ENS | PSL
Inria INVENTEURS DU MONDE NUMÉRIQUE
PR[AI]RIE PaRis Artificial Intelligence Research InstitutE
valeo.ai
NYU Center for Data Science

Large-Scale Unsupervised Object Discovery

Huy V. Vo, Elena Sizikova, Cordelia Schmid, Patrick Pérez, Jean Ponce

NeuRIPS 2021





Unsupervised Object Discovery (UOD)

Extract useful information from images in form of object concepts.

Many guises:

Unsupervised Object Discovery (UOD)

Extract useful information from images in form of object concepts.

Many guises:

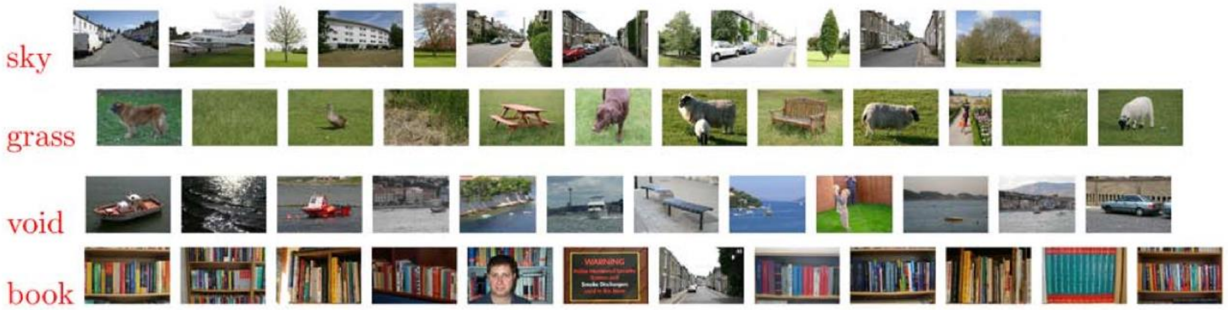
- Object categories discovery
  (Weber et al., CVPR'00, Grauman et al., CVPR'06, Russell et al., CVPR'06, Sivic et al., ICCV'05, Tang et al., CIVR'08, Tuytelaars et al., IJCV'10)

(Tuytelaars et al, IJCV'10)

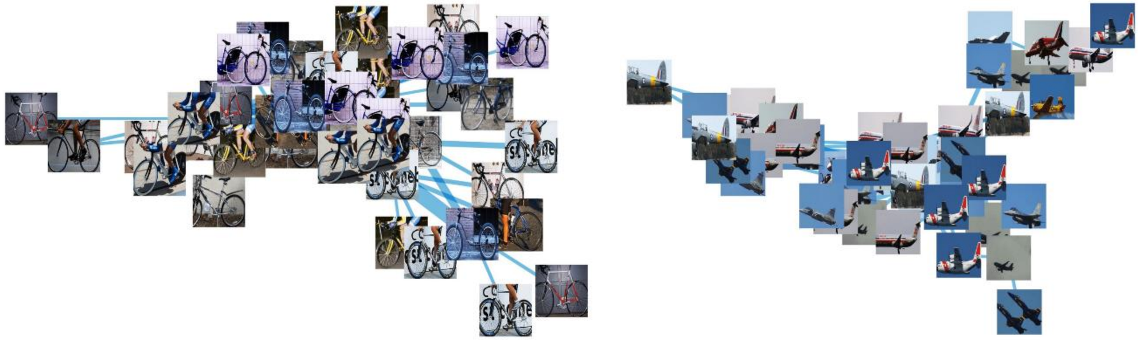
Unsupervised Object Discovery (UOD)

Extract useful information from images in form of object concepts.
Many guises:

- Object categories discovery
  (Weber et al., CVPR'00, Grauman et al., CVPR'06, Russell et al., CVPR'06, Sivic et al., ICCV'05, Tang et al., CIVR'08, Tuytelaars et al., IJCV'10)
- Finding image links and/or localize objects
  (Kim et al., NIPS'19, Cho et al., CVPR'15, Vo et al., CVPR'19 & ECCV'20)



(Tuytelaars et al, IJCV'10)



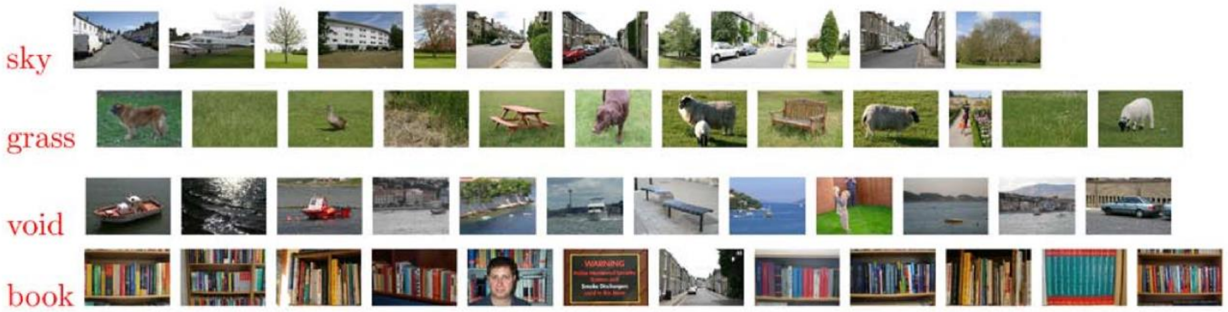
(Vo et al., CVPR'19)

Unsupervised Object Discovery (UOD)

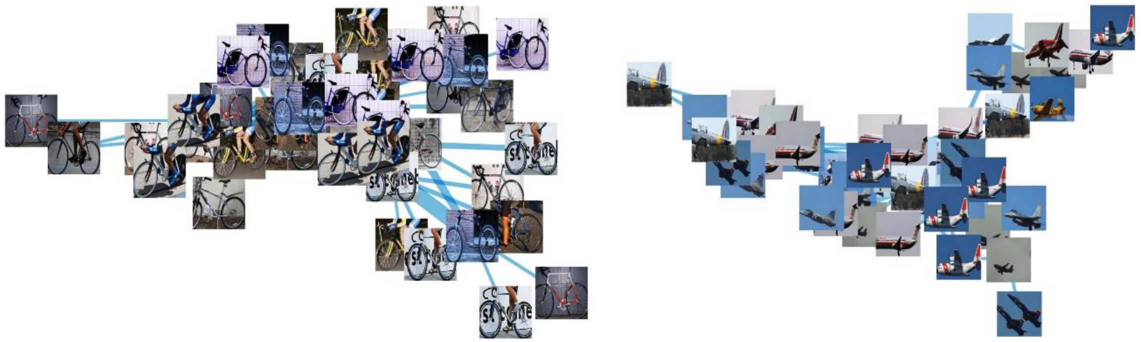
Extract useful information from images in form of object concepts.

