從 YOLOv4 到 YOLOv7 的發展過程

廖弘源

Scientist vs Engineer

Scientists use systematic methods to solve long-term problems

Engineers use existing methods to solve current problems

Outline of this talk

- MOST AI project: 2018/1 – 2021/12

  ``Smart City Traffic Flow Solutions''
- Object detection is the first step toward *EVERYTHING*

  - development history of object detectors

  - one-stage vs. two-stage object detectors
- Why PRN, CSPNet, and YOLOv4 were developed ?
- Why YOLOR, ELAN and then YOLOv7

The Objectives of this MOST sponsored Project

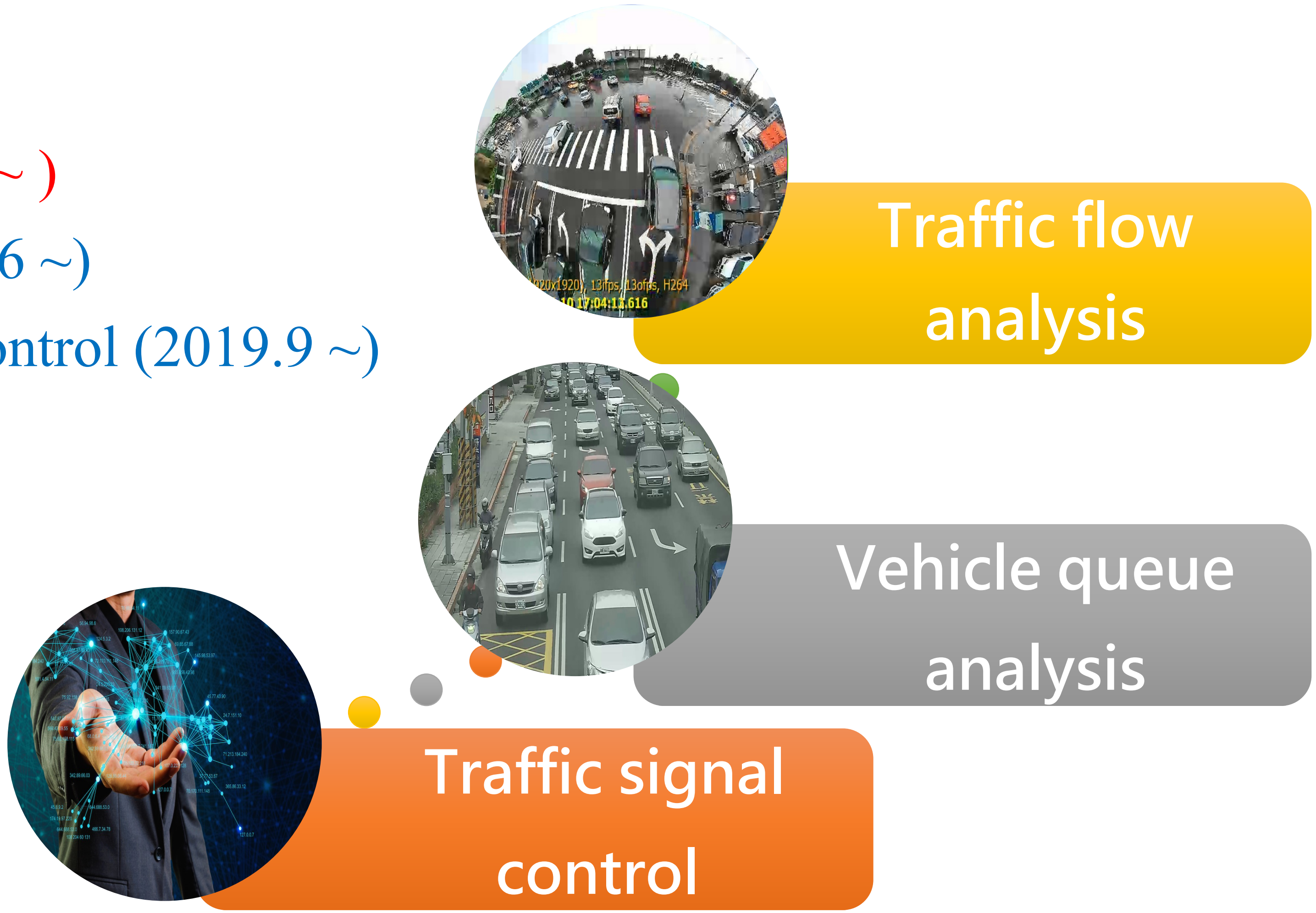
- Bring academic research and development capabilities into industry
- Commercialize artificial intelligence technology, let the products enter the international market
- Let the world see Taiwan

Consensuses of the team:(1) Select real-world issues given by industry; (2) Try to introduce science into problem-solving process

Smart City Traffic Flow Solutions

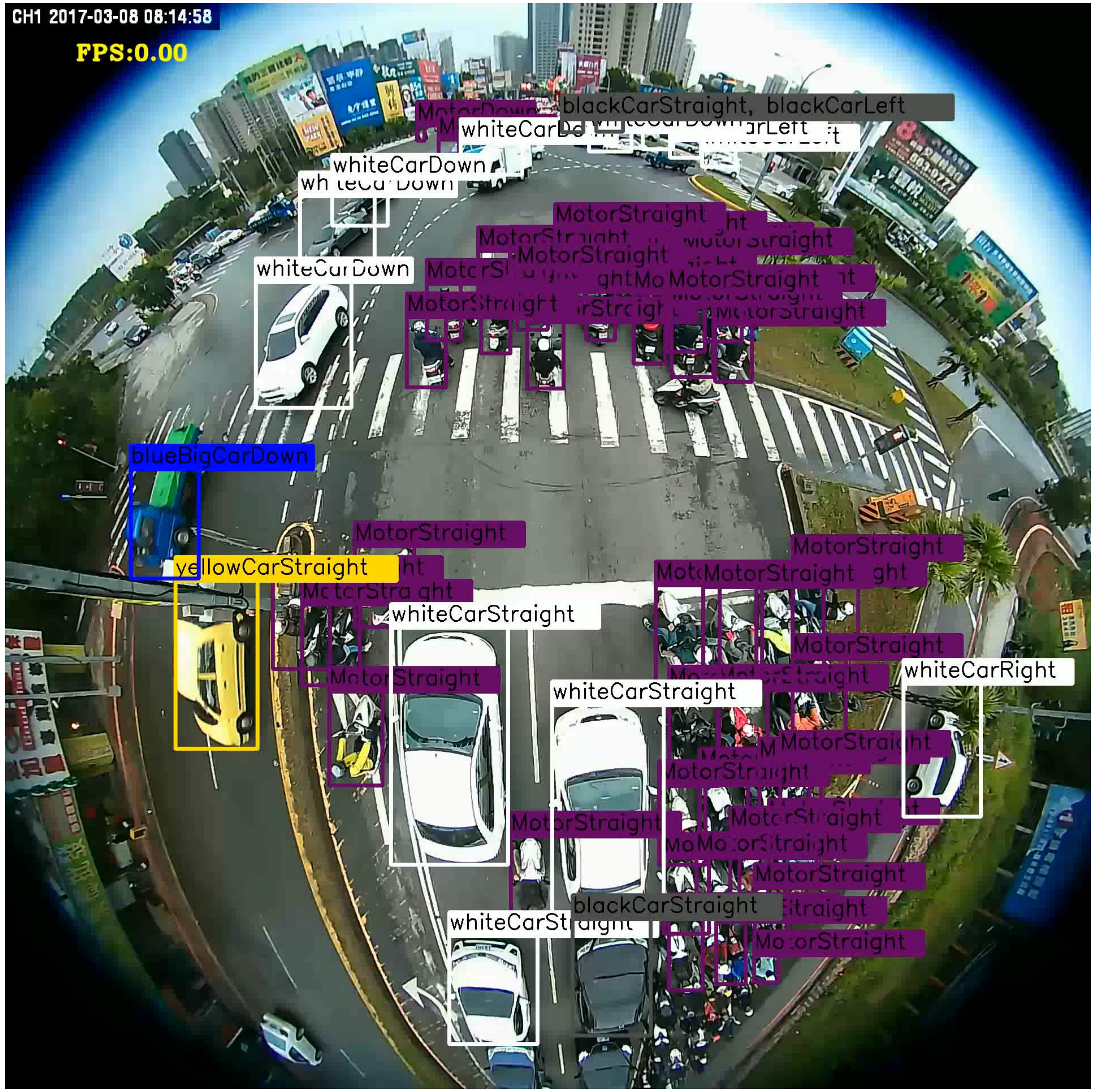
Development process and phased goals

- Traffic flow analysis (2018.6 ~)
- Vehicle queue analysis (2019.6 ~)
- Road network traffic signal control (2019.9 ~)



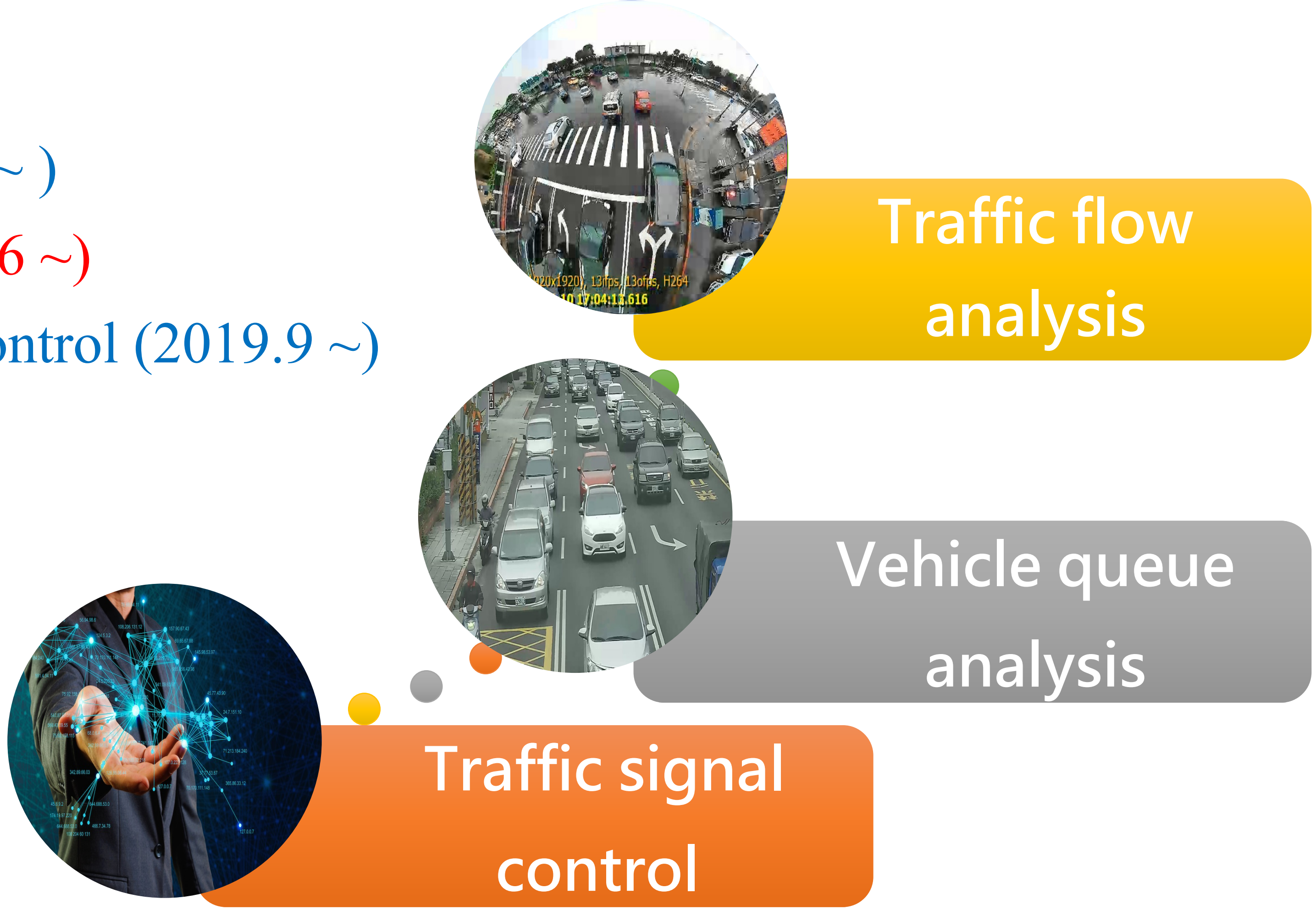
Demo video

- Nvidia Jetson TX2 real time
- Accomplished during 2018.6 ~ 2018.11

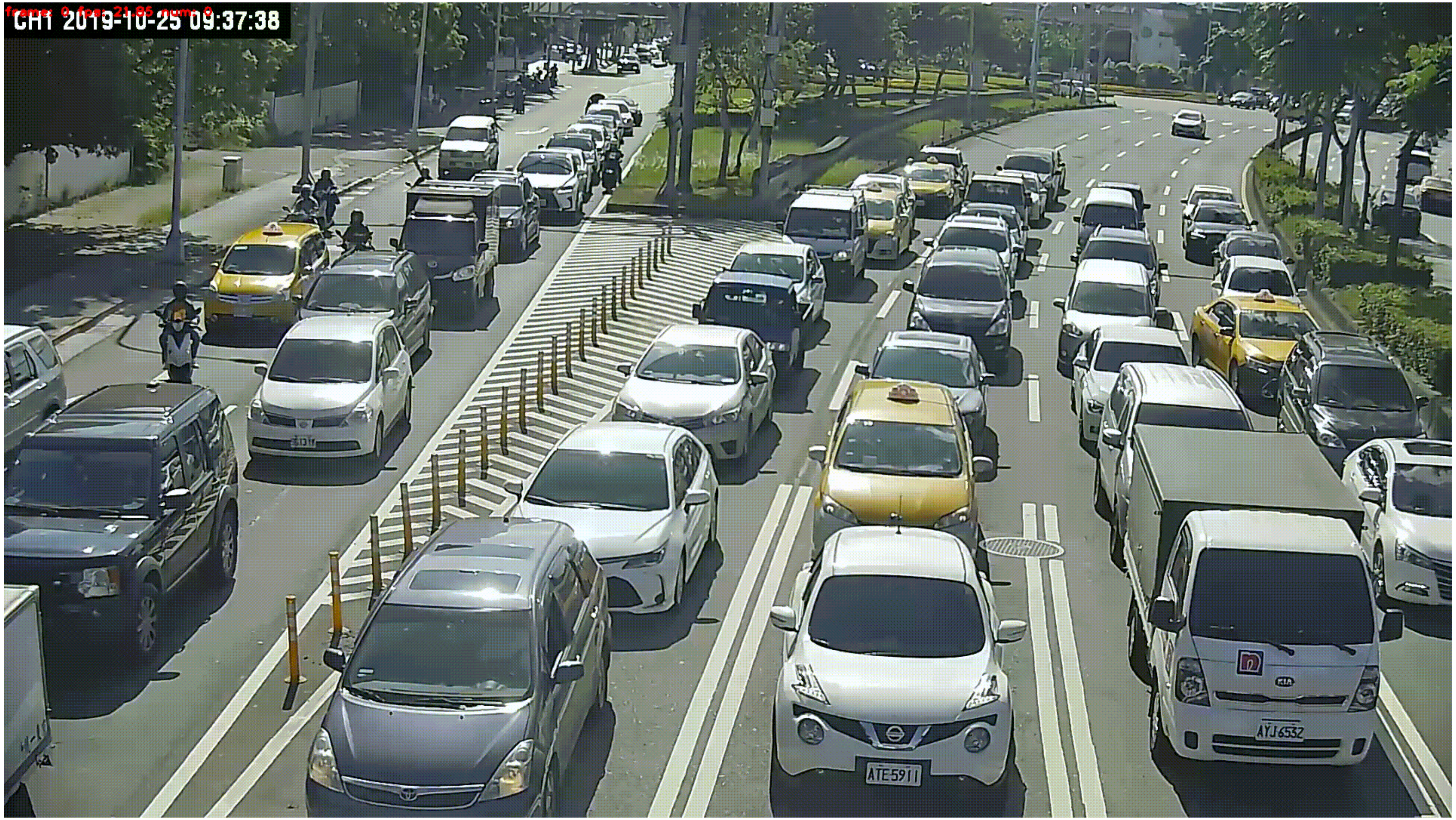


Development process and phased goals

- Traffic flow analysis (2018.6 ~)
- Vehicle queue analysis (2019.6 ~)
- Road network traffic signal control (2019.9 ~)



Vehicles detected by YOLOv4 during the day time

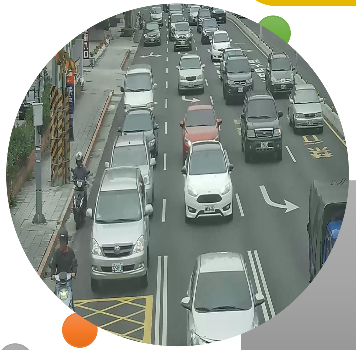


Vehicles detected by YOLOv4 during the night

Development process and phased goals

- Traffic flow analysis (2018.6 ~)
- Vehicle queue analysis (2019.6 ~)
- Road network traffic signal control (2019.9 ~)

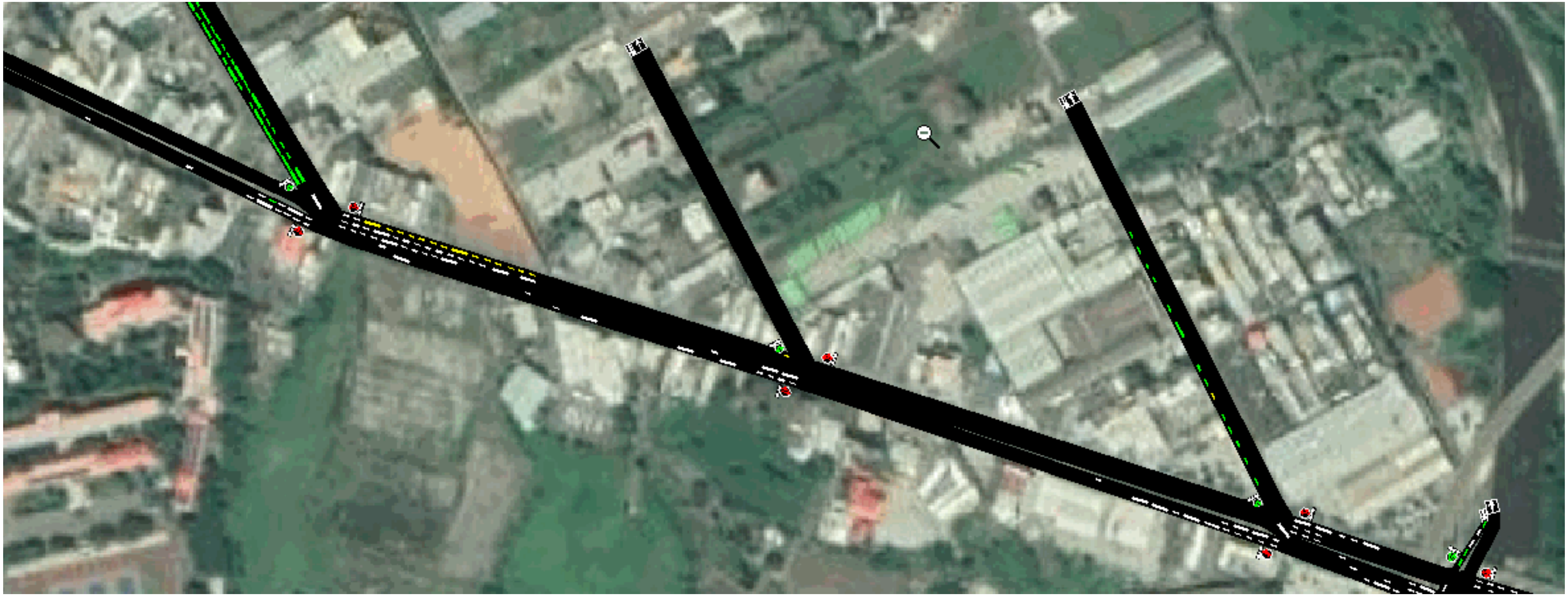
Traffic flow analysis



Vehicle queue analysis

Traffic signal control

Dynamic control of traffic signs



Outline of this talk

- MOST AI project: 2018/1 – 2021/12

  ``Smart City Traffic Flow Solutions’’
- Object detection is the first step toward *EVERYTHING*

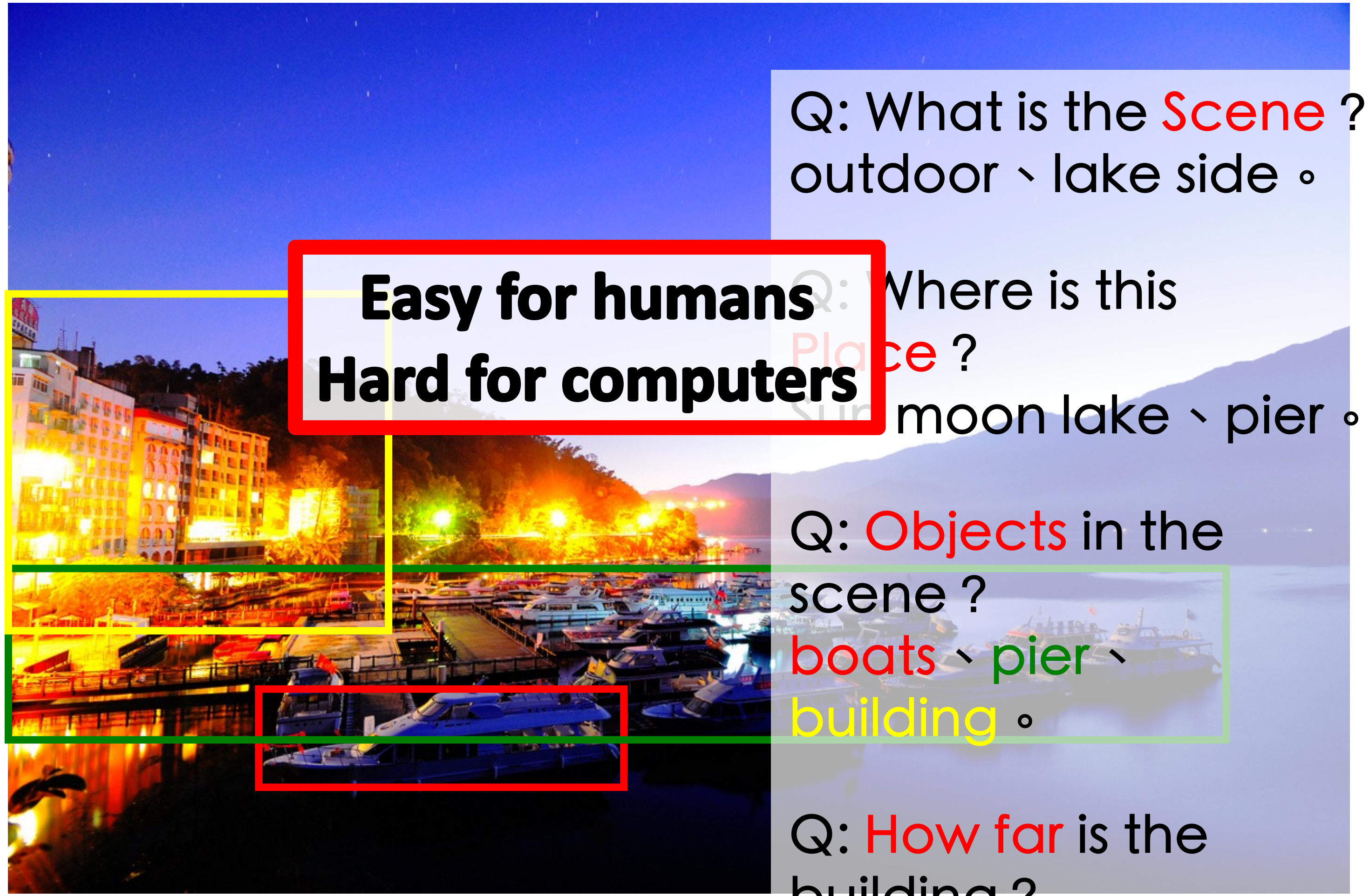
  - development history of object detectors

  - one-stage vs. two-stage object detectors
- Why PRN, CSPNet, and YOLOv4 were developed ?
- Why YOLOR, ELAN and then YOLOv7

物件偵測幾乎是所有領域相關研究
的第一步

以深度學習為基礎的物件偵測是如何開始的？

2010 以前處理一張未知影像到底有多難?



Easy for humans
Hard for computers

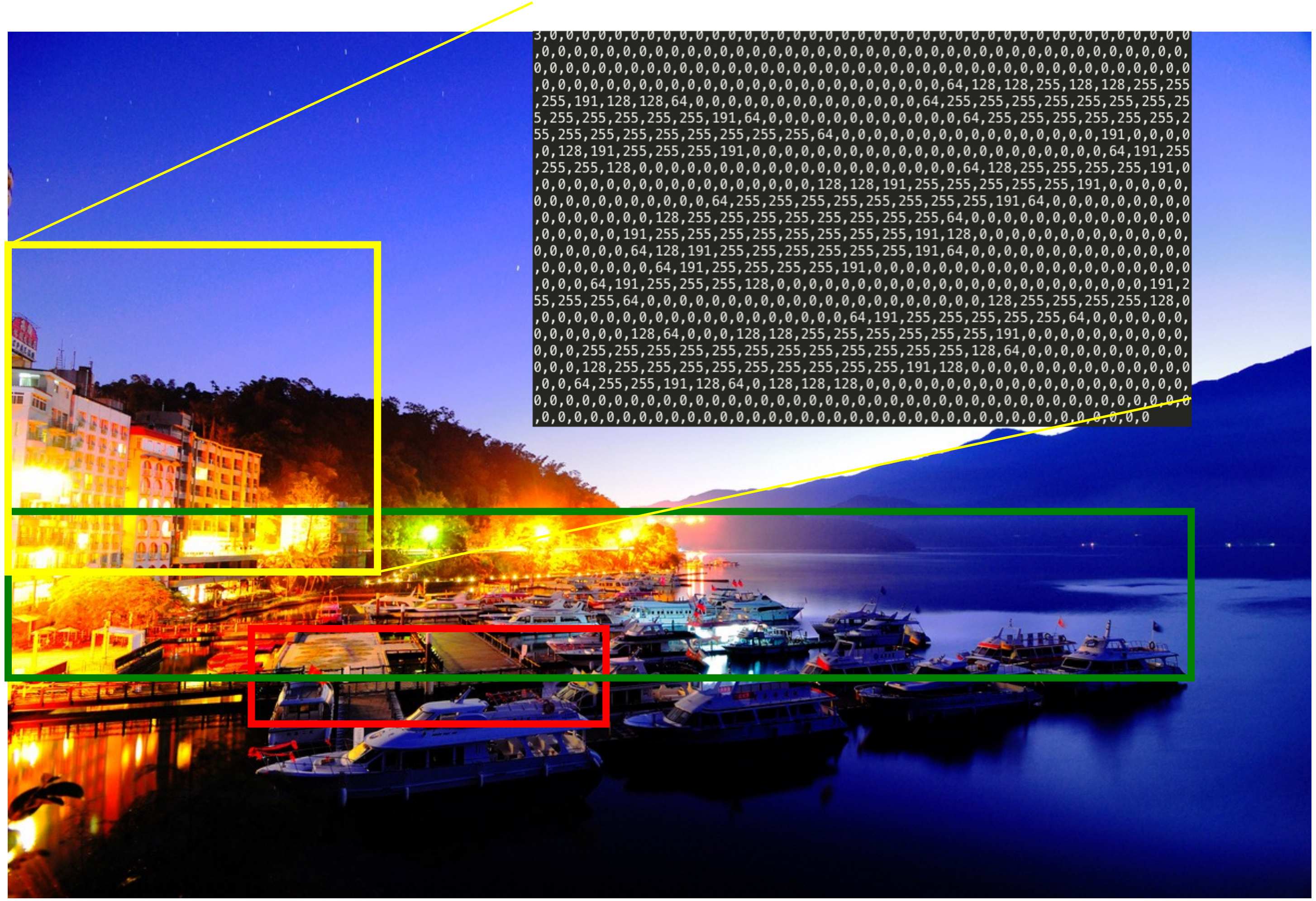
Q: What is the Scene ?
outdoor、lake side。

Q: Where is this Place ?
Sun moon lake、pier。

Q: Objects in the scene ?
boats、pier、building。

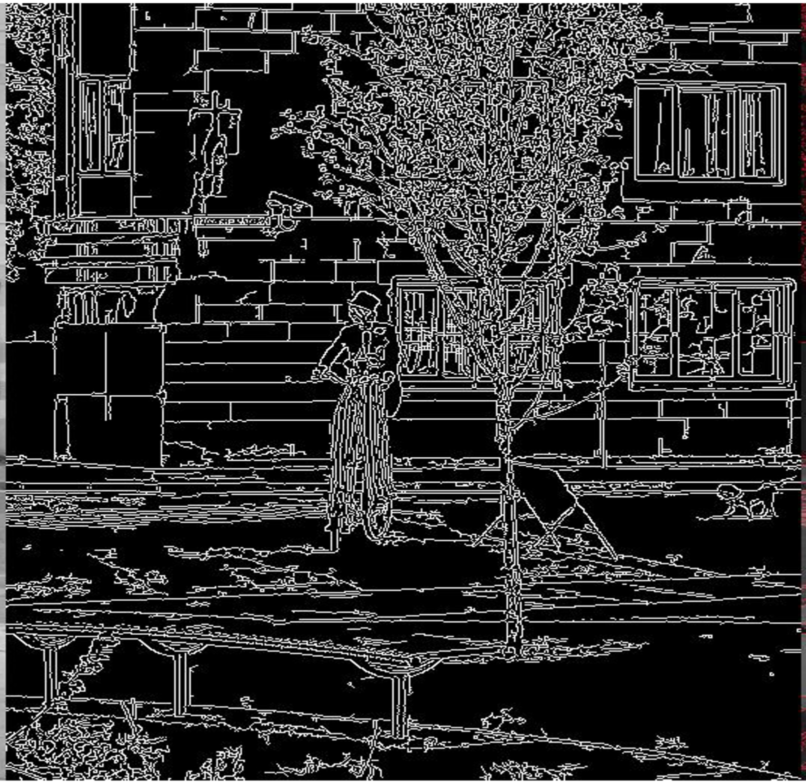
Q: How far is the building ?
50 meters。

2010 以前處理一張未知影像到底有多難?



