Image Recognition with Python

Computer Education

KMUTNB
INVENTION TO INNOVATION

LECTURER

NATTAPOL TONGYOUN

4th year Student

KMUTNB
INVENTION TO INNOVATION

LECTURER

NADA MUENCHOL

4th year Student

LET'S GO!

WORKSHOP 1 :

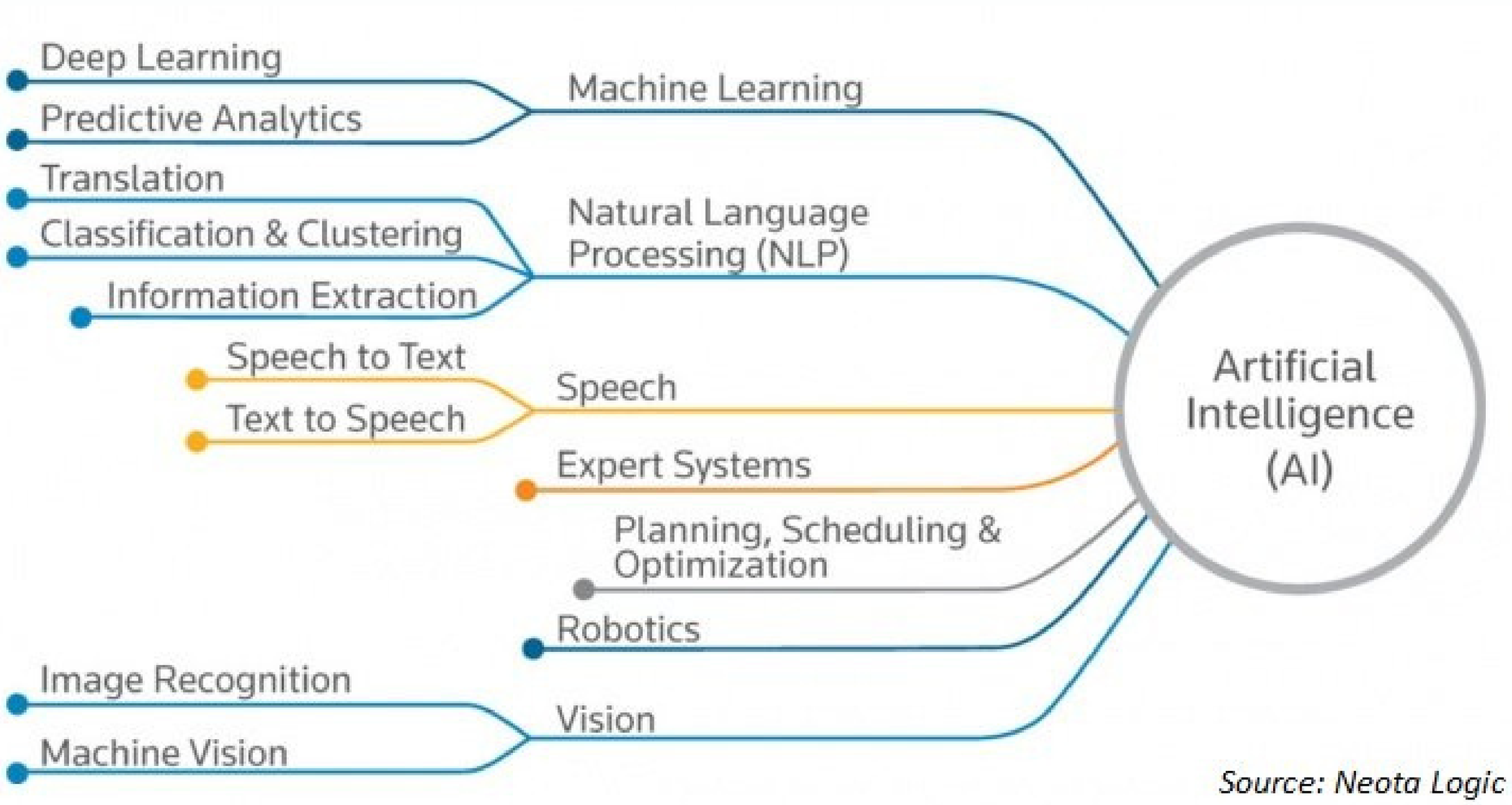
Artificial Intelligence

Artificial Intelligence

ปัญญาประดิษฐ์ (Artificial Intelligence) หรือ เอไอ (AI) หมายถึง ความฉลาดเทียมที่สร้างขึ้นให้กับสิ่งที่ไม่มีชีวิต ที่จะช่วยวิเคราะห์และตอบสนองพฤติกรรมของมนุษย์ได้ดียิ่งขึ้น

เหตุใดปัญญาประดิษฐ์ถึงมีความสำคัญ

- AI มีการเรียนรู้ซ้ำ ๆ ได้อย่างอัตโนมัติและศึกษาผ่านข้อมูลเหล่านั้น
- AI เพิ่มความชาญฉลาด
- AI เรียนรู้จากอัลกอริทึมการเรียนรู้แบบก้าวหน้า (progressive)
- AI จะวิเคราะห์ข้อมูลมากกว่าและลึกกว่า
- AI สามารถสร้างความแม่นยำอย่างเหลือเชื่อ
- AI สามารถใช้ประโยชน์อย่างสูงสุดจากข้อมูลที่มี


Deep Learning
Predictive Analytics
Machine Learning
Translation
Classification & Clustering
Information Extraction
Natural Language Processing (NLP)
Speech to Text
Text to Speech
Speech
Expert Systems
Planning, Scheduling & Optimization
Robotics
Image Recognition
Machine Vision
Vision
Artificial Intelligence (AI)
Source: Neota Logic

AI ในชีวิตประจำวัน

Computer Vision

Computer vision

เป็นแขนงหนึ่งของวิทยาการปัญญาประดิษฐ์หรือ AI ซึ่งทำการ **"ฝึกฝน"** คอมพิวเตอร์ และระบบให้สามารถเข้าใจและตอบสนองต่อข้อมูลภาพได้อย่างชาญฉลาด จากภาพหรือวิดีโอ

และแบบจำลอง deep learning นั้น อุปกรณ์ต่าง ๆ จะสามารถเรียนรู้ที่จะระบุและทราบถึงวัตถุต่าง ๆ จากนั้นจะสามารถทำการตอบสนองต่อสิ่งที่มัน "มองเห็น" ได้ต่อไป

เป้าหมายของ Computer Vision

- การตรวจจับ การแบ่งขอบเขต การระบุตำแหน่ง และการจดจำวัตถุ
- การประเมินผลสำหรับการแบ่งลักษณะของวัตถุในภาพหรือเปรียบเทียบวัตถุ
- การเปรียบเทียบวัตถุในมุมมองต่าง ๆ ของรูปภาพหรือวัตถุนั้น ๆ
- การเชื่อมโยงมุมมองต่าง ๆ ของรูปภาพ
- ในการค้นหาวัตถุเหล่านั้นด้วยรูปภาพ จำเป็นต้องมีฐานข้อมูลขนาดใหญ่

การประยุกต์ใช้งาน Computer Vision

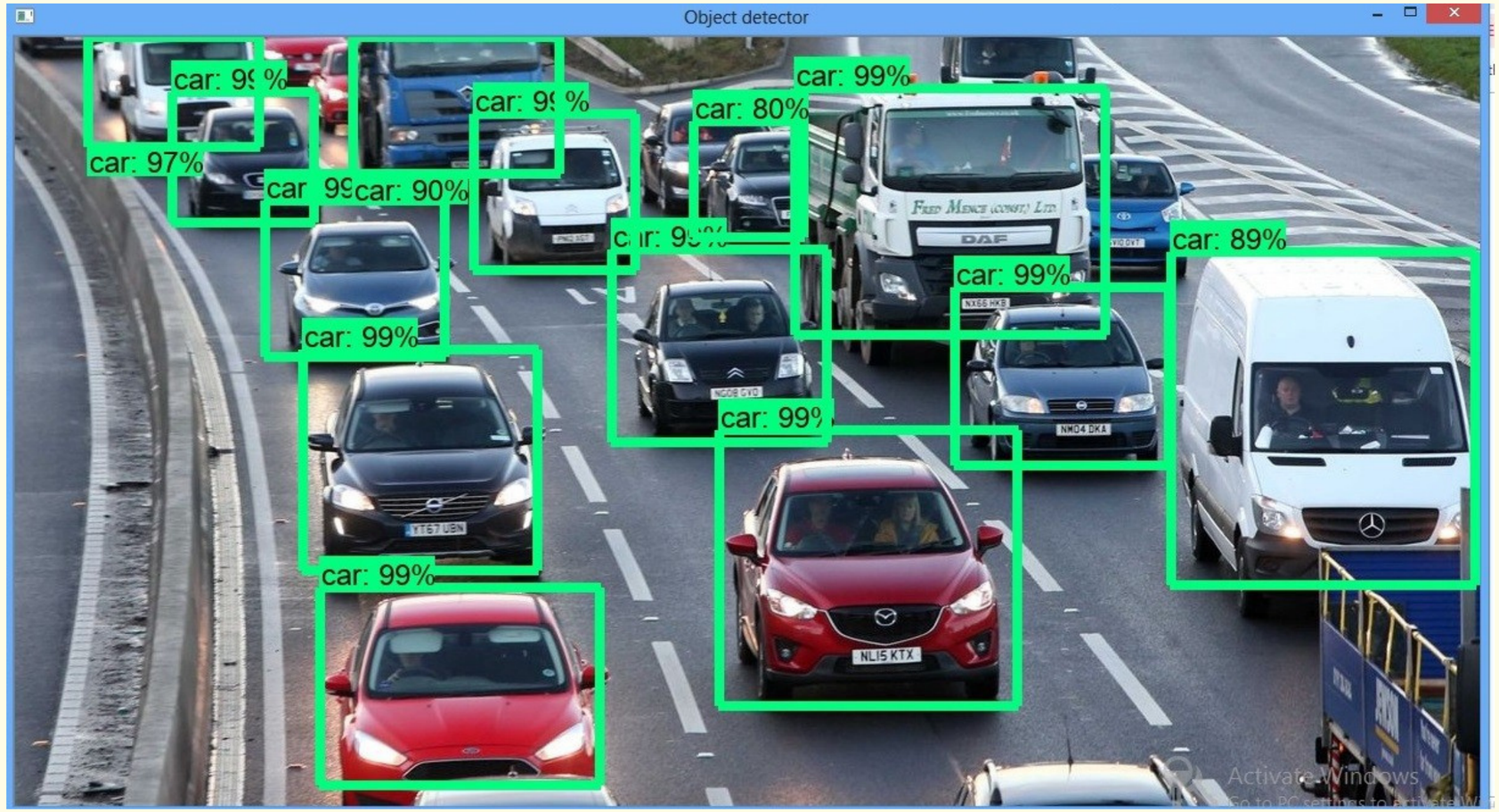
- Image segmentation
- Object detection
- Facial recognition
- Edge detection
- Pattern detection
- Image classification
- Feature matching

ตัวอย่างการใช้งาน Computer Vision

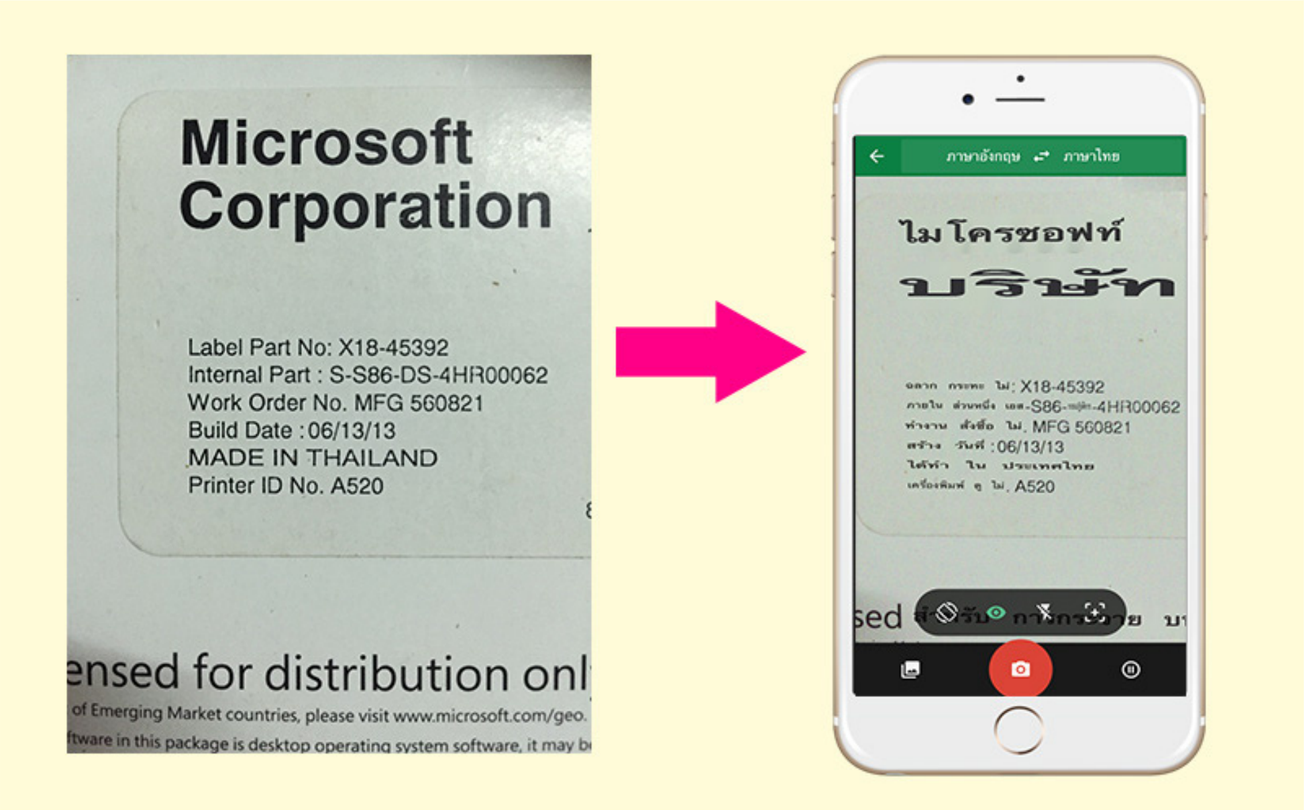
ระบบเก็บข้อมูลการจอดรถ

การนำไปใช้ประโยชน์ด้านการแพทย์

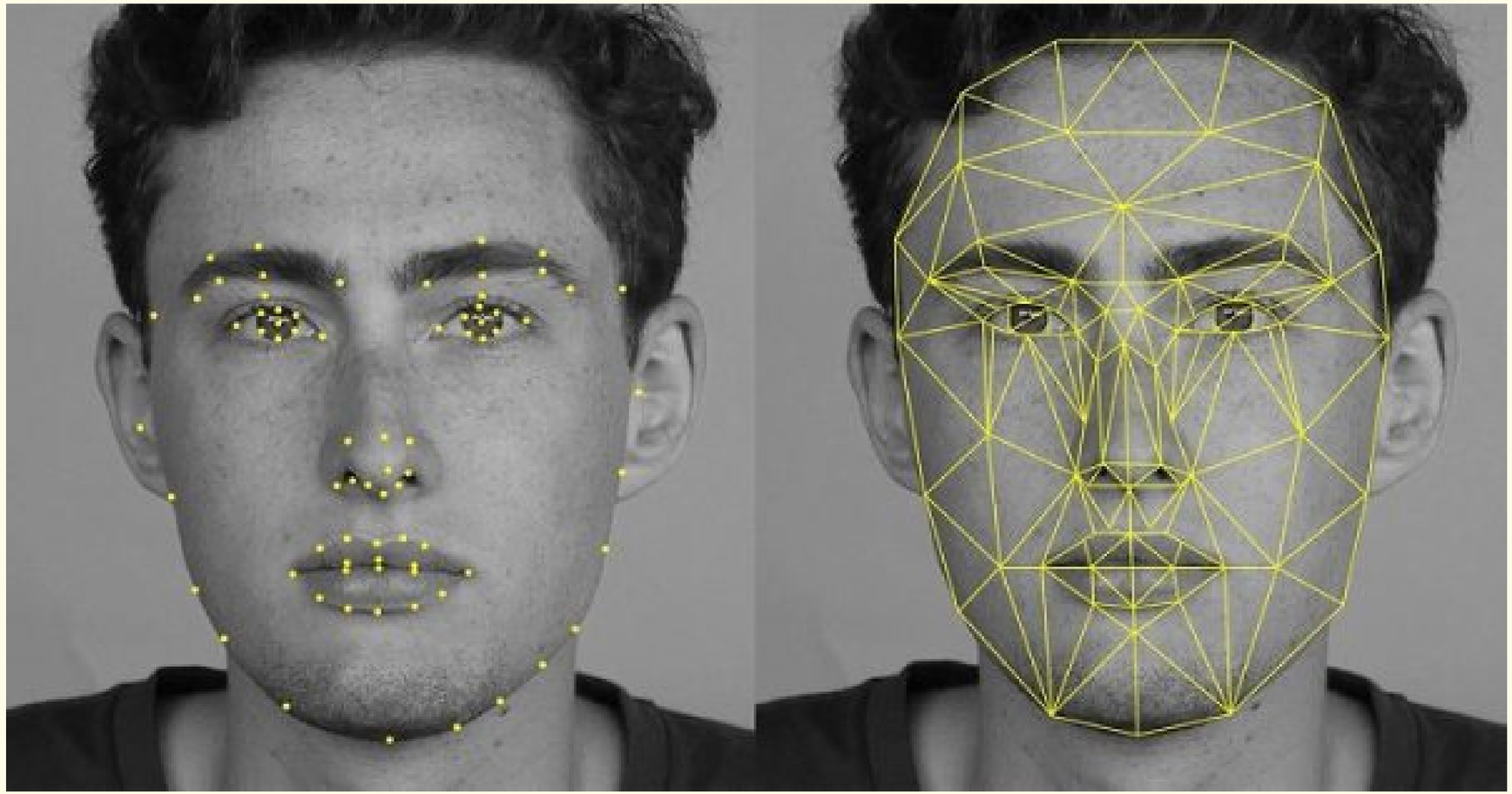
ระบบเก็บข้อมูลจราจร



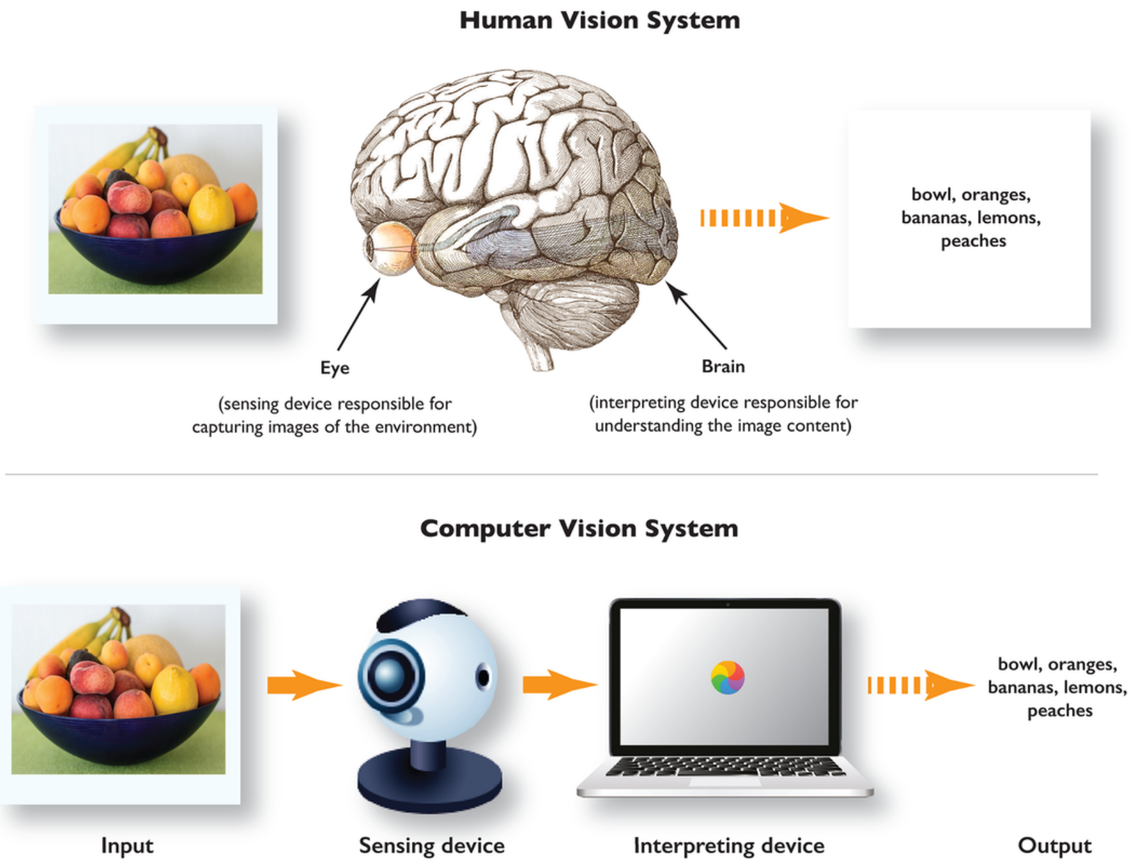
ระบบแปลภาษา



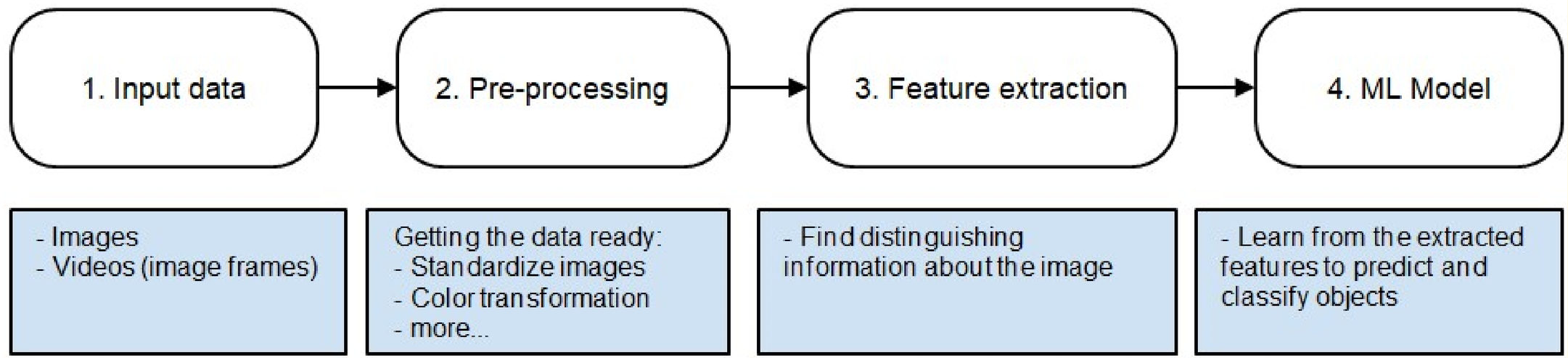
ระบบสแกนใบหน้า



เปรียบเทียบการทำงาน Computer Vision



Computer Vision pipeline

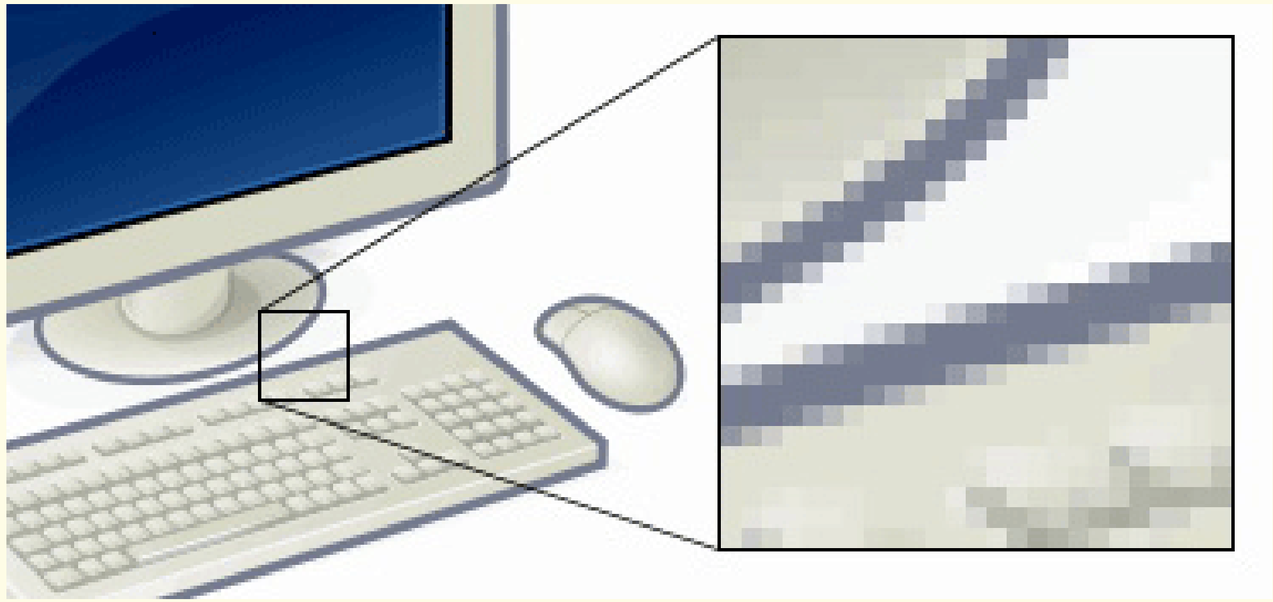


*Pipeline = กระบวนการทั้งหมดที่ถูกส่งต่อกันมาจากแผนกหนึ่งไปยังแผนกหนึ่ง

What is Pixel?

Pixel

เป็นหน่วยพื้นฐานซึ่งเล็กที่สุดของภาพดิจิตอล เทียบได้กับจุดสีของภาพ 1 จุด หลากหลายสี หลายๆจุดที่เรียงชิดติดกันถูกรวมกันทำให้เกิดเป็นภาพ

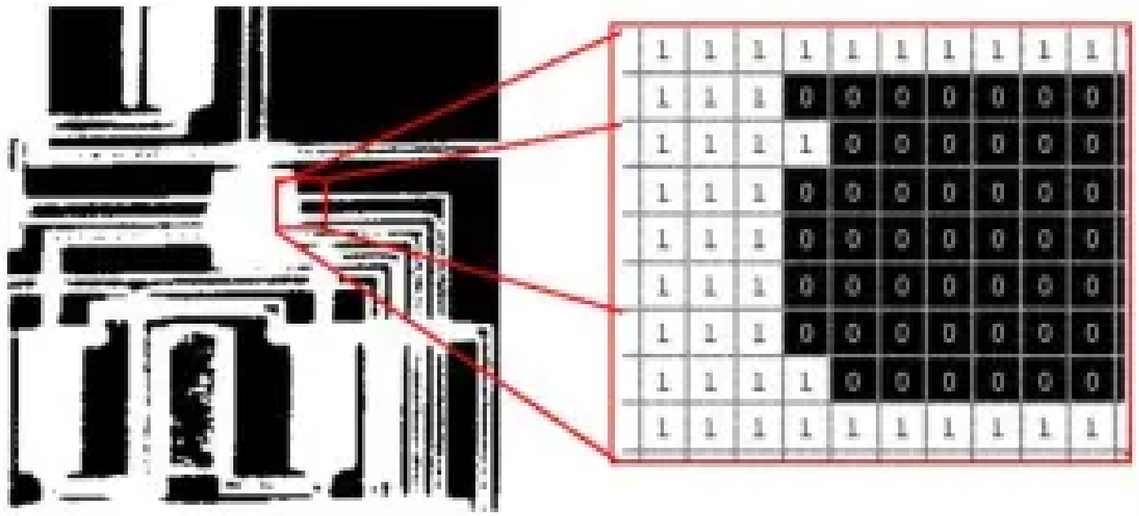


Digital Image

ภาพดิจิทัล เป็นการแสดงผลภาพในลักษณะสองมิติในหน่วยที่เรียกว่าพิกเซล ภาพดิจิทัลจะมีรูปแบบการเก็บเป็นเมทริกซ์ ซึ่งจะมีการจัดเก็บภาพแต่ละชนิดต่างกันซึ่งจะมีการจัดเก็บภาพแต่ละชนิดต่างกัน ขึ้นอยู่กับระบบสีของภาพดังกล่าว

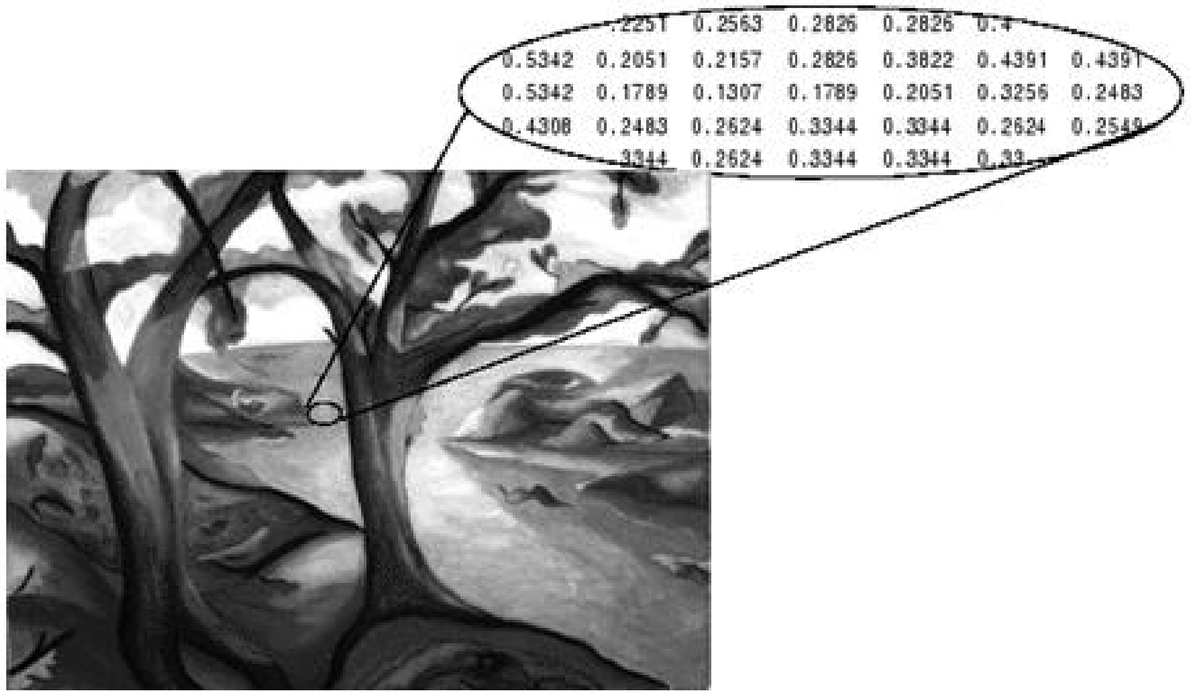
Binary Image

ไบนารีในทางดิจิตอลหมายถึงว่ามีเพียง 2 สถานะคือ 0 และ 1 ซึ่งภาพไบนารี่เป็นรูปที่ใช้เนื้อที่เพียง 1 บิต มีแค่ความเข้ม 2 ค่าเท่านั้นคือ 0 และ 1 เป็นรูปที่ใช้เนื้อที่เพียง 1 บิตหมายความว่าพิกเซลใดที่มีค่าเป็น 0 ก็จะหมายถึงว่าพิกเซลนั้นจะแสดงสีดำ พิกเซลใดที่มีค่าเป็น 1 จะหมายถึงว่าพิกเซลนั้นจะแสดงสีขาว



Grayscale Image

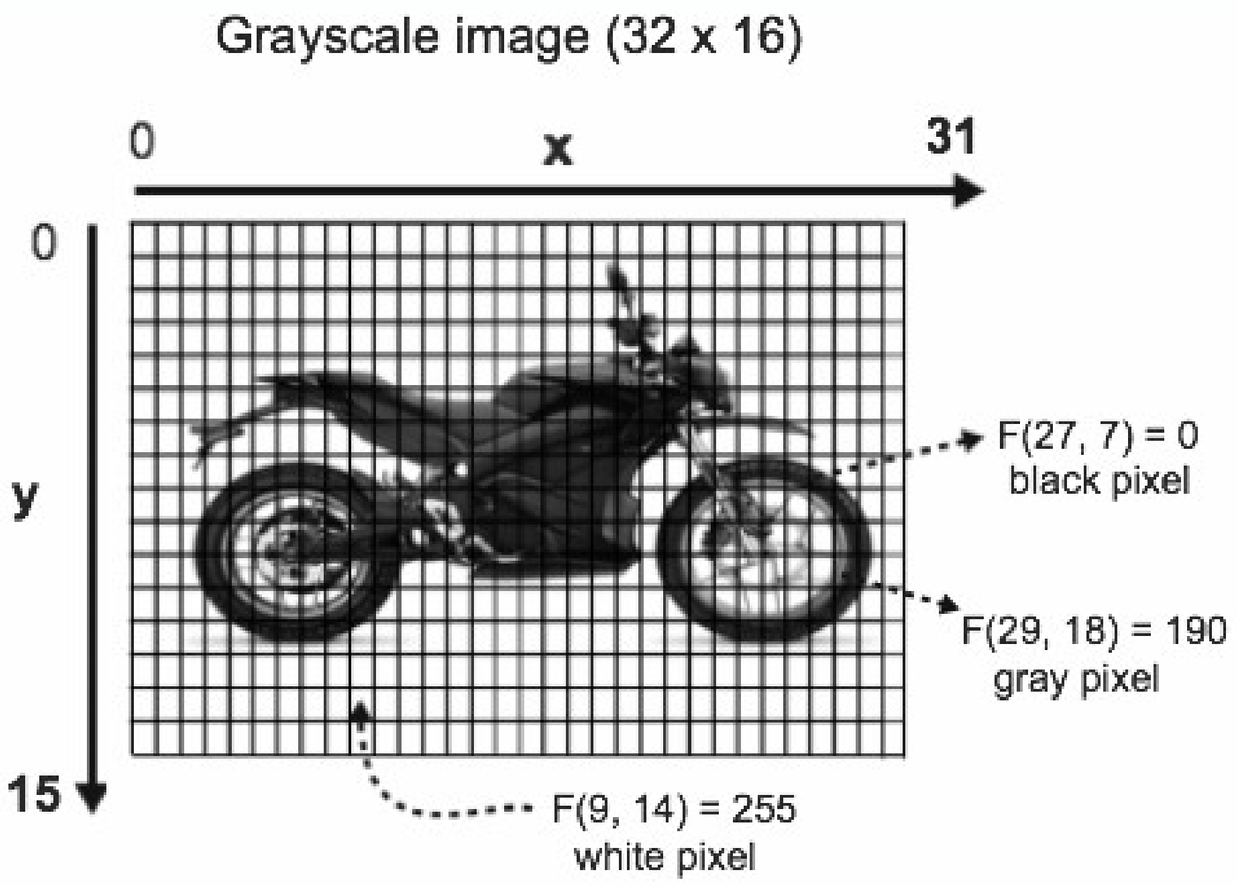
Grayscale Image เป็นรูปที่เก็บโดยใช้รูปแบบของอาร์เรย์ 2 มิติ โดยค่าที่เก็บจะมีค่าอยู่ในช่วง ๆ หนึ่ง ซึ่งระดับของสีขึ้นอยู่กับขนาดของบิตที่ใช้เก็บค่าสี



RGB Image

RGB” คือสีของรูปภาพที่เกิดจากการผสมสีของแสง 3 สี คือ แดง (Red) เขียว (Green) และน้ำเงิน (Blue)

ตัวอย่างการนำไปใช้



What we see

25	43	11	04	70	87	12	31	43	10	05	77	12	06	45	09	29	30	02
56	22	75	03	22	96	45	12	23	03	77	67	81	45	22	04	90	22	21
32	45	41	91	87	62	35	02	00	11	62	25	43	11	04	70	87	12	61
31	43	10	05	77	12	06	45	09	29	30	56	22	75	03	22	96	45	05
12	23	03	77	67	81	45	22	04	90	22	32	45	41	91	87	62	35	44
02	00	11	62	25	43	11	04	70	87	12	31	43	10	05	77	12	06	10
45	09	29	30	56	22	75	03	22	96	45	12	23	03	77	67	81	45	55
22	04	90	22	32	45	41	91	87	62	35	02	00	11	62	25	43	11	80
04	70	87	12	31	43	10	05	77	12	06	45	09	29	30	56	22	75	08
03	22	96	45	12	23	03	77	67	81	45	22	04	90	22	32	45	41	99
91	87	62	35	02	00	11	62	22	01	00	72	65	23	01	00	22	04	30
90	22	32	45	41	91	87	62	35	02	00	11	62	25	43	11	04	70	42
87	12	31	43	10	05	77	12	06	45	09	29	30	56	22	75	03	22	91
96	45	12	23	03	77	67	81	45	22	04	90	22	32	45	41	91	87	40
62	35	02	00	11	62	22	01	00	72	65	23	01	00	56	22	75	03	67
22	96	45	12	23	03	77	67	81	45	22	04	90	22	32	45	41	91	22

What computers see

Image Processing

กระบวนการจัดการและวิเคราะห์รูปภาพให้เป็นข้อมูลในแบบดิจิทัล โดยใช้คอมพิวเตอร์เพื่อให้ได้ข้อมูลที่ต้องการทั้งในเชิงคุณภาพและปริมาณ (ขนาด รูปร่าง) หลังจากนั้นสามารถนำข้อมูลไปวิเคราะห์และสร้างเป็นระบบ

เทคนิคที่ใช้ในการประมวลผลภาพ

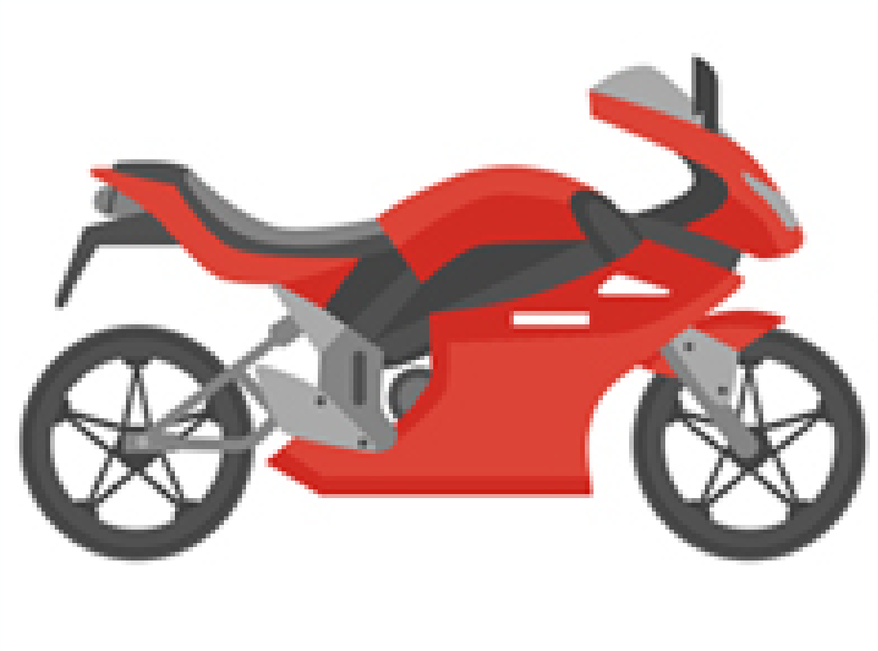
1. ปรับปรุงคุณภาพของภาพ (Image Enhancement)
2. การกรองภาพหรือการกำจัดสัญญาณรบกวนออกจากภาพ (Image Filters)
3. การซ้อนทับภาพ (Image Registration)
4. การคืนสภาพของภาพ (Image Restoration)
5. การแบ่งส่วนภาพ (Image Segmentation)

เทคนิคที่ใช้ในการประมวลผลภาพ

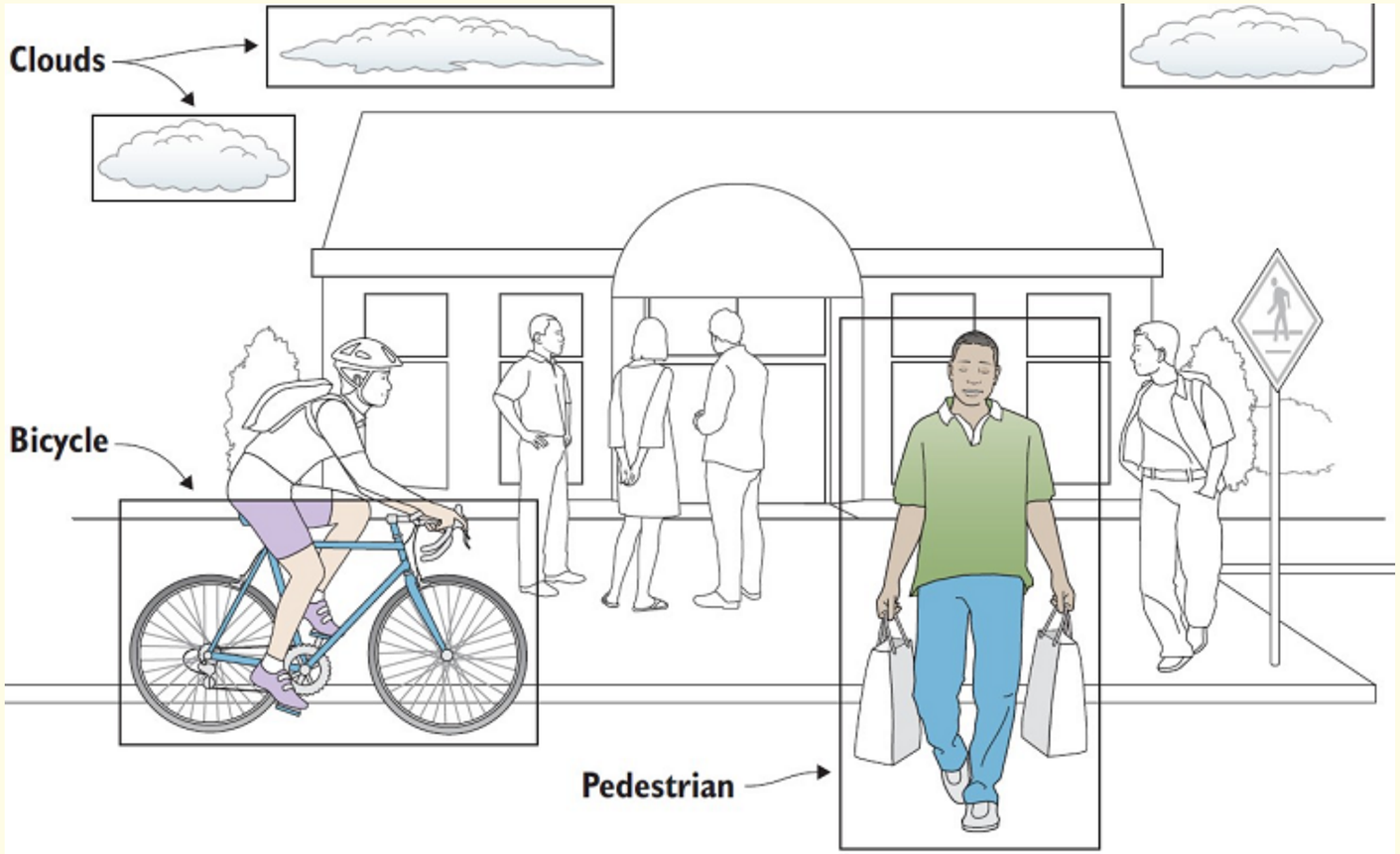
6. การหาขอบภาพในวัตถุ (Image Segmentation and EdgDeTection)

7. การบีบอัดภาพ (Image Compression)
8. การสร้างภาพ 3 มิติ (3D Image Reconstruction)

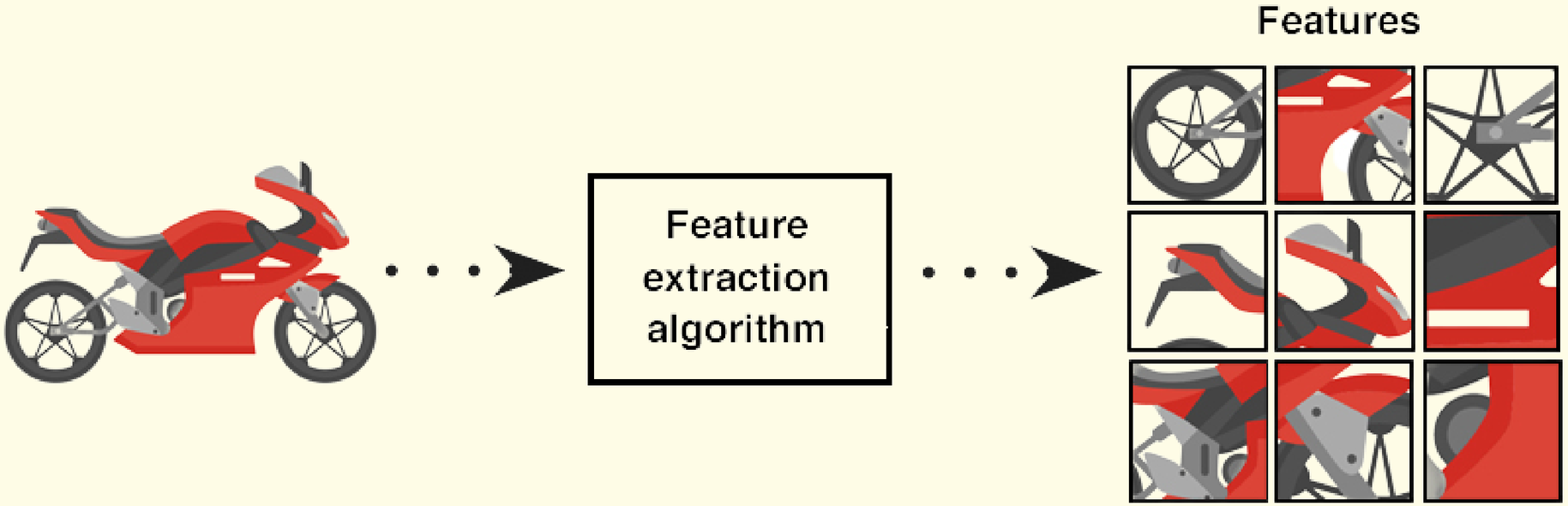
Input Image



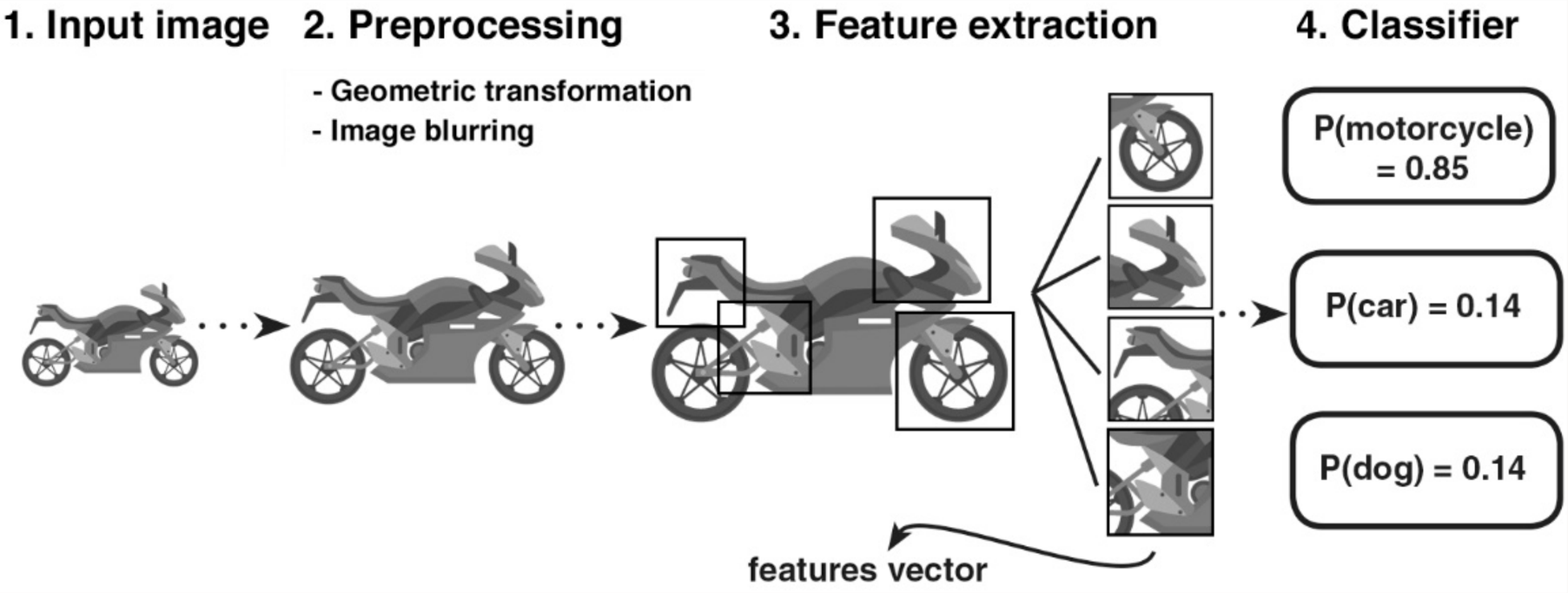
Pre-processing



Feature Extraction



Classifier



*Class = สิ่งที่ใช้แทนเพื่อให้เกิดเป็นผลลัพธ์สุกท้ายในการจำแนก

Classifier

01 Logistic Regression

02 Dicision Trees

03 Support vector Machines

04 Naive Bayes Classifier

05 K-Nearest Neighbour (KNN)

06 Convolutional Neural Network

Logistic Regression

การวิเคราะห์การถดถอยโลจิสติก เป็นการวิเคราะห์เพื่อทำนายโอกาสที่จะเกิด โดยอาศัยสมการโลจิสติกที่สร้างขึ้นจากชุดตัวแปรทำนาย ที่เป็นตัวแปรที่มีข้อมูลอยู่ ในระดับช่วงเป็นอย่างน้อย โดยที่ระหว่างตัวแปรทำนายจะต้องมีความสัมพันธ์กันต่ำและในการวิเคราะห์ จะต้องใช้ขนาดตัวแปรทำนายไม่ต่ำกว่า 30 ตัวแปร

Decision Trees

เป็นการเรียนรู้ที่ใช้จำแนกประเภทของตัวอย่าง โดยระบบจะเรียนรู้จากตัวอย่างที่สอนให้ แล้วเก็บความรู้ที่ได้ในรูปแบบของต้นไม้ตัดสินใจ ซึ่งอาจถูกแปลงไปเป็นอีกกฎอีกทีหนึ่ง นิยมใช้มากใน Machine Learning

Support Vector Machines

เป็นอัลกอริทึมที่สามารถนำมาช่วยแก้ปัญหาการจำแนกข้อมูล ใช้ในการวิเคราะห์ข้อมูลและจำแนกข้อมูล โดยอาศัยหลักการของการหาสัมประสิทธิ์ของสมการเพื่อสร้างเส้นแบ่งแยกกลุ่มข้อมูลที่ถูกป้อนเข้าสู่กระบวนการสอนให้ระบบเรียนรู้ โดยเน้นไปยังเส้นแบ่งแยกแยะกลุ่มข้อมูลได้ดีที่สุด

K-Nearest Neighbour

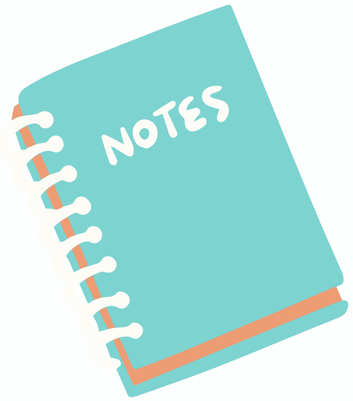
เป็นวิธีการจำแนกประเภทข้อมูลวิธีหนึ่ง โดยจัดว่าเป็นการจำแนกแบบมีผู้ฝึกสอน (Supervised Machine Learning Algorithm) หรือ ทราบคำตอบอยู่แล้ว จากนั้นจะใช้โมเดลในการจำแนกประเภทข้อมูลจากข้อมูลที่รู้คำตอบ

Naive Bayes Classifier

เป็น Model หนึ่งในการแบ่งกลุ่มที่ต้องการโดยใช้ความน่าจะเป็นที่ชื่อว่า Naive bayes ยกตัวอย่างเช่น ต้องการแบ่งกลุ่มว่าคนไข้ที่เข้ามานั้นเป็นไข้หวัดใหญ่หรือไม่ ซึ่งจะต้องถามอาการคนไข้มาให้ได้มากที่สุดว่าอาการเป็นอย่างไร แล้วถึงจะคาดคะเนจากข้อมูลอาการที่ได้ว่ามีความน่าจะเป็นไข้หวัดใหญ่เท่าไหร่ ซึ่งความน่าจะเป็นดังกล่าวจะใช้ Naive Bayes ในการหาค่าความน่าจะเป็น

Convolutional Neural Network (CNN)

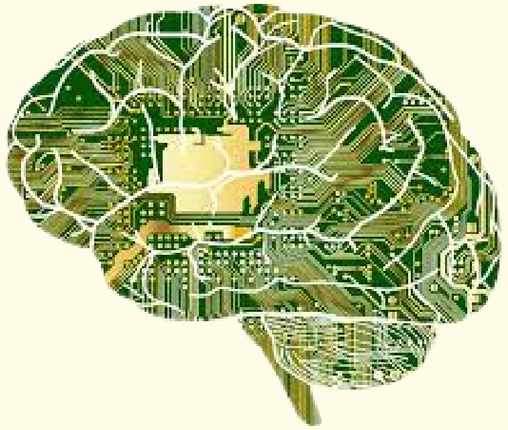
โครงข่ายประสาทแบบคอนโวลูชัน โดยจะจำลองการมองเห็นของมนุษย์ที่มองพื้นที่เป็นที่ย่อยๆ และนำกลุ่มของพื้นที่ย่อยๆ มาผสานกัน เพื่อดูว่าสิ่งที่เห็นอยู่เป็นอะไรกันแน่



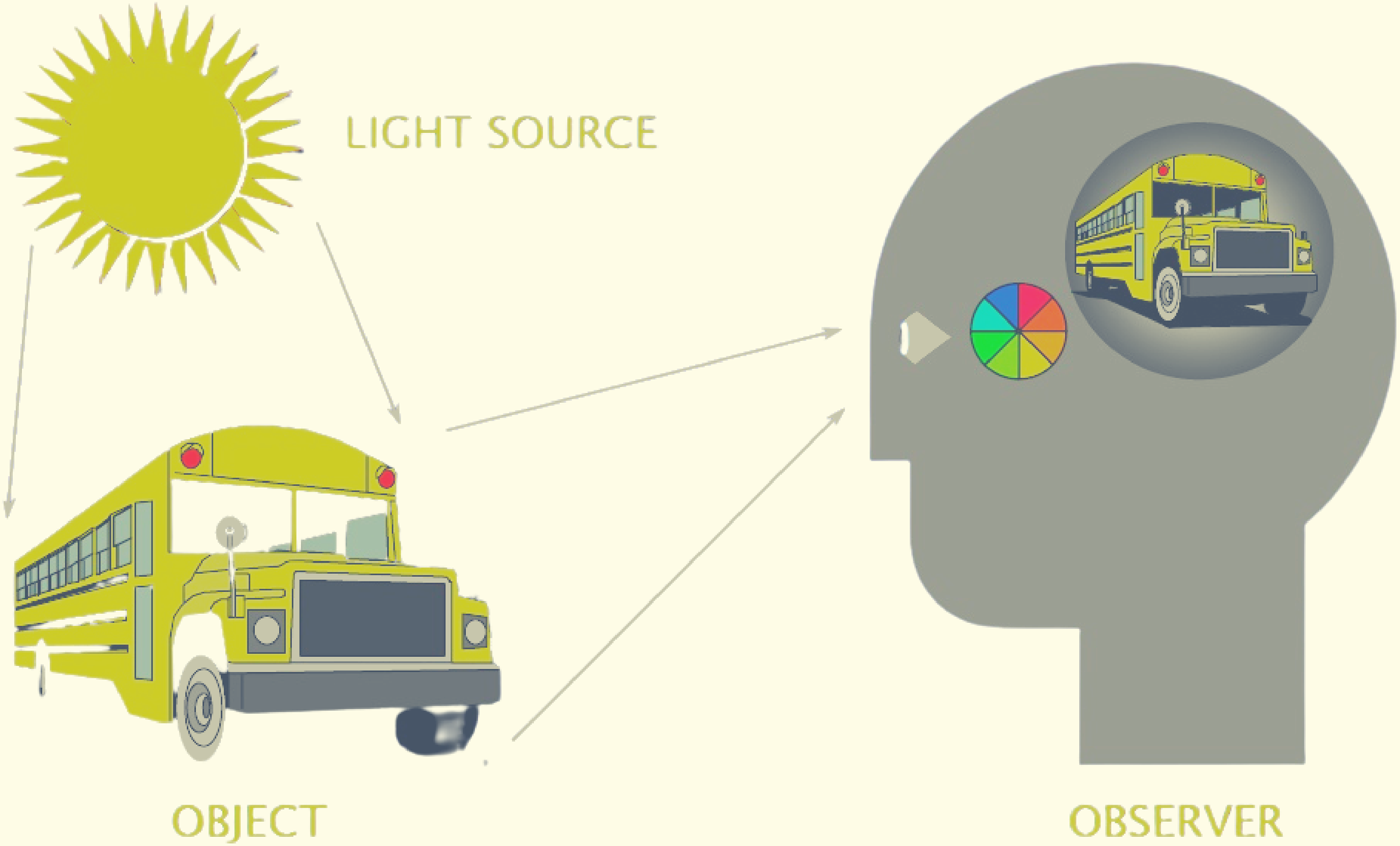
Neural Network

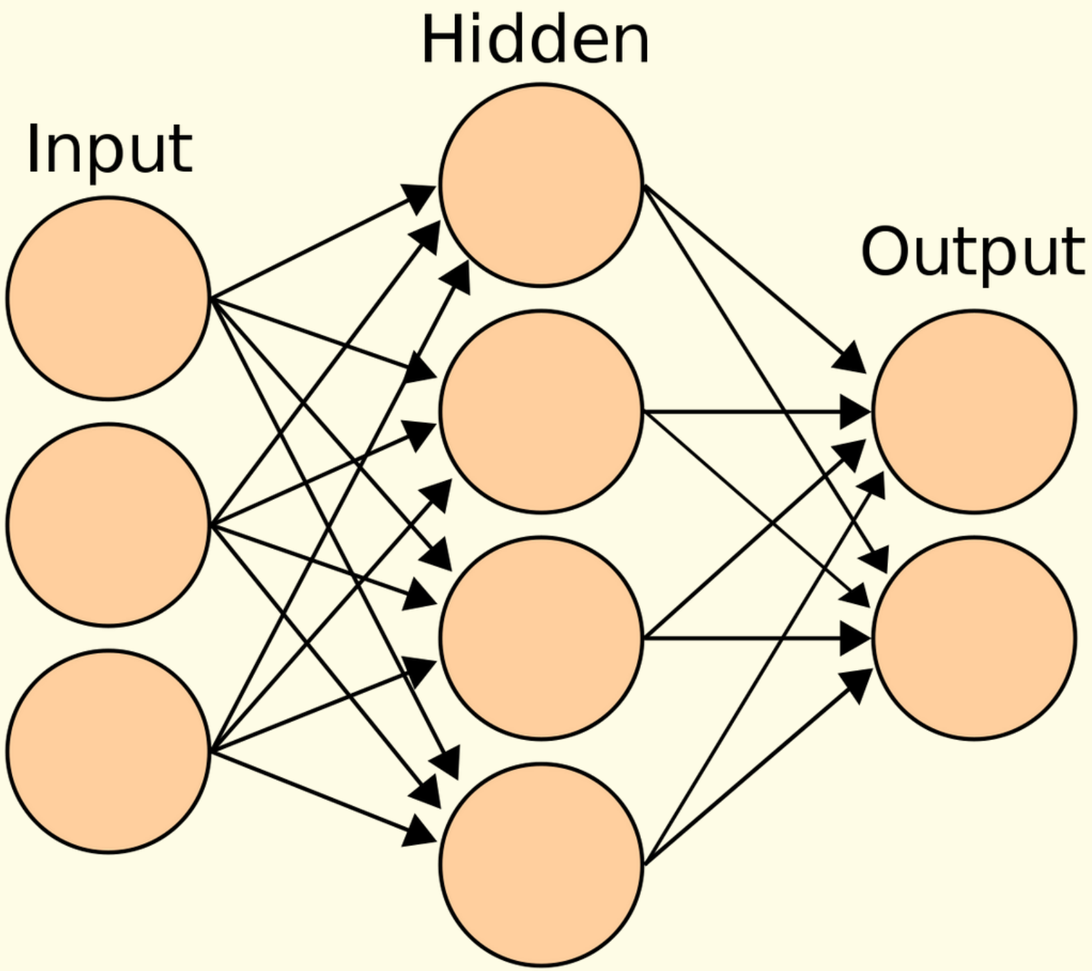
Neural Network

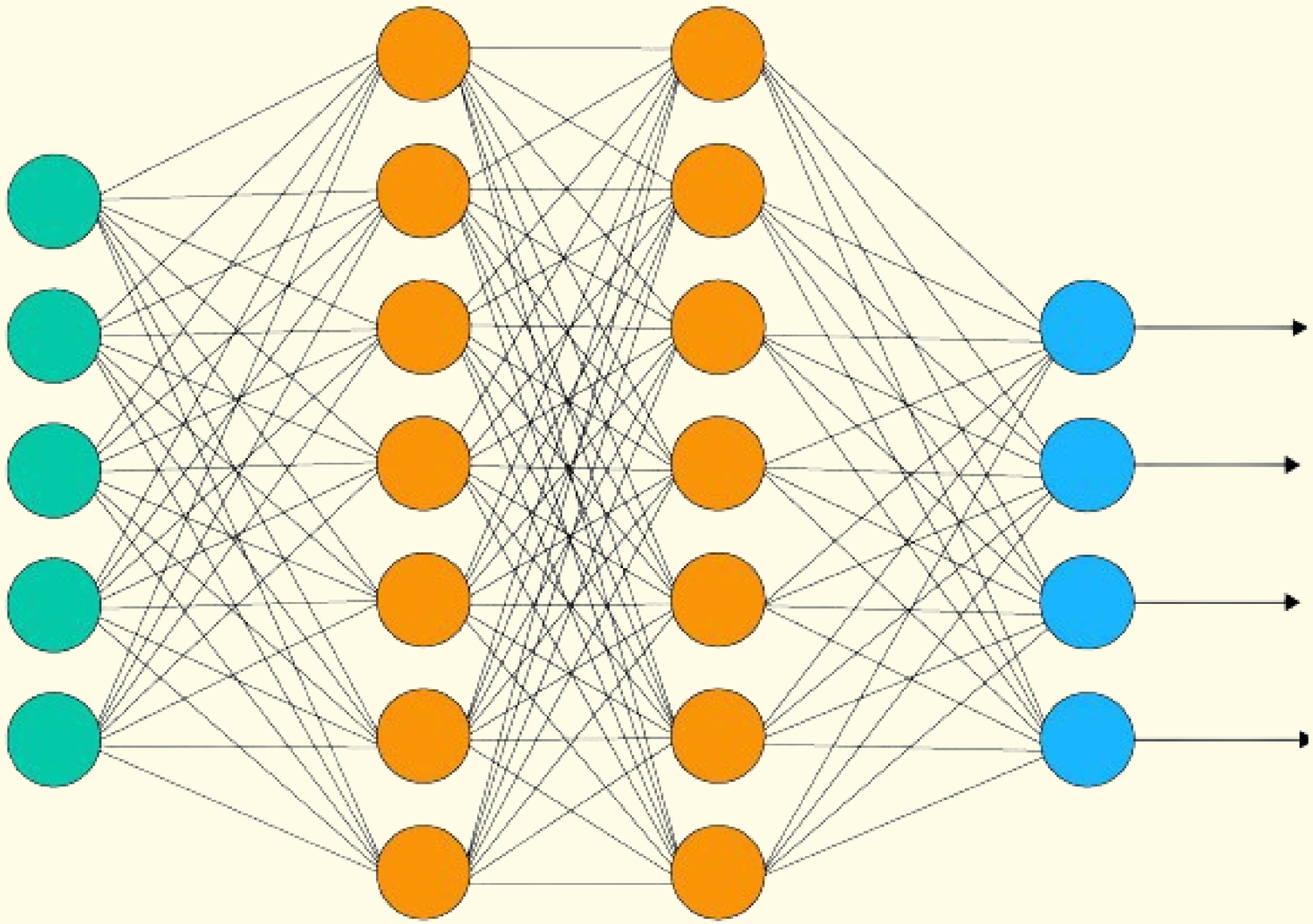
Neural Network หรือ Artificial Neural Network คือ โครงข่ายประสาทเทียม เป็นสาขาหนึ่งของปัญญาประดิษฐ์ Artificial Intelligence (AI) เป็นแนวคิดที่ออกแบบระบบโครงข่ายคอมพิวเตอร์ให้เลียนแบบการทำงานของสมองมนุษย์

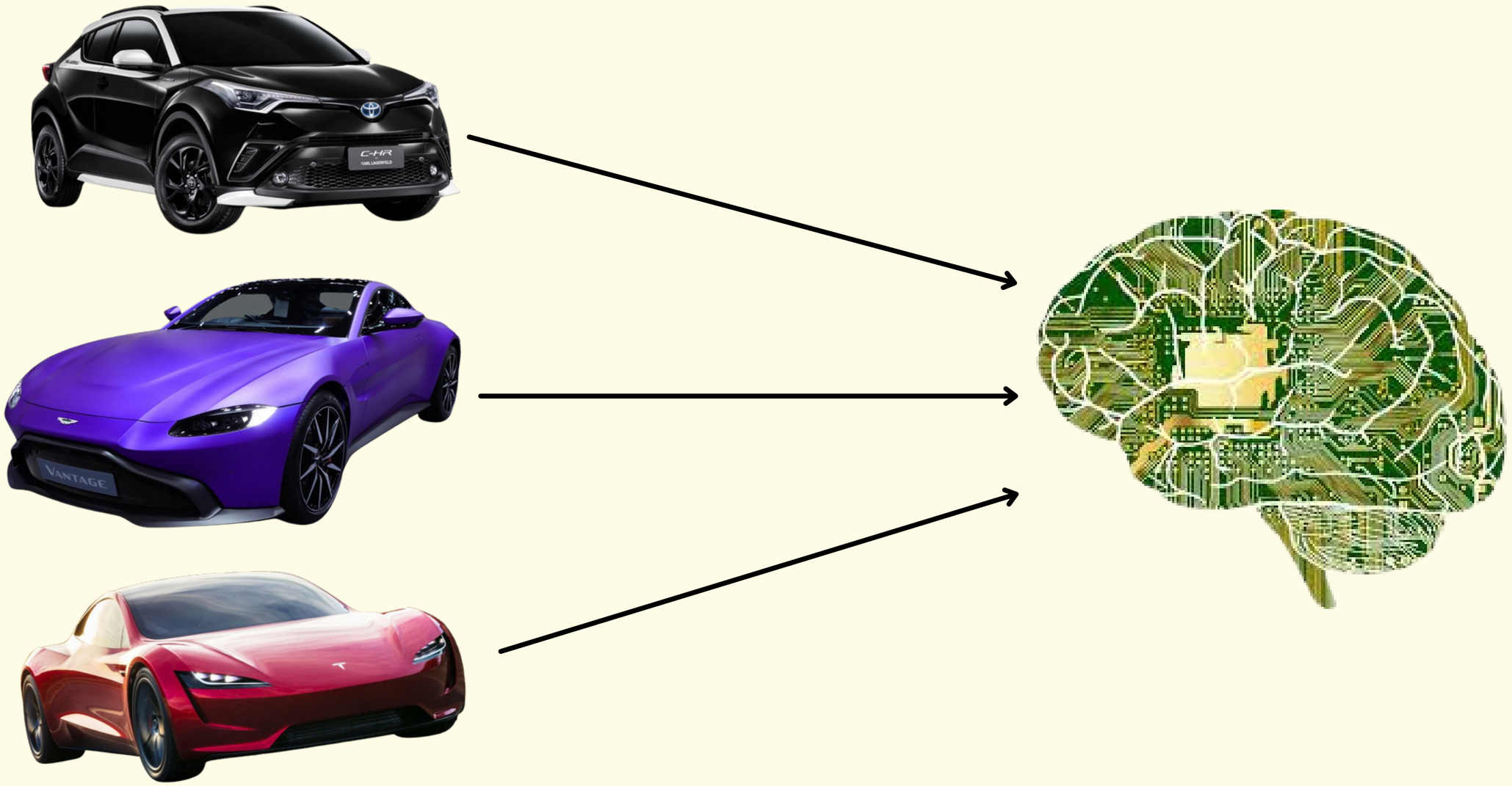


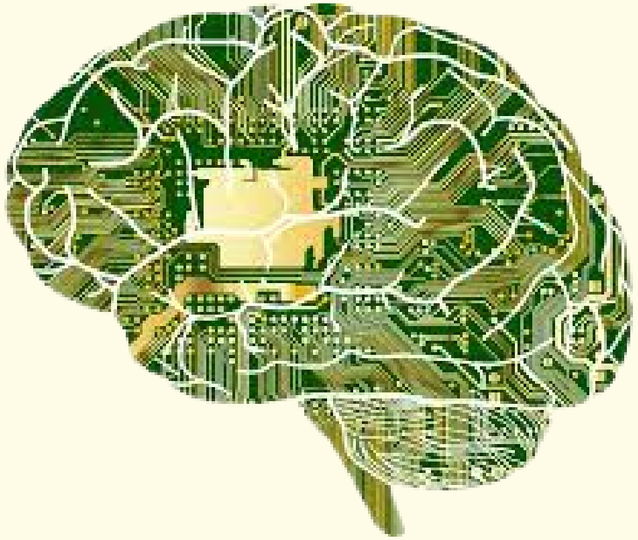
VANTAGE


LIGHT SOURCE
OBJECT
OBSERVER


Hidden
Input
Output




C-HR
VANTAGE

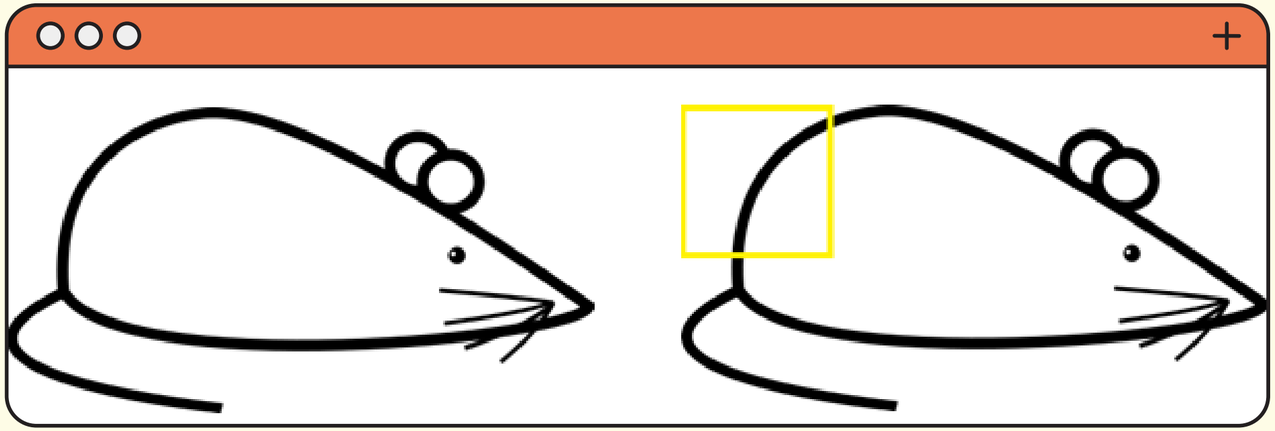


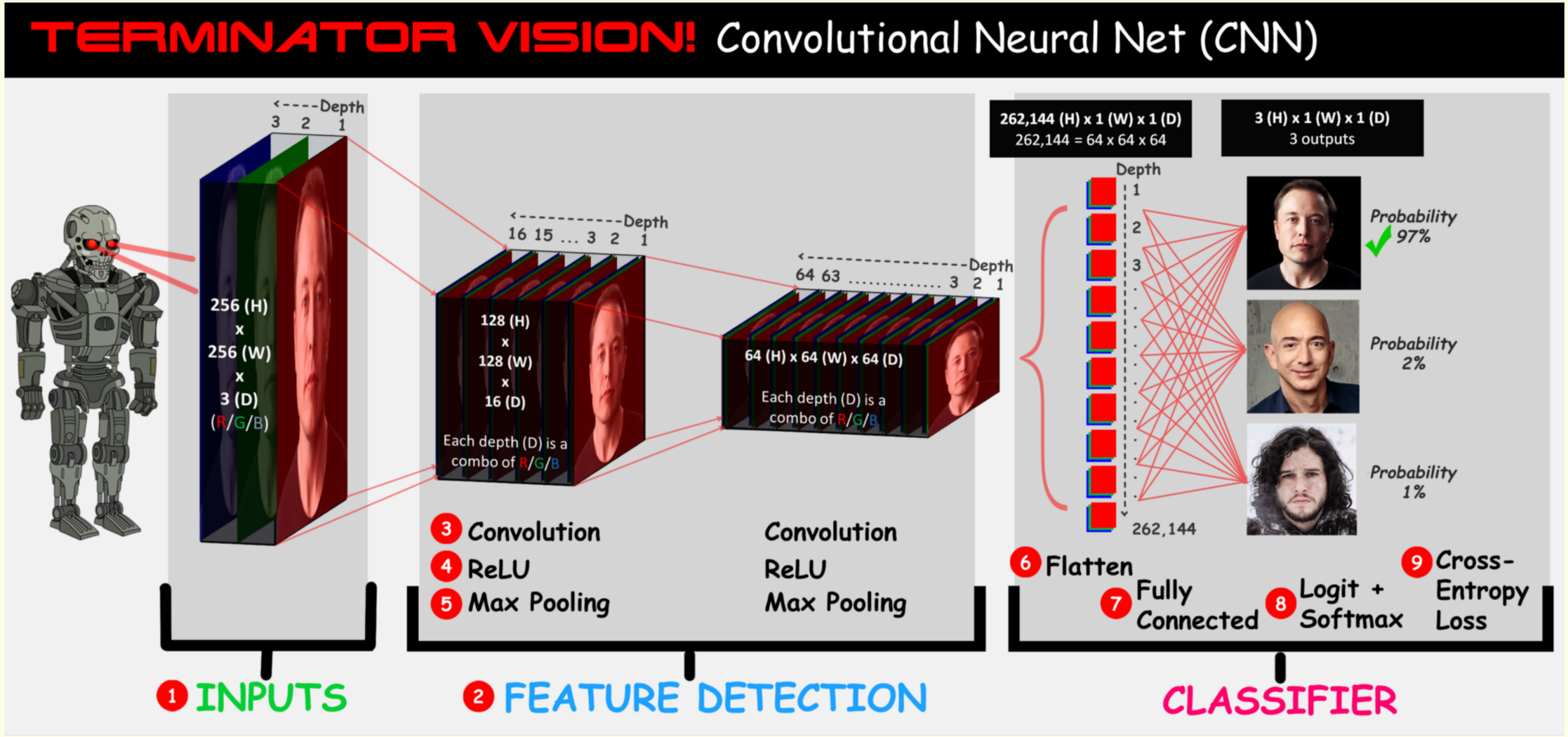
= Cars

WHAT IS CNN ?

