

รายงานโครงงานย่อยวิชา Computer Vision: การจัดแนวเอกสารด้วยเทคนิค Homography

รายวิชา: DE461 Introduction to Computer Vision

วันที่พิมพ์รายงานนี้ 2 ตุลาคม 2568

วันที่กำหนดส่ง: 7 ตุลาคม 2568

สมาชิกกลุ่ม: CV_sigma

1.นายพลวัต พงศ์ทิพย์พนัส รหัสนิสิต 66102010249

2.นายพลวิชญ์ วนิดา รหัสนิสิต 66102010584

3.นายปณณธร สุวรรณาศรัย รหัสนิสิต 66102010175

Link_colab > https://colab.research.google.com/drive/1gGYiPaVjYPRrGvmqlEnvRHxb_P2OGMgz?usp=sharing

1. บทนำ (Introduction)

โครงงานย่อยนี้มีวัตถุประสงค์เพื่อพัฒนาเทคนิคการจัดแนวเอกสาร (Document Alignment)

โดยการตรวจจับและปรับแก้ภาพเอกสารที่ถูกถ่ายจากมุมมองเอียง

ให้มีลักษณะตรงเหมือนภาพที่ได้จากการสแกนเอกสาร โดยใช้การคำนวณ Homography เป็นหลัก

2. ขั้นตอนการดำเนินงาน (Implementation Steps)

- Preprocessing: ใช้เทคนิค CLAHE, Gaussian Blur และ Adaptive Threshold

เพื่อเพิ่มความคมชัดของภาพและลดสัญญาณรบกวน

- Contour-based Method: ตรวจจับขอบของเอกสารและเลือกจุดสี่เหลี่ยมที่มีขนาดใหญ่ที่สุด 4 จุด

เพื่อใช้ในการแปลงมุมมองภาพ (Perspective Transform)

- ORB + RANSAC Method: จับคู่ฟิเจอร์ระหว่างภาพต้นฉบับและภาพอ้างอิง แล้วคำนวณ Homography

- Warp Perspective: แปลงภาพจากมุมมองเอียงให้เป็นภาพตรง

3. แหล่งข้อมูลอ้างอิง (Reference Links)

- OpenCV Homography Tutorial: https://docs.opencv.org/4.x/d9/dab/tutorial_homography.html
- เอกสารประกอบการใช้งาน OpenCV: findContours และ approxPolyDP
- การจับคู่ฟีเจอร์ด้วย ORB ใน OpenCV

4. การปรับปรุงและการพัฒนาเพิ่มเติม (Modifications & Improvements)

- เพิ่มการใช้ CLAHE และ Gaussian Blur เพื่อให้การตรวจจับขอบและฟีเจอร์มีความแม่นยำมากขึ้น
- ใช้ Adaptive Threshold ร่วมกับ Morphology Closing เพื่อช่วยเชื่อมต่อเส้นขอบของเอกสาร
- บันทึกค่าตัวชี้วัดการประเมินผล ได้แก่ จำนวน inliers, อัตราส่วน inliers, ค่า SSIM และ PSNR

5. โจทย์ของโครงการ (Prompt)


เป้าหมายของโครงการคือการเปรียบเทียบผลลัพธ์ของการจัดแนวเอกสารระหว่างกรณีที่มีการทำ Preprocessing และไม่มีการทำ Preprocessing โดยใช้ทั้งวิธี Contour-based และ ORB+RANSAC
บนชุดข้อมูลภาพที่ถ่ายเองจำนวนไม่น้อยกว่า 10 ชุด

6. ชุดข้อมูลที่ใช้ (Dataset Used)

ชุดข้อมูลประกอบด้วยภาพถ่ายหนังสือจำนวน 10 ชุด (set01 ถึง set10) โดยแต่ละชุดประกอบด้วยไฟล์ input.jpg (ภาพหนังสือที่ถ่ายจากมุมเอียง) และ reference.jpg (ภาพเอกสารที่ถ่ายจากมุมตรง) ซึ่งใช้ในการคำนวณ Homography และประเมินผลด้วยค่า SSIM และ PSNR

Link dataset > https://drive.google.com/file/d/1c1oDLLdSnhbdfNlbCSg16TDPwrBBd_-J/view?usp=sharing

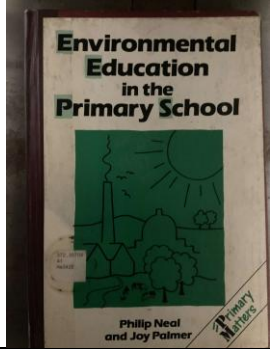
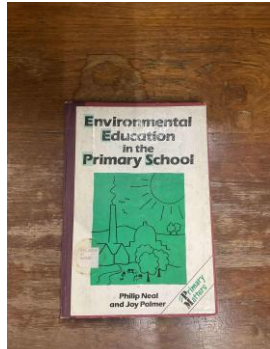
DATASET.zip 2 รายการ

