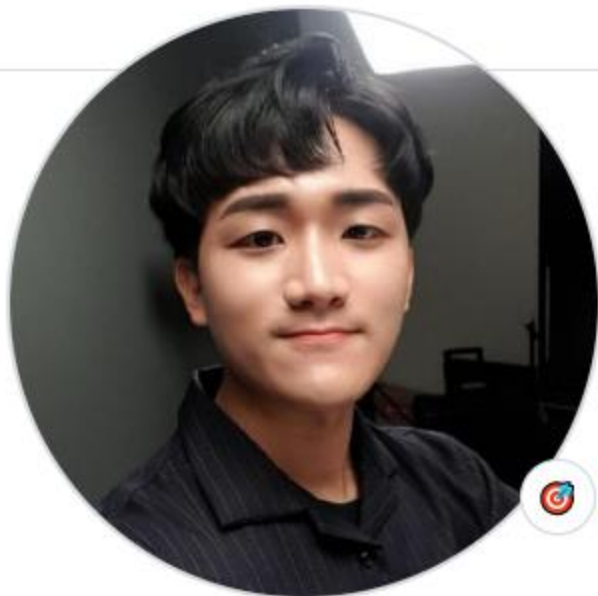


# 경험공유회: AI 비전공자로 AI 논문쓰기

거꾸로 학습하며 배운 AI

조상현 (OGQ GYN)

# 1. 소개



Sanghyun Jo

shjo-april

Deep Learning Research Engineer

## Work experience

[Researcher]

✓ OGQ GYN – Research Engineer (17.07 ~ Present)

1. 국내 최초 영상 보안 지능형CCTV 성능 인증 획득 (방화/쓰러짐, 날씨 환경 적용)
2. 드론 실종자 수색 관련 영상 분석 AI 모델 개발 (경찰청)
3. 수도시설 지능형 안전관리시스템 개발 (한국수자원공사)
4. 다중 언어 지원하는 Tagging AI 모델 개발 (네이버, 연합뉴스, 아프리카TV)
5. Quantization 기반 경량화 AI 개발 (아이닉스)

[Lecturer]

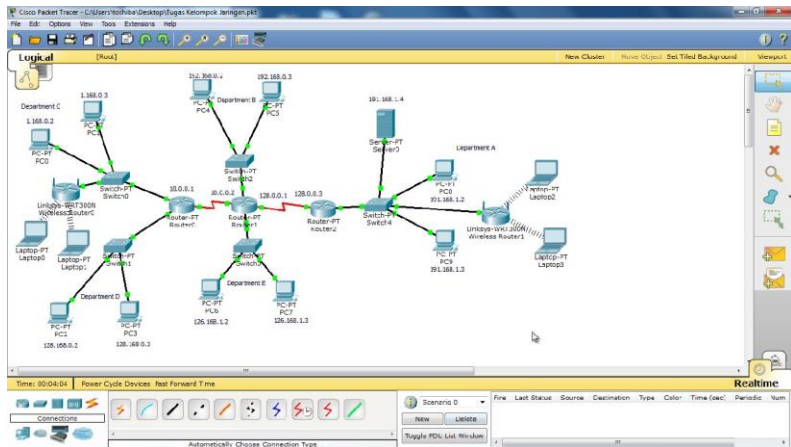
- ✓ Developers or MS/PhD Students – Tutoring (18.06 ~ Present)
- ✓ NEXTPAGE – Lecturer (18.10 ~ 20.09)
- ✓ EasyDeep – Lecturer & VOD (20.11 ~ 21.02)

[Reviewer] – ICIP 2022, IEEE TCSVT 2022, ICML 2022

## Papers

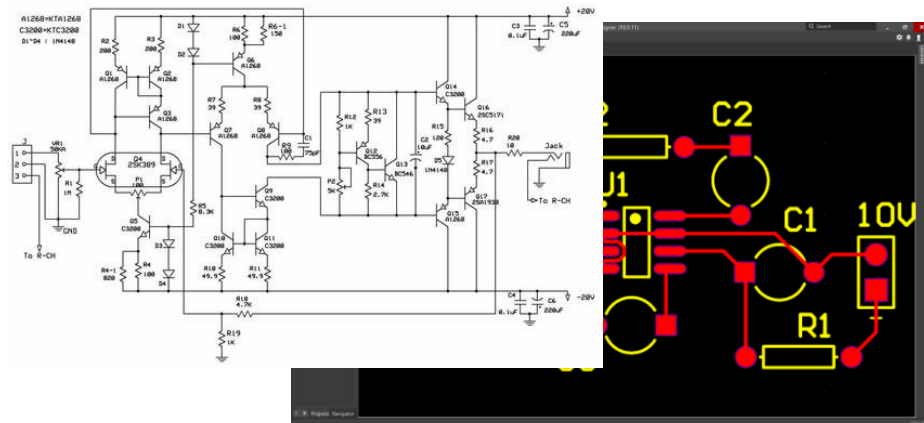
- ✓ **Puzzle-CAM: Improved localization via matching partial and full features**  
(Git stars: 150, Citations: 34, ICIP 2021)
- ✓ **RecurSeed and EdgePredictMix: Single-stage learning is sufficient for Weakly-Supervised Semantic Segmentation**  
(Git stars: 31, Citations: 1, Under Review, SOTA )

# 2. 성장 과정 - 입사 전 (15.03 ~ 17.06)



The screenshot displays a debugger window for the CPU main thread of the module 'ollydbg'. It shows a hex dump of memory addresses and the corresponding assembly instructions. A log data window below shows a breakpoint hit at address 0040103B.

Address	Hex dump	Command	Comment	Registers (FPU)
00401030	64 00	PUSH 0		EDI 00000000
00401022	8B 00	CALL <JMP.&KERNEL32.Get	Module: KERNEL	ECX 0012FFB4
00401023	C6E20000	MOV EDI, EAX		EDX 7C90EB94
0040102E	5A	CALL 0040F2F4		EBP 77FD0000
00401032	24E20000	POP EDI		ESP 0012FFC0
00401034	F8E20000	CALL 0040F2F8		EBP 0012FFB8
00401035	60 00	PUSH 0	Collydt	EESI 00000000
00401036	14F00000	CALL 0040E054	Collydt	EDI 00000000
00401040	59	POP ECX		EIP 0040103B




This section displays three book covers. On the left is the cover of 'Introduction to Algorithms, Third Edition' by Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. In the middle is the cover of '리버스 엔지니어링' (Reverse Engineering) by 원도우 실행 파일 구조와 원리로 배우는, Volume 1: 파일 구조 편. On the right is the cover of 'Practical Malware Analysis' by Michael S. Stalder, featuring an illustration of a baby in a crib.

## 2. 성장 과정 - 입사 후 (17.07 ~ Present)

Pinned

Customize your pins

 **PuzzleCAM** Public ⋮


[ICIP2021] Puzzle-CAM: Improved localization via matching partial and full features.

