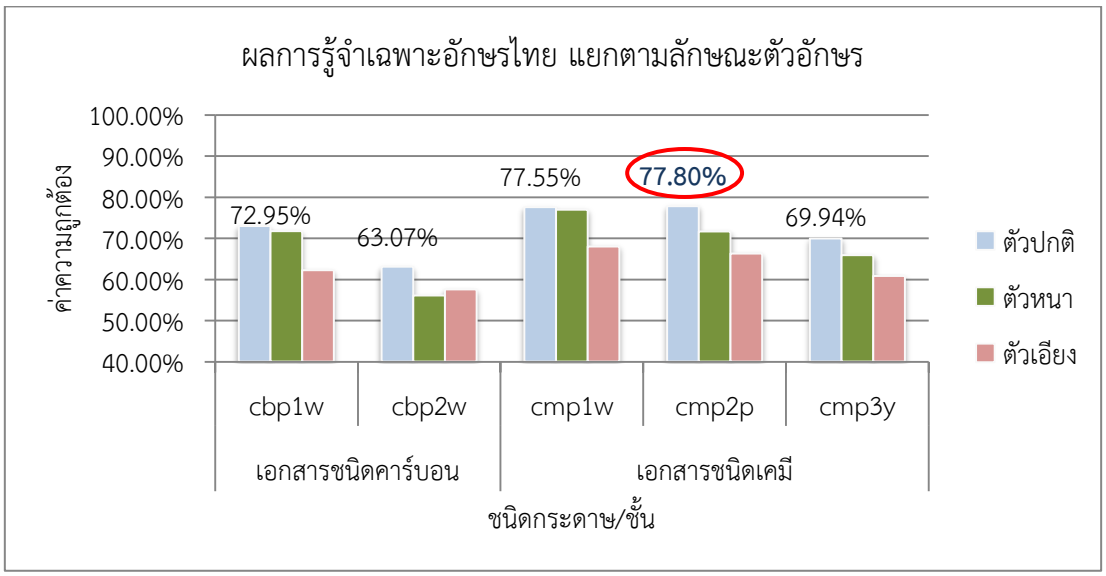
จากตาราง 4-2 ในส่วนของปัจจัยขนาดตัวอักษร สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-9 จากการดำเนินงาน เอกสารชนิดเคมี cmp1w ขนาดตัวอักษร 18 ให้ค่าความถูกต้องสูงสุด เท่ากับ 78.02% และเอกสารชนิดคาร์บอน cmp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 74.38% และเอกสารสำเนาชนิดเคมีทั้ง cmp2p และ cmp3y ให้ค่าความถูกต้องของตัวอักษรขนาด 18 สูงกว่าเอกสารสำเนาชนิดคาร์บอน cmp2w เท่ากับ 10.55% และ 6.02% ตามลำดับ



ภาพประกอบ 4-9 ผลการรู้จำเฉพาะอักษรไทย แยกตามขนาดตัวอักษร

4.2.3 ผลการรู้จำเฉพาะอักษรไทย แยกตามลักษณะตัวอักษร

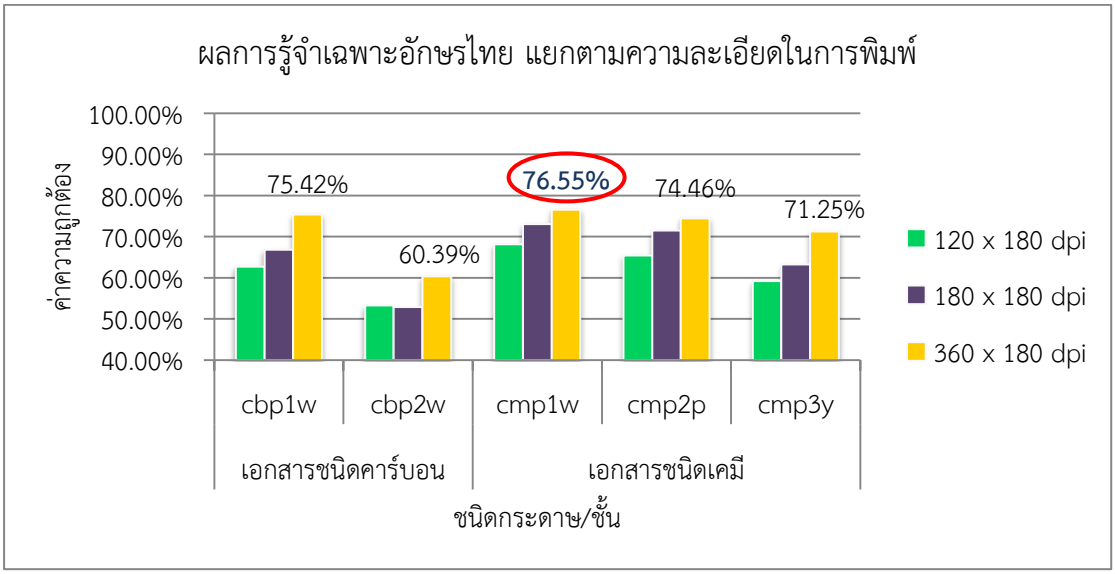
จากตาราง 4-2 ในส่วนของปัจจัยลักษณะตัวอักษร สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-10 จากการดำเนินงาน ลักษณะตัวอักษร ตัวปกติ ให้ค่าความถูกต้องสูงสุด ทั้งเอกสารชนิดคาร์บอน cbp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 72.95% และเอกสารชนิดเคมี cmp2p ให้ค่าความถูกต้องสูงสุดเท่ากับ 77.80% และเอกสารสำเนาชนิดเคมีทั้ง cmp1w และ cmp3y ให้ค่าความถูกต้องสูงกว่าเอกสารสำเนาชนิดคาร์บอน cbp2w เท่ากับ 14.73% และ 6.87% ตามลำดับ



ภาพประกอบ 4-10 ผลการรู้จำเฉพาะอักษรไทย แยกตามลักษณะตัวอักษร

4.2.4 ผลการรู้จำเฉพาะอักษรไทย แยกตามความละเอียดในการพิมพ์

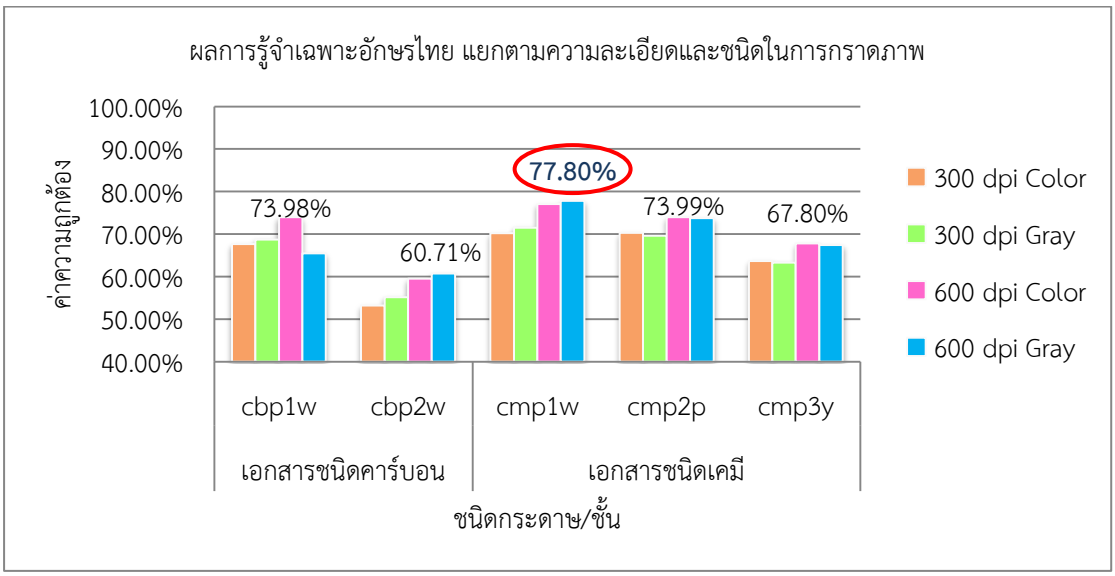
จากตาราง 4-2 ในส่วนของปัจจัยความละเอียดในการพิมพ์ สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-11 จากการดำเนินงาน การพิมพ์ที่ความละเอียด 360 x 180 dpi ให้ค่าความถูกต้องสูงสุด ทั้งเอกสารชนิดคาร์บอน cbp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 75.42% และเอกสารชนิดเคมี cmp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 76.55% ซึ่งได้ผลใกล้เคียงกัน และเอกสารชนิดเคมีทั้ง cmp2p และ cmp3y ให้ค่าความถูกต้องสูงกว่าเอกสารสำเนาชนิดคาร์บอน cbp2w เท่ากับ 14.07% และ 10.86% ตามลำดับ



ภาพประกอบ 4-11 ผลการรู้จำเฉพาะอักษรไทย แยกตามความละเอียดในการพิมพ์

4.2.5 ผลการรู้จำเฉพาะอักษรไทย แยกตามความละเอียดและชนิดในการกราดภาพ

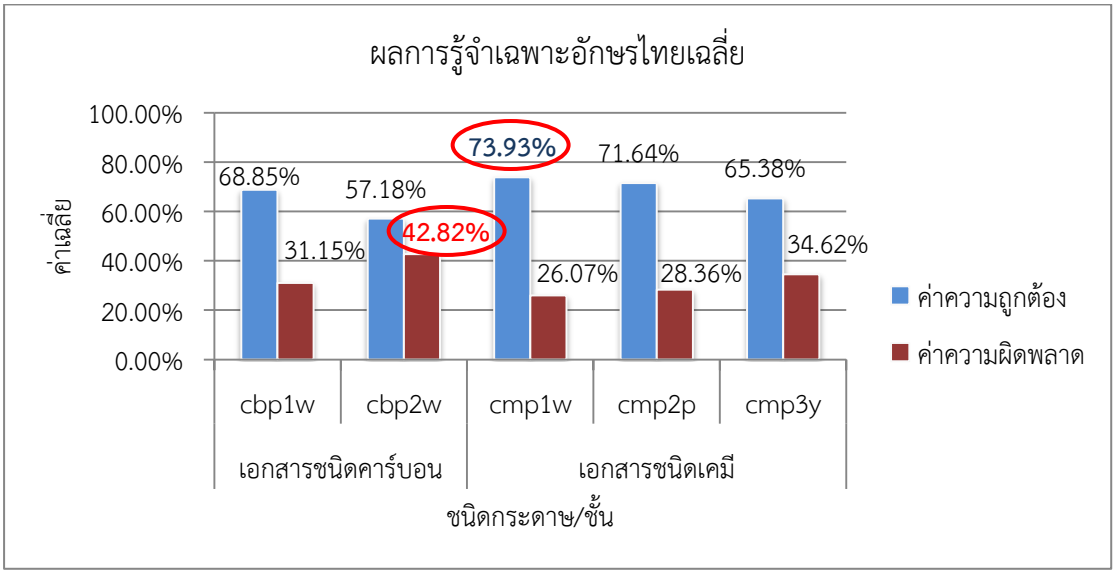
จากตาราง 4-2 ในส่วนของปัจจัยความละเอียดและชนิดในการกราดภาพ สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-12 จากการดำเนินงาน ความละเอียดและชนิดในการกราดภาพของเอกสารชนิดเคมี cmp1w ที่กราดภาพที่ความละเอียด 600 dpi ชนิด Gray ให้ค่าความถูกต้องสูงสุด เท่ากับ 77.80% และเอกสารชนิดคาร์บอน cbp1w ที่กราดภาพที่ความละเอียด 600 dpi ชนิด Color ให้ค่าความถูกต้องสูงสุดเท่ากับ 73.98% และเอกสารชนิดเคมีทั้ง cmp1w และ cmp3y ให้ค่าความถูกต้องสูงกว่าเอกสารชนิดคาร์บอน เท่ากับ 13.28% และ 7.09% ตามลำดับ



ภาพประกอบ 4-12 ผลการรู้จำเฉพาะอักษรไทย แยกตามความละเอียดและชนิดในการกราดภาพ

4.2.6 ผลการรู้จำเฉพาะอักษรไทยเฉลี่ย

จากตาราง 4-2 สำหรับเอกสารอักษรไทย สามารถสรุปเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-13 จากผลการทดลองได้ผลลัพธ์ค่าเฉลี่ยโดยรวม เอกสารชนิดเคมี cmp1w ให้ค่าความถูกต้องสูงสุดทุกปัจจัย เท่ากับ 85.33% และเอกสารชนิดคาร์บอน cbp1w ให้ค่าความถูกต้องเฉลี่ย เท่ากับ 81.57% และเอกสารสำเนาชนิดคาร์บอน cbp2w ให้ค่าความถูกต้องเฉลี่ยต่ำสุด เท่ากับ 75.42% ซึ่งยังต่ำกว่าเอกสารสำเนาชนิดเคมีทั้ง cmp2p และ cmp3y เท่ากับ 14.46% และ 8.20% ตามลำดับ

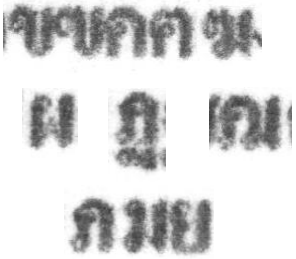


ภาพประกอบ 4-13 ผลการรู้จำเฉพาะอักษรไทยเฉลี่ย

4.2.7 ปัญหาการรู้จำเฉพาะอักษรไทย

การรู้จำเฉพาะอักษรไทยมีการแสดงผลที่ผิดพลาด ดังตาราง 4-3

ตาราง 4-3 ตัวอย่างการรู้จำที่ผิดพลาดเฉพาะอักษรไทย

อักษรไทย	ผลการรู้จำ	สาเหตุ
ข ฆ ค ศ ผ ฏ ภ ฌ ม	ร *บ *U *พ ศ ศ แ ฏ ก ณ ฆ	ส่วนหัวหาย มองเป็นเส้น เมื่อเป็นเส้นแล้วมีจุดเข้มมาทางซ้ายมองเป็นตัวอักษรที่คล้ายกัน 
ย	ช	หมึกเข้ม กระจายไป ส่วนหยักกับส่วนหลังตัวอักษรติดกัน
ขฆ ศฆ	พ ศพ	หมึกเข้ม กระจาย ตัวอักษรอยู่ติดกันแล้วมองเป็นตัวอื่น
ฆ ค ฆ ห ฬ ฮ	ซ ก พ ท พ 0 8 ช	ส่วนม้วนกลมขาดหายไปหรือวงกลมทึบกลายเป็นเส้น มองเป็นตัวคล้ายกัน

ชนิดแบบอักษร Angsana New จะรู้จำผิดพลาดมาก ได้แก่ ข ฆ ค ศ ฆ ณ ฏ ผ ภ ม ย ศ ษ ห ฬ ฮ เกิดจากความหนาของเส้นตัวอักษรทำให้ตัวอักษรติดกันได้ง่าย หรือวงกลมทึบกลายเป็นเส้น ทำให้ไปคล้ายกับตัวอักษรอื่น

ลักษณะตัวอักษรตัวหนา จะผิดพลาดมากที่สุดตามลำดับได้แก่ ศ ฆ ค ฮ ฌ ข ย เกิดจากความหนาของเส้นตัวอักษร สีเข้มทึบมองเป็นเส้น ตัวอักษรติดกัน หมึกพิมพ์กระจาย

4.3 ผลการรู้จำเฉพาะอักษรอังกฤษ

จากการทดลอง ประมวลผล เปรียบเทียบ วิเคราะห์ปัจจัยที่มีผลต่อการรู้จำเอกสาร ได้ผลการทดลอง ดังตาราง 4-4

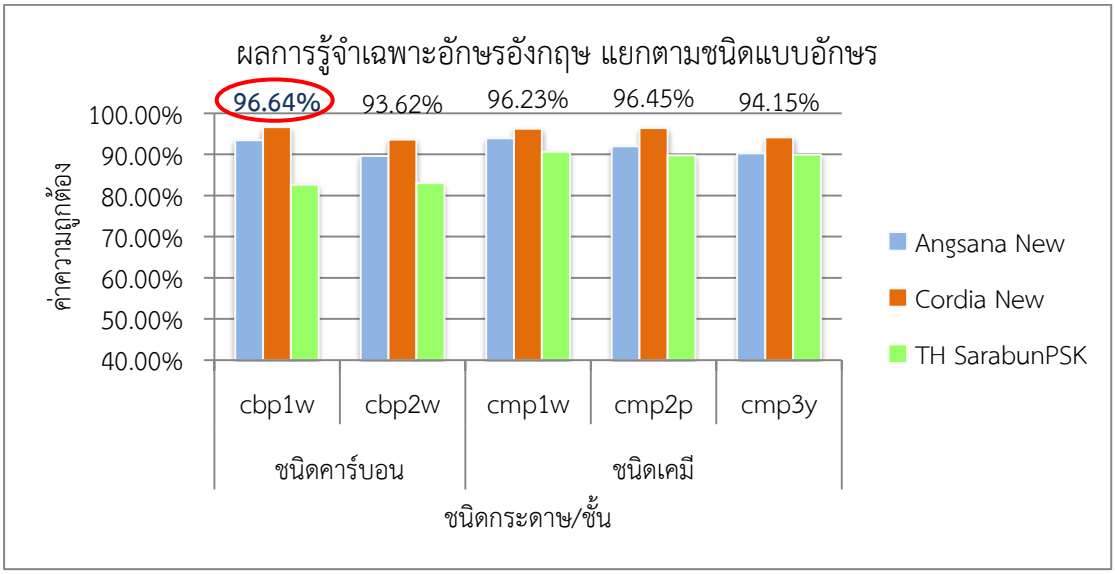
ตาราง 4-4 ผลการรู้จำเฉพาะอักษรอังกฤษ

ปัจจัย	ชนิดแบบฟอร์ม	ค่าความถูกต้อง (%)									
		เอกสารชนิดคาร์บอน				เอกสารชนิดเคมี					
		cbp1w		cbp2w		cmp1w		cmp2p		cmp3y	
		$\bar{x}$	S.D.	$\bar{x}$	S.D.	$\bar{x}$	S.D.	$\bar{x}$	S.D.	$\bar{x}$	S.D.
ชนิดแบบอักษร	Angsana New	93.48	0.48	89.63	0.75	93.91	0.39	91.99	2.73	90.27	1.39
	Cordia New	96.64	0.34	93.62	0.91	96.23	1.65	96.45	0.63	94.15	0.28
	TH SarabunPSK	82.66	15.04	83.04	1.67	90.63	3.27	89.77	2.90	89.92	1.82
ขนาดตัวอักษร	14	91.96	1.24	86.65	0.78	92.62	1.28	91.44	1.33	88.33	1.67
	16	93.67	0.94	89.03	0.78	94.42	1.17	93.42	1.00	92.93	0.95
	18	87.15	15.60	90.61	1.62	93.74	2.43	93.34	3.77	93.07	0.57
ลักษณะตัวอักษร	ตัวปกติ	91.62	5.05	87.71	2.36	93.15	2.86	94.09	0.36	91.08	1.11
	ตัวหนา	92.91	5.19	91.20	2.48	95.83	0.29	94.25	1.78	93.88	0.42
	ตัวเอียง	88.26	4.71	86.40	4.45	91.78	1.73	89.86	4.02	89.37	1.54
ความละเอียดในการพิมพ์	120 x 180 dpi	87.94	4.78	86.11	1.25	90.57	2.97	88.39	4.81	88.48	1.55
	180 x 180 dpi	90.57	4.85	87.24	2.41	93.59	0.81	93.19	1.27	90.46	1.29
	360 x 180 dpi	97.90	0.19	94.64	0.98	96.33	3.13	97.30	0.25	95.27	0.23
ความละเอียดและชนิดในการกราดภาพ	300 dpi Color	93.18	1.59	87.80	1.24	92.19	1.08	93.53	1.44	91.67	2.94
	300 dpi Gray	92.58	2.19	88.15	2.32	92.19	1.08	89.69	2.11	89.96	3.46
	600 dpi Color	94.37	1.20	89.36	2.58	94.93	0.91	93.97	1.69	92.25	2.17
	600 dpi Gray	83.59	17.18	89.75	2.03	95.04	1.35	93.75	1.58	91.90	2.28
ค่าเฉลี่ย		**91.16**	**5.04**	**88.81**	**1.79**	**93.57**	**1.65**	**92.78**	**1.98**	**91.44**	**1.48**

หมายเหตุ ตัวเลขสีน้ำเงิน หมายถึง ค่ามากที่สุด ตัวเลขสีแดง หมายถึง ค่าน้อยที่สุด

4.3.1 ผลการรู้จำเฉพาะอักษรอังกฤษ แยกตามชนิดแบบอักษร

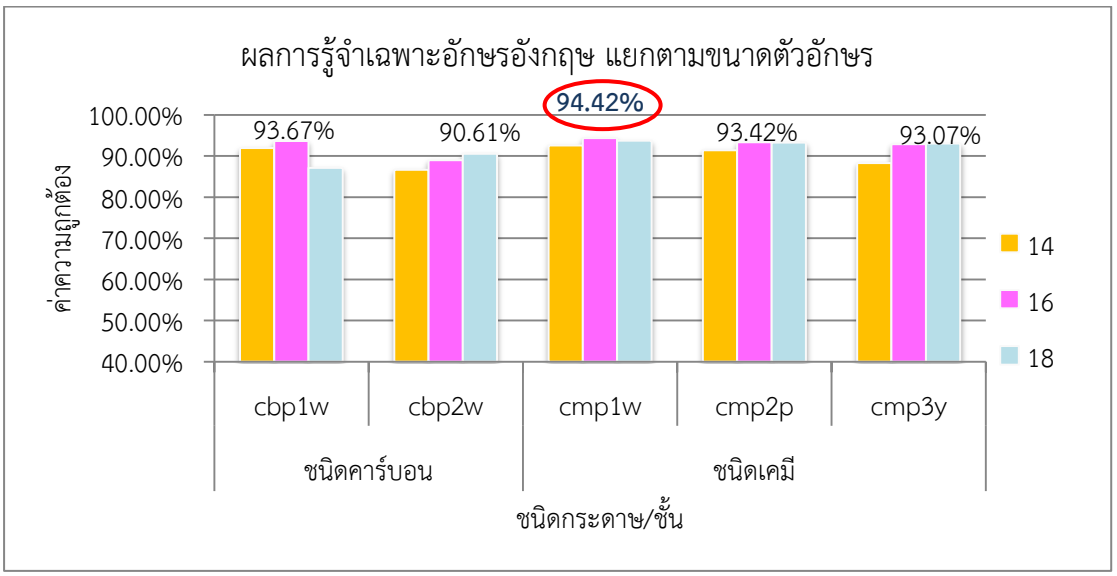
จากตาราง 4-4 ในส่วนของปัจจัยชนิดแบบอักษร สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-14 จากการดำเนินงาน ชนิดแบบอักษร Cordia New ให้ค่าความถูกต้องสูงสุด ทั้งเอกสารชนิดคาร์บอน cbp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 96.64% และเอกสารชนิดเคมี cmp2p ให้ค่าความถูกต้องสูงสุดเท่ากับ 96.45% ซึ่งได้ผลใกล้เคียงกัน และเอกสารสำเนาชนิดเคมีทั้ง cmp2p และ cmp3y ให้ค่าความถูกต้องสูงกว่าเอกสารสำเนาชนิดคาร์บอน เท่ากับ 2.83 และ 0.53% ตามลำดับ ไม่แตกต่างกันมากนัก



ภาพประกอบ 4-14 ผลการรู้จำเฉพาะอักษรอังกฤษ แยกตามชนิดแบบอักษร

4.3.2 ผลการรู้จำเฉพาะอักษรอังกฤษ แยกตามขนาดตัวอักษร

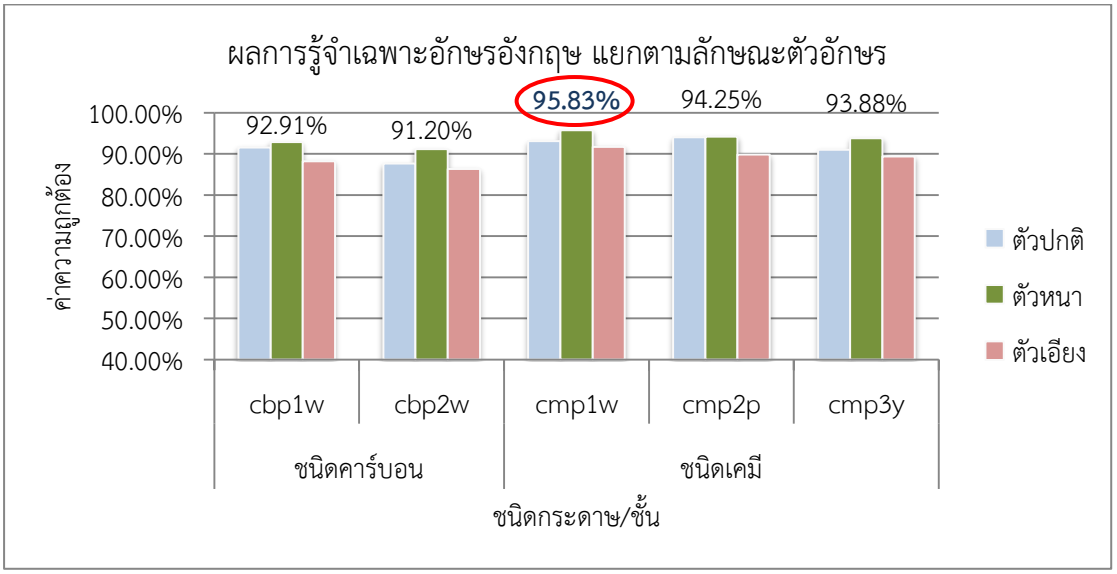
จากตาราง 4-4 ในส่วนของปัจจัยขนาดตัวอักษร สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-15 จากการดำเนินงาน ขนาดตัวอักษร 14 ให้ค่าความถูกต้องสูงสุด ทั้งเอกสารชนิดคาร์บอน cbp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 93.67% และเอกสารชนิดเคมี cmp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 94.42% ซึ่งได้ผลใกล้เคียงกัน ส่วนเอกสารสำเนาชนิดเคมีทั้ง cmp2p และ cmp3y ให้ค่าความถูกต้องของตัวอักษรขนาด 16 และ 18 ใกล้เคียงกันซึ่งมากกว่าเอกสารสำเนาชนิดคาร์บอน cmp2w เท่ากับ 2.81% และ 2.46% ตามลำดับ



ภาพประกอบ 4-15 ผลการรู้จำเฉพาะอักษรอังกฤษ แยกตามขนาดตัวอักษร

4.3.3 ผลการรู้จำเฉพาะอักษรอังกฤษ แยกตามลักษณะตัวอักษร

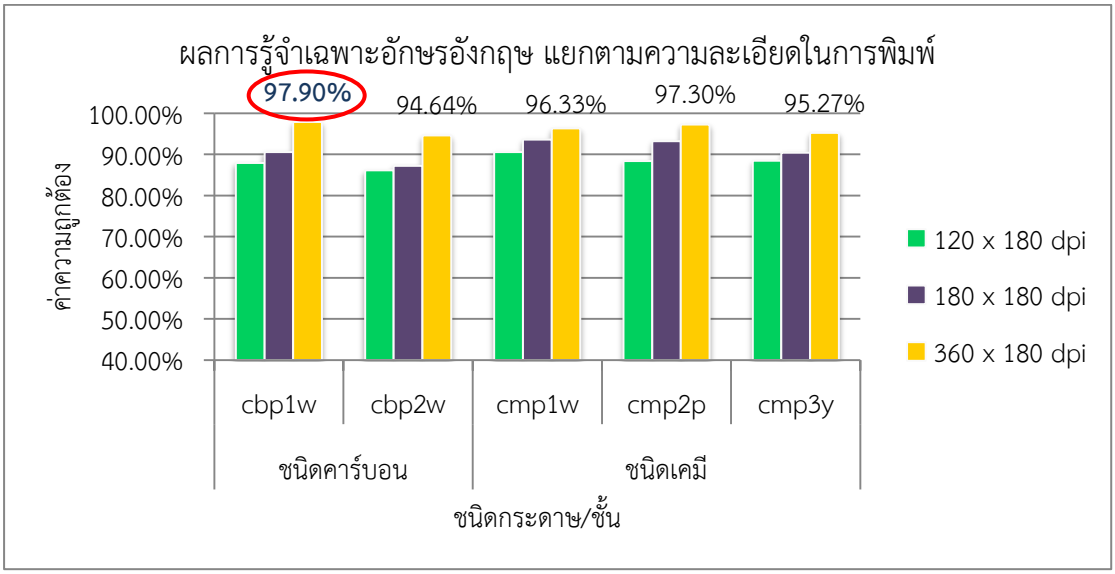
จากตาราง 4-4 ในส่วนของปัจจัยลักษณะตัวอักษร สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-16 จากการดำเนินงาน ลักษณะตัวอักษร ตัวหนา ให้ค่าความถูกต้องสูงสุด ทั้งเอกสารชนิดคาร์บอน cbp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 92.91% และเอกสารชนิดเคมี cmp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 95.83% ซึ่งได้ผลใกล้เคียงกัน และเอกสารสำเนาชนิดเคมีทั้ง cmp2p และ cmp3y ให้ค่าความถูกต้องสูงกว่าเอกสารสำเนาชนิดคาร์บอน เท่ากับ 3.05% และ 2.68% ตามลำดับ ไม่แตกต่างกันมากนัก



ภาพประกอบ 4-16 ผลการรู้จำเฉพาะอักษรอังกฤษ แยกตามลักษณะตัวอักษร

4.3.4 ผลการรู้จำเฉพาะอักษรอังกฤษ แยกตามความละเอียดในการพิมพ์

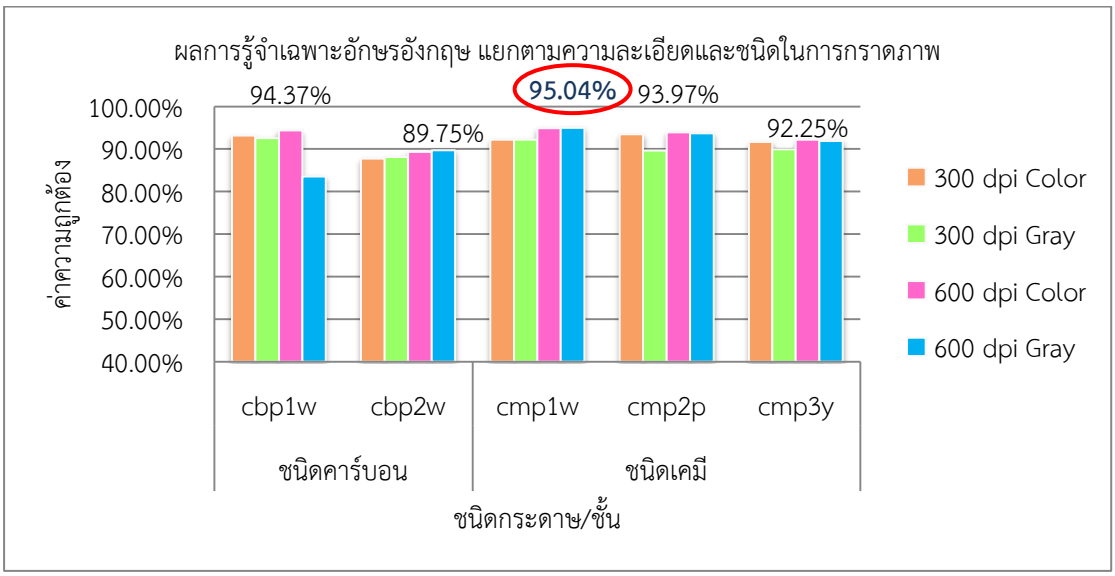
จากตาราง 4-4 ในส่วนของปัจจัยความละเอียดในการพิมพ์ สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-17 จากการดำเนินงาน การพิมพ์ที่ความละเอียด 360 x 180 dpi ให้ค่าความถูกต้องสูงสุด ทั้งเอกสารชนิดคาร์บอน cbp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 97.90% และเอกสารชนิดเคมี cmp2p ให้ค่าความถูกต้องสูงสุดเท่ากับ 97.30% ซึ่งได้ผลใกล้เคียงกัน และเอกสารชนิดเคมีทั้ง cmp1w และ cmp3y ให้ค่าความถูกต้องสูงกว่าเอกสารสำเนาชนิดคาร์บอน เท่ากับ 2.65% และ 0.63% ตามลำดับ ให้ค่าใกล้เคียงกัน