Many guises:

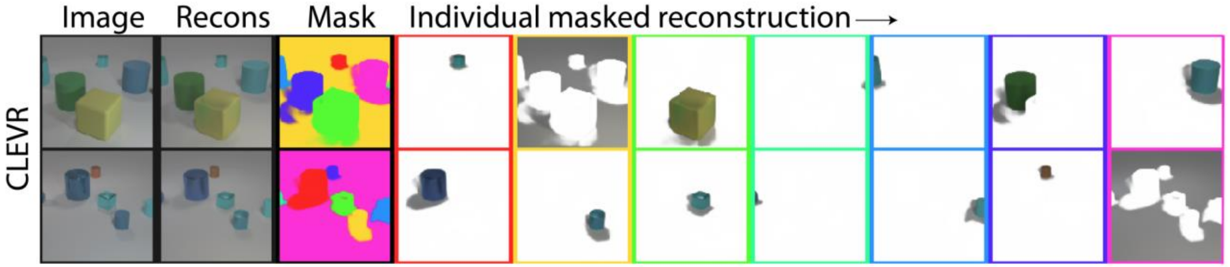
- Object categories discovery
  (Weber et al., CVPR'00, Grauman et al., CVPR'06, Russell et al., CVPR'06, Sivic et al., ICCV'05, Tang et al., CIVR'08, Tuytelaars et al., IJCV'10)
- Finding image links and/or localize objects
  (Kim et al., NIPS'19, Cho et al., CVPR'15, Vo et al., CVPR'19 & ECCV'20)
- Learn object representation and discover object masks
  (Burgess et al., 2019, Greff et al., ICML'19, Engelcke et al., ICLR'20, Locatello et al., NeuRIPS'20, Monnier et al., 2021)



(Tuytelaars et al, IJCV'10)



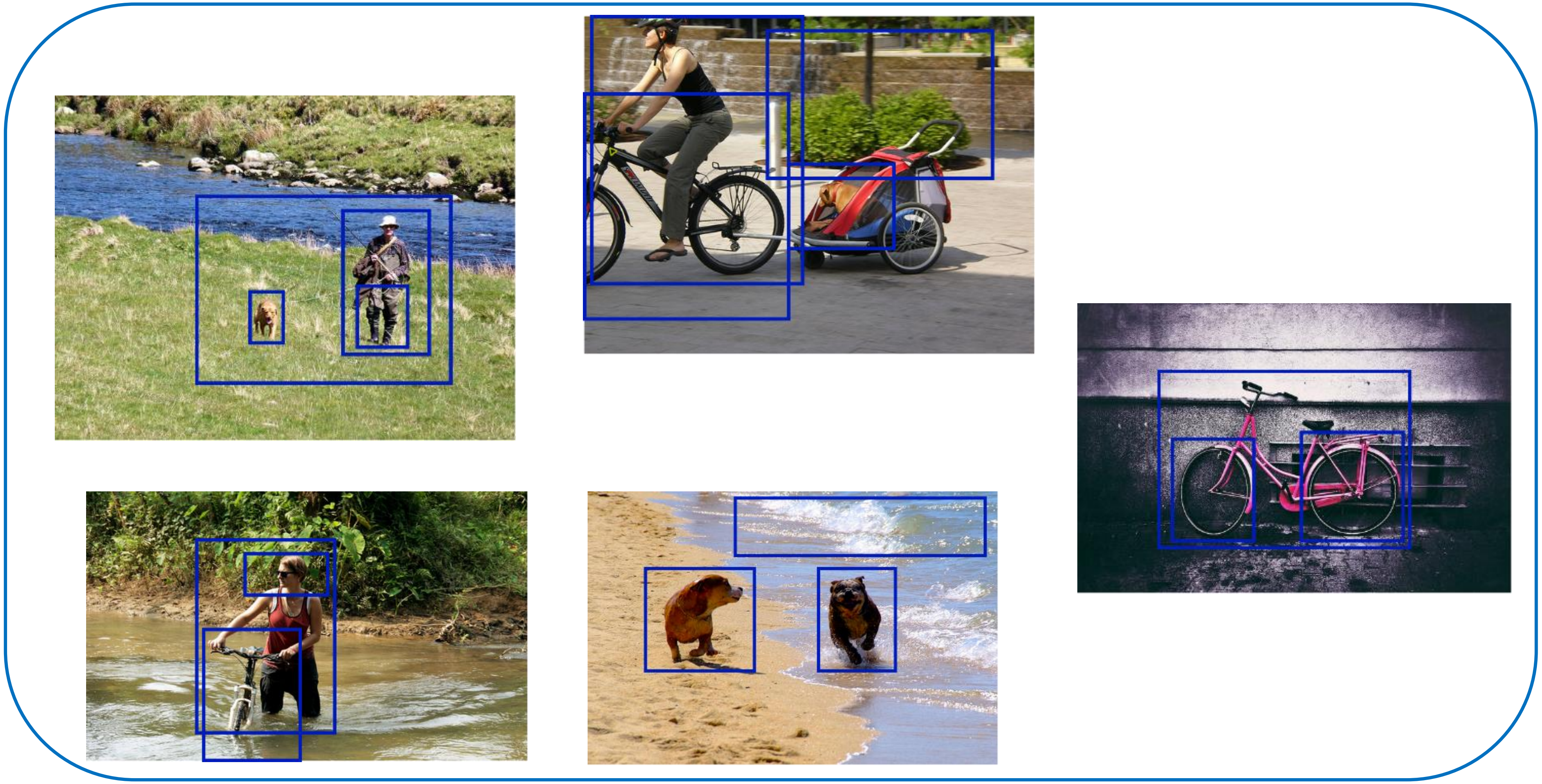
(Vo et al., CVPR'19)