常見的作法: edge detection, 希望為形成封閉曲線做準備,方便未來 segmentation 工作順利進行

original image

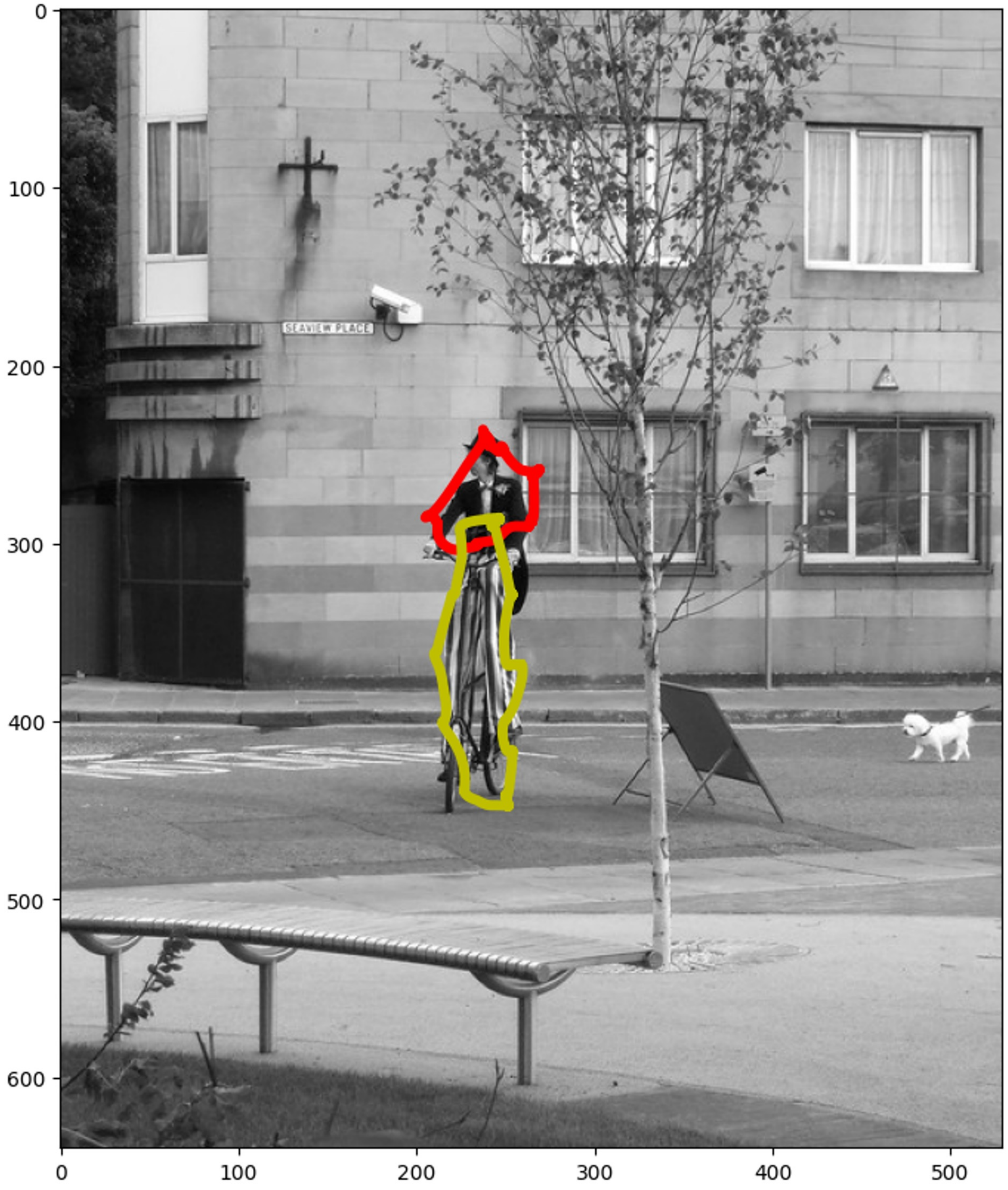


after canny edge detection

visualize canny edge

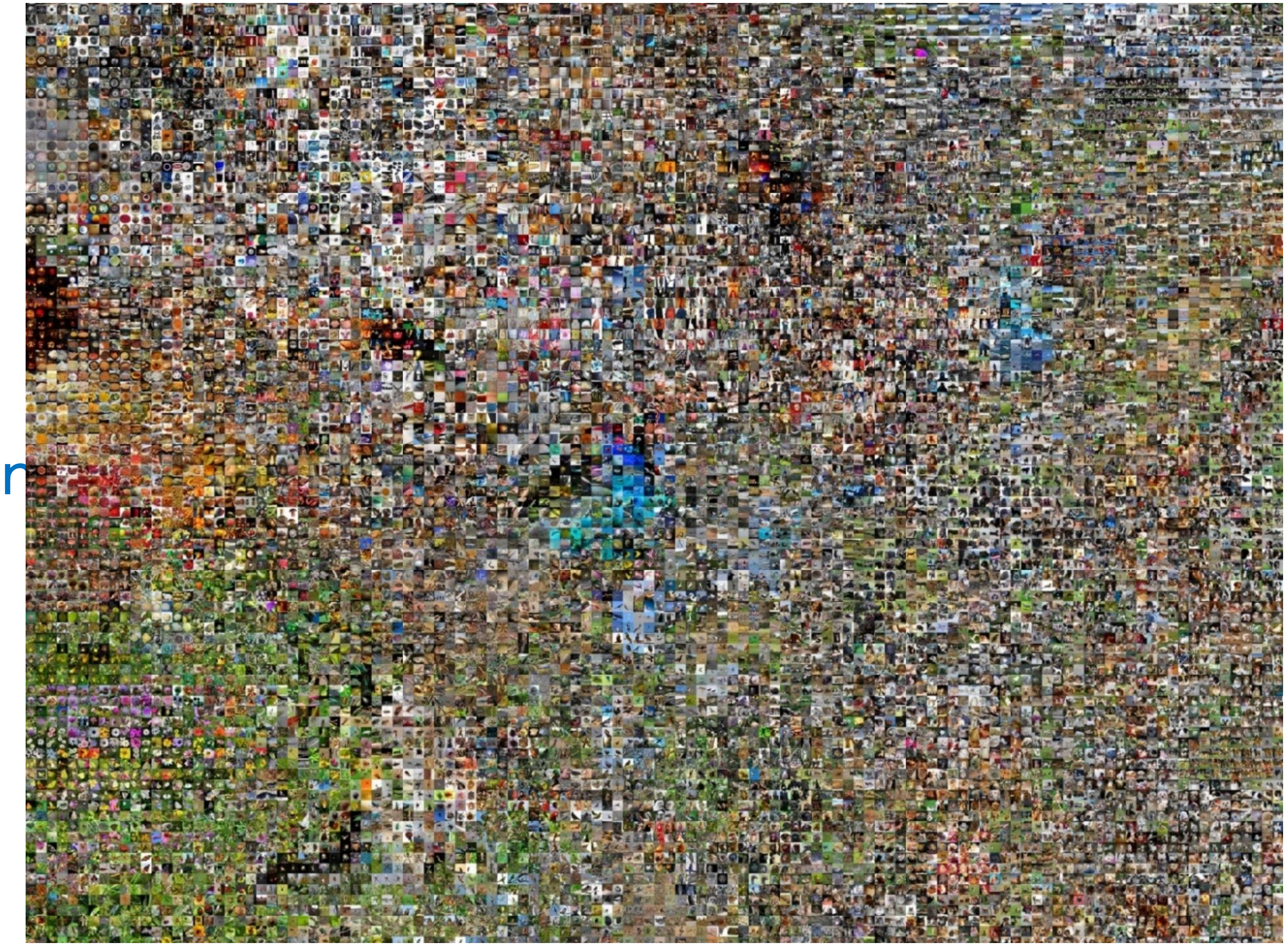
接續的作法: segmentation, 希望形成封閉曲線, 方便進行描述

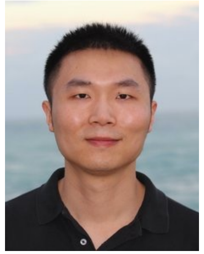
但前述兩做法均無法完美 -- ill-posed 問題



We were lucky that ImageNet was built in 2010. It's like a visual dictionary

- Start in 2007 at Princeton
- Debut in CVPR' 2009
- Stop in 2010
  - ➢Total classes : 21K
  - ➢Total images : 14M
- ILSVR Challenge: 2010 - presen





Jia Deng

Fei-Fei Li

How ImageNet Works ?

Visual based dictionary

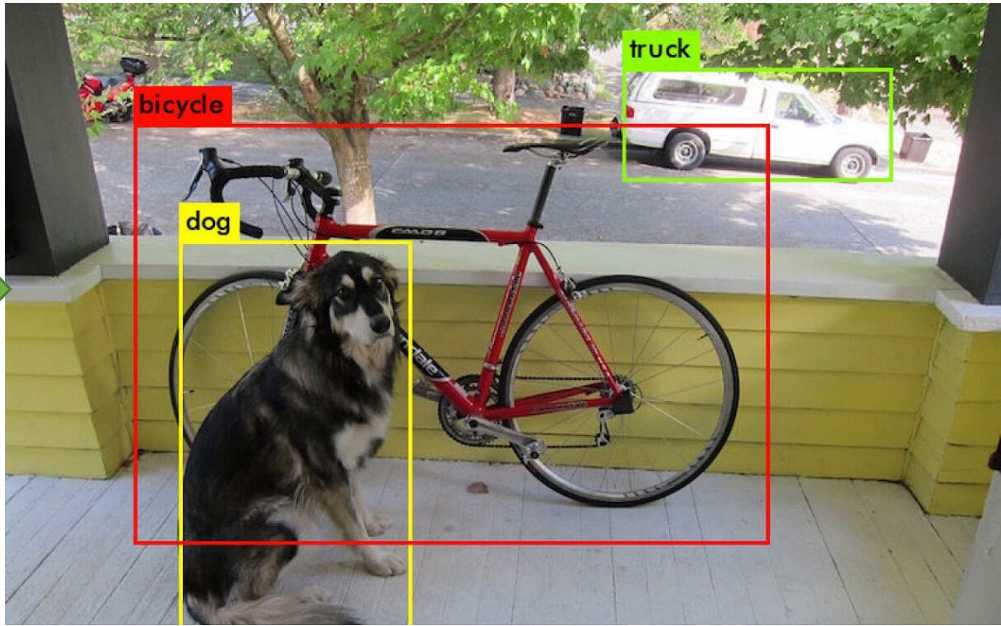
Training

unknown image

Inference

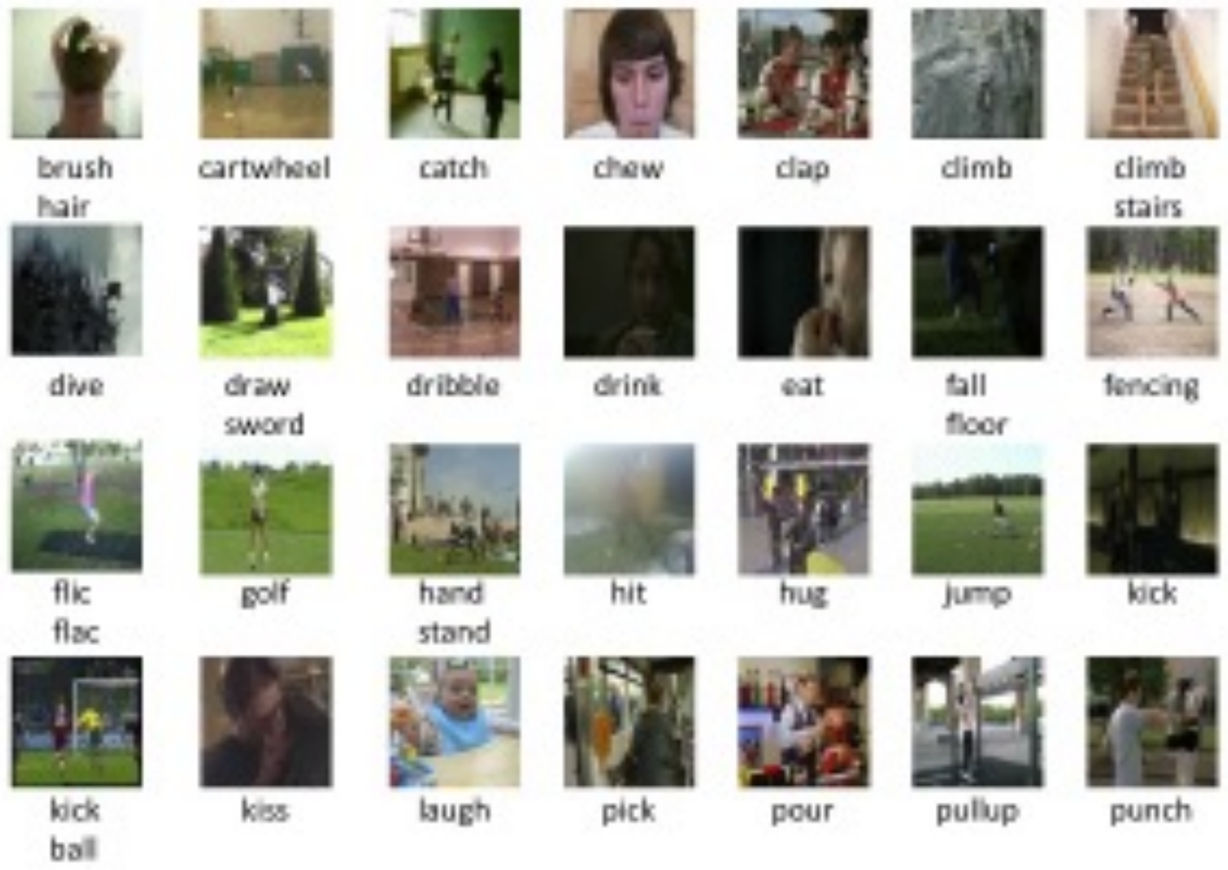
Feature extractor and Classifier

Inference



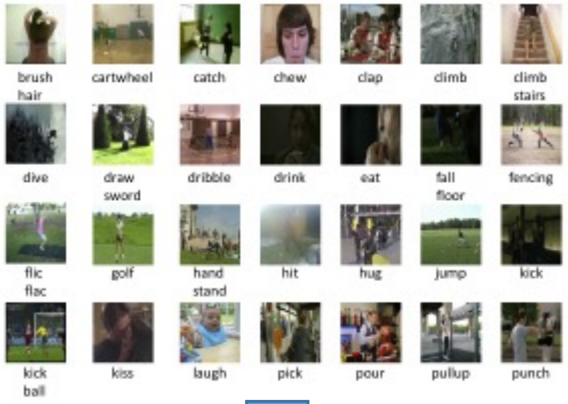
https://serre-lab.clps.brown.edu/resource/hmdb-a-large-human-motion-database/

HMDB51 (knowing the power of ImageNet, we are able to build an action dictionary too)



- HMDB51 – About 2GB for a total of 7,000 clips distributed in 51 action classes (model and kinetics dataset, 2017)

How Action Dictionary Works ?



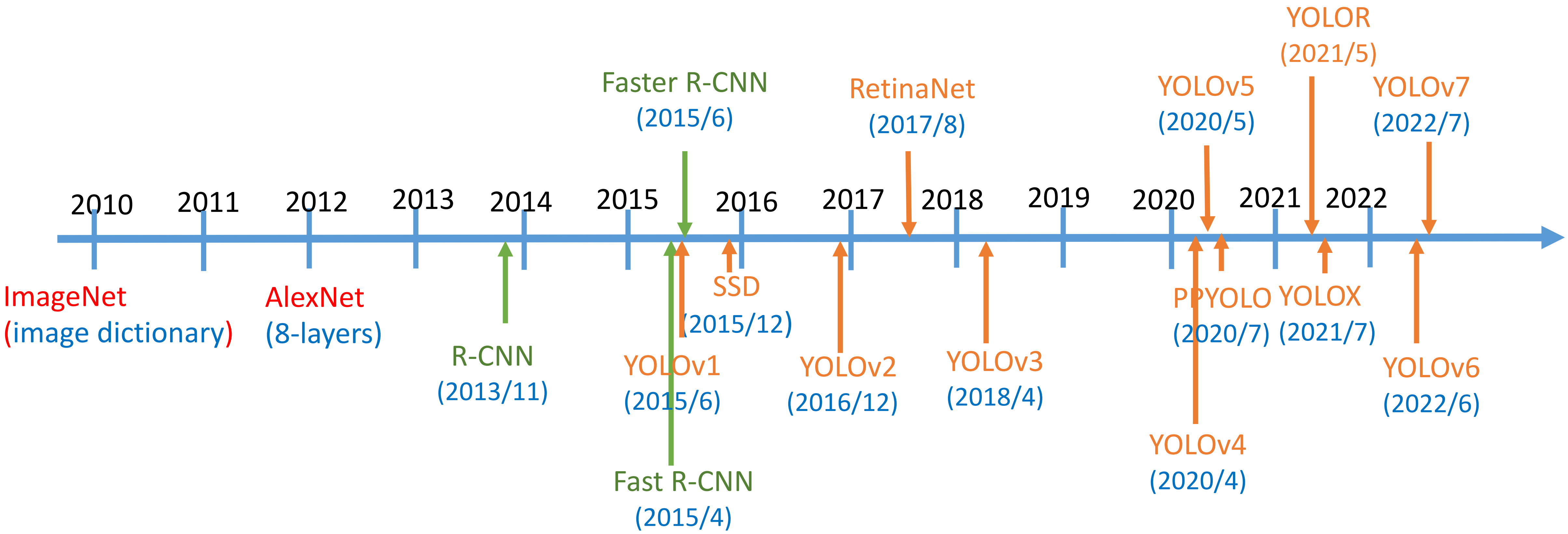
Training

Inference

Feature extractor and Classifier

Inference

Development History of one-stage and two-stage object detectors



Outline of this talk

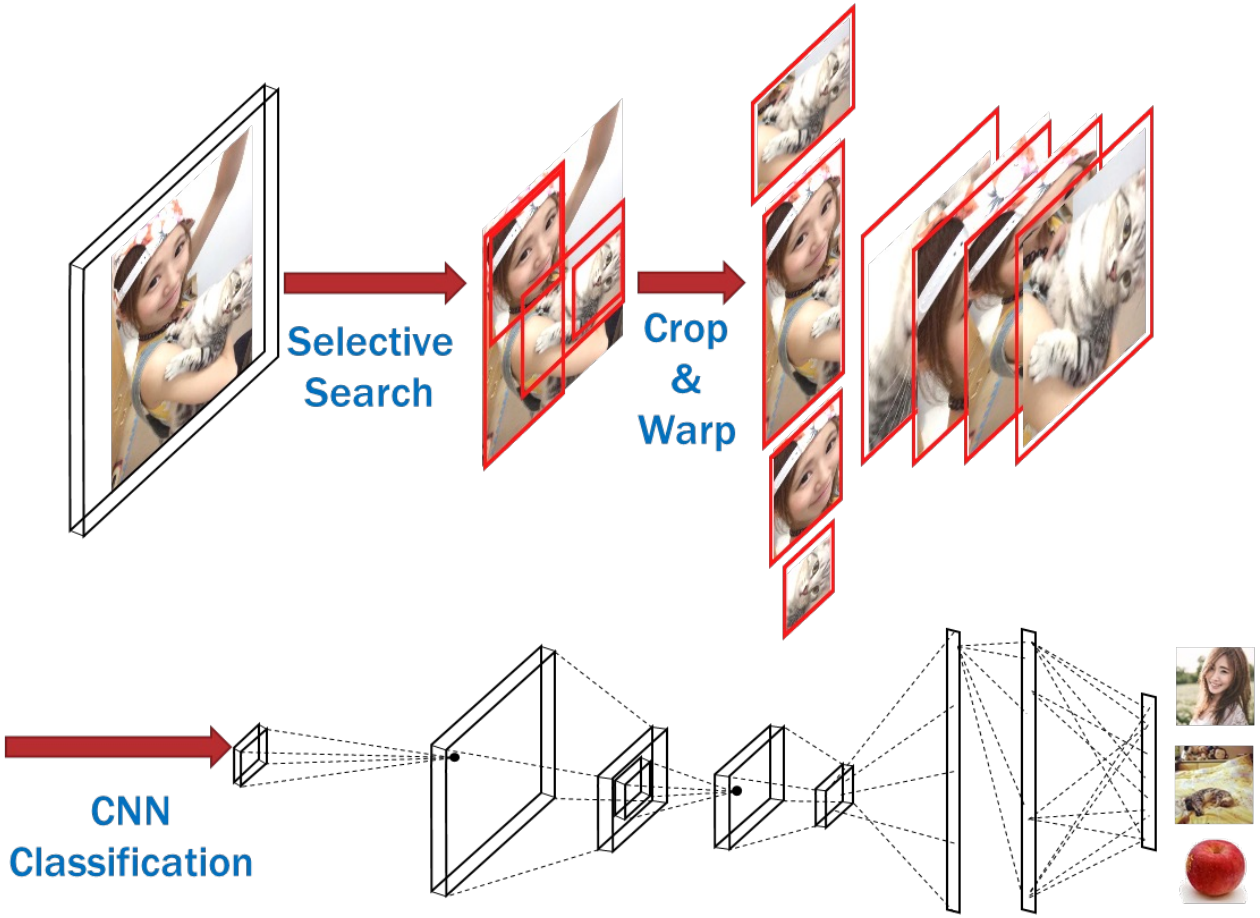
- MOST AI project: 2018/1 – 2021/12
  ``Smart City Traffic Flow Solutions’’
- Object detection is the first step toward *EVERYTHING*
  - development history of object detectors
  - one-stage vs. two-stage object detectors
- Why PRN, CSPNet, and YOLOv4 were developed ?
- Why YOLOR, ELAN and then YOLOv7

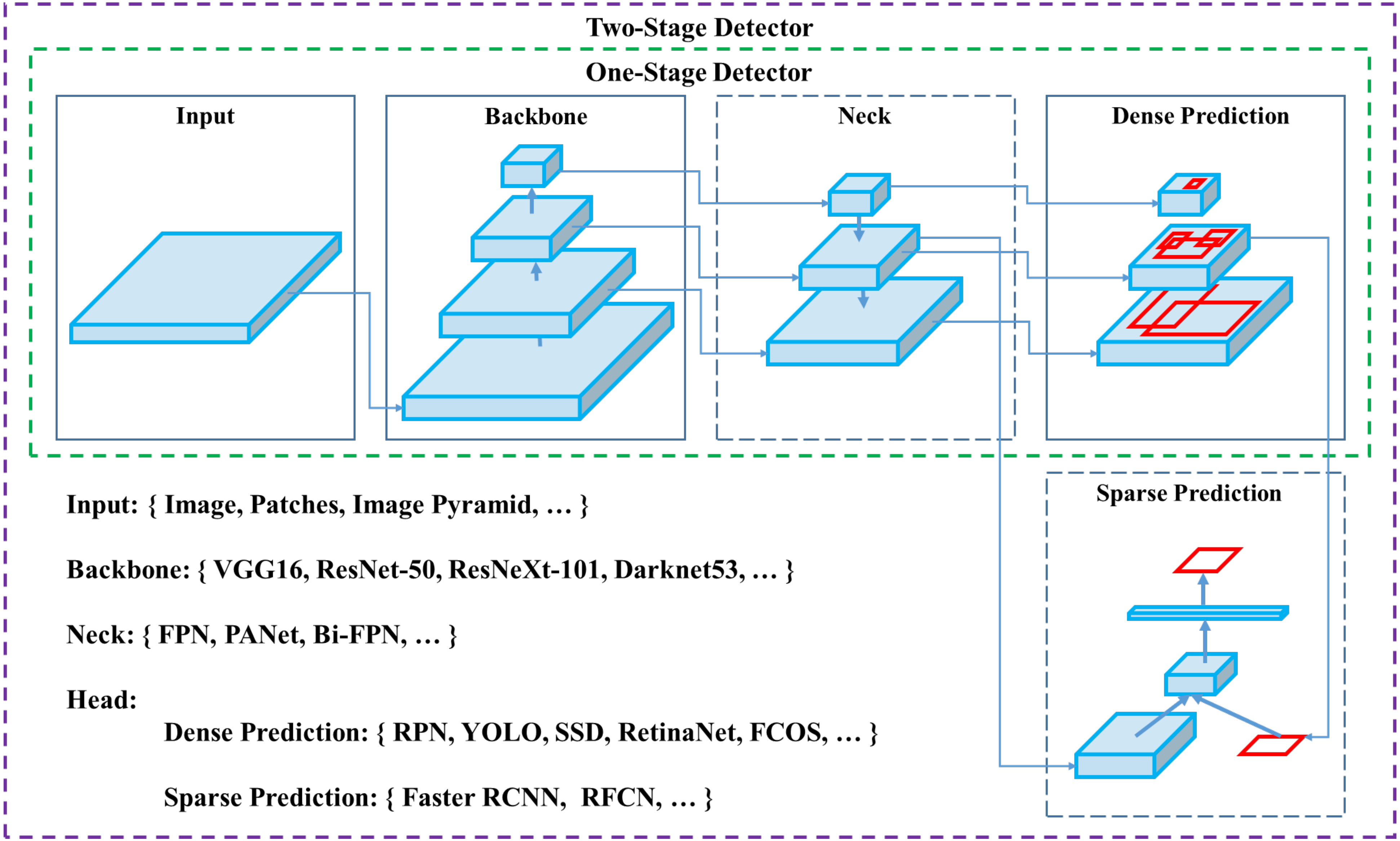
R-CNN series two-stage object detectors
R-CNN(2013/11), Fast R-CNN(2015/4), Faster R-CNN(2015/6)

Region Proposal

CNN Feature Extraction

SVM Classification




Two-Stage Detector
One-Stage Detector
Input
Backbone
Neck
Dense Prediction
Sparse Prediction
Input: { Image, Patches, Image Pyramid, … }
Backbone: { VGG16, ResNet-50, ResNeXt-101, Darknet53, … }
Neck: { FPN, PANet, Bi-FPN, … }
Head:
Dense Prediction: { RPN, YOLO, SSD, RetinaNet, FCOS, … }
Sparse Prediction: { Faster RCNN, RFCN, … }

Outline of this talk

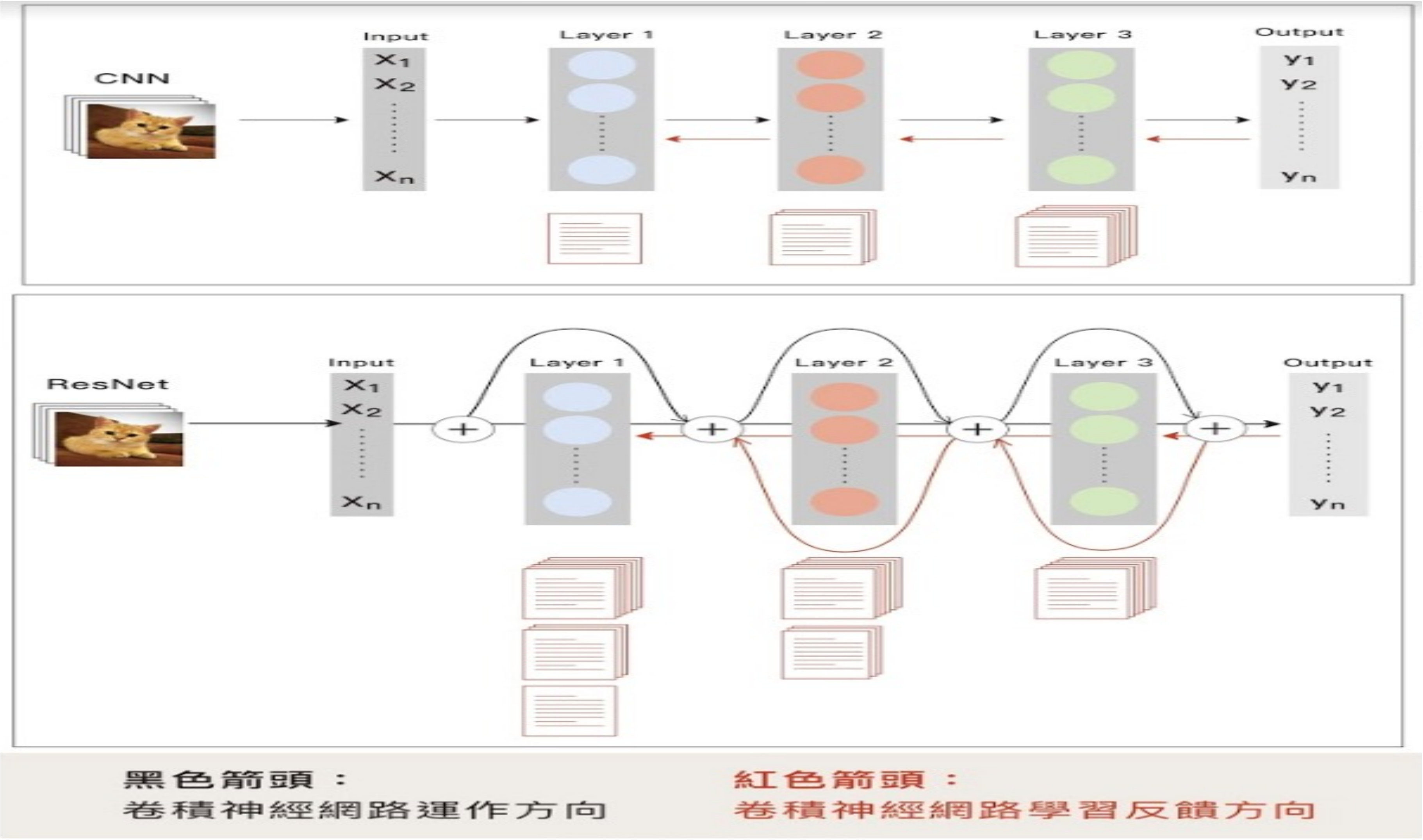
- MOST AI project: 2018/1 – 2021/12

  ``Smart City Traffic Flow Solutions''
- Object detection is the first step toward *EVERYTHING*

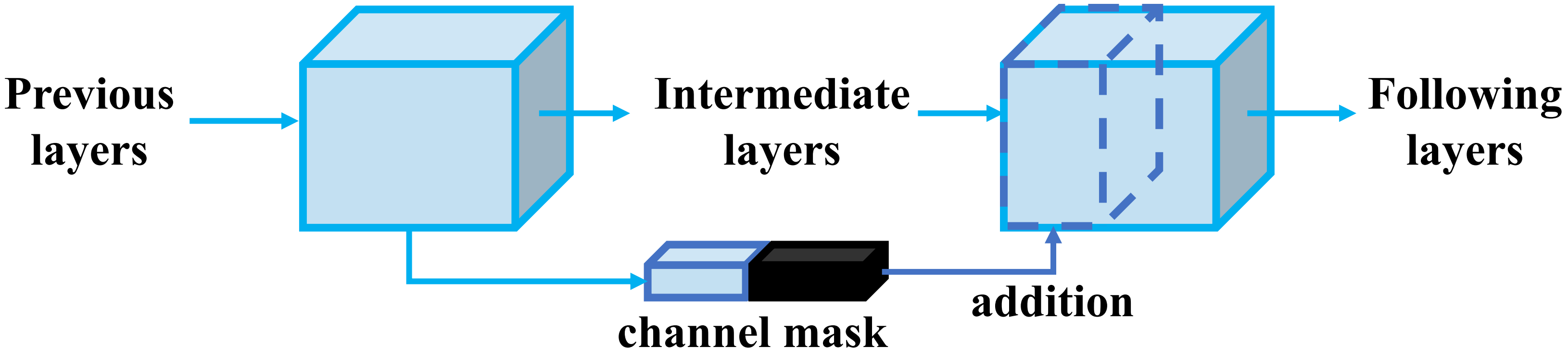
  - development history of object detectors

  - one-stage vs. two-stage object detectors
- Why PRN, CSPNet, and YOLOv4 were developed ?
- Why YOLOR, ELAN and then YOLOv7

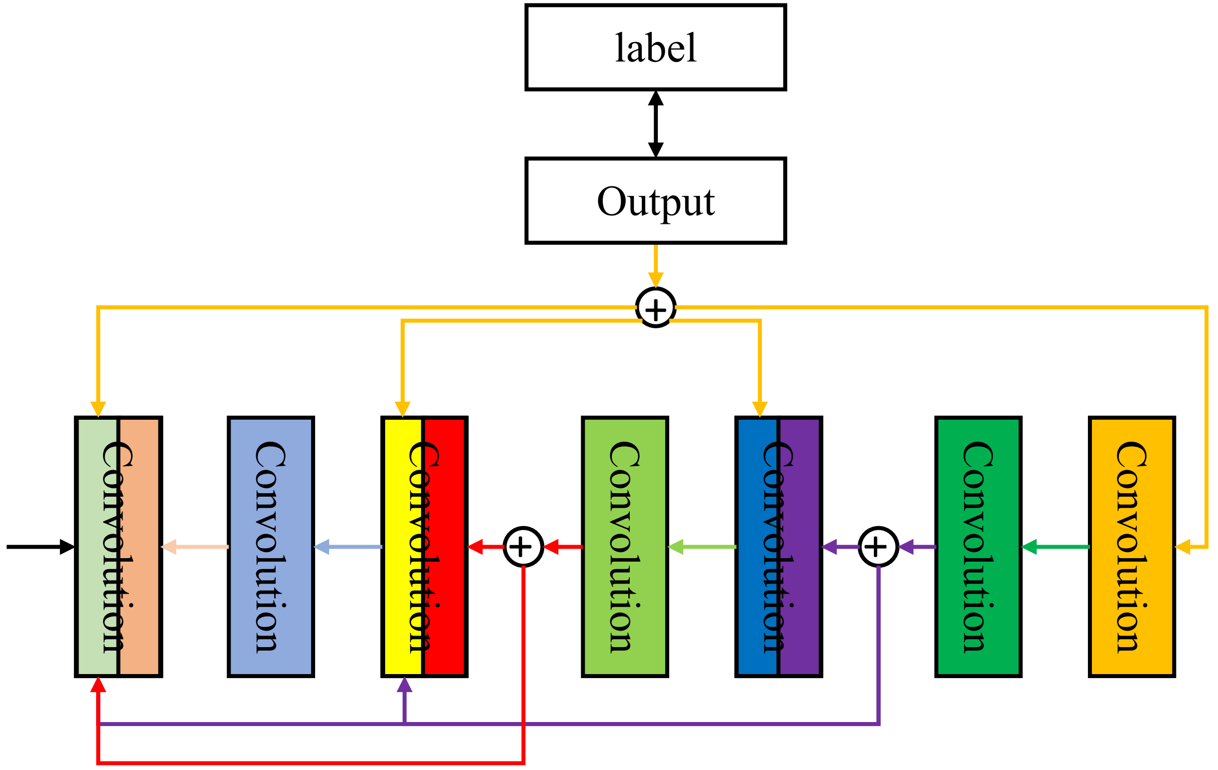
Redundancy Problem in Residual Net



Layer level design (partial residual network, May-June 2018)

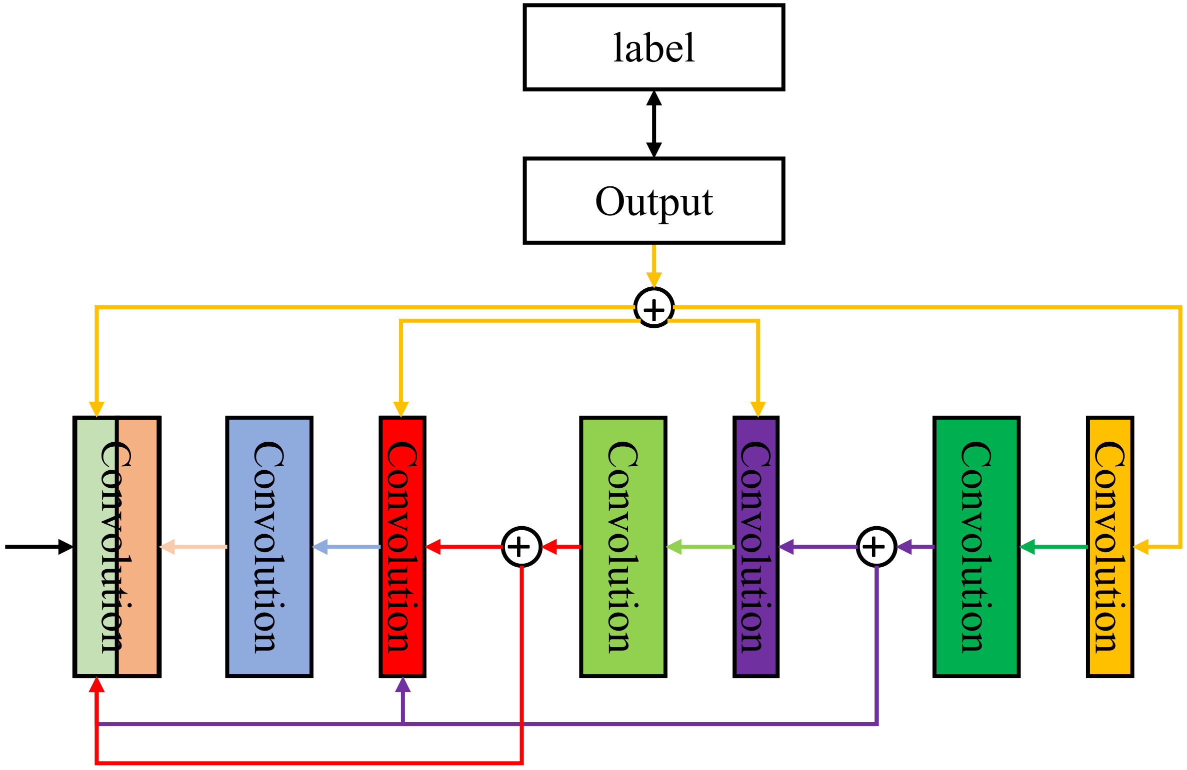


Backpropagation of Partial Residual Network (2018/6-12)



