

IMAGE RECOGNITION WITH PYTHON

Computer Education

LECTURER

**JIRAWAT
CHUAYKHUN**

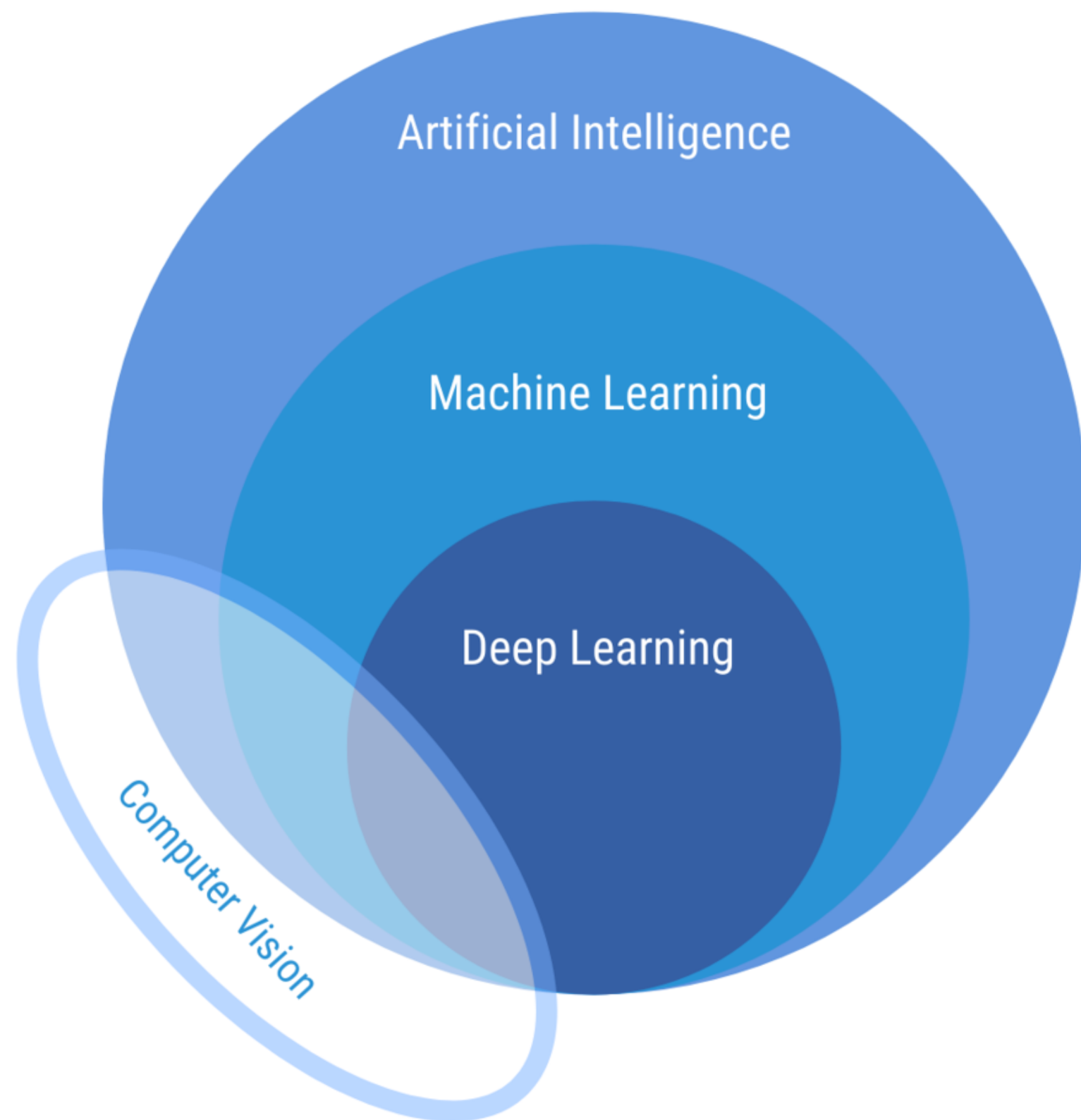
4th year Student



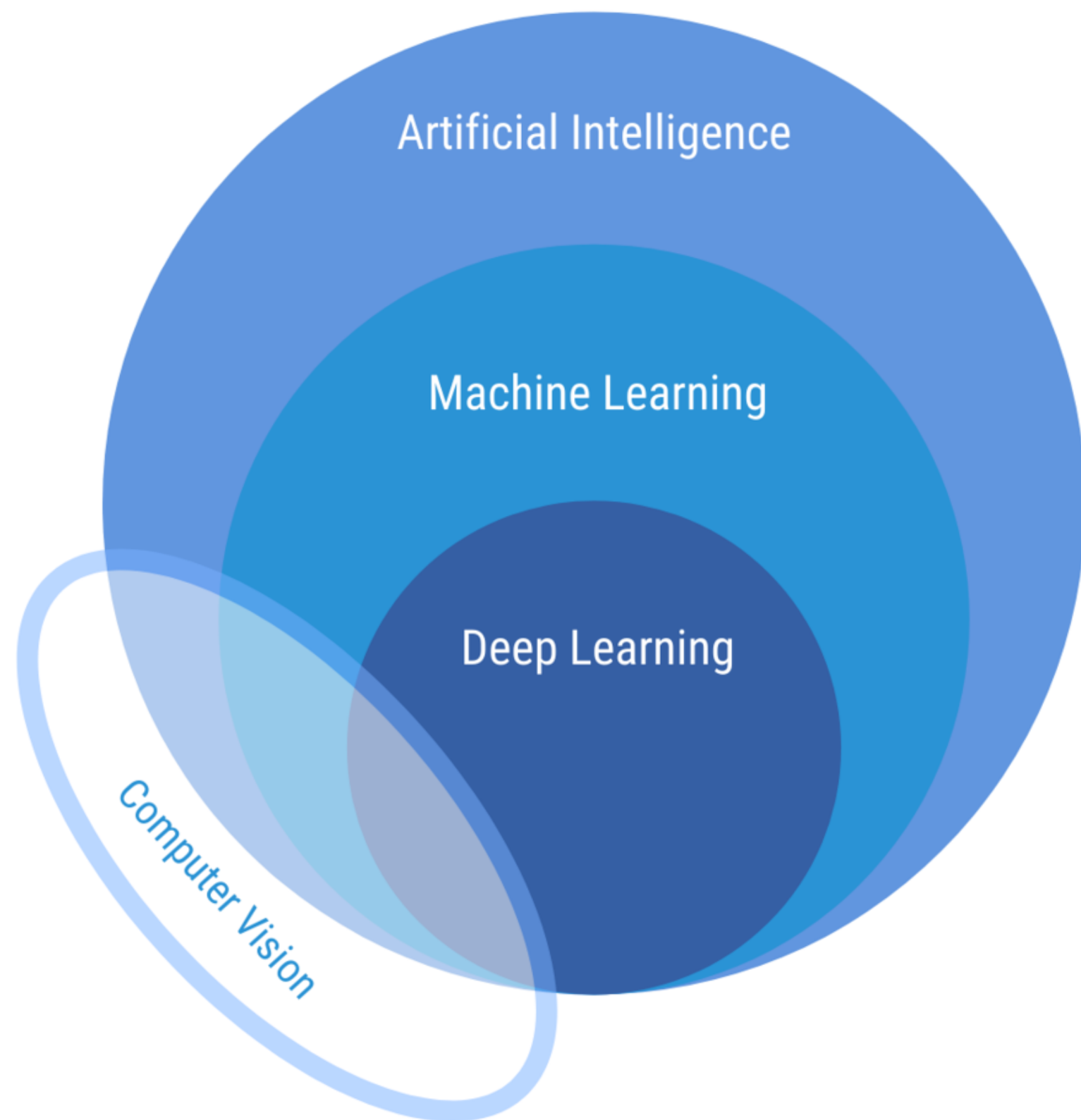
LECTURER
POONPIPAT
KONGTHONG

4th year Student



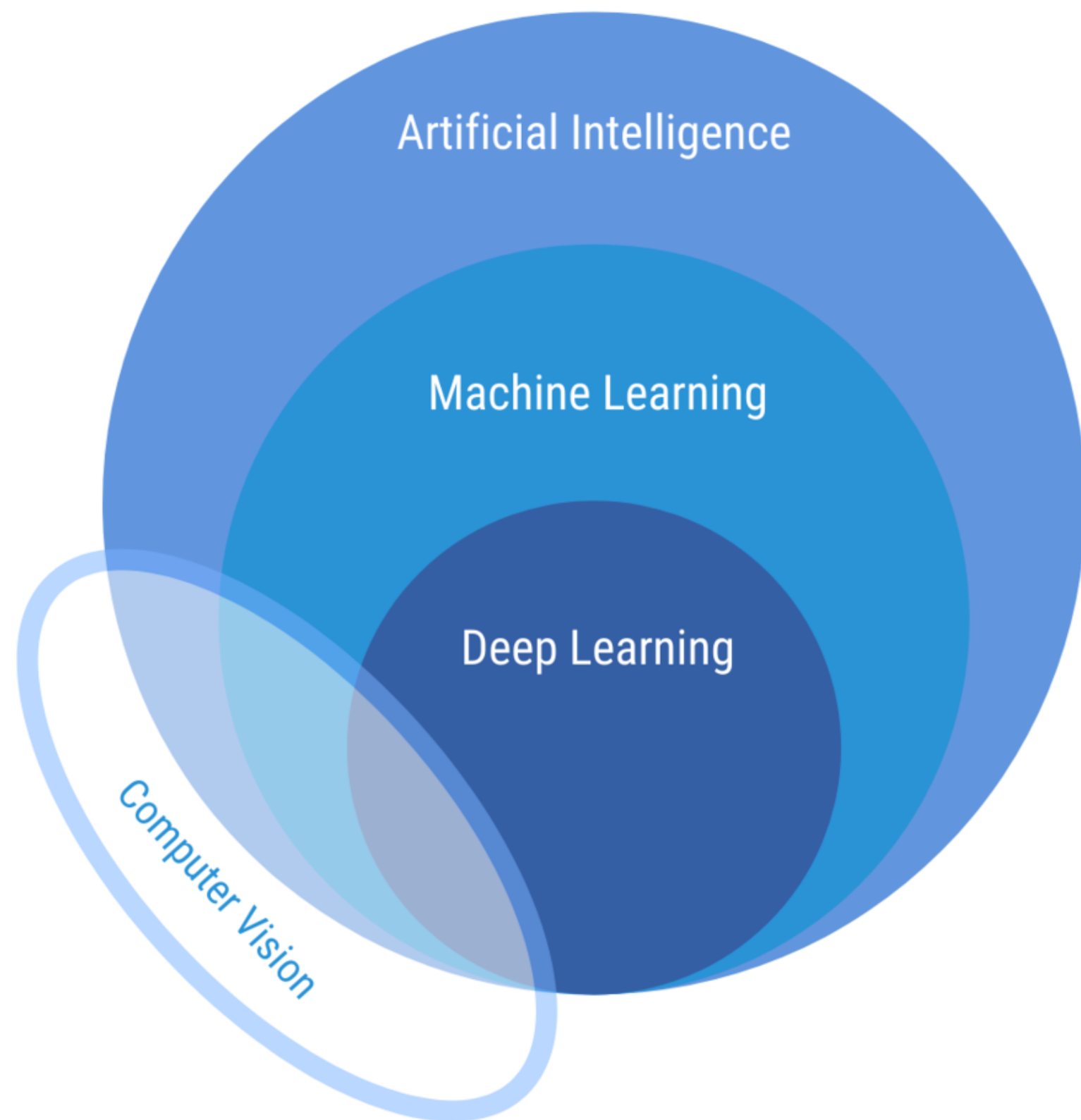


ปัญญาประดิษฐ์ หรือ (AI)
การรวมความฉลาดของมนุษย์สู่
เครื่องจักร (Machine)



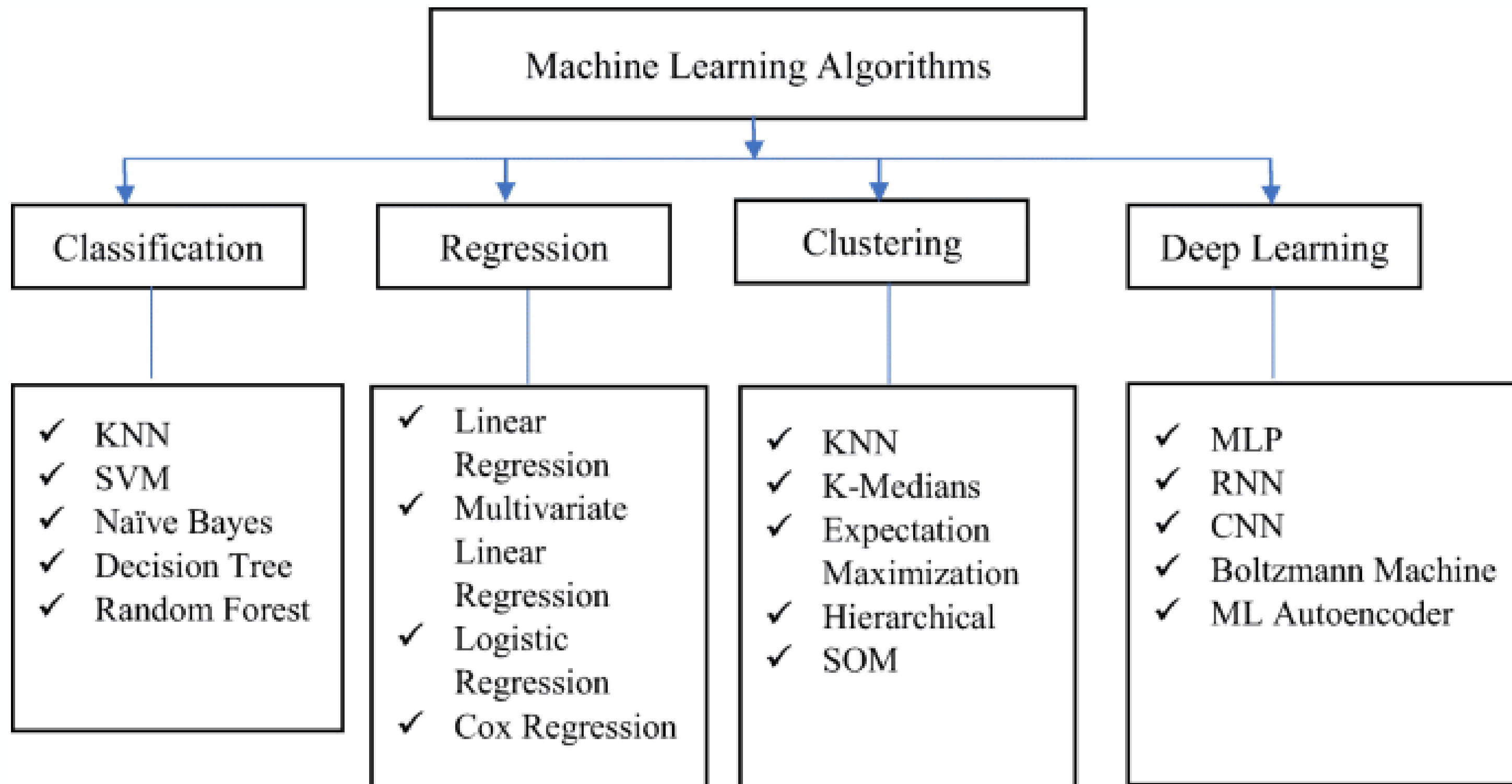
Machine Learning (ML)

การสอนให้ระบบคอมพิวเตอร์ทำการ
เรียนรู้ได้ด้วยตนเองโดยการใช้
'ข้อมูล'



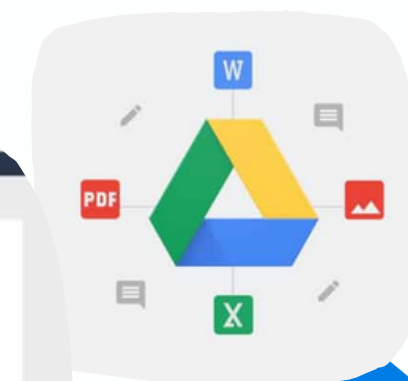
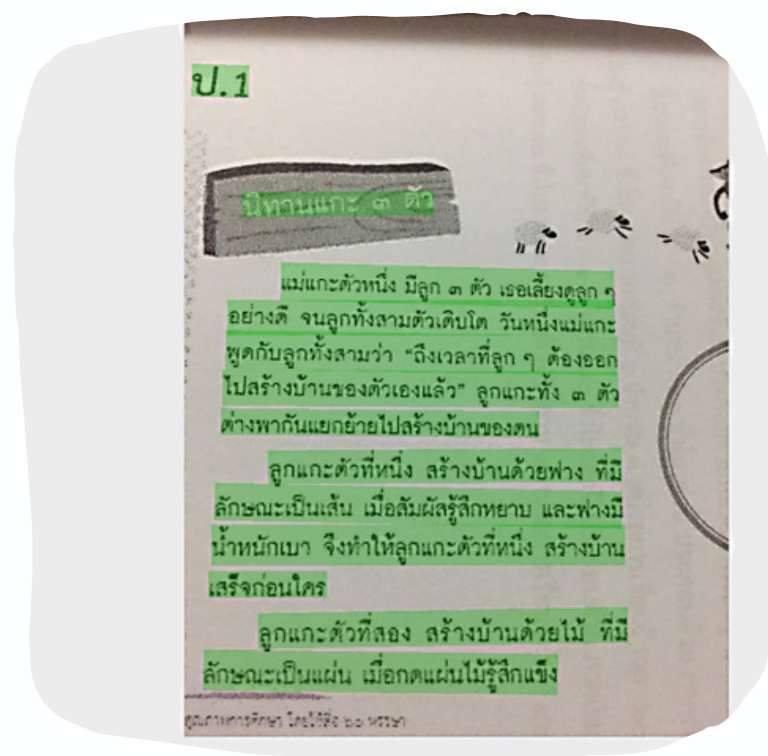
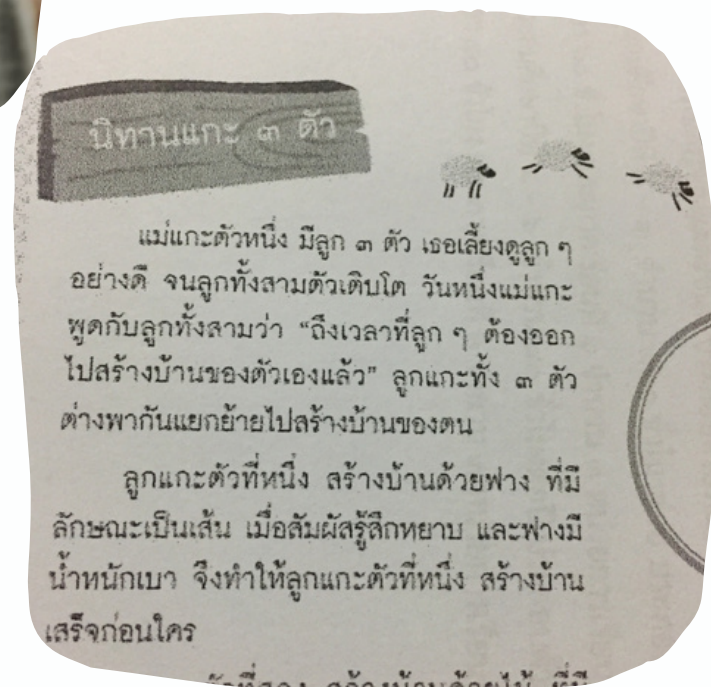
Deep Learning

อัลกอริทึมแบบระบบเรียนรู้เชิงลึก
การเรียนรู้แบบอัตโนมัติด้วยการ เลียน
แบบการทำงานของโครงข่ายประสาท
ของมนุษย์



Optical Character Recognition

การรู้จำอักขระด้วยแสง



1 เทคนิคการรู้จำตัวอักษรด้วยแสง คือ ?

2 ประเภทของเทคนิคการรู้จำตัวอักษรด้วยแสง

3 ขั้นตอนการประมวลผล

4 เครื่องมือ OCR Open source

เทคนิคการรู้จำตัวอักษรด้วยแสง คือ ?

สแกนเอกสาร



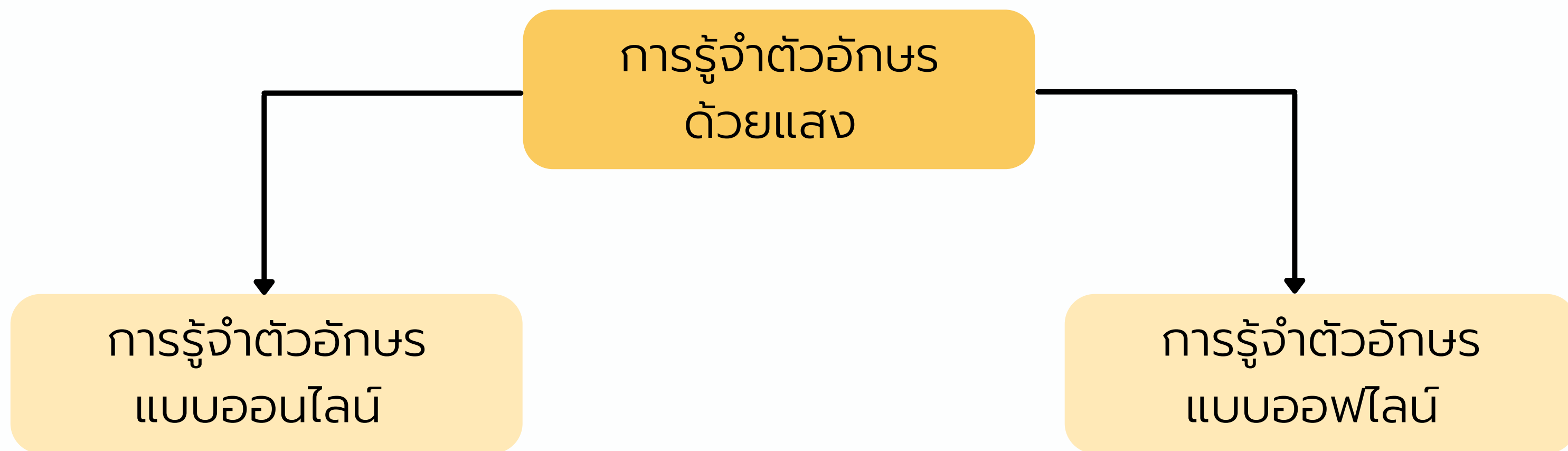
เอกสาร pdf



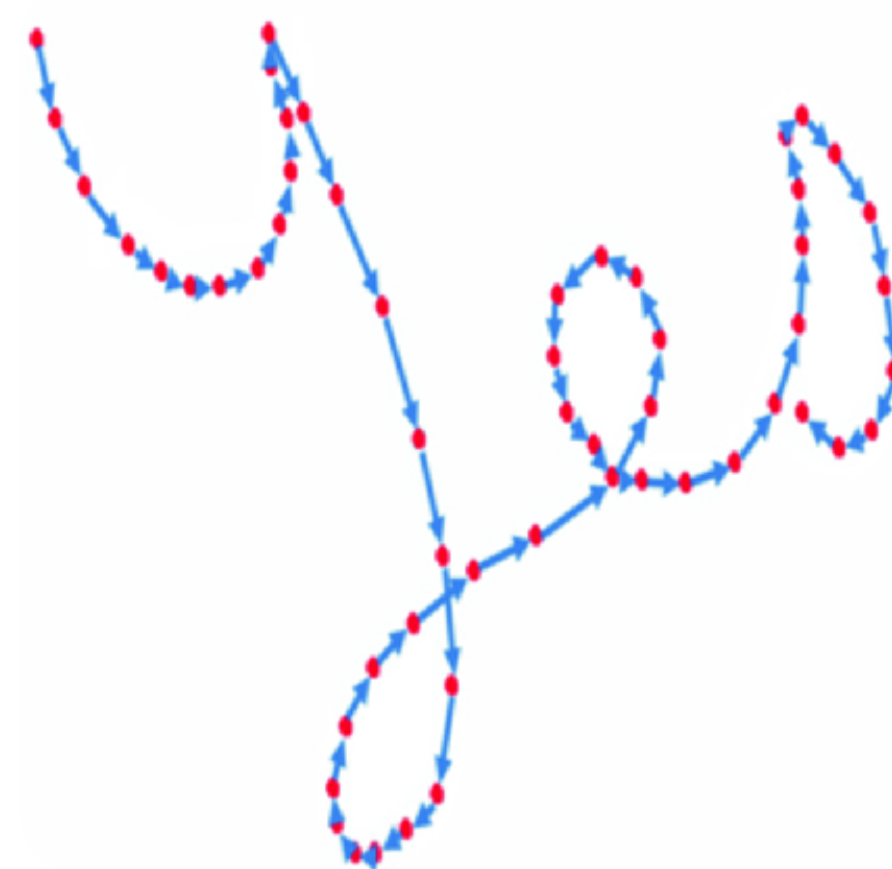
รูปภาพ



ประเภทของเทคนิค การรู้จำตัวอักษรด้วยแสง



การรู้จำตัวอักษรแบบออนไลน์

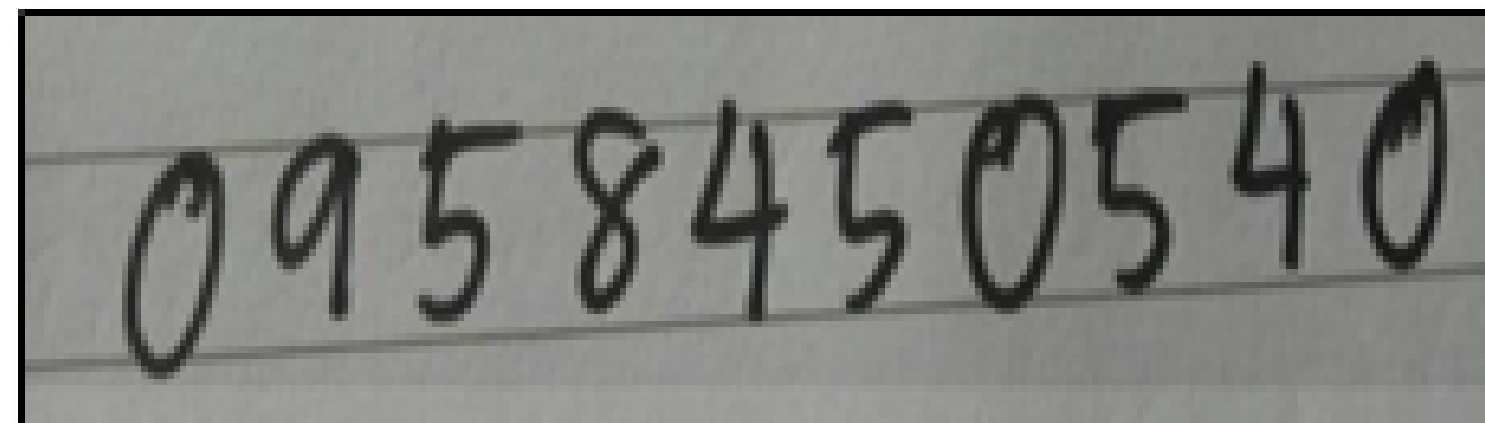


การเรียนรู้ตัวอักษรแบบออฟไลน์

ตัวอักษรแบบโดด (Single Character)

