PFDet: 2nd Place Solution to OIC Object Detection Track

Speaker: Yusuke Niitani

Takuya Akiba*, Tommi Kerola*, Yusuke Niitani*,
Toru Ogawa*, Shotaro Sano*, and Shuji Suzuki*
*: Equal Contribution

Preferred Networks, Inc., Japan

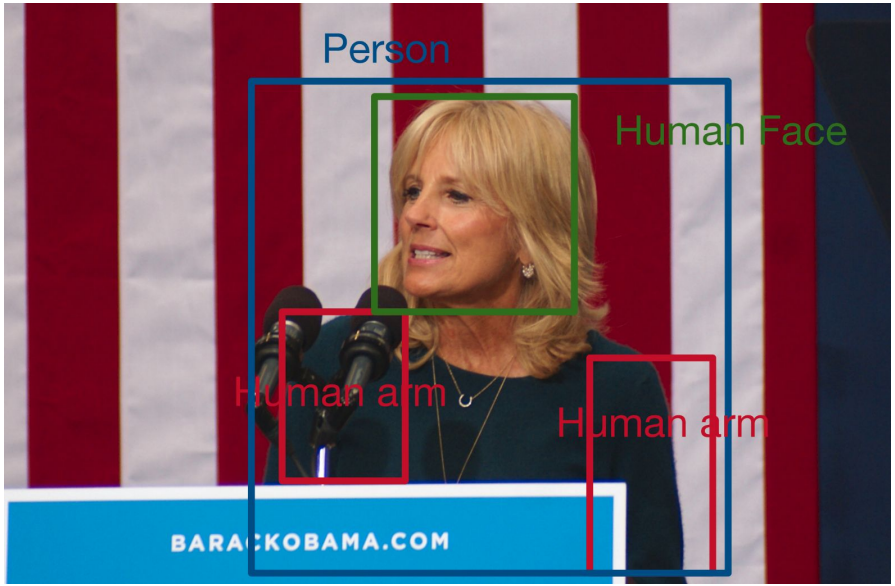
Outline

1. Sparsely verified labels
2. Huge class imbalance
3. Large number of images
4. Basic Architecture
5. Ensembling
6. Results