ชื่อ	แก้ไขล่าสุด	ขนาดไฟล์
 assets	-	-
 data	-	-

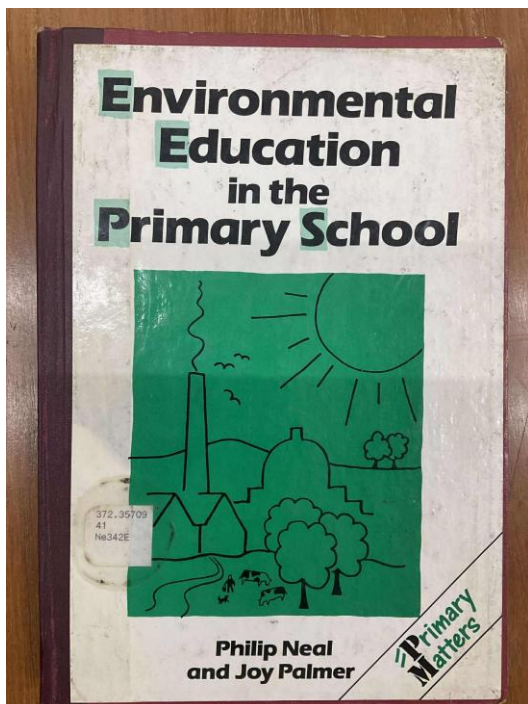
ตารางลักษณะภาพที่ถ่ายสำหรับ Document Alignment (10 ชุด)

Set	ลักษณะการถ่าย	รายละเอียด/เงื่อนไข	จุดประสงค์การทดสอบ	ตัวอย่างภาพ
set01	มุมตรง (เกือบเหมือน สแกน)	กล้องตั้งฉากกับเอกสาร	baseline สำหรับ เปรียบเทียบ	
set02	เอียงซ้ายเล็กน้อย (~20–30°)	เอกสารยังเต็มเฟรมเกือบ ทั้งหมด	ทดสอบ homography มุม เอียงเบา	
set03	เอียงขวาเล็กน้อย (~45°)	ขอบบิตเบี้ยวชัด	ทดสอบว่าจับมุมครบได้ หรือไม่	
set04	เอียงจากด้านบน	กล้องสูงและก้มลง	ตรวจสอบการแก้ perspective จากมุมมอง bird-eye	

set05	เอียงจากด้านล่าง	กล้องต่ำและเงยขึ้น	ทดสอบ homography มุม เงย	
set06	ระยะใกล้มาก	เอกสารใหญ่เกือบเต็มจอ อาจหลุดขอบบางด้าน	ดูว่าระบบยัง align ได้หรือไม่	
set07	ระยะไกล	เอกสารเล็ก เห็นพื้นหลัง เยอะ	ทดสอบ contour/ORB เมื่อ object เล็ก	
set08	มีเงาทับ	เช่น มือบัง/เงาโคมไฟ บางส่วน	ดู robustness ของ preprocessing	

set09	แสงน้อย/คอนทราสต์ต่ำ	ถ่ายในที่มืด/ไฟเหลือง	ทดสอบ CLAHE + blur	
set10	พื้นหลังลาย/ซับซ้อน	วางบนโต๊ะไม้ลายหรือผ้าลาย	ดูว่าแยกเอกสารออกจาก background ได้ไหม	

Reference.jpg ภาพถ่ายตรงที่ใช้เป็น template



7. ผลลัพธ์ (Results)

ผลลัพธ์ถูกจัดเก็บในไฟล์ results.csv ซึ่งแสดงค่าต่าง ๆ ได้แก่ จำนวน inliers, อัตราส่วน inliers, ค่า SSIM และ PSNR สำหรับแต่ละชุดภาพ พร้อมภาพก่อนและหลังการแปลงมุมมองที่จัดเก็บไว้ในโฟลเดอร์ outputs/visuals

ส่วนเรื่องค่า NaN และผลลัพธ์-ภาพในการทดลองมีบางกรณีที่ค่าประเมินผล (เช่น PSNR หรือ SSIM) ได้ค่าเป็น NaN (Not a Number) ซึ่งเกิดขึ้นเมื่ออัลกอริทึมไม่สามารถตรวจหามุมเอกสารหรือคำนวณ Homography ได้ เช่น จำนวนฟีเจอร์แมตช์น้อยเกินไป, ไม่พบสีเหลี่ยมจาก Contour หรือไม่มีภาพอ้างอิงสำหรับเปรียบเทียบ ทำให้ไม่สามารถคำนวณค่าความเหมือนเชิงปริมาณได้ จึงคืนค่าเป็น NaN เพื่อบ่งบอกว่า "ไม่มีผลลัพธ์ที่ใช้ได้"