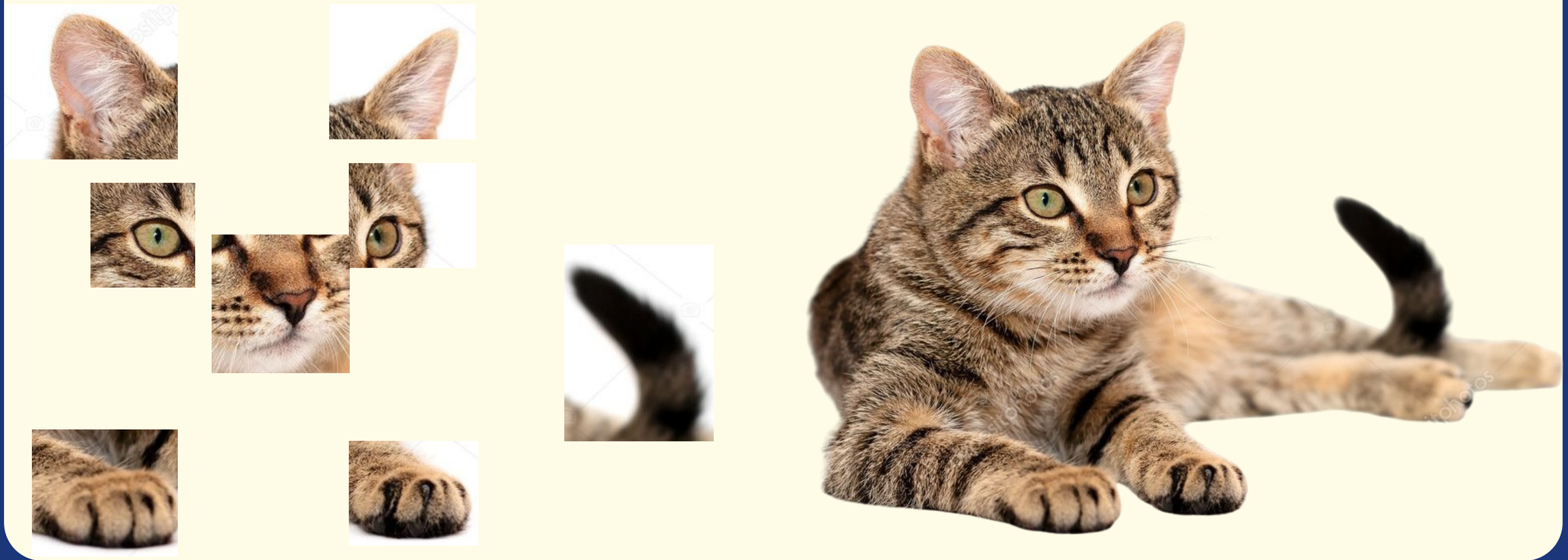
Convolutional Neural Network (CNN)

โครงข่ายประสาทเทียมหนึ่งในกลุ่ม bio-inspired โดยที่ CNN จะจำลองการมองเห็นของมนุษย์ที่มองพื้นที่เป็นที่ย่อย ๆ และนำกลุ่มของพื้นที่ย่อยๆมาผสานกันเพื่อดูว่าสิ่งที่เห็นอยู่เป็นอะไรกันแน่

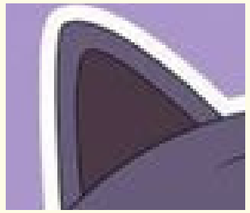


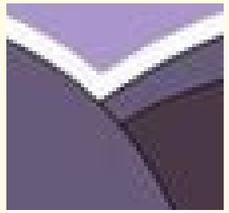

TERMINATOR VISION! Convolutional Neural Net (CNN)
Depth
3 2 1
256 (H)
x
256 (W)
x
3 (D)
(R/G/B)
Depth
16 15 ... 3 2 1
128 (H)
x
128 (W)
x
16 (D)
Each depth (D) is a combo of R/G/B
Depth
64 63 ..... 3 2 1
64 (H) x 64 (W) x 64 (D)
Each depth (D) is a combo of R/G/B
262,144 (H) x 1 (W) x 1 (D)
262,144 = 64 x 64 x 64
3 (H) x 1 (W) x 1 (D)
3 outputs
Depth
1
2
3
262,144
Probability 97%
Probability 2%
Probability 1%
3 Convolution
4 ReLU
5 Max Pooling
Convolution
ReLU
Max Pooling
6 Flatten
7 Fully Connected
8 Logit + Softmax
9 Cross-Entropy Loss
1 INPUTS
2 FEATURE DETECTION
CLASSIFIER

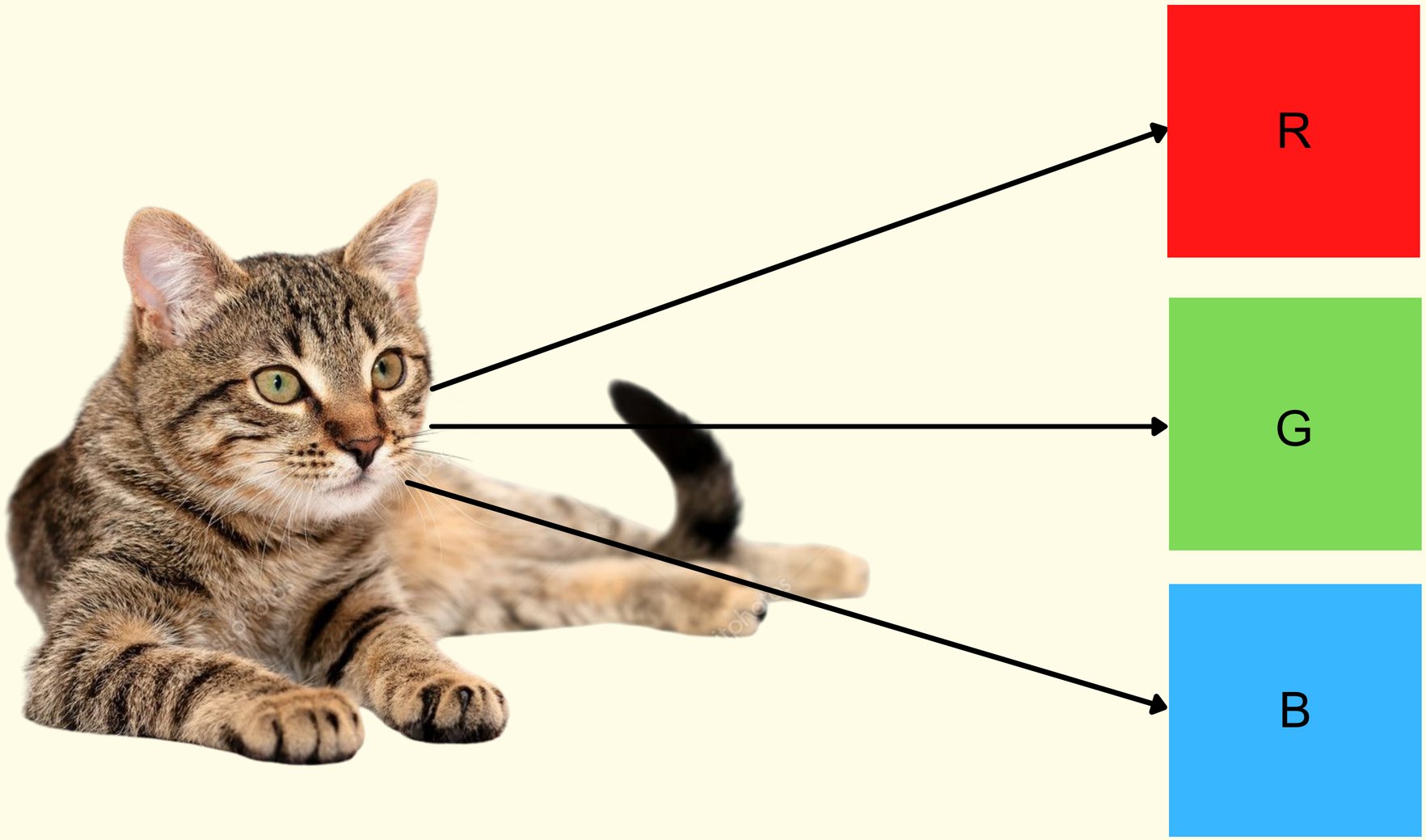
CNN



CNN



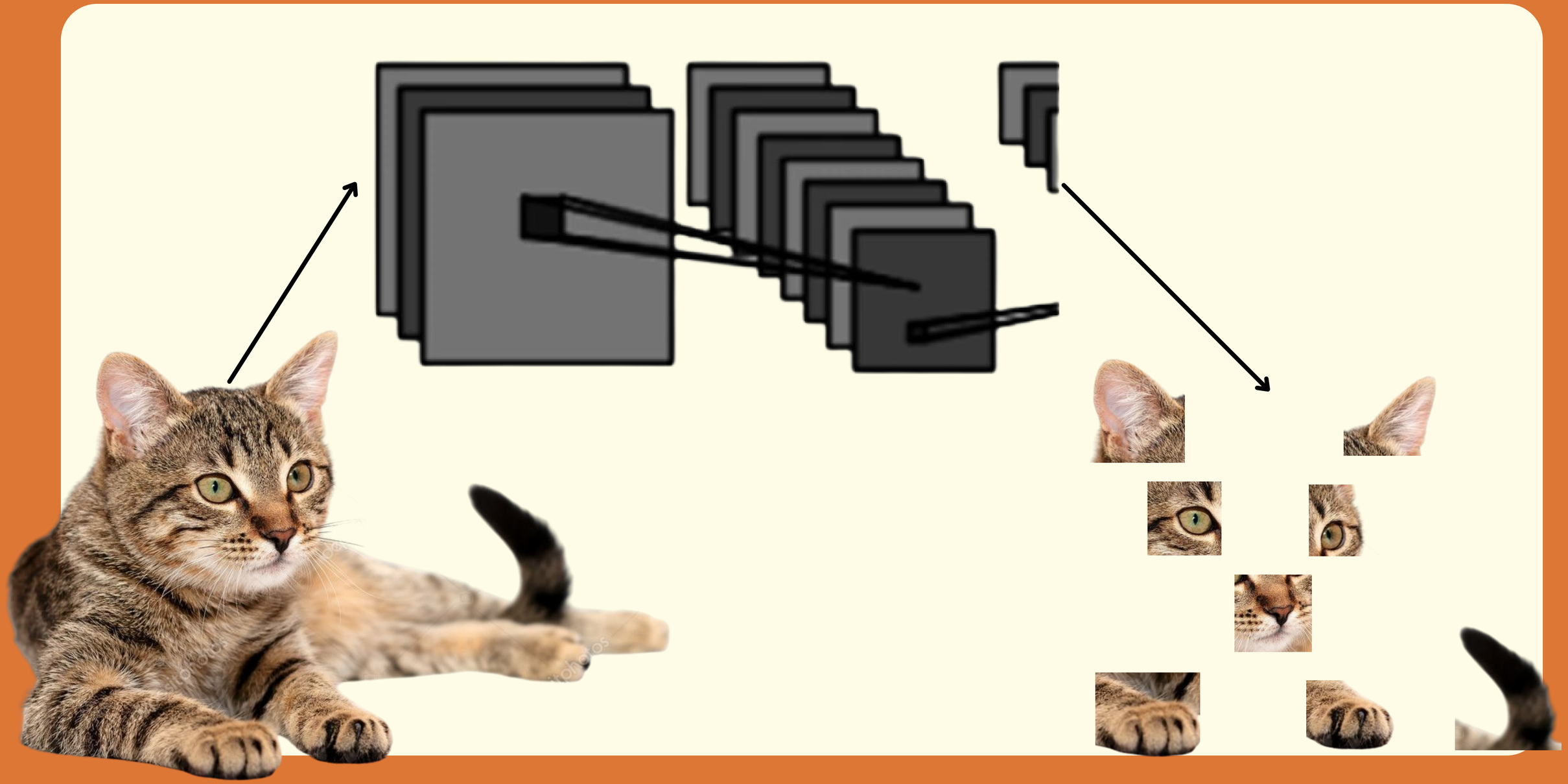


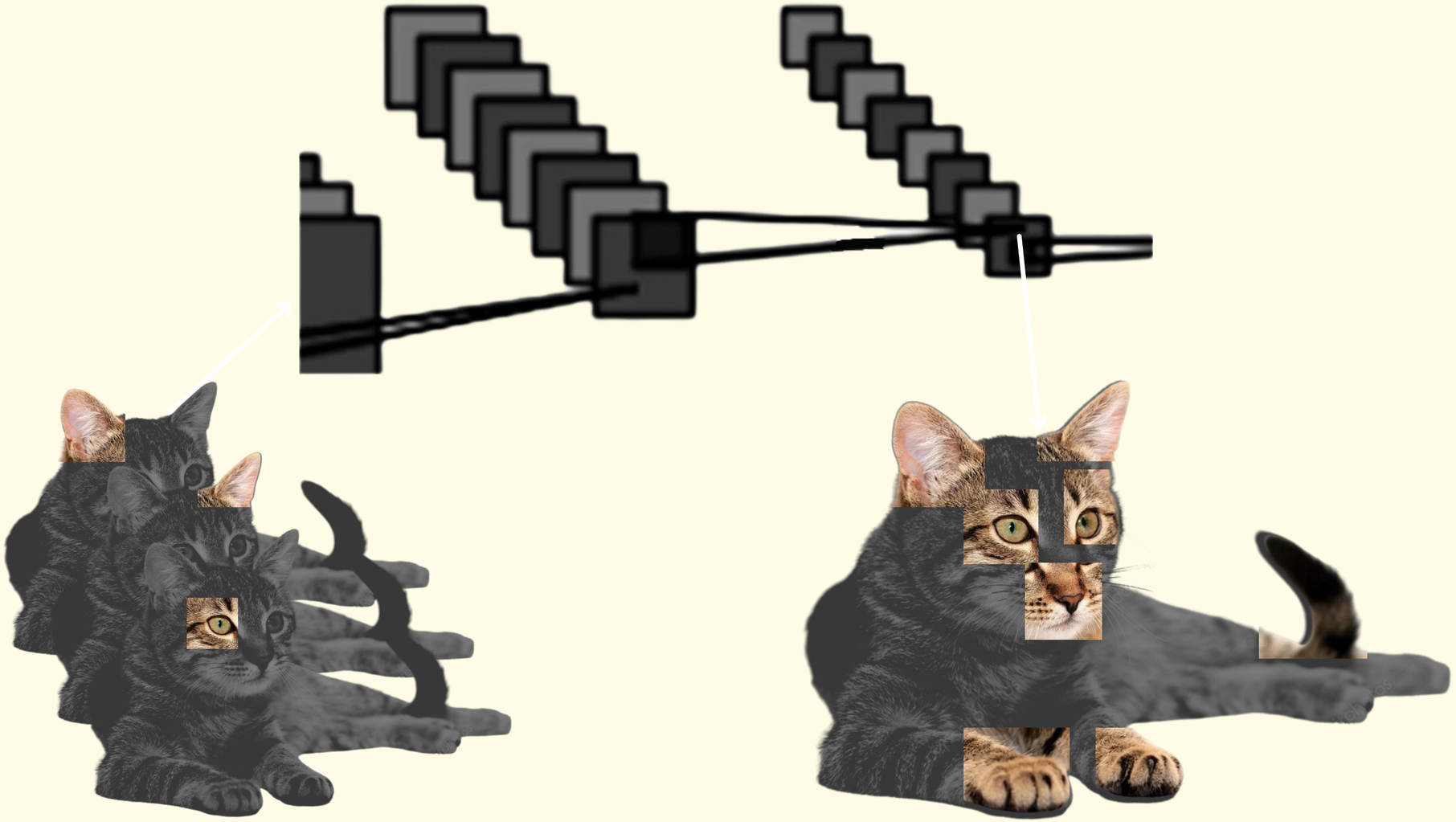

R
G
B

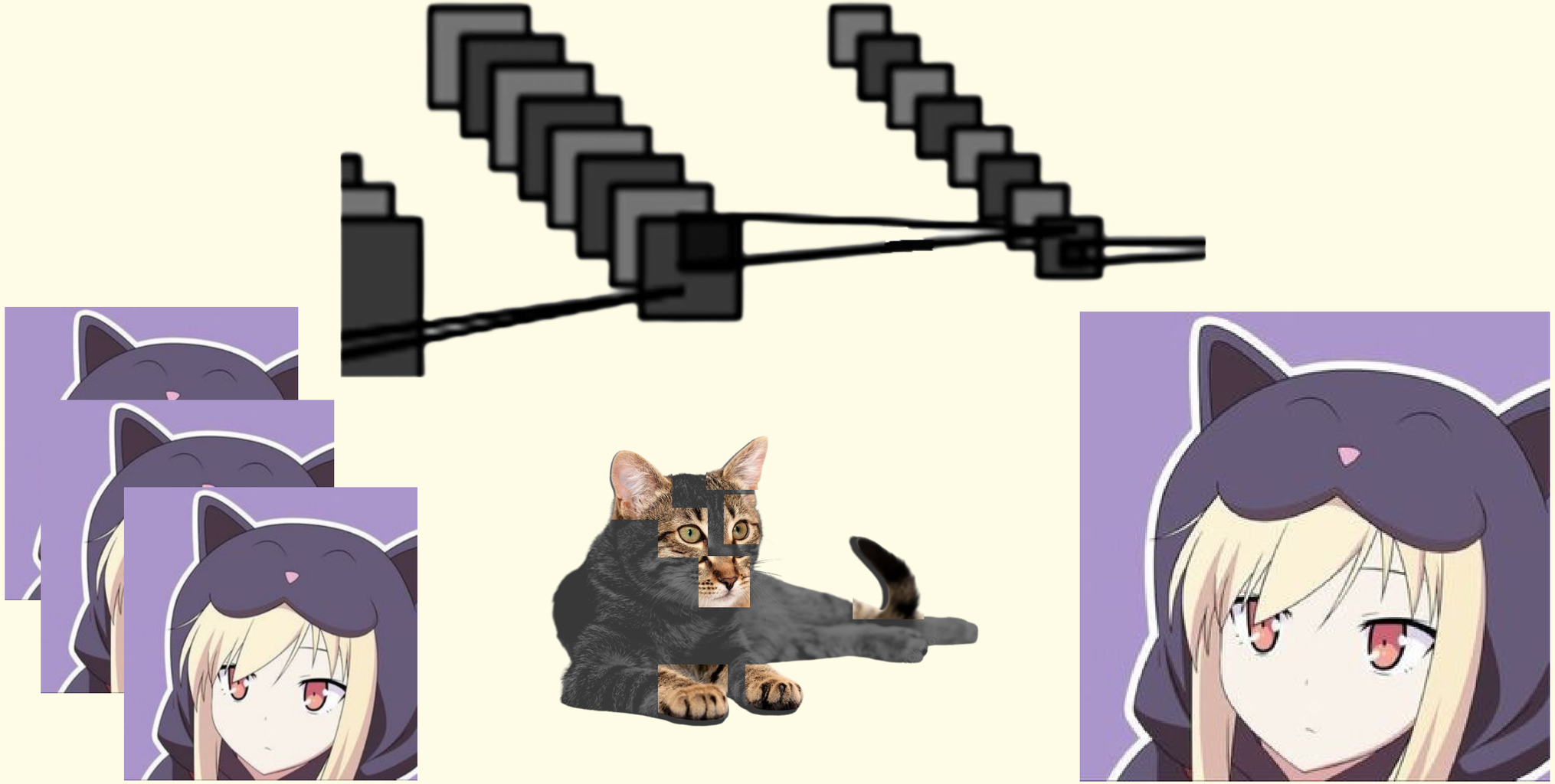
R

G

B







Cat = 25%

Human = 75%

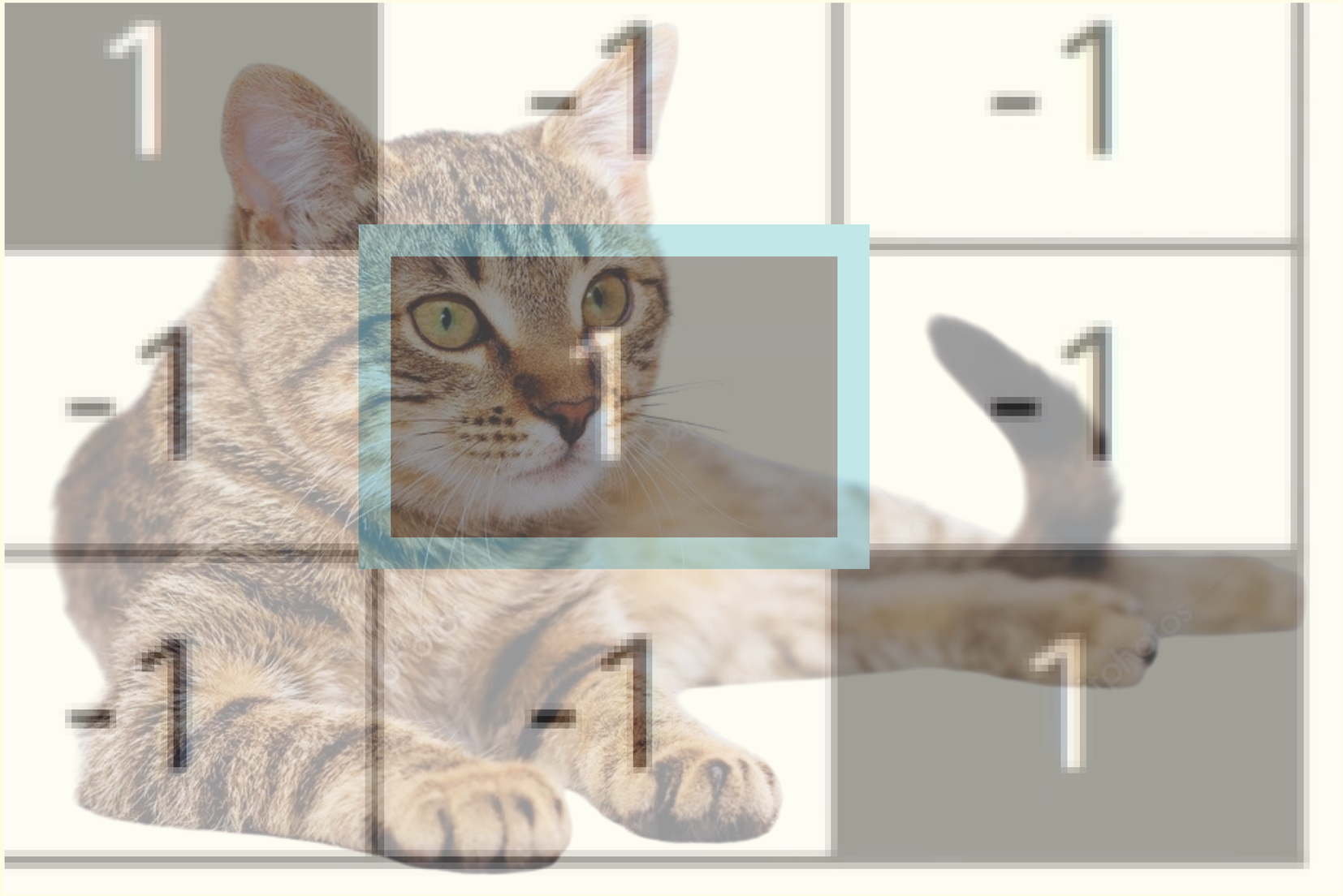
Cat = 100%

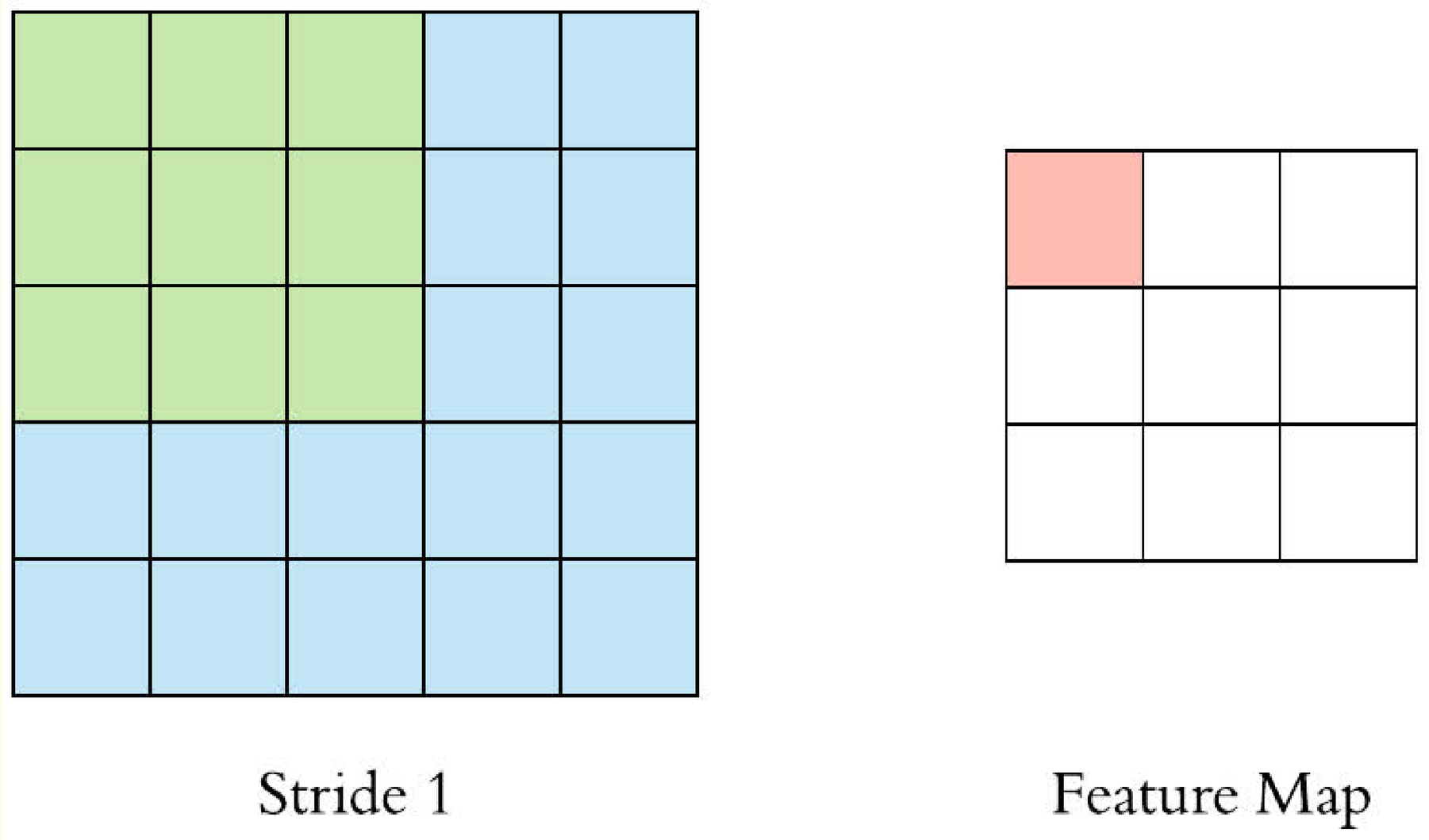
Human = 0%

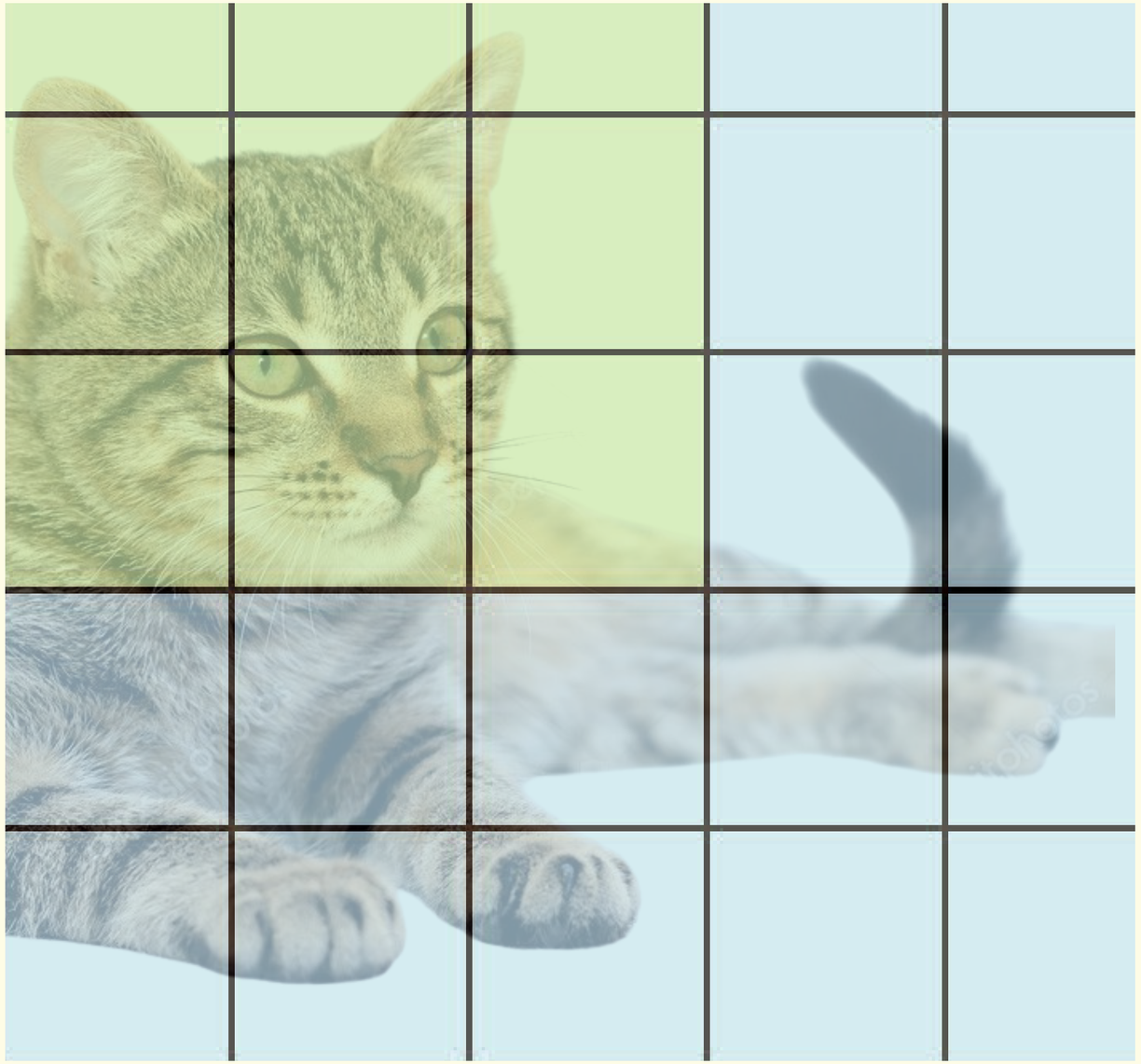
FEATURE EXTRACTION

Filter

1	-1	-1
-1	1	-1
-1	-1	1

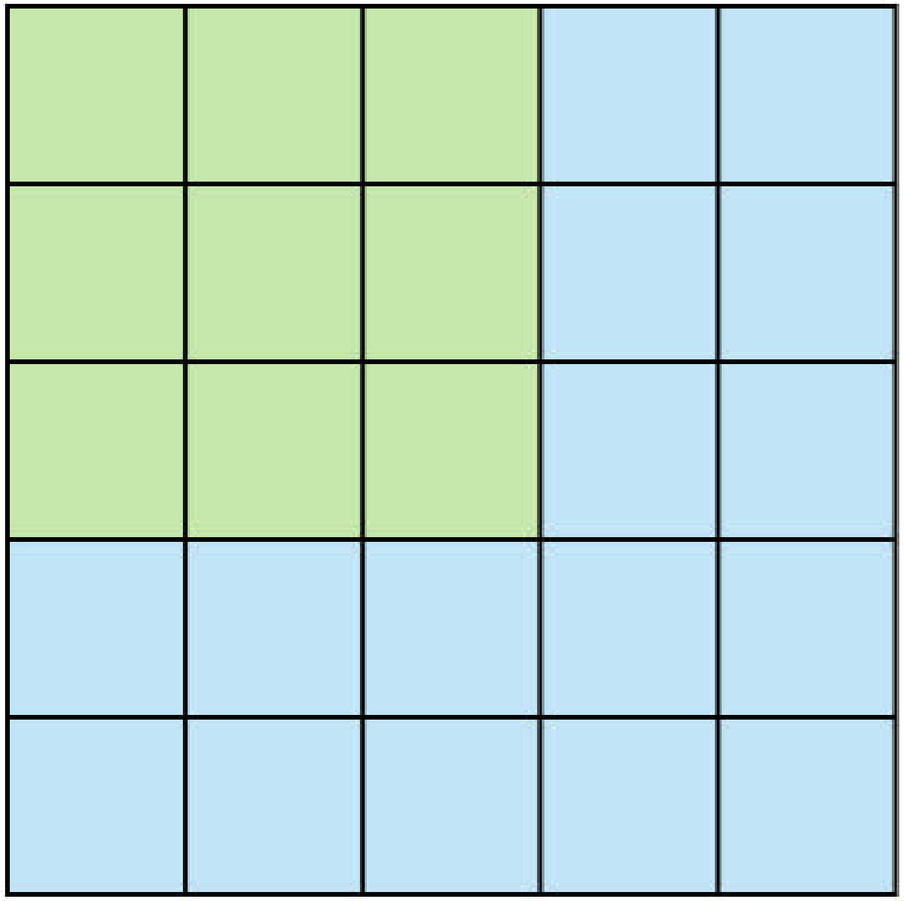

1
-1
-1
-1
1
-1
-1
-1
1


Stride 1
Feature Map

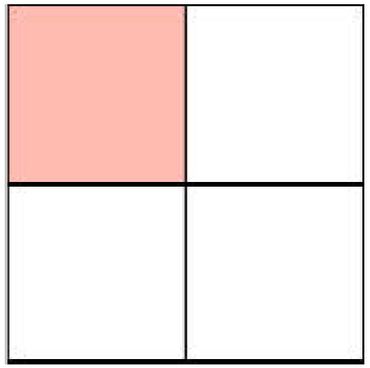


1x1	1x0	1x1	0	0
0x0	1x1	1x0	1	0
0x1	0x0	1x1	1	1
0	0	1	1	0
0	1	1	0	0

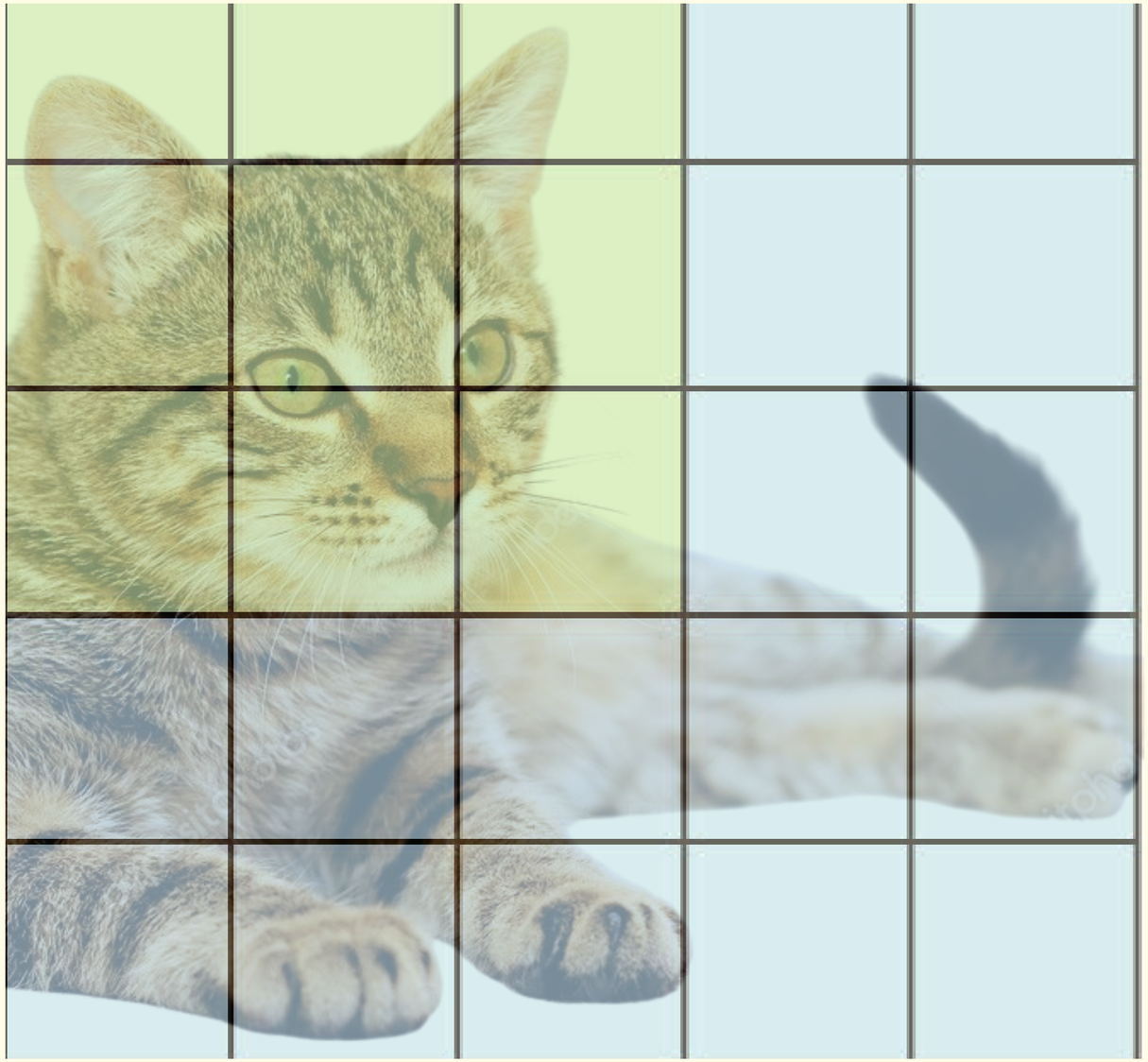
4		

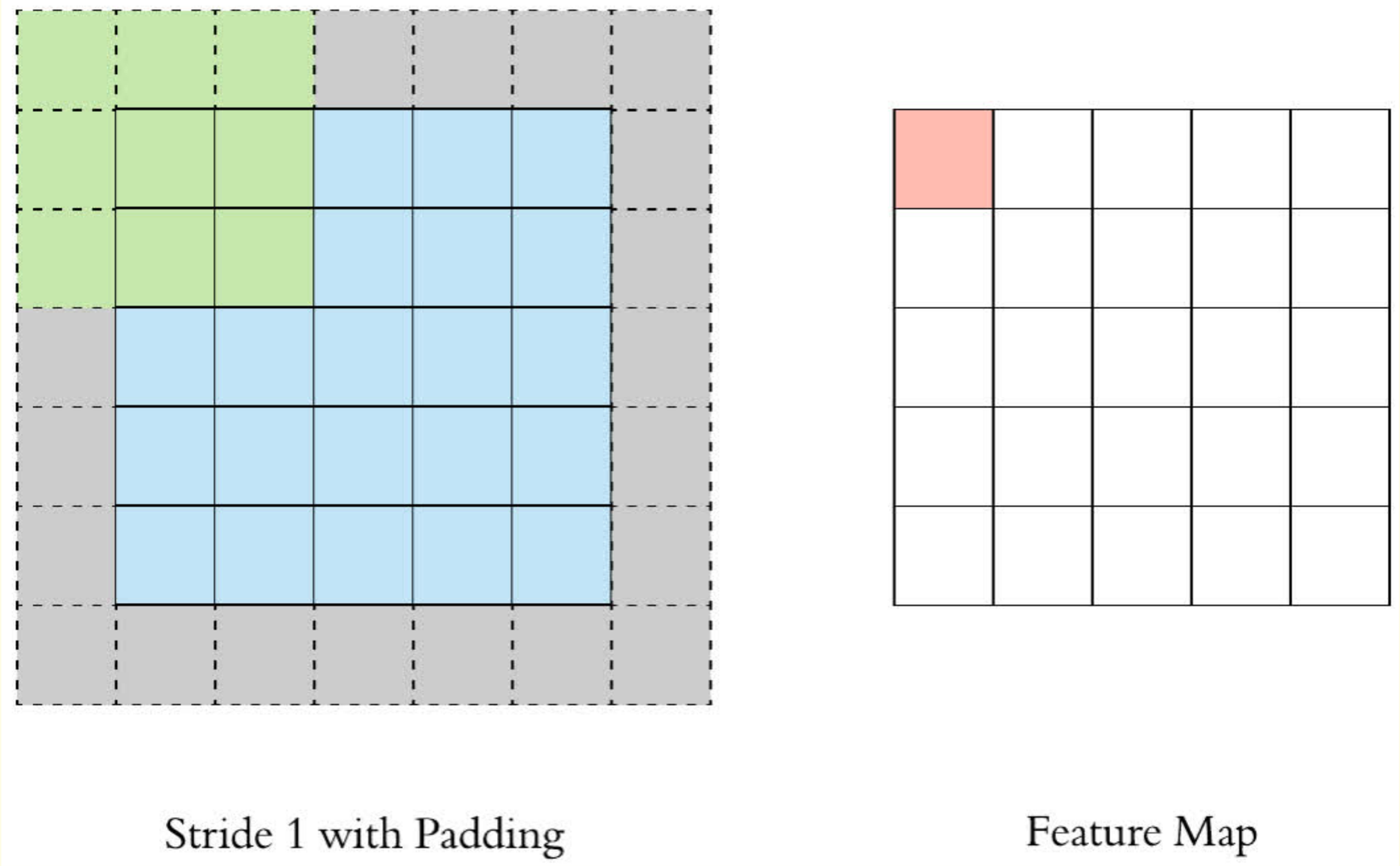


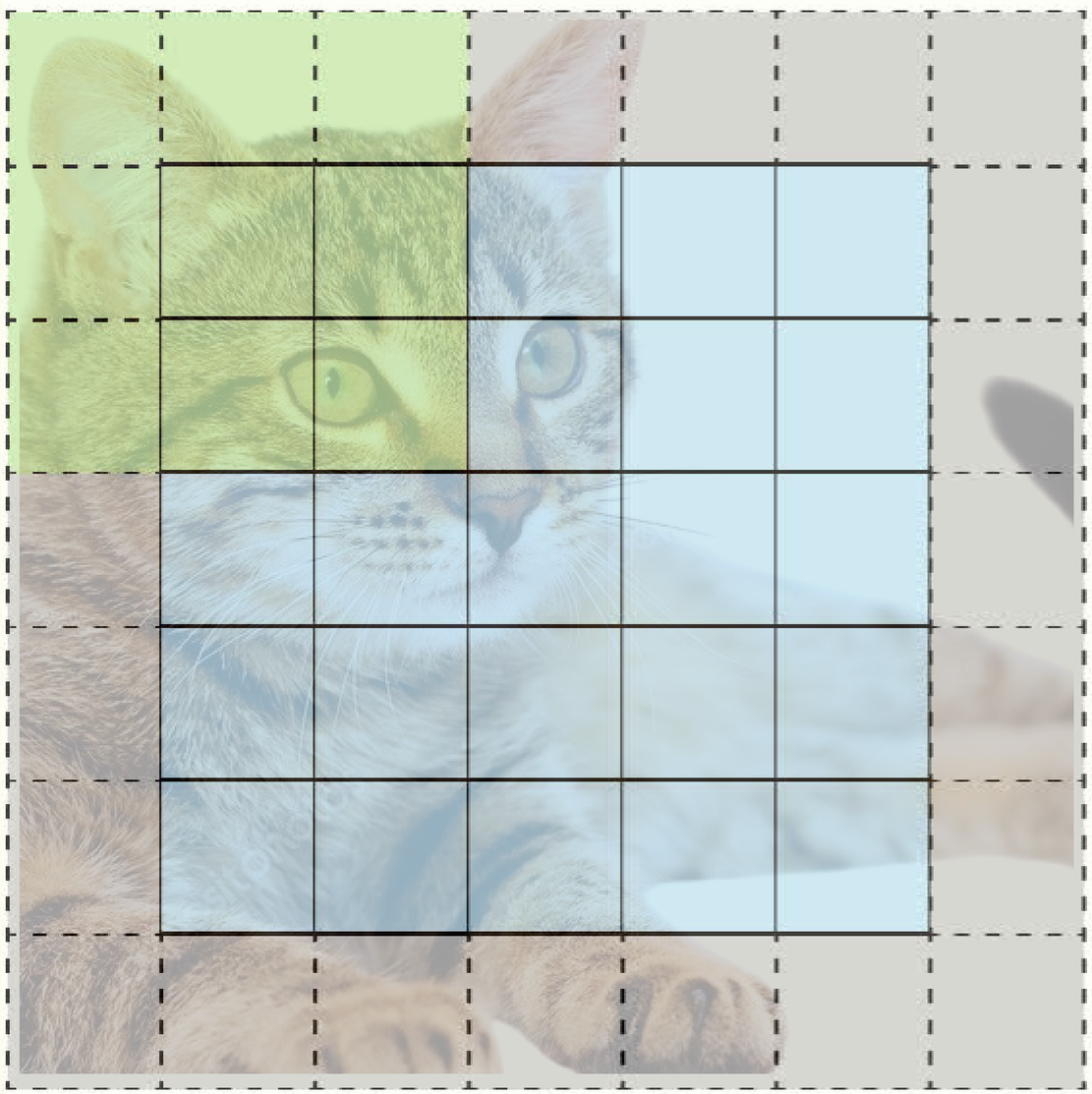
Stride 2



Feature Map

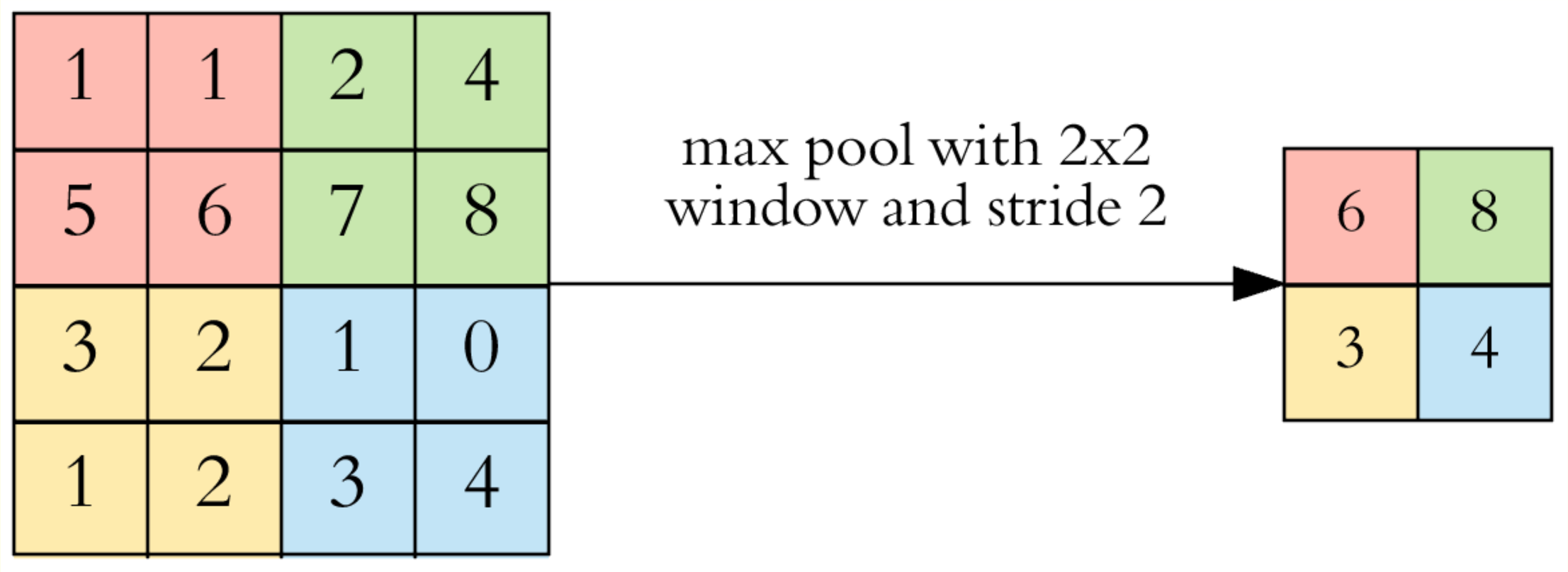


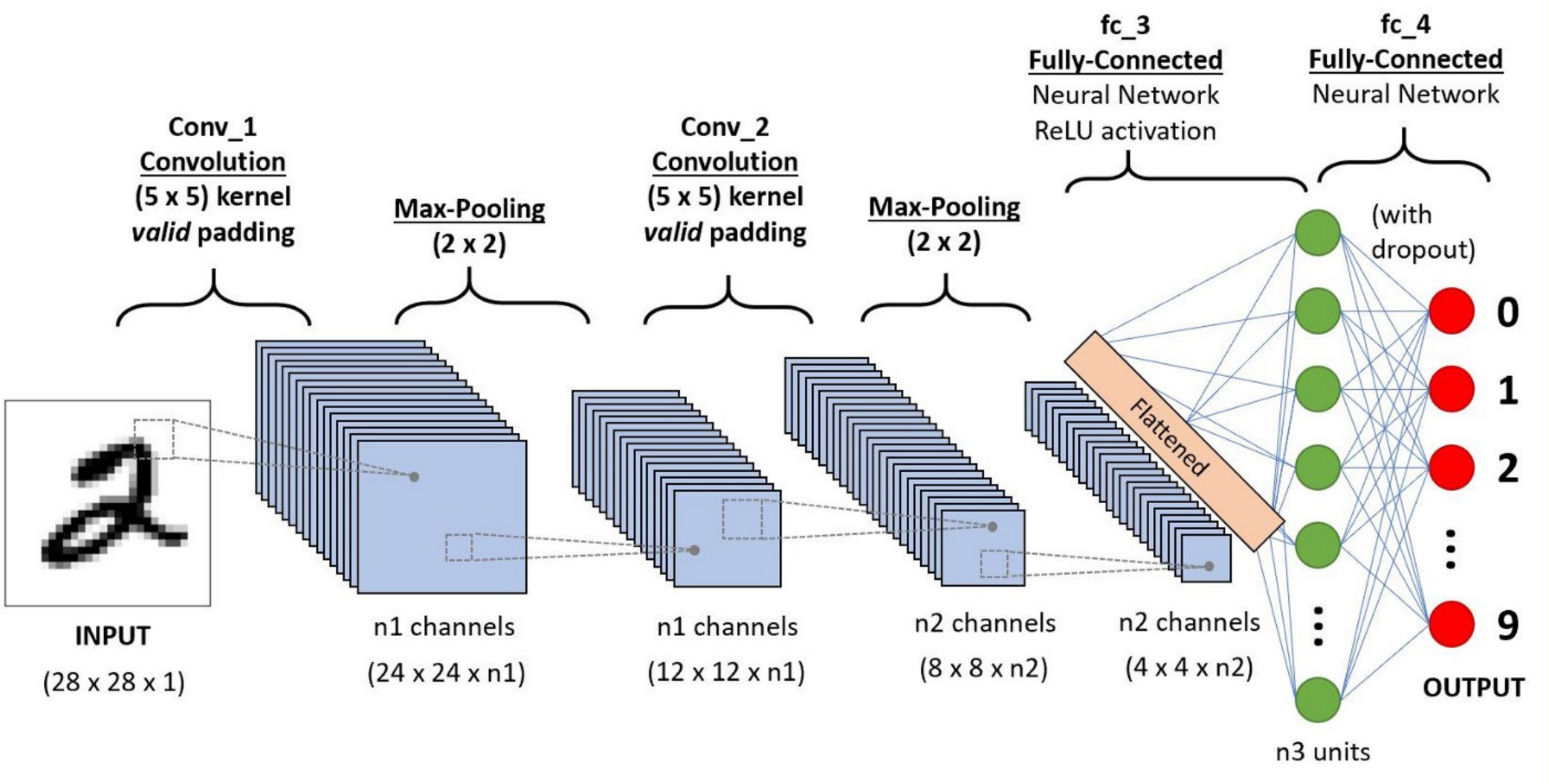

Stride 1 with Padding
Feature Map



Max Pooling

Max Pooling ✨


1 1 2 4
5 6 7 8
3 2 1 0
1 2 3 4
max pool with 2x2
window and stride 2
6 8
3 4


Conv_1
Convolution
(5 x 5) kernel
valid padding
Max-Pooling
(2 x 2)
Conv_2
Convolution
(5 x 5) kernel
valid padding
Max-Pooling
(2 x 2)
fc_3
Fully-Connected
Neural Network
ReLU activation
fc_4
Fully-Connected
Neural Network
(with dropout)
Flattened
INPUT
(28 x 28 x 1)
n1 channels
(24 x 24 x n1)
n1 channels
(12 x 12 x n1)
n2 channels
(8 x 8 x n2)
n2 channels
(4 x 4 x n2)
n3 units
0
1
2
9
OUTPUT

WORKSHOP 2:

IMAGE RECOGNITION WITH PYTHON

Computer Education

LECTURER

KETT KONGNAKORN

4th year Student

LECTURER

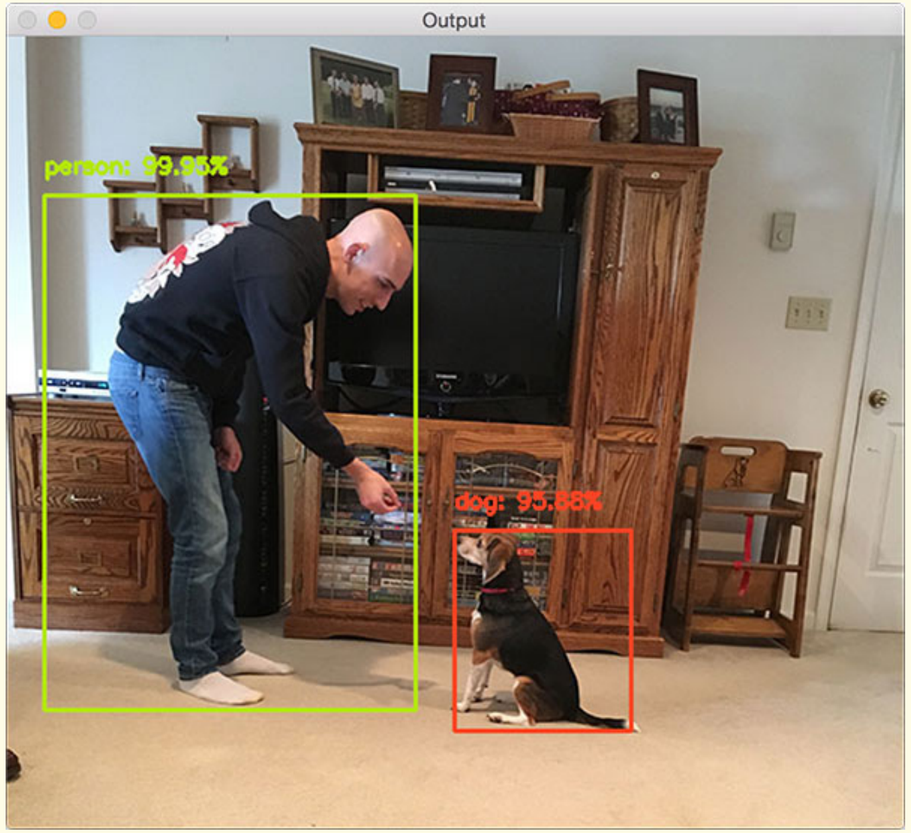
PIYADET KITPHADUNG

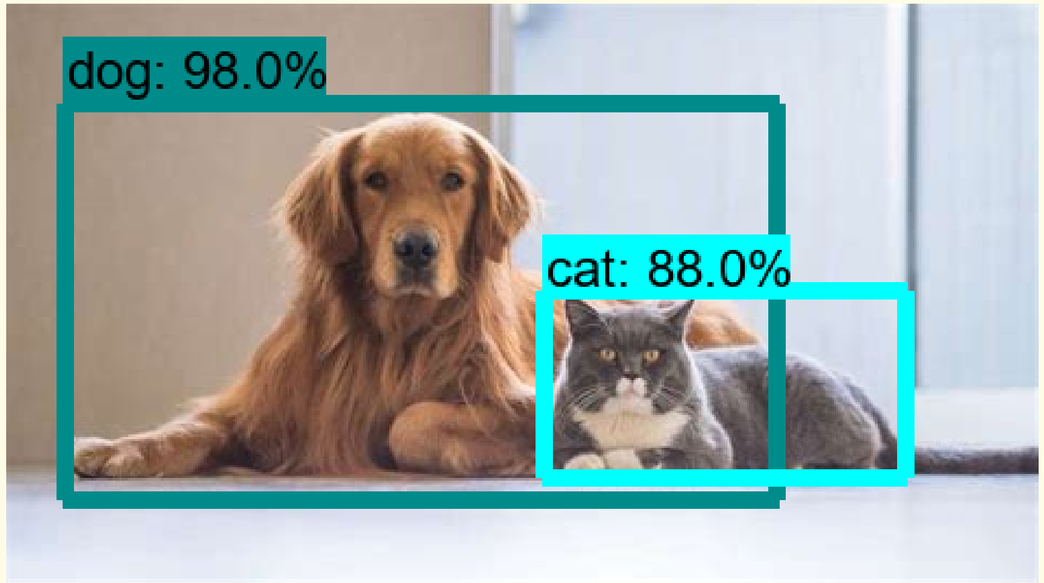
4th year Student

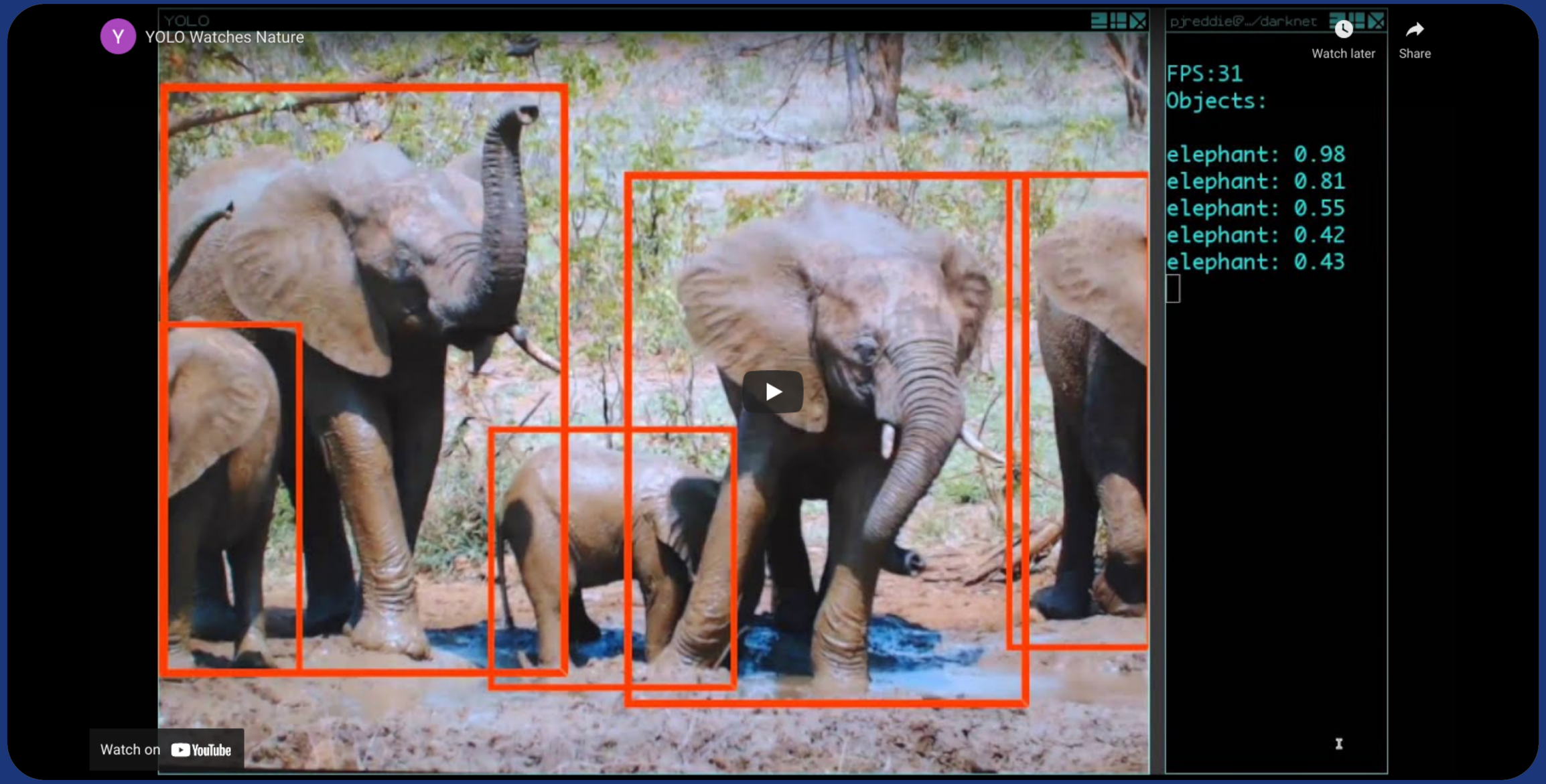
Object Detection

Object Classification VS Object Detection

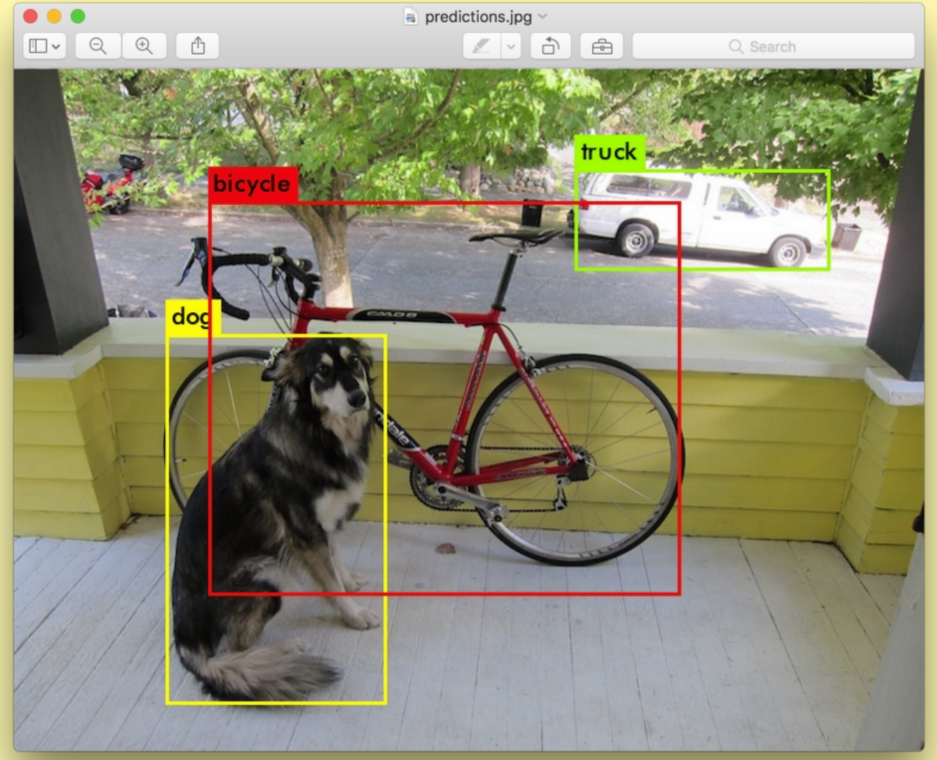
What is YOLO ?