Challenges of OIC

paper link: https://arxiv.org/abs/1809.00778

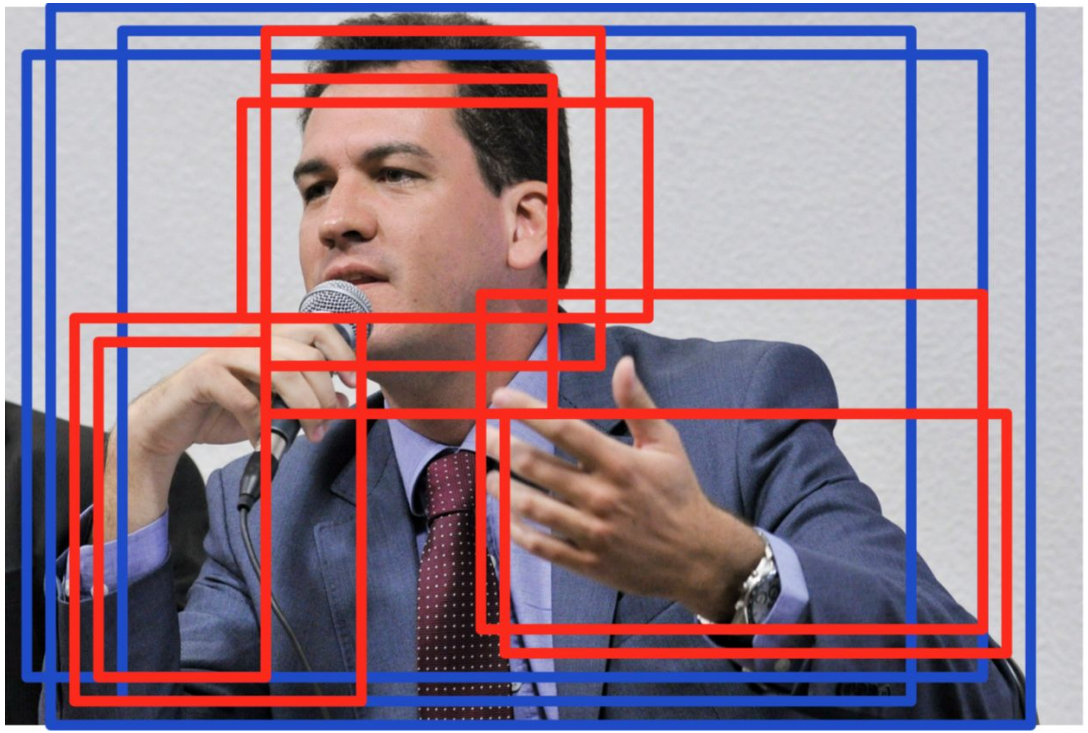
Sparsely Verified Labels



Verified

Human parts unverified

Falsely Labeled Instances

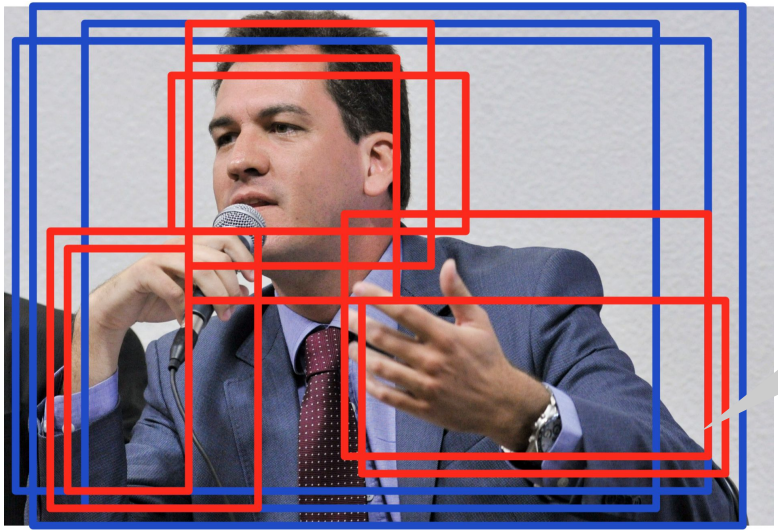


Positive

Negative

Co-occurrence Loss

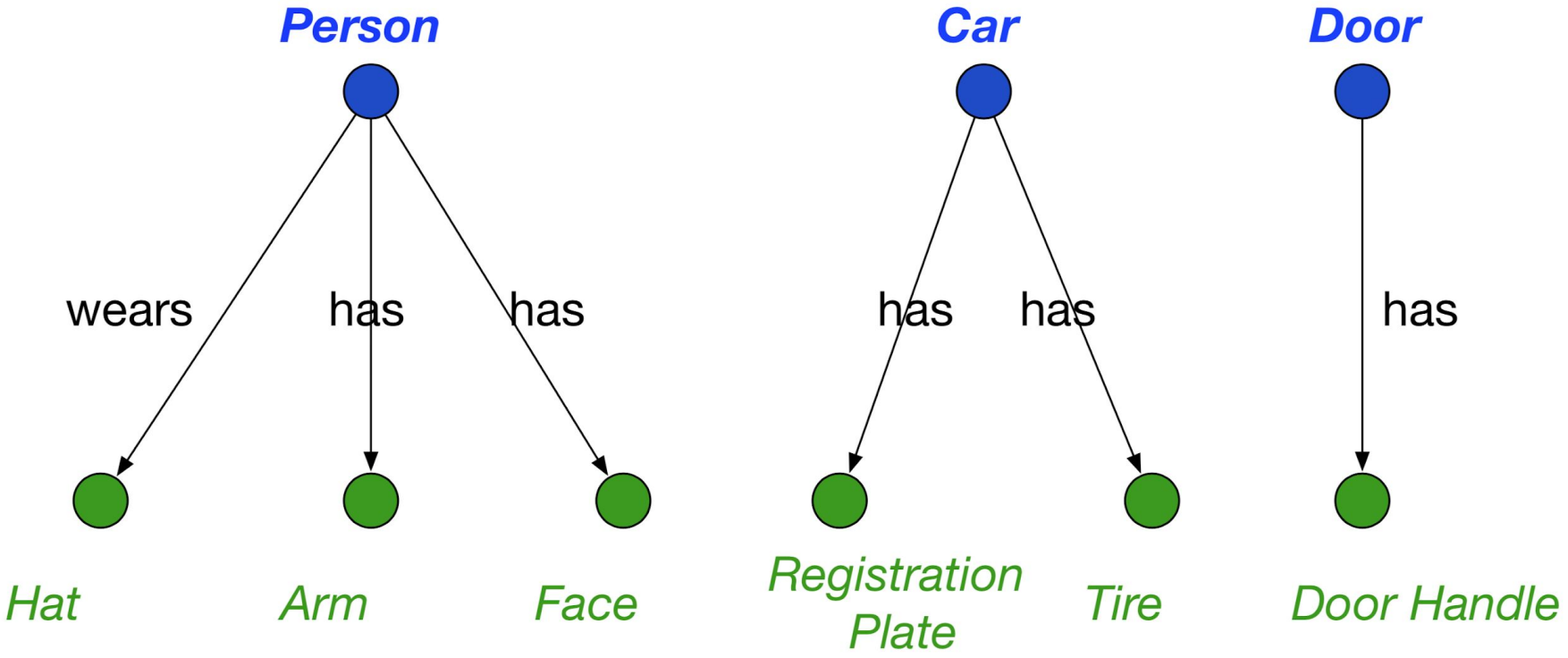
Ignore classification loss for a *part class* if a proposal is inside the *subject class*