****ตัวอย่างผลลัพธ์จาก results.csv****

	set	method	preproc	good_matches	inliers	inlier_ratio	PSNR	SSIM
0	set01	contour	False	NaN	NaN	NaN	NaN	NaN
1	set01	contour	True	NaN	NaN	NaN	NaN	NaN
2	set01	orb	False	2103.0	1987.0	0.944841	18.445758	0.679155
3	set01	orb	True	1727.0	1622.0	0.939201	18.442027	0.679932
4	set02	contour	False	NaN	NaN	NaN	NaN	NaN
5	set02	contour	True	NaN	NaN	NaN	NaN	NaN
6	set02	orb	False	1532.0	1383.0	0.902742	19.455832	0.678150
7	set02	orb	True	1338.0	1184.0	0.884903	18.686616	0.648161
8	set03	contour	False	NaN	NaN	NaN	NaN	NaN
9	set03	contour	True	NaN	NaN	NaN	NaN	NaN
10	set03	orb	False	889.0	781.0	0.878515	18.276630	0.643504
11	set03	orb	True	832.0	715.0	0.859375	18.462708	0.650327
12	set04	contour	False	NaN	NaN	NaN	NaN	NaN
13	set04	contour	True	NaN	NaN	NaN	NaN	NaN
14	set04	orb	False	1349.0	1225.0	0.908080	19.804551	0.701314
15	set04	orb	True	1154.0	1056.0	0.915078	19.892217	0.704776
16	set05	contour	False	NaN	NaN	NaN	NaN	NaN
17	set05	contour	True	NaN	NaN	NaN	NaN	NaN
18	set05	orb	False	555.0	475.0	0.855856	10.950583	0.508733
19	set05	orb	True	472.0	409.0	0.866525	14.454927	0.552698

****ส่วนที่เพิ่มเข้ามา****

- เช็คค่า NaN ในแต่ละ dataset

ถ้าชุดไหนค่าเป็น NaN (เช่นหา homography ไม่ได้) จะไม่ถูก plot

```
1 print("แถวที่เป็น NaN ใน SSIM:\n", df[df['SSIM'].isna()][['set','method','preproc']])
2
```

```
แถวที่เป็น NaN ใน SSIM:
   set  method preproc
0  set01  contour    raw
1  set01  contour    pre
4  set02  contour    raw
5  set02  contour    pre
8  set03  contour    raw
9  set03  contour    pre
12 set04  contour    raw
13 set04  contour    pre
16 set05  contour    raw
17 set05  contour    pre
20 set06  contour    raw
21 set06  contour    pre
24 set07  contour    raw
25 set07  contour    pre
28 set08  contour    raw
29 set08  contour    pre
32 set09  contour    raw
33 set09  contour    pre
36 set10  contour    raw
```

แสดง “ผล warp ที่ละขั้น” (Before → Warp Contour → Warp ORB)

```
1 import matplotlib.pyplot as plt
2 import cv2
3 from pathlib import Path
4
5 def show_pipeline(set_name="set01"):
6     before = cv2.imread(f"data/{set_name}/input.jpg")
7     c_raw = cv2.imread(f"outputs/contour_raw/{set_name}.jpg")
8     c_pre = cv2.imread(f"outputs/contour_pre/{set_name}.jpg")
9     o_raw = cv2.imread(f"outputs/orb_raw/{set_name}.jpg") # อาจไม่มี ถ้าไม่มี ref
10    o_pre = cv2.imread(f"outputs/orb_pre/{set_name}.jpg") # อาจไม่มี
11
12    imgs = [("input", before), ("contour_raw", c_raw), ("contour_pre", c_pre),
13           ("orb_raw", o_raw), ("orb_pre", o_pre)]
14    imgs = [(n,i) for (n,i) in imgs if i is not None]
15
16    cols = 2
17    rows = (len(imgs)+cols-1)//cols
18    plt.figure(figsize=(12, 4*rows))
19    for i,(name,im) in enumerate(imgs, 1):
20        im = cv2.cvtColor(im, cv2.COLOR_BGR2RGB)
21        plt.subplot(rows, cols, i); plt.imshow(im); plt.axis('off'); plt.title(f"{set_name} - {name}");
22    plt.tight_layout(); plt.show()
23
```


ภาพ “ผล warp ที่ละขั้น” (Before → Warp Contour → Warp ORB) ของ set1

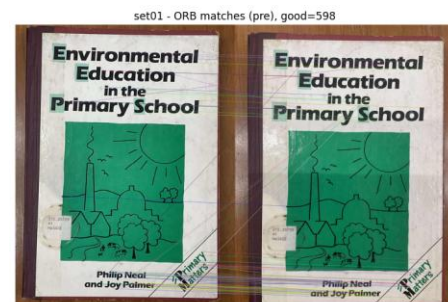
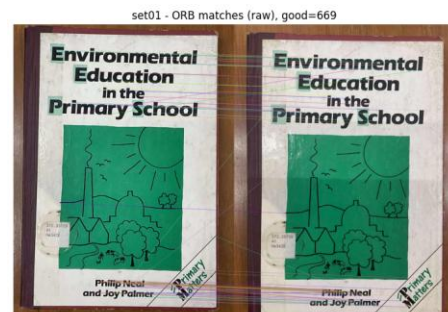