Output
person: 99.95%
dog: 95.88%

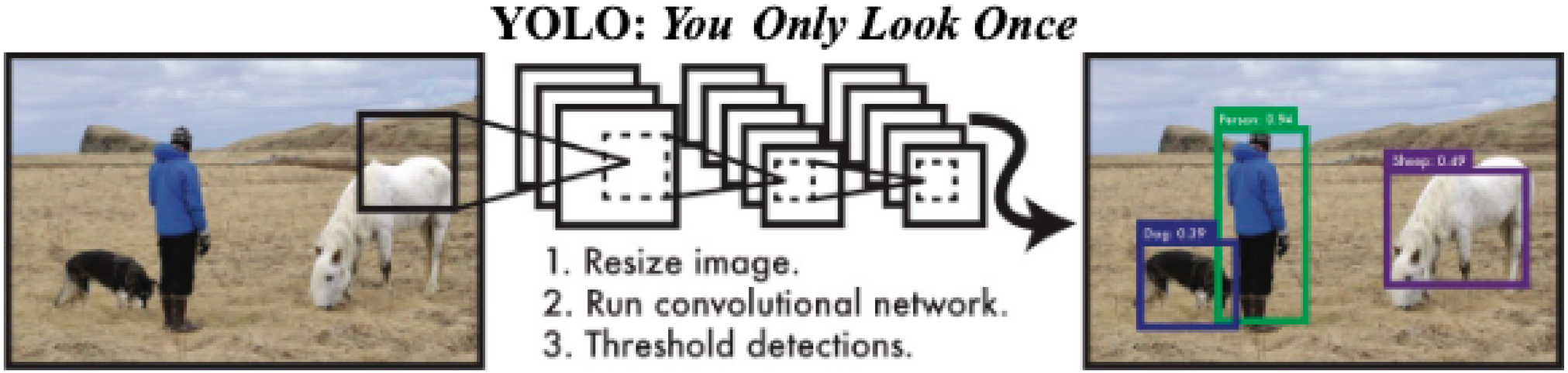

dog: 98.0%
cat: 88.0%


YOLO Watches Nature
Watch later
Share
YOLO
pjreddie@.../darknet
FPS:31
Objects:
elephant: 0.98
elephant: 0.81
elephant: 0.55
elephant: 0.42
elephant: 0.43
Watch on YouTube

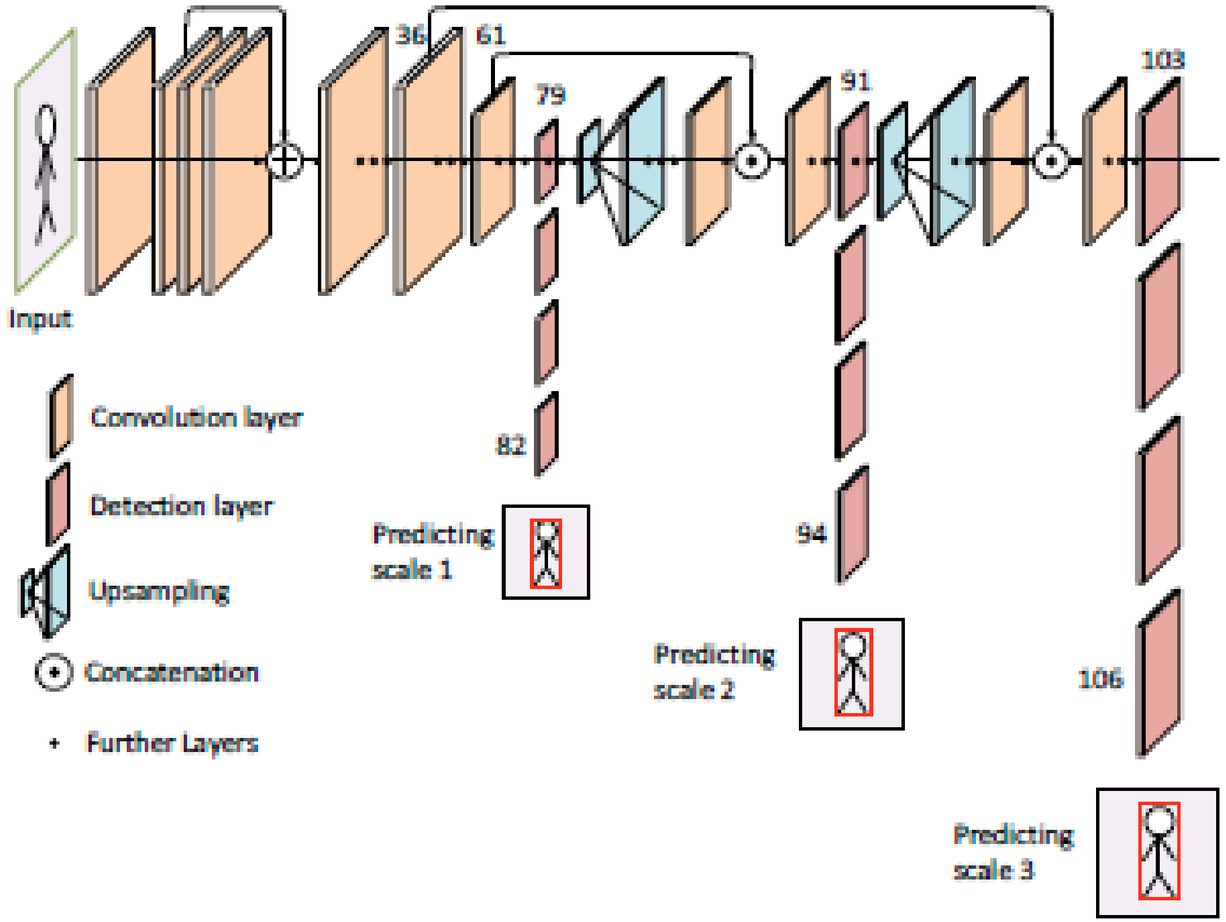
YOLO เป็นการรู้จำวัตถุที่ใช้วิธีการที่เรียกว่า “Fast single-shot detection” ซึ่งจะเป็นวิธีการที่สามารถตรวจจับวัตถุได้จากการส่งผ่านรูปภาพเข้าไปในระบบเพียงครั้งเดียว



With YOLO, you only look once at an image to perform detection

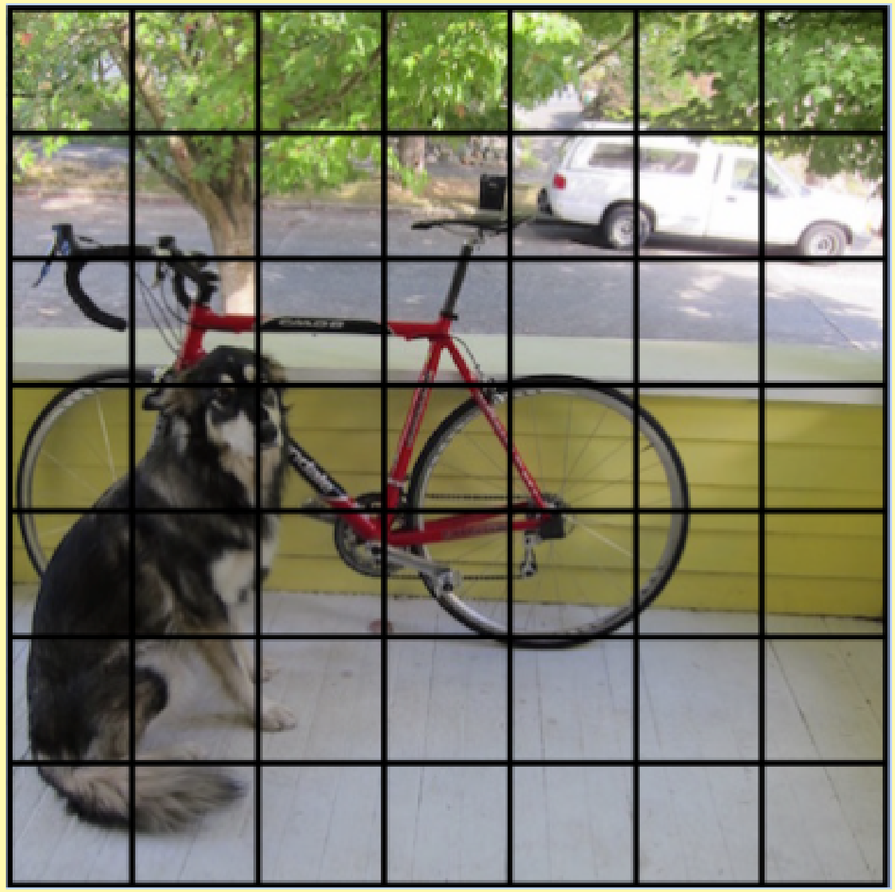


โครงสร้าง Yolov3



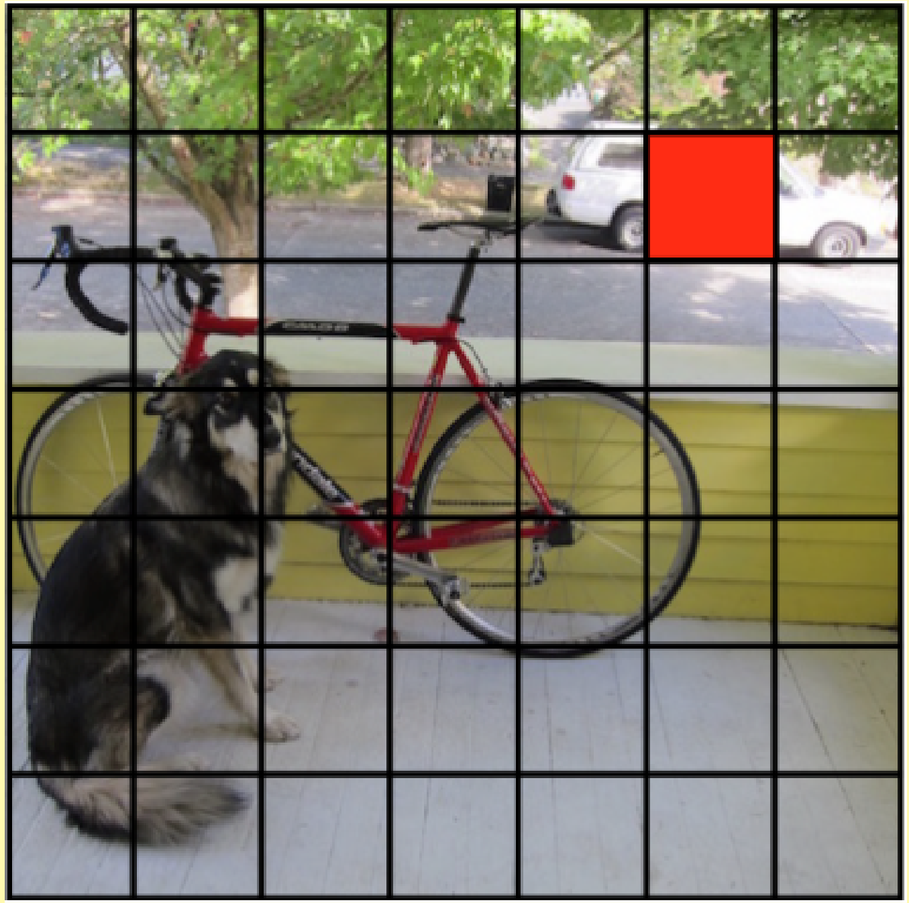
สักษณะการทำงานแบ่งเป็นชั้น ๆ
ชั้นของคอนโวลูชั่น
ชั้นของการเพิ่มอัตราสุ่ม
ชั้นของการทำนายผล

INPUT

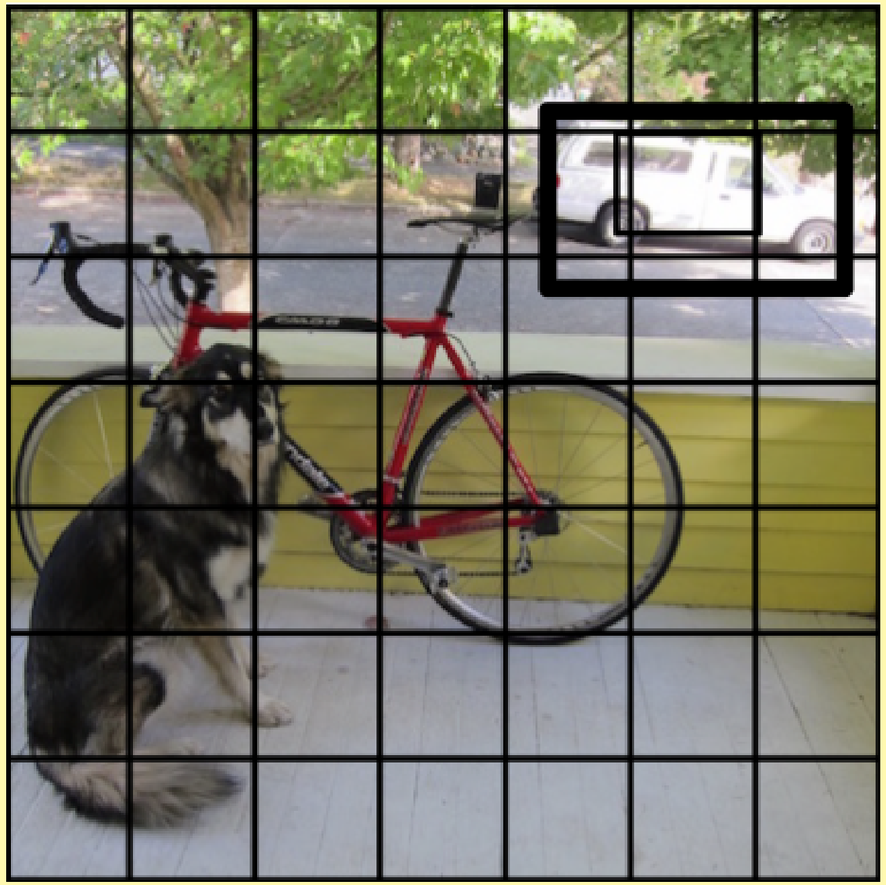


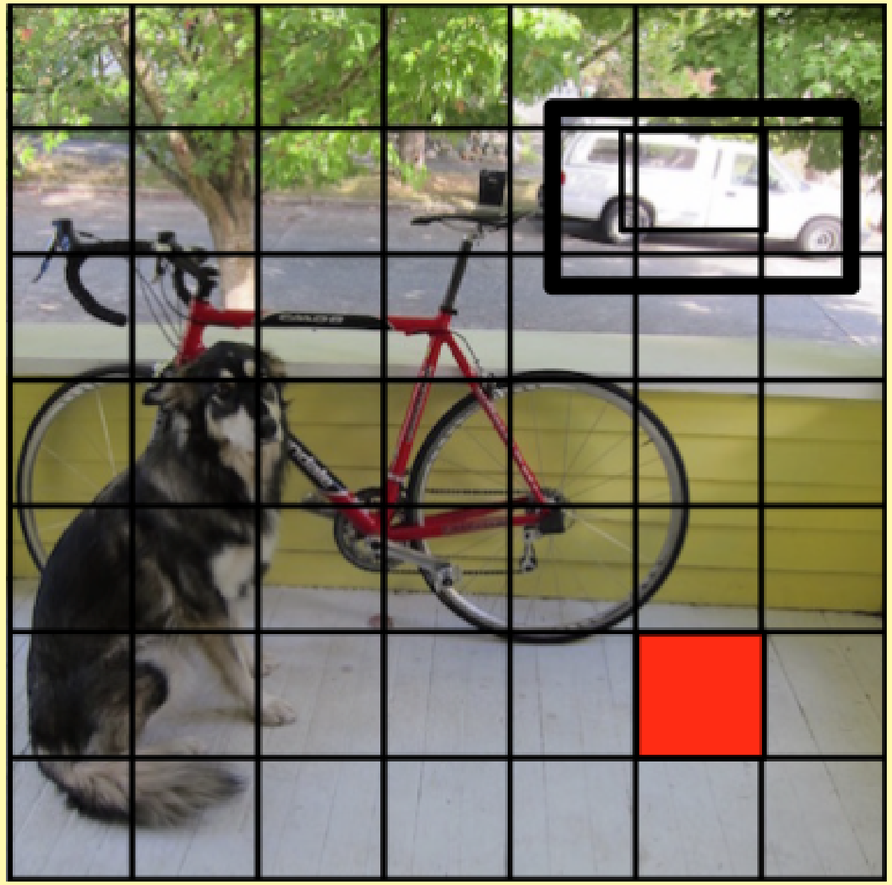
แบ่งภาพเป็นสี่เหลี่ยมตาราง

ในแต่ละตารางที่แบ่งออกมาจะทำการทำนายค่า **boxes** และค่า **confidences : P(Object)**

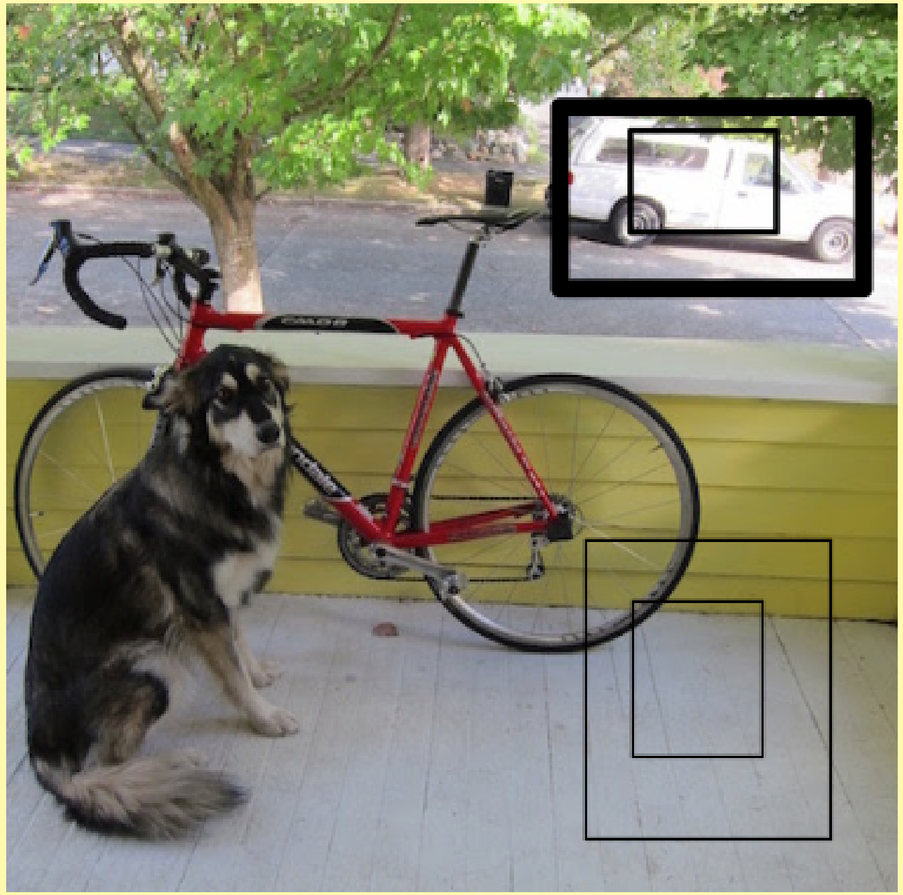


ในแต่ละตารางที่แบ่งออกมาจะทำการทำนายค่า **boxes** และค่า **confidences : P(Object)**

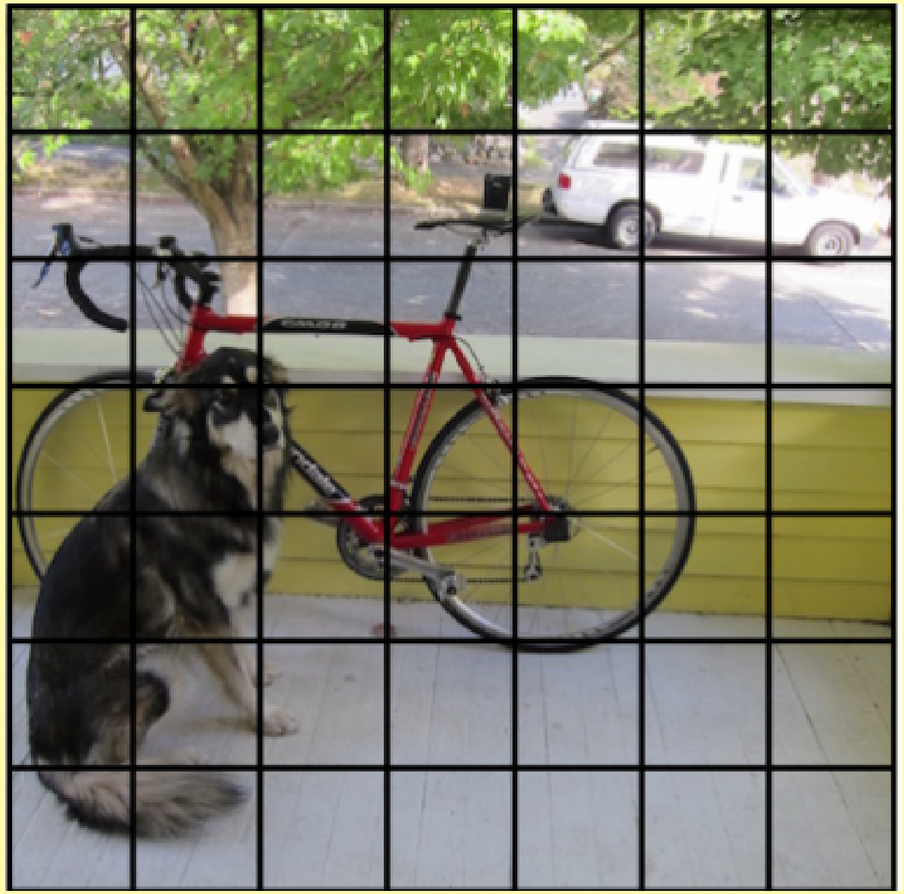


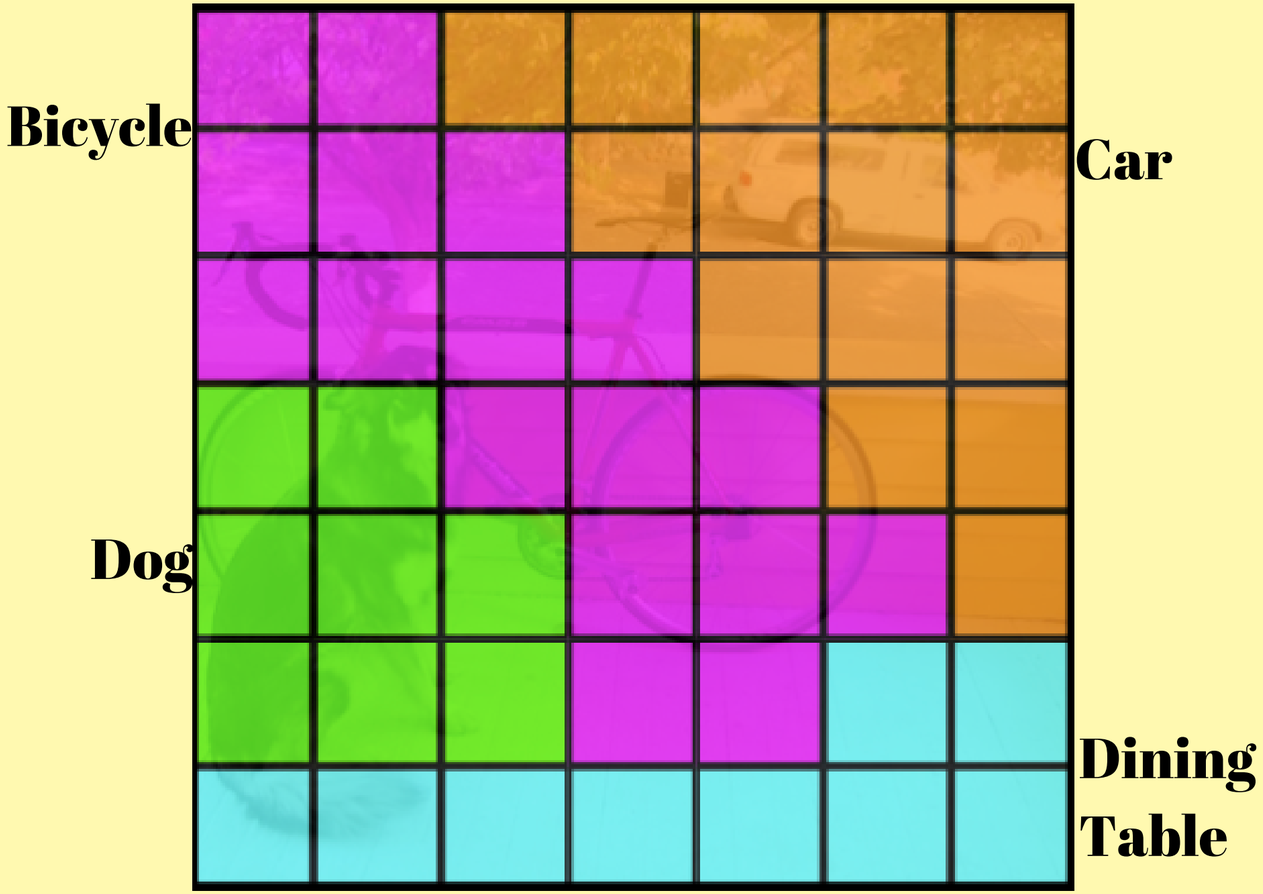


ในแต่ละตารางที่แบ่งออกมาจะทำการทำนายค่า **boxes** และค่า **confidences : P(Object)**



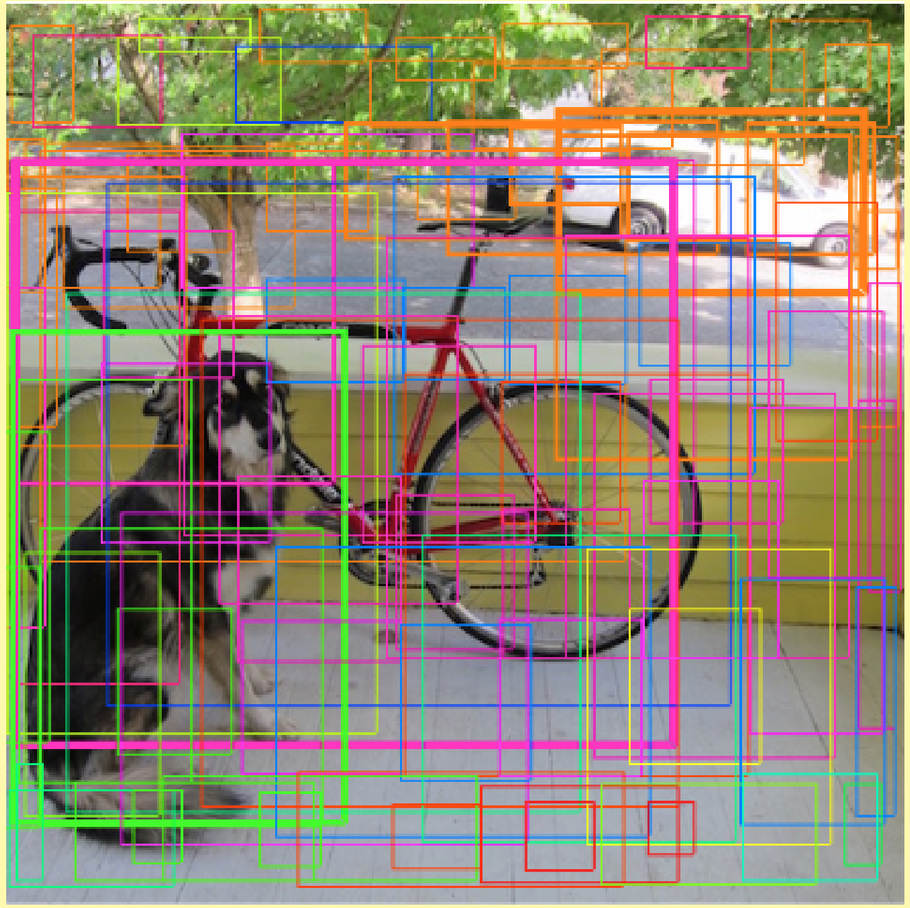
ทำนายความน่าจะเป็นของคลาสจากตารางทั้งหมด



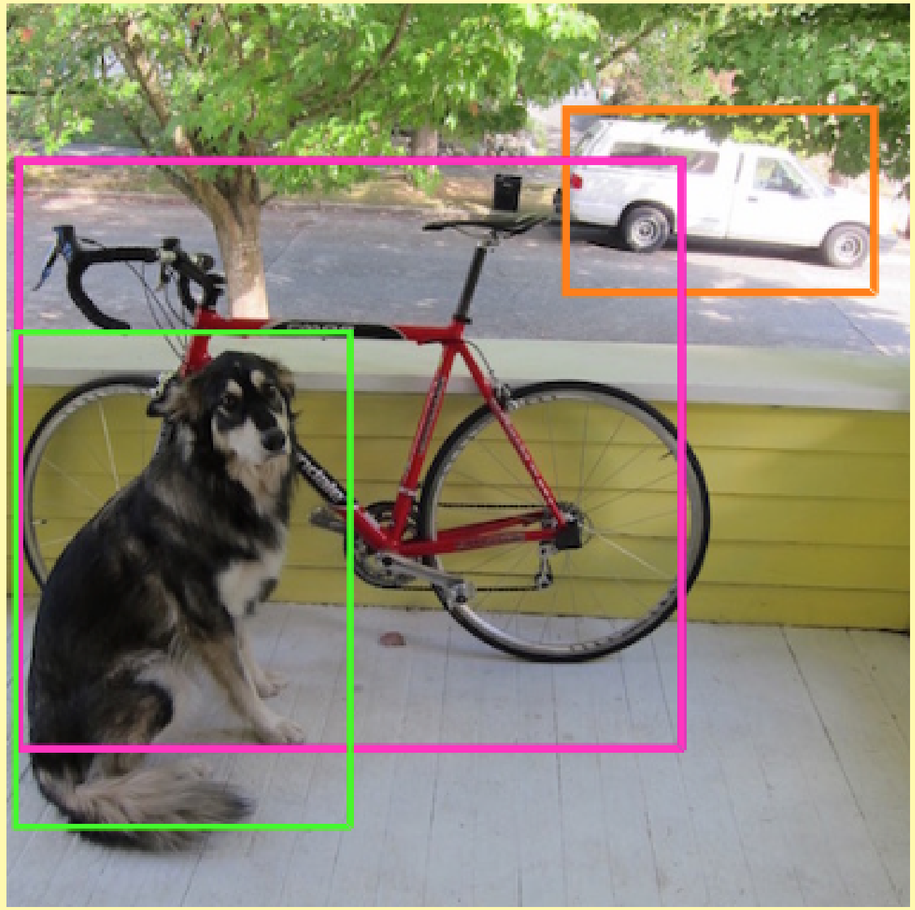


กำหนดเงื่อนไขบนวัตถุ :P(Car | Object)

รวมช่องตารางกับคลาสเข้าด้วยกัน

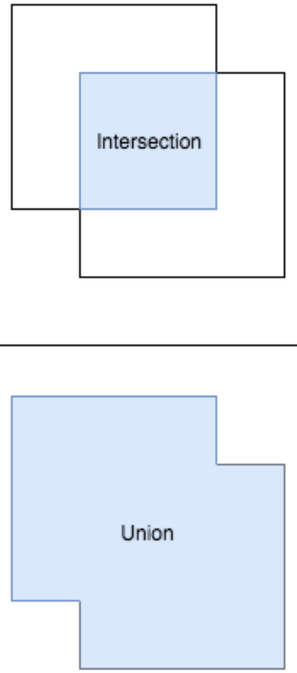


ทำการ **NMS and threshold detections**

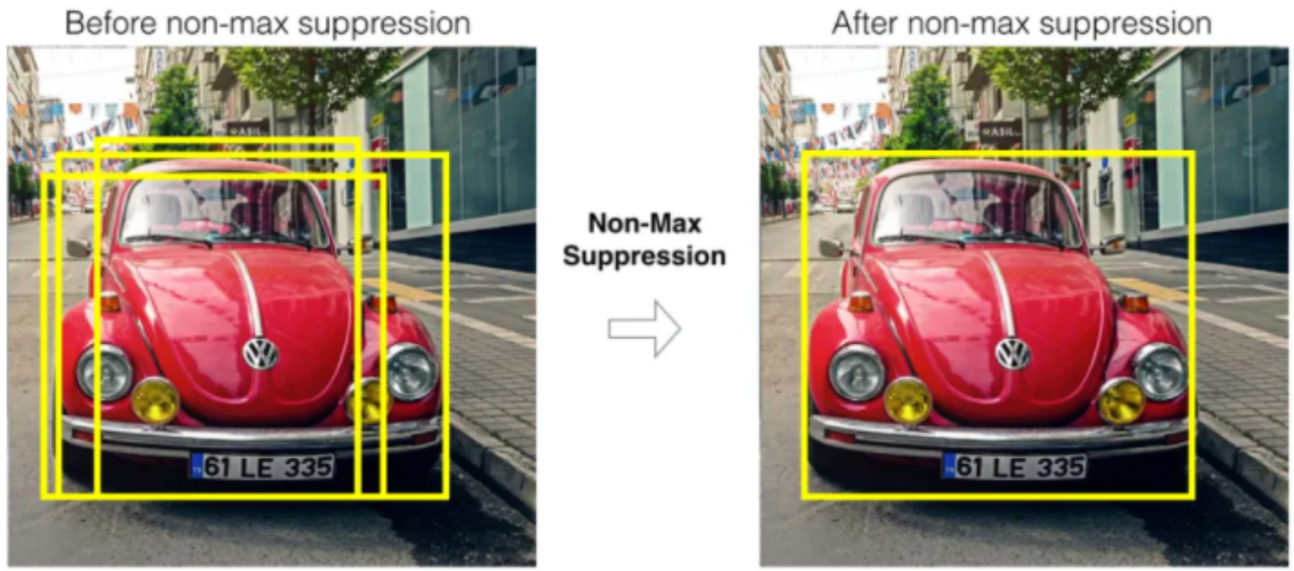


Non-maximum Suppression (NMS)

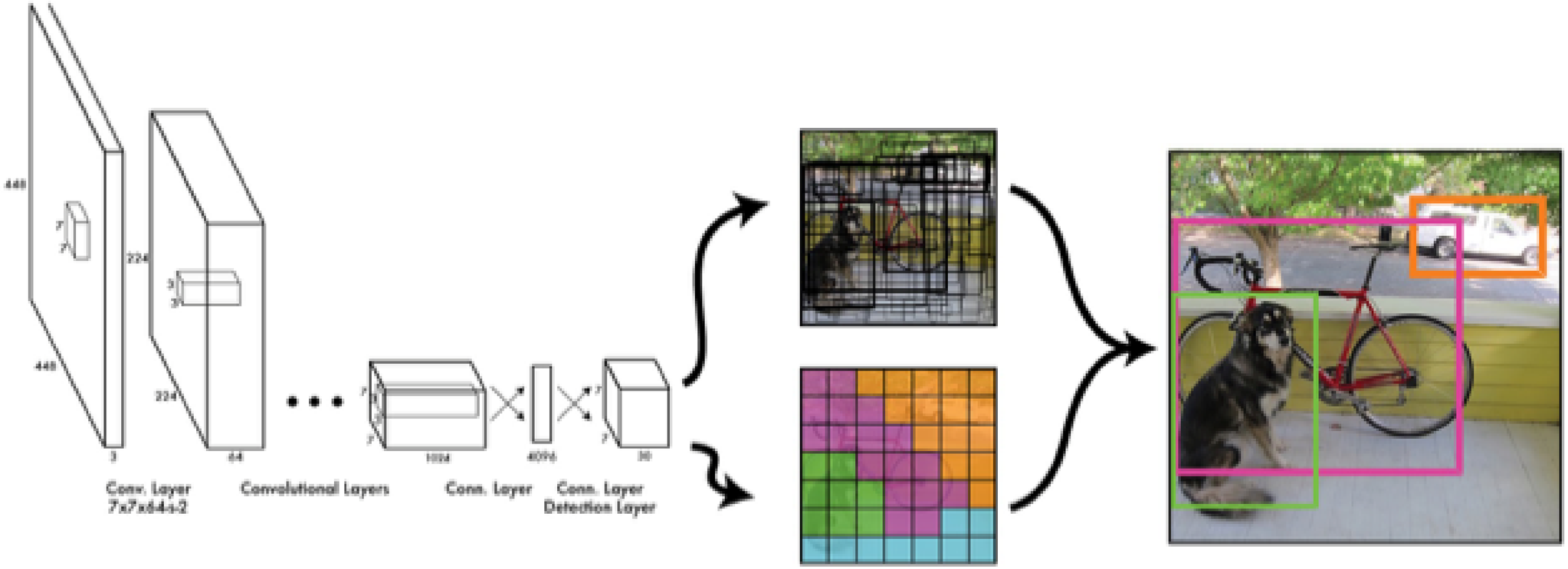
$$\text{IoU} = \frac{\text{Intersection}}{\text{Union}}$$



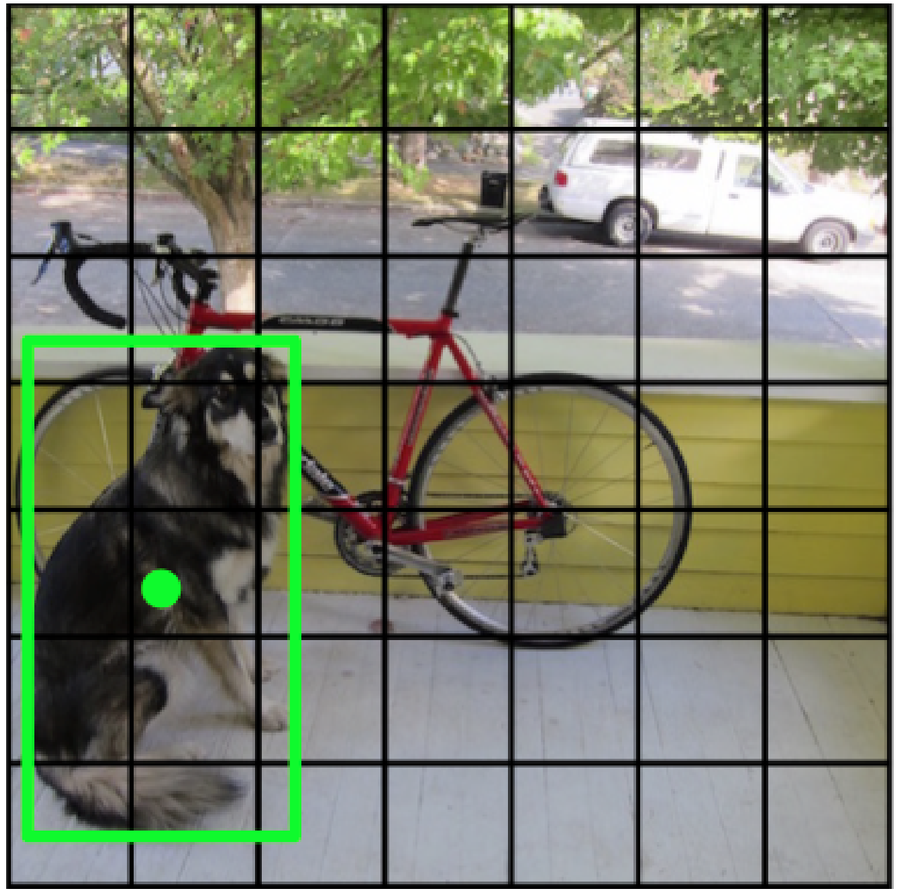
Intersection over Union

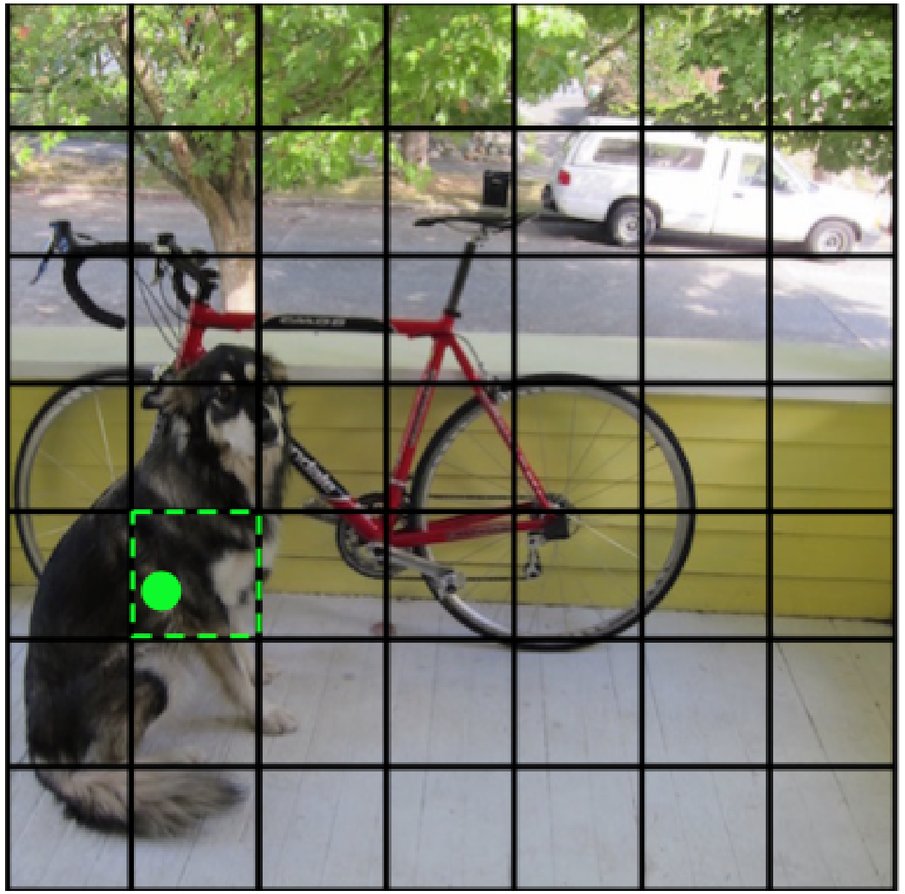


การรวมทั้งสองการตรวจจับวัตถุ



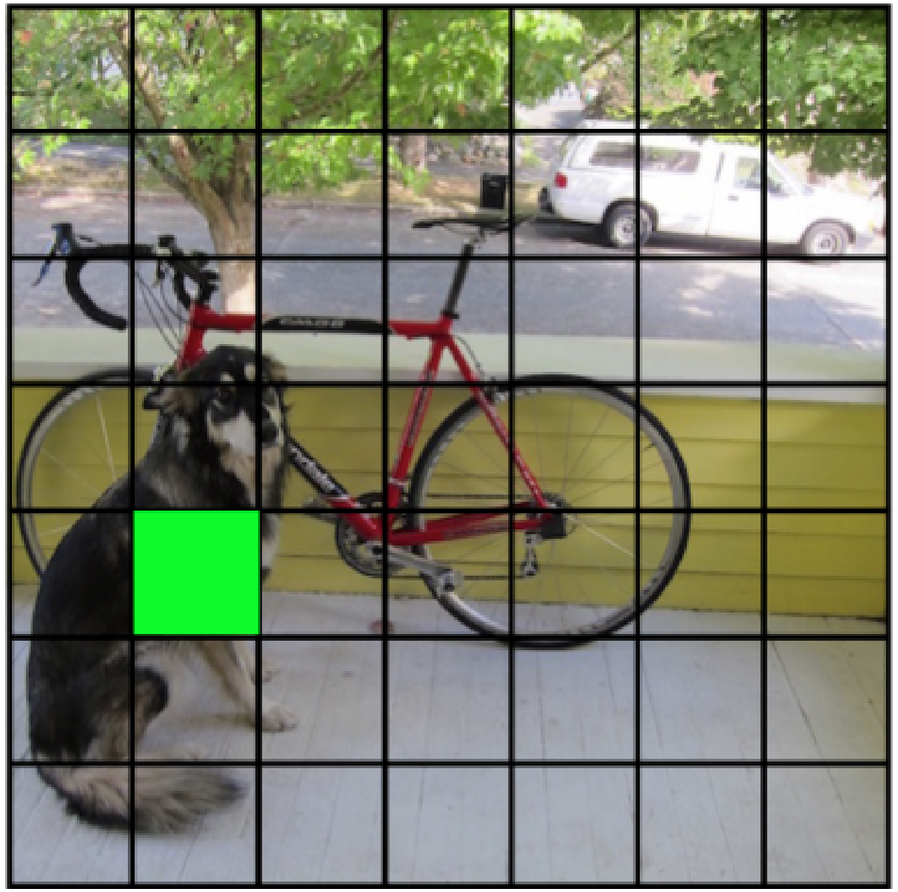
ในระหว่างการเทรนนั้นจะมีการจับคู่

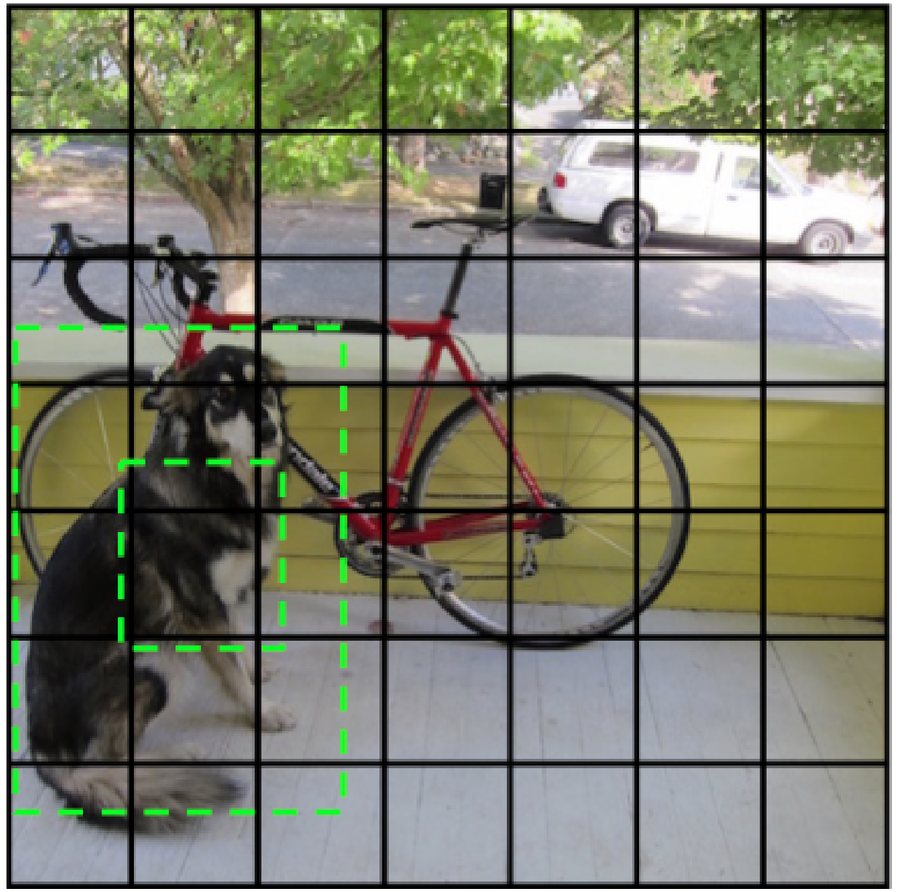




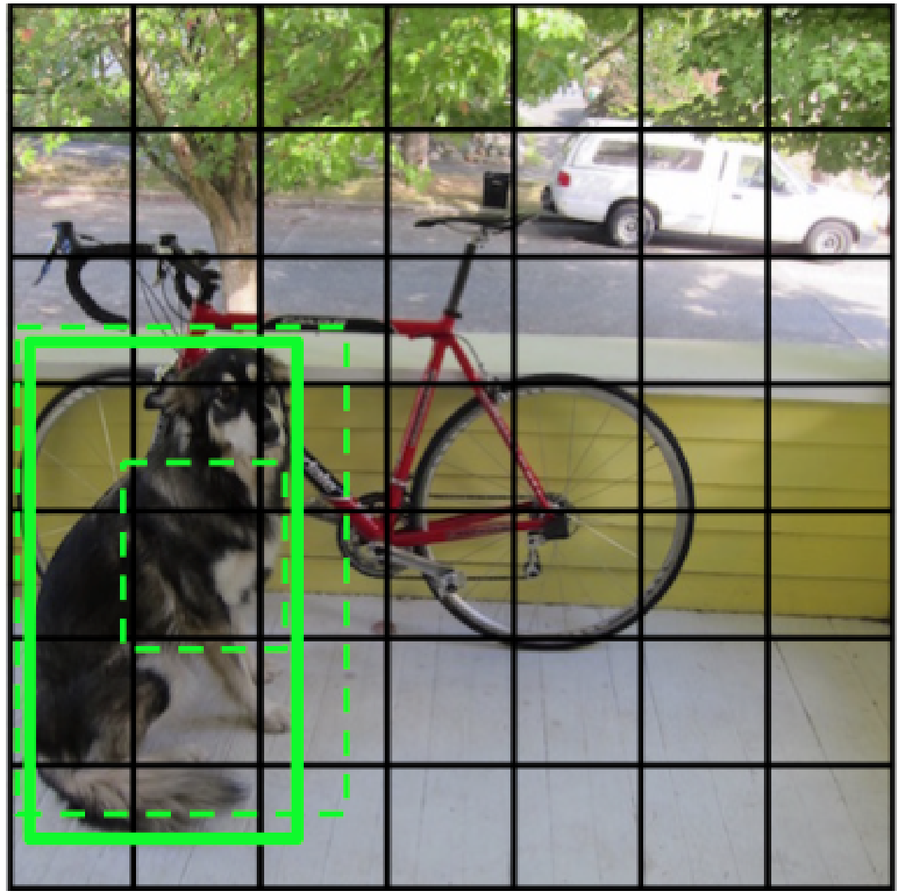
ปรับให้ช่องตารางนั้นเป็นคลาสหนึ่งคลาส

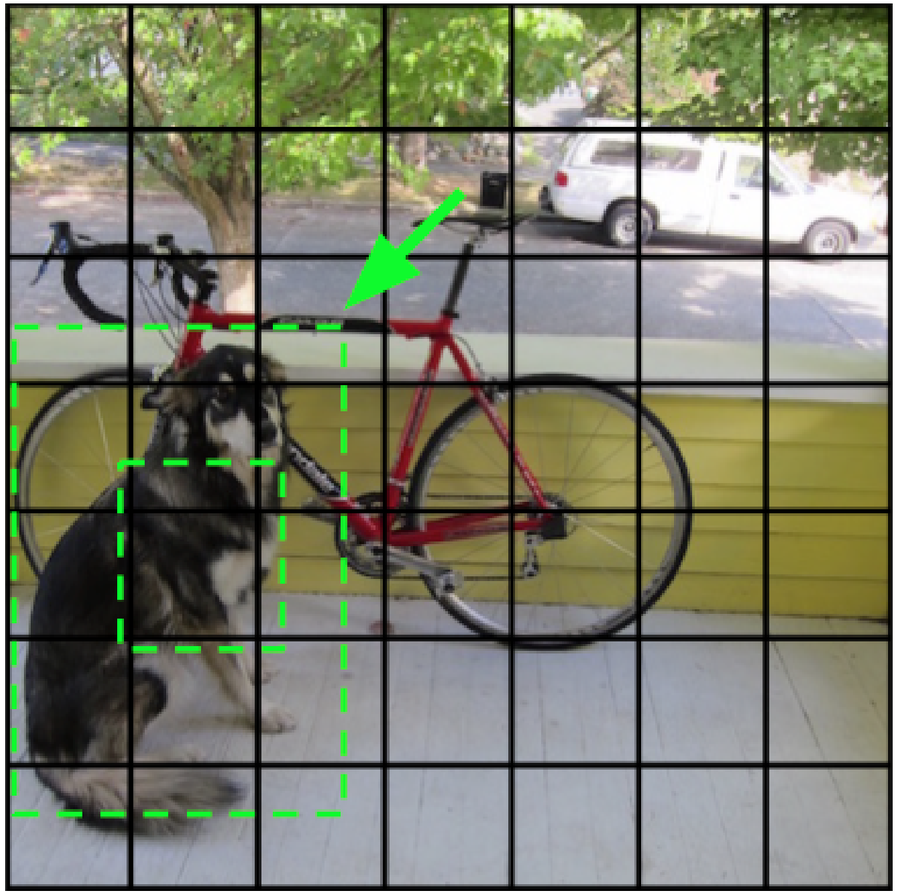
Dog = 1
Cat = 0
Bike = 0
...



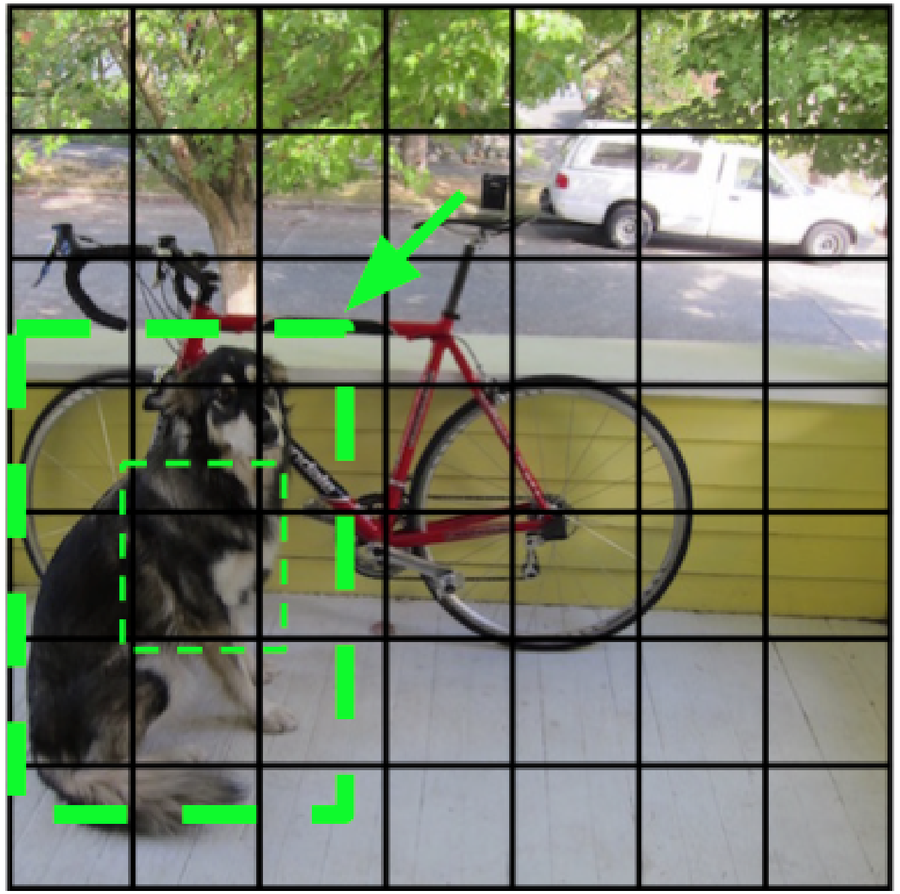


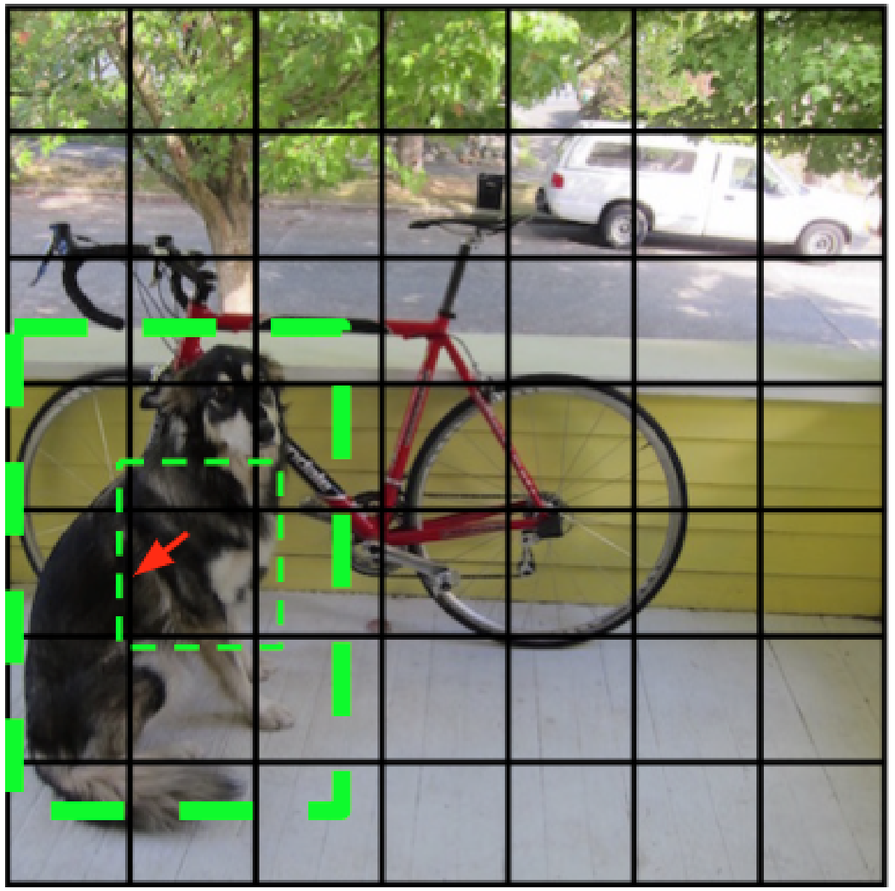
หาช่องตารางที่ดีที่สุด จากนั้นเพิ่มค่าคาดเดา