Ignore:
Face
Arm

Negative:
Car, ...

Label Relationship



Co-occurrence Loss Results

+22.7AP improvement on *"Human Parts"*

Table 4: Ablative study of co-occurrence loss on classes that can be ignored by the l[illegible]. The scores are AP calculated on the validation set of the dataset.

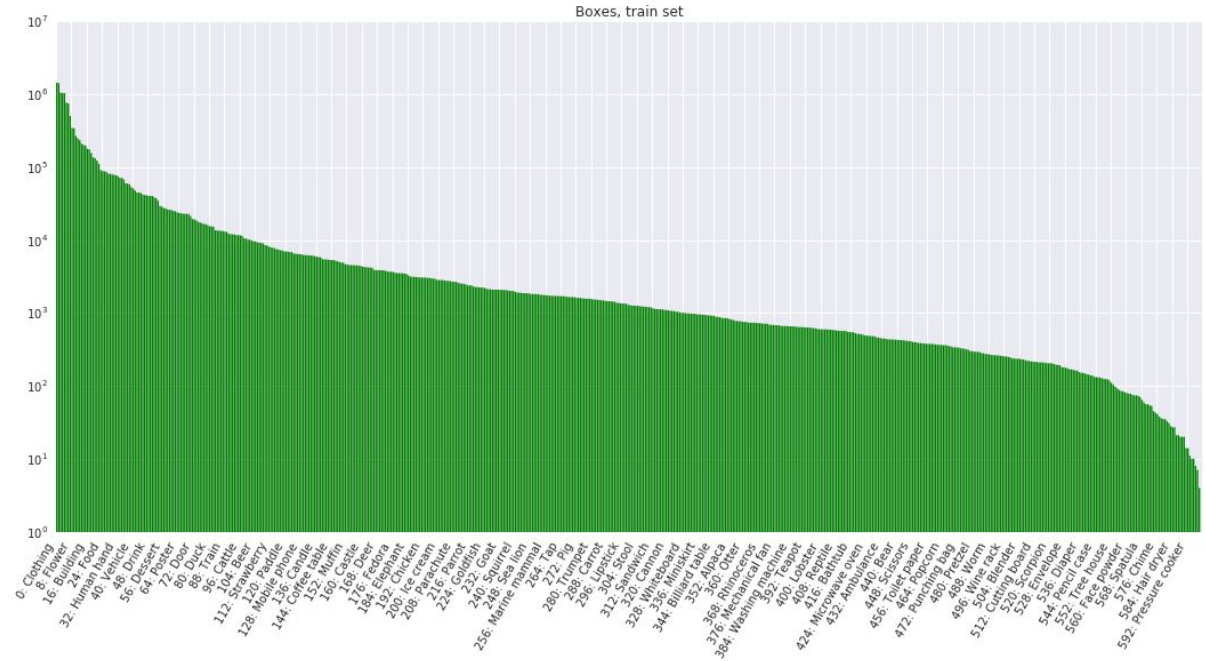
	Arm	Ear	Nose	Mouth	Hair	Eye	Beard	Face	Head	Foot	Leg	Hand	Glove	Hat	Dress	Fedora
Baseline	40.9	17.5	34.7	21.4	63.8	27.3	55.5	82.7	55.1	50.7	41.6	32.3	**63.4**	64.9	70.6	67.0
Co-occurrence	**55.2**	**62.6**	**69.6**	**55.2**	**74.7**	**64.0**	**76.8**	**91.4**	**78.9**	**59.5**	**54.4**	**53.6**	60.8	**69.0**	**73.9**	**70.3**