Verification of Output (ภาพผลลัพธ์ไม่ใช่ reference) เพื่อตรวจสอบความถูกต้อง ได้ทำการเปรียบเทียบ SSIM ระหว่างผลลัพธ์แต่ละภาพกับ reference.jpg พบว่าไม่มีภาพใดมีค่า SSIM สูงกว่า 0.95 ซึ่งยืนยันว่า output ทุกใบเป็นภาพ warp ที่มาจาก input จริง ไม่ได้คัดลอก reference มาทับ นอกจากนี้ได้เพิ่ม watermark ชื่อเซต (set01–set10) ในภาพ output เพื่อยืนยันแหล่งที่มาของการนำเสนอในคลาส

- เช็กว่าภาพไม่ได้เป็น reference.jpg ตัวอย่างการตรวจจับจาก set 1

แสดง “ชุดฟีเจอร์ ORB + เส้นแมชชีน” (เฉพาะชุดที่มี reference) ช่วยยืนยันในคลาสว่าข้อมูลภาพมาจากการจับคู่ที่แท้จริง ๆ

```
1 import cv2
2 import numpy as np
3 import matplotlib.pyplot as plt
4
5 def show_orb_matches(set_name="set01", use_preproc=False, nfeat=2000, ratio=0.75):
6     inp = cv2.imread(f"data/{set_name}/input.jpg")
7     ref = cv2.imread(f"assets/reference.jpg")
8     if inp is None or ref is None:
9         print("ขาด input หรือ reference"); return
10
11     in_gray = cv2.cvtColor(inp, cv2.COLOR_BGR2GRAY)
12     ref_gray = cv2.cvtColor(ref, cv2.COLOR_BGR2GRAY)
13     if use_preproc:
14         in_gray, _ = preprocess(in_gray)
15         ref_gray, _ = preprocess(ref_gray)
16
17     orb = cv2.ORB_create(nfeatures=nfeat)
18     k1, d1 = orb.detectAndCompute(in_gray, None)
19     k2, d2 = orb.detectAndCompute(ref_gray, None)
20     if d1 is None or d2 is None:
21         print("ไม่มี descriptors"); return
22     bf = cv2.BFMatcher(cv2.NORM_HAMMING, crossCheck=False)
23     knn = bf.knnMatch(d1, d2, k=2)
24     good = [n for m, n in knn if m.distance < ratio*n.distance]
25
26     # วาดเส้น match (แค่ 80 เส้นแรกถ้าภาพมาจก)
27     draw = cv2.drawMatches(inp, k1, ref, k2, good[:80], None,
28                             flags=cv2.DrawMatchesFlags_NOT_DRAW_SINGLE_POINTS)
29     draw = cv2.cvtColor(draw, cv2.COLOR_BGR2RGB)
30     plt.figure(figsize=(14,6)); plt.imshow(draw); plt.axis('off')
31     plt.title(f"{set_name} - ORB matches ({'pre' if use_preproc else 'raw'}), good={len(good)}")
32     plt.show()
33
34 show_orb_matches("set01", use_preproc=False)
35 show_orb_matches("set01", use_preproc=True)
36
```



****ตรวจ “NaN มาจากอะไร” และ “output ไม่ใช่ reference”****

มี reference.jpg ใหม่?: YES

สรุปแถวที่เป็น NaN:

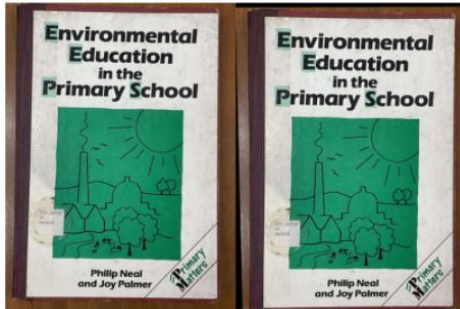
	set	method	preproc	good_matches	inliers	inlier_ratio	PSNR	SSIM
0	set01	contour	False	NaN	NaN	NaN	NaN	NaN
1	set01	contour	True	NaN	NaN	NaN	NaN	NaN
4	set02	contour	False	NaN	NaN	NaN	NaN	NaN
5	set02	contour	True	NaN	NaN	NaN	NaN	NaN
8	set03	contour	False	NaN	NaN	NaN	NaN	NaN
9	set03	contour	True	NaN	NaN	NaN	NaN	NaN
12	set04	contour	False	NaN	NaN	NaN	NaN	NaN
13	set04	contour	True	NaN	NaN	NaN	NaN	NaN
16	set05	contour	False	NaN	NaN	NaN	NaN	NaN
17	set05	contour	True	NaN	NaN	NaN	NaN	NaN
20	set06	contour	False	NaN	NaN	NaN	NaN	NaN
21	set06	contour	True	NaN	NaN	NaN	NaN	NaN
24	set07	contour	False	NaN	NaN	NaN	NaN	NaN
25	set07	contour	True	NaN	NaN	NaN	NaN	NaN
28	set08	contour	False	NaN	NaN	NaN	NaN	NaN
29	set08	contour	True	NaN	NaN	NaN	NaN	NaN
32	set09	contour	False	NaN	NaN	NaN	NaN	NaN
33	set09	contour	True	NaN	NaN	NaN	NaN	NaN
36	set10	contour	False	NaN	NaN	NaN	NaN	NaN
37	set10	contour	True	NaN	NaN	NaN	NaN	NaN