A B C D E F G H I J K
L M N O P Q R S ... Z

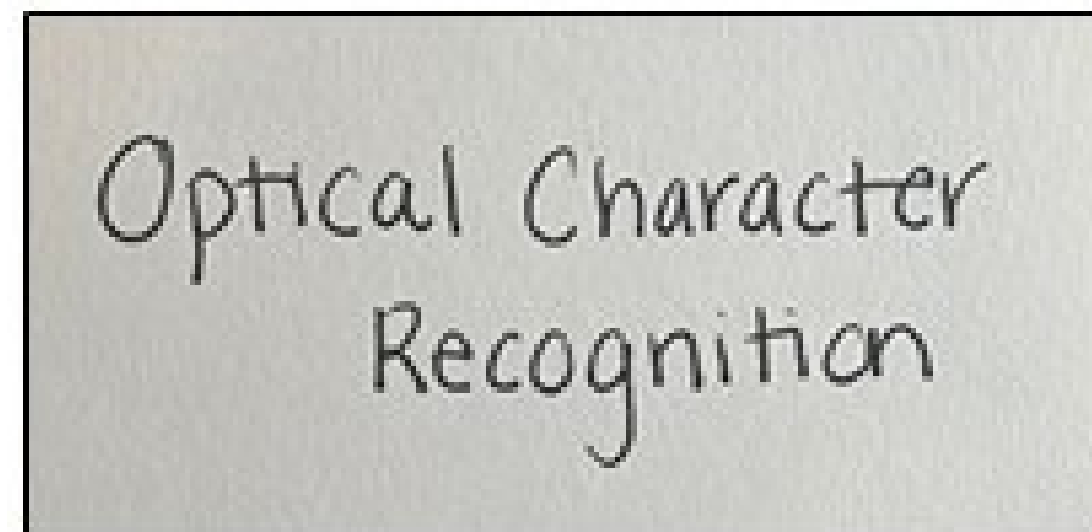
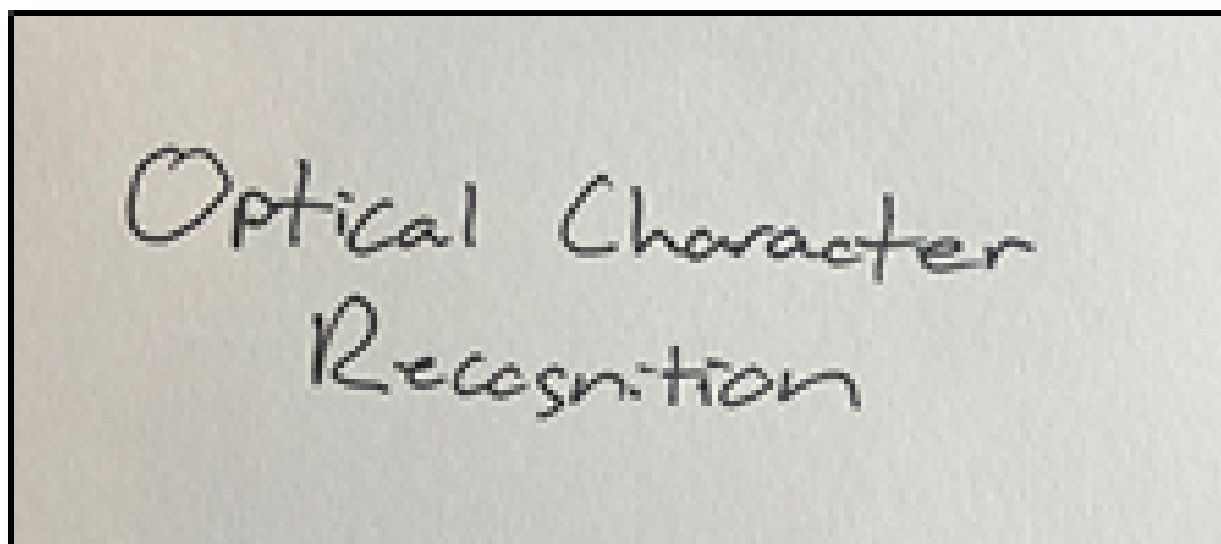
ตัวพิมพ์แบบฟอนต์เฉพาะ



ลายมือเขียนแบบตัวโดด

การเรียนรู้ตัวอักษรแบบออฟไลน์

ตัวอักษรลายมือแบบเขียนต่อเนื่อง (Script recognition)



ขั้นตอนการประมวลผลของ เทคนิคการรู้จำตัวอักษรด้วยแสง



1

ขั้นตอนการประมวลผลขั้นต้น (Pre-Processing)

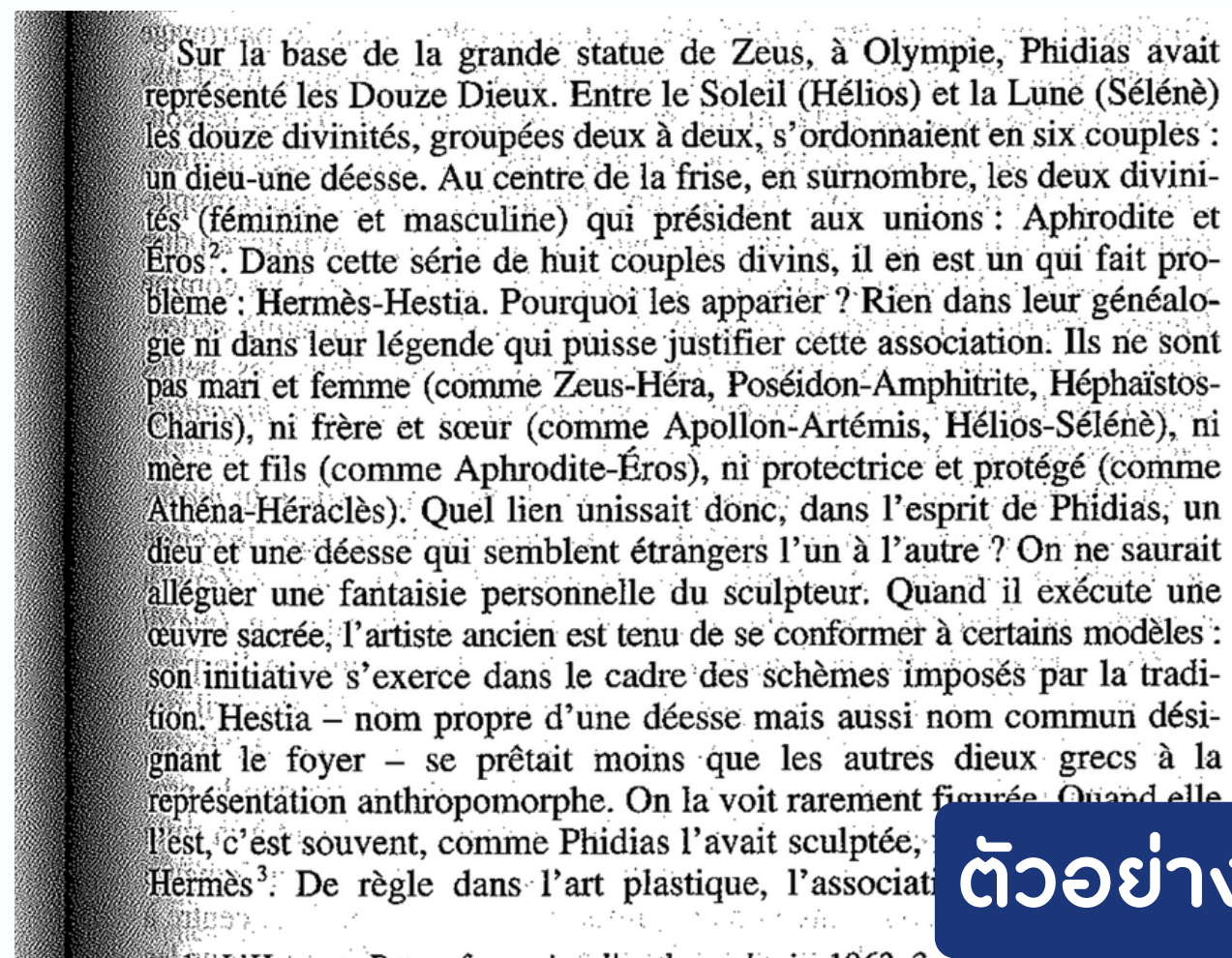
การลดสัญญาณรบกวน (Noise Filtering)

เป็นการกรองข้อมูลภาพที่มีสิ่งแทรกซ้อนที่ไม่ต้องการ
โดยมีจุดประสงค์เพื่อลดความผิดพลาด

1

ขั้นตอนการประมวลผลขั้นต้น (Pre-Processing)

Noise คือ จุดรบกวนในภาพที่เกิดจาก การปรับเพิ่มค่าความไวแสง



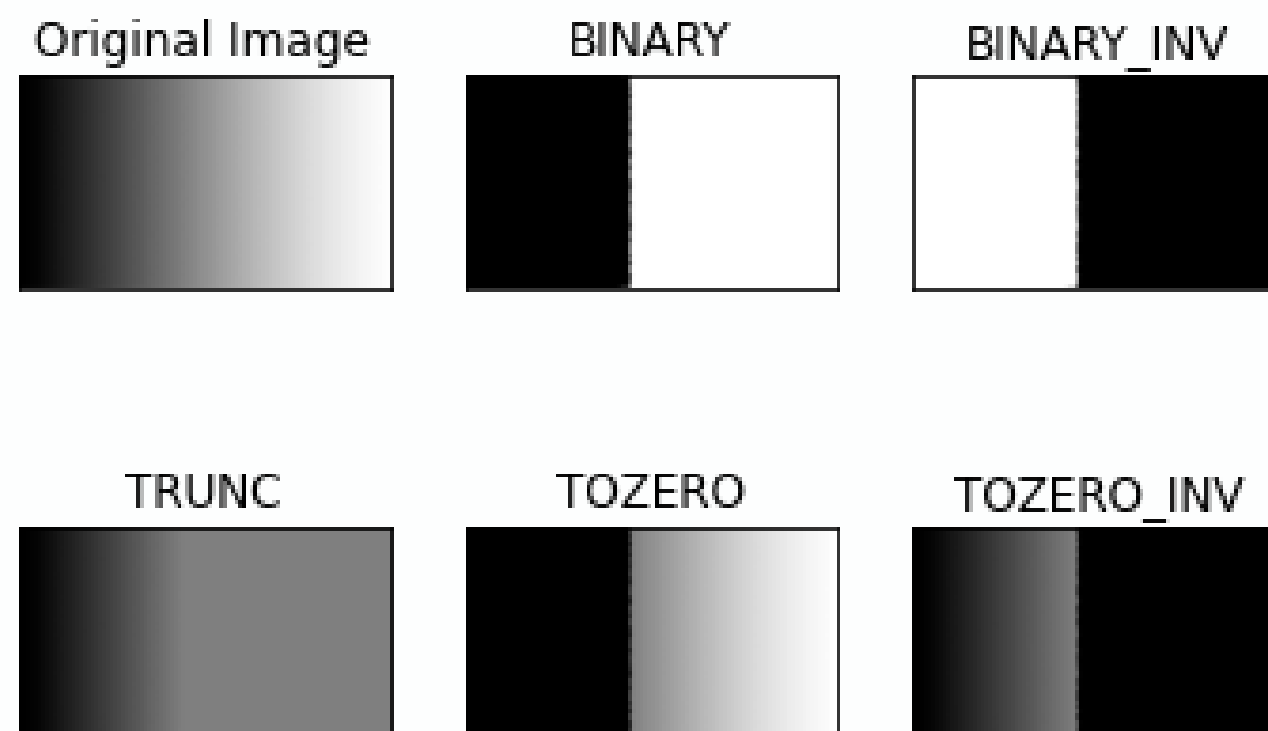
ตัวอย่างภาพที่เกิดจุดรบกวน

1

ขั้นตอนการประมวลผลขั้นต้น (Pre-Processing)

ฟังก์ชัน `cv2.threshold()`

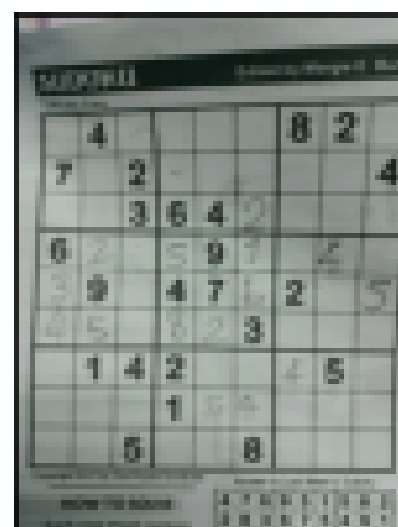
ใช้คัดกรองภาพโดยพิจารณาความสว่าง เพื่อที่จะคัดเอาหรือปรับค่าบางส่วน



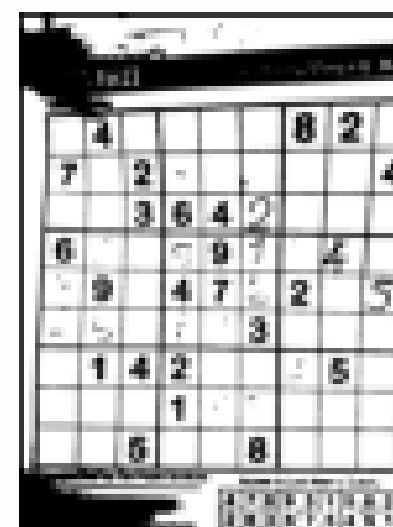
1

ขั้นตอนการประมวลผลขั้นต้น (Pre-Processing)

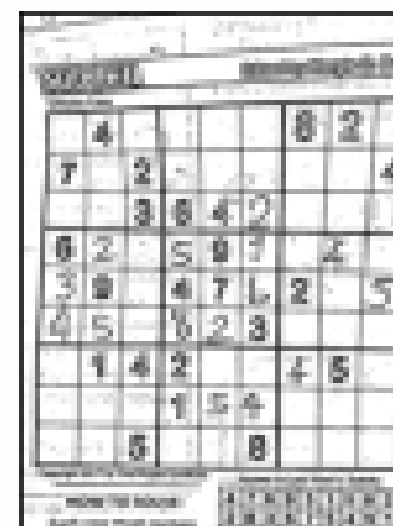
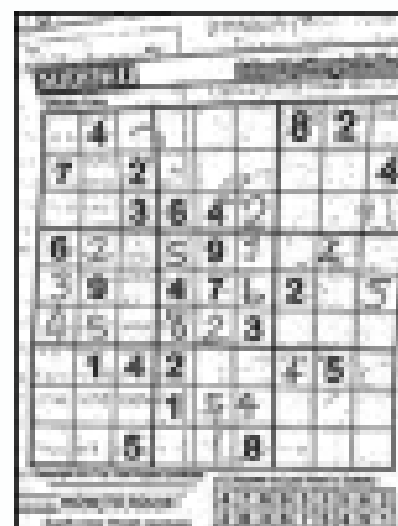
Original Image



Global Thresholding ($\tau = 127$)



Adaptive Mean Thresholding Adaptive Gaussian Thresholding



1

ขั้นตอนการประมวลผลขั้นต้น (Pre-Processing)

Original



Global Thresholding



Otsu_s_Thresholding



Gaussian + Otsu



1

ขั้นตอนการประมวลผลขั้นต้น (Pre-Processing)

Original



Global Thresholding



Otsu_s_Thresholding



Gaussian + Otsu



1

ขั้นตอนการประมวลผลขั้นต้น (Pre-Processing)

Skew correction

ตรวจหาค่าความเอียงและการหมุนกลับของภาพเอกสาร

Our last argument is how we want to approximate the contour. We use `cv2.CHAIN_APPROX_SIMPLE` to compress horizontal, vertical, and diagonal segments into their endpoints only. This saves both computation and memory. If we wanted all the points along the contour, without compression, we can pass in `cv2.CHAIN_APPROX_NONE`; however, retrieving all points along a contour is often unnecessary and is wasteful of resources.

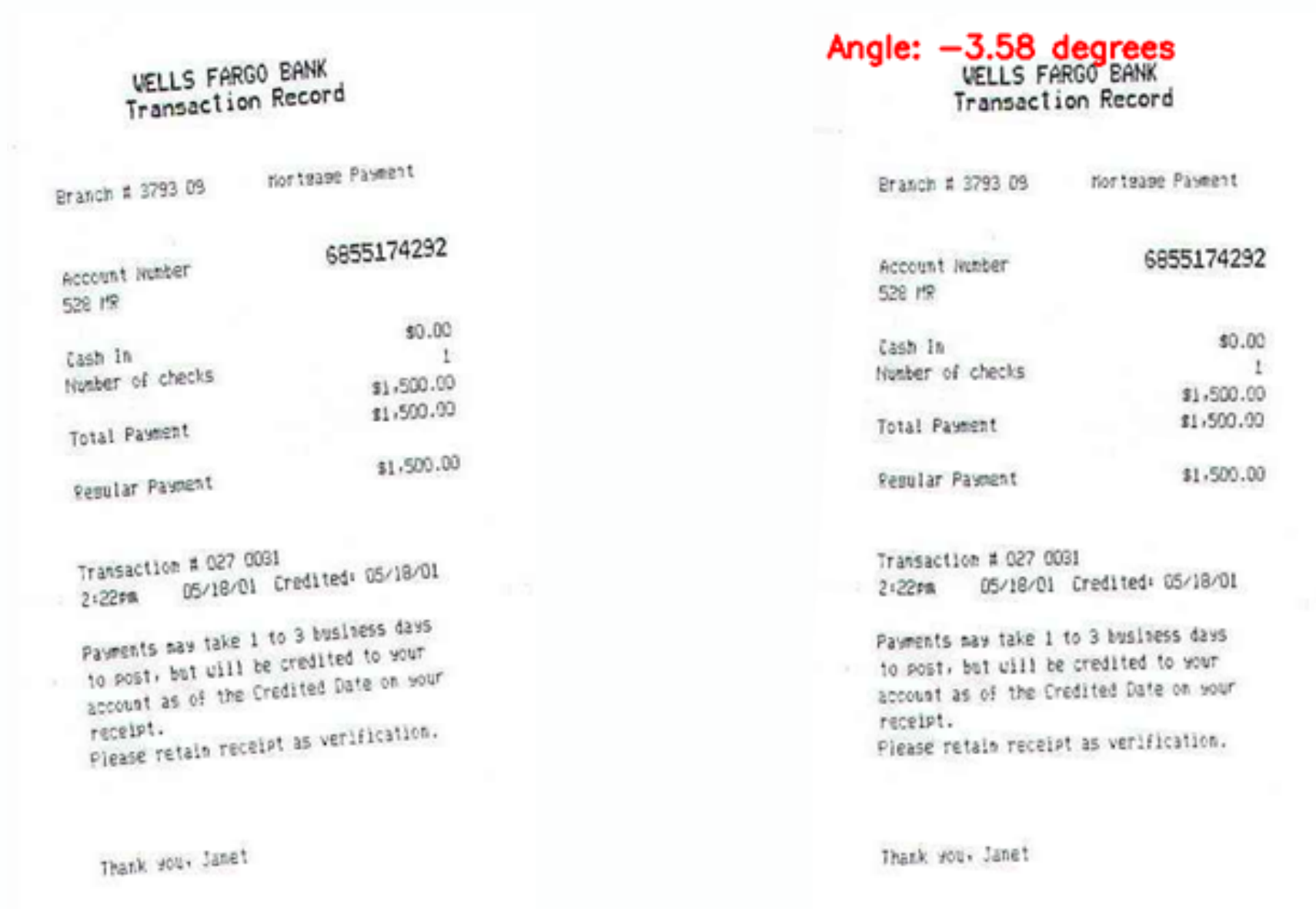
Our last argument is how we want to approximate the contour. We use `cv2.CHAIN_APPROX_SIMPLE` to compress horizontal, vertical, and diagonal segments into their endpoints only. This saves both computation and memory. If we wanted all the points along the contour, without compression, we can pass in `cv2.CHAIN_APPROX_NONE`; however, retrieving all points along a contour is often unnecessary and is wasteful of resources.

Our last argument is how we want to approximate the contour. We use `cv2.CHAIN_APPROX_SIMPLE` to compress horizontal, vertical, and diagonal segments into their endpoints only. This saves both computation and memory. If we wanted all the points along the contour, without compression, we can pass in `cv2.CHAIN_APPROX_NONE`; however, retrieving all points along a contour is often unnecessary and is wasteful of resources.

Our last argument is how we want to approximate the contour. We use `cv2.CHAIN_APPROX_SIMPLE` to compress horizontal, vertical, and diagonal segments into their endpoints only. This saves both computation and memory. If we wanted all the points along the contour, without compression, we can pass in `cv2.CHAIN_APPROX_NONE`; however, retrieving all points along a contour is often unnecessary and is wasteful of resources.

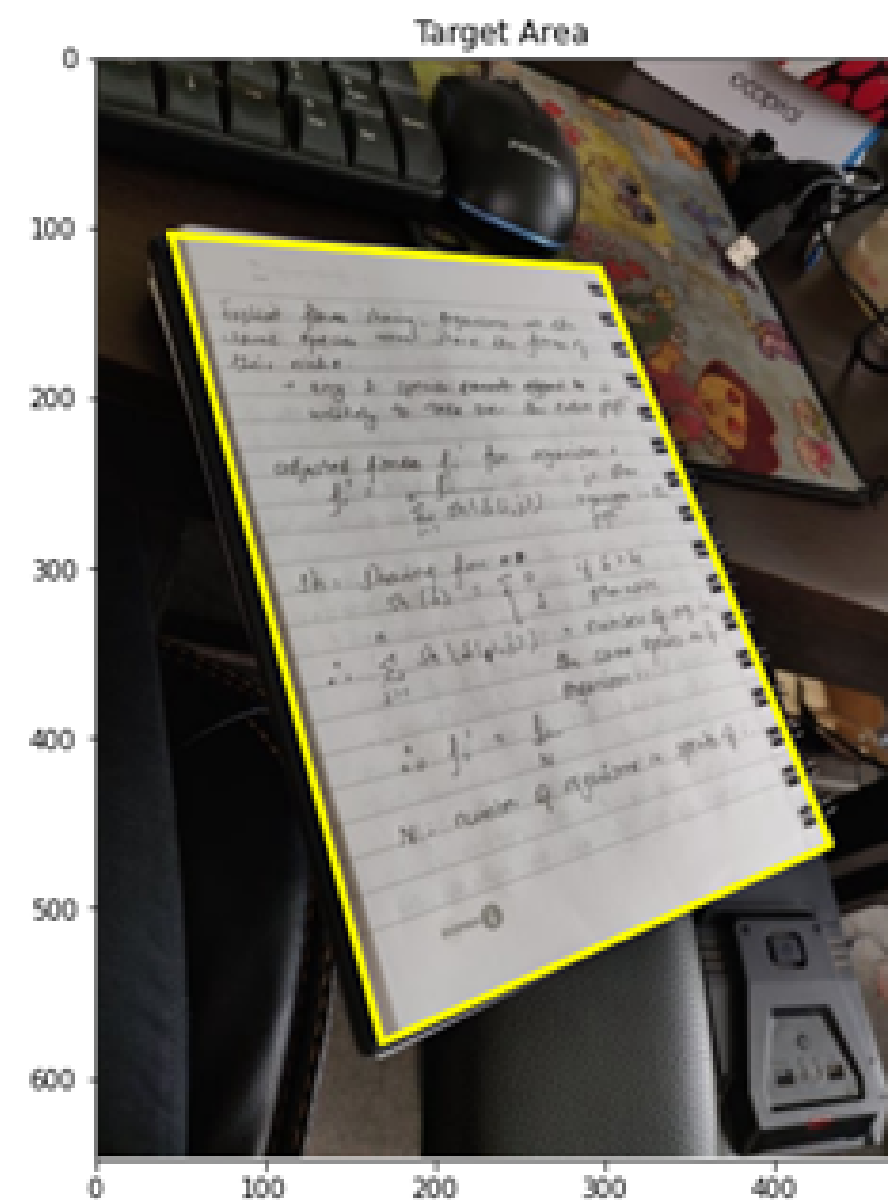
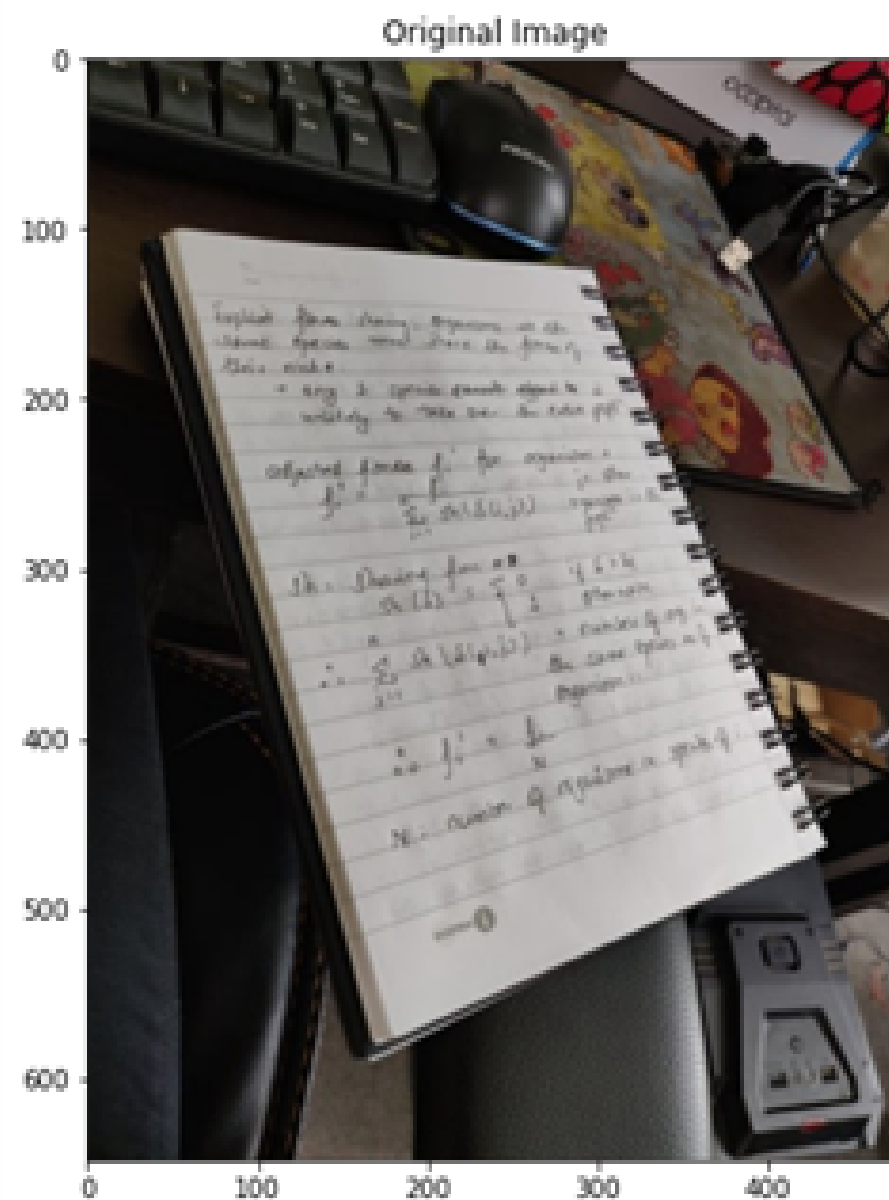
1