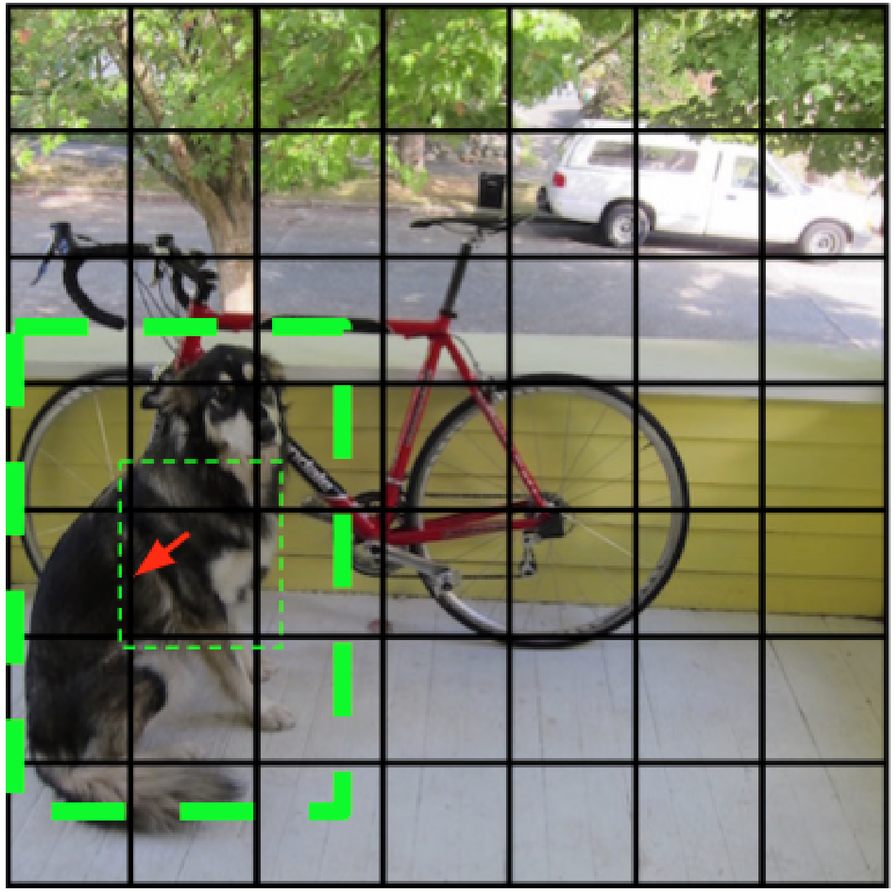


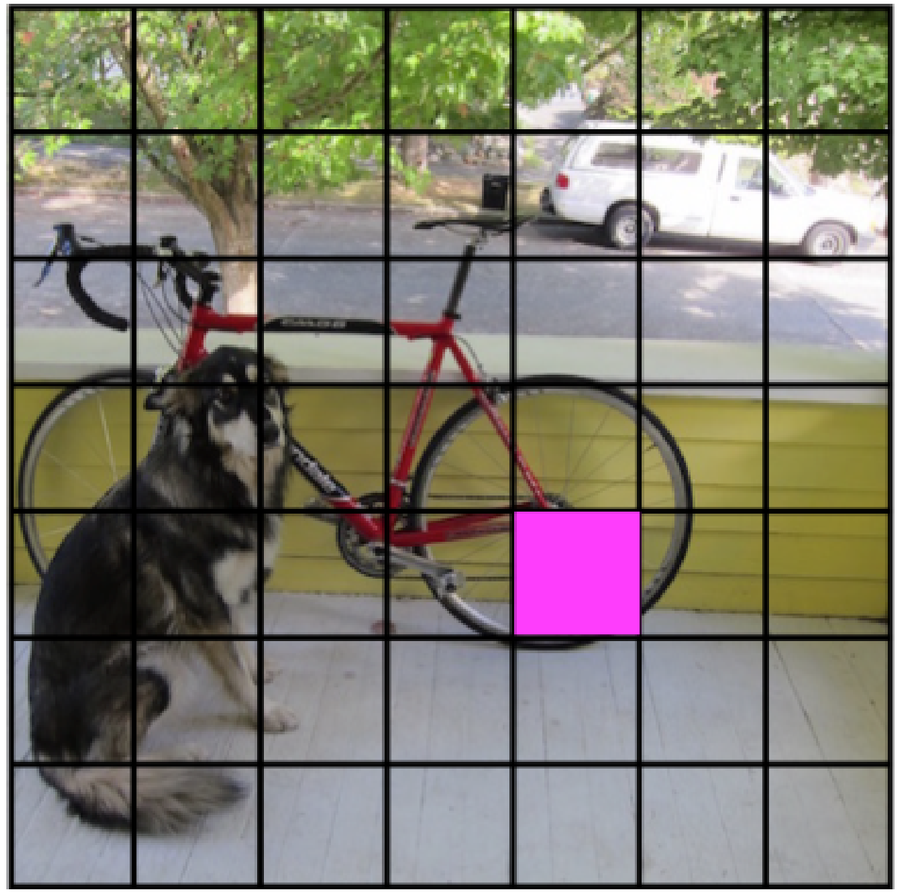
จากนั้นจะลดค่าคาดเดาของตารางอื่น



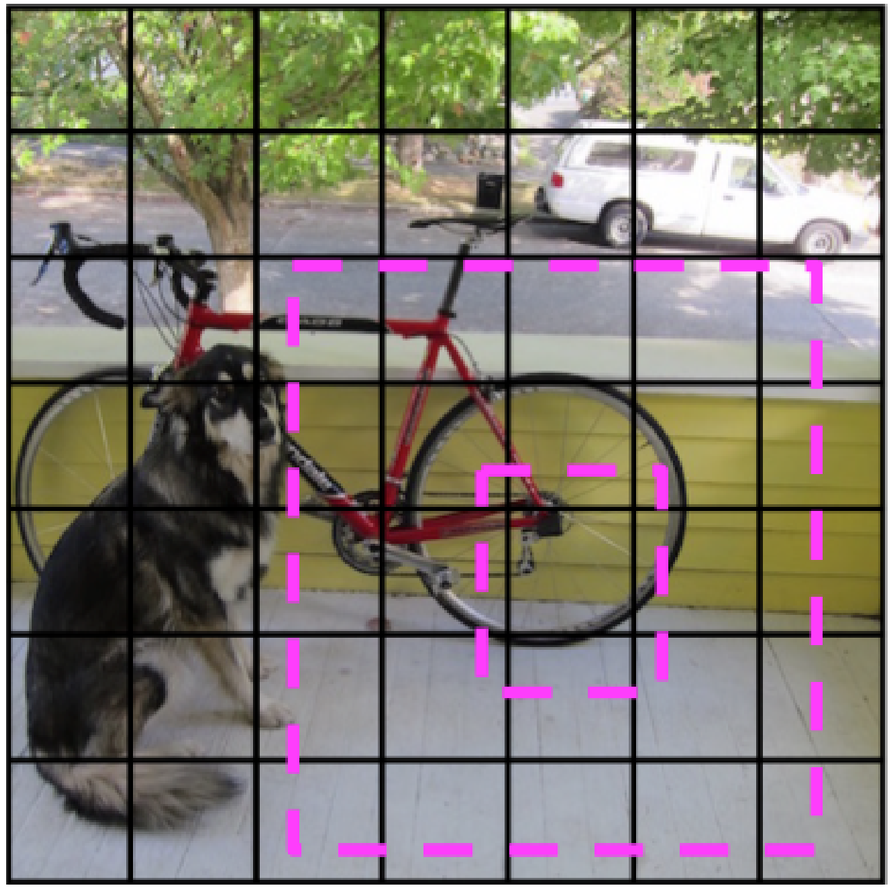


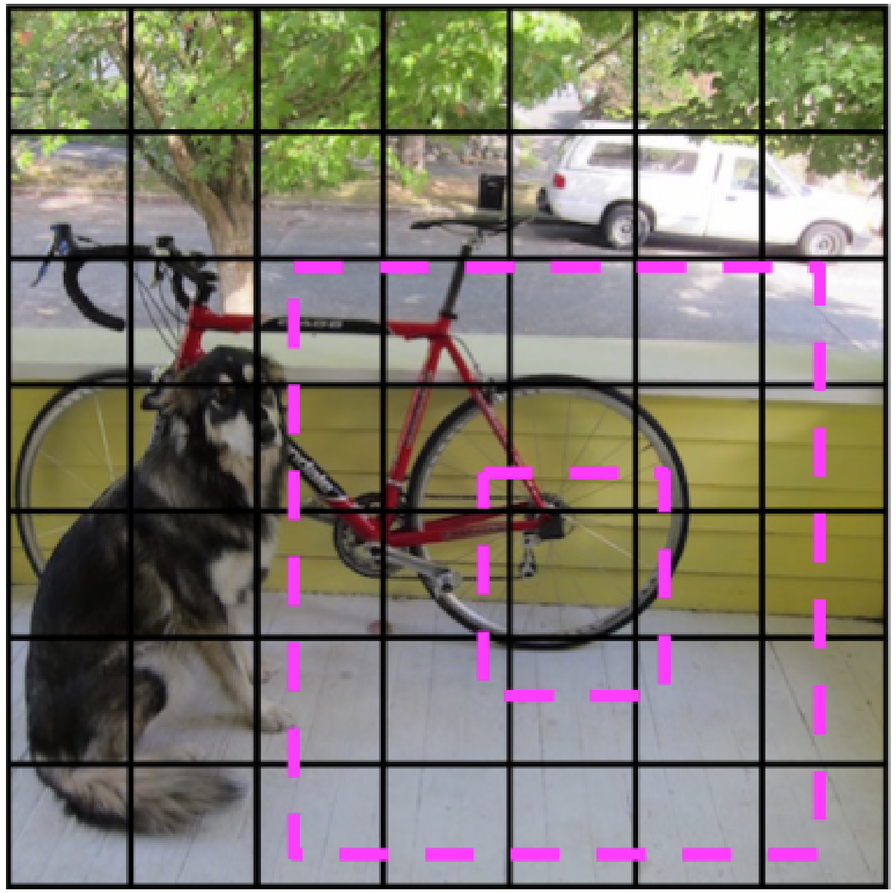
บางจุดไม่มีการอ้างอิงไปยังคลาสใด



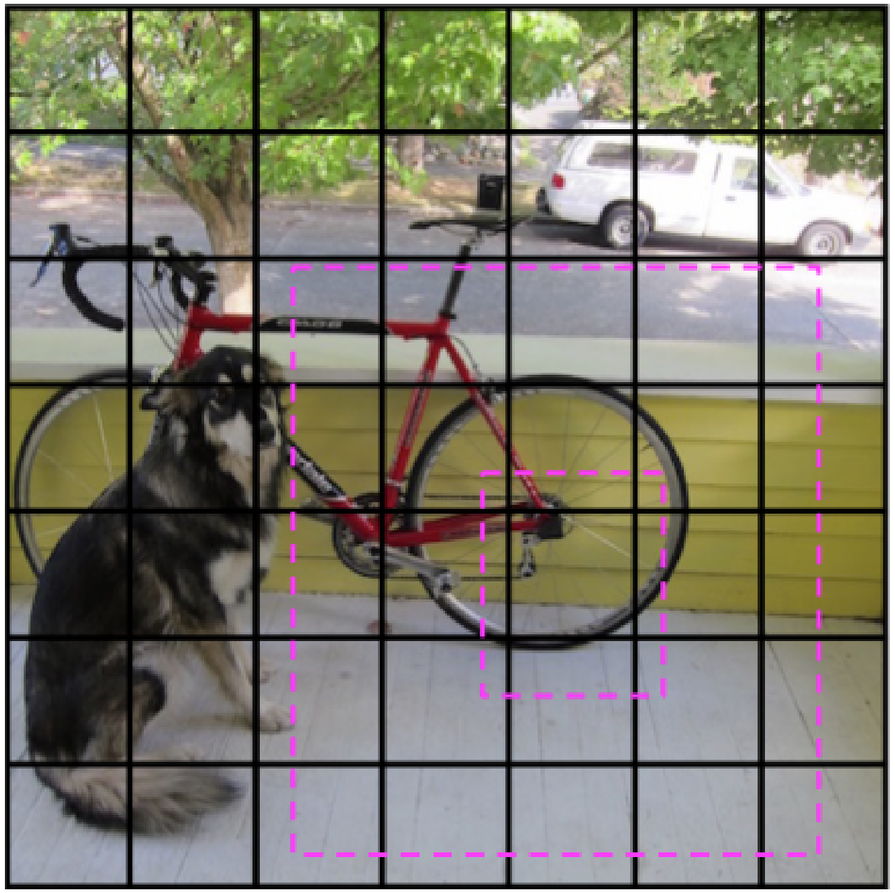


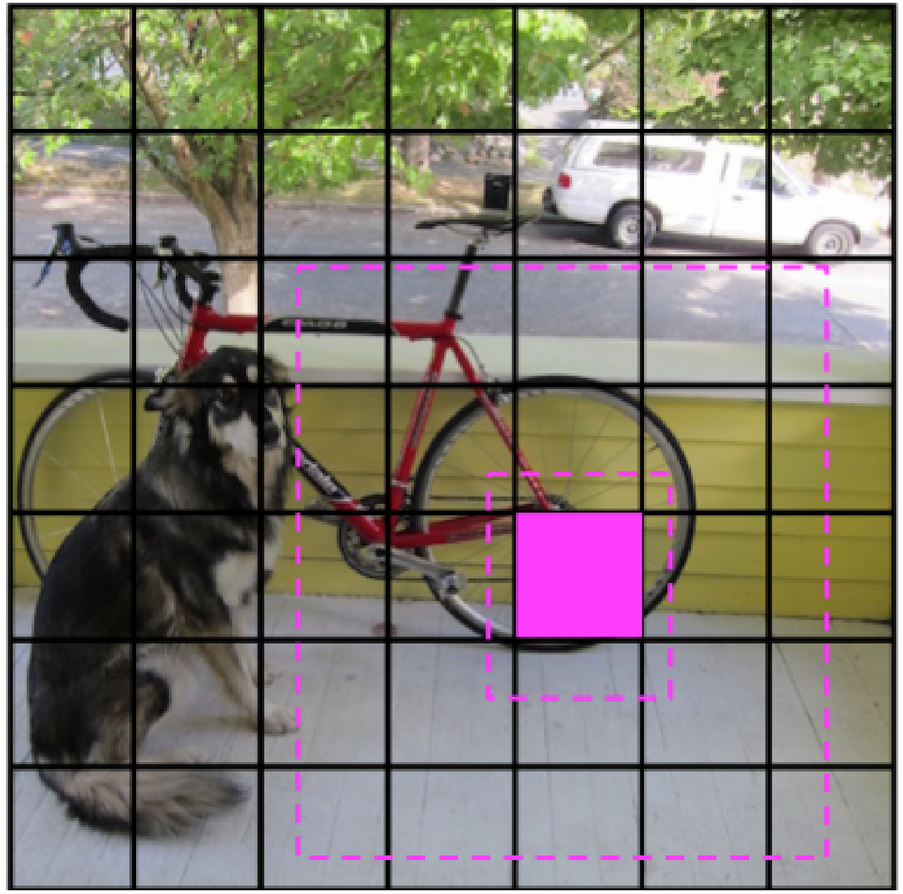
ลดค่า confidence ของกรอบนั้น ๆ



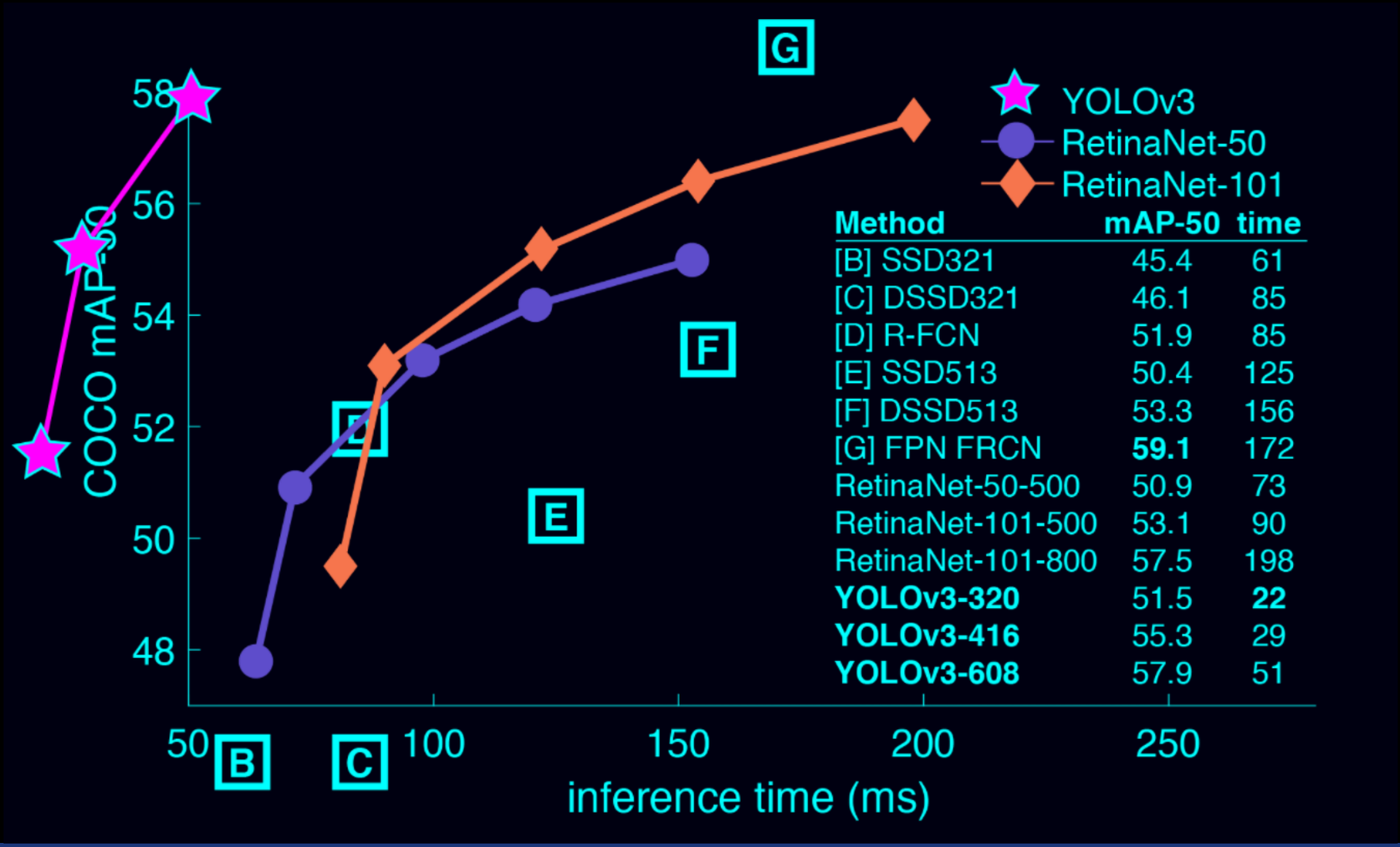


ลดค่า confidence ของกรอบนั้น ๆ





ปัจจุบัน YOLO มีหลายแบบจำลองให้เลือกใช้ (และมีโครงสร้างที่ไม่เหมือนกัน)

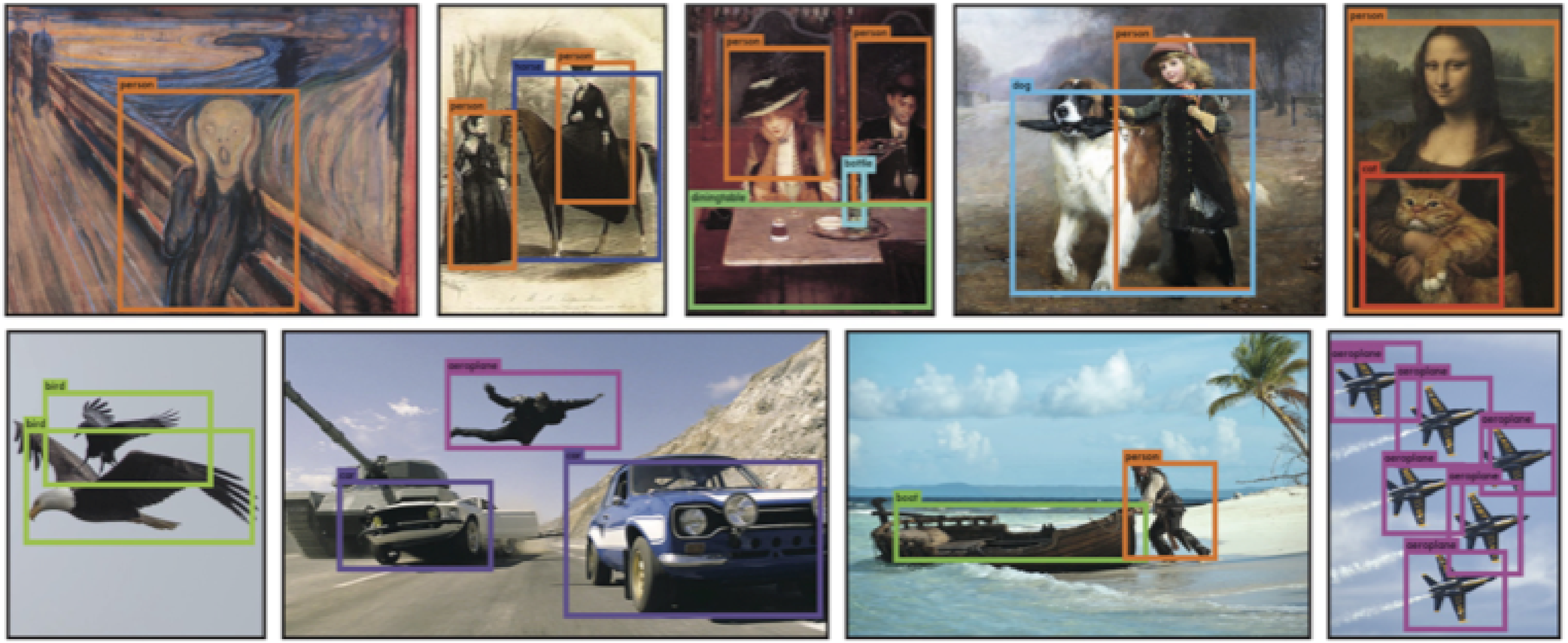


Method	mAP-50	time
[B] SSD321	45.4	61
[C] DSSD321	46.1	85
[D] R-FCN	51.9	85
[E] SSD513	50.4	125
[F] DSSD513	53.3	156
[G] FPN FRCN	**59.1**	172
RetinaNet-50-500	50.9	73
RetinaNet-101-500	53.1	90
RetinaNet-101-800	57.5	198
YOLOv3-320	51.5	**22**
YOLOv3-416	55.3	29
YOLOv3-608	57.9	51

YOLO

Model	Train	Test	mAP	FLOPS	FPS	Cfg	Weights
YOLOv3-320	COCO trainval	test-dev	51.5	38.97 Bn	45	cfg	weights
YOLOv3-416	COCO trainval	test-dev	55.3	65.86 Bn	35	cfg	weights
YOLOv3-608	COCO trainval	test-dev	57.9	140.69 Bn	20	cfg	weights
YOLOv3-tiny	COCO trainval	test-dev	33.1	5.56 Bn	220	cfg	weights
YOLOv3-spp	COCO trainval	test-dev	60.6	141.45 Bn	20	cfg	weights

YOLOv3-tiny คือ Layers น้อยที่สุด (อันบนสุด) จะสังเกตุว่าความเร็วค่อนข้างดีมาก 220 FPS
แต่ถ้าอยากให้แบบจำลองดีขึ้น อาจเลือกใช้ YOLOv3-spp ก็จะถูกต้องขึ้นมาก แต่ความเร็วก็จะตกลง

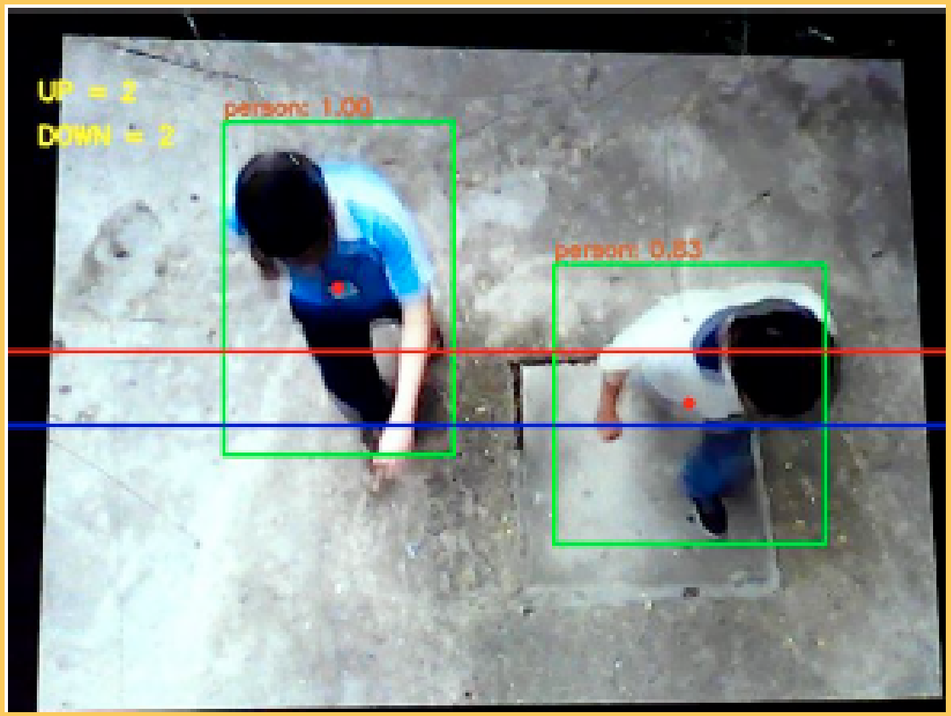

person
person
person
person
person
person
person
person
dog
person
person
cat
bird
bird
aeroplane
car
car
person
boat
aeroplane
aeroplane
aeroplane
aeroplane
aeroplane
aeroplane

ตัวอย่างการนำไปใช้งาน

การตรวจจับหมวกนิรภัยและการใช้อาวุธปืน
เพื่อเตือนภัยเหตุโจรกรรม
จากภาพกล้องวงจรปิดแบบเวลาจริง

จากงานวิจัย
คุณยงยุทธ ละมูลมอญ
และ
คุณธนาสัย สุคนธ์พันธ์

ตัวอย่างการนำไปใช้งาน



การสร้างระบบตรวจนับบุคคลแบบเวลาจริง
ราคาประหยัด
บน Raspberry Pi โดยประยุกต์อัลกอริทึม
Tiny YOLO V3

จากงานวิจัย
คุณปราโมทย์ ปัญญาโต
และ
คุณนลิน สีดาห้าว

Tools Anaconda

ANACONDA®

https://www.anaconda.com/products/individual#Downloads

Anaconda Installers

Windows

Python 3.8

64-Bit Graphical Installer (477 MB)

32-Bit Graphical Installer (409 MB)

MacOS

Python 3.8

64-Bit Graphical Installer (440 MB)

64-Bit Command Line Installer (433 MB)

Linux

Python 3.8

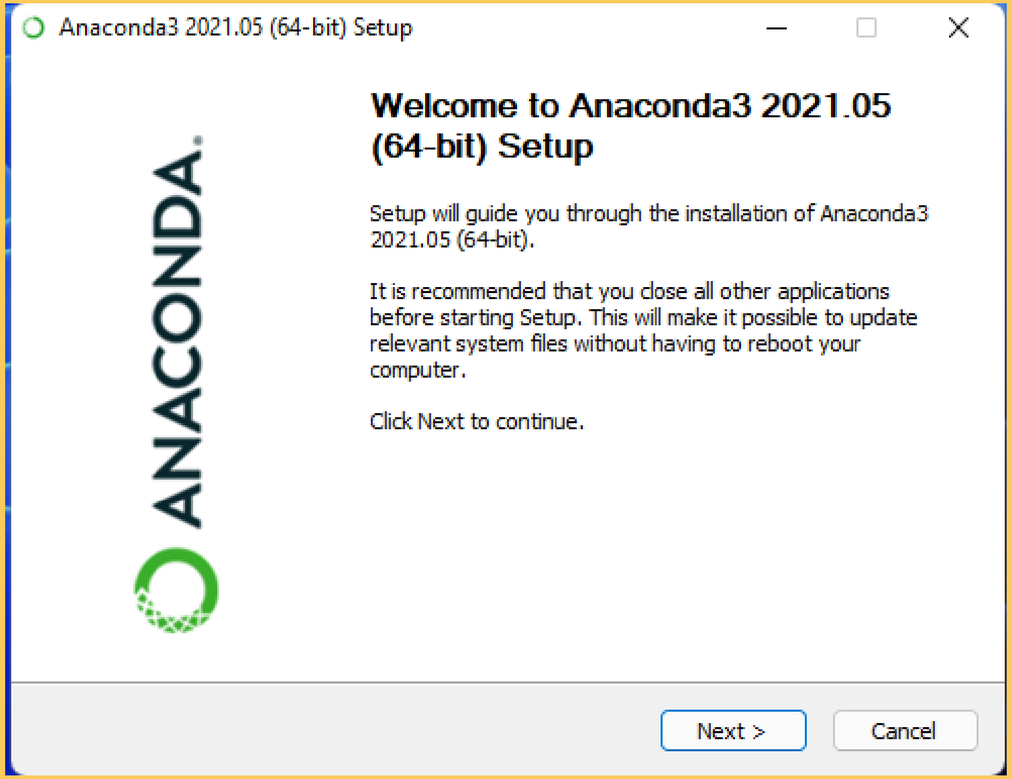
64-Bit (x86) Installer (544 MB)

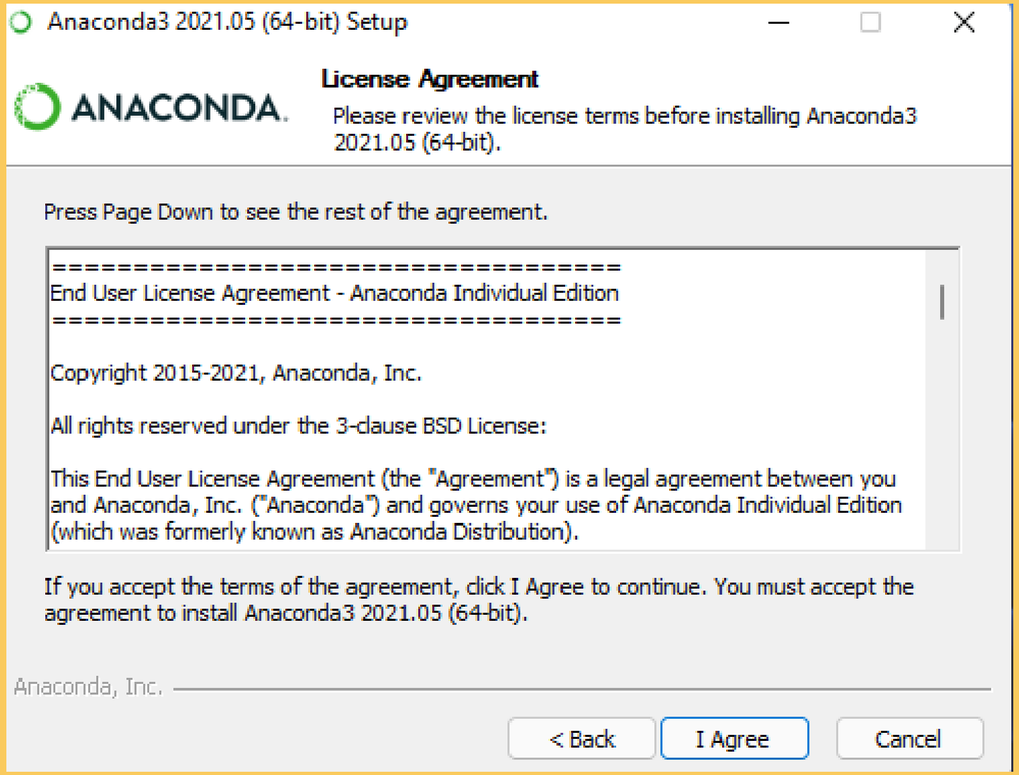
64-Bit (Power8 and Power9) Installer (285 MB)

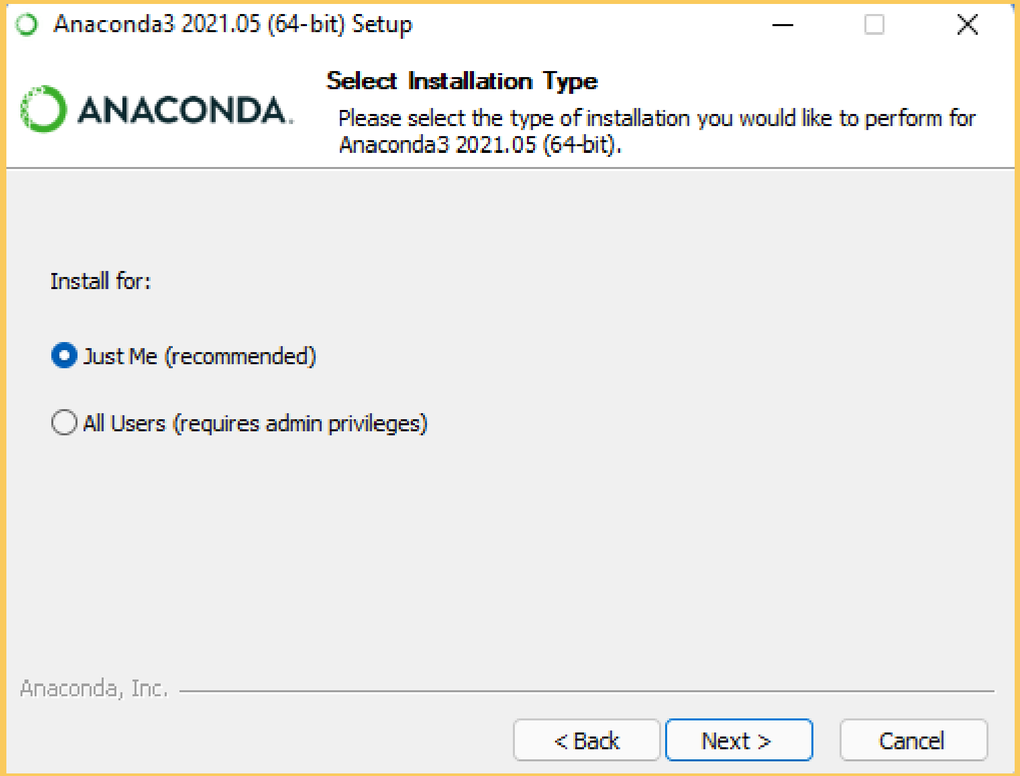
64-Bit (AWS Graviton2 / ARM64) Installer (413 M)

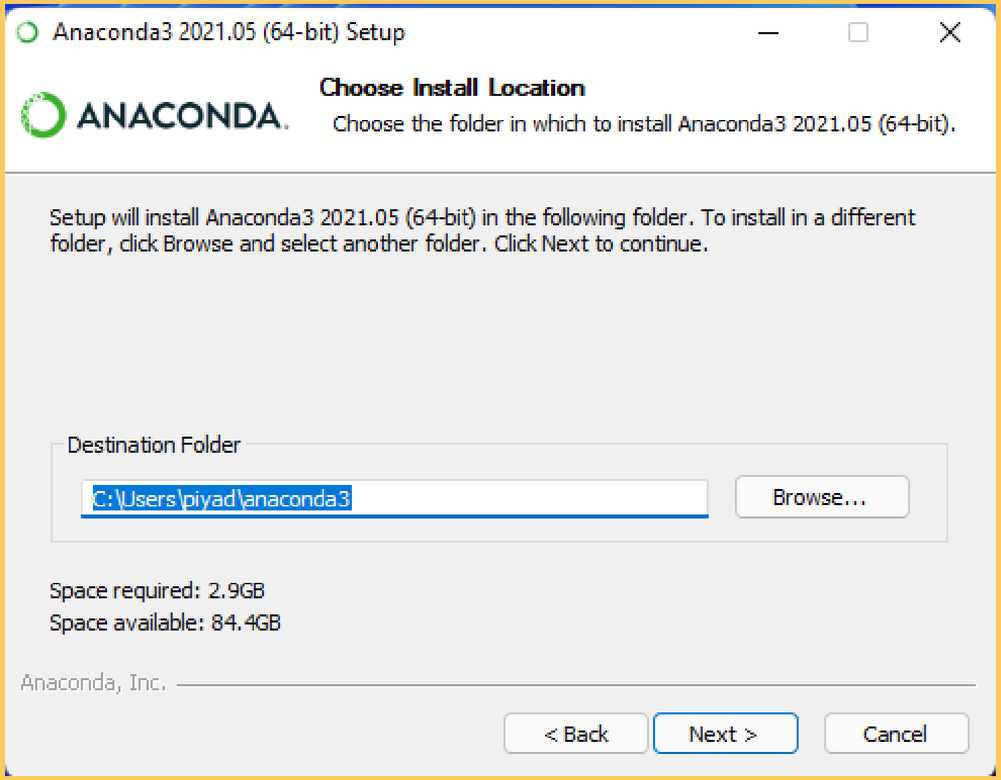
64-bit (Linux on IBM Z & LinuxONE) Installer (292 M)

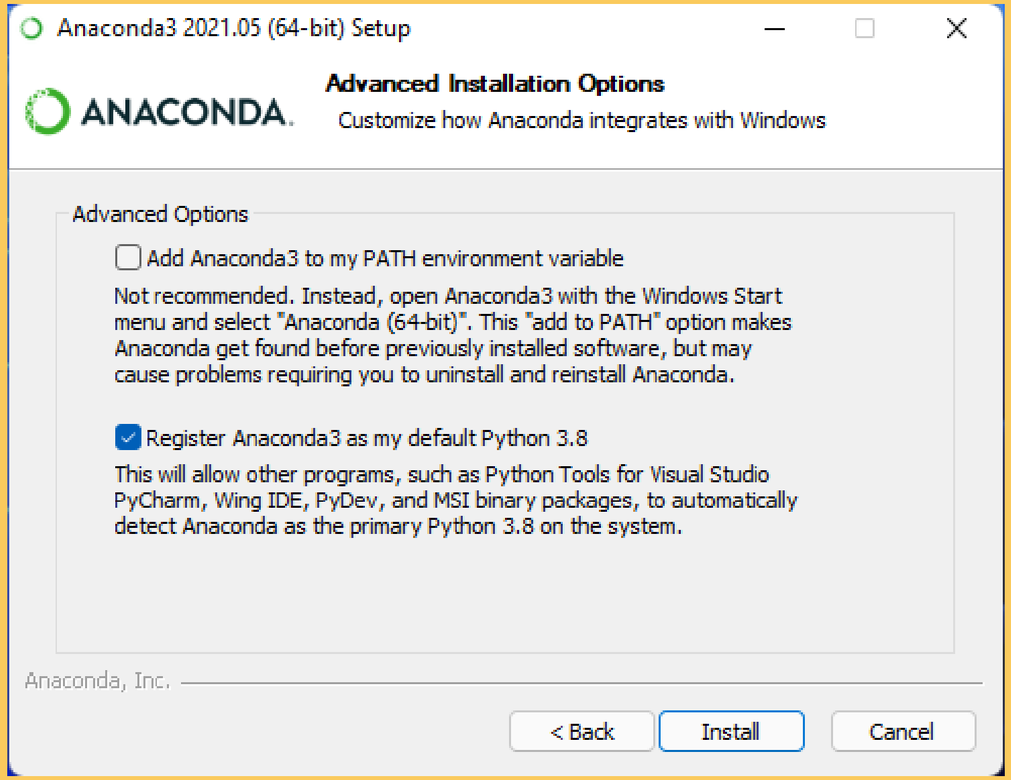
Anaconda3-2021.0....exe
66.3/477 MB, 1 min left

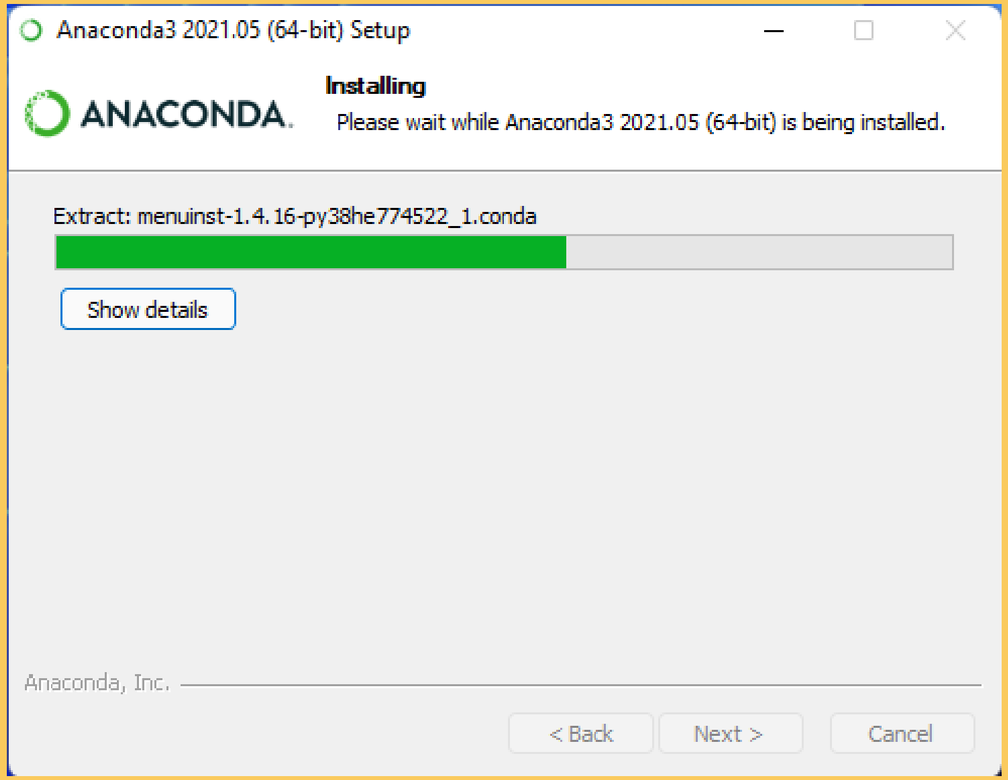

Anaconda3 2021.05 (64-bit) Setup
ANACONDA.
Welcome to Anaconda3 2021.05 (64-bit) Setup
Setup will guide you through the installation of Anaconda3 2021.05 (64-bit).
It is recommended that you close all other applications before starting Setup. This will make it possible to update relevant system files without having to reboot your computer.
Click Next to continue.
Next >
Cancel

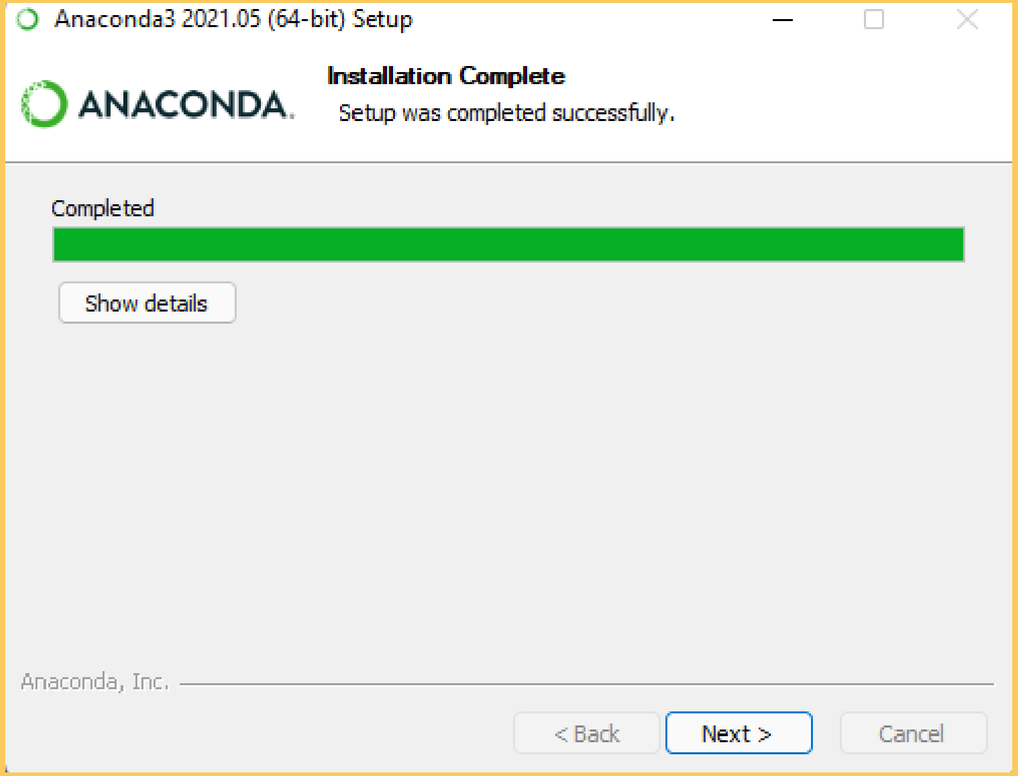

Anaconda3 2021.05 (64-bit) Setup
ANACONDA.
License Agreement
Please review the license terms before installing Anaconda3 2021.05 (64-bit).
Press Page Down to see the rest of the agreement.
==========
End User License Agreement - Anaconda Individual Edition
==========
Copyright 2015-2021, Anaconda, Inc.
All rights reserved under the 3-clause BSD License:
This End User License Agreement (the "Agreement") is a legal agreement between you and Anaconda, Inc. ("Anaconda") and governs your use of Anaconda Individual Edition (which was formerly known as Anaconda Distribution).
If you accept the terms of the agreement, click I Agree to continue. You must accept the agreement to install Anaconda3 2021.05 (64-bit).
Anaconda, Inc.
< Back
I Agree
Cancel

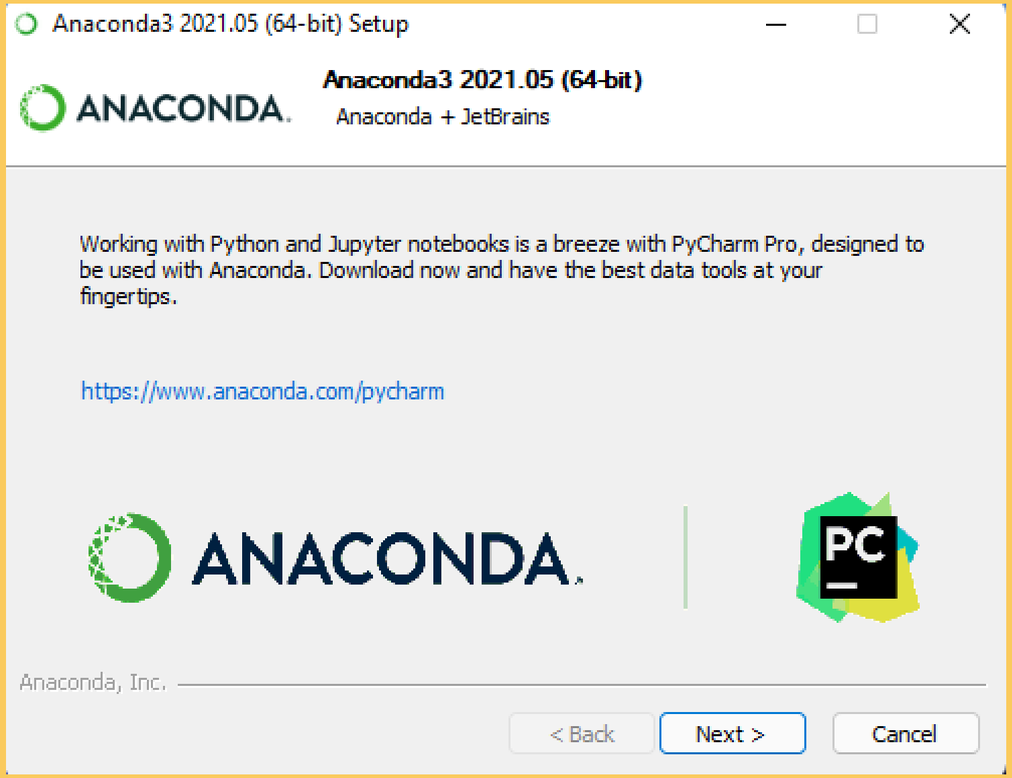

Anaconda3 2021.05 (64-bit) Setup
ANACONDA.
Select Installation Type
Please select the type of installation you would like to perform for Anaconda3 2021.05 (64-bit).
Install for:
Just Me (recommended)
All Users (requires admin privileges)
Anaconda, Inc.
< Back
Next >
Cancel

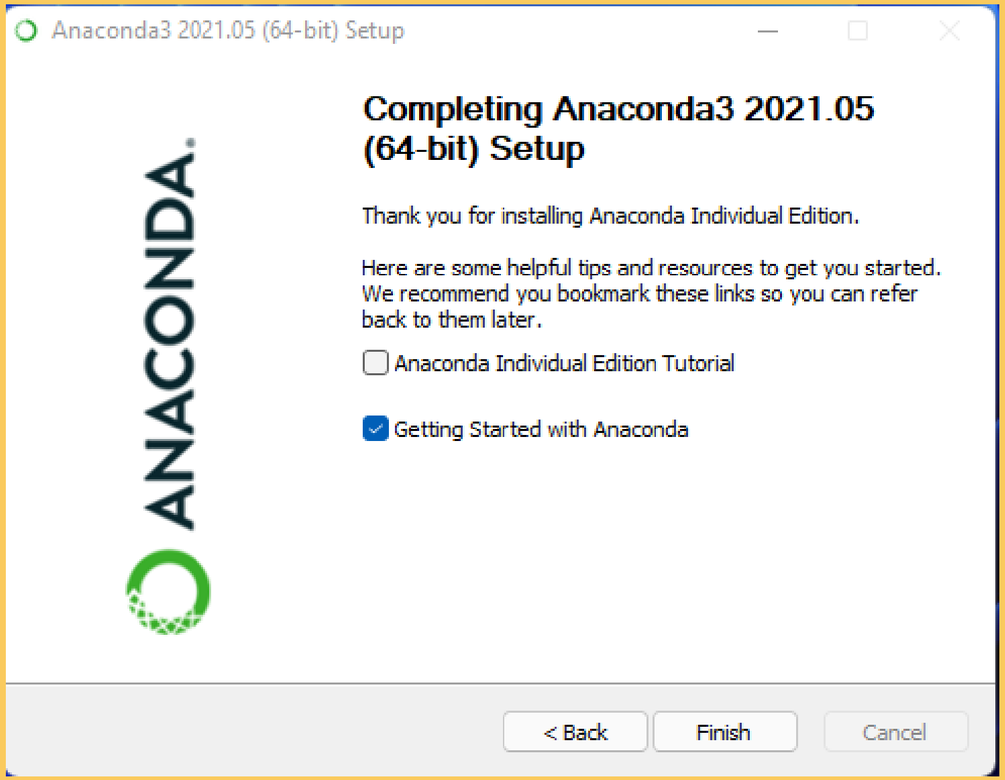

Anaconda3 2021.05 (64-bit) Setup
ANACONDA.
Choose Install Location
Choose the folder in which to install Anaconda3 2021.05 (64-bit).
Setup will install Anaconda3 2021.05 (64-bit) in the following folder. To install in a different folder, click Browse and select another folder. Click Next to continue.
Destination Folder
C:\Users\piyad\anaconda3
Browse...
Space required: 2.9GB
Space available: 84.4GB
Anaconda, Inc.
< Back
Next >
Cancel


Anaconda3 2021.05 (64-bit) Setup
ANACONDA.
Advanced Installation Options
Customize how Anaconda integrates with Windows
Advanced Options
Add Anaconda3 to my PATH environment variable
Not recommended. Instead, open Anaconda3 with the Windows Start menu and select "Anaconda (64-bit)". This "add to PATH" option makes Anaconda get found before previously installed software, but may cause problems requiring you to uninstall and reinstall Anaconda.
Register Anaconda3 as my default Python 3.8
This will allow other programs, such as Python Tools for Visual Studio PyCharm, Wing IDE, PyDev, and MSI binary packages, to automatically detect Anaconda as the primary Python 3.8 on the system.
Anaconda, Inc.
< Back
Install
Cancel


Anaconda3 2021.05 (64-bit) Setup
ANACONDA.
Installing
Please wait while Anaconda3 2021.05 (64-bit) is being installed.
Extract: menuinst-1.4.16-py38he774522_1.conda
Show details
Anaconda, Inc.
< Back
Next >
Cancel

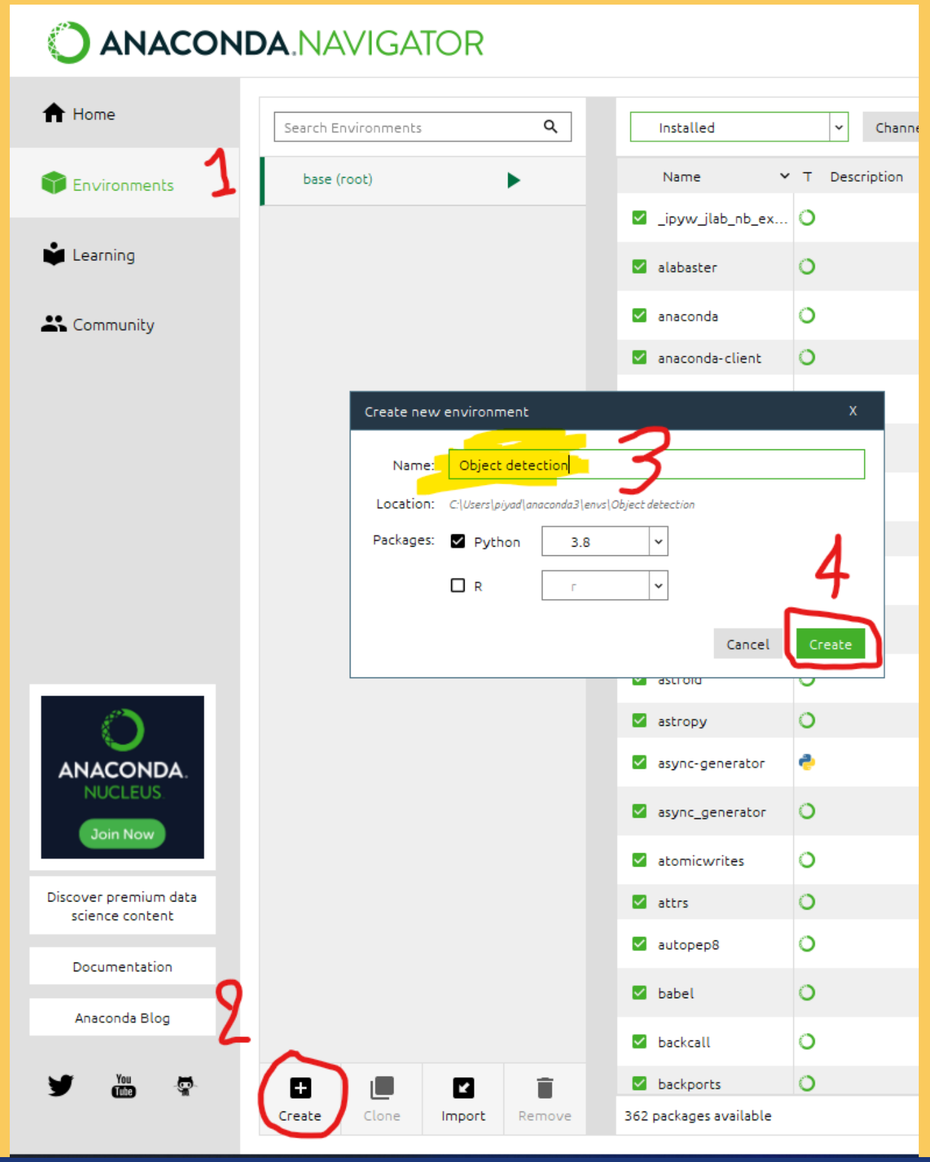

Anaconda3 2021.05 (64-bit) Setup
ANACONDA.
Installation Complete
Setup was completed successfully.
Completed
Show details
Anaconda, Inc.
< Back
Next >
Cancel

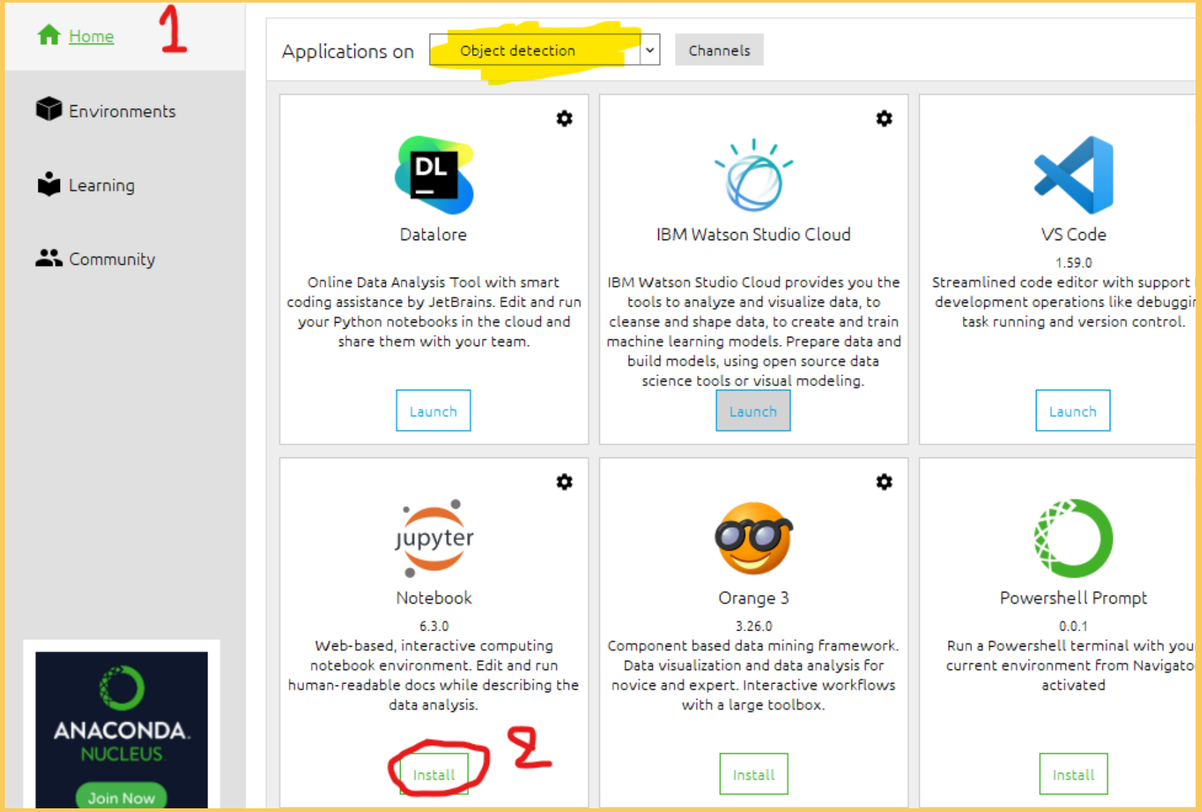

Anaconda3 2021.05 (64-bit) Setup
ANACONDA.
Anaconda3 2021.05 (64-bit)
Anaconda + JetBrains
Working with Python and Jupyter notebooks is a breeze with PyCharm Pro, designed to be used with Anaconda. Download now and have the best data tools at your fingertips.
https://www.anaconda.com/pycharm
ANACONDA.
PC
Anaconda, Inc.
< Back
Next >
Cancel


Anaconda3 2021.05 (64-bit) Setup
ANACONDA.
Completing Anaconda3 2021.05 (64-bit) Setup
Thank you for installing Anaconda Individual Edition.
Here are some helpful tips and resources to get you started. We recommend you bookmark these links so you can refer back to them later.
Anaconda Individual Edition Tutorial
Getting Started with Anaconda
< Back
Finish
Cancel

Anaconda3 (64-bit)
New
Anaconda Navigator (anaconda3)
New
Anaconda Powershell Prompt (anaconda3)
New
Anaconda Prompt (anaconda3)
New
Jupyter Notebook (anaconda3)
New
Reset Spyder Settings (anaconda3)
New
Spyder (anaconda3)
New

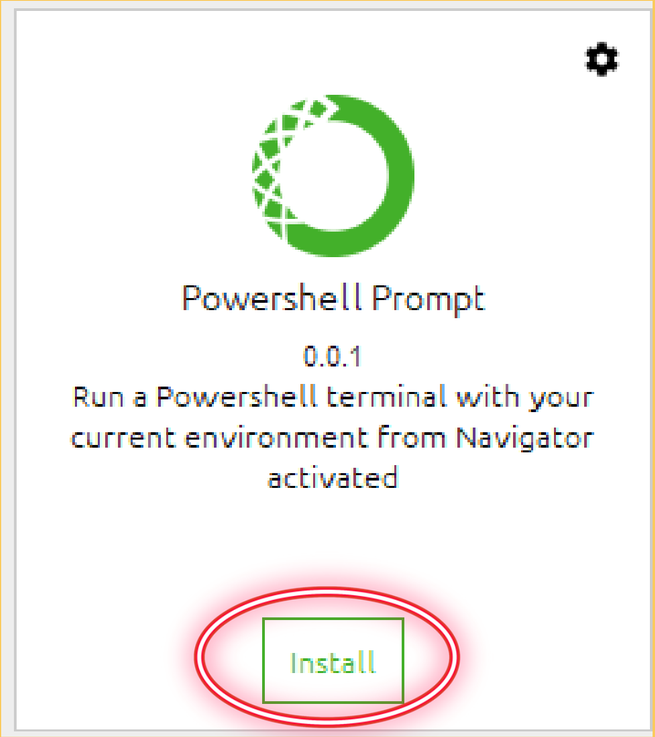
สร้าง Envoriment สำหรับใช้การทำงาน

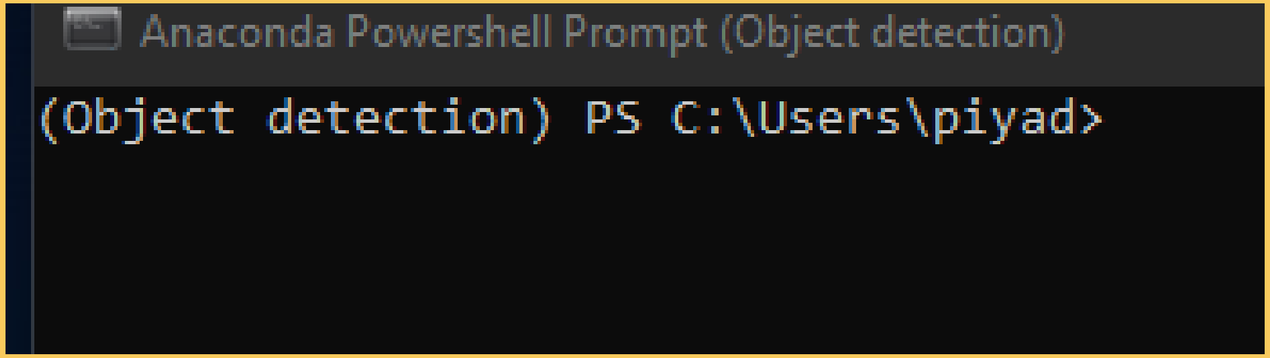




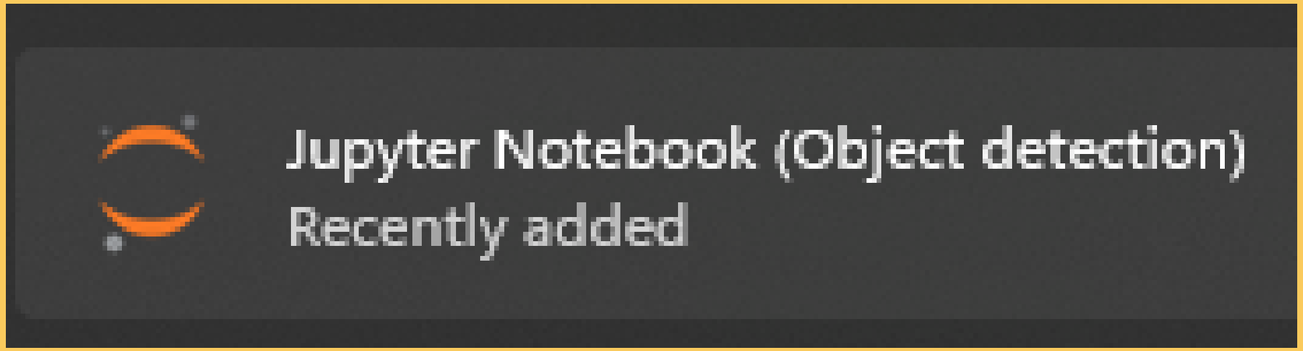
Install application **notebook** on **C:\Users\piyad\anaconda3\envs\Object detection**

สร้าง Envoriment สำหรับใช้การทำงาน

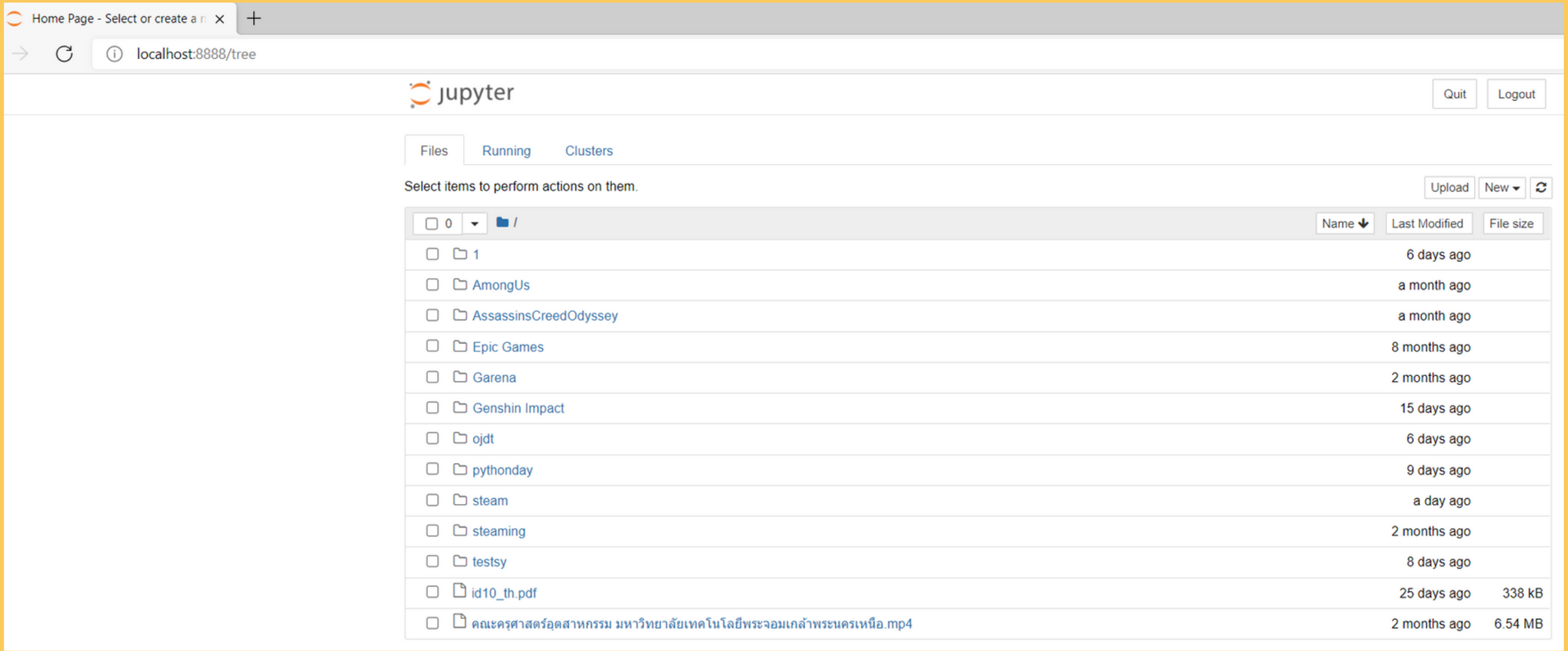




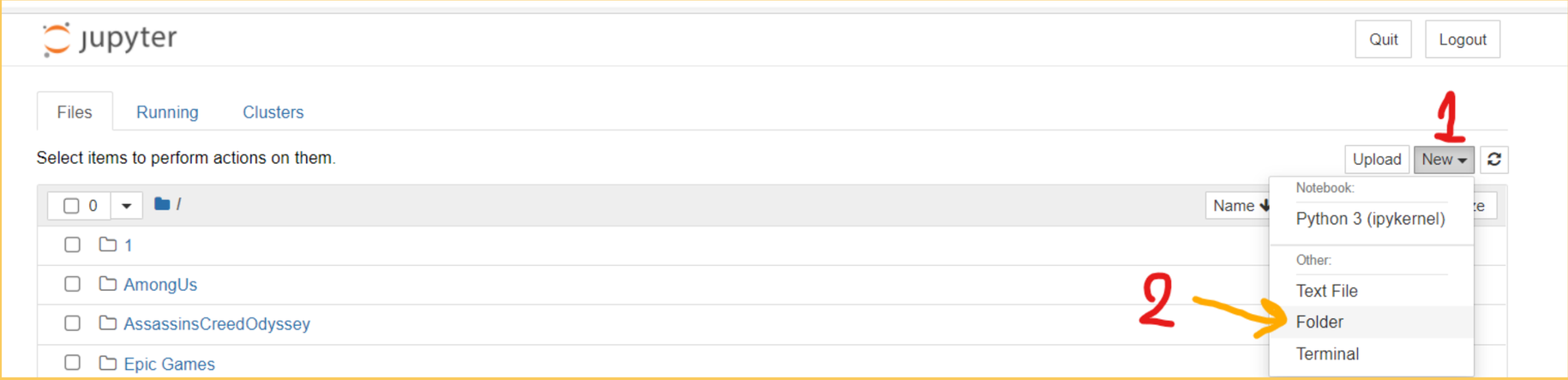
เข้าใช้งาน jupyter Notebook

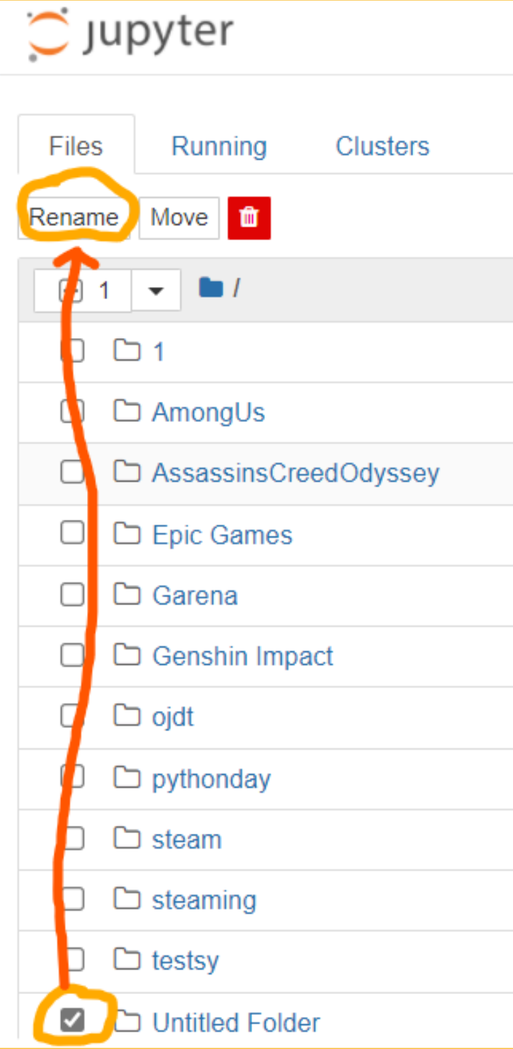


หน้าตาของ Jupyter Notebook



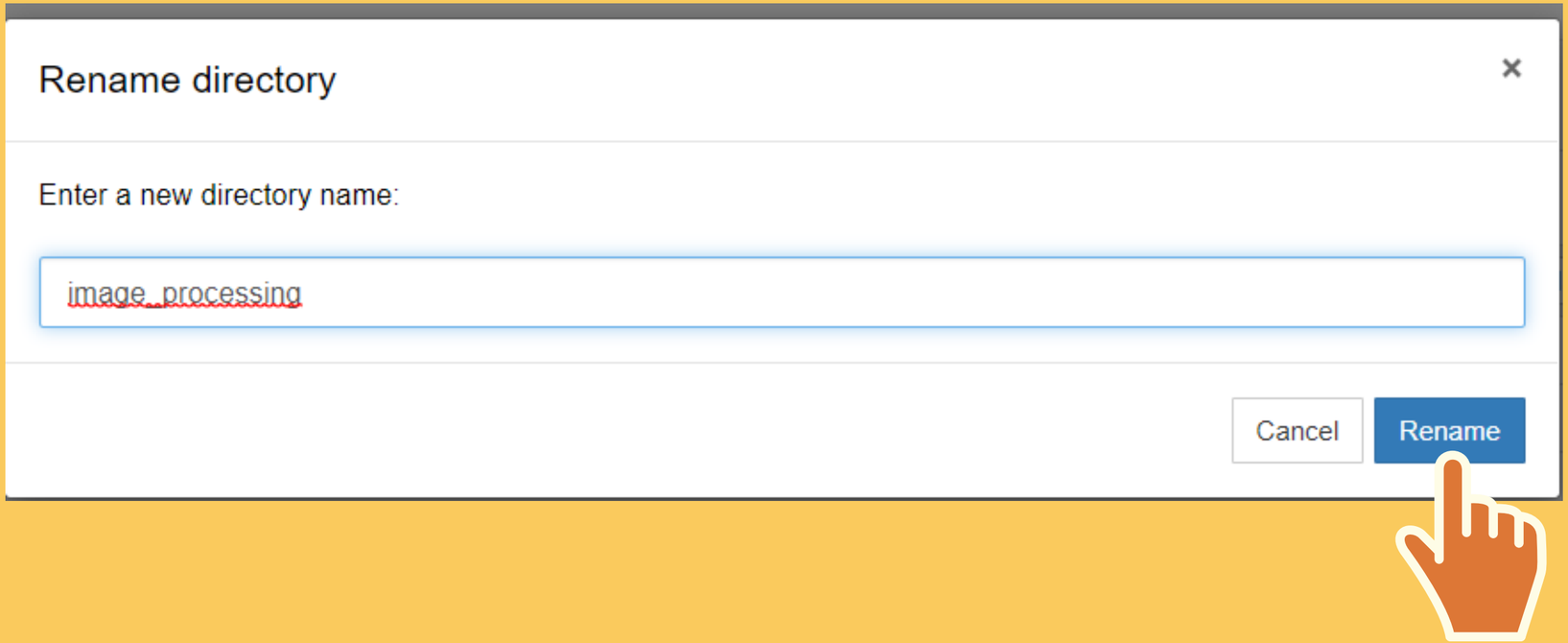
สร้างโฟลเดอร์ เพื่อเก็บชิ้นงาน





เปลี่ยนชื่อโฟล์เดอร์

ตามที่ต้องการแต่ขอแนะนำให้ใช้คำว่า image_processing


Rename directory
Enter a new directory name:
image_processing
Cancel
Rename

image_processing

OpenCV

OpenCV (Open source Computer Vision) เป็นไลบรารีฟังก์ชันการเขียนโปรแกรม (Library of Programming Functions) โดยส่วนใหญ่จะมุ่งเป้าไปที่การแสดงผลด้วยคอมพิวเตอร์แบบเรียลไทม์ (Real-Time Computer Vision)

Opencv

https://pypi.org/project/opencv-python/

Search projects

Help Sponsors Log in Register

opencv-python 4.5.3.56

```
pip install opencv-python
```

Latest version

Released: Jul 11, 2021

Wrapper package for OpenCV python bindings.