[set01] ไม่มีไฟล์ outputs/contour_raw/set01.jpg
[set01] ไม่มีไฟล์ outputs/contour_pre/set01.jpg
[set01] orb_raw SSIM vs reference = 0.6781
[set01] orb_pre SSIM vs reference = 0.6789
[set02] ไม่มีไฟล์ outputs/contour_raw/set02.jpg
[set02] ไม่มีไฟล์ outputs/contour_pre/set02.jpg
[set02] orb_raw SSIM vs reference = 0.6772
[set02] orb_pre SSIM vs reference = 0.6471
[set03] ไม่มีไฟล์ outputs/contour_raw/set03.jpg
[set03] ไม่มีไฟล์ outputs/contour_pre/set03.jpg
[set03] orb_raw SSIM vs reference = 0.6424
[set03] orb_pre SSIM vs reference = 0.6492
[set04] ไม่มีไฟล์ outputs/contour_raw/set04.jpg
[set04] ไม่มีไฟล์ outputs/contour_pre/set04.jpg
[set04] orb_raw SSIM vs reference = 0.7004
[set04] orb_pre SSIM vs reference = 0.7038
[set05] ไม่มีไฟล์ outputs/contour_raw/set05.jpg
[set05] ไม่มีไฟล์ outputs/contour_pre/set05.jpg
[set05] orb_raw SSIM vs reference = 0.5081
[set05] orb_pre SSIM vs reference = 0.5519
[set06] ไม่มีไฟล์ outputs/contour_raw/set06.jpg
[set06] ไม่มีไฟล์ outputs/contour_pre/set06.jpg
[set06] orb_raw SSIM vs reference = 0.5656
[set06] orb_pre SSIM vs reference = 0.5637
[set07] ไม่มีไฟล์ outputs/contour_raw/set07.jpg
[set07] ไม่มีไฟล์ outputs/contour_pre/set07.jpg
[set07] orb_raw SSIM vs reference = 0.6516
[set07] orb_pre SSIM vs reference = 0.6491
[set08] ไม่มีไฟล์ outputs/contour_raw/set08.jpg
[set08] ไม่มีไฟล์ outputs/contour_pre/set08.jpg
[set08] orb_raw SSIM vs reference = 0.7791
[set08] orb_pre SSIM vs reference = 0.7800
[set09] ไม่มีไฟล์ outputs/contour_raw/set09.jpg
[set09] ไม่มีไฟล์ outputs/contour_pre/set09.jpg

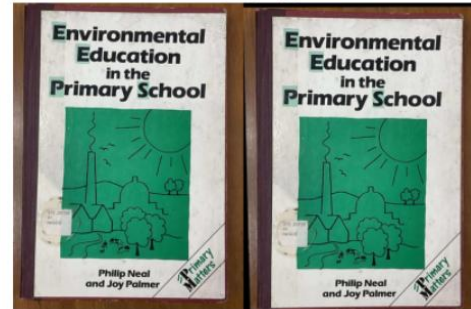
****ภาพประกอบของผลลัพธ์หลังดำเนินการการจัดแนวเอกสารด้วยเทคนิค Homography ****