ภาพประกอบ 4-17 ผลการรู้จำเฉพาะอักษรอังกฤษ แยกตามความละเอียดในการพิมพ์

4.3.5 ผลการรู้จำเฉพาะอักษรอังกฤษ แยกตามความละเอียดและชนิดในการกราดภาพ

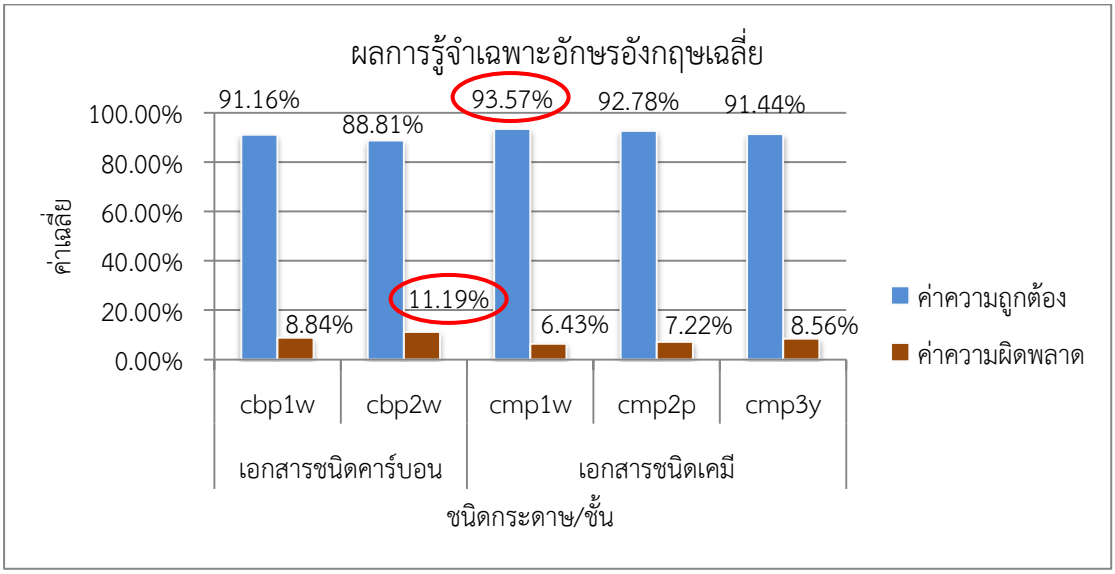
จากตาราง 4-4 ในส่วนของปัจจัยความละเอียดและชนิดในการกราดภาพ สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-18 จากการดำเนินงาน ความละเอียดและชนิดในการกราดภาพของเอกสารชนิดเคมี cmp1w ที่กราดภาพที่ความละเอียด 600 dpi ชนิด Gray ให้ค่าความถูกต้องสูงสุด เท่ากับ 95.04% และเอกสารชนิดคาร์บอน cbp1w ที่กราดภาพที่ความละเอียด 600 dpi ชนิด Color ให้ค่าความถูกต้องสูงสุดเท่ากับ 94.37% ซึ่งได้ผลใกล้เคียงกัน และเอกสารชนิดเคมีทั้ง cmp1w และ cmp3y ให้ค่าความถูกต้องสูงกว่าเอกสารสำเนาชนิดคาร์บอน เท่ากับ 4.22% และ 2.50% ตามลำดับ



ภาพประกอบ 4-18 ผลการรู้จำเฉพาะอักษรอังกฤษ แยกตามความละเอียดและชนิดในการกราดภาพ

4.3.6 ผลการรู้จำเฉพาะอักษรอังกฤษเฉลี่ย

จากตาราง 4-4 สำหรับเอกสารอักษรอังกฤษ สามารถสรุปเป็นกราฟเปรียบเทียบได้ดัง ภาพประกอบ 4-19 จากผลการทดลองได้ผลลัพธ์ค่าเฉลี่ย เอกสารชนิดเคมี cmp1w ให้ค่าความถูกต้องเฉลี่ยสูงสุด เท่ากับ 93.57% และเอกสารชนิดคาร์บอน cbp1w ให้ค่าความถูกต้องเฉลี่ยเพียง 91.16% และเอกสารสำเนาชนิดคาร์บอน cbp2w ให้ค่าความถูกต้องเฉลี่ยต่ำสุด เท่ากับ 88.81% ซึ่งยังต่ำกว่าเอกสารสำเนาชนิดเคมีทั้ง cmp2p และ cmp3y เท่ากับ 3.97% และ 2.63% ตามลำดับ ไม่แตกต่างกันมากนัก

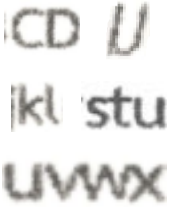


ภาพประกอบ 4-19 ผลการรู้จำเฉพาะอักษรอังกฤษเฉลี่ย

4.3.7 ปัญหาการรู้จำเฉพาะอักษรอังกฤษ

การรู้จำเฉพาะอักษรอังกฤษมีการแสดงผลที่ผิดพลาด ดังตาราง 4-5

ตาราง 4-5 ตัวอย่างการรู้จำที่ผิดพลาดเฉพาะอักษรอังกฤษ

อักษรอังกฤษ	ผลการรู้จำ	สาเหตุ
a f g	O Q / H § $ q e	ส่วนหัวหาย มองเป็นเส้น เมื่อเป็นเส้นแล้วมีจุดเข้มมาทางซ้าย มองเป็นตัวอักษรที่คล้ายกัน
h L	fr I	เส้นขาดหาย มองเป็นตัวอื่น
CD I J kl s uvw vw x	W U Id 5 wM wv m < >(m k	หมึกเข้ม กระจายไป ตัวอักษรติดกัน มองเป็นตัวอื่นที่คล้ายกัน 

อักษรอังกฤษรู้จำผิดพลาดมากคือลักษณะตัวอักษรตัวเอียง ดังตาราง 4-5 เกิดจากความหนาของเส้นตัวอักษร สีเข้มทึบมองเป็นเส้น ตัวอักษรติดกัน หมึกพิมพ์กระจาย

ชนิดแบบอักษร TH SarabunPSK จะรู้จำผิดพลาดมากตามลำดับ ได้แก่ I J g a l w x D C f v เกิดจากความบางของเส้นตัวอักษร ทำให้เส้นตัวอักษรขาด โดยเฉพาะลักษณะตัวเอียง

ส่วนลักษณะตัวหนาเมื่อพิมพ์ด้วยกระดาษชนิดคาร์บอน สำเนาที่ 2 เส้นตัวอักษรก็จะหนาหมึกเข้มทำให้ตัวอักษรติดกันได้ง่าย หรือวงกลมทึบกลายเป็นเส้น ทำให้ไปคล้ายตัวอักษรอื่น

4.4 ผลการรู้จำเฉพาะเลขไทย

จากการทดลอง ประมวลผล เปรียบเทียบ วิเคราะห์ปัจจัยที่มีผลต่อการรู้จำเอกสารดอทเมทริกซ์ ชนิดคาร์บอน และชนิดเคมี ของเอกสารเลขไทย ได้ผลการทดลอง ดังตาราง 4-6

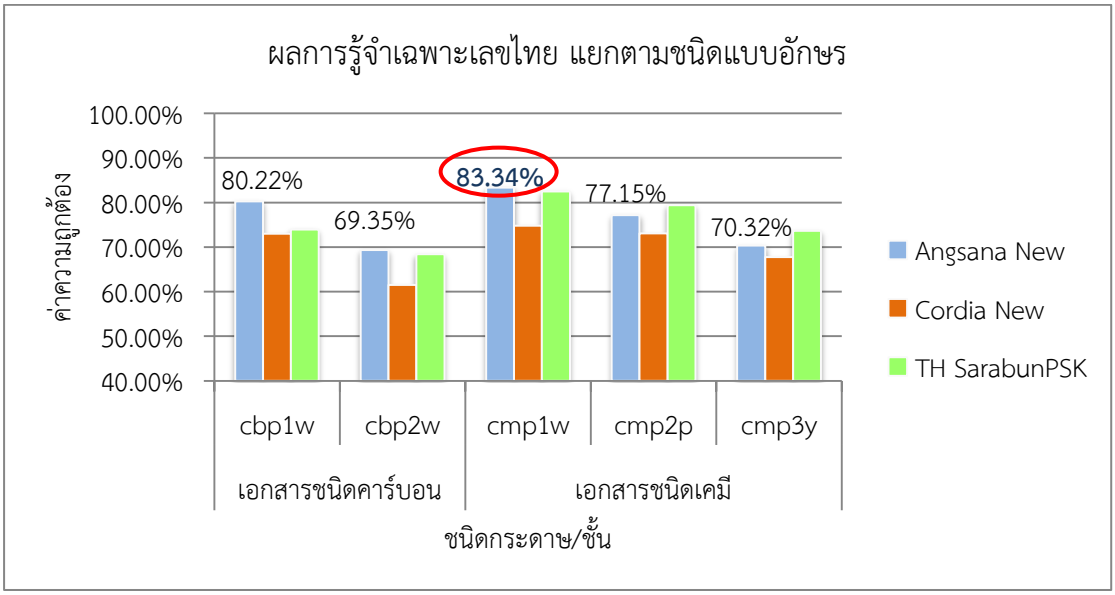
ตาราง 4-6 ผลการรู้จำเฉพาะเลขไทย

ปัจจัย	ชนิดแบบฟอร์ม	ค่าความถูกต้อง (%)									
		เอกสารชนิดคาร์บอน				เอกสารชนิดเคมี					
		cbp1w		cbp2w		cmp1w		cmp2p		cmp3y	
		$\bar{x}$	S.D.	$\bar{x}$	S.D.	$\bar{x}$	S.D.	$\bar{x}$	S.D.	$\bar{x}$	S.D.
ชนิดแบบอักษร	Angsana New	80.22	3.40	69.35	3.85	83.34	3.03	77.15	2.70	70.32	1.45
	Cordia New	72.98	1.26	61.49	0.78	74.76	1.13	73.06	1.00	67.73	3.07
	TH SarabunPSK	73.95	12.95	68.42	3.48	82.46	2.40	79.40	1.95	73.66	1.71
ขนาดตัวอักษร	14	70.41	2.79	55.42	3.05	76.25	1.70	71.05	1.35	61.89	3.89
	16	81.46	0.94	71.71	2.85	81.49	0.33	79.39	0.47	74.52	0.94
	18	75.28	12.95	72.13	2.39	82.82	2.87	79.16	4.68	62.06	15.74
ลักษณะตัวอักษร	ตัวปกติ	71.42	5.24	62.70	2.55	75.59	2.53	74.14	0.93	69.71	2.83
	ตัวหนา	76.49	3.22	65.18	3.55	80.57	1.05	75.49	2.81	68.71	0.27
	ตัวเอียง	79.23	4.06	74.83	6.15	85.19	84.40	80.19	79.97	73.29	2.64
ความละเอียดในการพิมพ์	120 x 180 dpi	71.28	3.29	61.81	0.83	76.03	3.62	72.32	5.22	63.29	3.06
	180 x 180 dpi	74.73	3.04	63.80	5.95	80.97	1.07	77.45	1.27	71.57	1.70
	360 x 180 dpi	75.53	22.42	64.66	3.98	77.84	0.89	73.23	3.00	71.71	2.29
ความละเอียดและชนิดในการกราดภาพ	300 dpi Color	76.07	7.32	63.45	10.89	79.27	2.66	76.69	5.32	60.76	12.86
	300 dpi Gray	77.12	7.54	65.64	7.45	78.66	4.24	73.97	5.45	69.04	9.14
	600 dpi Color	79.54	6.07	68.80	10.06	81.39	3.76	77.74	4.38	62.44	13.23
	600 dpi Gray	70.13	12.96	67.79	9.84	81.43	4.09	77.73	5.89	72.39	6.81
ค่าเฉลี่ย		**75.36**	**6.84**	**66.07**	**4.85**	**79.88**	**7.49**	**76.14**	**7.90**	**68.32**	**5.10**

หมายเหตุ ตัวเลขสีน้ำเงิน หมายถึง ค่ามากที่สุด ตัวเลขสีแดง หมายถึง ค่าน้อยที่สุด

4.4.1 ผลการรู้จำเฉพาะเลขไทย แยกตามชนิดแบบอักษร

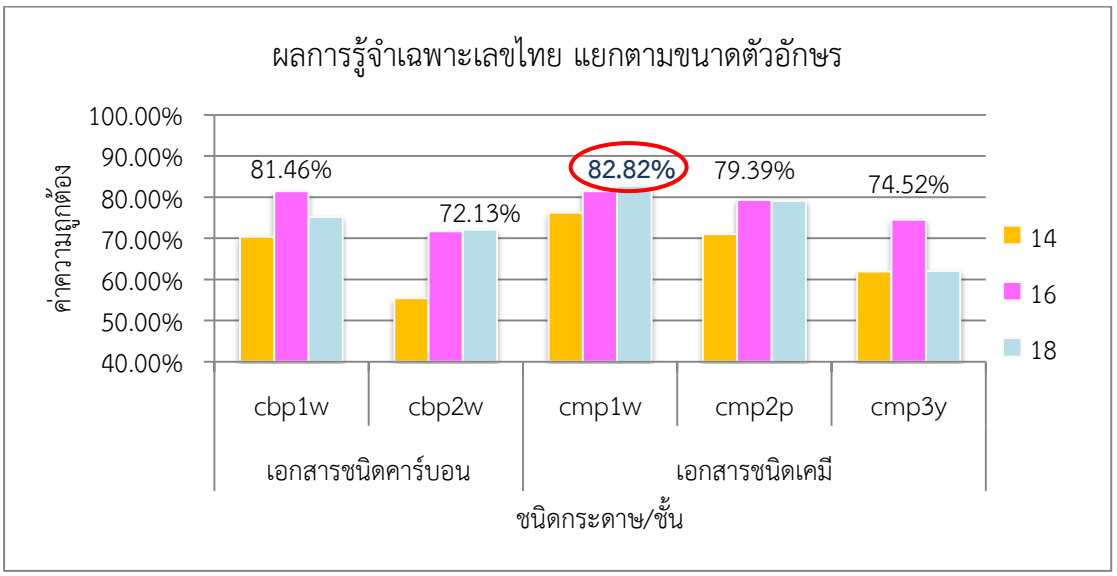
จากตาราง 4-6 ในส่วนของปัจจัยชนิดแบบอักษร สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-20 ชนิดแบบอักษร Angsana New ให้ค่าความถูกต้องสูงสุด ทั้งเอกสารชนิดคาร์บอน cbp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 80.22% และเอกสารชนิดเคมี cmp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 83.34% และเอกสารสำเนาชนิดเคมีทั้ง cmp2p และ cmp3y ให้ค่าความถูกต้องสูงกว่าเอกสารสำเนาชนิดคาร์บอน เท่ากับ 10.05% และ 4.31% ตามลำดับ



ภาพประกอบ 4-20 ผลการรู้จำเฉพาะเลขไทย แยกตามชนิดแบบอักษร

4.4.2 ผลการรู้จำเฉพาะเลขไทย แยกตามขนาดตัวอักษร

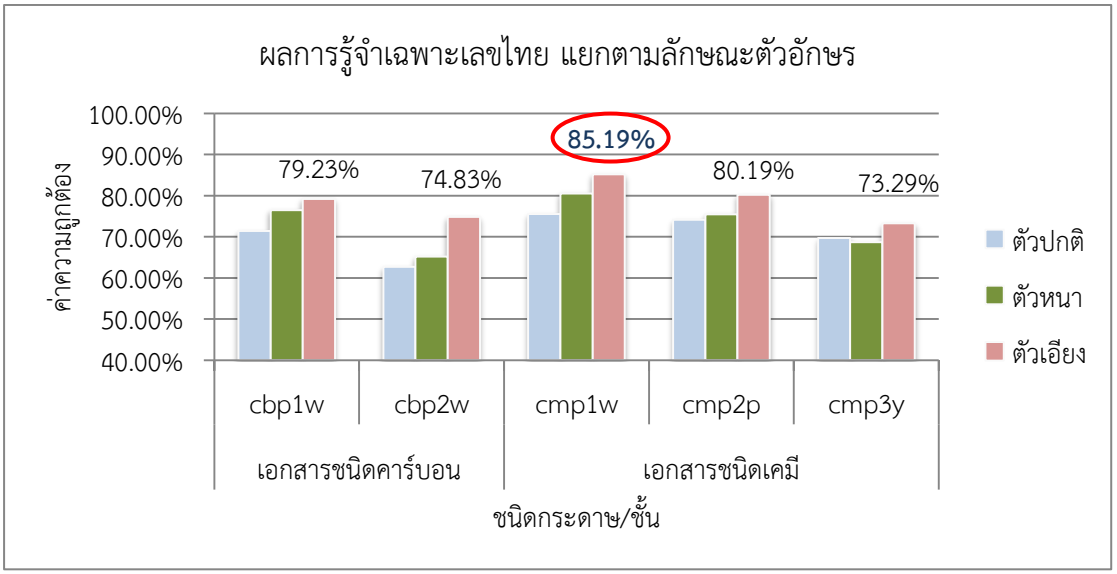
จากตาราง 4-6 ในส่วนของปัจจัยขนาดตัวอักษร สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-21 เอกสารชนิดเคมี cmp1w ขนาดตัวอักษร 18 ให้ค่าความถูกต้องสูงสุด เท่ากับ 82.82% และเอกสารชนิดคาร์บอน cmp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 81.46% ซึ่งใกล้เคียงกัน และเอกสารสำเนาชนิดเคมีทั้ง cmp2p และ cmp3y ให้ค่าความถูกต้องสูงกว่าเอกสารสำเนาชนิดคาร์บอน cbp2w เท่ากับ 7.26% และ 2.39% ตามลำดับ



ภาพประกอบ 4-21 ผลการรู้จำเฉพาะเลขไทย แยกตามขนาดตัวอักษร

4.4.3 ผลการรู้จำเฉพาะเลขไทย แยกตามลักษณะตัวอักษร

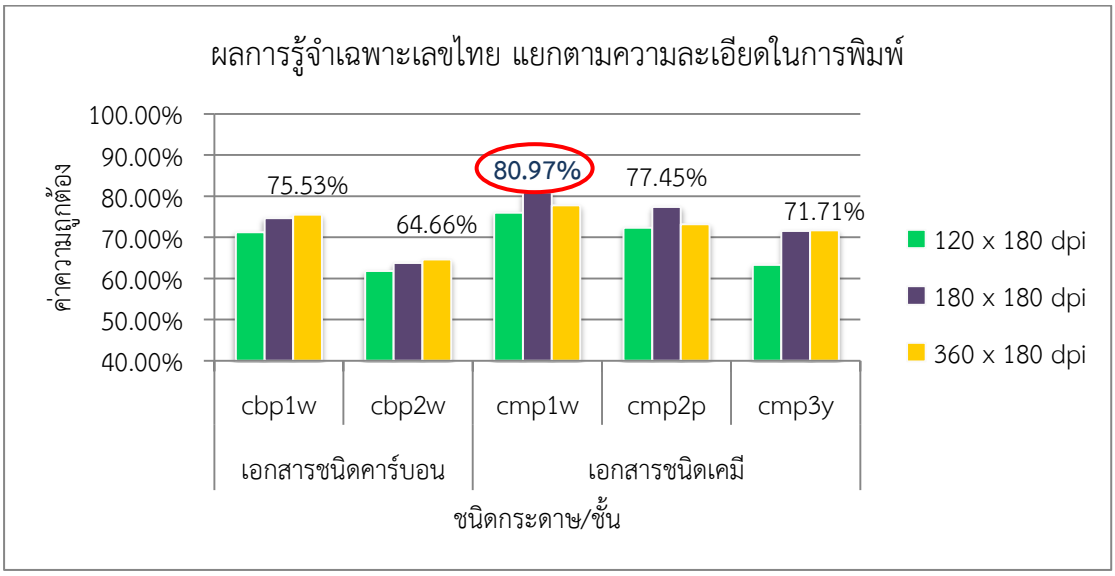
ตาราง 4-6 ในส่วนของปัจจัยลักษณะตัวอักษร สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-22 ลักษณะตัวอักษร ตัวเอียง ให้ค่าความถูกต้องสูงสุดทุกชนิดเอกสาร ทั้งเอกสารชนิดเคมี cmp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 85.19% และเอกสารชนิดคาร์บอน cbp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 79.23% แตกต่างกันเท่ากับ 5.96%



ภาพประกอบ 4-22 ผลการรู้จำเฉพาะเลขไทย แยกตามลักษณะตัวอักษร

4.4.4 ผลการรู้จำเฉพาะเลขไทย แยกตามความละเอียดในการพิมพ์

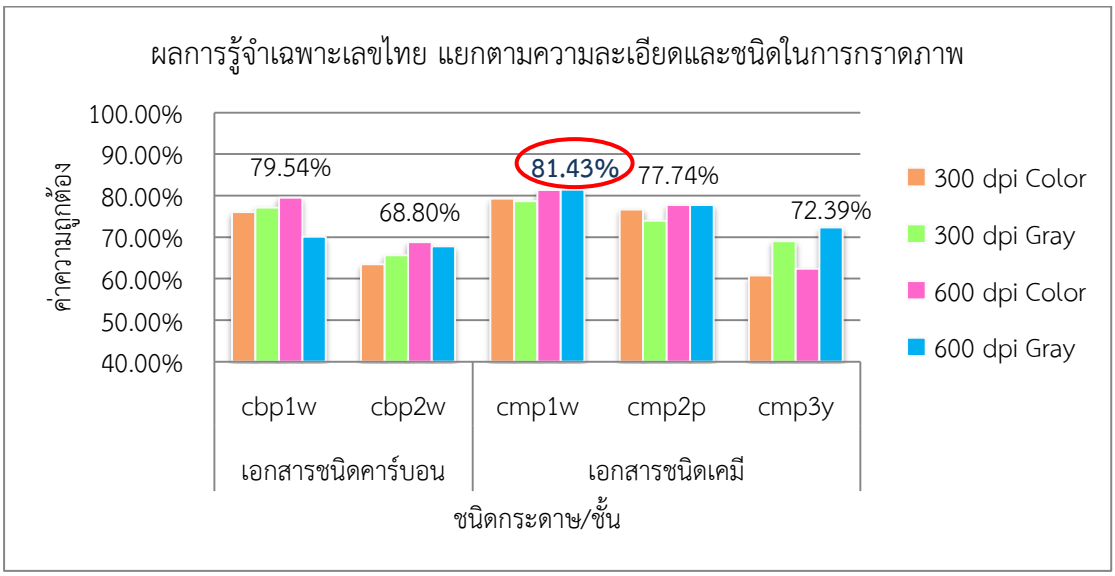
ตาราง 4-6 ในส่วนของปัจจัยความละเอียดในการพิมพ์ สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-23 เอกสารชนิดเคมี cmp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 80.97% ที่การพิมพ์ที่ความละเอียด 180 x 180 dpi เอกสารชนิดคาร์บอน cbp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 75.53% และเอกสารชนิดเคมีทั้ง cmp2p และ cmp3y ให้ค่าความถูกต้องสูงกว่าเอกสารสำเนาชนิดคาร์บอน cbp2w เท่ากับ 12.79% และ 7.04% ตามลำดับ



ภาพประกอบ 4-23 ผลการรู้จำเฉพาะเลขไทย แยกตามความละเอียดในการพิมพ์

4.4.5 ผลการรู้จำเฉพาะเลขไทย แยกตามความละเอียดและชนิดในการกราดภาพ

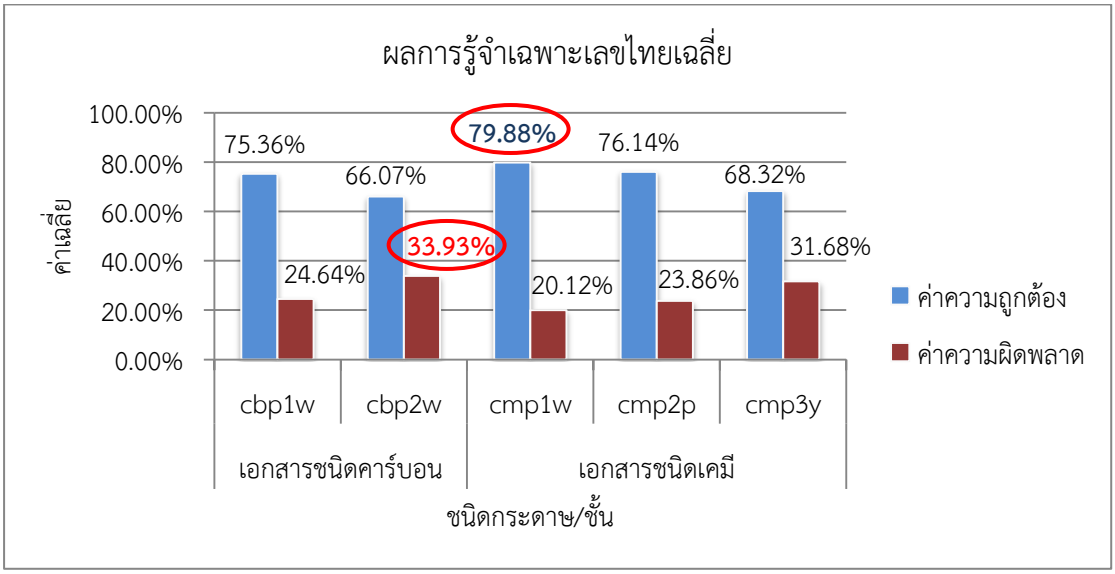
ตาราง 4-6 ในส่วนของปัจจัยความละเอียดและชนิดในการกราดภาพ สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-24 ความละเอียดและชนิดในการกราดภาพของเอกสารชนิดเคมี cmp1w ที่กราดภาพที่ความละเอียด 600 dpi ชนิด Gray ให้ค่าความถูกต้องสูงสุด เท่ากับ 81.43% และเอกสารชนิดคาร์บอน cbp1w ที่กราดภาพที่ความละเอียด 600 dpi ชนิด Color ให้ค่าความถูกต้องสูงสุดเท่ากับ 79.54% ซึ่งใกล้เคียงกันทั้งเอกสารชนิดเคมี และเอกสารชนิดคาร์บอน และเอกสารสำเนาชนิดเคมีทั้ง cmp2p และ cmp3y ให้ค่าความถูกต้องสูงกว่าเอกสารชนิดคาร์บอน cbp2w เท่ากับ 8.94% และ 3.59% ตามลำดับ



ภาพประกอบ 4-24 ผลการรู้จำเฉพาะเลขไทย แยกตามความละเอียดและชนิดในการกราดภาพ

4.4.6 ผลการรู้จำเฉพาะเลขไทยเฉลี่ย

จากตาราง 4-6 สำหรับเอกสารเลขไทย สามารถสรุปเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-25 จากผลการทดลองได้ผลลัพธ์ค่าเฉลี่ย เอกสารชนิดเคมี cmp1w ให้ค่าความถูกต้องเฉลี่ยสูงสุด เท่ากับ 79.88% และเอกสารชนิดคาร์บอน cbp1w ให้ค่าความถูกต้องเฉลี่ยเท่ากับ 75.36% และเอกสารสำเนาชนิดคาร์บอน cbp1w ให้ค่าความถูกต้องเฉลี่ยต่ำสุด เท่ากับ 66.07% ซึ่งยังต่ำกว่าเอกสารสำเนาชนิดเคมีทั้ง cmp2p และ cmp3y เท่ากับ 10.06% และ 2.25% ตามลำดับ

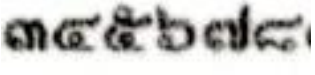


ภาพประกอบ 4-25 ผลการรู้จำเฉพาะเลขไทยเฉลี่ย

4.4.7 ปัญหาการรู้จำเฉพาะเลขไทย

การรู้จำเฉพาะเลขไทยมีการแสดงผลที่ผิดพลาด ดังตาราง 4-7

ตาราง 4-7 ตัวอย่างการรู้จำที่ผิดพลาดเฉพาะเลขไทย

เลขไทย	ผลการรู้จำ	สาเหตุ
๑	•	หมึกเข้มทึบเป็นรูปคล้ายวงกลม แล้วมองเป็นจุด
๓ ๔ ๕ ๖ ๗ ๘	»)CD *or* *fr* b *» rf « OJ	หมึกเข้ม กระจาย ตัวอักษรติดกัน แล้วมองเป็นตัวอักษร หรือสัญลักษณ์อื่นที่คล้ายกัน 

เลขไทยรู้จำผิดพลาดมากคือตัวที่พิมพ์ที่ความละเอียด 600 dpi ดังตาราง 4-7 ทำให้เพิ่มความหนาของตัวอักษรที่พิมพ์ด้วยกระดาษชนิดคาร์บอนสำเนาที่ 2 (cbp2w) สีเข้มทึบมองเป็นเส้น ตัวอักษรติดกัน หมึกพิมพ์กระจาย

ชนิดแบบอักษร Cordia New จะรู้จำผิดพลาดมาก ได้แก่ ๑ ๕ ๘ เกิดจากหมึกเข้มทึบ ความหนาของเส้นตัวอักษรทำให้ตัวอักษรติดกันได้ง่าย หรือวงกลมทึบกลายเป็นเส้น ทำให้ไปคล้ายกับตัวอักษรอื่น

4.5 ผลการรู้จำเฉพาะเลขอารบิค

จากการทดลอง ประมวลผล เปรียบเทียบ วิเคราะห์ปัจจัยที่มีผลต่อการรู้จำเอกสารดอทเมทริกซ์ ชนิดคาร์บอน และชนิดเคมี ของเอกสารเลขอารบิค ได้ผลการทดลอง ดังตาราง 4-8

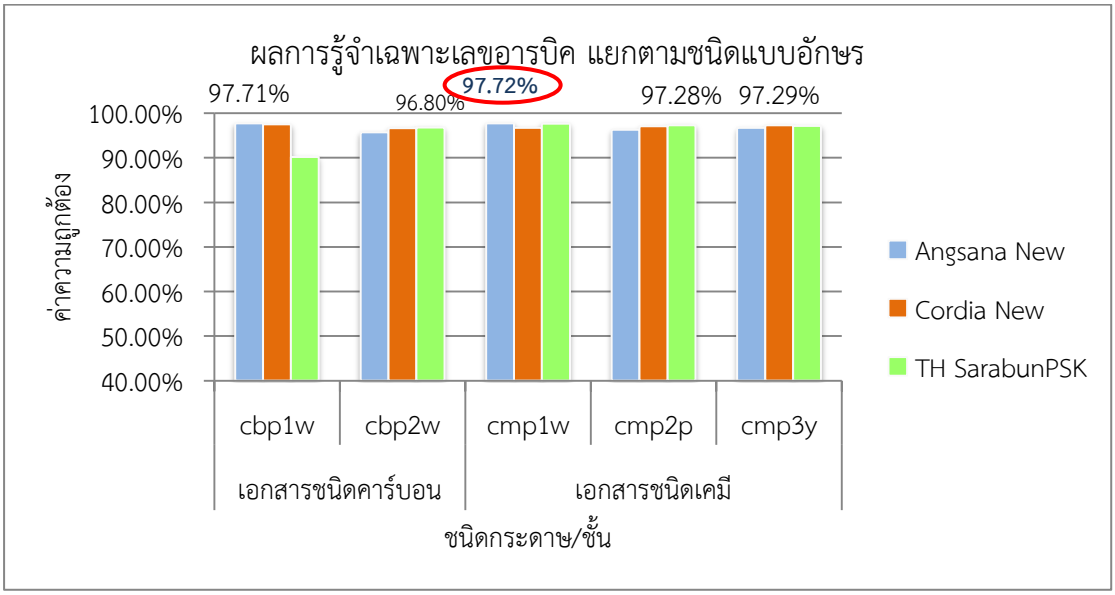
ตาราง 4-8 ผลการรู้จำเฉพาะเลขอารบิค

ปัจจัย	ชนิดแบบฟอร์ม	ค่าความถูกต้อง (%)									
		เอกสารชนิดคาร์บอน				เอกสารชนิดเคมี					
		cbp1w		cbp2w		cmp1w		cmp2p		cmp3y	
		$\bar{x}$	S.D.	$\bar{x}$	S.D.	$\bar{x}$	S.D.	$\bar{x}$	S.D.	$\bar{x}$	S.D.
ชนิดแบบอักษร	Angsana New	97.71	0.64	95.71	1.20	**97.72**	0.31	96.26	2.76	96.67	0.85
	Cordia New	97.48	0.22	96.65	0.98	96.73	1.42	97.04	0.17	97.29	0.29
	TH SarabunPSK	**90.15**	16.29	96.80	2.49	97.65	1.77	97.28	2.12	97.10	0.58
ขนาดตัวอักษร	14	97.70	0.54	95.52	1.98	97.76	0.46	97.03	1.61	96.41	0.77
	16	98.08	0.37	96.98	0.65	**98.37**	0.11	97.98	0.14	97.55	0.24
	18	**89.56**	16.48	96.66	2.07	96.01	1.60	95.57	3.31	97.11	0.20
ลักษณะตัวอักษร	ตัวปกติ	95.15	5.20	94.23	4.76	96.36	1.64	97.60	0.43	97.36	0.12
	ตัวหนา	95.95	5.39	96.28	2.08	**98.58**	0.44	97.59	1.95	97.24	0.20
	ตัวเอียง	94.24	5.35	**93.47**	4.93	97.19	0.17	95.40	2.61	96.47	0.44
ความละเอียดในการพิมพ์	120 x 180 dpi	**94.93**	4.96	96.56	0.98	96.66	1.73	95.28	4.67	96.96	0.32
	180 x 180 dpi	95.48	5.41	95.80	3.92	**98.29**	0.52	97.90	0.61	97.29	0.75
	360 x 180 dpi	96.85	0.21	95.75	0.80	95.89	3.19	96.67	0.35	96.26	0.61
ความละเอียดและชนิดในการกราดภาพ	300 dpi Color	97.58	0.57	94.37	1.88	96.98	1.74	97.66	0.33	96.83	1.11
	300 dpi Gray	97.72	0.09	96.27	1.34	96.67	2.17	94.39	3.62	96.85	0.67
	600 dpi Color	98.00	0.29	97.64	0.33	97.80	0.54	97.65	0.67	97.27	0.40
	600 dpi Gray	**87.15**	19.32	97.28	0.42	**98.06**	0.48	97.74	0.64	97.13	0.56
ค่าเฉลี่ย		**95.23**	**5.08**	**96.00**	**1.92**	**97.29**	**1.14**	**96.82**	**1.62**	**96.99**	**0.51**