Navigation

- Project description
- Release history
- Download files

Project links

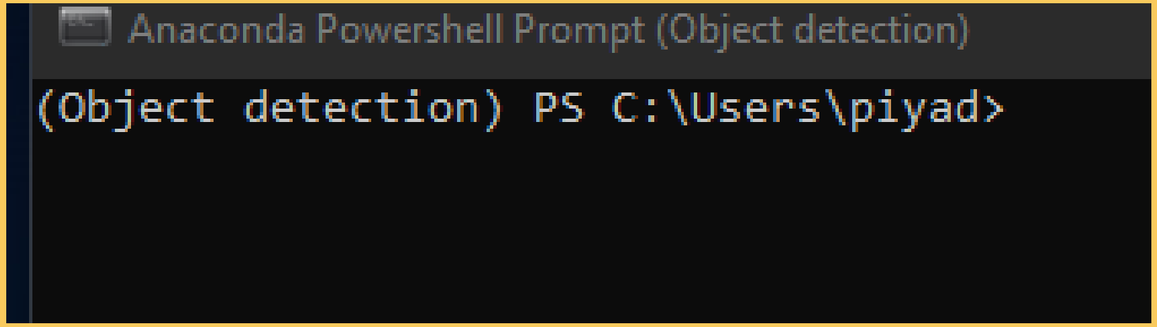
Project description

downloads 119M

OpenCV on Wheels

Pre-built CPU-only OpenCV packages for Python.

Check the manual build section if you wish to compile the bindings from source to enable additional modules such as CUDA.



ขั้นตอนติดตั้ง

เปิด PowerShell

```
Anaconda Powershell Prompt (Object detection)
(Object detection) PS C:\Users\piyad> pip install opencv-python
Collecting opencv-python
  Using cached opencv_python-4.5.3.56-cp38-cp38-win_amd64.whl (34.9 MB)
```

ขั้นตอนติดตั้ง

ป้อนคำสั่ง

pip install opencv-python

NumPy

เป็นชื่อของ library ที่ใช้ในการคำนวนทางคณิตศาสตร์ในภาษา Python ซึ่งภายในถูกเขียนด้วยภาษา C จึงทำงานได้เร็วและมีประสิทธิภาพ โดย NumPy มีความสามารถในการจัดการกับอาเรย์หลายมิติและข้อมูลแบบเมทริกซ์ เราสามารถติดตั้ง NumPy ได้ง่ายผ่าน package installer ของ Python ด้วยคำสั่ง

เป็นแก่นสำคัญของ Deep Learning แล้วยังเป็นแก่นสำคัญของการประยุกต์ใช้ Python ในงานคำนวณต่างๆ ทั้งด้านวิทยาศาสตร์ วิศวกรรมศาสตร์ สถิติ ธุรกิจ กราฟฟิก ฯลฯ อีกด้วย ดังนั้นเรียนรู้ Numpy ไว้มีประโยชน์ต่อชีวิตนักเขียนโปรแกรมมืออาชีพสายคำนวนมากๆ

NUMPY

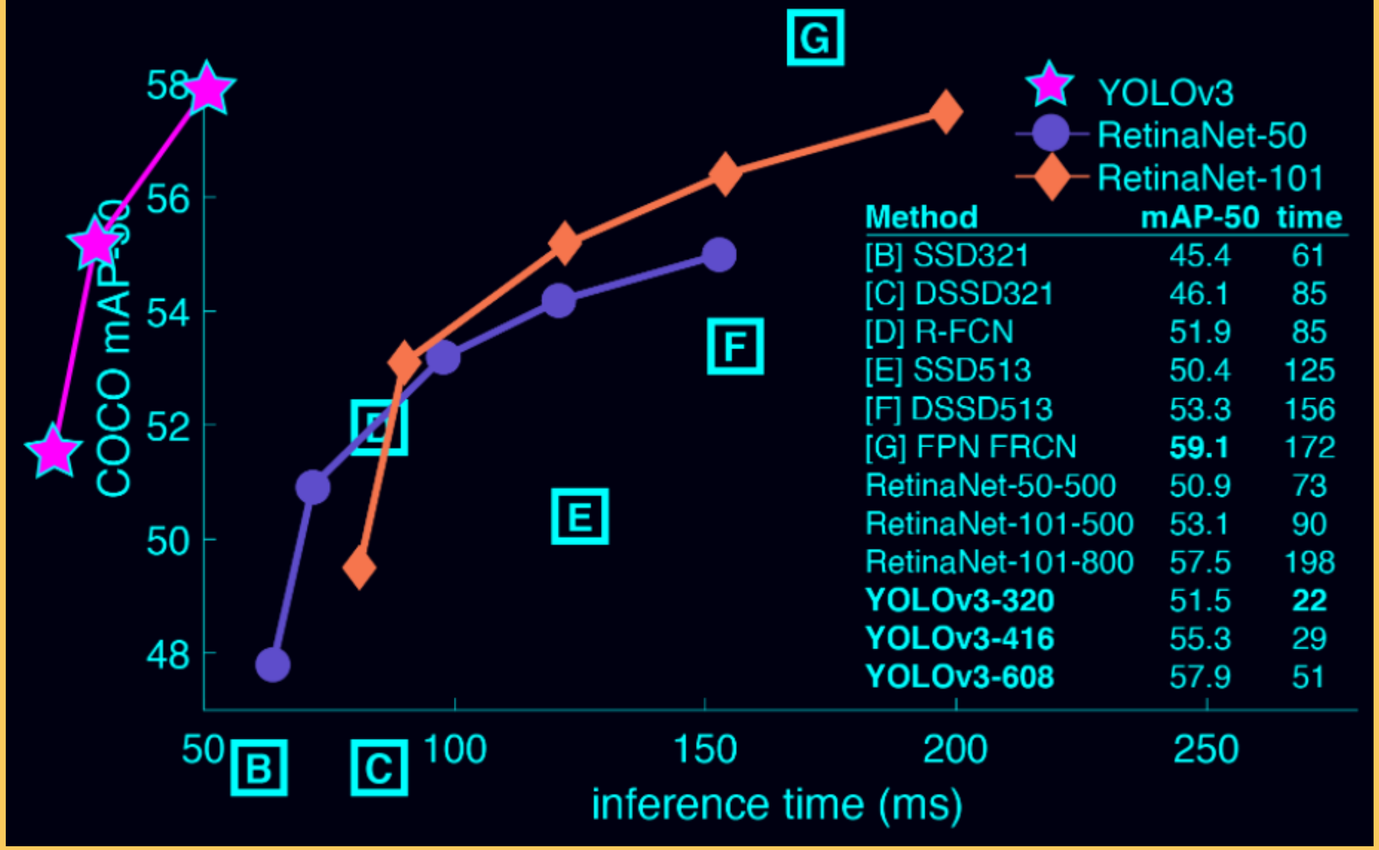
https://numpy.org/install/

YOLO

https://pjreddie.com/darknet/yolo/

Comparison to Other Detectors

YOLOv3 is extremely fast and accurate. In mAP measured at .5 IOU YOLOv3 is on par with Focal Loss but about 4x faster. Moreover, you can easily tradeoff between speed and accuracy simply by changing the size of the model, no retraining required!



Method	mAP-50	time
[B] SSD321	45.4	61
[C] DSSD321	46.1	85
[D] R-FCN	51.9	85
[E] SSD513	50.4	125
[F] DSSD513	53.3	156
[G] FPN FRCN	**59.1**	172
RetinaNet-50-500	50.9	73
RetinaNet-101-500	53.1	90
RetinaNet-101-800	57.5	198
YOLOv3-320	51.5	**22**
YOLOv3-416	55.3	29
YOLOv3-608	57.9	51

SSD321	COCO trainval	test-dev	45.4	-	16		link
DSSD321	COCO trainval	test-dev	46.1	-	12		link
R-FCN	COCO trainval	test-dev	51.9	-	12		link
SSD513	COCO trainval	test-dev	50.4	-	8		link
DSSD513	COCO trainval	test-dev	53.3	-	6		link
FPN FRCN	COCO trainval	test-dev	59.1	-	6		link
Retinanet-50-500	COCO trainval	test-dev	50.9	-	14		link
Retinanet-101-500	COCO trainval	test-dev	53.1	-	11		link
Retinanet-101-800	COCO trainval	test-dev	57.5	-	5		link
YOLOv3-320	COCO trainval	test-dev	51.5	38.97 Bn	45	cfg	weights
YOLOv3-416	COCO trainval	test-dev	55.3	65.86 Bn	35	cfg	weights
YOLOv3-608	COCO trainval	test-dev	57.9	140.69 Bn	20	cfg	weights
YOLOv3-tiny	COCO trainval	test-dev	33.1	5.56 Bn	220	cfg	weights
YOLOv3-spp	COCO trainval	test-dev	60.6	141.45 Bn	20	cfg	weights

YOLOv3-tiny	COCO trainval	test-dev	33.1	5.56 Bn	220	cfg	weights

ดาวน์โหลดทั้งไฟล์ cfg
ลงในโฟลเดอร์ที่สร้างไว้

Why GitHub? Team Enterprise Explore Marketplace Pricing / Sign

pjreddie / darknet

Notifications Star 21.2k

Code Issues 1.7k Pull requests 160 Actions Projects Wiki Security Insight

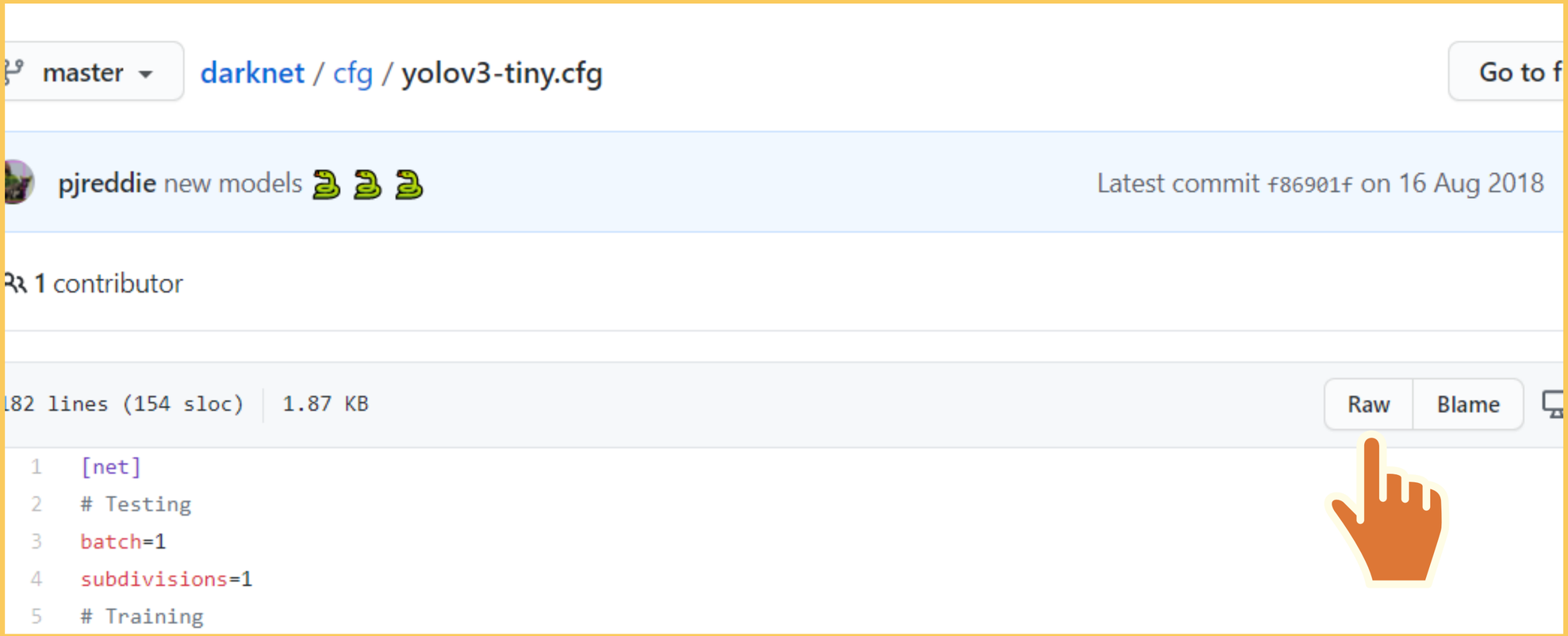
master darknet / cfg / yolov3-tiny.cfg Go to fil

pjreddie new models Latest commit f86901f on 16 Aug 2018

1 contributor

182 lines (154 sloc) 1.87 KB Raw Blame

```
[net]
# Testing
batch=1
```


master
darknet / cfg / yolov3-tiny.cfg
Go to f
pjreddie new models
Latest commit f86901f on 16 Aug 2018
1 contributor
182 lines (154 sloc)
1.87 KB
Raw
Blame
[net]
# Testing
batch=1
subdivisions=1
# Training

กดปุ่ม Raw

https://raw.githubusercontent.com/pj...

```
[net]
# Testing
batch=1
subdivisions=1
# Training
# batch=64
# subdivisions=2
width=416
height=416
channels=3
momentum=0.9
decay=0.0005
angle=0
saturation = 1.5
exposure = 1.5
hue=.1

learning_rate=0.001
burn_in=1000
max_batches = 500200
policy=steps
steps=400000,450000
scales=.1,.1

[convolutional]
batch_normalize=1
filters=16
size=3
stride=1
pad=1
activation=leaky

[maxpool]
size=2
stride=2

[convolutional]
```

ทำการคลิกขวาเพื่อบันทึก
บันทึกลงในโฟล์เดอร์งาน

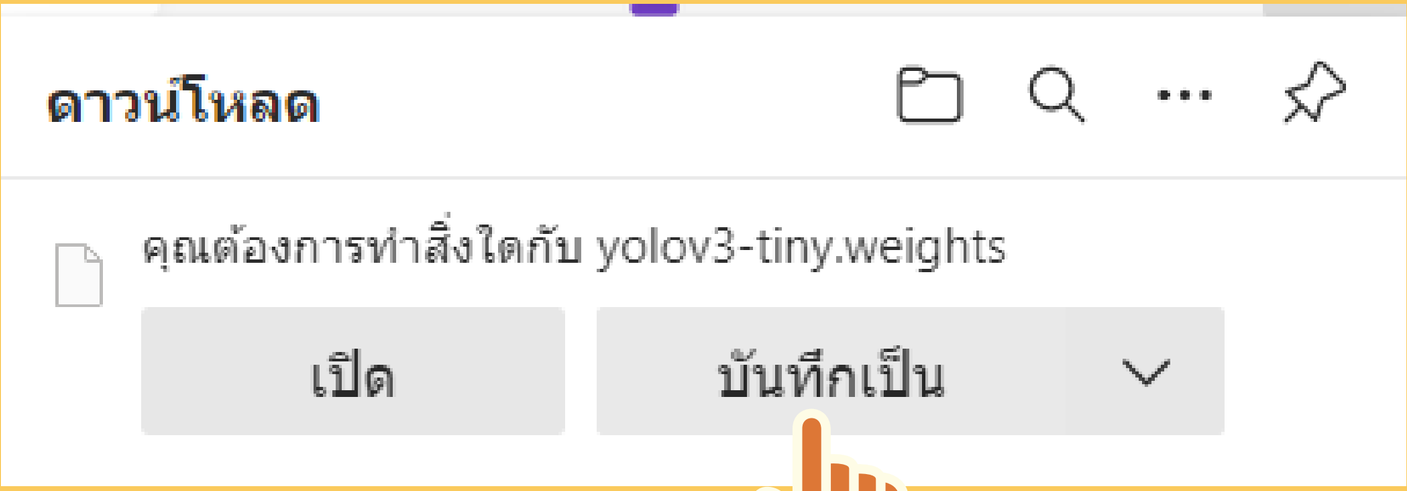
jupyter yolov3.cfg 11 นาทีที่แล้ว

File Edit View Language

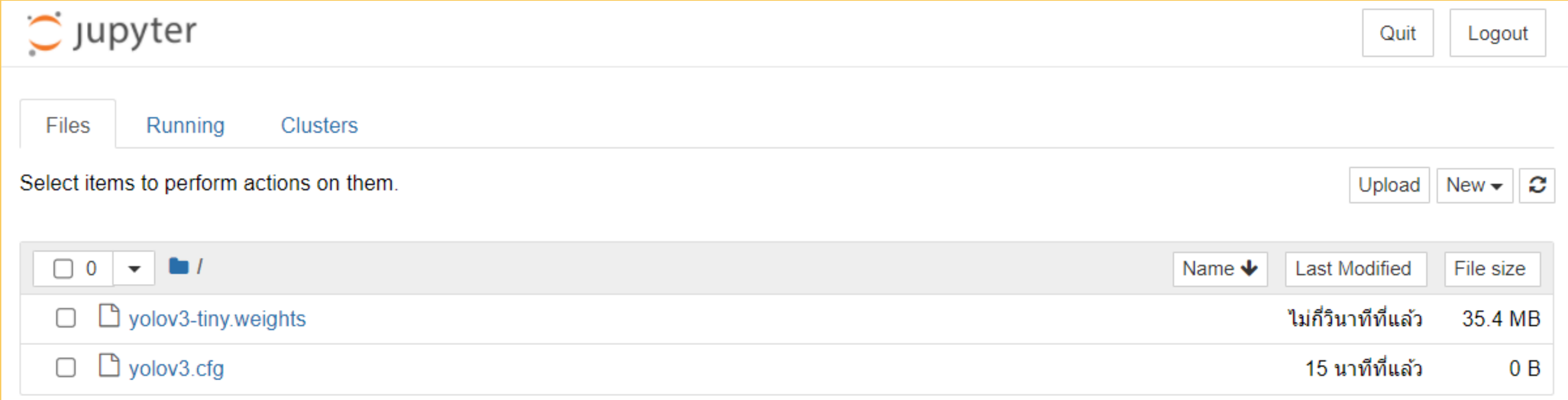
```
[net]
# Testing
batch=1
subdivisions=1
# Training
# batch=64
# subdivisions=2
width=416
height=416
channels=3
momentum=0.9
decay=0.0005
angle=0
saturation = 1.5
exposure = 1.5
hue=.1

learning_rate=0.001
burn_in=1000
max_batches = 500200
policy=steps
steps=400000,450000
```

YOLOv3-tiny	COCO trainval	test-dev	33.1	5.56 Bn	220	cfg	weights

ดาวน์โหลดทั้งไฟล์ weights
ลงในโฟลเดอร์ที่สร้างไว้



ดาวน์โหลดทั้งไฟล์ weights
ลงในโฟลเดอร์ที่สร้างไว้

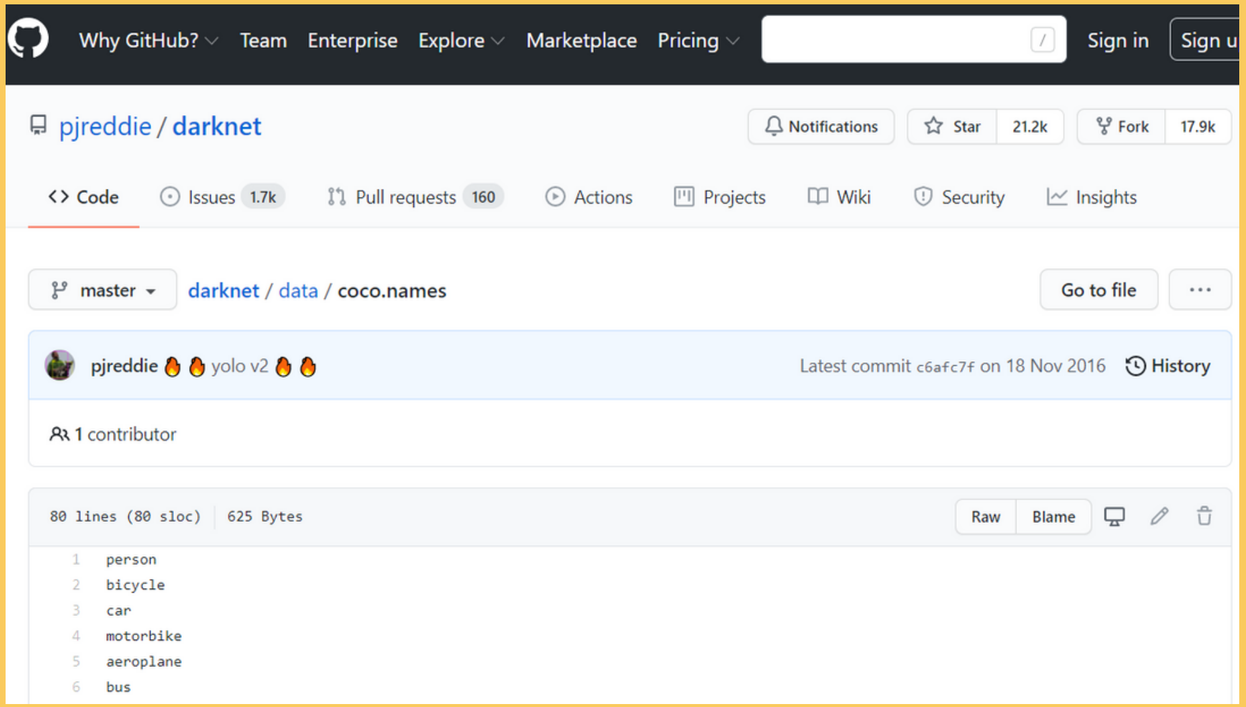


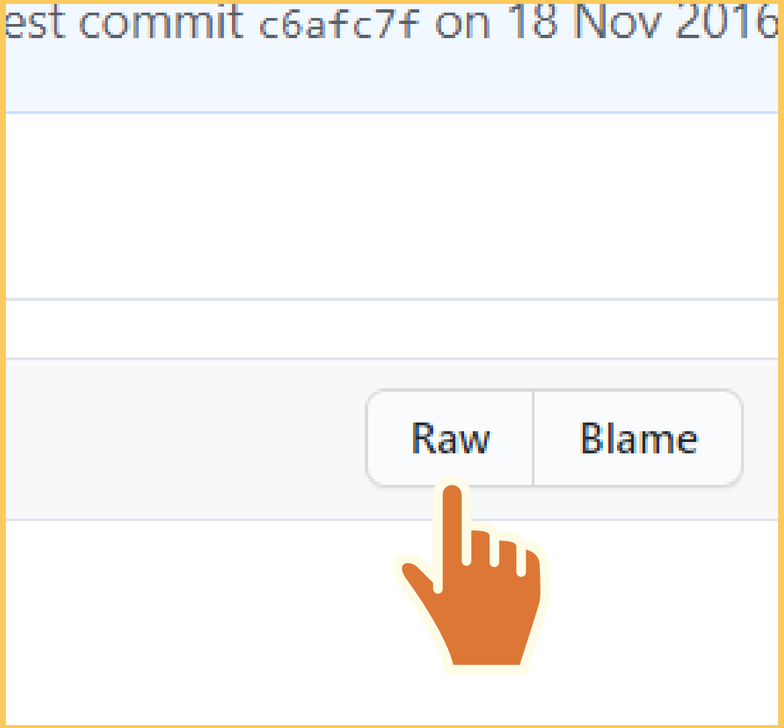
มี 2 ไฟล์ในโฟลเดอร์คือ cfg และ weights

COCO

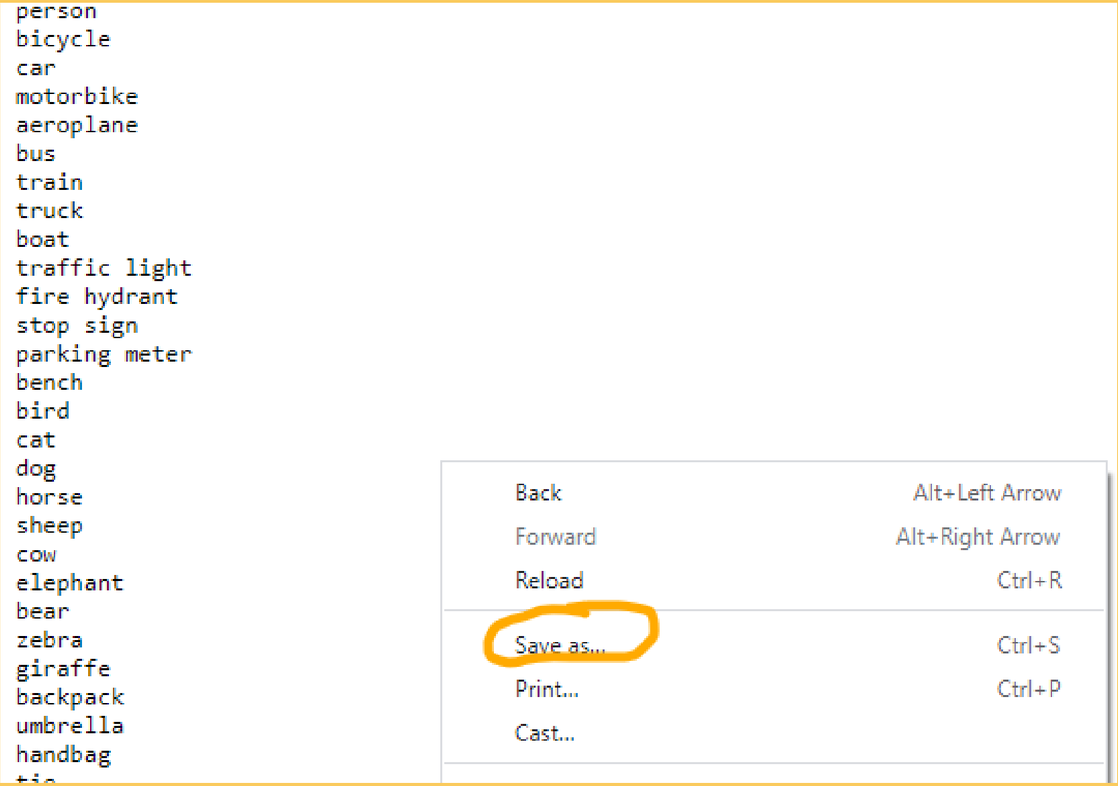
https://github.com/pjreddie/darknet/blob/master/data/coco.names

COCO
Common Objects in Context

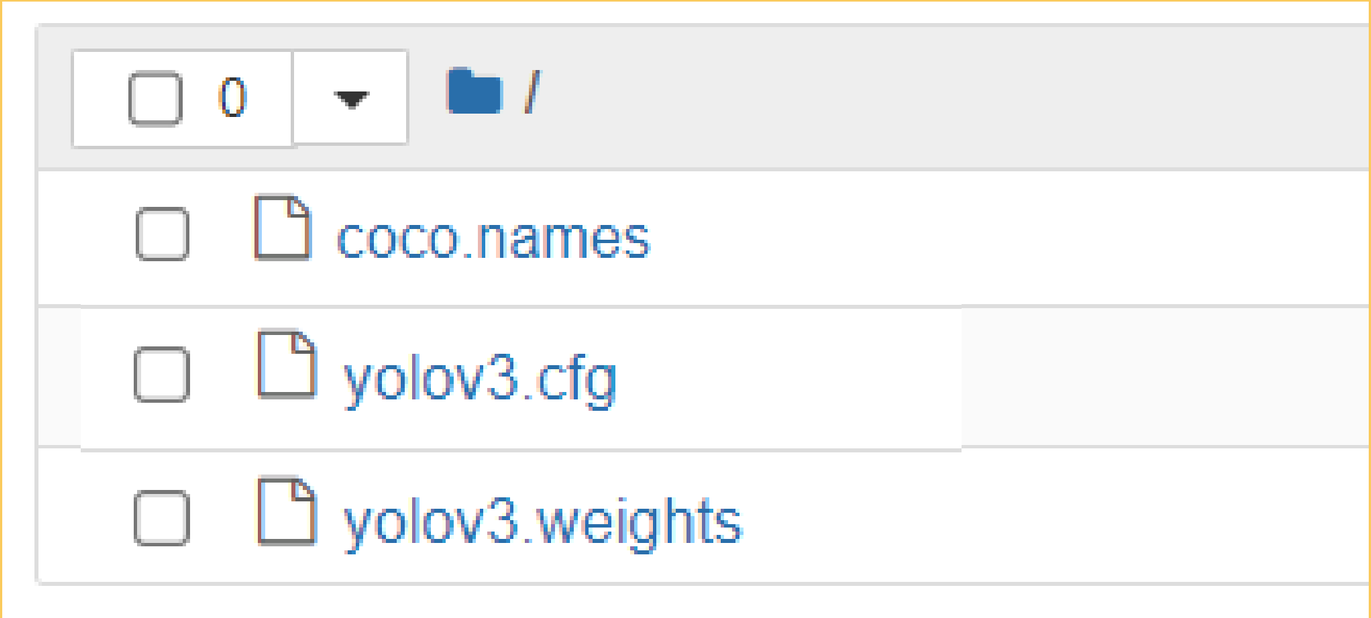

Why GitHub?
Team
Enterprise
Explore
Marketplace
Pricing
Sign in
Sign u
pjreddie / darknet
Notifications
Star 21.2k
Fork 17.9k
Code
Issues 1.7k
Pull requests 160
Actions
Projects
Wiki
Security
Insights
master
darknet / data / coco.names
Go to file
pjreddie yolo v2
Latest commit c6afc7f on 18 Nov 2016
History
1 contributor
80 lines (80 sloc) 625 Bytes
Raw
Blame
person
bicycle
car
motorbike
aeroplane
bus


est commit c6afc7f on 18 Nov 2016
Raw
Blame

กดปุ่ม Raw



ทำการคลิกขวาเพื่อบันทึก
บันทึกลงในโฟล์เดอร์งาน



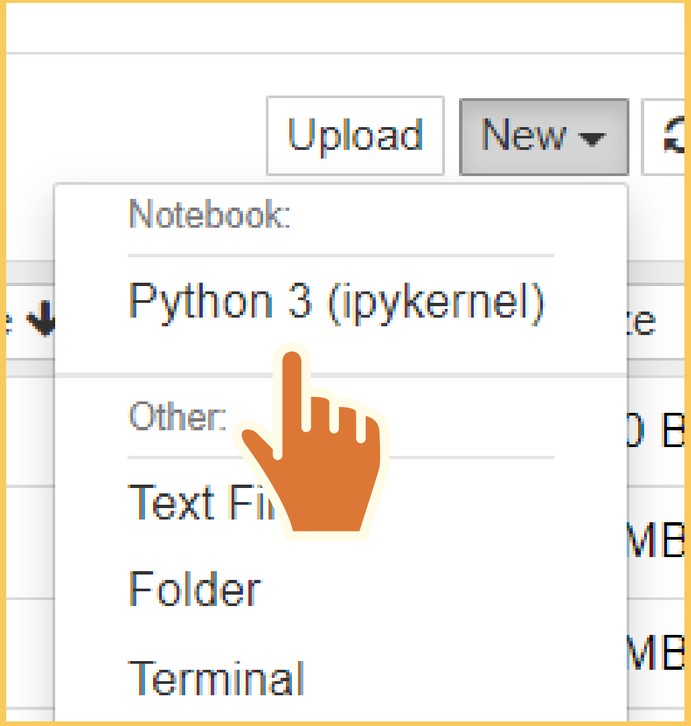
มี 3 ไฟล์ในโฟลเดอร์คือ cfg ,weights และ names

ดาวน์โหลดภาพ

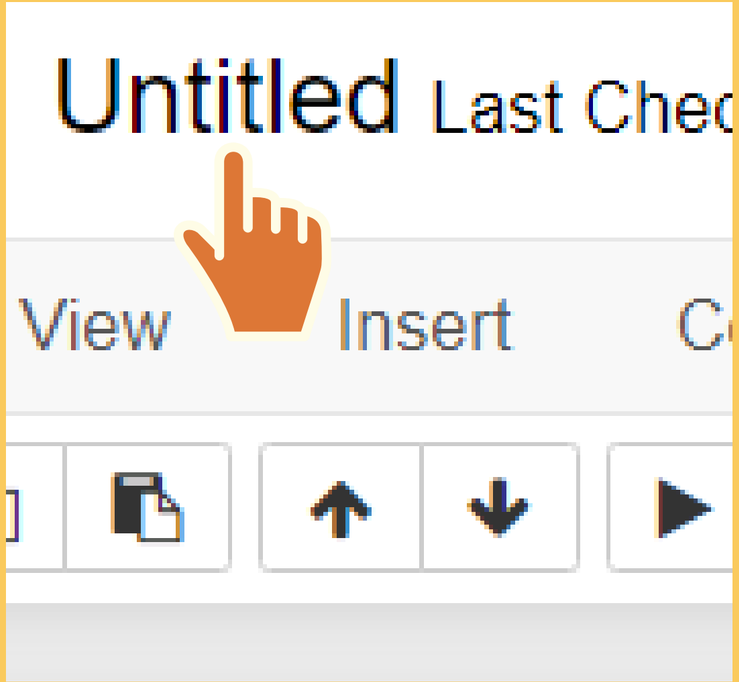
https://bit.ly/2VPF2eo

- coco.names
- example.png
- yolov3-tiny.weights
- yolov3.cfg

มี 4 ไฟล์ในโฟลเดอร์คือ
cfg ,weights ,names และไฟล์ภาพ

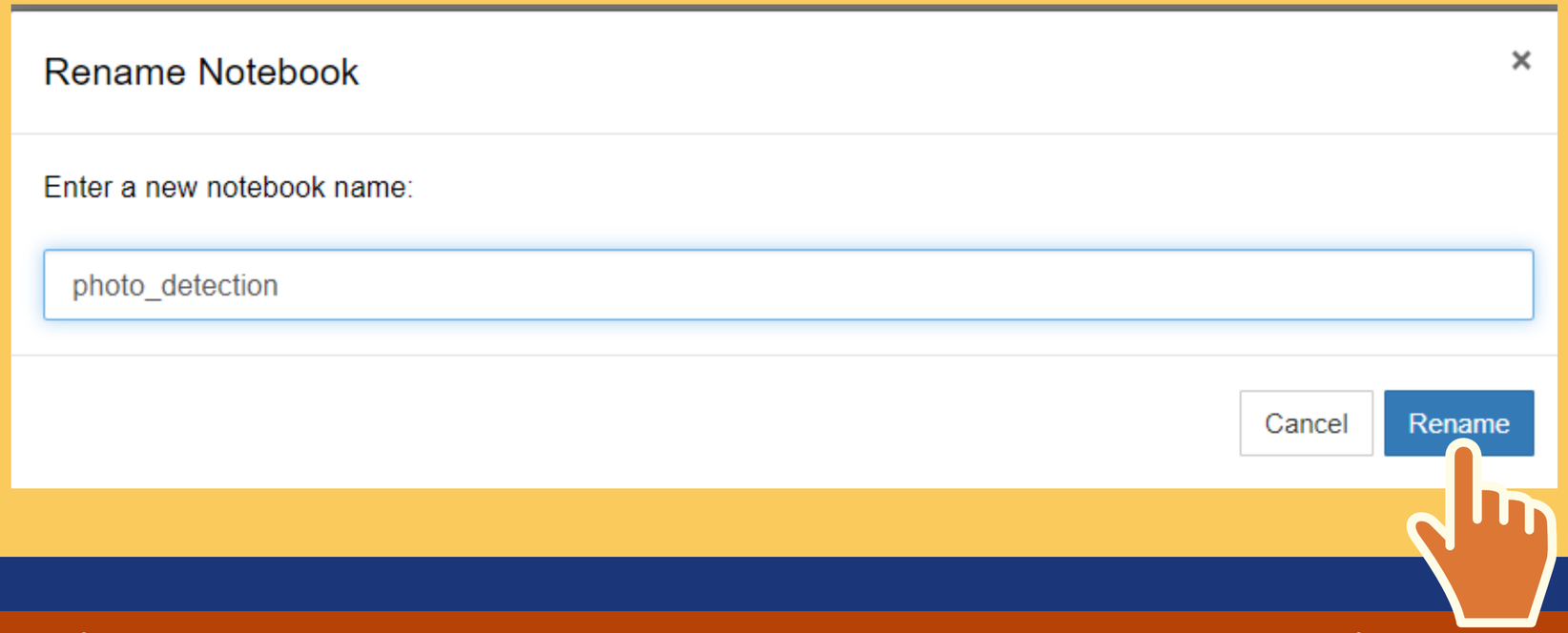


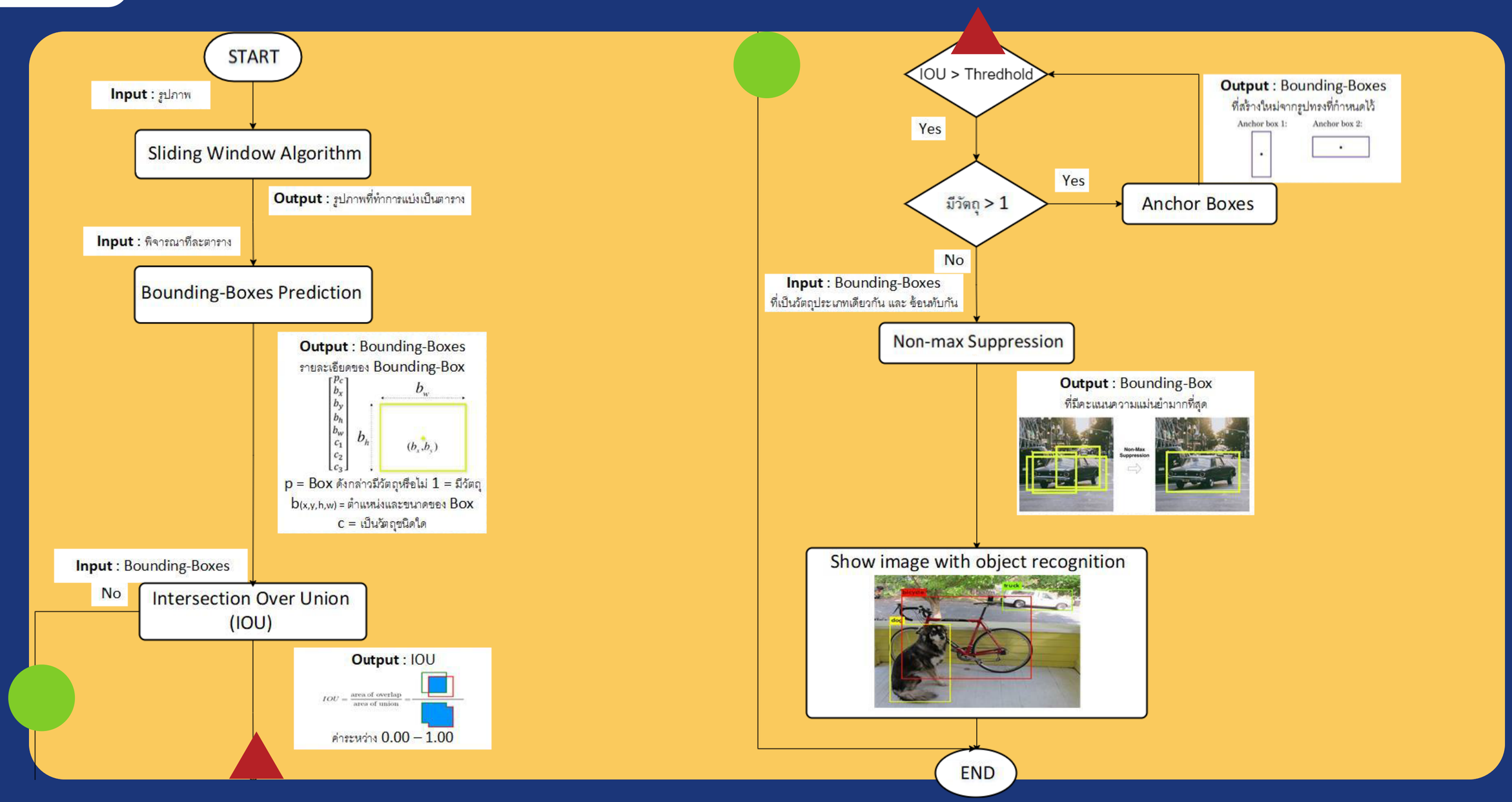
เลือก Python 3



กดเพื่อเปลี่ยนชื่อ

เปลี่ยนชื่อเป็น photo_detection เสร็จสิ้นกดปุ่ม Rename




START
Input : รูปภาพ
Sliding Window Algorithm
Output : รูปภาพที่ทำการแบ่งเป็นตาราง
Input : พิจารณาทีละตาราง
Bounding-Boxes Prediction
Output : Bounding-Boxes
รายละเอียดของ Bounding-Box
p = Box ดังกล่าวมีวัตถุหรือไม่ 1 = มีวัตถุ
b(x,y,h,w) = ตำแหน่งและขนาดของ Box
c = เป็นวัตถุชนิดใด
Input : Bounding-Boxes
No
Intersection Over Union (IOU)
Output : IOU
ค่าระหว่าง 0.00 – 1.00
IOU > Thredhold
Yes
มีวัตถุ > 1
Yes
Anchor Boxes
Output : Bounding-Boxes
ที่สร้างใหม่จากรูปทรงที่กำหนดไว้
Anchor box 1:
Anchor box 2:
No
Input : Bounding-Boxes
ที่เป็นวัตถุประเภทเดียวกัน และ ซ้อนทับกัน
Non-max Suppression
Output : Bounding-Box
ที่มีคะแนนความแม่นยำมากที่สุด
Non-Max Suppression
Show image with object recognition
bicycle
truck
dog
END

สรุป

- Yolo คืออะไร
- กระบวนการทำงาน Yolo
- การติดตั้งเครื่องมือ
    - python
    - Jupiter notebook
    - Library

เริ่มเขียน code กันเถอะ

```
In [2]: import numpy as np
        import time
        import cv2
```

นำเข้า Libraly ต่าง ๆที่ได้ติดตั้งไปก่อนหน้านี้

```
INPUT_FILE='example.png'
LABELS_FILE='coco.names'
CONFIG_FILE='yolov3.cfg'
WEIGHTS_FILE='yolov3-tiny.weights'
CONFIDENCE_THRESHOLD=0.3
```

INPUT_FILE คือ ไฟล์ที่ต้องการนำมา Detection

LABELS_FILE คือ ไฟล์ข้อมูลแอตทริบิวต์ต่าง ๆ ของวัตถุ

CONFIG_FILE คือ ไฟล์กำหนดค่าของโมเดล YOLO

WEIGHT_FILE คือ ไฟล์ที่ผ่านการ train model YOLO

CONFIDENCE_THRESHOLD คือ เกณฑ์ความน่าจะเป็นวัตถุในภาพ

```
LABELS = open(LABELS_FILE).read().strip().split("\n")
```

คำสั่งให้นำ LABEL FILE ซึ่งตอนนี้เป็น COCO.NAMES ให้ระบบนำไปแกะออกมานำไปบรรจุใน LABEL โดยให้เรียงเป็นบรรทัดเอาไว้ (มองไม่เห็น)

```
np.random.seed(4)
COLORS = np.random.randint(0, 255, size=(len(LABELS), 3),
    dtype="uint8")
```

คำสั่งให้สร้างค่า seed ที่ถูกสุ่มขึ้นมาจากนั้นนำไปสุ่มสีที่แตกต่างกัน สำหรับนำไปแสดงเป็นกรอบรอบชิ้นวัตถุที่ถูก Detection

```
net = cv2.dnn.readNetFromDarknet(CONFIG_FILE, WEIGHTS_FILE)
```

คำสั่งให้สร้างตัวแปร net ขึ้นมาเพื่อนำค่า CONFIG และ WEIGHT ส่งไปยังเซิร์ฟเวอร์ของ Darknet เพื่อนำมาคำนวณในโปรแกรมต่อไป

```
image = cv2.imread(INPUT_FILE)
(H, W) = image.shape[:2]
```

คำสั่งสร้างตัวแปร image ขึ้นมาเพื่อค่าของรูปภาพตัวอย่างที่นำเข้ามา

คำสั่งสร้างตัวแปร (H, W) เพื่อเก็บความสูงและความกว้างของภาพไว้

```
ln = net.getLayerNames()
ln = [ln[i[0] - 1] for i in net.getUnconnectedOutLayers()]
```

คำสั่งกำหนดให้ Output เฉพาะส่วนที่ต้องการจาก YOLO

```
blob = cv2.dnn.blobFromImage(image, 1 / 255.0, (416, 416),
    swapRB=True, crop=False)
net.setInput(blob)
start = time.time()
layerOutputs = net.forward(ln)
end = time.time()
```

สร้างตัวแปร blob นำไปลดขนาดรูปภาพที่นำเข้ามาทดสอบให้เหลือ 416*416 ให้สลับสีฟ้ากับน้ำเงิน ให้ส่งค่าไป net.setInput เริ่มนับเวลาเพื่อที่จะคำนวณว่า YOLO ใช้เวลาในการคำนวณไปเท่าไหร่

```
print("[INFO] YOLO took {:.6f} seconds".format(end - start))
```

แสดงค่าเวลาที่ถูกใช้ไปในการคำนวณ YOLO ขึ้นมาในช่อง OUTPUT

```
boxes = []
confidences = []
classIDs = []
```

ประการตัวแปร Array ทั้งหมด 3 ตัวคือ boxes, confidences, classIDs

```
for output in layerOutputs:
```

คำสั่ง วนซ้ำแต่ละเอาต์พุตของเลเยอร์

```
for detection in output:
```

คำสั่ง วนซ้ำการตรวจจับแต่ละครั้ง

```
scores = detection[5:]
classID = np.argmax(scores)
confidence = scores[classID]
```

คำสั่ง แยก class ID และ ค่า confidence (เช่น ความน่าจะเป็น) ของการตรวจจับวัตถุปัจจุบัน

```
if confidence > CONFIDENCE_THRESHOLD:
```

หากค่า confidence ที่คาดการณ์ว่าตรวจจับมีค่ามากกว่าเกณฑ์ความน่าจะเป็นวัตถุในภาพ

```
box = detection[0:4] * np.array([W, H, W, H])
(centerX, centerY, width, height) = box.astype("int")
```

ปรับขนาด bounding box รูปโดยสัมพันธ์กับขนาดของรูปภาพ YOLO จะส่งตำแหน่งตรงกลางbounding box (x, y) ของกรอบ ล้อมรอบตามด้วยความกว้างและความสูงของกล่อง

```
x = int(centerX - (width / 2))
y = int(centerY - (height / 2))
```

ใช้จุดตำแหน่งตรงกลาง bounding box (x, y) - ที่มีการ detection เพื่อหามุมด้านบนและด้านซ้ายของ bounding box

```
boxes.append([x, y, int(width), int(height)])
confidences.append(float(confidence))
classIDs.append(classID)
```

อัปเดตรายการ bounding box confidence และ classID

```
idxs = cv2.dnn.NMSBoxes(boxes, confidences, CONFIDENCE_THRESHOLD,
    CONFIDENCE_THRESHOLD)
```

คำสั่งสร้างตัวแปร idxs เก็บค่า boxes, confidence, CONFIDENCE_THRESHOLD, CONFIDENCE_THRESHOLD

```
if len(idxs) > 0:
```

คำสั่งถ้ามีการตรวจสอบว่ามีการตรวจพบ idxs อย่างน้อยหนึ่งรายการ

```
for i in idxs.flatten():
```

คำสั่ง วนรอบซ้ำดัชนีที่เราเก็บไว้ใน idxs

```
(x, y) = (boxes[i][0], boxes[i][1])
(w, h) = (boxes[i][2], boxes[i][3])
```

แยกค่า bounding box กล่องขอบเขต

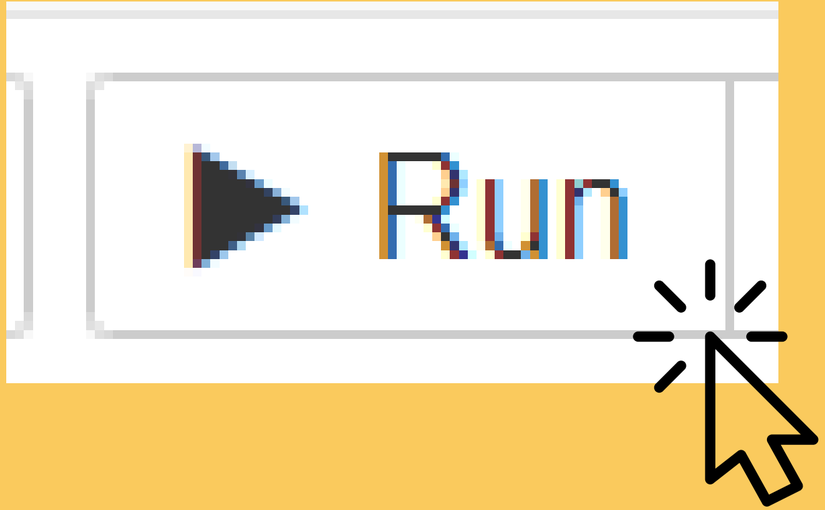
```
color = [int(c) for c in COLORS[classIDs[i]]]

cv2.rectangle(image, (x, y), (x + w, y + h), color, 2)
text = "{}: {:.4f}".format(LABELS[classIDs[i]], confidences[i])
cv2.putText(image, text, (x, y - 5), cv2.FONT_HERSHEY_SIMPLEX
    0.5, color, 2)
```

****

```
cv2.imshow('Image',image)
cv2.waitKey(0)
cv2.destroyAllWindows()
```

คำสั่ง แสดง output image

ลองกดทดสอบกัน



ภาพ ที่ผ่านการ detection

Workshop 1 ❤

ให้ผู้อบรมเลือกภาพในโฟลเดอร์ตาม
ลิ้งค์ด้านล่าง จากนั้นไปแก้โค๊ดส่วนนึง
เพื่อให้ YOLO ทำการ Detection

https://bit.ly/3iPakL6

Workshop 1 ❤

ผู้เข้าอบรมต้องส่งงานโดยการ Print Screen จากนั้น Save ส่งภาพมาตามฟอร์มลิ้งค์ด้านล่าง

https://bit.ly/3iLl0Kz

Workshop 2 ❤

ให้ผู้อบรมเลือกวัตถุที่อยู่ใกล้ๆตัวมา
เพื่อให้ YOLO ทำการ Detectionโดย
ให้หาวัตถุมาไว้ในภาพมากกว่า2ชิ้น

Workshop 2 ❤

ผู้เข้าอบรมต้องส่งงานโดยการ Print Screen จากนั้น Save ส่งภาพมาตามฟอร์มลิ้งค์ด้านล่าง

https://bit.ly/39tMHSI

IMAGE RECOGNITION WITH PYTHON

Computer Education

LECTURER

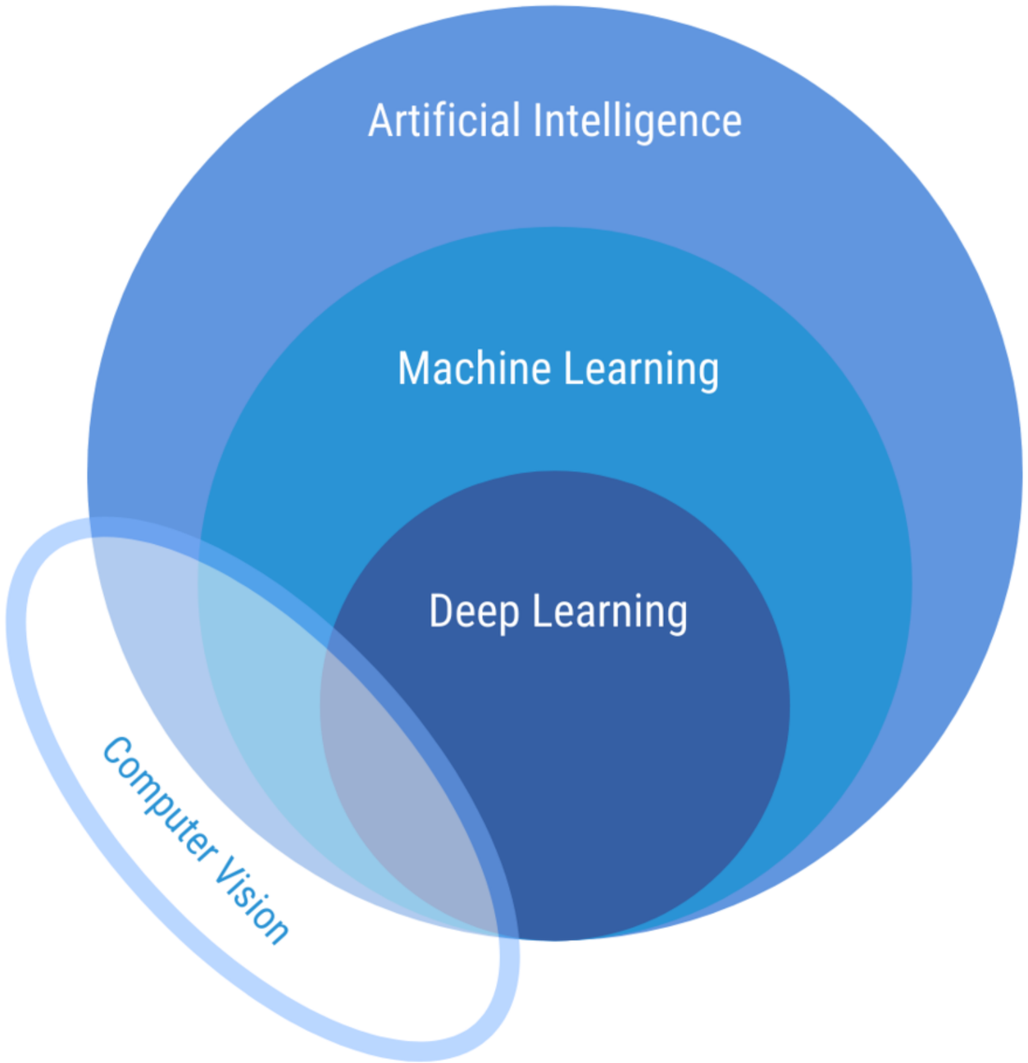
JIRAWAT CHUAYKHUN

4th year Student

LECTURER

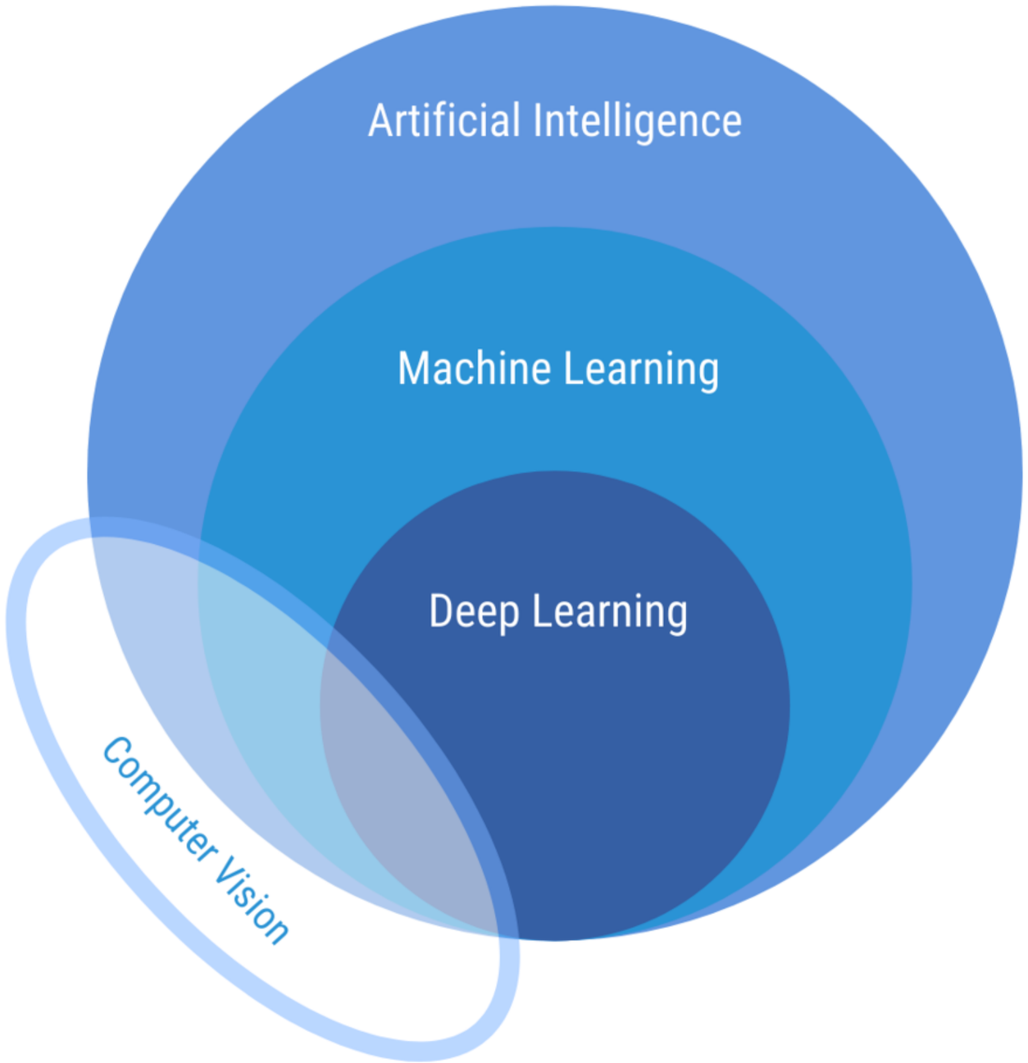
POONPIPAT KONGTHONG

4th year Student



ปัญญาประดิษฐ์ หรือ (AI)

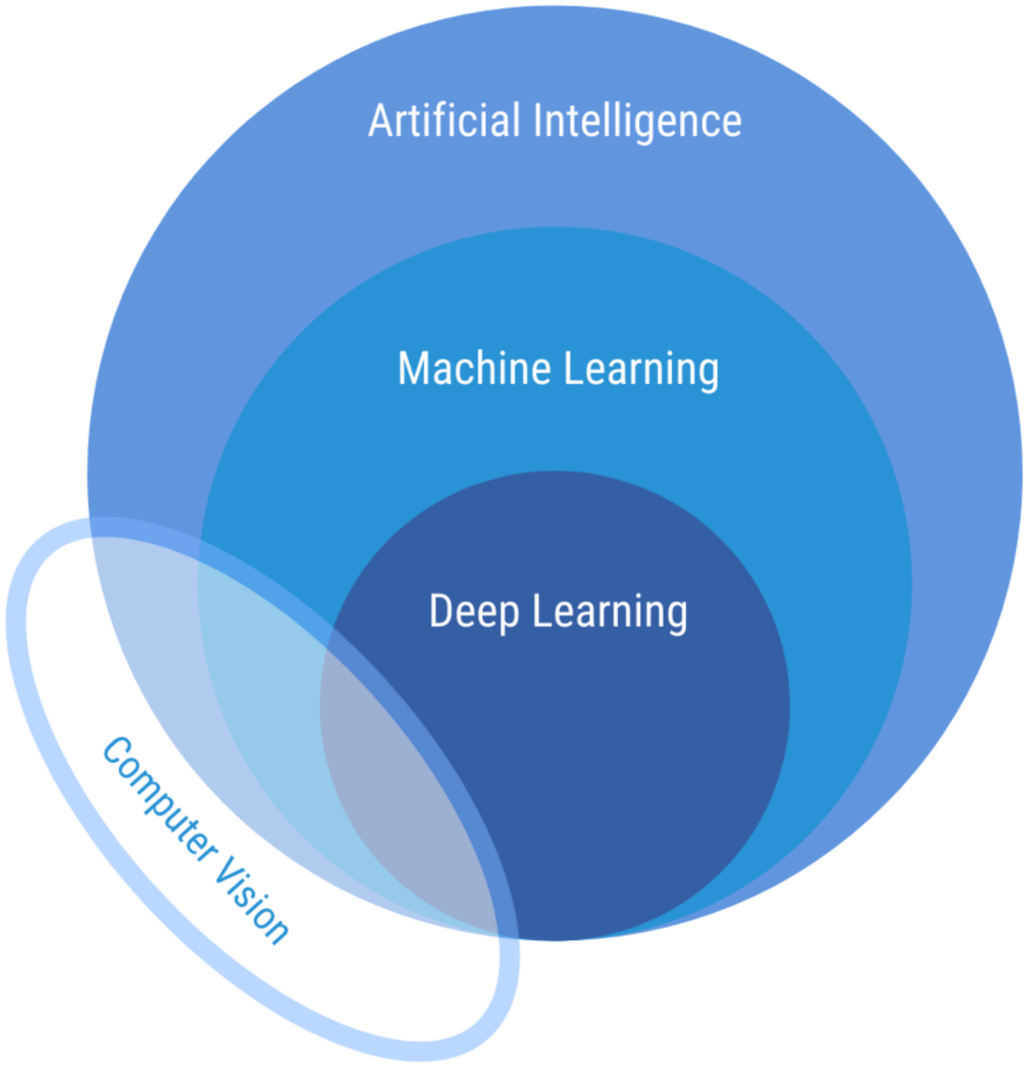
การรวมความฉลาดของมนุษย์สู่เครื่องจักร (Machine)



Machine Learning (ML)

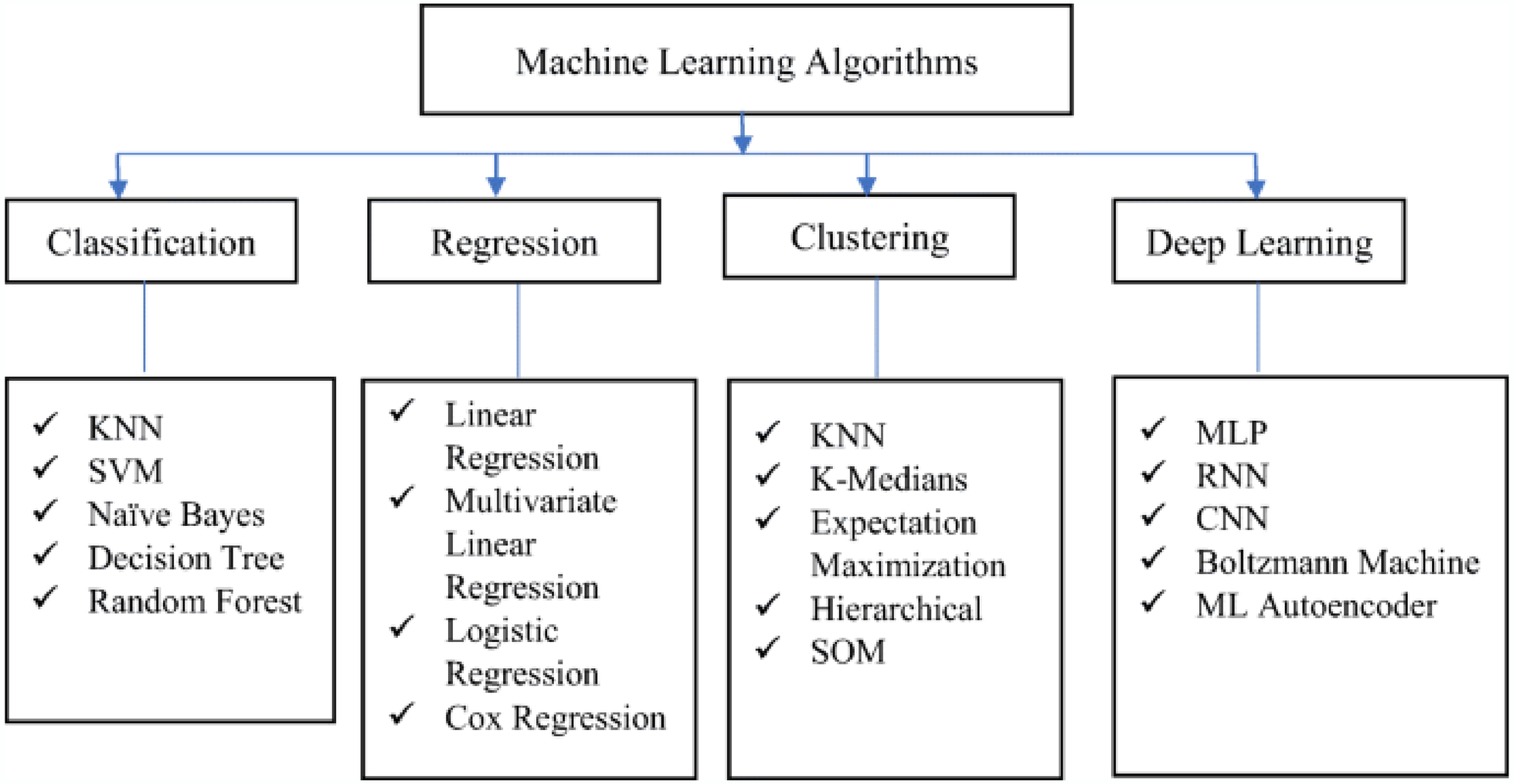
การสอนให้ระบบคอมพิวเตอร์ทำการเรียนรู้ได้ด้วยตนเองโดยการใช้

'ข้อมูล'



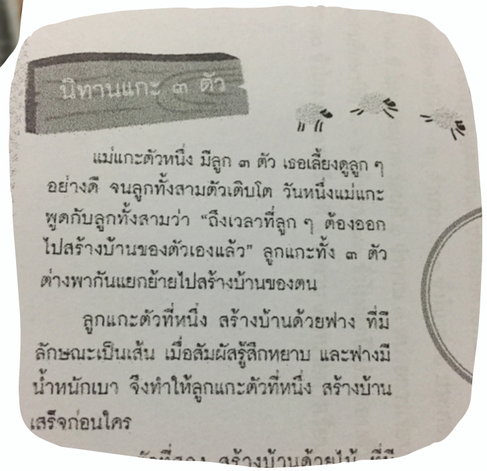
Deep Learning

อัลกอริทึมแบบระบบเรียนรู้เชิงลึก การเรียนรู้แบบอัตโนมัติด้วยการ เลียนแบบการทำงานของโครงข่ายประสาทของมนุษย์


Machine Learning Algorithms
Classification
Regression
Clustering
Deep Learning
✓ KNN
✓ SVM
✓ Naïve Bayes
✓ Decision Tree
✓ Random Forest
✓ Linear Regression
✓ Multivariate Linear Regression
✓ Logistic Regression
✓ Cox Regression
✓ KNN
✓ K-Medians
✓ Expectation Maximization
✓ Hierarchical
✓ SOM
✓ MLP
✓ RNN
✓ CNN
✓ Boltzmann Machine
✓ ML Autoencoder

Optical Character Recognition

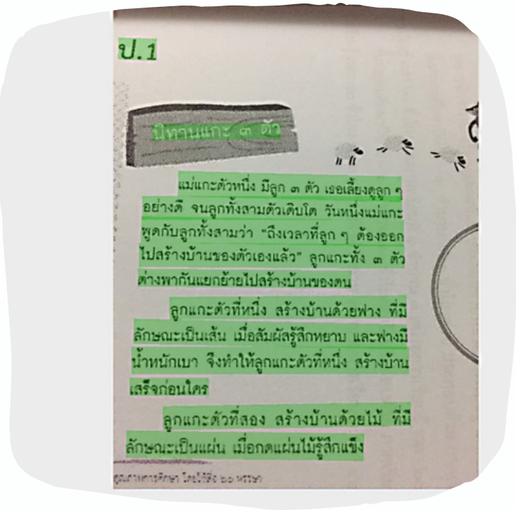
การรู้จำอักขระด้วยแสง



นิทานแกะ ๓ ตัว

แม่แกะตัวหนึ่ง มีลูก ๓ ตัว เธอเลี้ยงดูลูก ๆ อย่างดี จนลูกทั้งสามตัวเติบโต วันหนึ่งแม่แกะพูดกับลูกทั้งสามว่า "ถึงเวลาที่ลูก ๆ ต้องออกไปสร้างบ้านของตัวเองแล้ว" ลูกแกะทั้ง ๓ ตัว ต่างพากันแยกย้ายไปสร้างบ้านของตน

ลูกแกะตัวที่หนึ่ง สร้างบ้านด้วยฟาง ที่มีลักษณะเป็นเส้น เมื่อสัมผัสรู้สึกหยาบ และฟางมีน้ำหนักเบา จึงทำให้ลูกแกะตัวที่หนึ่ง สร้างบ้านเสร็จก่อนใคร



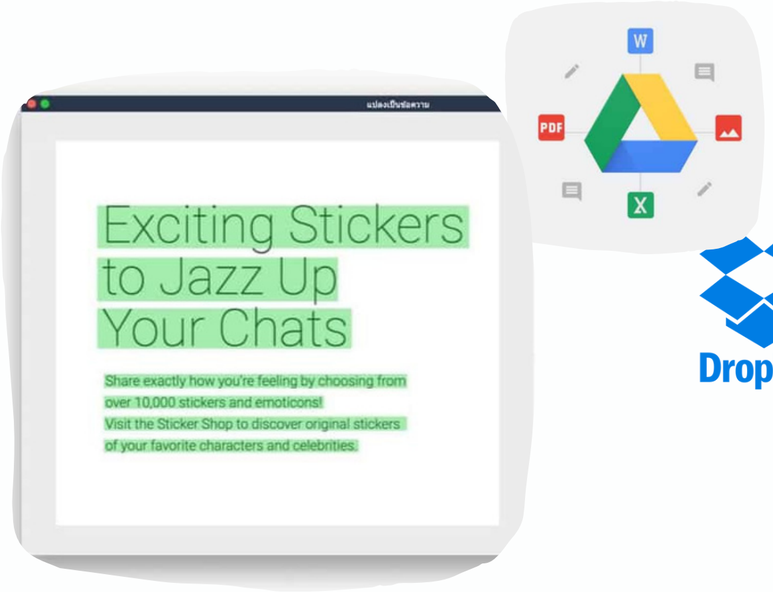
ป.1

นิทานแกะ ๓ ตัว

แม่แกะตัวหนึ่ง มีลูก ๓ ตัว เธอเลี้ยงดูลูก ๆ อย่างดี จนลูกทั้งสามตัวเติบโต วันหนึ่งแม่แกะพูดกับลูกทั้งสามว่า "ถึงเวลาที่ลูก ๆ ต้องออกไปสร้างบ้านของตัวเองแล้ว" ลูกแกะทั้ง ๓ ตัว ต่างพากันแยกย้ายไปสร้างบ้านของตน

ลูกแกะตัวที่หนึ่ง สร้างบ้านด้วยฟาง ที่มีลักษณะเป็นเส้น เมื่อสัมผัสรู้สึกหยาบ และฟางมีน้ำหนักเบา จึงทำให้ลูกแกะตัวที่หนึ่ง สร้างบ้านเสร็จก่อนใคร

ลูกแกะตัวที่สอง สร้างบ้านด้วยไม้ ที่มีลักษณะเป็นแผ่น เมื่อกดแผ่นไม้รู้สึกแข็ง



1 เทคนิคการรู้จำตัวอักษรด้วยแสง คือ ?

