Kaggle -人工智慧的最佳試煉場

深度學習與Kaggle實戰

王淳恆, Andrew Wang
沐恩生醫

2019/01/05

Chief Data Scientist (需配合出差美國日本等辦公室, 年薪600~1200萬)

本公司其他工作

擅長工具：Python

工作技能：不拘

其他條件：[待遇]

年薪：600萬~1200萬（含配股獎金）

[職務需求]

年資 (參考二者年資)

1. 5年以上資料分析專案或專業顧問經驗
2. 2年以上資料科學相關 PhD、PostDoc、教授等研究經驗

需具備條件 (至少其中四項)：

1. 熟悉機器學習/深度學習演算法及統計方法
2. 具備商業分析專案經驗
3. 具備工業設備分析專案經驗
4. Kaggle 分析競賽排名至少前 5%
5. 資料分析相關學術論文發表 Best Paper Award
6. 英文語言能力佳 (多益860分以上) 或母語者

Chief Data Scientist (需配合出差美國日本等辦公室, 年薪600~1200萬)

資料科學家

擅長工具：Python

工作技能：不拘

其他條件：[待遇]

年薪：600萬~1200萬（含配股獎金）

年薪:600萬~1200萬

年資 (參考二者年資)

1. 5年以上資料分析專案或專業顧問經驗
2. 2年以上資料科學相關 PhD、PostDoc、教授等研究經驗

需具備條件 (至少其中四項)：

1. 熟悉機器學習/深度學習演算法及統計方法
2. 具備商業分析專案經驗
3. 具備工業設備分析專案經驗
4. Kaggle 分析競賽排名至少前 5%
5. 資料分析相關學術論文發表 Best Paper Award
6. 英文語言能力佳 (多益860分以上) 或母語者

Kaggle分析競賽排名至少前5%

What is **Kaggle**?

Kaggle

From Wikipedia, the free encyclopedia

Kaggle is a platform for predictive modelling and analytics competitions in which companies and researchers post data and statisticians and data miners compete to produce the best models for predicting and describing the data. The crowdsourcing approach relies on the fact that there are countless strategies that can be applied to any predictive modelling task and it is impossible to know at the outset which technique or analyst will be most effective.

On 8 March 2017, Google announced that they were acquiring Kaggle.[1] They will join the Google Cloud team and continue to be a distinct brand.[2]

Kaggle是目前最大的人工智慧與資料科學競賽平台

2017 Google 併購

BOSCH

Bosch Production Line Performance

BOSCH改善生產線效能

Overview | Data | Kernels | Discussion | Leaderboard | Rules | Late Submission

Overview

- Description
- Evaluation
- Prizes
- Timeline
- Ieee Bigdata 2016

A good chocolate soufflé is decadent, delicious, and delicate. But, it's a challenge to prepare. When you pull a disappointingly deflated dessert out of the oven, you instinctively retrace your steps to identify at what point you went wrong. Bosch, one of the world's leading manufacturing companies, has an imperative to ensure that the recipes for the production of its advanced mechanical components are of the highest quality and safety standards. Part of doing so is closely monitoring its parts as they progress through the manufacturing processes.

Mercedes-Benz Greener Manufacturing

Can you cut the time a Mercedes-Benz spends on the test bench?

$25,000
Prize Money

Daimler

賓士加速測試流程

To ensure the safety and reliability of each and every unique car configuration before they hit the road, Daimler's engineers have developed a robust testing system. But, optimizing the speed of their testing system for so many possible feature combinations is complex and time-consuming without a powerful algorithmic approach. As one of the world's biggest manufacturers of premium cars, safety and efficiency are paramount on Daimler's production lines.

PAKDD 2014 - ASUS Malfunctional Components Prediction

華碩筆電故障元件預估

Overview | Data | Kernels | Discussion | Leaderboard | Rules | Late Submission

Overview

- Description
- Evaluation
- Prizes
- Timeline
- About The Sponsor Asus
- Winners

The goal of PAKDD 2014 competition is to predict future malfunctional components of ASUS notebooks from historical data. This will help estimate how many products will require maintenance or repair services.

ASUS has provided information on its laptop shipments as well as the laptops requiring maintenance or repair services. Participants will use this information to estimate how many of each module of a specific model will require maintenance or repair services.

Acknowledgements

The organizers of PAKDD would like to thank ASUS for sponsorship of this competition.

Research Prediction Competition

WSDM - KKBox's Music Recommendation Challenge

Can you build the best music recommendation system?

$5,000
Prize Money

KKBOX ·

KKBox音樂推薦

Overview | Data | Kernels | Discussion | Leaderboard | Rules | Late Submission

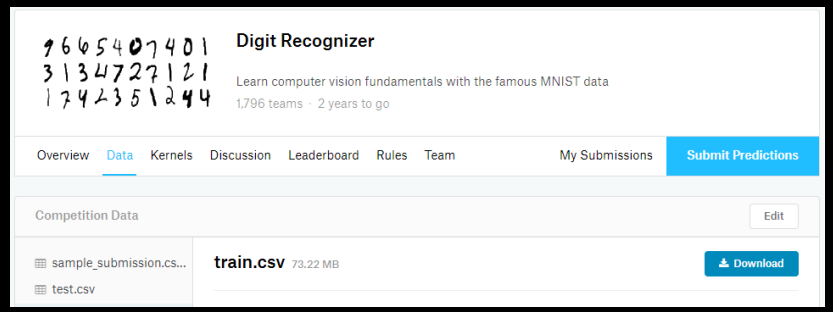
Overview

- Description
- Evaluation
- Prizes
- Timeline

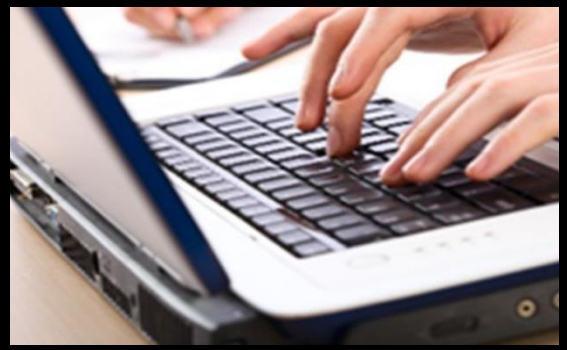
The 11th ACM International Conference on Web Search and Data Mining (WSDM 2018) is challenging you to build a better music recommendation system using a donated dataset from KKBOX. WSDM (pronounced "wisdom") is one of the the premier conferences on web inspired research involving search and data mining. They're committed to publishing original, high quality papers and presentations, with an emphasis on practical but principled novel models.

How to Kaggle?

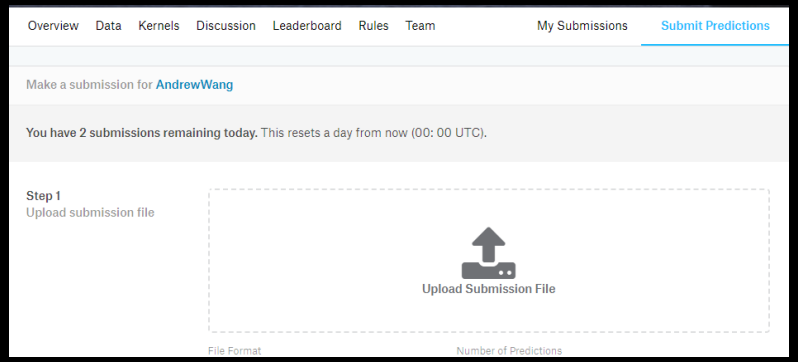
Download Data



Train & Predict



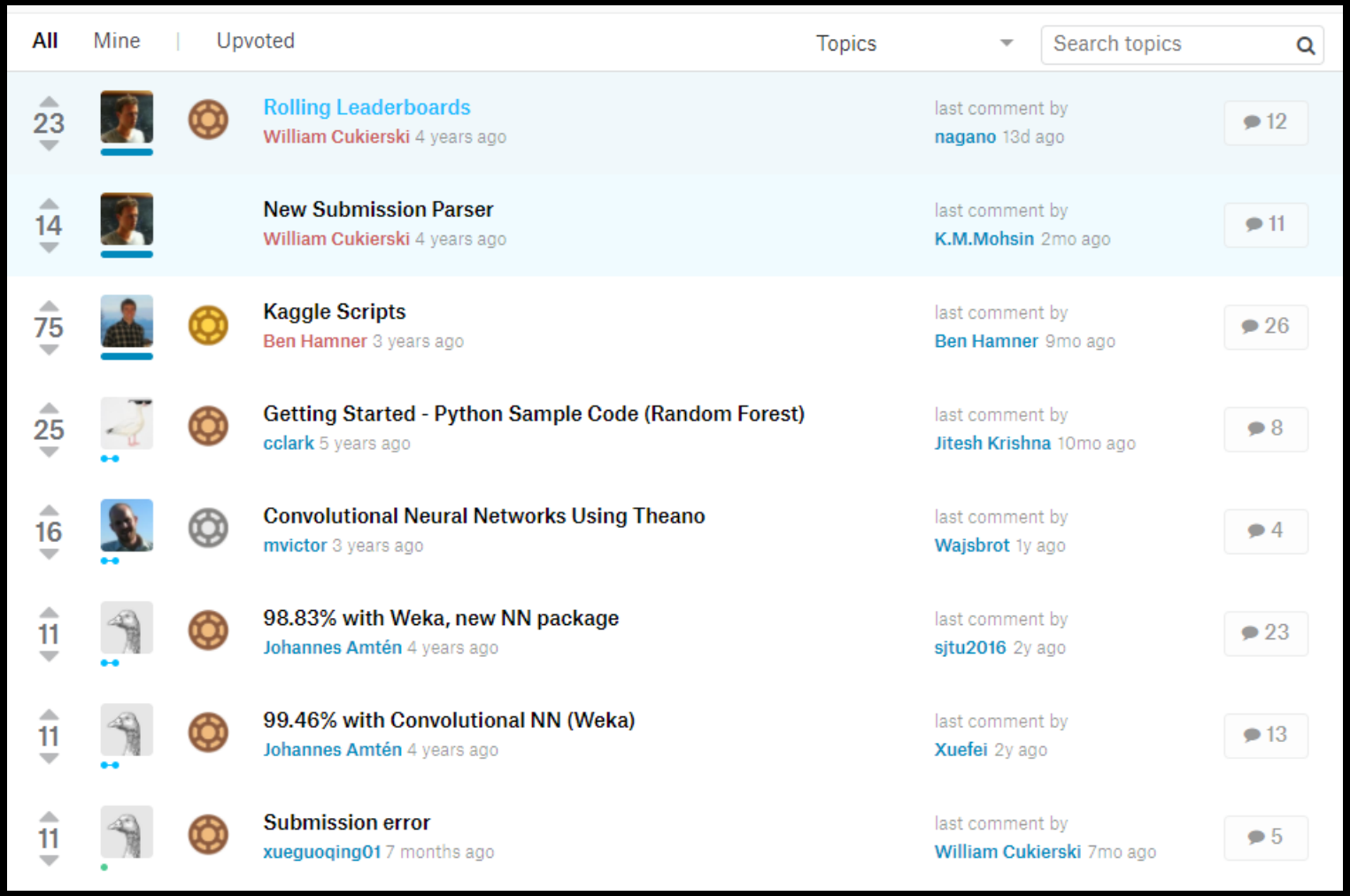
Submit



Leaderboard in Kaggle

#	△1w	Team Name	Kernel	Team Members	Score	Entries	Last
1	—	Srisailam Varakala			1.00000	2	2d
2	—	Deep-DIT			1.00000	2	1mo
3	—	Amorth			1.00000	1	1mo
4	—	linbo_iacas			1.00000	1	1mo
5	—	Alpha Zero			1.00000	2	14d
6	—	bestLoveForMxy			1.00000	9	8d
7	new	Georgi Pamukov			1.00000	1	5d
8	▼1	Aaron Sun			0.99985	2	2mo
9	▼1	Lorenzo Ridolfi			0.99957	3	1mo
10	▼1	Honey Comb			0.99957	3	1mo
11	▼1	Adarsh Verma			0.99957	6	1mo
12	▲50	dncc_3			0.99957	10	4d
13	▼2	DerekGrant			0.99942	5	2mo

Discussion in Kaggle



深度學習與Kaggle實戰

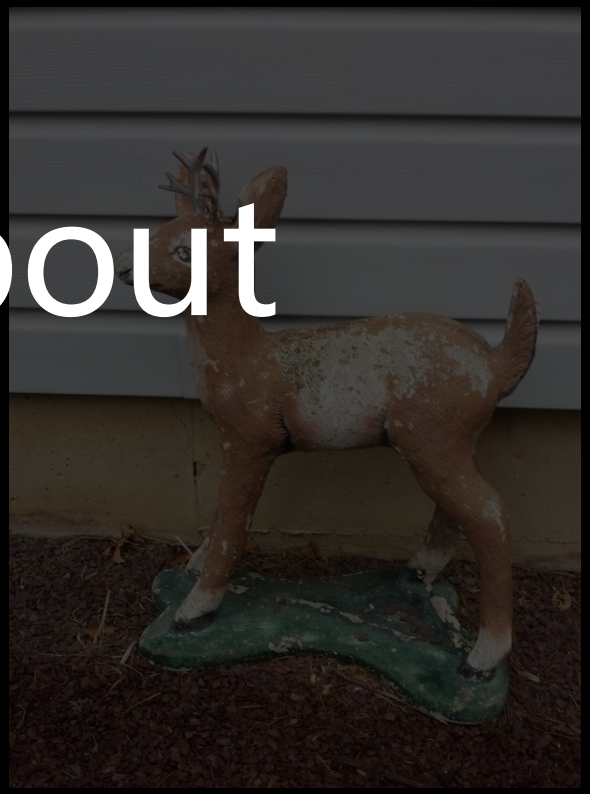
- Camera Model Identification
  - 相機型號識別競賽 銀牌經驗分享
- Passenger Screening Algorithm Challenge
  - 機場乘客檢查競賽 銀牌經驗分享
- RSNA Pneumonia Detection Challenge
  - 肺炎檢測競賽 銀牌經驗分享
- Practical Application in Kaggle

Outline

- Camera Model Identification
  - 相機型號識別競賽 銀牌經驗分享
- Passenger Screening Algorithm Challenge
  - 機場乘客檢查競賽 銀牌經驗分享
- RSNA Pneumonia Detection Challenge
  - 肺炎檢測挑戰 銀牌經驗分享
- Practical Application in Kaggle

Which picture is took by iPhone 6, HTC One or Samsung Galaxy S4?