หมายเหตุ ตัวเลขสีน้ำเงิน หมายถึง ค่ามากที่สุด ตัวเลขสีแดง หมายถึง ค่าน้อยที่สุด

4.5.1 ผลการรู้จำเฉพาะเลขอารบิค แยกตามชนิดแบบอักษร

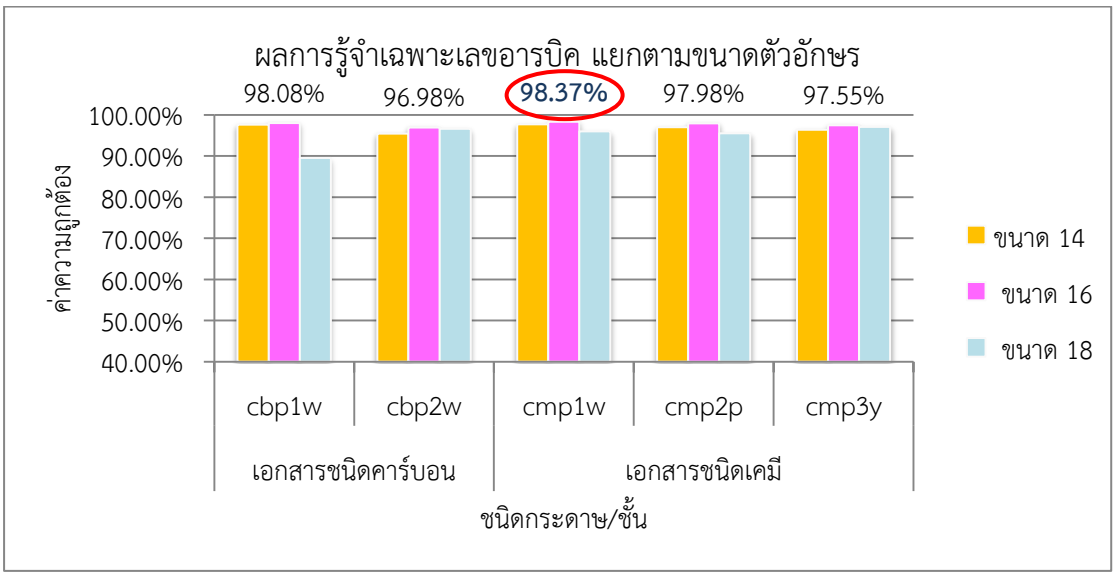
จากตาราง 4-8 ในส่วนของปัจจัยชนิดแบบอักษร สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-26 ชนิดแบบอักษร Angsana New ให้ค่าความถูกต้องสูงสุด ทั้งเอกสารชนิดคาร์บอน cbp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 97.71% และเอกสารชนิดเคมี cmp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 97.72% และเอกสารสำเนาชนิดเคมีทั้ง cmp2p และ cmp3y ให้ค่าความถูกต้องสูงกว่าเอกสารสำเนาชนิดคาร์บอน เท่ากับ 0.48% และ 0.30% ตามลำดับ



ภาพประกอบ 4-26 ผลการรู้จำเฉพาะเลขอารบิค แยกตามชนิดแบบอักษร

4.5.2 ผลการรู้จำเฉพาะเลขอารบิค แยกตามขนาดตัวอักษร

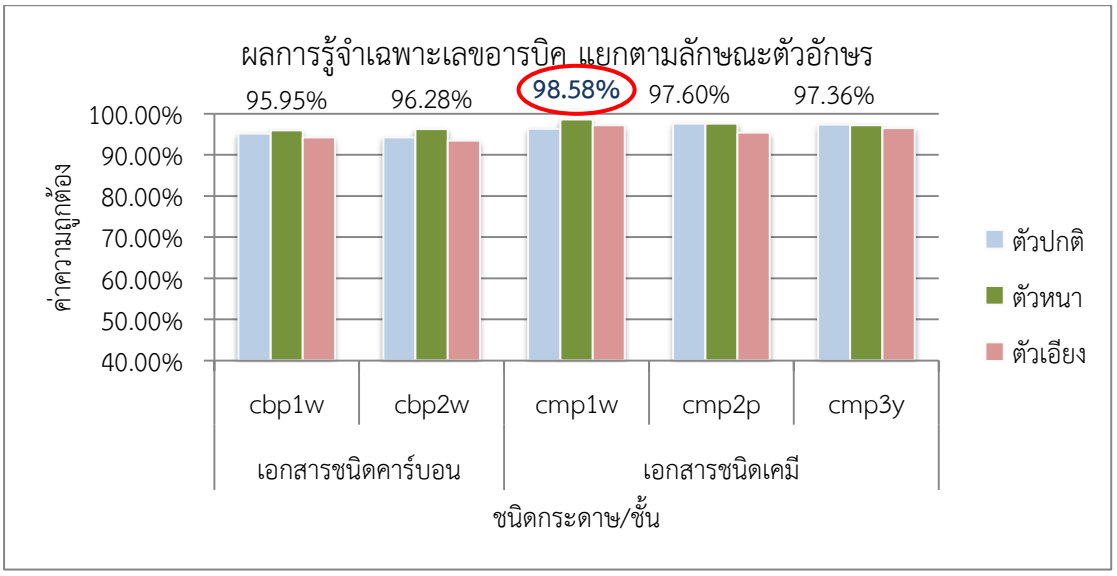
ตาราง 4-8 ในส่วนของปัจจัยขนาดตัวอักษร สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-27 เอกสารชนิดเคมี cmp1w ขนาดตัวอักษร 16 ให้ค่าความถูกต้องสูงสุด เท่ากับ 98.37% และเอกสารชนิดคาร์บอน cmp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 98.08% ซึ่งใกล้เคียงกัน และเอกสารสำเนาชนิดเคมีทั้ง cmp2p และ cmp3y ให้ค่าความถูกต้องสูงกว่าเอกสารสำเนาชนิดคาร์บอน cbp2w เท่ากับ 1.00% และ 0.56% ตามลำดับ



ภาพประกอบ 4-27 ผลการรู้จำเฉพาะเลขอารบิค แยกตามขนาดตัวอักษร

4.5.3 ผลการรู้จำเฉพาะเลขอารบิค แยกตามลักษณะตัวอักษร

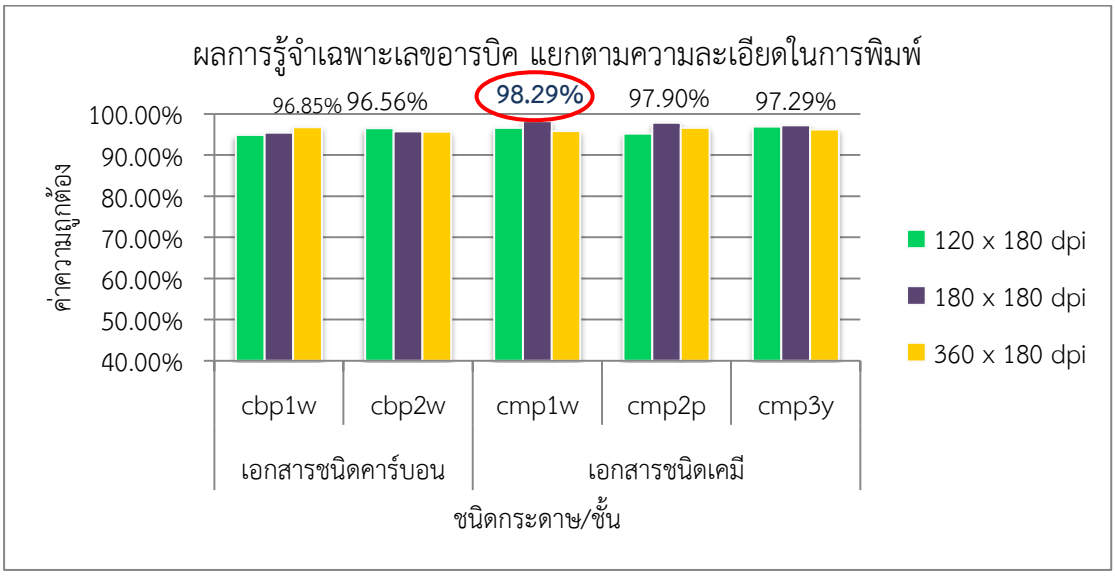
ตาราง 4-8 ในส่วนของปัจจัยลักษณะตัวอักษร สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-28 ลักษณะตัวอักษร ตัวหนา ให้ค่าความถูกต้องสูงสุด ทั้งเอกสารชนิดคาร์บอน cbp2w ให้ค่าความถูกต้องสูงสุดเท่ากับ 96.28% และเอกสารชนิดเคมี cmp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 98.58% และเอกสารสำเนาชนิดเคมีทั้ง cmp2p และ cmp3y ให้ค่าความถูกต้องสูงกว่าเอกสารสำเนาชนิดคาร์บอน เท่ากับ 1.32% และ 1.08% ตามลำดับ



ภาพประกอบ 4-28 ผลการรู้จำเฉพาะเลขอารบิค แยกตามลักษณะตัวอักษร

4.5.4 ผลการรู้จำเฉพาะเลขอารบิค แยกตามความละเอียดในการพิมพ์

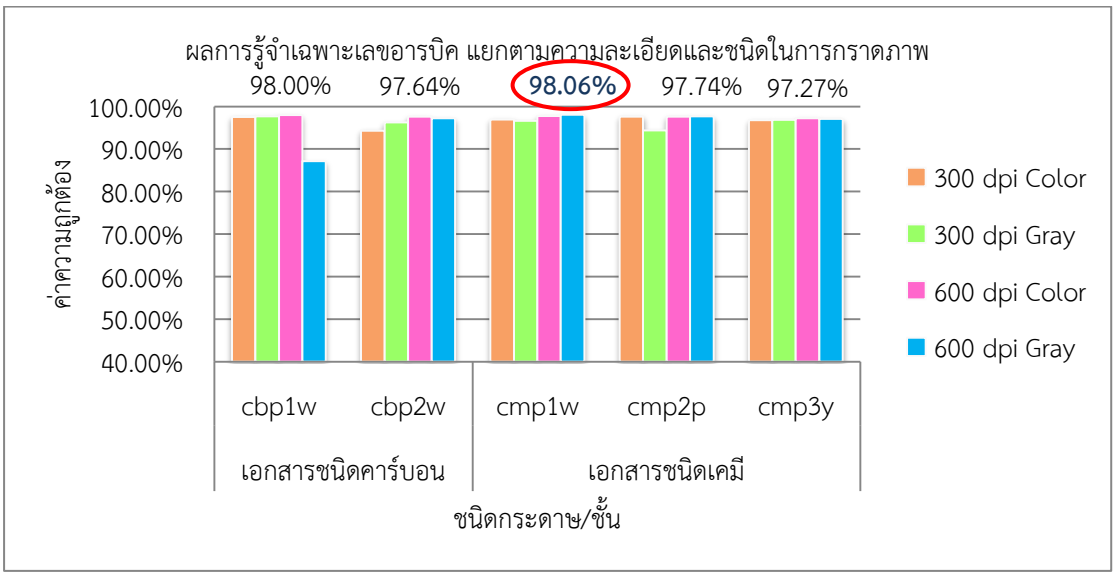
ตาราง 4-8 ในส่วนของปัจจัยความละเอียดในการพิมพ์ สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-29 เอกสารชนิดเคมี cmp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 98.29% ที่การพิมพ์ที่ความละเอียด 180 x 180 dpi เอกสารชนิดคาร์บอน cbp1w ให้ค่าความถูกต้องสูงสุดเท่ากับ 96.85% ซึ่งได้ผลใกล้เคียงกัน และเอกสารชนิดเคมีทั้ง cmp2p และ cmp3y ให้ค่าความถูกต้องสูงกว่าเอกสารสำเนาชนิดคาร์บอน เท่ากับ 1.34% และ 0.73% ตามลำดับ



ภาพประกอบ 4-29 ผลการรู้จำเฉพาะเลขอารบิค แยกตามความละเอียดในการพิมพ์

4.5.5 ผลการรู้จำเฉพาะเลขอารบิค แยกตามความละเอียดและชนิดในการกราดภาพ

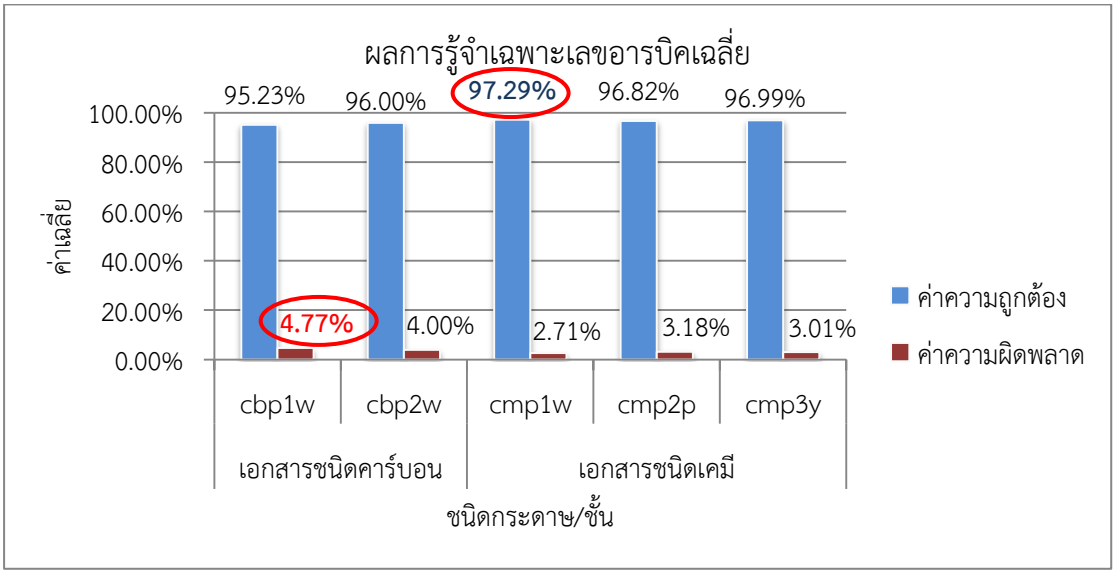
ตาราง 4-8 ในส่วนของปัจจัยความละเอียดและชนิดในการกราดภาพ สามารถแสดงเป็นกราฟเปรียบเทียบได้ดังภาพประกอบ 4-30 ความละเอียดและชนิดในการกราดภาพของเอกสารชนิดเคมี cmp1w ที่กราดภาพที่ความละเอียด 600 dpi ชนิด Gray ให้ค่าความถูกต้องสูงสุด เท่ากับ 98.06% และเอกสารชนิดคาร์บอน cbp1w ที่กราดภาพที่ความละเอียด 600 dpi ชนิด Color ให้ค่าความถูกต้องสูงสุดเท่ากับ 98.00% ซึ่งใกล้เคียงกันเท่ากับ 0.06%



ภาพประกอบ 4-30 ผลการรู้จำเฉพาะเลขอารบิค แยกตามความละเอียดและชนิดในการกราดภาพ

4.5.6 ผลการรู้จำเฉพาะเลขอารบิคเฉลี่ย

จากตาราง 4-8 สำหรับเอกสารสำหรับเลขอารบิค สามารถสรุปเป็นกราฟเปรียบเทียบได้ดัง ภาพประกอบ 4-31 จากผลการทดลองได้ผลลัพธ์ค่าเฉลี่ย เอกสารชนิดเคมี cmp1w ให้ค่าความถูกต้องเฉลี่ยสูงสุด เท่ากับ 97.29% และเอกสารชนิดคาร์บอน cbp2w ให้ค่าความถูกต้องเฉลี่ย 96.00% และเอกสารสำเนาชนิดคาร์บอน cbp1w ให้ค่าความถูกต้องเฉลี่ยต่ำสุด เท่ากับ 95.23% ซึ่งยังต่ำกว่าเอกสารสำเนาชนิดเคมีทั้ง cmp2p และ cmp3y เท่ากับ 0.82% และ 0.99% ตามลำดับ ไม่แตกต่างกันมาก

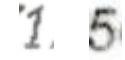


ภาพประกอบ 4-31 ผลการรู้จำเฉพาะเลขอารบิคเฉลี่ย

4.5.7 ปัญหาการรู้จำเฉพาะเลขอารบิค

การรู้จำเฉพาะเลขอารบิคมีการแสดงผลที่ผิดพลาดดังตาราง 4-9

ตาราง 4-9 ตัวอย่างการรู้จำที่ผิดพลาดเฉพาะเลขอารบิค

เลขไทย	ผลการรู้จำ	สาเหตุ
1 5	I (ตัวไอ) i 6	หมึกเข้ม กระจาย ตัวอักษรติดกัน แล้วมองเป็นตัวอักษรที่คล้ายกัน 

เลขอารบิครู้จำผิดพลาดมากคือลักษณะตัวอักษรตัวเอียง ได้แก่ 1 และ 5 ดังตาราง 4-9 ทำให้เพิ่มความหนาของตัวอักษรที่พิมพ์ด้วยกระดาษชนิดคาร์บอนสำเนาที่ 2 (cbp2w) สีเข้มทึบ มองเป็นเส้น ตัวอักษรติดกัน หมึกพิมพ์กระจาย

ชนิดแบบอักษร TH SarabunPSK จะรู้จำผิดพลาดมากที่สุด ได้แก่ 1 แต่ 5 จะผิดพลาดกับชนิดแบบอักษร Cordia New เกิดจากหมึกเข้มทึบ ความหนาของเส้นตัวอักษรทำให้ตัวอักษรติดกันได้ง่าย ทำให้ไปคล้ายกับตัวอักษรอื่น

4.6 ปัจจัยที่ทำให้ผลการรู้จำดีขึ้นได้มากที่สุด

ปัจจัยที่ทำให้ผลการรู้จำดีขึ้นได้มากที่สุดวิเคราะห์จากการนำผลการรู้จำโดยรวมของเอกสารดอทเมทริกซ์ชนิดคาร์บอน และชนิดเคมีของเอกสารทั้งหมด (อักษรไทย อักษรอังกฤษ เลขไทย และเลขอารบิค) เปรียบเทียบกับค่าเฉลี่ย ดังตาราง 4-10

ตาราง 4-10 เปรียบเทียบผลการรู้จำโดยรวมกับค่าเฉลี่ย

ปัจจัย	ชนิดแบบฟอร์ม	ค่าความถูกต้อง (%)														
		เอกสารชนิดคาร์บอน						เอกสารชนิดเคมี								
		cbp1w			cbp2w			cmp1w			cmp2p			cmp3y		
		$\bar{x}$	ผลต่าง	S.D.	$\bar{x}$	ผลต่าง	S.D.	$\bar{x}$	ผลต่าง	S.D.	$\bar{x}$	ผลต่าง	S.D.	$\bar{x}$	ผลต่าง	S.D.
ชนิดแบบอักษร	Angsana New	82.11	0.55	1.97	74.08	-1.34	2.56	83.72	**-1.62**	1.93	81.12	**-2.53**	2.81	77.57	**-2.39**	1.45
	Cordia New	86.57	5.00	1.30	79.01	**3.59**	1.53	90.00	**4.66**	2.75	86.83	**3.18**	0.74	82.99	**3.04**	1.70
	TH SarabunPSK	76.14	**-5.43**	13.41	73.41	**-2.01**	2.45	84.61	-0.72	3.34	83.43	-0.22	2.26	79.63	-0.33	1.28
ขนาดตัวอักษร	14	78.93	**-2.64**	2.19	68.50	**-6.92**	1.57	81.69	**-3.64**	1.63	79.48	**-4.17**	1.35	74.05	**-5.90**	1.87
	16	86.08	4.51	1.79	78.04	2.62	2.32	87.61	**2.28**	1.90	86.15	**2.50**	0.99	82.38	2.43	1.41
	18	79.80	-1.76	13.91	79.97	**4.55**	2.50	86.80	1.47	3.48	85.75	2.10	3.49	83.75	**3.79**	1.18
ลักษณะตัวอักษร	ตัวปกติ	82.98	1.42	4.39	76.85	**1.43**	1.49	85.91	0.58	3.40	86.45	**2.80**	0.44	81.56	**1.61**	1.17
	ตัวหนา	83.64	2.08	4.42	75.77	0.35	3.00	87.58	**2.25**	1.13	84.28	0.63	2.38	81.13	1.18	1.11
	ตัวเอียง	78.19	**-3.38**	3.73	75.05	**-0.37**	5.18	82.61	**-2.72**	2.48	80.65	**-3.00**	3.37	77.49	**-2.46**	2.05
ความละเอียดในการพิมพ์	120 x 180 dpi	77.56	**-4.00**	3.95	72.59	-2.83	1.20	81.40	**-3.93**	3.22	78.99	**-4.66**	3.90	75.66	**-4.30**	2.08
	180 x 180 dpi	80.39	-1.17	3.93	72.59	**-2.83**	3.25	84.91	-0.43	1.68	83.80	0.15	1.21	78.82	-1.14	1.43
	360 x 180 dpi	86.23	4.66	2.90	78.84	**3.42**	2.50	87.01	**1.68**	3.58	86.30	**2.65**	1.55	83.99	**4.04**	1.55
ความเอียดและชนิดในการกราดภาพ	300 dpi Color	82.31	0.75	4.57	73.13	**-2.29**	5.56	83.23	**-2.10**	2.71	83.50	-0.15	3.91	79.34	-0.62	5.76
	300 dpi Gray	82.71	1.14	5.03	74.36	-1.06	5.94	83.48	-1.85	3.02	81.22	**-2.43**	3.61	78.39	**-1.56**	5.47
	600 dpi Color	85.57	4.00	4.14	77.02	1.60	6.79	87.23	1.90	3.78	85.32	**1.67**	3.82	81.36	**1.41**	4.66
	600 dpi Gray	75.83	**-5.73**	14.95	77.50	**2.08**	6.33	87.54	**2.20**	3.67	85.13	1.48	4.40	81.15	1.20	5.13
ค่าเฉลี่ย		**81.57**		**5.41**	75.42		**3.39**	85.33		**2.73**	**83.65**		**2.51**	**79.95**		**2.46**

หมายเหตุ ตัวเลขสีน้ำเงิน หมายถึง ค่ามากที่สุด ตัวเลขสีแดง หมายถึง ค่าน้อยที่สุด
ผลต่าง = $\bar{x}$ – ค่าเฉลี่ย

จากตาราง 4-10 ผลต่าง คือความแตกต่างของค่าเฉลี่ยของแต่ละปัจจัยลบด้วยค่าเฉลี่ยโดยรวมทั้งหมด ทำให้ทราบว่าปัจจัยไหนมีข้อผิดพลาดมากหรือน้อย แสดงว่าปัจจัยนั้น ๆ จะมีผลต่อค่าความถูกต้องและค่าความผิดพลาดของการรู้จำเอกสารดอทเมทริกซ์

จึงสรุปได้ว่า ปัจจัยที่มีผลต่อการรู้จำมากที่สุด ตามลำดับ ได้แก่

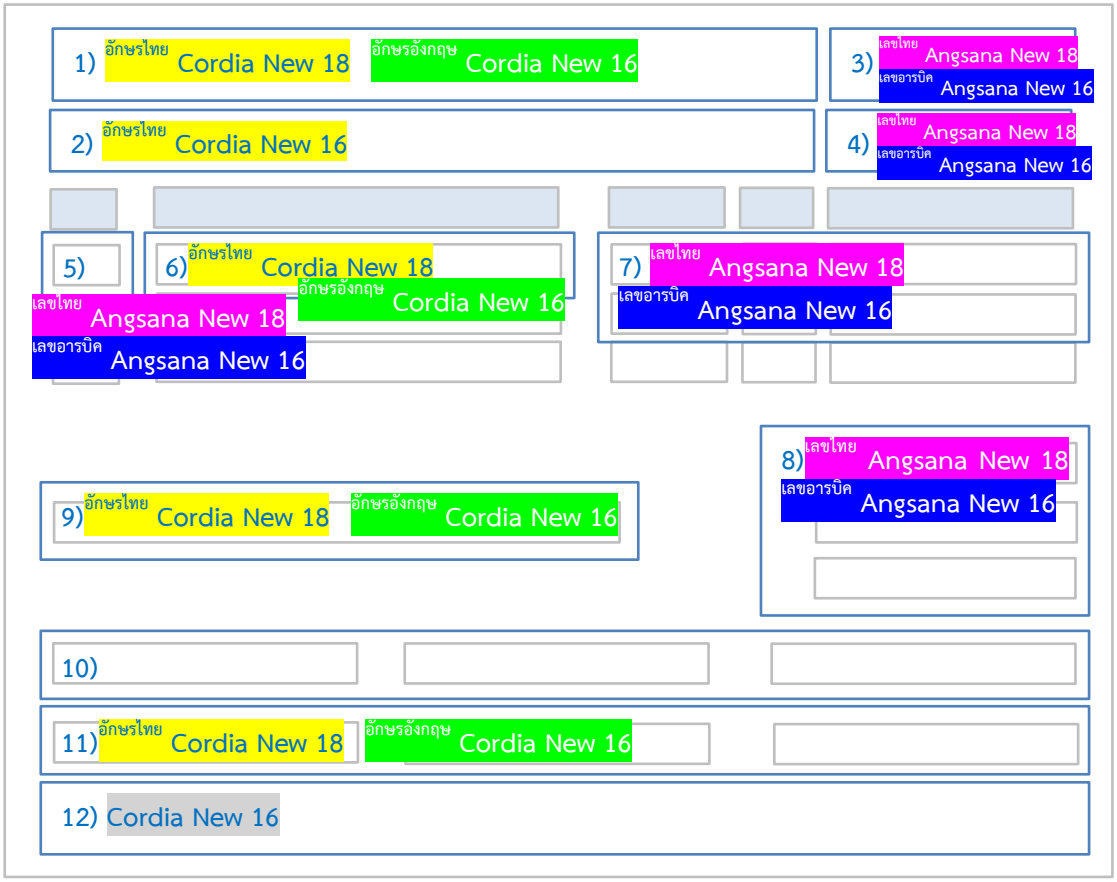
ลำดับ 1 ชนิดแบบอักษร

ลำดับ 2 ขนาดตัวอักษร

ลำดับ 3 ความละเอียดในการพิมพ์

4.6.1 ข้อแนะนำในการออกแบบฟอร์มที่ดี

ตัวอย่างการออกแบบฟอร์มที่ดี ดังภาพประกอบ 4-32



ภาพประกอบ 4-32 ตัวอย่างการออกแบบฟอร์มที่ดี

1) หัวกระดาษ เป็นชื่อหน่วยงาน จะเป็นตัวอักษร
2) ที่อยู่ เบอร์โทรศัพท์ อีเมล จะเป็นตัวอักษรและตัวเลข
3) เลขที่เอกสาร จะเป็นตัวอักษรและตัวเลข หรือเป็นตัวเลขอย่างเดียว
4) วัน/เดือน/ปี จะเป็นตัวอักษรและตัวเลข หรือเป็นตัวเลขอย่างเดียว
5) ลำดับที่ จะเป็นตัวเลข
6) รายการสินค้า รหัสสินค้า ชื่อสินค้า
7) จำนวนหน่วย ราคาต่อหน่วย จำนวนเงิน จะเป็นตัวเลข
8) มูลค่าสินค้า ภาษีมูลค่าเพิ่ม จำนวนเงินรวม จะเป็นตัวเลข

9) จำนวนเงินรวมเป็นตัวอักษร (Bath Text) จะเป็นตัวอักษร

10) ลงชื่อ ผู้ส่งสินค้า ผู้รับสินค้า ผู้รับเงิน เป็นลายมือชื่อ หรือลายเซ็น

11) ชื่อ ผู้ส่งสินค้า ผู้รับสินค้า ผู้รับเงิน จะเป็นตัวอักษร

12) หมายเหตุ จะเป็นตัวอักษรและตัวเลข

ส่วนที่เป็นตัวอักษร 1) 2) 6) 9) 11) และ 12) ถ้าเป็นอักษรไทยควรใช้ชนิดแบบอักษร Cordia New ขนาดตัวอักษร 18 ถ้าเป็นอักษรอังกฤษควรใช้ชนิดแบบอักษร Cordia New ขนาดตัวอักษร 16 แต่ถ้าเป็นตัวอักษรทั้งไทยและอังกฤษควรเลือกใช้ชนิดแบบอักษร Cordia New เพราะให้ค่าความถูกต้องสูงสุด

ถ้าเป็นแบบฟอร์มทางราชการที่ถูกบังคับชนิดแบบอักษร TH SarabunPSK ต้องเลือกขนาดตัวอักษรให้ใหญ่เพราะเป็นอันดับ 2 รองลงมาจากชนิดแบบอักษร จะทำให้การรู้จำดีขึ้นได้มากที่สุด แต่ถ้าไม่ใช่หนังสือราชการควรเลือก Cordia New

อีเมลจะเป็นอักษรอังกฤษ ควรใช้ชนิดแบบอักษร Cordia New ขนาดตัวอักษร 16

ส่วนที่เป็นตัวเลข 3) 4) 5) 7) และ 8) ถ้าเป็นเลขไทยควรใช้ชนิดแบบอักษร Angsana New ขนาดตัวอักษร 18 ถ้าเป็นเลขอารบิคควรใช้ชนิดแบบอักษร Angsana New ขนาดตัวอักษร 16 ถ้าจำเป็นต้องใช้ทั้งเลขไทยและเลขอารบิคควรเลือกใช้ชนิดแบบอักษร Angsana New เพราะให้ค่าความถูกต้องสูงสุด

เบอร์โทรศัพท์จะเป็นตัวเลขอารบิคควรใช้ชนิดแบบอักษร Angsana New ขนาดตัวอักษร 16

เลขที่เอกสารควรเลือกใช้เขอารบิค แต่ถ้าจำเป็นต้องใช้เลขไทยควรใช้ชนิดแบบอักษร Angsana New ขนาดตัวอักษร 18

การเลือกใช้ตัวเลขควรเลือกตัวเลขอารบิคเพราะให้ค่าความถูกต้องสูงกว่าเลขไทย แต่ถ้าจำเป็นต้องใช้เลขไทยก็ต้องปรับให้ขนาดใหญ่ขึ้น

ที่อยู่ จะเป็นตัวอักษรและตัวเลข ควรเลือกใช้ชนิดแบบอักษร Cordia New ขนาดตัวอักษร 16

รายการสินค้า รหัสสินค้า ชื่อสินค้า ควรเป็นอักษรไทย หรืออักษรอังกฤษ หรือเลขอารบิค หรือเลขไทย อย่างใดอย่างหนึ่งจะทำให้การรู้จำดีที่สุด สำหรับอักษรไทยควรใช้ชนิดแบบอักษร Cordia New ขนาดตัวอักษร 18 อักษรอังกฤษควรใช้ชนิดแบบอักษร Cordia New ขนาดตัวอักษร 16 ตัวเลขอารบิคควรใช้ชนิดแบบอักษร Angsana New ขนาดตัวอักษร 16 กรณีที่มีทั้งอักษรไทย อักษรอังกฤษ และตัวเลข ควรใช้ชนิดแบบอักษร Cordia New ขนาดตัวอักษร 16

ตัวเลขที่ไม่ควรใช้เป็นรหัสสินค้าได้แก่ อักษร s ดังตัวอย่างในตาราง 4-5 เลข ๑ ดังตัวอย่างในตาราง 4-7 เลข 1 ดังตัวอย่างในตาราง 4-9 เพราะทำการรู้จำแล้วผิดพลาดบ่อย