2 ประเภทของเทคนิคการรู้จำตัวอักษรด้วยแสง

3 ขั้นตอนการประมวลผล

4 เครื่องมือ OCR Open source

เทคนิคการรู้จำตัวอักษรด้วยแสง คือ ?

สแกนเอกสาร

เอกสาร pdf

รูปภาพ

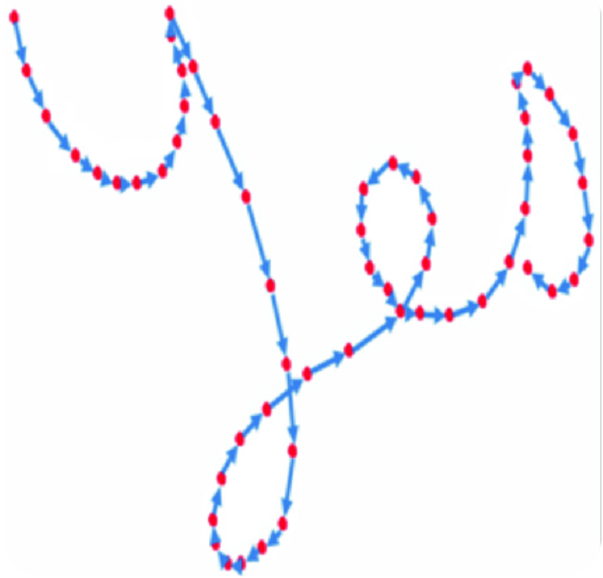
ประเภทของเทคนิค การรู้จำตัวอักษรด้วยแสง

การรู้จำตัวอักษร
ด้วยแสง

การรู้จำตัวอักษร
แบบออนไลน์

การรู้จำตัวอักษร
แบบออฟไลน์

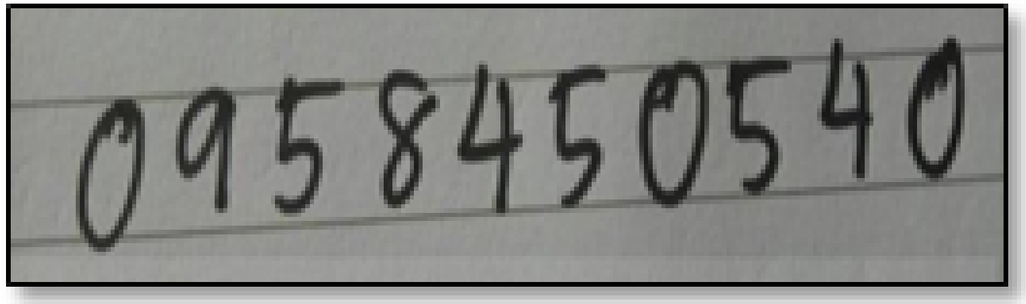
การรู้จำตัวอักษรแบบออนไลน์



การเรียนรู้ตัวอักษรแบบออฟไลน์

ตัวอักษรแบบโดด (Single Character)

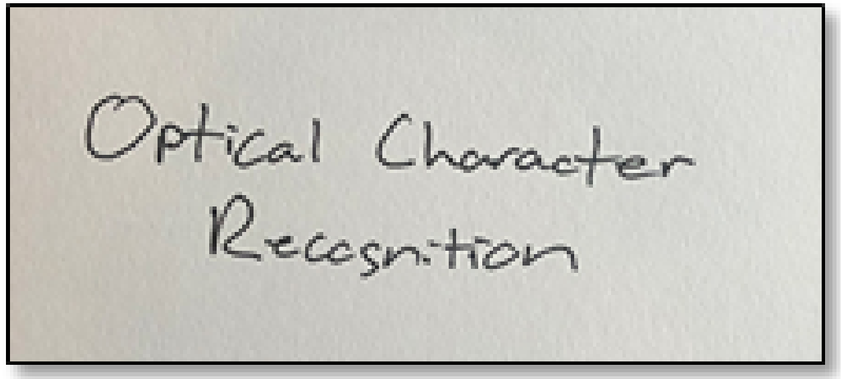
ตัวพิมพ์แบบฟอนต์เฉพาะ



ลายมือเขียนแบบตัวโดด

การเรียนรู้ตัวอักษรแบบออฟไลน์

ตัวอักษรลายมือแบบเขียนต่อเนื่อง (Script recognition)



ขั้นตอนการประมวลผลของเทคนิคการรู้จำตัวอักษรด้วยแสง

1. ขบวนการประมวลผลขั้นต้น (Pre-Processing)
2. การรู้จำ (Recognition)
3. ขบวนการประมวลผลขั้นสุดท้าย (Post-Processing)

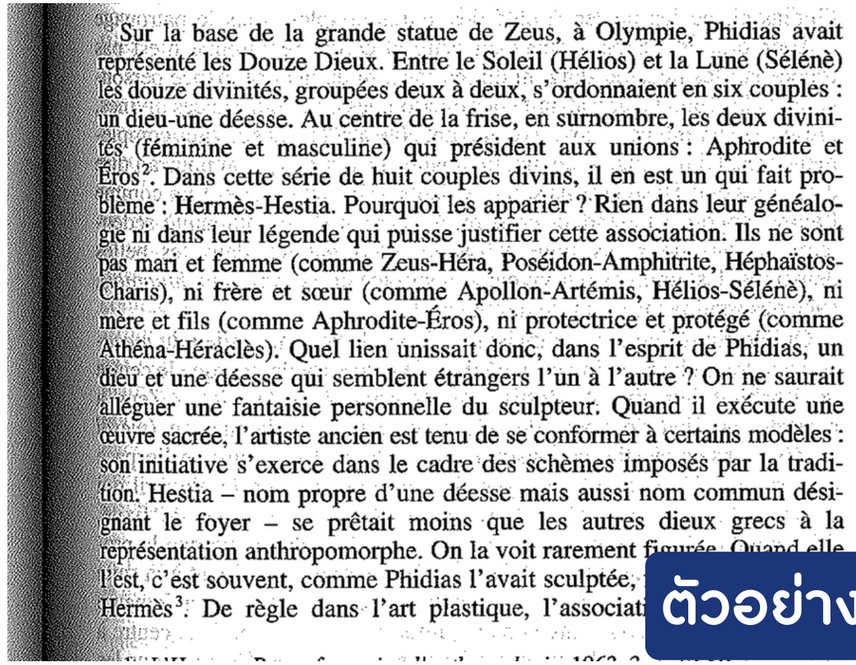
1 ขบวนการประมวลผลขั้นต้น (Pre-Processing)

การลดสัญญาณรบกวน (Noise Filtering)

เป็นการกรองข้อมูลภาพที่มีสิ่งแทรกซ้อนที่ไม่ต้องการ
โดยมีจุดประสงค์เพื่อลดความผิดพลาด

1 ขบวนการประมวลผลขั้นต้น (Pre-Processing)

Noise คือ จุดรบกวนในภาพที่เกิดจาก การปรับเพิ่มค่าความไวแสง



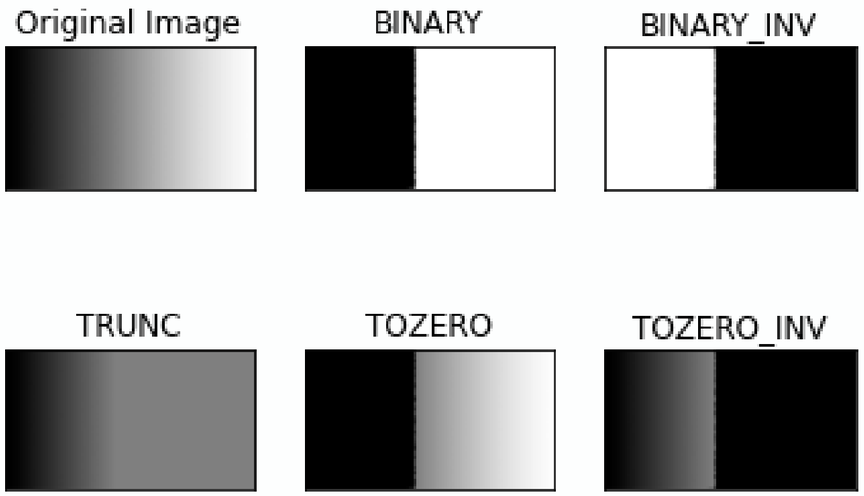
Sur la base de la grande statue de Zeus, à Olympie, Phidias avait représenté les Douze Dieux. Entre le Soleil (Hélios) et la Lune (Sélénè) les douze divinités, groupées deux à deux, s'ordonnaient en six couples : un dieu-une déesse. Au centre de la frise, en surnombre, les deux divinités (féminine et masculine) qui président aux unions : Aphrodite et Éros[2]. Dans cette série de huit couples divins, il en est un qui fait problème : Hermès-Hestia. Pourquoi les apparier ? Rien dans leur généalogie ni dans leur légende qui puisse justifier cette association. Ils ne sont pas mari et femme (comme Zeus-Héra, Poséidon-Amphitrite, Héphaïstos-Charis), ni frère et sœur (comme Apollon-Artémis, Hélios-Sélénè), ni mère et fils (comme Aphrodite-Éros), ni protectrice et protégé (comme Athéna-Héraclès). Quel lien unissait donc, dans l'esprit de Phidias, un dieu et une déesse qui semblent étrangers l'un à l'autre ? On ne saurait alléguer une fantaisie personnelle du sculpteur. Quand il exécute une œuvre sacrée, l'artiste ancien est tenu de se conformer à certains modèles : son initiative s'exerce dans le cadre des schèmes imposés par la tradition. Hestia – nom propre d'une déesse mais aussi nom commun désignant le foyer – se prêtait moins que les autres dieux grecs à la représentation anthropomorphe. On la voit rarement figurée. Quand elle l'est, c'est souvent, comme Phidias l'avait sculptée, [illegible] Hermès[3]. De règle dans l'art plastique, l'associati[illegible]

ตัวอย่างภาพที่เกิดจุดรบกวน

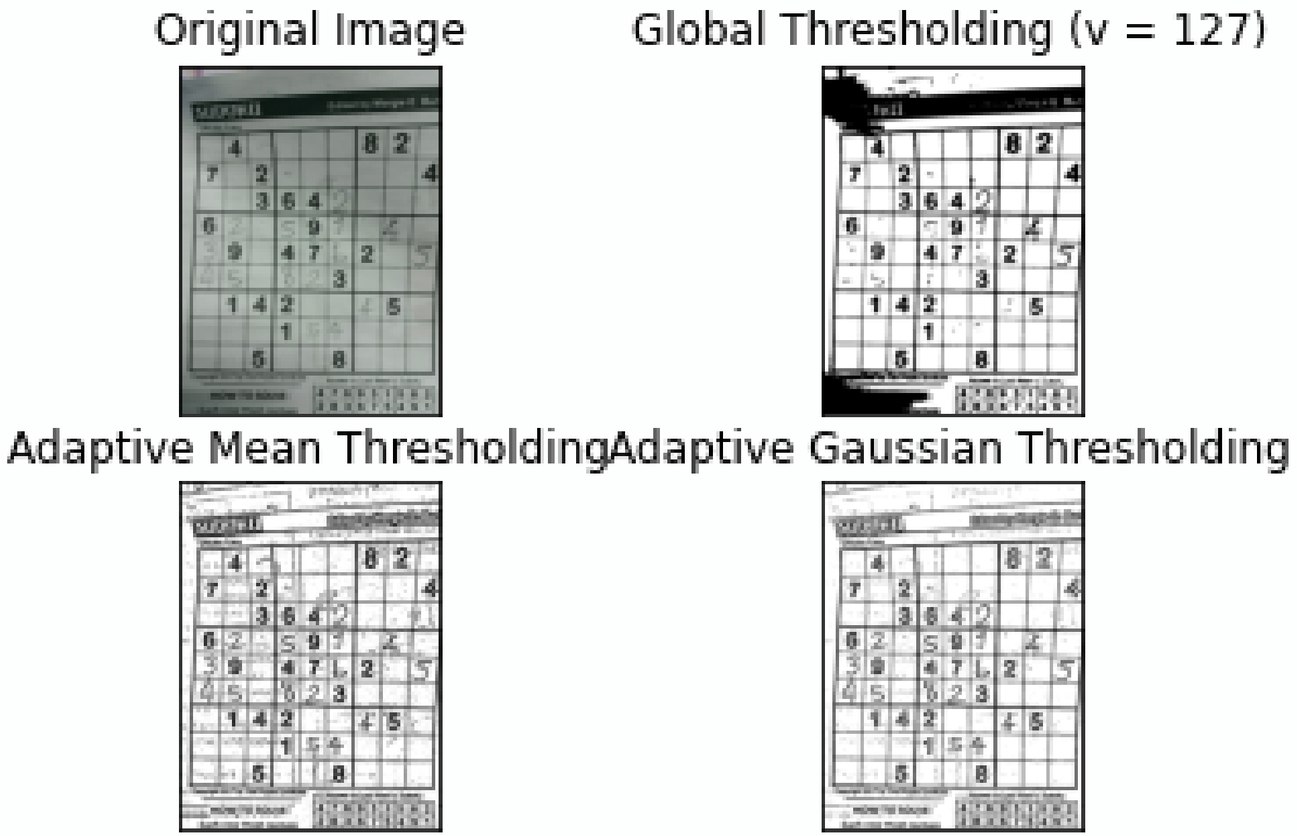
1 ขบวนการประมวลผลขั้นต้น (Pre-Processing)

ฟังก์ชัน cv2.threshold()

ใช้คัดกรองภาพโดยพิจารณาความสว่าง เพื่อที่จะคัดเอาหรือปรับค่าบางส่วน



1 ขบวนการประมวลผลขั้นต้น (Pre-Processing)



1 ขบวนการประมวลผลขั้นต้น (Pre-Processing)

Original

Global Thresholding

Otsu_s_Thresholding

Gaussian + Otsu

1 ขบวนการประมวลผลขั้นต้น (Pre-Processing)

Original

Global Thresholding

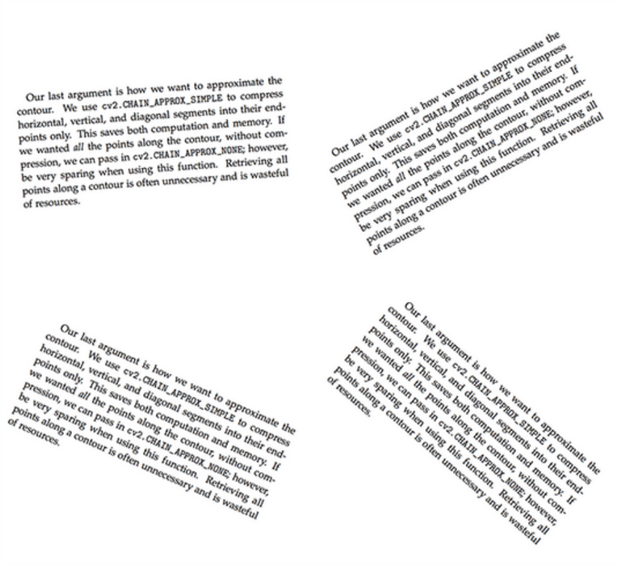
Otsu_s_Thresholding

Gaussian + Otsu

1 ขบวนการประมวลผลขั้นต้น (Pre-Processing)

Skew correction

ตรวจหาค่าความเอียงและการหมุนกลับของภาพเอกสาร



1 ขบวนการประมวลผลขั้นต้น (Pre-Processing)

WELLS FARGO BANK
Transaction Record

Branch # 3793 09 Mortgage Payment

Account Number 6855174292
528 MR

Cash In $0.00
Number of checks 1
$1,500.00
Total Payment $1,500.00

Regular Payment $1,500.00

Transaction # 027 0091
2:22PM 05/18/01 Credited: 05/18/01

Payments may take 1 to 3 business days to post, but will be credited to your account as of the Credited Date on your receipt.
Please retain receipt as verification.

Thank you, Janet

WELLS FARGO BANK
Transaction Record

Branch # 3793 09 Mortgage Payment

Account Number 6855174292
528 MR

Cash In $0.00
Number of checks 1
$1,500.00
Total Payment $1,500.00

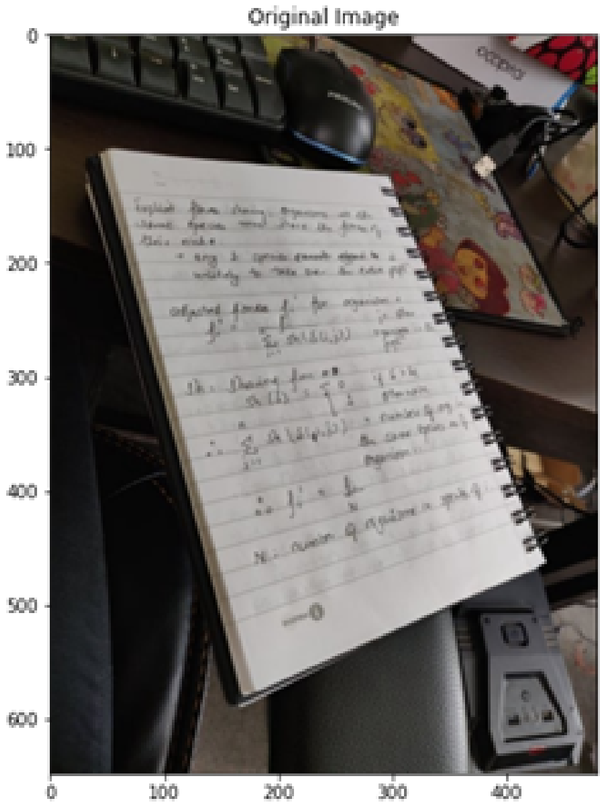
Regular Payment $1,500.00

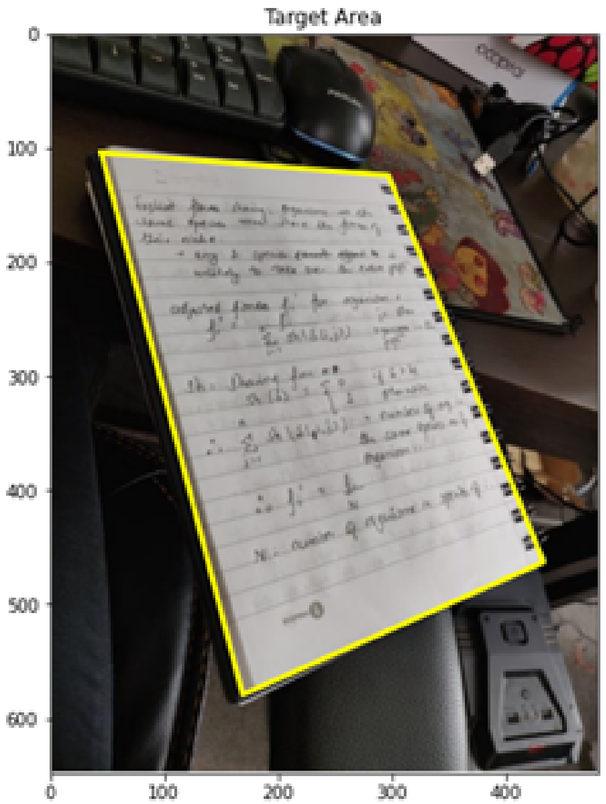
Transaction # 027 0091
2:22PM 05/18/01 Credited: 05/18/01

Payments may take 1 to 3 business days to post, but will be credited to your account as of the Credited Date on your receipt.
Please retain receipt as verification.

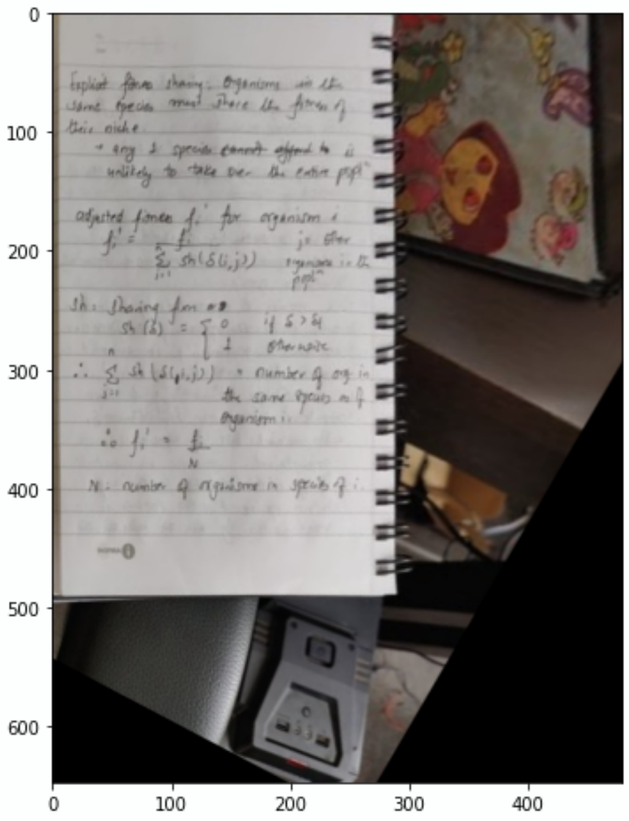
Thank you, Janet

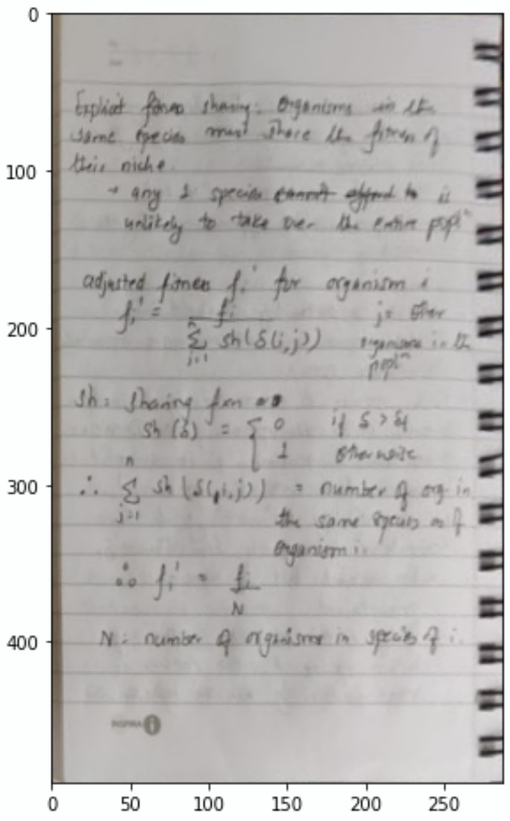
1 ขบวนการประมวลผลขั้นต้น (Pre-Processing)





1 ขบวนการประมวลผลขั้นต้น (Pre-Processing)





2 การรู้จำ (Recognition)

1
วิธีการวิเคราะห์ทางโครงสร้าง
(Structural Analysis)

2
แกรมมาแมชชีน
(formal grammar machine)

3
วิธีการรู้จำด้วยการเปรียบเทียบคู่รูปแบบ
(Template Matching)

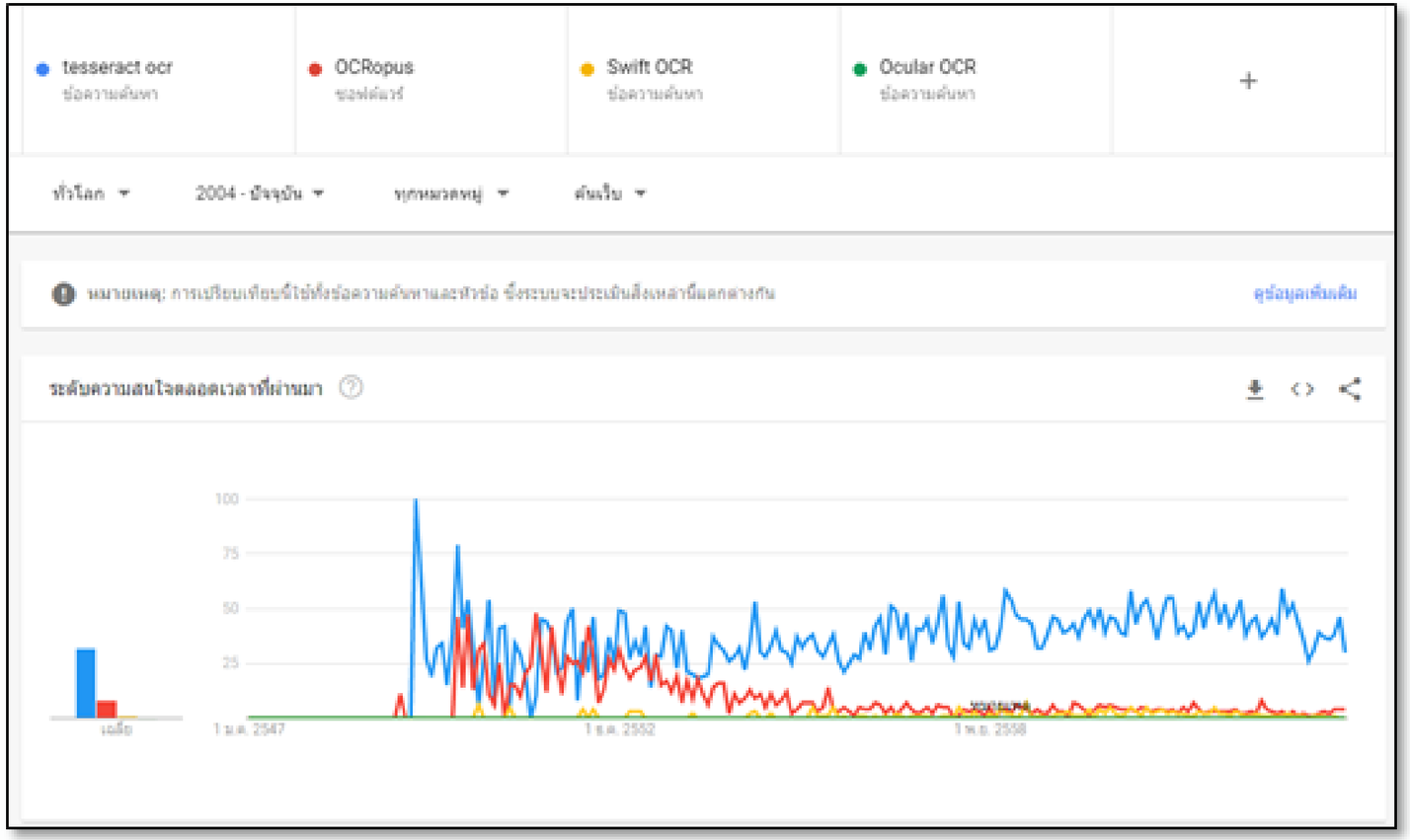
4
โครงข่ายประสาทเทียม
(Neural Network)

3 ขบวนการประมวลผลขั้นสุดท้าย (Post-Processing)

การประมวลผลขั้นตอนสุดท้าย คือ
การตรวจสอบและแก้ไขผลจากการรู้จำรูปตัวอักษร
ที่ผ่านกระบวนการรู้จำมีผลเป็นรหัสตัวอักษร

เช่น ไวยากรณ์ภาษา

เครื่องมือ OCR Open source



เครื่องมือ OCR Open source Tools

1

Tesseract OCR

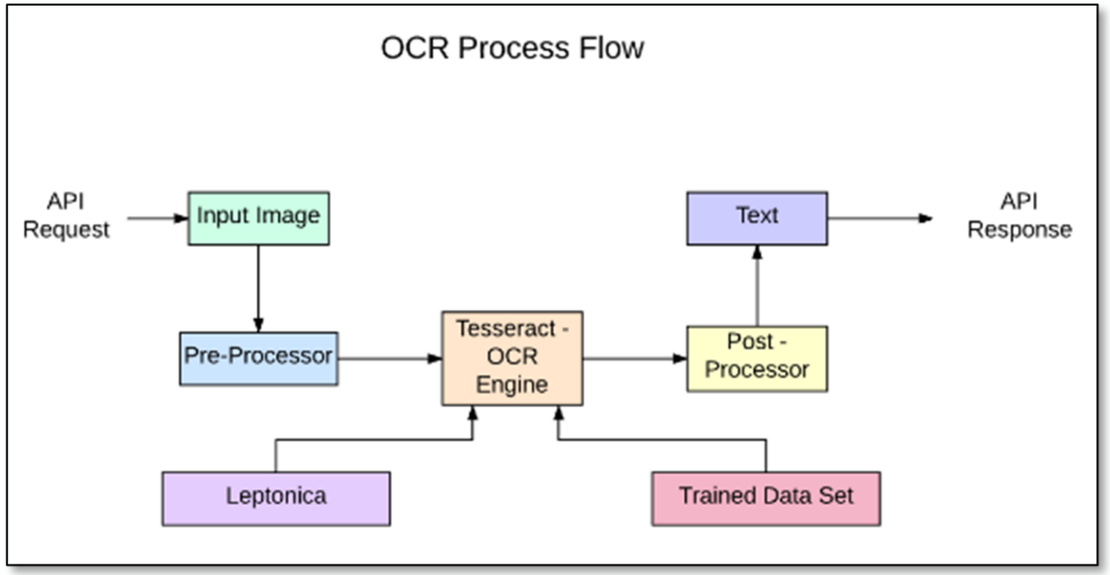
| Python

2

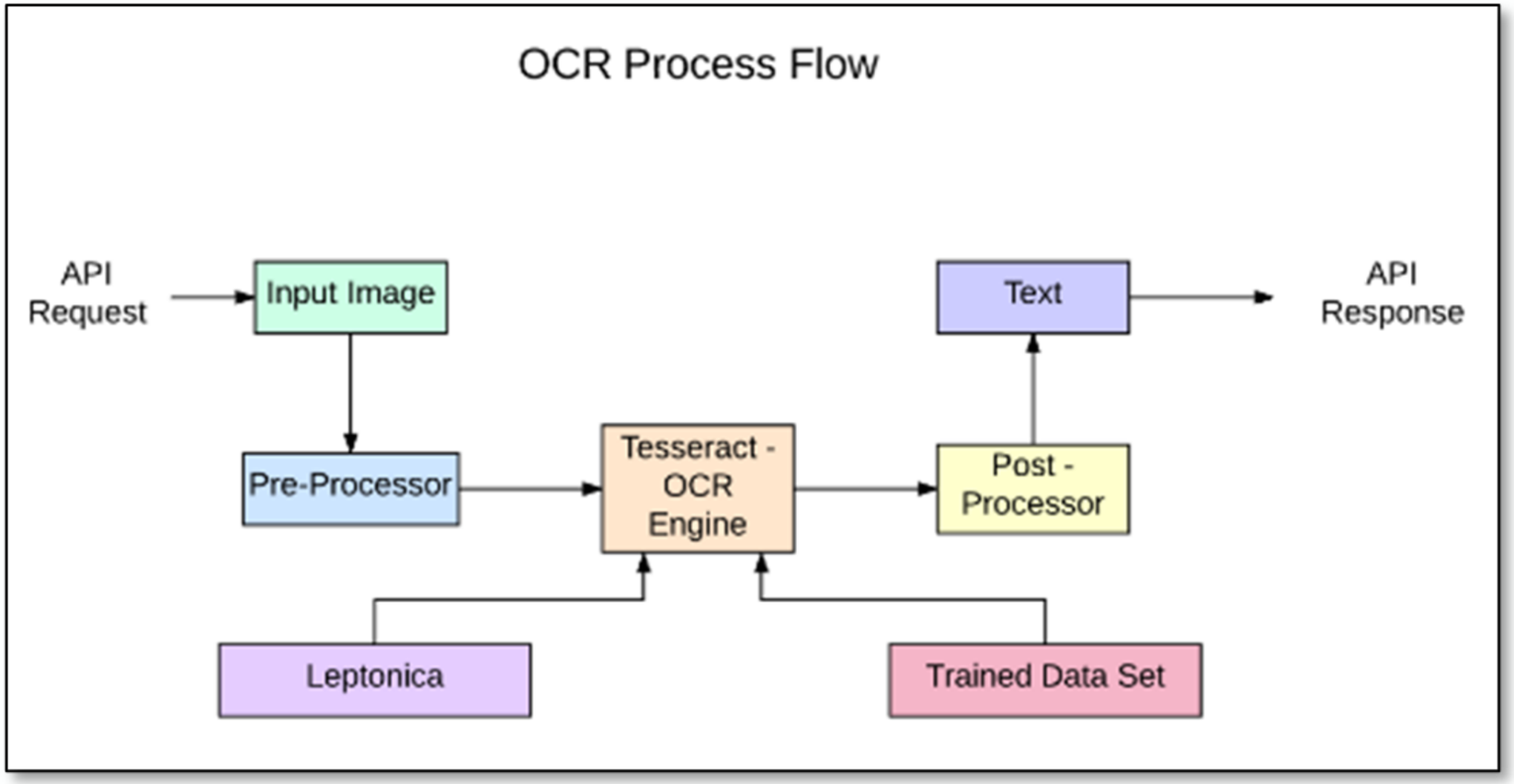
OCROPUS

Tesseract OCR คืออะไร ?

Tesseract เป็นเอ็นจิ้นการรู้จำข้อความ Open source (OCR) ซึ่งอยู่ภายใต้ลิขสิทธิ์ Apache 2.0 ปล่อยอย่างเป็นทางการ ปัจจุบัน คือ 4.1.1



Tesseract OCR คืออะไร ?



Tesseract OCR

1984
PhD project by HP

2005
Open Source by HP

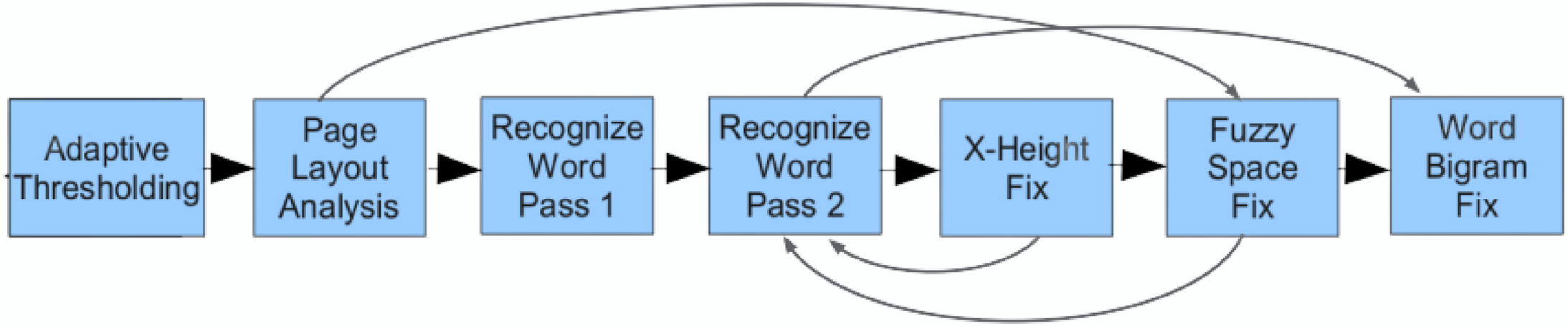
2006
Take Over by Google
version 1

2006
Expanded to 6 langs
version 2

2010
Layout analysis
60 langs
version 3

2016
LSTM
100 langs
version 4

โครงสร้างการประมวลผล

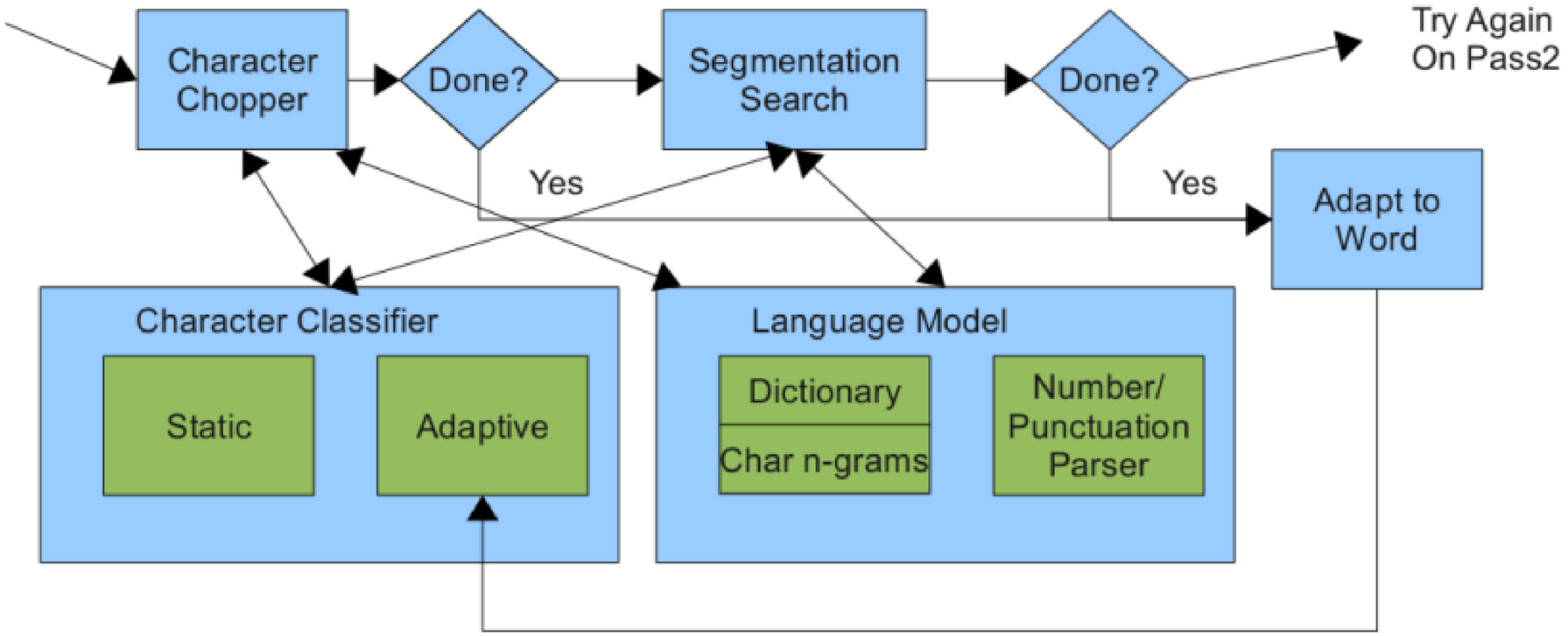


1 2 3 4 5 6 7

Preprocess | Recoginize | Postprocess

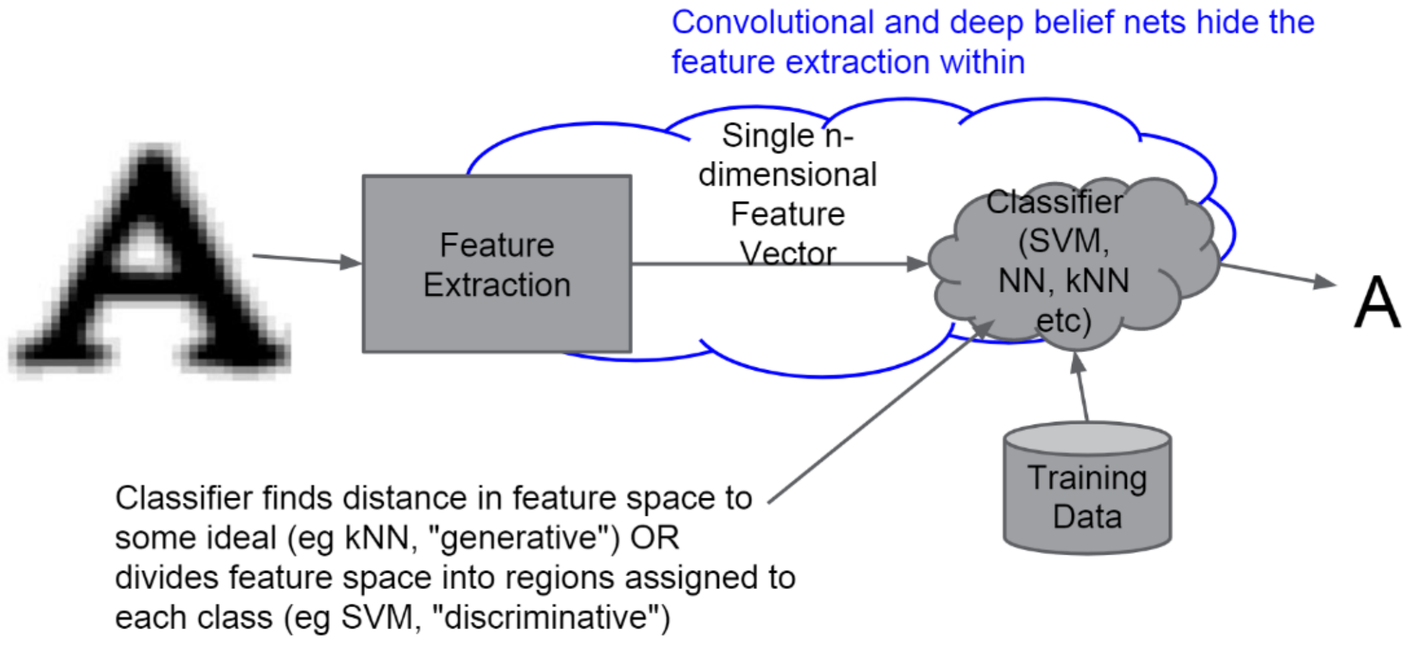
โครงสร้างการประมวลผล

Tesseract Word Recognizer

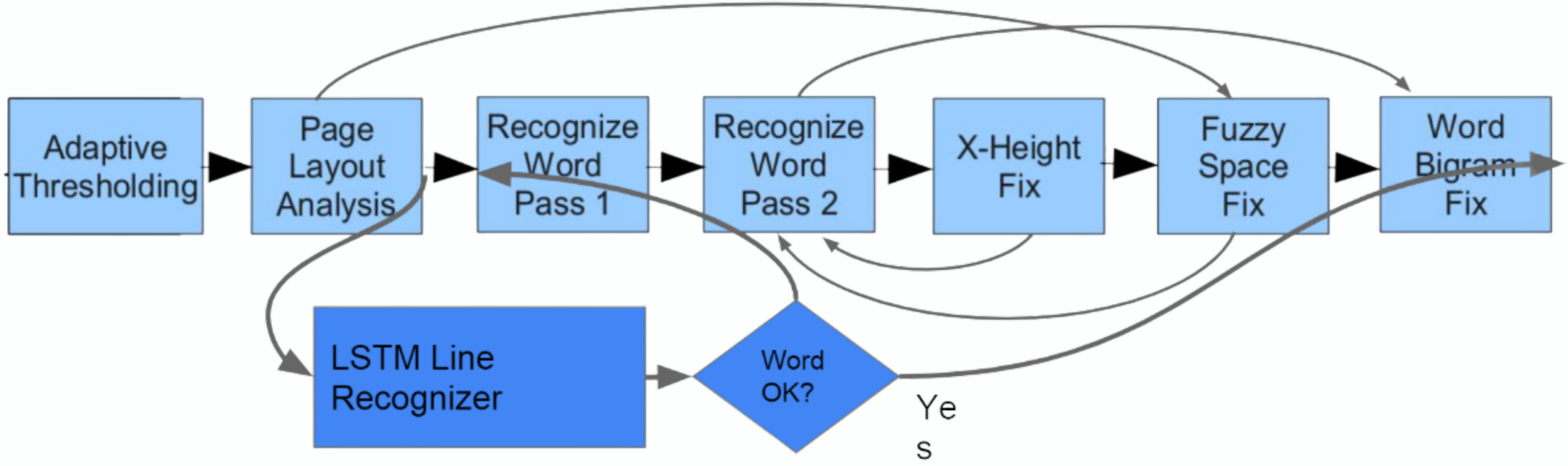


โครงสร้างการประมวลผล

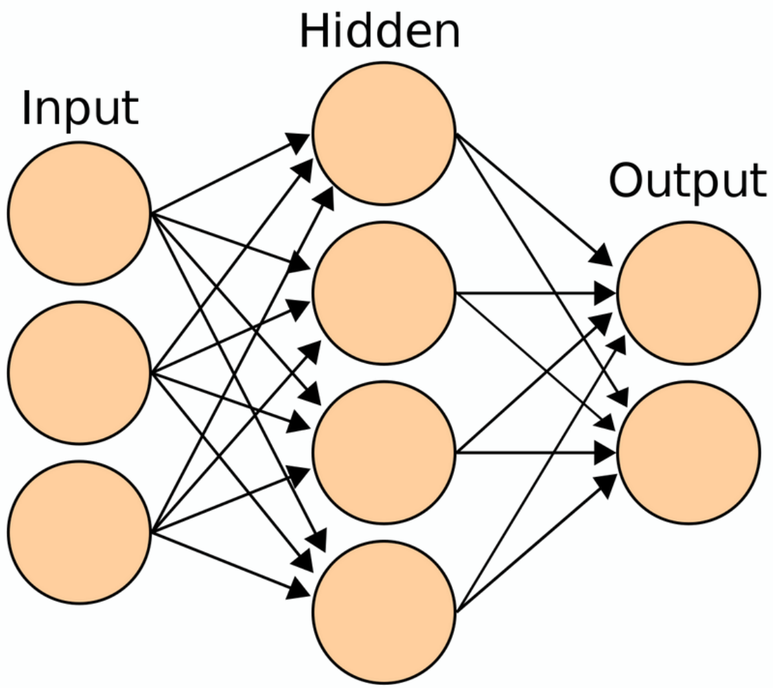
Background: Classical character classification



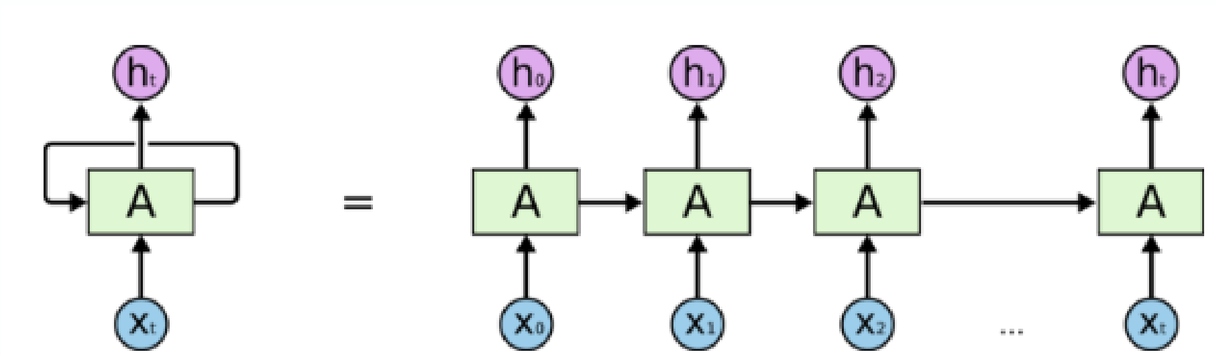
โครงสร้างการประมวลผล 4.0



ANN กับ RNN

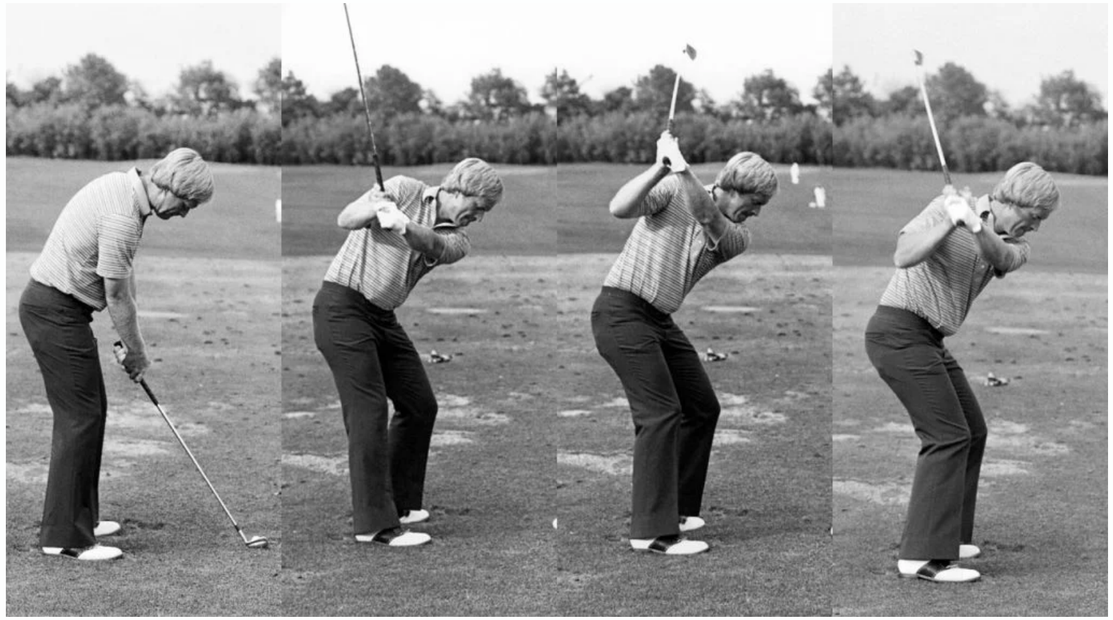


Artificial Neural Network



Recurrent Neural Network

RNN จะใช้งานกับข้อมูลที่มีลักษณะเป็นลำดับ (sequence)



Video
(Sequence of images)

RNN จะใช้งานกับข้อมูลที่มีลักษณะเป็นลำดับ (sequence)

Good Morning.

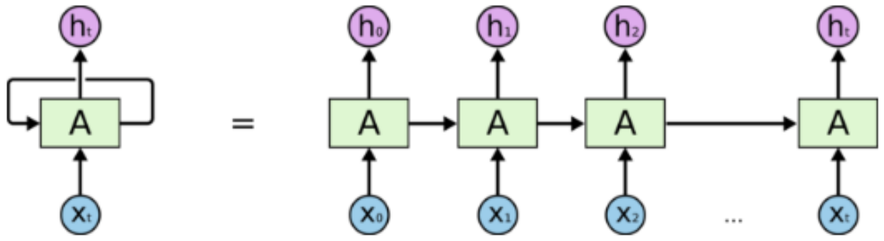
Nice to meet you.

Words

(Sequence of words)

RNN LSTM

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



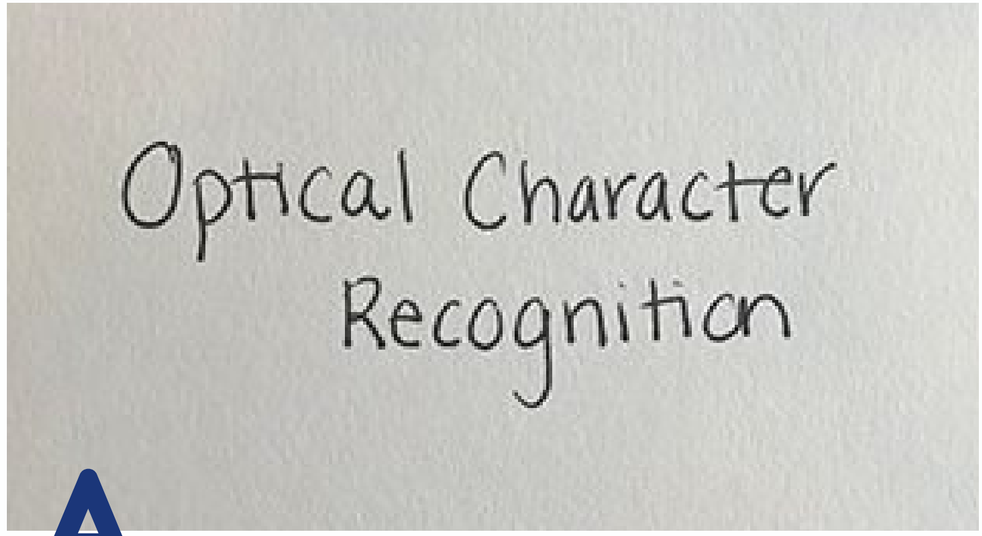
An unrolled recurrent neural network.

ข้อจำกัดของ Tesseract OCR

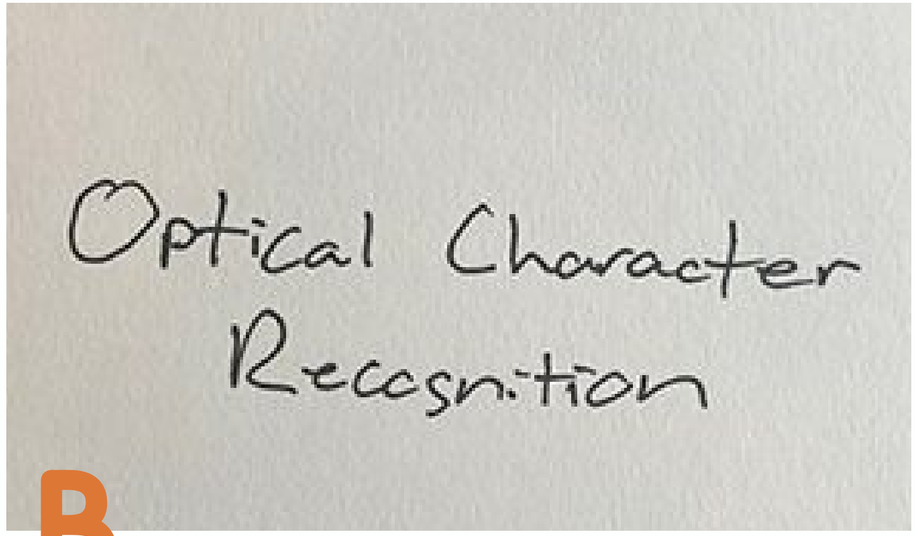
- ประมวลผลไม่ดีนักกับภาพที่ได้รับผลกระทบจากสิ่งของ มุมมองที่บิดเบี้ยว และพื้นหลังที่ซับซ้อน

ข้อจำกัดของ Tesseract OCR

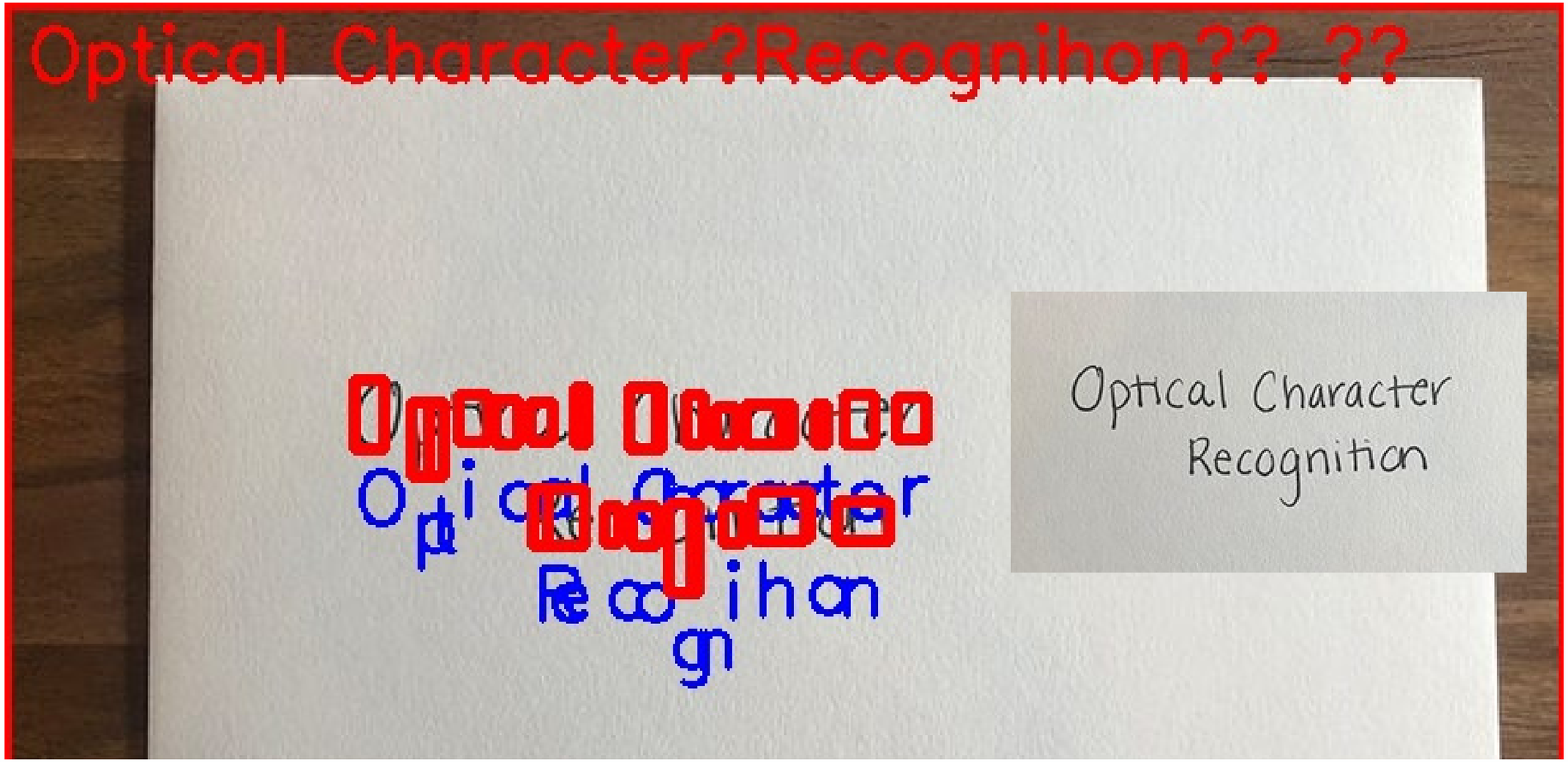
- ไม่สามารถประมวลผลลายมือแบบต่อเนื่องได้



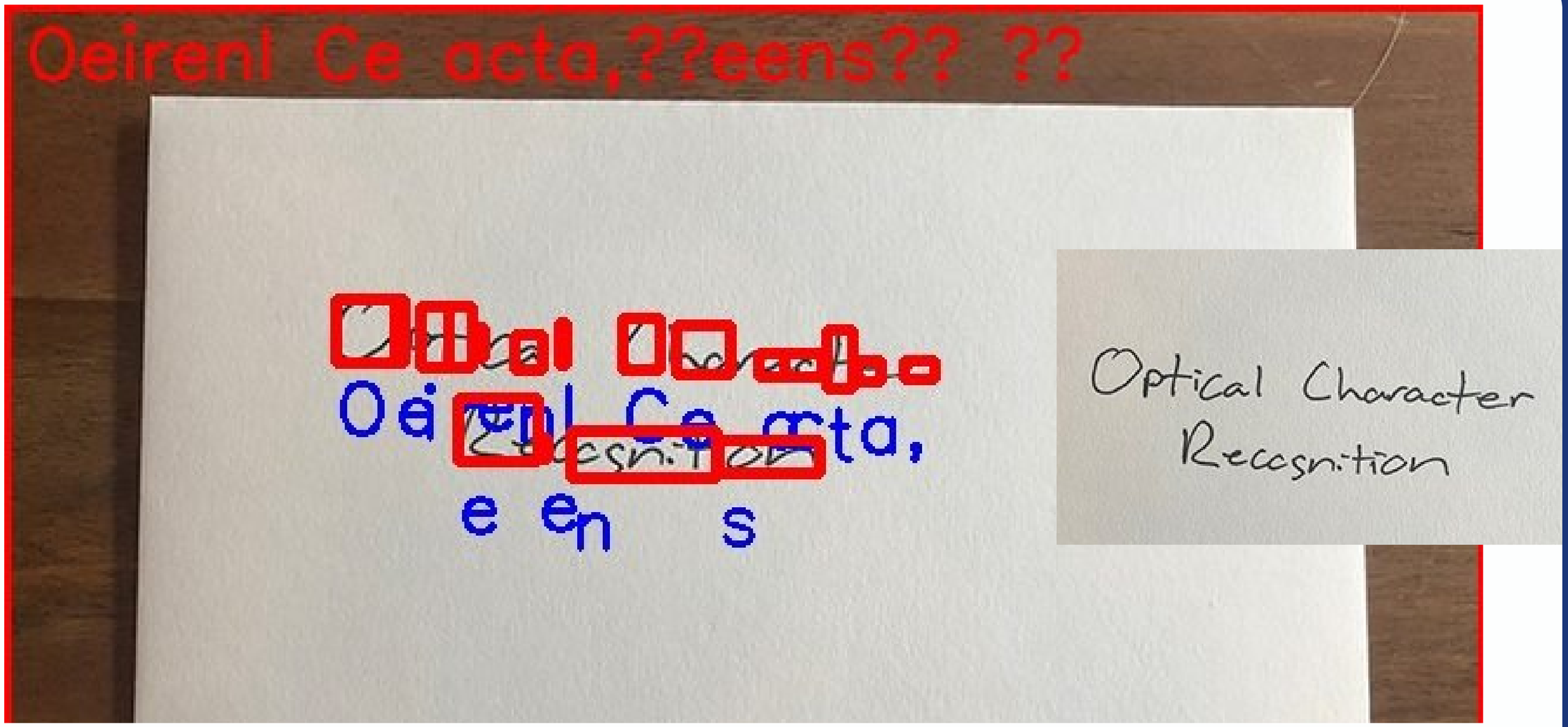
A



B


Optical Character?Recognihon?? ??
Optical Character
Recognition

A



B

ข้อจำกัดของ Tesseract OCR

- อาจพบว่าการประมวลผลไม่ถูกต้อง และแสดงผลเป็นเอาต์พุต
- หากเอกสารมีภาษาอื่นนอกเหนือจากที่ระบุในอาร์กิวเมนต์ ผลลัพธ์อาจไม่ดี
- รูปภาพคุณภาพต่ำอาจทำให้ OCR มีความแม่นยำต่ำ

ข้อจำกัดของ Tesseract OCR

วารสารวิจัยและส่งเสริมวิชาการเกษตร 33(1): 66-76

ผลการวิจัย

ข้อมูลพื้นฐานของเกษตรกรชาวสวนยางพารา

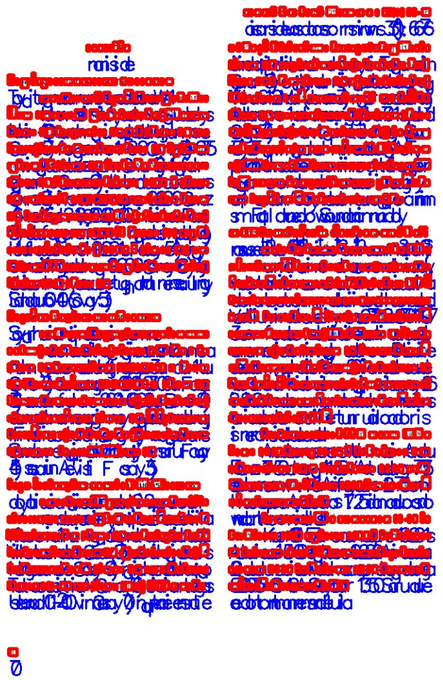
ผลการวิเคราะห์ข้อมูลทั่วไปของกลุ่มตัวอย่างทั้งหมด 400 คน พบว่า กลุ่มตัวอย่างส่วนใหญ่เป็นเพศชาย (ร้อยละ 67) มีอายุระหว่าง 41-50 ปี (ร้อยละ 34) การศึกษาอยู่ในระดับประถมศึกษา(ร้อยละ 56) มีภูมิลำเนาอยู่ในภาคใต้ (ร้อยละ 94) มีอาชีพเป็นเจ้าของสวนยาง (ร้อยละ 63) มีประสบการณ์กรีดยาง 10-20 ปี (ร้อยละ 46) และมีรายได้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

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

การประเมินแนวคิดเกี่ยวกับเครื่องกรีดยางพาราอัตโนมัติ

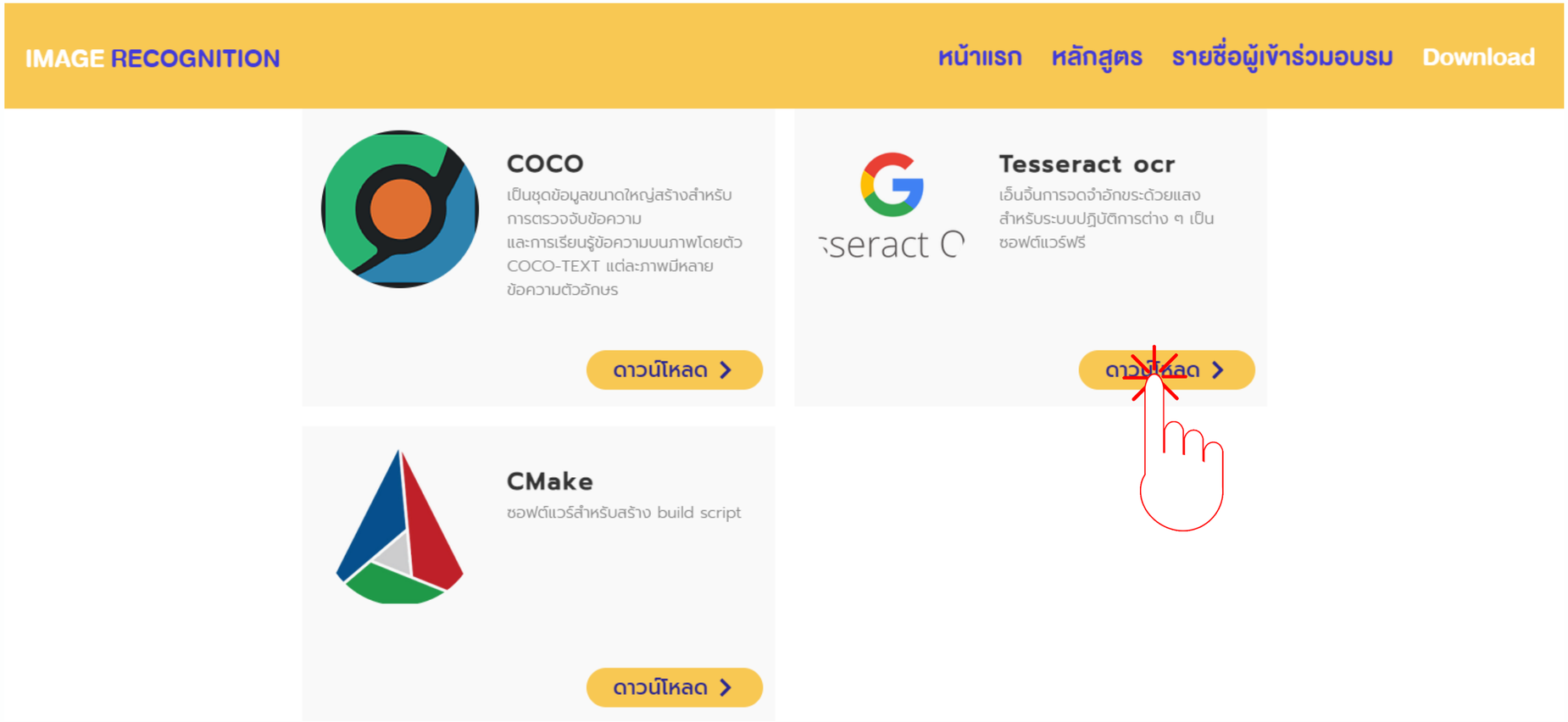
แนวคิดเกี่ยวกับเครื่องกรีดยางพาราอัตโนมัติเป็นเครื่องกรีดยางที่ไม่ต้องอาศัยคนกรีด การประเมินแนวคิด

naniside??Toyaitugmuoanunsnsyradauara Ws1??Hanr



naniside??Toyaitugmuoanunsnsyradauara Ws1??Hanr

ติดตั้ง Tesseract



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.