● Python ☆ 150 🍷 43

 **RecurSeed\_and\_EdgePredictMix** Public ⋮

[Under Review] RecurSeed and EdgePredictMix: Single-stage learning is sufficient for Weakly-Supervised Semantic Segmentation

● Python ☆ 31 🍷 3

 **Tensorflow\_RetinaFace** Public ⋮

[Re-implementation] RetinaFace: Single-stage Dense Face Localisation in the Wild

● Python ☆ 26 🍷 11

 **Tensorflow\_Improving\_Pairwise\_Ranking\_for\_Multi-label\_Image\_Classification** Public ⋮


[Re-implementation] Improving Pairwise Ranking for Multi-label Image Classification (CVPR2017)

● Python ☆ 17 🍷 4

 **Tensorflow\_GIoU** Public ⋮

[Re-implementation] Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression (CVPR2019)

● Python ☆ 13 🍷 2

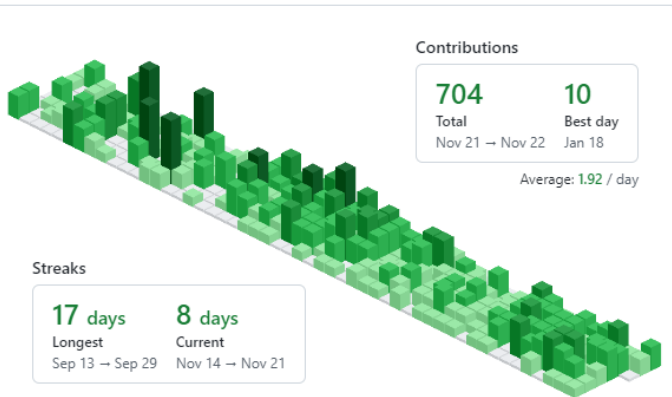
 **Tensorflow\_FixMatch** Public ⋮

[Re-implementation] FixMatch: Simplifying Semi-Supervised Learning with Consistency and Confidence

● Python ☆ 14 🍷 4

704 contributions in the last year

Contribution settings ▾ 2D 3D



## 2. 성장 과정 - 입사 후 (17.07 ~ Present)

localhost:5000

### SummaRizer with Keywords (SRK)

- CVPR2022  CVPR2021  CVPR2020  CVPR2019  CVPR2018  CVPR2017
- NIPS2021  NIPS2020  NIPS2019  NIPS2018  NIPS2017
- ICLR2022  ICLR2021  ICLR2020  ICLR2019  ICLR2018
- ICML2022  ICML2021  ICML2020  ICML2019  ICML2018  ICML2017
- ICCV2021  ICCV2019  ICCV2017
- ECCV2020  ECCV2018
- WACV2022  WACV2021  WACV2020

Keywords: weakly supervised, semantic, segmentation

PDF filename: 220809\_WSSS.pdf

Send

Created by Sanghyun Jo. © 2022.

For any issues, please contact [shjo.april@gmail.com](mailto:shjo.april@gmail.com).



#### Doodle It Yourself: Class Incremental Learning by Drawing a Few Sketches

The human visual system is remarkable in learning new visual concepts from just a few examples. This is precisely the goal behind few-shot class incremental learning (FSCIL), where the emphasis is additionally placed on ensuring the model does not suffer from "forgetting". In this paper, we push the boundary further for FSCIL by addressing two key questions that bottleneck its ubiquitous application (i) can the model learn from diverse modalities other than just photo (as humans do), and (ii) what if photos are not readily accessible (due to ethical and privacy constraints). Our key

innovati  
where th  
framewo  
graph at  
text in t



#### Distilling Image Dehazing With Heterogeneous Task Imitation

InProcee  
author  
title  
booktit  
month  
year  
pages  
}

State-of-the-art deep dehazing models are often difficult in training. Knowledge distillation paves a way to train a student network assisted by a teacher network. However, most knowledge distill methods are used for image classification and segmentation as well as object detection, and few investigate distilling image restoration and use different task for knowledge transfer. In this paper, we propose a knowledge-distill dehazing network which distills image dehazing with the heterogeneous task imitation. In our network, the teacher is an off-the-shelf auto-encoder network and is used for image reconstruction. The dehazing network is trained assisted by the teacher network with the process-oriented learning mechanism. The student network imitates the task of image reconstruction in the teacher network. Moreover, we design a spatial-weighted channel-attention residual block for the student image dehazing network to adaptively learn the content-aware channel level attention and pay more attention to the features for dense hazy regions reconstruction. To evaluate the effectiveness of the proposed method, we compare our method with several state-of-the-art methods on two synthetic and real-world datasets, as well as real hazy images.

https://o

```
InProceedings{Hong_2020_CVPR,  
author = {Hong, Ming and Xie, Yuan and Li, Cuihua and Qu, Yanyun},  
title = {Distilling Image Dehazing With Heterogeneous Task Imitation},  
booktitle = {Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)},  
month = {June},  
year = {2020}  
}
```

[https://openaccess.thecvf.com/content\\_CVPR\\_2020/papers/Hong\\_Distilling\\_Image\\_DeHazing\\_With\\_Heterogeneous\\_Task\\_Imitation\\_CVPR\\_2020\\_paper.pdf](https://openaccess.thecvf.com/content_CVPR_2020/papers/Hong_Distilling_Image_DeHazing_With_Heterogeneous_Task_Imitation_CVPR_2020_paper.pdf)

## 2. 성장 과정 - 입사 후 (17.07 ~ Present)

### Writing: 매 주 1회 이메일 작성 (Business)

Q1. What's your thought about the potential with the logistics business in line with reflecting this comment? ↵

- The growth rate of the logistics business will significantly decrease. The most governments have decided to mitigate the regulation of Covid-19 to protect existing companies from the pressure of inflation, causing a critical damage to the **delivery** industry. Moreover, the war occurred by Russia has accelerated the negative effect globally. For example, Amazon announced the **revocation** of contract and the sublease **to offset the current headwind**. As a result, numerous incidents explicitly explain why the logistics business has the low potential. ↵

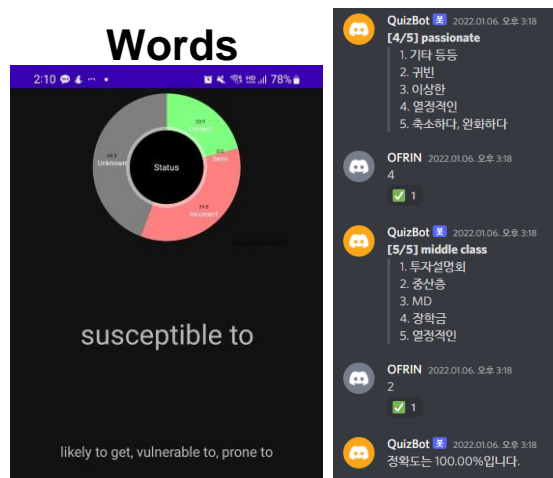
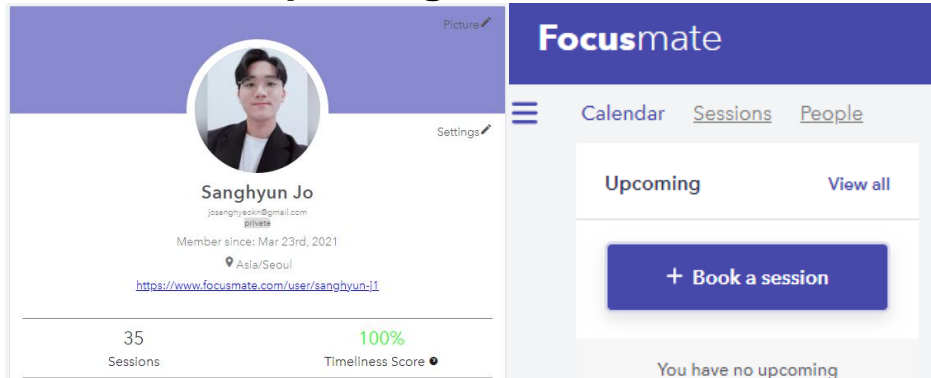
Q2. There are the puts and takes to lay groundwork for a new business. What if our company decides to jump in, which segment we need to lean into? ↵