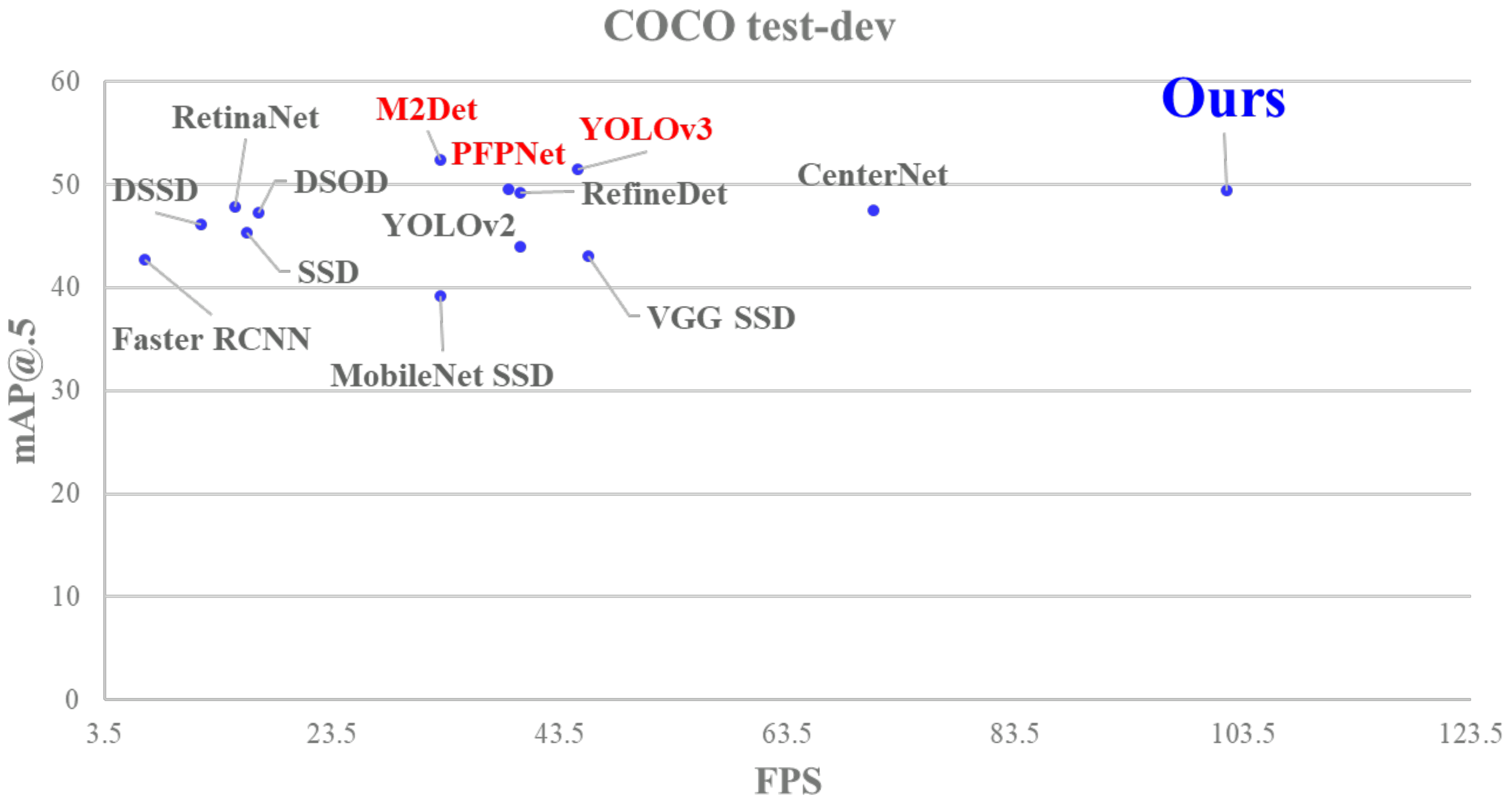
Same amount of parameters, computations, connections, increase the amount of gradient combinations and variability

Variation of Partial Residual Network



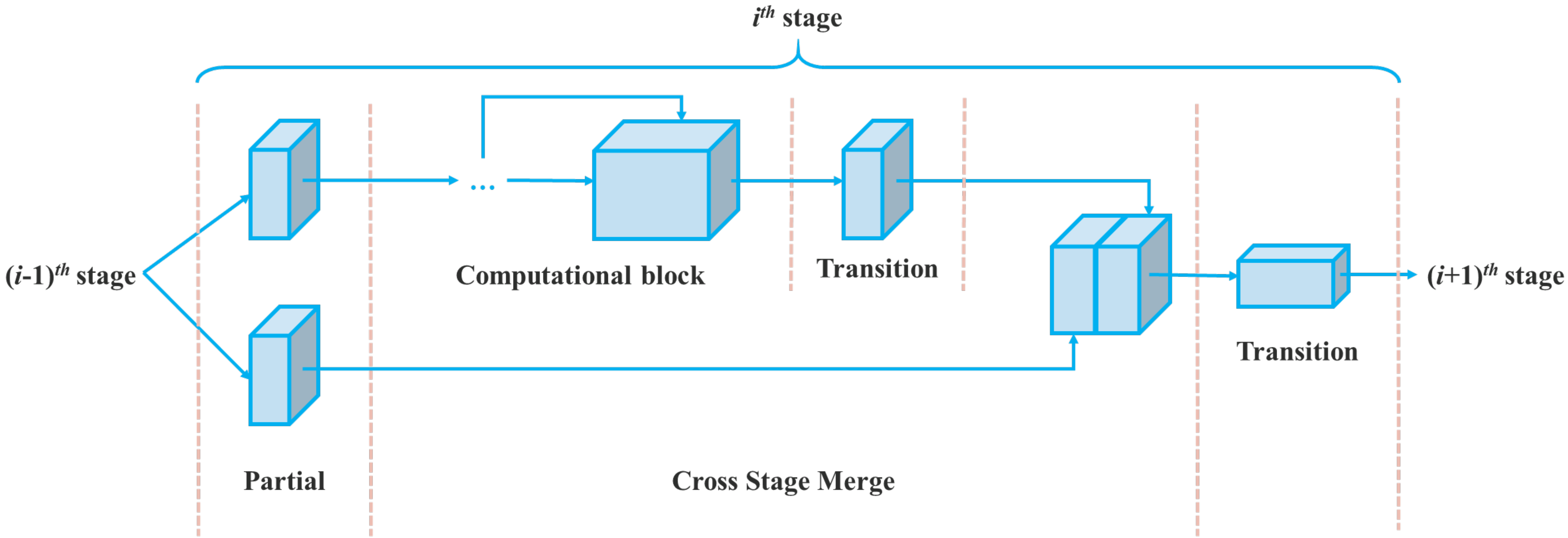
Cut down computations and parameters, maintain connections, increase gradient combinations

Comparison: performance of Partial Residual Net



- Ours (49.5%)
- Faster RCNN [NIPS2015]
- DSSD
- SSD [ECCV2016]
- DSOD [ICCV2017]
- MobileNet SSD
- PFPNet [ECCV2018] (49.6%)
- RefineDet [CVPR2018]
- YOLOv2 [CVPR2017]
- YOLOv3 (51.5%)
- VGG SSD [ECCV2016]
- CenterNet
- M2Det [AAAI2019] (52.4%)
- RetinaNet [ICCV2017]

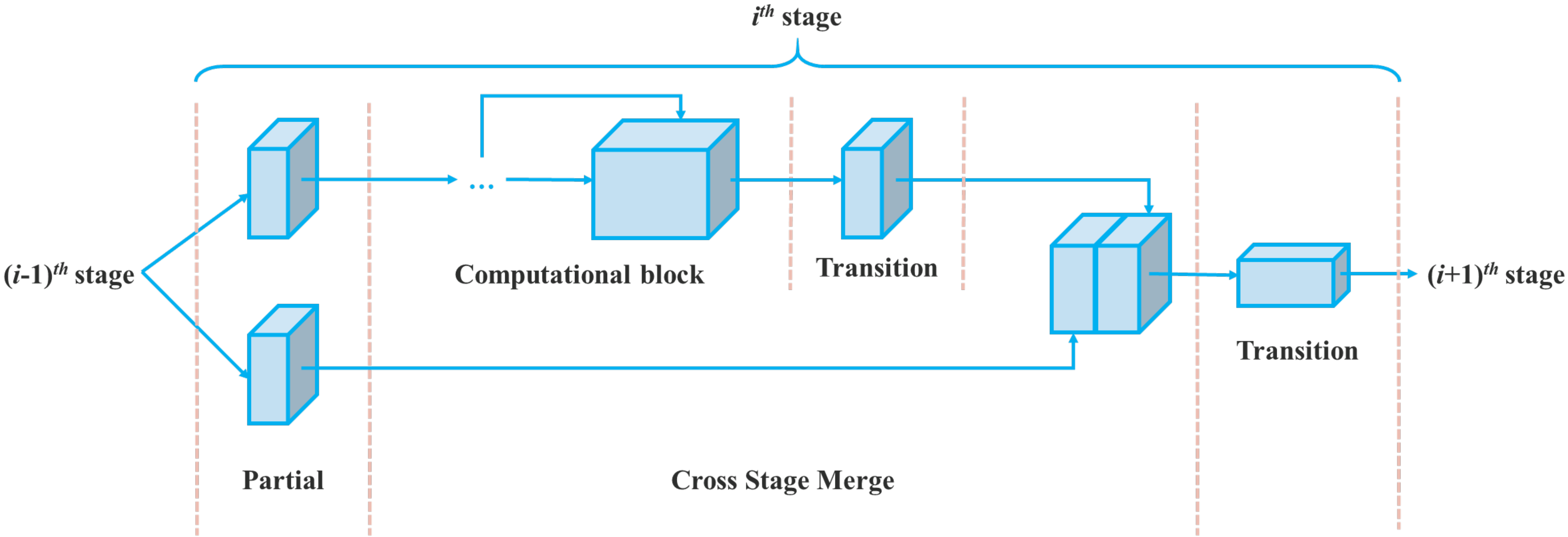
Stage level design (CSPNet, June 2019)



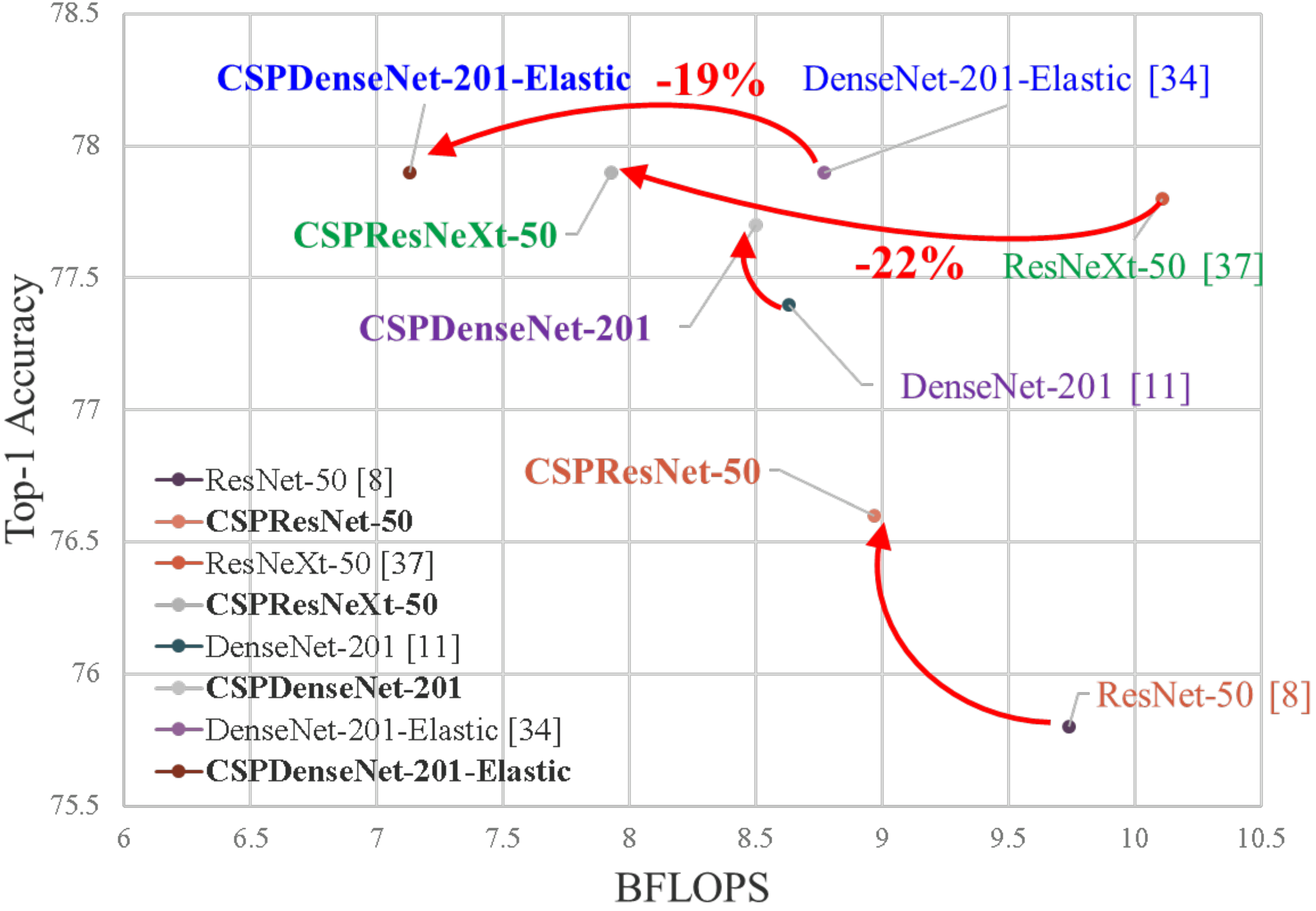
New Architecture of YOLOv4 - CSPNet

Table 1: Parameters of neural networks for image classification.

Backbone model	Input network resolution	Receptive field size	Parameters	Average size of layer output (WxHxC)	BFLOPs (512x512 network resolution)	FPS (GPU RTX 2070)
CSPResNext50	512x512	425x425	20.6 M	**1058 K**	31 (15.5 FMA)	62
CSPDarknet53	512x512	725x725	**27.6 M**	950 K	52 (26.0 FMA)	**66**
EfficientNet-B3 (ours)	512x512	**1311x1311**	12.0 M	668 K	11 (5.5 FMA)	26

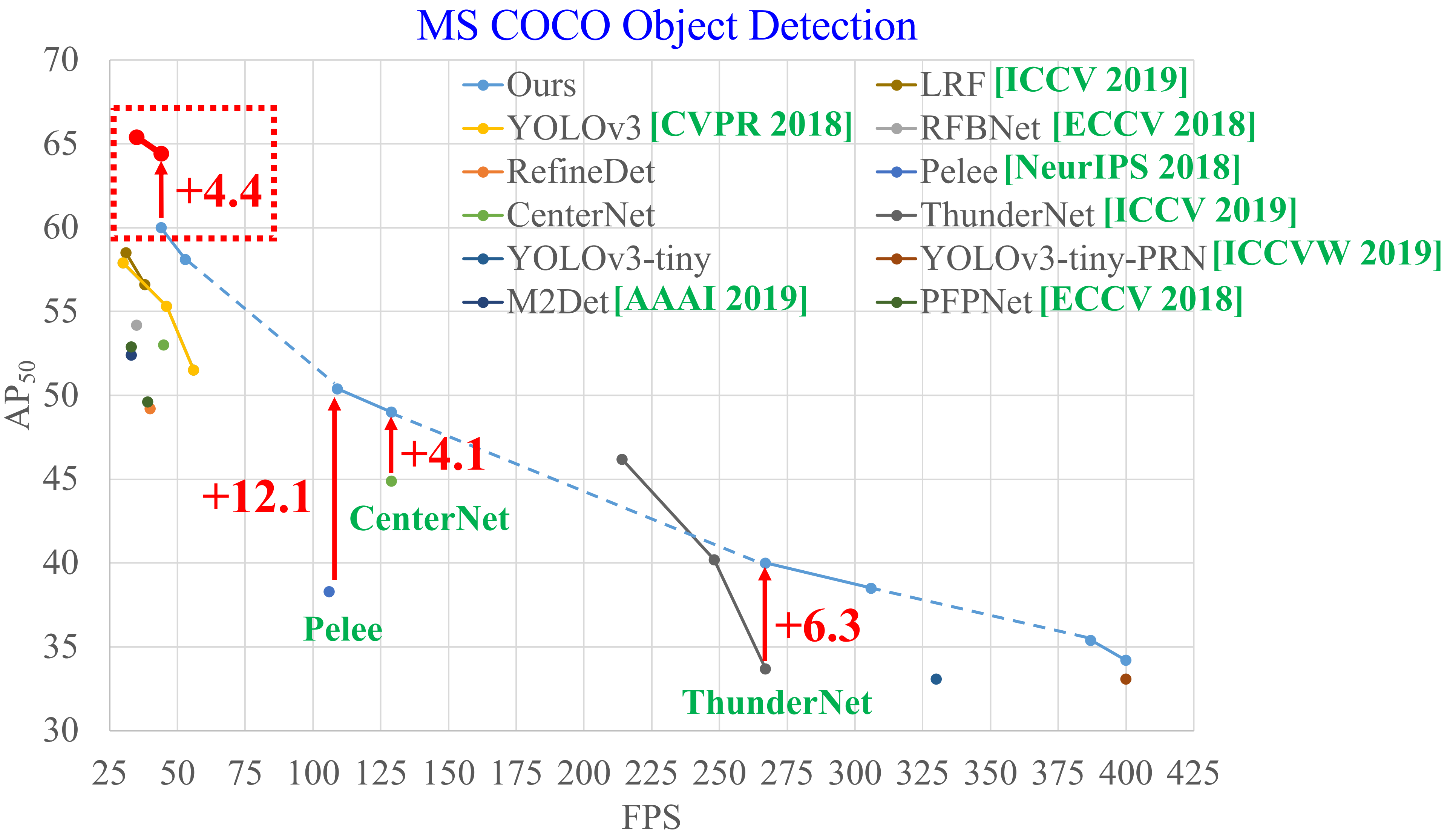


CSPNet (1/2)

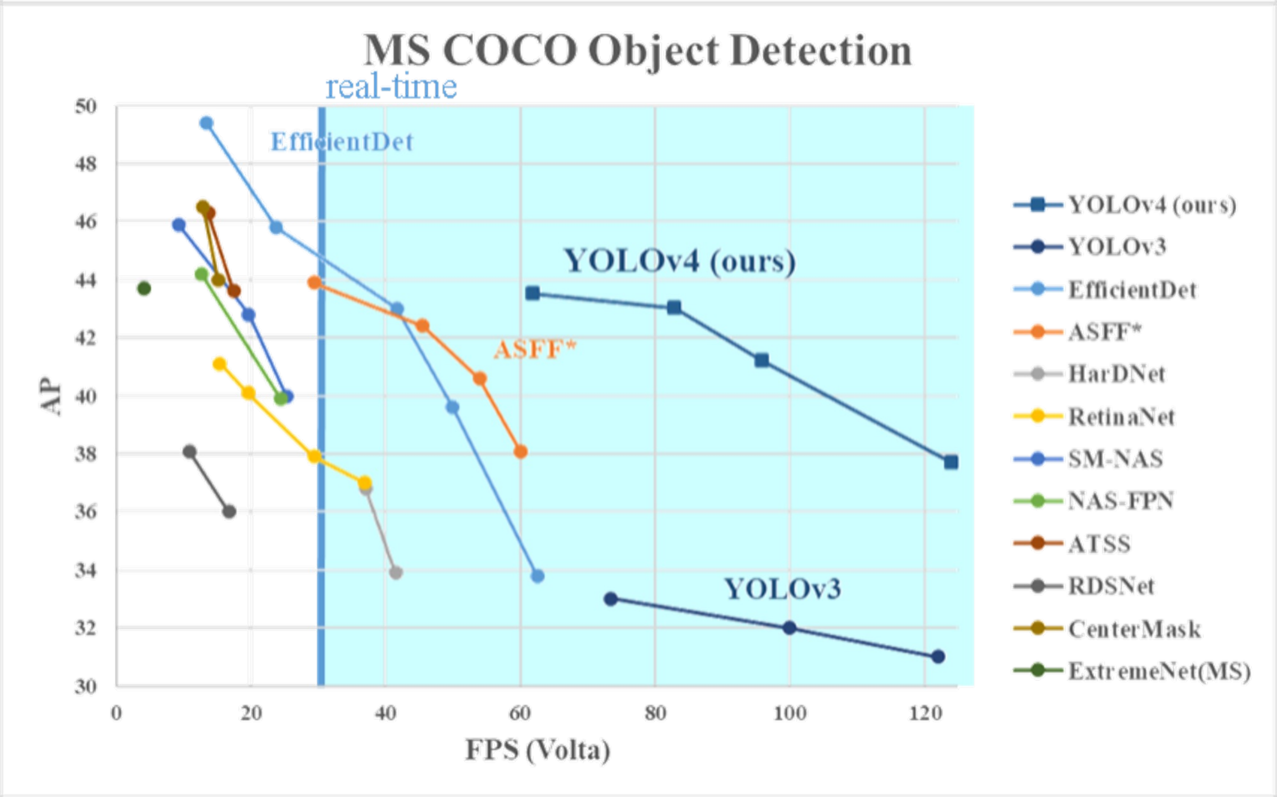


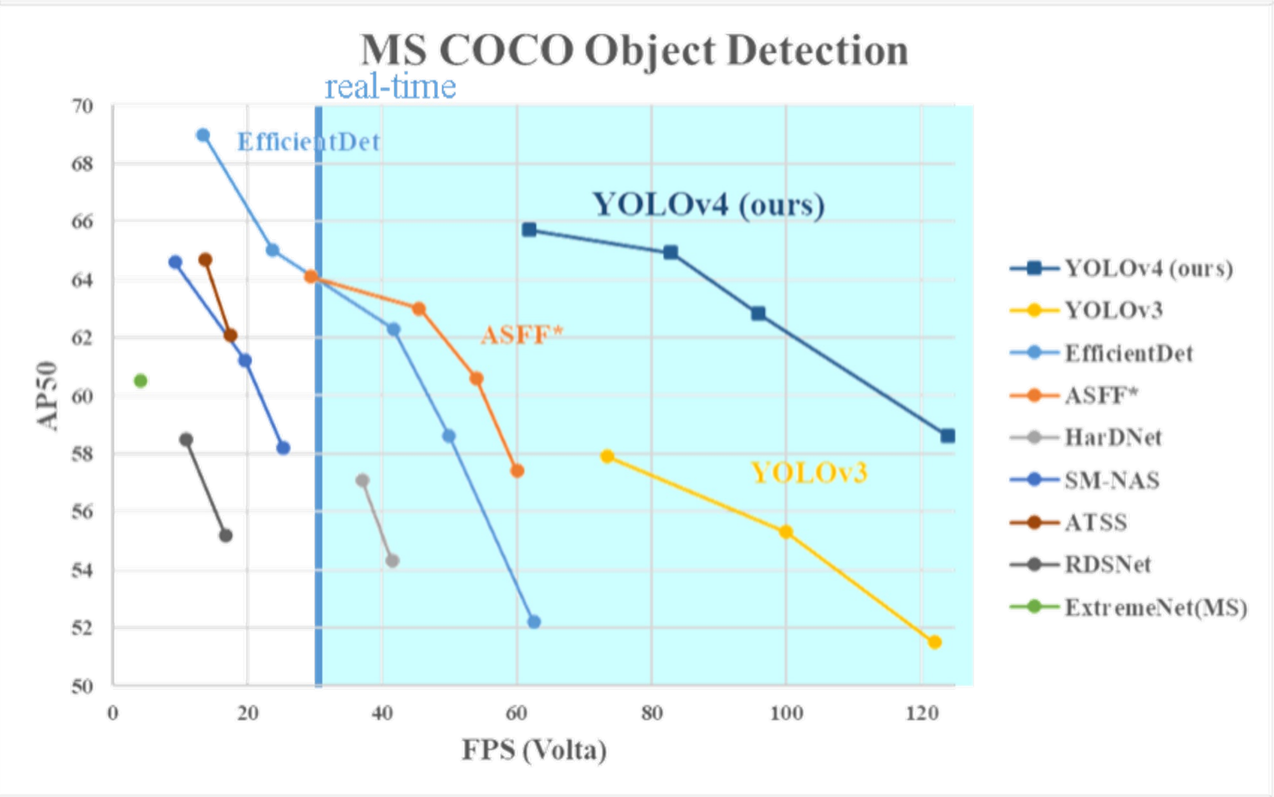
CSPDenseNet-201-Elastic
-19%
DenseNet-201-Elastic [34]
CSPResNeXt-50
-22%
ResNeXt-50 [37]
CSPDenseNet-201
DenseNet-201 [11]
CSPResNet-50
ResNet-50 [8]
ResNet-50 [8]
CSPResNet-50
ResNeXt-50 [37]
CSPResNeXt-50
DenseNet-201 [11]
CSPDenseNet-201
DenseNet-201-Elastic [34]
CSPDenseNet-201-Elastic
Top-1 Accuracy
78.5
78
77.5
77
76.5
76
75.5
6
6.5
7
7.5
8
8.5
9
9.5
10
10.5
BFLOPS

CSPNet (2/2)

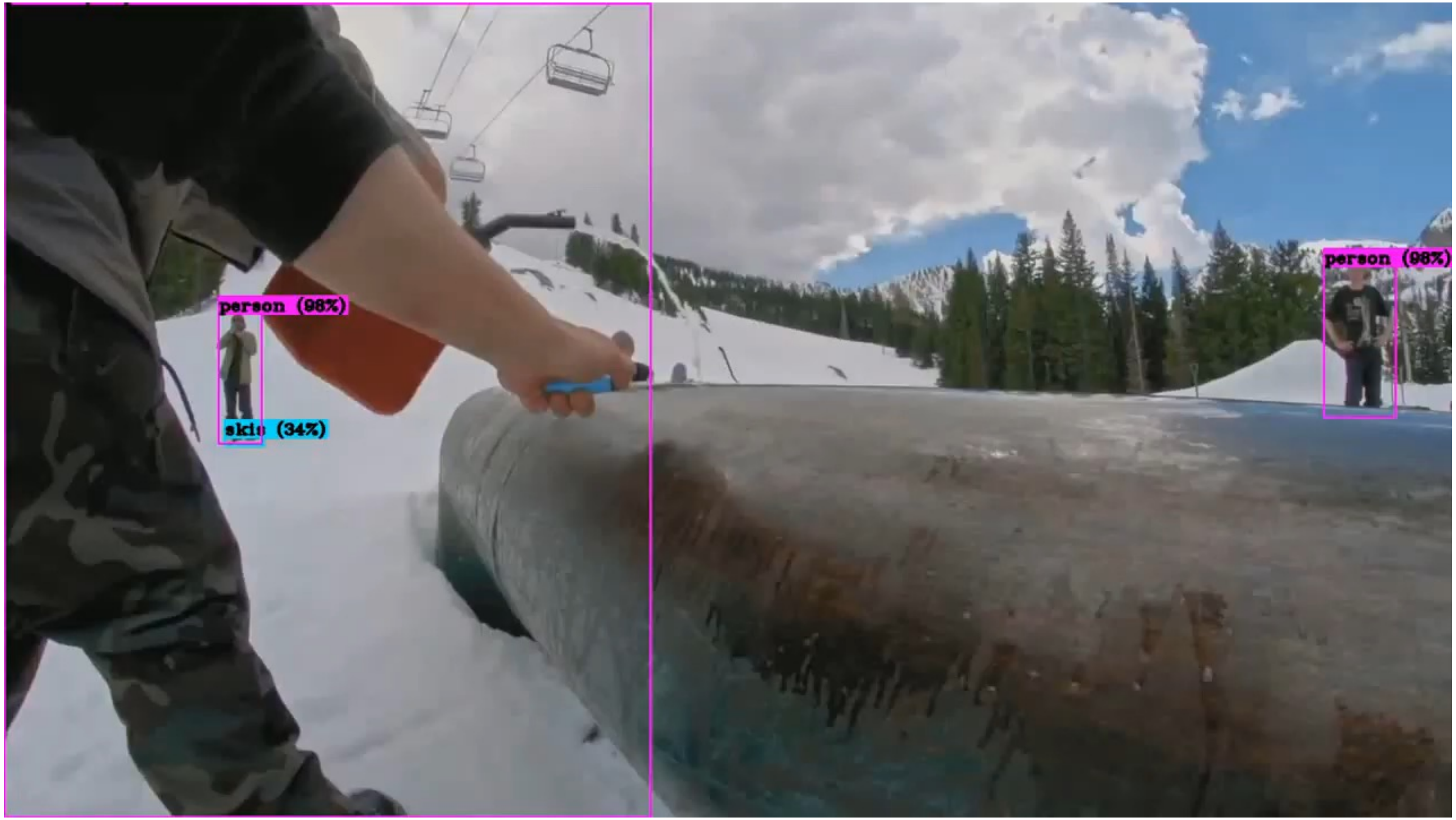

MS COCO Object Detection
Ours
YOLOv3 [CVPR 2018]
RefineDet
CenterNet
YOLOv3-tiny
M2Det [AAAI 2019]
LRF [ICCV 2019]
RFBNet [ECCV 2018]
Pelee [NeurIPS 2018]
ThunderNet [ICCV 2019]
YOLOv3-tiny-PRN [ICCVW 2019]
PFPNet [ECCV 2018]
+4.4
+12.1
+4.1
+6.3
Pelee
CenterNet
ThunderNet
AP50
FPS
70
65
60
55
50
45
40
35
30
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425

Comparison on Volta GPU





YOLOv4 Object Detection demo



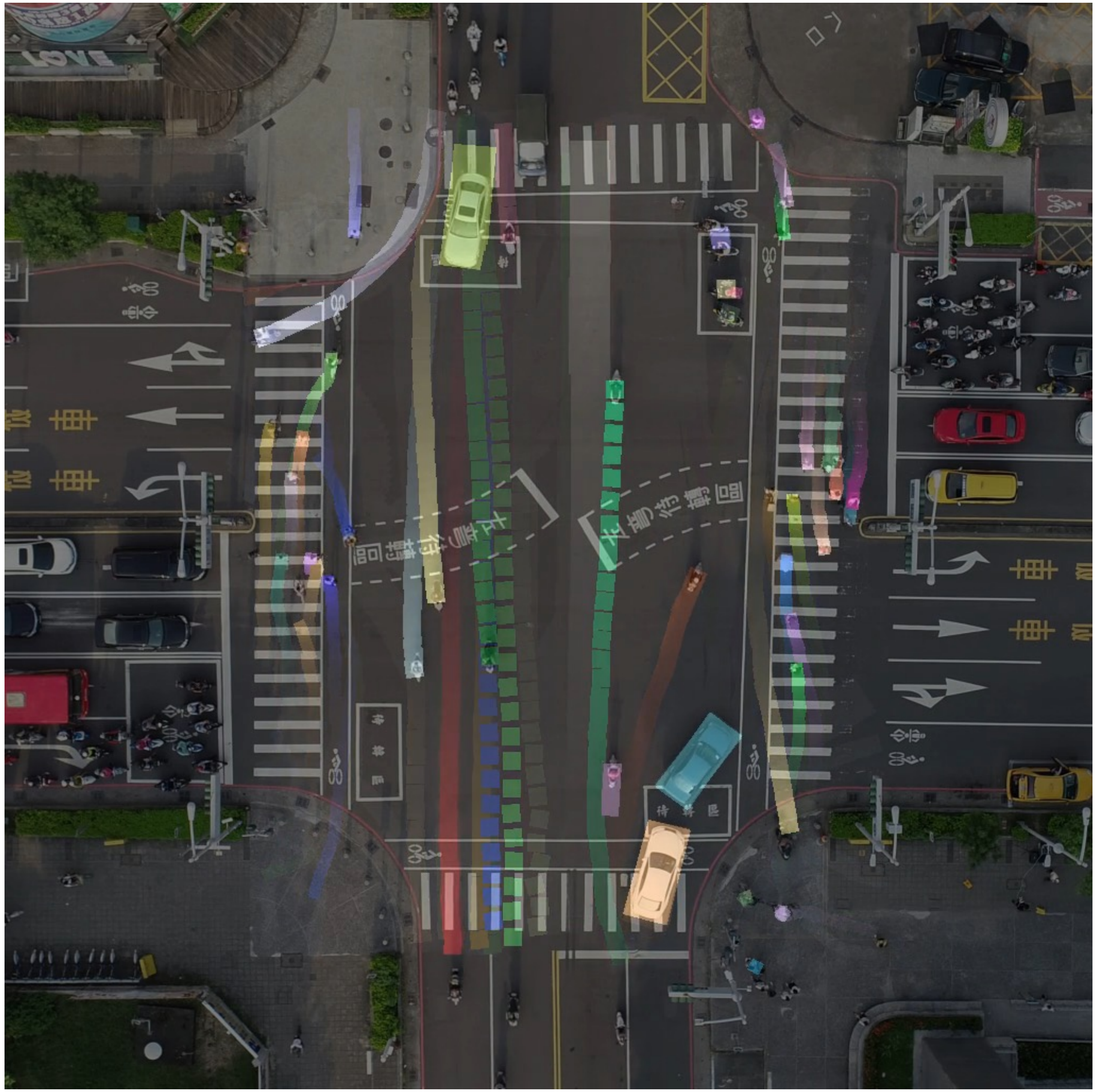
https://drive.google.com/file/d/1RFo4xOYceAYTwT8UhH-KmaeO5CEASHjv/view