	Footwe.	Sandal	Boot	Sports.	Coat	Sock	Glasse.	Belt	Helmet	Jeans	High h.	Scarf	Swimwe.	Earrin.	Bicycl.	Shorts
Baseline	61.9	53.6	**61.6**	52.9	58.0	**70.6**	74.9	**66.8**	80.2	62.7	76.6	71.6	**63.4**	82.0	75.1	69.7
Co-occurrence	**68.5**	**58.9**	57.9	**61.2**	**73.3**	67.1	**85.4**	61.9	**82.4**	**77.6**	**78.8**	**75.8**	**63.4**	**86.1**	**75.8**	**75.4**

	Baseba.	Minisk.	Cowboy.	Goggles	Jacket	Shirt	Sun ha.	Suit	Trouse.	Brassi.	Tie	Licens.	Wheel	Tire	Handle	Average
Baseline	**67.2**	**62.5**	65.0	79.3	69.5	70.9	61.3	83.7	62.5	**82.6**	84.7	72.1	48.3	49.4	41.1	61.1
Co-occurrence	62.2	58.7	**73.3**	**86.7**	**74.3**	**81.6**	**66.4**	**87.0**	**69.8**	74.5	**91.5**	**74.6**	**66.4**	[illegible]	**46.2**	**70.3**