- Our company has to **lean into** B2B logistics revenue stream with Korean Conglomerates Manufacturing. The conglomerates already **laid the groundwork** for manufacturing. Therefore, our company not only conducts **opportunities to experiment with** the superiority of a new business but also minimizes **the barrier to entry** a new business by leveraging existing manufacturing infrastructures. ↵

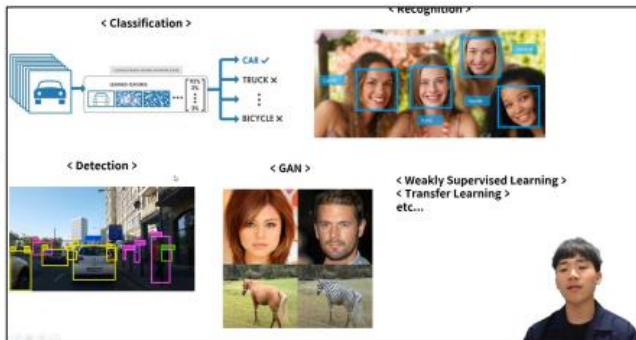
Q3. If we decided to avoid business risk in the logistics and invest in the logistics company by equity investment method investment, which domestic logistics company would be a promising option? ↵

- To my knowledge, CJ Express is the most promising option to avoid business **a risk** in the logistics. Samsung SDS relies on the logistics BPO as shown in the first figure. By contrast, a category of the revenue in CJ Express includes global, CL, construction, and delivery. The diversification of the revenue **is factored into minimizing** a risk due to an unstable market. ↵

### Speaking: Focusmate

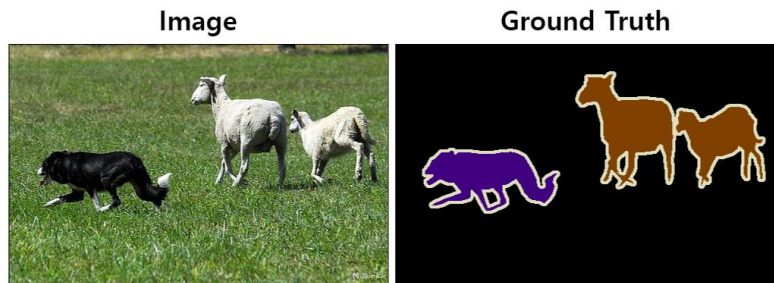


## 2. 성장 과정 - 입사 후 (17.07 ~ Present)



## 2. 성장 과정 - 입사 후 (17.07 ~ Present)

Fully-Supervised Semantic Segmentation



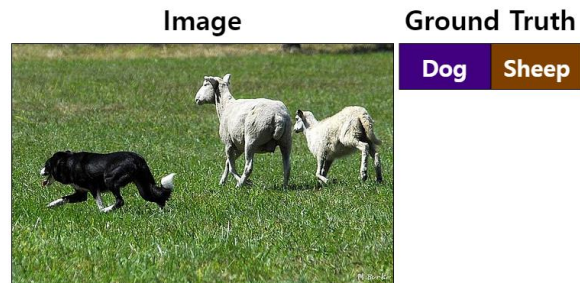
Training

Inference

FSSS Prediction



Weakly-Supervised Semantic Segmentation



Training

Inference

WSSS Prediction



# 3. 논문 – Puzzle-CAM (20.12 ~ 21.01)

2021  
01

SUN	MON	TUE	WED	THU	FRI	SAT
27	28	29	30	31	01 신년	02
03	04	05	06	07	08	09
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31						



The screenshot shows a LaTeX editor interface with a file explorer on the left containing files like 'images', 'IEEEbib.bst', 'refs.bib', 'spconf.sty', 'Template\_example.pdf', and 'Template.tex'. The main window displays the source code of a paper, with line numbers 96 to 133 visible. The code discusses 'Weakly-supervised semantic segmentation (WSSS)' and 'Puzzle-CAM'. On the right side, there is a 'Track changes' panel showing a list of updates with 'Re-open' and 'Delete' buttons for each entry.

# 3. 논문 – Puzzle-CAM (20.12 ~ 21.01)

**E-mail:** shjo.april@gvnetworks.com **Affiliation:** GYNetworks  
**Name:** Sanghyun Jo **URL:** https://github.com/OFRIN  
**Default Category:** eess.IV **Country:** Korea (South)  
**Groups:** eess and cs **Career Status:** Other

[Change User Information](#) | [Change Password](#) | [Change Email](#) | [Disable MathJax \(What is MathJax?\)](#)

## Article Submissions







 Update  Delete  Unsubmit

Identifier	Type	Title	Status	Actions	Expires
------------	------	-------	--------	---------	---------

[START NEW SUBMISSION](#)

Would you like to replace or take another action on one of your previously-announced articles?  
Use the link icons under "Articles You Own" instead.

## Articles You Own

 Replace  Withdraw  Cross list  Journal ref  Annotate  Link code & data

Identifier	Primary Category	Title	Actions	Author
2204.06754	cs.CV	RecurSeed and CertainMix for Weakly Supervised Semantic Segmentation	     	Y
2101.11253	cs.CV	Puzzle-CAM: Improved localization via matching partial and full features	     	Y

# 3. 논문 – Puzzle-CAM (20.12 ~ 21.01)

## Motivation

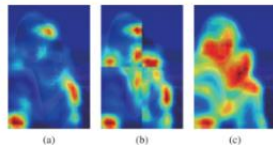


Fig 1. Illustration of the motivation of this work.

### Observation

- ✓ The conventional CAMs activates part of objects for classification (a)
- ✓ The merged CAMs from image patches scatter the attention (b)
- ✓ Puzzle-CAM suppresses the attention on discriminative region of the object (c)

### Process

1. Tiling an image to image patches to divide into the attention.
2. Merging the feature maps from the network to produce the reconstructed features.
3. Matching partial and full features with reconstructing regularization.

## Puzzle-CAM

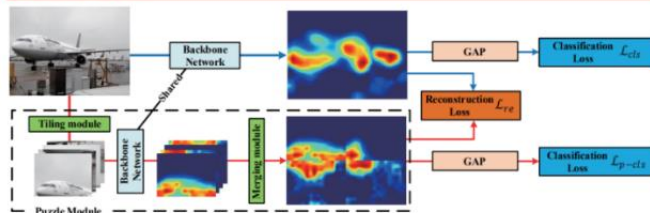


Fig 2. Illustration of the Puzzle-CAM architecture.