รหัสสินค้าไม่ควรรวมตัวอักษรกับตัวเลข ๐ (เลขศูนย์ไทย) กับ o (อักษรอังกฤษ โอ) ไว้ด้วยกันเพราะผิดพลาดบ่อย ดังตัวอย่างในตาราง 4-9

เพื่อป้องกันการผิดพลาดของการรู้จำ โปรแกรม ABBYY Find Reader สามารถกำหนดพื้นที่ได้ทันทีว่าบริเวณไหนควรเป็นตัวอักษรไทย ตัวอักษรอังกฤษ หรือตัวเลข ก่อนการแปลงไฟล์ได้ เช่น ๐ (เลขศูนย์ไทย) กับ o (อักษรอังกฤษ โอ) โดยการสร้างฟอร์มเฉพาะ ว่าบริเวณพื้นที่ไหน

ควรวิเคราะห์อย่างไร มีให้เลือกได้หลายประเภท ได้แก่ ข้อความ ตาราง รูปภาพ ภาพพื้นหลัง บาร์โค๊ด หรือจะไม่กำหนดก็ได้ ดัง

ข้อควรระวังในการทำการรู้จำ ได้แก่ ลักษณะของเอกสาร ถ้าเอกสารไม่ชัด สีตัวอักษรไม่เข้ม ความเข้มของผ้าหมึกมากหรือน้อยเกินไป คุณภาพของแรงกดของหัวพิมพ์ต่ำ กระดาษมีความชื้น ทำให้ผลการรู้จำผิดพลาด หรือไม่สามารถทำการรู้จำได้

ควรพิมพ์เอกสารด้วยกระดาษชนิดเคมีเพราะสามารถรู้จำได้ดีกว่าเอกสารชนิดคาร์บอน

ชนิดแบบอักษรที่ไม่ควรใช้คือ TH SarabunPSK ดังตาราง 4-1 เพราะทำการรู้จำแล้วผิดพลาดมากที่สุด ทั้งอักษรไทย อักษรอังกฤษ เลขไทย และเลขอารบิค

บทที่ 5

บทสรุปและข้อเสนอแนะ

บทนี้จะกล่าวถึงสรุปผลการวิจัย ได้แก่ สรุปผลการรู้จำโดยรวม สรุปผลการรู้จำเฉพาะอักษรไทย สรุปผลการรู้จำเฉพาะอักษรอังกฤษ สรุปผลการรู้จำเฉพาะเลขไทย สรุปผลการรู้จำเฉพาะเลขอารบิค รวมทั้งประโยชน์ที่ได้รับจากงานวิจัย และข้อเสนอแนะในการดำเนินการวิจัย

5.1 สรุปผลการวิจัย

จากการดำเนินการวิจัย ประมวลผล เปรียบเทียบ วิเคราะห์ปัจจัยที่มีผลต่อการรู้จำเอกสารดอทเมทริกซ์ ชนิดคาร์บอน และชนิดเคมี เพื่อนำไปสู่การออกแบบฟอร์มที่ดีที่สุด ให้สามารถทำการรู้จำได้ดีขึ้น ถูกต้องแม่นยำขึ้นสรุปได้ดังนี้

5.1.1 สรุปผลการรู้จำโดยรวม

ควรเลือกใช้กระดาษชนิดเคมี ชนิดแบบอักษร Cordia New ขนาดตัวอักษร 16 ลักษณะตัวอักษร ตัวหนา ความละเอียดในการพิมพ์ 360 x 180 dpi และกราดภาพที่ความละเอียด 600 dpi ชนิด Gray เพราะให้ค่าความถูกต้องสูงสุด เพื่อให้ได้มาซึ่งแบบฟอร์มที่เหมาะสมสำหรับการนำไปทำการรู้จำมากที่สุด ทำให้การจัดการงานด้านเอกสารสมบูรณ์แบบและถูกต้องมากที่สุด

5.1.2 สรุปผลการรู้จำเฉพาะอักษรไทย

ควรเลือกใช้กระดาษชนิดเคมี ชนิดแบบอักษร Cordia New ขนาดตัวอักษร 18 ลักษณะตัวอักษร ตัวปกติ ความละเอียดในการพิมพ์ 360 x 180 dpi และกราดภาพที่ความละเอียด 600 dpi ชนิด Gray เพราะให้ค่าความถูกต้องสูงสุด เพื่อให้ได้มาซึ่งแบบฟอร์มที่เหมาะสมสำหรับการนำไปทำการรู้จำมากที่สุด ทำให้การจัดการงานด้านเอกสารสมบูรณ์แบบและถูกต้องมากที่สุด

5.1.3 สรุปผลการรู้จำเฉพาะอักษรอังกฤษ

ควรเลือกใช้กระดาษชนิดเคมี ชนิดแบบอักษร Cordia New ขนาดตัวอักษร 16 ลักษณะตัวอักษร ตัวหนา ความละเอียดในการพิมพ์ 360 x 180 dpi และกราดภาพที่ความละเอียด 600 dpi ชนิด Gray เพราะให้ค่าความถูกต้องสูงสุด เพื่อให้ได้มาซึ่งแบบฟอร์มที่เหมาะสมสำหรับการนำไปทำการรู้จำมากที่สุด ทำให้การจัดการงานด้านเอกสารสมบูรณ์แบบและถูกต้องมากที่สุด

5.1.4 สรุปผลการรู้จำเฉพาะเลขไทย

ควรเลือกใช้กระดาษชนิดเคมี ชนิดแบบอักษร Angsana New ขนาดตัวอักษร 18 ลักษณะตัวอักษร ตัวเอียง ความละเอียดในการพิมพ์ 180 x 180 dpi และกราดภาพที่ความละเอียด

600 dpi ชนิด Gray เพราะให้ค่าความถูกต้องสูงสุด เพื่อให้ได้มาซึ่งแบบฟอร์มที่เหมาะสมสำหรับการนำไปทำการรู้จำมากที่สุด ทำให้การจัดการงานด้านเอกสารสมบูรณ์แบบและถูกต้องมากที่สุด

5.1.5 สรุปผลการรู้จำเฉพาะเลขอารบิค

ควรเลือกใช้กระดาษชนิดเคมี ชนิดแบบอักษร Angsana New ขนาดตัวอักษร 16 ลักษณะตัวอักษร ตัวหนา ความละเอียดในการพิมพ์ 180 x 180 dpi และกราดภาพที่ความละเอียด 600 dpi ชนิด Gray เพราะให้ค่าความถูกต้องสูงสุด เพื่อให้ได้มาซึ่งแบบฟอร์มที่เหมาะสมสำหรับการนำไปทำการรู้จำมากที่สุด ทำให้การจัดการงานด้านเอกสารสมบูรณ์แบบและถูกต้องมากที่สุด

5.2 ประโยชน์ที่ได้จากงานวิจัย

จากการดำเนินการวิจัยทำให้ทราบถึงปัจจัยที่มีผลกระทบต่อการรู้จำตัวอักษรสำหรับเอกสารที่พิมพ์ด้วยเครื่องพิมพ์ดอทเมทริกซ์ได้แก่ กระดาษที่ใช้ ชนิดแบบอักษร ขนาดตัวอักษร ลักษณะตัวอักษร ความละเอียดในการพิมพ์ และความละเอียดในชนิดและการกราดภาพ ทำให้ทราบแนวทางในการออกแบบฟอร์มเอกสารให้เหมาะสมสามารถนำไปรู้จำแล้วได้ผลลัพธ์ดีที่สุด สามารถนำไปใช้งานได้อย่างทันท่วงทีและมีประสิทธิภาพ

5.3 ข้อเสนอแนะในการดำเนินการวิจัย

การศึกษาผลกระทบจากแรงกดของหัวพิมพ์ที่แตกต่างกัน ของเครื่องพิมพ์ดอทเมทริกซ์ที่คุณภาพ อายุการใช้งานต่างกัน กระดาษต่อเนื่องชนิดเคมีที่จำนวนชั้นเท่ากันและสีแตกต่างกัน การออกแบบโปรแกรมในการนับตัวอักษรถูก/ผิดที่สามารถนับได้อัตโนมัติผ่านข้อจำกัดต่างๆ ได้เพื่อผู้ใช้ไม่ต้องเปรียบเทียบตัวอักษรเองในกรณีที่ผิดพลาด การพัฒนานวัตกรรมที่สามารถพิมพ์เอกสาร การศึกษาการกราดภาพ และทำการรู้จำและแสดงผลได้ในเครื่องเดียว

บรรณานุกรม

ABBYY, "What Technology lies behind OCR?" [Online], 2015. http://finereader.abbyy.com/about_ocr/whatis_ocr

Bayer T. A. and Mogg-Schneider,H.U., 1997."A Generic System for Processing Invoices" Proceedings of the Fourth International Conference on Document Analysis and Recognition Vol.2, Ulm, 18-20 Aug. 1997, pp 740-744.

Berrin A. Yanikoglu, 2000. "Pitch-based segmentation and recognition of dot-matrix text" International Journal on Document Analysis and Recognition Volume 3, Issue 1, 21 Apr. 2000, pp 34-39.

C. H. Chen and J. L. DeCurtins, 1992. "A Segmentation-free Approach to OCR" Proceedings, 1992., IEEE Workshop on Applications of Computer Vision, Palm Springs, CA, 30 Nov.-2 Dec. 1992, pp 190-196.

Enrico Sorio, Alberto Bartoli, Giorgio Davanzo and Eric Medvet, 2012. "A Domain Knowledge-based Approach for Automatic Correction of Printed Invoices" International Conference on Information Society (i-Society), London,25-28 Jun. 2012, pp 151-155.

Iuliu Konya, Stefan Eickeler and Christoph Seibert, 2011. "Character enhancement for historical newspapers printed using hot metal typesetting" 2011 International Conference on Document Analysis and Recognition (ICDAR), Beijing, 18-21 Sep. 2011, pp 936-940.

Lee Dar-Shyang and Smith Ray, 2012. "Improving Book OCR by Adaptive Language and Image Models" 10th IAPR International Workshop on Document Analysis Systems (DAS), Gold Cost, QLD, 27-29 Mar. 2012, pp 155-119.

MING Delie, LIU Jian and TIAN Jinwen, 2002. "The Design and Implementation of a Chinese Financial Invoice Recognition System" Image Processing and Multimedia Communications 4th EURASIP-IEEE Region 8 International Symposium on VIPromCom, 16-19 Jun. 2002, pp 79-82.

SHUNJI MORI, CHING Y. SUEN and KAZUHIKO YAMAMOTO, 1992. “Historical Review of OCR Research and Development” Proceeding of the IEEE Vol.80, No. 7, July. 1992, pp 1029-1058.

Tanushyam Chattopdhyyay, Priyanka Sinha and Provot Biswas, 2011. “Performance of Document Image OCR Systems for Recognizing Video Texts on Embedded Platform” International Conference on Computational Intelligence and Communication Systems, IEEE, 2011, pp 606-610.

วรพจน์ พีระวิทย์, 2548. การเพิ่มประสิทธิภาพของระบบรู้จำอักษรไทย โดยการซ่อมแซมตัวอักษรไม่สมบูรณ์. วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต มหาวิทยาลัยเกษตรศาสตร์.

ภาคผนวก

ภาคผนวก ก

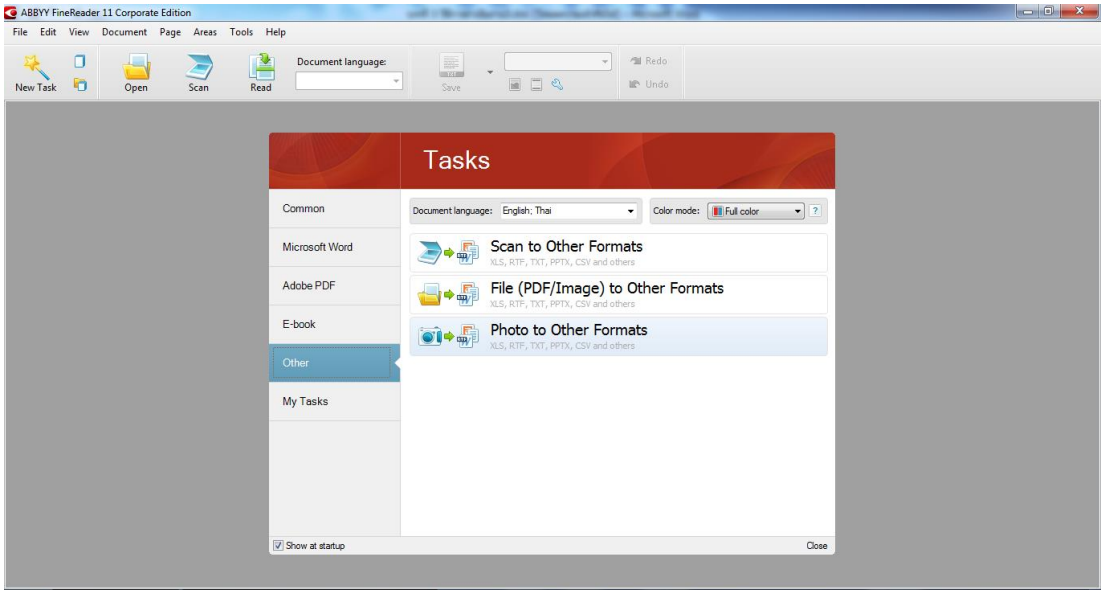
คู่มือการใช้โปรแกรม ABBYY Fine Reader 11

ในส่วนนี้จะกล่าวถึงขั้นตอนการรู้จำโปรแกรม ABBYY Fine Reader 11 โดยการรู้จำทั้งหน้าเอกสาร และการสร้างฟอร์มเฉพาะสำหรับการรู้จำโดยการกำหนดบริเวณพื้นที่

1. การรู้จำด้วยโปรแกรม ABBYY Fine Reader 11 โดยการรู้จำทั้งหน้าเอกสาร

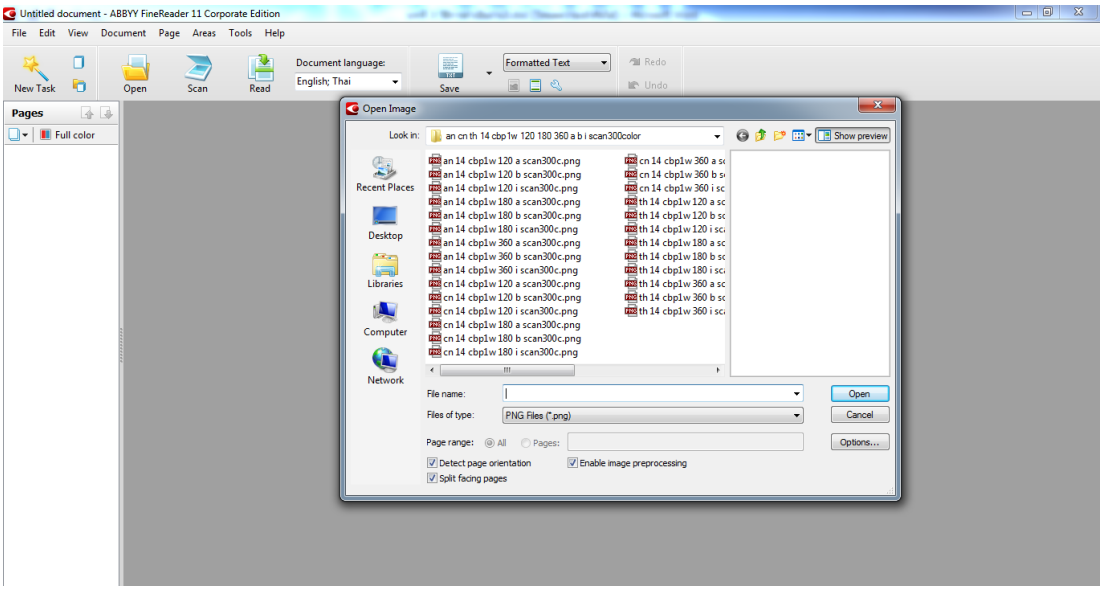
ขั้นตอนการรู้จำด้วยโปรแกรม ABBYY Fine Reader 11 โดยการรู้จำทั้งหน้าเอกสาร ดังนี้

1) เปิดโปรแกรม ABBYY Fine Reader 11 แล้วคลิกเมนู Other เลือก Photo to Other Formats ดังภาพประกอบ ก-1



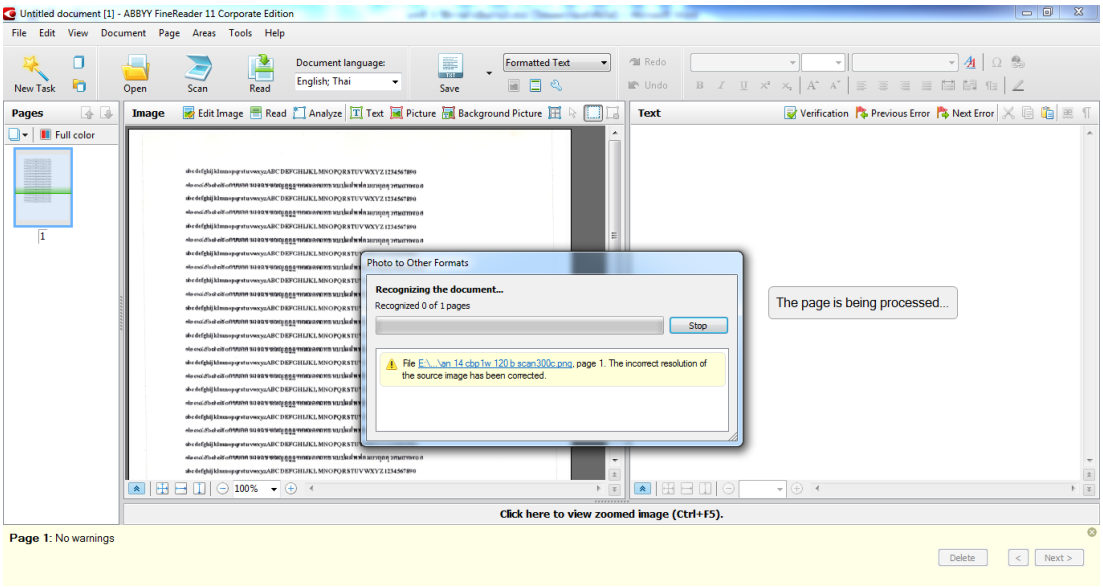
ภาพประกอบ ก-1 หน้าต่างโปรแกรม ABBYY Fine Reader 11

2) จะปรากฏหน้าต่างการทำงานและหน้าต่าง Open Image ให้คลิกเลือกไฟล์รูปภาพที่ต้องการ แล้วคลิกปุ่ม Open ดังภาพประกอบ ก-2



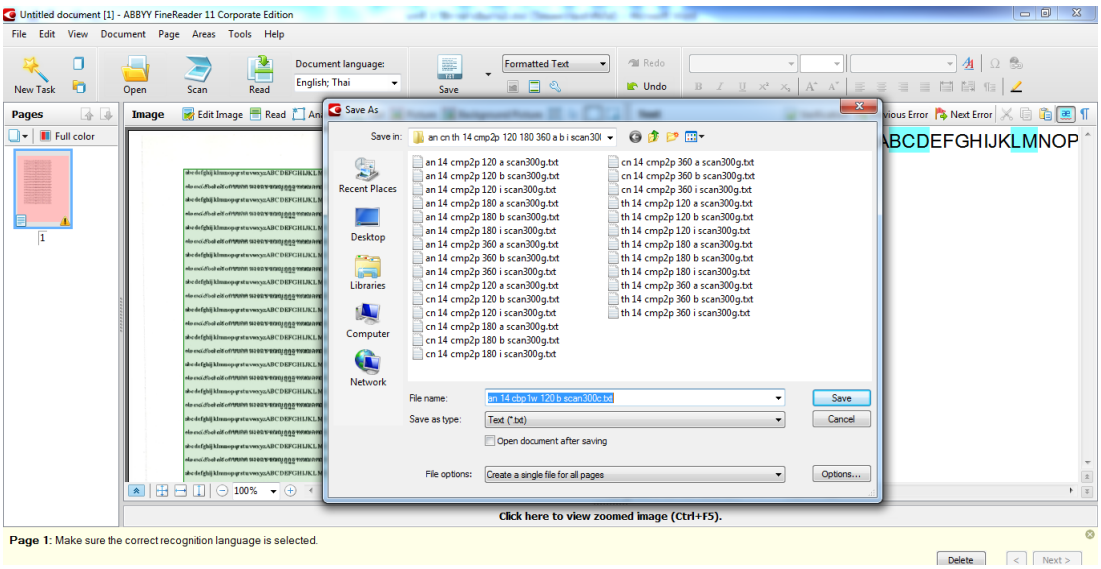
ภาพประกอบ ก-2 หน้าต่าง Open Image

3) โปรแกรมก็จะอ่านไฟล์ภาพเอกสาร แล้วแปลงไฟล์ ดังภาพประกอบ ก-3



ภาพประกอบ ก-3 หน้าต่างการแปลงไฟล์รูปภาพเป็นตัวอักษร

4) โปรแกรมก็จะเริ่มแปลงไฟล์และแสดงเป็นแถบสีคลุมพื้นที่การแปลงไฟล์ แยกเป็น 3 แถบสี ได้แก่ พื้นที่สีแดงคือพื้นที่ที่แปลงเป็นรูปภาพ พื้นที่สีน้ำเงินแปลงเป็นตาราง และพื้นที่สีเขียวแปลงเป็นตัวอักษร เมื่อโปรแกรมทำงานเสร็จ ก็จะขึ้นหน้าต่างให้บันทึก เลือกบันทึกเป็นไฟล์ .txt แล้วคลิกปุ่ม Save ดังภาพประกอบ ก-4

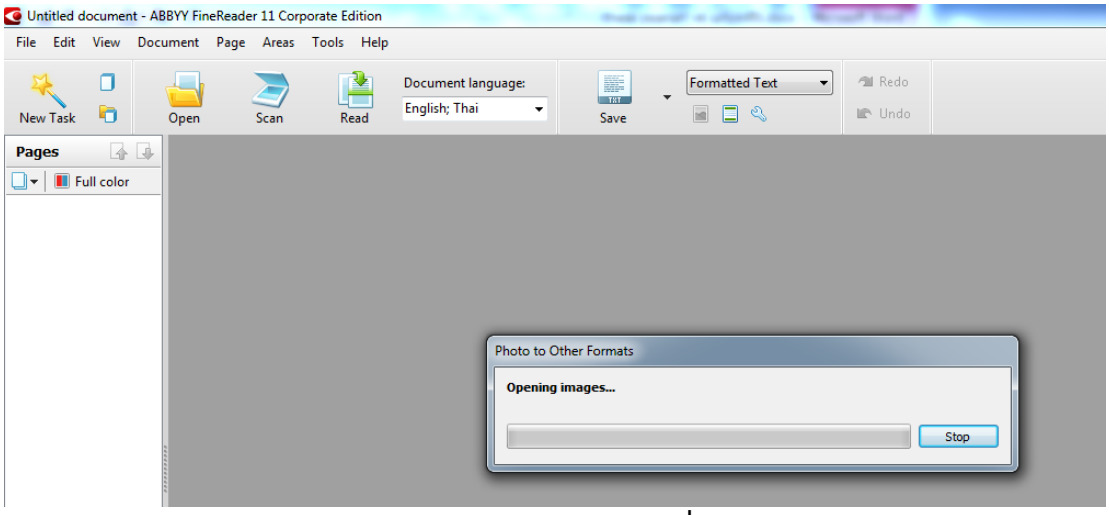


ภาพประกอบ ก-4 หน้าต่างการบันทึกไฟล์ .txt

2. การสร้างฟอร์มเฉพาะสำหรับการรู้จำ

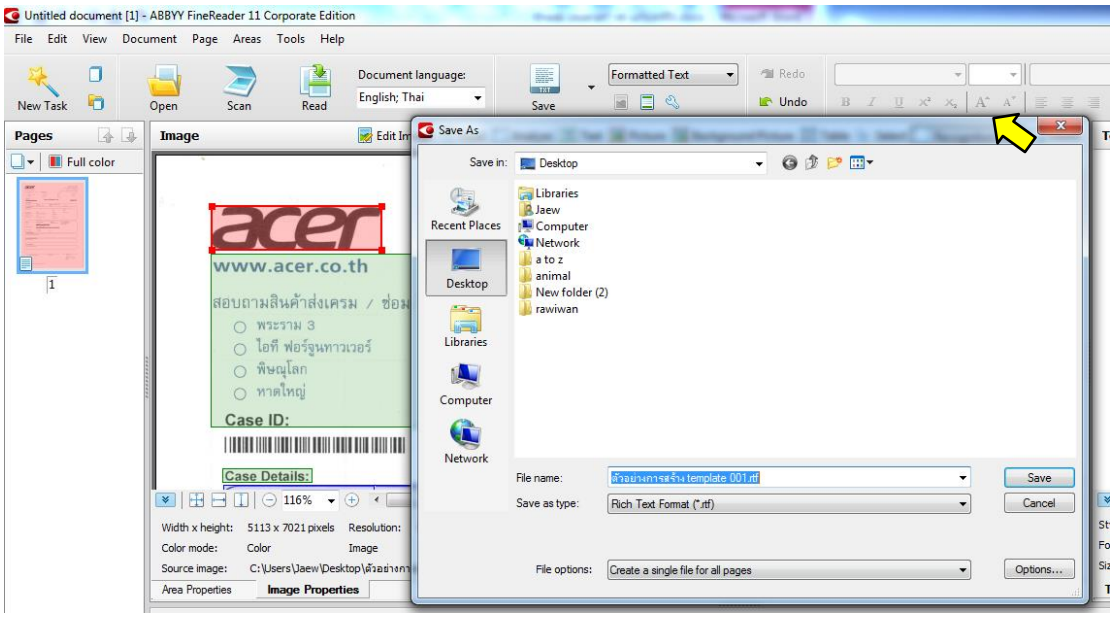
โปรแกรม ABBYY Fine Reader สามารถสร้างแบบฟอร์มเฉพาะสำหรับการรู้จำได้ดังนี้

1) เปิดโปรแกรม ABBYY Fine Reader
2) เปิดไฟล์เอกสารที่ต้องการทำการรู้จำ ดังภาพประกอบ ก-5

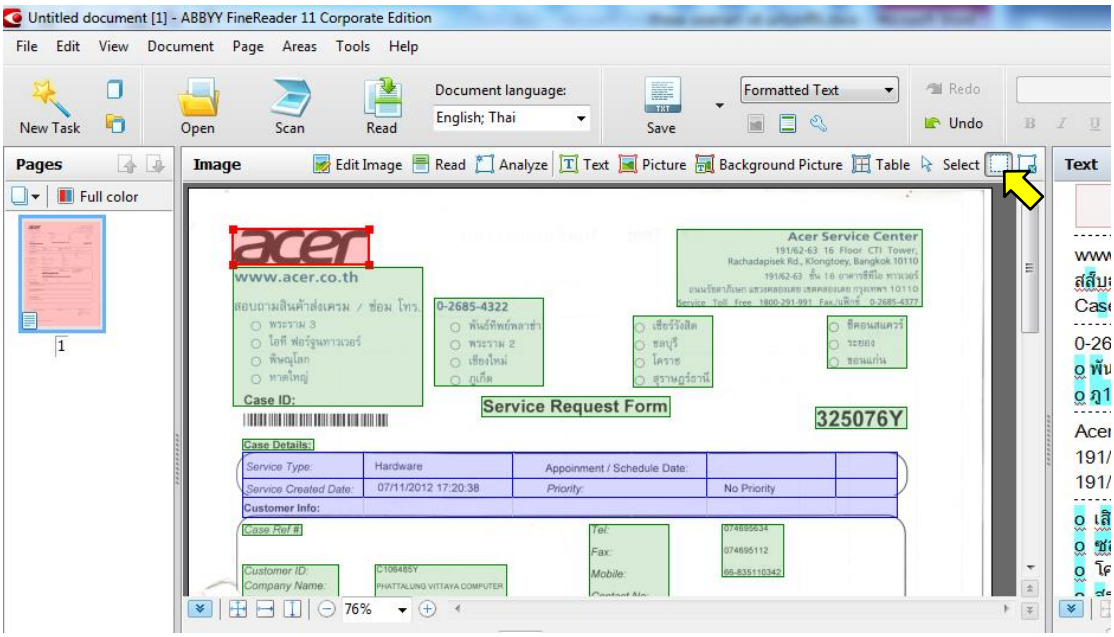


ภาพประกอบ ก-5 เปิดไฟล์เอกสารที่ต้องการรู้จำ

โปรแกรมจะเลือกพื้นที่ให้อัตโนมัติว่าบริเวณไหนควรเป็นพื้นที่อะไร ดังภาพประกอบ ก-6 และ ภาพประกอบ ก-7 เช่น ข้อความ บาร์โค๊ด รูปภาพ เท่าที่โปรแกรมสามารถเลือกได้ขึ้นอยู่กับคุณภาพความชัดเจนของเอกสาร แล้วแปลงไฟล์ให้อัตโนมัติ และแสดงหน้าต่างให้บันทึกไฟล์ตามที่ต้องการ กรณีนี้ไม่ต้องบันทึกให้ปิดหน้าต่างการบันทึกไฟล์ แล้วตรวจสอบว่าพื้นที่ไหนเราต้องการ พื้นที่ไหนเราไม่ต้องการก็สามารถลบออกได้ ถ้าต้องการพื้นที่อื่นอีกก็กำหนดเพิ่มได้

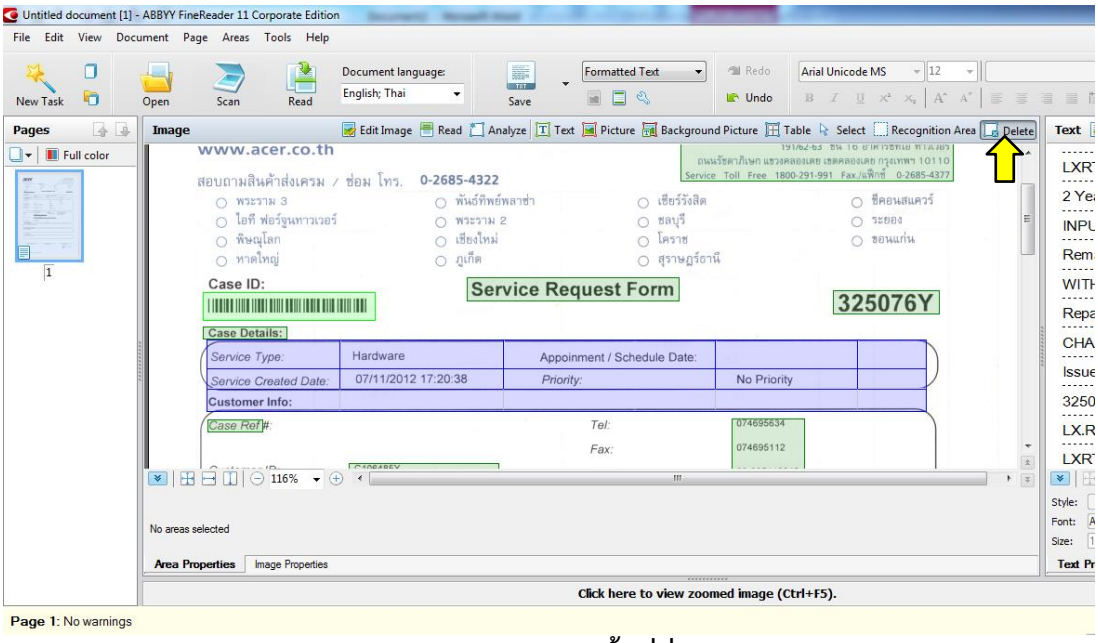


ภาพประกอบ ก-6 การรู้จำและบันทึกไฟล์อัตโนมัติ



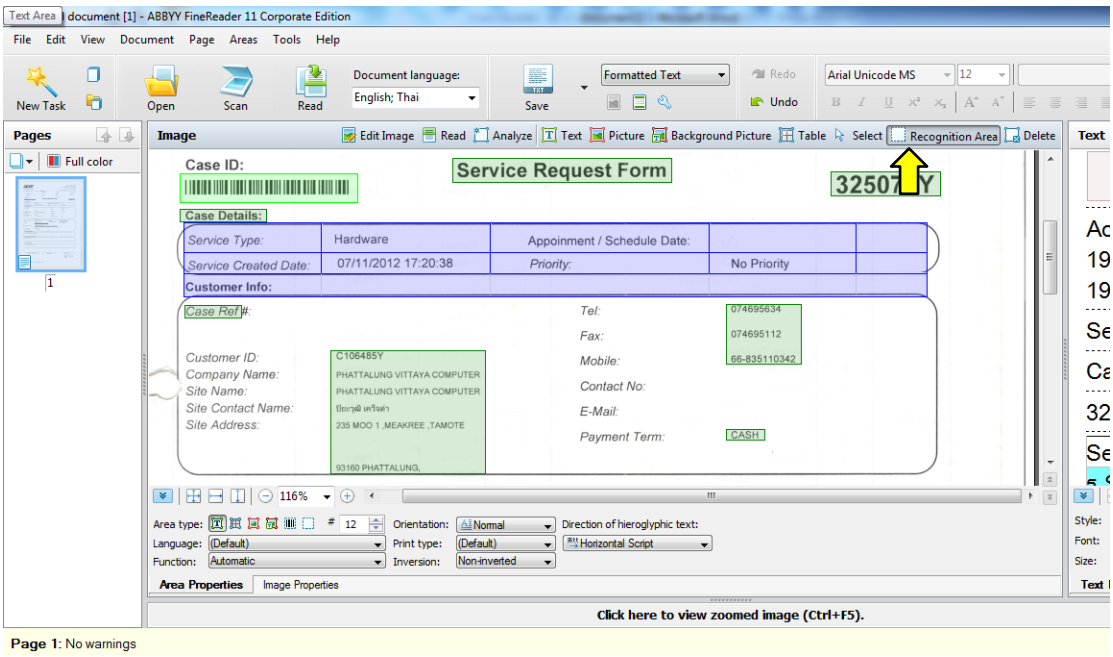
ภาพประกอบ ก-7 โปรแกรมเลือกการวิเคราะห์ประเภทให้อัตโนมัติ