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)

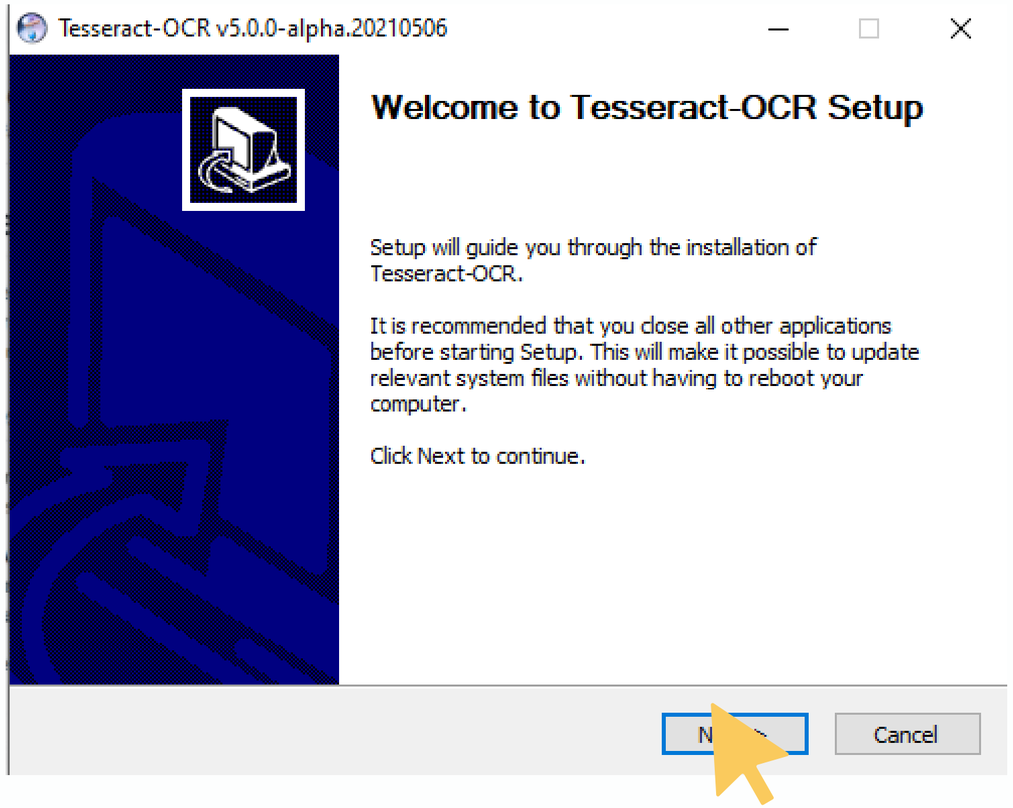
0 Contributors · 9 Issues · 1k Stars · 234 Forks

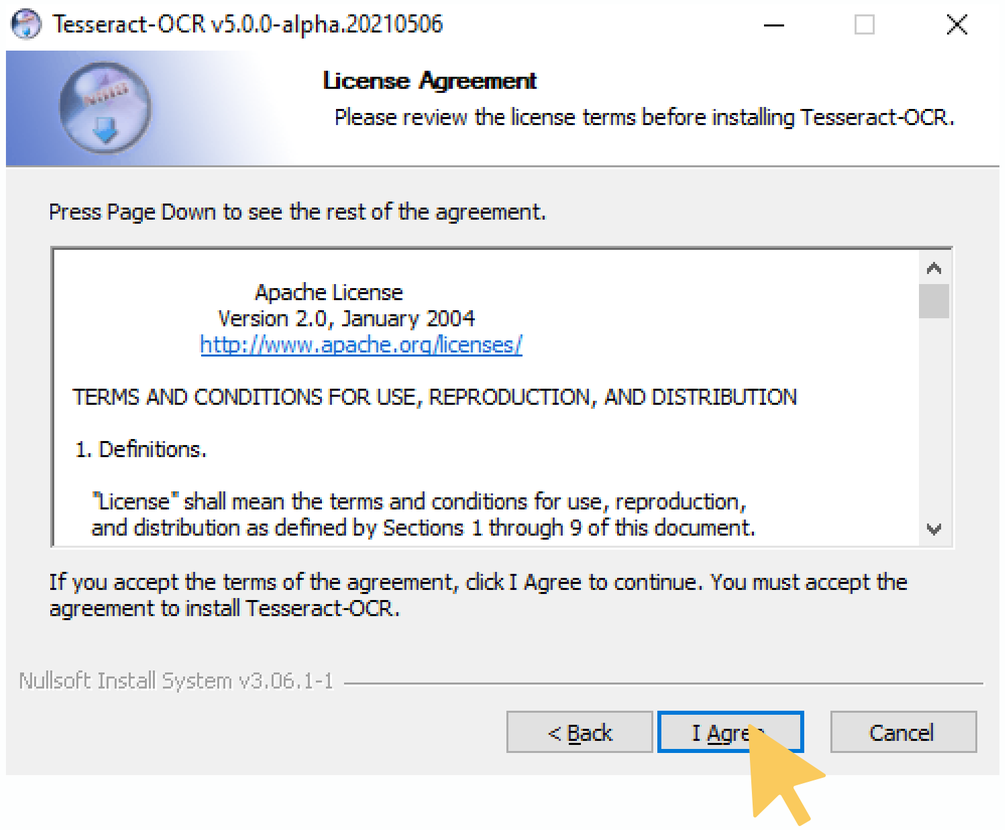
Home · UB-Mannheim/tesseract Wiki

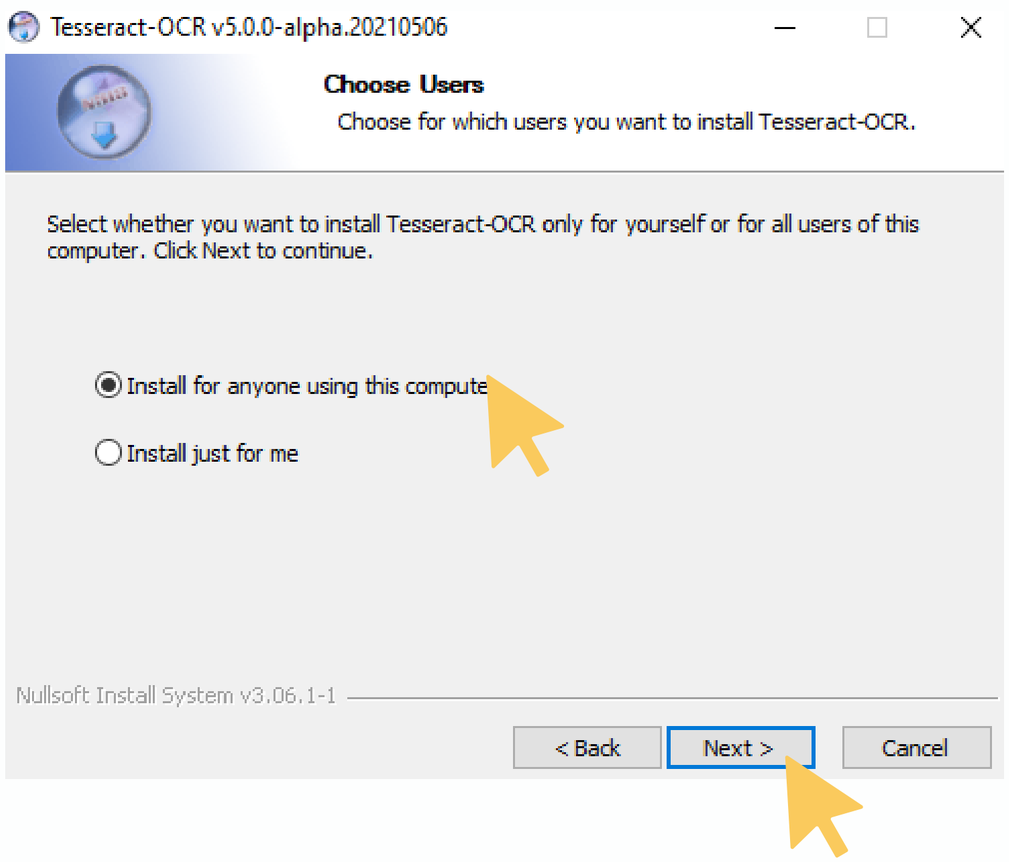
Tesseract Open Source OCR Engine (main repository) - Home · UB-Mannheim/tesseract Wiki

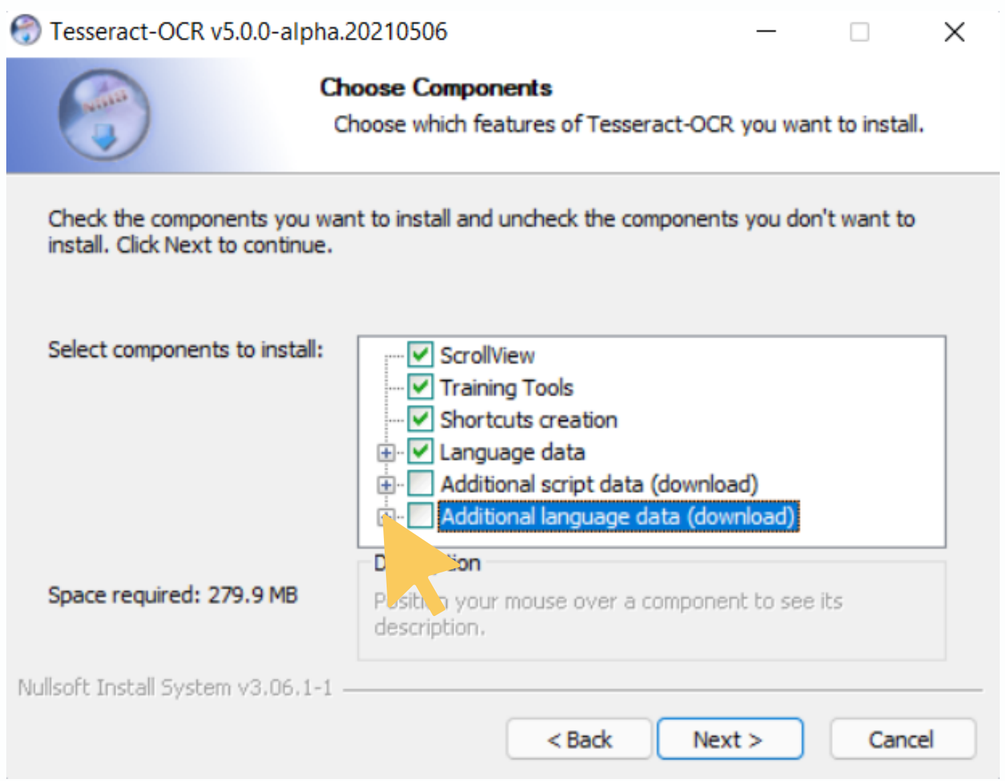
GitHub

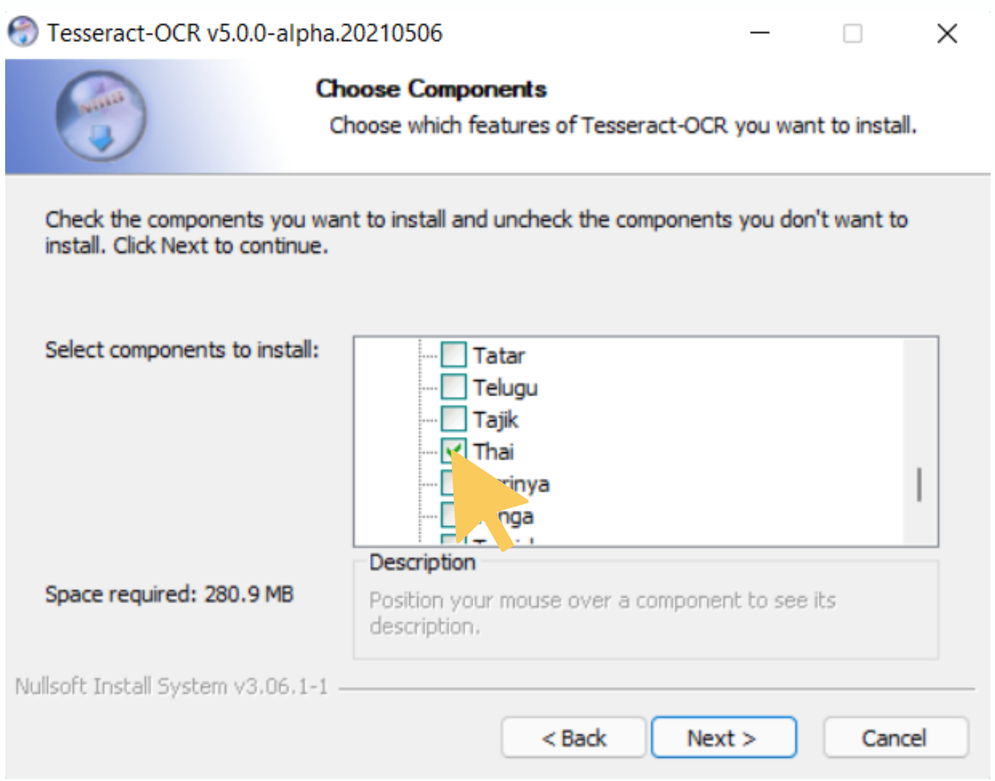
- tesseract-ocr-w32-setup-v5.0.0-alpha.20210506.exe (32 bit) and
- tesseract-ocr-w64-setup-v5.0.0-alpha.20210506.exe (64 bit) resp.

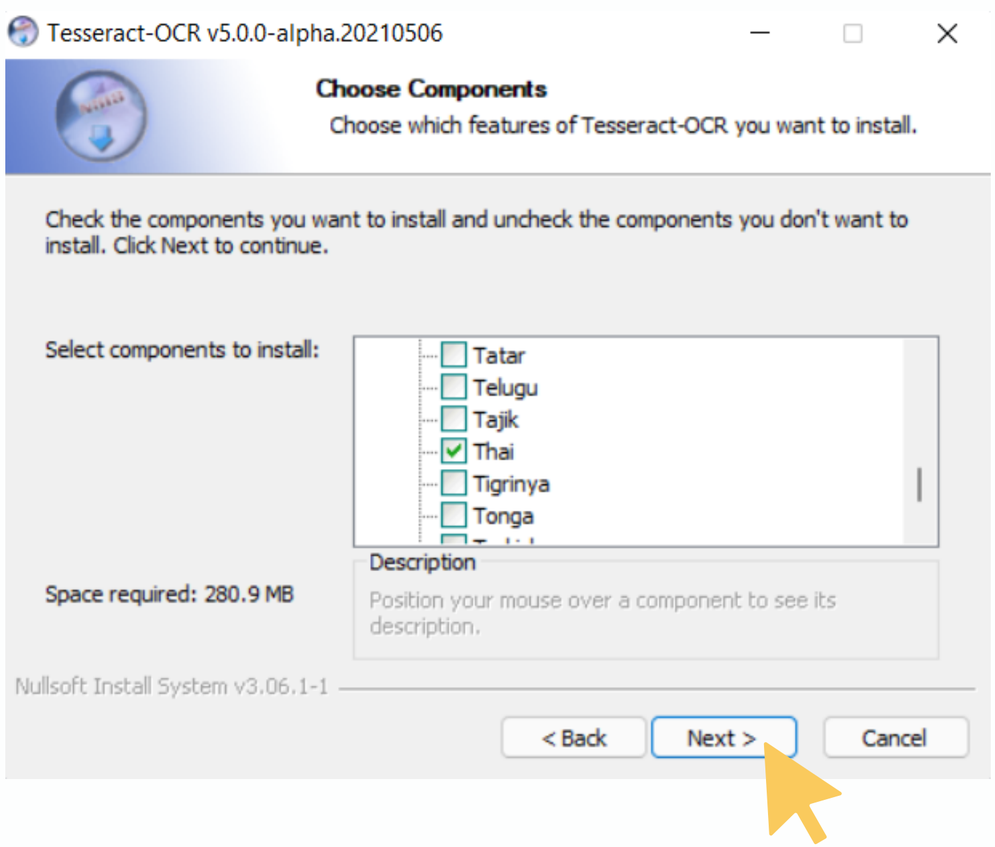

Tesseract-OCR v5.0.0-alpha.20210506
Welcome to Tesseract-OCR Setup
Setup will guide you through the installation of Tesseract-OCR.
It is recommended that you close all other applications before starting Setup. This will make it possible to update relevant system files without having to reboot your computer.
Click Next to continue.
Cancel

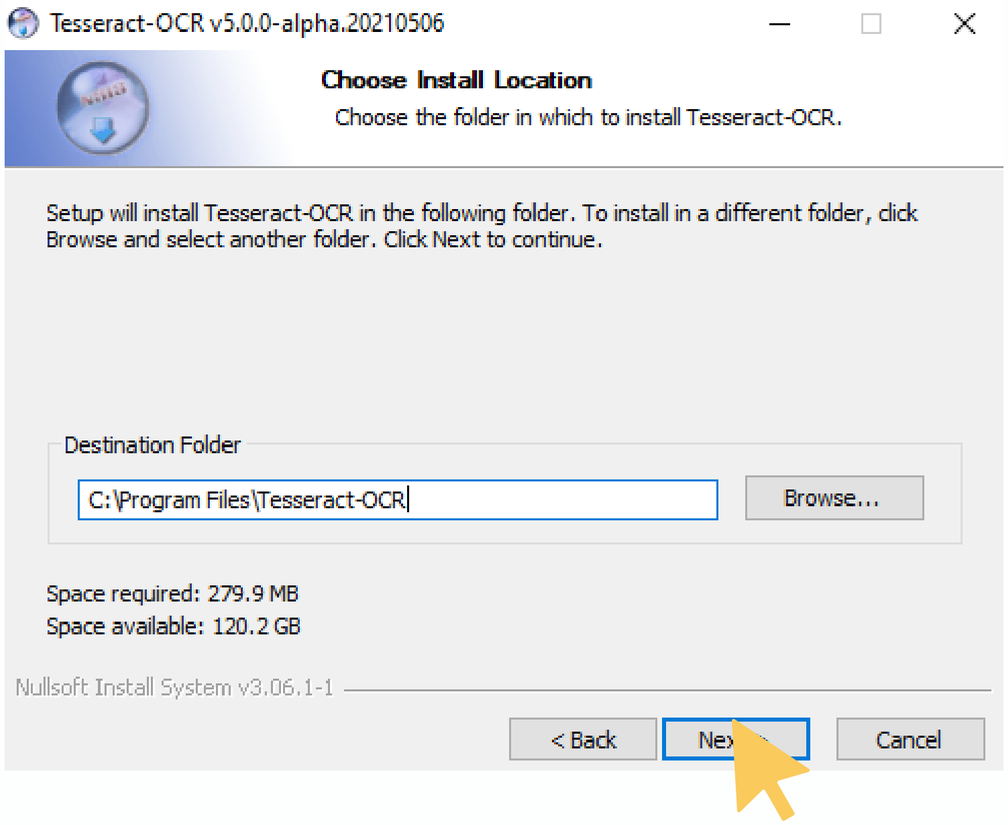

Tesseract-OCR v5.0.0-alpha.20210506
License Agreement
Please review the license terms before installing Tesseract-OCR.
Press Page Down to see the rest of the agreement.
Apache License
Version 2.0, January 2004
http://www.apache.org/licenses/
TERMS AND CONDITIONS FOR USE, REPRODUCTION, AND DISTRIBUTION
1. Definitions.
"License" shall mean the terms and conditions for use, reproduction, and distribution as defined by Sections 1 through 9 of this document.
If you accept the terms of the agreement, click I Agree to continue. You must accept the agreement to install Tesseract-OCR.
Nullsoft Install System v3.06.1-1
< Back
Cancel

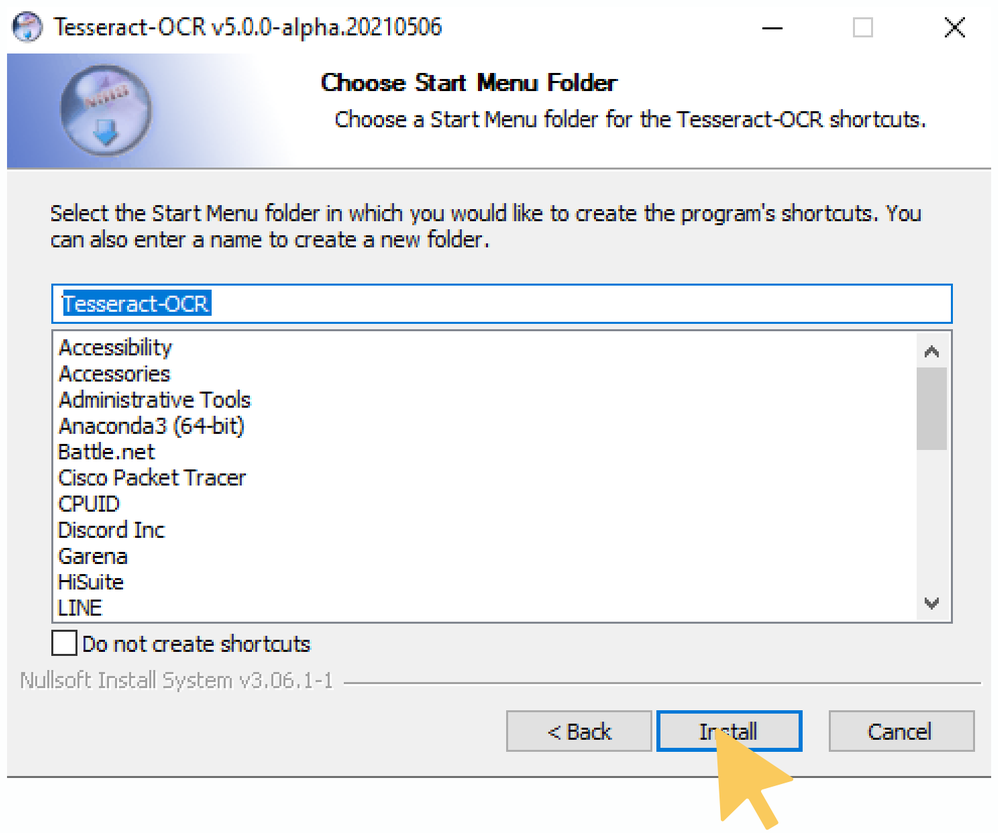

Tesseract-OCR v5.0.0-alpha.20210506
Choose Users
Choose for which users you want to install Tesseract-OCR.
Select whether you want to install Tesseract-OCR only for yourself or for all users of this computer. Click Next to continue.
Install for anyone using this compute
Install just for me
Nullsoft Install System v3.06.1-1
< Back
Next >
Cancel

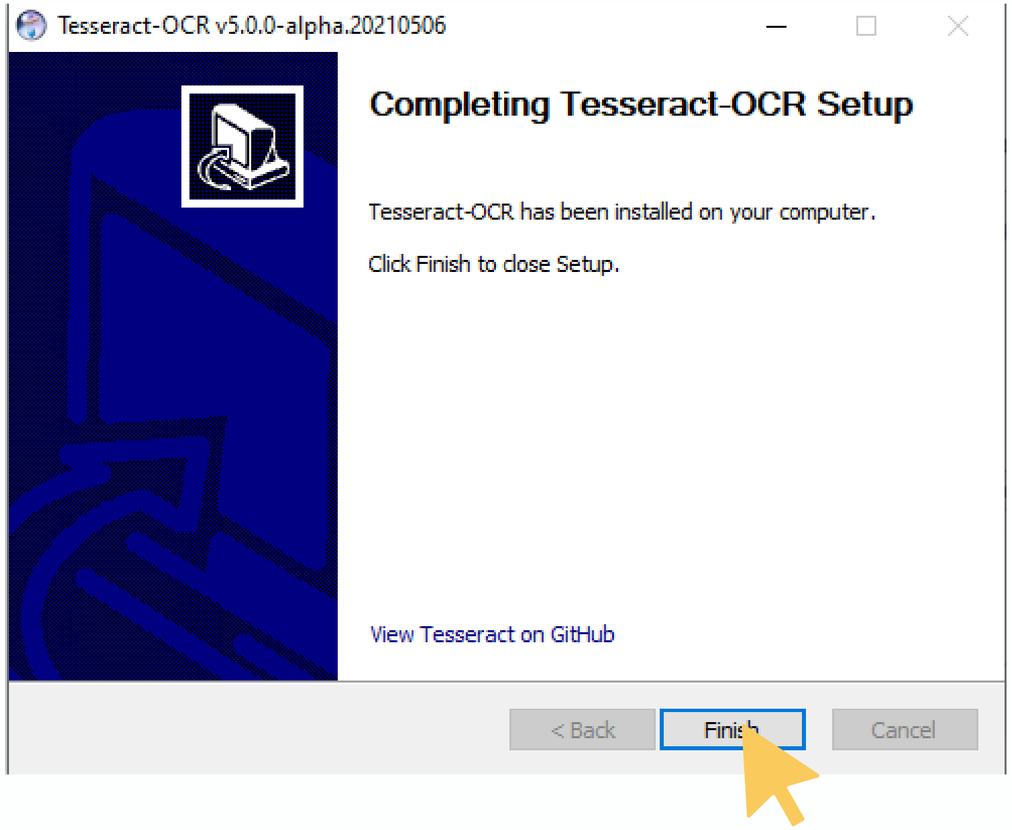

Tesseract-OCR v5.0.0-alpha.20210506
Choose Components
Choose which features of Tesseract-OCR you want to install.
Check the components you want to install and uncheck the components you don't want to install. Click Next to continue.
Select components to install:
ScrollView
Training Tools
Shortcuts creation
Language data
Additional script data (download)
Additional language data (download)
Space required: 279.9 MB
Position your mouse over a component to see its description.
Nullsoft Install System v3.06.1-1
< Back
Next >
Cancel


Tesseract-OCR v5.0.0-alpha.20210506
Choose Components
Choose which features of Tesseract-OCR you want to install.
Check the components you want to install and uncheck the components you don't want to install. Click Next to continue.
Select components to install:
Tatar
Telugu
Tajik
Thai
Space required: 280.9 MB
Description
Position your mouse over a component to see its description.
Nullsoft Install System v3.06.1-1
< Back
Next >
Cancel

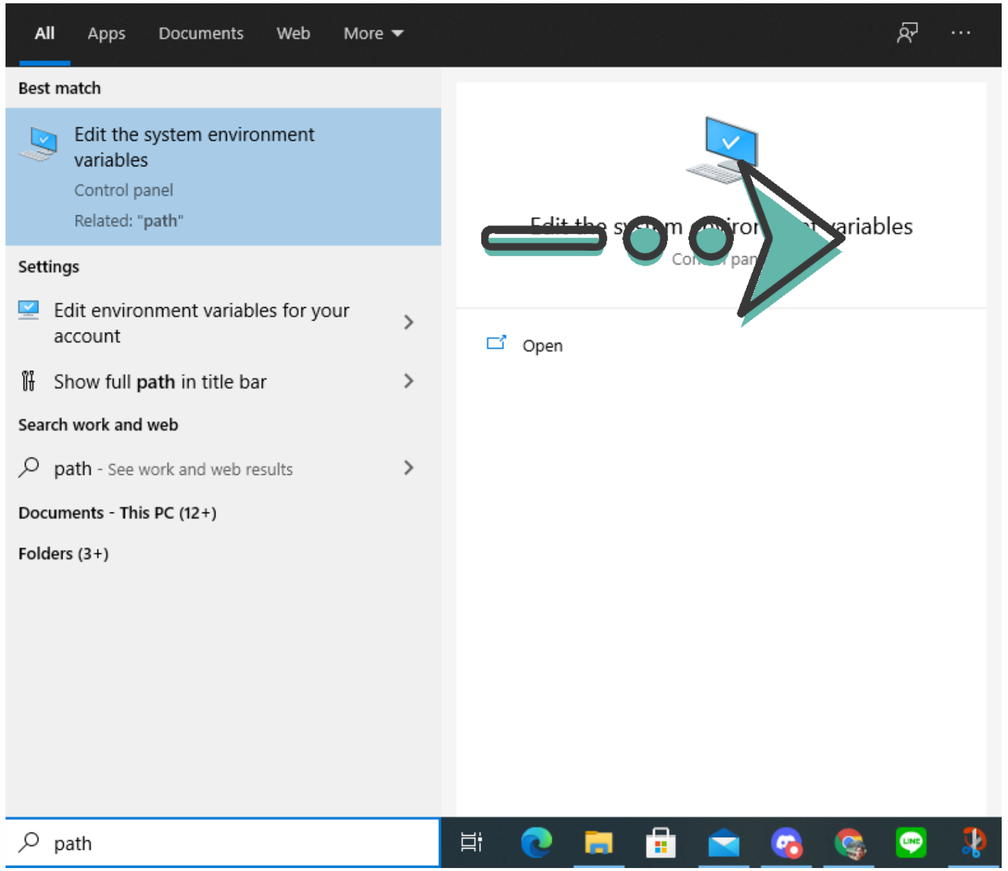

Tesseract-OCR v5.0.0-alpha.20210506
Choose Components
Choose which features of Tesseract-OCR you want to install.
Check the components you want to install and uncheck the components you don't want to install. Click Next to continue.
Select components to install:
Tatar
Telugu
Tajik
Thai
Tigrinya
Tonga
Space required: 280.9 MB
Description
Position your mouse over a component to see its description.
Nullsoft Install System v3.06.1-1
< Back
Next >
Cancel

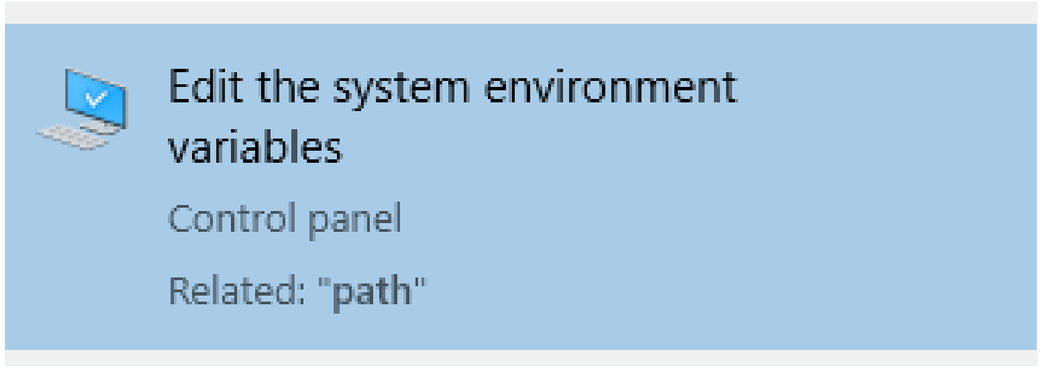

Tesseract-OCR v5.0.0-alpha.20210506
Choose Install Location
Choose the folder in which to install Tesseract-OCR.
Setup will install Tesseract-OCR in the following folder. To install in a different folder, click Browse and select another folder. Click Next to continue.
Destination Folder
C:\Program Files\Tesseract-OCR
Browse...
Space required: 279.9 MB
Space available: 120.2 GB
Nullsoft Install System v3.06.1-1
< Back
Next >
Cancel


Tesseract-OCR v5.0.0-alpha.20210506
Choose Start Menu Folder
Choose a Start Menu folder for the Tesseract-OCR shortcuts.
Select the Start Menu folder in which you would like to create the program's shortcuts. You can also enter a name to create a new folder.
Tesseract-OCR
Accessibility
Accessories
Administrative Tools
Anaconda3 (64-bit)
Battle.net
Cisco Packet Tracer
CPUID
Discord Inc
Garena
HiSuite
LINE
Do not create shortcuts
Nullsoft Install System v3.06.1-1
< Back
Install
Cancel

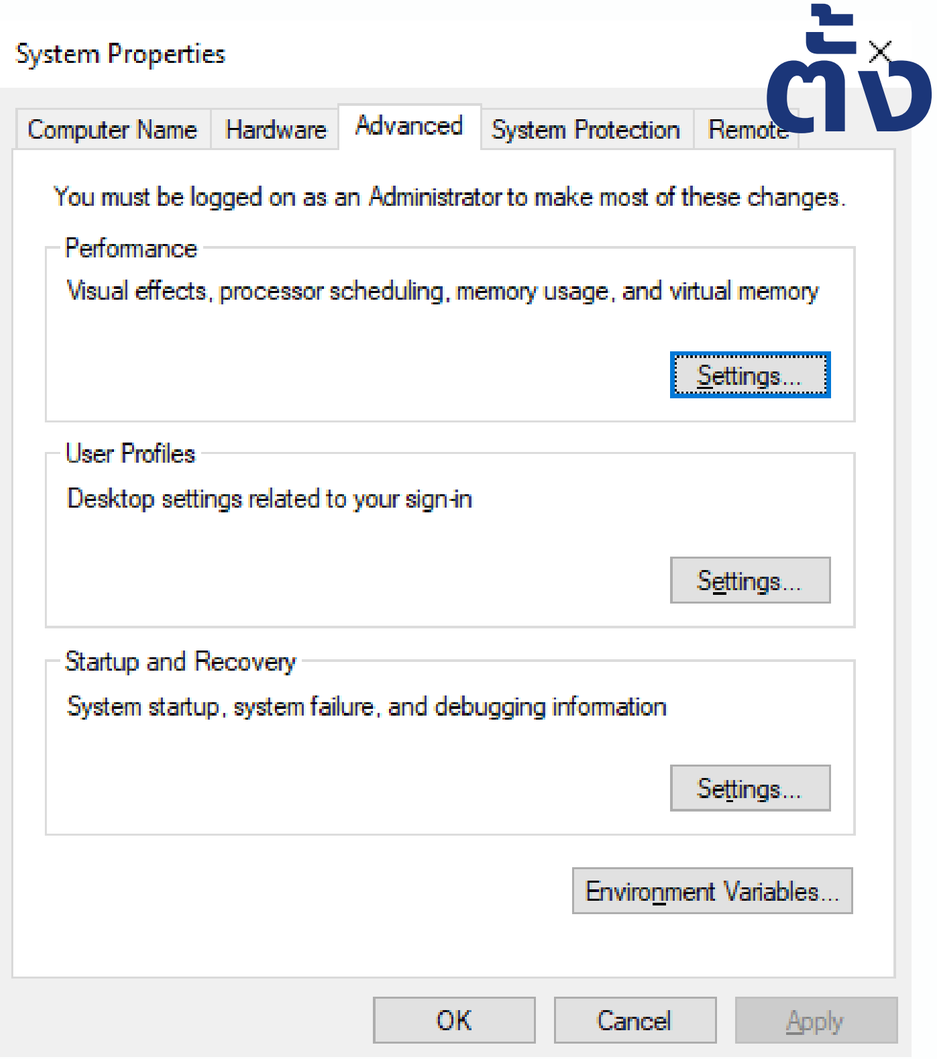

Tesseract-OCR v5.0.0-alpha.20210506
Completing Tesseract-OCR Setup
Tesseract-OCR has been installed on your computer.
Click Finish to close Setup.
View Tesseract on GitHub
< Back
Finish
Cancel

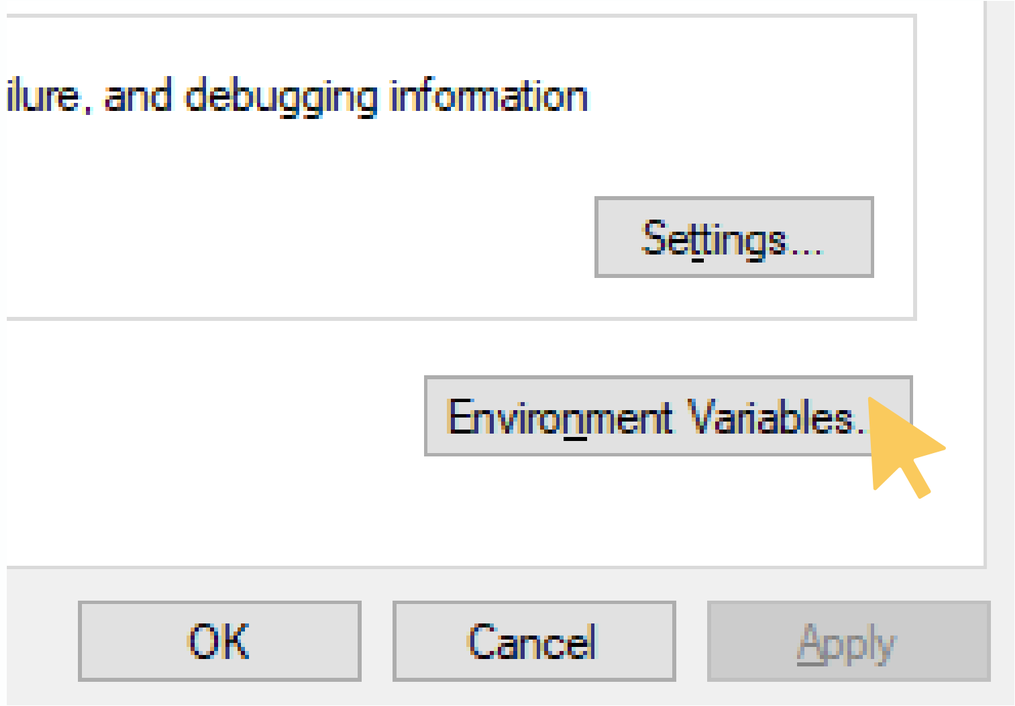
ตั้งค่า Path


All
Apps
Documents
Web
More
Best match
Edit the system environment variables
Control panel
Related: "path"
Settings
Edit environment variables for your account
Show full path in title bar
Search work and web
path - See work and web results
Documents - This PC (12+)
Folders (3+)
Open
path

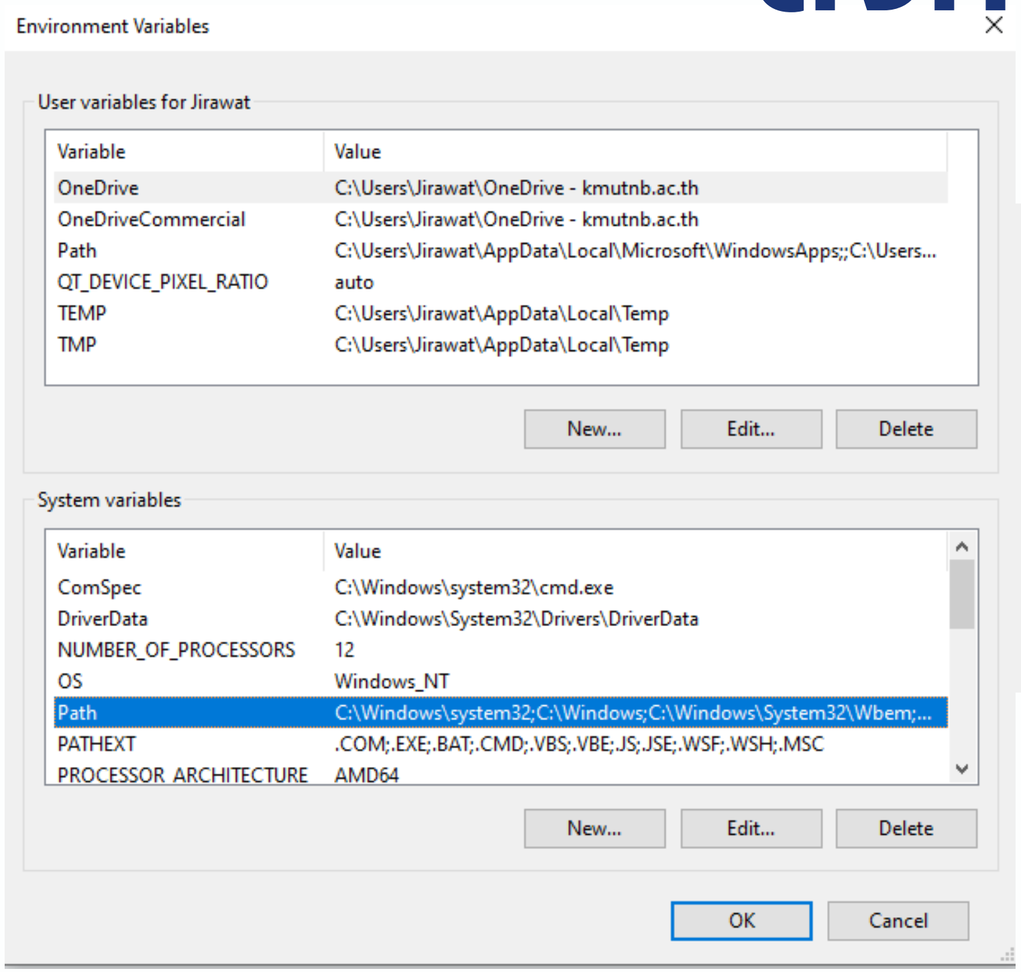

Edit the system environment variables
Control panel
Related: "path"

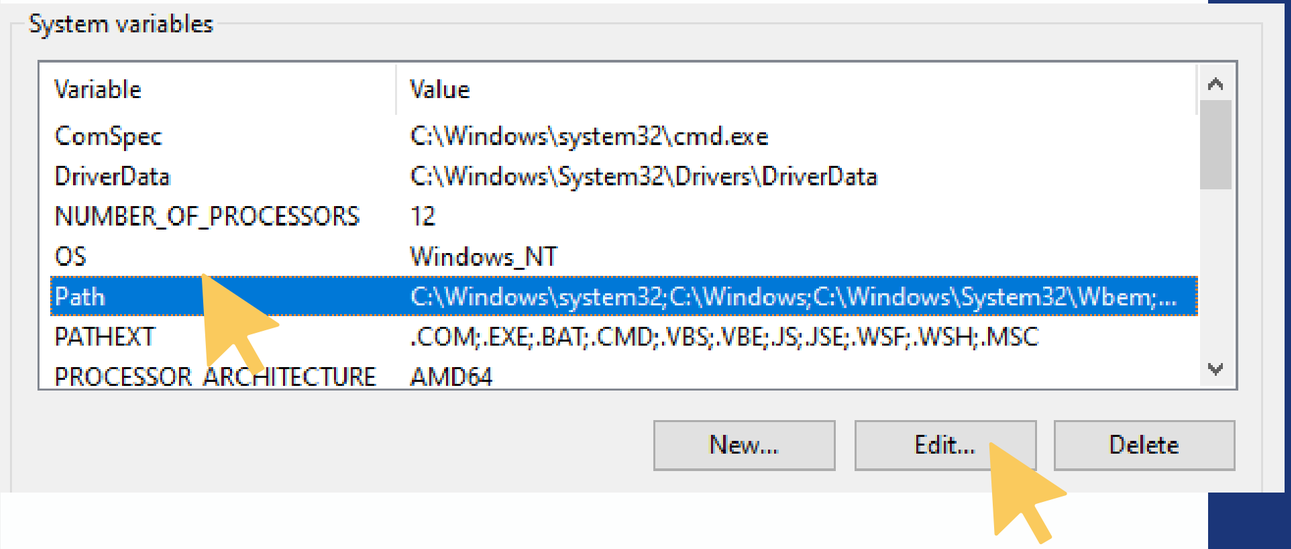
ตั้งค่า Path


System Properties
Computer Name
Hardware
Advanced
System Protection
Remote
You must be logged on as an Administrator to make most of these changes.
Performance
Visual effects, processor scheduling, memory usage, and virtual memory
Settings...
User Profiles
Desktop settings related to your sign-in
Settings...
Startup and Recovery
System startup, system failure, and debugging information
Settings...
Environment Variables...
OK
Cancel
Apply

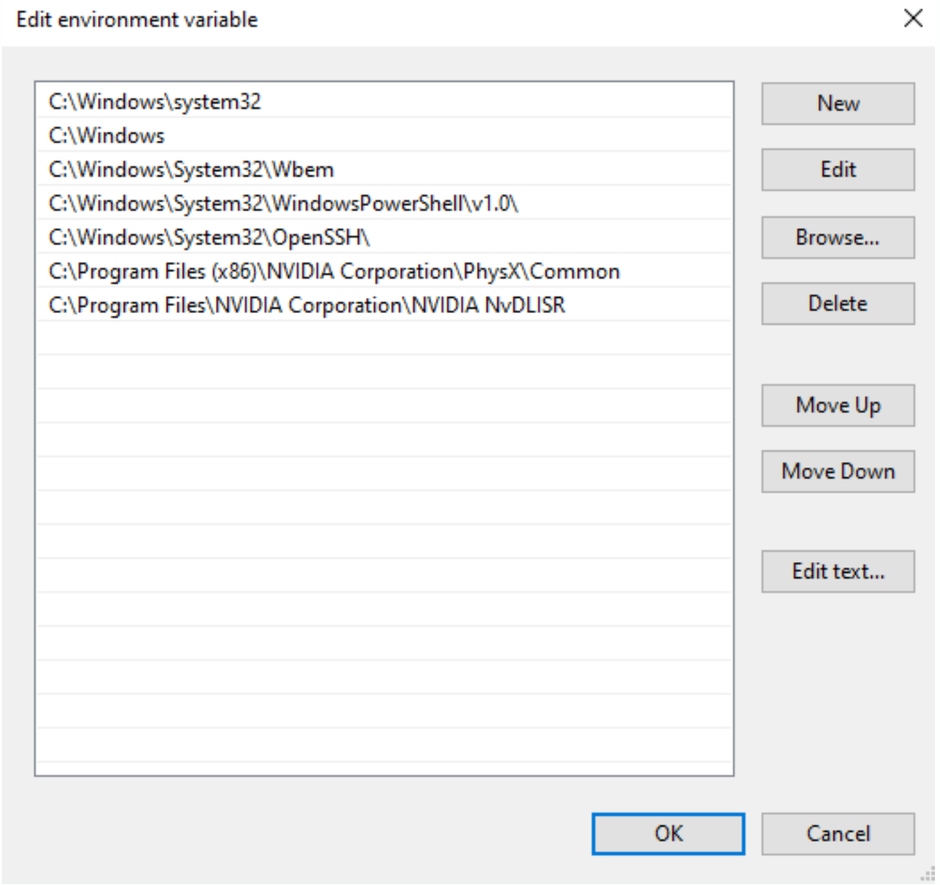

ilure, and debugging information
Settings...
Environment Variables...
OK
Cancel
Apply

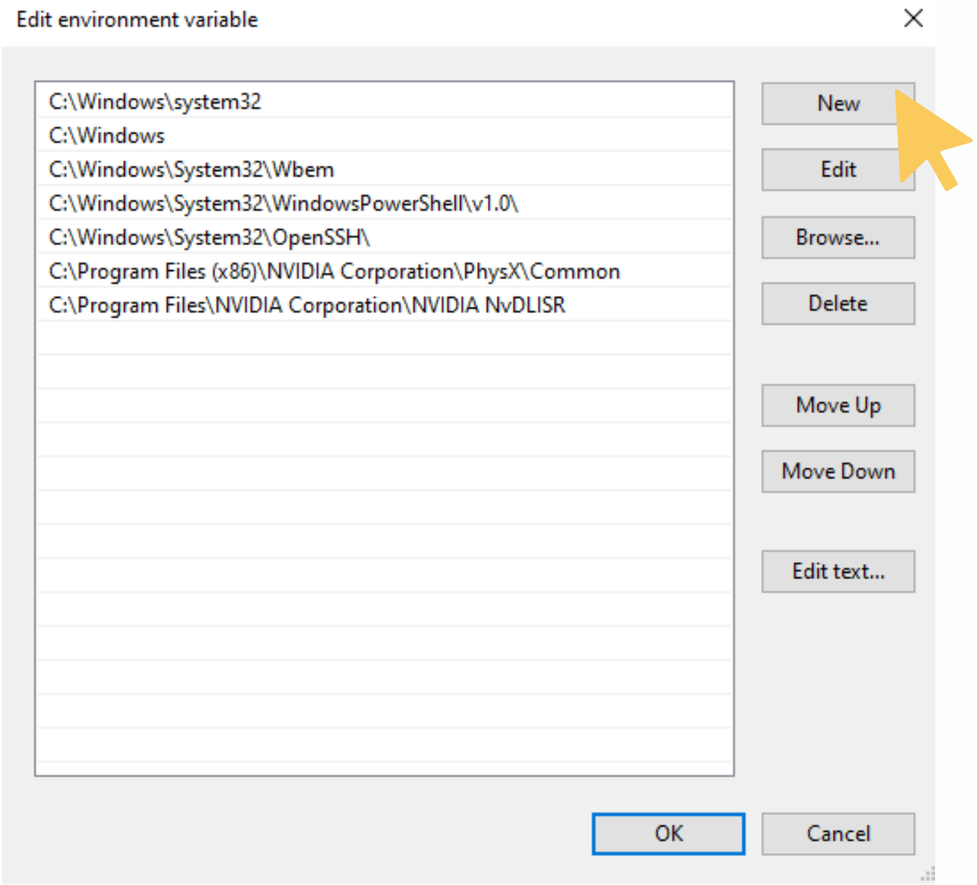
ตั้งค่า Path


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

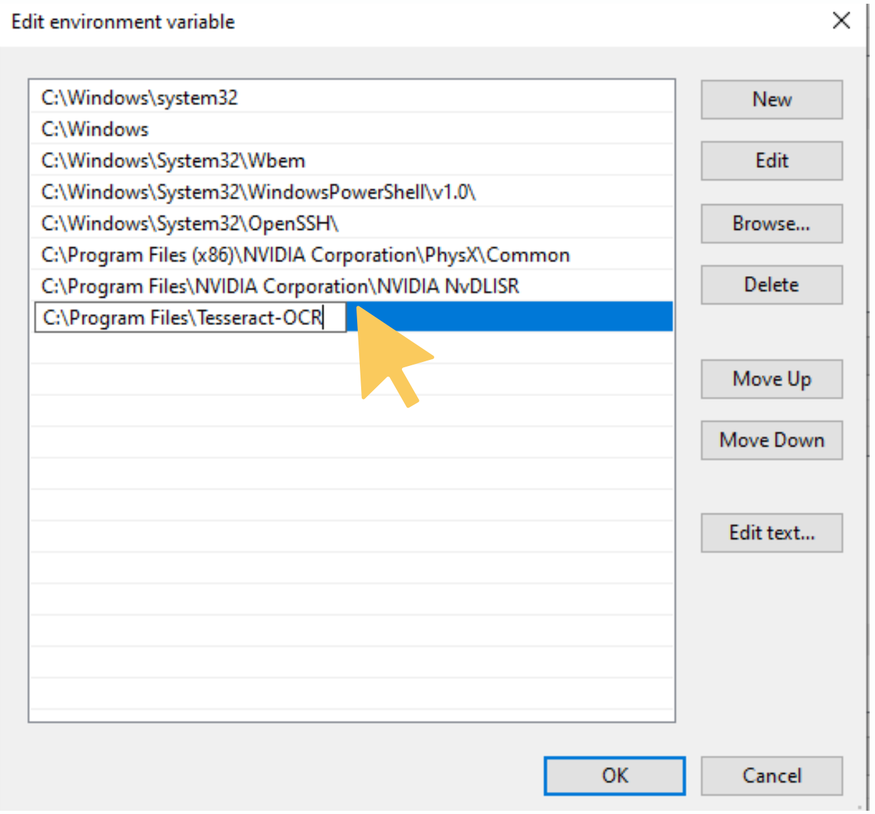

System variables
Variable
Value
ComSpec
C:\Windows\system32\cmd.exe
DriverData
C:\Windows\System32\Drivers\DriverData
NUMBER_OF_PROCESSORS
12
OS
Windows_NT
Path
C:\Windows\system32;C:\Windows;C:\Windows\System32\Wbem;...
PATHEXT
.COM;.EXE;.BAT;.CMD;.VBS;.VBE;.JS;.JSE;.WSF;.WSH;.MSC
PROCESSOR_ARCHITECTURE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

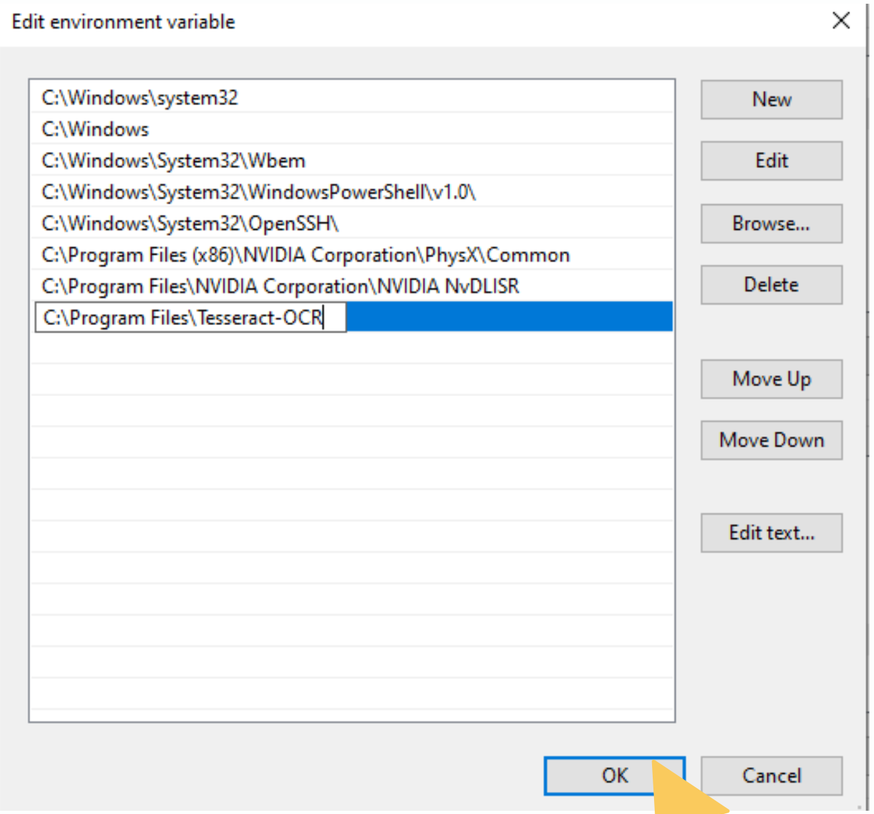
ตั้งค่า Path


Edit environment variable
C:\Windows\system32
C:\Windows
C:\Windows\System32\Wbem
C:\Windows\System32\WindowsPowerShell\v1.0\
C:\Windows\System32\OpenSSH\
C:\Program Files (x86)\NVIDIA Corporation\PhysX\Common
C:\Program Files\NVIDIA Corporation\NVIDIA NvDLISR
New
Edit
Browse...
Delete
Move Up
Move Down
Edit text...
OK
Cancel


Edit environment variable
C:\Windows\system32
C:\Windows
C:\Windows\System32\Wbem
C:\Windows\System32\WindowsPowerShell\v1.0\
C:\Windows\System32\OpenSSH\
C:\Program Files (x86)\NVIDIA Corporation\PhysX\Common
C:\Program Files\NVIDIA Corporation\NVIDIA NvDLISR
New
Edit
Browse...
Delete
Move Up
Move Down
Edit text...
OK
Cancel

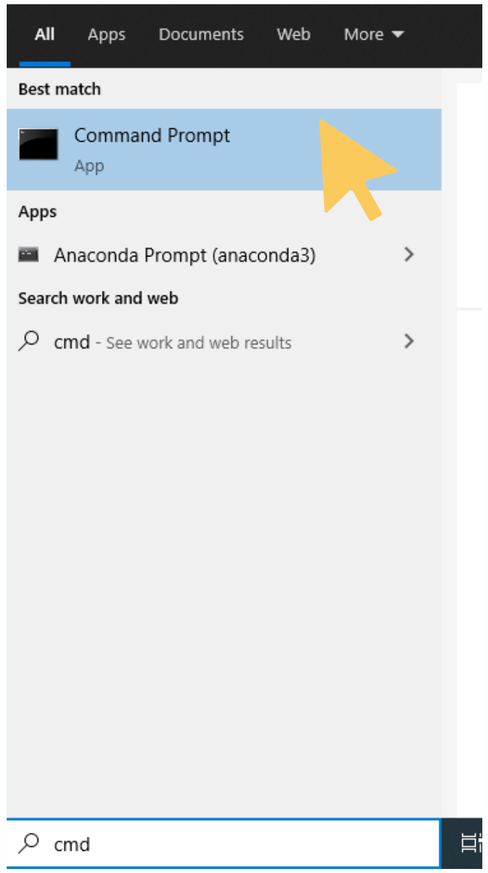
ตั้งค่า Path

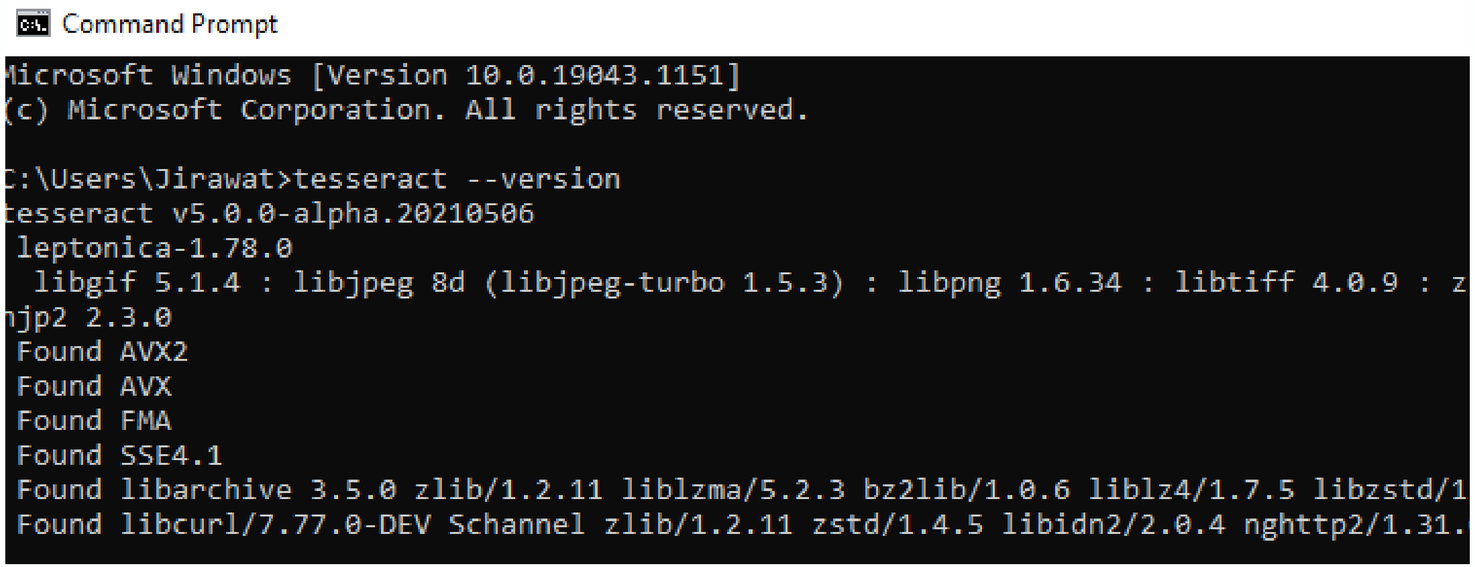




C:\Program Files\Tesseract-OCR

ตรวจสอบการติดตั้ง





tesseract --version

ติดตั้งส่วนเสริม Jupyter

jupyter

Logout

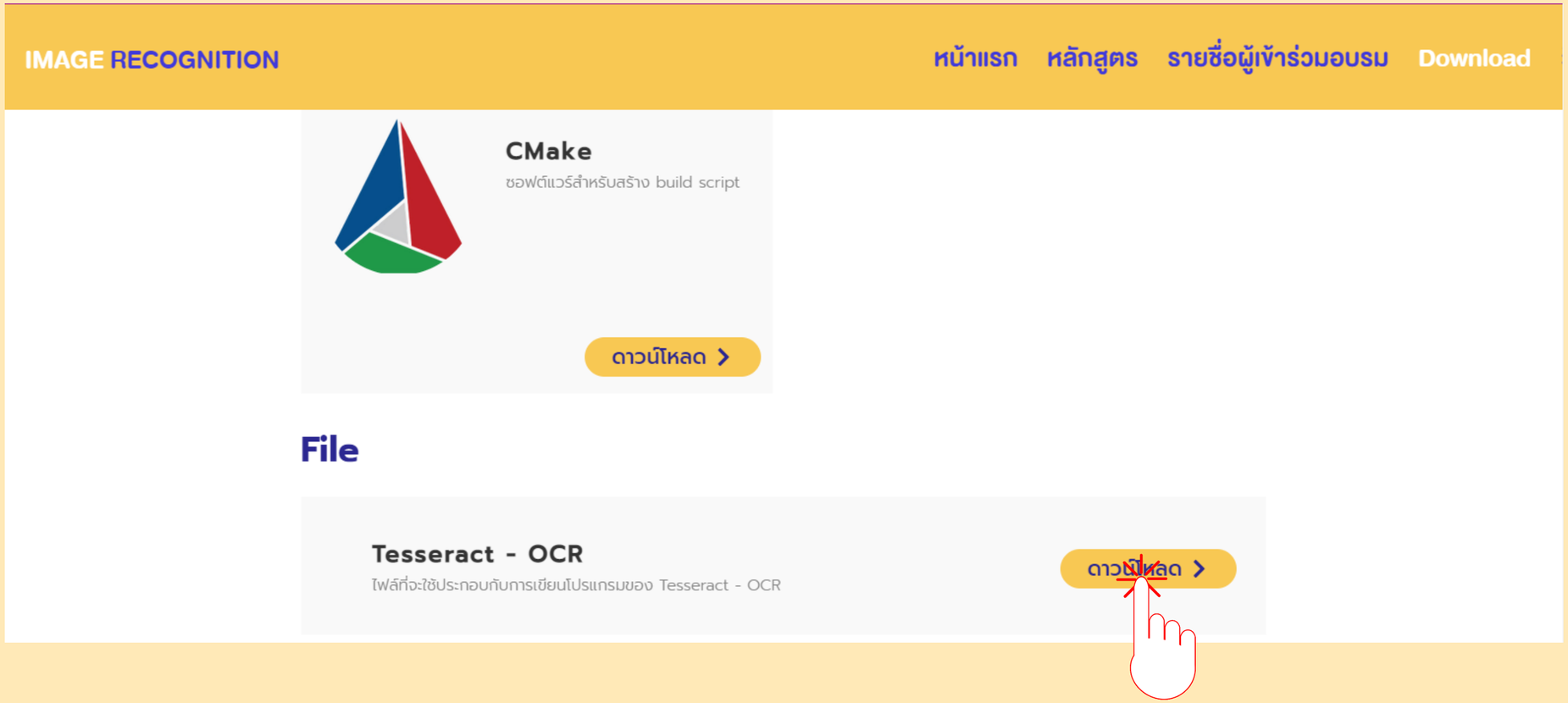
```
Windows PowerShell
Copyright (C) Microsoft Corporation. All rights reserved.

Try the new cross-platform PowerShell https://aka.ms/pscore6

PS C:\Users\Jirawat> cd D:
PS D:\> cd python3
PS D:\python3> pip install opencv-python
```

>pip install opencv-python

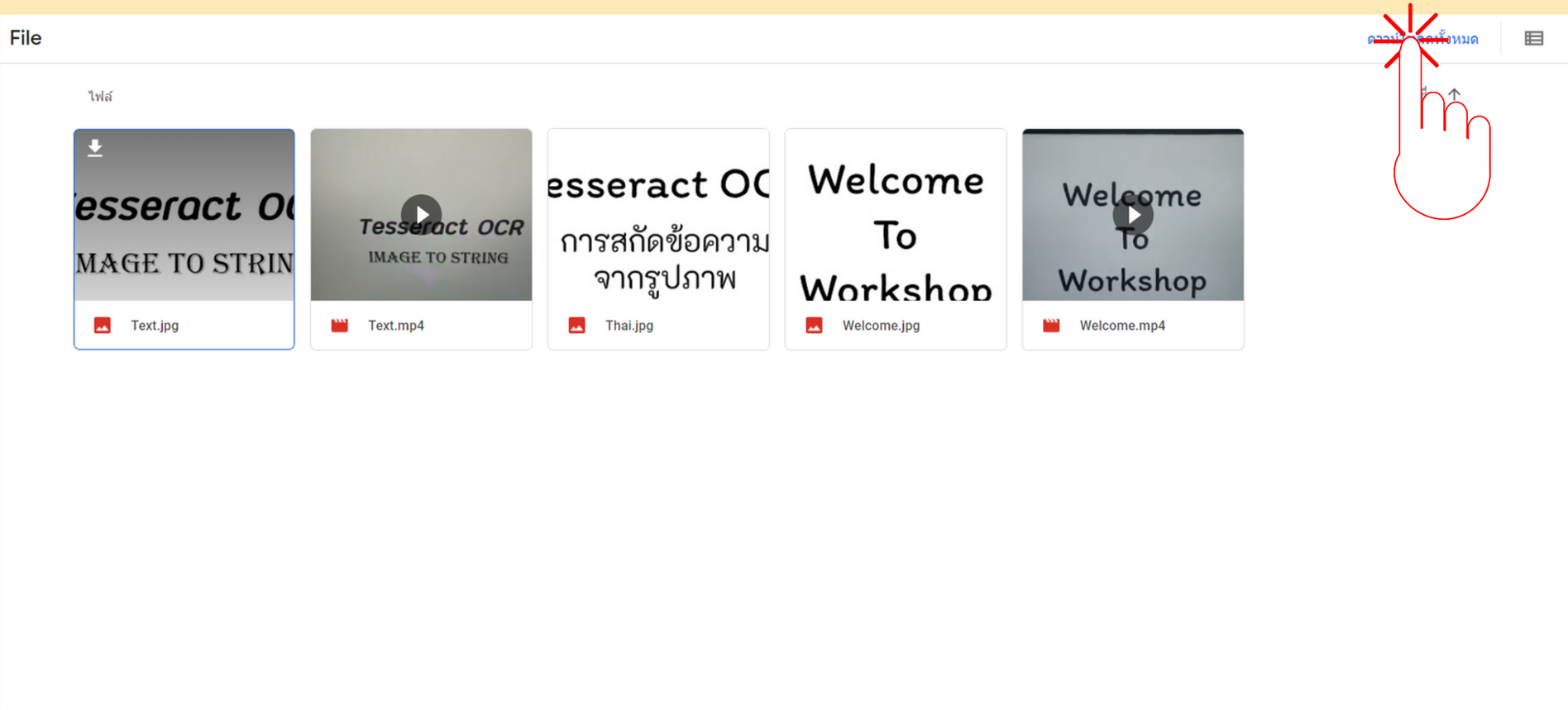
```
PS D:\python3> pip install pytesseract
Requirement already satisfied: pytesseract in c:\users\jirawat\anaconda3\lib\site-packages (0.3.8)
Requirement already satisfied: Pillow in c:\users\jirawat\anaconda3\lib\site-packages (from pytesseract) (8.2.0)
PS D:\python3>
```