### Proposed Losses

$$L_{cls} = \ell_{cls}(\hat{Y}^s, Y)$$

$$L_{p-cls} = \ell_{cls}(\hat{Y}^{re}, Y)$$

$$L_{re} = \|A^s - A^{re}\|_1$$

$$L = L_{cls} + L_{p-cls} + \alpha L_{re}$$

- ✓ Ablation study: using different loss functions

$L_{cls}$	$L_{p-cls}$	$L_{re}$	mIoU (%)
✓			47.82
✓	✓		47.70
✓		✓	49.21
✓	✓	✓	<b>51.53</b>

## Quantitative Results

- ✓ Ablation study: using different backbones

Method	Backbone	CAM (%)	CAM +RW (%)	CAM+RW +dCRF (%)
AffinityNet [4]	ResNet-50	47.82	58.10	59.70
Puzzle-CAM	ResNet-50	51.53	64.16	64.70
Puzzle-CAM	ResNeSt-50	57.59	69.48	69.91
Puzzle-CAM	ResNeSt-101	61.85	71.92	72.46
Puzzle-CAM	ResNeSt-269	<b>62.45</b>	<b>74.14</b>	<b>74.67</b>

- ✓ Comparison with existing state-of-the-art methods

Method	Backbone	Supervision	val	test
AffinityNet [4]	Wide-ResNet-38	$\mathcal{I}$	61.7	63.7
DSRG [12]	ResNet-101	$\mathcal{I} + \mathcal{S}$	61.4	63.2
SeeNet [13]	ResNet-101	$\mathcal{I} + \mathcal{S}$	63.1	62.8
IRNet [4]	ResNet-50	$\mathcal{I}$	63.5	64.8
FickleNet [6]	ResNet-101	$\mathcal{I} + \mathcal{S}$	64.9	65.3
ICD [17]	ResNet-101	$\mathcal{I}$	64.1	64.3
SEAM [5]	Wide-ResNet-38	$\mathcal{I}$	64.5	65.7
Ours (Puzzle-CAM)	ResNeSt-101	$\mathcal{I}$	66.9	67.7
Ours (Puzzle-CAM)	ResNeSt-269	$\mathcal{I}$	<b>71.9</b>	<b>72.2</b>

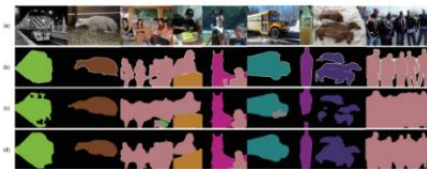
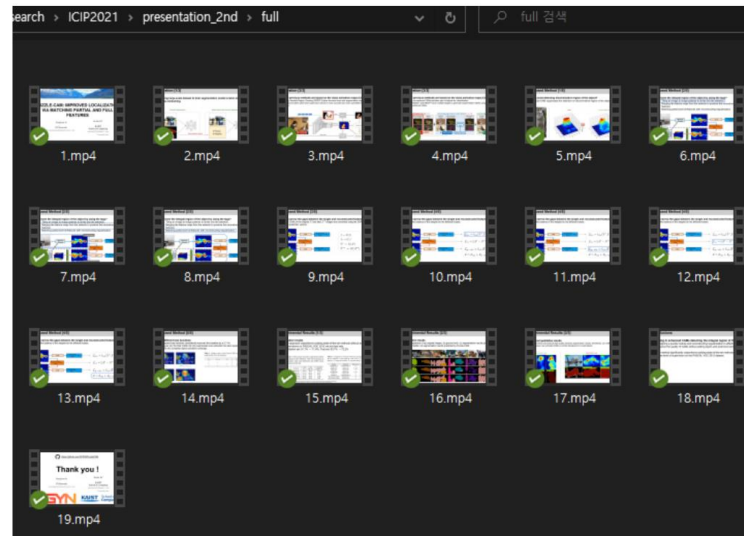
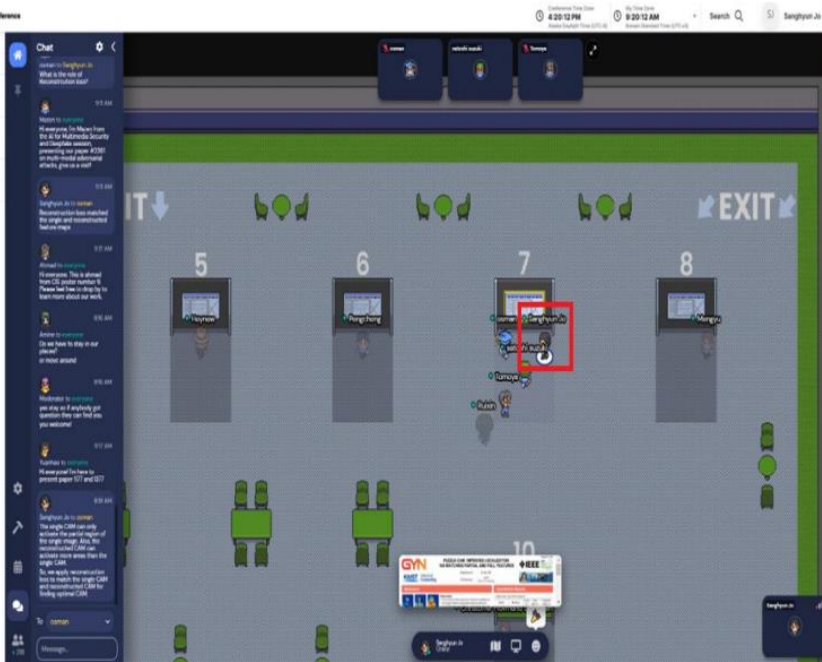


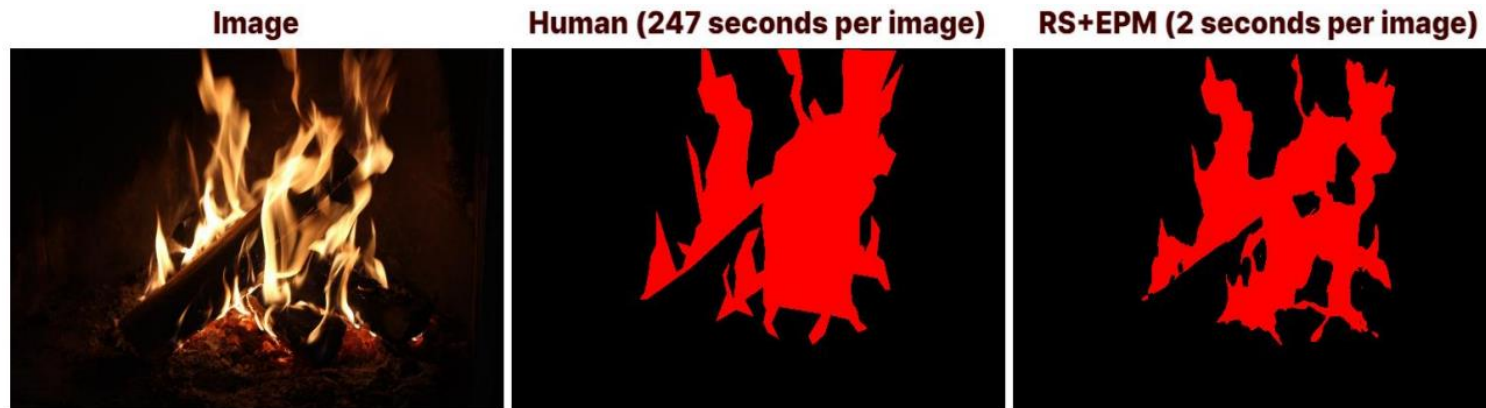
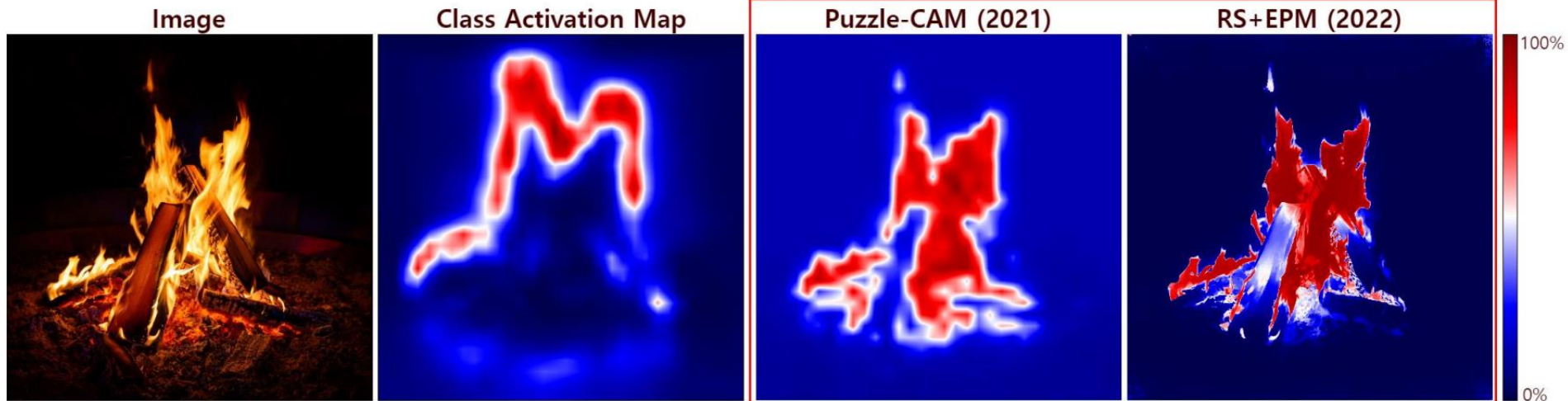
Fig 3. (a): original images. (b): ground truth. (c): segmentation results predicted by AffinityNet. (d): segmentation results predicted by Puzzle-CAM.



