JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Exemplar-FreeSOLO: Enhancing Unsupervised Instance Segmentation with Exemplars

Taoseef Ishtiak[1] , Qing En[1], Yuhong Guo[1,2]

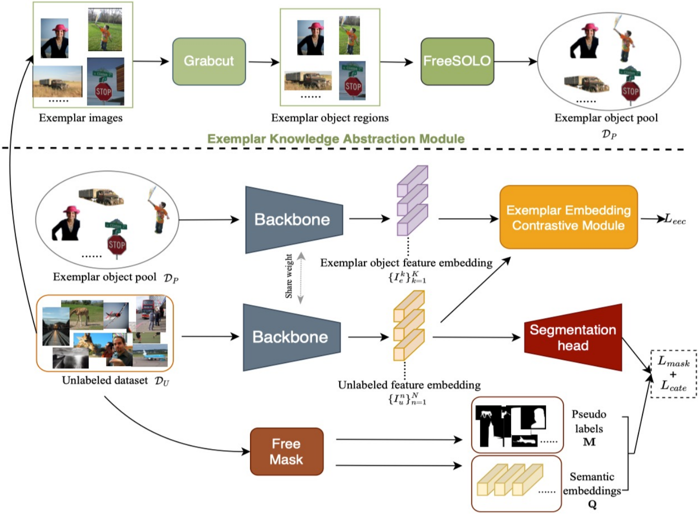
[1]Carleton University, Ottawa, Canada
[2]Canada CIFAR AI Chair, Amii, Canada

taoseefishtiak@cmail.carleton.ca
QingEn@cunet.carleton.ca
YuhongGuo@carleton.ca

WED-PM-292

Quick preview of Exemplar-FreeSOLO

- The novel Exemplar-FreeSOLO is proposed to enhance unsupervised instance segmentation through exemplar knowledge extraction and contrastive embedding learning.



Outline

- Research significance
- Motivation
- Method
- Experimental results
- Conclusions

Outline

- Research significance
- Motivation
- Method
- Experimental results
- Conclusions

Semantic segmentation

- Semantic or instance segmentation
    - Image semantic or instance segmentation
    - Video object or instance segmentation
- Recognizing, understanding what's in the image in pixel level

Semantic segmentation

Instance segmentation

Image

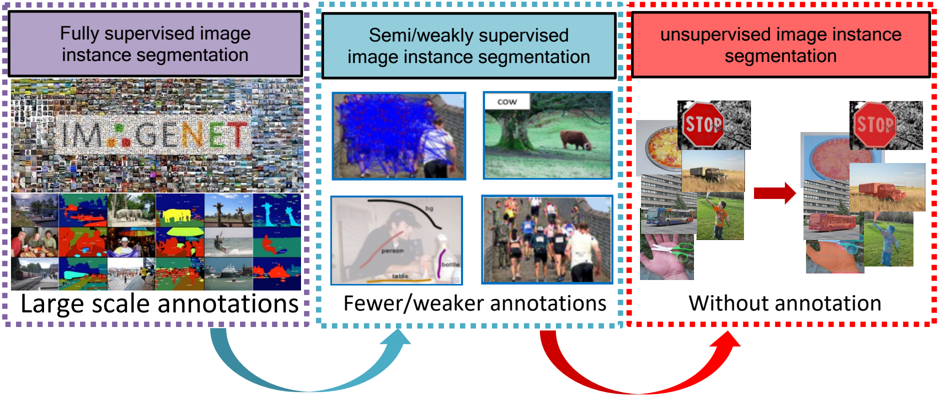
Semantic segmentation

Instance segmentation

Video

Unsupervised image instance segmentation

- Directly exploit existing unannotated images while being able to continuously upgrade the effectiveness of the segmentation models with incoming data

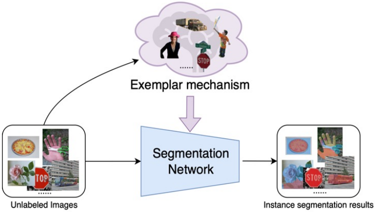


Outline

- Research significance
- Motivation
- Method
- Experimental results
- Conclusions

Motivations

- <u>Challenges</u>
- Heavily influenced by the noisy pseudo-labels
  - Prone to generating a large number of false positive regions.
- Difficult to learn discriminative information
  - Tends to lead to fragmentation problems.
- <u>Our solutions</u>
- Excavate information from unlabeled data through an exemplar mechanism
- Produces top-down knowledge guidance and enhances the discriminability of the segmentation model.

