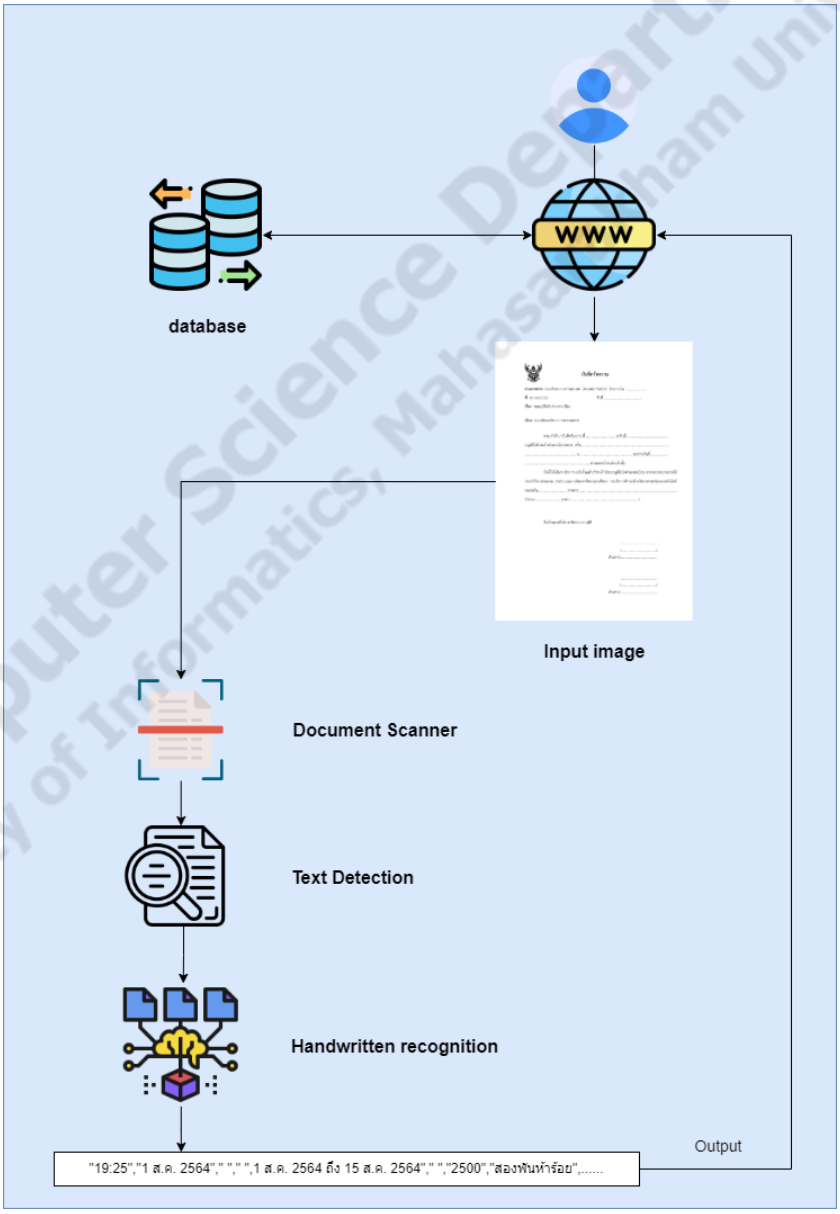
บทที่ 3

วิธีดำเนินงานวิจัย

ในบทนี้จะอธิบายถึงชุดข้อมูลที่ใช้ในโครงงานนี้ และขั้นตอนในการดำเนินงานของการแปลงแบบฟอร์มอัตโนมัติ ดังนี้

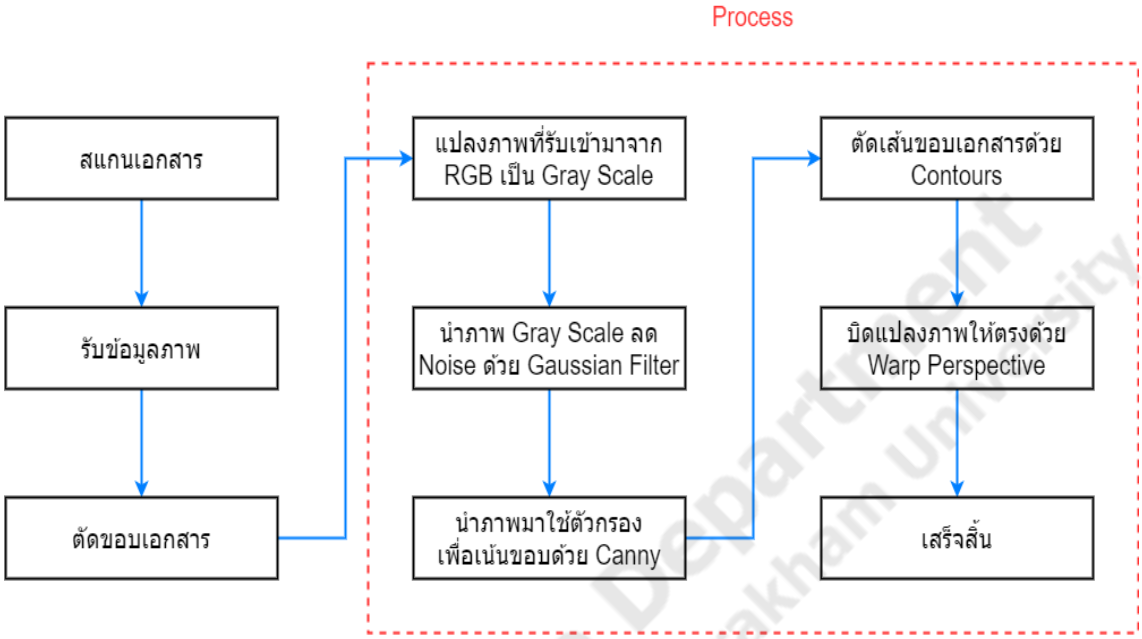
3.1 กรอบการดำเนินงาน



ภาพประกอบที่ 3.1 กรอบการดำเนินงานของระบบ

Computer Science Department
Faculty of Informatics, Mahasarakham University

ส่วนที่ 1: การสแกนหาเอกสาร (Document Scanner)



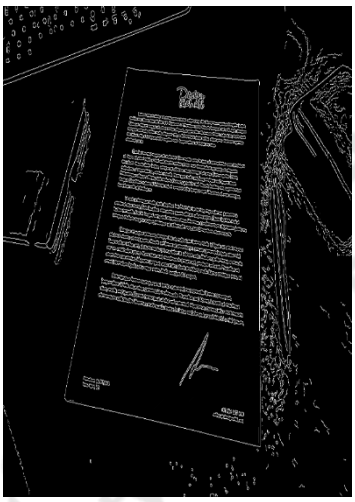
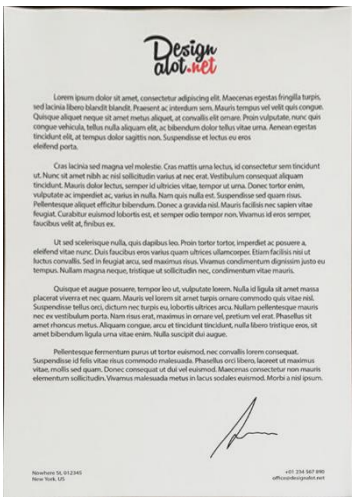
ภาพประกอบที่ 3.2 ขั้นตอนในการทำ Document Scanner

ขั้นตอนในการทำ Document Scanner จะมีรายละเอียดดังนี้

รับข้อมูลภาพเข้ามา (Input) โดยข้อมูลภาพที่รับเข้ามาจะเป็นภาพ RGB หลังจากรับ Input เข้ามาแล้วจะเริ่มทำขั้นตอนในการตัดขอบเอกสารมีการประมวลผม (Process) ดังนี้ เริ่มจากการแปลงข้อมูลภาพที่รับเข้ามาจาก RGB ให้เป็นภาพระดับเทา (Gray Scale) หลังจากได้ภาพระดับเทา (Gray Scale) มาแล้วทำการลดสัญญาณรบกวน (Noise) ด้วยตัวกรองเกาส์เซียน (Gaussian Filter) หลังจากได้ภาพระดับเทา (Gray Scale) ที่ทำการลดสัญญาณรบกวน (Noise) แล้วทำการนำภาพใช้ตัวกรองเพื่อค้นหาหรือเน้นส่วนขอบด้วยตัวกรอง Canny หลังจากใช้ตัวกรอง Canny ในการเน้นส่วนขอบแล้วทำการตัดเส้นขอบของเอกสารด้วยการวาดเส้นเค้าโครงลงบนภาพซึ่งมี Draw Contours และ Find Contours หลังจากได้เค้าโครงเส้นขอบของเอกสารแล้วทำการบิดแปลงภาพให้ตรงด้วย Warp Perspective

Computer Science Department Faculty of Informatics Mahasarakham University

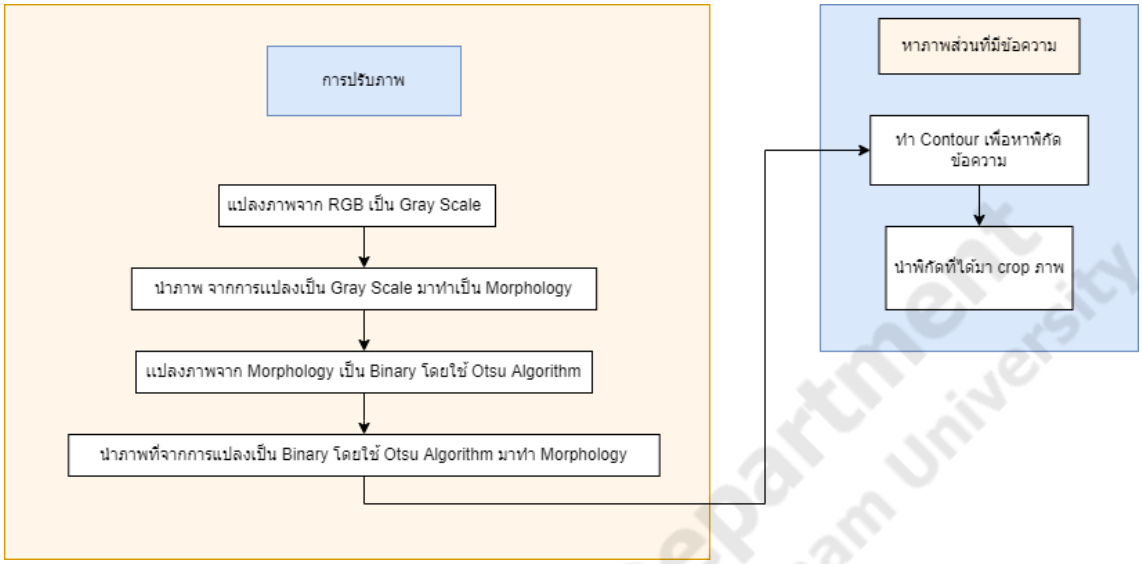
ตารางที่ 3.1 รูปผลลัพธ์ของแต่ละขั้นตอนในการทำ Document Scanner

Input	Gray Scale	Gaussian Filter
Canny	**Contours**	**Perspective**
		

ส่วนที่ 2: การตัดภาพ (Crop)

เมื่อผ่าน การสแกนหาเอกสารในรูปภาพ (Document Scanner) แล้วเราจะทำการตัดเอาภาพส่วนที่เราต้องการเพื่อ นำเข้าไปสู่กระบวนการการ การตรวจจับข้อความ (Text Detection) ในขั้นตอนต่อไป

ส่วนที่ 3: การตรวจจับข้อความ (Text Detection)