3) การลบพื้นที่ที่ไม่ต้องการ ดังภาพประกอบ ก-8 คลิกปุ่ม Delete Area แล้วเลื่อนไปคลิกบริเวณพื้นที่ที่ไม่ต้องการ โปรแกรมก็จะยกเลิกการเลือกพื้นที่นั้น



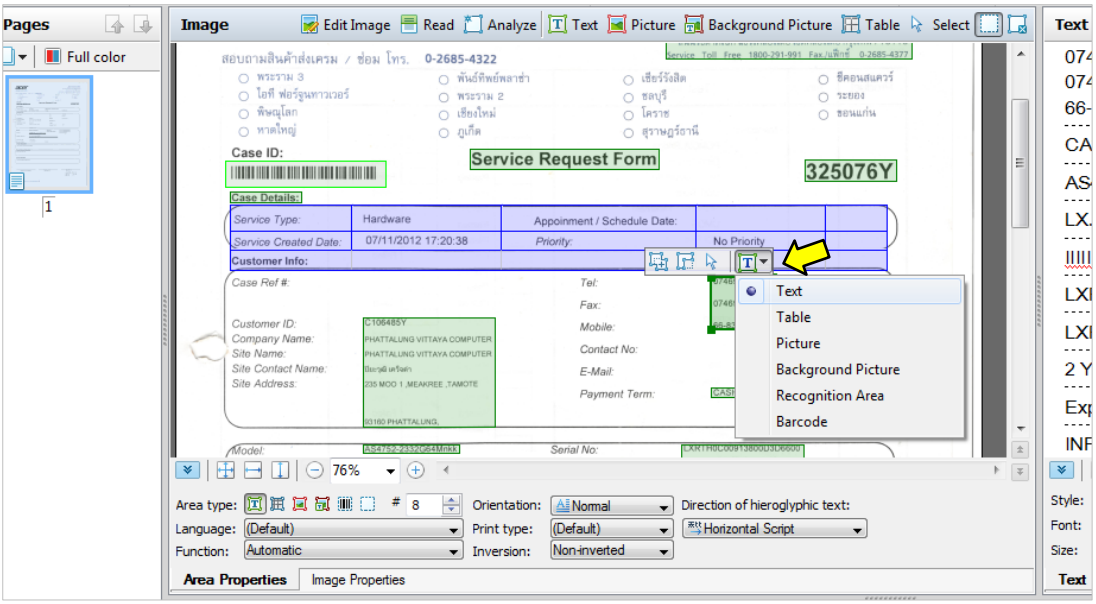
ภาพประกอบ ก-8 การลบพื้นที่ที่ไม่ต้องการ

4) การเลือกพื้นที่ที่ต้องการเพิ่ม ดังภาพประกอบ ก-9 คลิกปุ่ม Draw Recognition Area แล้วเลือกพื้นที่ที่ต้องการ



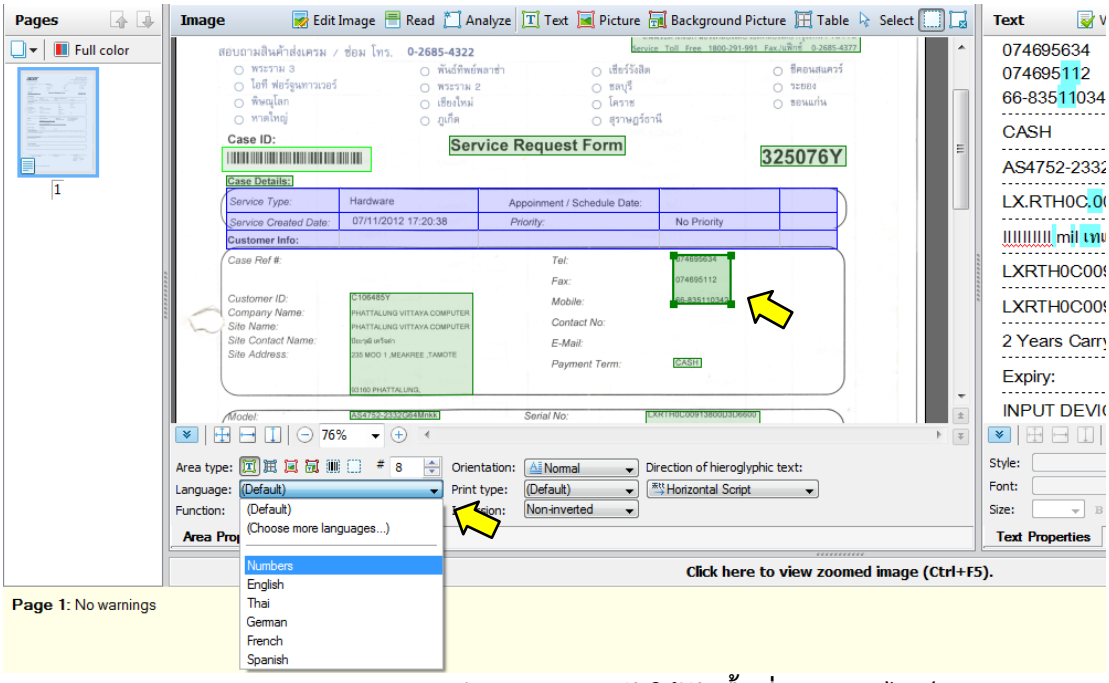
ภาพประกอบ ก-9 การเพิ่มพื้นที่ที่ต้องการ

5) กำหนดพื้นที่การแปลงไฟล์ให้วิเคราะห์ตามที่ต้องการ ดังภาพประกอบ ก-10 คลิกเลือกพื้นที่ที่ต้องการ → คลิกลูกศรลง เลือกการวิเคราะห์ที่ต้องการ



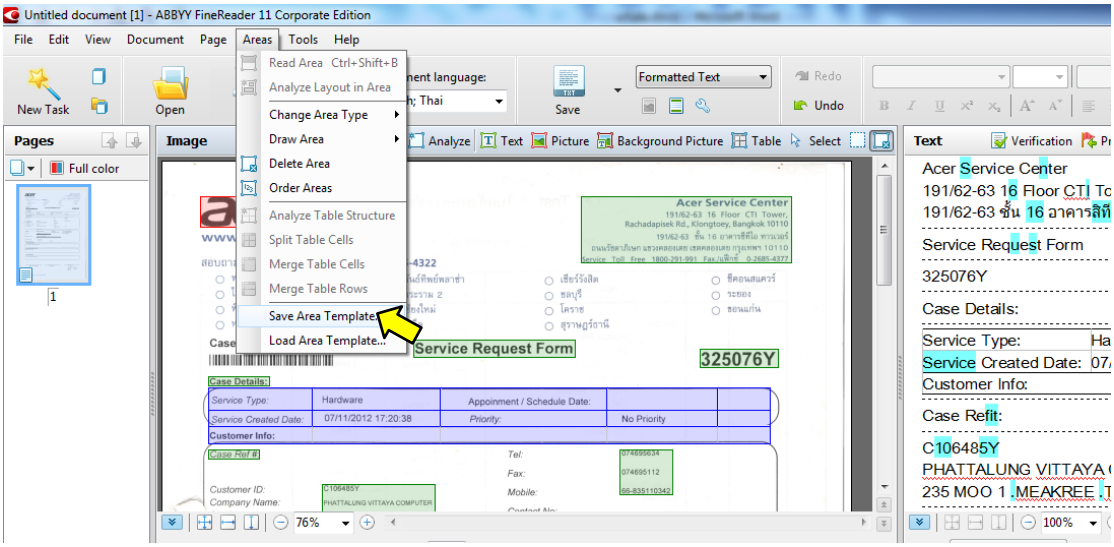
ภาพประกอบ ก-10 การกำหนดพื้นที่การแปลงไฟล์ให้วิเคราะห์ตามที่ต้องการ

6) กำหนดคุณสมบัติให้กับพื้นที่ ดัง**ผิดพลาด! ไม่พบแหล่งการอ้างอิง** คลิกถบ Area Properties กำหนดคุณสมบัติที่ต้องการจาก Area Type จนครบทั้งเอกสาร



ภาพประกอบ ก-11 การกำหนดคุณสมบัติให้กับพื้นที่การแปลงไฟล์

7) บันทึกแบบฟอร์มเฉพาะ ดังภาพประกอบ ก-12 คลิกเมนู Area → Save Area Template เมื่อต้องการเรียกใช้แบบฟอร์มเฉพาะ คลิกเมนู Area → Load Area Template



ภาพประกอบ ก-12 การบันทึกแบบฟอร์มเฉพาะ

ภาคผนวก ข

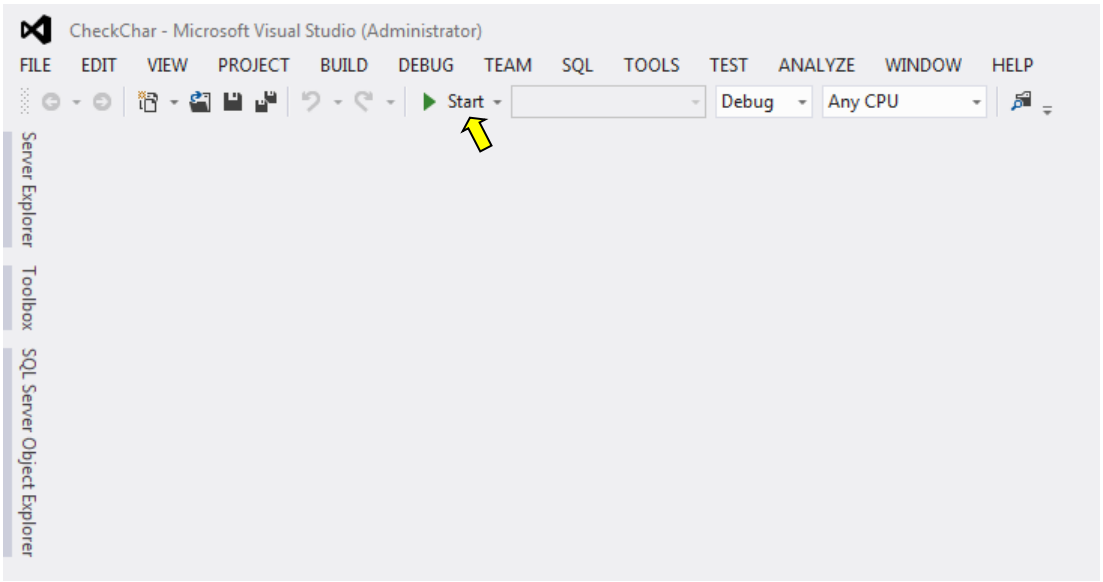
คู่มือการใช้โปรแกรม Check Character

ในส่วนนี้จะกล่าวถึงขั้นตอนการใช้โปรแกรม Check Character ที่ผู้วิจัยได้พัฒนาขึ้นสำหรับการปรียบเทียบนับตัวอักษรถูก/ผิด

1. ขั้นตอนการวัดผลการรู้จำ

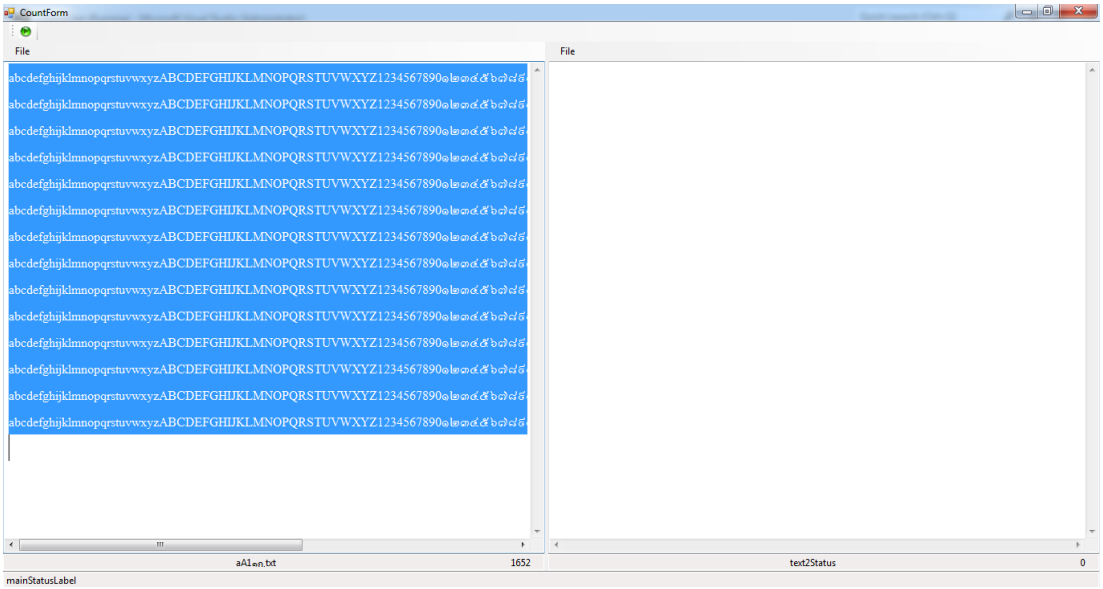
การวัดผลการรู้จำแบ่งเป็น 5 ขั้นตอน ดังนี้

1) เปิดโปรแกรม Check Character คลิกปุ่ม Start ดังภาพประกอบ ข-1



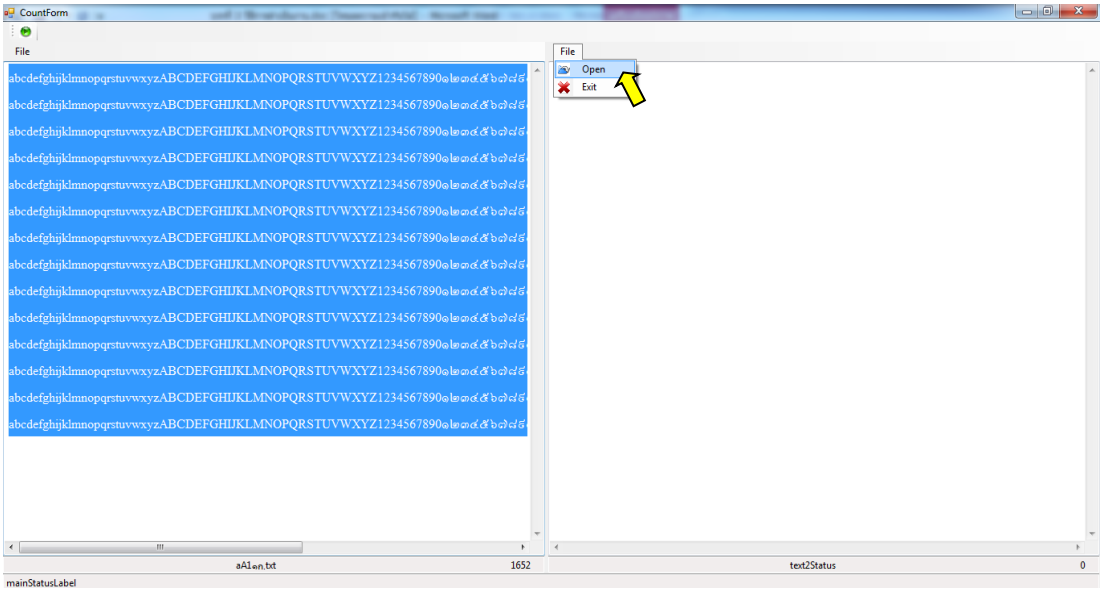
ภาพประกอบ ข-1 หน้าต่างโปรแกรม Check Character

โปรแกรมจะแสดงหน้าต่าง Count Form ดังภาพประกอบ ข-2



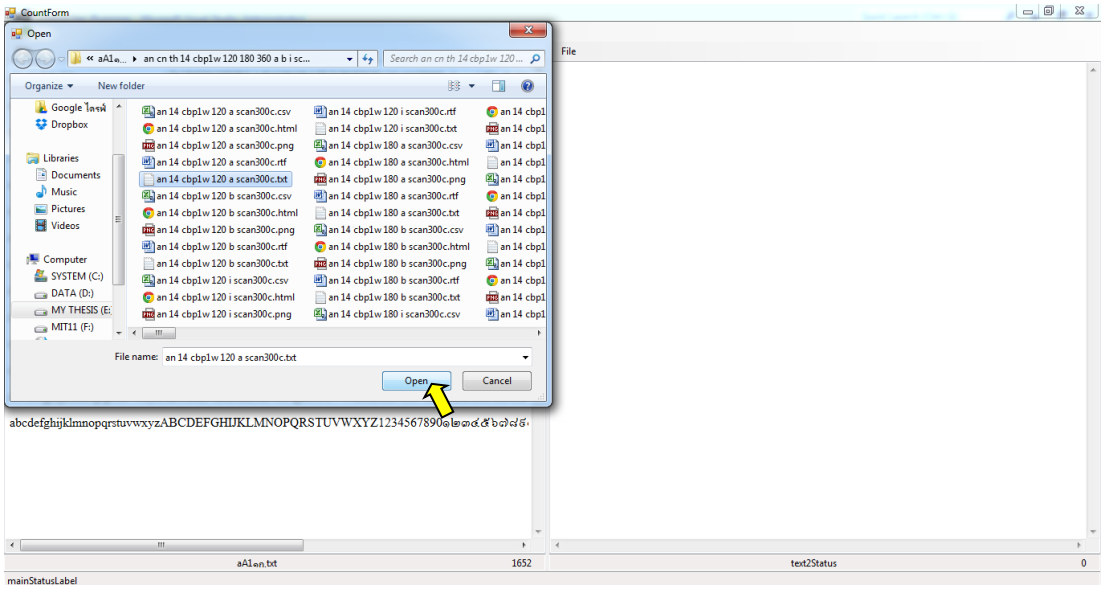
ภาพประกอบ ข-2 หน้าต่าง Count Form

2) คลิกเมนู File, Open เพื่อเปิดไฟล์ .txt ที่ต้องการ ดังภาพประกอบ ข-3



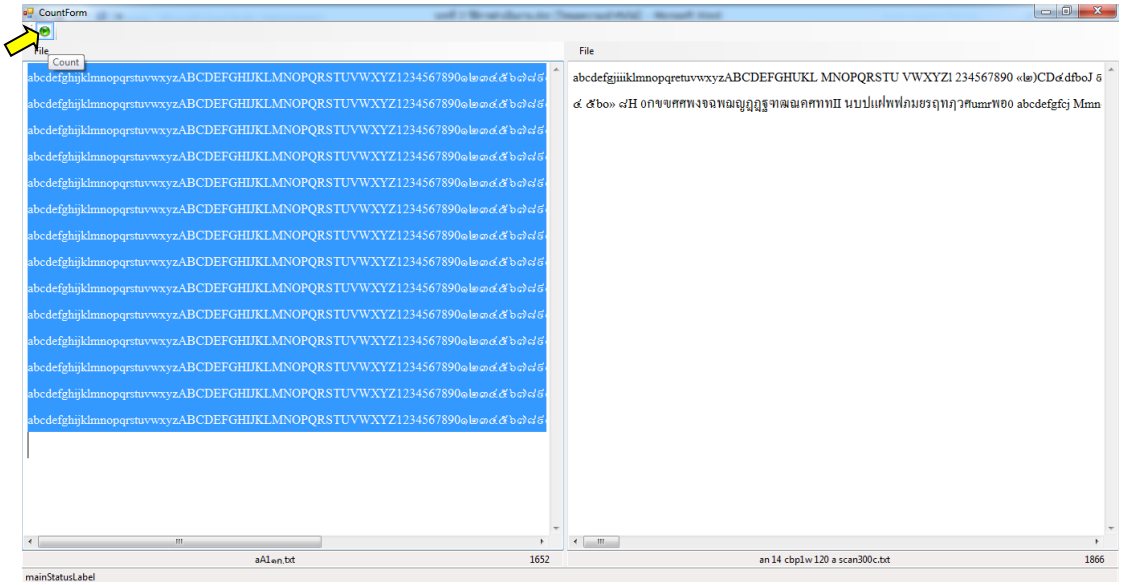
ภาพประกอบ ข-3 เมนู Open เปิดไฟล์ .txt ที่ต้องการ

3) คลิกเลือกไฟล์ที่ต้องการ แล้วคลิกปุ่ม Open ดังภาพประกอบ ข-4



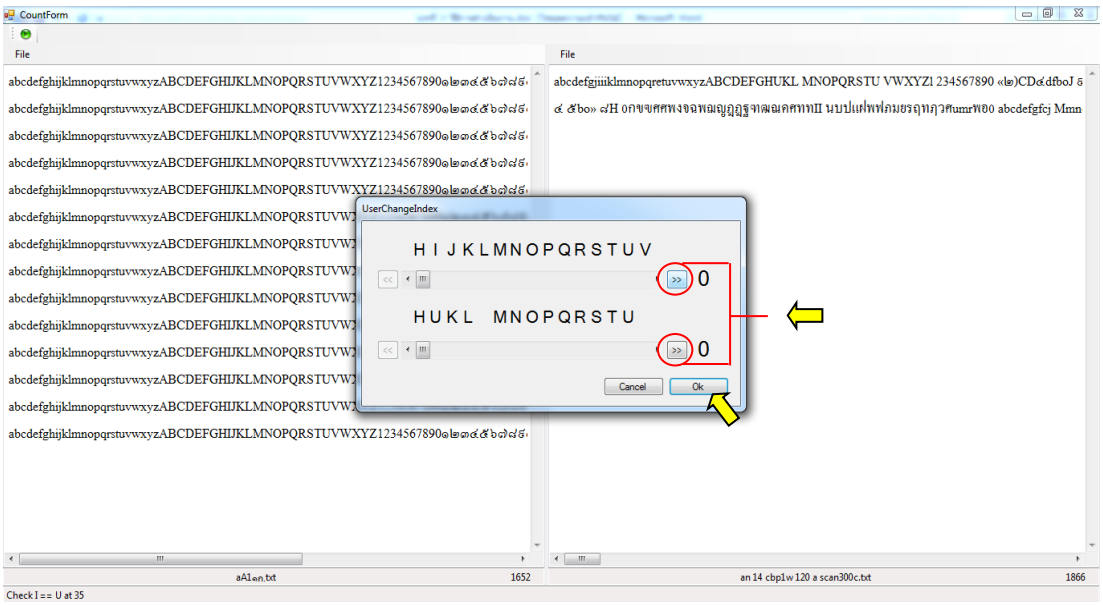
ภาพประกอบ ข-4 เปิดไฟล์ .txt ที่ต้องการ

4) คลิกปุ่ม Count เพื่อเริ่มนับคำถูก/ผิด ดังภาพประกอบ ข-5



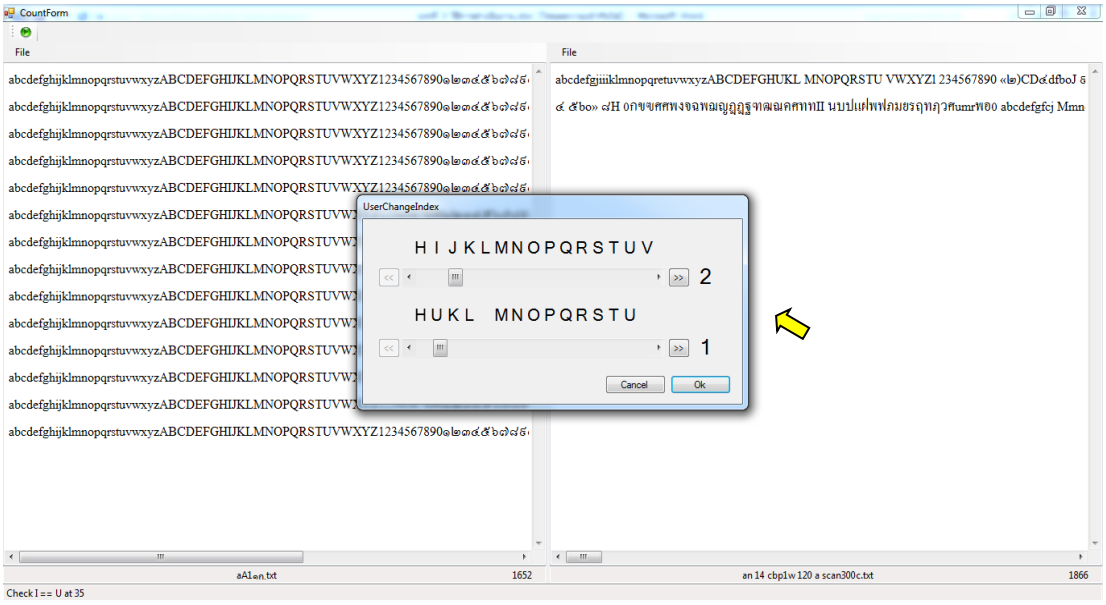
ภาพประกอบ ข-5 การนับคำถูก/ผิด

โปรแกรมจะนับตัวอักษรถูก/ผิด อัตโนมัติ ตัวอักษรที่ผิดซึ่งสามารถเกิดจากกรณีซ้ำ ขาดหาย เกิน หรือช่องว่าง กรณีนี้ตัวอักษรตัว I และ J อยู่ติดกัน แปลงไฟล์ออกมาเป็นตัว U แสดงหน้าต่าง User Change Index ให้เลื่อนกดเริ่มเปรียบเทียบตัวอักษรตัวถัดไป ดังภาพประกอบ ข-6



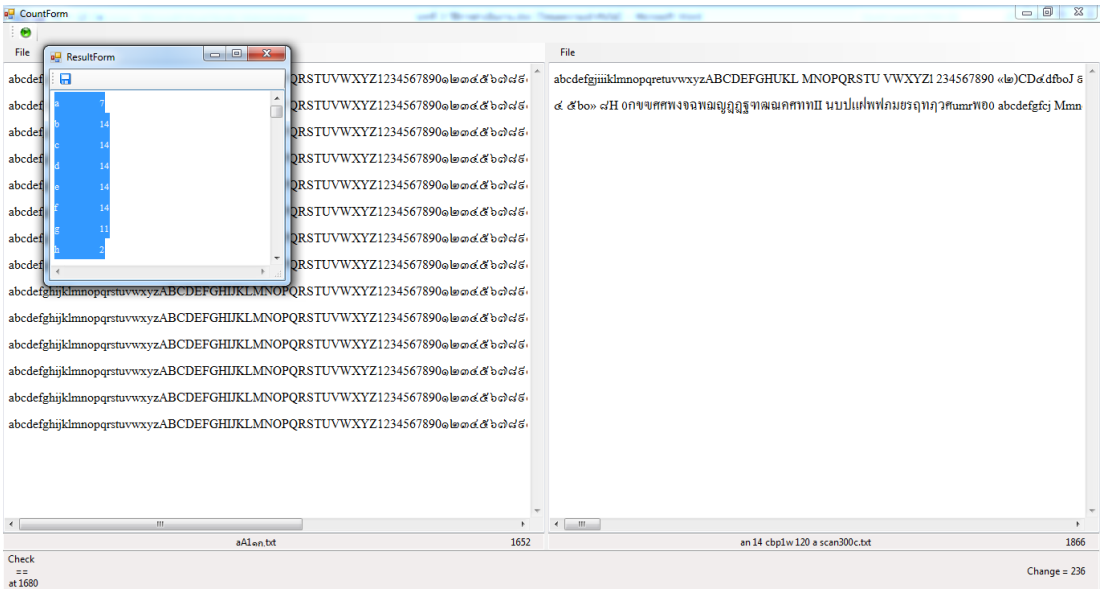
ภาพประกอบ ข-6 การเลื่อนนับตัวอักษรถูก/ผิด

การนับตัวอักษรถูก/ผิด ตัวนับจะแสดงค่าต้นแบบข้างบนกับค่าเปรียบเทียบข้างล่าง ถ้าไม่ตรงกัน เราสามารถขยับให้ตรงกับจุดเริ่มต้นใหม่ (ข้ามตัวอักษรที่ผิดพลาด) สามารถเลื่อนนับต่อไปได้ไม่จำกัด ในตัวอย่างนี้เลื่อนกดเริ่มเปรียบเทียบตัวอักษรตัวถัดไปคือตัว K แล้วคลิกปุ่ม OK โปรแกรมก็จะนับคำถูก/ผิด อัตโนมัติต่อไป ดังภาพประกอบ ข-7



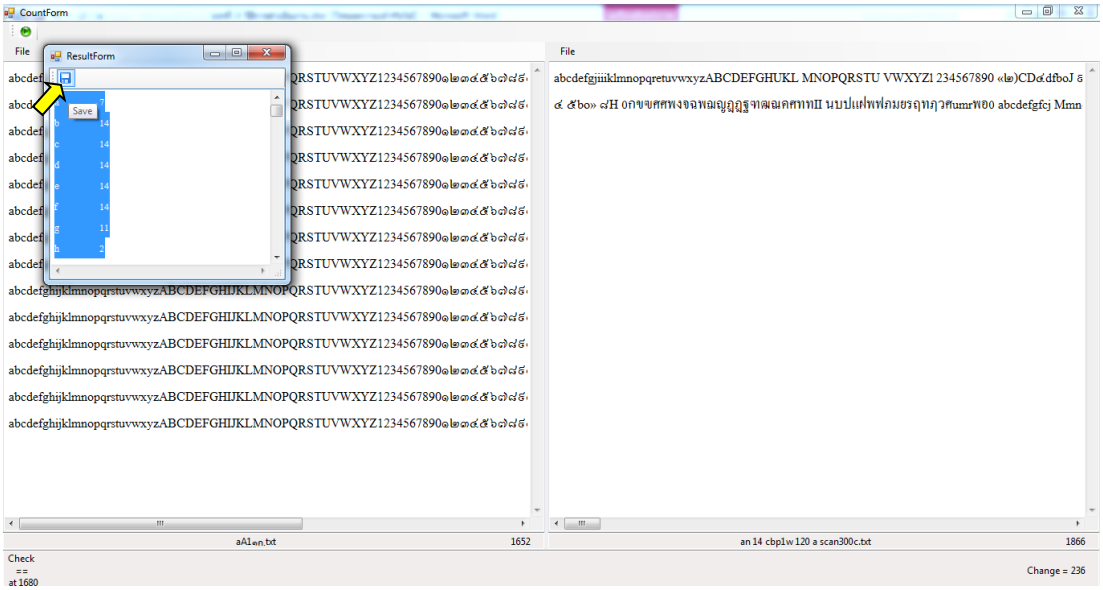
ภาพประกอบ ข-7 การเปรียบเทียบตัวอักษร

เมื่อนับคำถูก/ผิดครบทั้งหมด ก็จะปรากฏหน้าต่าง Result Form แสดงผลการนับตัวอักษรที่ถูก ดังภาพประกอบ ข-8