(Greff et al., ICML'19)

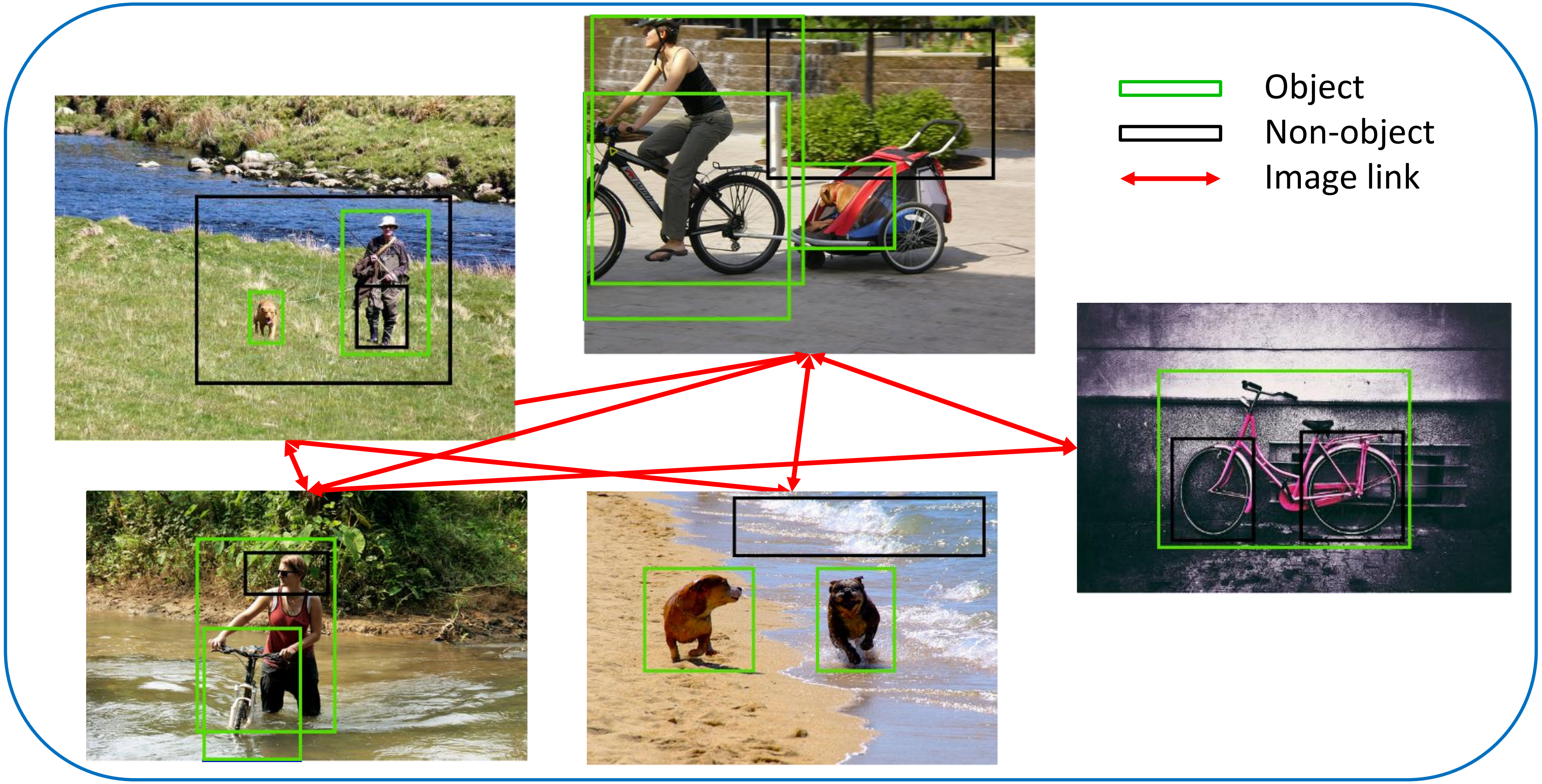
Prior work

Finding frequently appearing salient visual patterns: Analyse image similarities, region salient scores *(Kim et al., NeuRIPS'09; Cho et al., CVPR'15; Vo et al., CVPR'19, ECCV'20)*.