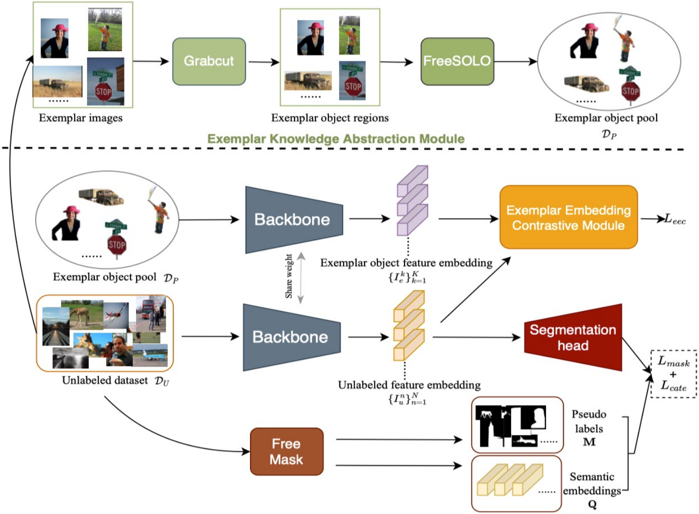


Outline

- Research significance and challenges
- Motivation
- Method
- Experimental results
- Conclusions

Overview of Exemplar-FreeSOLO

- The novel Exemplar-FreeSOLO is proposed to enhance unsupervised instance segmentation through exemplar knowledge extraction and contrastive embedding learning.

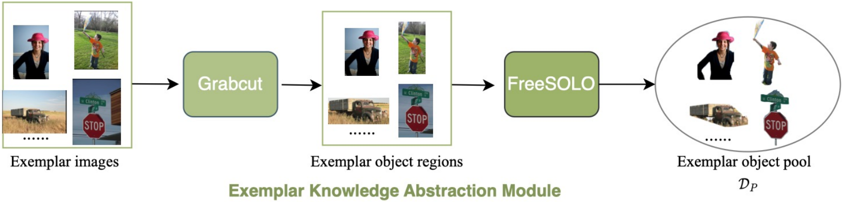


Exemplar Knowledge Abstraction Module

- Extract exemplar objects from the selected exemplar images and build an exemplar pool of objects:

$$\mathcal{D}_P = \{X_e^k\}_{k=1}^K$$

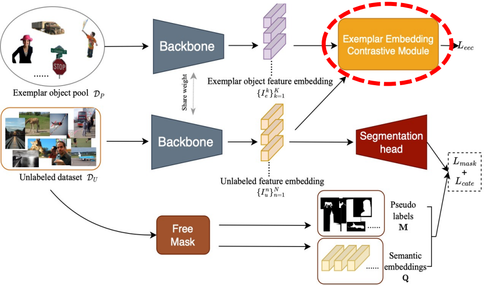
- Representatives or pivots for the corresponding hidden categories of the extracted objects.

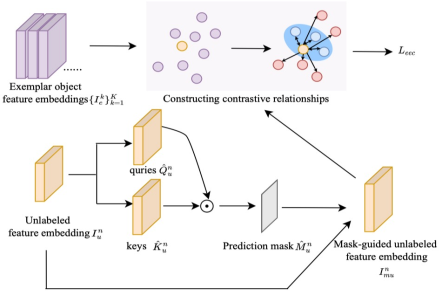


Architecture of the proposed Exemplar Knowledge Abstraction Module

Exemplar Embedding Contrastive Module

- Construct contrastive relationships between the exemplar embeddings and the unlabeled image embeddings
- Boost the discriminative capability of the instance segmentation network with an additional contrastive embedding loss during the self-training process





Architecture of the proposed Exemplar Embedding Contrastive Module

Exemplar Embedding Contrastive Module

- Select positive and negative samples:

$$I_{e,pos}^{n} \in \{I_e^k \in \mathcal{I}_P : \mathrm{sim}(I_{mu}^n, I_e^k) > \alpha\}$$

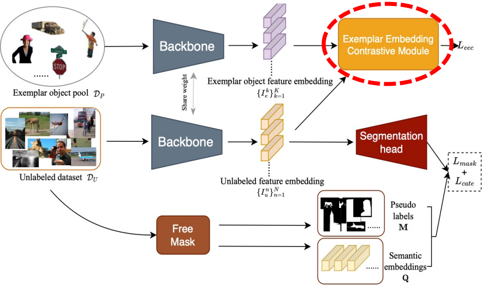
$$\mathcal{I}_{e,neg}^{n} = \{I_e^k \in \mathcal{I}_P : \mathrm{sim}(I_{mu}^n, I_e^k) < \beta\}$$