ภาพประกอบที่ 3.3 ขั้นตอนการตรวจจับข้อความ

ขั้นตอนนี้คือขั้นตอนการเตรียมภาพก่อนเข้าไปสู่การรู้จำตัวอักษร โดยนำภาพที่ได้จากการ Crop เข้าสู่กระบวนการ Pre-process เพื่อกำจัดแสงเงาและสิ่งรบกวนต่างๆบนภาพที่อาจส่งผลต่อการรู้จำตัวอักษร โดยขั้นตอนแรกของการ Pre-process หลังจากนำไฟล์ภาพเข้ามาแล้วจะนำภาพที่เข้ามามาแปลงภาพ จาก RGB เป็นภาพ Gray Scale เมื่อแปลงภาพมาให้อยู่ในระดับ Gray Scale แล้วจะนำภาพที่ได้มาทำ MORPH_GRADIENT เพื่อเอาเส้นขอบเค้าโครงของภาพ และนำภาพที่ได้จากการทำ MORPH_GRADIENT แล้วนำไปแปลงเป็นภาพสองระดับด้วย Thresholding โดยใช้ Otsu Algorithm จากนั้นนำภาพขาวดำไปทำ.MORPH_CLOSE เพื่อทำให้เส้นขอบเค้าโครงของภาพพองขึ้น และจึงนำภาพนั้นไปหาพิกัดของบรรทัดและนำไป Crop เพื่อเป็น Input ของกระบวนการการรู้จำลายมือ (Handwritten recognition) ต่อไป

ส่วนที่ 4: การรู้จำลายมือเขียน (Handwritten recognition)

ขั้นตอนนี้คือการทำรู้จำตัวอักษร โดยภาพที่ถูกนำเข้าจะต้องเข้าสู่กระบวนการทำ Pre-processing ก่อนเพื่อเตรียมภาพก่อนเข้าสู่ Model ที่เป็นกระบวนการในการรู้จำ จึงได้ผลลัพธ์ออกมา

3.2 การสแกนหาเอกสาร (Document Scanner)

ขั้นตอนในการทำ Document Scanner มีรายละเอียดดังนี้

Computer Science Department
Faculty of Science, Maha Sarakham University

3.2.1 การรับข้อมูลภาพเข้ามา (Input Image)

การรับข้อมูลภาพเข้ามาโดยข้อมูลภาพที่รับเข้ามาจะอยู่ในรูปแบบไฟล์ภาพ PNG, JPG, BMP, TIFF เป็นต้นโดยข้อมูลภาพที่รับเข้ามาจะเป็นภาพสี RGB

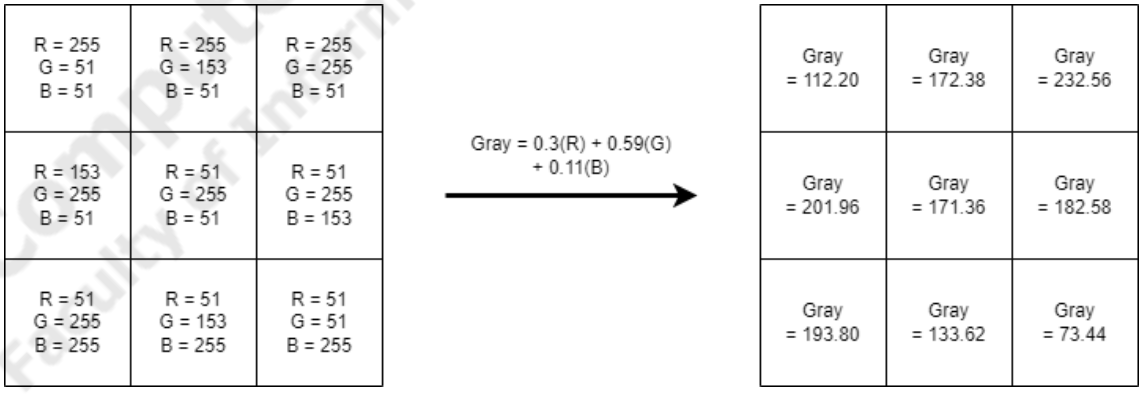
3.2.2 การแปลงภาพสี RGB ไปเป็นภาพระดับเทา (Gray Scale)

ภาพระดับสีเทา (Gray Scale) คือภาพขาว - ดำ - เทา โดยจะมีระดับความเข้มจากการแปลงสีเท่าคือ 0–255 (8-bit)

ภาพสี RGB เป็น Gray Scale สมการทางคณิตศาสตร์คือ

$$\text{Gray} = 0.3(\text{R}) + 0.59(\text{G}) + 0.11(\text{B}) \quad (3.1)$$

Gray คือ ค่าความเข้มของสีเทามีค่า 0–255
R(Red) คือ ค่าความเข้มของสีแดงมีค่า 0–255
G(Green) คือ ค่าความเข้มของสีเขียวมีค่า 0–255
B(Blue) คือ ค่าความเข้มของสีน้ำเงินมีค่า 0–255
ตัวอย่างสมการของการแปลงภาพ RGB เป็น Gray Scale



ภาพประกอบที่ 3.4 ก่อนและหลังจากค่า RGB ไปเป็น Gray Scale

จากตัวอย่างข้างบนจะเห็นได้ว่าค่าสีตรงจุด Pixel นั้นๆ จะมีค่า RGB ตามที่แสดงหลังจากใช้สมการแปลงเป็น Gray Scale จะได้ค่า RGB ใหม่เช่น

Computer Science Department
Faculty of Informatics, Mahasarakham University

Gray(1, 0)

R = 255, G = 153, B = 51

Gray = 0.3(255) + 0.59(153) + 0.11(51)

Gray = 172.38

Gray(1, 1)

R = 51, G = 255 , B = 51

Gray = 0.3(51) + 0.59(255) + 0.11(51)

Gray = 171.36

ภาพประกอบที่ 3.5 ข้อมูลภาพที่รับเข้ามาเป็น RGB

ภาพประกอบที่ 3.6 ภาพหลังจากแปลง RGB เป็นภาพระดับเทา

Computer Science Department
Faculty of Informatics Mahasarakham University

3.2.3 การลดสัญญาณรบกวน (Noise) ภาพระดับเทาด้วยตัวกรองเกาส์เซียน (Gaussian Filter)

เป็นตัวกรองสัญญาณรบกวนที่มีคุณลักษณะคล้ายระฆังคว่ำ ใช้สำหรับลดสัญญาณรบกวน (Noise) และลบความคมชัดของรูปภาพ

สมการทางคณิตศาสตร์ของตัวกรองเกาส์เซียน (Gaussian Filter)
ตัวกรองเกาส์เซียนแบบ 1 มิติ มีสมการดังนี้

$$G(x) = \sqrt{\frac{a}{x}} * \ e^{-a*x^2} \qquad (3.2)$$

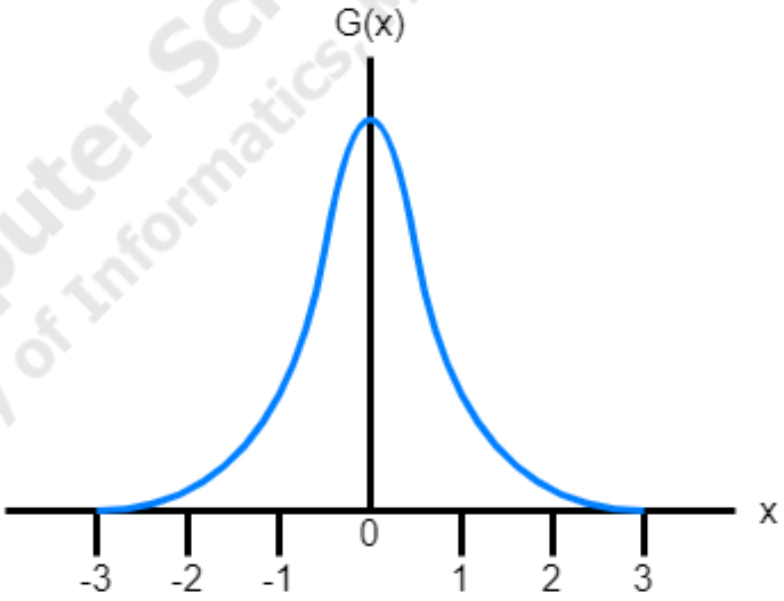
หรือเขียนได้จากการใช้พารามิเตอร์ส่วนเบี่ยงเบน คือ

$$G(x) = \sqrt{\frac{a}{x}} * \ e^{-a*x^2} \qquad (3.3)$$

เมื่อ x คือ ค่าตัวแปรในแกน x

σ คือ ค่าส่วนเบี่ยงเบนมาตรฐาน

ตัวอย่างการแทนค่าสมการเกาส์เซียนแบบ 1 มิติ โดยแทนค่า x = 3 จะได้ดังรูป



ภาพประกอบที่ 3.7 ตัวกรองสัญญาณรบกวนเกาส์เซียนแบบ 1 มิติ

$$\mathrm{G(x,y)} \ = \ \frac{1}{2\pi\sigma^2} \ e^{-\frac{(x^2+y^2)}{2\sigma^2}} \qquad (3.4)$$

Computer Science Department
Faculty of Informatics, Mahasarakham University

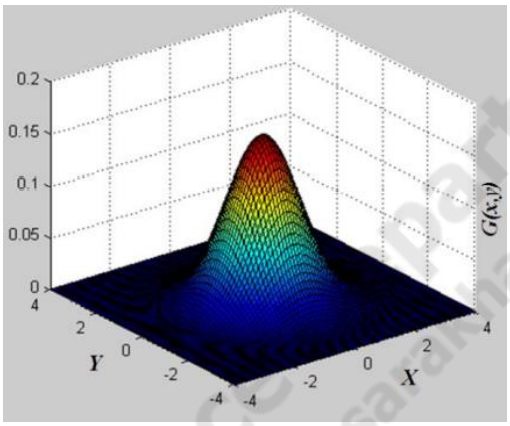
เมื่อ $2\pi\sigma^2$ คือ ผลรวมของตัวกรองเกาส์เซียน

x คือ ค่าตัวแปรในแกน x

y คือ ค่าตัวแปรในแกน y

σคือ ค่าส่วนเบี่ยงเบนมาตรฐาน

ตัวอย่างการแทนค่าสมการเกาส์เซียนแบบ 2 มิติ โดยแทนค่า x = -4, y = 4 แล้วเมื่อทำการ plot กราฟจะได้ดังรูป



ภาพประกอบที่ 3.8 ตัวกรองสัญญาณรบกวนเกาส์เซียนแบบ 2 มิติ

ภาพประกอบที่ 3.9 ภาพระดับเทา

Computer Science Department
Faculty of Informatics
Mahasarakham University

ภาพประกอบที่ 3.10 ภาพระดับเทาหลังใช้ตัวกรองสัญญาณรบกวนเกาส์เซียน

3.2.4 การทำตรวจจับขอบ (Edge Detection) โดยใช้ตัวกรอง Canny

Canny Edge Detection เป็นการหาเส้นรอบวัตถุที่อยู่ในภาพ โดยจะทำการแปลงภาพระดับเทา (Gray Scale) เป็นภาพไบนารี (Binary Image) ของเส้นรอบวัตถุ และมีอัลกอริทึมหลายขั้นตอนในการหาเส้นรอบวัตถุเนื่องจากขอบภาพเกิดจากความแตกต่างของความเข้มแสงจากจุดหนึ่งไปยังอีกจุดหนึ่ง หากความต่างนี้มีค่ามากจะทำให้เห็นขอบภาพได้อย่างชัดเจน แต่ถ้าความต่างนี้มีค่าน้อยจะทำให้เห็นขอบภาพได้ไม่ชัดเจน