ขั้นตอนการประมวลผลขั้นต้น (Pre-Processing)



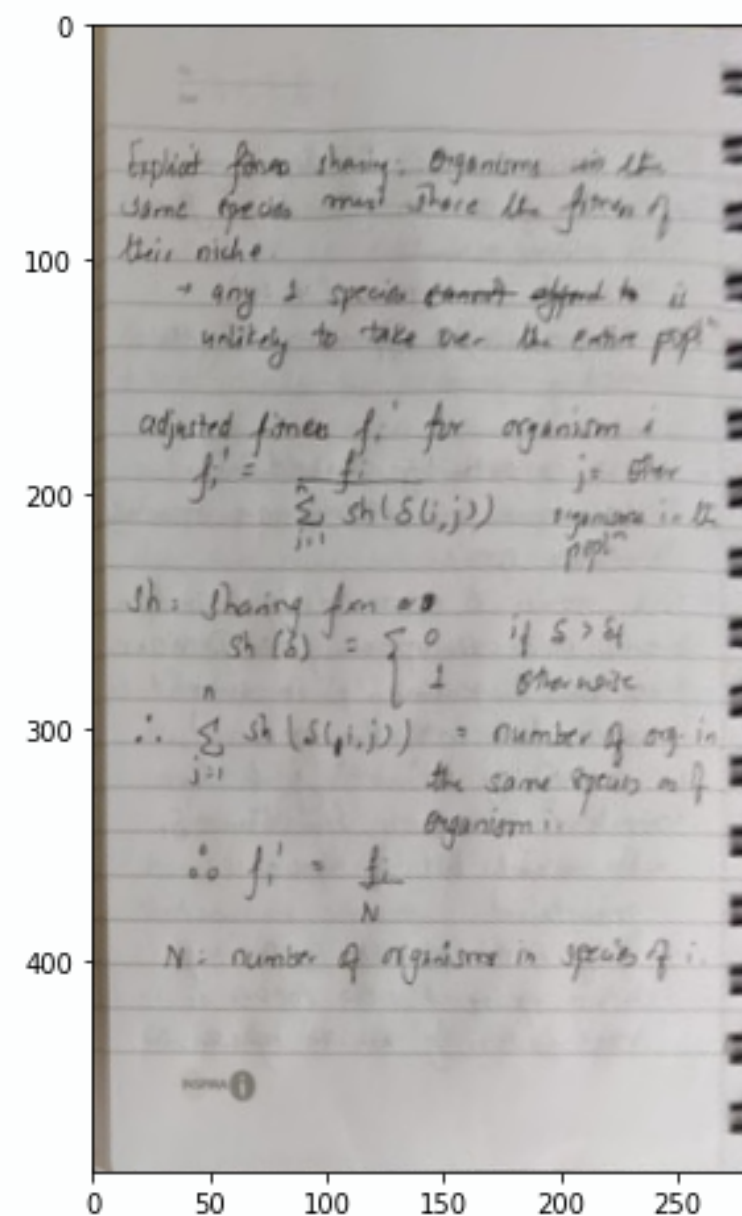
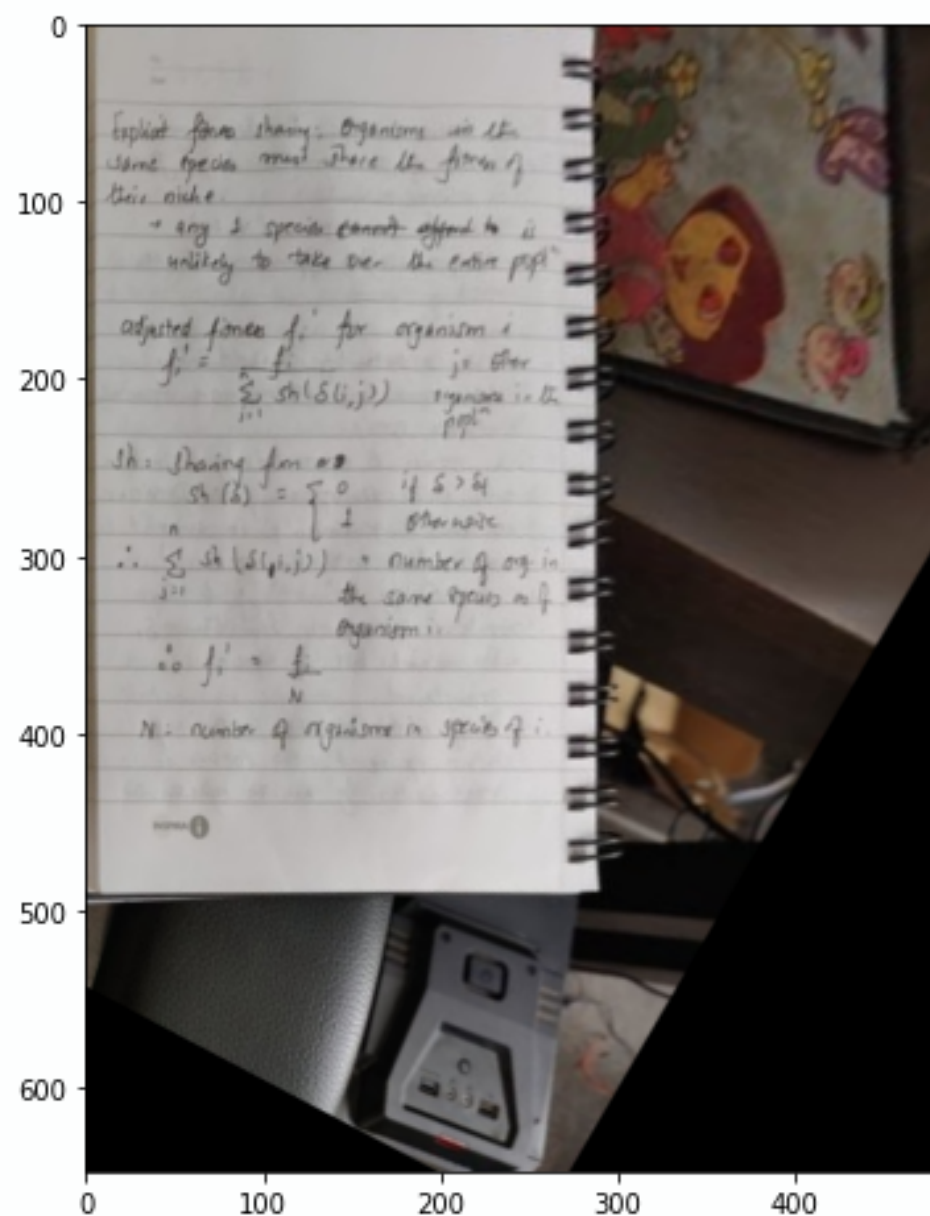
1

ขั้นตอนการประมวลผลขั้นต้น (Pre-Processing)



1

ขั้นตอนการประมวลผลขั้นต้น (Pre-Processing)



2 การรู้จำ (Recognition)

1

วิธีการวิเคราะห์ทางโครงสร้าง
(Structural Analysis)

2

แกรมมาแมชชีน
(formal grammar machine)

3

วิธีการรู้จำด้วยการเปรียบเทียบคู่รูปแบบ
(Template Matching)

4

โครงข่ายประสาทเทียม
(Neural Network)

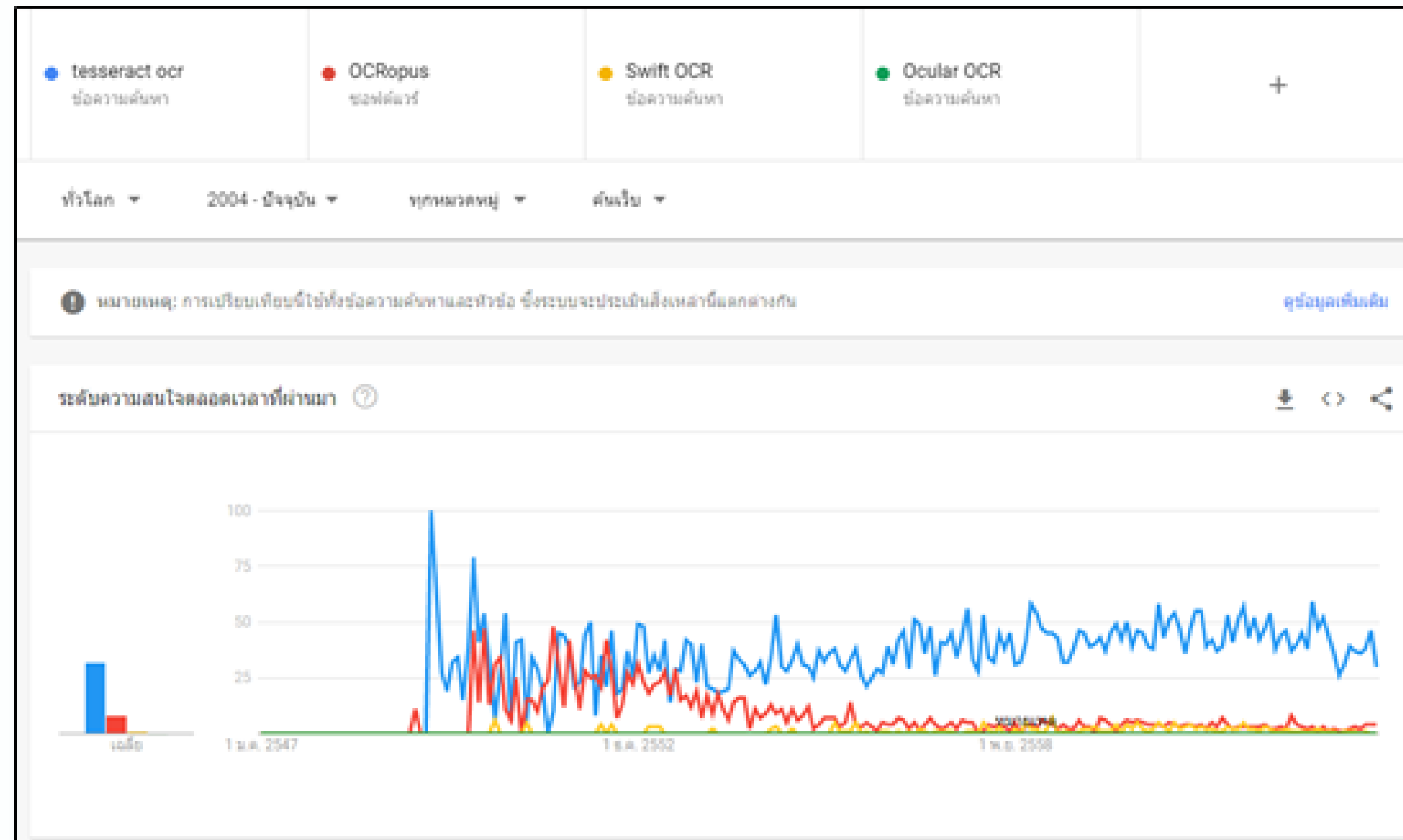
3

ขั้นตอนการประมวลผลขั้นสุดท้าย (Post-Processing)

การประมวลผลขั้นตอนสุดท้าย คือ
การตรวจสอบและแก้ไขผลจากการรู้จำรูปตัวอักษร
ที่ผ่านกระบวนการรู้จำมีผลเป็นรหัสตัวอักษร

เช่น ไวยากรณ์ภาษา

เครื่องมือ OCR Open source



เครื่องมือ OCR Open source Tools

1



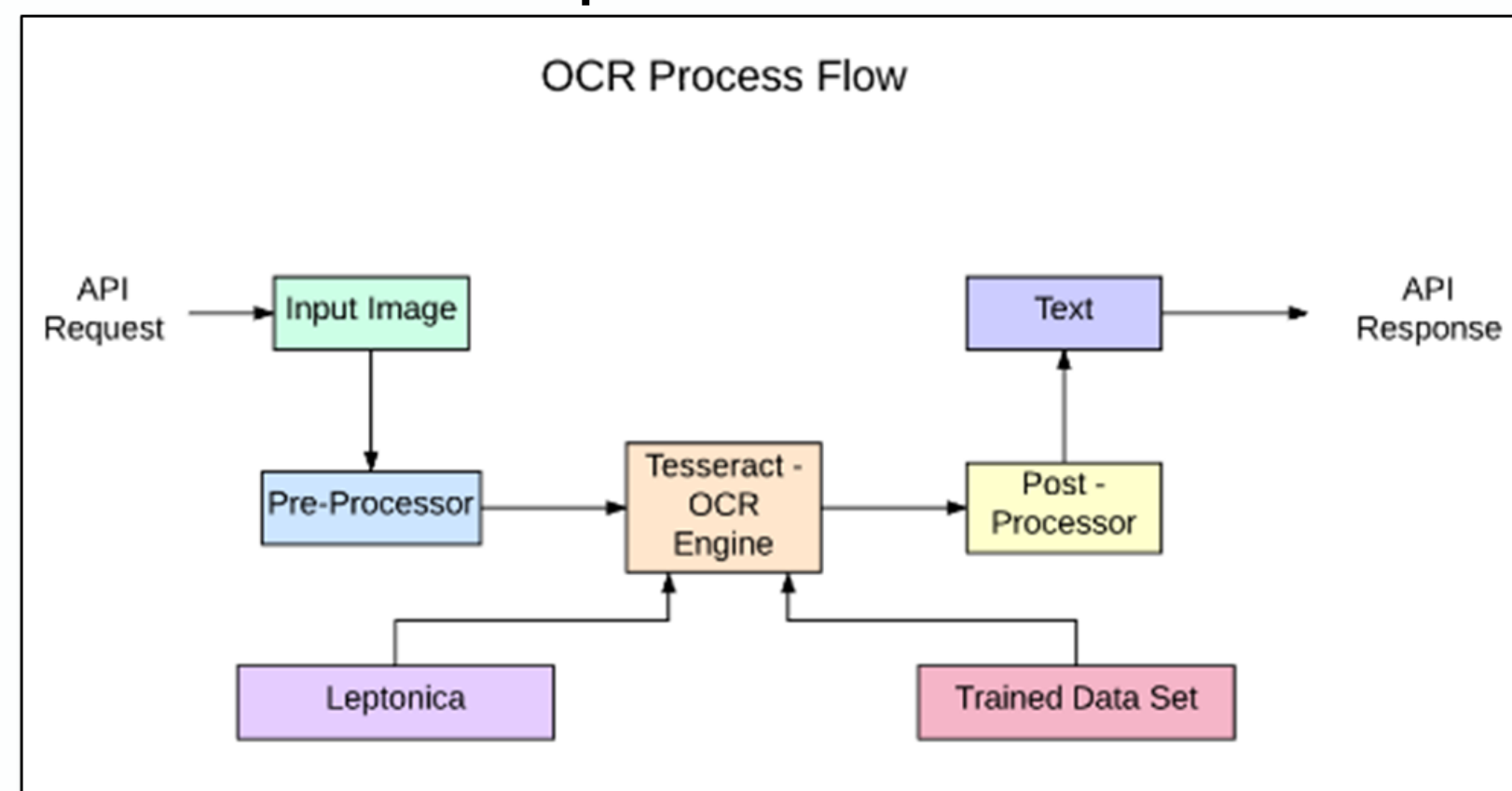
Tesseract OCR
| Python

2

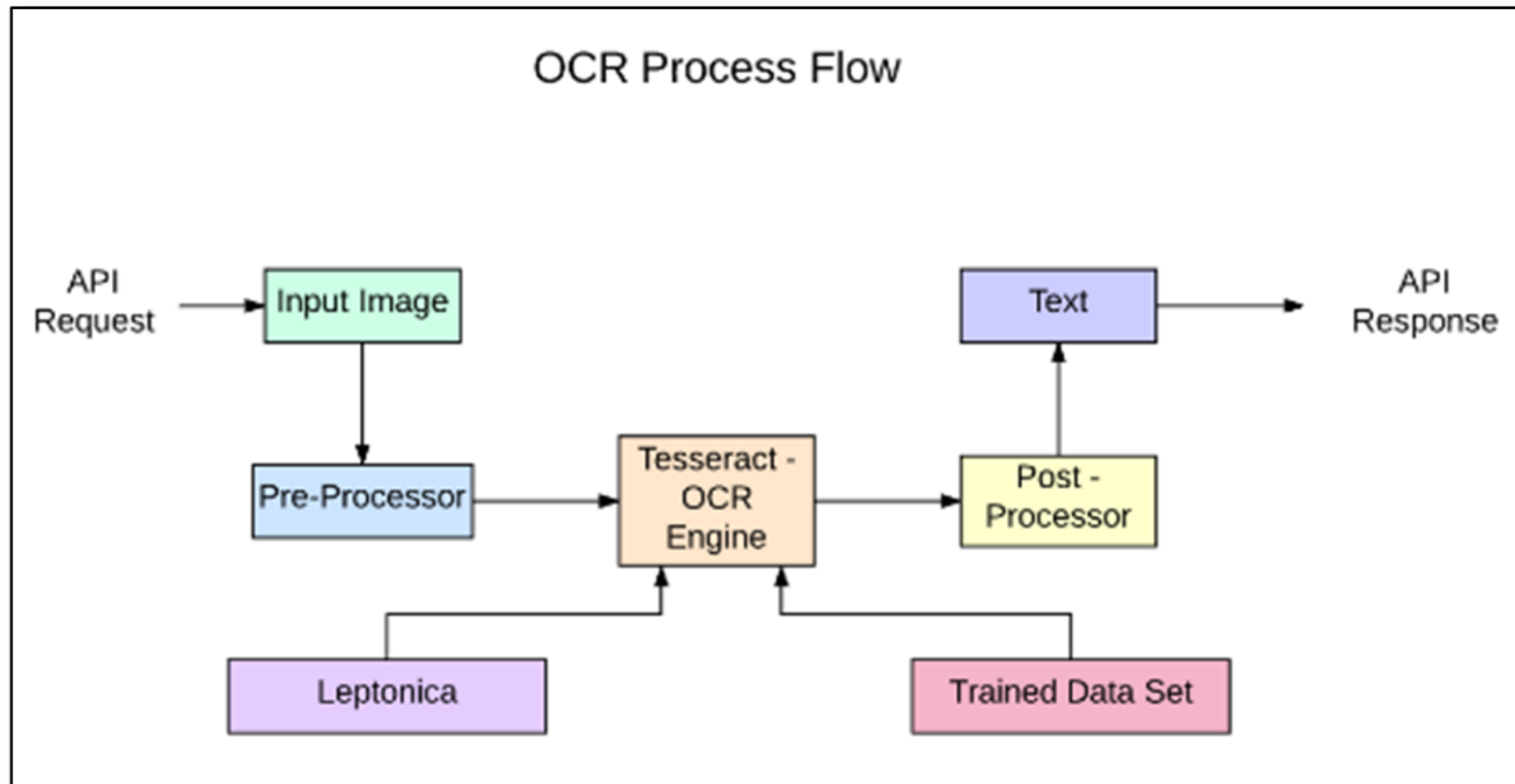


Tesseract OCR คืออะไร ?

Tesseract เป็นเอ็นจินการรู้จำข้อความ Open source (OCR) ซึ่งอยู่ภายใต้ลิขสิทธิ์ Apache 2.0 ปล่อย่อยอย่างเป็นทางการ ปัจจุบัน คือ 4.1.1



Tesseract OCR คืออะไร ?