智慧交通

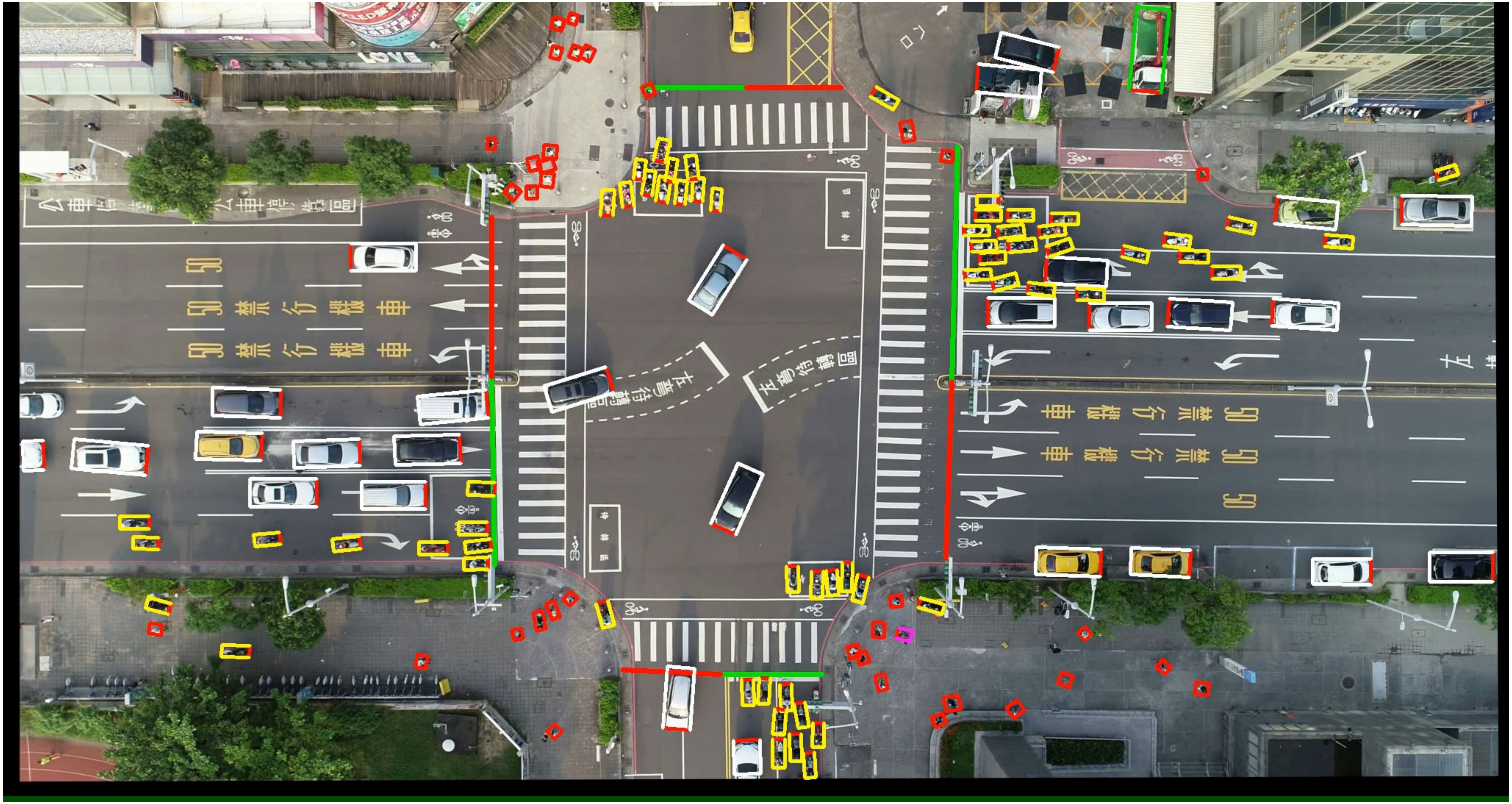

國道1號增設台74線系統交流道工程(第186標)
施工現況
台74線東行線


第186標11012空拍紀錄影片
國道1號增設台74線系統交流道工程(第186標)
car 0.62
car 0.40
van 0.53
van 0.81
car 0.87
car 0.88
truck 0.90
car 0.70
施工現況
西行線車道封閉管制作業
活動申請許可編號AB2109210002
18:40 / 20:34

Tracking of vehicles and pedestrians



Full range vehicle tracking

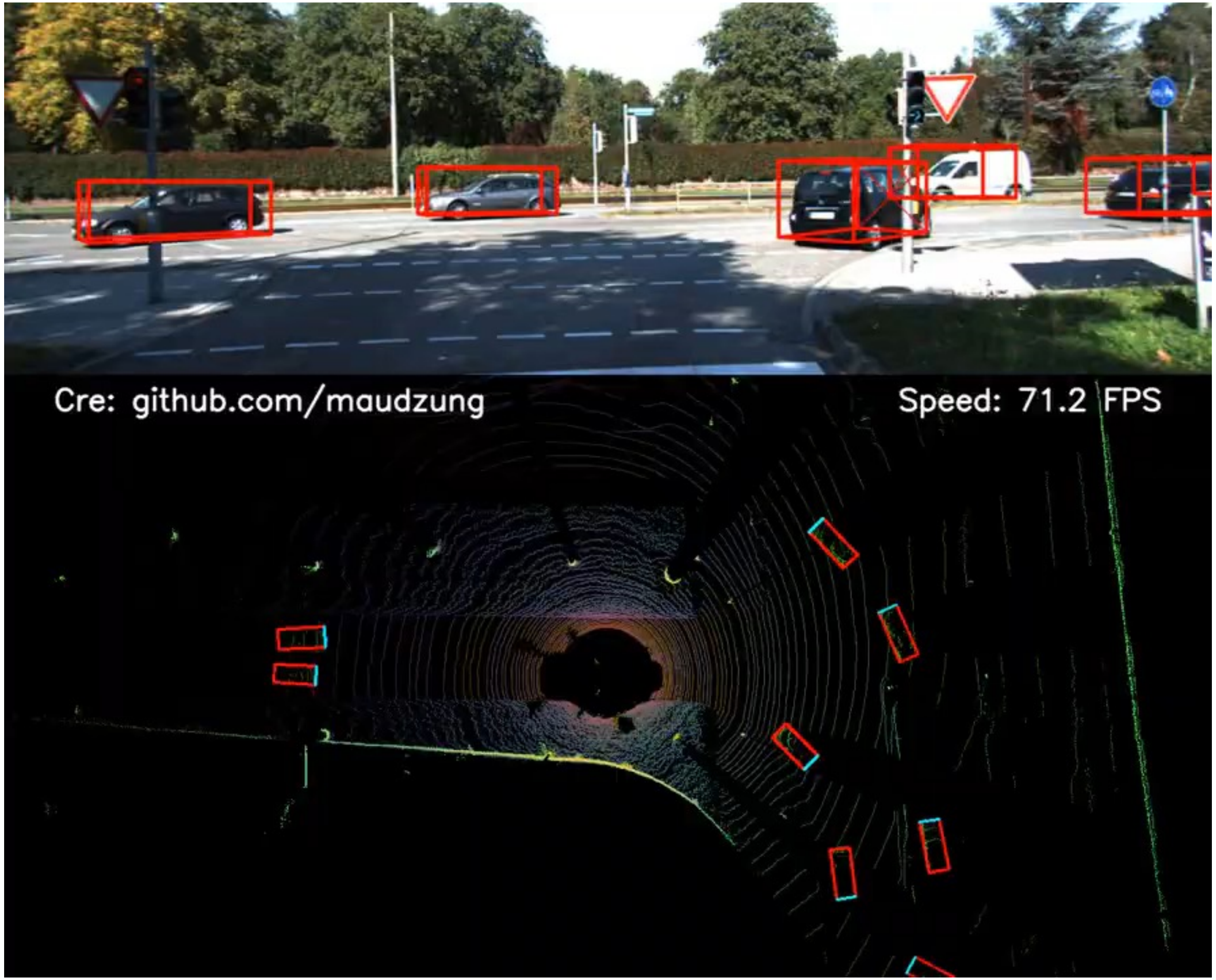


A group of people crossing a street in a city.

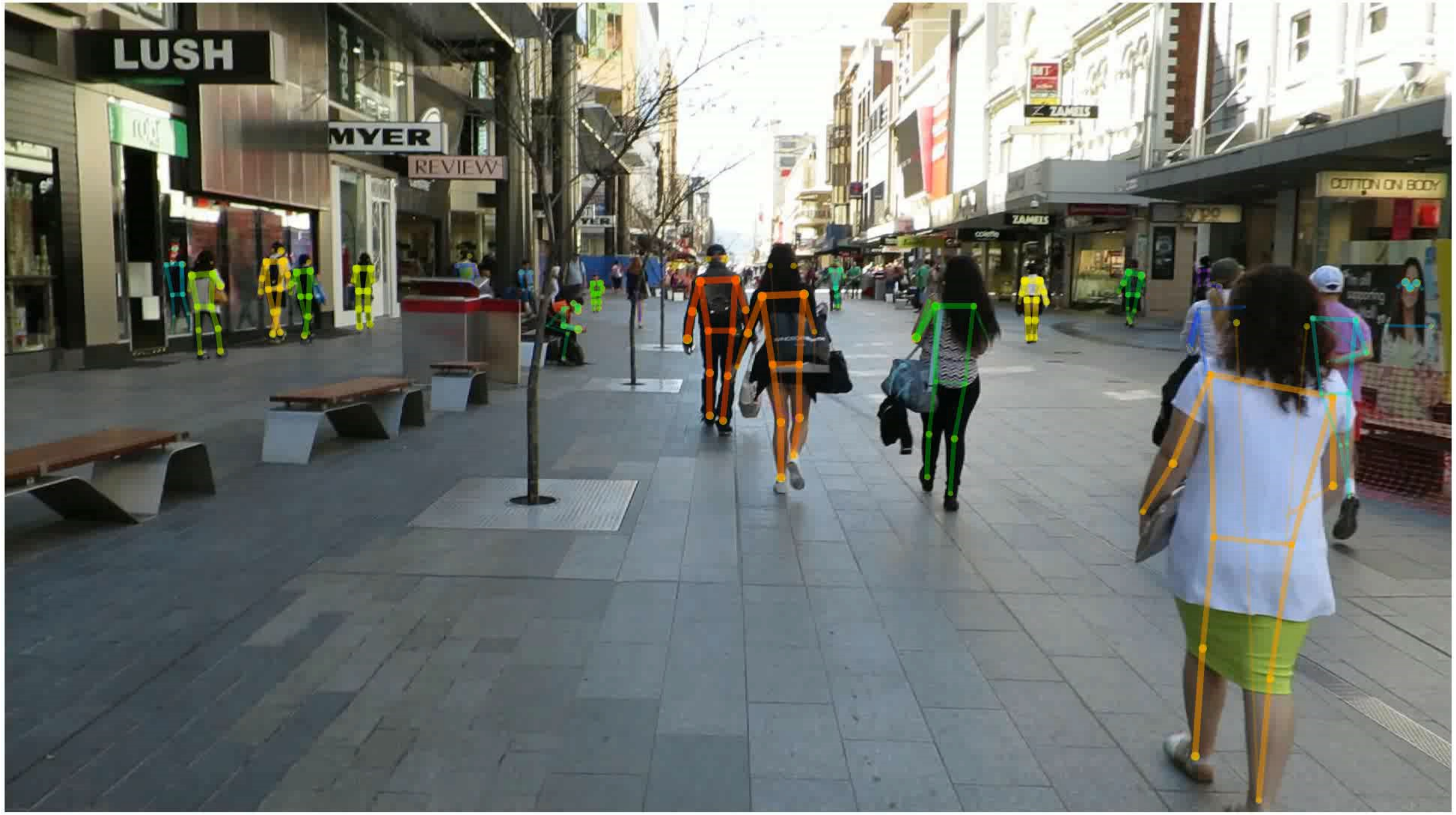

2021/06/21 09:05:35


2021/06/21 08:33:37

3D perception (YOLOv4 + Complex YOLO)



Behavior analysis (YOLOv4 + AlphaPose)

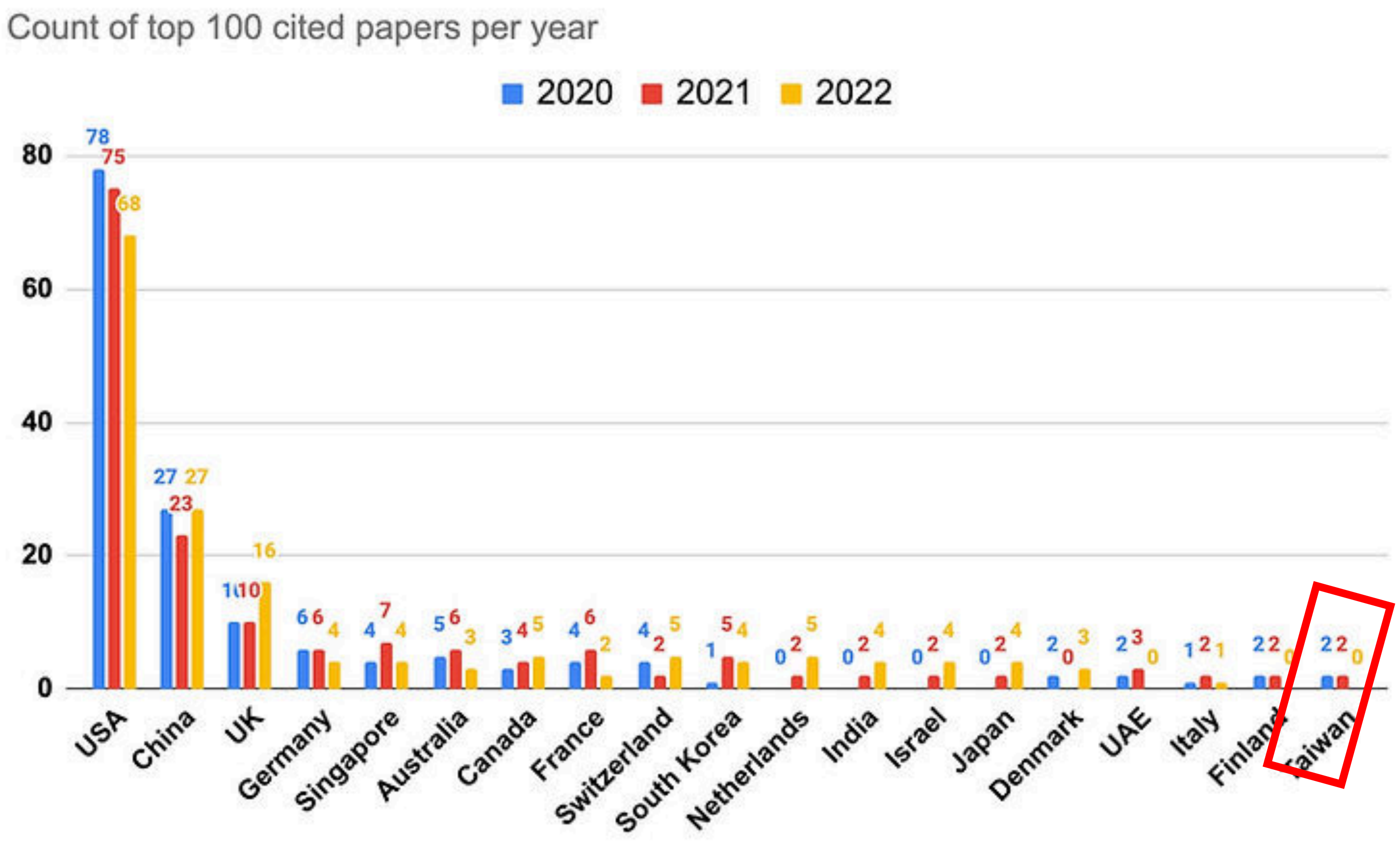


The impact of related publications

- Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv:2004.10934vl*, number arXiv preprint arXiv:2004.10934, April 2020 (Google citation: **16307**, as of July 8, 2024)
- C. Y. Wang, H. Y. Mark Liao, Y. H. Wu, P. Y. Chen, J. W. Hsieh, and I. H. Yeh, "CSPNet: A New Backbone that can Enhance Learning Capability of CNN," *IEEE International Conference on Computer Vision and Pattern Recognition Workshop (CVPRW) on ``Low power computer vision"*, IEEE, Seattle, USA, June 2020 (Google citation: **3995**, as of July 8, 2024)
- Chien-Yao Wang, Alexey Bochkovskiy, and Hong-Yuan Mark Liao, "Scaled-YOLOv4: Scaling Cross Stage Partial Network," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2021 (Google citation: **1609**, as of July 8, 2024)

Report of Sergi Castella Sape, March 2, 2023, ``Must Read: the 100 most cited AI papers in 2022''

Listed top 100 cited papers in 2020, 2021, and 2022



2020

#3 YOLOv4

#68 Scaled-YOLOv4

2020

1 An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale -> (From Google, 11914 citations) The first work showing how a plain Transformer could do great in Computer Vision.

2 Language Models are Few-Shot Learners -> (From OpenAI, 8070 citations) GPT-3. This paper does not need further explanation at this stage.

3 YOLOv4: Optimal Speed and Accuracy of Object Detection -> (From Academia Sinica, Taiwan, 8014 citations) Robust and fast object detection sells like hotcakes.

4 Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer -> (From Google, 5906 citations) A rigorous study of transfer learning with Transformers, resulting in the famous T5.

5 Bootstrap your own latent: A new approach to self-supervised Learning -> (From DeepMind and Imperial College, 2873 citations) Showing that negatives are not even necessary for representation learning.

Outline of this talk

- MOST AI project: 2018/1 – 2021/12
  ``Smart City Traffic Flow Solutions''
- Object detection is the first step toward *EVERYTHING*
  - development history of object detectors
  - one-stage vs. two-stage object detectors
- Why PRN, CSPNet, and YOLOv4 were developed ?
- Development process of YOLOv7 ?

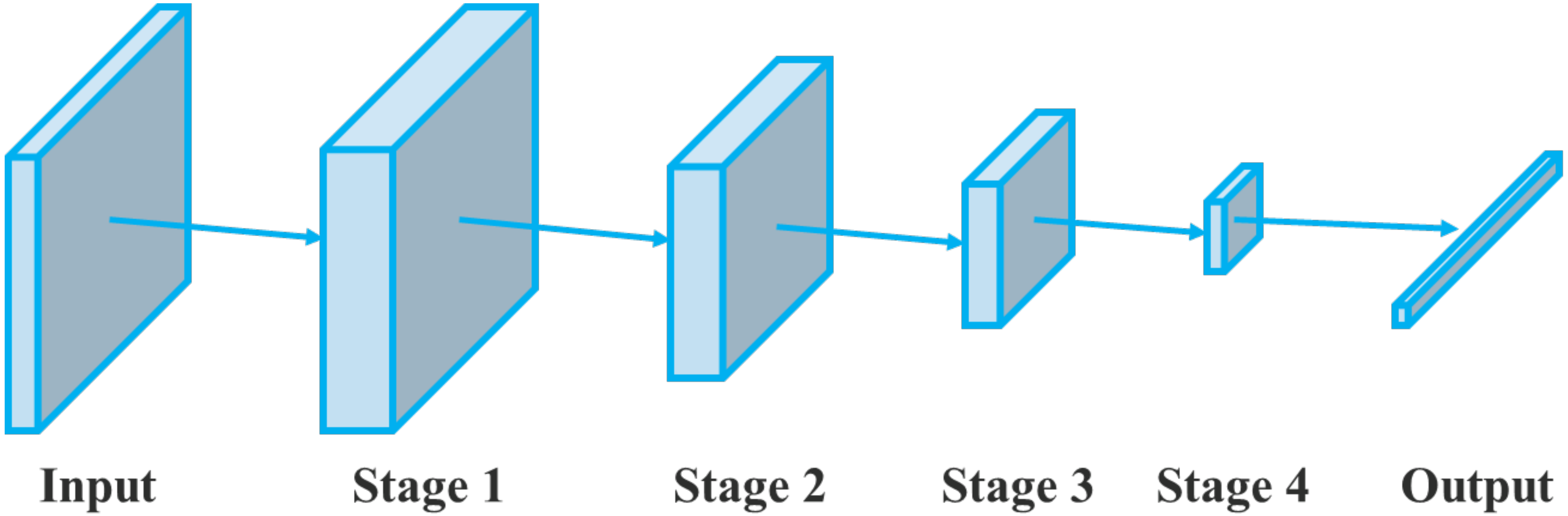
The impact of related publications

- Chien-Yao Wang , Alexey Bochkovskiy, and Hong-Yuan Mark Liao, "YOLOv7:Trainable Bag-of-freebies sets new state-of-the-art for real-time object detectors," *CVPR*, June 2023 (Google citation: 5918, as of July 8, 2024)
- C. Y. Wang, I. H. Yeh, H. Y. Mark Liao, "You Only Learn One Representation: Unified Network for Multiple Tasks," *Journal of Information Science and Engineering*, Vol.39(3), May 2023 (invited paper, Google citation: 603, as of July 8, 2024)
- C. Y. Wang, H. Y. Mark Liao, and I. H. Yeh, ``Designing Network Design Strategies Through Gradient Path Analysis," *Journal of Information Science, Vol.39(4), July* 2023 (invited paper, Google citation: 178, as of July 8, 2024)

From YOLOv4 to YOLOv7

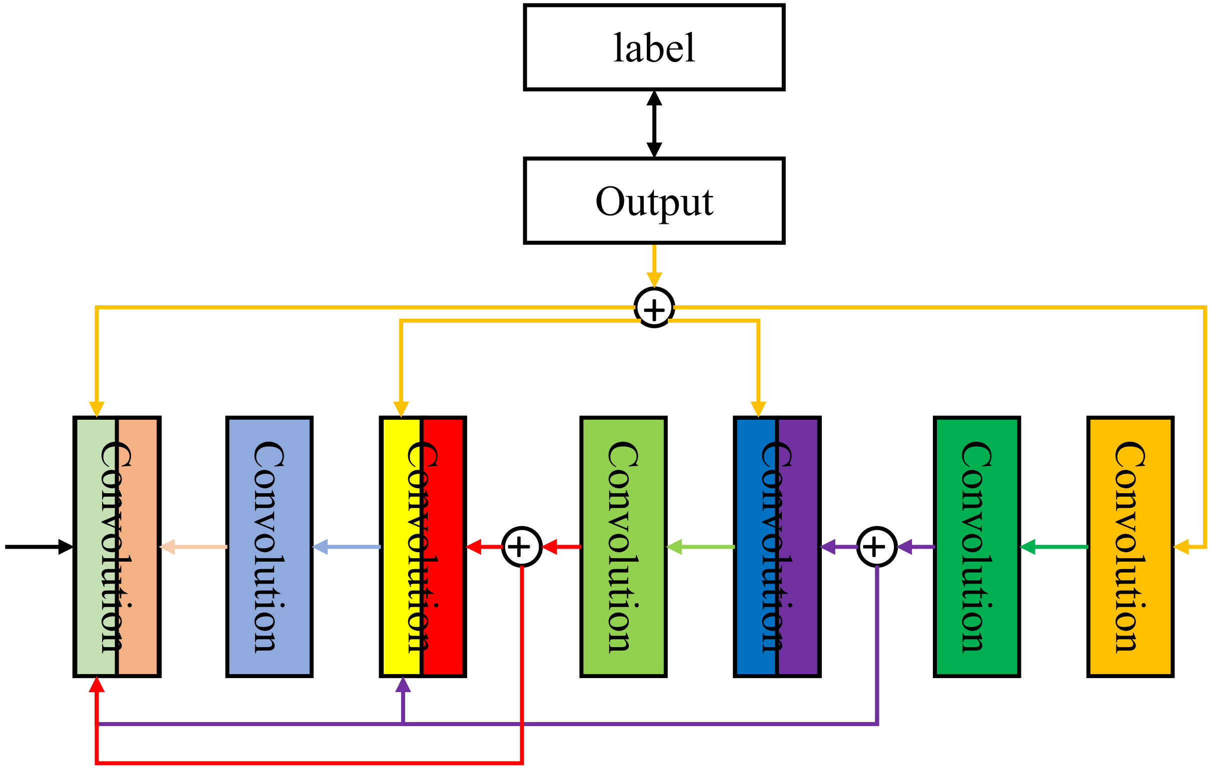
- Partial Residual Network (PRN) – layer-level design (2018)
- CSPNet – stage-level design (2019)
- ELAN, E-ELAN – network-level design (2021-2022)

Convolutional Deep NN has *many* stages

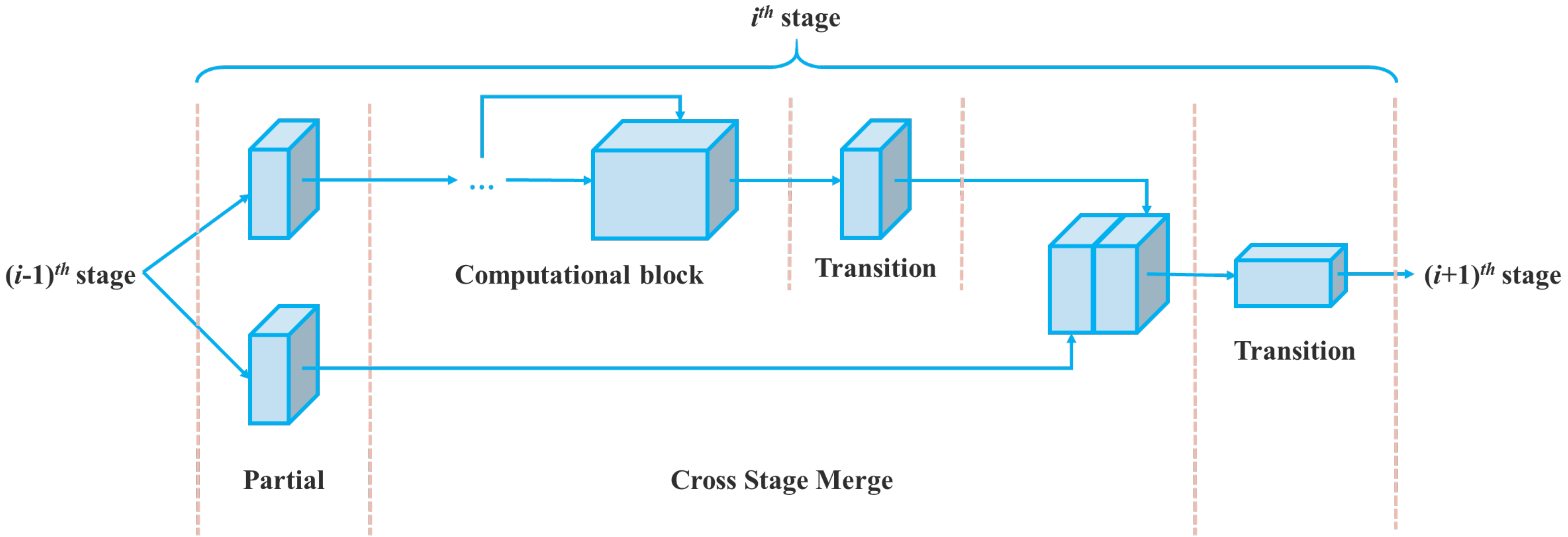


Partial Residual Network (2018/6-12)

layer-level design



Stage level design (CSPNet, June 2019)



What’s next after Stage-level Design ?

- Designing speedy and accurate object detector one has to consider the following:

1、network architecture；
2、feature integration method；
3、detection method；
4、loss function；
5、label assignment method；
6、training method。

YOLOv7 Research Contributions

1. Network architecture optimization
2. Training process optimization

Network architecture optimization:

How model scaling affect design of architecture? (1/4)

Local GPU(~ 50 layers)
e.g. Notebook

Mobile GPU
(< 10 layers)

Cloud GPU(> 100 layers)

Network architecture optimization:
How model scaling affect design of architecture? (2/4)

- Possible *scaling factors* include:
  - resolution (size of input image)
  - depth (no. of layers)
  - width (no. of channels)
  - stage (no. of feature pyramids)

to achieve a good *trade-off* for the amount of network parameters, computation, inference speed, and accuracy.

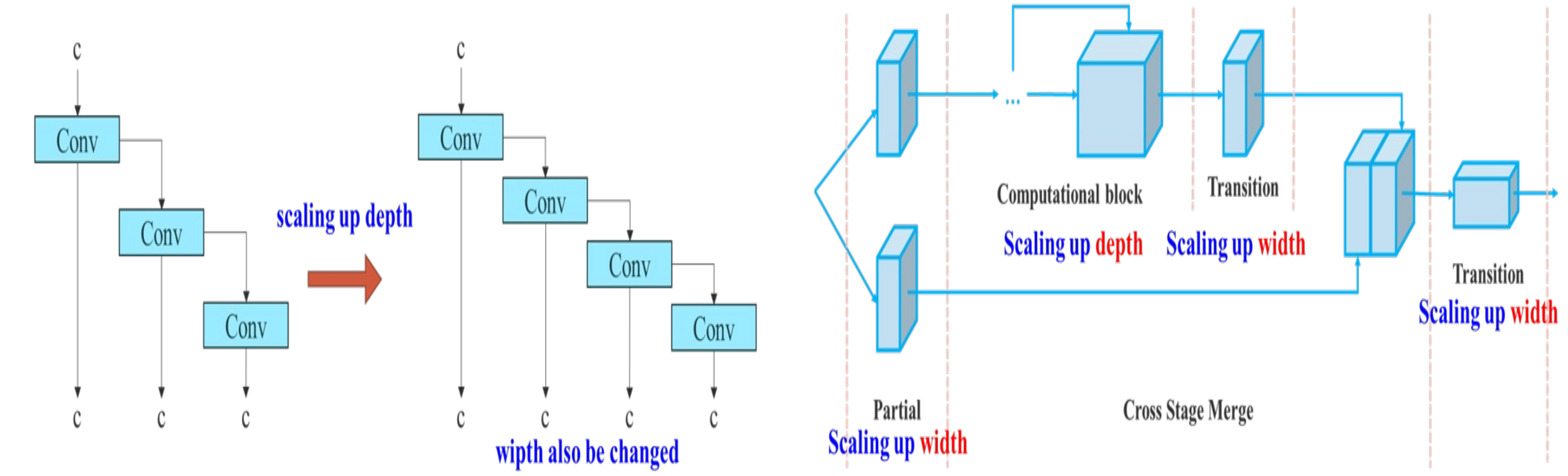
Network architecture optimization: How model scaling affect design of architecture? (3/4)

- Almost all model scaling methods analyze *individual* scaling factor *independently*
- *The problem comes*: In all *concatenation-based* models, when *depth* is scaled, the *width* of some layers will change too
- YOLOv7 is also *concatenation-based*, need *new compound scaling* method

Network architecture optimization:
How model scaling affect design of architecture? (4/4)

- For a concatenation-based architecture (ex. CSPNet, DenseNet)

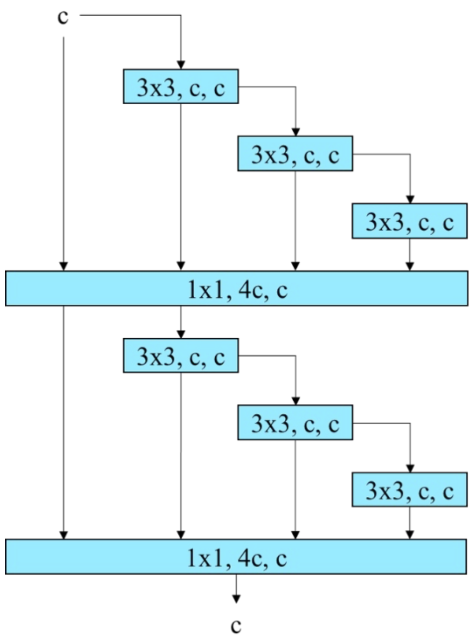
  - scaling up or down on depth, in-degree of subsequent transition layer will increase or decrease,



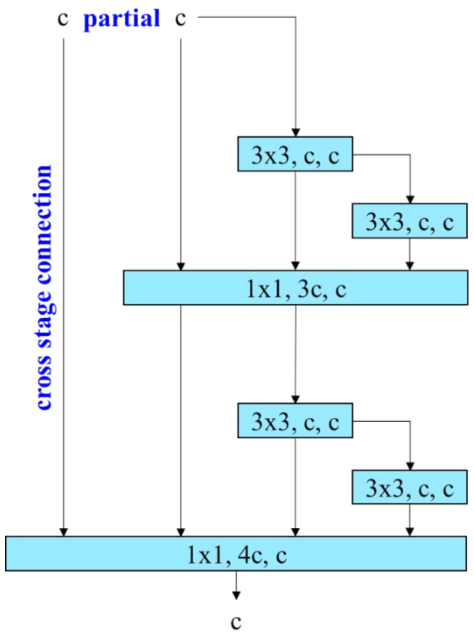
(a) concatenation-based model (b) scaled-up concatenation-based model (c) compound scaling up depth and width for concatenation-based model

Network architecture optimization:
Efficient Layer Aggregation Networks (ELAN) (1/4)

- CSPVoVNet(for tiny v4)是由 VoVNet 變化而來，其架構考慮參數量、運算量、計算密度，還分析了梯度路徑(gradient path)，使不同層的權重學到更多樣性的特徵，使推論速度更快、準確率越高。

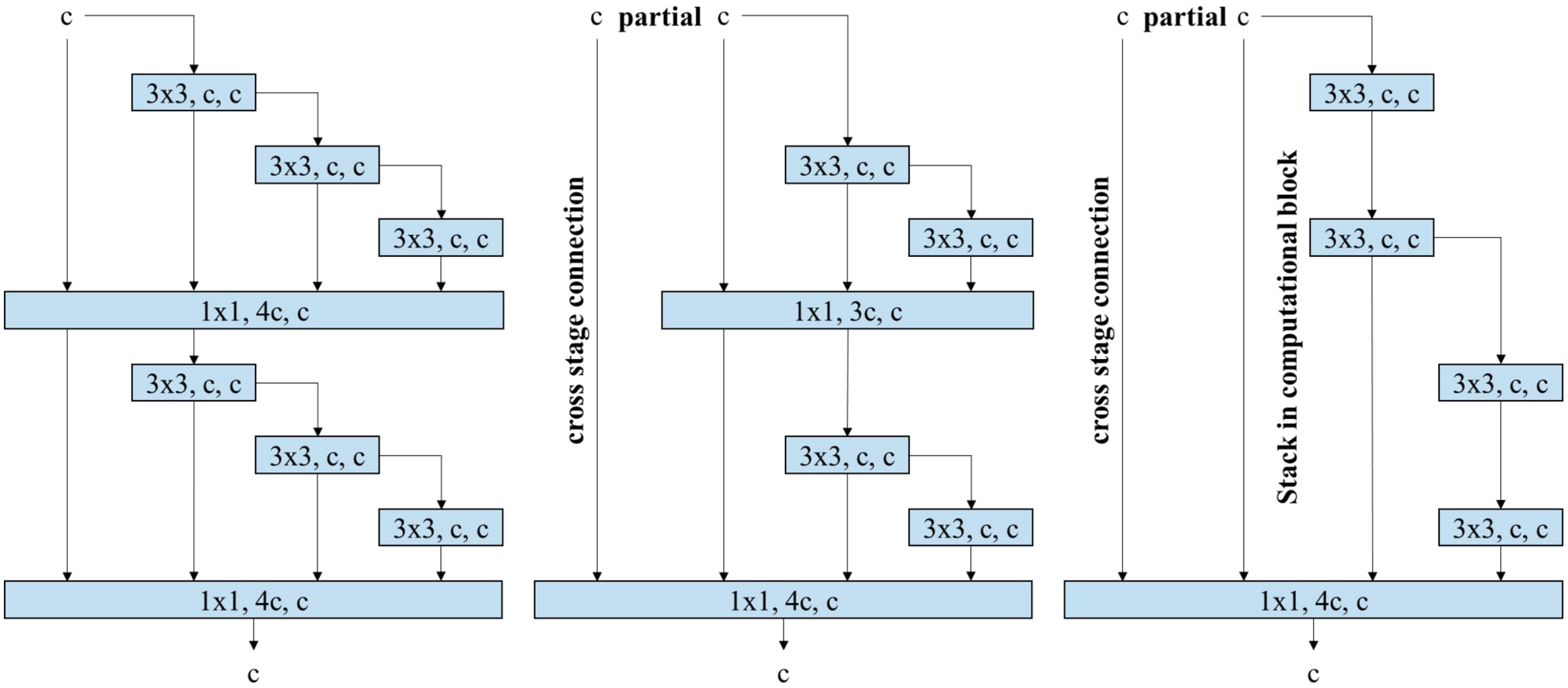


VoVNet



CSPVoVNet

Network architecture optimization:
Efficient Layer Aggregation Networks (ELAN) (2/4)