on 47 *part classes* average **+9.2AP** improvement

Huge Class Imbalance

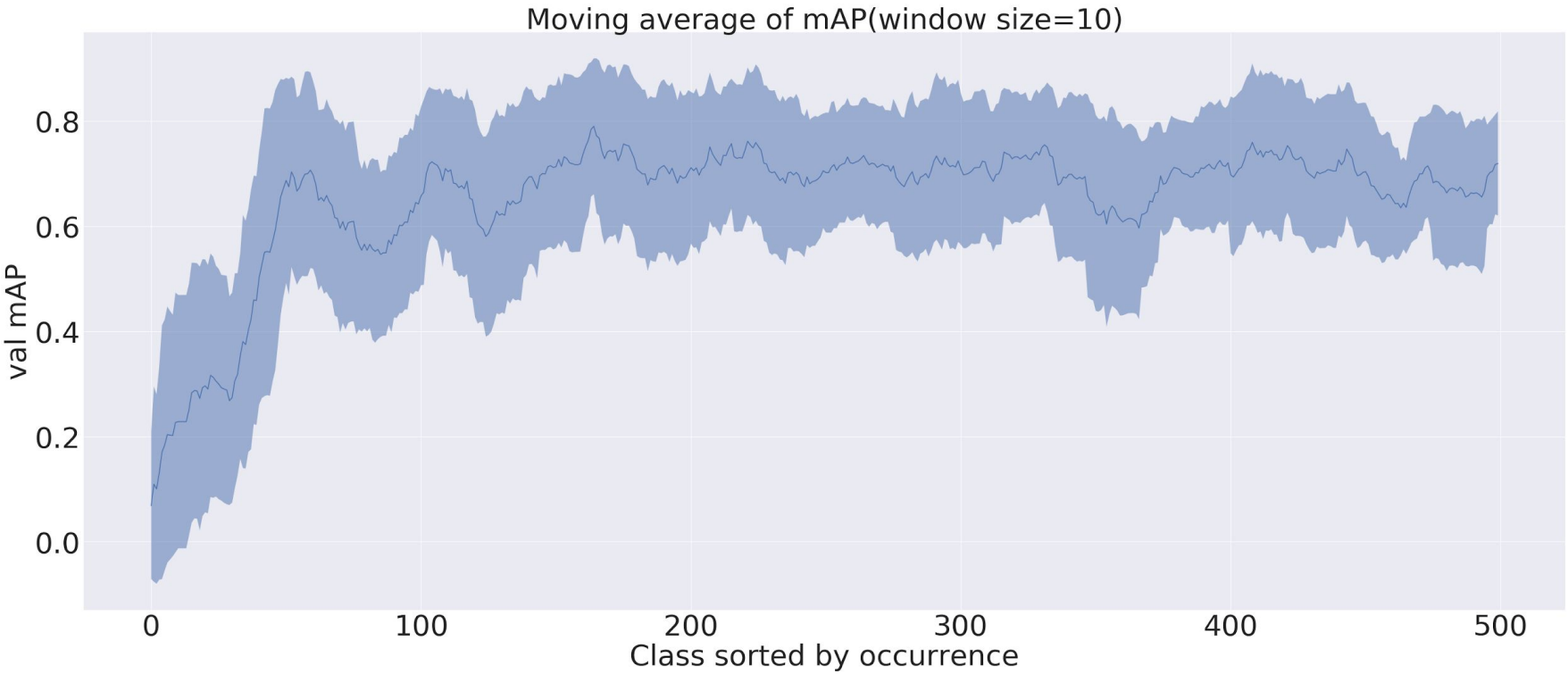


***Pressure Cooker*: 17 images**
***Person*: 800k images**

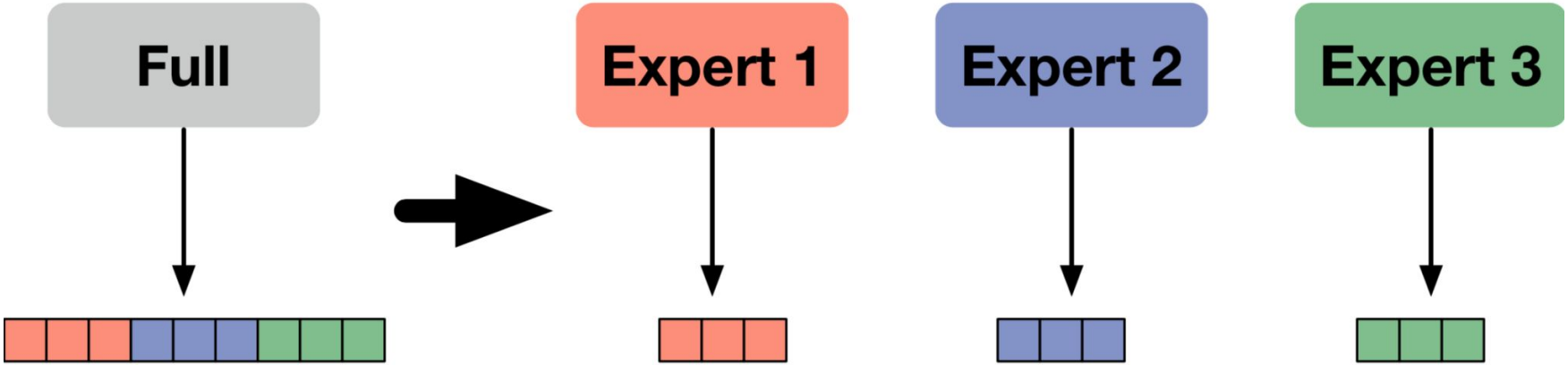
238 classes appear in <1000 images

Image from *https://storage.googleapis.com/openimages/web/factsfigures.html*

Class-wise Performance



Expert model ensemble

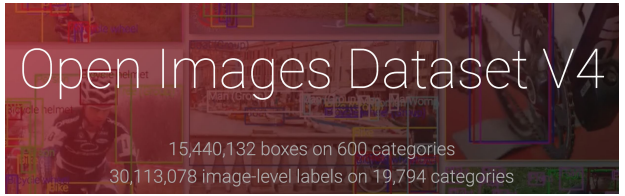


	Index 11-100	Index 101-250	Index 251-350
Full	51.9	70.5	**70.9**
Class10 experts	**65.6**	**73.1**	66.3
Class40 experts	61.0	66.3	50.9

Large Number of Images

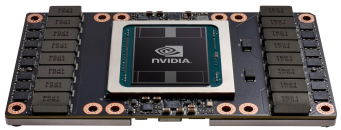
	MS COCO	Open Images
# of classes	80	500
# of images	0.12M	1.7M

Increase by more than **x14**



High Scalability

HW: In-house GPU Cluster



V100 (32GB) x512
Infiniband

SW: ChainerMN

Chainer MN

Scalability Results

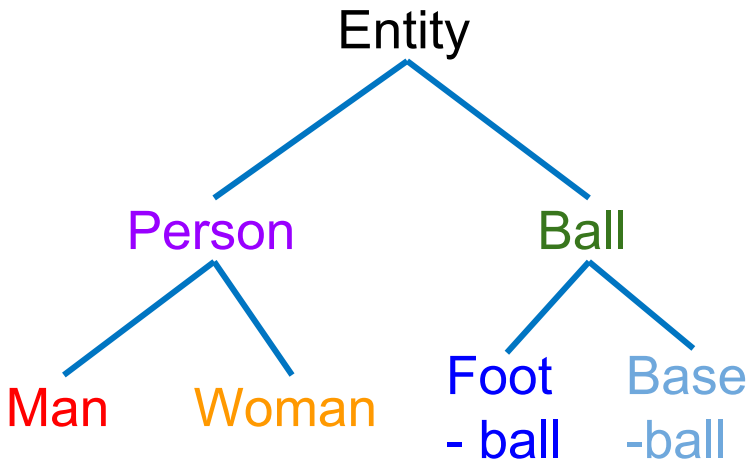
- Training 16 epochs finished in 33 hours
- Scaling efficiency is 83% compared to 8GPUs