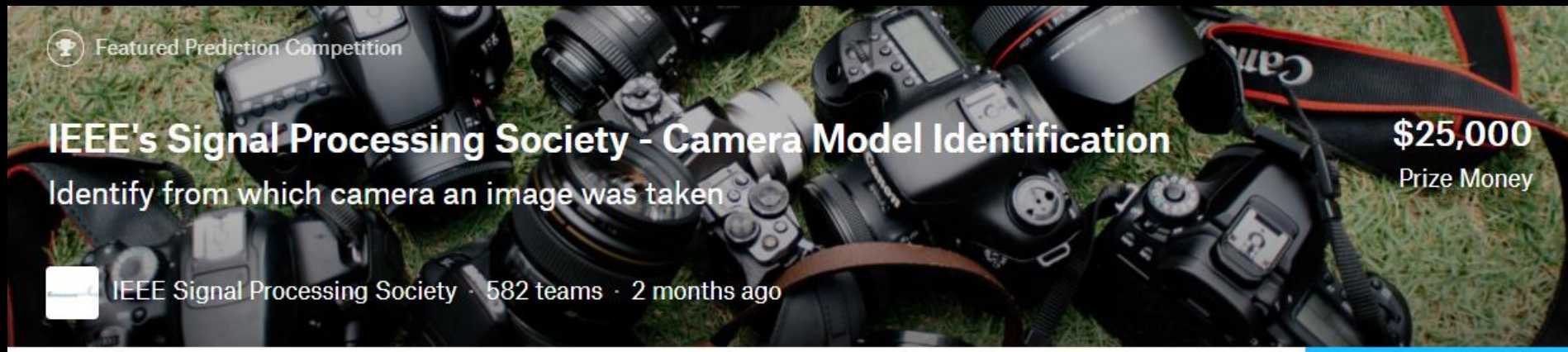
Which picture is took by iPhone 6, HTC One or Samsung Galaxy S4?



Accuracy about

97%

Camera Model Identification
- IEEE's Signal Processing Society



Overview Data Kernels Discussion Leaderboard Rules Team My Submissions **Late Submission**

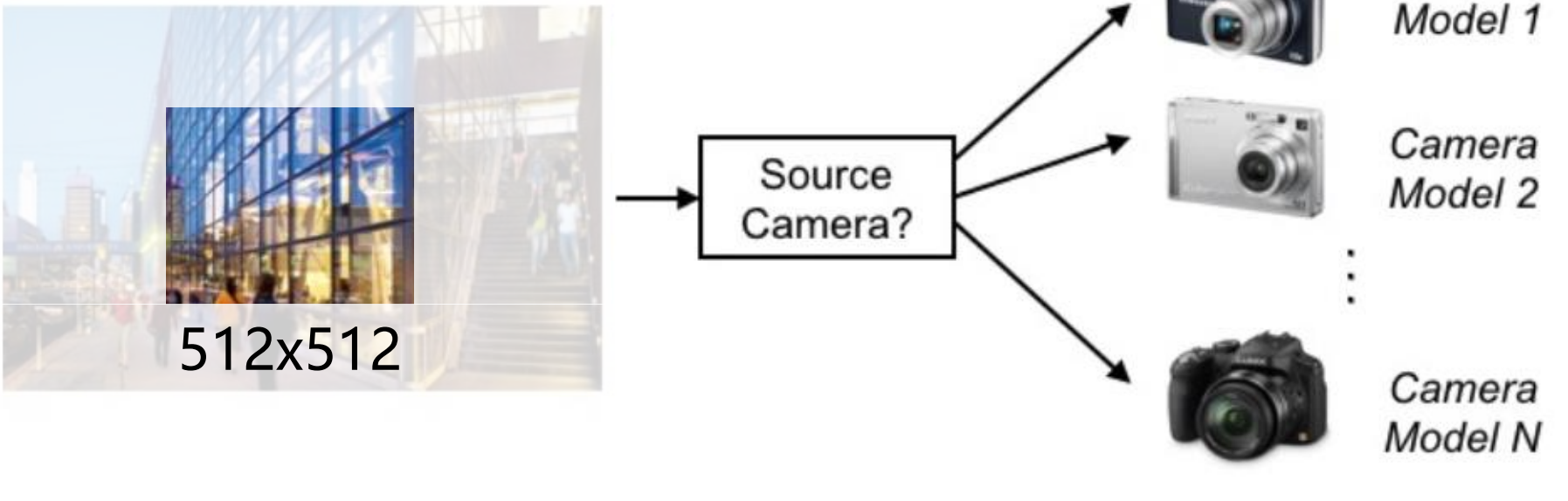
Overview

Description

Evaluation

Finding footage of a crime caught on tape is an investigator's dream. But even with crystal clear, damning evidence, one critical question always remains–is the footage real?

Camera Model Identification



Camera Model Identification

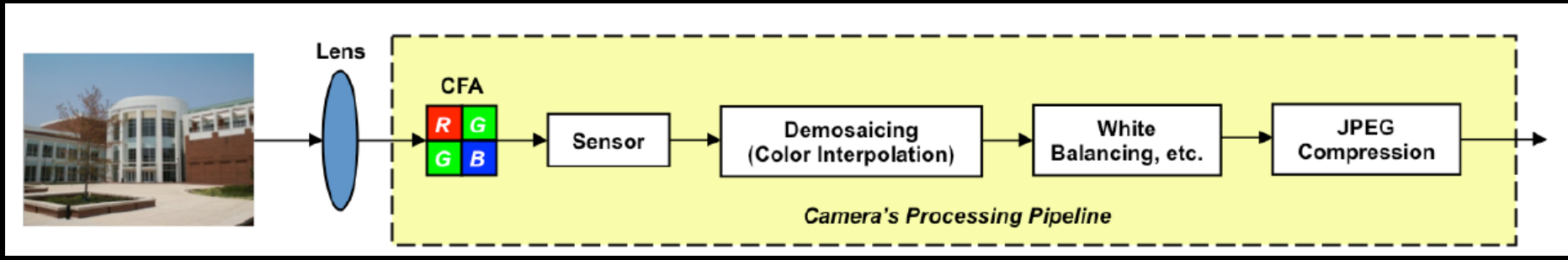
- 10 camera models,
  - Sony NEX-7
  - Motorola Moto X / Nexus 6 / DROID MAXX
  - LG Nexus 5x
  - Apple iPhone 6 /Apple iPhone 4s
  - HTC One M7
  - Samsung Galaxy S4 / Galaxy Note 3

Camera Model Identification

- Manipulated
  - JPEG compression
    - quality factor = 70
    - quality factor = 90
  - resizing (via bicubic interpolation)
    - factor of 0.5
    - factor of 0.8
    - factor of 1.5
    - factor of 2.0
  - gamma correction
    - gamma = 0.8
    - gamma = 1.2

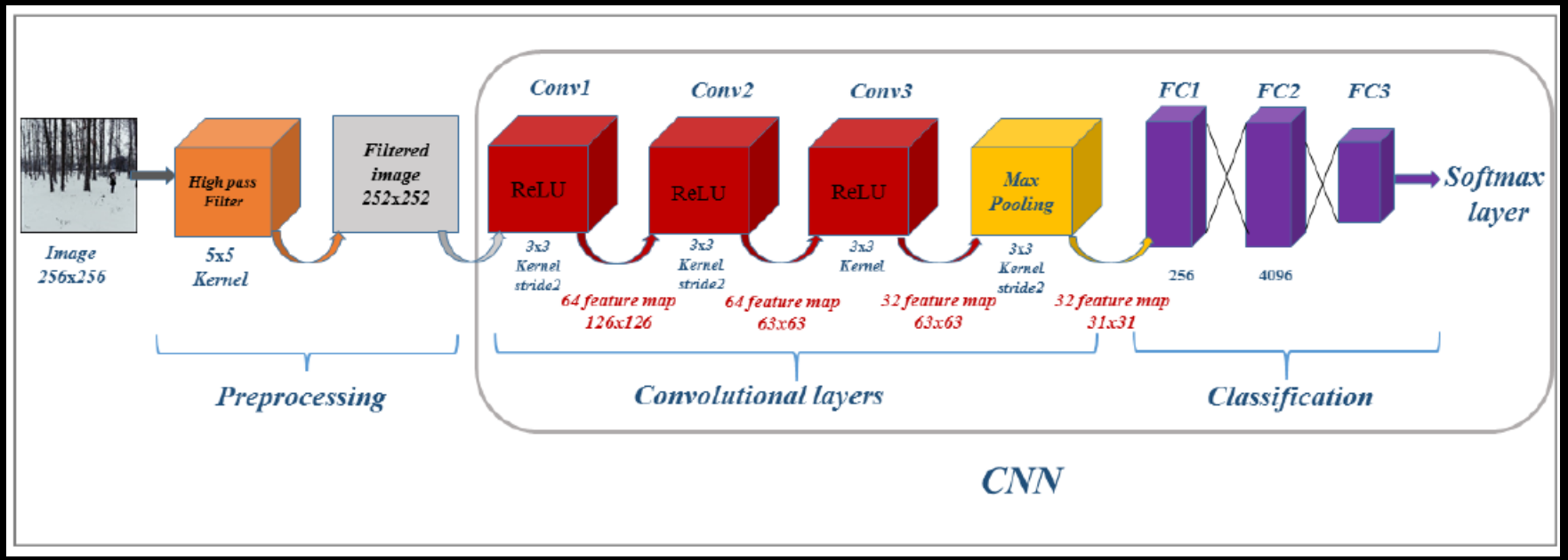
Camera Model Identification

- The different Camera Model has different processing pipeline
- The different processing pipeline has the different noise



Camera Model Identification

- Paper: Use High-Pass filter to extract noise pattern, then use CNN to identify

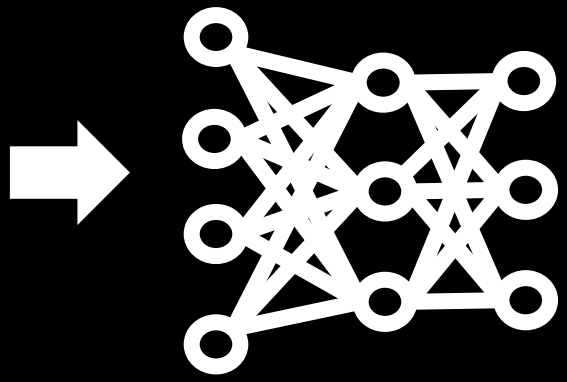


ref: Camera Model Identification With The Use of Deep Convolutional Neural Networks