สมการสำหรับใช้ตัวกรองเป็นดังนี้

$$\text{Edge_Gradient(G)} = \sqrt{G_x^2 + G_y^2} \quad \textbf{(3.5)}$$

$$\text{Angle}(\theta) = tan^{-1}(\frac{G_y}{G_x}) \quad \textbf{(3.6)}$$

โดยที่

Gx คือ จุดมีการประมาณอนุพันธ์ในแนวนอน

Gy คือ จุดมีการประมาณอนุพันธ์ในแนวตั้ง

Computer Science Department
Faculty of Informatics, Mahasarakham University

ชนิดของ Sobel Mask ขนาด 3x3
ตัวทางซ้ายคือแกน X และทางขวาคือแกน Y

1/4

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

1/4

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

ภาพประกอบที่ 3.11 Sobel Mask ขนาด 3x3 แกน X และแกน Y

ตัวอย่างการคำนวณสมการ

1	2	3
4	5	6
7	8	9

หา Gradient ที่ตำแหน่ง Pixel (2, 2)

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

Gx Mask

หา Gx = 1/4 [(-1*1) + (-2*4) + (-1*7) + (1*3) + (2*6) + (1*9)]

Gx = 2

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

Gy Mask

หา Gy = 1/4 [(-1*1) + (-2*2) + (-1*3) + (1*7) + (2*8) + (1*9)]

Gy = 6

ภาพประกอบที่ 3.12 การคำนวณหา Gradient และ Gx Mask กับ Gy Mask

Computer Science Department
Faculty of Informatics, Mahasarakham University

จะได้ว่า

$$G(2,2) = \begin{bmatrix} Gx \\ Gy \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$$

ดังนั้น Gradient Direction เท่ากับ

$$\text{Angle}(2,2) = tan^{-1}\left(\frac{G_y}{G_x}\right) = tan^{-1}\left(\frac{6}{2}\right) = 71.51$$

ดังนั้น Edge Magnitude เท่ากับ

$$\text{Edge_Gradient(G)} = \sqrt{G_x^2 + G_y^2} = \sqrt{2^2 + 6^2} = \sqrt{40}$$

ในที่นี้ Threshold1 และ Threshold2 คือค่าจุดเปลี่ยนสำหรับคัดกรองว่าตรงไหนควรจะเป็นขอบโดยภายในกระบวนการของวิธีนี้จะมีการคำนวณค่าความน่าจะเป็นขอบของแต่ละจุดอยู่จากนั้นค่านั้นจะถูกนำมาพิจารณาโดยดูค่ากรองต่ำสุดและสูงสุดดังนี้

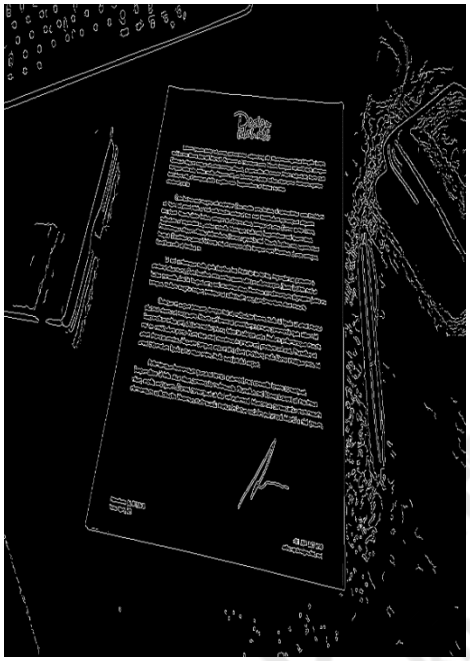
ตารางที่ 3.2 จุดต่ำกว่าและสูงกว่าของค่ากรองที่ตั้งไว้

<table>
<tr><td colspan="2">จุดที่ค่าต่ำกว่าค่ากรองต่ำสุด</td><td>0 (ไม่ใช่ขอบ)</td></tr>
<tr><td rowspan="2">จุดที่ค่าอยู่ระหว่างค่ากรองต่ำสุดและสูงสุด</td><td>ถ้าไม่มีจุดที่ค่าสูงกว่าค่ากรองสูงสุดอยู่ข้างๆ</td><td rowspan="2">255 (เป็นขอบ)</td></tr>
<tr><td>ถ้ามีจุดที่ค่าสูงกว่าค่ากรองสูงสุดอยู่ข้างๆ</td></tr>
<tr><td colspan="2">จุดที่ค่าสูงกว่าค่ากรองสูงสุด</td></tr>
</table>

Threshold1 และ Threshold2 ที่ใส่ลงไปนี้ตัวหนึ่งจะเป็นค่ากรองสูงสุดอีกตัวเป็นค่ากรองต่ำสุดจะใส่ตัวไหนเป็นค่ามากกว่าก็ได้ ใส่สลับกันก็ไม่มีผลต่างกัน

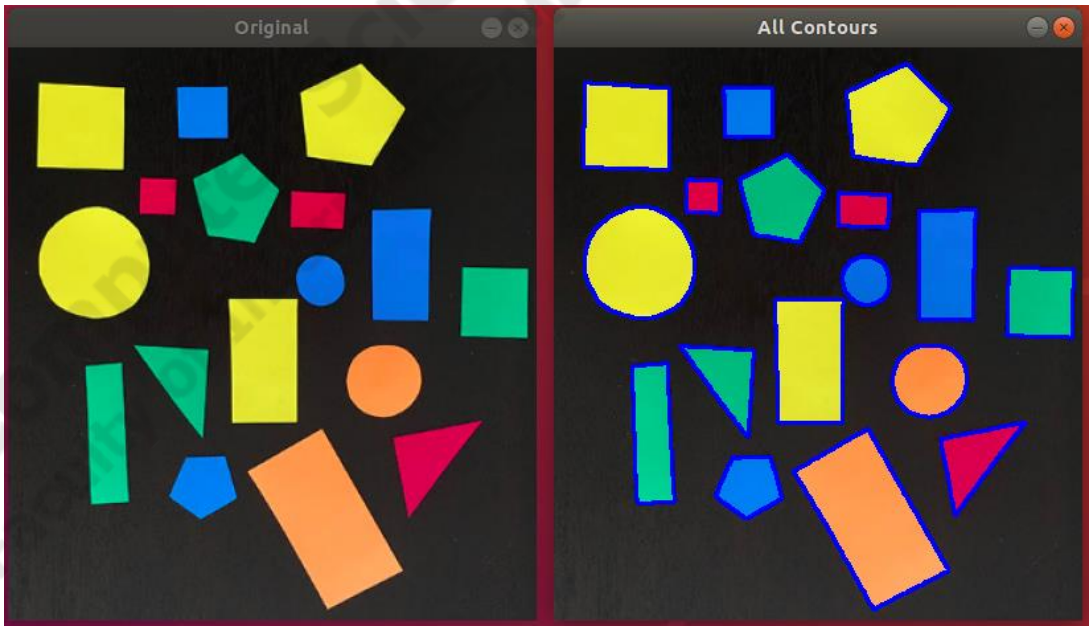
ภาพประกอบที่ 3.13 ภาพระดับเทาหลังใช้ตัวกรองสัญญาณรบกวนเกาส์เซียน

Computer Science Department
Faculty of Informatics, Mahasarakham University



ภาพประกอบที่ 3.14 ภาพระดับเทาหลังใช้ตัวกรอง Canny

3.2.5 การหาเส้นเค้าโครง (Contours) ของเอกสาร



ภาพประกอบที่ 3.15 การทำงานของ Draw Contours

วิธีดำเนินการ

กำหนดตัวแปรแต่ละตัว

f_{iy} คือ ค่าของพิกเซลที่ตำแหน่ง (i,j)

i คือ ค่าของตำแหน่งพิกเซลในแนวแกน x

y คือ ค่าของตำแหน่งพิกเซลในแนวแกน y

เงื่อนไขของการ Contour คือจะกำหนดหมายเลขเฉพาะให้กับเส้นขอบใหม่

เงื่อนไขที่ 1 ให้นำ N เข้า และตั้งค่า i2,y2

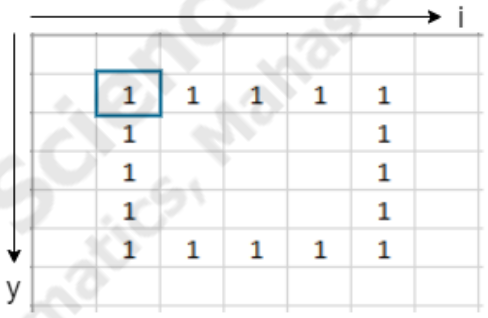
else if เส้นขอบด้านในของวัตถุ

ให้นำ N เข้า และตั้งค่า i2.y2 เป็น i,y+1

else if $f_{iy} > 1$

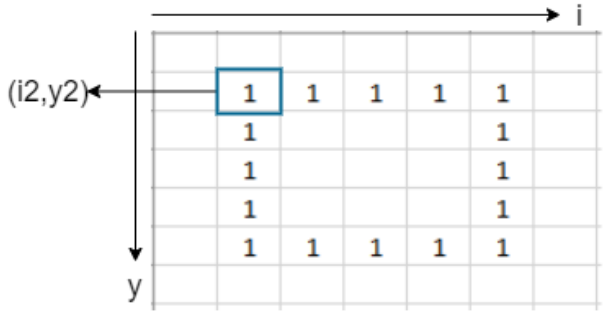
ให้ N = N

ขั้นตอนที่ 1 เริ่มสแกนรูปภาพจากซ้ายไปขวาจนกว่าจะพบพิกเซลวัตถุ โดยหากค่าพิกเซลที่พบมีค่ามากกว่าหรือเท่ากับ 1 แปลว่าพิกเซลนี้อาจเป็นขอบภาพของวัตถุดังภาพ



ภาพประกอบที่ 3.16 ตัวอย่างการ Contour ขั้นตอนที่ 1

ขั้นตอนที่ 2 หลังจากได้จุดที่น่าจะเป็นขอบภาพแล้วจะทำตามขั้นตอนดังเงื่อนไขที่ 1



ภาพประกอบที่ 3.17 ตัวอย่างการ Contour ขั้นตอนที่ 2

Computer Science Department
Faculty of Informatics, Mahasarakham University

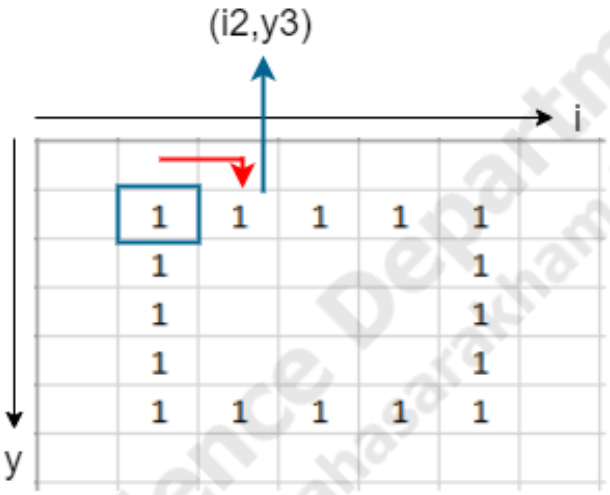
ขั้นตอนที่ 3 เริ่มหาเส้นขอบจาก (i2,y2) โดยจะดูรอบๆตามเข็มนาฬิกาของพิกเซลในพื้นที่ใกล้เคียงของ (i,y) โดยมีเงื่อนไขการหาดังนี้