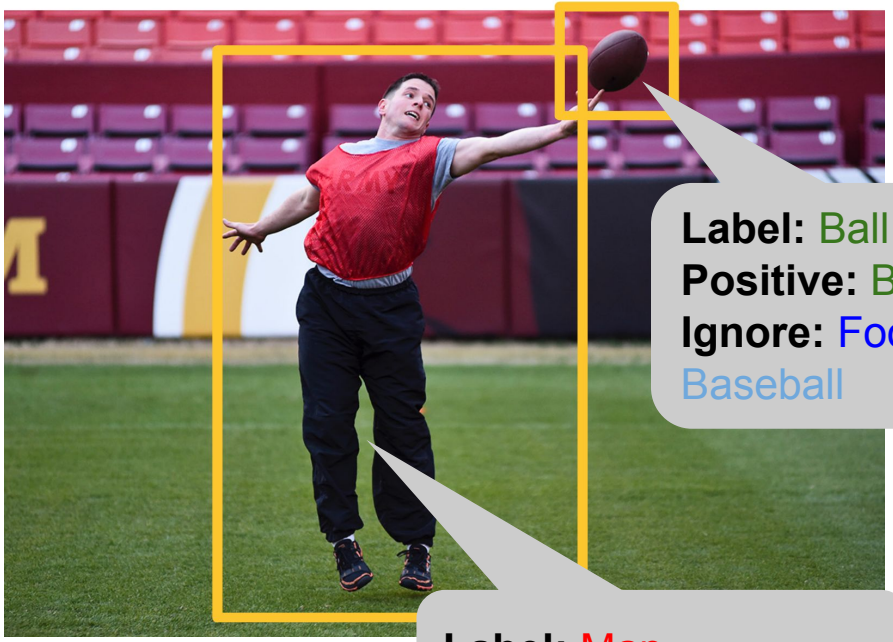
Basic Architecture

- FPN (SENet-154 and SE-ResNeXt-101)
- Multi-node BN
- NMW [1]
- Global Context:
  - Additional FPN block
  - PSP
  - Context head [2]

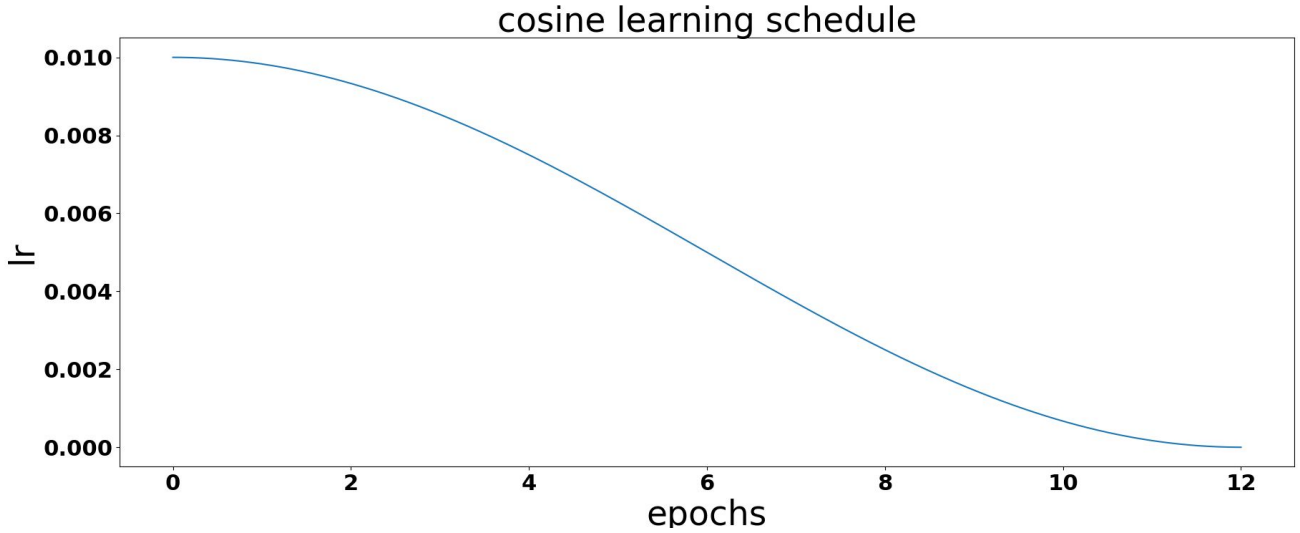
[1]: H. Zhou et. al., Cad: Scale invariant framework for real-time object detection. ICCVW 2017
[2]: Y. Zhu et. al., Couplenet: Coupling global structure with local parts for object detection. ICCV2017