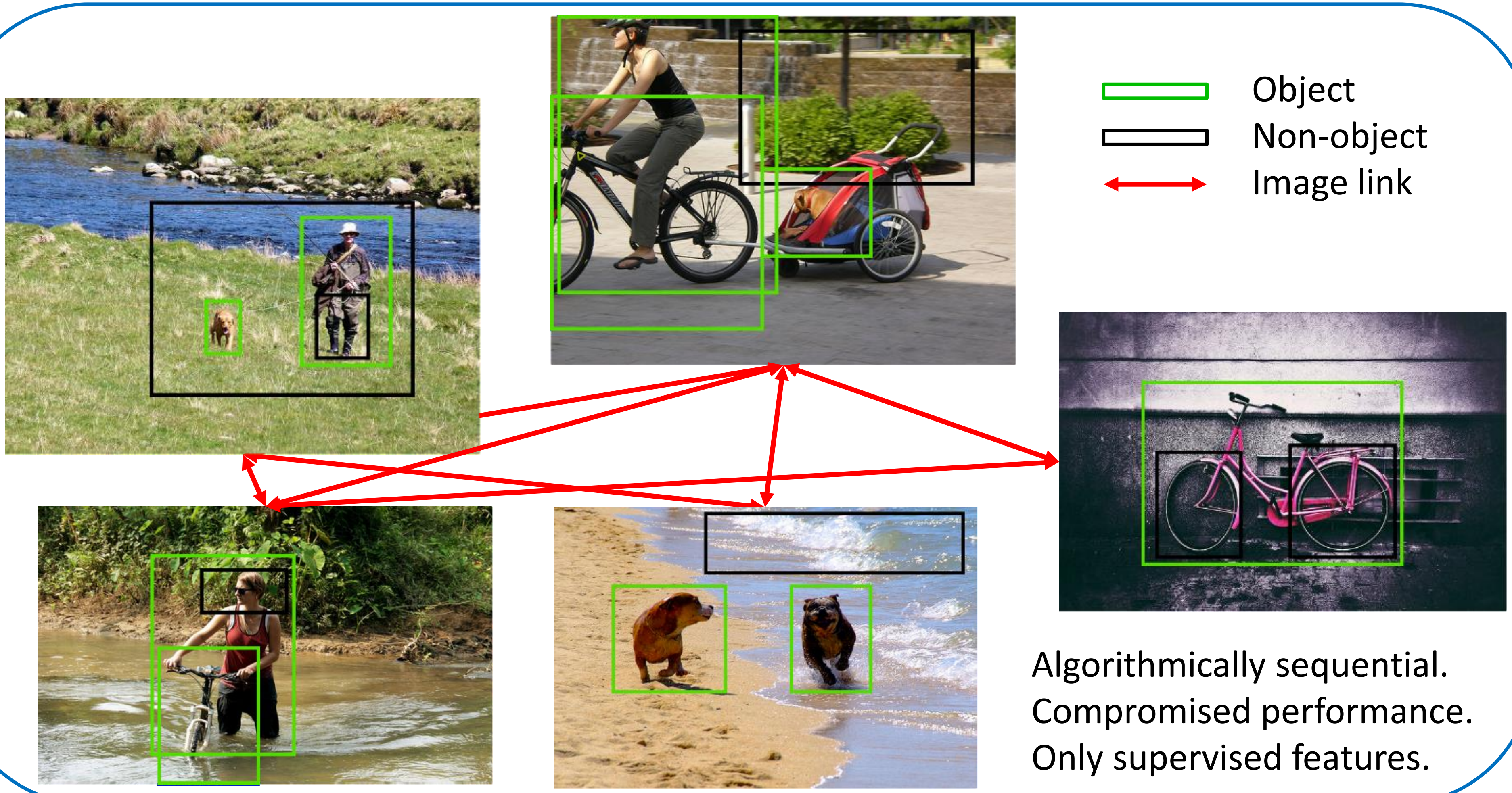
Prior work

Finding frequently appearing salient visual patterns: Analyse image similarities, region salient scores *(Kim et al., NeuRIPS'09; Cho et al., CVPR'15; Vo et al., CVPR'19, ECCV'20)*.



Prior work

Finding frequently appearing salient visual patterns: Analyse image similarities, region salient scores *(Kim et al., NeuRIPS'09; Cho et al., CVPR'15; Vo et al., CVPR'19, ECCV'20)*.



Algorithmically sequential.
Compromised performance.
Only supervised features.

Proposed Approach

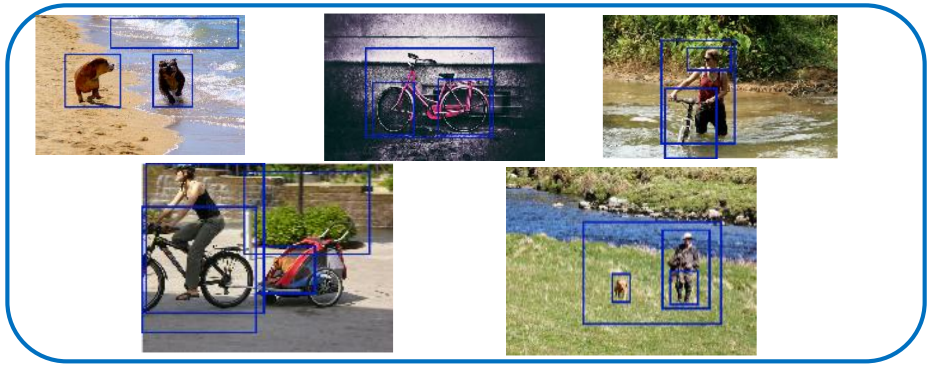
UOD can naturally be formulated as a ranking problem.

Completely unsupervised pipeline.

Proposed Approach

UOD can naturally be formulated as a ranking problem.

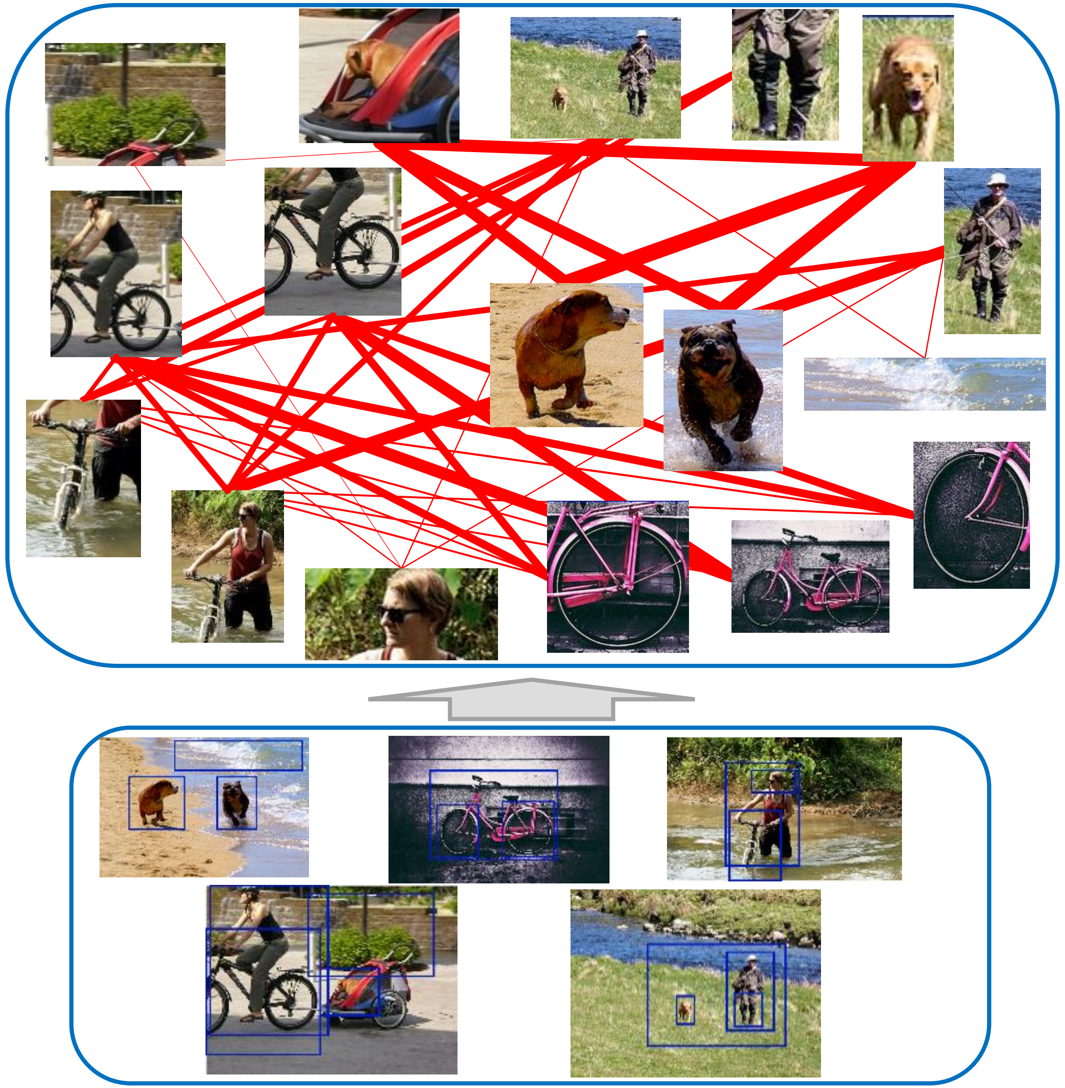
Completely unsupervised pipeline.