เงื่อนไขที่ 2 if pixel != 0

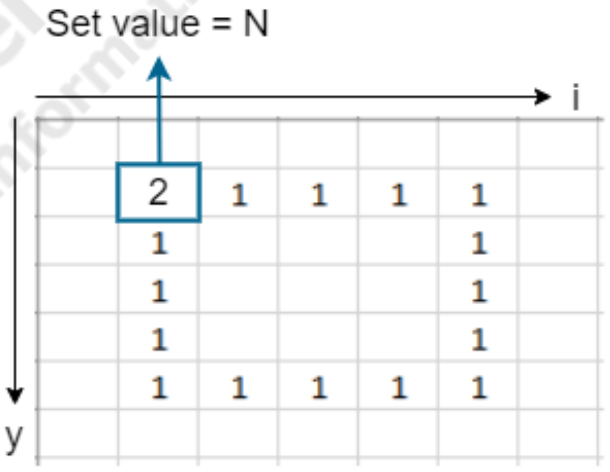
Set ค่า N = 2 และ จะเปลี่ยนตำแหน่งปัจจุบันเป็น (i2,y2+1)

else if หากพบ pixel ที่ไม่ใช่ 1

ให้ $f_{iy} = -2$



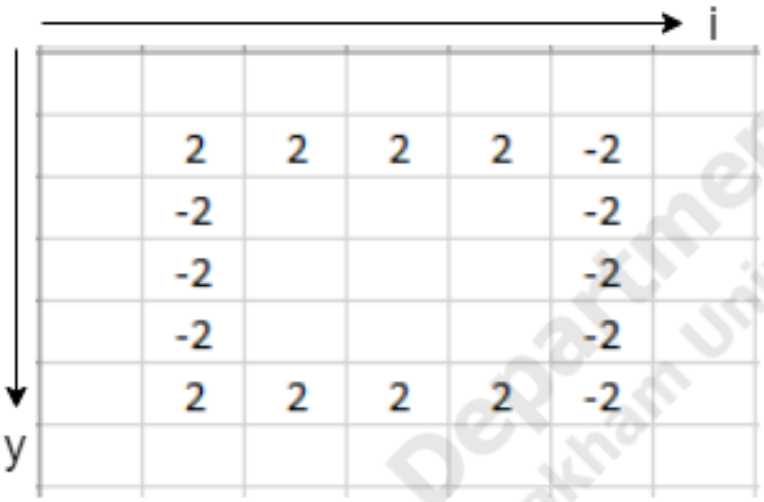
ภาพประกอบที่ 3.18 ตัวอย่างการ Contour ขั้นตอนที่ 3.1



ภาพประกอบที่ 3.19 ตัวอย่างการ Contour ขั้นตอนที่ 3.2

Computer Science Department
Faculty of Informatics, Mahasarakham University

เงื่อนไขที่ 3 if $f_{iy} != 1$ ให้ตั้งค่า |N $= f_{iy}$| และเริ่มสแกนพิกเซลถัดไป $|f_{iy+1}|$ โดยจะทำซ้ำขั้นตอนข้างต้นไปเรื่อยๆจนครบทุกพิกเซลและทำการหยุดเมื่อถึงมุมขวาล่างของภาพซึ่งจะแสดงผลลัพธ์ดังภาพประกอบที่ 3.21



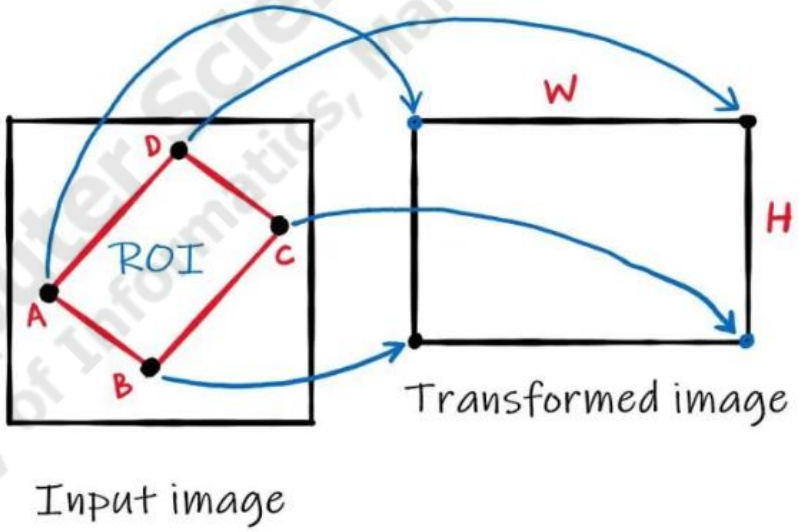
ภาพประกอบที่ 3.20 ภาพผลลัพธ์หลังการทำ Contour จนครบ

ภาพประกอบที่ 3.21 ภาพตัวอย่างก่อนการทำ Contour

Computer Science Department
Faculty of Informatics Mahasarakham University

ภาพประกอบที่ 3.22 ภาพตัวอย่างหลังการทำ Contour

3.2.6 การบิดแปลงภาพให้ตรงด้วย Warp Perspective



ภาพประกอบที่ 3.23 การทำงานของ Warp Perspective

ขั้นตอนการทำงาน

การทำงานของ Warp Perspective การใช้จะคล้ายกับ Warp Affine แต่จะเป็นการคูณด้วยอาเรย์ขนาด 3×3 แทน ซึ่งจะทำให้แปลงรูปได้หลากหลายขึ้นกว่าเดิมสมการของ Warp Perspective ถ้ามีจุดพิกัดอยู่จุดหนึ่งในรูปภาพ 2 มิติ พิกัด (x1, y1) อาจจะเขียนเป็นแมทริกซ์ได้ว่า

Computer Science Department
Faculty of Informatics, Mahasarakham University

$$\begin{bmatrix} x1 \\ y1 \\ 1 \end{bmatrix}$$

ภาพประกอบที่ 3.24 เมทริกซ์ 2 มิติของพิกัดในรูปภาพ

โดยที่ 1 นี้เป็นแกน z ซึ่งจะให้เป็น 1 เสมอ เพราะจริงๆแล้วภาพมีแค่ ๒ มิติ ไม่ได้ต้องใช้มิติที่ ๓ แต่เพิ่มเข้ามาเพื่อใช้แทนมิติที่เป็นค่าคงที่

ถ้ามี n จุดก็เอาพิกัดของแต่ละจุดมาเขียนรวมกันเป็นเมทริกซ์แบบนี้

$$X = \begin{bmatrix} x1 & x2 & \dots & xn \\ y1 & y2 & \dots & yn \\ 1 & 1 & \dots & 1 \end{bmatrix}$$

ภาพประกอบที่ 3.25 เมทริกซ์พิกัดของรูปภาพที่รับเข้ามา

เมทริกซ์ที่ใช้บิดแปลงภาพด้วย Warp Perspective คือเมทริกซ์ขนาด 3×3

$$M = \begin{bmatrix} ax & bx & cx \\ by & by & cy \\ 1 & 1 & 1 \end{bmatrix}$$

ภาพประกอบที่ 3.26 เมทริกซ์บิดแปลงรูปภาพขนาด 3×3

เมทริกซ์ของพิกัดที่จะได้หลังแปลงนั้นจะได้มาจากการเอาเมทริกซ์การแปลงมาคูณเข้ากับเมทริกซ์พิกัดเดิม

$$Xnew = M * X$$

$$= \begin{bmatrix} ax & bx & cx \\ by & by & cy \\ 1 & 1 & 1 \end{bmatrix} * \begin{bmatrix} x1 & x2 & \dots & xn \\ y1 & y2 & \dots & yn \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} ax(x1) + bx(y1) + cx(1) & ax(x2) + bx(y2) + cx(1) & \dots & ax(xn) + bx(yn) + cx(1) \\ ay(x1) + by(y1) + cy(1) & ay(x2) + by(y2) + cy(1) & \dots & ay(xn) + by(yn) + cy(1) \\ 1 & 1 & & 1 \end{bmatrix}$$

ภาพประกอบที่ 3.27 การคำนวณเมทริกซ์ของพิกัดใหม่

ก็จะได้เมทริกซ์ของพิกัดใหม่

Computer Science Department
Faculty of Informatics Mahasarakham University

$$Xnew = \begin{bmatrix} x1new & x2new & \dots & xn(new) \\ y1new & y2new & \dots & yn(new) \\ 1 & 1 & \dots & 1 \end{bmatrix}$$

ภาพประกอบที่ 3.28 เมทริกซ์ของพิกัดใหม่

ค่าพิกัดใหม่ในแต่ละจุดเป็น

$$x_{i^*} = a_x x_i + b_x y_i + cx \qquad (3.7)$$
$$y_{i^*} = a_y x_i + b_y y_i + cx$$

ตัวอย่างการบิดแปลงเมทริกซ์

$$X = \begin{bmatrix} 4 & 7 & 6 & 3 \\ 1 & 2 & 4 & 3 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$M = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix}$$

Xnew = M * X

$$= \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix} * \begin{bmatrix} 4 & 7 & 6 & 3 \\ 1 & 2 & 4 & 3 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

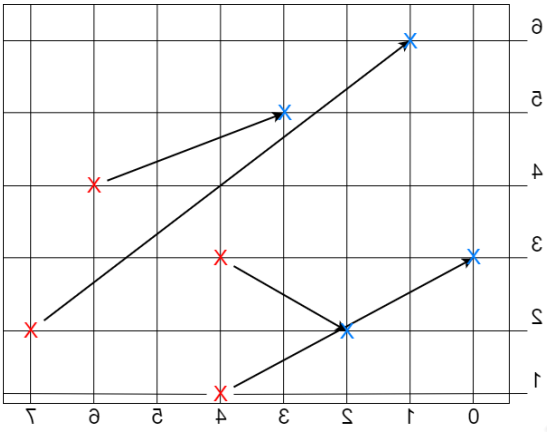
$$= \begin{bmatrix} (0\times4+1\times1+-1\times1) & (0\times7+1\times2+-1\times1) & (0\times6+1\times4+-1\times1) & (0\times3+1\times3+-1\times1) \\ (1\times4+0\times1+-1\times1) & (1\times7+0\times2+-1\times1) & (1\times6+0\times4+-1\times1) & (1\times3+0\times3+-1\times1) \\ (1\times4+1\times1+1\times1) & (1\times7+1\times2+1\times1) & (1\times6+1\times4+1\times1) & (1\times3+1\times3+1\times1) \end{bmatrix}$$

$$Xnew = \begin{bmatrix} 0 & 1 & 3 & 2 \\ 3 & 6 & 5 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