Hierarchy-aware Sigmoid Loss



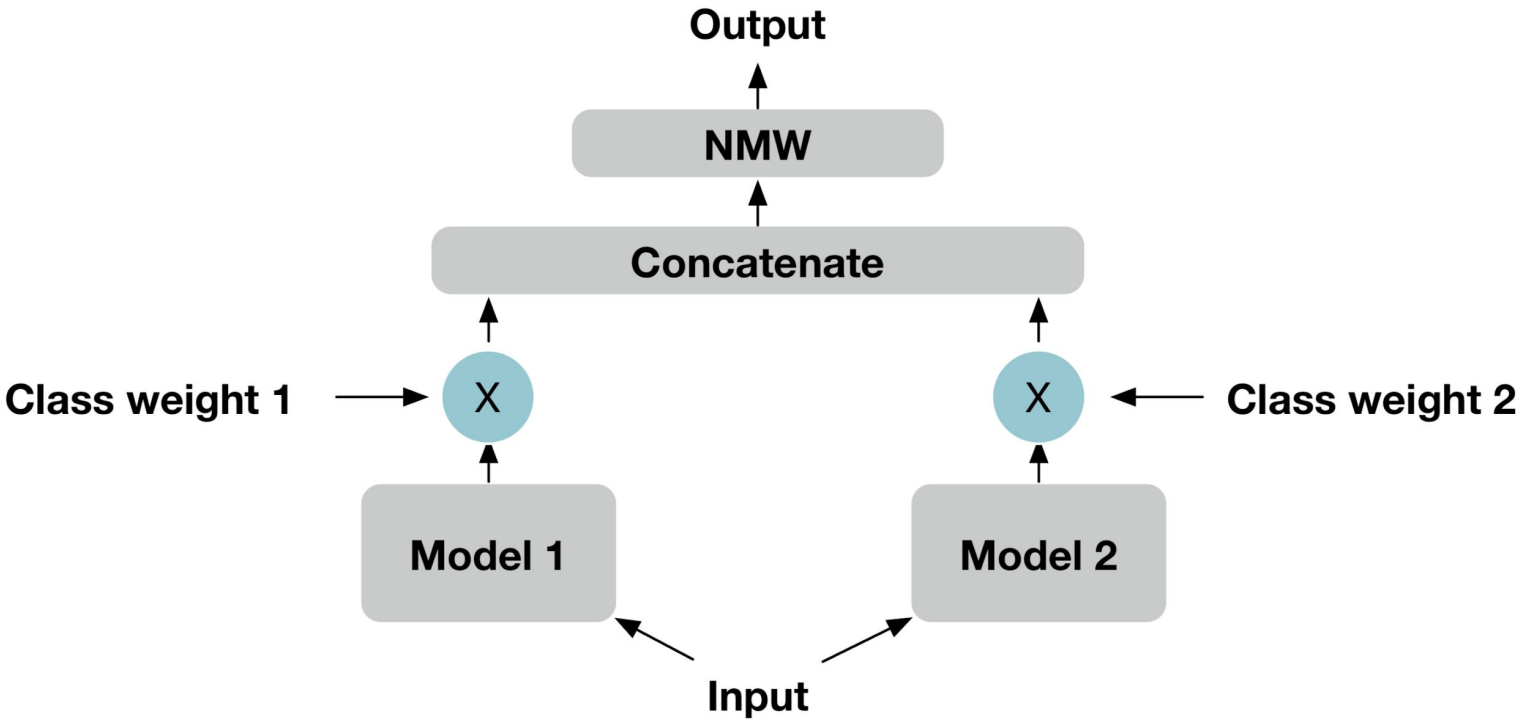


Cosine LR Schedule

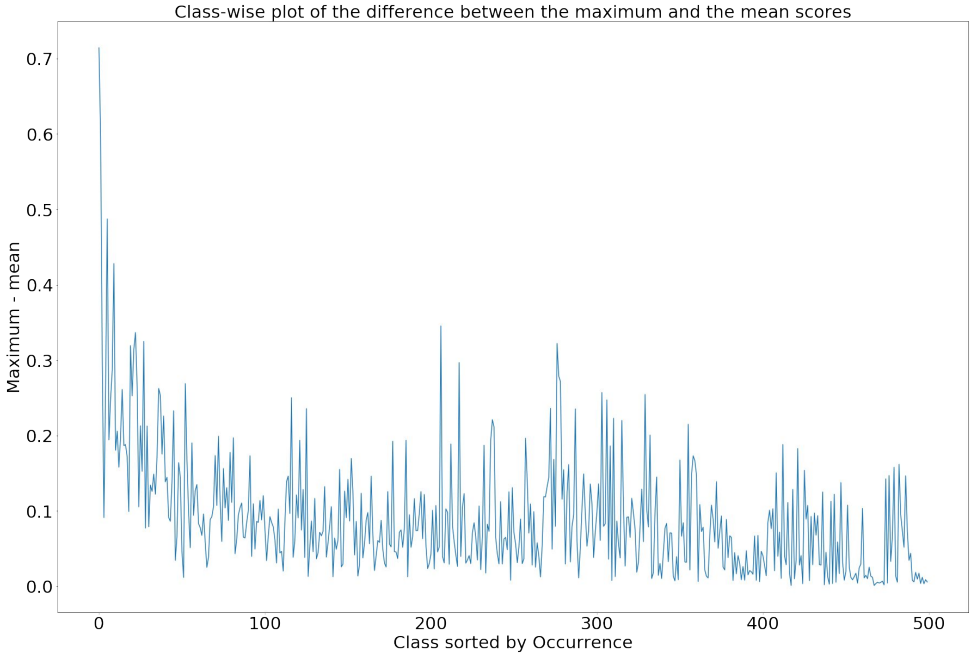


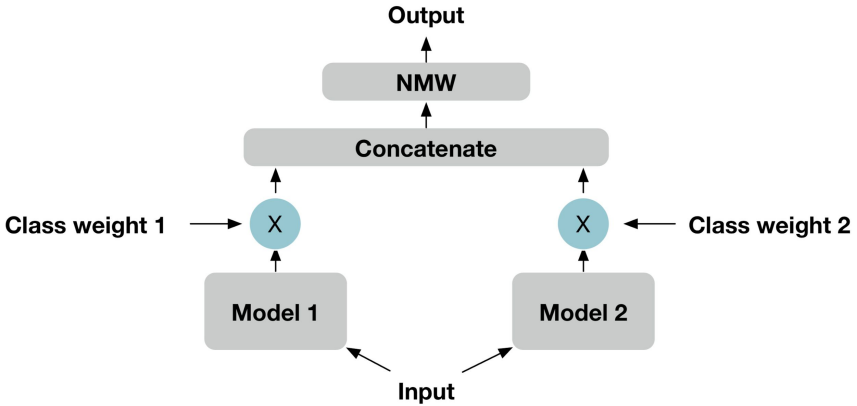
$$\eta = \eta_0 \frac{\cos(\%\ \text{of progress} \times \pi) + 1}{2}$$

Ensemble


Output
NMW
Concatenate
Class weight 1
X
X
Class weight 2
Model 1
Model 2
Input

Ensemble: Class-wise Weight

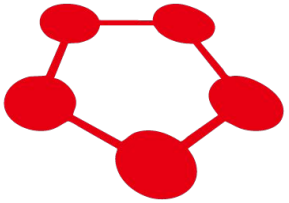




SW Setup

CuPy

GPU Array Library

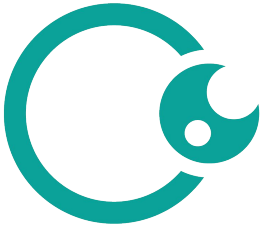


Chainer

Define-by-run DL framework

Chainer **MN**

Multi-Node Distributed DL



Chainer **CV**

CV with DL made easy

Results: Single model

	validation mAP
Baseline (FPN with SE-ResNeXt-101)	60.0
+ multi-scale training	60.3 (+0.3)
+ PSP and add BN to head	60.4 (+0.1)
+ Cosine Annealing	63.4 (+3.0)
+ Add FPN scale	64.5 (+1.1)
+ Co-occurrence loss	65.2 (+0.7)
+ 16 epochs	65.8 (+0.6)
+ Context head	66.0 (+0.2)
+ SENet-154 and additional anchors	67.5 (+1.5)

Results: Ensemble

	val mAP	Public LB	Private LB
Single best model	69.95	55.81	53.43
+ class20 experts	71.73	59.34	55.87
+ class10 experts	72.33	60.19	56.61
+ All the others except COCO	73.98	61.83	57.97
+ COCO	74.07	62.34	58.48
+ class-weight ensemble		**62.88**	58.63
Competition winner		61.71	**58.66**

Conclusion

- Scalability
- Expert models
- Co-occurrence Loss

We are hiring!!!
https://www.preferred-networks.jp/en/jobs

Team members
with Chainer T-shirts