Tesseract OCR

1984

PhD project by
HP

2005

Open Source
by HP

2006

Take Over by
Google

version 1

2006

Expanded to 6
langs

version 2

2010

Layout analysis
60 langs

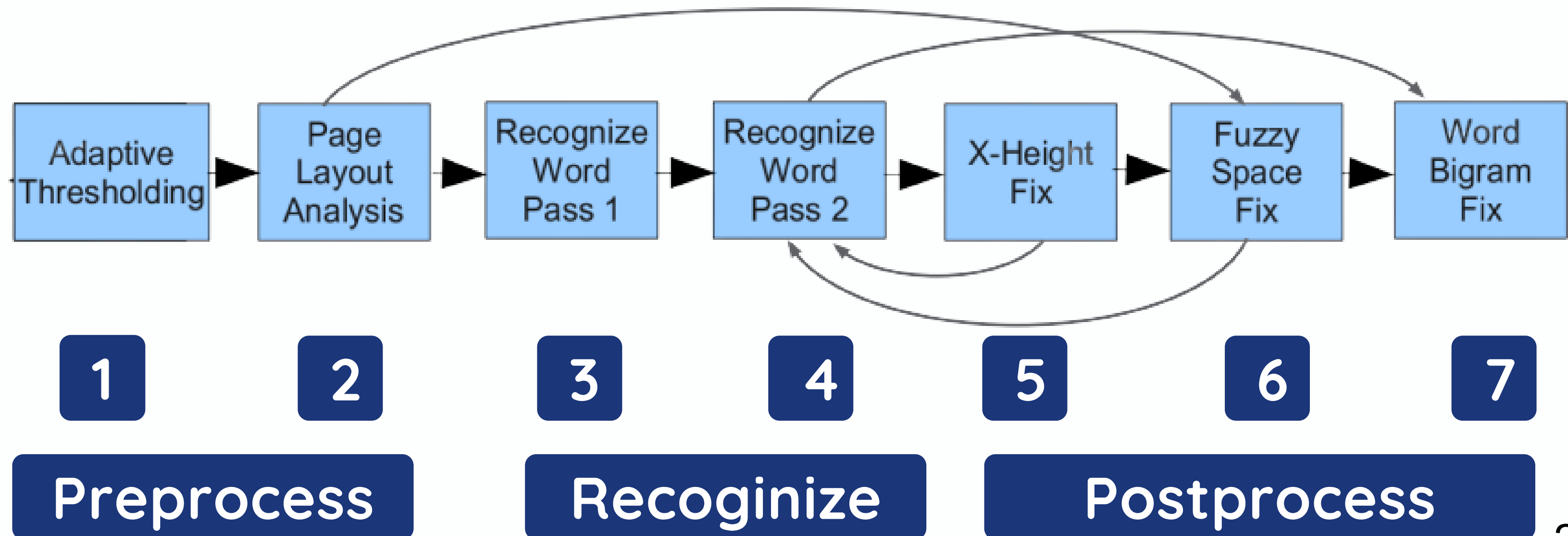
version 3

2016

LSTM
100 langs

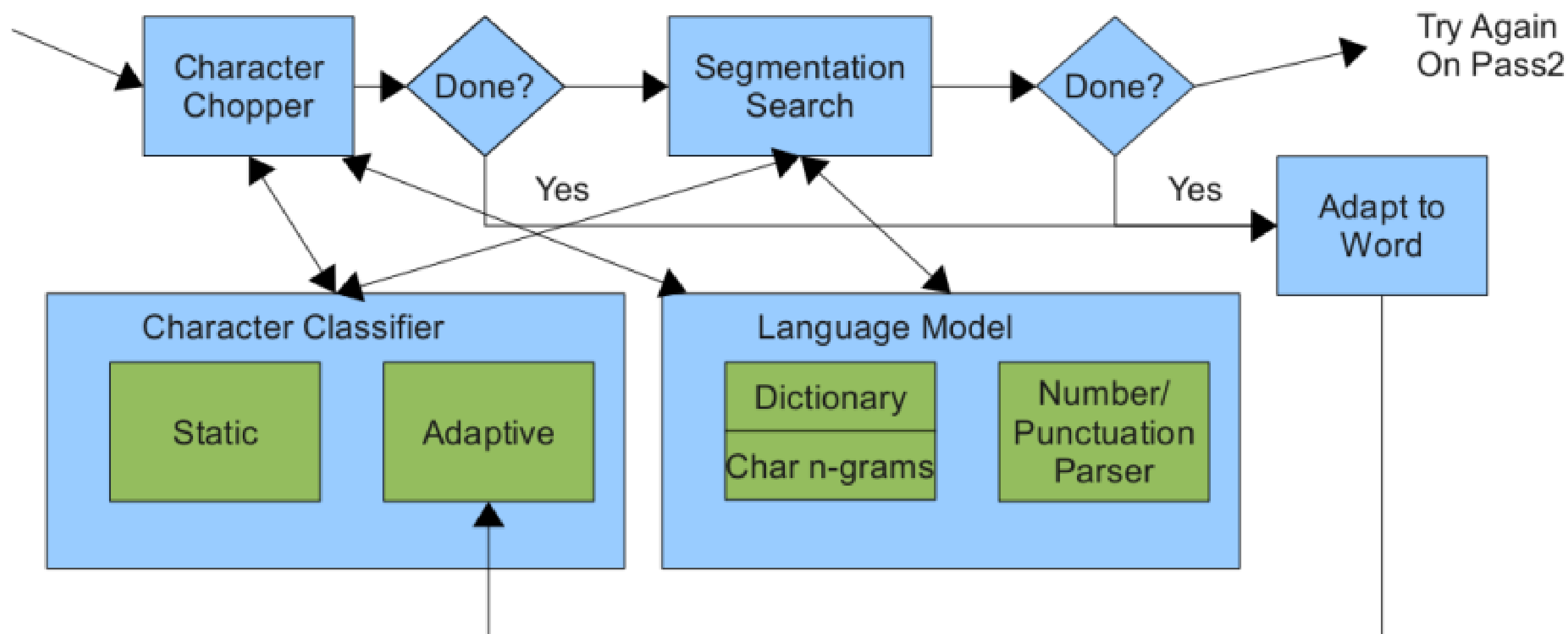
version 4

โครงสร้างการประมวลผล



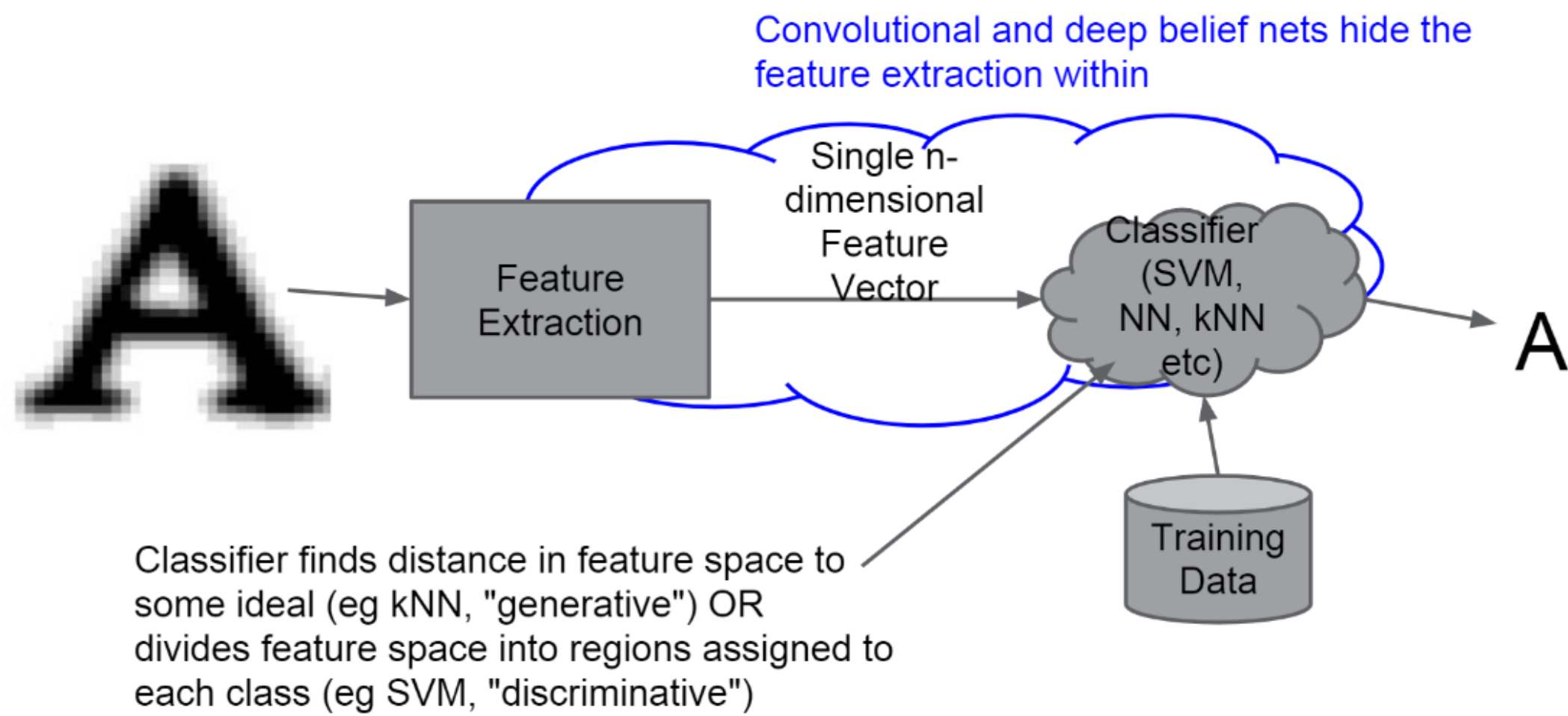
โครงสร้างการประมวลผล

Tesseract Word Recognizer

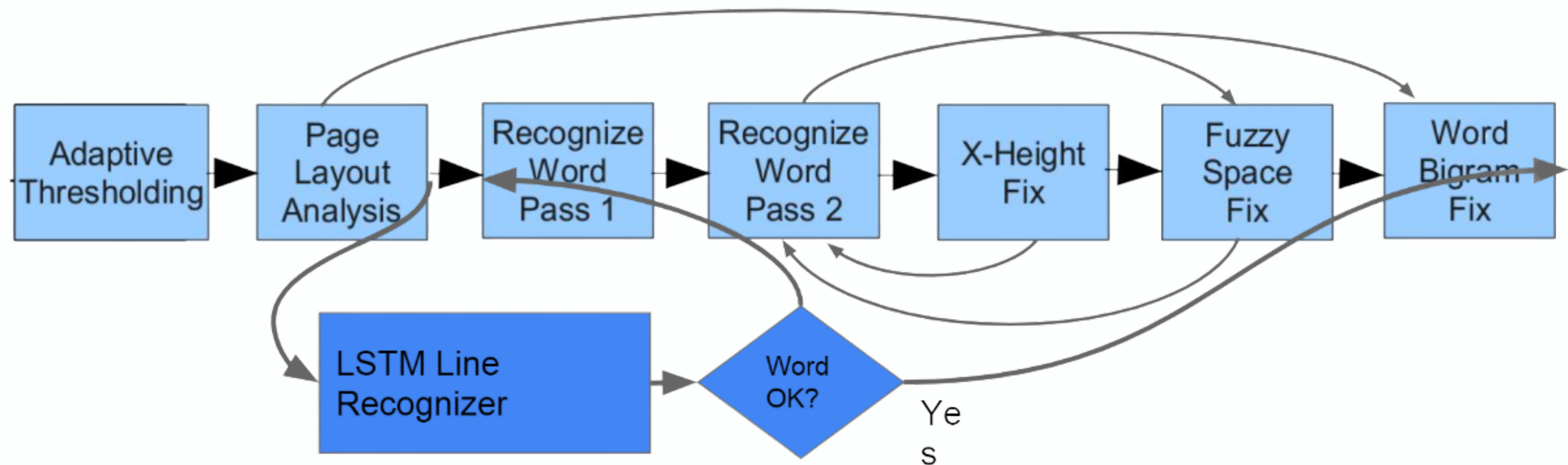


โครงสร้างการประมวลผล

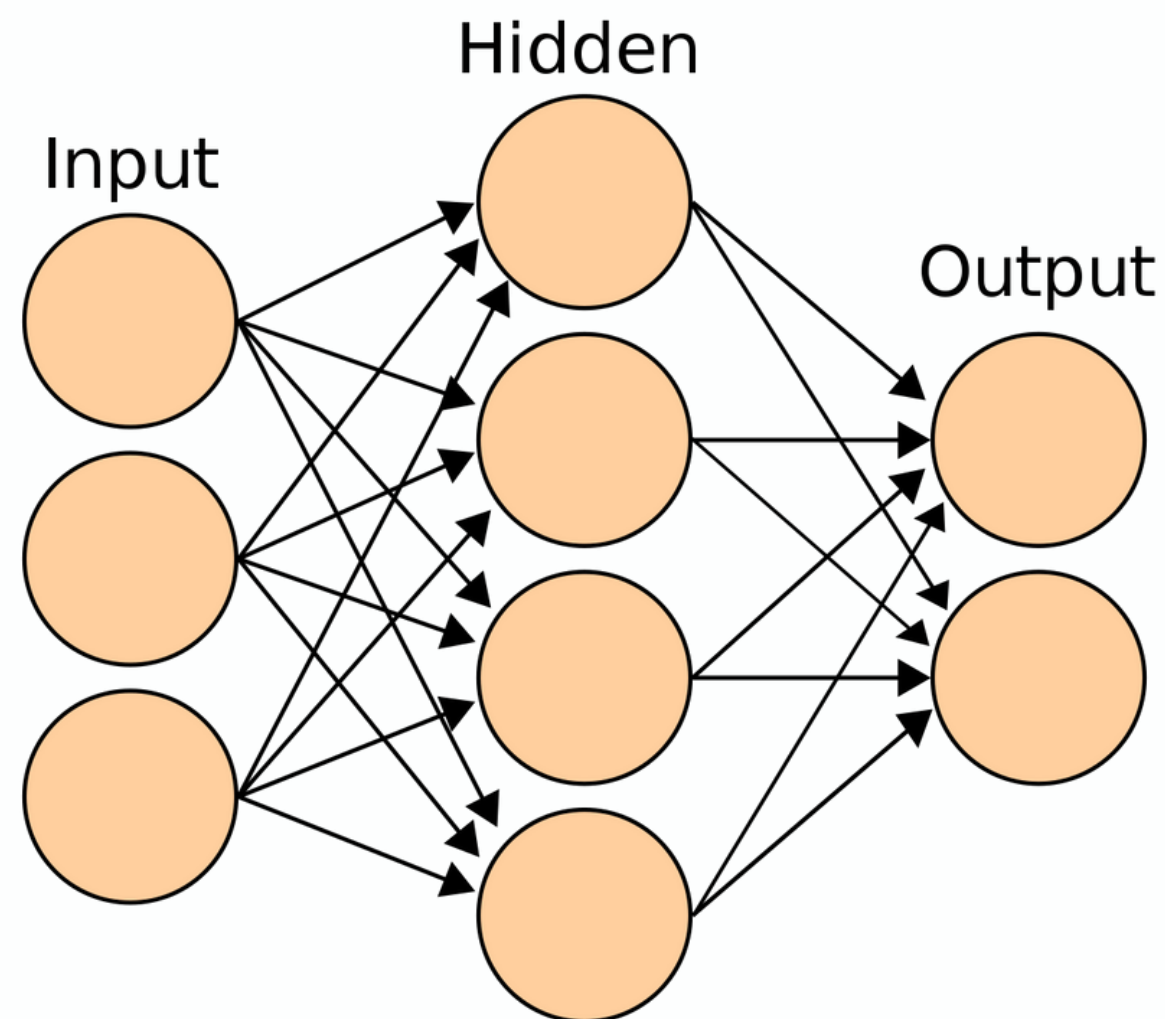
Background: Classical character classification



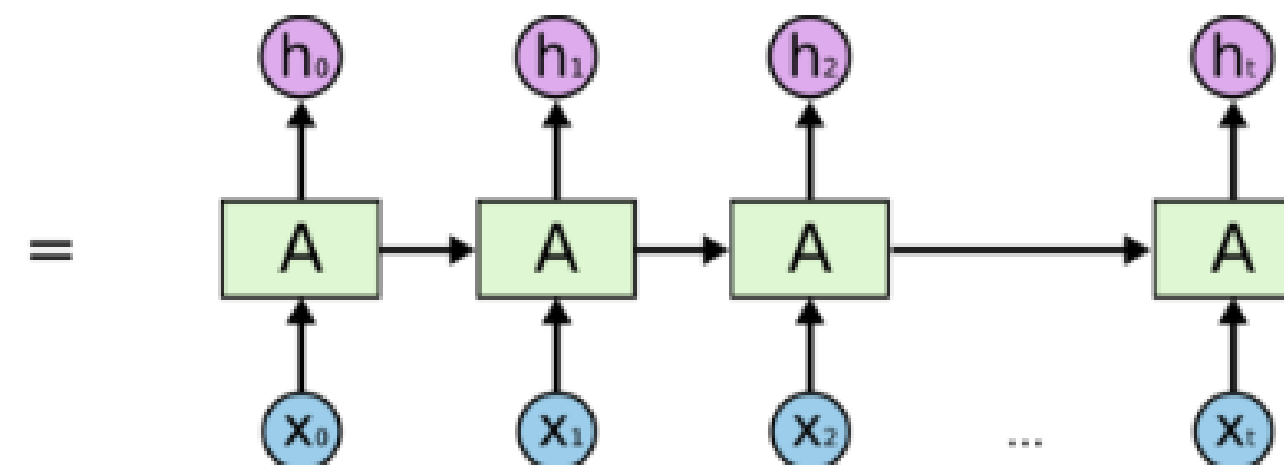
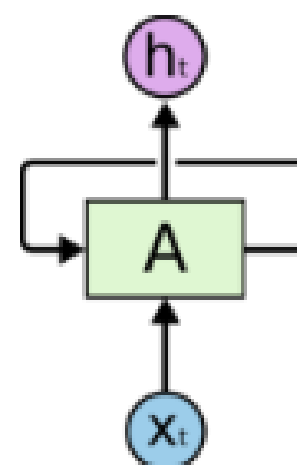
โครงสร้างการประมวลผล 4.0



ANN กับ RNN

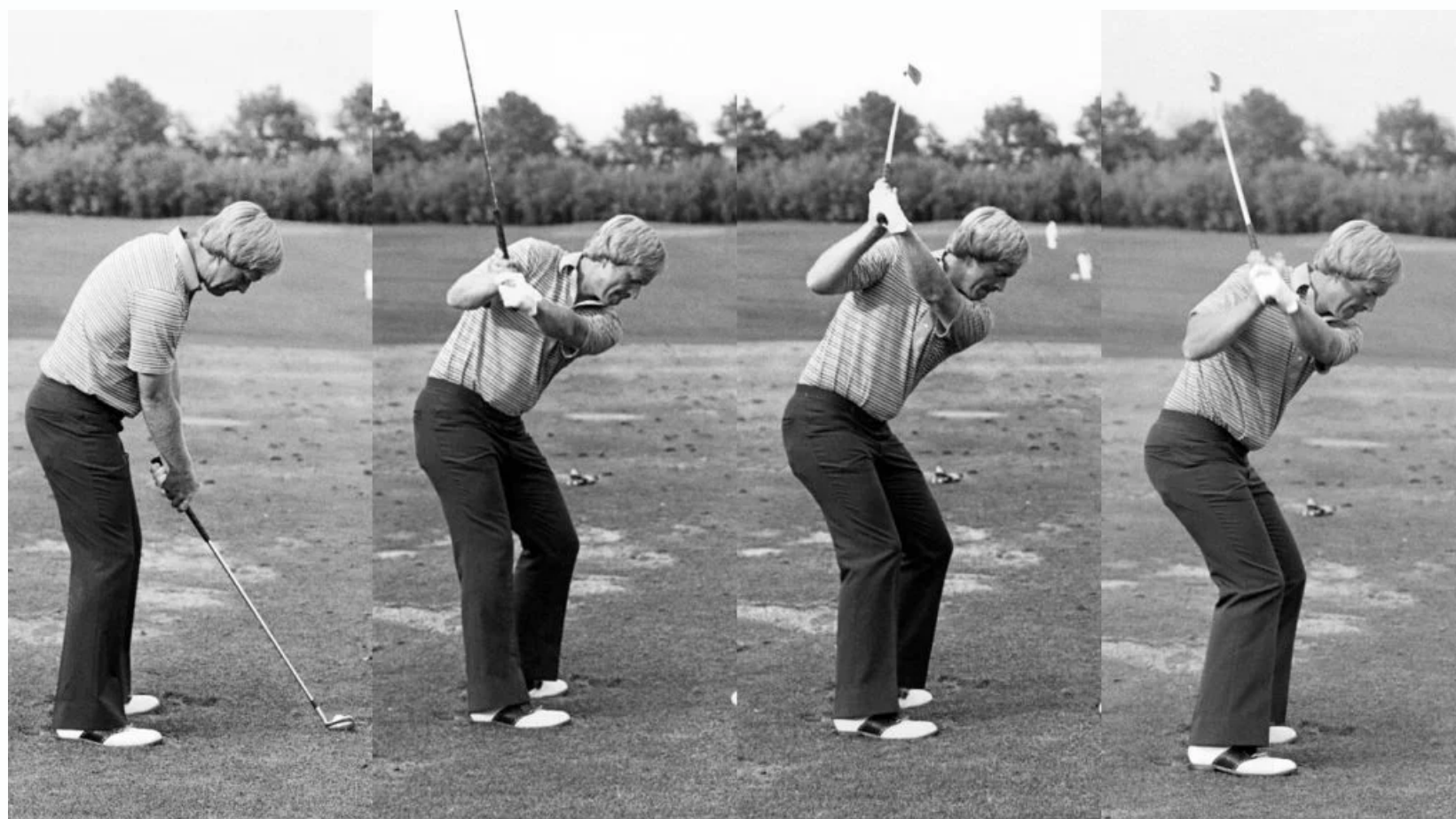


Artificial Neural Network



Recurrent Neural Network

RNN จะทำงานกับข้อมูลที่มีลักษณะเป็นลำดับ (sequence)



Video
(Sequence of images)

RNN จะทำงานกับข้อมูลที่มีลักษณะเป็นลำดับ (sequence)

Good Morning.

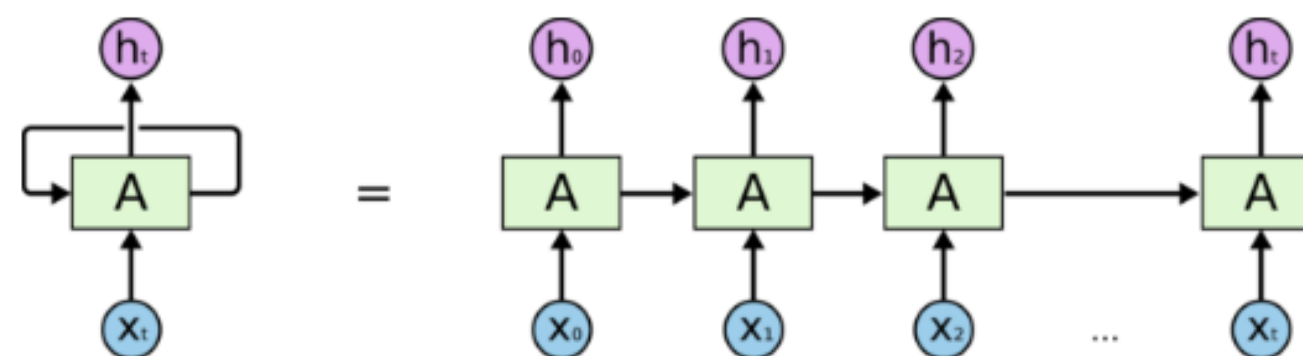
Nice to meet you.

Words

(Sequence of words)

RNN LSTM

LSTM นั้นจัดว่าเป็นโครงข่ายประเภท
Recurrent Neural Network (RNN)
นั่นคือ NN ที่มีการนำเอา Output
ของมันเองก่อนหน้านี้กลับมาใช้ใหม่



An unrolled recurrent neural network.

ข้อจำกัดของ Tesseract OCR

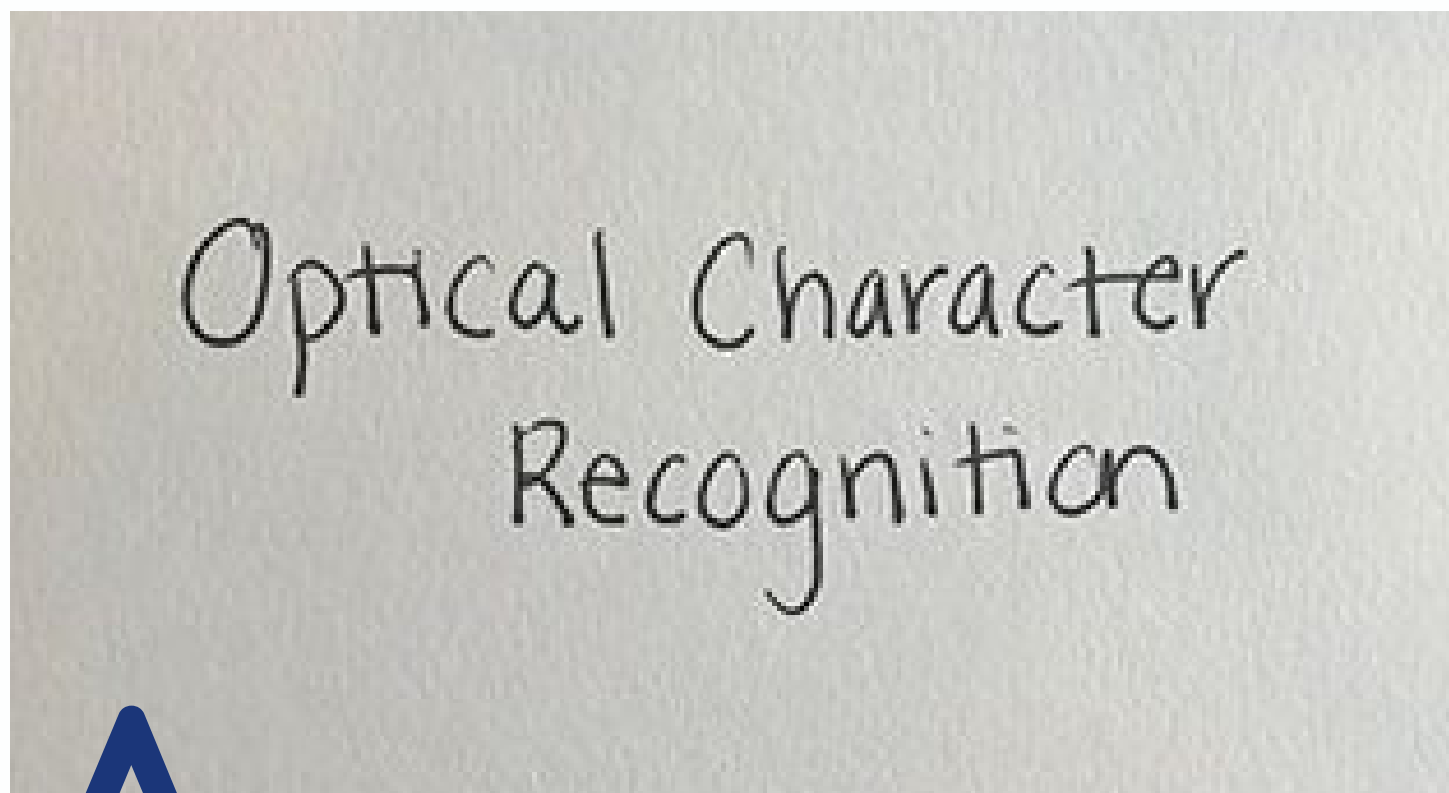
- ประมวลผลไม่ดีนักกับภาพที่ได้รับผลกระทบจากสิ่งของ
มุมมองที่บิดเบี้ยว และพื้นหลังที่ซับซ้อน



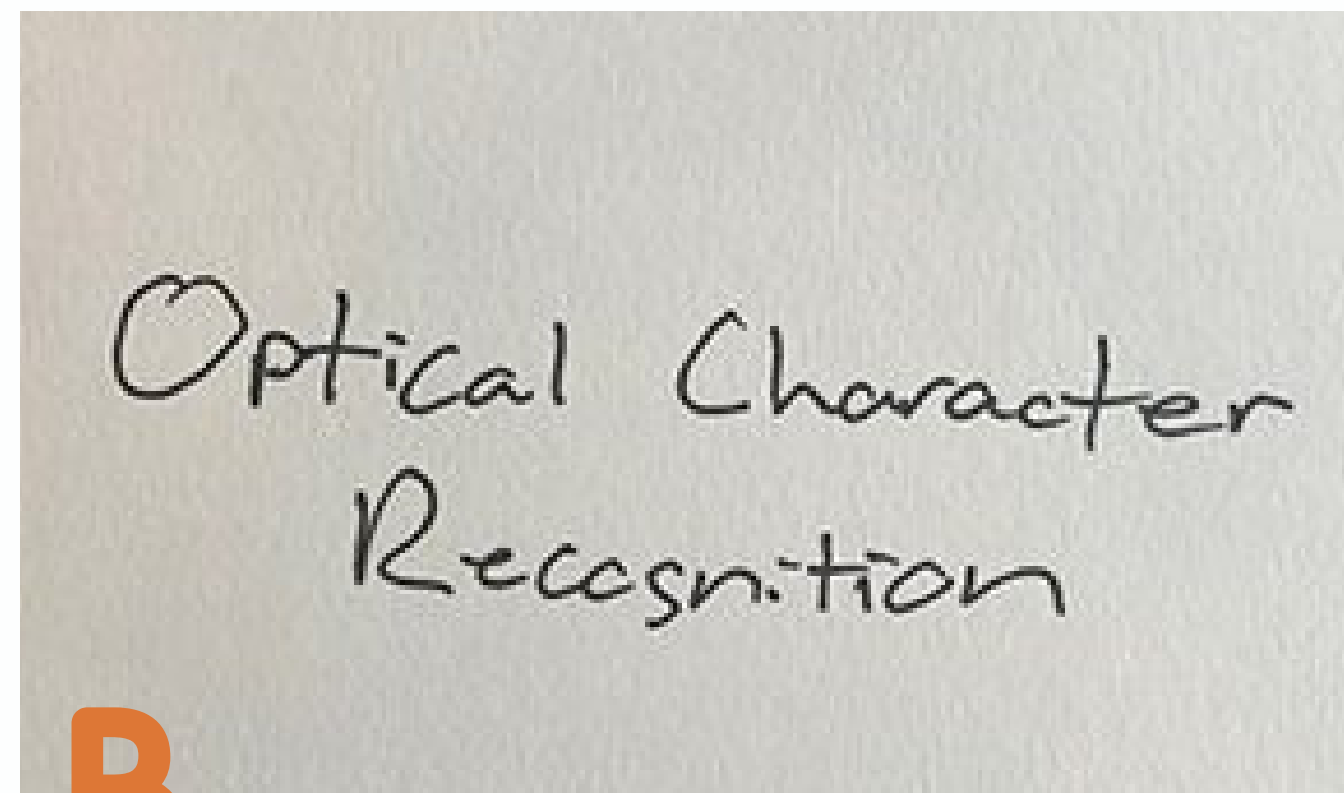
??4 bl?:??mn lees?rath??Ae?We, Pe?a wees?sss hese??_????A??Feed maal?ol?? ??