VoVNet

CSPVoVNet

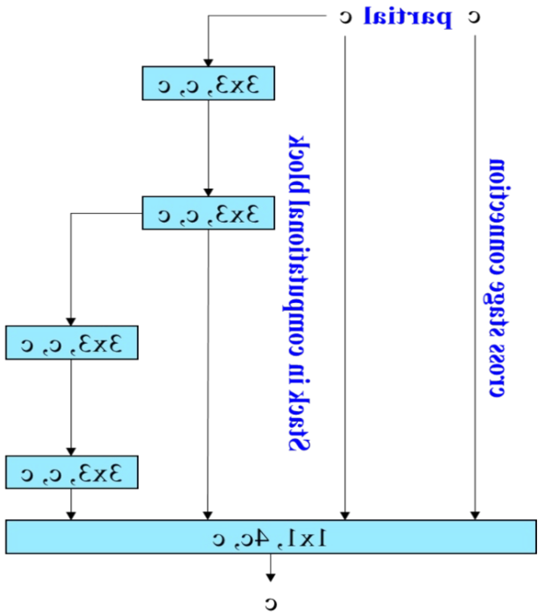
ELAN (March 2021)
Network-level design

Network architecture optimization:
Efficient Layer Aggregation Networks (ELAN) (3/4)

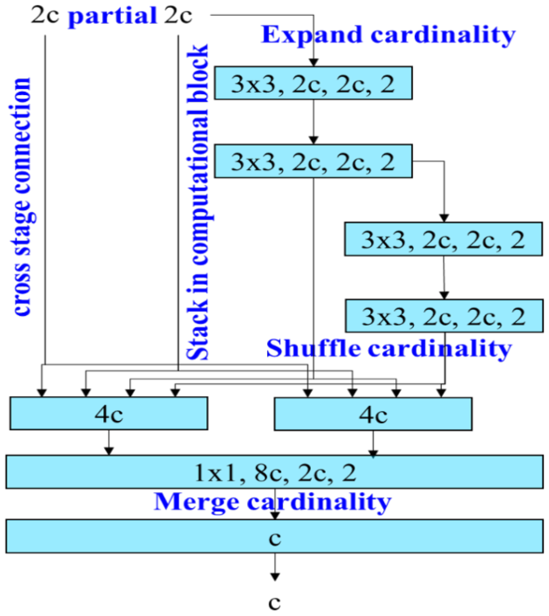
- In YOLOv7(July 2022), we consider the amount of memory the *gradient path takes*
- CONCLUSION: *the shorter the gradient path, the more powerful the network will be able to learn*

Network architecture optimization:
Extended Efficient Layer Aggregation Networks (E-ELAN) (4/4)

- ELAN (正常 v7) 控制最短和最長的 gradient path *讓網絡更有效的學習和收斂*。由於 model scaling會*破壞穩定的狀態*，YOLOv7 提出了 Extended ELAN (**適用大型 v7**)，透過 expand, shuffle, merge cardinality 等方式達到不破壞梯度路徑的目的，並增強網路學習能力



ELAN



E-ELAN

YOLOv7 Research Contributions

1. Network architecture optimization
2. Training process optimization –

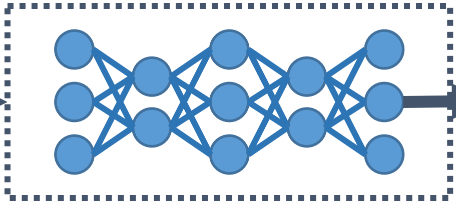
Introduction of implicit knowledge (YOLOR)

How YOLOR assist the training process ?(1/7)

→ What is this?	→ A Shiba Inu.
→ Where is the Shiba Inu?	→ In a room.
→ Where is she?	→ In a room.
→ What is she doing?	→ LOL.
→ What is her name?	→ A I.
→ Do you love her?	→ Yes! Sure! Of course!

Same Input + Different Objectives = Different Answers
(Single model for multiple tasks)

How YOLOR assist the training process ?(2/7)



Extracted Representation

What is this?

Where is the Shiba Inu?

What is she doing?

What is her name?

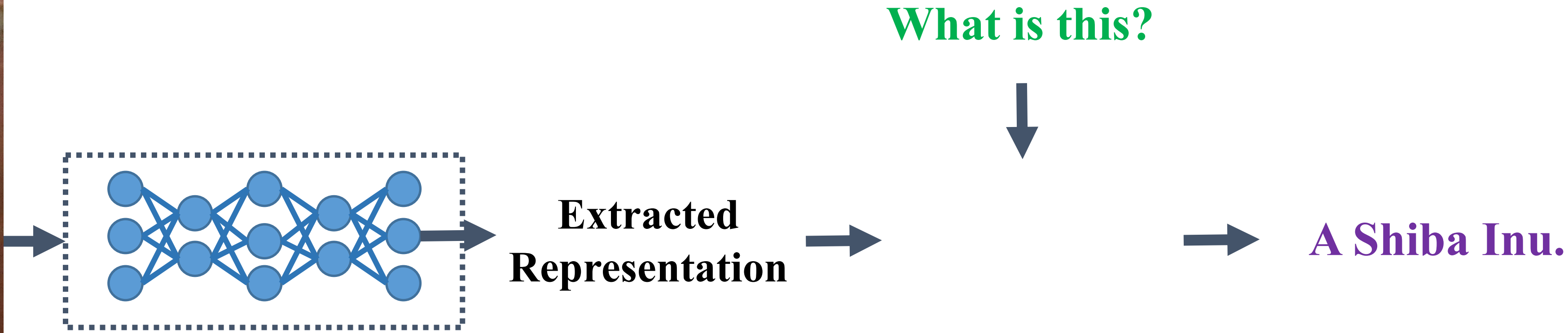
A Shiba Inu.

In a room.

LOL.

A I.

How YOLOR assist the training process ?(3/7)



Difficulty 1 – **training did not prepare multi-tasks**

Difficulty 2 – **no single representation can answer all questions**

Difficulty 3 – **different tasks may influence each other**

How YOLOR assist the training process ?(4/7)

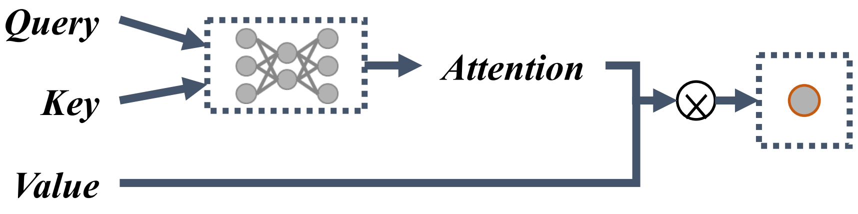
Classification

shiba

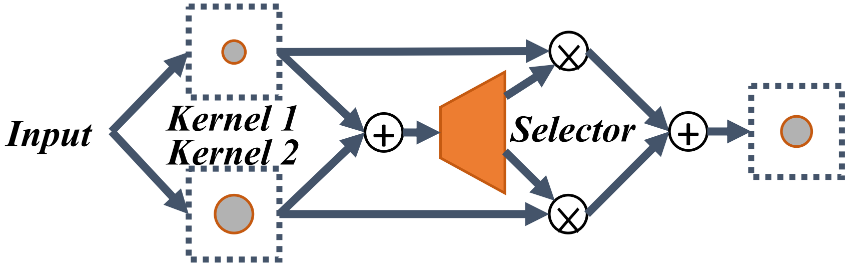
Identification

acoshiba kabosumama marutaro

How YOLOR assist the training process ?(5/7)



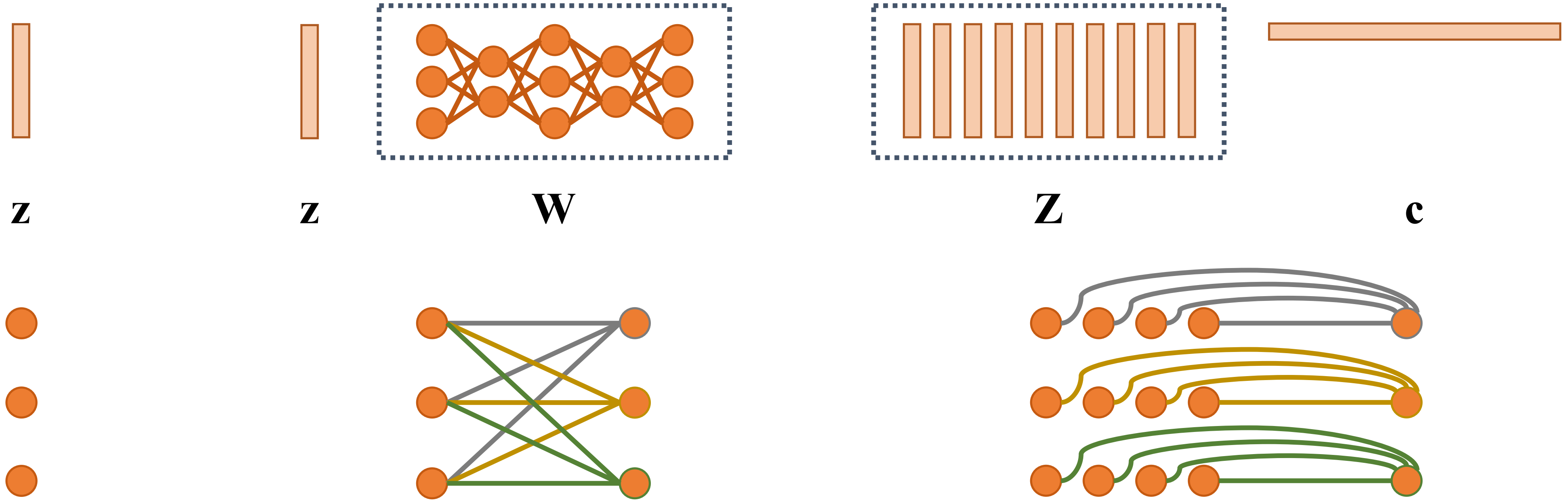
Transformer



Selective Kernel

Model Explicit Knowledge

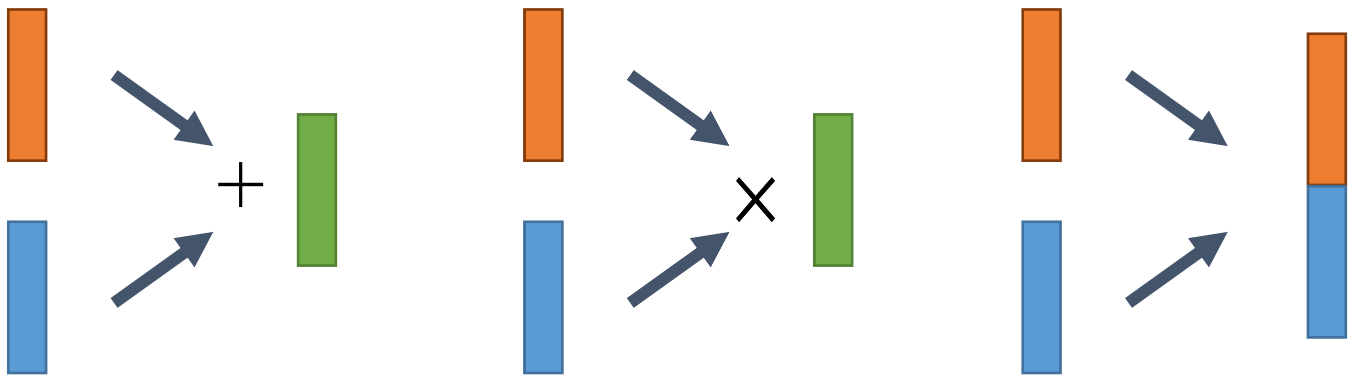
How YOLOR assist the training process ?(6/7)



(a) Vector (single task) (b) Neural network (single task) (c) Matrix factorization (multi-task)

Model Implicit knowledge

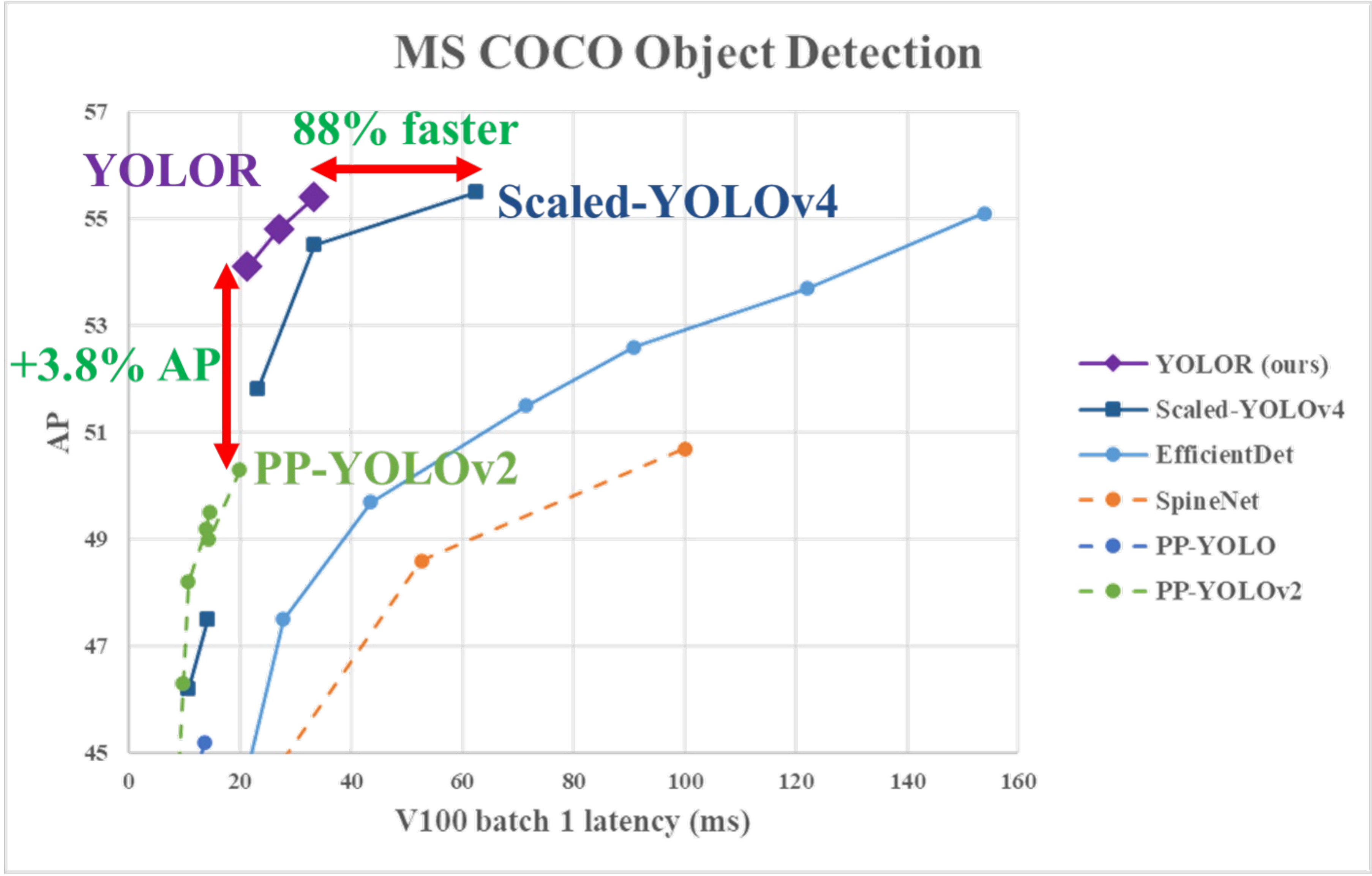
How YOLOR assist the training process ?(7/7)



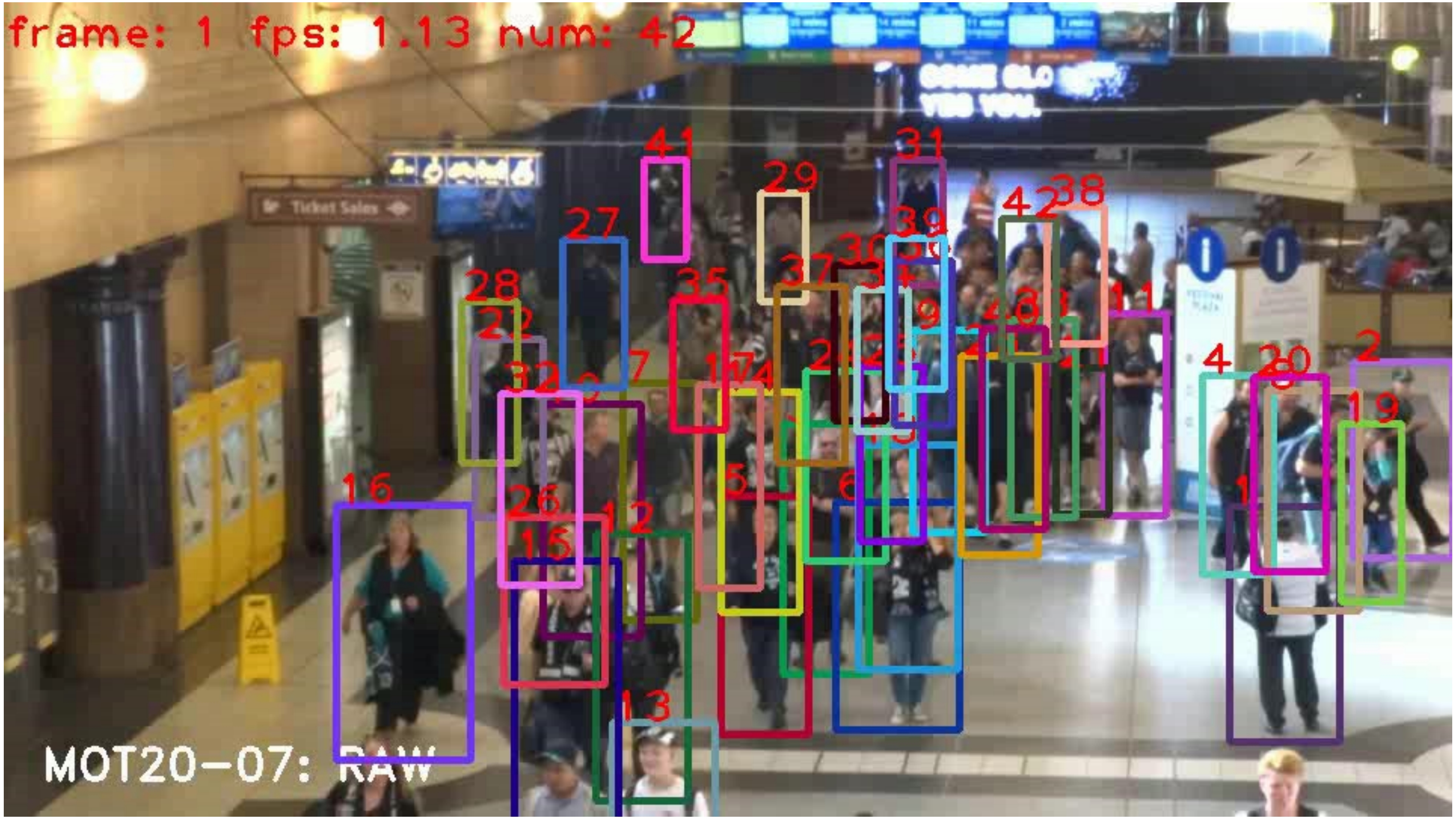
(a) addition (b) multiplication (c) concatenation

Combine explicit knowledge and implicit knowledge

Performance



YOLOR - MOT

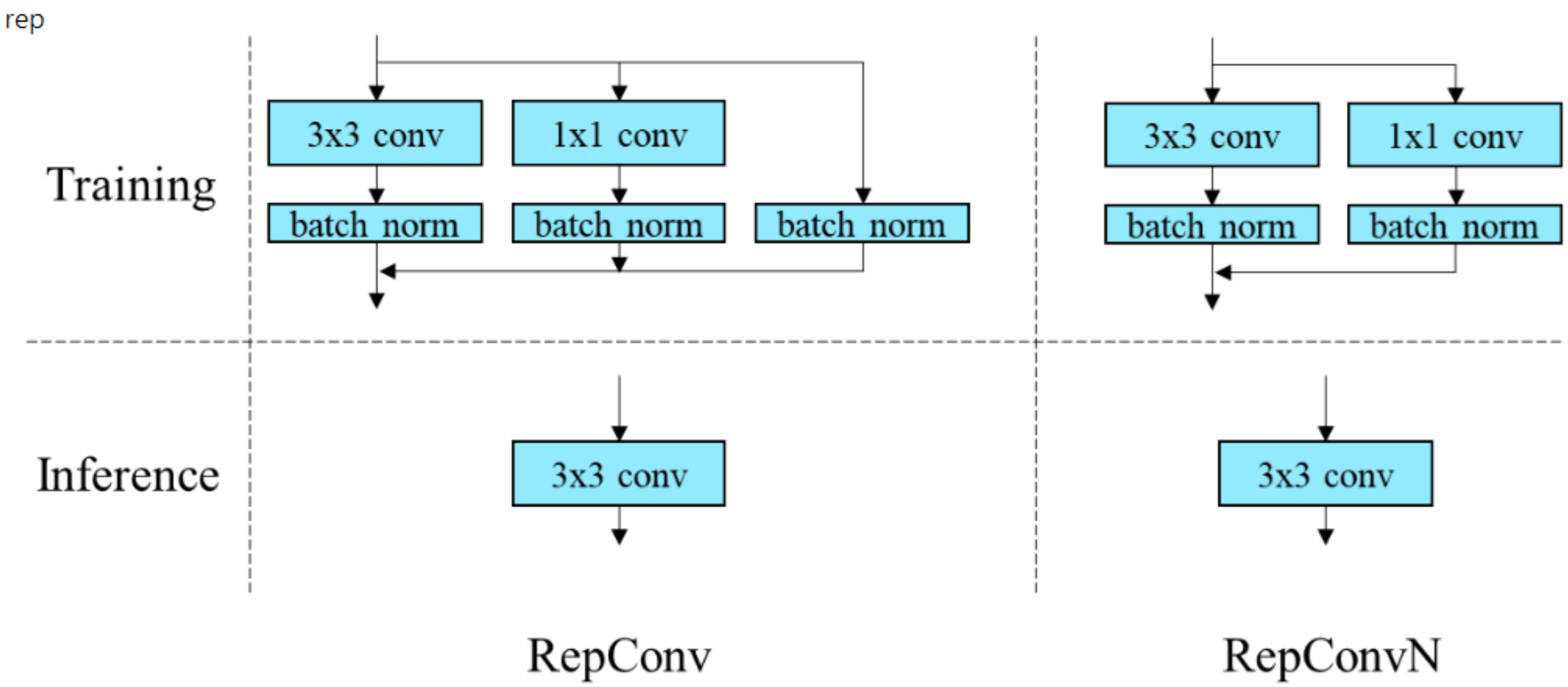


YOLOv7 Research Contributions

1. Network architecture optimization
2. Training process optimization – model reparameterization

Model Reparameterization(1/4)

- *RepConv* combines 3x3 convolution, 1x1 convolution, and identity connection in one convolutional layer
- *RepConv* achieves excellent performance on *VGG* (PlainNet)



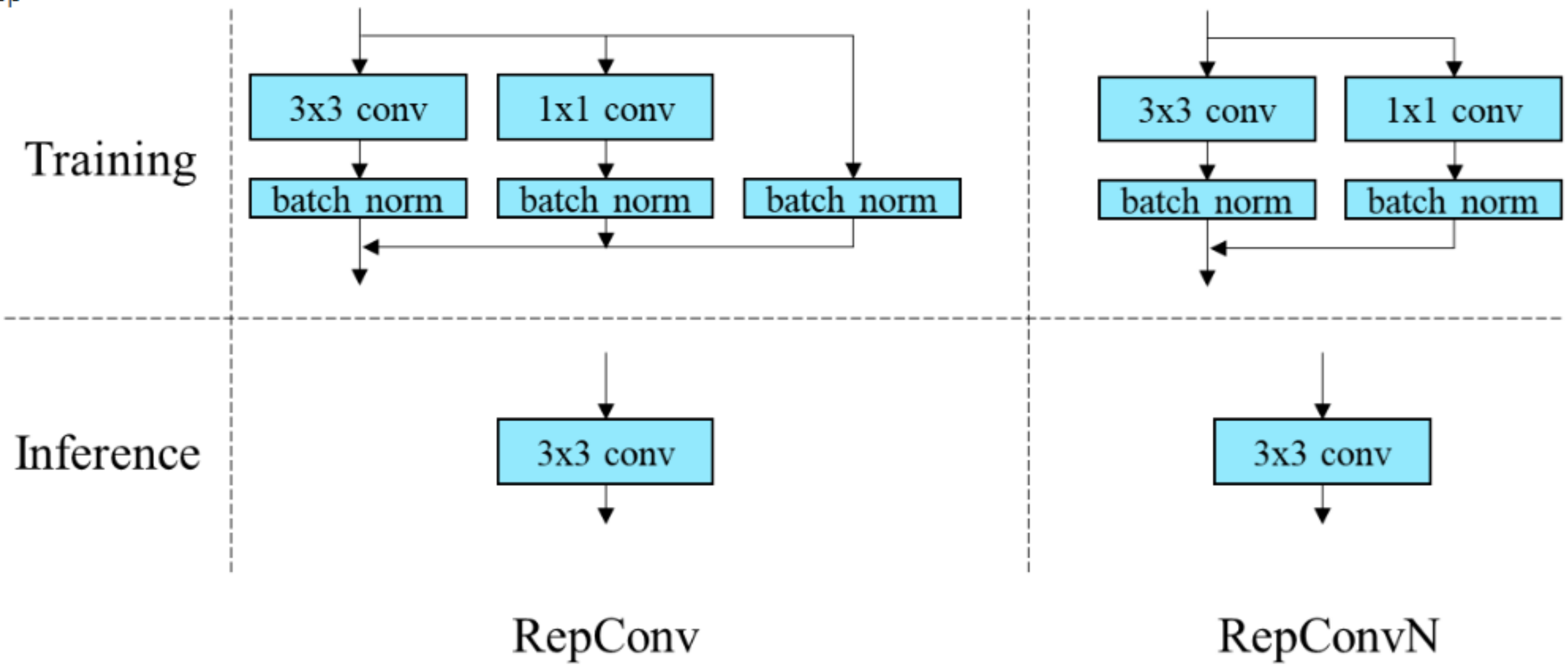
Model Reparameterization(2/4)

- But, when applied to ResNet (with residual) or DenseNet (with concatenation), accuracy will be *significantly dropped*
- *Solution is re-parameterized convolution*

Model Reparameterization(3/4)

- *identity connection* in *RepConv* destroys the *residual* in ***ResNet*** and the *concatenation* in *DenseNet*, and these designs provide more *diversity of gradients* for different feature maps

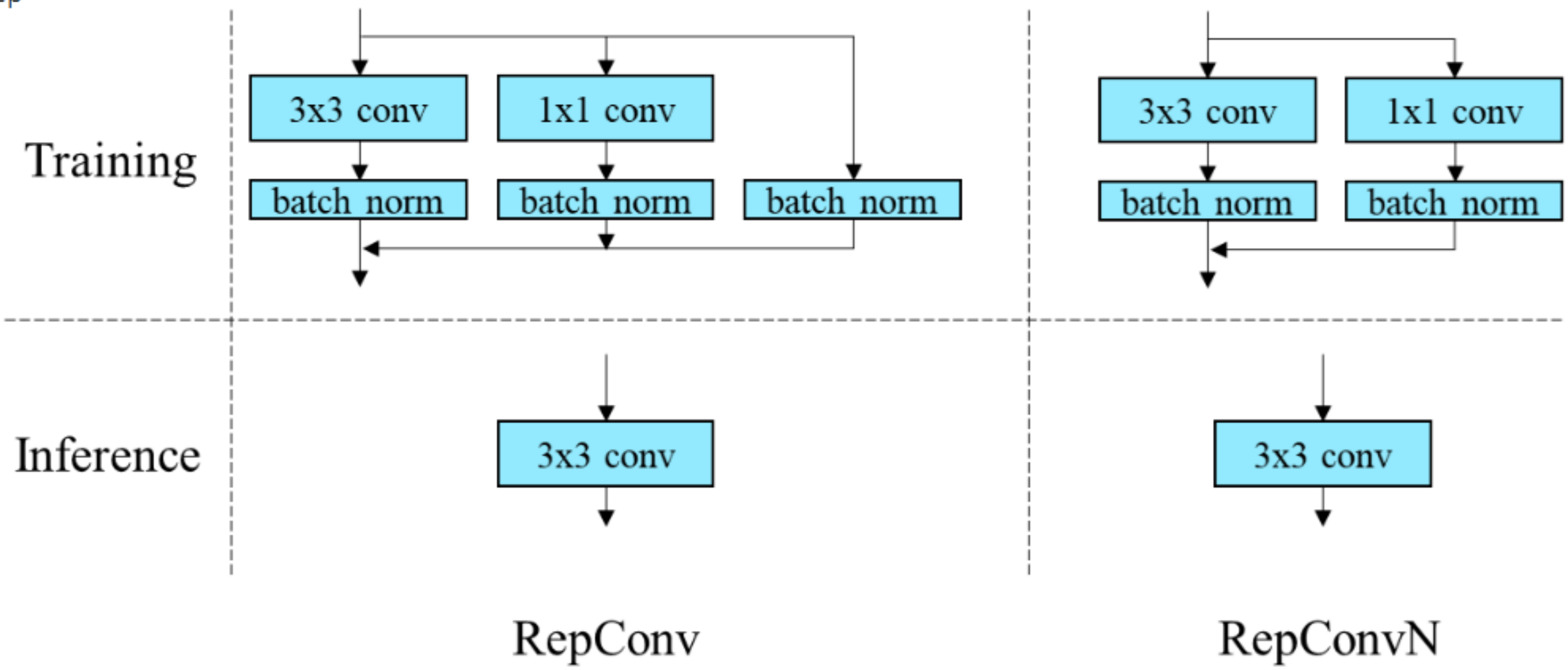
rep



Model Reparameterization(4/4)

SOLUTION: use *RepConvN* (RepConv *without* identity connection) to the training process (make use of ***masking*** technique)

rep

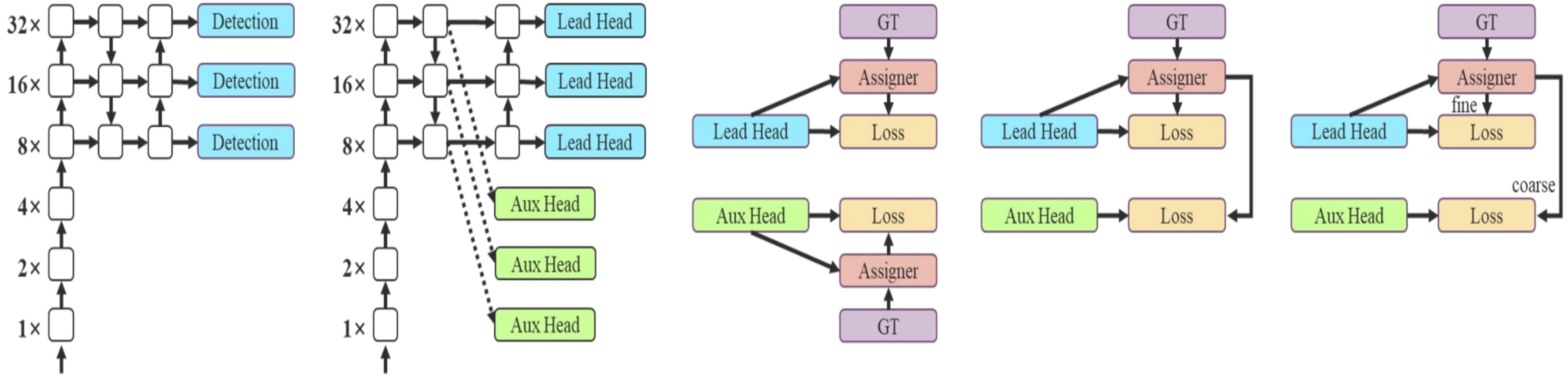


YOLOv7 Research Contributions

1. Network architecture optimization
2. Training process optimization – Dynamic label assignment

Dynamic label assignment strategy(1/2)