ภาพประกอบที่ 3.29 ตัวอย่างการคำนวณหาเมทริกซ์พิกัดใหม่

การใช้ Warp Perspective จะเป็นการย้ายจุดทุกจุดพิกเซลบนภาพจากพิกัดเก่าไปพิกัดใหม่

Computer Science Department
Faculty of Informatics, Mahasarakham University



ภาพประกอบที่ 3.30 ตัวอย่างการแปลงพิกัดเก่าไปยังพิกัดใหม่

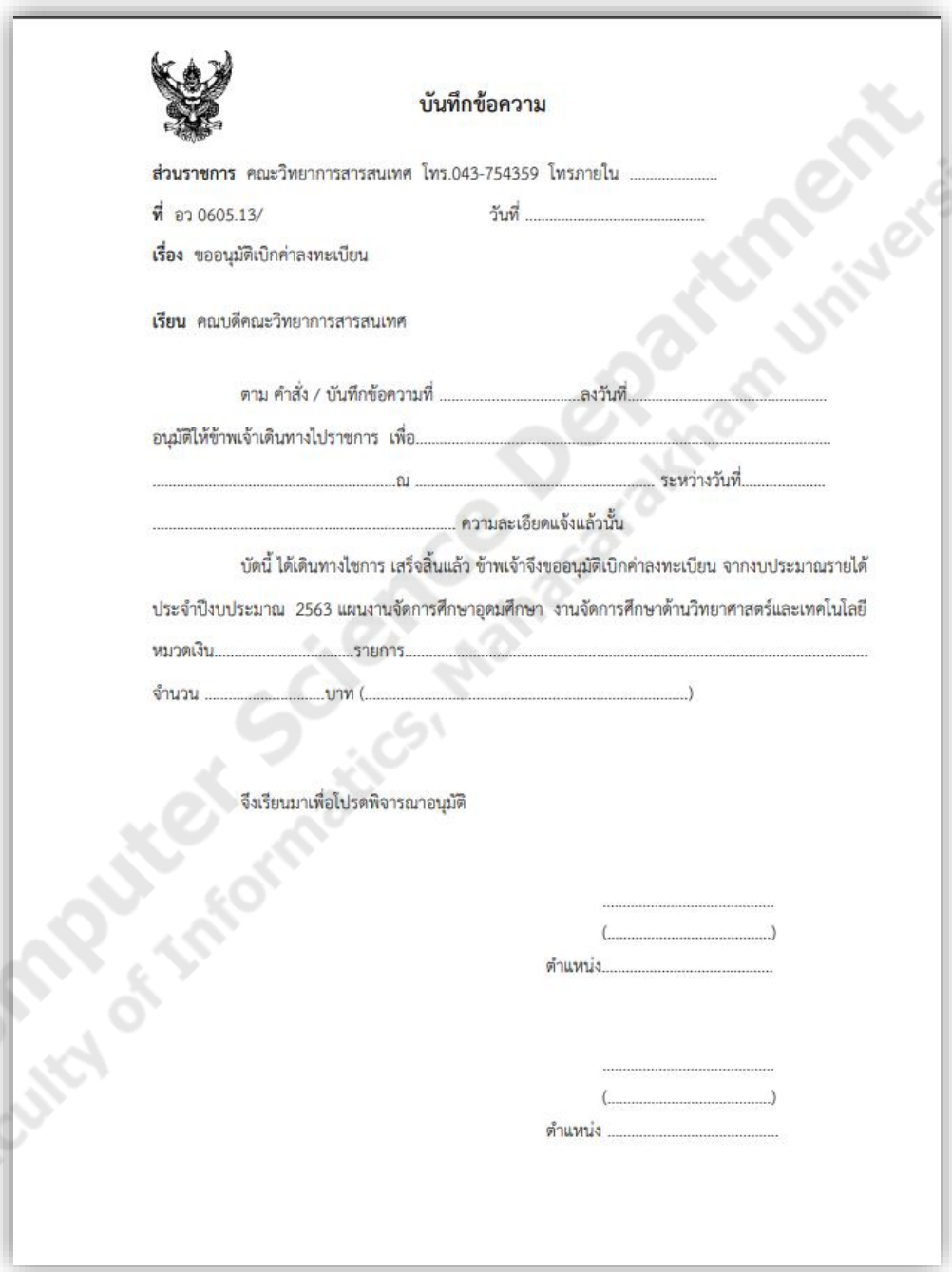
ภาพประกอบที่ 3.31 ภาพตัวอย่างก่อนการทำ Warp Perspective

ภาพประกอบที่ 3.32 ภาพตัวอย่างหลังการทำ Warp Perspective

Computer Science Department
Faculty of Informatics, Mahasarakham University

3.3 การตรวจจับข้อความ (Text Detection)

ภาพที่นำเข้า เป็นภาพสี RGB ที่ได้มาจากการ การสแกนหาเอกสาร (Document Scanner)



บันทึกข้อความ

ส่วนราชการ คณะวิทยาการสารสนเทศ โทร.043-754359 โทรภายใน .....

ที่ อว 0605.13/ วันที่ .....

เรื่อง ขออนุมัติเบิกค่าลงทะเบียน

เรียน คณบดีคณะวิทยาการสารสนเทศ

ตาม คำสั่ง / บันทึกข้อความที่ .....ลงวันที่.....

อนุมัติให้ข้าพเจ้าเดินทางไปราชการ เพื่อ.....

.....ณ ..... ระหว่างวันที่.....

..... ความละเอียดแจ้งแล้วนั้น

บัดนี้ ได้เดินทางไชการ เสร็จสิ้นแล้ว ข้าพเจ้าจึงขออนุมัติเบิกค่าลงทะเบียน จากงบประมาณรายได้ ประจำปีงบประมาณ 2563 แผนงานจัดการศึกษาอุดมศึกษา งานจัดการศึกษาด้านวิทยาศาสตร์และเทคโนโลยี หมวดเงิน.....รายการ.....

จำนวน .....บาท (.....)

จึงเรียนมาเพื่อโปรดพิจารณาอนุมัติ

.....

(.....)

ตำแหน่ง.....

.....

(.....)

ตำแหน่ง .....

Computer Science Department
Faculty of Informatics, Mahasarakham University

ภาพประกอบที่ 3.33 ตัวอย่างภาพที่ผ่านการสแกน

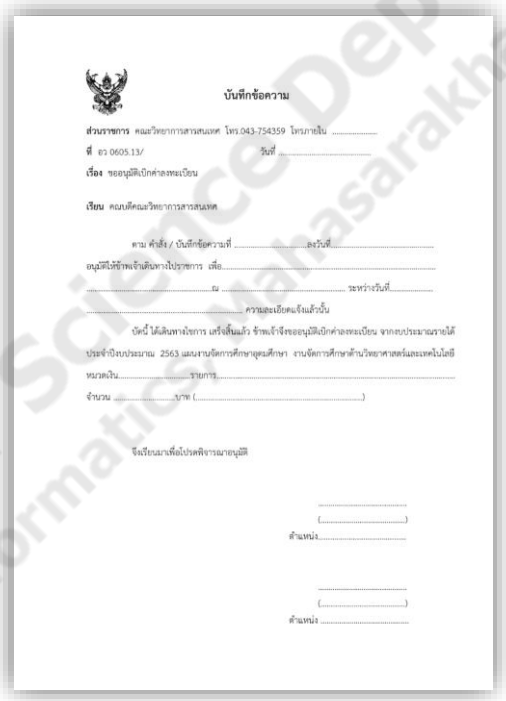
3.3.1 การเตรียมพร้อมภาพ (Pre-Processing)

โดยจะประกอบไปด้วยขั้นตอนในการการแปลงภาพให้เป็นระดับเทาแปลงให้เป็นไบนารี ขาว-ดำ และตัดรูปภาพเฉพาะตรงส่วนที่มีข้อความอยู่ โดยขั้นตอนและวิธีการต่างๆสามารถอธิบายได้ดังต่อไปนี้

การหาพิกัดที่ต้องการตัด

เป็นการหาค่าความกว้างและความสูงเพื่อระบุตำแหน่งที่จะเอามาตัด crop รูปภาพให้เหลือเฉพาะส่วนที่ต้องการ โดยมีขั้นตอนและวิธีดำเนินการ ดังนี้

ขั้นตอนที่ 1 เมื่อภาพที่ Input มาทำการ reshape จะได้ค่าความกว้างและความสูง ซึ่งจากภาพประกอบที่ 3.34 จะมีค่าความกว้างและความสูงคือ (1357,1920)



บันทึกข้อความ

ส่วนราชการ คณะวิทยาการสารสนเทศ โทร.043-754359 โทรภายใน .....

ที่ อว 0605.13/ วันที่ .....

เรื่อง ขออนุมัติเบิกค่าลงทะเบียน

เรียน คณบดีคณะวิทยาการสารสนเทศ

ตาม คำสั่ง / บันทึกข้อความที่ .....ลงวันที่.....

อนุมัติให้ข้าพเจ้าเดินทางไปราชการ เพื่อ.....

ณ ..... ระหว่างวันที่.....

..... ความละเอียดแล้วนั้น

บัดนี้ ได้เดินทางไปราชการ เสร็จสิ้นแล้ว ข้าพเจ้าจึงขออนุมัติเบิกค่าลงทะเบียน จากงบประมาณรายได้ประจำปีงบประมาณ 2563 แผนงานจัดการศึกษาอุดมศึกษา งานจัดการศึกษาด้านวิทยาศาสตร์และเทคโนโลยี หมวดเงิน.....รายการ.....

จำนวน .....บาท (.....)

จึงเรียนมาเพื่อโปรดพิจารณาอนุมัติ

.....

(.....)

ตำแหน่ง.....

.....

(.....)

ตำแหน่ง .....

ภาพประกอบที่ 3.34 ตัวอย่างเอกสารหลัง reshape

ขั้นตอนที่ 2 หลังจากได้ค่าความกว้างและความสูงแล้ว เราจะทำการระบุตำแหน่งที่ต้องการตัด

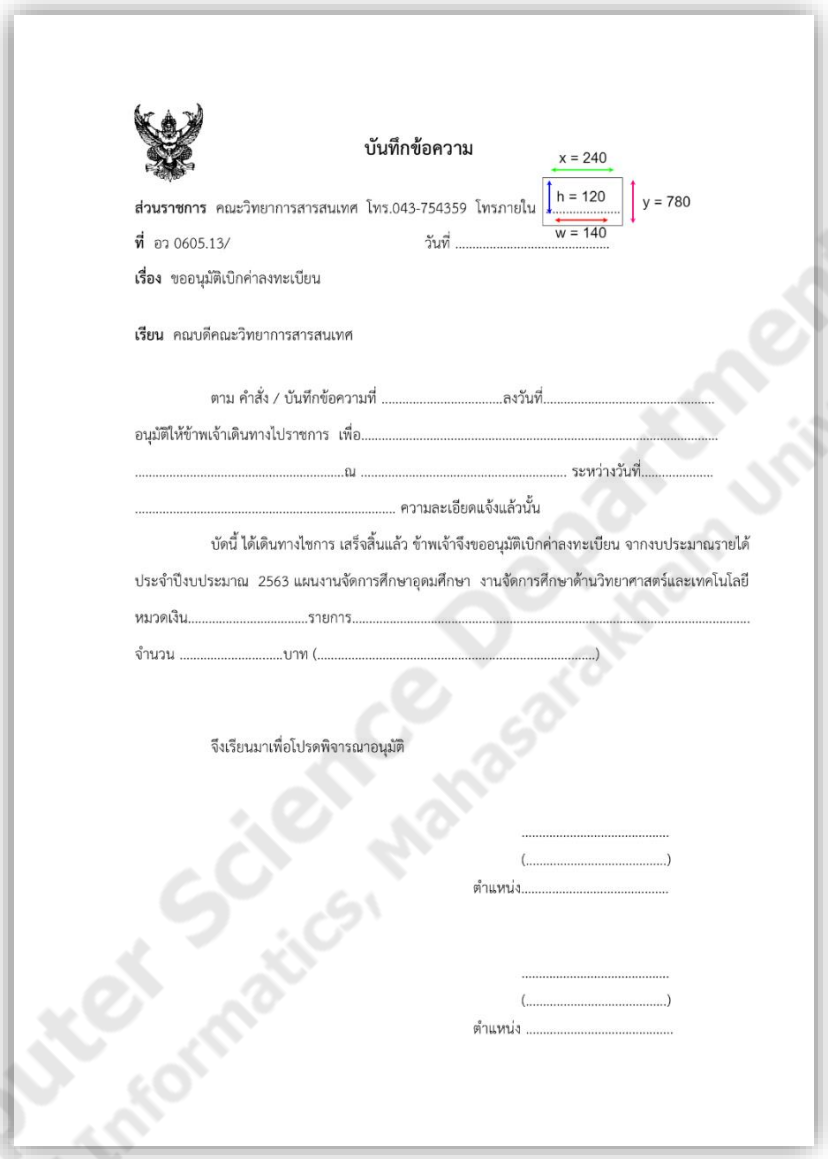
x คือ ตำแหน่งแนวนอนบนรูปภาพ

y คือ ตำแหน่งแนวตั้งบนรูปภาพ

w คือ ความกว้างของกรอบที่เราต้องการ crop

h คือ ความสูงของกรอบที่เราต้องการ crop

Computer Science Department Faculty of Informatics Mahasarakham University



บันทึกข้อความ

x = 240

ส่วนราชการ คณะวิทยาการสารสนเทศ โทร.043-754359 โทรภายใน h = 120 y = 780

w = 140

ที่ อว 0605.13/ วันที่ .....

เรื่อง ขออนุมัติเบิกค่าลงทะเบียน

เรียน คณบดีคณะวิทยาการสารสนเทศ

ตาม คำสั่ง / บันทึกข้อความที่ .....ลงวันที่.....