ข้อจำกัดของ Tesseract OCR

- ไม่สามารถประมวลผลลายมือแบบต่อเนื่องได้



A



B

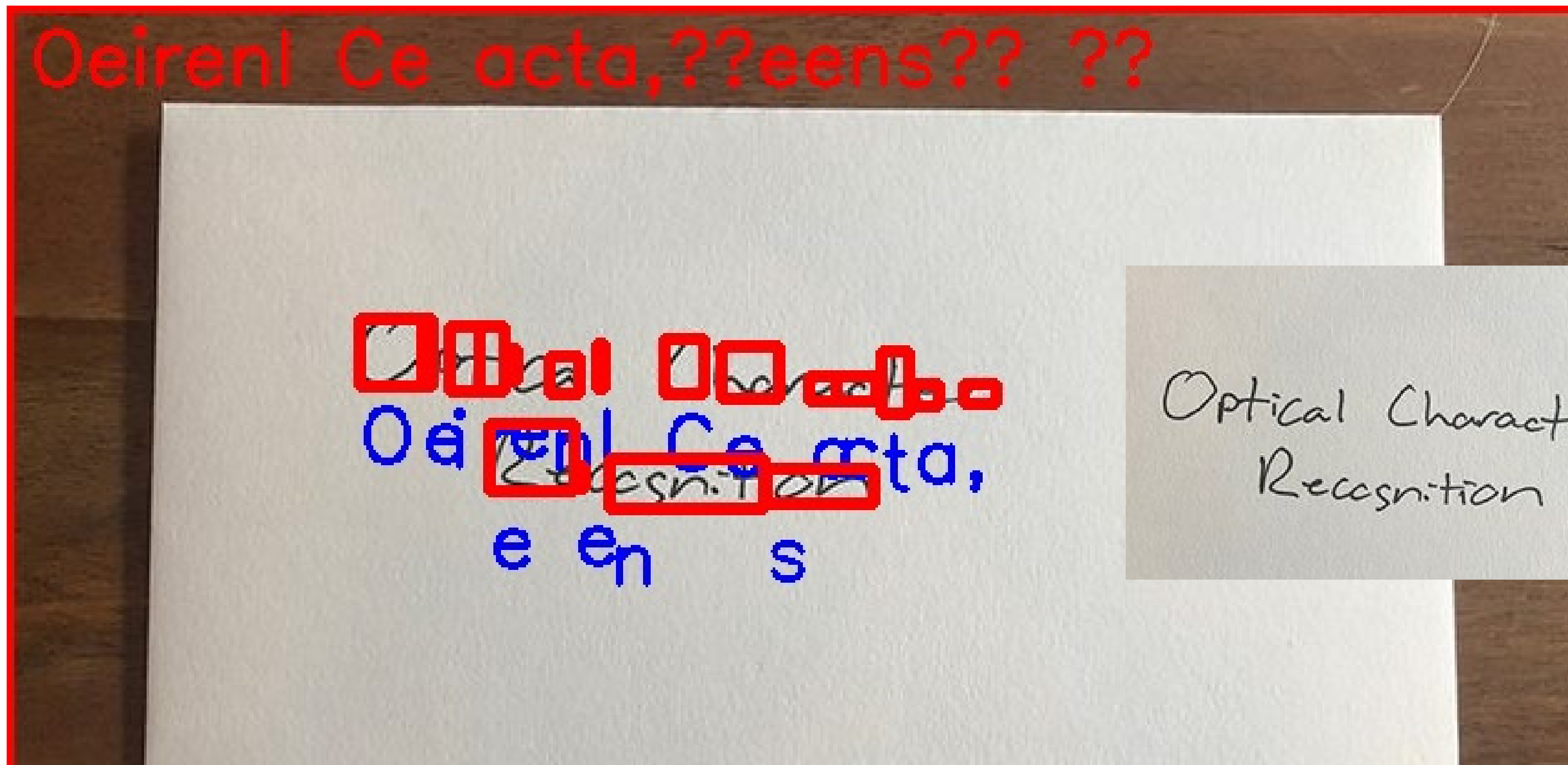
Optical Character? Recognihon?? ??

Optical Character
Recognihon

Optical Character
Recognition

A

39



B

40

ข้อจำกัดของ Tesseract OCR

- อาจพบว่าการประมวลผลไม่ถูกต้อง และแสดงผลเป็นเอาต์พุต
- หากเอกสารมีภาษาอื่นนอกเหนือจากที่ระบุในอาร์กิวเมนต์ ผลลัพธ์อาจไม่ดี
- รูปภาพคุณภาพต่ำอาจทำให้ OCR มีความแม่นยำต่ำ

ข้อจำกัดของ Tesseract OCR

วารสารวิจัยและส่งเสริมวิชาการเกษตร 33(1): 66-76

naniside??ToyaitugmuoanunsnsyradauaraWs1??Hanr

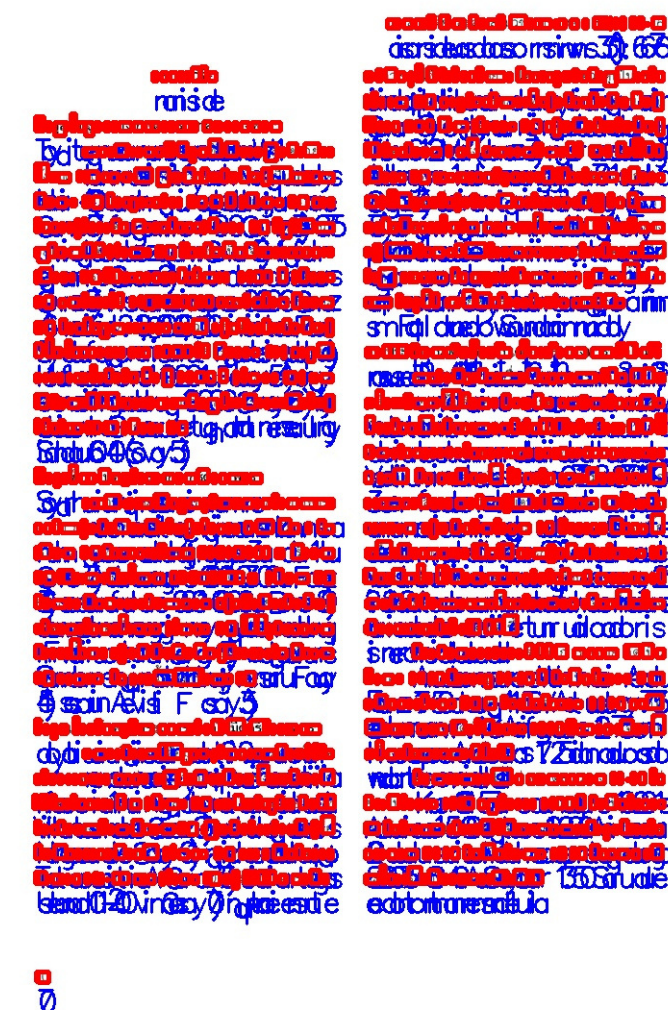
ผลการวิจัย

ข้อมูลพื้นฐานของเกษตรกรชาวสวนยางพารา

ผลการวิเคราะห์ข้อมูลทั่วไปของกลุ่มตัวอย่างทั้งหมด 400 คน พบว่า กลุ่มตัวอย่างส่วนใหญ่เป็นเพศชาย (ร้อยละ 67) มีอายุระหว่าง 41-50 ปี (ร้อยละ 34) การศึกษาอยู่ในระดับประถมศึกษา (ร้อยละ 56) มีภูมิลำเนาอยู่ในภาคใต้ (ร้อยละ 94) มีอาชีพเป็นเจ้าของสวนยาง (ร้อยละ 63) มีประสบการณ์กรีดยาง 10-20 ปี (ร้อยละ 46) และมีรายได้ 10,000-20,000 บาทต่อเดือน (ร้อยละ 51) ส่วนข้อมูลยางพารา พบว่า กลุ่มตัวอย่างส่วนใหญ่มีพื้นที่สวนยางพารา 10-20 ไร่ (ร้อยละ 54) อายุต้นยางพาราที่กรีดยางส่วนใหญ่ 10-20 ปี (ร้อยละ 66) และสัดส่วนรายได้เจ้าของสวนยางกับลูกจ้างกรีดยางส่วนใหญ่ มีสัดส่วน 60:40 (ร้อยละ 53)

ชนิดวัสดุที่ใช้ทำมีดกรีดยาง โดยระบุชนิดวัสดุเป็นหลัก (ร้อยละ 64) อายุมีดกรีดยางที่กลุ่มตัวอย่างส่วนใหญ่ใช้งาน 6-12 เดือน (ร้อยละ 34) กลุ่มตัวอย่างส่วนใหญ่ลับมีดกรีดยาง 1 ครั้ง สามารถกรีดยางได้ 701 ต้นขึ้นไป (ร้อยละ 35) และเคยพบปัญหาการใช้มีดกรีดยาง (ร้อยละ 78) ซึ่งปัญหาที่กลุ่มตัวอย่างพบโดยส่วนใหญ่ คือ ปัญหาคมมีดกรีดยางบิ่นง่าย นอกจากนี้การเลือกใช้อายุมีดกรีดยาง กลุ่มตัวอย่างจะคำนึงถึงความคมของมีดกรีดยางเป็นสำคัญ รองลงมา คือ อายุการใช้งาน ขนาด รูปร่าง น้ำหนัก ราคา วัสดุที่ใช้ และยี่ห้อหรือแหล่งผลิต ตามลำดับ

การประเมินแนวคิดเกี่ยวกับเครื่องกรีดยางพาราอัตโนมัติ
แนวคิดเกี่ยวกับเครื่องกรีดยางพาราอัตโนมัติเป็น
เครื่องกรีดยางที่ไม่ต้องอาศัยคนกรีดยาง การประเมินแนวคิด



naniside??ToyaitugmuoanunsnsyradauaraWs1??Hanr

ติดตั้ง Tesseract

IMAGE RECOGNITION

หน้าแรก [หลักสูตร](#) [รายชื่อผู้เข้าร่วมอบรม](#) [Download](#)



COCO

เป็นชุดข้อมูลขนาดใหญ่สร้างสำหรับการตรวจจับข้อความ และการเรียนรู้ข้อความบนภาพโดยตัว COCO-TEXT แต่ละภาพมีหลายข้อความตัวอักษร

[ดาวน์โหลด >](#)



Tesseract ocr

เอ็นจินการจดจำอักขระด้วยแสง สำหรับระบบปฏิบัติการต่าง ๆ เป็นซอฟต์แวร์ฟรี

[ดาวน์โหลด >](#)



CMake

ซอฟต์แวร์สำหรับสร้าง build script

[ดาวน์โหลด >](#)

ติดตั้ง Tesseract

Home

Stefan Weil edited this page on Jun 28 · 70 revisions

Tesseract at UB Mannheim

The Mannheim University Library (UB Mannheim) uses Tesseract to perform OCR (optical character recognition) of historical German newspapers ([Allgemeine Preußische Staatszeitung](#), [Deutscher Reichsanzeiger](#)). The latest results with OCR from more than 360,000 scans are available [online](#).

Tesseract installer for Windows

Normally we run Tesseract on Debian GNU Linux, but there was also the need for a Windows version. That's why we have built a Tesseract installer for Windows.

WARNING: Tesseract should be either installed in the directory which is suggested during the installation or in a new directory. The uninstaller removes the whole installation directory. If you installed Tesseract in an existing directory, that directory will be removed with all its subdirectories and files.

The latest installers can be downloaded here:

- [tesseract-ocr-w32-setup-v5.0.0-alpha.20210506.exe](#) (32 bit) and
- [tesseract-ocr-w64-setup-v5.0.0-alpha.20210506.exe](#) (64 bit) resp.



- [tesseract-ocr-w32-setup-v5.0.0-alpha.20210506.exe](#) (32 bit) and
- [tesseract-ocr-w64-setup-v5.0.0-alpha.20210506.exe](#) (64 bit) resp.

We don't provide an installer for Tesseract 4.1.0 because we think that the latest version 5.0.0-alpha is better for most Windows users in many aspects (functionality, speed, stability). Version 4.1 is only needed for people who develop software based on the Tesseract API and who need 100 % API compatibility with version 4.0.

There are also [older versions](#) available.

In addition, we also provide [documentation](#) which was generated by Doxygen.

UB-Mannheim/ tesseract



Tesseract Open Source OCR Engine (main repository)

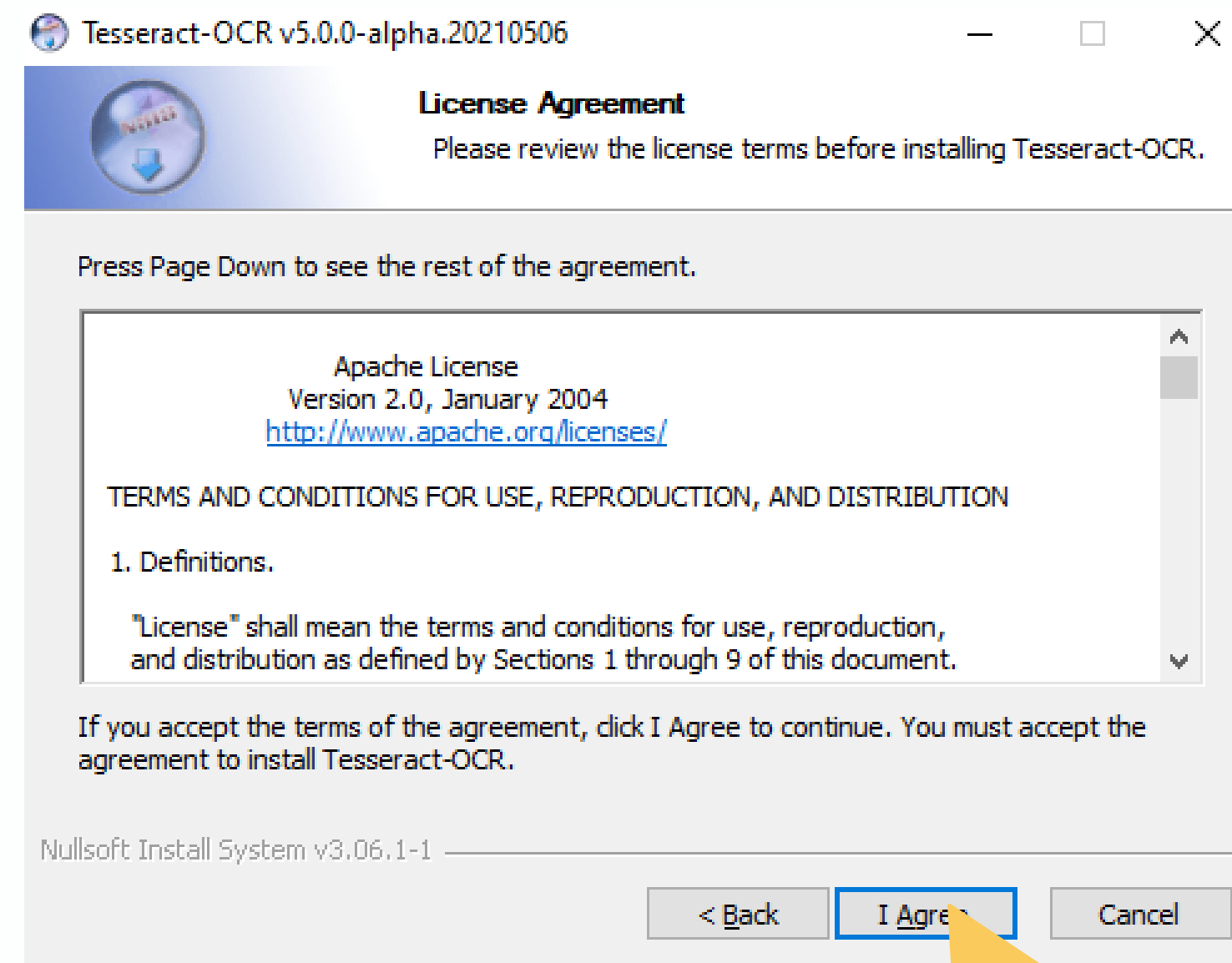
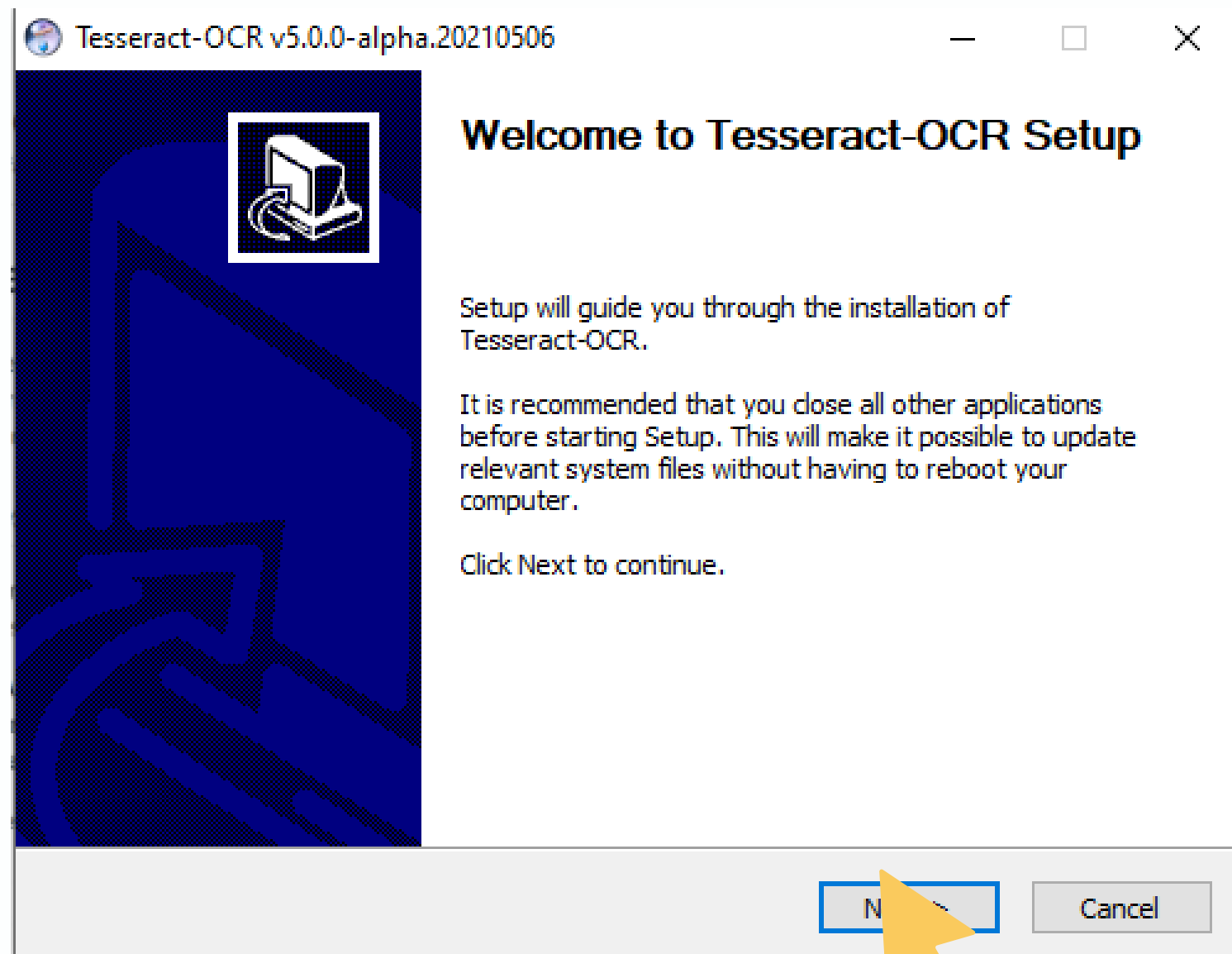
0 Contributors 9 Issues 1k Stars 236 Forks

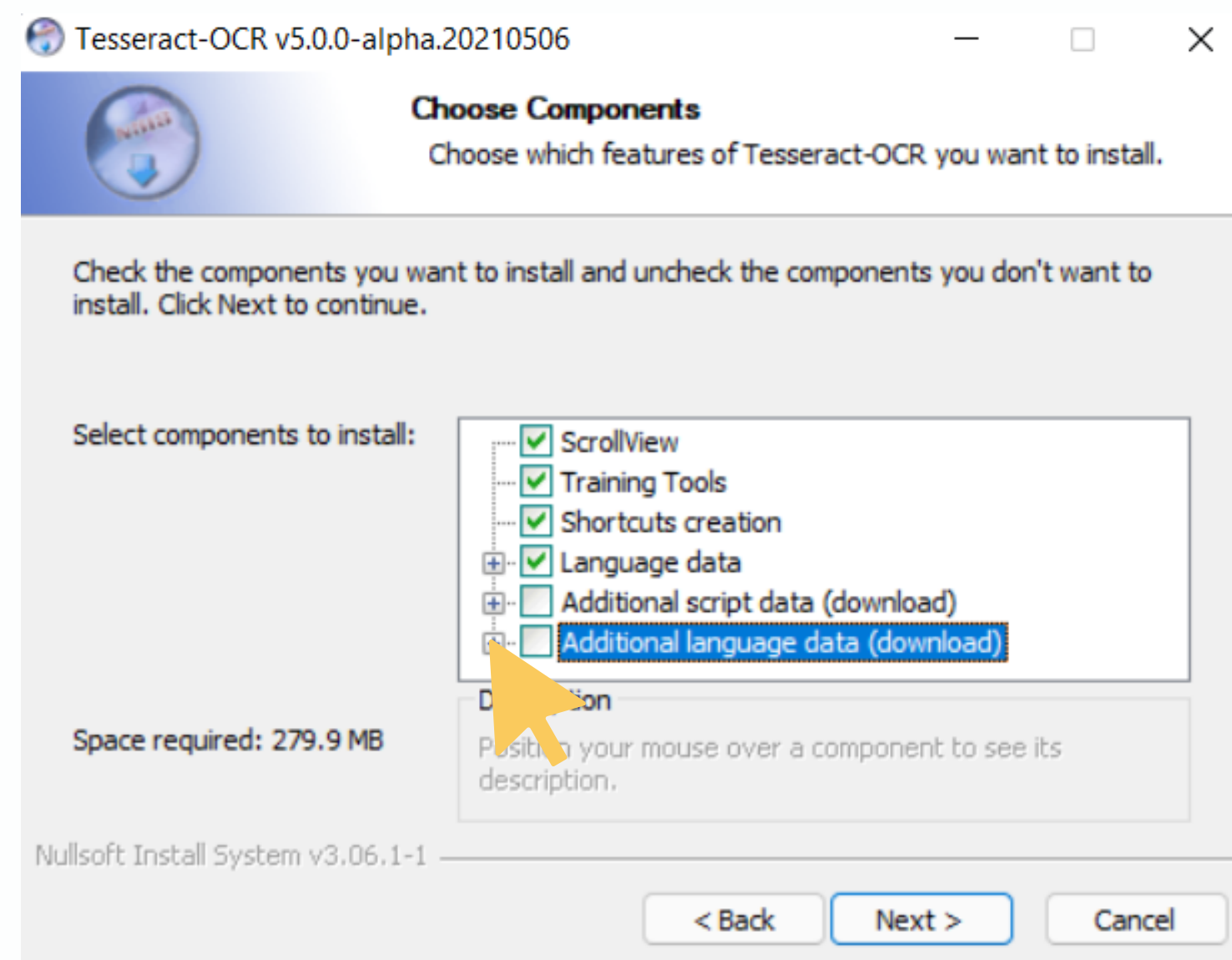
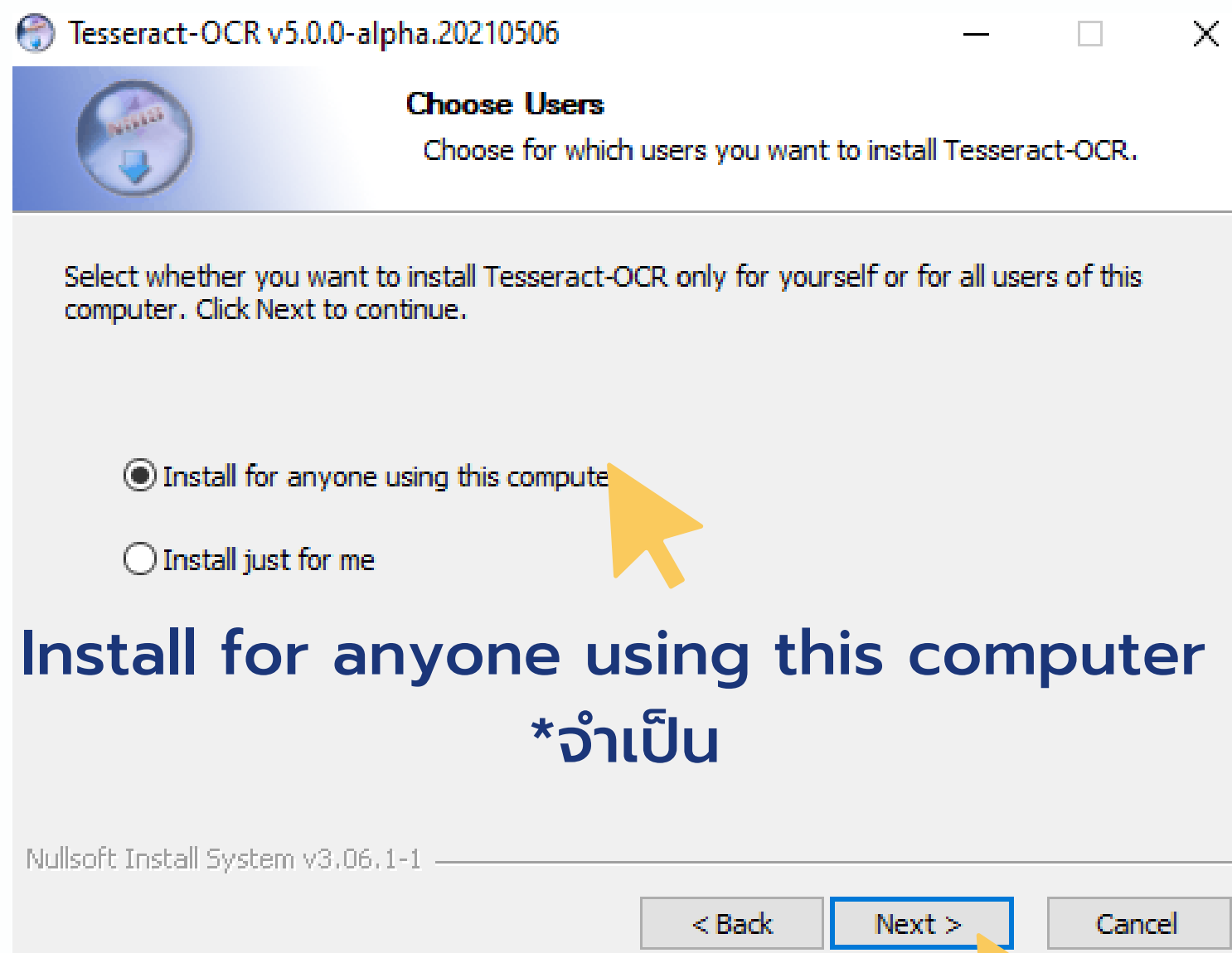