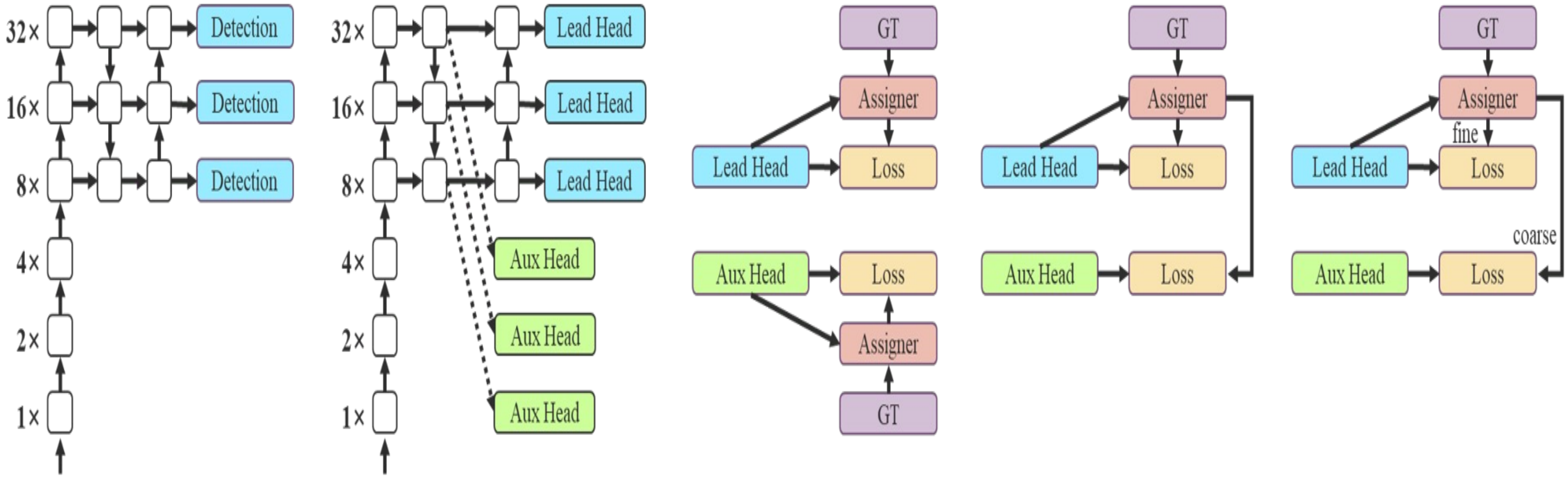
- Deep supervision is often used in training deep networks
- *Main concept*: add extra auxiliary head in middle layers, and use *shallow network weights with assistant loss* as guide



(a) Normal model (b) Model with auxiliary head (c) Independent assigner (d) Lead guided assigner (e) Coarse-to-fine lead guided assigner

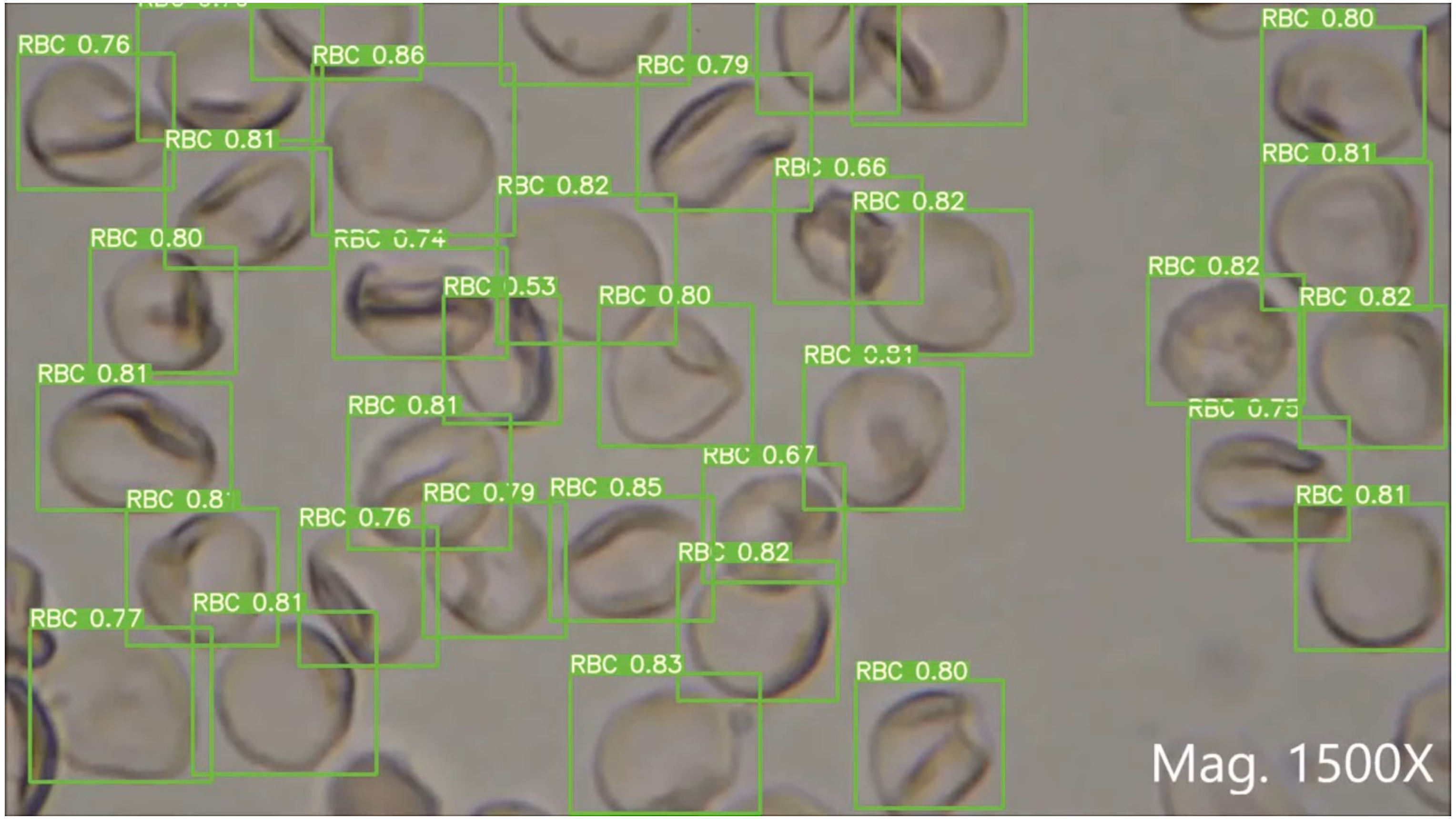
Dynamic label assignment strategy(2/2)

- Proposed new method: Use lead head prediction as guidance to generate coarse-to-fine hierarchical labels, such as (d) and (e)



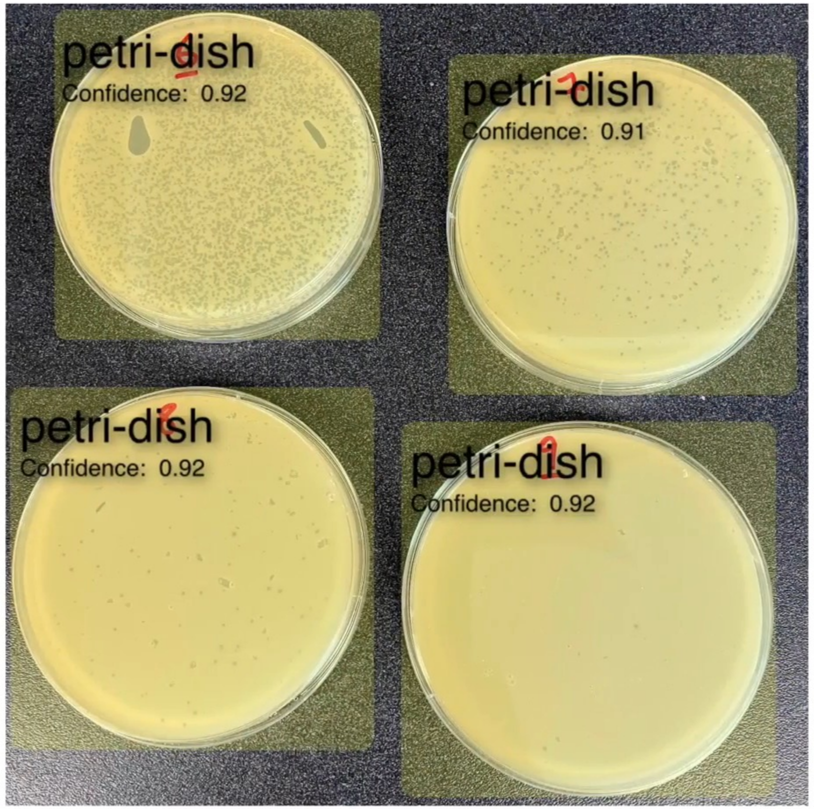
(a) Normal model (b) Model with auxiliary head (c) Independent assigner (d) Lead guided assigner (e) Coarse-to-fine lead guided assigner

YOLOv7 及 YOLOv9 的應用



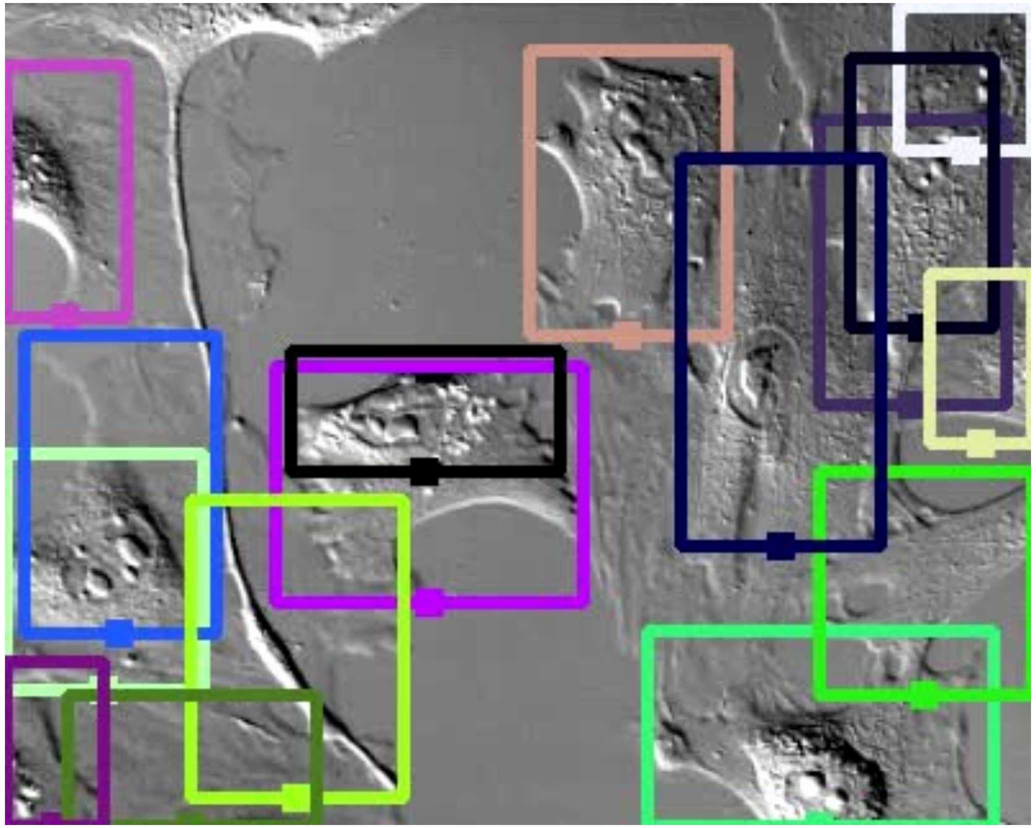
血球偵測

https://www.youtube.com/watch?v=1Al4s5x7EAU



噬菌體培養皿監測

https://imgur.com/1u8ybg5

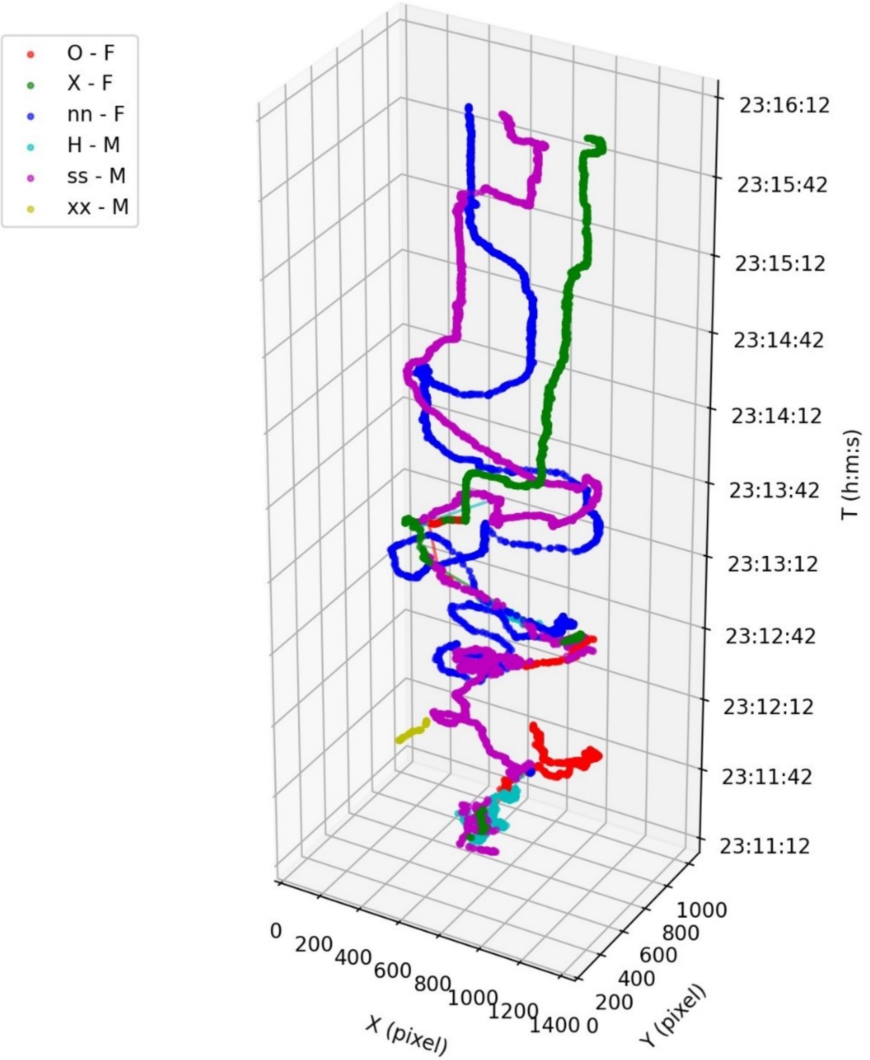


細胞有絲分裂檢測, 追蹤和譜系

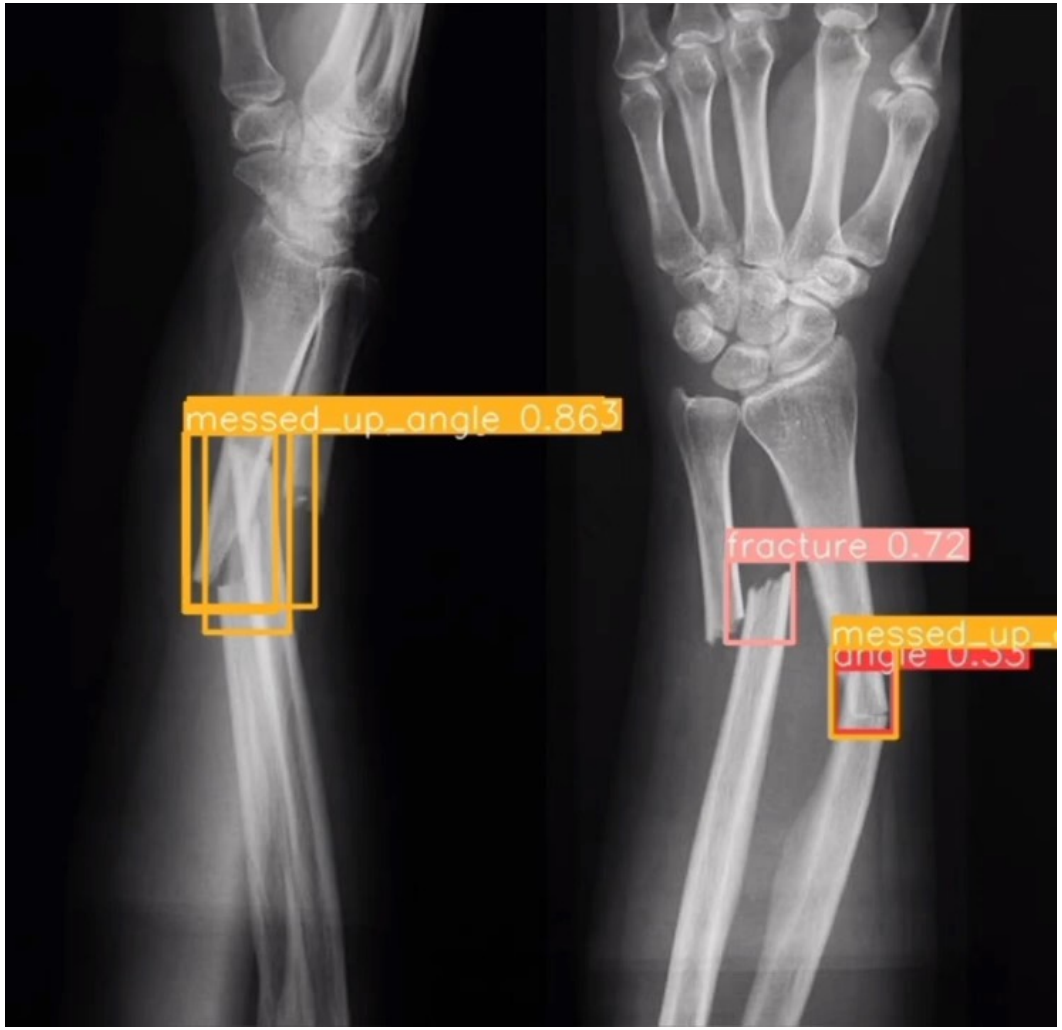
https://github.com/CIVA-Lab/EDNet-YOLO

Burying Beetle Tracking – Results

Tracked videos



(o, x, nn, represent different females and H, ss, xx, represent different males in the group)

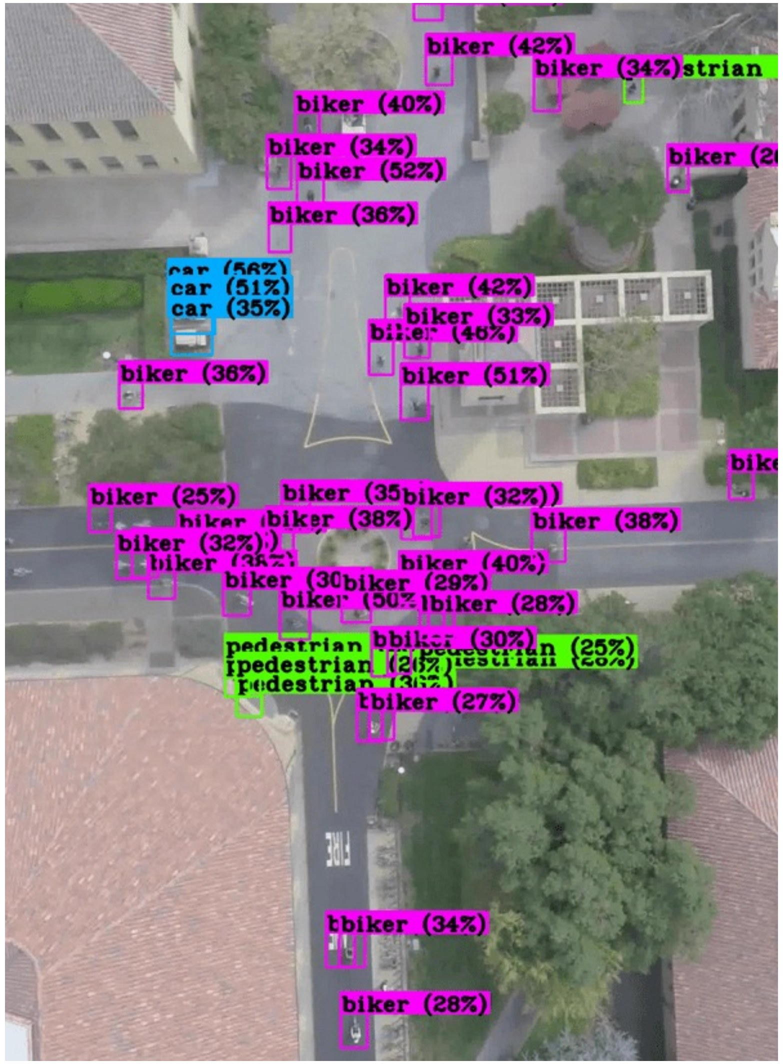


骨折檢測

https://www.youtube.com/watch?v=sKIeHLPVHBQ

國防

https://learnopencv.com/fine-tuning-yolov9/



https://github.com/soumyadbanik/object-detection-on-aerial-videos

Drone Tracking

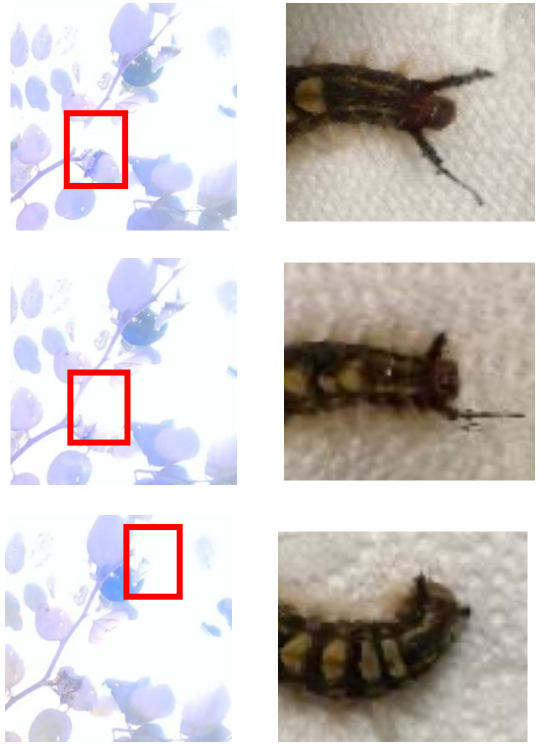
https://www.youtube.com/watch?v=kLVupa1nkZs

Multiple Caterpillar Tracking via YOLOv7 + SORT

Detection and tracking at 15-20 cm (YOLOv7+SORT)

Adaptive Photonics Lab, NYCU

Laser Pest Control:
Multi-object Tracking + Elimination



生命科學的應用

從生物影像或影片出發

生醫所陳建璋老師
老鼠實驗

陳老師實驗室的大方向

能否使用人工智慧技術分辨老鼠是不是有慢性疼痛？

腳爪是否因為有害的熱 化學或機械刺激而縮回

行為分析:嗅聞 行走 側走 走走停停 跳躍 爬行 梳理毛髮 直立 搖頭和舔

- 老鼠的``XX”疾病有那些症狀？
- 資訊所的人能利用 AI 幫忙處理那些問題？
  - 資訊所的人希望能 win-win
  - 找出最重要的切入點,共同設計實驗

能否使用人工智慧技術來區分不同神經病變的疾病小鼠？

- 野生型
- 慢性疼痛
- 帕金森症
- 阿茲海默症
- 亨丁頓舞蹈症
- 肌萎縮性脊髓側索硬化症

- 行為: 焦慮 抑鬱 身體或肌肉異常震動
- 嗅聞 行走 側走 走走停停 跳躍 爬行 梳理毛髮 直立 搖頭和舔

- 陳老師說上述的異常狀況可能反映在
  - 步伐
  - 身體某些部位的疼痛
  - 老鼠臉部表情的呆滯
  - etc.

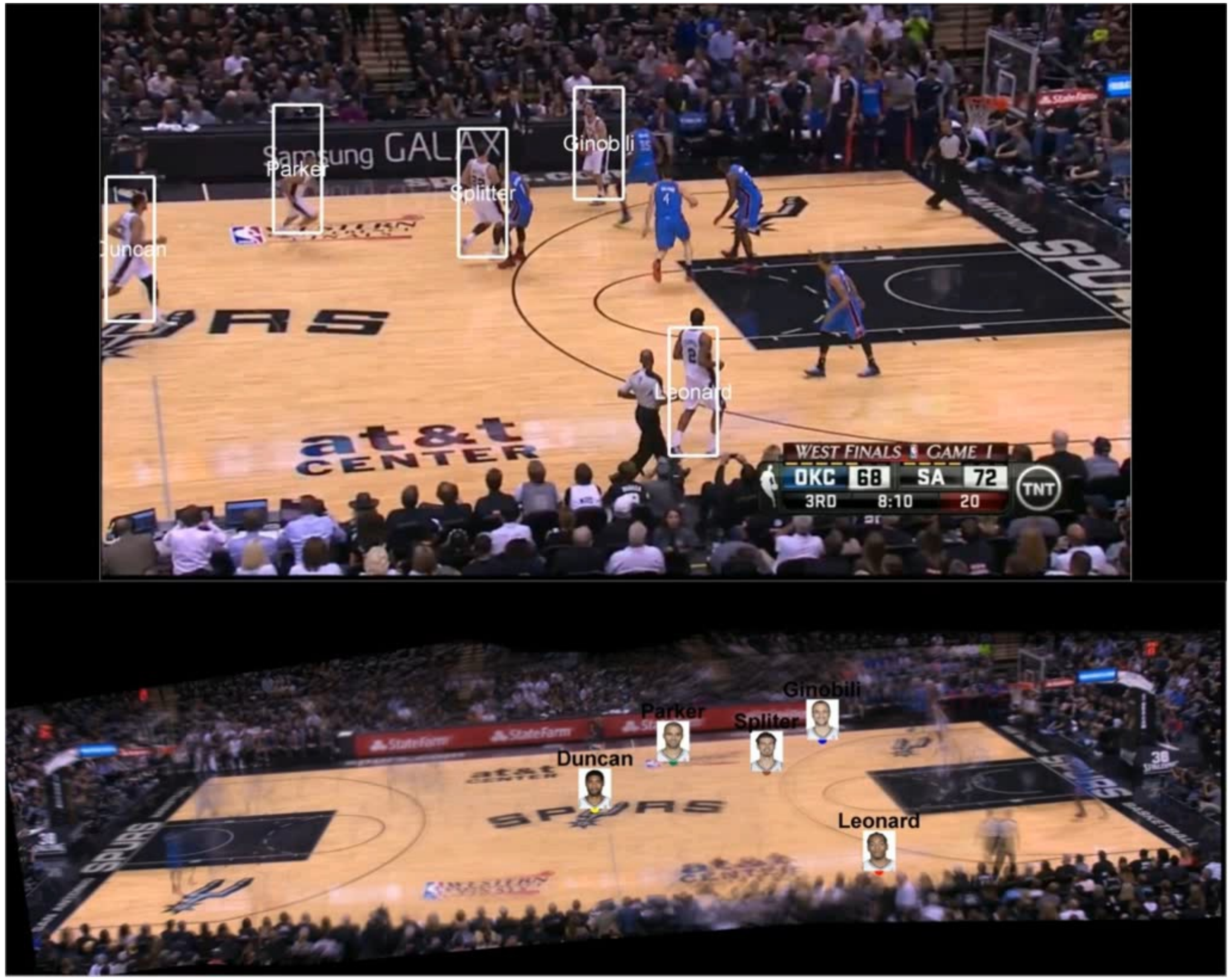
如何設計實驗？

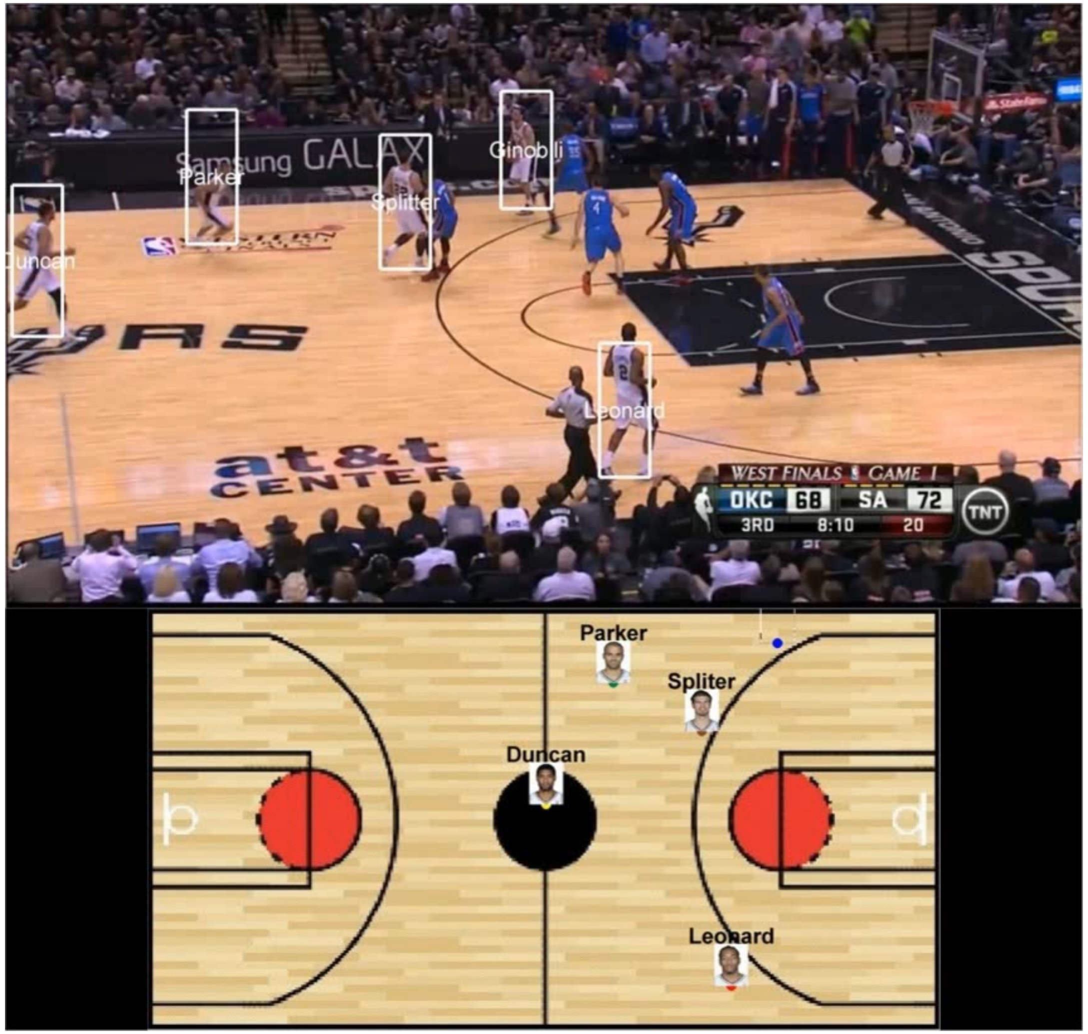
- 步伐
- 身體某些部位的疼痛
- 老鼠臉部表情的呆滯
- etc.

針對步伐如何設計實驗？

- 過去已經有的類似技術(籃球戰術分析)
- 針對老鼠步伐分析需要什麼設備?

將公牛隊使用 wind wheel 戰術的片段找出來


Parker
Splitter
Ginobili
Duncan
Leonard
WEST FINALS GAME 1
OKC 68 SA 72
3RD 8:10 20
TNT
Parker
Ginobili
Spliter
Duncan
Leonard


Parker
Splitter
Leonard
WEST FINALS GAME 1
OKC 68
SA 72
3RD 8:10 20
TNT
Parker
Spliter
Duncan
Leonard

- Tactics Analysis based on *spatiotemporal trajectories* of 5 offense players

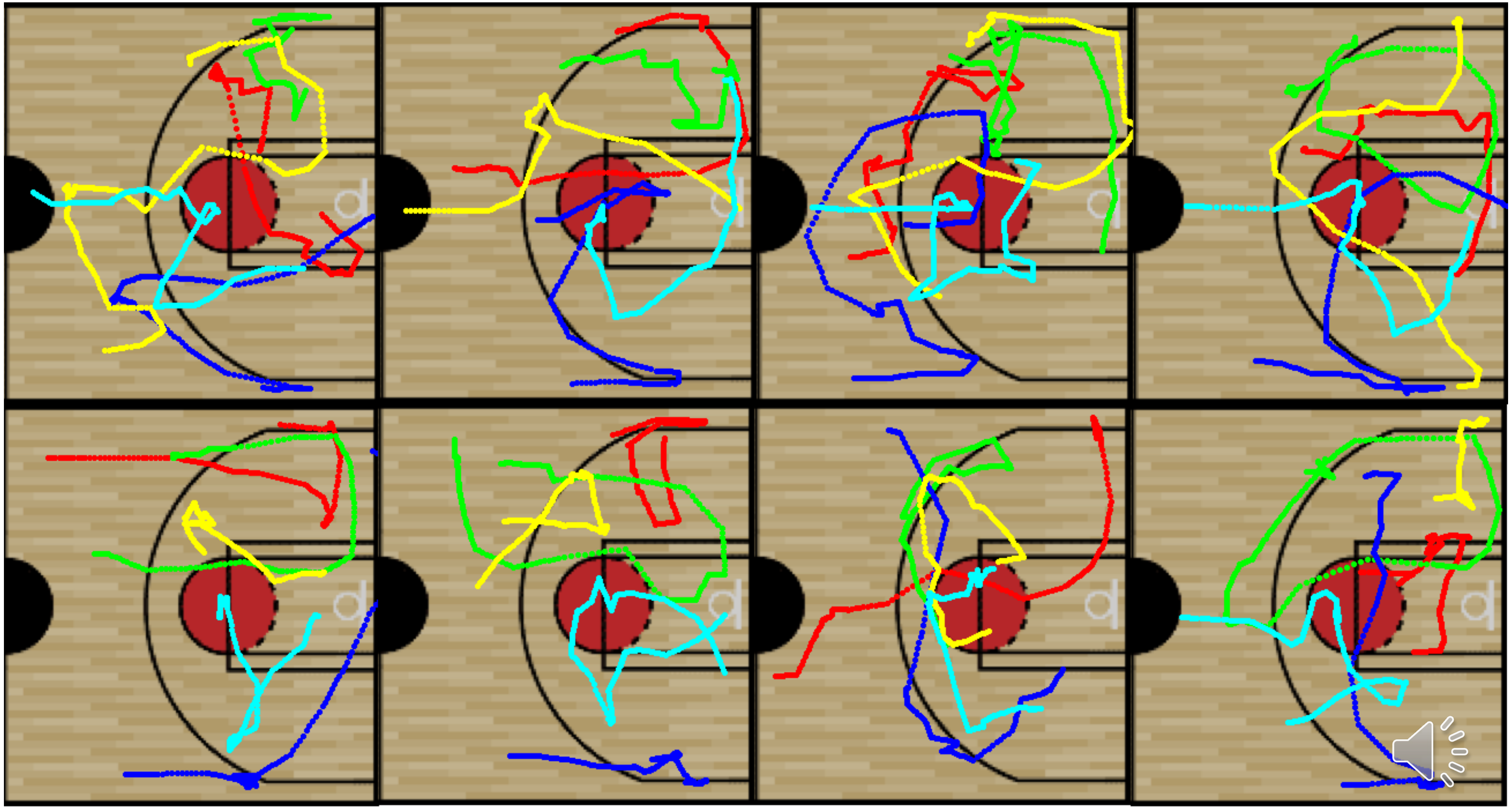
A Two-Stage Un-supervised Clustering for Tactic Analysis

- Stage-1: Un-supervised clustering of all available tactics based on their mutual distances

- Stage-2: Un-supervised clustering of all tactics clustered into the same cluster in Stage-1 (try to separate the role of each offense player)

Second-stage: how to model an offense strategy ?