Proposed Approach

UOD can naturally be formulated as a ranking problem.

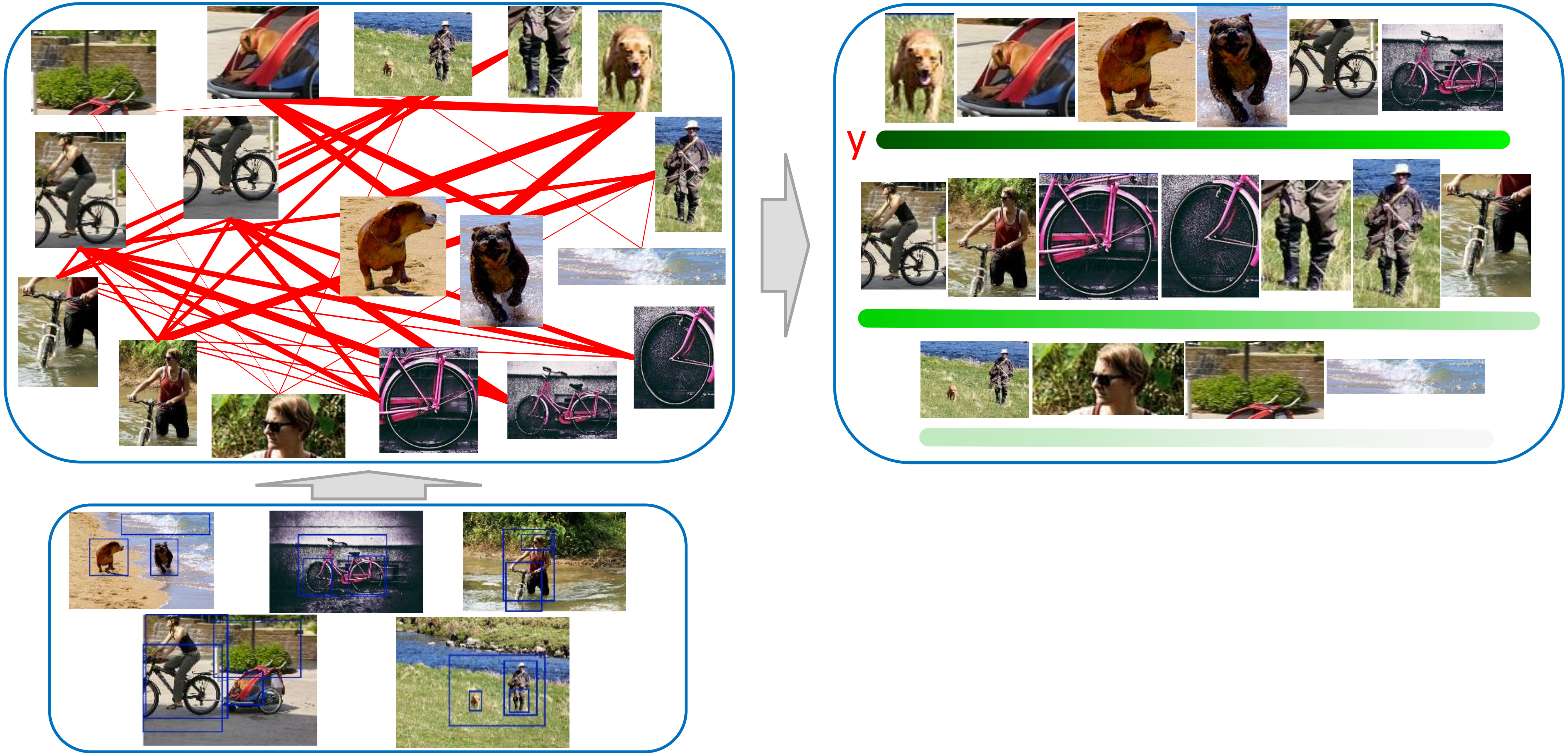
Completely unsupervised pipeline.



Proposed Approach

UOD can naturally be formulated as a ranking problem.