Home · [UB-Mannheim/tesseract Wiki](#)

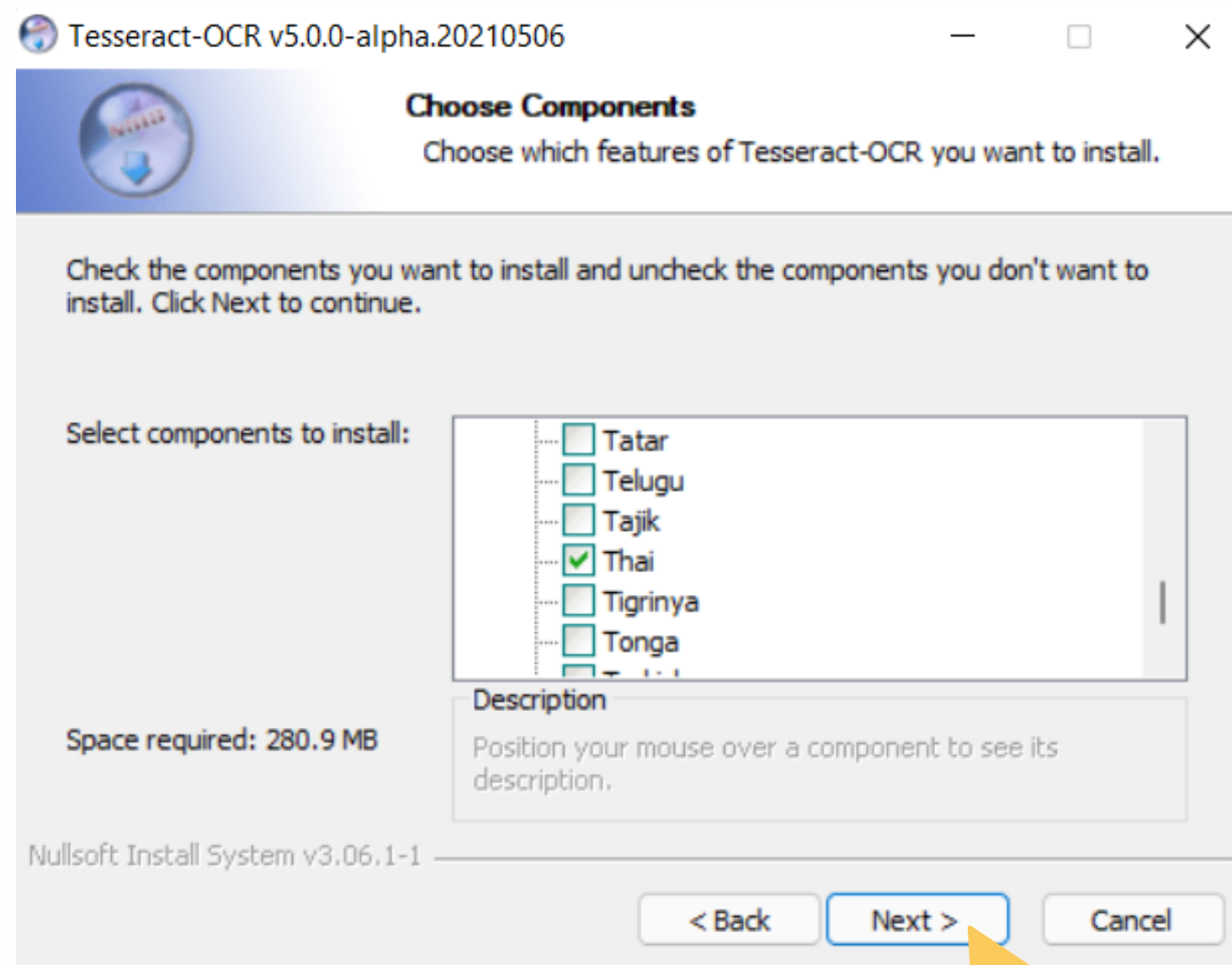
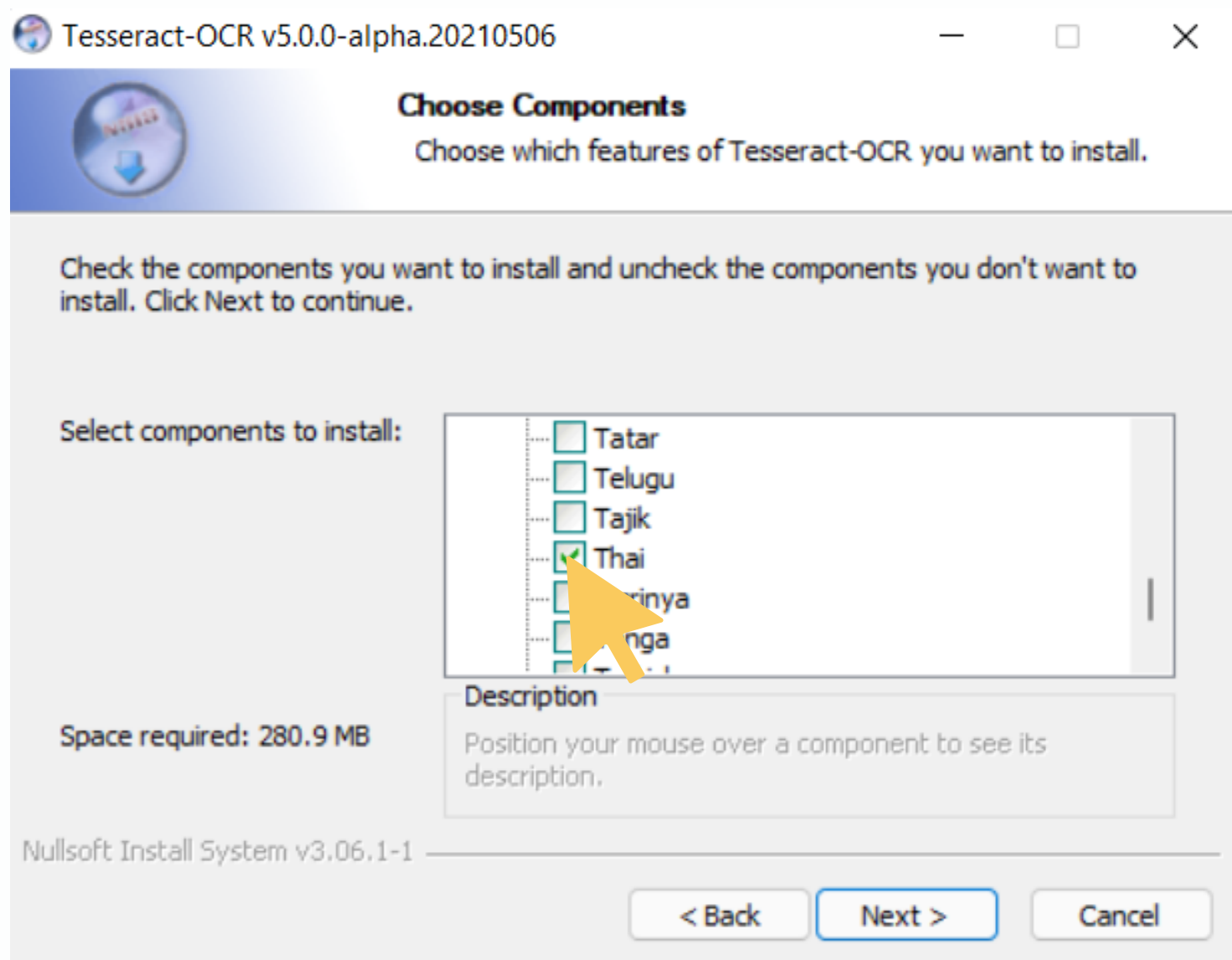
Tesseract Open Source OCR Engine (main repository) - Home
· [UB-Mannheim/tesseract Wiki](#)

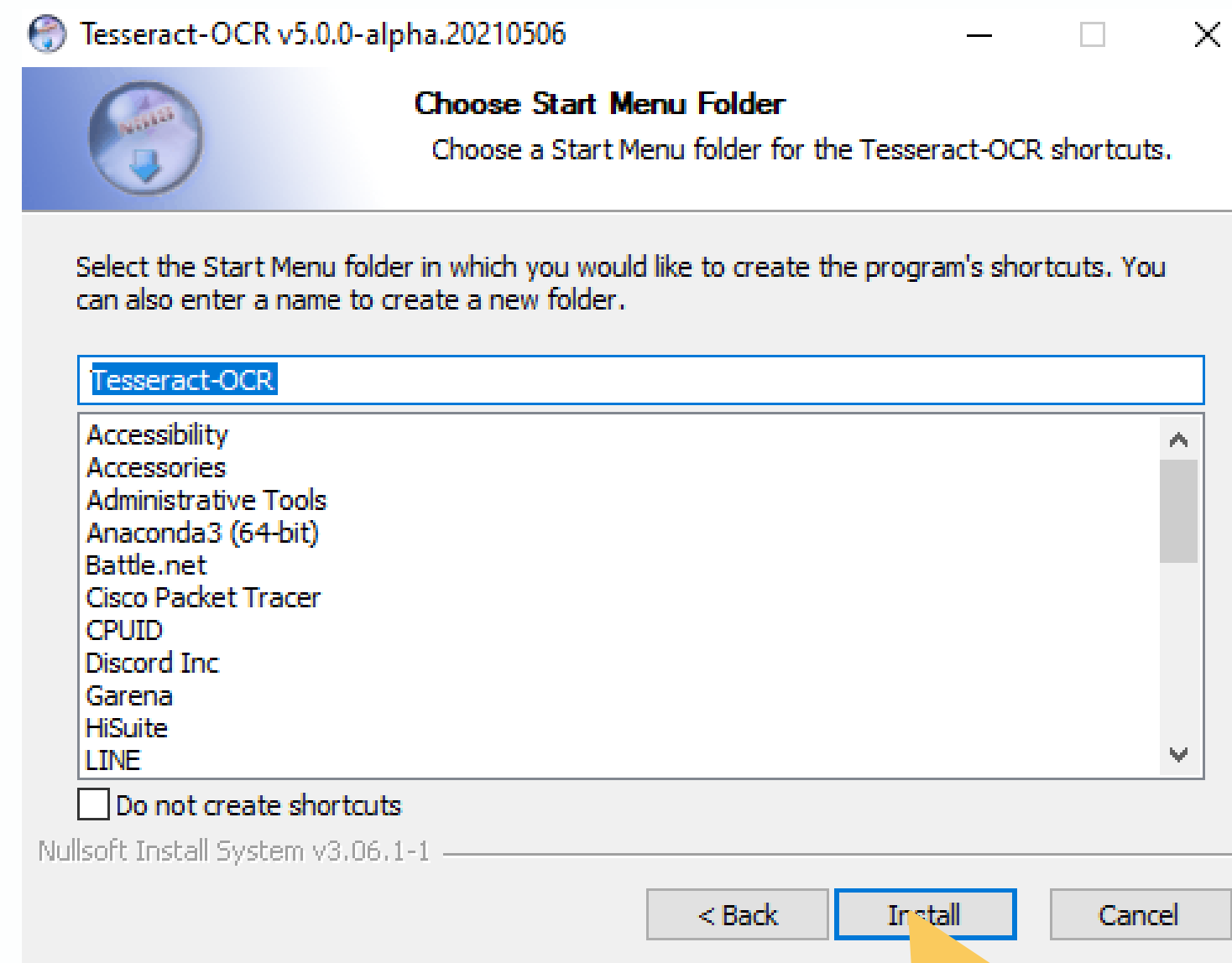
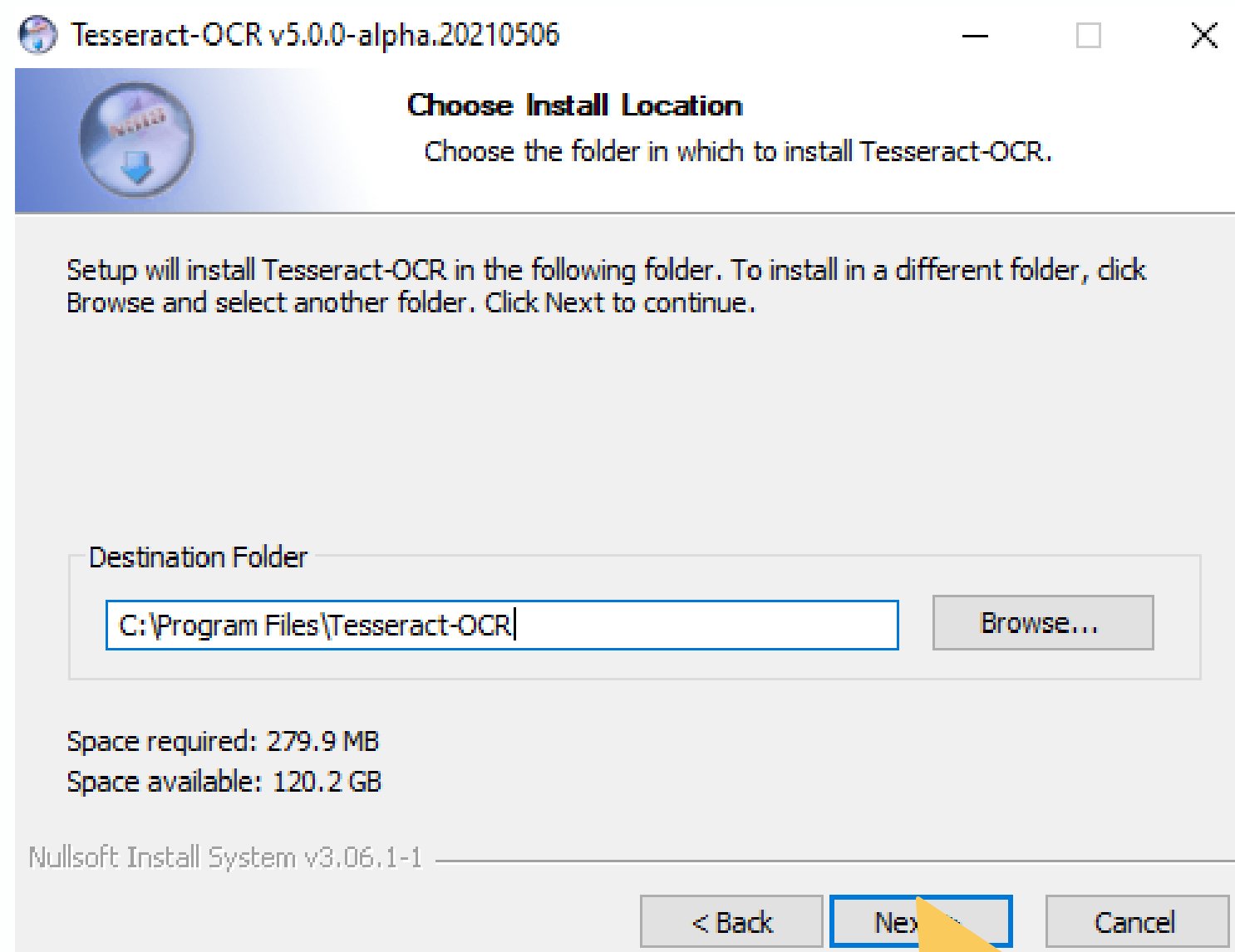
[GitHub](#)

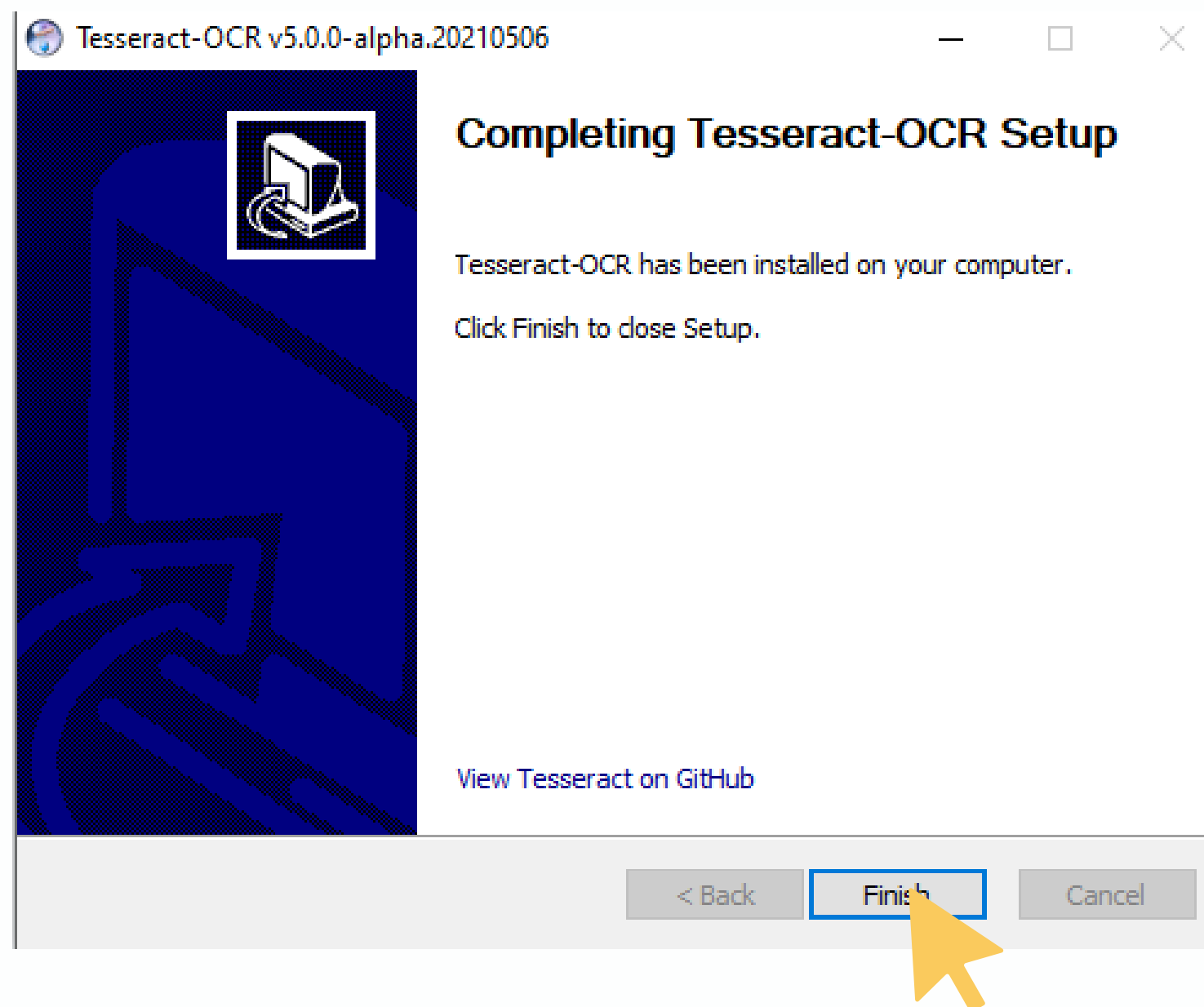




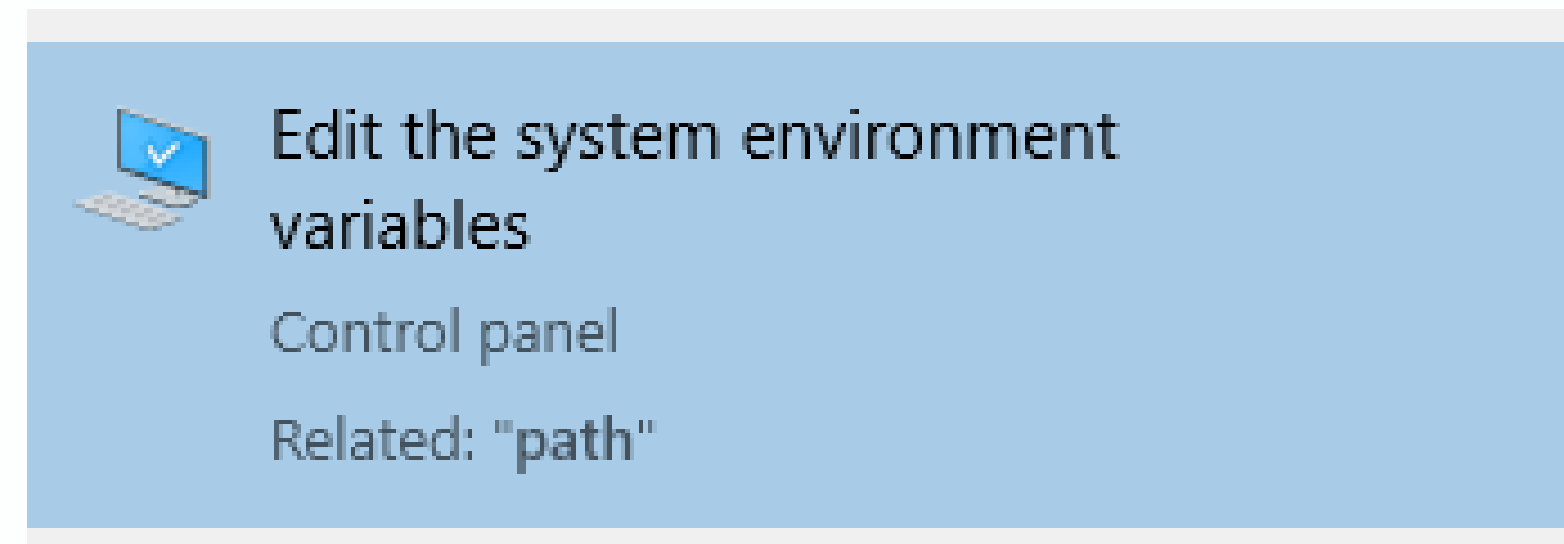
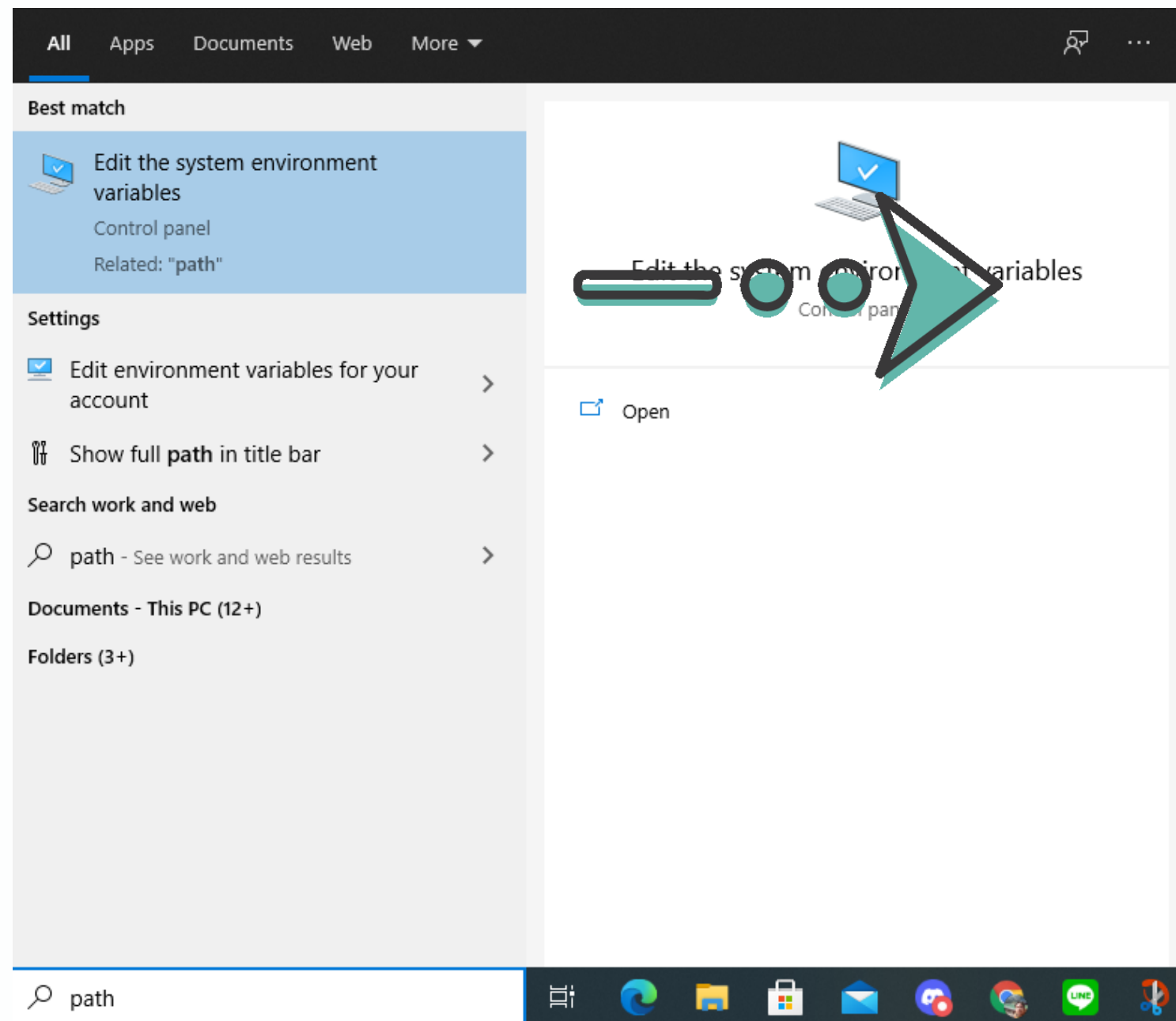
กดเครื่องหมาย + เพื่อเลือกภาษาที่จะติดตั้งเพิ่ม