- 8 different trajectory sets of right hawk, each consists of 5 trajectories generated by 5 offense players



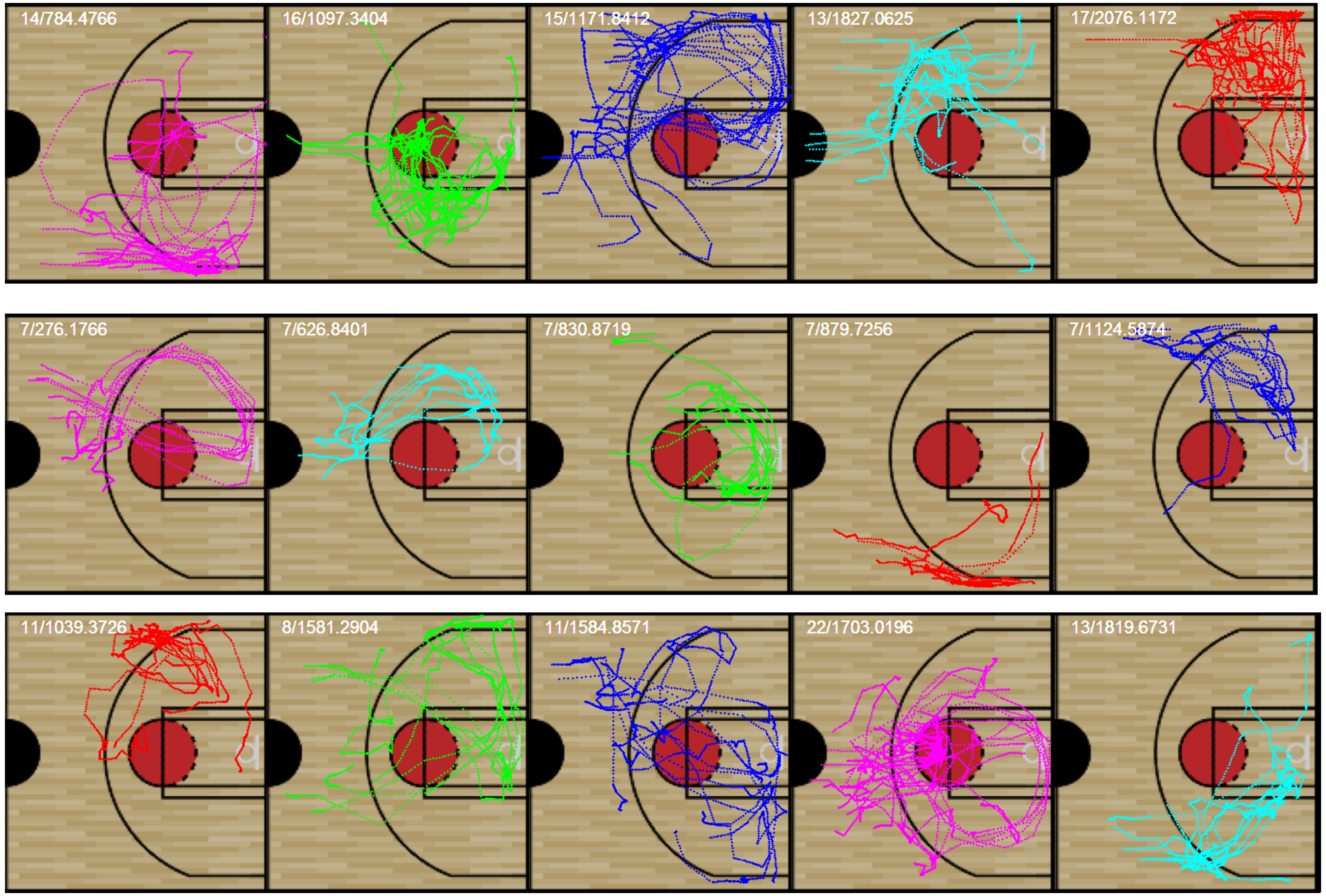
Clustering by Trajectory Distance

- Based on *the distance between trajectories,* one can separate each group of *tactics* into five group of trajectories, each corresponds to a role (an offense player)

Hawk

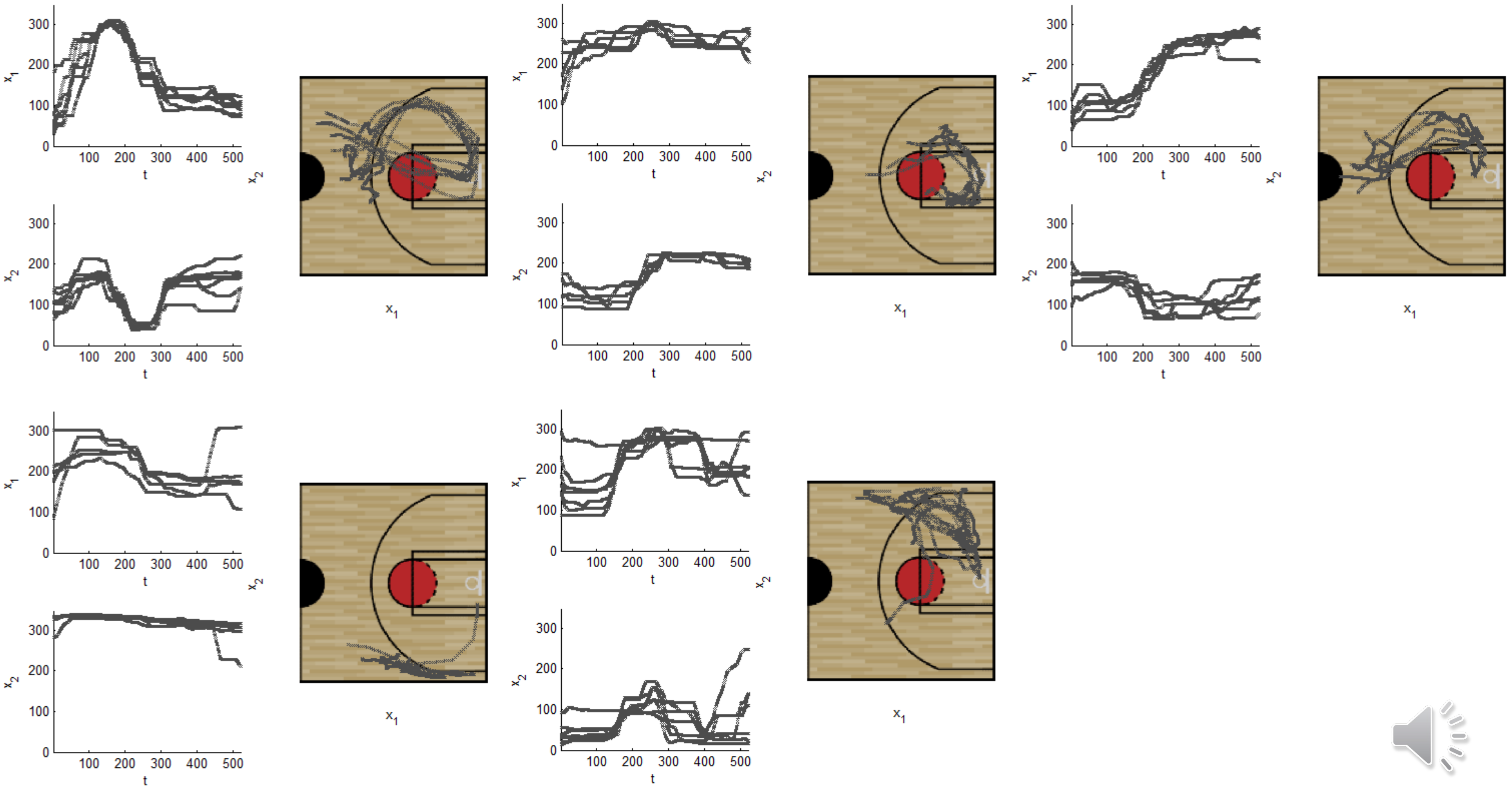
Wing Wheel

Princeton

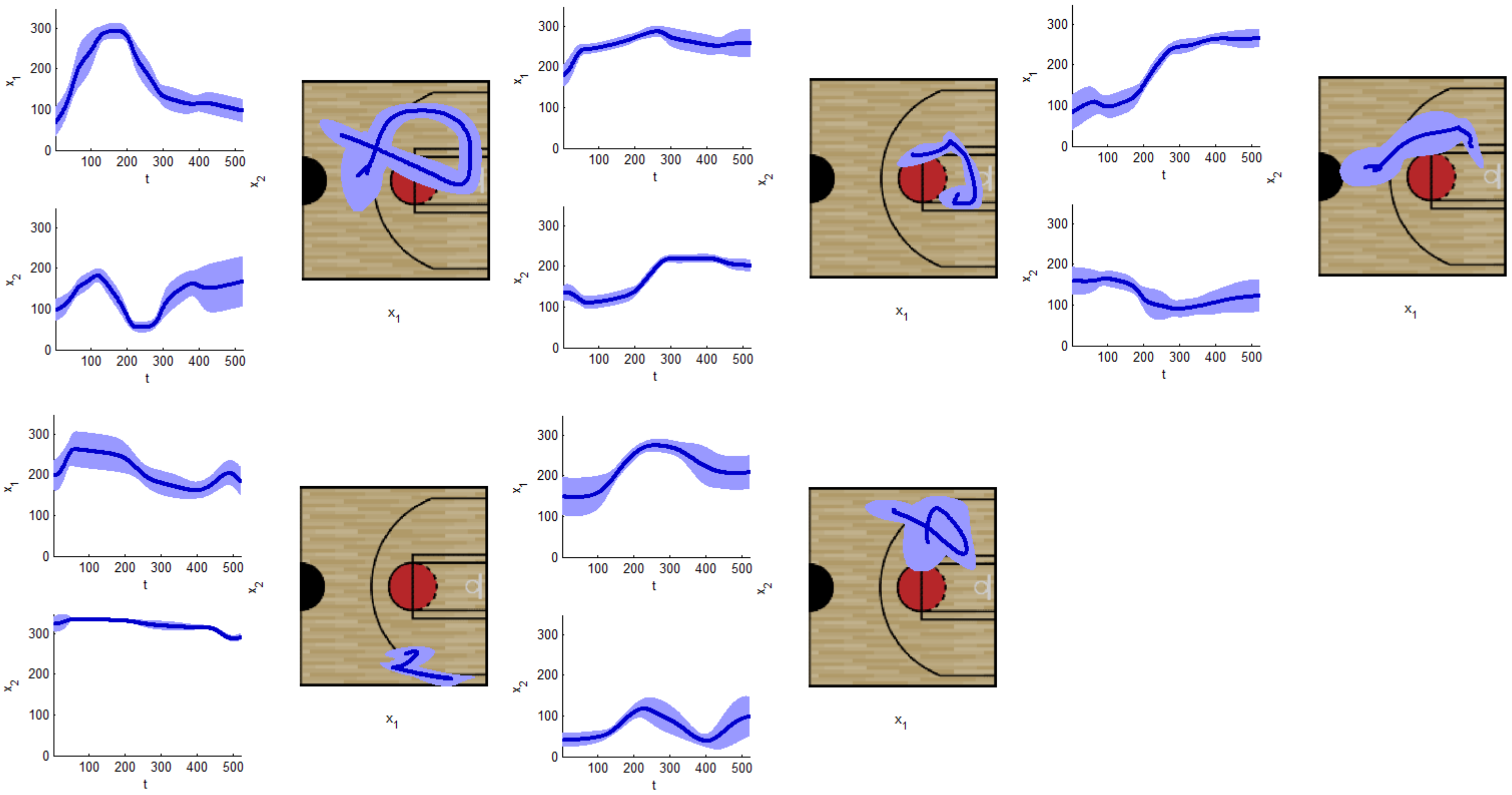


Temporal Alignment

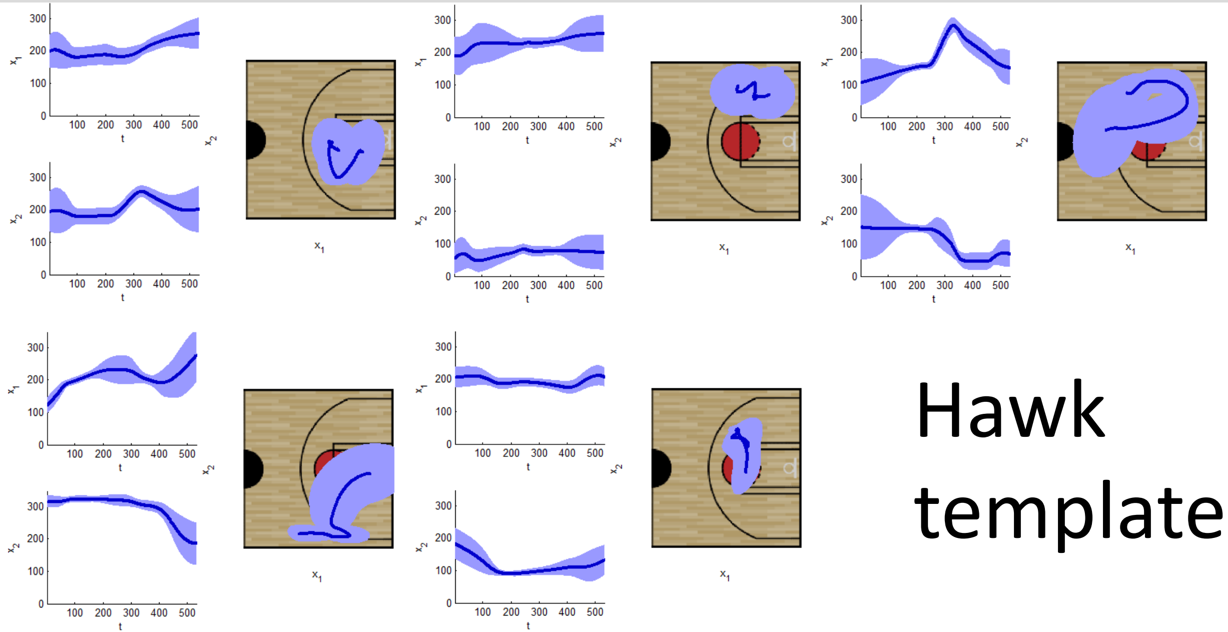
For each role, we use the *velocities along x- and y-direction*, respectively, to model it (use DTW to solve the alignment problem)



The Built Model



Demo _ Classification

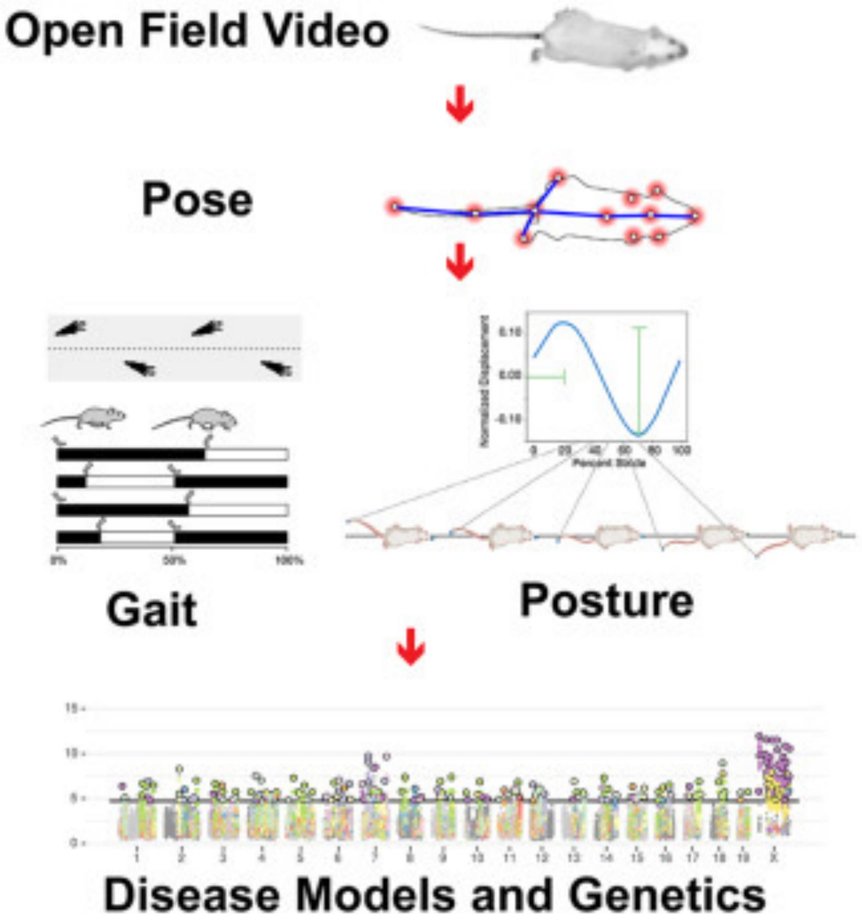


Hawk template

針對步伐如何設計實驗？

- 過去已經有的類似技術(籃球戰術分析)
- 針對老鼠步伐分析需要什麼設備?

根據老鼠的走路分析是否為自閉症

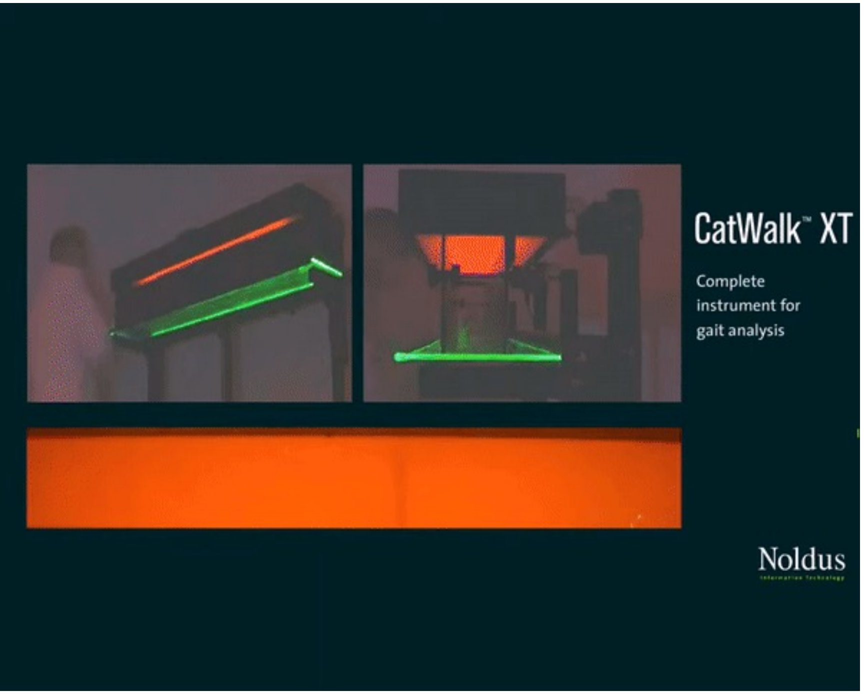


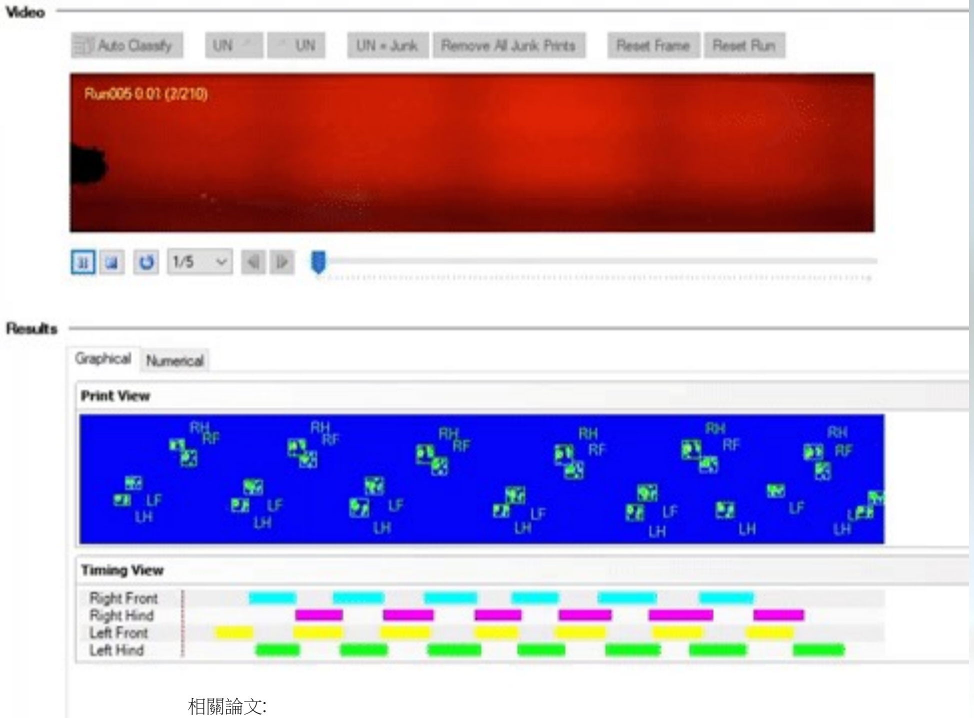
[Stride-level analysis of mouse open field behavior using deep-learning-based pose estimation - ScienceDirect]

Keith Sheppard, Justin Gardin, Gautam S. Sabnis, Asaf Peer, Megan Darrell, Sean Deats, Brian Geuther, Cathleen M. Lutz, Vivek Kumar, "*Stride-level analysis of mouse open field behavior using deep-learning-based pose estimation*", Cell Reports, Volume 38, Issue 2, 2022,

老鼠腳印壓力檢測系統 Catwalk XT

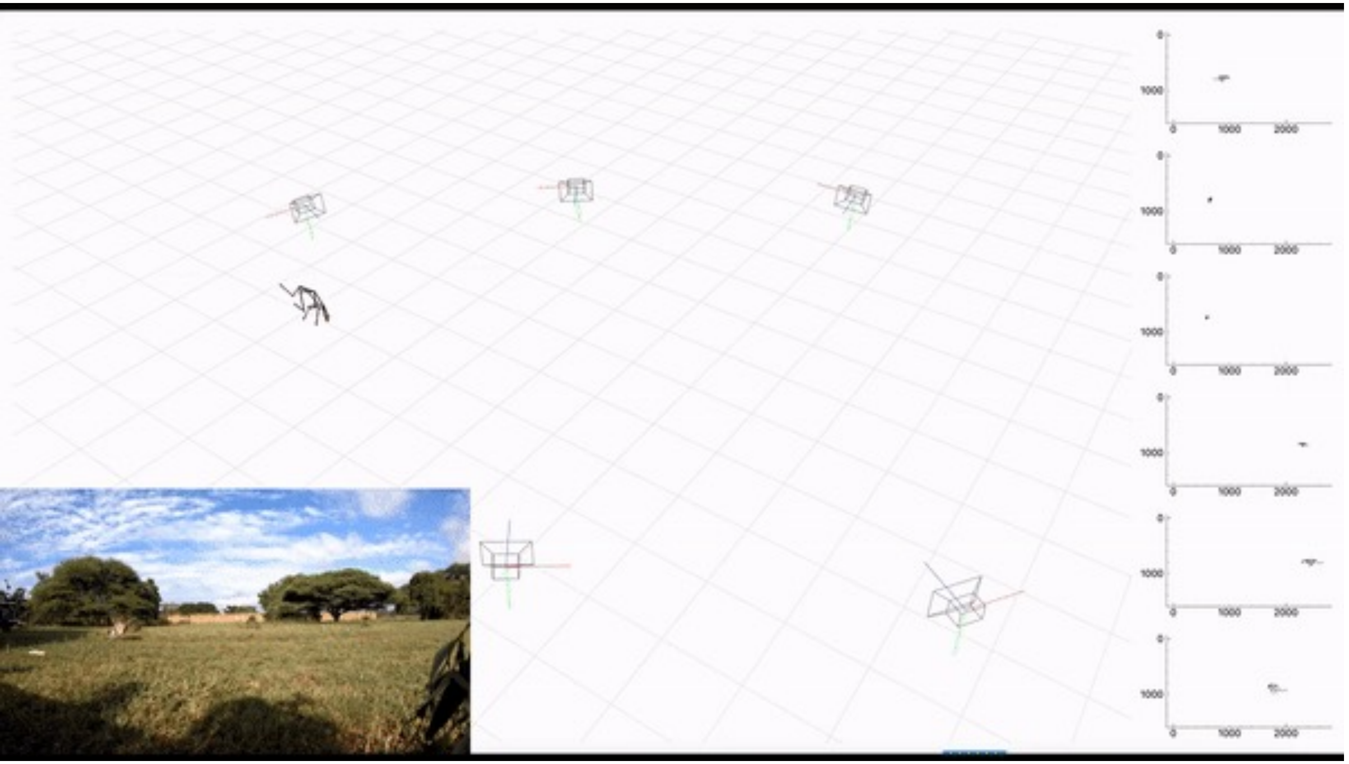
- 機器大小:130 x 68 x 152 cm
- 108年決標價格200萬
- 原理: 玻璃走道，檢測老鼠的軌跡，腳印重量會影響綠光LED折射，有判定腳印輕重的功能。(關節炎)





相關論文:
1. Ritter J, Menger M, Herath SC, Histing T, Kolbenschlag J, Daigeler A, Heinzel JC, Prahm C. Translational evaluation of gait behavior in rodent models of arthritic disorders with the CatWalk device - a narrative review. Front Med (Lausanne). 2023 Oct 6;10:1255215. doi: 10.3389/fmed.2023.1255215. PMID: 37869169; PMCID: PMC10587608.
2. Kristina Ängeby Möller, Cecilia Aulin, Azar Baharpoor, Camilla I Svensson, “*Pain behaviour assessments by gait and weight bearing in surgically induced osteoarthritis and inflammatory arthritis*”, Physiology & Behavior, Volume 225, 2020.

3D-DeepLabCut



[DeepLabCut/docs/Overviewof3D.md at main · DeepLabCut/DeepLabCut (github.com)](https://github.com/DeepLabCut/DeepLabCut/blob/main/docs/Overviewof3D.md)

AcinoSet: A 3D Pose Estimation Dataset and Baseline Models for Cheetahs in the Wild, ICRA, 2021

如何設計實驗？

- 步伐
- 身體某些部位的疼痛
  % 臉部表情
  % 肢體動作異常
- etc.

老鼠臉部顯出疼痛的表情

如何設計偵測疼痛的實驗？

- 過去已經有的類似技術(人臉部表情分析)
- 老鼠臉部表情分析因鼠臉五官表情變化時變異很小,很難偵測,它未若人臉可細分至七個或更多表情
- 老鼠疼痛實驗只要分痛或不痛即可

人臉表情分析

Automatic Audience Emotion Reader

用AI幫助講者知道聽眾反應

How to make a good speech ?

How to make a good speech?

So the system has to sense the reaction of the audience – from facial emotion !

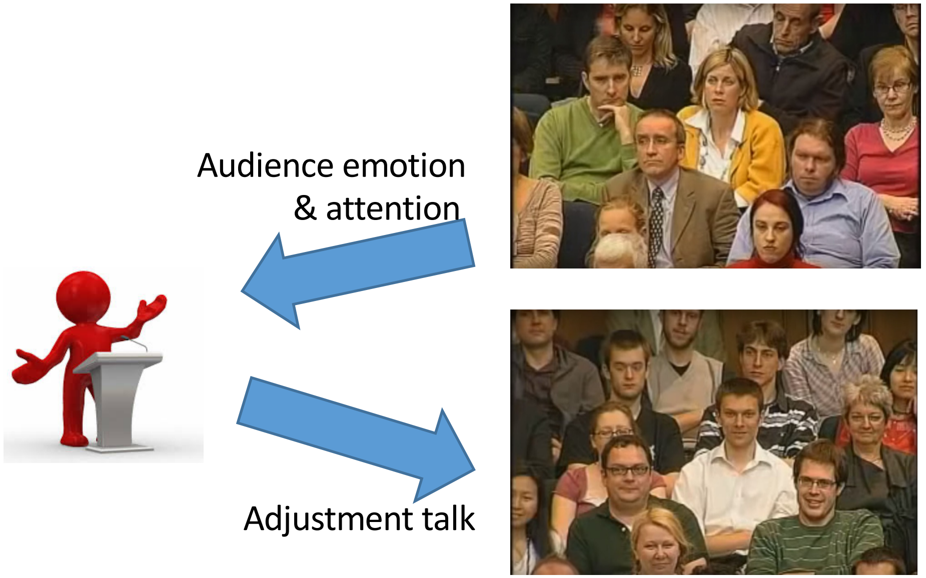
You

Audience emotion
& attention

Your audience

What an ``Automatic Audience Emotion Reader'' should do ?

- ``read'' the emotions from audience
- ``inform timely to the speaker'' to say something to attract the audience



Related works

The Facial Action Coding System (FACS)

- Based on Facial muscle anatomy
- Coding the face muscle to action unit
- Emotional Facial Action Coding System

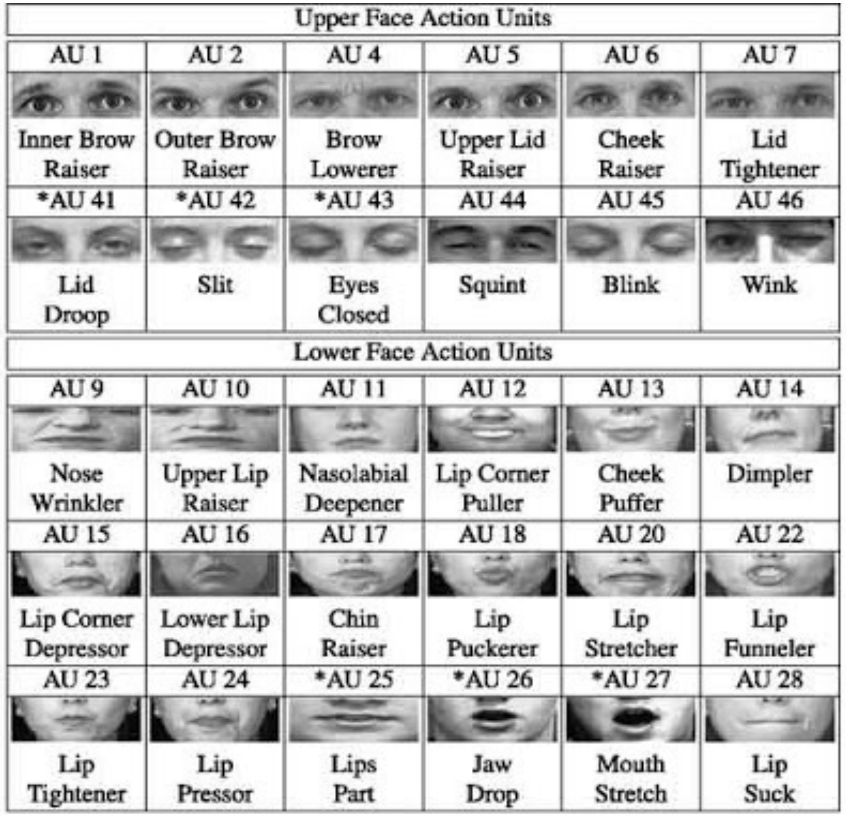
P. Ekman, W.V. Friesen, "Facial Action Coding System: A Technique for the Measurement of Facial Movement, Consulting Psychologists Press", Palo Alto, 1978.