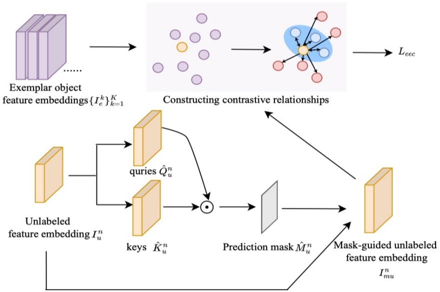
- The exemplar embedding contrastive loss is defined as follows:

$$L_{eec} = -\sum_n \log \frac{pos(n)}{pos(n) + neg(n)}$$

$$pos(n) = \exp(\langle I_{mu}^n, I_{e,pos}^n \rangle / \tau)$$

$$neg(n) = \frac{1}{|\mathcal{I}_{e,neg}^n|} \sum_{I_e \in \mathcal{I}_{e,neg}^n} \exp(\langle I_{mu}^n, I_e \rangle / \tau)$$





Architecture of the proposed Exemplar Embedding Contrastive Module

- The overall loss function:

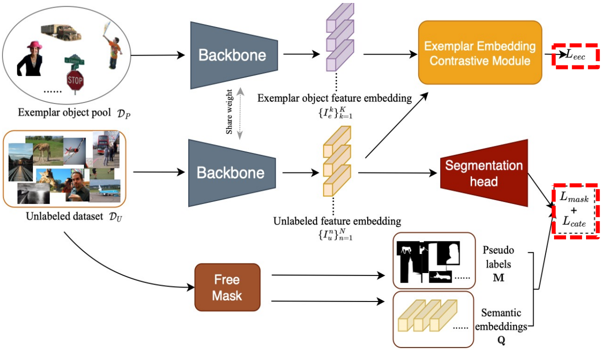
$$L_{total} = L_{mask} + L_{cate} + \lambda_{eec} L_{eec}$$

- The mask segmentation loss L_{mask} :

$$L_{\text{mask}} = \gamma L_{avg\_proj}(\mathbf{M}^*, \mathbf{M}) + L_{max\_proj}(\mathbf{M}^*, \mathbf{M}) + L_{pairwise}(\mathbf{M}^*)$$

- The category loss L_{cate}:

$$L_{cate} = L_{focal}(\mathbf{M}^*, \mathbf{M}) + \mu L_{sem}(\mathbf{Q}^*, \mathbf{Q})$$



Outline

- Research significance
- Motivation
- Method
- Experimental results
- Conclusions

Experimental results

- Class-agnostic instance segmentation results on MS COCO val2017: **4.4 AP improvements over FreeSOLO**.

Model	AP_{50}	AP_{75}	AP	AR_1	AR_{10}	AR_{100}
w/anns:						
MCG [1]	4.6	0.8	1.6	1.9	7.4	18.2
COB [31]	8.8	1.9	3.3	2.9	10.1	22.7
w/o anns:						
FreeSOLO [48]	9.8	2.9	4.0	4.1	10.5	12.7
Exemplar-FreeSOLO	13.2	6.3	8.4	7.3	15.8	15.5

- Unsupervised class-agnostic object detection results on MS COCO val2017: **Excellent performance with 12.6 AP.**

Model	AP_{50}	AP_{75}	AP	AR_1	AR_{10}	AR_{100}
UP-DETR [9]	0.0	0.0	0.0	0.0	0.0	0.4
SS [43]	0.5	0.1	0.2	0.2	1.5	10.9
DETReg [2]	3.1	0.6	1.0	0.6	3.6	12.7
FreeSOLO [48]	12.2	4.2	5.5	4.6	11.4	15.3
Exemplar-FreeSOLO	17.9	8.6	12.6	8.2	13.0	17.9

Experimental results

- Unsupervised instance segmentation results on UVO val: **4.4 AP improvement over FreeSOLO**.