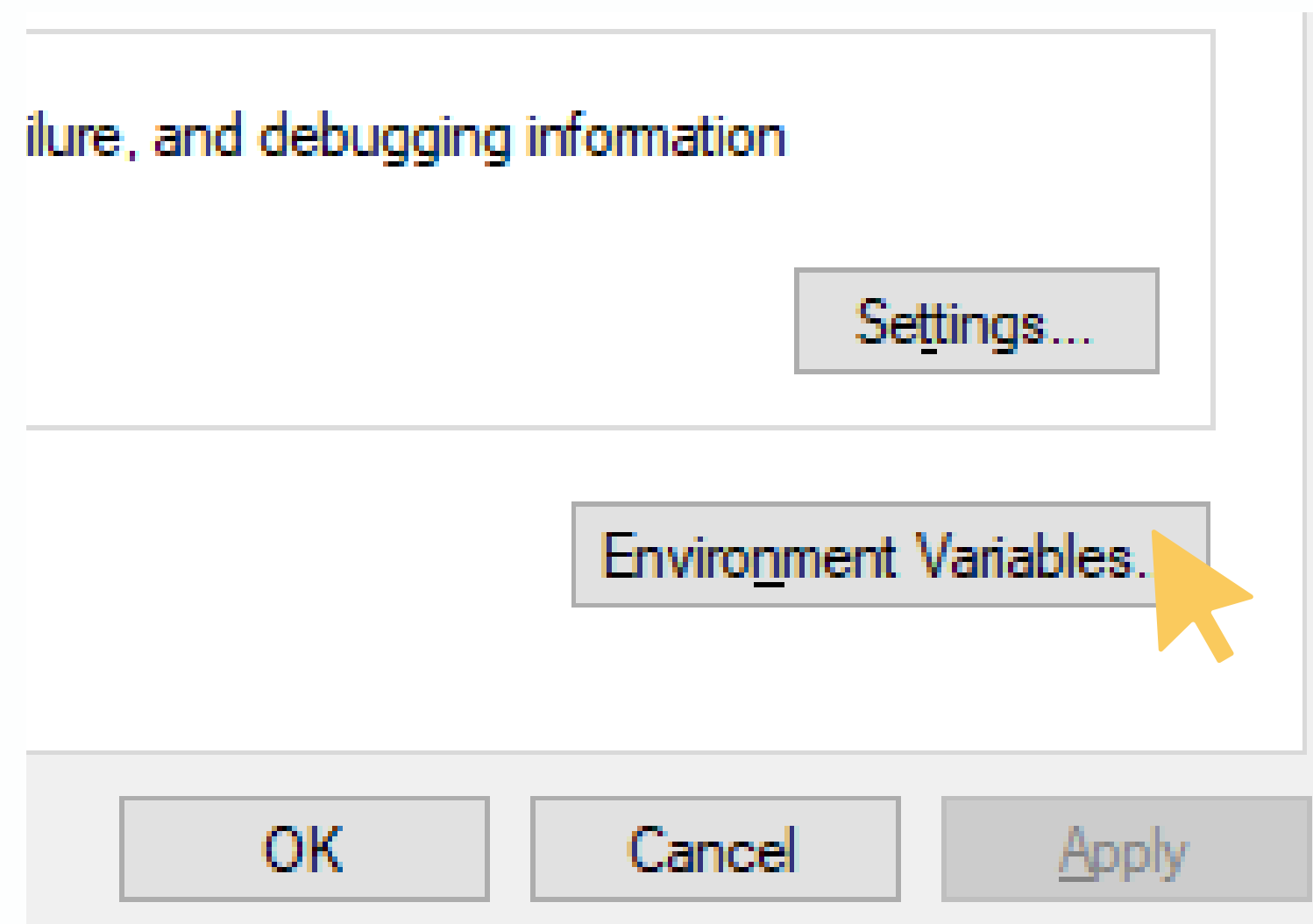
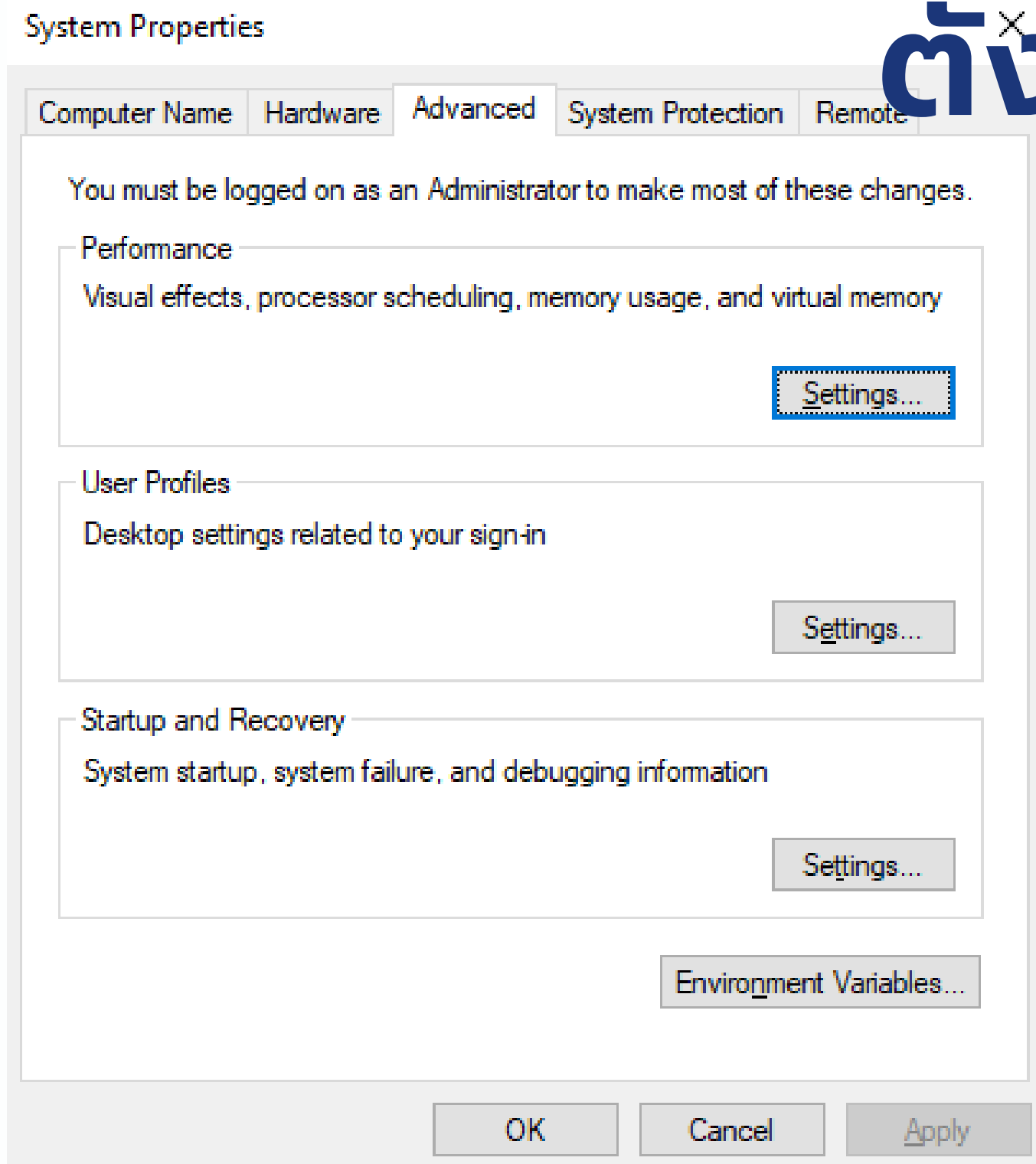




ตั้งค่า Path



ตั้งค่า Path



ตั้งค่า Path

Environment Variables

User variables for Jirawat

Variable	Value
OneDrive	C:\Users\Jirawat\OneDrive - kmutnb.ac.th
OneDriveCommercial	C:\Users\Jirawat\OneDrive - kmutnb.ac.th
Path	C:\Users\Jirawat\AppData\Local\Microsoft\WindowsApps;;C:\Users...
QT_DEVICE_PIXEL_RATIO	auto
TEMP	C:\Users\Jirawat\AppData\Local\Temp
TMP	C:\Users\Jirawat\AppData\Local\Temp

New... Edit... Delete

System variables

Variable	Value
ComSpec	C:\Windows\system32\cmd.exe
DriverData	C:\Windows\System32\Drivers\DriverData
NUMBER_OF_PROCESSORS	12
OS	Windows_NT
Path	C:\Windows\system32;C:\Windows;C:\Windows\System32\Wbem;...
PATHEXT	.COM;.EXE;.BAT;.CMD;.VBS;.VBE;.JS;.JSE;.WSF;.WSH;.MSC
PROCESSOR_ARCHITECTURE	AMD64

New... Edit... Delete

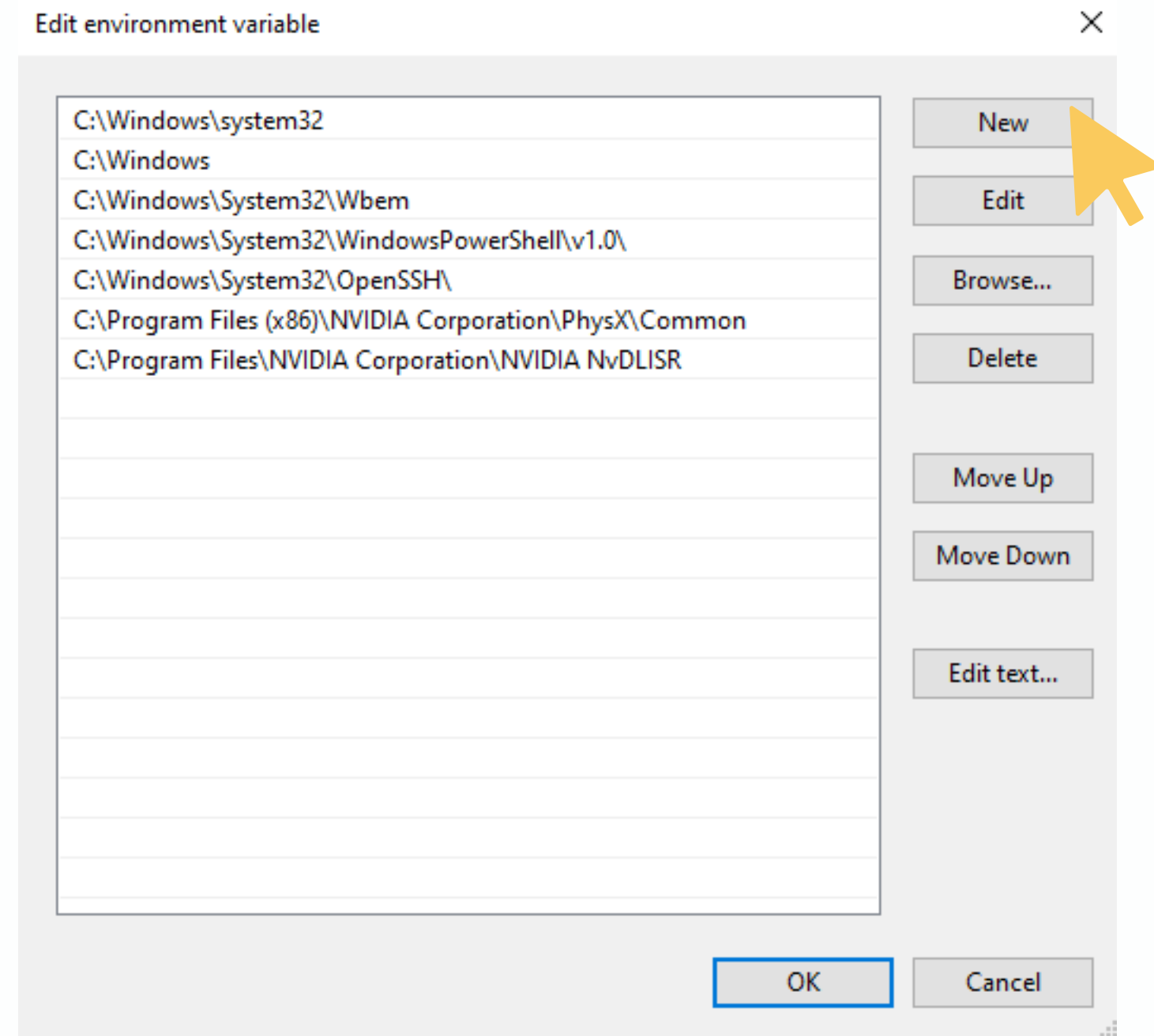
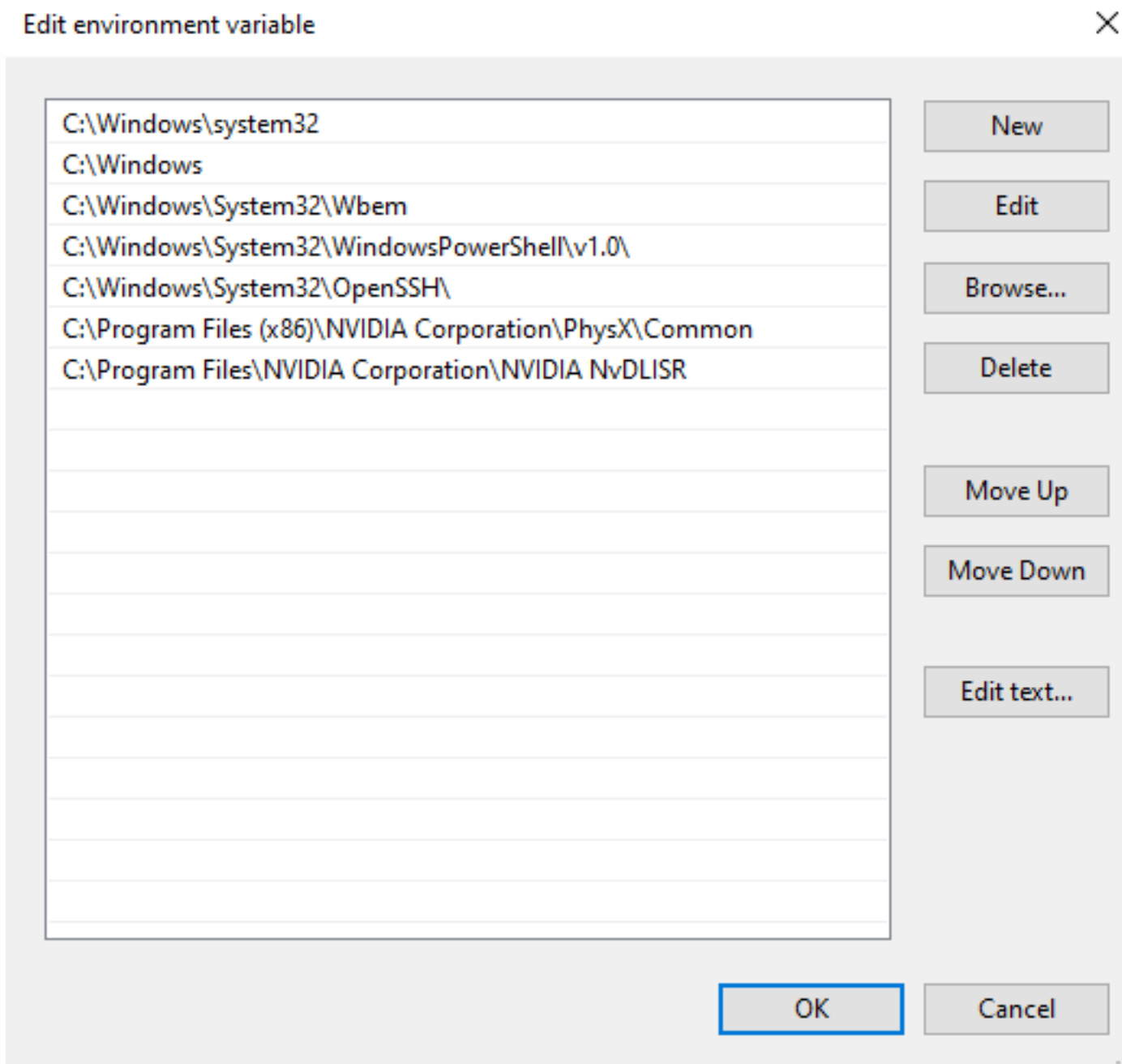
OK Cancel

System variables

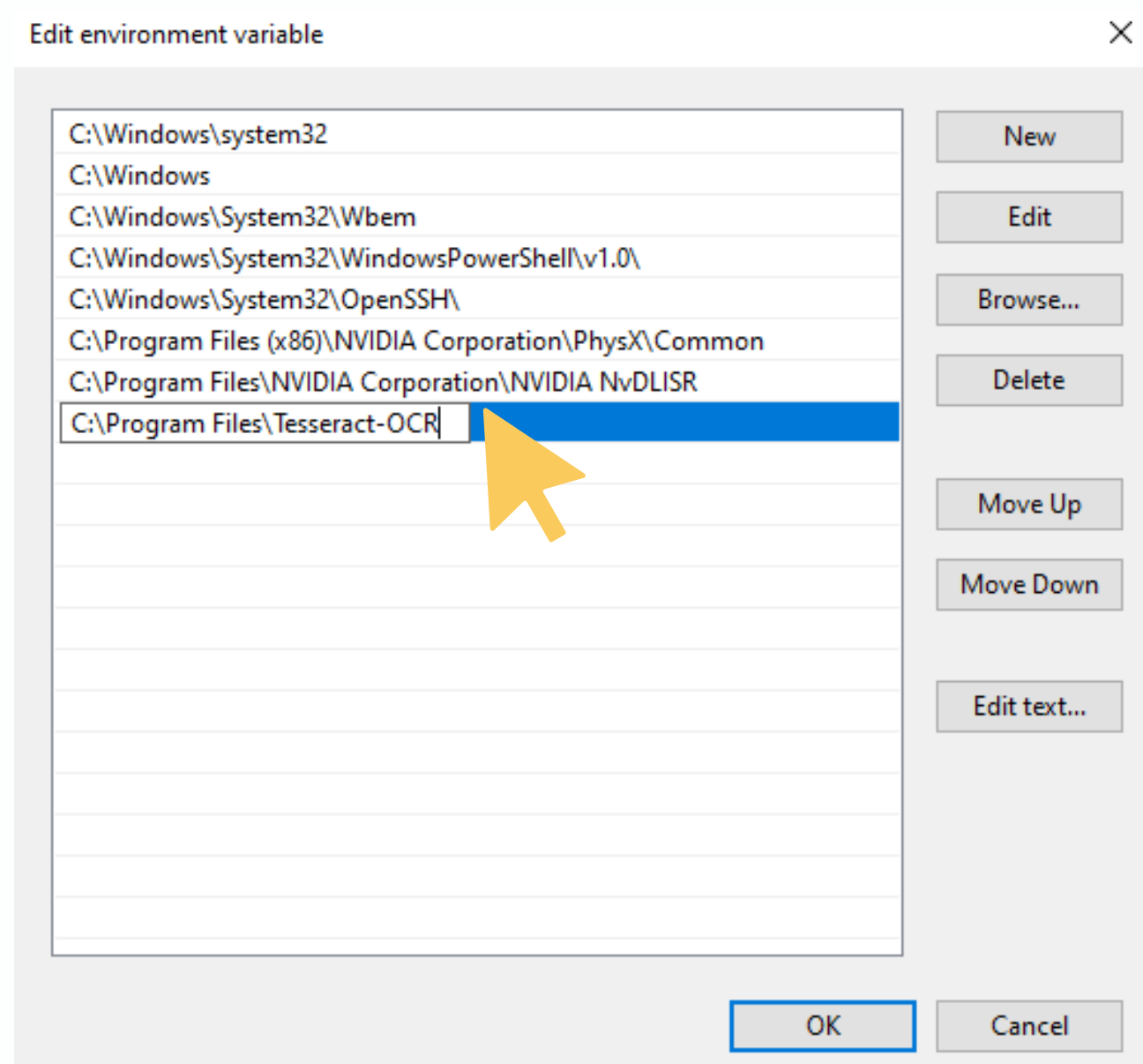
Variable	Value
ComSpec	C:\Windows\system32\cmd.exe
DriverData	C:\Windows\System32\Drivers\DriverData
NUMBER_OF_PROCESSORS	12
OS	Windows_NT
Path	C:\Windows\system32;C:\Windows;C:\Windows\System32\Wbem;...
PATHEXT	.COM;.EXE;.BAT;.CMD;.VBS;.VBE;.JS;.JSE;.WSF;.WSH;.MSC
PROCESSOR_ARCHITECTURE	AMD64

New... Edit... Delete

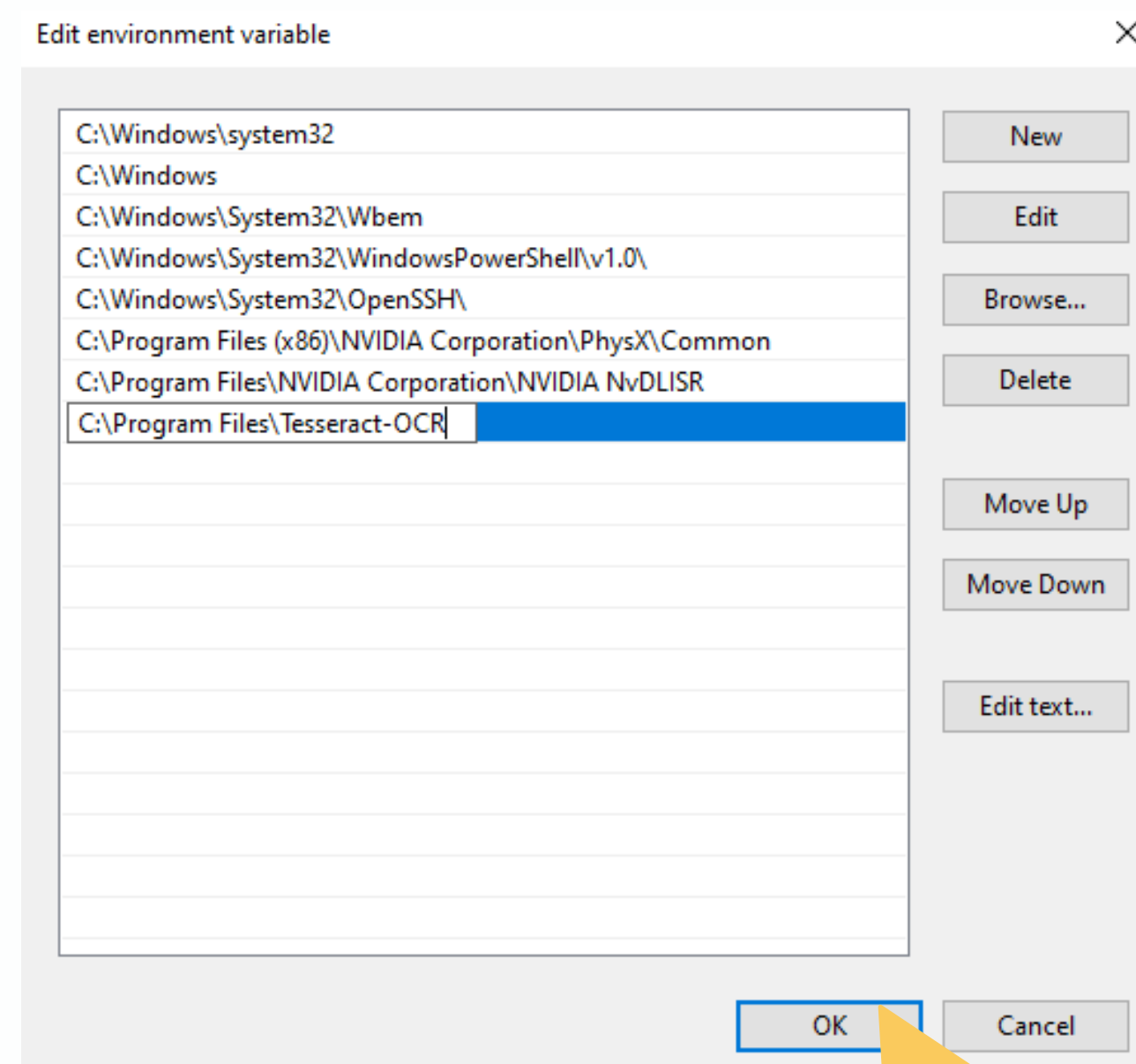
ตั้งค่า Path



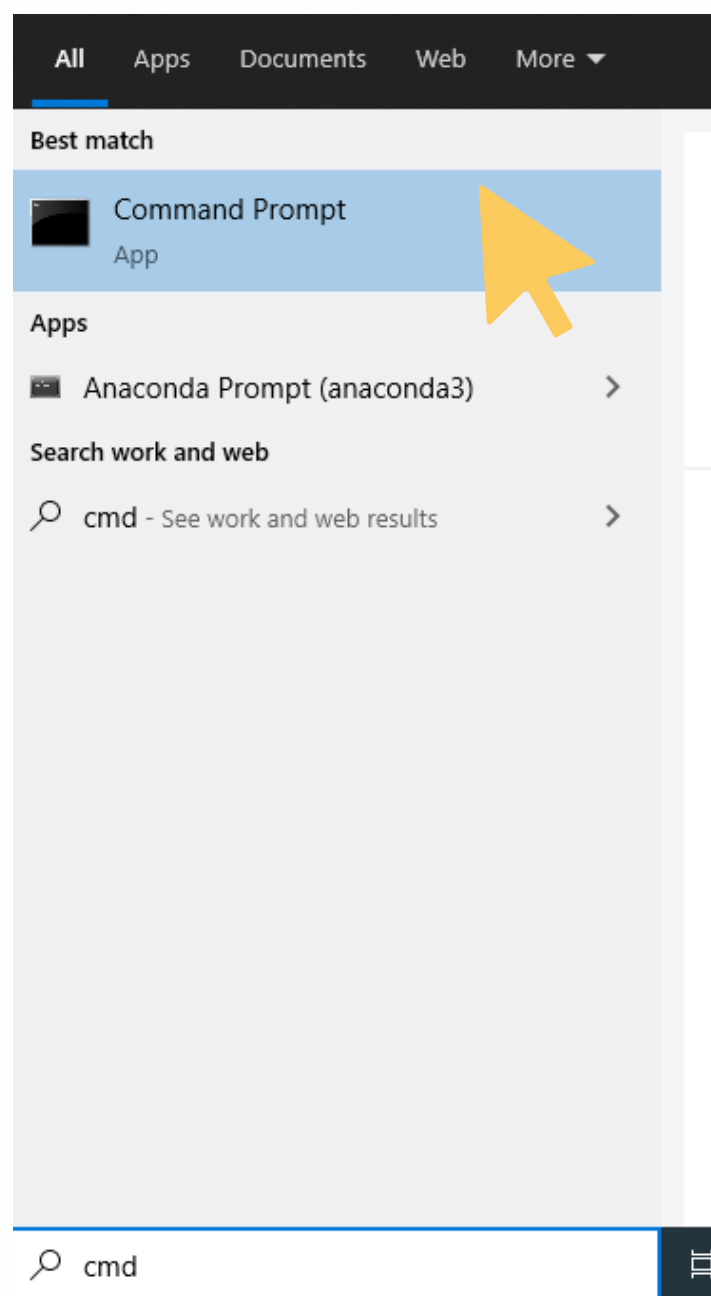
ตั้งค่า Path



C:\Program Files\Tesseract-OCR



ตรวจสอบการติดตั้ง



Command Prompt


```
Microsoft Windows [Version 10.0.19043.1151]
(c) Microsoft Corporation. All rights reserved.

C:\Users\Jirawat>tesseract --version
tesseract v5.0.0-alpha.20210506
leptonica-1.78.0
libgif 5.1.4 : libjpeg 8d (libjpeg-turbo 1.5.3) : libpng 1.6.34 : libtiff 4.0.9 : z
hjp2 2.3.0
Found AVX2
Found AVX
Found FMA
Found SSE4.1
Found libarchive 3.5.0 zlib/1.2.11 liblzma/5.2.3 bz2lib/1.0.6 liblz4/1.7.5 libzstd/1
Found libcurl/7.77.0-DEV Schannel zlib/1.2.11 zstd/1.4.5 libidn2/2.0.4 nghttp2/1.31.
```

tesseract --version



ดาวน์โหลดไฟล์ที่จะใช้งาน


[หน้าแรก](#)
[หลักสูตร](#)
[Download](#)
[Workshop](#)
[แบบประเมิน](#)
[วิทยากร](#)