Camera Model Identification

- High pass filter could be learned by CNN? Yes

InceptionResNetV2



iPhone 6

ref: Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning

ImageNet

ImageNet Dataset

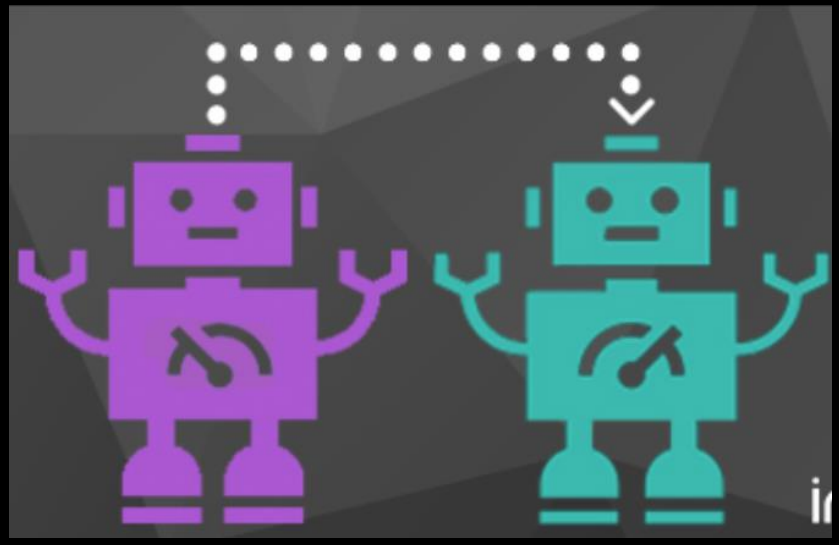
120萬張圖片
1000類

IMAGENET

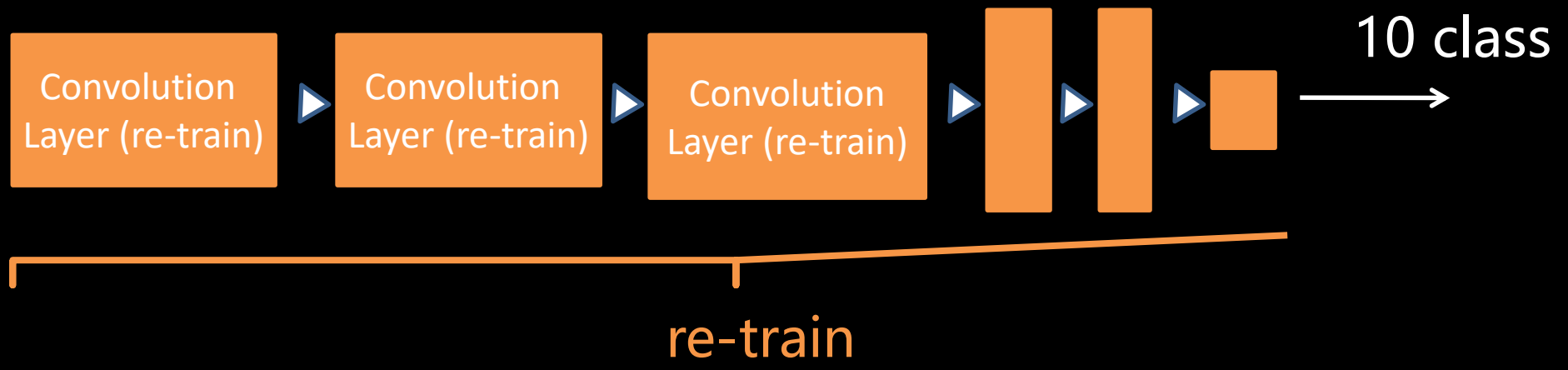
Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Fei-Fei, L. (2015). Imagenet large scale visual recognition challenge. *arXiv preprint arXiv:1409.0575*. [web]

Transfer Learning

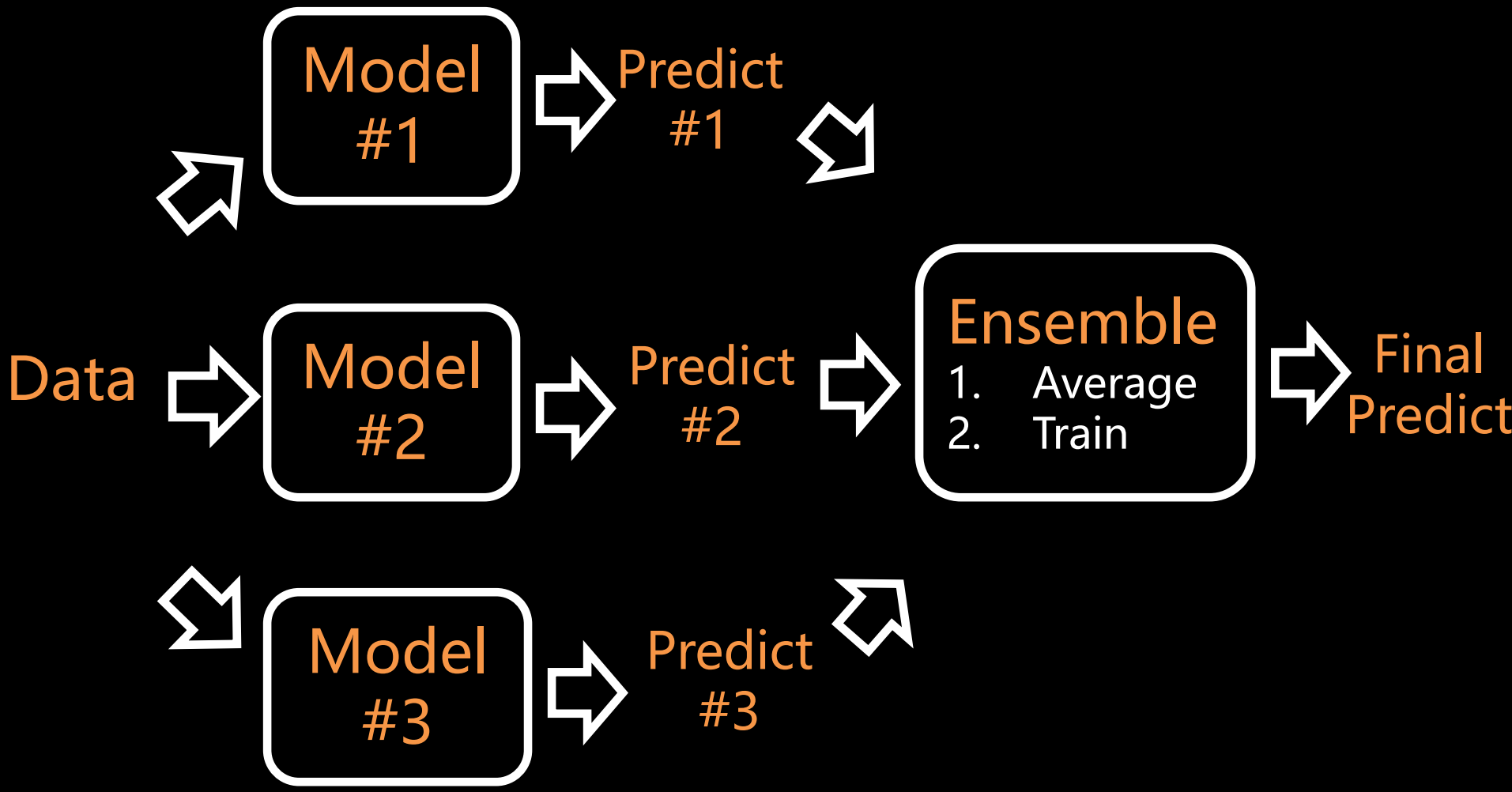
- Transfer learning involves taking a pre-trained neural network and adapting the neural network to a new, different data set.



Transfer Learning



Ensemble


Model #1
Predict #1
Data
Model #2
Predict #2
Ensemble
1. Average
2. Train
Final Predict
Model #3
Predict #3

Camera Model Identification
- IEEE's Signal Processing Society



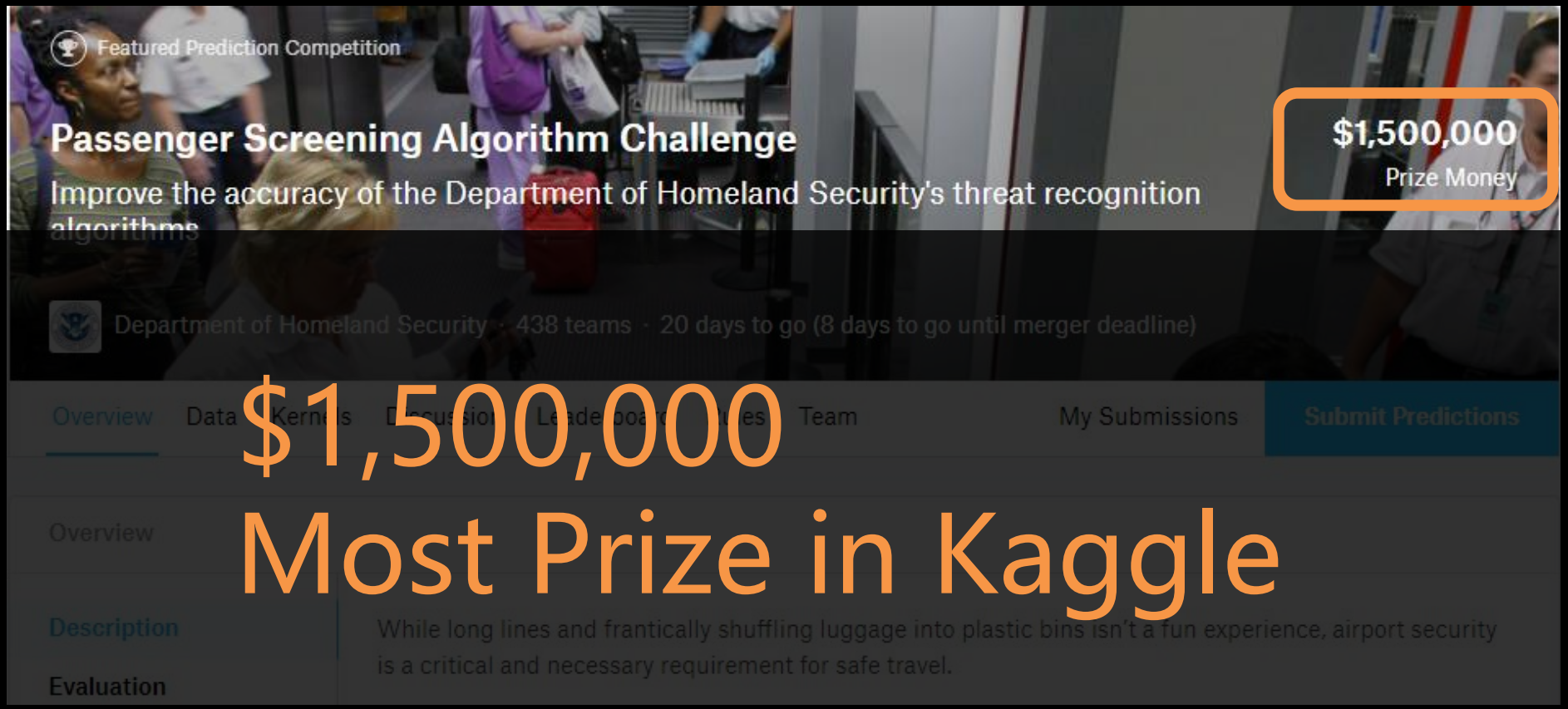
Win the Silver Medal
(36/582)

with outrunner and yutengtung

Outline

- Camera Model Identification
  - 相機型號識別競賽 銀牌經驗分享
- Passenger Screening Algorithm Challenge
  - 機場乘客檢查競賽 銀牌經驗分享
- RSNA Pneumonia Detection Challenge
  - 肺炎檢測挑戰 銀牌經驗分享
- Medical Application in Kaggle