Model	AP_{50}	AP_{75}	AP
w/anns:			
SOLOv2 [49] w/COCO	38.0	20.9	21.4
Mask R-CNN [16] w/COCO	31.0	14.2	15.9
SOLOv2 [49] w/LVIS	14.8	5.9	7.1
Mask R-CNN [16] w/LVIS	18.1	4.1	6.8
w/o anns:			
FreeSOLO [48]	12.7	3.0	4.8
Exemplar-FreeSOLO	14.2	7.3	9.2

- Multi-object discovery results on PASCAL VOC: **significantly outperforms the existing state-of-the-art methods**

Model	AP_{50}	AP_{75}	AP
Kim et al. [23]	9.5	-	2.5
DDT+ [53]	8.7	-	3.0
rOSD [44]	13.1	-	4.3
LOD [45]	13.9	-	4.5
LOST [40]	19.8	-	6.7
FreeSOLO [48]	24.5	7.2	10.2
Exemplar-FreeSOLO	26.8	8.2	12.6

Experimental results

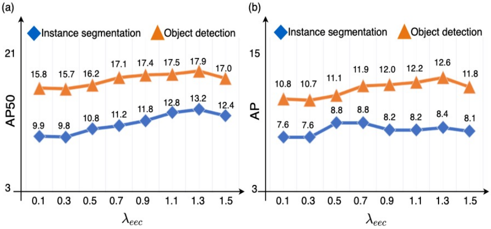
- The proposed exemplar mechanism contributes **2.2 and 4.4 AP improvements** for segmentation and detection respectively.

Model	AP_{50}	AP_{75}	AP	AP_s	AP_m	AP_l
Segmentation:						
vanilla FreeSOLO	9.8	2.9	4.0	3.6	13.5	10.8
Semi-super-box	11.2	3.5	5.8	3.0	7.7	19.2
Semi-super-mask	12.5	3.9	6.2	3.5	8.2	19.7
Exemplar-FreeSOLO	13.2	6.3	8.4	5.5	16.6	22.2
Detection:						
vanilla FreeSOLO	12.2	4.2	5.5	5.1	13.8	16.8
Semi-super-box	13.4	3.9	7.1	4.7	12.1	15.9
Semi-super-mask	14.3	4.7	8.2	4.8	12.6	16.7
Exemplar-FreeSOLO	17.9	8.6	12.6	6.8	15.9	19.9

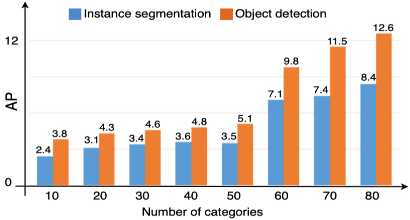
- Even with only one exemplar in each category, **very impressive results** can still be obtained.

	Num	AP_{50}	AP_{75}	AP	AP_s	AP_m	AP_l
Segmentation	1	13.2	6.3	8.4	5.5	16.6	22.2
	3	13.5	6.4	8.7	5.5	16.8	22.4
	5	13.6	6.6	8.8	5.7	16.8	22.5
	7	13.6	6.7	8.8	5.6	16.8	22.8
	9	13.8	6.9	8.9	5.8	16.8	23.2
Detection	1	17.9	8.6	12.6	6.8	15.9	19.9
	3	18.5	8.8	12.7	6.8	16.2	20.2
	5	18.8	9.1	12.8	7.2	16.5	20.6
	7	18.8	9.2	12.8	7.4	16.5	20.8
	9	18.8	9.5	13.0	7.7	16.7	20.9

- The best results for the segmentation and detection task are obtained with $\lambda_{eec} = 1.3$



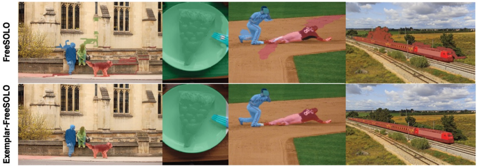
- The performance of the proposed framework gradually **improves** as the number of classes increases.



Experimental results

- Qualitative Results on MS COCO

- Exemplar-FreeSOLO can segment the corresponding targets in complex scenes more accurately than FreeSOLO



Outline

- Research significance
- Our proposed new experimental scenarios
- Method
- Conclusions

Conclusions

- The novel ExemplarFreeSOLO to fully exploring an effective exemplar mechanism for unsupervised instance segmentation.
- The exemplar knowledge abstraction module (EKA) to acquire beneficial top-down guidance knowledge.
- The exemplar embedding contrastive module (EEC) to enhance the discriminative capability of the instance segmentation network.
- Remarkable performance on three datasets.

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Thanks For Your Attention!

Taoseef Ishtiak[1], Qing En[1], Yuhong Guo[1,2]

[1]Carleton University, Ottawa, Canada
[2]Canada CIFAR AI Chair, Amii, Canada

taoseefishtiak@cmail.carleton.ca
QingEn@cunet.carleton.ca
YuhongGuo@carleton.ca