อนุมัติให้ข้าพเจ้าเดินทางไปราชการ เพื่อ.....

.....ณ ..... ระหว่างวันที่.....

..... ความละเอียดแจ้งแล้วนั้น

บัดนี้ ได้เดินทางไชการ เสร็จสิ้นแล้ว ข้าพเจ้าจึงขออนุมัติเบิกค่าลงทะเบียน จากงบประมาณรายได้ประจำปีงบประมาณ 2563 แผนงานจัดการศึกษาอุดมศึกษา งานจัดการศึกษาด้านวิทยาศาสตร์และเทคโนโลยี หมวดเงิน.....รายการ.....

จำนวน .....บาท (.....)

จึงเรียนมาเพื่อโปรดพิจารณาอนุมัติ

.....

(.....)

ตำแหน่ง.....

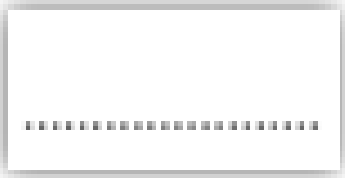
.....

(.....)

ตำแหน่ง .....

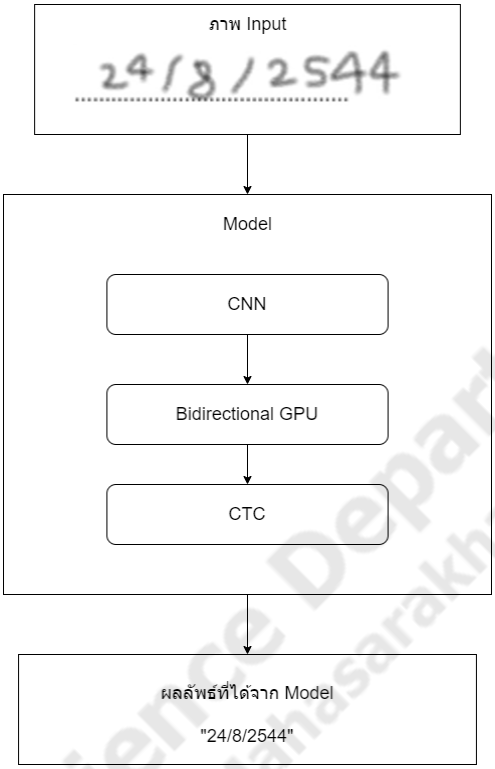
Computer Science Department
Faculty of Informatics, Mahasarakham University

ภาพประกอบที่ 3.35 ตัวอย่างการระบุตำแหน่งที่ต้องการตัด



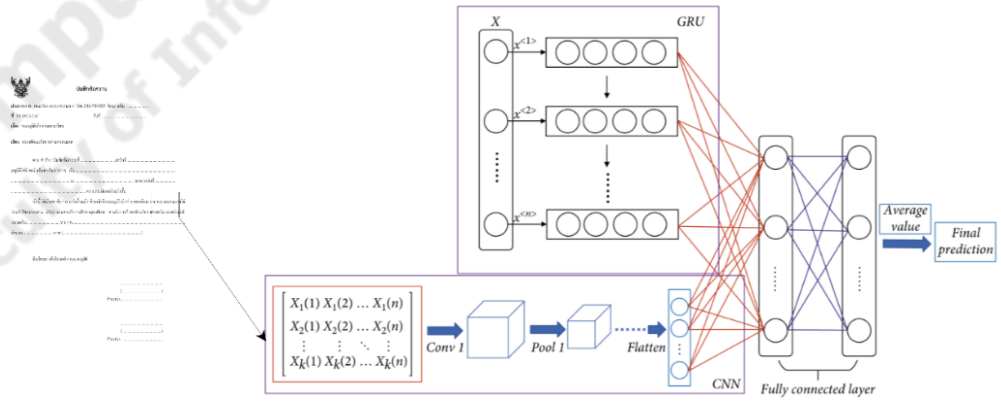
ภาพประกอบที่ 3.36 ผลลัพธ์การตัด crop

3.4 การรู้จำลายมือ (Handwritten recognition)



ภาพประกอบที่ 3.37 ขั้นตอนการทำงานของการรู้จำลายมือ

เมื่อผ่านกระบวนการ Pre-Processing เสร็จแล้วผลลัพธ์ที่ได้จากขั้นตอน Pre-Processing จะถูกส่งเข้าสู่กระบวนการรู้จำตัวลายมือ โดยจะมีรายละเอียดการทำงาน ดังนี้



ภาพประกอบที่ 3.38 Flow การทำงานการรู้จำลายมือ

ที่มา: https://www.hindawi.com/journals/mpe/2020/1428104/

Computer Science Department
Faculty of Informatics, Mahasarakham University

เมื่อทำการ input ภาพที่ต้องการเข้ามา

Convolution Neural Network (CNN) จะทำการนำบางส่วนของภาพที่ input เข้ามาทำที่ละส่วนๆ โดนเมื่อนำภาพบ้างส่วนที่เข้ามาแล้วทำการทำ Convolution โดนนำค่าบนภาพมาทำ Convolution จากนั้นเมื่อได้ค่าที่ทำ Convolution นำมาทำ Max Polling จะจากนั้นจะทำการ dropout เพื่อลดจำนวน Node ออกตามที่เรากำหนด เพื่อลดการเกิด Over Fitting ของข้อมูล โดยค่าที่ใช้กำหนด Dropout เช่น 0.2 หมายความว่า 20% ของจำนวน Node ทั้งหมดจะไม่มีการ Update ทำให้ช่วยลดการเกิด Over Fitting ของข้อมูล และ จะทำการ Reshape รูปภาพ โดยจะส่งข้อมูลเข้าทีละ Time-Step เอามาต่อกันเป็น 1 แถวยาวจากนั้นจะส่งให้ข้อมูลให้

Bidirectional GRU เพื่อทำการประมวลผล ผลลัพธ์ที่ได้จาก CNN จะส่งข้อมูลเข้า GRU ครั้งละ 1 Time Step เข้าสู่ Algorithm GRU จนครบทุก Time Step ซึ่งการ ทำงานของ GRU สามารถอธิบายได้ดังภาพประกอบที่ 3.50 ขั้นตอน

Fully Connected เป็นชั้นสุดท้ายของ CNN ที่เราทำการตัดออกมาทำการต่อเข้ากับ Bidirectional GRU เพื่อทำการทำนายผลว่าภาพที่ออกมาจาก ขั้นตอนก่อนหน้าควรที่จะเป็นตัวอักษรตัวใด โดยผ่านกระบวนการ 2 ขั้นตอน

1. Soft Max Function ทำหน้าที่หาระยะห่างระหว่างคำตอบ กับผลเฉลย โดยตัวเลขที่ได้มาจาก Function Soft Max จะผ่านการทำ Normalization ให้ค่าอยู่ระหว่าง 0-1

2. Cross Entropy เป็นฟังก์ชันของการตัดสินใจว่าภาพนี้ควรจะเป็นตัวอักษรตัวใด หากมีค่าเป็น 0 ถึง 1 ตรงกับภาพใดที่ใช้สำหรับการเรียนรู้ คอมพิวเตอร์จะทำนายว่าข้อมูลนั้นถูกต้อง และจะทำนายผลออกมา

Connectionist Temporal Classification (CTC) ทำหน้าที่ในการจัดเรียงตัวอักษรให้เป็นคำ โดยตัดช่องว่างและการตัดอักขระที่มีลักษณะติดกัน จะทำการตัดตัวที่ซ้ำออก

โครงข่ายประสาทเทียมแบบคอนโวลูชัน Convolution Neural Network (CNN)

ชั้น Convolution เป็นขั้นตอนในการสกัดคุณลักษณะของภาพ โดยการทำงานจะทำการแยกองค์ประกอบของภาพ เช่น สี หรือรูปร่าง

วิธีการคำนวณ

คำนวณโดยเอาค่าในภาพมาคูณกับค่าใน Window ที่ตำแหน่งตรงกัน ทำจนครบขนาดของ Window แล้วหาผลรวมของทุกตำแหน่งและแทนค่าในตำแหน่ง ดังภาพประกอบที่ 3.70 ค่าใหม่หลังจากการคำนวณ ในรอบที่ 1

Computer Science Department Faculty of Informatics, Mahasarakham University

7	4	2	2	7	8
4	6	4	3	4	6
1	5	7	1	8	2
7	8	8	1	2	4
1	5	1	5	8	1
2	1	4	7	6	7

ภาพประกอบที่ 3.39 ภาพต้นฉบับก่อนทำ Convolution ขนาด 6 x 6

1	0	1
0	1	0
1	0	1

ภาพประกอบที่ 3.40 Filter สำหรับการทำ Convolution ขนาด 3 x 3

1	0	1
0	1	0
1	0	1

*

7	4	2	2	7	8
4	6	4	3	4	6
1	5	7	1	8	2
7	8	8	1	2	4
1	5	1	5	8	1
2	1	4	7	6	7

ภาพประกอบที่ 3.41 การทำ Convolution ในรอบแรก

Computer Science Department
Faculty of Informatics, Mahasarakham University

ผลลัพธ์ ในรอบที่ 1

ตำแหน่งที่ 1,1 = 7 x 1 = 7

ตำแหน่งที่ 1,2 = 4 x 0 = 0

ตำแหน่งที่ 1,3 = 2 x 1 = 2

ตำแหน่งที่ 2,1 = 4 x 0 = 0

ตำแหน่งที่ 2,2 = 6 x 1 = 6

ตำแหน่งที่ 2,3 = 4 x 0 = 0

ตำแหน่งที่ 3,1 = 1 x 1 = 1

ตำแหน่งที่ 3,2 = 5 x 0 = 0

ตำแหน่งที่ 3,3 = 7 x 1 = 7

ผลรวมของทุกตำแหน่ง = 7+0+2+0+6+0+1+0+7 = 23

23			

ภาพประกอบที่ 3.42 ค่าใหม่หลังจากการคำนวณ ในรอบแรก

1	0	1
0	1	0
1	0	1

*

7	4	2	2	7	8
4	6	4	3	4	6
1	5	7	1	8	2
7	8	8	1	2	4
1	5	1	5	8	1
2	1	4	7	6	7

ภาพประกอบที่ 3.43 การทำ Convolution ในรอบที่สอง

Computer Science Department
Faculty of Informatics, Mahasarakham University

ผลลัพธ์ ในรอบที่ 2

ตำแหน่งที่ 1,1 = 4 x 1 = 4

ตำแหน่งที่ 1,2 = 2 x 0 = 0

ตำแหน่งที่ 1,3 = 2 x 1 = 2

ตำแหน่งที่ 2,1 = 6 x 0 = 0

ตำแหน่งที่ 2,2 = 4 x 1 = 4

ตำแหน่งที่ 2,3 = 3 x 0 = 0

ตำแหน่งที่ 3,1 = 5 x 1 = 5

ตำแหน่งที่ 3,2 = 7 x 0 = 0

ตำแหน่งที่ 3,3 = 1 x 1 = 1

ผลรวมของทุกตำแหน่ง = 4+0+2+0+4+0+5+0+1 = 16

23	16		

ภาพประกอบที่ 3.44 ค่าใหม่หลังจากการคำนวณ ในรอบที่สอง

23	16	27	17
28	25	19	22
18	24	25	11
26	18	25	27

ภาพประกอบที่ 3.45 ค่าหลังจากการคำนวณ Convolution จนครบทั้งหมด

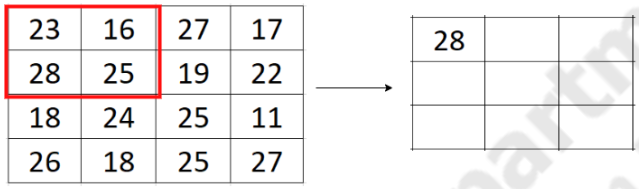
ชั้นพูล (Pooling layer) หลังจากที่ทำ Convolution เสร็จแล้ว จะทำการลดขนาดของข้อมูล แต่รายละเอียด Input ยังคงเดิม ซึ่งจะใช้ Max Pooling ในการคำนวณ