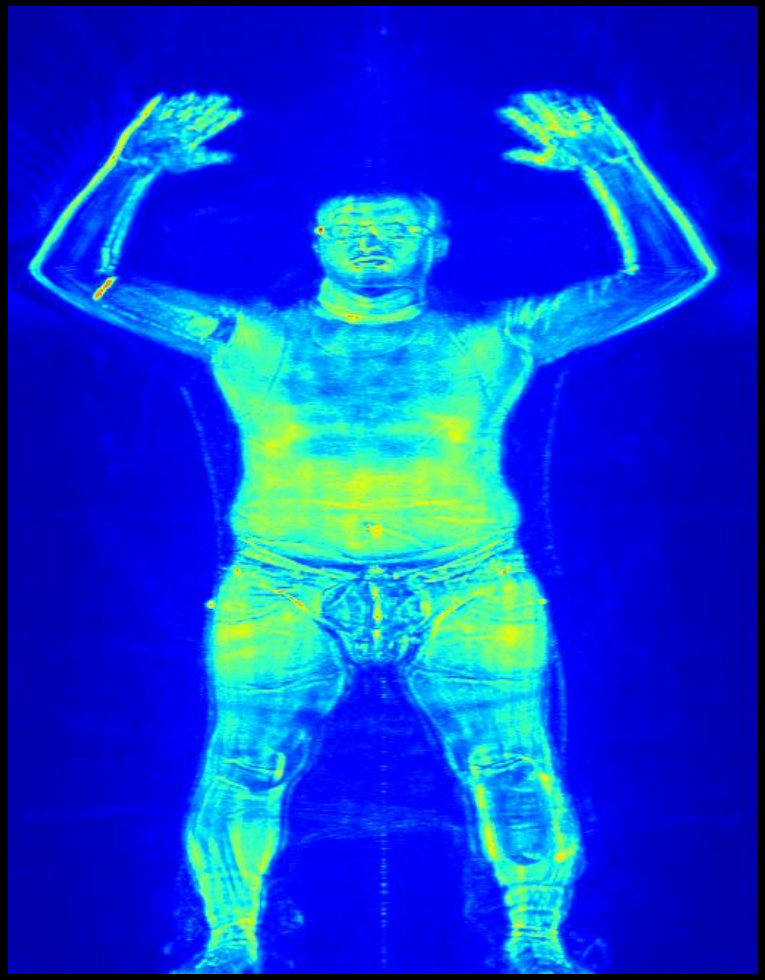
Passenger Screening Algorithm Challenge



$1,500,000
Most Prize in Kaggle

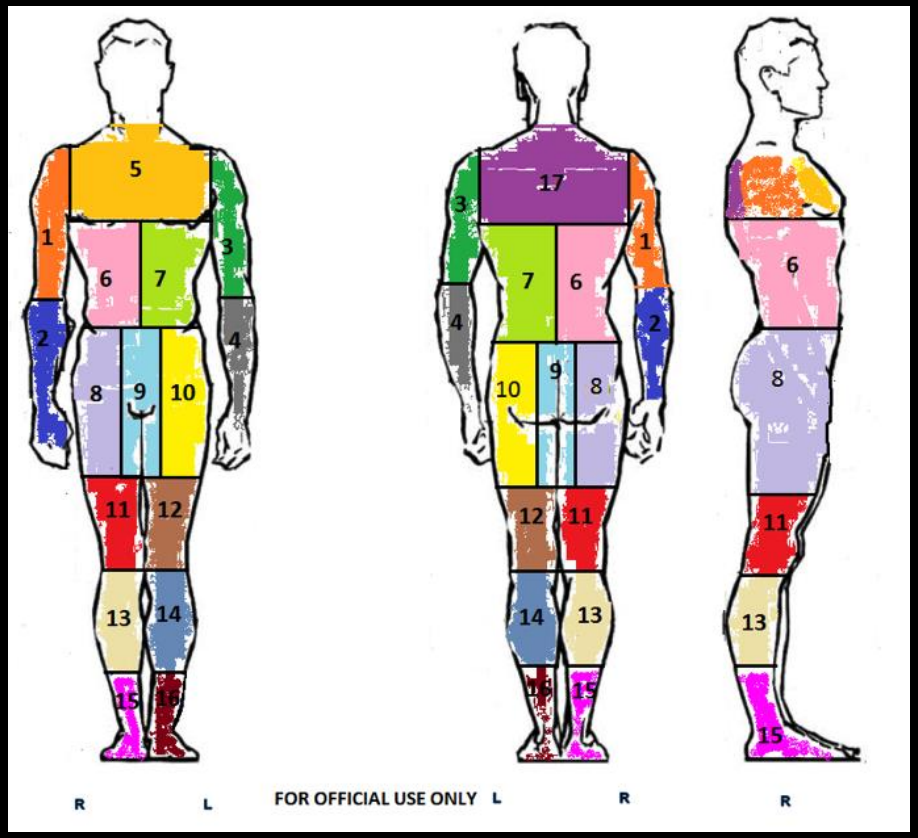
Passenger Screening Algorithm Challenge

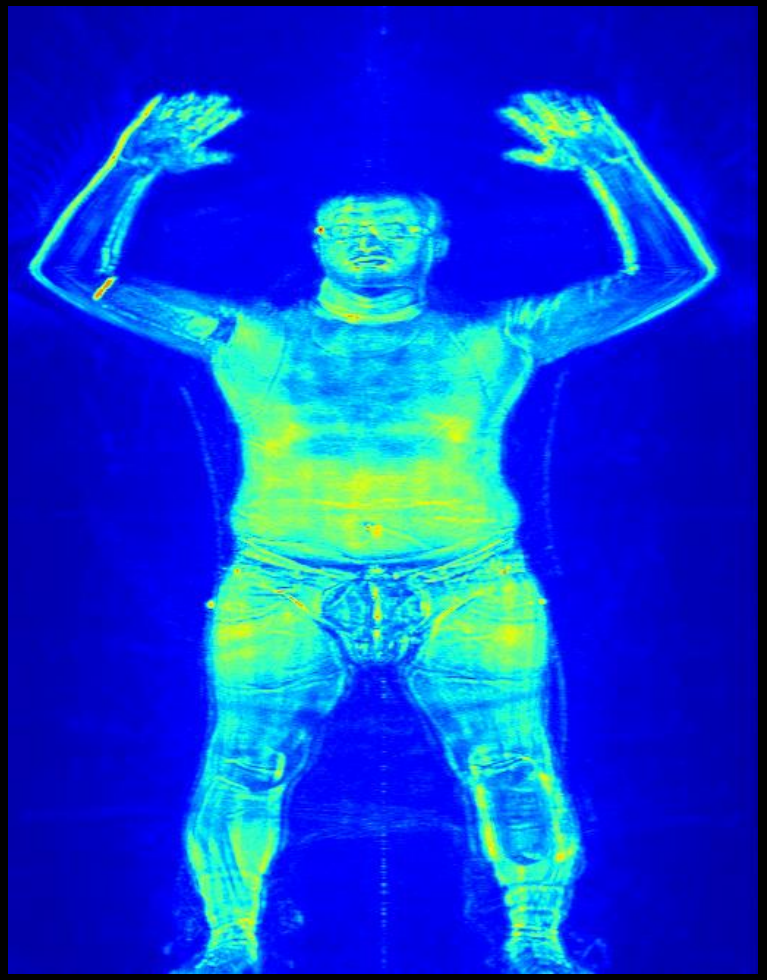
- Improve the accuracy of the Department of Homeland Security's threat recognition algorithms



Passenger Screening
Algorithm Challenge

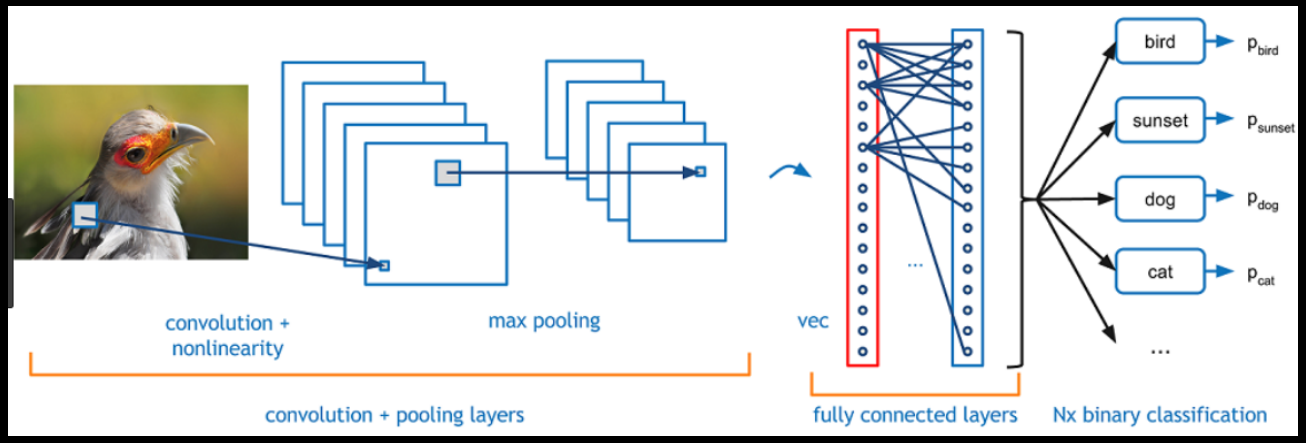
- 17 Class





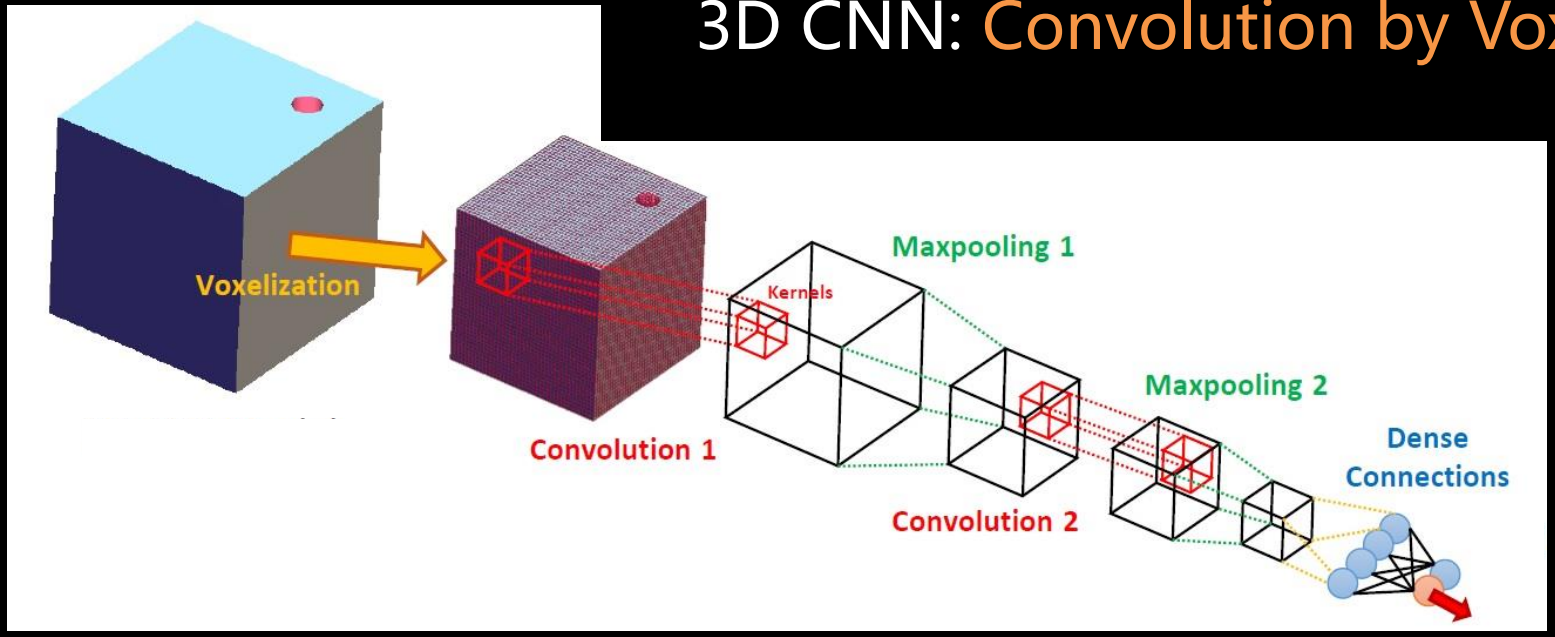
Win the Silver Medal
Rank 35 (/518, Top 7%)

#	△pub	Team Name	Kernel	Team Members	Score	Entries	Last
1	▲138	idle_speculation			0.02417	2	3d
2	▲73	serg14			0.02659	2	6d
3	▲71	David O. | Thomas A.			0.03042	4	4d
33	▲69	Alon Daks			0.15457	1	9d
34	▲52	alanpo			0.15786	2	8d
35	▲11	AndrewWang			0.15917	4	8d
36	▲47	Roland Luethy			0.15945	2	10d
37	▲63	ploider			0.16394	2	9d
38	▲107	Bastiaan Bergman			0.16403	2	8d