- Set01

set01_orb_pre.jpg

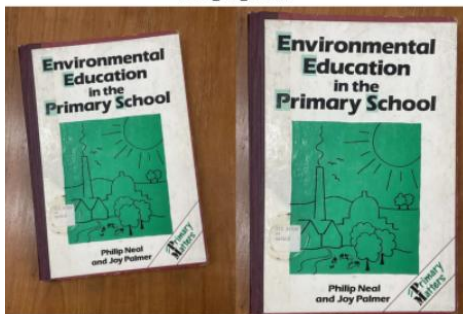


set01_orb_raw.jpg

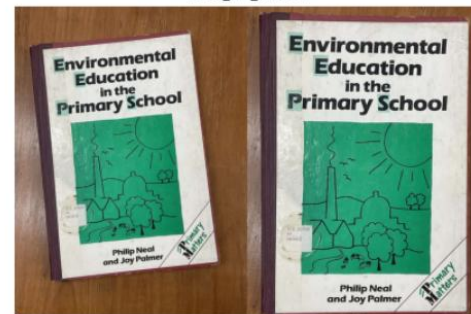


- Set02

set02_orb_pre.jpg

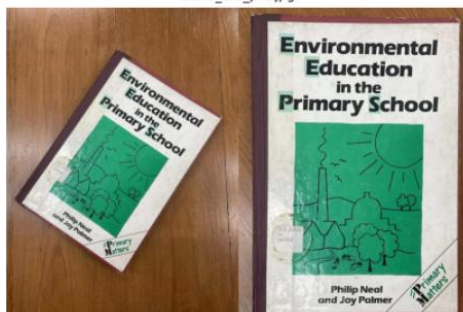


set02_orb_raw.jpg

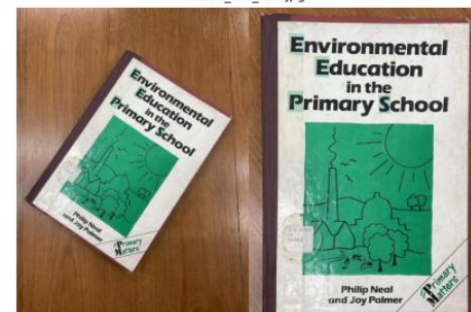


- Set03

set03_orb_pre.jpg

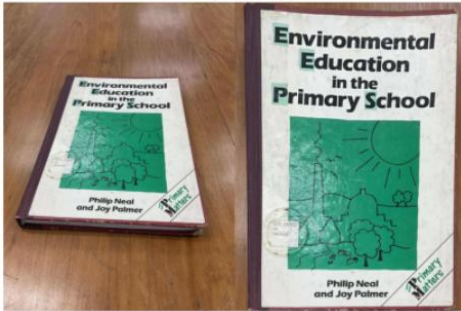


set03_orb_raw.jpg

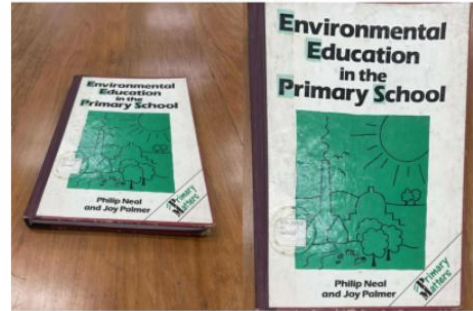


- Set04

set05_orb_pre.jpg

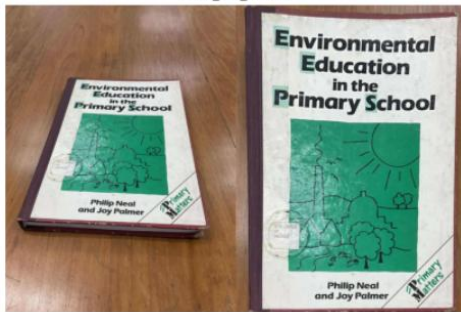


set05_orb_raw.jpg

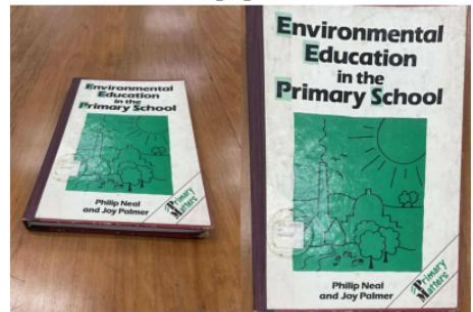


- Set05

set05_orb_pre.jpg

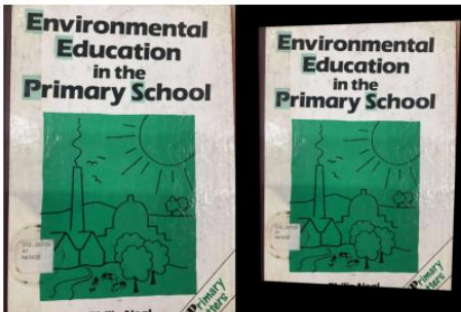


set05_orb_raw.jpg

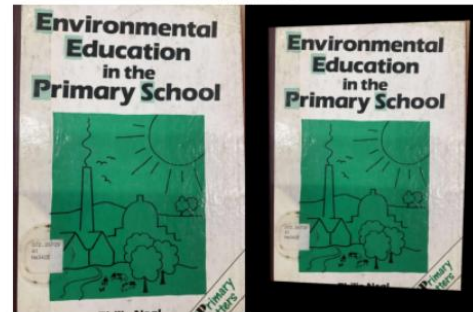


- Set06

set06_orb_pre.jpg

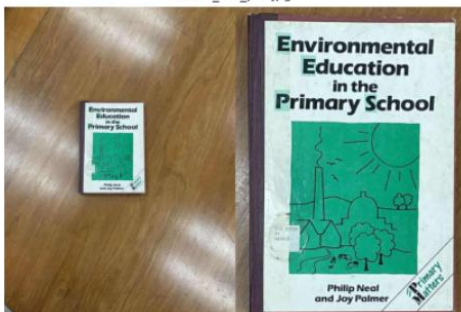