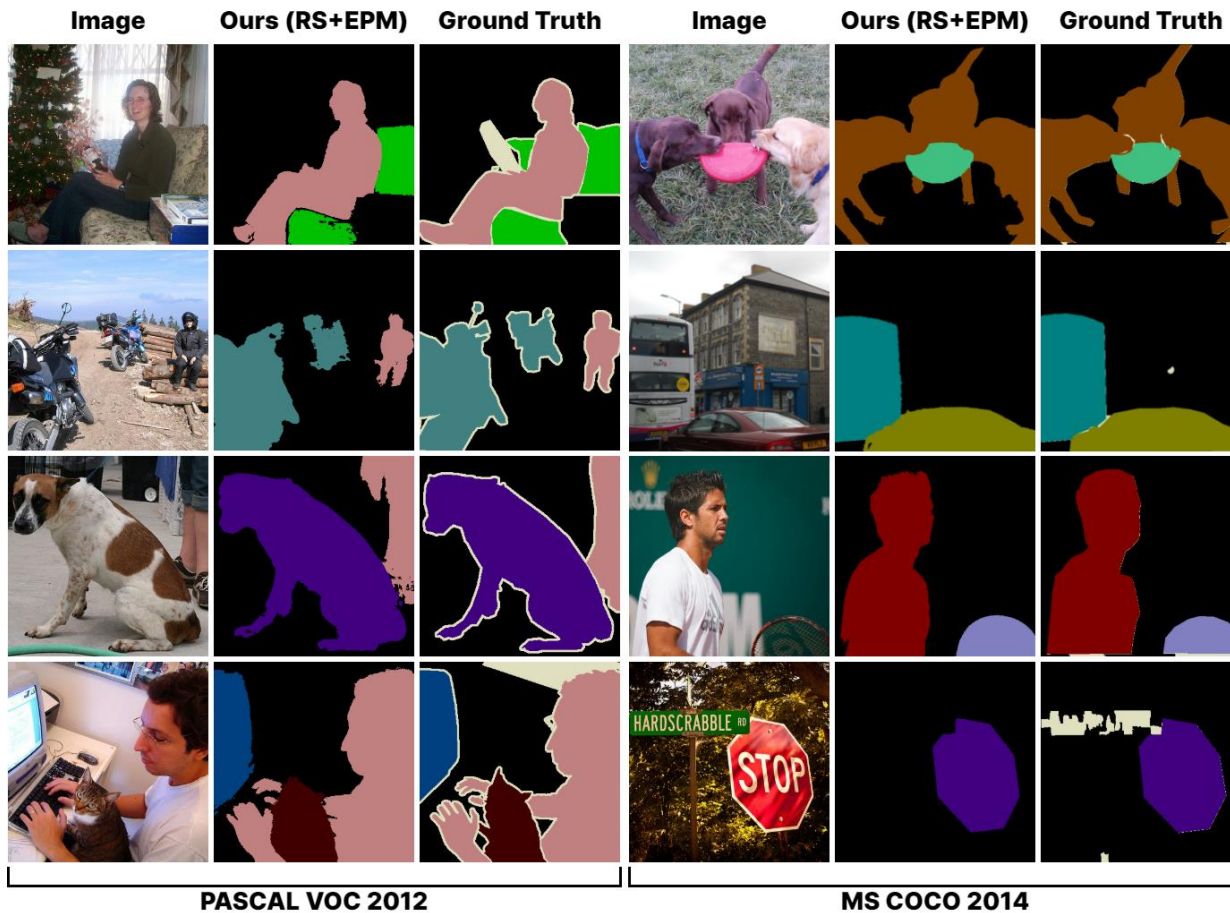
# 3. 논문 – Puzzle-CAM (20.12 ~ 21.01)



### 3. 논문 – RS+EPM (21.11 ~ Present)



### 3. 논문 – RS+EPM (21.11 ~ Present)



### 3. 논문 – RS+EPM (21.11 ~ Present)

**Table 2:** Performance comparison of WSSS methods in terms of mIoU (%) on PASCAL VOC 2012 and COCO 2014. \* and † indicate the backbone of VGG-16 and ResNet-50, respectively. Sup., supervision;  $\mathcal{I}$ , image-level class labels;  $\mathcal{S}$ , the saliency supervision;  $\mathcal{D}$ , the external dataset.

Method	Backbone	Sup.	VOC		COCO
			<i>val</i>	<i>test</i>	
Single stage:					
EM ICCV'15 [17]	VGG16	$\mathcal{I}$	38.2	39.6	-
RRM AAAI'20 [19]	WR38	$\mathcal{I}$	62.6	62.9	-
SSSS CVPR'20 [20]	WR38	$\mathcal{I}$	62.7	64.3	-
AFA CVPR'22 [43]	MiT-B1	$\mathcal{I}$	66.0	66.3	38.9
Ours (single-stage, RS)	R50	$\mathcal{I}$	66.5	67.9	40.0
Ours (single-stage, RS+EPM)	R50	$\mathcal{I}$	<b>69.5</b>	<b>70.6</b>	<b>42.2</b>
PSA CVPR'18 [5]	WR38	$\mathcal{I}$	61.7	63.7	-
IRNet CVPR'19 [31]	R50	$\mathcal{I}$	63.5	64.8	-
SEAM CVPR'20 [29]	WR38	$\mathcal{I}$	64.5	65.7	31.9
AdvCAM CVPR'21 [7]	R101	$\mathcal{I}$	68.1	68.0	-
CSE ICCV'21 [54]	WR38	$\mathcal{I}$	68.4	68.2	36.4
CPN ICCV'21 [30]	WR38	$\mathcal{I}$	67.8	68.5	-
RIB NIPS'21 [32]	R101	$\mathcal{I}$	68.3	68.6	43.8
ReCAM CVPR'22 [55]	R101	$\mathcal{I}$	68.5	68.4	-
ADEHE CVPR'22 [41]	R101	$\mathcal{I}$	68.6	68.9	-
AMR AAAI'22 [56]	R101	$\mathcal{I}$	68.8	69.1	-
URN AAAI'22 [57]	R101	$\mathcal{I}$	69.5	69.7	40.7
SIPE CVPR'22 [13]	R101	$\mathcal{I}$	68.8	69.7	40.6
AMN CVPR'22 [36]	R101	$\mathcal{I}$	69.5	69.6	44.7
MCTformer CVPR'22 [39]	WR38	$\mathcal{I}$	71.9	71.6	42.0
SANCE CVPR'22 [42]	R101	$\mathcal{I}$	70.9	72.2	44.7†
Ours (multi-stage, RS)	R101	$\mathcal{I}$	72.8	72.8	45.8
Ours (multi-stage, RS+EPM)	R101	$\mathcal{I}$	<b>74.4</b>	<b>73.6</b>	<b>46.4</b>