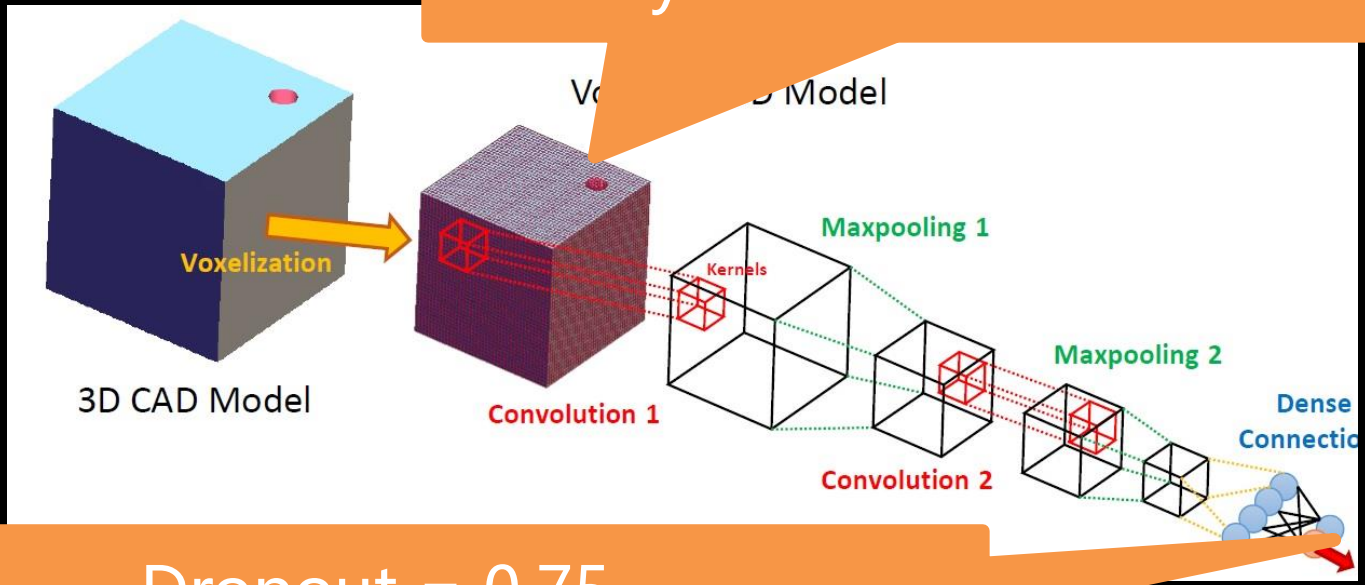
Modify 2D CNN to 3D CNN

2D CNN: Convolution by Pixel
3D CNN: Convolution by Voxel

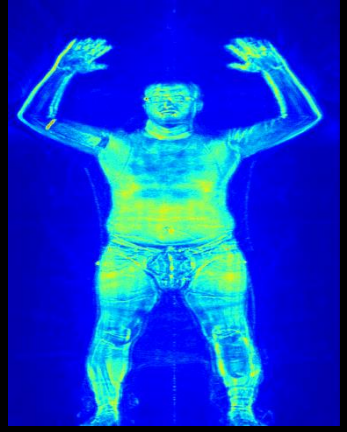


3D CNN

16 feature in first layer and x2 to next layer



Dropout = 0.75
Full connection 2048 and 512 feature



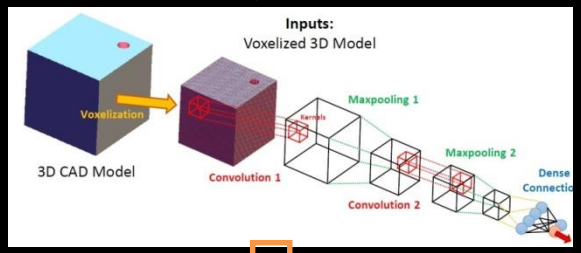
512x512x660

128x128x165

Stride 4

3D Augmentation

3D shift ± 5 pixel
3D rotate ± 5 degree



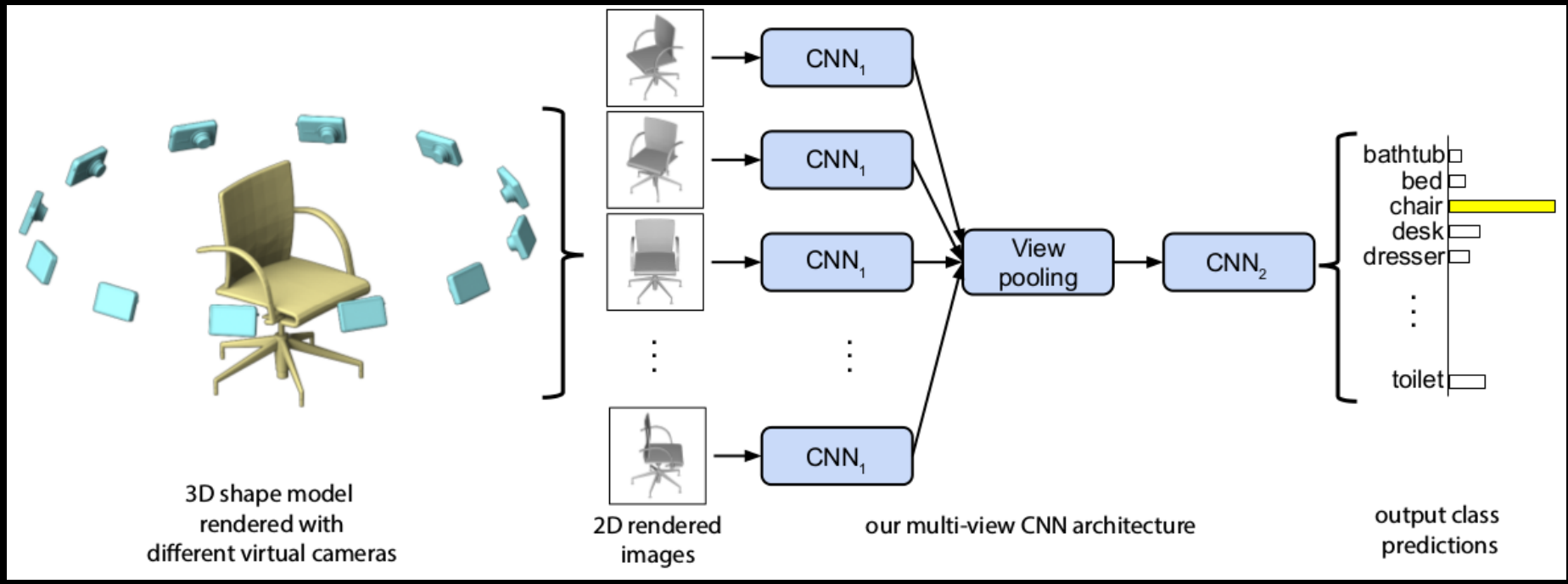
3D CNN

Multiple Labels
(17 Labels)

3D CNN is Hard to Train

- Only 1147 data size, easy overfitting
- Can't use transfer learning
- Train 400 epochs in about 50 hours
- 3D Augmentation need powerful CPU
- More feature number could better, but limited by GPU memory
  (16 CPU Cores with 2x 1080 Ti GPUs)

Multi-View CNN

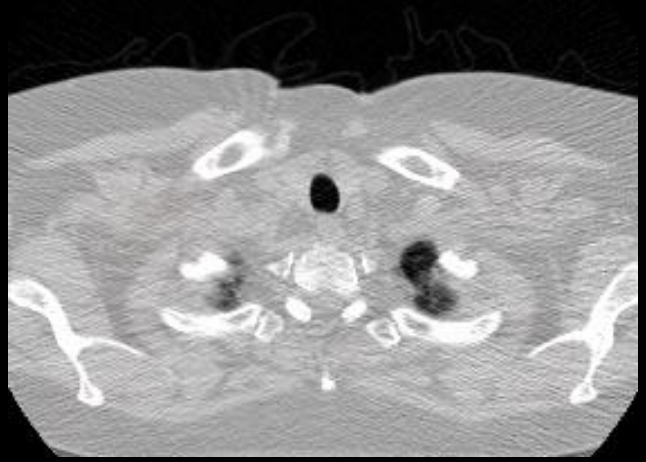


Could use transfer learning for small data set size in Multi-View CNN

Ref: http://vis-www.cs.umass.edu/mvcnn/

The Application of 3D CNN in Medical AI

- Lung Cancer



More Detail

- My Medium : https://goo.gl/skJReo

Outline

- Camera Model Identification
  - 相機型號識別競賽 銀牌經驗分享
- Passenger Screening Algorithm Challenge
  - 機場乘客檢查競賽 銀牌經驗分享
- RSNA Pneumonia Detection Challenge
  - 肺炎檢測挑戰 銀牌經驗分享
- Practical Application in Kaggle

RSNA Pneumonia Detection Challenge

Featured Prediction Competition

RSNA Pneumonia Detection Challenge

Can you build an algorithm that automatically detects potential pneumonia cases?

$30,000
Prize Money

Radiological Society of North America · 380 teams · 4 hours ago

Overview | Data | Kernels | Discussion | Leaderboard | Rules | Team | My Submissions | Late Submission

Overview

Description

Evaluation

In this competition, you're challenged to build an algorithm to detect a visual signal for pneumonia in medical images. Specifically, your algorithm needs to automatically locate lung opacities on chest radiographs.