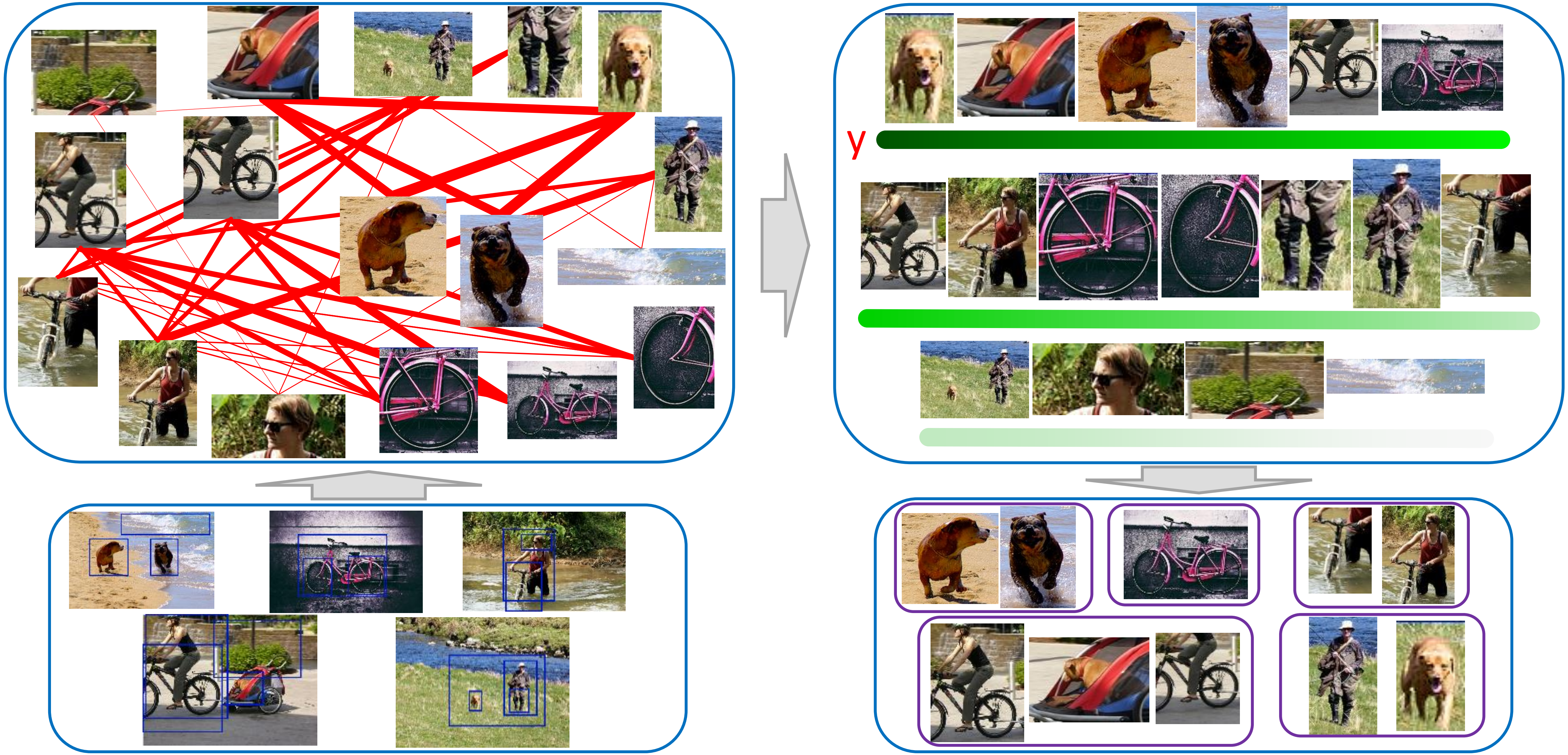
Completely unsupervised pipeline.



Proposed Approach

UOD can naturally be formulated as a ranking problem.

Completely unsupervised pipeline.



Ranking Formulations

Input: Symmetric, weight matrix W, number of nodes N.

Output: Nodes' importance.

Quadratic optimization (Q)

- $y_i \in [0,1]$ is the importance of node i, $y = (y_1, y_2, \ldots, y_N)^T$.
- Total importance of edges of node i: $c_i = \sum_j y_i W_{ij} y_j$.
- Optimization problem:

$$y^* = \underset{y \geq 0, \|y\| \leq 1}{\operatorname{argmax}} \sum_i c_i = y^T W y$$

- y^* is the eigenvector associated to the largest eigenvalue of W.

Edmund Landau, Deutsches Wochenschach 1895.

Ranking Formulations

Input: Symmetric, weight matrix W, number of nodes N.

Output: Nodes' importance.

PageRank

- $A = \text{normalize\_columns}(W)$.
- PageRank matrix: $P = (1-\beta)A + \beta u \mathbb{1}_N^T$ where $\beta \geq 0$ and $u \in \boldsymbol{R}^N$ such that $\sum_i u_i = 1$.
- PageRank vector: Eigenvector associated to the largest eigenvalue of the PageRank matrix.
- u can be used to inject prior knowledge to the model, e.g., prioritize some nodes (personalized PageRank).

Brin and Page, Computer Network 1998; Page et al., Technical Report, 1999.

Ranking Formulations

Input: Symmetric, weight matrix W, number of nodes N.

Output: Nodes' importance.

Hybrid formulation (LOD)

- (Q)'s solution can be used as prior.
- Build u using (Q)'s solution.
- LOD: Personalized PageRank with

$$P = (1-\beta)A + \beta u(Q)\mathbb{1}_N^T$$

Q, PageRank and LOD are all eigenvector problems.

Distributed Implementation

Power iterations:

- Input: matrix W ,$x^{(0)}$, norm L_p, number of iterations T.
- Repeat: $x^{(t+1)} = \frac{1}{\|Wx^{(t)}\|} W x^{(t)}$

W is divided into chunks of rows of blocks $W^1, W^2, \ldots, W^u$

Distributed Implementation

Power iterations:

- Input: matrix W, $x^{(0)}$, norm L_p, number of iterations T.
- Repeat: $x^{(t+1)} = \frac{1}{\|Wx^{(t)}\|} Wx^{(t)}$

W is divided into chunks of rows of blocks $W^1, W^2, \ldots, W^u$

$$Wx = \begin{pmatrix} Wx_1 \\ Wx_2 \\ \vdots \\ Wx_k \\ Wx_{k+1} \\ \vdots \\ \vdots \\ Wx_{n-1} \\ Wx_n \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} & \cdots & S_{1n} \\ S_{21} & S_{22} & \cdots & S_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ S_{k1} & S_{k2} & \cdots & S_{kn} \\ S_{(k+1)1} & S_{(k+1)2} & \cdots & S_{(k+1)n} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ S_{(n-1)1} & S_{(n-1)2} & \cdots & S_{(n-1)n} \\ S_{n1} & S_{n2} & \cdots & S_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \\ x_{k+1} \\ \vdots \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix}$$