วิธีการคำนวณ จะทำการหนดค่า Stride = 1 และ Window ขนาด 2 x 2 โดย Max Pooling เป็นการนำค่ามากที่สุดที่อยู่ในบริเวณของ Window มาเป็นค่าของรูปแบบใหม่ จากนั้นขยับ Window ตามค่า Stride ดังนี้

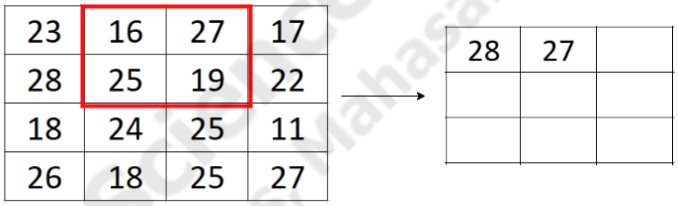
Computer Science Department
Faculty of Information, Mahasarakham University

23	16	27	17
28	25	19	22
18	24	25	11
26	18	25	27

ภาพประกอบที่ 3.46 ภาพก่อนการทำ Max Pooling



ภาพประกอบที่ 3.47 ภาพการทำ Max Pooling ครั้งแรก



ภาพประกอบที่ 3.48 ภาพการทำ Max Pooling ครั้งที่สอง

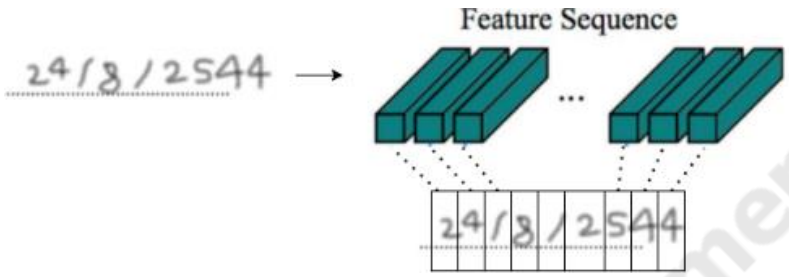
28	27	27
28	25	25
26	25	27

ภาพประกอบที่ 3.49 ภาพหลังจากการทำ Max Pooling จนครบแล้ว

Dropout เป็นขั้นตอนในการลดจำนวน Node ออกตามจำนวน เพื่อลดการเกิด Overfitting ของข้อมูล โดยค่าที่ใช้กำหนด Dropout เช่น 0.2 หมายความว่า 20% ของจำนวน Node ทั้งหมด จะไม่มีการ Update ค่า Weight ซึ่งค่าที่ใช้ใน Dropout เกิดจากการสุ่มความน่าจะเป็นขึ้นมา

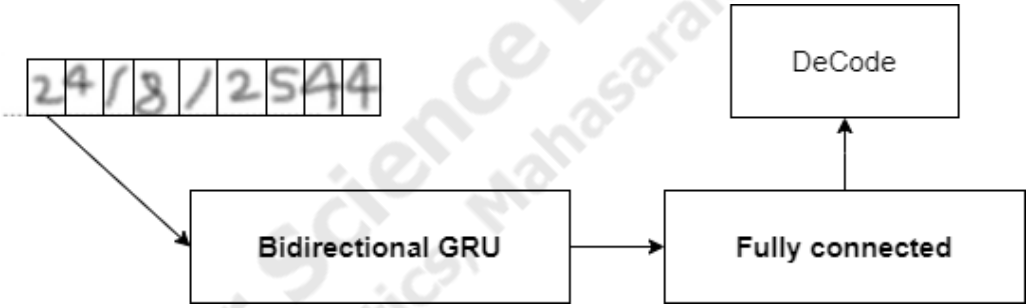
Computer Science Department
Faculty of Informatics, Mahasarakham University

Reshape เป็นการแปลงรูปภาพ โดยจะส่งข้อมูลเข้าทีละ Time-Step เอามาเรียงต่อกันเป็นแถวยาว จากนั้นส่งข้อมูลให้ Bidirectional GRU เพื่อทำการประมวลผลในขั้นตอนต่อไป



ภาพประกอบที่ 3.50 ตัวอย่าง Time-Step ของข้อมูล

Bidirectional GRU จากการ Reshape ผลลัพธ์ที่ได้จาก CNN (ไม่รวม Layer Fully Connected) โดนจะส่งข้อมูลเข้า GRU ครั้งละ 1 Time Step เข้าสู่ Algorithm GRU จนครบทุก Time Step ซึ่งการทำงานของ GRU สามารถอธิบายได้ดังภาพประกอบที่ 3.50



ภาพประกอบที่ 3.51 ตัวอย่างการ Feed ข้อมูลเข้าทีละ Time Step

ชั้นฟูลลีคอนเนก (Fully connected layer)

เป็นชั้นสุดท้ายของ CNN ที่เราทำการตัดออกมาทำการต่อเข้ากับ Bidirectional GRU เพื่อทำการทำนายผลว่าภาพที่ออกมาจาก ขั้นตอนก่อนหน้าควรที่จะเป็นตัวอักษรตัวใด โดยผ่านกระบวนการ 2 ขั้นตอน ดังนี้

Soft Max Function เป็นการหาระยะห่างระหว่างคำตอบกับผลเฉลย โดยตัวเลขที่ได้มากจาก Function Soft Max จะผ่านการทำ Normalization ให้ค่าอยู่ระหว่าง 0 – 1 ดังสมการ

$$\boldsymbol{\sigma(x_j) = \frac{e^{x_j}}{\sum_{i=1}^{k} x^{x_j}} f\ or\ j = 1, \dots, k} \qquad \textbf{(3.17)}$$

โดยที่ $\boldsymbol{x_j}$ คือ ค่าที่ Input

e คือ ตัวแปรที่มีค่าเท่ากับ 2.71828182846

$\boldsymbol{e^{x_j}}$ คือ ค่าที่ได้จากการนำ e มายกกำลังด้วย $\boldsymbol{x_j}$

ตัวอย่างการคำนวณ $\boldsymbol{x_j}$ = (7.0,3.0,6.0)

ขั้นตอนที่ 1 หาค่า $\boldsymbol{e^{x_j}}$ แต่ละค่า

$$e^{7.0} = 2.718^{7.0} = 1{,}095.8375$$

$$e^{3.0} = 2.718^{3.0} = 20.0793$$

$$e^{6.0} = 2.718^{6.0} = 403.1778$$

ขั้นตอนที่ 2

หาค่า $\sum_{i=1}^{k} e^{x_j} = 1{,}095.8375 + 20.0793 + 403.1778$

$$= 1{,}519.0946$$

ขั้นตอนที่ 3 หาค่า Soft Max โดยแทนค่าในสมการ จะได้

$$\sigma(7.0) = \frac{1095.8375}{1519.0946} = 0.7213$$

$$\sigma(3.0) = \frac{20.0793}{1519.0946} = 0.0132$$

$$\sigma(6.0) = \frac{403.1778}{1519.0946} = 0.2654$$

ซึ่งเมื่อนำค่า Soft Max ทุกค่ามารวมกันจะได้ผลลัพธ์เท่ากับ 1

Cross Entropy เป็นฟังก์ชันของการตัดสินใจว่าภาพนี้ควรจะเป็นตัวอักษรใด โดยหากค่าที่ผ่านการ Soft Max มาแล้วค่าจริงที่แสดงว่าอักษรนั้นอยู่ Class Lable ไหน มีค่าเป็น 1 ตรงกัน คอมพิวเตอร์จะทำนายว่าข้อมูลนั้นถูกต้อง และจะทำนายผลออกมาซึ่งสามารถหาได้ดังสมการ

$$\text{Cross Entropy} = -\sum_i' Y * \log\log(Y_i) \qquad \textbf{(3.18)}$$

โดยที่ Y_i' คือ ค่าจริงที่แสดงว่าอักษรนั้นอยู่ Class Label ไหน

Y_i คือ ค่าที่ผ่าน Function Soft Max

Connectionist Temporal Classification (CTC)

เป็นขั้นตอนในการจัดเรียงตัวอักษรให้เป็นคำ โดยตัดช่องว่างและการตัดอักขระที่มีลักษณะติดกัน จะทำการตัดตัวที่ซ้ำออก โดยคำนวณได้จากสมการ

Computer Science Department
Faculty of Informatics, Mahasarakham University

$$-\sum_{(x,z)\in S} \log_{eP}(z|x) \quad (3.19)$$

โดยที่ x คือ Training Simple

s คือ Training Data

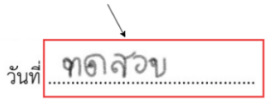
z คือ Generated Sequence

3.5 ชุดข้อมูลที่ใช้ในการทดลอง

ชุดข้อมูลที่ใช้ในการสร้างโมเดลรู้จำ มากกว่า 200 ภาพจะแบ่งออกเป็นชุดสอน 60% ชุดวาลิเดท 20% และชุดทดสอบ 20% ซึ่งนำมาตัดเป็นข้อมูล 1 บรรทัด

บันทึกข้อความ

ภาพประกอบที่ 3.52 ตัวอย่างชุดข้อมูลเอกสาร เบิกค่าลงทะเบียน (ชุดข้อมูล 1 ภาพ)



ภาพประกอบที่ 3.53 ตัวอย่างชุดข้อมูลเอกสาร เบิกค่าลงทะเบียน (ชุดข้อมูล 1 บรรทัด)

3.6 การวัดประสิทธิภาพ

การวัดประสิทธิภาพมี 2 อย่างคือ

1) การวัดประสิทธิภาพการตัดข้อความว่าตัดตรงหรือไม่ตรง

2) การวัดการรู้จำตัวอักษรโดยใช้ Levenshtein Edit Distance ซึ่งจะแสดงรายละเอียดดังนี้

Computer Science Department
Faculty of Informatics, Mahasarakham University

3.6.1 การวัดประสิทธิภาพการตรวจจับข้อความ Confusion Matrix ด้วยค่าความถูกต้อง (Accuracy)

Confusion Matrix เป็นการประเมินผลลัพธ์ของการทำนาย (Prediction) โดยมีแนวคิดว่าสิ่งที่โมเดล (Model) ทำนายกับผลลัพธ์ที่เกิดขึ้นจริงมีสักส่วนเป็นอย่างไรซึ่งมีตารางการวัดประสิทธิภาพดังตารางดังนี้

ตารางที่ 3.3 ตารางของ Confusion Matrix

Predicted / Actually	Actually Positive	Actually Negative
Predicted Positive	True Positive (TP)	False Positive (FP)
Predicted Negative	False Negative (FN)	True Negative (TN)

หรือกล่าวได้ว่า Accuracy = ผลรวมของตัวเลขบนเส้นทแยงมุมในตาราง Confusion Matrix / จำนวนทั้งหมดดังสมการ

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad \textbf{(3.20)}$$

โดยที่

True Positive (TP) คือ สิ่งที่ผลเฉลยว่า “จริง” และทำนายมีค่า “จริง”
True Negative (TN) คือ สิ่งที่ผลเฉลยว่า “ไม่จริง” และทำนายมีค่า “ไม่จริง”
False Positive (FP) คือ สิ่งที่ผลเฉลยว่า “จริง” แต่ทำนายมีค่า “ไม่จริง”
False Negative (FN) คือ สิ่งที่ผลเฉลยว่า “ไม่จริง” แต่ทำนายมีค่า “จริง”

ตัวอย่างการทำ Confusion Matrix

ตารางที่ 3.4 ตัวอย่างชุดข้อมูลอบรมและสัมมนา

Expected	Predicted
อบรม	อบรม
อบรม	สัมมนา
อบรม	อบรม

Computer Science Department
Faculty of Informatics, Mahasarakham University

ตารางที่ 3.4 ตัวอย่างชุดข้อมูลอบรมและสัมมนา (ต่อ)