Data Set

YOLO - Real-Time Object Detection
 ไฟล์ที่จะใช้ประกอบกับการเขียนโปรแกรมของ YOLO

ดาวน์โหลด >

Tesseract - OCR
 ไฟล์ที่จะใช้ประกอบกับการเขียนโปรแกรมของ Tesseract - OCR


✖

ดาวน์โหลด >

Face Recognition
 ไฟล์ที่จะใช้ประกอบกับการเขียนโปรแกรมของ Face Recognition

ดาวน์โหลด >

ดาวน์โหลดไฟล์ที่จะใช้งาน


[หน้าแรก](#)
[หลักสูตร](#)
[Download](#)
[Workshop](#)
[แบบประเมิน](#)
[วิทยากร](#)

Data Set

YOLO - Real-Time Object Detection
 ไฟล์ที่จะใช้ประกอบกับการเขียนโปรแกรมของ YOLO

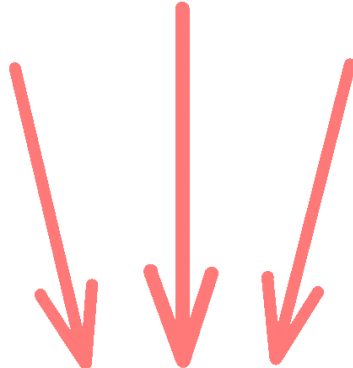
ดาวน์โหลด >


Tesseract - OCR
 ไฟล์ที่จะใช้ประกอบกับการเขียนโปรแกรมของ Tesseract - OCR

ดาวน์โหลด >

Face Recognition
 ไฟล์ที่จะใช้ประกอบกับการเขียนโปรแกรมของ Face Recognition

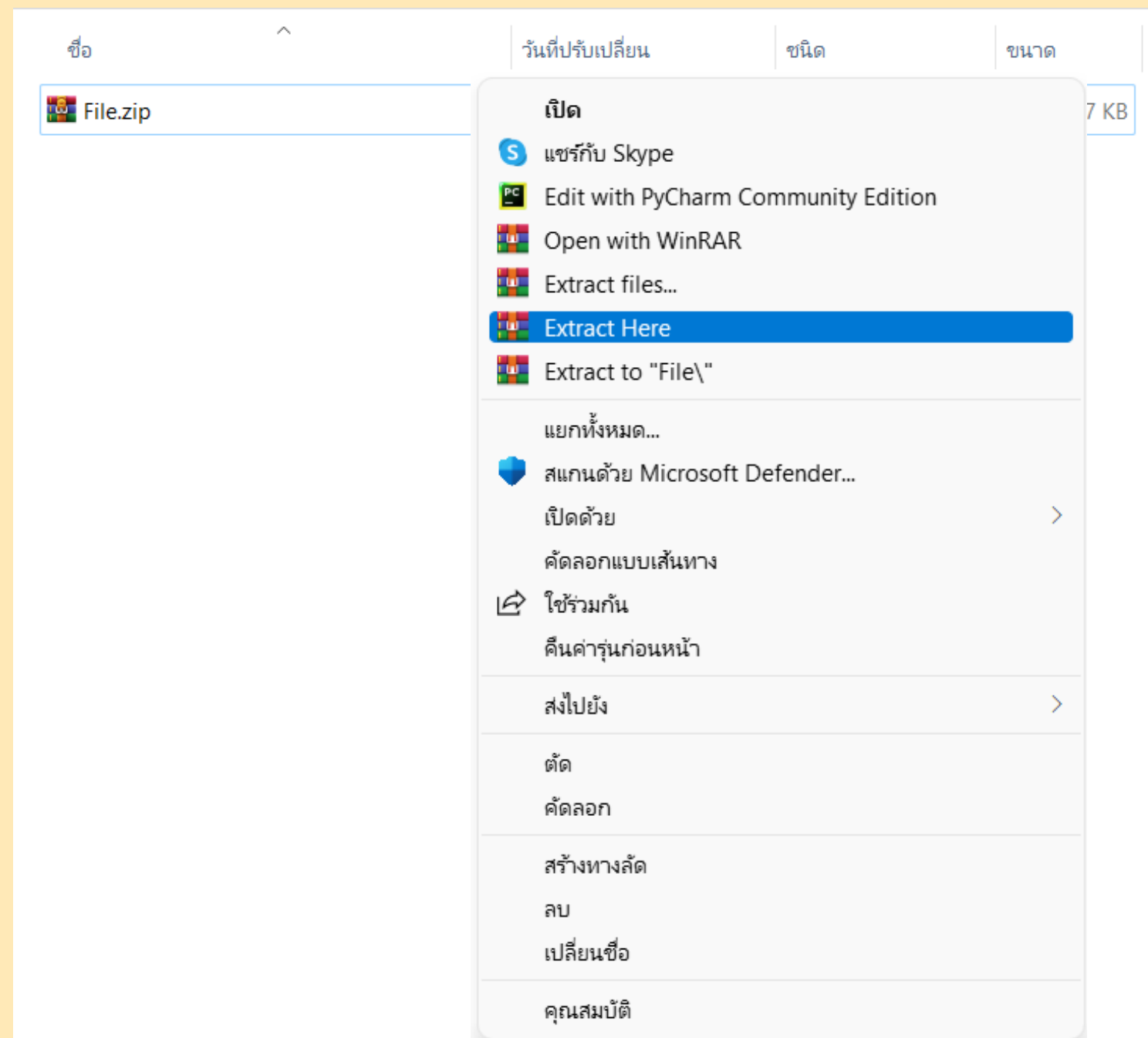
ดาวน์โหลด >



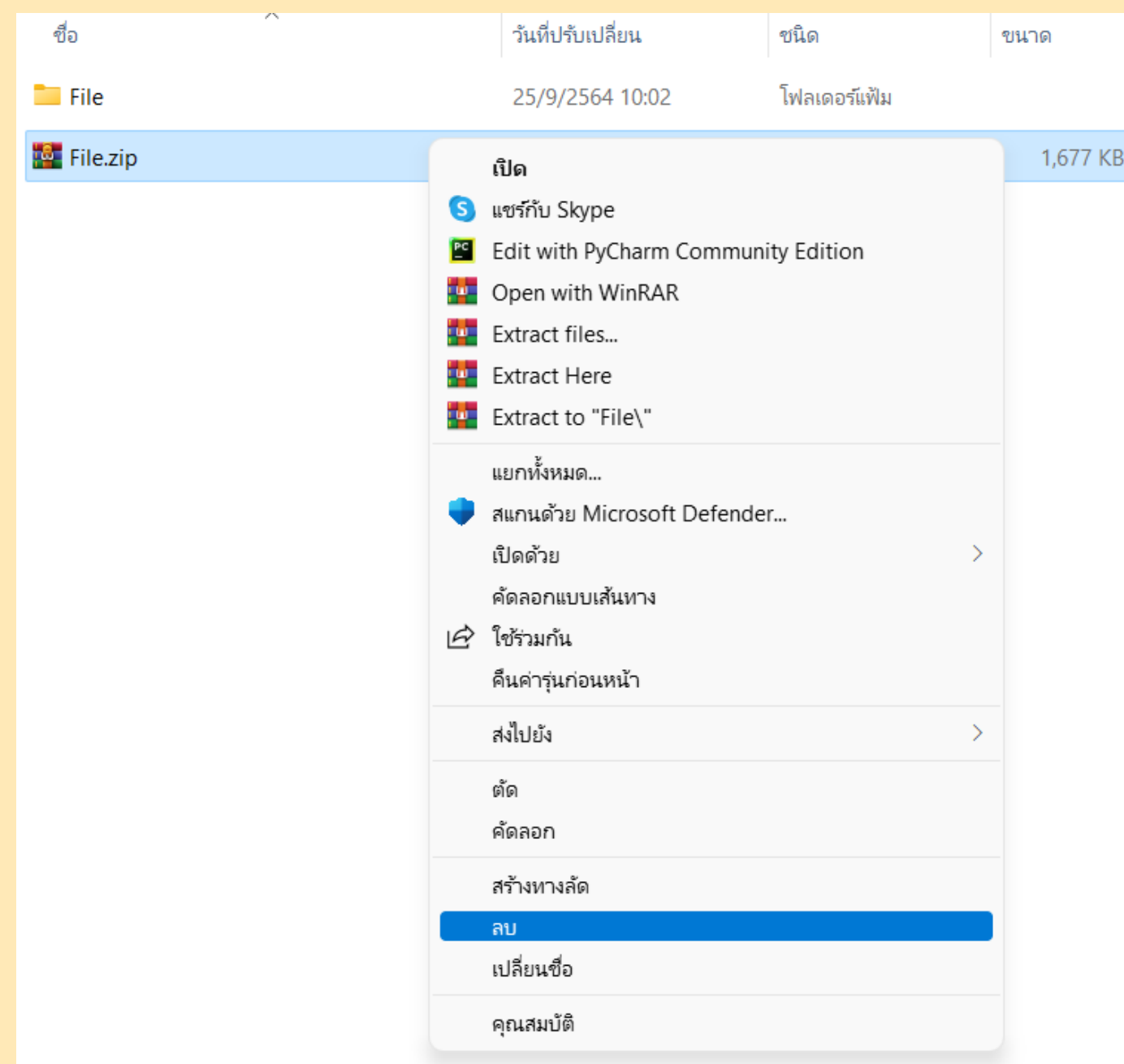

 File.zip

แสดงทั้งหมด

ดาวน์โหลดไฟล์ที่จะใช้งาน

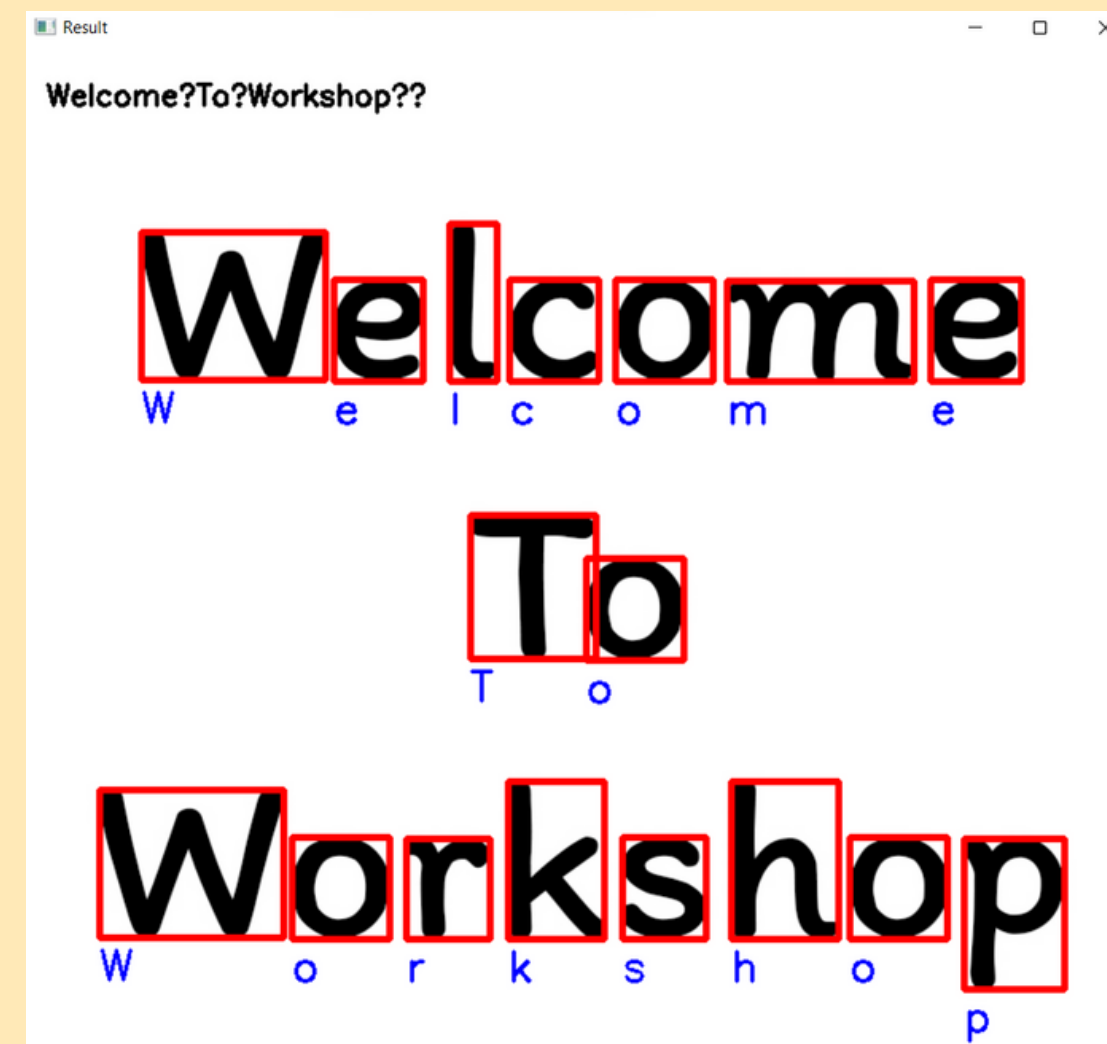


ดาวน์โหลดไฟล์ที่จะใช้งาน



Example 1

การสกัดข้อความ
จากรูปภาพ



Workshop 1

โปรแกรมสกัด ข้อความจาก รูปภาพ



Workshop1.txt - แผ่นจดบันทึก

ไฟล์ แก้ไข รูปแบบ ดู วิธีใช้

Welcome
To
Workshop

สิ่งที่ต้องเตรียม

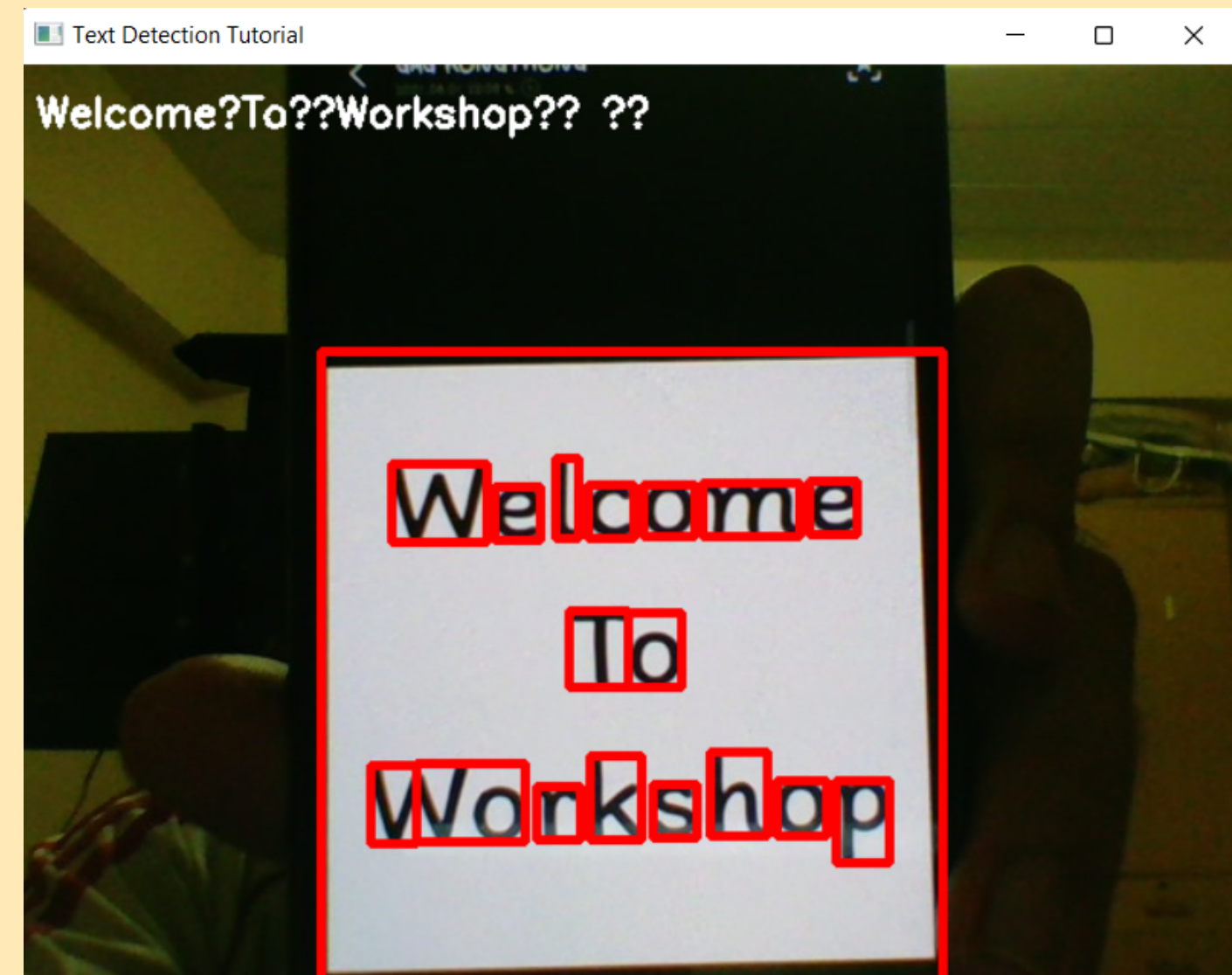
- 1 เครื่องคอมพิวเตอร์
- 2 Web Browser
- 3 ไฟล์รูปภาพ Welcome.jpg
- 4 ไฟล์รูปภาพ Text.jpg

โค้ดเพิ่มเติม

```
# นำข้อความที่อ่านได้เก็บเข้าตัวแปร imgcharsplit ในรูปแบบ List
imgcharsplit = imgchar.splitlines()
# แสดงข้อมูลใน List
print(imgcharsplit)
# เปิดไฟล์ Workshop1.txt ถ้ายังไม่มีจะสร้างโดยอัตโนมัติ
with open('Workshop1.txt', 'w') as result:
    # วนลูปอ่านข้อมูลใน List และนำไปเก็บไว้ในตัวแปร line
    for line in imgcharsplit:
        # ตรวจสอบว่าข้อมูลที่อ่านมาเป็นค่าว่างหรือไม่ ถ้าไม่ใช่ ก็เข้าเงื่อนไข
        if(line != ""):
            result.write(line) # เขียนข้อมูลลงในไฟล์ Workshop1.txt
            result.write('\n') # ขึ้นบรรทัดใหม่
```

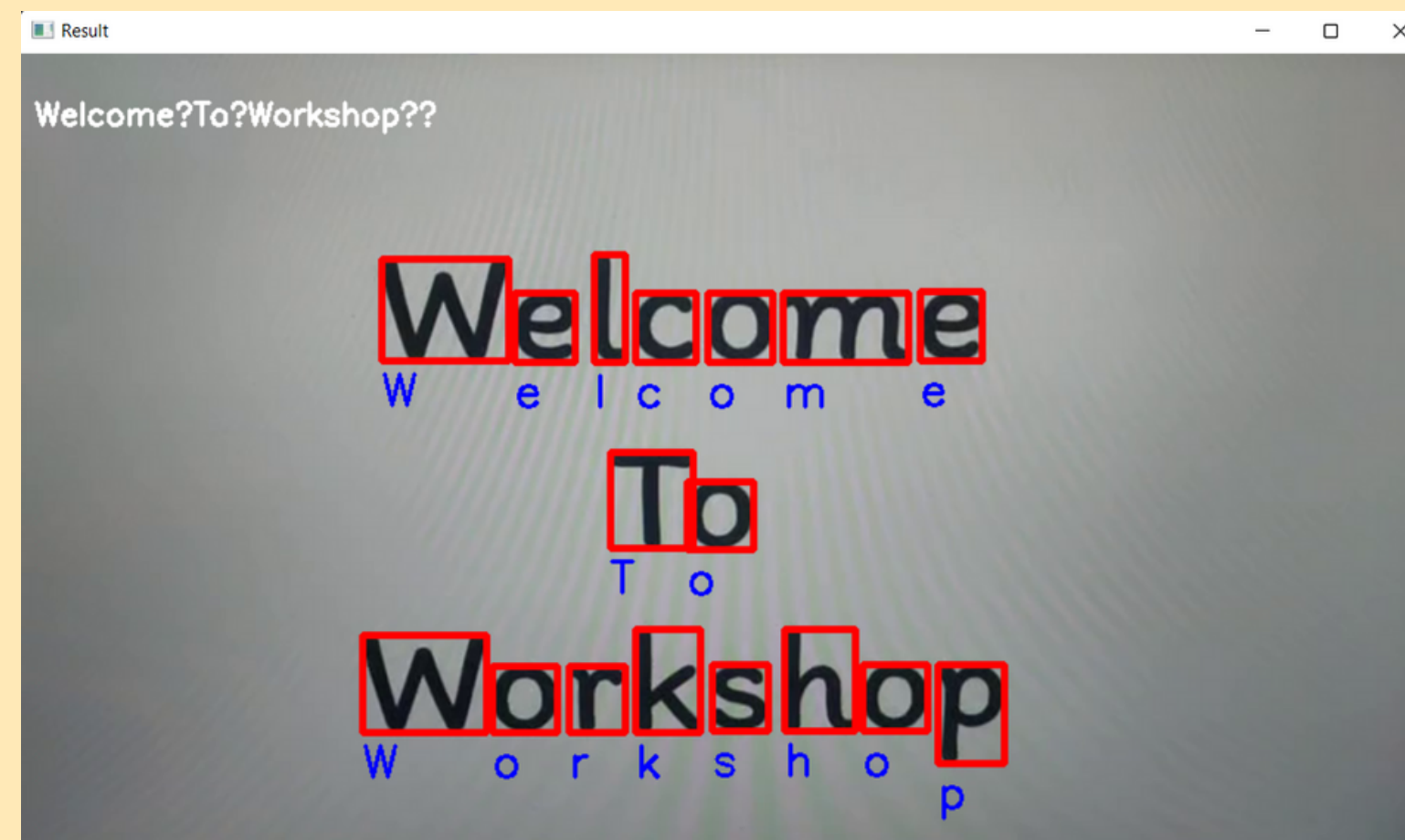

Example 2

การสกัดข้อความ
แบบตามเวลาจริง



Workshop 2

โปรแกรมสกัด ข้อความแบบตาม วิดีโอ



สิ่งที่ต้องเตรียม

- 1 เครื่องคอมพิวเตอร์
- 2 Web Browser
- 3 ไฟล์รูปภาพ Welcome.mp4
- 4 ไฟล์รูปภาพ Text.mp4

โค้ดเพิ่มเติม

```
# cv2.VideoCapture(ไฟล์วิดีโอ)
cap = cv2.VideoCapture("./File/Welcome.mp4")
# ตรวจสอบว่ากล้องพร้อมใช้งานหรือไม่
if not cap.isOpened():
    # ถ้าไม่จะแสดงข้อความ Cannot open video
    print("Cannot open video")
```