>pip install pytesseract

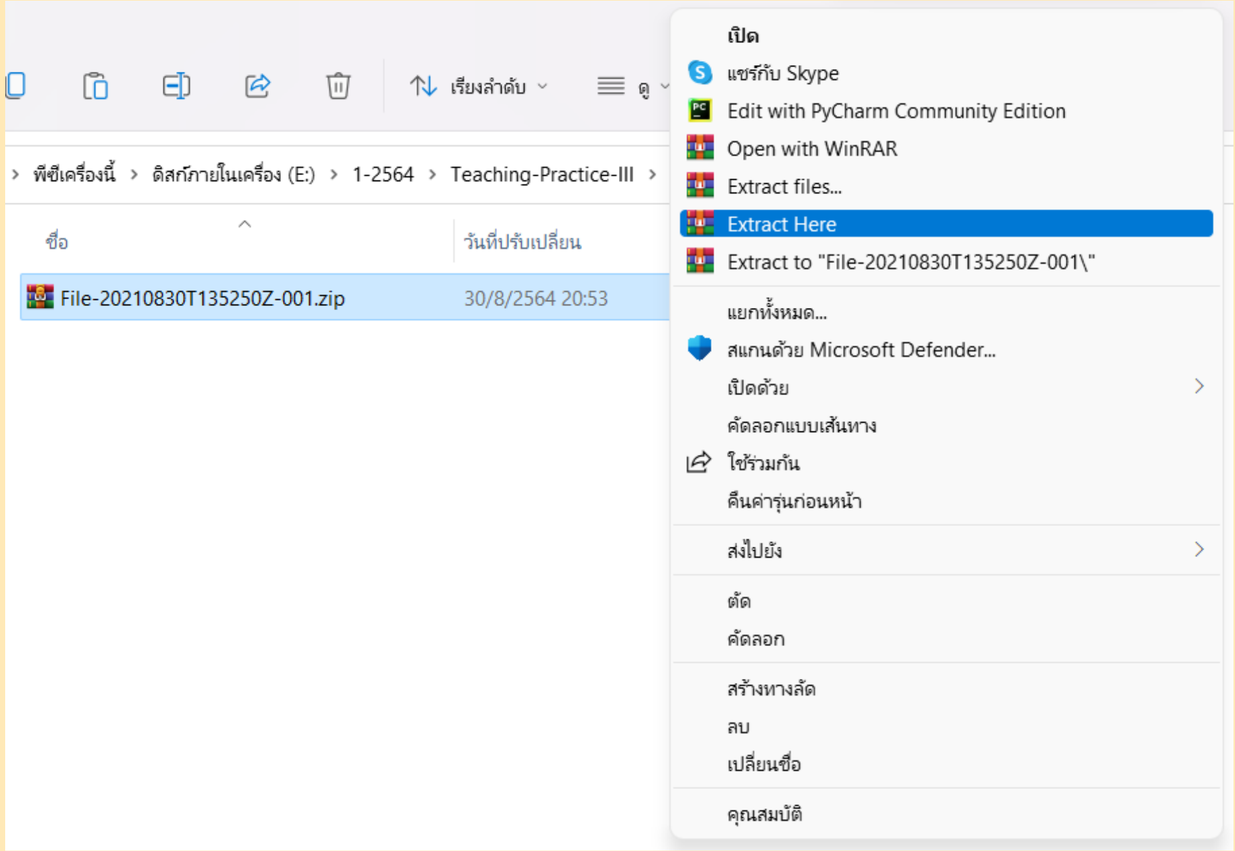
ดาวน์โหลดไฟล์ที่จะใช้งาน



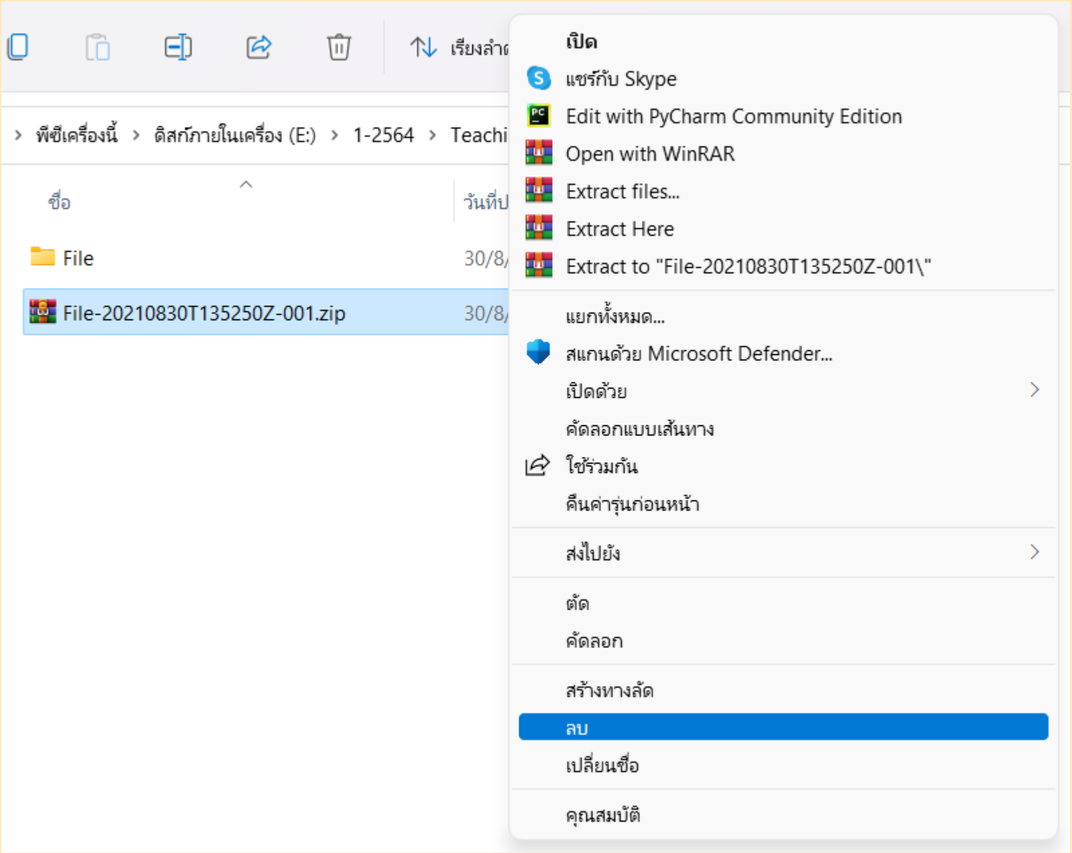
ดาวน์โหลดไฟล์ที่จะใช้งาน



ดาวน์โหลดไฟล์ที่จะใช้งาน

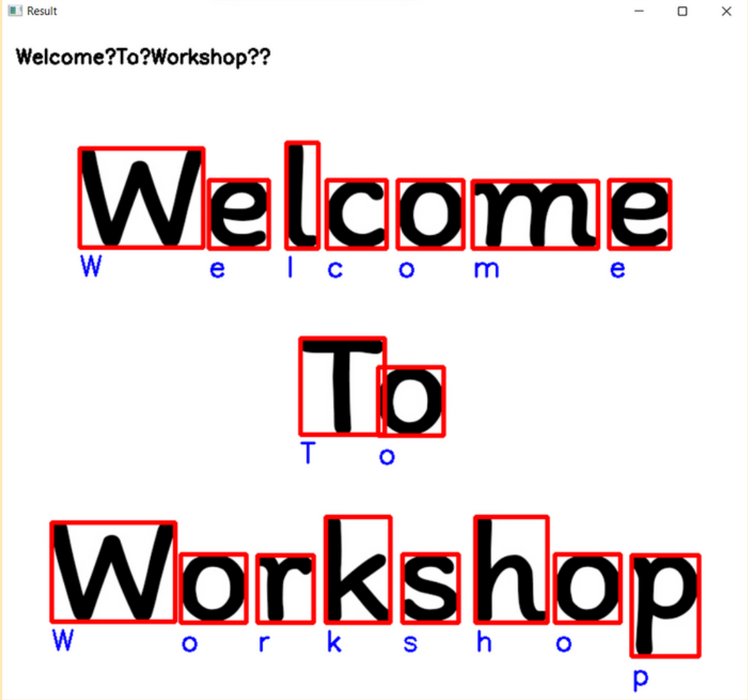


ดาวน์โหลดไฟล์ที่จะใช้งาน



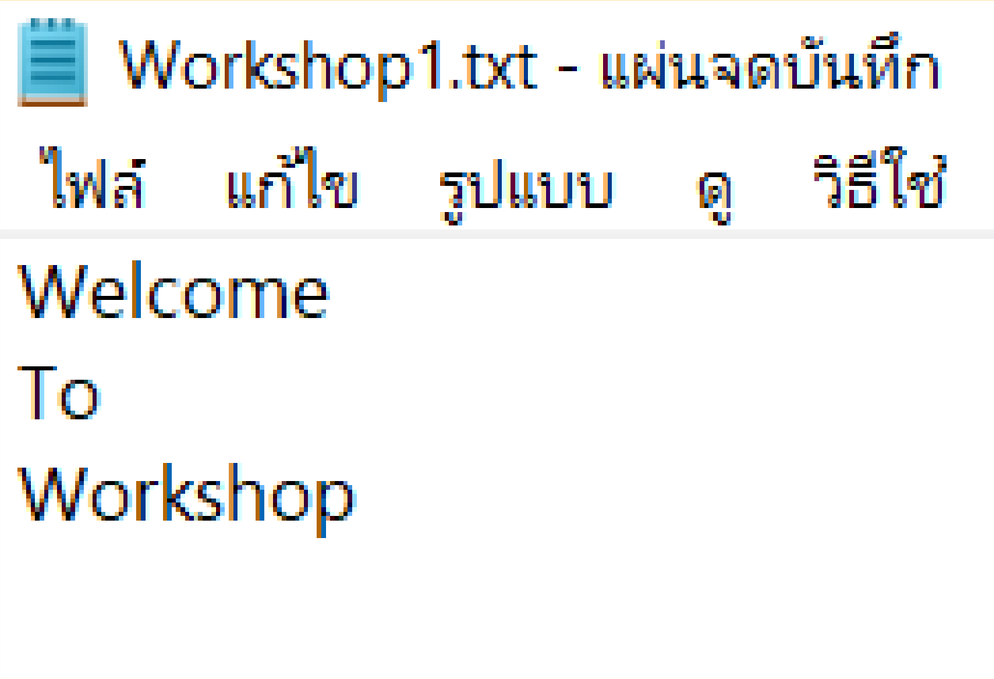
Example 1

การสกัดข้อความจากรูปภาพ



Workshop 1

โปรแกรมสกัดข้อความจากรูปภาพ



สิ่งที่ต้องเตรียม

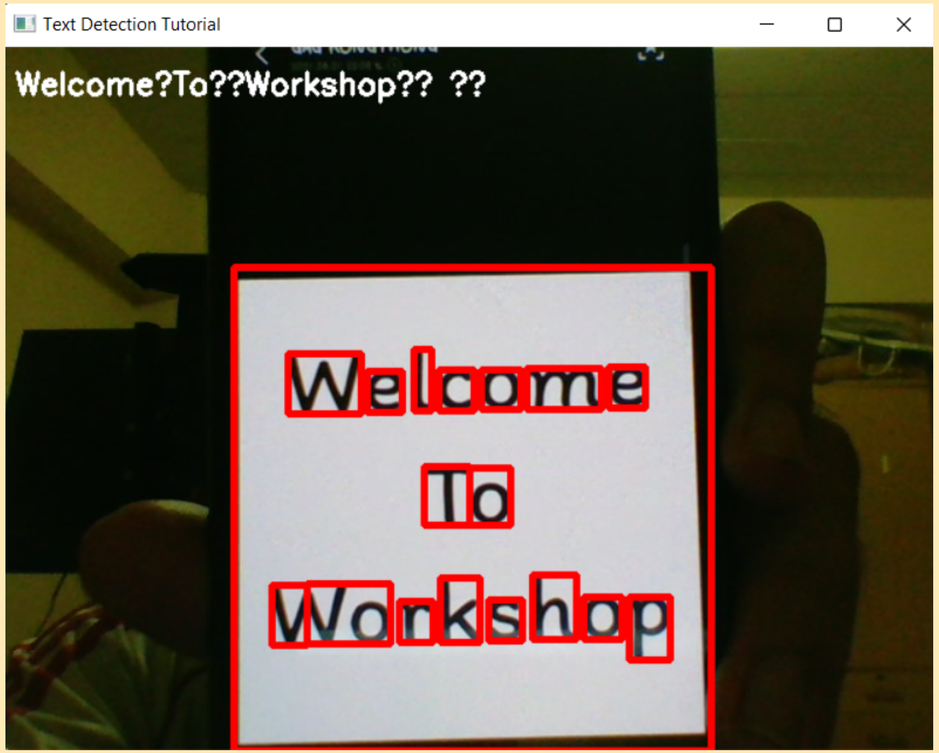
1. เครื่องคอมพิวเตอร์
2. Web Browser
3. ไฟล์รูปภาพ Welcome.jpg
4. ไฟล์รูปภาพ Text.jpg

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
```

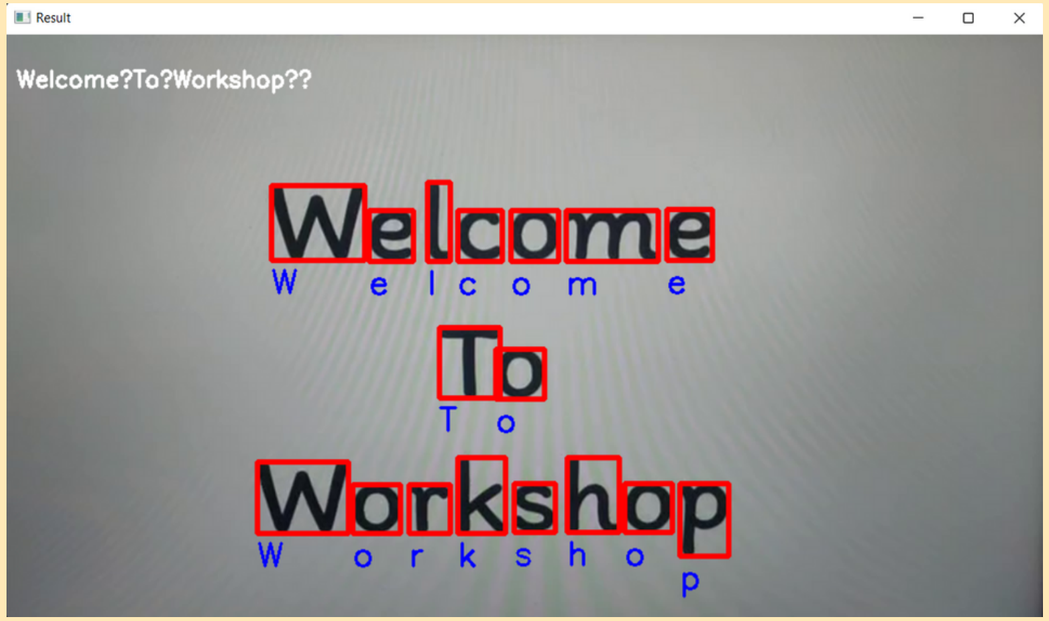
Example 2

การสกัดข้อความแบบตามเวลาจริง



Workshop 2

โปรแกรมสกัดข้อความแบบตามวิดีโอ



สิ่งที่ต้องเตรียม

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

โค้ดเพิ่มเติม

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

Image Recognition With Python

Face Recognition

LECTURER

PANUMAS SEEDAENG

4th year Student

KMUTNB
INVENTION TO INNOVATION

LECTURER

PHANUWIT KHWANNU

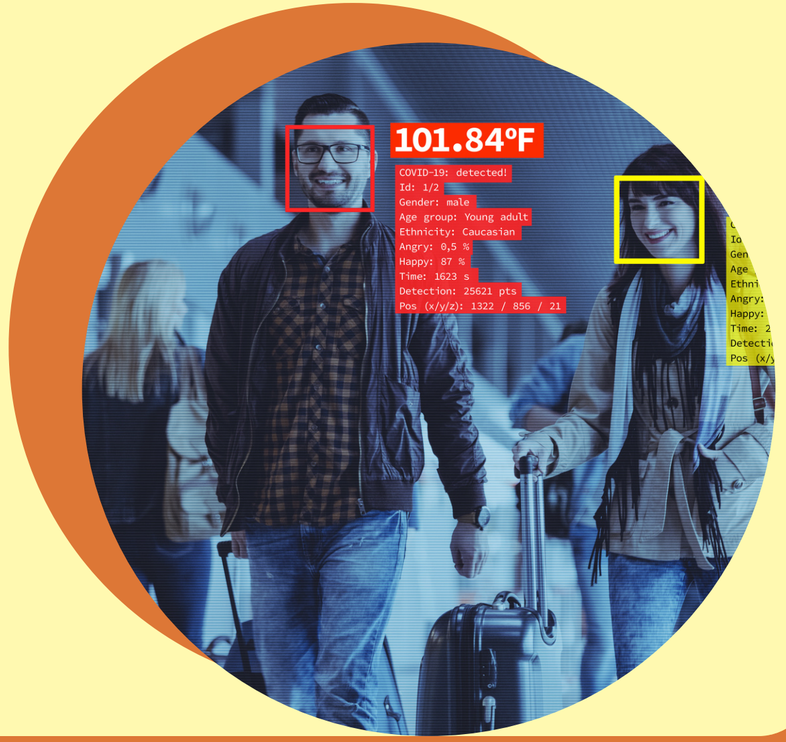
4th year Student

Facebook's facial recognition tools raise privacy concerns
Watch later
Share
Watch on YouTube

Face Recognition

ระบบการจำใบหน้า หรือ ระบบการจดจำใบหน้า

คือ ระบบการตรวจหาใบหน้าของมนุษย์
และปรับภาพใบหน้าโดยอัตโนมัติ



ระบบการรู้จำใบหน้า

- กรอบจะปรากฏขึ้นบนใบหน้าที่ถูกตรวจจับ
- โฟกัสสีโดยอัตโนมัติ
- ค่าการวัดแสงโดยอัตโนมัติ
- วัดลักษณะเฉพาะบนใบหน้า
    - ระยะห่างระหว่างดวงตา
    - ความกว้างของจมูก
    - ความลึกเบ้าตา
    - เป็นลักษณะสำคัญ

- เมื่อบันทึกด้วยคุณภาพแบบ HD เทคโนโลยีการบีบอัดจะจัดสรรความจุของข้อมูลให้ลดลง
- แต่ได้ข้อมูลที่เป็นประโยชน์มากขึ้นเพื่อปรับคุณภาพของภาพ
- ข้อมูลที่ได้จะถูกนำไปเปรียบเทียบกับข้อมูลตัวอย่างที่เก็บบันทึกไว้

ประวัติของการจดจำใบหน้า

1960s

จุดกำเนิดของเทคโนโลยี Facial Recognition สามารถจำแนกภาพ ใบหน้าแบบ Manua ผ่านการใช้งาน RAND Tablet สามารถใช้ในการบันทึกและจำแนกอัตลักษณ์ต่างๆ บนใบหน้า อาทิ ตา หู จมูก ปาก และไรผม

1970s

พัฒนาความแม่นยำให้มีความแม่นยำมากขึ้น โดยระบุอัตลักษณ์เจาะจง 21 จุด อาทิ สีผม ความหนาของริมฝีปาก โดยสามารถจำแนกใบหน้าได้โดยอัตโนมัติ แต่ระบบยังต้องใช้การคำนวณวัดและระบุตำแหน่งจากผู้ใช้เช่นเดิม

1980s / 90s

ใช้การวิเคราะห์ องค์ประกอบหลักกับใช้เทคนิคพีชคณิตเชิงเส้นทั่วไปเพื่อแก้ไขจุดบกพร่องเดิมที่ต้องคำนวณวัดค่าและระบุตำแหน่งแบบ Manua

1990s / 2000s

โปรแกรม FERET สร้างฐานข้อมูลภาพใบหน้า ในชุดทดสอบมีภาพนิ่ง 2,413 ภาพ คิดเป็น 856 คน การทำภาพทดสอบสำหรับการจดจำใบหน้าได้สร้างแรงบันดาลใจให้เกิดนวัตกรรมและส่งผลให้เทคโนโลยีการจดจำใบหน้ามีประสิทธิภาพมากขึ้น

2000s

มีคนกังวลเรื่องสิทธิส่วนบุคคล (Privacy) และความปลอดภัย รัฐบาลสหรัฐฯ จึงบังคับใช้กฎหมายสำหรับข้อมูลที่ใช้ในการปรับใช้เทคโนโลยีการจดจำใบหน้า

2006

อัลกอริธึมการจดจำใบหน้าล่าสุดที่มีอยู่ มีการใช้ภาพใบหน้าความละเอียดสูงการสแกนใบหน้า 3 มิติและภาพม่านตาในการทดสอบ ผลการวิจัยระบุว่าอัลกอริธึมใหม่มีความแม่นยำมากกว่าอัลกอริทึมการจดจำใบหน้าที่ผ่านมา

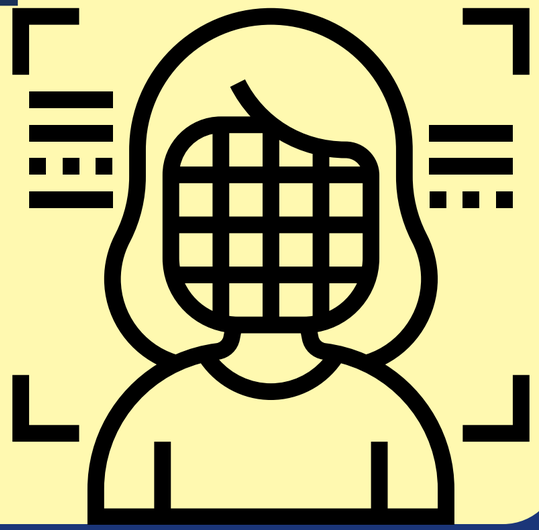
2010-ปัจจุบัน

เริ่มใช้ฟังก์ชันการจดจำใบหน้าที่ช่วยระบุผู้คนที่ใบหน้าอาจมีรูปภาพที่ผู้ใช้ร่วมกับการจดจำใบหน้าให้เข้ากับชีวิตประจำวันของเรา

ระบบจดจำใบหน้าในปัจจุบัน

เทคโนโลยีการจดจำใบหน้าก้าวหน้าอย่างรวดเร็วตั้งแต่ปี 2010 เป็นต้นมา เริ่มจากแอปเปิ้ลเปิดตัว iPhone X - ผู้ใช้ iPhone คนแรกสามารถปลดล็อคด้วย FaceID เป็นการใช้ในแพร่หลายไปยังสถานที่ใหญ่ ๆ อย่างสนามกีฬา คอนเสิร์ต หรือเรื่องประวันแบบการเข้าใช้งานแอปพลิเคชัน ยืนยันตัวตนการมาทำงานหรือมาเรียน

การตรวจหาใบหน้า



การตรวจหาตำแหน่งของใบหน้า

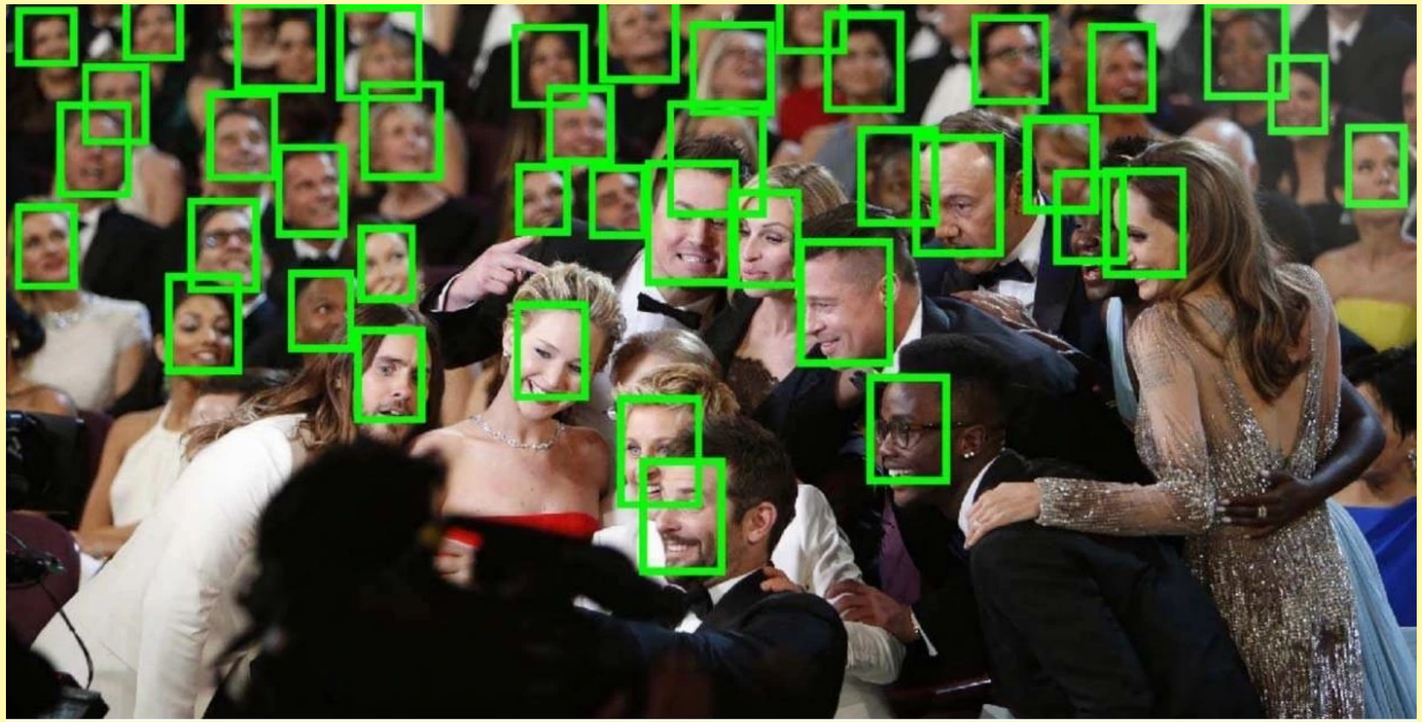
เป็นการตรวจหาใบหน้าทั่ว ๆ ไป ยังไม่สามารถแยกแยะได้ว่า หน้านั้นเป็นหน้าของใครตำแหน่งของ จมูก และ ตาทั้งสองข้าง ในรูปของเมตริกซ์ แนวตั้งและแนวนอน

วิธีการทำงานของ Face Recognition

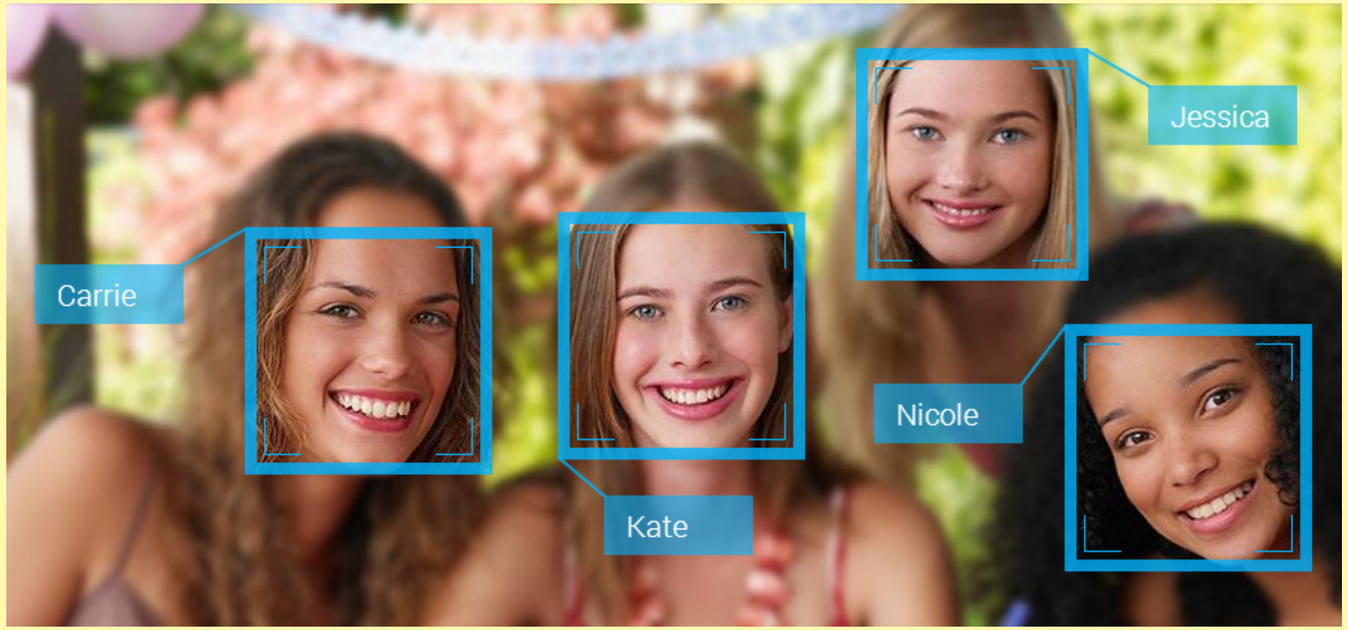
- Face Detection
- Face Recognition

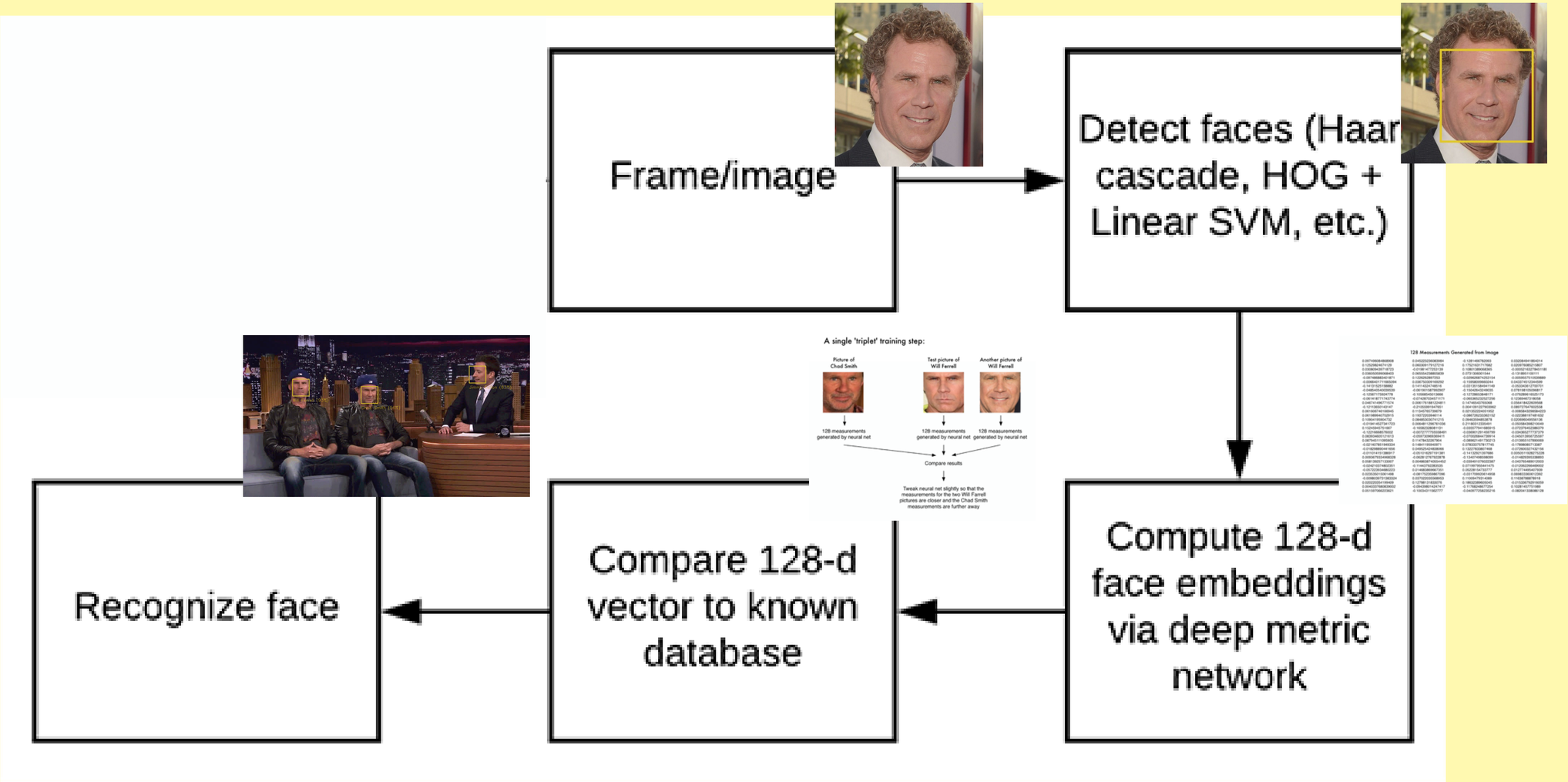
การตรวจจับใบหน้า Face Detection

ค้นหาใบหน้าของบุคคลจากภาพหรือวีดีโอ โดยไม่ได้ระบุบละเอียด

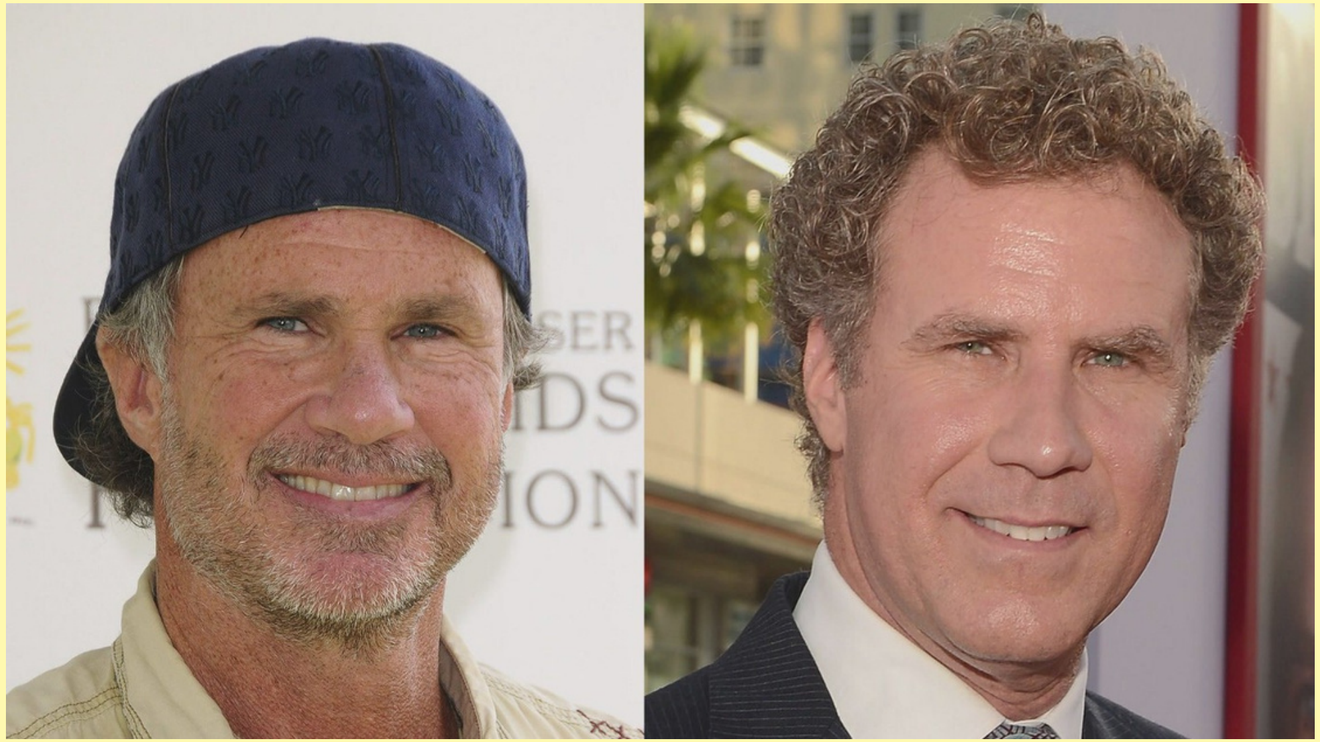


การจดจำใบหน้า Face Recognition
เปรียบเทียบกับข้อมูลใน data base ที่มีอยู่ เพื่อให้สามารถระบุได้ว่าบุคคลผู้นั้นเป็นใคร

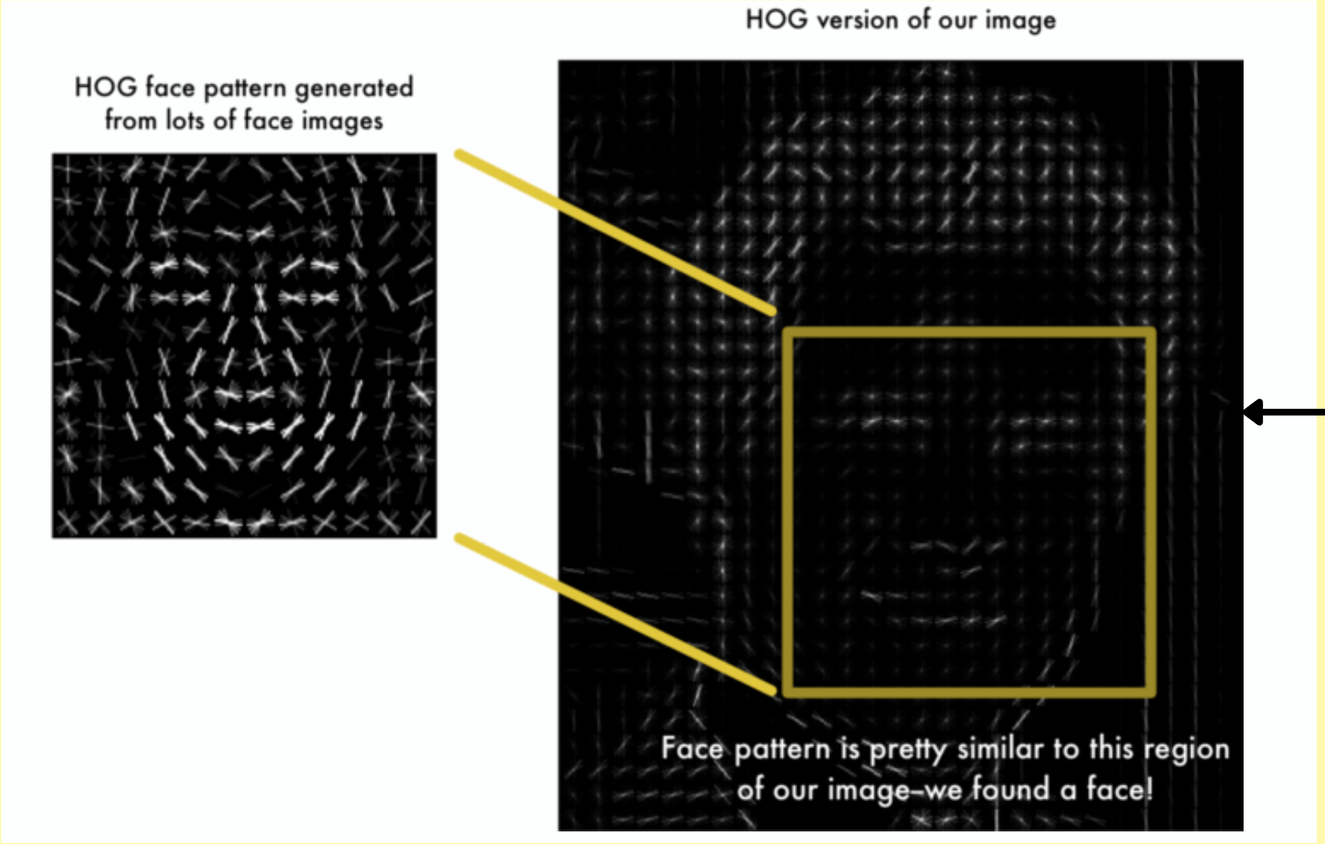


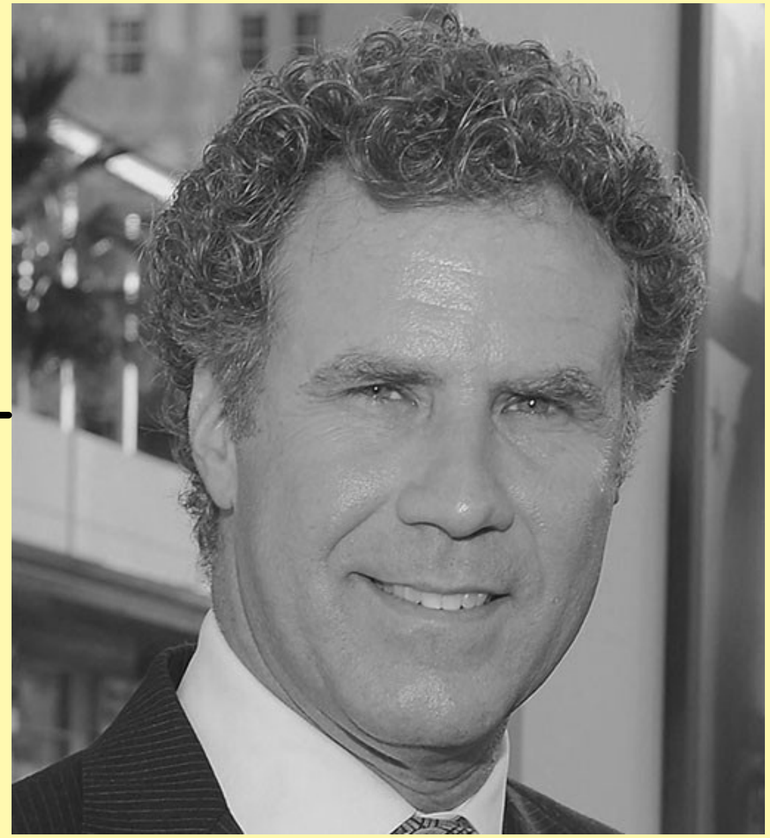

Frame/image
Detect faces (Haar cascade, HOG + Linear SVM, etc.)
Compute 128-d face embeddings via deep metric network
Compare 128-d vector to known database
Recognize face
A single 'triplet' training step:
Picture of Chad Smith
Test picture of Will Ferrell
Another picture of Will Ferrell
128 measurements generated by neural net
Compare results
Tweak neural net slightly so that the measurements for the two Will Ferrell pictures are closer and the Chad Smith measurements are further away
128 Measurements Generated from Image

Dataset (input)

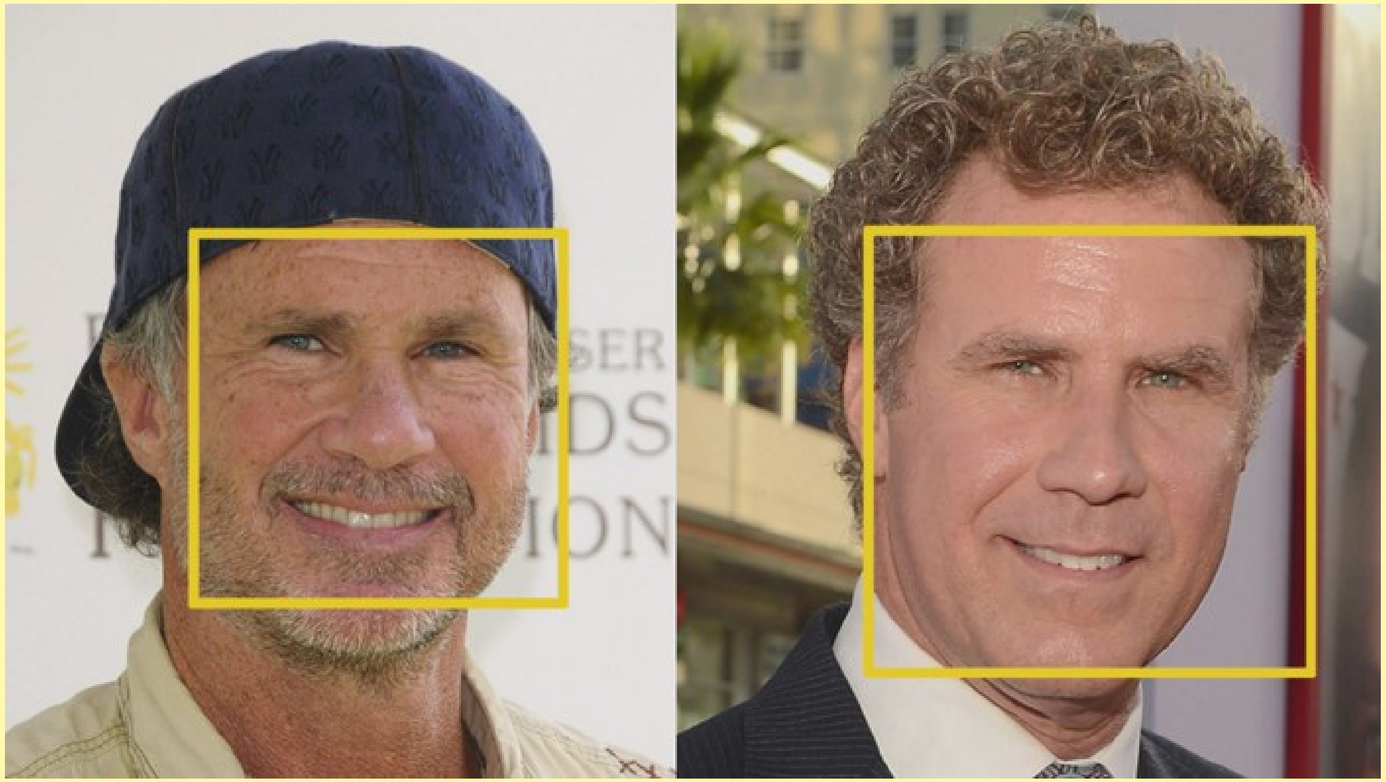


Histogram of oriented gradients (HOG)

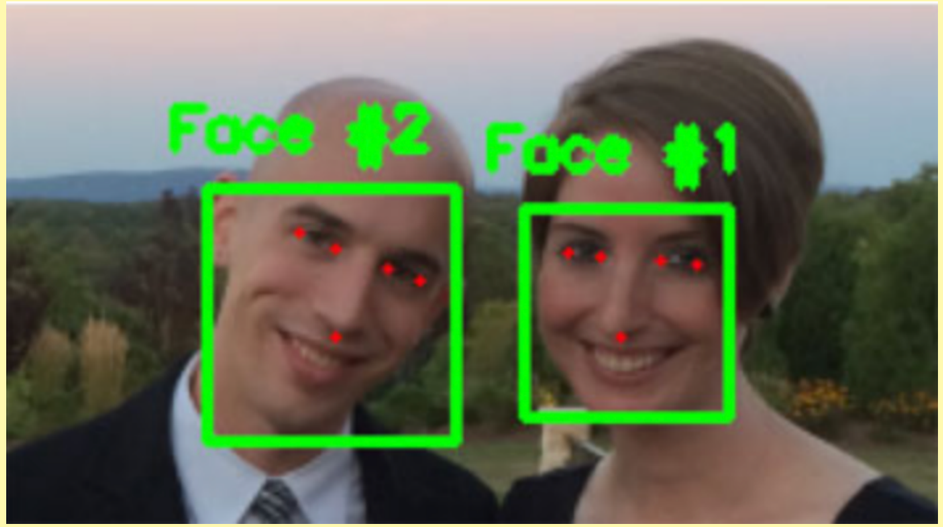


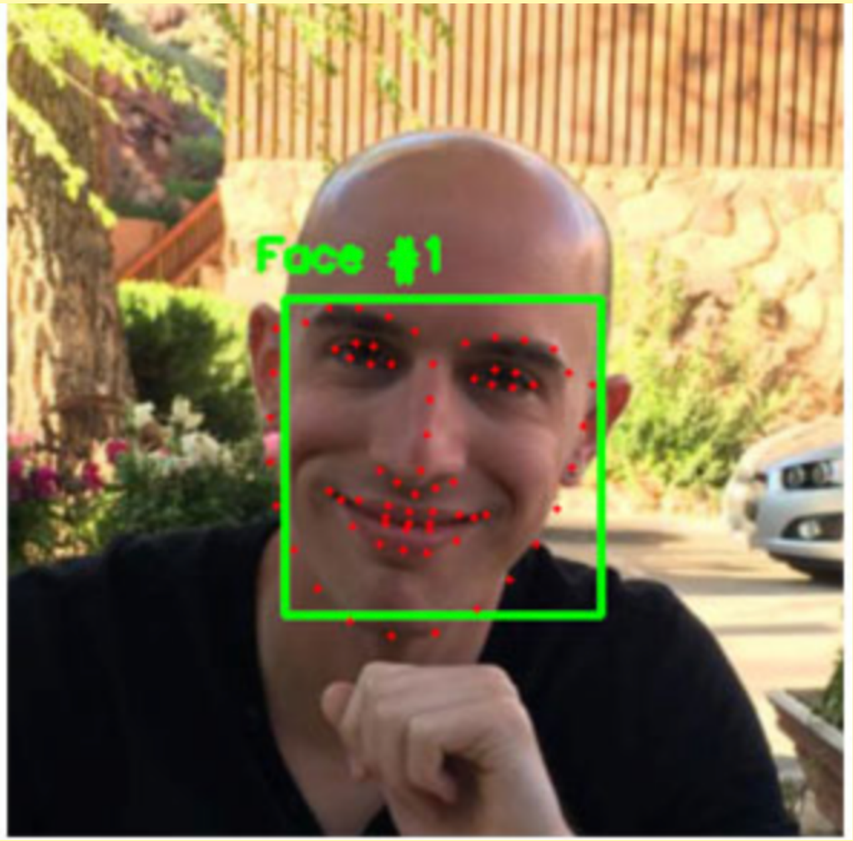


Histogram of oriented gradients (HOG)



Histogram of oriented gradients (HOG)

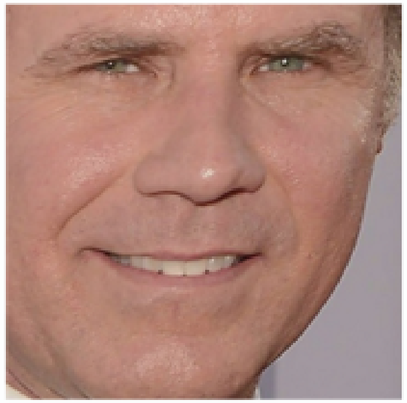




A comparison of the 5-point facial landmarks (left) and the dlib 68-point facial landmarks (right).

Histogram of oriented gradients (HOG)

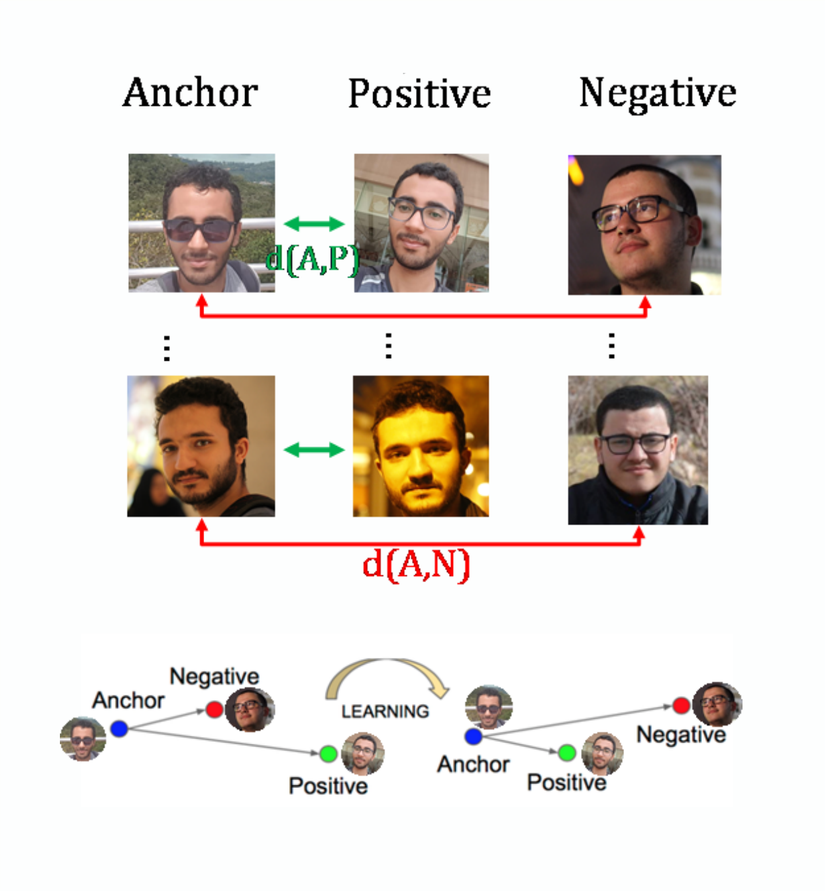
Input Image

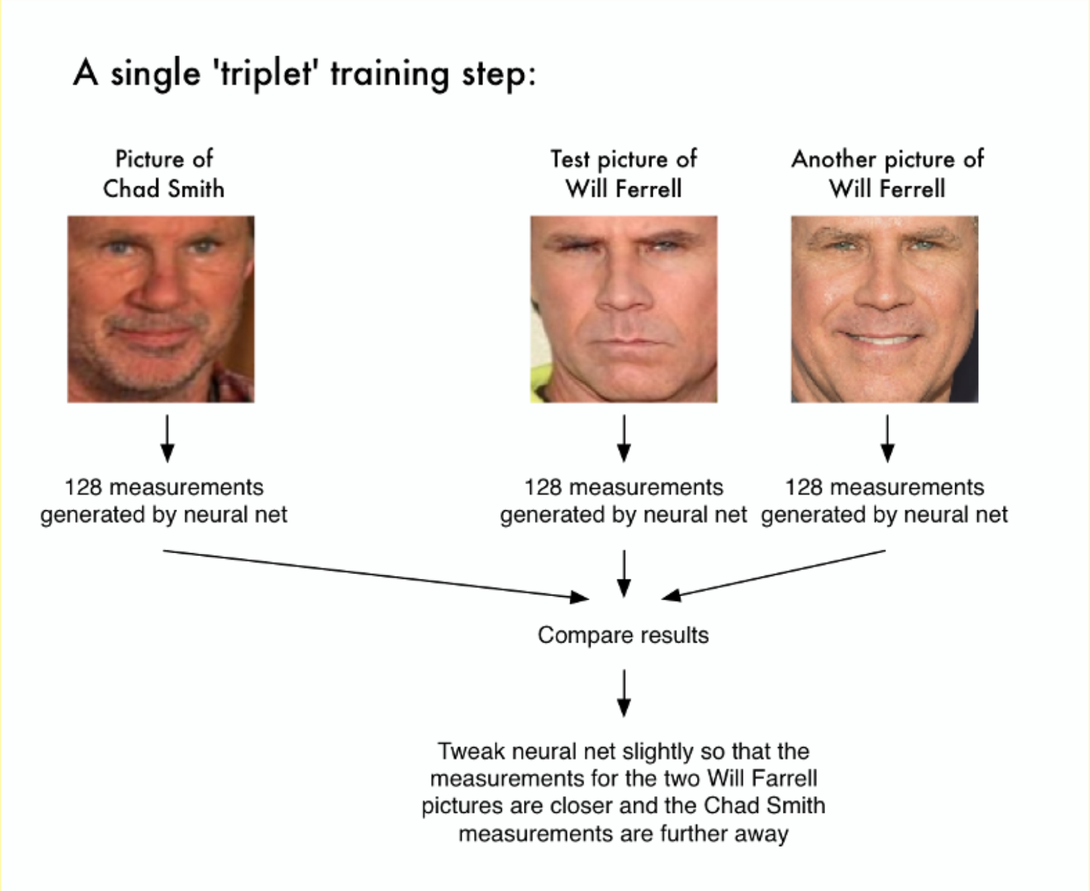


128 Measurements Generated from Image

0.097496084868908	0.045223236083984	-0.1281466782093	0.032084941864014
0.12529824674129	0.060309179127216	0.17521631717682	0.020976085215807
0.030809439718723	-0.01981477253139	0.10801389068365	-0.00052163278451189
0.036050599068403	0.065554238855839	0.0731306001544	-0.1318951100111
-0.097486883401871	0.1226262897253	-0.029626874253154	-0.0059557510539889
-0.0066401711665094	0.036750309169292	-0.15958009660244	0.043374512344599
-0.14131525158882	0.14114324748516	-0.031351584941149	-0.053343612700701
-0.048540540039539	-0.061901587992907	-0.15042643249035	0.078198105096817
-0.12567175924778	-0.10568545013666	-0.12728653848171	-0.076289616525173
-0.061418771743774	-0.074287034571171	-0.065365232527256	0.12369467318058
0.046741496771574	0.006176188122481 1	0.14746543765068	0.056418422609568
-0.12113650143147	-0.21055991947851	0.004109122790396 2	0.089727647602558
0.061606746166945	0.11345765739679	0.021352224051952	-0.0085843298584223
0.061989940702915	0.19372203946114	-0.086726233363152	-0.022388197481632
0.10904195904732	0.084853030741215	0.09463594853878	0.020696049556136
-0.019414527341723	0.0064811296761036	0.21180312335491	-0.050584398210049
0.15245945751667	-0.16582328081131	-0.035577941685915	-0.072376452386379
-0.12216668576002	-0.0072777755558491	-0.036901291459799	-0.034365277737379
0.083934605121613	-0.059730969369411	-0.070026844739914	-0.045013956725597
0.087945111095905	0.11478432267904	-0.089621491730213	-0.013955107890069
-0.021407851949334	0.14841195940971	0.078333757817745	-0.17898085713387
-0.018298890441656	0.049525424838066	0.13227833807468	-0.072600327432156
-0.011014151386917	-0.051016297191381	-0.14132921397686	0.0050511928275228
0.009367933496832 8	-0.062812767922878	-0.13407498598099	-0.014829395338893
0.058139257133007	0.0048638740554452	-0.039491076022387	-0.043765489012003
-0.024210374802351	-0.11443792283535	0.071997955441475	-0.012062266469002
-0.057223934680223	0.014683869667351	0.05228154733777	0.012774495407939
0.023535015061498	-0.081752359867096	-0.031709920614958	0.069833360612392
-0.0098039731383324	0.037022035568953	0.11009479314089	0.11638788878918
0.020220354199409	0.12788131833076	0.18632389605045	-0.015336792916059
0.0040337680839002	-0.094398014247417	-0.11768248677254	0.10281457751989
0.0515970662236 21	-0.10034311562777	-0.040977258235216	-0.082041338086128

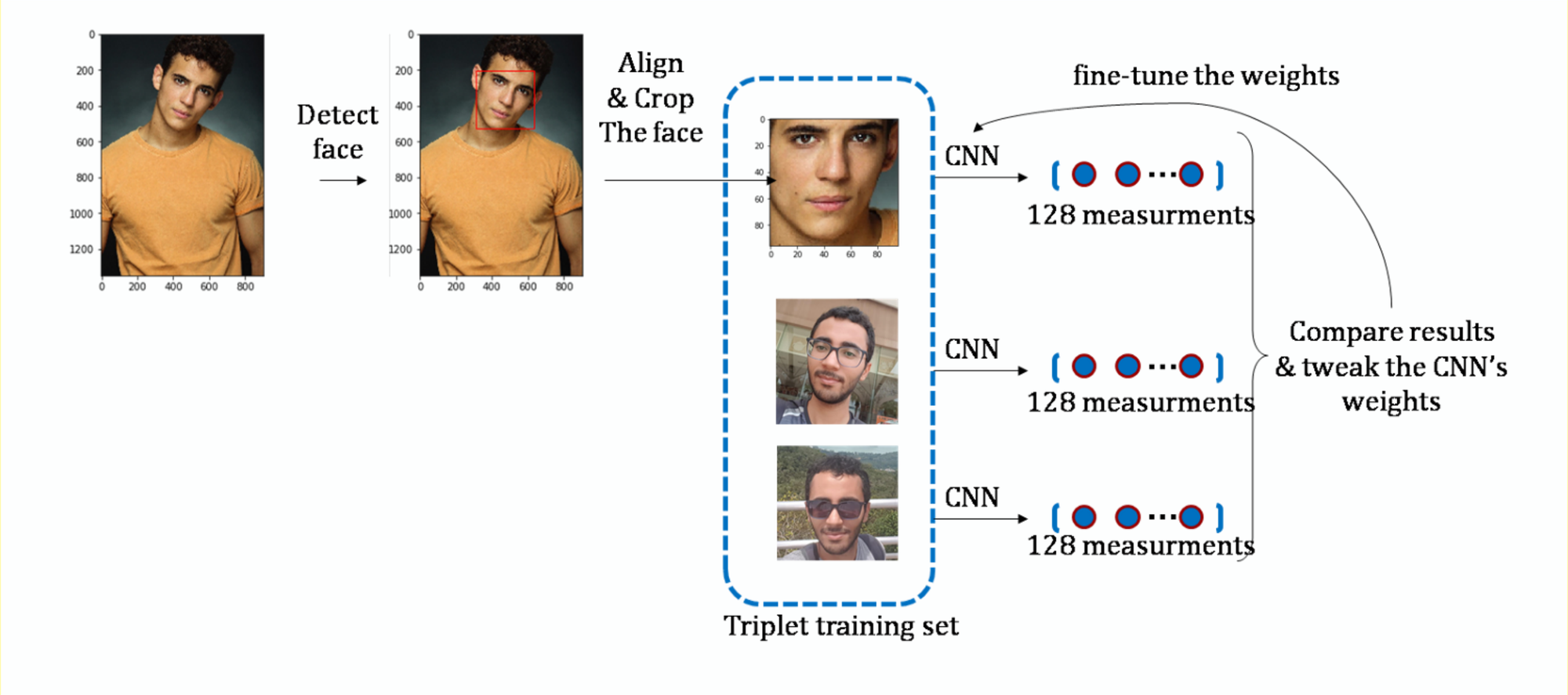
Triplet loss





Output

face_locations : CNN



ประโยชน์ของ Face Recognition

- จะจดจำใบหน้าของบุคคลที่ปรากฏบนกล้อง CCTV
  - ระบบรักษาความปลอดภัย
  - อุปกรณ์สแกนใบหน้าควบคุมการเปิดปิดประตู
  - ระบบลงเวลาพนักงาน
  - ระบบสแกนใบหน้าตามอาคาร

ติดตั้งส่วนเสริมการใช้งาน

เข้าไปที่ anacode และพิมพ์คำสั่งติดตั้ง

```
conda install -c anaconda cmake
conda install -c conda-forge dlib
conda install -c conda-forge opencv
conda install -c conda-forge face_recognition
```

Work shop 1
face_rec_img

เขียนโปรแกรม

Work shop 1 face_rec_img

เรียกใช้งาน library

```
import face_recognition

import os

import cv2
```

สร้างตัวแปรและกำหนดค่าตัวแปร

```
KNOWN_FACES_DIR = 'known_faces'
UNKNOWN_FACES_DIR = 'unknown_faces'
TOLERANCE = 0.6
MODEL = 'cnn'
```

สร้างตัวแปรและกำหนดค่าตัวแปร

```
print('Loading known faces...')
known_faces = []
known_names = []
```

ดึงภาพใบหน้าที่รู้จักมาเก็บในตัวแปร

```
for name in os.listdir(KNOWN_FACES_DIR):

    # Next we load every file of faces of known person
    for filename in
os.listdir(f'{KNOWN_FACES_DIR}/{name}'):

        # Load an image

Image=face_recognition.load_image_file(f'{KNOWN_FACES_
DIR}/{name}/{filename}')
```

ดึงภาพใบหน้าที่รู้จักมาเก็บในตัวแปร

```
# Get 128-dimension face encoding
# Always returns a list of found faces, for this purpose we take first face only (assuming one face per image as you can't be twice on one image)
encoding = face_recognition.face_encodings(image)[0]
```

ดึงภาพใบหน้าที่รู้จักมาเก็บในตัวแปร

```
# Append encodings and name
known_faces.append(encoding)
known_names.append(name)
```

Processing รูปภาพที่ไม่รู้จัก

```
print('Processing unknown faces...')
for filename in os.listdir(UNKNOWN_FACES_DIR):
```

Processing รูปภาพที่ไม่รู้จัก

```
print(f'Filename {filename}', end='')
image = face_recognition.load_image_file(f'{UNKNOWN_FACES_DIR}/{filename}')
locations = face_recognition.face_locations(image, model=MODEL)
encodings = face_recognition.face_encodings(image, locations)
image = cv2.cvtColor(image, cv2.COLOR_RGB2BGR)
```

เปรียบเทียบใบหน้ากับระบบ

```
for face_encoding, face_location in zip(encodings, locations):
        results = face_recognition.compare_faces(known_faces,face_encoding,
TOLERANCE)
    match = None
    if True in results:
        match = known_names[results.index(True)]
        print(f' - {match} from {results}')
```

สร้างกรอบสำหรับรูปภาพ

```
top_left = (face_location[3], face_location[0])
bottom_right = (face_location[1], face_location[2])
color = [0, 255, 0]
cv2.rectangle(image, top_left, bottom_right, color, 3)
```

สร้างกรอบสำหรับข้อความ

```
top_left = (face_location[3], face_location[2])
bottom_right = (face_location[1], face_location[2] + 22)
cv2.rectangle(image, top_left, bottom_right, color, cv2.FILLED)
cv2.putText(image, match, (face_location[3] + 10, face_location[2] + 15),
cv2.FONT_HERSHEY_SIMPLEX, 0.5, (200, 200, 200), 2)
```

แสดงผลการทำงานรูปภาพ

```
cv2.imshow(filename, image)
cv2.waitKey(10000)
#cv2.destroyWindow(filename)
```

กิจกรรมที่ 1

ให้ผู้เรียนทำการตรวจหาใบหน้าตนเองด้วยรูปภาพ

1. นำรูปภาพของตนเองเข้าสู่ไฟล์ know_faces
2. อัปโหลดรูปภาพของตนเองเข้าทดสอบที่โฟนเดอร์ unknow_faces
3. เปลี่ยนค่า TOLERANCE
4. การ Print screen ภาพเพื่อส่ง
5. อัพโหลดรูปที่ลิงค์ https://forms.gle/m8LN2EGKjJuGAQbGA

Work shop 2
face_rec_camera

เปลี่ยนเป็นการจากภาพเป็น camera

```
#UNKNOWN_FACES_DIR = 'unknown_faces'
video = cv2.VideoCapture(0) # could put in a filename
```

การดึงภาพจากวิดิณอมาประมวลผล

```
#for filename in os.listdir(UNKNOWN_FACES_DIR):
while True:
    #print(f'Filename {filename}', end='')
    #image = face_recognition.load_image_file(f'{UNKNOWN_FACES_DIR}/{filename}')
    ret, image = video.read()
    locations = face_recognition.face_locations(image, model=MODEL)
    encodings = face_recognition.face_encodings(image, locations)
    #image = cv2.cvtColor(image, cv2.COLOR_RGB2BGR)
```

การแสดงผล

```
cv2.imshow(filename, image)
if cv2.waitKey(1) & 0xFF == ord("q"):
    break

#cv2.waitKey(10000)
```