Featured Prediction Competition

RSNA Pneumon

Can you build an al

$30,000
Prize Money

RSNA Radiological Societ

Overview Data Kern

Late Submission

Overview

Description

Evaluation

Timeline

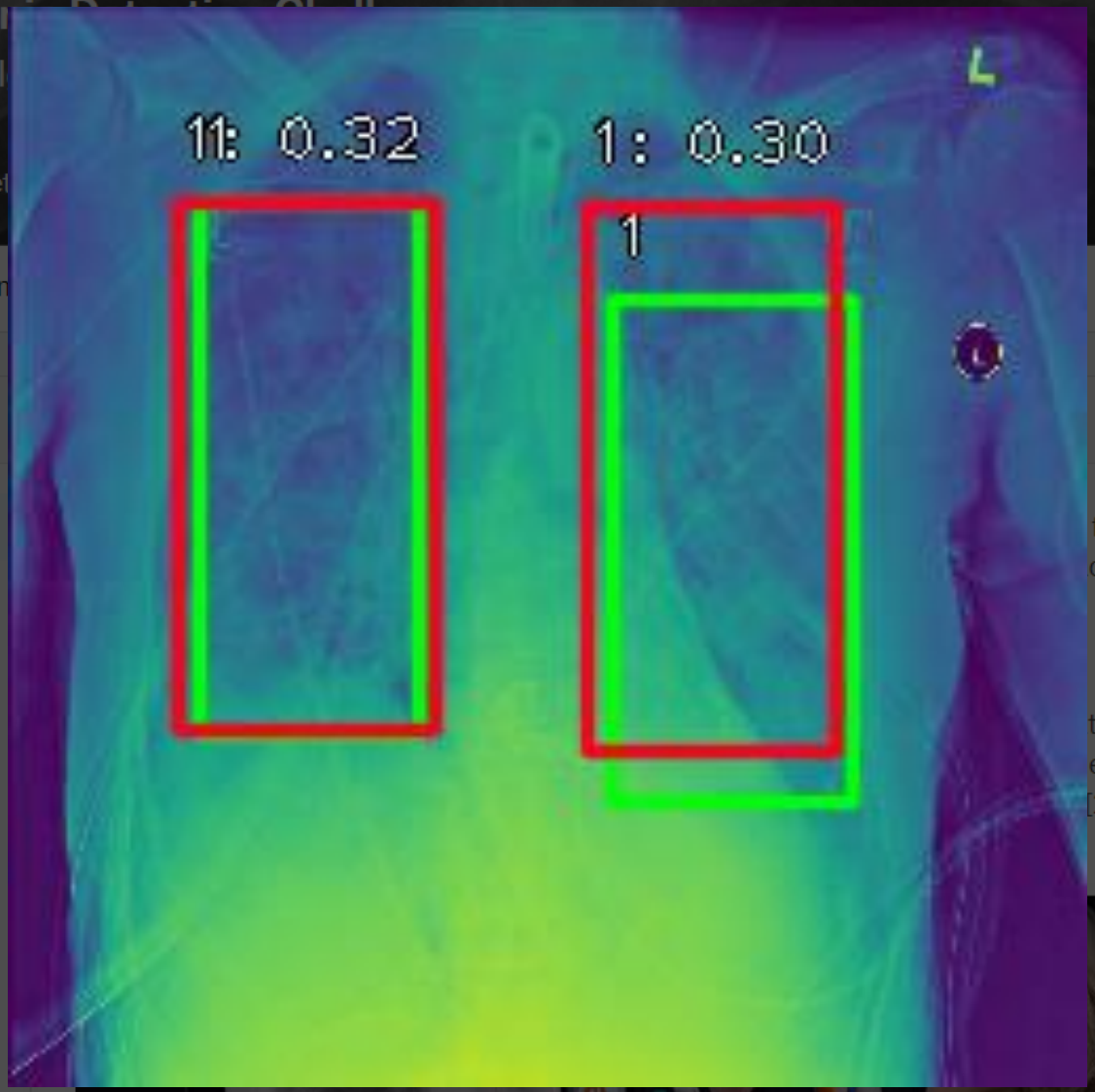
Prizes

Getting Started

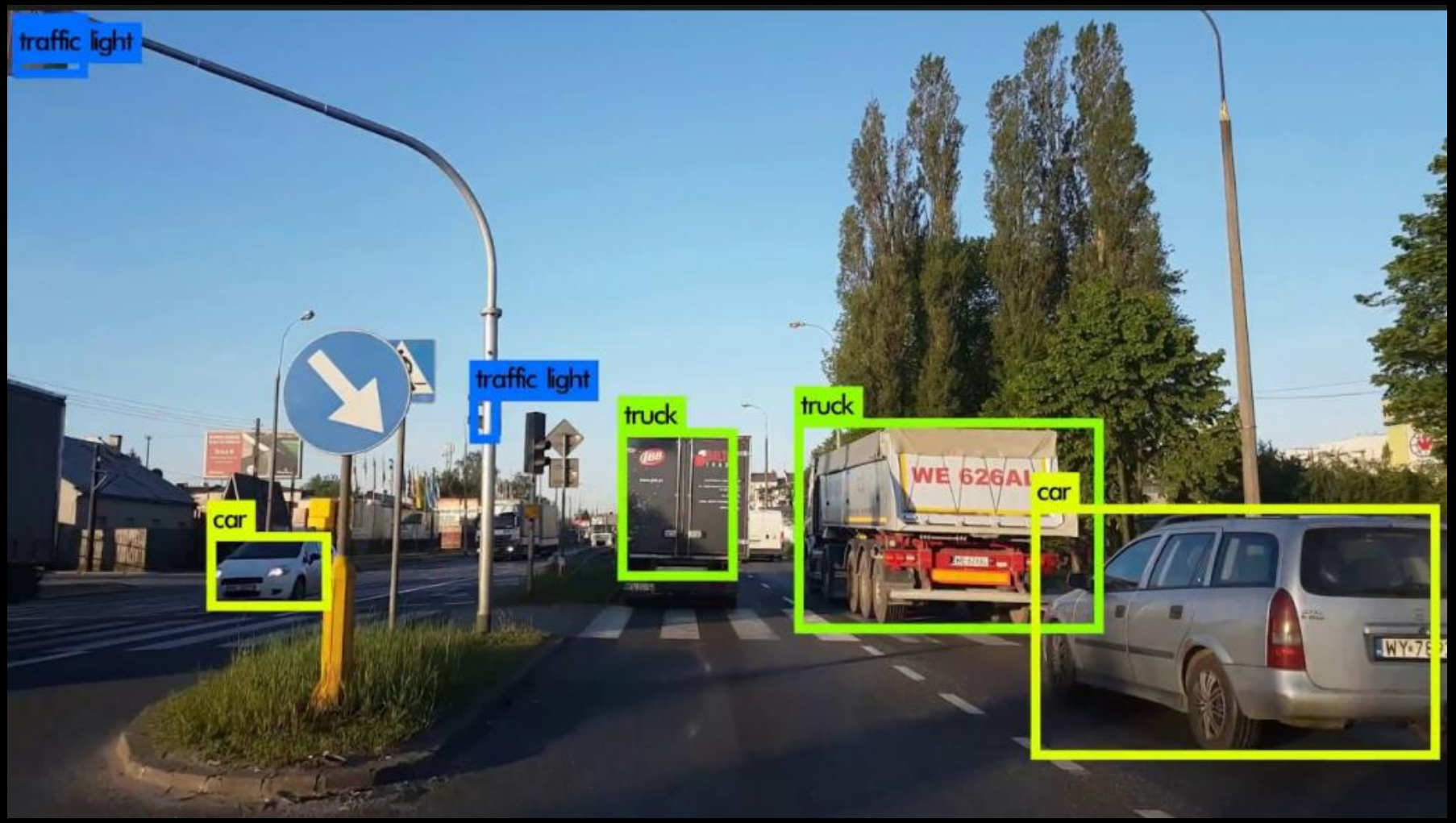
Acknowledgements

for pneumonia in
cities on chest

tionally. In 2015,
eumonia accounts for
[2], keeping the



Object Detection (物件偵測)



花蓮青農數西瓜

某些場合數人頭


5349
99%
5381
5345
99%
5384
99%
5408
99%
5439
99%
5446
99%
5479
99%
5502
100%
5494
99%
5587
99%
5584
99%
5607
99%
5625
100%
5629
99%
5676
99%
5707
99%
5717
99%
5755
99%
5817
99%
5815
99%
5816

選擇演算法

- 準確度 或 預測速度

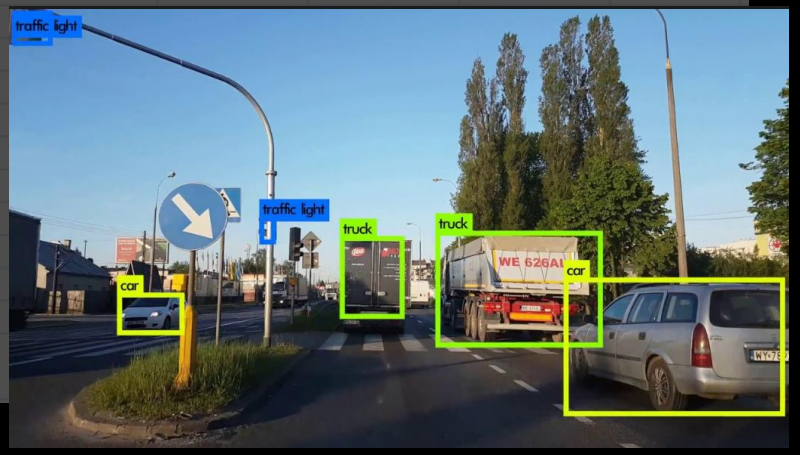
Method	VOC 2007 test	VOC 2012 tes	COCO (AP @[0.5:0.95	COCO (AP @0.5	COCO (AP @[0.5:0.95]) (small)	COCO (AP @[0.5:0.95]) (mid)	COCO (AP @[0.5:0.95]) (large)	time (fps)
One-Stage								
YOLO	52.7/63.4	57.9/NA	NA					45/155
YOLOv2	78.6	73.4	21.6					40
YOLOv3 608x608 (Darknet-53)			33	57.9	18.3	35.4	41.9	
SSD	77.2/79.8	75.8/78.5	25.1/28.8					46/19
DSSD	81.5	80	33.2					5.5
RON	81.3	80.7	27.4					15
RetinaNet800 (ResNet-101-FPN)	NA	N	39.1	59.1	21.8	42.7	50.2	5
RetinaNet800 (ResNeXt-101-FPN)			40.8	61.1	24.1	44.2	51.2	
RefineDet512+ (ResNet-101)			41.8	62.9	25.6	45.1	54.1	
CornerNet			42.1	57.8	20.8	44.8	56.7	
Two-Stage								
RCNN	66	NA	NA					47s
Fast RCNN	77	(wth coco data)	NA					0.5s
Faster RCNN	73.2	70.4	NA					200ms
RFCN	79.5	77.6	29.9					170ms
FPN	NA	NA	36.2	59.1	18.2	39	48.2	6
Faster-RCNN (ResNet-101)	85.6	83.8	37.4	59				
Mask RCNN	NA	NA	38.2					2.5
R-FCN with Deformable Convolutional Networks			37.5	58	19.4	40.1	52.5	
Deformable R-FCN, ResNet-v1-101			35.7	56.8	15.2	38.8	51.5	
Deformable FPN + OHEM, ResNet-v1-101			41.2	63.5	24.3	44.9	54.4	
Deformable FPN + OHEM + Soft NMS + multi-scale testing, ResNet-v1-101			44.4	65.5	30.8	47.3	56.4	

選擇演算法

速度:

1. SSD
2. YOLO3

Method	VOC 2007 test	VOC 2012 tes	COCO (AP @[0.5:0.95	COCO (AP @0.5	COCO (AP @[0.5:0.95]) (small)	COCO (AP @[0.5:0.95]) (mid)	COCO (AP @[0.5:0.95]) (large)	time (fps)
One-Stage								
YOLO	52.7/63.4	57.9/NA	NA					45/155
YOLOv2	78.6	73.4	21.6					40
YOLOv3 608x608 (Darknet-53)			33	57.9	18.3	35.4	41.9	
SSD	77.2/79.8	75.8/78.5	25.1/28.8					46/19
DSSD	81.5	80	33.2					5.5
RON	81.3	80.7	27.4					15
RetinaNet800 (ResNet-101-FPN)	NA	N	39.1	59.1	21.8	42.7	50.2	5
RetinaNet800 (ResNeXt-101-FPN)			40.8	61.1	24.1	44.2	51.2	
RefineDet512+ (ResNet-101)			41.8	62.9	25.6	45.1	54.1	
CornerNet			42.1	57.8	20.8	44.8	56.7	
Two-Stage								
RCNN	66	NA	NA					
Fast RCNN	77	(wth coco data)	NA					
Faster RCNN	73.2	70.4	NA					
RFCN	79.5	77.6	29.9					
FPN	NA	NA	36.2					
Faster-RCNN (ResNet-101)	85.6	83.8	37.4					
Mask RCNN	NA	NA	38.2					
R-FCN with Deformable Convolutional Networks			37.5					
Deformable R-FCN, ResNet-v1-101			35.7					
Deformable FPN + OHEM, ResNet-v1-101			41.2					
Deformable FPN + OHEM + Soft NMS + multi-scale testing, ResNet-v1-101			44.4					

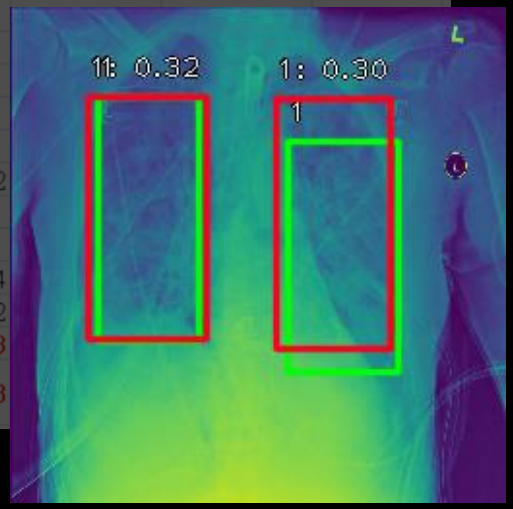


選擇演算法

準確度:

1. RetinaNet
2. Mask R-CNN
3. Deformable FCN

Method	VOC 2007 test	VOC 2012 tes	COCO (AP @[0.5:0.95	COCO (AP @0.5	COCO (AP @[0.5:0.95]) (small)	COCO (AP @[0.5:0.95]) (mid)	COCO (AP @[0.5:0.95]) (large)	time (fps)
One-Stage								
YOLO	52.7/63.4	57.9/NA	NA					45/155
YOLOv2	78.6	73.4	21.6					40
YOLOv3 608x608 (Darknet-53)			33	57.9	18.3	35.4	41.9	
SSD	[illegible]	[illegible]	25.1/28.8					46/19
DSSD	[illegible]	[illegible]	33.2					5.5
RON	[illegible]	[illegible]	27.4					15
RetinaNet800 (ResNet-101-FPN)	NA	N	39.1	59.1	21.8	42.7	50.2	5
RetinaNet800 (Re[illegible]t-101-FPN)			[illegible]	61.1	24.1	44.2	51.2	
RefineDet512+ (Res[illegible]t-101)			[illegible]	62.9	25.6	45.1	54.1	
CornerNet			[illegible]	57.8	20.8	44.8	56.7	
Two-Stage								
RCNN	[illegible]	NA	NA					
Fast RCNN	[illegible]	[illegible]vth pco[illegible]	NA					
Faster RCNN	73.2	70.4	NA					
RFCN	79.5	77.6	29.9					
FPN	NA	NA	36.2	59.1	18.2			
Faster-RCNN (ResNet-101)	85.6	83.8	37.4	59				
Mask RCNN	NA	NA	38.2					
R-FCN with Deformable Convolutional Networks			37.5	58	19.4			
Deformable R-FCN, ResNet-v1-101			35.7	56.8	15.2			
Deformable FPN + OHEM, ResNet-v1-101			41.2	63.5	24.3			
Deformable FPN + OHEM + Soft NMS + multi-scale testing, ResNet-v1-101			44.4	65.5	30.8			



肺炎偵測的演算法

準確度:

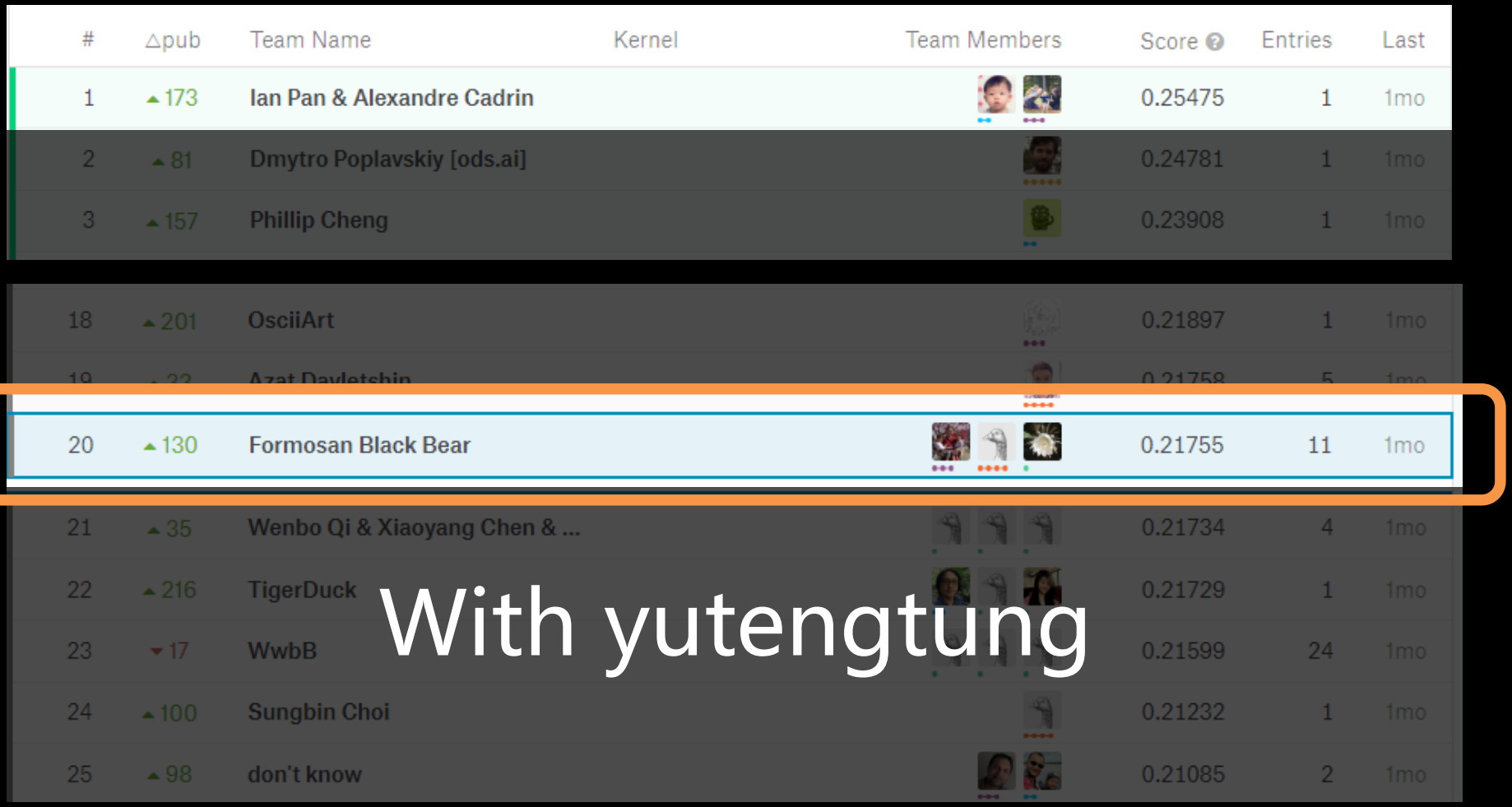
1. RetinaNet
2. Mask R-CNN

結合

3. Deformable FCN

Method	VOC 2007 test	VOC 2012 tes	COCO (AP @[0.5:0.95	COCO (AP @0.5	COCO (AP @[0.5:0.95]) (small)	COCO (AP @[0.5:0.95]) (mid)	COCO (AP @[0.5:0.95]) (large)	time (fps)
One-Stage								
YOLO	52.7/63.4	57.9/NA	NA					45/155
YOLOv2	78.6	73.4	21.6					40
YOLOv3 608x608 (Darkn			33	57.9	18.3	35.4	41.9	
SSD	2/79.8	75.8/78.5	25.1/28.8					46/19
DSSD	81.5	80	33.2					5.5
RON	81.3	80.7	27.4					15
RetinaNet800 (ResNet-101 FPN)		N	39	59.1	21.8	42.7	50.2	5
RetinaNet800 (ResNeXt-1 -FPN)			0.8	61.1	24.1	44.2	51.2	
RefineDet512+ (ResNet-1)			41.8	62.9	25.6	45.1	54.1	
CornerNet			42.1	57.8	20.8	4	56.7	
Two-Stage								
RCNN	66	NA	NA					47s
Fast RCNN								0.5s
Faster RCNN	73.	70.4	NA					200ms
RFCN	79.	77.6	29					170ms
FPN	NA	NA	36.2	59.1	18.2	39	48.2	6
Faster-RCNN (ResNet-101)	85.6	83.8	37.4	59				
Mask RCNN	NA	NA	38.2					2.5
R-FCN with Deformable Convolutional Networks			37.5	58	19.4	40.1	52.5	
Deformable R-FCN, ResNet-v1-101			35.7	56.8	15.2	38.8	51.5	
Deformable FPN + OHEM, ResNet-v1-101			41.2	63.5	24.3	44.9	54.4	
Deformable FPN + OHEM + Soft NMS + multi-scale testing, ResNet-v1-101			44.4	65.5	30.8	47.3	56.4	

Win the Silver Medal
Rank 20 (/1499, Top 2%)



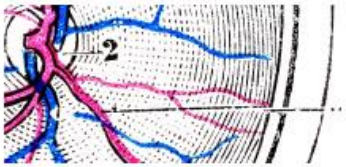
#	△pub	Team Name	Kernel	Team Members	Score	Entries	Last
1	▲173	Ian Pan & Alexandre Cadrin			0.25475	1	1mo
2	▲81	Dmytro Poplavskiy [ods.ai]			0.24781	1	1mo
3	▲157	Phillip Cheng			0.23908	1	1mo
18	▲201	OsciiArt			0.21897	1	1mo
20	▲130	Formosan Black Bear			0.21755	11	1mo
21	▲35	Wenbo Qi & Xiaoyang Chen & ...			0.21734	4	1mo
22	▲216	TigerDuck			0.21729	1	1mo
23	▼17	WwbB			0.21599	24	1mo
24	▲100	Sungbin Choi			0.21232	1	1mo
25	▲98	don't know			0.21085	2	1mo

With yutengtung

Outline

- Camera Model Identification
  - 相機型號識別競賽 銀牌經驗分享
- Passenger Screening Algorithm Challenge
  - 機場乘客檢查競賽 銀牌經驗分享
- RSNA Pneumonia Detection Challenge
  - 肺炎檢測挑戰 銀牌經驗分享
- Practical Application in Kaggle

Diabetic Retinopathy Detection



Diabetic Retinopathy Detection

Identify signs of diabetic retinopathy in eye images

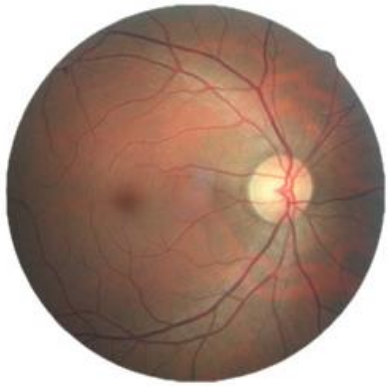
$100,000 · 661 teams · 3 years ago

Overview | Data | Kernels | Discussion | Leaderboard | Rules | Team | My Submissions | Late Submission

Overview

- Description
- Evaluation
- Prizes
- References
- Timeline

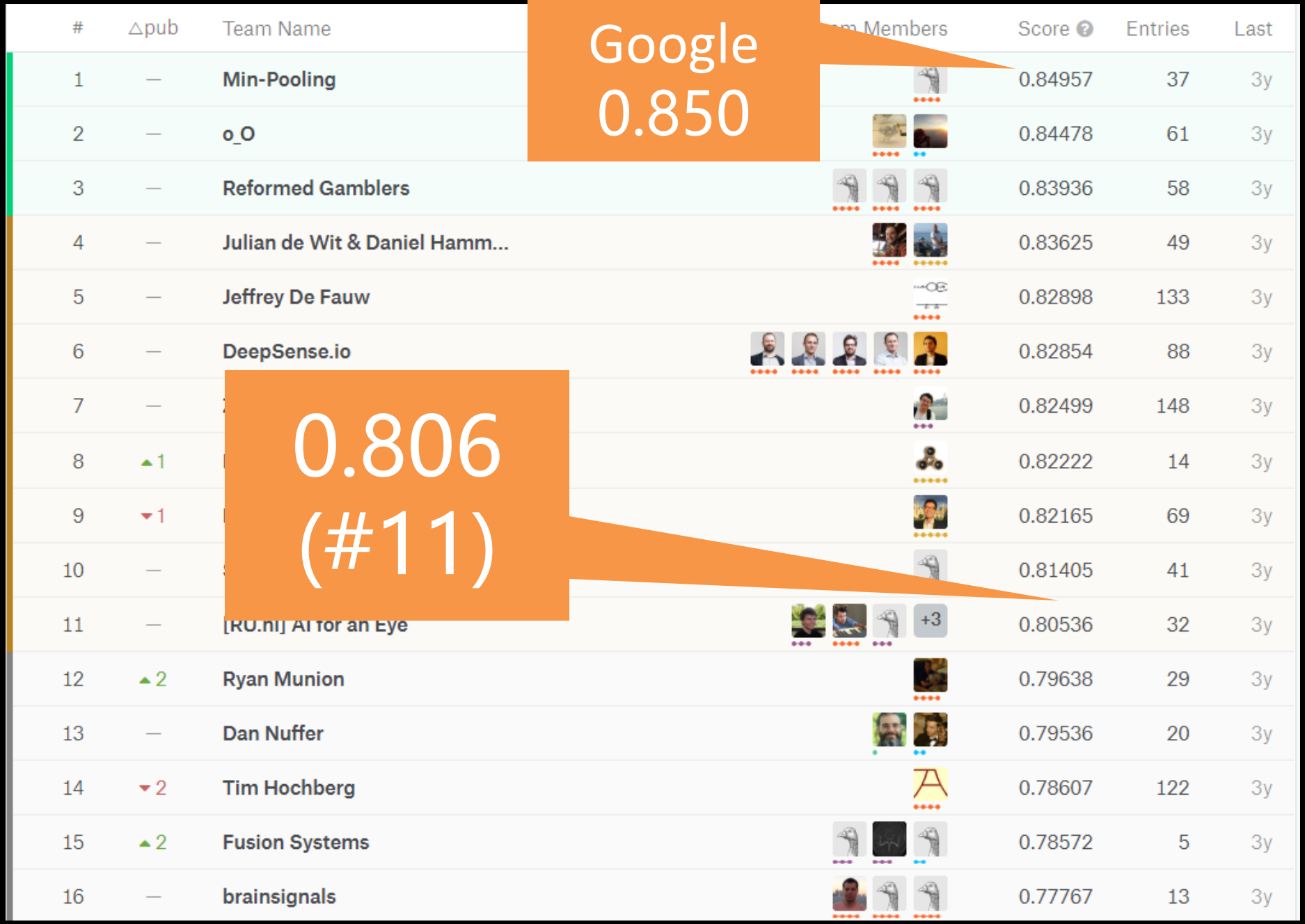
Diabetic retinopathy is the leading cause of blindness in the working-age population of the developed world. It is estimated to affect over 93 million people.



The US Center for Disease Control and Prevention estimates that 29.1 million people in the US have diabetes and the World Health Organization estimates that 347 million people have the disease worldwide. Diabetic Retinopathy (DR) is an eye disease associated with long-standing diabetes. Around 40% to 45% of Americans with diabetes have some stage of the disease. Progression to vision impairment can be slowed or averted if DR is detected in time, however this can be difficult as the disease often shows few symptoms until it is too late to provide effective treatment.