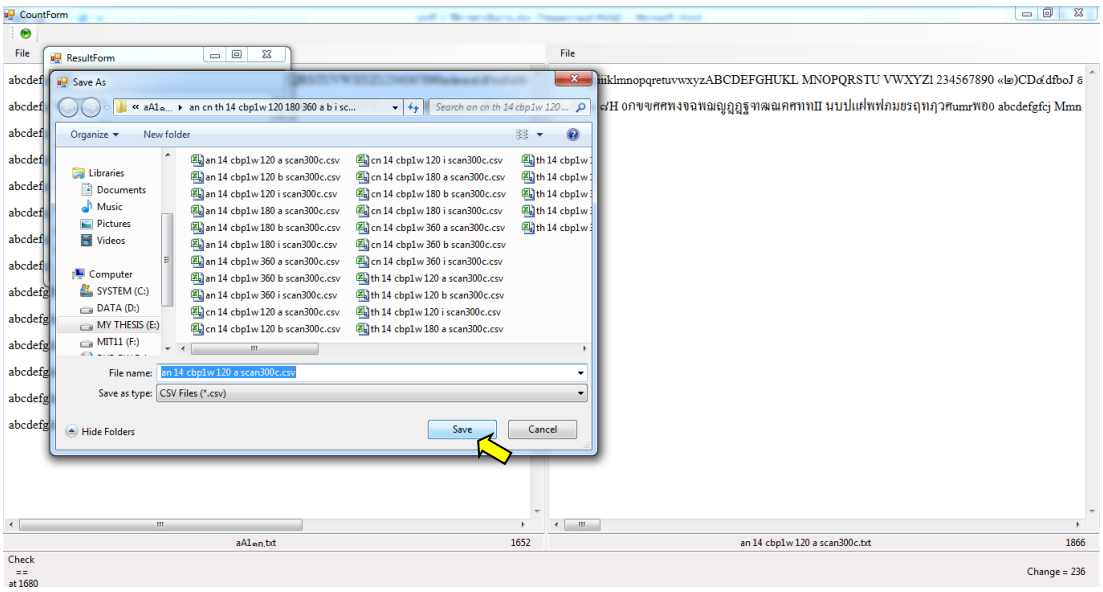


ภาพประกอบ ข-8 การแสดงผลการนับตัวอักษรที่ถูก

5) คลิกปุ่ม Save บันทึกการนับตัวอักษรที่ถูกเป็นไฟล์ .csv ดังภาพประกอบ ข-9 และ ภาพประกอบ ข-10



ภาพประกอบ ข-9 ปุ่มบันทึกผลการนับตัวอักษรที่ถูก



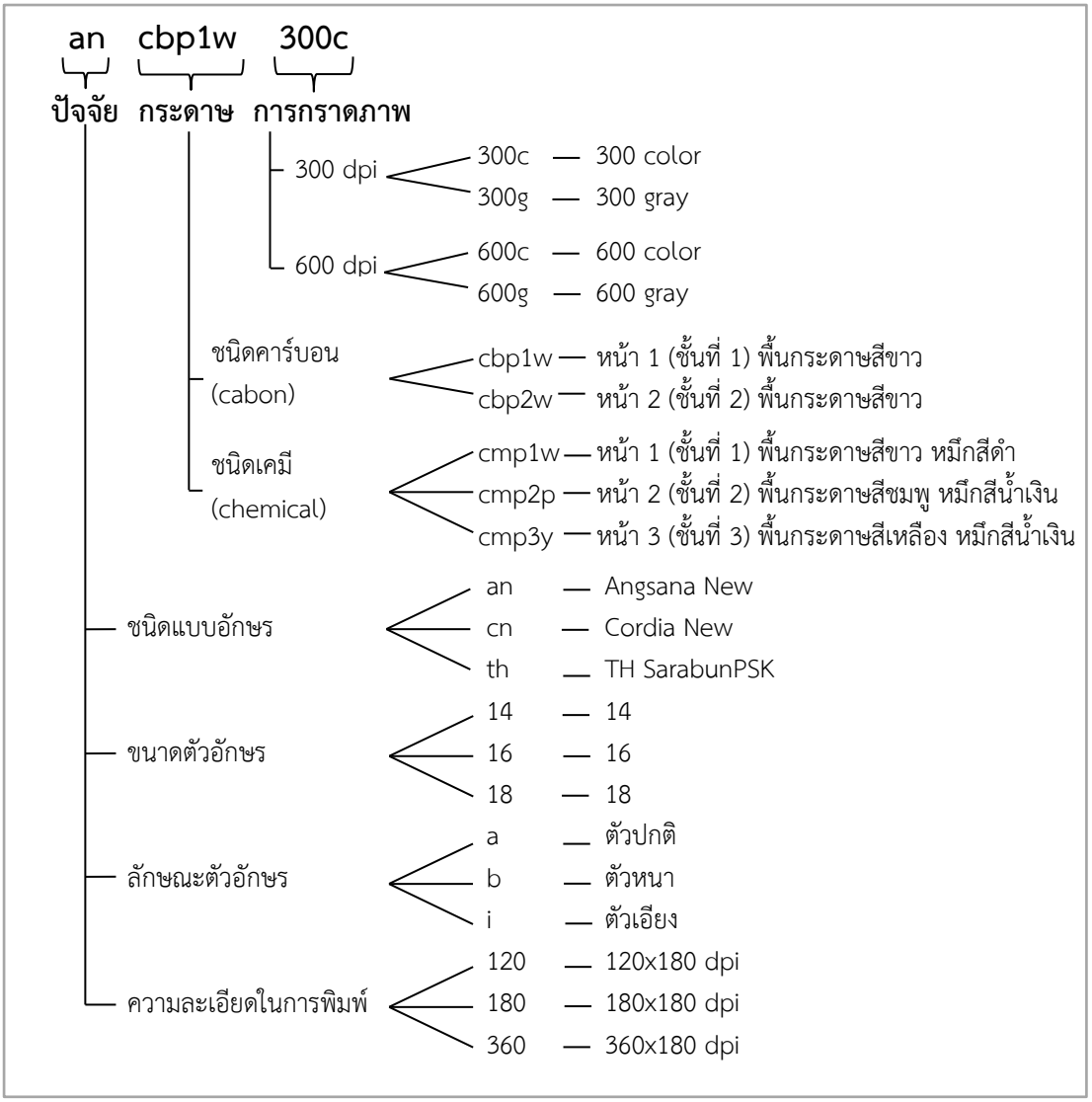
ภาพประกอบ ข-10 บันทึกการวัดผลการรู้จำ

ภาคผนวก ค

เปรียบเทียบผลการรู้จำ

ในส่วนนี้จะกล่าวถึงผลการรู้จำของเอกสารชนิดคาร์บอน และชนิดเคมีแยกตามปัจจัยชนิดแบบอักษร ขนาดตัวอักษร ลักษณะตัวอักษร ความละเอียดในการพิมพ์ และความละเอียดและชนิดในการกราดภาพ

มีวิธีการตั้งชื่อไฟล์เพื่อให้สะดวกต่อการตั้งชื่อ การจัดเก็บ ให้เข้าใจง่ายด้วยการตั้งรูปแบบรหัสชื่อไฟล์ให้มีความหมายตรงกับลักษณะไฟล์ แบ่งเป็น 3 กลุ่ม เรียงตามลำดับดังภาพประกอบ ค-1



ภาพประกอบ ค-1 การตั้งชื่อไฟล์เอกสาร

ตัวอย่างรหัส

an cbp1w 300c คือไฟล์เอกสาร ชนิดแบบอักษร Angsana New ที่พิมพ์ด้วยกระดาษชนิดคาร์บอน ชั้นที่ 1 พื้นกระดาษสีขาว ความละเอียดและชนิดในการกราดภาพ 300 dpi ชนิด Color ทั้งขนาดตัวอักษร 14 16 และ 18 ลักษณะตัวอักษรตัวปกติ ตัวหนา และตัวเอียง ความละเอียดในการพิมพ์ 180 x 120 dpi 180 x 180 dpi และ 180 x 360 dpi

1. เปรียบเทียบผลการรู้จำเอกสารชนิดคาร์บอนและชนิดเคมี

การเปรียบเทียบผลการรู้จำจะมีเอกสารกระดาษชนิดคาร์บอนและชนิดเคมี โดยเปรียบเทียบตามชนิดแบบอักษร ขนาดตัวอักษร ลักษณะตัวอักษร ความละเอียดในการพิมพ์ และความละเอียดและชนิดในการกราดภาพ ดังตาราง ค-1 ถึง ตาราง ค-10

ตาราง ค-1 ผลการรู้จำเอกสารชนิดคาร์บอน ตามชนิดแบบอักษร

ไฟล์เอกสาร	อักษรไทย				อักษรอังกฤษ				เลขไทย				เลขอารบิค			
	ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด	
	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%
an cbp1w 300c	10,873	62.53	6,515	37.47	18,345	93.33	1,311	6.67	2,930	77.51	850	22.49	3,665	96.96	115	3.04
an cbp1w 300g	11,119	63.95	6,269	36.05	18,259	92.89	1,397	7.11	2,914	77.09	866	22.91	3,682	97.41	98	2.59
an cbp1w 600c	12,107	69.63	5,281	30.37	18,480	94.02	1,176	5.98	3,123	82.62	657	17.38	3,710	98.15	70	1.85
an cbp1w 600g	12,132	69.77	5,256	30.23	18,414	93.68	1,242	6.32	3,162	83.65	618	16.35	3,717	98.33	63	1.67
cn cbp1w 300c	12,591	72.41	4,797	27.59	18,993	96.63	663	3.37	2,717	71.88	1,063	28.12	3,681	97.38	99	2.62
cn cbp1w 300g	12,892	74.14	4,496	25.86	18,903	96.17	753	3.83	2,768	73.23	1,012	26.77	3,697	97.80	83	2.20
cn cbp1w 600c	13,623	78.35	3,765	21.65	19,051	96.92	605	3.08	2,822	74.66	958	25.34	3,679	97.33	101	2.67
cn cbp1w 600g	13,582	78.11	3,806	21.89	19,039	96.86	617	3.14	2,727	72.14	1,053	27.86	3,682	97.41	98	2.59
th cbp1w 300c	11,832	68.05	5,556	31.95	17,611	89.60	2,045	10.40	2,979	78.81	801	21.19	3,720	98.41	60	1.59
th cbp1w 300g	11,834	68.05	5,554	31.94	17,431	88.68	2,225	11.32	3,063	81.03	717	18.97	3,702	97.94	78	2.06
th cbp1w 600c	12,859	73.95	4,529	26.05	18,116	92.17	1,540	7.83	3,075	81.35	705	18.65	3,724	98.52	56	1.48
th cbp1w 600g	8,458	48.64	8,930	51.36	11,836	60.22	7,820	39.78	2,064	54.60	1,716	45.40	2,484	65.71	1,296	34.29
an cbp2w 300c	8,291	47.68	9,097	52.32	17,527	89.17	2,129	10.83	2,493	65.95	1,287	34.05	3,557	94.10	223	5.90
an cbp2w 300g	8,712	50.10	8,676	49.90	17,459	88.82	2,197	11.18	2,516	66.56	1,264	33.44	3,611	95.53	169	4.47
an cbp2w 600c	9,926	57.09	7,462	42.91	17,730	90.20	1,926	9.80	2,804	74.18	976	25.82	3,657	96.75	123	3.25
an cbp2w 600g	9,884	56.84	7,504	43.16	17,756	90.33	1,900	9.67	2,672	70.69	1,108	29.31	3,647	96.48	133	3.52
cn cbp2w 300c	10,425	59.96	6,963	40.04	18,384	93.53	1,272	6.47	2,303	60.93	1,477	39.07	3,623	95.85	157	4.15
cn cbp2w 300g	10,496	60.36	6,892	39.64	18,155	92.36	1,501	7.64	2,318	61.32	1,462	38.68	3,620	95.77	160	4.23
cn cbp2w 600c	11,097	63.82	6,291	36.18	18,541	94.33	1,115	5.67	2,368	62.65	1,412	37.35	3,684	97.46	96	2.54
cn cbp2w 600g	11,429	65.73	5,959	34.27	18,527	94.26	1,129	5.74	2,309	61.08	1,471	38.92	3,687	97.54	93	2.46
th cbp2w 300c	9,060	52.10	8,328	47.90	15,864	80.71	3,792	19.29	2,399	63.47	1,381	36.53	3,521	93.15	259	6.85
th cbp2w 300g	9,586	55.13	7,802	44.87	16,366	83.26	3,290	16.74	2,610	69.05	1,170	30.95	3,686	97.51	94	2.49
th cbp2w 600c	10,017	57.61	7,371	42.39	16,421	83.54	3,235	16.46	2,630	69.58	1,150	30.42	3,731	98.70	49	1.30
th cbp2w 600g	10,356	59.56	7,032	40.44	16,639	84.65	3,017	15.35	2,706	71.59	1,074	28.41	3,698	97.83	82	2.17

ตาราง ค-2 ผลการรู้จำเอกสารชนิดเคมี ตามชนิดแบบอักษร

ไฟล์เอกสาร	อักษรไทย				อักษรอังกฤษ				เลขไทย				เลขอารบิค			
	ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด	
	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%
an cmp1w 300c	11,475	65.99	5,913	34.01	18,411	93.67	1,245	6.33	3,063	81.03	717	18.97	3,683	97.43	97	2.57
an cmp1w 300g	11,539	66.36	5,849	33.64	18,382	93.52	1,274	6.48	3,043	80.50	737	19.50	3,685	97.49	95	2.51
an cbp1w 600c	12,677	72.91	4,711	27.09	18,551	94.38	1,105	5.62	3,273	86.59	507	13.41	3,704	97.99	76	2.01
an cmp1w 600g	12,592	72.42	4,796	27.58	18,494	94.09	1,162	5.91	3,222	85.24	558	14.76	3,707	97.99	73	1.93
cn cmp1w 300c	13,022	74.89	4,366	25.11	18,453	93.88	1,203	6.12	2,859	75.63	921	24.37	3,577	94.63	203	5.37
cn cmp1w 300g	13,523	77.77	3,865	22.23	18,920	96.26	736	3.74	2,862	75.71	918	24.29	3,685	97.49	95	2.51
cn cmp1w 600c	14,166	81.47	3,222	18.53	19,155	97.45	501	2.55	2,773	73.36	1,007	26.64	3,672	97.14	108	2.86
cn cmp1w 600g	14,307	82.28	3,081	17.72	19,130	97.32	526	2.68	2,810	74.34	970	25.66	3,692	97.67	88	2.33
th cmp1w 300c	12,129	69.76	5,259	30.24	17,501	89.04	2,155	10.96	3,067	81.14	713	18.86	3,737	98.86	43	1.14
th cmp1w 300g	12,251	70.46	5,137	29.54	17,062	86.80	2,594	13.20	3,015	79.76	765	20.24	3,592	95.03	188	4.97
th cmp1w 600c	13,383	76.97	4,005	23.03	18,274	92.97	1,382	7.03	3,184	84.23	596	15.77	3,715	98.28	65	1.72
th cmp1w 600g	14,307	82.28	3,081	17.72	18,419	93.71	1,237	6.29	3,202	84.71	578	15.29	3,721	98.44	59	1.56
an cmp2p 300c	11,241	64.65	6,147	35.35	18,145	92.31	1,511	7.69	2,920	77.25	860	22.75	3,691	97.65	89	2.35
an cmp2p 300g	10,989	63.20	6,399	36.80	17,304	88.03	2,352	11.97	2,771	73.31	1,009	26.69	3,482	92.12	298	7.88
an cmp2p 600c	11,938	68.66	5,450	31.34	18,423	93.73	1,233	6.27	2,978	78.78	802	21.22	3,689	97.59	91	2.41
an cmp2p 600g	12,168	69.98	5,220	30.02	18,452	93.87	1,204	6.13	2,996	79.26	784	20.74	3,692	97.67	88	2.33
cn cmp2p 300c	13,193	75.36	4,285	24.64	18,797	96.96	597	3.04	2,720	72.67	1,033	27.33	3,665	96.96	115	3.04
cn cmp2p 300g	13,103	75.87	4,195	24.13	19,059	95.63	859	4.37	2,747	71.96	1,060	28.04	3,665	96.96	115	3.04
cn cmp2p 600c	13,655	78.53	3,733	21.47	19,050	96.92	606	3.08	2,770	73.28	1,010	26.72	3,665	96.96	115	3.04
cn cmp2p 600g	13,413	77.14	3,975	22.86	18,924	96.28	732	3.72	2,809	74.31	971	25.69	3,678	97.30	102	2.70
th cmp2p 300c	12,313	70.81	5,075	29.19	17,947	91.31	1,709	8.69	3,030	80.16	750	19.84	3,719	98.39	61	1.61
th cmp2p 300g	12,151	69.88	5,237	30.12	16,790	85.42	2,866	14.58	2,897	76.64	883	23.36	3,557	94.10	223	5.90
th cmp2p 600c	13,005	74.79	4,383	25.21	17,938	91.26	1,718	8.74	3,068	81.16	712	18.84	3,719	98.39	61	1.61
th cmp2p 600g	12,887	74.11	4,501	25.89	17,906	91.10	1,750	8.90	3,010	79.63	770	20.37	3,714	98.25	66	1.75
an cmp3y 300c	10,400	59.81	6,988	40.19	17,858	90.85	1,798	9.15	2,683	70.98	1,097	29.02	3,638	96.24	142	3.76
an cmp3y 300g	10,073	57.93	7,315	42.07	17,341	88.22	2,315	11.78	2,657	70.29	1,123	29.71	3,618	95.71	162	4.29
an cmp3y 600c	10,853	62.42	6,535	37.58	17,950	91.32	1,706	8.68	2,583	68.33	1,197	31.67	3,688	97.57	92	2.43
an cmp3y 600g	10,847	62.38	6,541	37.62	17,821	90.66	1,835	9.34	2,710	71.69	1,070	28.31	3,673	97.17	107	2.83
cn cmp3y 300c	11,884	68.35	5,504	31.65	18,449	93.86	1,207	6.14	2,502	66.19	1,278	33.81	3,679	97.33	101	2.67
cn cmp3y 300g	11,587	66.64	5,801	33.36	18,560	94.42	1,096	5.58	2,425	64.15	1,355	35.85	3,662	96.88	118	3.12
cn cmp3y 600c	12,895	74.16	4,493	25.84	18,545	94.35	1,111	5.65	2,664	70.48	1,116	29.52	3,685	97.49	95	2.51
cn cmp3y 600g	12,726	73.19	4,662	26.81	18,471	93.97	1,185	6.03	2,650	70.11	1,130	29.89	3,685	97.49	95	2.51
th cmp3y 300c	10,945	62.95	6,443	37.05	17,749	90.30	1,907	9.70	2,713	71.77	1,067	28.23	3,664	96.93	116	3.07
th cmp3y 300g	11,378	65.44	6,010	34.56	17,148	87.24	2,508	12.76	2,747	72.67	1,033	27.33	3,703	97.96	77	2.04
th cmp3y 600c	11,619	66.82	5,769	33.18	17,904	91.09	1,752	8.91	2,828	74.81	952	25.19	3,657	96.75	123	3.25
th cmp3y 600g	11,608	66.76	5,780	33.24	17,897	91.05	1,759	8.95	2,849	75.37	931	24.63	3,657	96.75	123	3.25

ตาราง ค-3 ผลการรู้จำเอกสารชนิดคาร์บอน ตามขนาดตัวอักษร

ไฟล์เอกสาร	อักษรไทย				อักษรอังกฤษ				เลขไทย				เลขอารบิค			
	ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด	
	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%
14 cbp1w 300c	10,153	58.39	7,235	41.61	18,001	91.58	1,655	8.42	2,556	67.62	1,224	32.38	3,664	96.93	116	3.07
14 cbp1w 300g	10,299	59.23	7,089	40.77	17,764	90.37	1,892	9.63	2,586	68.41	1,194	31.59	3,696	97.78	84	2.22
14 cbp1w 600c	11,314	65.07	6,074	34.93	18,285	93.03	1,371	6.97	2,742	72.54	1,038	27.46	3,700	97.88	80	2.12
14 cbp1w 600g	11,337	65.20	6,051	34.80	18,253	92.86	1,403	7.14	2,762	73.07	1,018	26.93	3,712	98.20	68	1.80
16 cbp1w 300c	12,342	70.98	5,046	29.02	18,321	93.21	1,335	6.79	3,040	80.42	740	19.58	3,703	97.96	77	2.04
16 cbp1w 300g	12,362	71.10	5,026	28.90	18,206	92.62	1,450	7.38	3,074	81.32	706	18.68	3,690	97.62	90	2.38
16 cbp1w 600c	13,449	77.35	3,939	22.65	18,620	94.73	1,036	5.27	3,126	82.70	654	17.30	3,717	98.33	63	1.67
16 cbp1w 600g	13,578	78.09	3,810	21.91	18,504	94.14	1,152	5.86	3,076	81.38	704	18.62	3,720	98.41	60	1.59
18 cbp1w 300c	12,801	73.62	4,587	26.38	18,627	94.76	1,029	5.24	3,030	80.16	750	19.84	3,699	97.86	81	2.14
18 cbp1w 300g	13,184	75.82	4,204	24.18	18,623	94.74	1,033	5.26	3,085	81.61	695	18.39	3,695	97.75	85	2.25
18 cbp1w 600c	13,826	79.51	3,562	20.49	18,742	95.35	914	4.65	3,152	83.39	628	16.61	3,696	97.78	84	2.22
18 cbp1w 600g	9,257	53.24	8,131	46.76	12,532	63.76	7,124	36.24	2,115	55.95	1,665	44.05	2,451	64.84	1,329	35.16
14 cbp2w 300c	7,346	42.25	10,042	57.75	16,978	86.38	2,678	13.62	1,923	50.87	1,857	49.13	3,516	93.02	264	6.98
14 cbp2w 300g	7,584	43.62	9,804	56.38	16,875	85.85	2,781	14.15	2,161	57.17	1,619	42.83	3,587	94.89	193	5.11
14 cbp2w 600c	8,020	46.12	9,368	53.88	17,038	86.68	2,618	13.32	2,162	57.20	1,618	42.80	3,681	97.38	99	2.62
14 cbp2w 600g	8,305	47.76	9,083	52.24	17,238	87.70	2,418	12.30	2,134	56.46	1,646	43.54	3,659	96.80	121	3.20
16 cbp2w 300c	9,680	55.67	7,708	44.33	17,429	88.67	2,227	11.33	2,645	69.97	1,135	30.03	3,648	96.51	132	3.49
16 cbp2w 300g	10,065	57.88	7,323	42.12	17,319	88.11	2,337	11.89	2,593	68.60	1,187	31.40	3,642	96.35	138	3.65
16 cbp2w 600c	11,047	63.53	6,341	36.47	17,603	89.56	2,053	10.44	2,801	74.10	979	25.90	3,686	97.51	94	2.49
16 cbp2w 600g	11,245	64.67	6,143	35.33	17,648	89.78	2,008	10.22	2,803	74.15	977	25.85	3,688	97.57	92	2.43
18 cbp2w 300c	10,750	61.82	6,638	38.18	17,368	88.36	2,288	11.64	2,627	69.50	1,153	30.50	3,537	93.57	243	6.43
18 cbp2w 300g	11,145	64.10	6,243	35.90	17,786	90.49	1,870	9.51	2,690	71.16	1,090	28.84	3,688	97.57	92	2.43
18 cbp2w 600c	11,973	68.86	5,415	31.14	18,051	91.83	1,605	8.17	2,839	75.11	941	24.89	3,705	98.02	75	1.98
18 cbp2w 600g	12,119	69.70	5,269	30.30	18,036	91.76	1,620	8.24	2,750	72.75	1,030	27.25	3,685	97.49	95	2.51

ตาราง ค-4 ผลการรู้จำเอกสารชนิดเคมี ตามขนาดตัวอักษร

ไฟล์เอกสาร	อักษรไทย				อักษรอังกฤษ				เลขไทย				เลขอารบิค			
	ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด	
	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%
14 cmp1w 300c	11,340	65.22	6,048	34.78	17,946	91.30	1,710	8.70	2,881	76.22	899	23.78	3,691	97.65	89	2.35
14 cmp1w 300g	11,270	64.81	6,118	35.19	18,034	91.75	1,622	8.25	2,792	73.86	988	26.14	3,674	97.20	106	2.80
14 cmp1w 600c	11,877	68.31	5,511	31.69	18,456	93.89	1,200	6.11	2,934	77.62	846	22.38	3,701	97.91	79	2.09
14 cmp1w 600g	12,140	69.82	5,248	30.18	18,382	93.52	1,274	6.48	2,922	77.30	858	22.70	3,715	98.28	65	1.72
16 cmp1w 300c	12,695	73.01	4,693	26.99	18,357	93.39	1,299	6.61	3,066	81.11	714	18.89	3,715	98.28	65	1.72
16 cmp1w 300g	13,054	75.07	4,334	24.93	18,363	93.42	1,293	6.58	3,096	81.90	684	18.10	3,724	98.52	56	1.48
16 cmp1w 600c	13,992	80.47	3,396	19.53	18,739	95.33	917	4.67	3,078	81.43	702	18.57	3,715	98.28	65	1.72
16 cmp1w 600g	14,120	81.21	3,268	18.79	18,774	95.51	882	4.49	3,081	81.51	699	18.49	3,719	98.39	61	1.61
18 cmp1w 300c	12,591	72.41	4,797	27.59	18,062	91.89	1,594	8.11	3,042	80.48	738	19.52	3,591	95.00	189	5.00
18 cmp1w 300g	12,989	74.70	4,399	25.30	17,967	91.41	1,689	8.59	3,032	80.21	748	19.79	3,564	94.29	216	5.71
18 cmp1w 600c	14,357	82.57	3,031	17.43	18,785	95.57	871	4.43	3,218	85.13	562	14.87	3,675	97.22	105	2.78
18 cmp1w 600g	14,324	82.38	3,064	17.62	18,887	96.09	769	3.91	3,231	85.48	549	14.52	3,686	97.51	94	2.49
14 cmp2p 300c	10,819	62.22	6,569	37.78	18,068	91.92	1,588	8.08	2,669	70.61	1,111	29.39	3,678	97.30	102	2.70
14 cmp2p 300g	10,934	62.88	6,454	37.12	17,583	89.45	2,073	10.55	2,629	69.55	1,151	30.45	3,579	94.68	201	5.32
14 cmp2p 600c	11,507	66.18	5,881	33.82	18,136	92.27	1,520	7.73	2,751	72.78	1,029	27.22	3,708	98.10	72	1.90
14 cmp2p 600g	11,228	64.57	6,160	35.43	18,110	92.13	1,546	7.87	2,694	71.27	1,086	28.73	3,706	98.04	74	1.96
16 cmp2p 300c	12,832	73.80	4,556	26.20	18,465	93.94	1,191	6.06	2,986	78.99	794	21.01	3,702	97.94	78	2.06
16 cmp2p 300g	12,736	73.25	4,652	26.75	18,067	91.92	1,589	8.08	3,026	80.05	754	19.95	3,699	97.86	81	2.14
16 cmp2p 600c	13,337	76.70	4,051	23.30	18,476	94.00	1,180	6.00	3,001	79.39	779	20.61	3,703	97.96	77	2.04
16 cmp2p 600g	13,340	76.72	4,048	23.28	18,440	93.81	1,216	6.19	2,991	79.13	789	20.87	3,711	98.17	69	1.83
18 cmp2p 300c	13,006	74.80	4,382	25.20	18,618	94.72	1,038	5.28	3,042	80.48	738	19.52	3,695	97.75	85	2.25
18 cmp2p 300g	12,663	72.83	4,725	27.17	18,799	87.71	2,415	12.29	2,733	72.30	1,047	27.70	3,426	90.63	354	9.37
18 cmp2p 600c	13,754	79.10	3,634	20.90	17,241	95.64	857	4.36	3,064	81.06	716	18.94	3,662	96.88	118	3.12
18 cmp2p 600g	13,900	79.94	3,488	20.06	18,732	95.30	924	4.70	3,130	82.80	650	17.20	3,667	97.01	113	2.99
14 cmp3y 300c	9,313	53.56	8,075	46.44	17,351	88.27	2,305	11.73	2,216	58.62	1,564	41.38	3,613	95.58	167	4.42
14 cmp3y 300g	9,414	54.14	7,974	45.86	16,904	86.00	2,752	14.00	2,211	58.49	1,569	41.51	3,632	96.08	148	3.92
14 cmp3y 600c	10,090	58.03	7,298	41.97	17,642	89.75	2,014	10.25	2,491	65.90	1,289	34.10	3,681	97.38	99	2.62
14 cmp3y 600g	9,922	57.06	7,466	42.94	17,551	89.29	2,105	10.71	2,440	64.55	1,340	35.45	3,651	96.59	129	3.41
16 cmp3y 300c	11,415	65.65	5,973	34.35	18,346	93.34	1,310	6.66	2,818	74.55	962	25.45	3,693	97.70	87	2.30
16 cmp3y 300g	11,674	67.14	5,714	32.86	17,989	91.52	1,667	8.48	2,802	74.13	978	25.87	3,674	97.20	106	2.80
16 cmp3y 600c	12,431	71.49	4,957	28.51	18,348	93.35	1,308	6.65	2,782	73.60	998	26.40	3,689	97.59	91	2.41
16 cmp3y 600g	12,384	71.22	5,004	28.78	18,384	93.53	1,272	6.47	2,865	75.79	915	24.21	3,693	97.70	87	2.30
18 cmp3y 300c	12,501	71.89	4,887	28.11	18,359	93.40	1,297	6.60	1,863	49.10	666	17.57	3,675	97.22	105	2.78
18 cmp3y 300g	11,950	68.73	5,438	31.27	18,156	92.37	1,500	7.63	2,816	74.50	964	25.50	3,677	97.28	103	2.72
18 cmp3y 600c	12,846	73.88	4,542	26.12	18,409	93.66	1,247	6.34	1,815	47.83	714	18.84	3,660	96.83	120	3.17
18 cmp3y 600g	12,875	74.05	4,513	25.95	18,254	92.87	1,402	7.13	2,904	76.83	876	23.17	3,671	97.12	109	2.88

ตาราง ค-5 ผลการรู้จำเอกสารชนิดคาร์บอน ตามลักษณะตัวอักษร

ไฟล์เอกสาร	อักษรไทย				อักษรอังกฤษ				เลขไทย				เลขอารบิค			
	ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด	
	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%
a cbp1w 300c	12,584	72.37	4,804	27.63	18,509	94.16	1,147	5.84	2,763	73.10	1,017	26.90	3,688	97.57	92	2.43
a cbp1w 300g	12,736	73.25	4,652	26.75	18,237	92.78	1,419	7.22	2,750	72.75	1,030	27.25	3,693	97.70	87	2.30
a cbp1w 600c	13,482	77.54	3,906	22.46	18,737	95.32	919	4.68	2,872	75.98	908	24.02	3,703	97.96	77	2.04
a cbp1w 600g	11,938	68.66	5,450	31.34	16,553	84.21	3,103	15.79	2,414	63.86	1,366	36.14	3,302	87.35	478	12.65
b cbp1w 300c	12,529	72.06	4,859	27.94	18,739	95.33	917	4.67	2,949	78.02	831	21.98	3,715	98.28	65	1.72
b cbp1w 300g	12,526	72.04	4,862	27.96	18,780	95.54	876	4.46	2,910	76.98	870	23.02	3,735	98.81	45	1.19
b cbp1w 600c	13,265	76.29	4,123	23.71	18,796	95.62	860	4.38	2,991	79.13	789	20.87	3,736	98.84	44	1.16
b cbp1w 600g	11,559	66.48	5,829	33.52	16,733	85.13	2,923	14.87	2,716	71.85	1,064	28.15	3,322	87.88	458	12.12
i cbp1w 300c	10,183	58.56	7,205	41.44	17,701	90.05	1,955	9.95	2,914	77.09	866	22.91	3,663	96.90	117	3.10
i cbp1w 300g	10,583	60.86	6,805	39.14	17,576	89.42	2,080	10.58	3,085	81.61	695	18.39	3,653	96.64	127	3.36
i cbp1w 600c	11,842	68.10	5,546	31.90	18,114	92.16	1,542	7.84	3,157	83.52	623	16.48	3,674	97.20	106	2.80
i cbp1w 600g	10,675	61.39	6,713	38.61	16,003	81.42	3,653	18.58	2,823	74.68	957	25.32	3,259	86.22	521	13.78
a cbp2w 300c	10,098	58.07	7,290	41.93	17,422	88.63	2,234	11.37	2,248	59.47	1,532	40.53	3,585	94.84	195	5.16
a cbp2w 300g	10,604	60.98	6,784	39.02	17,422	88.63	2,234	11.37	2,346	62.06	1,434	37.94	3,658	96.77	122	3.23
a cbp2w 600c	11,225	64.56	6,163	35.44	17,565	89.36	2,091	10.64	2,472	65.40	1,308	34.60	3,702	97.94	78	2.06
a cbp2w 600g	11,290	68.66	5,450	31.34	17,549	84.21	3,103	15.79	2,454	63.86	1,366	36.14	3,700	87.35	478	12.65
b cbp2w 300c	8,929	51.35	8,459	48.65	17,302	88.02	2,354	11.98	2,311	61.14	1,469	38.86	3,525	93.25	255	6.75
b cbp2w 300g	9,436	54.27	7,952	45.73	17,791	90.51	1,865	9.49	2,400	63.49	1,380	36.51	3,650	96.56	130	3.44
b cbp2w 600c	10,094	58.05	7,294	41.95	18,212	92.65	1,444	7.35	2,612	69.10	1,168	30.90	3,693	97.70	87	2.30
b cbp2w 600g	10,548	60.66	6,840	39.34	18,399	93.61	1,257	6.39	2,532	66.98	1,248	33.02	3,689	97.59	91	2.41
i cbp2w 300c	8,749	50.32	8,639	49.68	17,051	86.75	2,605	13.25	2,636	69.74	1,144	30.26	3,591	95.00	189	5.00
i cbp2w 300g	8,754	50.35	8,634	49.65	16,767	85.30	2,889	14.70	2,698	71.38	1,082	28.62	3,609	95.48	171	4.52
i cbp2w 600c	9,721	68.10	5,546	31.90	16,915	92.16	1,542	7.84	2,718	83.52	623	16.48	3,677	97.20	106	2.80
i cbp2w 600g	10,675	61.39	6,713	38.61	16,003	81.42	3,653	18.58	2,823	74.68	957	25.32	3,259	86.22	521	13.78