Related works

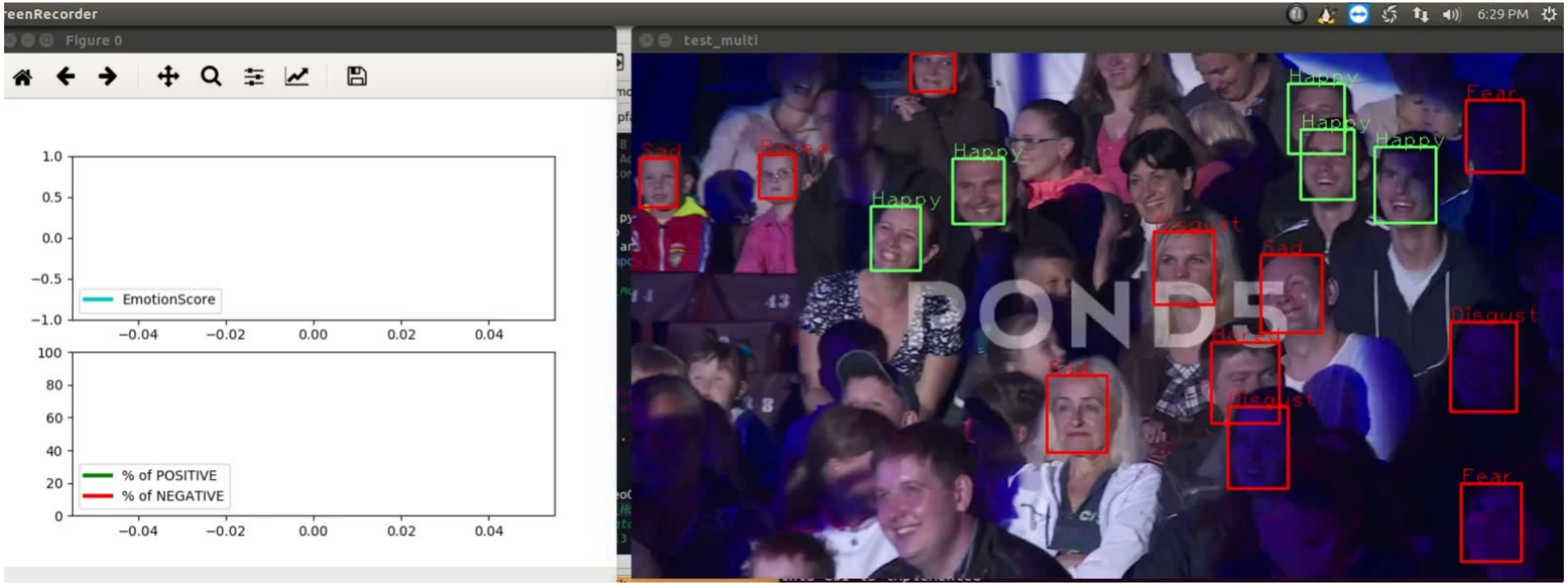
Related works

Related works

- Based on the combination of FACS codes, people can easily compose an ``emotion’’



Emotion	Action Units
Happiness	6+12
Sadness	1+4+15
Surprise	1+2+5B+26
Fear	1+2+4+5+7+20+26
Anger	4+5+7+23
Disgust	9+15+16
Contempt	R12A+R14A


reenRecorder
6:29 PM
Figure 0
test_multi
1.0
0.5
0.0
−0.5
−1.0
EmotionScore
−0.04
−0.02
0.00
0.02
0.04
100
80
60
40
20
0
% of POSITIVE
% of NEGATIVE
Happy
Fear
Sad
Disgust

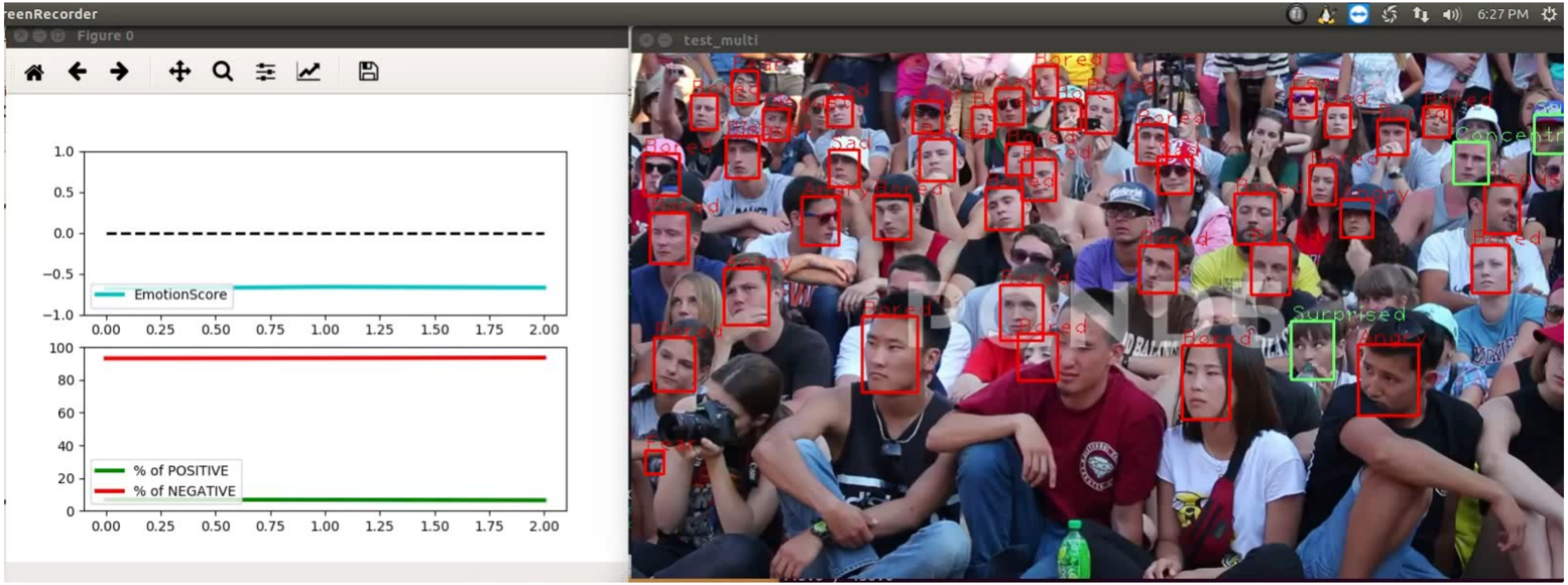


Figure 0
test_multi
6:27 PM
1.0
0.5
0.0
−0.5
−1.0
EmotionScore
0.00 0.25 0.50 0.75 1.00 1.25 1.50 1.75 2.00
100
80
60
40
20
0
% of POSITIVE
% of NEGATIVE
0.00 0.25 0.50 0.75 1.00 1.25 1.50 1.75 2.00
Bored
Sad
Surprised
Angry

針對老鼠表情如何分析？

- 老鼠臉部可用的資訊很少
- 只要分痛苦與否即可
- 人工智慧時代要建老鼠臉部表情字典

是否可以依靠偵測老鼠的表情辨識疼痛?

Images Generate from GPT4o

Animation example

The locations of Tom and Jerry's faces are detected using a trained YOLO model and the facial expression is classified into one of the following 4 : angry, happy, sad or surprised.

https://www.youtube.com/watch?v=qWu9L-J4HCM

Real world hamster emotion examples

(表情偵測的示意圖)

(表情偵測的示意圖)

Angry

https://www.youtube.com/watch?v=HuI6i_IL6Y0

Surprise

https://www.youtube.com/watch?v=XZsFbB3lV1A

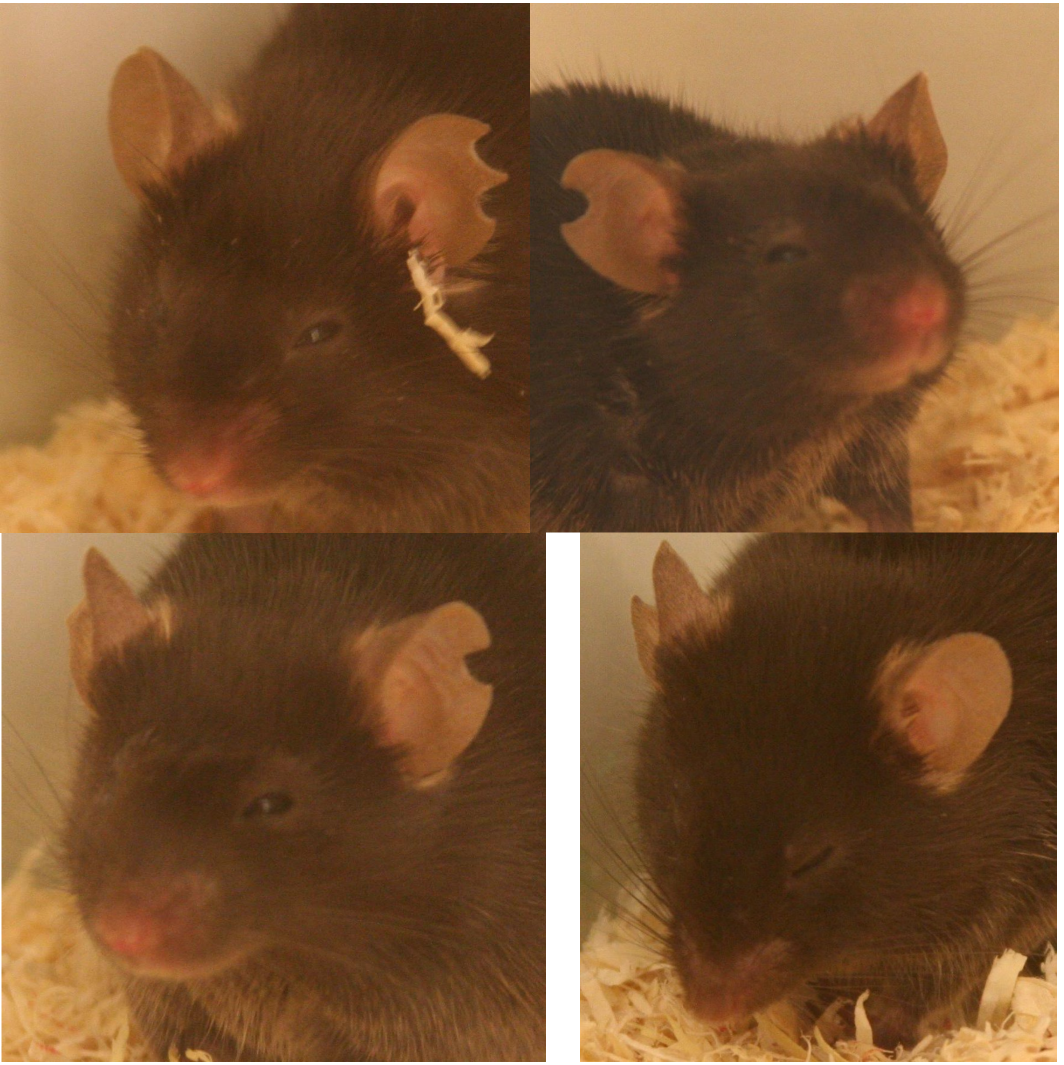
Real world fancy mouse face examples (老鼠臉部偵測的示意圖)

https://www.youtube.com/watch?v=qlZB3ZqaVY0&t=864s

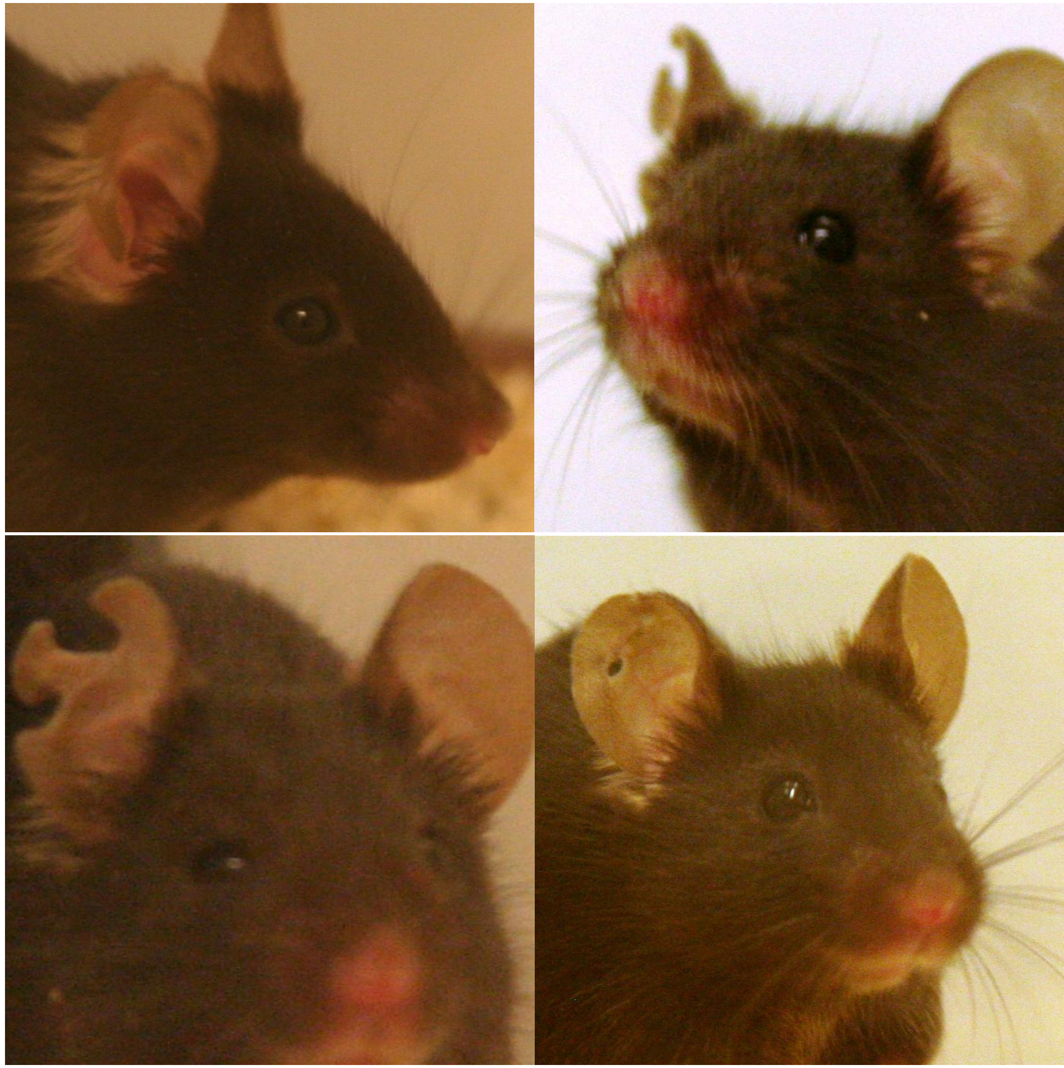
https://depositonce.tu-berlin.de/items/b321032f-f924-4f51-af5a-2baa0b0da63d

Black Mice Dataset

- A large set of images of C57BL/6JRj mice and a corresponding “well-being” label.



有疼痛(被去勢後30分鐘)的老鼠



無疼痛的老鼠(對照組)

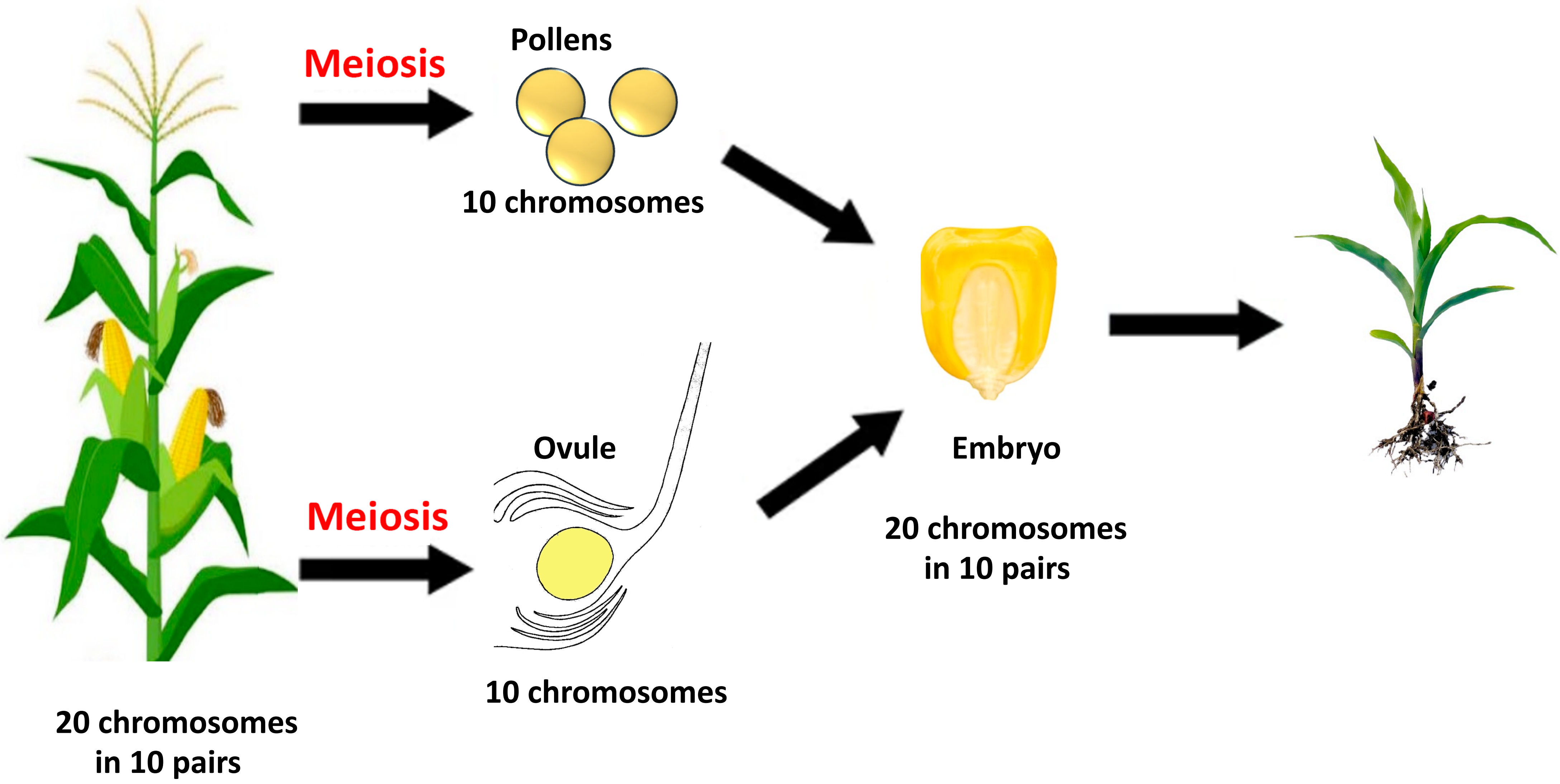
3D

減數分裂時染色體如何重組

植微所王中茹老師

Meiosis produces haploid gametes for sexual reproduction

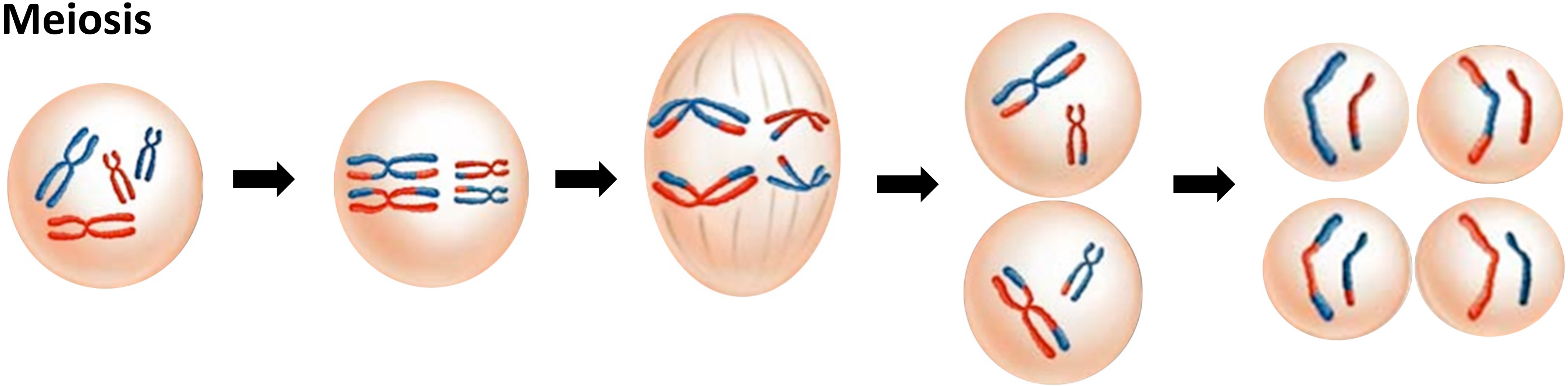
減數分裂:有性生殖的關鍵過程，產生單倍體配子



Homologous chromosomes recombine to shuffle genetic material

同源染色體配對，遺傳物質得以交換、重組

Meiosis

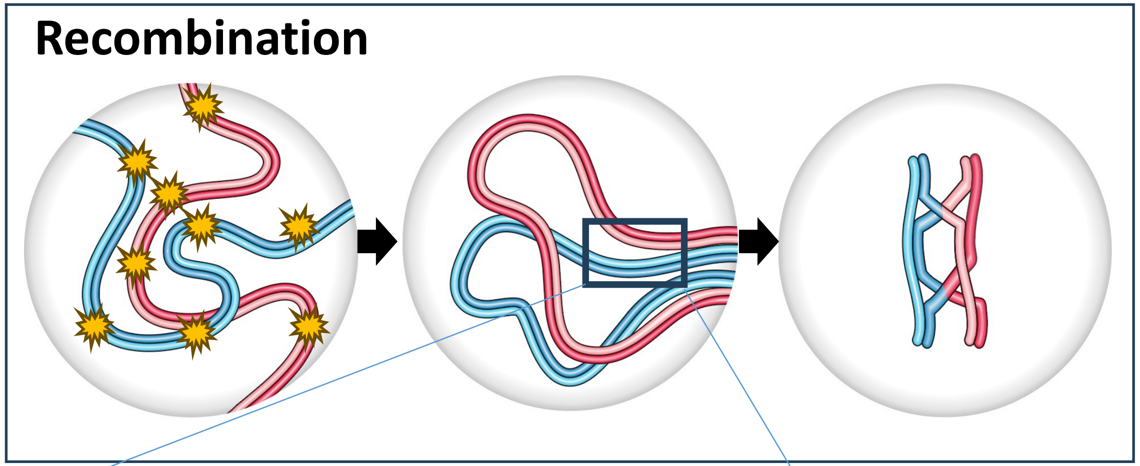


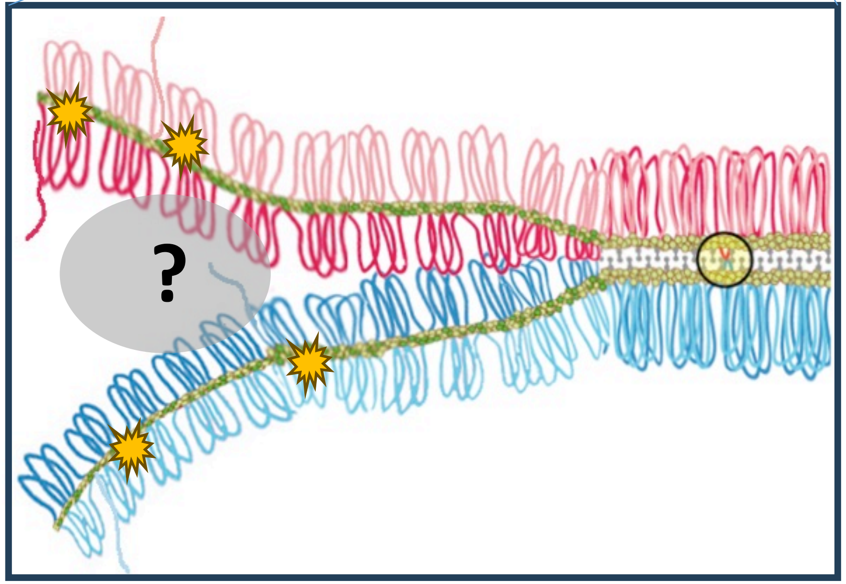
Why study meiosis and recombination?

- Essential for sexual reproduction
- Shuffle genetic material
- Important for agricultural breeding: produce beneficial traits
- Meiotic defects cause miscarriage, birth defects, and infertility

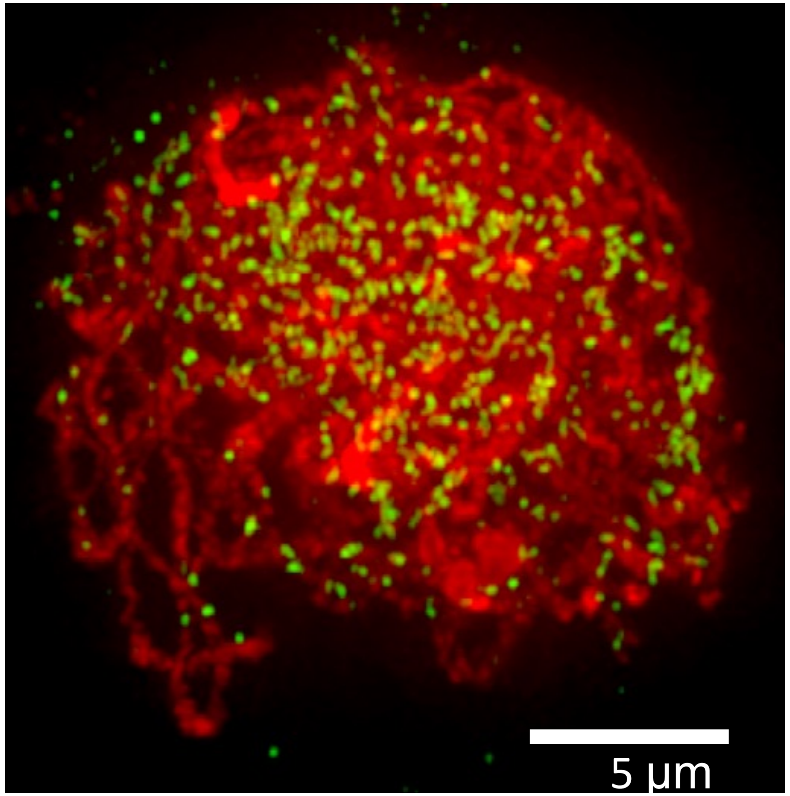
How is homologous recombination processed and achieved?
同源染色體怎麼重組？

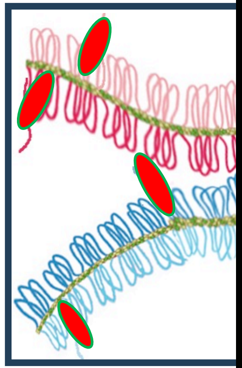
- Hundreds of initial sites
- Below resolution limit
- Within 3D nuclei
- Multiple steps
- Complicated pathways





Recombination events labeled/painted

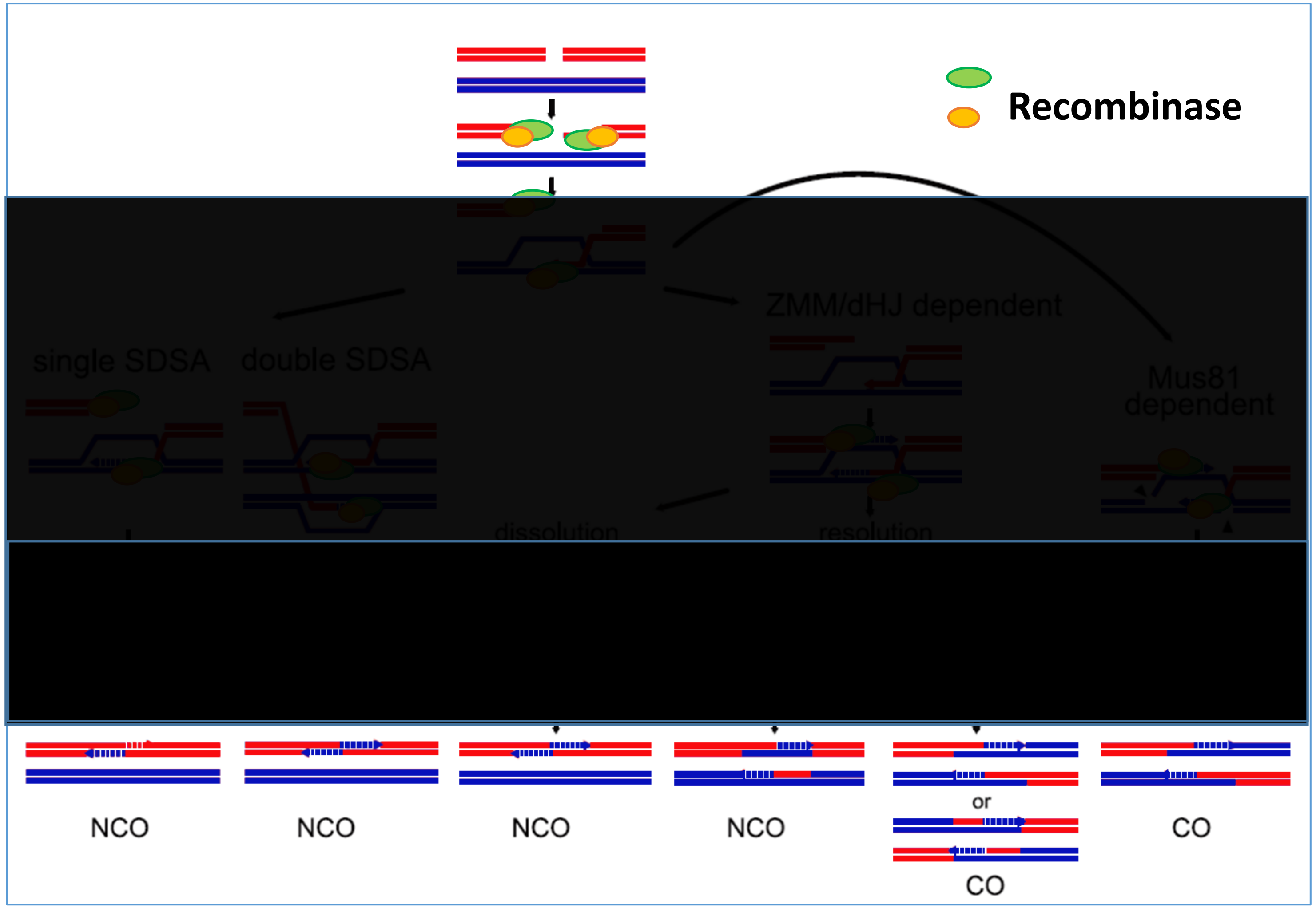


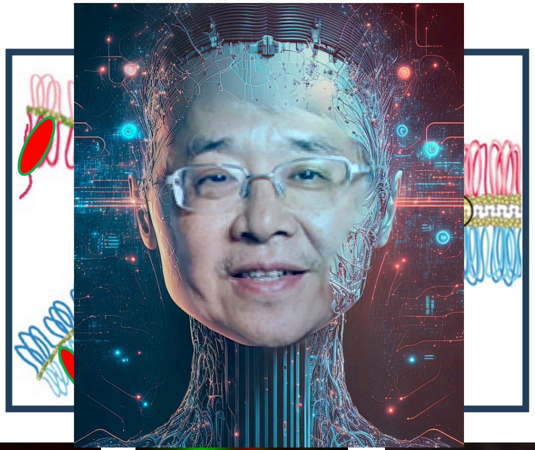


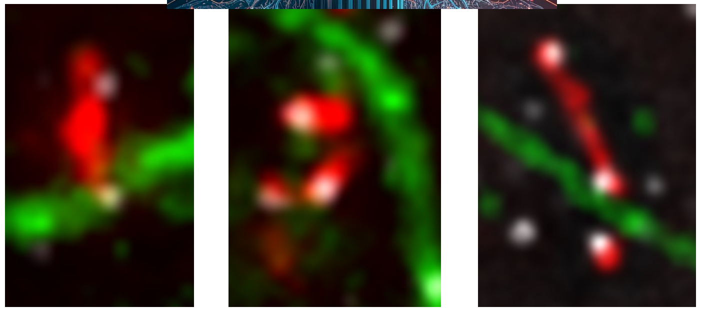

10.00 μm
X
Z

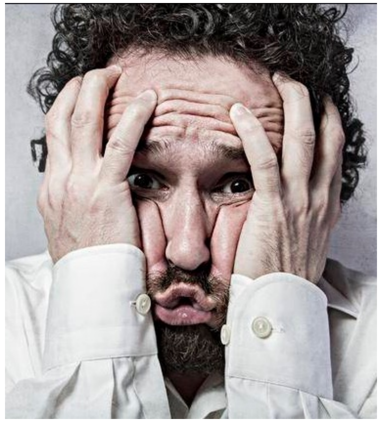
Identifying/validating possible recombination intermediates has been challenging

百年挑戰：重組中間體的辨識與鑑定







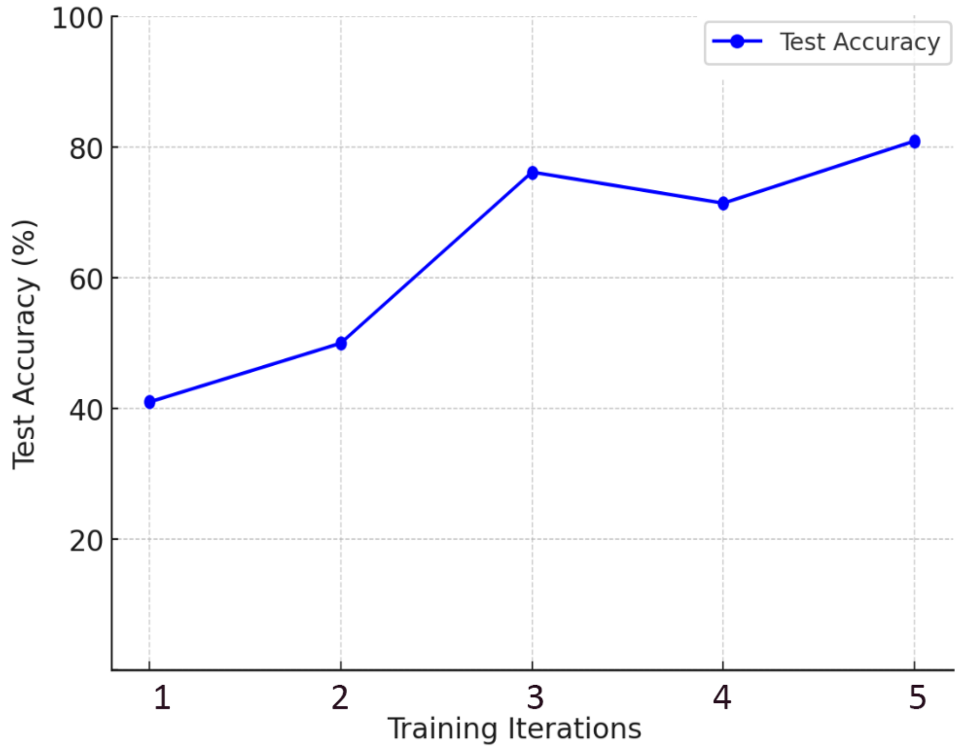


初步監督式學習學辨識3D Recombination intermediates (四種)

正確率達81%

Training	Data	Flip-augmentation	Axis signal	Batch norm	Test accuracy
I	167	x	o	x	41.38
	167	x	x	x	48.28
II	249	o	x	x	50.00
III	202	x	x	o	76.19
IV	202	o	x	o	71.43
V	202	o	o	o	80.95

終極目標:分類distinct intermediates with accuracy > 90.0



合作前進的關鍵：

AI 與生物學者的來回討論溝通 to align terms and share domain knowledge

1.定義問題 (Define Problem)

2.建立資料集 (Build Dataset)

3.訓練模型 (Train Model)

4.得到大量information，建立新的 hypothesis，設計實驗證明

Thanks for listening