# 3. 논문 – RS+EPM (21.11 ~ Present)

Computer Vision

## Weakly-Supervised Semantic Segmentation

84 papers with code • 3 benchmarks • 5 datasets

The semantic segmentation task is to assign a label from a label set to each pixel in an image. In the case of fully supervised setting, the dataset consists of images and their corresponding pixel-level class-specific annotations (expensive pixel-level annotations). However, in the weakly-supervised setting, the dataset consists of images and corresponding annotations that are relatively easy to obtain, such as tags/labels of objects present in the image.

(Image credit: [Weakly-Supervised Semantic Segmentation Network with Deep Seeded Region Growing](#))

### Benchmarks

Add a Result

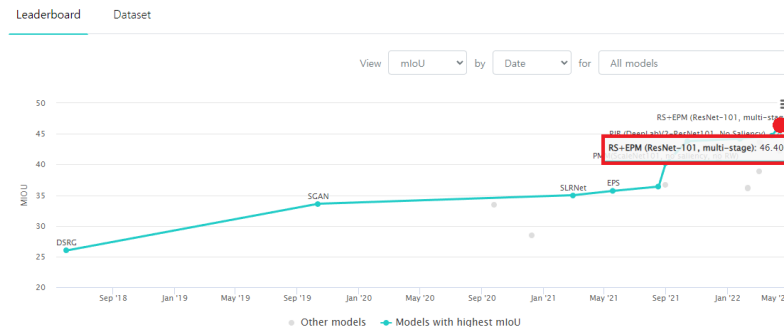
These leaderboards are used to track progress in Weakly-Supervised Semantic Segmentation

Trend	Dataset	Best Model	Paper	Code	Compare
	PASCAL VOC 2012 val	🏆 RS+EPM (ResNet-101, multi-stage)			See all
	PASCAL VOC 2012 test	🏆 RS+EPM (ResNet-101, multi-stage)			See all
	COCO 2014 val	🏆 RS+EPM (ResNet-101, multi-stage)			See all

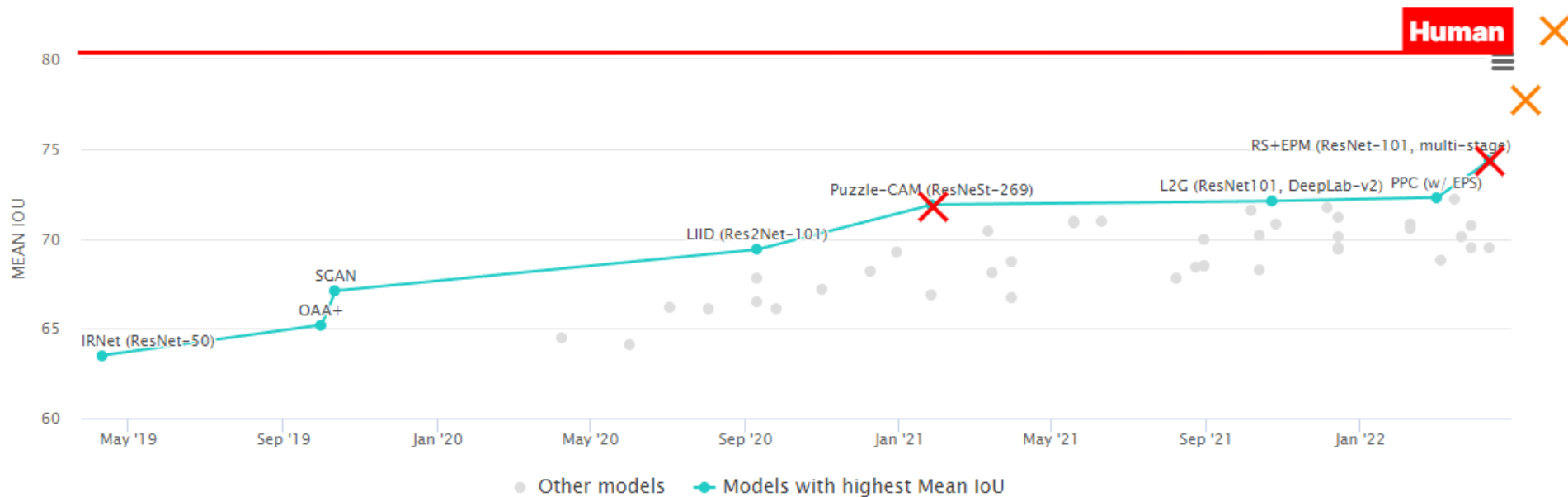
## Weakly-Supervised Semantic Segmentation on PASCAL VOC 2012 val



## Weakly-Supervised Semantic Segmentation on COCO 2014 val



### 3. 논문 – RS+EPM (21.11 ~ Present)



# 4. AI 발전 방향 예측

## 1. Content Generation (Image/Video/Audio)

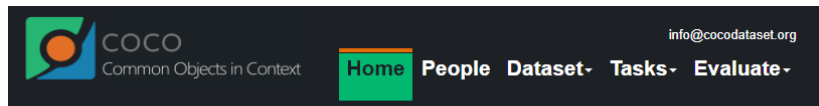
✓ DALLE-2, Stable Diffusion (CVPR 2022), Novel AI



## Visual Object Classes Challenge 2012 (VOC2012)



[click on an image to see the annotation]



### News

- We are pleased to announce the [LVIS 2021 Challenge and Workshop](#) to be held at ICCV.
- Please note that there will not be a COCO 2021 Challenge, instead, we encourage people to participate in the LVIS 2021 Challenge.
- We have partnered with the team behind the open-source tool [FiftyOne](#) to make it easier to download, visualize, and evaluate COCO
- [FiftyOne](#) is an open-source tool facilitating visualization and access to COCO data resources and serves as an evaluation tool for model analysis on COCO.