ตาราง ค-6 ผลการรู้จำเอกสารชนิดเคมี ตามลักษณะตัวอักษร

ไฟล์เอกสาร	อักษรไทย				อักษรอังกฤษ				เลขไทย				เลขอารบิค			
	ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด	
	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%
a cmp1w 300c	12,634	72.66	4,754	27.34	17,879	90.96	1,777	9.04	2,805	74.21	975	25.79	3,600	95.24	180	4.76
a cmp1w 300g	12,940	74.42	4,448	25.58	17,772	90.42	1,884	9.58	2,749	72.72	1,031	27.28	3,578	94.66	202	5.34
a cmp1w 600c	14,042	80.76	3,346	19.24	18,762	95.45	894	4.55	2,929	77.49	851	22.51	3,693	97.70	87	2.30
a cmp1w 600g	14,324	82.38	3,064	17.62	18,829	95.79	827	4.21	2,946	77.94	834	22.06	3,698	97.83	82	2.17
b cmp1w 300c	12,951	74.48	4,437	25.52	18,765	95.47	891	4.53	3,038	80.37	742	19.63	3,720	98.41	60	1.59
b cmp1w 300g	13,102	75.35	4,286	24.65	18,819	95.74	837	4.26	2,993	79.18	787	20.82	3,707	98.07	73	1.93
b cmp1w 600c	13,762	79.15	3,626	20.85	18,867	95.99	789	4.01	3,083	81.56	697	18.44	3,734	98.78	46	1.22
b cmp1w 600g	13,715	78.88	3,673	21.12	18,896	96.13	760	3.87	3,068	81.16	712	18.84	3,745	99.07	35	0.93
i cmp1w 300c	11,041	63.50	6,347	36.50	17,721	90.16	1,935	9.84	3,146	83.23	634	16.77	3,677	97.28	103	2.72
i cmp1w 300g	11,271	64.82	6,117	35.18	17,773	90.42	1,883	9.58	3,178	84.07	602	15.93	3,677	97.28	103	2.72
i cmp1w 600c	12,422	71.44	4,966	28.56	18,351	93.36	1,305	6.64	3,218	85.13	562	14.87	3,664	96.93	116	3.07
i cmp1w 600g	12,545	72.15	4,843	27.85	18,318	93.19	1,338	6.81	3,220	85.19	560	14.81	3,677	97.28	103	2.72
a cmp2p 300c	13,617	78.31	3,771	21.69	18,539	94.32	1,117	5.68	2,820	74.60	960	25.40	3,693	97.70	87	2.30
a cmp2p 300g	13,640	78.44	3,748	21.56	18,391	93.56	1,265	6.44	2,768	73.23	1,012	26.77	3,668	97.04	112	2.96
a cmp2p 600c	13,653	78.52	3,735	21.48	18,544	94.34	1,112	5.66	2,779	73.52	1,001	26.48	3,689	97.59	91	2.41
a cmp2p 600g	13,202	75.93	4,186	24.07	18,505	94.14	1,151	5.86	2,843	75.21	937	24.79	3,707	98.07	73	1.93
b cmp2p 300c	11,794	67.83	5,594	32.17	18,722	95.25	934	4.75	2,787	73.73	993	26.27	3,728	98.62	52	1.38
b cmp2p 300g	12,013	69.09	5,375	30.91	18,010	91.63	1,646	8.37	2,739	72.46	1,041	27.54	3,578	94.66	202	5.34
b cmp2p 600c	12,932	74.37	4,456	25.63	18,758	95.43	898	4.57	2,947	77.96	833	22.04	3,724	98.52	56	1.48
b cmp2p 600g	13,121	75.46	4,267	24.54	18,613	94.69	1,043	5.31	2,941	77.80	839	22.20	3,725	98.54	55	1.46
i cmp2p 300c	11,246	64.68	6,142	35.32	17,890	91.02	1,766	8.98	3,090	81.75	690	18.25	3,654	96.67	126	3.33
i cmp2p 300g	10,680	61.42	6,708	38.58	16,490	83.89	3,166	16.11	2,881	76.22	899	23.78	3,458	91.48	322	8.52
i cmp2p 600c	12,013	69.09	5,375	30.91	18,109	92.13	1,547	7.87	3,090	81.75	690	18.25	3,660	96.83	120	3.17
i cmp2p 600g	12,145	69.85	5,243	30.15	18,164	92.41	1,492	7.59	3,031	80.19	749	19.81	3,652	96.61	128	3.39
a cmp3y 300c	11,821	67.98	5,567	32.02	17,905	91.09	1,751	8.91	2,649	70.08	1,131	29.92	3,676	97.25	104	2.75
a cmp3y 300g	12,055	69.33	5,333	30.67	17,597	89.52	2,059	10.48	2,492	65.93	1,288	34.07	3,677	97.28	103	2.72
a cmp3y 600c	12,369	71.14	5,019	28.86	18,094	92.05	1,562	7.95	2,648	70.05	1,132	29.95	3,686	97.51	94	2.49
a cmp3y 600g	12,398	71.30	4,990	28.70	18,017	91.66	1,639	8.34	2,751	72.78	1,029	27.22	3,682	97.41	98	2.59
b cmp3y 300c	11,234	64.61	6,154	35.39	18,468	93.96	1,188	6.04	2,598	68.73	1,182	31.27	3,667	97.01	113	2.99
b cmp3y 300g	10,981	63.15	6,407	36.85	18,332	93.26	1,324	6.74	2,607	68.97	1,173	31.03	3,679	97.33	101	2.67
b cmp3y 600c	11,869	68.26	5,519	31.74	18,511	94.17	1,145	5.83	2,601	68.81	1,179	31.19	3,684	97.46	96	2.54
b cmp3y 600g	11,769	67.68	5,619	32.32	18,499	94.11	1,157	5.89	2,583	68.33	1,197	31.67	3,672	97.14	108	2.86
i cmp3y 300c	10,174	58.51	7,214	41.49	17,683	89.96	1,973	10.04	2,651	70.13	1,129	29.87	3,638	96.24	142	3.76
i cmp3y 300g	10,002	57.52	7,386	42.48	17,120	87.10	2,536	12.90	2,730	72.22	1,050	27.78	3,627	95.95	153	4.05
i cmp3y 600c	11,129	64.00	6,259	36.00	17,794	90.53	1,862	9.47	2,826	74.76	954	25.24	3,660	96.83	120	3.17
i cmp3y 600g	11,014	63.34	6,374	36.66	17,673	89.91	1,983	10.09	2,875	76.06	905	23.94	3,661	96.85	119	3.15

ตาราง ค-7 ผลการรู้จำเอกสารชนิดคาร์บอน ตามความละเอียดในการพิมพ์

ไฟล์เอกสาร	อักษรไทย				อักษรอังกฤษ				เลขไทย				เลขอารบิค			
	ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด	
	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%
120 cbp1w 300c	10,596	60.94	6,792	39.06	17,612	89.60	2,044	10.40	2,667	70.56	1,113	29.44	3,669	97.06	111	2.94
120 cbp1w 300g	10,909	62.74	6,479	37.26	17,633	89.71	2,023	10.29	2,704	71.53	1,076	28.47	3,677	97.28	103	2.72
120 cbp1w 600c	11,785	67.78	5,603	32.22	17,995	91.55	1,661	8.45	2,854	75.50	926	24.50	3,700	97.88	80	2.12
120 cbp1w 600g	10,345	59.50	7,043	40.50	15,900	80.89	3,756	19.11	2,553	67.54	1,227	32.46	3,308	87.51	472	12.49
180 cbp1w 300c	11,277	64.86	6,377	35.14	18,255	92.87	1,249	7.13	2,876	76.08	943	23.92	3,697	97.80	81	2.20
180 cbp1w 300g	11,555	66.45	5,820	33.55	17,988	91.51	1,461	8.49	2,911	77.01	847	22.99	3,711	98.17	74	1.83
180 cbp1w 600c	12,558	72.22	5,216	27.78	18,550	94.37	1,102	5.63	2,856	75.56	821	24.44	3,725	98.54	53	1.46
180 cbp1w 600g	11,097	63.82	6,626	36.18	16,415	83.51	3,214	16.49	2,656	70.26	1,114	29.74	3,303	87.38	479	12.62
360 cbp1w 300c	12,666	72.84	4,722	27.16	19,248	97.92	408	2.08	2,951	70.24	829	21.93	3,653	96.64	127	3.36
360 cbp1w 300g	12,902	74.20	4,486	25.80	19,189	97.62	467	2.38	2,920	77.25	860	22.75	3,672	97.14	108	2.86
360 cbp1w 600c	13,458	77.40	3,930	22.60	19,267	98.02	389	1.98	3,047	80.61	733	19.39	3,659	96.80	121	3.20
360 cbp1w 600g	13,429	77.23	3,959	22.77	19,272	98.05	384	1.95	2,798	74.02	982	25.98	3,659	96.80	121	3.20
120 cbp2w 300c	8,918	51.29	8,470	48.71	17,129	87.14	2,527	12.86	2,377	62.88	1,403	37.12	3,629	96.01	151	3.99
120 cbp2w 300g	8,945	51.44	8,443	48.56	16,578	84.34	3,078	15.66	2,308	61.06	1,472	38.94	3,611	95.53	169	4.47
120 cbp2w 600c	9,523	54.77	7,865	45.23	16,931	86.14	2,725	13.86	2,345	62.04	1,435	37.96	3,694	97.72	86	2.28
120 cbp2w 600g	9,667	55.60	7,721	44.40	17,062	86.80	2,594	13.20	2,316	61.27	1,464	38.73	3,666	96.98	114	3.02
180 cbp2w 300c	8,413	48.38	9,546	51.62	16,456	83.72	2,953	16.28	2,084	55.13	1,757	44.87	3,402	90.00	371	10.00
180 cbp2w 300g	8,916	51.28	8,750	48.72	17,289	87.96	2,274	12.04	2,451	64.84	1,370	35.16	3,661	96.85	107	3.15
180 cbp2w 600c	9,644	55.46	8,229	44.54	17,314	88.09	2,080	11.91	2,582	68.31	1,342	31.69	3,722	98.47	59	1.53
180 cbp2w 600g	9,812	56.43	7,826	43.57	17,529	89.18	1,944	10.82	2,530	66.93	1,325	33.07	3,700	97.88	78	2.12
360 cbp2w 300c	9,793	56.32	7,595	43.68	18,373	93.47	1,283	6.53	2,277	60.24	1,503	39.76	3,581	94.74	199	5.26
360 cbp2w 300g	9,921	57.06	7,467	42.94	18,517	94.21	1,139	5.79	2,364	62.54	1,416	37.46	3,611	95.53	169	4.47
360 cbp2w 600c	10,908	62.73	6,480	37.27	18,785	95.57	871	4.43	2,605	68.92	1,175	31.08	3,637	96.22	143	3.78
360 cbp2w 600g	11,380	65.45	6,008	34.55	18,736	95.32	920	4.68	2,531	66.96	1,249	33.04	3,649	96.53	131	3.47

ตาราง ค-8 ผลการรู้จำเอกสารชนิดเคมี ตามความละเอียดในการพิมพ์

ไฟล์เอกสาร	อักษรไทย				อักษรอังกฤษ				เลขไทย				เลขอารบิค			
	ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด	
	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%
120 cmp1w 300c	11,464	65.93	5,924	34.07	17,646	89.77	2,010	10.23	2,839	75.11	941	24.89	3,692	97.67	88	2.33
120 cmp1w 300g	11,061	63.61	6,327	36.39	17,043	86.71	2,613	13.29	2,695	71.30	1,085	28.70	3,558	94.13	222	5.87
120 cmp1w 600c	12,381	71.20	5,007	28.80	18,284	93.02	1,372	6.98	2,973	78.65	807	21.35	3,667	97.01	113	2.99
120 cmp1w 600g	12,515	71.97	4,873	28.03	18,234	92.77	1,422	7.23	2,989	79.07	791	20.93	3,698	97.83	82	2.17
180 cmp1w 300c	12,130	69.76	5,913	30.24	18,278	92.99	1,900	7.01	3,024	80.00	758	20.00	3,732	98.73	154	1.27
180 cmp1w 300g	12,348	71.01	4,953	28.99	18,247	92.83	1,285	7.17	3,027	80.08	753	19.92	3,687	97.54	95	2.46
180 cmp1w 600c	13,032	74.95	4,369	25.05	18,489	94.06	996	5.94	3,096	81.90	757	18.10	3,725	98.54	60	1.46
180 cmp1w 600g	13,287	76.41	4,308	23.59	18,573	94.49	1,050	5.51	3,095	81.88	786	18.12	3,717	98.33	55	1.67
360 cmp1w 300c	12,119	69.70	5,269	30.30	18,010	91.63	1,646	8.37	2,896	76.61	884	23.39	3,444	91.11	336	8.89
360 cmp1w 300g	13,235	76.12	4,153	23.88	19,234	97.85	422	2.15	2,976	78.73	804	21.27	3,687	97.54	93	2.46
360 cmp1w 600c	13,931	80.12	3,457	19.88	19,247	97.92	409	2.08	2,946	77.94	834	22.06	3,676	97.25	104	2.75
360 cmp1w 600g	13,958	80.27	3,430	19.73	19,246	97.91	410	2.09	2,951	78.07	829	21.93	3,692	97.67	88	2.33
120 cmp2p 300c	11,434	65.76	5,954	34.24	17,769	90.40	1,887	9.60	2,867	75.85	913	24.15	3,694	97.72	86	2.28
120 cmp2p 300g	10,692	61.49	6,696	38.51	15,959	81.19	3,697	18.81	2,444	64.66	1,336	35.34	3,338	88.31	442	11.69
120 cmp2p 600c	11,787	67.79	5,601	32.21	17,848	90.80	1,808	9.20	2,773	73.36	1,007	26.64	3,671	97.12	109	2.88
120 cmp2p 600g	11,606	66.75	5,782	33.25	17,920	91.17	1,736	8.83	2,851	75.42	929	24.58	3,704	97.99	76	2.01
180 cmp2p 300c	12,170	69.99	5,600	30.01	18,416	93.69	1,188	6.31	2,915	77.12	909	22.88	3,689	97.59	83	2.41
180 cmp2p 300g	12,184	70.07	5,497	29.93	17,959	91.37	1,602	8.63	2,898	76.67	950	23.33	3,674	97.20	98	2.80
180 cmp2p 600c	12,686	72.96	4,941	27.04	18,537	94.31	1,144	5.69	2,999	79.34	787	20.66	3,718	98.36	62	1.64
180 cmp2p 600g	12,692	72.99	5,019	27.01	18,358	93.40	1,234	6.60	2,899	76.69	861	23.31	3,722	98.47	62	1.53
360 cmp2p 300c	12,248	70.44	5,140	29.56	19,143	97.39	513	2.61	2,653	70.19	1,127	29.81	3,648	96.51	132	3.49
360 cmp2p 300g	12,704	73.06	4,684	26.94	19,075	97.04	581	2.96	2,691	71.19	1,089	28.81	3,671	97.12	109	2.88
360 cmp2p 600c	13,416	77.16	3,972	22.84	19,183	97.59	473	2.41	2,841	75.16	939	24.84	3,657	96.75	123	3.25
360 cmp2p 600g	13,423	77.20	3,965	22.80	19,097	97.16	559	2.84	2,887	76.38	893	23.62	3,640	96.30	140	3.70
120 cmp3y 300c	9,887	56.86	7,501	43.14	17,496	89.01	2,160	10.99	2,305	60.98	1,475	39.02	3,672	97.14	108	2.86
120 cmp3y 300g	9,695	55.76	7,693	44.24	16,943	86.20	2,713	13.80	2,329	61.61	1,451	38.39	3,647	96.48	133	3.52
120 cmp3y 600c	10,823	62.24	6,565	37.76	17,618	89.63	2,038	10.37	2,376	62.86	1,404	37.14	3,670	97.09	110	2.91
120 cmp3y 600g	10,782	62.01	6,606	37.99	17,507	89.07	2,149	10.93	2,560	67.72	1,220	32.28	3,671	97.12	109	2.88
180 cmp3y 300c	10,562	60.74	6,897	39.26	17,806	90.59	1,794	9.41	2,654	70.21	1,130	29.79	3,636	96.19	136	3.81
180 cmp3y 300g	10,787	62.04	6,847	37.96	17,418	88.61	2,170	11.39	2,646	70.00	1,148	30.00	3,699	97.86	87	2.14
180 cmp3y 600c	11,279	64.87	6,371	35.13	17,986	91.50	1,591	8.50	2,763	73.10	1,021	26.90	3,692	97.67	86	2.33
180 cmp3y 600g	11,339	65.21	6,493	34.79	17,912	91.13	1,723	8.87	2,759	72.99	1,072	27.01	3,683	97.43	89	2.57
360 cmp3y 300c	12,323	70.87	5,065	29.13	18,698	95.13	958	4.87	2,730	72.22	1,050	27.78	3,635	96.16	145	3.84
360 cmp3y 300g	11,629	66.88	5,759	33.12	18,691	95.09	965	4.91	2,600	68.78	1,180	31.22	3,607	95.42	173	4.58
360 cmp3y 600c	12,840	73.84	4,548	26.16	18,792	95.60	864	4.40	2,810	74.34	970	25.66	3,656	96.72	124	3.28
360 cmp3y 600g	12,764	73.41	4,624	26.59	18,724	95.26	932	4.74	2,702	71.48	1,078	28.52	3,656	96.72	124	3.28

ตาราง ค-9 ผลการรู้จำเอกสารชนิดคาร์บอน ตามความละเอียดและชนิดในการกราดภาพ

ไฟล์เอกสาร	อักษรไทย				อักษรอังกฤษ				เลขไทย				เลขอารบิค			
	ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด	
	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%
14 cbp1w 300c	10,153	58.39	7,235	41.61	18,001	91.58	1,655	8.42	2,556	67.62	1,224	32.38	2,556	67.62	1,224	32.38
16 cbp1w 300c	12,342	70.98	5,046	29.02	18,321	93.21	1,335	6.79	3,040	80.42	740	19.58	3,040	80.42	740	19.58
18 cbp1w 300c	12,801	73.62	4,587	26.38	18,627	94.76	1,029	5.24	3,030	80.16	750	19.84	3,030	80.16	750	19.84
14 cbp1w 300g	10,299	59.23	7,089	40.77	17,764	90.37	1,892	9.63	2,586	68.41	1,194	31.59	2,586	68.41	1,194	31.59
16 cbp1w 300g	12,362	71.10	5,026	28.90	18,206	92.62	1,450	7.38	3,074	81.32	706	18.68	3,074	81.32	706	18.68
18 cbp1w 300g	13,184	75.82	4,204	24.18	18,623	94.74	1,033	5.26	3,085	81.61	695	18.39	3,085	81.61	695	18.39
14 cbp1w 600c	11,314	65.07	6,074	34.93	18,285	93.03	1,371	6.97	2,742	72.54	1,038	27.46	2,742	72.54	1,038	27.46
16 cbp1w 600c	13,449	77.35	3,939	22.65	18,620	94.73	1,036	5.27	3,126	82.70	654	17.30	3,126	82.70	654	17.30
18 cbp1w 600c	13,826	79.51	3,562	20.49	18,742	95.35	914	4.65	3,152	83.39	628	16.61	3,152	83.39	628	16.61
14 cbp1w 600g	11,337	65.20	6,051	34.80	18,253	92.86	1,403	7.14	2,762	73.07	1,018	26.93	2,762	73.07	1,018	26.93
16 cbp1w 600g	13,578	78.09	3,810	21.91	18,504	94.14	1,152	5.86	3,076	81.38	704	18.62	3,076	81.38	704	18.62
18 cbp1w 600g	9,257	53.24	8,131	46.76	12,532	63.76	7,124	36.24	2,115	55.95	1,665	44.05	2,115	55.95	1,665	44.05
14 cbp2w 300c	7,346	42.25	10,042	57.75	16,978	86.38	2,678	13.62	1,923	50.87	1,857	49.13	1,923	50.87	1,857	49.13
16 cbp2w 300c	9,680	55.67	7,708	44.33	17,429	88.67	2,227	11.33	2,645	69.97	1,135	30.03	2,645	69.97	1,135	30.03
18 cbp2w 300c	10,750	61.82	6,638	38.18	17,368	88.36	2,288	11.64	2,627	69.50	1,153	30.50	2,627	69.50	1,153	30.50
14 cbp2w 300g	7,584	43.62	9,804	56.38	16,875	85.85	2,781	14.15	2,161	57.17	1,619	42.83	2,161	57.17	1,619	42.83
16 cbp2w 300g	10,065	57.88	7,323	42.12	17,319	88.11	2,337	11.89	2,593	68.60	1,187	31.40	2,593	68.60	1,187	31.40
18 cbp2w 300g	11,145	64.10	6,243	35.90	17,786	90.49	1,870	9.51	2,690	71.16	1,090	28.84	2,690	71.16	1,090	28.84
14 cbp2w 600c	8,020	46.12	9,368	53.88	17,038	86.68	2,618	13.32	2,162	57.20	1,618	42.80	2,162	57.20	1,618	42.80
16 cbp2w 600c	11,047	63.53	6,341	36.47	17,603	89.56	2,053	10.44	2,801	74.10	979	25.90	2,801	74.10	979	25.90
18 cbp2w 600c	11,973	68.86	5,415	31.14	18,051	91.83	1,605	8.17	2,839	75.11	941	24.89	2,839	75.11	941	24.89
14 cbp2w 600g	8,305	47.76	9,083	52.24	17,238	87.70	2,418	12.30	2,134	56.46	1,646	43.54	2,134	56.46	1,646	43.54
16 cbp2w 600g	11,245	64.67	6,143	35.33	17,648	89.78	2,008	10.22	2,803	74.15	977	25.85	2,803	74.15	977	25.85
18 cbp2w 600g	12,119	69.70	5,269	30.30	18,036	91.76	1,620	8.24	2,750	72.75	1,030	27.25	2,750	72.75	1,030	27.25

ตาราง ค-10 ผลการรู้จำเอกสารชนิดเคมี ตามความละเอียดและชนิดในการกราดภาพ

ไฟล์เอกสาร	อักษรไทย				อักษรอังกฤษ				เลขไทย				เลขอารบิค			
	ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด		ค่าความถูกต้อง		ค่าความผิดพลาด	
	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%	ตัวอักษร	%
14 cmp1w 300c	11,340	65.22	6,048	34.78	17,946	91.30	1,710	8.70	2,881	76.22	899	23.78	2,881	76.22	899	23.78
16 cmp1w 300c	12,695	73.01	4,693	26.99	18,357	93.39	1,299	6.61	3,066	81.11	714	18.89	3,066	81.11	714	18.89
18 cmp1w 300c	12,591	72.41	4,797	27.59	18,062	91.89	1,594	8.11	3,042	80.48	738	19.52	3,042	80.48	738	19.52
14 cmp1w 300g	11,270	64.81	6,118	35.19	18,034	91.75	1,622	8.25	2,792	73.86	988	26.14	2,792	73.86	988	26.14
16 cmp1w 300g	13,054	75.07	4,334	24.93	18,363	93.42	1,293	6.58	3,096	81.90	684	18.10	3,096	81.90	684	18.10
18 cmp1w 300g	12,989	74.70	4,399	25.30	17,967	91.41	1,689	8.59	3,032	80.21	748	19.79	3,032	80.21	748	19.79
14 cmp1w 600c	11,877	68.31	5,511	31.69	18,456	93.89	1,200	6.11	2,934	77.62	846	22.38	2,934	77.62	846	22.38
16 cmp1w 600c	13,992	80.47	3,396	19.53	18,739	95.33	917	4.67	3,078	81.43	702	18.57	3,078	81.43	702	18.57
18 cmp1w 600c	14,357	82.57	3,031	17.43	18,785	95.57	871	4.43	3,218	85.13	562	14.87	3,218	85.13	562	14.87
14 cmp1w 600g	12,140	69.82	5,248	30.18	18,382	93.52	1,274	6.48	2,922	77.30	858	22.70	2,922	77.30	858	22.70
16 cmp1w 600g	14,120	81.21	3,268	18.79	18,774	95.51	882	4.49	3,081	81.51	699	18.49	3,081	81.51	699	18.49
18 cmp1w 600g	14,324	82.38	3,064	17.62	18,887	96.09	769	3.91	3,231	85.48	549	14.52	3,231	85.48	549	14.52
14 cmp2p 300c	10,819	62.22	6,569	37.78	18,068	91.92	1,588	8.08	2,669	70.61	1,111	29.39	2,669	70.61	1,111	29.39
16 cmp2p 300c	12,832	73.80	4,556	26.20	18,465	93.94	1,191	6.06	2,986	78.99	794	21.01	2,986	78.99	794	21.01
18 cmp2p 300c	13,006	74.80	4,382	25.20	18,618	94.72	1,038	5.28	3,042	80.48	738	19.52	3,042	80.48	738	19.52
14 cmp2p 300g	10,934	62.88	6,454	37.12	17,583	89.45	2,073	10.55	2,629	69.55	1,151	30.45	2,629	69.55	1,151	30.45
16 cmp2p 300g	12,736	73.25	4,652	26.75	18,067	91.92	1,589	8.08	3,026	80.05	754	19.95	3,026	80.05	754	19.95
18 cmp2p 300g	12,663	72.83	4,725	27.17	17,241	87.71	2,415	12.29	2,733	72.30	1,047	27.70	2,733	72.30	1,047	27.70
14 cmp2p 600c	11,507	66.18	5,881	33.82	18,136	92.27	1,520	7.73	2,751	72.78	1,029	27.22	2,751	72.78	1,029	27.22
16 cmp2p 600c	13,337	76.70	4,051	23.30	18,476	94.00	1,180	6.00	3,001	79.39	779	20.61	3,001	79.39	779	20.61
18 cmp2p 600c	13,754	79.10	3,634	20.90	18,799	95.64	857	4.36	3,064	81.06	716	18.94	3,064	81.06	716	18.94
14 cmp2p 600g	11,228	64.57	6,160	35.43	18,110	92.13	1,546	7.87	2,694	71.27	1,086	28.73	2,694	71.27	1,086	28.73
16 cmp2p 600g	13,340	76.72	4,048	23.28	18,440	93.81	1,216	6.19	2,991	79.13	789	20.87	2,991	79.13	789	20.87
18 cmp2p 600g	13,900	79.94	3,488	20.06	18,732	95.30	924	4.70	3,130	82.80	650	17.20	3,130	82.80	650	17.20
14 cmp3y 300c	9,313	53.56	8,075	46.44	17,351	88.27	2,305	11.73	2,216	58.62	1,564	41.38	2,216	58.62	1,564	41.38
16 cmp3y 300c	11,415	65.65	5,973	34.35	18,346	93.34	1,310	6.66	2,818	74.55	962	25.45	2,818	74.55	962	25.45
18 cmp3y 300c	12,501	71.89	4,887	28.11	18,359	93.40	1,297	6.60	1,863	49.10	666	17.57	1,863	49.10	666	17.57
14 cmp3y 300g	9,414	54.14	7,974	45.86	16,904	86.00	2,752	14.00	2,211	58.49	1,569	41.51	2,211	58.49	1,569	41.51
16 cmp3y 300g	11,674	67.14	5,714	32.86	17,989	91.52	1,667	8.48	2,802	74.13	978	25.87	2,802	74.13	978	25.87
18 cmp3y 300g	11,950	68.73	5,438	31.27	18,156	92.37	1,500	7.63	2,816	74.50	964	25.50	2,816	74.50	964	25.50
14 cmp3y 600c	10,090	58.03	7,298	41.97	17,642	89.75	2,014	10.25	2,491	65.90	1,289	34.10	2,491	65.90	1,289	34.10
16 cmp3y 600c	12,431	71.49	4,957	28.51	18,348	93.35	1,308	6.65	2,782	73.60	998	26.40	2,782	73.60	998	26.40
18 cmp3y 600c	12,846	73.88	4,542	26.12	18,409	93.66	1,247	6.34	1,815	47.83	714	18.84	1,815	47.83	714	18.84
14 cmp3y 600g	9,922	57.06	7,466	42.94	17,551	89.29	2,105	10.71	2,440	64.55	1,340	35.45	2,440	64.55	1,340	35.45
16 cmp3y 600g	12,384	71.22	5,004	28.78	18,384	93.53	1,272	6.47	2,865	75.79	915	24.21	2,865	75.79	915	24.21
18 cmp3y 600g	12,875	74.05	4,513	25.95	18,254	92.87	1,402	7.13	2,904	76.83	876	23.17	2,904	76.83	876	23.17

ภาคผนวก ง

ผลงานตีพิมพ์และเผยแพร่

ภาพประกอบ ง-1 เกียรติบัตรรับรองการนำเสนอผลงาน

ภาพประกอบ ง-2 หน้าปกหนังสือเรื่องเต็มประชุมวิชาการ

การประชุมวิชาการทางวิศวกรรมศาสตร์มหาวิทยาลัยสงขลานครินทร์ ครั้งที่ 11
19-20 มิถุนายน 2558

การวิเคราะห์ปัจจัยในการพิมพ์เอกสารสำเนาด้วยเครื่องพิมพ์ดอทเมทริกซ์เพื่อการรู้จำในระบบสำนักงานอัตโนมัติ

Factors analysis in the copied documents printed with dot matrix printers for recognition in office automation.

ปิยะนาถ เควิจดำ[1]* สมชัย หลิมติโรรัตน์[2]
[1]โครงการจัดการศึกษาพิเศษ หลักสูตรวิทยาศาสตร์มหาบัณฑิต สาขาวิชาการจัดการเทคโนโลยีสารสนเทศ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยสงขลานครินทร์ อ.หาดใหญ่ จ.สงขลา 90112
[2]ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยสงขลานครินทร์ อ.หาดใหญ่ จ.สงขลา 90112
*อีเมล์ของผู้แต่งหลัก: piyanartq@gmail.com
Piyanart Quedum[1]*, Somchai Limsiroratana[2]
[1]Special Program in Management of Information Technology, Faculty of Engineering, Prince of Songkla University, Hat Yai, Songkhla 90112
[2] Department of Computer Engineering, Faculty of Engineering, Prince of Songkla University, Hat Yai, Songkhla 90112
*E-mail of corresponding author: piyanartq@gmail.com

บทคัดย่อ

เอกสารสำเนาที่พิมพ์ด้วยเครื่องพิมพ์ดอทเมทริกซ์ ยังคงมีใช้กันอย่างแพร่หลาย และเมื่อนำข้อมูลกลับเข้าสู่กระบวนการทางธุรกิจจะมี 2 ทางเลือกคือ ใช้คนป้อนข้อมูลซึ่งควบคุมได้ยากและมีข้อจำกัดเนื่องจากความเหนื่อยล้า และใช้การรู้จำอักขระด้วยแสงซึ่งทำให้เป็นระบบอัตโนมัติได้ดีกว่า แต่ก็มีข้อผิดพลาดอยู่บ้างเนื่องจากหลายปัจจัย งานวิจัยนี้ได้ศึกษาวิเคราะห์ปัจจัยต่างๆ จากตัวอย่างกระดาษต่อเนื่องทั้งแบบชนิดคาร์บอนและชนิดเคมี โดยวิเคราะห์ปัจจัยทั้งชนิดแบบอักษร ขนาดตัวอักษร ลักษณะตัวอักษร ความละเอียดในการพิมพ์ ความละเอียดและชนิดในการกราดภาพ เพื่อให้ได้วิธีที่ดีที่สุด เหมาะสมที่สุด สำหรับการออกแบบแบบฟอร์มทางธุรกิจที่มีประสิทธิภาพและประสิทธิผลในระบบสำนักงานอัตโนมัติ

คำหลัก การรู้จำ การรู้จำอักขระด้วยแสง ดอทเมทริกซ์

Abstract

Copies documents printed with dot matrix printers still widely used. When extract data back into business process, there are two choices, the human input method which is hard to control and limited by fatigue, and the optical character recoginition (OCR) method which better for automation systems that is not very accurate due to many factros. This research analyze on two type, carbon and chemicals continuous papers. The font, font size, font styles, printed resolution and scanned resolutions factors are analyze to obtain the best approach suitable to design the most effectively and effectiveness business form used in office automation system.