set06_orb_raw.jpg

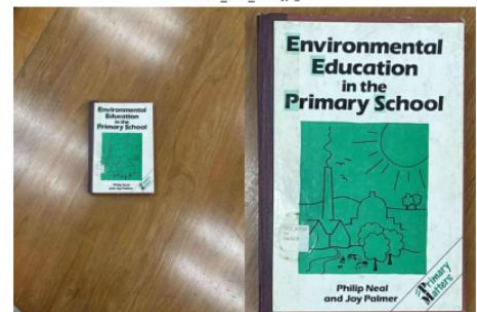


- Set07

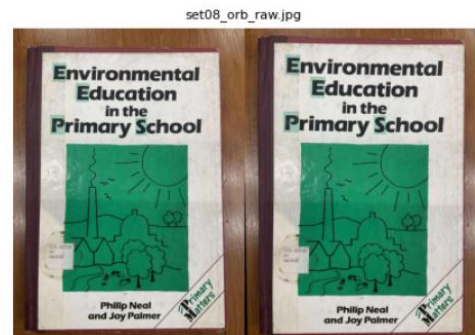
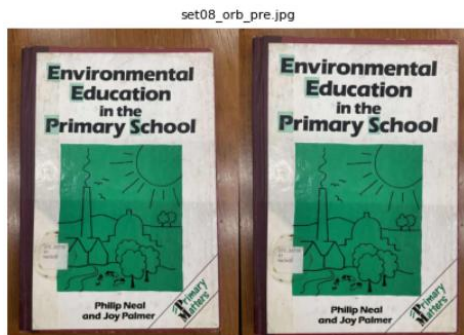
set07_orb_pre.jpg



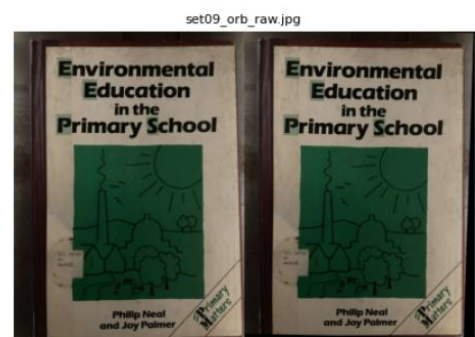
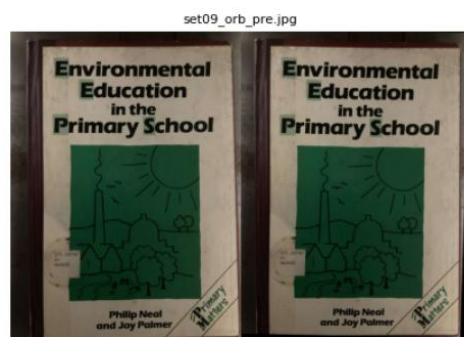
set07_orb_raw.jpg



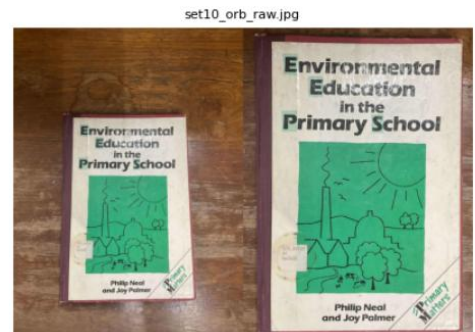
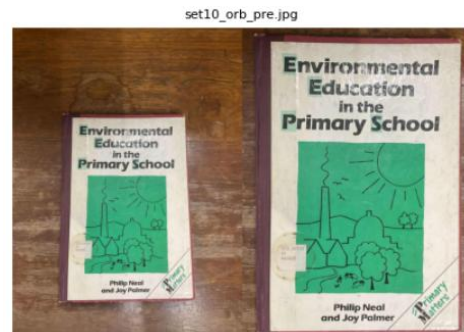
- Set08



- Set09



- Set10



8. การเปรียบเทียบประสิทธิภาพ (Performance Comparison)

จากการเปรียบเทียบระหว่างกรณีที่มีและไม่มีการทำ Preprocessing พบว่า การทำ Preprocessing ช่วยเพิ่มจำนวน inliers และอัตราส่วน inliers รวมถึงปรับปรุงค่า SSIM และ PSNR ได้อย่างมีนัยสำคัญ โดยเฉพาะในกรณีที่ภาพมีแสงน้อยหรือมีเงาทาบ

9. การอภิปรายผลและข้อจำกัด (Discussion & Limitations)