Distributed Implementation

Power iterations:

- Input: matrix W,$x^{(0)}$, norm L_p, number of iterations T.
- Repeat: $x^{(t+1)} = \frac{1}{\|Wx^{(t)}\|} W x^{(t)}$

W is divided into chunks of rows of blocks $W^1, W^2, \dots, W^u$

$$Wx = \begin{pmatrix} Wx_1 \\ Wx_2 \\ \vdots \\ \boxed{\begin{matrix} Wx_k \\ Wx_{k+1} \\ \vdots \end{matrix}} \\ \vdots \\ Wx_{n-1} \\ Wx_n \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} & \cdots & S_{1n} \\ S_{21} & S_{22} & \cdots & S_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ S_{k1} & S_{k2} & \cdots & S_{kn} \\ S_{(k+1)1} & S_{(k+1)2} & \cdots & S_{(k+1)n} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ S_{(n-1)1} & S_{(n-1)2} & \cdots & S_{(n-1)n} \\ S_{n1} & S_{n2} & \cdots & S_{nn} \end{pmatrix} \boxed{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \\ x_{k+1} \\ \vdots \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix}}$$

Distributed Implementation

Power iterations:

- Input: matrix $W, x^{(0)}$, norm L_p, number of iterations T.
- Repeat: $x^{(t+1)} = \frac{1}{\|Wx^{(t)}\|} Wx^{(t)}$

W is divided into chunks of rows of blocks $W^1, W^2, \dots, W^u$

$$Wx = \begin{pmatrix} Wx_1 \\ Wx_2 \\ \vdots \\ Wx_k \\ Wx_{k+1} \\ \vdots \\ \vdots \\ Wx_{n-1} \\ Wx_n \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} & \cdots & S_{1n} \\ S_{21} & S_{22} & \cdots & S_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ S_{k1} & S_{k2} & \cdots & S_{kn} \\ S_{(k+1)1} & S_{(k+1)2} & \cdots & S_{(k+1)n} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ S_{(n-1)1} & S_{(n-1)2} & \cdots & S_{(n-1)n} \\ S_{n1} & S_{n2} & \cdots & S_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \\ x_{k+1} \\ \vdots \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix}$$

Distributed Implementation

Power iterations:

- Input: matrix W, $x^{(0)}$, norm L_p, number of iterations T.
- Repeat: $x^{(t+1)} = \frac{1}{\|Wx^{(t)}\|} Wx^{(t)}$

W is divided into chunks of rows of blocks $W^1, W^2, \dots, W^u$

$$Wx = \begin{pmatrix} Wx_1 \\ Wx_2 \\ \vdots \\ Wx_k \\ Wx_{k+1} \\ \vdots \\ \vdots \\ Wx_{n-1} \\ Wx_n \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} & \cdots & S_{1n} \\ S_{21} & S_{22} & \cdots & S_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ S_{k1} & S_{k2} & \cdots & S_{kn} \\ S_{(k+1)1} & S_{(k+1)2} & \cdots & S_{(k+1)n} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ S_{(n-1)1} & S_{(n-1)2} & \cdots & S_{(n-1)n} \\ S_{n1} & S_{n2} & \cdots & S_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \\ x_{k+1} \\ \vdots \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} \qquad \Longrightarrow \qquad Wx = \begin{pmatrix} W^1 x \\ W^2 x \\ \vdots \\ W^u x \end{pmatrix}$$

Qualitative Comparisons

Datasets: COCO trainval (COCO120K, ~120k images), COCO_20k (~20k images), OpenImages (Op1.7M, ~1.7 million images), Op50k(~50k images).
Metrics: Correct Localization (CorLoc), Average Precision (AP).

Method	Single-object				Multi-object							
	CorLoc				AP50				AP@[50:95]			
	C20K	C120K	Op50K	Op1.7M	C20K	C120K	Op50K	Op1.7M	C20K	C120K	Op50K	Op1.7M
EB [83]	28.8	29.1	32.7	32.8	4.86	4.91	5.46	5.49	1.41	1.43	1.53	1.53
Wei [71]	38.2	38.3	34.8	34.8	2.41	2.44	1.86	1.86	0.73	0.74	0.6	0.6
Kim [32]	35.1	34.8	37.0	-	3.93	3.93	4.13	-	0.96	0.96	0.98	-
Vo [67]	**48.5**	48.5	48.0	47.8	5.18	5.03	4.98	4.88	1.62	1.6	1.58	1.57
Ours (LOD+Self [18])	41.1	42.4	**49.5**	**49.4**	4.56	4.90	6.37	**6.28**	1.29	1.37	1.87	**1.86**
Ours (LOD)	**48.5**	**48.6**	48.1	47.7	**6.63**	**6.64**	**6.46**	**6.28**	**1.98**	**2.0**	**1.88**	1.83

[18] Gidaris et al., CVPR'21; [32] Kim et al., NIPS'09; [67] Vo et al., ECCV'20; [71] Wei et al., PR'19; [83] Zitnick et al., ECCV'14; Symonian et al., ICLR'15; Lin et al., ECCV'14; Krasin et al., 2017.

Qualitative Comparisons

Datasets: COCO trainval (COCO120K, ~120k images), COCO_20k (~20k images), OpenImages (Op1.7M, ~1.7 million images), Op50k(~50k images).

Metrics: Correct Localization (CorLoc), Average Precision (AP).