### What is COCO?



COCO is a large-scale object detection, segmentation, and captioning dataset. COCO has several features:

- ✓ Object segmentation
- ✓ Recognition in context
- ✓ Superpixel stuff segmentation
- ✓ 330K images (>200K labeled)
- ✓ 1.5 million object instances
- ✓ 80 object categories
- ✓ 91 stuff categories
- ✓ 5 captions per image
- ✓ 250,000 people with keypoints

### Collaborators

Tsung-Yi Lin Google Brain  
Genevieve Patterson MSR, Trash TV  
Matteo R. Ronchi Caltech  
Yin Cui Google  
Michael Maire TTI-Chicago  
Serge Belongie Cornell Tech  
Lubomir Bourdev WaveOne, Inc.  
Ross Girshick FAIR  
James Hays Georgia Tech  
Pietro Perona Caltech  
Deva Ramanan CMU  
Larry Zitnick FAIR  
Piotr Dollár FAIR

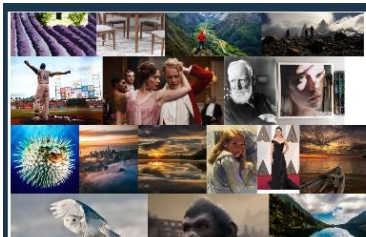
### Sponsors



# 4. AI 발전 방향 예측

## 1. Content Generation (Image/Video/Audio)

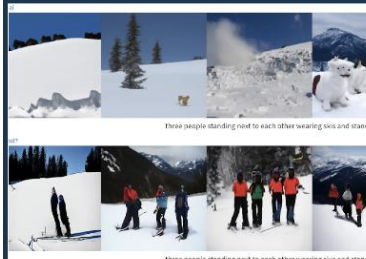
✓ DALLE-2, Stable Diffusion (CVPR 2022), Novel AI



### LAION-Aesthetics

by: Christoph Schuhmann, 8 Aug, 2022

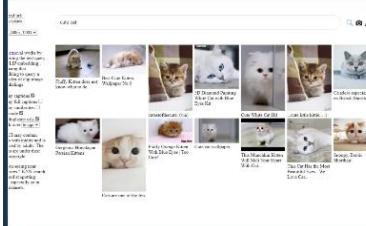
We present LAION-Aesthetics, several collections of subsets from LAION 5B with high visual quality. To create LAION-Aesthetics we trained several lightweight models that predict the rating people gave when they were asked "How much do you like this image on a scale from 1 to 10?". LAION-Aesthetics ...



### LAION-5B: A NEW ERA OF OPEN LARGE-SCALE MULTI-MODAL DATASETS

by: Romain Beaumont, 3 Mar, 2022

We present a dataset of 5,85 billion CLIP-filtered image-text pairs, 14x bigger than LAION-400M, previously the biggest openly accessible image-text dataset in the world. Authors: Christoph Schuhmann, Richard Vencu, Romain Beaumont, Theo Coombes, Cade Gordon, Aarush Katta, Robert Kaczmarczyk, Jenia ...



### LAION-400-MILLION OPEN DATASET

by: Christoph Schuhmann, 8 Aug, 2021

We present LAION-400M: 400M English (image, text) pairs Concept and Content The LAION-400M dataset is entirely openly, freely accessible. WARNING: be aware that this large-scale dataset is non-curated. It was built for research purposes to enable testing model training on larger scale for broad rese...



### Laion coco: 600M synthetic captions from Laion2B-en

by: Christoph Schuhmann, Andreas Köpf, Richard Vencu, Theo Coombes, Romain Beaumont, 9 Sep, 2022

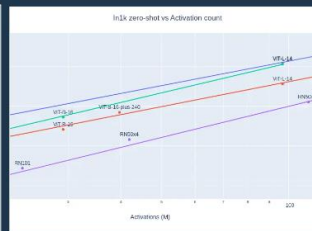
Author: Christoph Schuhmann, Andreas Köpf, Theo Coombes, Richard Vencu, Benjamin Trom, Romain Beaumont We present LAION-COCO, the world's largest dataset of 600M generated high-quality captions for publicly available web-images Laion5B has five billion natural captions. They provide a lot of infor...



### Laion translated: 3B captions translated to English from laion5B

by: Marianna Nezhurina, Romain Beaumont, Richard Vencu and Christoph Schuhmann, 9 Sep, 2022

Author: Marianna Nezhurina Romain Beaumont Richard Vencu Christoph Schuhmann Laion5B dataset was automatically collected from a section of the human web (common crawl). Can models generate different and interesting data compared to what humans write? That's a question we are interested in investigat...



### Large scale openCLIP: L/14, H/14 and g/14 trained on LAION-2B

by: Romain Beaumont, 9 Sep, 2022

We trained three large CLIP models with OpenCLIP: ViT-L/14, ViT-H/14 and ViT-g/14 (ViT-g/14 was trained only for about a third the epochs compared to the rest). The H/14 model achieves 78.0% zero shot top-1 accuracy on ImageNet and 73.4% on zero-shot image retrieval at Recall@5 on MS COCO. As of Sep...



# 4. AI 발전 방향 예측

## 1. Content Generation (Image/Video/Audio)

✓ DALLE-2, Stable Diffusion (CVPR 2022), Novel AI



# 4. AI 발전 방향 예측

## 1. Content Generation (Image/Video/Audio)

✓ DALLE-2, Stable Diffusion (CVPR 2022), Novel AI



Input images



in the Acropolis



swimming



sleeping



in a doghouse



in a bucket



getting a haircut



Input images



worn by a bear



in the jungle



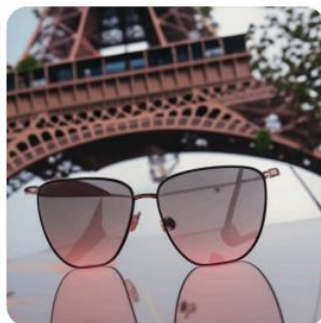
on red fabric



at Mt. Fuji



on top of snow



with Eiffel Tower

# 4. AI 발전 방향 예측

## 1. Content Generation (Image/Video/Audio)

✓ DALLE-2, Stable Diffusion (CVPR 2022), Imagen, Novel AI

Input Image



Edited Image



Target Text:

“A bird spreading wings”

Input Image



Edited Image



“A person giving the thumbs up”

Input Image



Edited Image



“A goat jumping over a cat”



Target Text:



“A sitting dog”



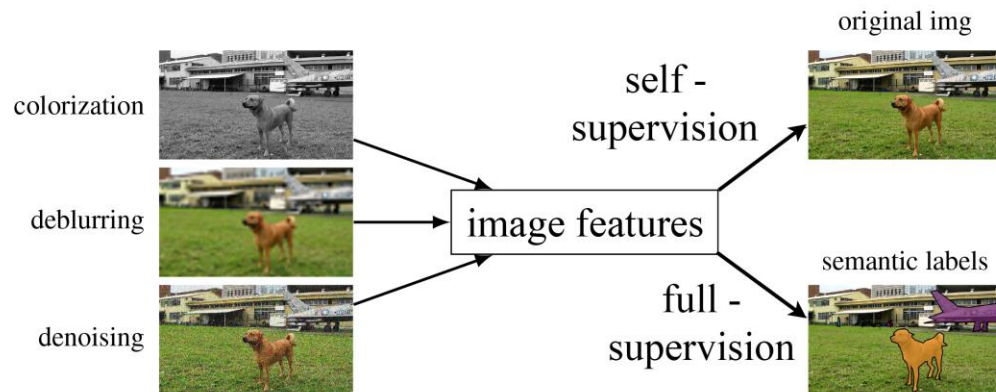
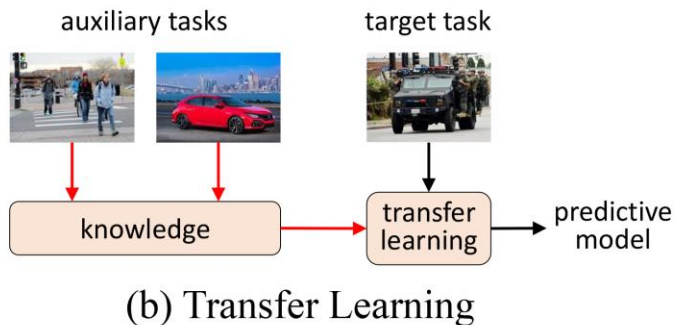
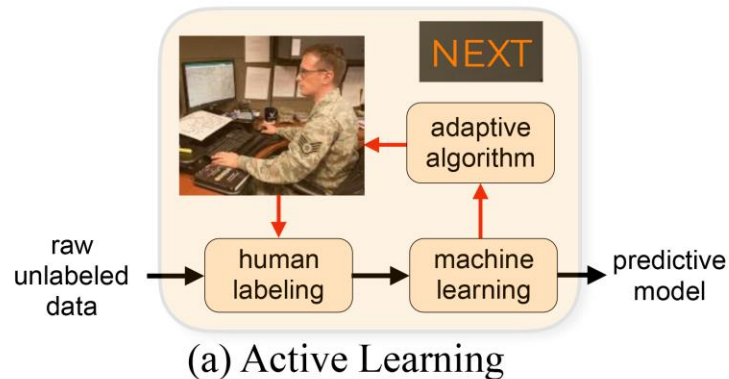
“Two kissing parrots”