- การทำ Preprocessing มีผลชัดเจนในภาพที่มีความคมชัดต่ำหรือมีสัญญาณรบกวน
- วิธี Contour-based ให้ผลลัพธ์ที่ดีเมื่อขอบเอกสารชัดเจน แต่ไม่สามารถใช้งานได้ดีในกรณีที่พื้นหลังซับซ้อน
- วิธี ORB+RANSAC ให้ผลลัพธ์ที่ดีเมื่อภาพอ้างอิงและภาพต้นฉบับมีพีเจอร์ชัดเจน แต่จะมีข้อจำกัดเมื่อมีแสงสะท้อนหรือพีเจอร์น้อย
- ข้อจำกัดเพิ่มเติม ได้แก่ ภาพที่มีมุมเอียงมากเกินไป, เอกสารบางส่วนหายไป หรือพื้นหลังมีลักษณะใกล้เคียงกับเอกสาร

10. สรุปผลการดำเนินงาน (Conclusion)

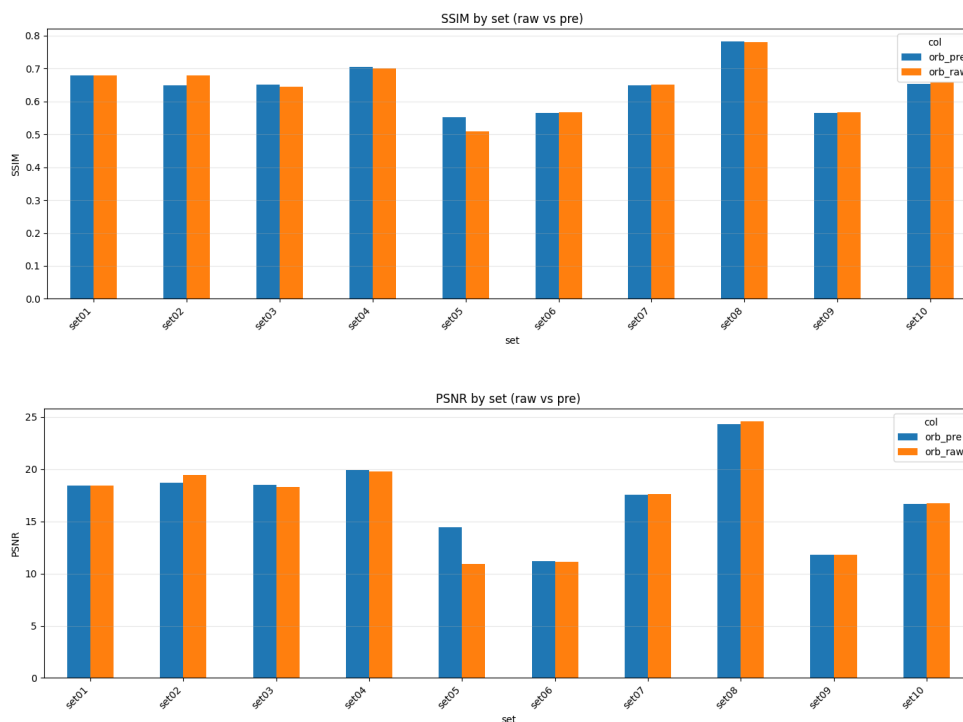
โครงการนี้นำเสนอการจัดแนวเอกสารด้วยเทคนิค Homography

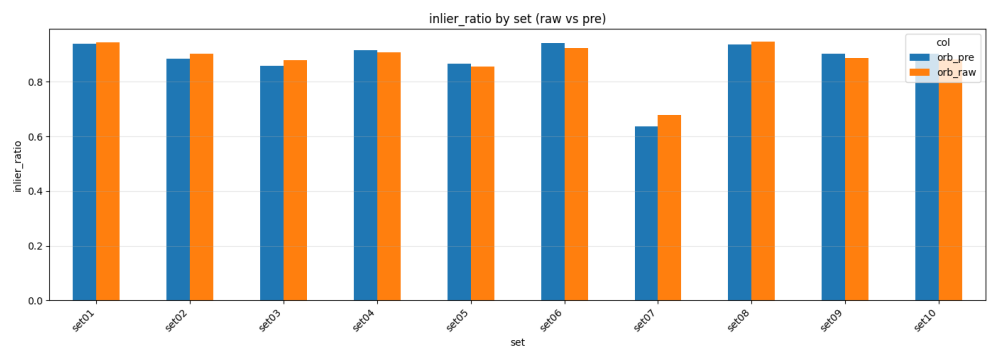
โดยเปรียบเทียบผลลัพธ์ระหว่างกรณีที่มีและไม่มีการทำ Preprocessing ด้วยวิธี Contour-based และ ORB+RANSAC

ผลการทดลองแสดงให้เห็นว่า Preprocessing มีส่วนช่วยให้ผลลัพธ์ดีขึ้นอย่างชัดเจน และ pipeline

ที่พัฒนาขึ้นสามารถนำไปประยุกต์ใช้ในงานจริง เช่น การสแกนเอกสารผ่านกล้องโทรศัพท์มือถือ

กราฟสรุปผลจาก outputs/results.csv





ขอบคุณครับ