Keywords: recognition, OCR, dot matrix

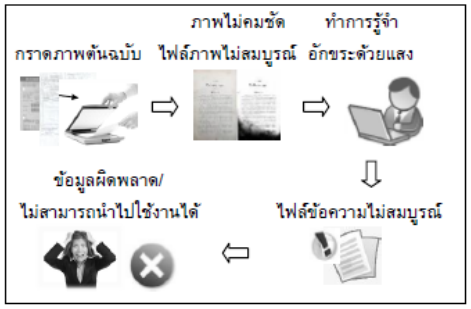
1. บทนำ

การจัดการเอกสาร มีความสำคัญมากสำหรับธุรกิจห้างร้าน หน่วยงาน ทั้งของรัฐ เอกชน และอื่นๆเอกสารที่พิมพ์ด้วยเครื่องพิมพ์ดอทเมทริกซ์ โดยเฉพาะเอกสารสำเนาจะมีลักษณะที่ไม่สมบูรณ์ เช่น ตัวอักษรไม่คมชัดเพราะผ้าหมึกจาง ตัวอักษรติดกัน ตัวอักษรขาดเนื่องจากหัวเข็มหัก เมื่อนำเอกสารกระดาษไปกราดภาพ (Scan) และจัดการให้เป็นไฟล์เอกสารด้วยการรู้จำอักขระด้วยแสง (OCR : Optical Character Recognition) ภาพที่ได้จากการกราดภาพไม่ชัดเนื่องจากการตั้งค่าไม่เหมาะสม ดังรูปที่ 1

รูปที่ 1 ตัวอย่างไฟล์ภาพเอกสารดอทเมทริกซ์ที่ไม่สมบูรณ์

ภาพประกอบ ง-3 เนื้อหาบทความวิชาการหน้า 117

จากรูปที่ 2 เป็นกระบวนการนำข้อมูลกลับเข้าสู่กระบวนการธุรกิจ ซึ่งจะเกิดกรณีผิดพลาดต่างๆ ในแต่ละขั้นตอน เมื่อนำไปแปลงเป็นไฟล์ข้อความโดยการรู้จำอักขระด้วยแสง มักจะเจอปัญหาไฟล์ข้อความที่ได้ไม่ถูกต้องครบถ้วน ไม่สมบูรณ์ ข้อมูลผิดพลาดหรือไม่สามารถนำไปใช้งานได้



รูปที่ 2 ปัญหาของการรู้จำจากไฟล์ภาพที่ไม่สมบูรณ์

เอกสารที่พิมพ์ด้วยเครื่องพิมพ์ดอทเมทริกซ์ ที่เป็นเอกสารสำเนา ที่พิมพ์ได้หลายชั้นทำให้ประหยัดทั้งเวลาและต้นทุน โดยใช้กระดาษชนิดคาร์บอน หรือชนิดเคมี ซึ่งมีความหลากหลาย ทั้งกระดาษ สีพื้นของกระดาษ สีของหมึกพิมพ์ ความคมชัดของผ้าหมึก ความละเอียดในการพิมพ์

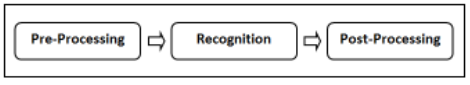
งานวิจัยนี้จะวิเคราะห์ปัจจัยในการพิมพ์เอกสารสำเนา ด้วยเครื่องพิมพ์ดอทเมทริกซ์ ที่ใช้กระดาษชนิดคาร์บอน และชนิดเคมีที่มีผลต่อการรู้จำอักขระด้วยแสง เช่น สีตัวอักษร สีกระดาษ ขนาดตัวอักษร ความละเอียดในการพิมพ์ เป็นต้น เพื่อให้ได้มาซึ่งแนวทางในการออกแบบฟอร์มที่ดีเหมาะสมกับการนำไปทำการรู้จำและลดข้อผิดพลาดที่อาจจะเกิดขึ้นได้

2. ทฤษฎีและหลักการ

2.1 ความรู้พื้นฐานเกี่ยวกับการรู้จำอักขระด้วยแสง

การรู้จำอักขระด้วยแสง หมายถึง การแปลงเอกสารที่อยู่ในรูปแบบของไฟล์ภาพ ให้เป็นเอกสารที่อยู่ในรูปแบบของไฟล์ข้อความโดยอัตโนมัติ ช่วยให้สะดวกในการจัดการข้อมูลงานเอกสาร และการนำไฟล์ข้อความไปใช้

วิธีการรู้จำอักขระด้วยแสง มี 3 ขั้นตอน คือ Pre-Processing , Recognition และ Post-Processing ตามลำดับ ดังรูปที่ 3



รูปที่ 3 กระบวนการทำงานของการรู้จำอักขระด้วยแสง

2.1.1 การประมวลผลขั้นต้น (Pre-Processing)

เป็นการจัดการกับไฟล์ภาพ ก่อนที่จะนำไปวิเคราะห์ข้อมูลตัวอักษร ให้มีความถูกต้องเพื่อการรู้จำมากที่สุด ซึ่งในขั้นตอนนี้จะเป็นการวิเคราะห์สิ่งแปลกปน สัญญาณรบกวน จุดดำ จุดขาว ที่ติดมาในภาพ เพื่อแยกแยะให้ได้ว่าไม่ใช่ตัวอักษร และไม่นำไปรู้จำ

2.1.2 การรู้จำ (Recognition)

เป็นการวิเคราะห์ภาพ ที่ได้จากการประมวลผลขั้นต้นว่าเป็นตัวอักษรอะไร ทำให้เราได้มาซึ่งตัวอักษรหรือข้อความที่เราต้องการ ซึ่งมีวิธีในการรู้จำอยู่หลายวิธี เช่น การรู้จำโดยใช้หลักสถิติ การรู้จำโดยอาศัยโครงข่ายประสาทเทียม [1] เป็นต้น

2.1.3 การประมวลผลขั้นปลาย (Post-Processing)

เป็นการนำรูปตัวอักษรที่ได้จากการรู้จำ ไปเข้ารหัสแสดงผลลัพธ์เป็นตัวอักษรออกมา ซึ่งในการรู้จำอาจจะไม่ถูกต้องทั้งหมด การประมวลผล หรือการตรวจสอบก็จะเพิ่มประสิทธิภาพในการทำงานของการรู้จำให้ทำงานได้ดีขึ้น ในการตรวจสอบนี้จะทำการแก้ไขความถูกต้องของภาษา ตัวสะกด ไวยากรณ์ การนำพจนานุกรมมาช่วยในการตรวจสอบคำศัพท์ หาคำผิด และแก้ไขให้ถูกต้อง หรือแสดงให้เห็นตัวอักษรที่ผิด หรือแก้ไขให้อัตโนมัติ

2.2 การเลือกใช้โปรแกรมการอ่านอักขระด้วยแสง

การเลือกใช้โปรแกรมการรู้จำอักขระด้วยแสง มีความสำคัญมาก เพราะมีผลโดยตรงต่อการวิจัย ควรใช้โปรแกรมที่มีความแม่นยำสูงช่วยให้การทำงานมีประสิทธิภาพ สามารถนำข้อมูลไปใช้ได้ดีที่สุดในการรู้จำ โปรแกรมรู้จำเท่าที่ได้ศึกษามีตัวอย่าง ตามตารางที่ 1 ดังนี้

ตารางที่ 1 ข้อดี/ข้อเสีย ของโปรแกรมรู้จำอักขระด้วยแสง

โปรแกรมรู้จำ	*ข้อดี*	*ข้อเสีย*
ABBYY Fine Reader	ความสามารถครบถ้วนประสิทธิภาพดี สามารถรู้จำการวางผังหน้ากระดาษ	ซอฟต์แวร์ต้องซื้อยากที่จะไปรวมกับโปรแกรมประยุกต์อื่น
GOCR	ไม่ต้องสอนสามารถที่จะเพิ่มชนิดแบบอักษรได้ สามารถเปิดแฟ้มข้อมูลได้หลายรูปแบบ	ไม่รองรับลักษณะตัวอักษรเอียง ไม่รู้จำการวางผังหน้ากระดาษ ตัวอักษรต้องแยกจากันโดยเด็ดขาด
Tesseract	ความแม่นยำดี สามารถไปรวมกับโปรแกรมประยุกต์อื่นได้	ไม่รู้จำการวางผังหน้ากระดาษ นำมาใช้ยาก

ซึ่งมีงานวิจัยได้ทดสอบความแม่นยำในการรู้จำ โดยใช้โปรแกรม ABBYY Fine Reader พบว่ามีความถูกต้อง 97 – 99% ใช้งานได้ง่าย มีประสิทธิภาพ [2] ใช้เทคโนโลยีการรู้จำ Integrity, Purposefulness and Adaptability (IPA) เป็นเทคโนโลยีที่อยู่เบื้องหลัง หลักการทั้ง 3 อย่างนี้ทำให้สามารถมีการรู้จำคล้ายมนุษย์ [3] และนอกจากนี้มีงานวิจัยที่ได้ศึกษาการรู้จำชนิดแบบ

ภาพประกอบ ง-4 เนื้อหาบทความวิชาการหน้า 118

อักษร (Font) แบบเก่าๆ โดยเลือกใช้ ABBYY Fine Reader ในการศึกษาปัจจัยที่มีผลกระทบต่อการอ่านอักขระด้วยแสงเช่นกัน [4]

งานวิจัยนี้เลือกใช้ ABBYY Fine Reader เนื่องจากมีความแม่นยำในการรู้จำและให้ค่าความถูกต้อง 97 – 99%

3. วิธีการดำเนินการวิจัย

3.1 รวบรวมเอกสารตัวอย่างการวิจัย

ตัวอย่างการวิจัยได้แก่ เอกสารสำเนาที่พิมพ์ด้วยเครื่องพิมพ์ดอทเมทริกซ์ใช้กระดาษต่อเนื่องชนิดคาร์บอนซึ่งมีกระดาษเป็นสีขาว ตัวอักษรสีดำ และชนิดเคมีซึ่งมีกระดาษสีชมพู ตัวอักษรสีน้ำเงิน ตามปัจจัยที่สามารถควบคุมได้ ดังนี้

3.1.1 ชนิดแบบอักษร มี 3 ปัจจัย ได้แก่ Angsana New (นิยมใช้ทั่วไป), Cordia New (นิยมใช้ทั่วไป) และ TH SarabunPSK (ใช้ในราชการ)

3.1.2 ขนาดตัวอักษร มี 3 ปัจจัย ได้แก่ ขนาด 14 (ถ้าเล็กกว่าภาษาไทยจะใช้ไม่ได้), 16 และ 18 (ถ้าใหญ่กว่าก็ใหญ่เกินไป)

3.1.3 ลักษณะตัวอักษร มี 3 ปัจจัย ได้แก่ ตัวปกติ ตัวหนา และตัวเอียง

3.1.4 ความละเอียดในการพิมพ์ มี 3 ปัจจัย ได้แก่ 120 x 180 dpi, 180 x 180 dpi และ 360 x 180 dpi (เครื่องพิมพ์ที่ใช้พิมพ์มี 3 ความละเอียด)

3.1.5 ความละเอียดและชนิดในการกราดภาพ มี 4 ปัจจัย ได้แก่ ความละเอียด 300 dpi (เพียงพอที่จะใช้ในการรู้จำอักขระด้วยแสง ไม่เร็วไม่ช้าเกินไป) มี 2 แบบ คือ ชนิด Color และ Gray ความละเอียด 600 dpi มี 2 แบบ คือ ชนิด Color และ Gray บันทึกเป็นไฟล์รูปภาพแบบ PNG (Portable Network Graphic)

3.1.6 ตัวอักษรที่ใช้ทดสอบการรู้จำ มีทั้งภาษาไทย ภาษาอังกฤษตัวพิมพ์ใหญ่ และตัวพิมพ์เล็ก

จากปัจจัยชนิดแบบอักษร (3) ขนาดตัวอักษร (3) ลักษณะอักษร (3) ความละเอียดในการพิมพ์ (3) เมื่อนำมาพิมพ์ลงบนกระดาษสำเนา 2 ชนิดตามที่ได้กำหนดไว้คือ คาร์บอน และเคมี ทำให้สามารถคำนวณเป็นจำนวนรูปแบบได้ทั้งหมด 3 x 3 x 3 x 3 x 2 = 162 หน้า

จากปัจจัยความละเอียดและชนิดในการกราดภาพ (4) จะได้ไฟล์รูปภาพทั้งหมด 162 x 4 = 648 ไฟล์

3.2 การทำการรู้จำ

นำไฟล์ภาพที่ได้จากการกราดภาพ มาทำการรู้จำด้วยโปรแกรม ABBYY Fine Reader 11 ระบบปฏิบัติการ Windows 7 บันทึกเป็นไฟล์ข้อความ .txt

3.3 วัดผลการรู้จำ

นำไฟล์ .txt ที่ได้จากการอ่านอักขระด้วยแสงมานับตัวอักษรถูก/ผิด ด้วยโปรแกรมที่พัฒนาขึ้น บันทึกเป็นไฟล์ .csv

3.4 หาค่าความถูกต้องและค่าความผิดพลาดของการรู้จำ

นำไฟล์ .csv ที่ได้จากการวัดผลการรู้จำ ด้วยการนับตัวอักษรถูก/ผิด ประมวลผล วิเคราะห์ และเปรียบเทียบ ด้วยโปรแกรม Microsoft Excel

4. ผลการดำเนินงาน

จากการดำเนินการวิจัย ประมวลผล เปรียบเทียบ วิเคราะห์ปัจจัยที่มีผลต่อการรู้จำเอกสารสำเนาดอทเมทริกซ์ ชนิดคาร์บอน และชนิดเคมี ได้ผลการทดลองดังตารางที่ 2 ตารางที่ 3 และตารางที่ 4 ดังนี้

ตารางที่ 2 ผลการการรู้จำอักขระด้วยแสงแยกตามปัจจัย (ภาษาไทยและภาษาอังกฤษ)

ปัจจัย	ชนิดแบบฟอร์ม	ค่าความถูกต้อง			
		เอกสารสำเนาชนิดคาร์บอน		เอกสารสำเนาชนิดเคมี	
		$\bar{x}$ (%)	S.D. (%)	$\bar{x}$ (%)	S.D. (%)
ชนิดแบบอักษร	Angsana New	74.08	2.56	81.21	2.79
	Cordia New	**79.01**	1.53	**86.83**	0.74
	TH Sarabun PSK	73.41	2.45	83.43	2.26
ขนาดตัวอักษร	14	68.50	1.57	79.56	1.31
	16	78.04	2.32	**86.15**	0.99
	18	**79.97**	2.50	85.75	3.49
ลักษณะตัวอักษร	ตัวปกติ	**77.41**	1.80	**86.45**	0.44
	ตัวหนา	75.77	3.00	84.36	2.32
	ตัวเอียง	73.32	1.64	80.65	3.37
ความละเอียดในการพิมพ์	120 x 180	72.59	1.20	78.99	3.90
	180 x 180	76.29	3.29	**86.31**	1.17
	360 x 180	**77.63**	2.22	86.16	1.24
ความละเอียดและชนิดในการกราดภาพ	300 Color	73.13	5.56	83.61	3.71
	300 Gray	74.36	5.94	81.22	3.61
	600 Color	77.02	6.79	**85.32**	3.82
	600 Gray	**77.50**	6.33	85.13	4.40
ค่าเฉลี่ย		**75.50**	**3.17**	**83.82**	**2.47**

จากตารางที่ 2 ชนิดแบบอักษร เป็นข้อมูลที่ได้จากการนำค่าความถูกต้องของการรู้จำอักขระด้วยแสง จากชนิดแบบอักษร Angsana New, Cordia New และ TH SarabunPSK ทั้งขนาดตัวอักษร 14 16 และ 18 ลักษณะตัวอักษรตัวปกติ ตัวหนา และตัวเอียง ที่พิมพ์ที่ความละเอียด 180 x 120 dpi, 180 x 180 dpi และ 360 x 180 dpi และการกราดภาพด้วยความละเอียด 300 และ 600 dpi ทั้งชนิด Color และ Gray

ขนาดตัวอักษร เป็นข้อมูลที่ได้จากการนำค่าความถูกต้องของการรู้จำอักขระด้วยแสง จากขนาดตัวอักษร 14 16 และ 18 ทั้ง

ภาพประกอบ ง-5 เนื้อหาบทความวิชาการหน้า 119

ชนิดแบบอักษร Angsana New, Cordia New และ TH SarabunPSK ลักษณะตัวอักษรตัวปกติ ตัวหนา และตัวเอียง ที่พิมพ์ที่ความละเอียด 180 x 120 dpi, 180 x180 dpi และ 360 x 180 dpi ที่กราดภาพด้วยความละเอียด 300 และ 600 dpi ทั้งชนิด Color และ Gray

ลักษณะตัวอักษร เป็นข้อมูลที่ได้จากการนำค่าความถูกต้องของการรู้จำอักขระด้วยแสง จากลักษณะตัวอักษรตัวปกติ ตัวหนา และตัวเอียง ทั้งชนิดแบบอักษร Angsana New, Cordia New และ TH SarabunPSK ขนาดตัวอักษร 14 16 และ 18 ที่พิมพ์ที่ความละเอียด 180 x 120 dpi, 180 x180 dpi และ 360 x 180 dpi และกราดภาพด้วยความละเอียด 300 และ 600 dpi ทั้งชนิด Color และ Gray

ความละเอียดในการพิมพ์ เป็นข้อมูลที่ได้จากการนำค่าความถูกต้องของการรู้จำอักขระด้วยแสง จากความละเอียดในการพิมพ์ที่ 180 x 120 dpi, 180 x180 dpi และ 360 x 180 dpi ทั้งชนิดแบบอักษร Angsana New, Cordia New และ TH SarabunPSK ลักษณะตัวอักษรตัวปกติ ตัวหนา และตัวเอียง ทั้งขนาดตัวอักษร 14 16 และ 18 ที่กราดภาพด้วยความละเอียด 300 และ 600 dpi ทั้งชนิด Color และ Gray

ความละเอียดและชนิดในการกราดภาพ เป็นข้อมูลที่ได้จากการนำค่าความถูกต้องของการการรู้จำอักขระด้วยแสง จากความละเอียดในการกราดภาพ 300 และ 600 dpi ทั้งชนิด Color และ Gray ทั้งชนิดแบบอักษร Angsana New, Cordia New และ TH SarabunPSK ลักษณะตัวอักษรตัวปกติ ตัวหนา และตัวเอียง ขนาดตัวอักษร 14 16 และ 18 และความละเอียดในการพิมพ์ 180 x 120 dpi, 180 x 180 dpi และ 360 x 180 dpi

ถ้าต้องการจะเลือกชนิดแบบอักษร ที่มีความเหมาะสมที่สุดกับแบบฟอร์มควรเลือกชนิดแบบอักษร Cordia New เพราะให้ค่าความถูกต้องสูงสุด ทั้งเอกสารสำเนาชนิดคาร์บอน เท่ากับ 79.01% และเอกสารสำเนาชนิดเคมี เท่ากับ 86.83%

ถ้าต้องการเลือกขนาดตัวอักษร ที่มีความเหมาะสมที่สุดกับแบบฟอร์ม สำหรับเอกสารสำเนาชนิดคาร์บอนควรเลือกขนาดตัวอักษร 18 เพราะให้ค่าความถูกต้อง 79.97% แต่ถ้าสำหรับเอกสารสำเนาชนิดเคมีควรเลือกขนาดตัวอักษร 16 เพราะให้ค่าความถูกต้องสูงสุดเท่ากับ 86.15%

ถ้าต้องการเลือกลักษณะตัวอักษร ที่มีความเหมาะสมที่สุดกับแบบฟอร์มควรเลือกลักษณะตัวอักษรตัวปกติ เพราะให้ค่าความถูกต้องสูงสุด ทั้งเอกสารสำเนาชนิดคาร์บอน เท่ากับ 77.41% และเอกสารสำเนาชนิดเคมี เท่ากับ 86.45% ไม่ควรใช้ลักษณะตัวอักษรตัวเอียง

ถ้าต้องการเลือกความละเอียดในการพิมพ์ ที่มีความเหมาะสมที่สุดกับแบบฟอร์ม สำหรับเอกสารสำเนาชนิดคาร์บอนควรเลือกความละเอียดในการพิมพ์เท่ากับ 360 x 180 dpi เพราะให้ค่าความถูกต้องสูงสุดเท่ากับ 77.63% สำหรับเอกสารสำเนาชนิดเคมีควรเลือกความละเอียดในการพิมพ์เท่ากับ 180 x 180 dpi เพราะให้ค่าความถูกต้องสูงสุดเท่ากับ 86.31%

ถ้าต้องการเลือกความละเอียดและชนิดในการกราดภาพ ที่มีความเหมาะสมที่สุดกับแบบฟอร์ม สำหรับเอกสารสำเนาชนิดคาร์บอนควรเลือกกราดภาพที่ความละเอียด 600 dpi ชนิด Gray เพราะให้ค่าความถูกต้องสูงสุดเท่ากับ 77.50% สำหรับเอกสารสำเนาชนิดเคมีควรเลือกกราดภาพที่ความละเอียด 600 dpi ชนิด Color เพราะให้ค่าความถูกต้องสูงสุดเท่ากับ 85.32%

ตารางที่ 3 ผลการรู้จำอักขระด้วยแสง แยกตามปัจจัย (ภาษาอังกฤษ)

ปัจจัย	ชนิดแบบฟอร์ม	ค่าความถูกต้อง			
		เอกสารสำเนาชนิดคาร์บอน		เอกสารสำเนาชนิดเคมี	
		$\bar{x}$ (%)	S.D. (%)	$\bar{x}$ (%)	S.D. (%)
ชนิดแบบอักษร	Angsana New	90.54	0.77	92.52	2.72
	Cordia New	**94.11**	0.90	**96.54**	0.52
	TH Sarabun PSK	87.07	1.75	92.26	2.49
ขนาดตัวอักษร	14	88.08	0.87	92.35	1.38
	16	**92.12**	0.87	**95.44**	0.60
	18	91.52	1.70	93.55	3.68
ลักษณะตัวอักษร	ตัวปกติ	91.56	1.38	94.85	0.46
	ตัวหนา	**92.09**	2.18	**95.04**	1.92
	ตัวเอียง	88.55	0.67	91.44	3.28
ความละเอียดในการพิมพ์	120 x 180	88.66	0.97	90.12	4.63
	180 x 180	90.14	2.26	95.07	0.91
	360 x 180	**92.91**	0.84	**96.14**	0.28
ความละเอียดและชนิดในการกราดภาพ	300 Color	89.44	2.20	94.41	1.42
	300 Gray	89.99	2.34	90.97	3.32
	600 Color	91.34	2.54	**94.94**	1.51
	600 Gray	**91.52**	2.04	94.77	1.47
ค่าเฉลี่ย		***90.60***	***1.52***	***93.78***	***1.91***

จากตารางที่ 3 เป็นผลการทดลองแยกเฉพาะภาษาอังกฤษ ได้ผลคือ เอกสารสำเนาชนิดเคมีจะให้ค่าความถูกต้องมากกว่าเอกสารสำเนาชนิดคาร์บอนทุกปัจจัย ให้ค่าความถูกต้องสูงสุด ดังนี้ ชนิดแบบอักษร Cordia New เท่ากับ 96.54% ขนาดตัวอักษร 16 เท่ากับ 95.44% ลักษณะตัวอักษรตัวหนา เท่ากับ 95.04% ความละเอียดในการพิมพ์ 360 x 180 dpi เท่ากับ 96.14% ความละเอียดในการกราดภาพ 600 dpi ชนิด Color เท่ากับ 94.94% ส่วนเอกสารสำเนาชนิดคาร์บอนให้ค่าความถูกต้องสูงสุด ดังนี้ ชนิดแบบอักษร Cordia New เท่ากับ 94.11% ขนาดตัวอักษร 16 เท่ากับ 92.12% ลักษณะตัวอักษรตัวหนา เท่ากับ 92.09% ความละเอียดในการพิมพ์ 360 x 180 dpi เท่ากับ 92.91% ความละเอียด

ภาพประกอบ ง-6 เนื้อหาบทความวิชาการหน้า 120

ในการกราดภาพ 600 dpi ชนิด Gray เท่ากับ 91.52%

ตารางที่ 4 ผลการรู้จำอักขระด้วยแสง แยกตามปัจจัย (ภาษาไทย)

ปัจจัย	ชนิดแบบฟอร์ม	ค่าความถูกต้อง			
		เอกสารสำเนาชนิดคาร์บอน		เอกสารสำเนาชนิดเคมี	
		$\bar{x}$ (%)	S.D. (%)	$\bar{x}$ (%)	S.D. (%)
ชนิดแบบอักษร	Angsana New	55.86	4.56	68.68	2.98
	Cordia New	**62.29**	2.33	**76.07**	1.27
	TH Sarabun PSK	58.30	3.24	73.65	2.25
ขนาดตัวอักษร	14	46.81	2.43	65.41	1.49
	16	62.45	4.04	75.88	1.49
	18	**67.19**	3.48	**77.11**	3.54
ลักษณะตัวอักษร	ตัวปกติ	**61.11**	2.80	**77.15**	0.90
	ตัวหนา	57.71	3.94	72.54	3.37
	ตัวเอียง	56.46	2.92	68.71	3.58
ความละเอียดในการพิมพ์	120 x 180	54.80	1.79	66.67	3.12
	180 x 180	**60.96**	4.51	**76.62**	1.66
	360 x 180	60.70	3.75	75.11	2.38
ความละเอียดและชนิดในการกราดภาพ	300 Color	55.07	10.18	71.66	6.28
	300 Gray	57.06	9.95	70.42	5.57
	600 Color	61.16	11.54	**74.66**	6.43
	600 Gray	**61.97**	11.13	74.46	7.70
ค่าเฉลี่ย		***58.74***	***5.16***	***72.80***	***3.37***

จากตารางที่ 4 เป็นผลการทดลองแยกเฉพาะภาษาไทย ได้ผลคือ เอกสารสำเนาชนิดเคมีจะให้ค่าความถูกต้องมากกว่าเอกสารสำเนาชนิดคาร์บอนทุกปัจจัย ให้ค่าความถูกต้องสูงสุด ดังนี้ ชนิดแบบอักษร Cordia New เท่ากับ 76.07% ขนาดตัวอักษร 18 เท่ากับ 77.11% ลักษณะตัวอักษรตัวปกติ เท่ากับ 77.15% ความละเอียดในการพิมพ์ 180 x 180 dpi เท่ากับ 76.62% ความละเอียดในการกราดภาพ 600 dpi ชนิด Color เท่ากับ 74.66% ส่วนเอกสารสำเนาชนิดคาร์บอนให้ค่าความถูกต้องสูงสุด ดังนี้ ชนิดแบบอักษร Cordia New เท่ากับ 62.29% ขนาดตัวอักษร 18 เท่ากับ 67.19% ลักษณะตัวอักษรตัวปกติเท่ากับ 61.11% ความละเอียดในการพิมพ์ 180 x 180 dpi เท่ากับ 60.96% ความละเอียดในการกราดภาพ 600 dpi ชนิด Gray เท่ากับ 61.97%

ข้อควรระวังในการรู้จำอักขระด้วยแสง ได้แก่ ลักษณะของเอกสาร ถ้าเอกสารไม่ชัด สีตัวอักษรไม่เข้ม คุณภาพแรงกดของหัวพิมพ์ต่ำ ความเข้มของผ้าหมึกน้อย หรือความเข้มของผ้าหมึกมากเกินไปและกระดาษมีความชื้น ทำให้ผลการรู้จำอักขระด้วยแสงผิดพลาด หรือไม่สามารถทำการรู้จำอักขระด้วยแสง ได้

5. สรุปผลการดำเนินงาน

จากการดำเนินงานสรุปได้ว่า สำหรับตัวอักษรทั้งภาษาอังกฤษและภาษาไทย เอกสารสำเนาชนิดเคมีให้ค่าความถูกต้องเฉลี่ยเท่ากับ 83.82% มากกว่าเอกสารสำเนาชนิดคาร์บอน ซึ่งให้ค่าความถูกต้องเฉลี่ยเพียง 75.50% โดยควรเลือกใช้ชนิดแบบอักษร Cordia New ขนาด 16 ลักษณะตัวอักษรตัวปกติ ความละเอียดในการพิมพ์ 180 x 180 dpi ความละเอียดในการกราดภาพ 300 dpi ชนิด Color เพื่อให้ได้มาซึ่งแบบฟอร์มที่เหมาะสมสำหรับนำไปทำการรู้จำมากที่สุด ทำให้การจัดการงานด้านเอกสารสมบูรณ์แบบและถูกต้องมากที่สุด

สำหรับตัวอักษรภาษาอังกฤษ เอกสารสำเนาชนิดเคมี ให้ค่าความถูกต้องเฉลี่ยเท่ากับ 93.78% มากกว่าเอกสารสำเนาชนิดคาร์บอน ซึ่งให้ค่าความถูกต้องเฉลี่ยเพียง 90.60% โดยควรเลือกใช้ชนิดแบบอักษร Cordia New ขนาด 16 ลักษณะตัวอักษรตัวหนา ความละเอียดในการพิมพ์ 360 x 180 dpi ความละเอียดในการกราดภาพ 600 dpi ชนิด Color เพื่อให้ได้มาซึ่งแบบฟอร์มที่เหมาะสมสำหรับนำไปทำการรู้จำมากที่สุด ทำให้การจัดการงานด้านเอกสารสมบูรณ์แบบและถูกต้องมากที่สุด

สำหรับตัวอักษรภาษาไทย เอกสารสำเนาชนิดเคมี ให้ค่าความถูกต้องเฉลี่ยเท่ากับ 72.80% มากกว่าเอกสารสำเนาชนิดคาร์บอน ซึ่งให้ค่าความถูกต้องเฉลี่ยเพียง 58.74% โดยควรเลือกใช้ชนิดแบบอักษร Cordia New ขนาด 18 ลักษณะตัวอักษรตัวปกติ ความละเอียดในการพิมพ์ 180 x 180 dpi ความละเอียดในการกราดภาพ 600 dpi ชนิด Color เพื่อให้ได้มาซึ่งแบบฟอร์มที่เหมาะสมสำหรับนำไปทำการรู้จำมากที่สุด ทำให้การจัดการงานด้านเอกสารสมบูรณ์แบบและถูกต้องมากที่สุด

เอกสารอ้างอิง

[1] Berrin A. Yanikoglu, 2000. Pitch-based segmentation and recognition of dot-matrix text. International Journal on Document Analysis and Recognition Volume 3, Issue 1, 21 Apr. 2000: 34-39.

[2] Tanushyam Chattopdhyyay, Priyanka Sinha and Provot Biswas, 2011. Performance of Document Image OCR Systems for Recognizing Video Texts on Embedded Platform. International Conference on Computational Intelligence and Communication Systems, IEEE, 2011: 606-610.

[3] ABBYY, What Technology lies behind OCR? [Online], 2015. http://finereader.abbyy.com/about_ocr/whatis_ocr

[4] Iuliu Konya, Stefan Eickeler and Christoph Seibert, 2011. Character enhancement for historical newspapers printed using hot metal typesetting. 2011 International Conference on Document Analysis and Recognition (ICDAR), Beijing, 18-21 Sep. 2011: 936-940.

ภาพประกอบ ง-7 เนื้อหาบทความวิชาการหน้า 121

ประวัติผู้เขียน

ชื่อ สกุล ปิยะนาถ เคว็จดำ
รหัสประจำตัวนักศึกษา 5510121046
วุฒิการศึกษา

วุฒิ	ชื่อสถาบัน	ปีที่สำเร็จการศึกษา
บริหารธุรกิจบัณฑิต (การตลาด)	มหาวิทยาลัยรามคำแหง	2542

การตีพิมพ์เผยแพร่ผลงาน

ปิยะนาถ เคว็จดำ, สมชัย หลิมศิโรรัตน์ 2558. การวิเคราะห์ปัจจัยในการพิมพ์เอกสารสำเนาด้วยเครื่องพิมพ์ดอทเมทริกซ์เพื่อการรู้จำในระบบสำนักงานอัตโนมัติ. การประชุมวิชาการทางวิศวกรรมศาสตร์ มหาวิทยาลัยสงขลานครินทร์ ครั้งที่ 11, 117-121, ภูเก็ต, ประเทศไทย, 19 – 20 มิถุนายน 2558.