Currently, detecting DR is a time-consuming and manual process that requires a trained clinician to examine and evaluate digital

Ref: https://www.kaggle.com/c/diabetic-retinopathy-detection

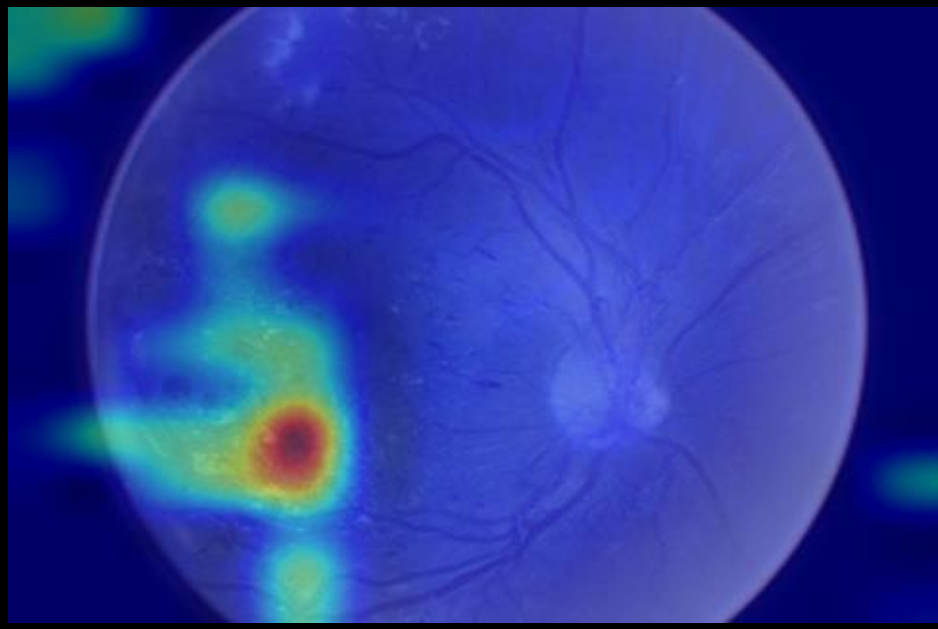


#	△pub	Team Name	Members	Score	Entries	Last
1	—	Min-Pooling		0.84957	37	3y
2	—	o_O		0.84478	61	3y
3	—	Reformed Gamblers		0.83936	58	3y
4	—	Julian de Wit & Daniel Hamm...		0.83625	49	3y
5	—	Jeffrey De Fauw		0.82898	133	3y
6	—	DeepSense.io		0.82854	88	3y
7	—			0.82499	148	3y
8	▲1			0.82222	14	3y
9	▼1			0.82165	69	3y
10	—			0.81405	41	3y
11	—	[RUnl] AI for an Eye	+3	0.80536	32	3y
12	▲2	Ryan Munion		0.79638	29	3y
13	—	Dan Nuffer		0.79536	20	3y
14	▼2	Tim Hochberg		0.78607	122	3y
15	▲2	Fusion Systems		0.78572	5	3y
16	—	brainsignals		0.77767	13	3y

Grad-CAM in Proliferative DR for Explanation (XAI)

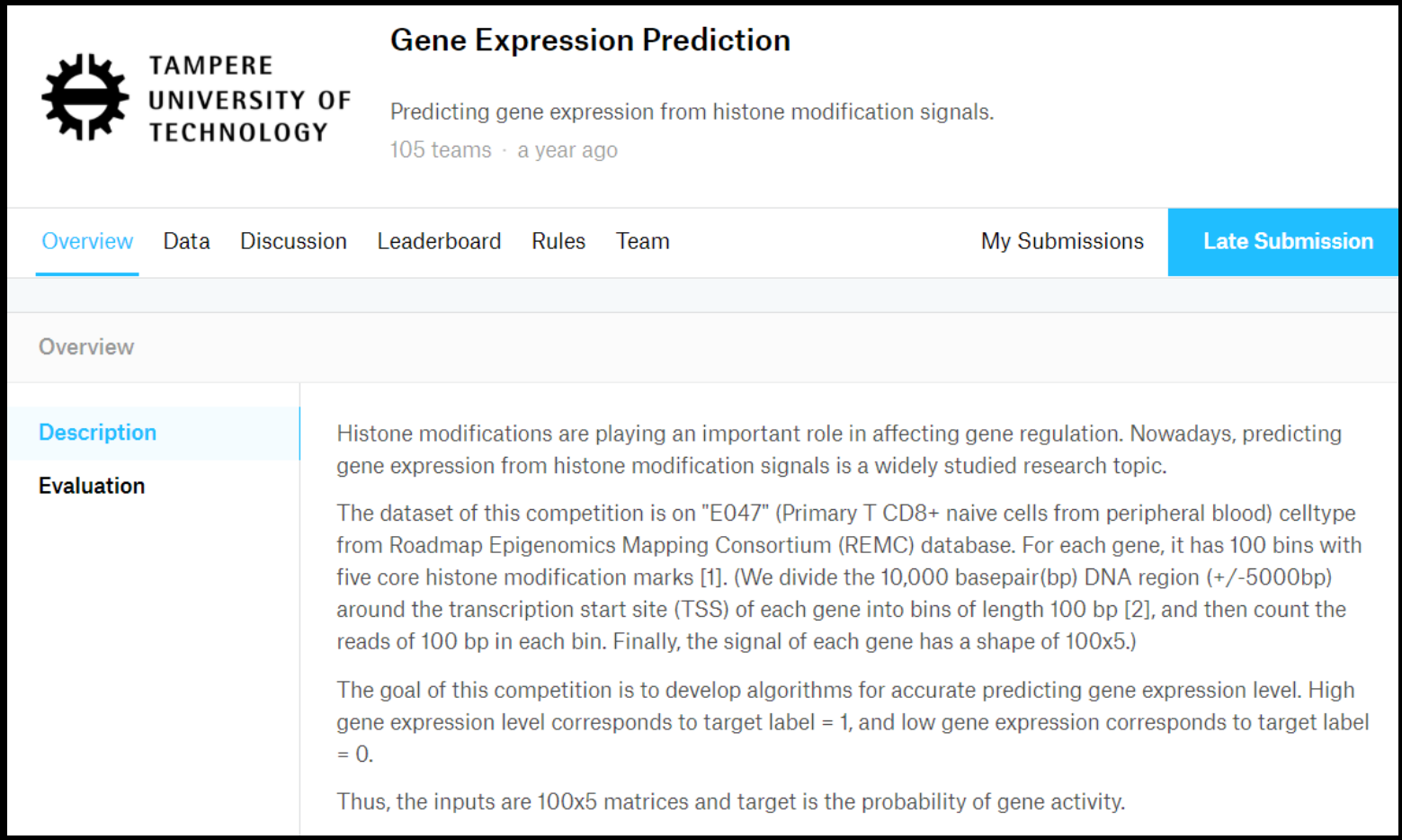
Input

Grad-CAM



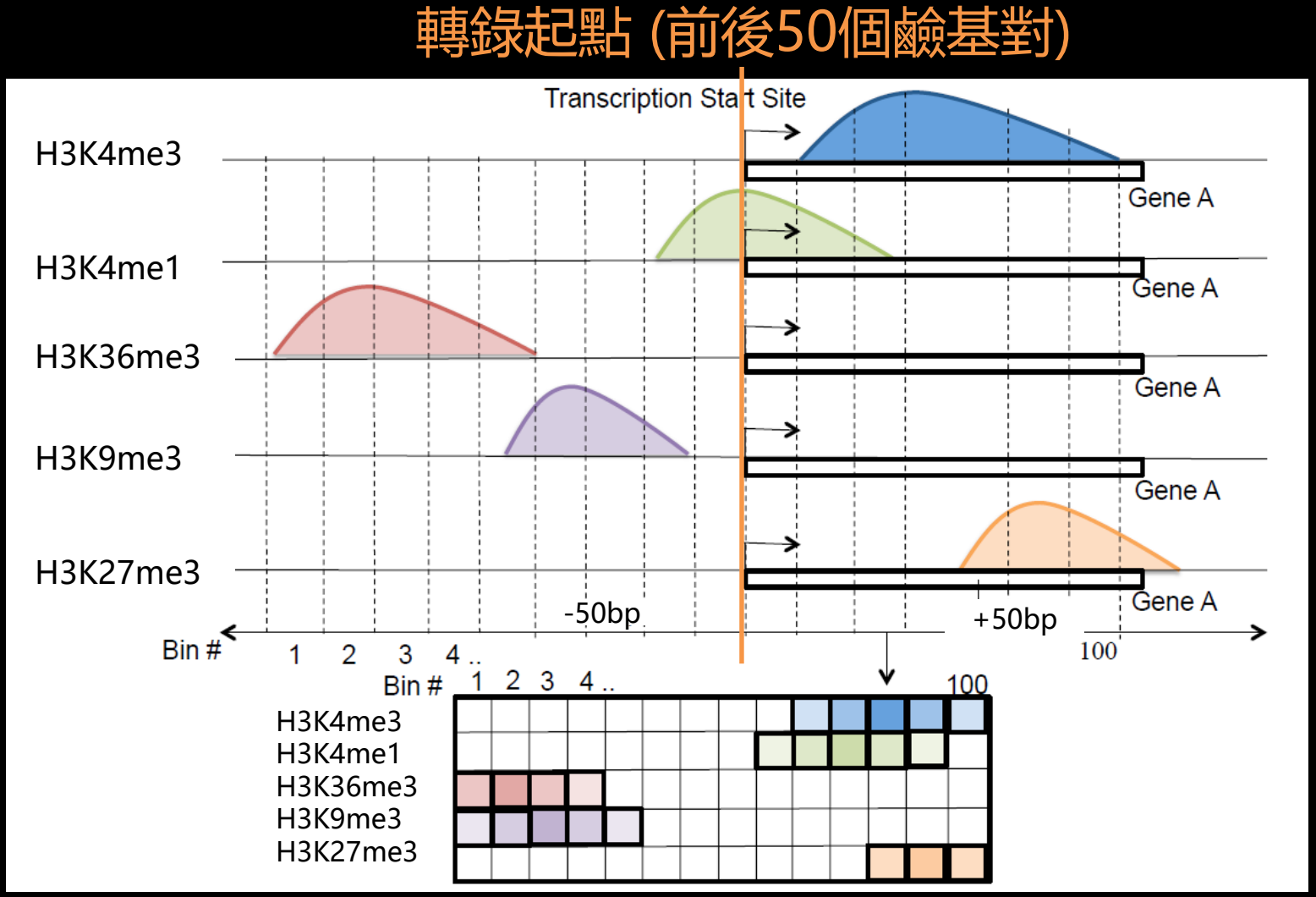
No DR

Gene Expression Prediction



https://www.kaggle.com/c/gene-expression-prediction

HM Signal and Gene Expression



Ensemble DeepChrome and werDNAnet with 5-fold

Submission and Description	Private Score
submission_merge_deep_chrome_conv2_chrome_.csv 7 days ago by AndrewWang	0.92942

TAMPERE UNIVERSITY OF TECHNOLOGY

Gene Expression Prediction

Predicting gene expression from histone modification signals.

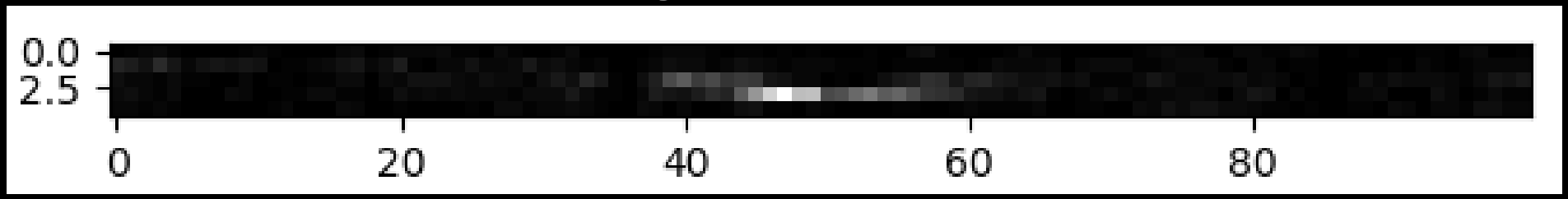
105 teams · a year ago

#2

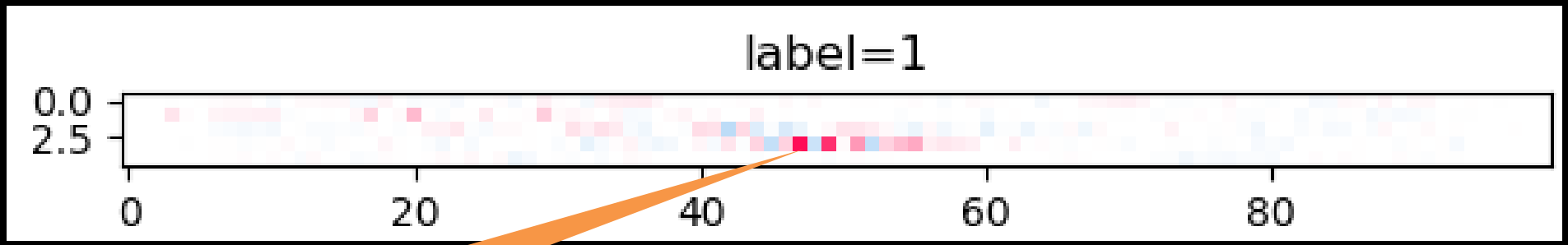
#	△pub	Team Name	Kernel	Team Members	Score
1	—	AndrewChang			0.94798
2	▲8	Group 32 (262847, 270715, 27...			0.92931
3	▲20	Group 36			0.92801
4	▼1	Group 40			0.92787
5	▼1	group28_tut (267926, 267071, ...			0.92777
6	▲7	richpiana			0.92706
7	▲19	Group 5 TUT/BINF			0.92596

SHAP of Gene Expression (XAI)

Histone Modifications Signal:



SHAP Value:



H3K9me3

王淳恆 Andrew Wang

- Muen Biomedical (沐恩生醫)

- Expertise
  - Technical Manager in Mediatek
  - Medical AI (2018~)
  - Deep Learning in Image (2015~)
  - Finance Analysis using AI (2005~)
  - Algorithm and Chip Design in wireless communication and image processing (2001~)

王淳恆 Andrew Wang