Expected	Predicted
สัมมนา	อบรม
สัมมนา	อบรม
สัมมนา	สัมมนา

ตารางที่ 3.5 ตัวอย่างการคำนวณ Confusion Matrix

Confusion Matrix	อบรม	สัมมนา
อบรม	2	2
สัมมนา	1	1

ตัวอย่างการหา Accuracy ของคำว่า “อบรม”

จากสูตร $\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$

แทนค่า $\text{Accuracy} = \frac{2 + 1}{2 + 1 + 2 + 1}$

จะได้ค่า $\text{Accuracy} = 0.5$ ดังนั้น “อบรม” มีค่าความถูกต้องอยู่ที่ 50%

3.6.1 การวัดประสิทธิภาพการรู้จำตัวอักษร Character Error Rate (CER) โดยใช้ Levenshtein Edit Distance

Levenshtein Edit Distance เป็นขั้นตอนการวัดเพื่อหาค่าความแตกต่างของข้อความทั้งสองชุดระหว่างชุดแรกที่เป็นผลเฉลย และชุดที่สองที่เป็นชุดผลการทำนาย โดยจะนับจำนวนครั้งของการแทรกเข้า ลบออก และการแทนที่ ดังสมการนี้

โดยที่

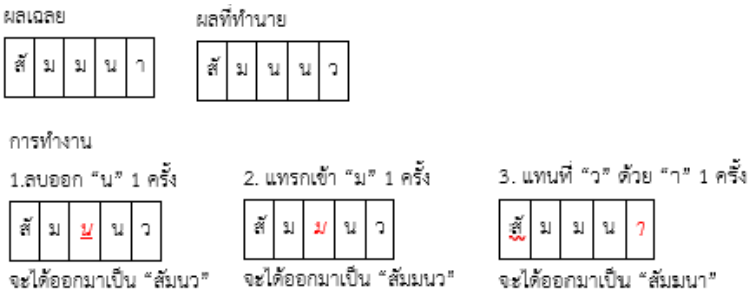
S คือ จำนวนตัวอักขระที่ถูกแทนที่

D คือ จำนวนตัวอักขระที่ถูกลบออก

I คือ จำนวนตัวอักขระที่ถูกแทรกเข้า

N คือ จำนวนตัวอักขระของผลเฉลย

Computer Science Department Faculty of Informatics Mahasarakham University



ภาพประกอบที่ 3.54 การวัดประสิทธิภาพการรู้จำตัวอักษรคำที่ 1

จากสมการ

$$CER = \frac{S + D + I}{N} * 100$$

แทนค่า $\quad CER = \frac{1+1+1}{5} * 100$

จะได้ $\quad CER = 60\%$

ดังนั้น Character Error Rate ของคำว่า “สัมมนา” มีค่าผิดพลาดอยู่ที่ 60%

ภาพประกอบที่ 3.55 การวัดประสิทธิภาพการรู้จำตัวอักษรคำที่ 2

จากสมการ

$$CER = \frac{S + D + I}{N} * 100$$

แทนค่า $\quad CER = \frac{1+2+2}{11} * 100$

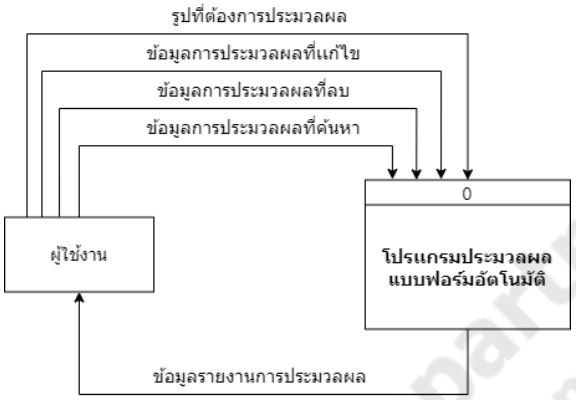
จะได้ $\quad CER = 45.45\%$

ดังนั้น Character Error Rate ของคำว่า “สัมมนา” มีค่าผิดพลาดอยู่ที่ 45.45%

Computer Science Department
Faculty of Informatics Mahasarakham University

3.7 ผลการออกแบบและพัฒนาระบบ

3.7.1 Context Diagram



ภาพประกอบที่ 3.56 ภาพ Context Diagram

3.7.2 External Entity Description

ตารางที่ 3.6 ตารางแสดง External Entity Description

Name	Description	Input Data Flow	Output Data Flow
ผู้ใช้งาน	ผู้ใช้บริการของระบบสามารถส่งรูปภาพเพื่อเข้ามาประมวลผลได้	- รูปภาพที่จะประมวลผล	- รายงานผลลัพธ์การประมวลผล

3.7.3 Data Store Description

ตารางที่ 3.7 ตารางแสดง Data Store Description

ID	Name	Description	Data Structure
D1	แฟ้มข้อมูลการประมวล	ฐานข้อมูลสำหรับเก็บข้อมูลในการประมวลผล	รหัสการประมวลผลข้อความ+เวลาโทรทำนาย+วันที่ทำนาย+บันทึกข้อความที่ทำนาย+วันที่บันทึกข้อความที่ทำนาย+เดินทางไปราชการที่ทำนาย+เดินทางไปราชการที่ทำนาย - (ต่อ)+ณ ที่ทำนาย+ระหว่างที่ที่ทำนาย+ระหว่างที่ที่ทำนาย – (ต่อ)+หมวดเงินที่ทำนาย+รายการที่ทำนาย+จำนวนเงินที่ทำนาย+จำนวนเงินที่เป็นภาษาไทยที่ทำนาย+ชื่อนามสกุลหนึ่งที่ทำนาย+ตำแหน่งหนึ่งที่ทำนาย+ชื่อนามสกุลสองที่ทำนาย+ตำแหน่งสองที่ทำนาย

Computer Science Department
Faculty of Informatics Mahasarakham University

3.7.4 Data Structure Description

ตารางที่ 3.8 ตารางแสดง Data Structure Description

Name	Description	Source	Destination	Data Structure
ข้อมูลรูปภาพที่ต้องประมวล	ไฟล์รูปภาพที่ผู้ใช้ส่งเข้ามาประมวลผล	ผู้ใช้งาน	Process 1.0 การประมวลผล	รหัสการประมวลผลข้อความ+เวลาโทรทำนาย+วันที่ทำนาย+บันทึกข้อความที่ทำนาย+วันที่บันทึกข้อความที่ทำนาย+เดินทางไปราชการที่ทำนาย+เดินทางไปราชการที่ทำนาย-(ต่อ)+ณ ทำนาย+ระหว่างที่ที่ทำนาย+ระหว่างที่ที่ทำนาย–(ต่อ)+หมวดเงินที่ทำนาย+รายการที่ทำนาย+ จำนวนเงินที่ทำนาย+ จำนวนเงินที่เป็นภาษาไทยที่ทำนาย+ชื่อนามสกุลหนึ่งที่ทำนาย+ตำแหน่งหนึ่งที่ทำนาย+ชื่อนามสกุลสองที่ทำนาย+ตำแหน่งสองที่ทำนาย
		Process 1.0 การประมวลผล	D1 แฟ้มข้อมูลการประมวลผล	
รายงานข้อมูลการประมวณผล	ผลลัพธ์ข้อความที่ผ่านการประมวลผล	ผู้ใช้งาน	Process 1.0 การประมวลผล	รหัสการประมวลผลข้อความ+เวลาโทรทำนาย+วันที่ทำนาย+บันทึกข้อความที่ทำนาย+วันที่บันทึกข้อความที่ทำนาย+เดินทางไปราชการที่ทำนาย+เดินทางไปราชการที่ทำนาย-(ต่อ)+ณ ทำนาย+ระหว่างที่ที่ทำนาย+ระหว่างที่ที่ทำนาย–(ต่อ)+หมวดเงินที่ทำนาย+รายการที่ทำนาย+ จำนวนเงินที่ทำนาย+ จำนวนเงินที่เป็นภาษาไทยที่ทำนาย+ชื่อนามสกุลหนึ่งที่ทำนาย+ตำแหน่งหนึ่งที่ทำนาย+ชื่อนามสกุลสองที่ทำนาย+ตำแหน่งสองที่ทำนาย
		Process 1.0 การทำนายผล	D1 แฟ้มข้อมูลการประมวลผล	

Computer Science Department Faculty of Informatics, Mahasarakham University

ตารางที่ 3.8 ตารางแสดง Data Structure Description – (ต่อ)

Name	Description	Source	Destination	Data Structure
ข้อมูลการประมวลผลที่แก้ไข	ข้อมูลการประมวลผล	ผู้ใช้งาน	Process 1.1 แก้ไขข้อมูลการประมวลผล	รหัสการประมวลผลข้อความ+เวลาโทรทำนาย+วันที่ทำนาย+บันทึกข้อความที่ทำนาย+วันที่บันทึกข้อความที่ทำนาย+เดินทางไปราชการที่ทำนาย+เดินทางไปราชการที่ทำนาย-(ต่อ)+ณ ทำนาย+ระหว่างที่ที่ทำนาย+ระหว่างที่ที่ทำนาย–(ต่อ)+หมวดเงินที่ทำนาย+รายการที่ทำนาย+ จำนวนเงินที่ทำนาย+ จำนวนเงินที่เป็นภาษาไทยที่ทำนาย+ชื่อนามสกุลหนึ่งที่ทำนาย+ตำแหน่งหนึ่งที่ทำนาย+ชื่อนามสกุลสองที่ทำนาย+ตำแหน่งสองที่ทำนาย
		Process 1.1 แก้ไขข้อมูลการประมวลผล	D1 แฟ้มข้อมูลการประมวลผล	
ข้อมูลการประมวลผลที่ลบ	ข้อมูลการประมวลผล	ผู้ใช้งาน	Process 1.2 ลบข้อมูลการประมวลผล	ชื่อผู้อัพโหลดข้อมูล
		Process 1.2 ลบข้อมูลการประมวลผล	D1 แฟ้มข้อมูลการประมวลผล	
ข้อมูลการประมวลผลที่ค้นหา	ข้อมูลการประมวลผล	ผู้ใช้งาน	Process 1.3 ค้นหาข้อมูลการประมวลผล	ชื่อผู้อัพโหลดข้อมูล

Computer Science Department Faculty of Informatics Mahasarakham University

3.7.5 รายละเอียดตารางข้อมูล (Data Table Description)

ตารางที่ 3.9 ตารางแสดงข้อมูลการทำนาย (Predicts)

Attirbute Name	Type	Size	Description	Key type	References
dID	Int	11	รหัสการประมวลผล	primary key	10
call time	varchar	200	เวลาที่โทร		17:55
deta1	varchar	100	วันที่		1 ส.ค. 2564
message	varchar	200	บันทึกข้อความ		ร้องขอ
deta2	varchar	100	วันที่		1 ส.ค. 2564
For	varchar	200	เพื่อ		ไปราชการ
For	varchar	200	เพื่อ – (ต่อ)		เพื่อนำเสนองาน
on	varchar	200	ณ		ต่างจังหวัด
during the day	varchar	200	ระหว่างวันที่		1 ส.ค. 2564 ถึง
during the day – (ต่อ)	varchar	200	ระหว่างวันที่ – (ต่อ)		15 ส.ค. 2564
money category	varchar	100	หมวดเงิน		ออกราชการ
list	varchar	200	รายการ		งานวิจัย
money	Int	11	จำนวนเงิน		20000
Thai language money	varchar	200	จำนวนเงินภาษาไทย		สองหมื่น
name1	varchar	200	ชื่อ		อนพัช
position1	varchar	200	ตำแหน่ง		นักศึกษา
name2	varchar	200	ชื่อ		พงศ์นรา
position2	varchar	200	ตำแหน่ง		นักศึกษา

Computer Science Department Faculty of Informatics, Mahasarakham University