“A children’s drawing of a waterfall”

# 4. AI 발전 방향 예측

## 2. Data-efficient Training



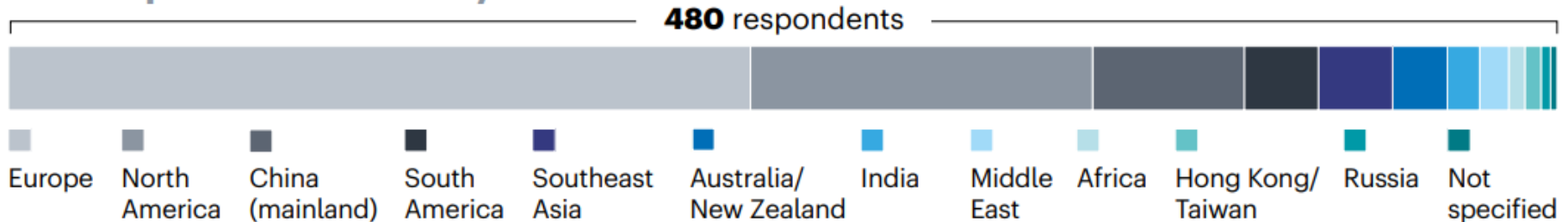
## 4. AI 발전 방향 예측

### 3. Ethical Issues

# FACIAL RECOGNITION: A SURVEY ON ETHICS

*Nature* surveyed\* nearly 500 researchers who work in facial recognition, computer vision and artificial intelligence about ethical issues relating to facial-recognition research. They are split on whether certain types of this research are ethically problematic and what should be done about concerns.

#### Who responded to the survey?



# 4. AI 발전 방향 예측

## 3. Ethical Issues

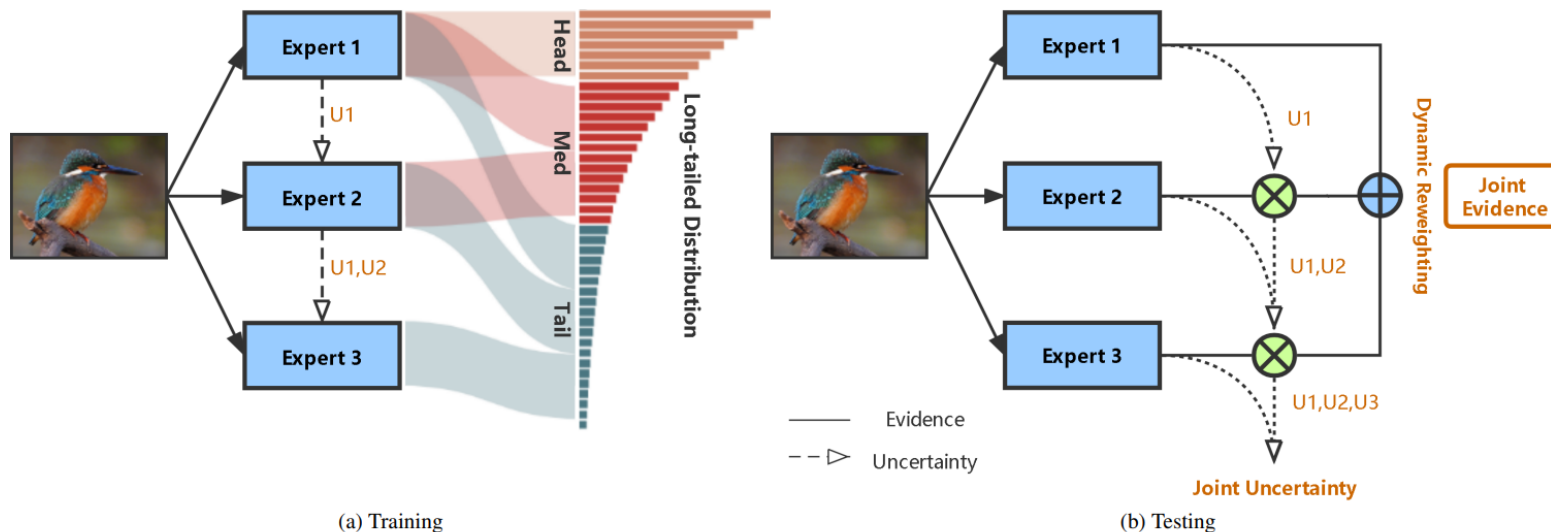


Figure 1. Overview of the proposed Trustworthy Long-Tailed Classification (TLC).  $U_1$ ,  $U_2$  and  $U_3$  are the uncertainties of expert 1, 2 and 3 respectively. In training (a), we provide an example of collaborating in different class groups for multiple experts. TLC dynamically assigns averagely more experts to the samples in tail classes than those in head classes. This assignment is achieved automatically by identifying hard samples with uncertainty. In testing (b), the joint uncertainty is formed with the Dempster's rule, and the joint evidence is obtained by uncertainty-based dynamic reweighting.

# 5. 채용

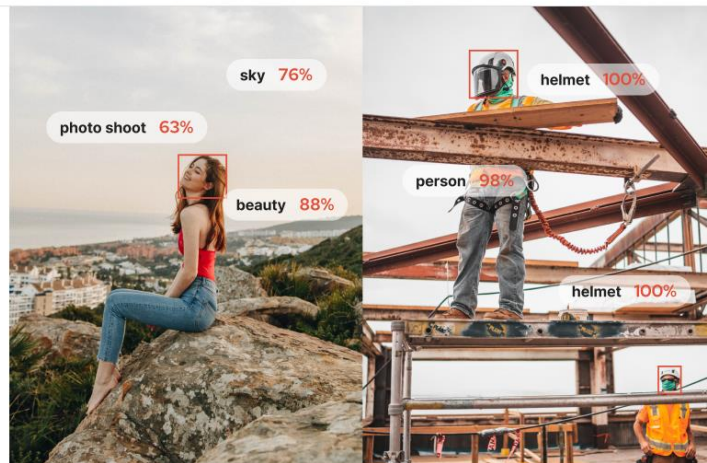


OGQ GYN DEEP MANAGER

## OGQ GYN Deep Manager

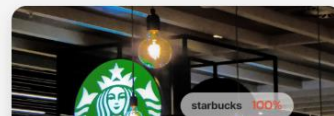
OGQ GYN의 AI 기술 솔루션  
딥매니저를 직접 체험해 보세요.

딥매니저는 OGQ GYN의 딥러닝 기반의 이미지 인식  
기술을 활용한 AI 솔루션입니다.



### for Image

이미지를 분석해서 이미지에 알맞은 단어를 달아주고,  
이미지 내 객체에 대한 위치를 나타내는 APIs입니다.



# Q & A

## Puzzle-CAM



## RS+EPM