Method	Single-object				Multi-object							
	CorLoc				AP50				AP@[50:95]			
	C20K	C120K	Op50K	Op1.7M	C20K	C120K	Op50K	Op1.7M	C20K	C120K	Op50K	Op1.7M
EB [83]	28.8	29.1	32.7	32.8	4.86	4.91	5.46	5.49	1.41	1.43	1.53	1.53
Wei [71]	38.2	38.3	34.8	34.8	2.41	2.44	1.86	1.86	0.73	0.74	0.6	0.6
Kim [32]	35.1	34.8	37.0	-	3.93	3.93	4.13	-	0.96	0.96	0.98	-
Vo [67]	**48.5**	48.5	48.0	47.8	5.18	5.03	4.98	4.88	1.62	1.6	1.58	1.57
Ours (LOD+Self [18])	41.1	42.4	**49.5**	**49.4**	4.56	4.90	6.37	**6.28**	1.29	1.37	1.87	**1.86**
Ours (LOD)	**48.5**	**48.6**	48.1	47.7	**6.63**	**6.64**	**6.46**	**6.28**	**1.98**	**2.0**	**1.88**	1.83

[18] Gidaris et al., CVPR'21; [32] Kim et al., NIPS'09; [67] Vo et al., ECCV'20; [71] Wei et al., PR'19; [83] Zitnick et al., ECCV'14; Symonian et al., ICLR'15; Lin et al., ECCV'14; Krasin et al., 2017.

Qualitative Comparisons

Datasets: COCO trainval (COCO120K, ~120k images), COCO_20k (~20k images), OpenImages (Op1.7M, ~1.7 million images), Op50k(~50k images).
Metrics: Correct Localization (CorLoc), Average Precision (AP).

Method	Single-object				Multi-object							
	CorLoc				AP50				AP@[50:95]			
	C20K	C120K	Op50K	Op1.7M	C20K	C120K	Op50K	Op1.7M	C20K	C120K	Op50K	Op1.7M
EB [83]	28.8	29.1	32.7	32.8	4.86	4.91	5.46	5.49	1.41	1.43	1.53	1.53
Wei [71]	38.2	38.3	34.8	34.8	2.41	2.44	1.86	1.86	0.73	0.74	0.6	0.6
Kim [32]	35.1	34.8	37.0	-	3.93	3.93	4.13	-	0.96	0.96	0.98	-
Vo [67]	**48.5**	48.5	48.0	47.8	5.18	5.03	4.98	4.88	1.62	1.6	1.58	1.57
Ours (LOD+Self [18])	41.1	42.4	**49.5**	**49.4**	4.56	4.90	6.37	**6.28**	1.29	1.37	1.87	**1.86**
Ours (LOD)	**48.5**	**48.6**	48.1	47.7	**6.63**	**6.64**	**6.46**	**6.28**	**1.98**	**2.0**	**1.88**	1.83

[18] Gidaris et al., CVPR'21; [32] Kim et al., NIPS'09; [67] Vo et al., ECCV'20; [71] Wei et al., PR'19; [83] Zitnick et al., ECCV'14; Symonian et al., ICLR'15; Lin et al., ECCV'14; Krasin et al., 2017.

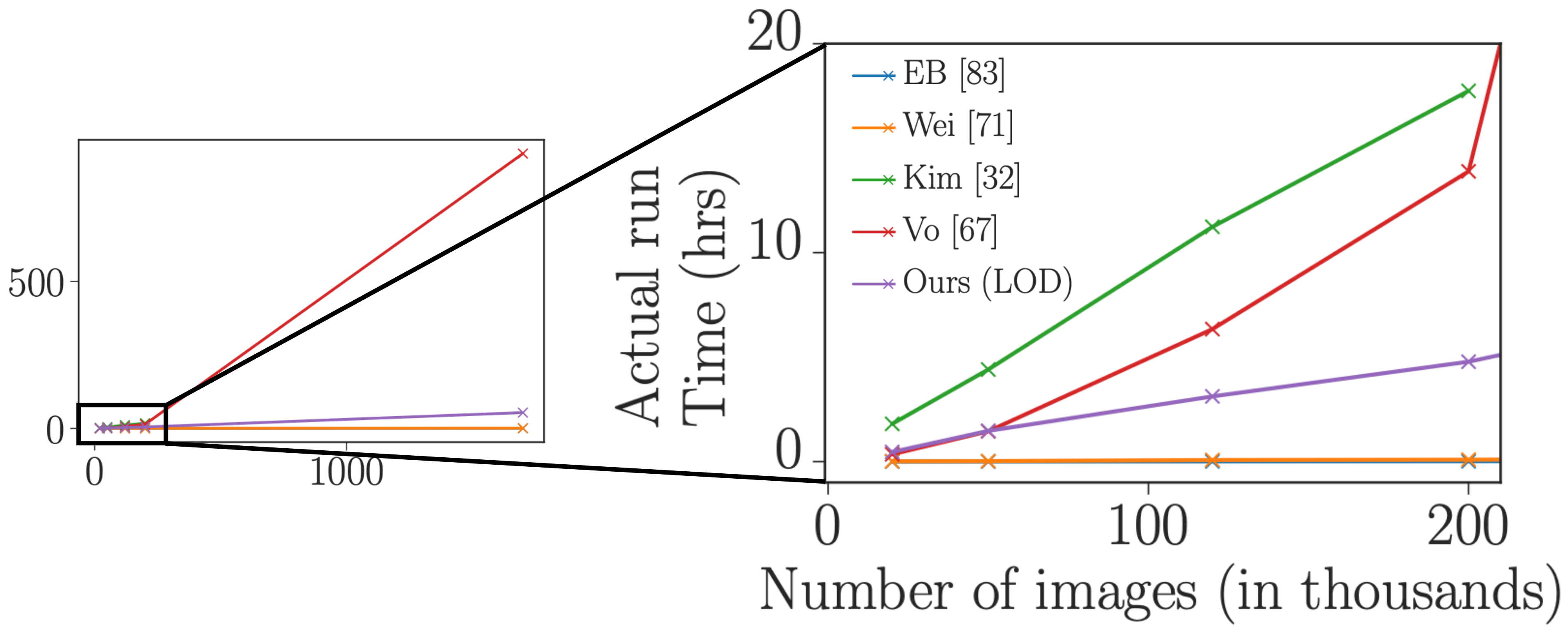
Qualitative Comparisons

Datasets: COCO trainval (COCO120K, ~120k images), COCO_20k (~20k images), OpenImages (Op1.7M, ~1.7 million images), Op50k(~50k images).
Metrics: Correct Localization (CorLoc), Average Precision (AP).

Method	Single-object				Multi-object							
	CorLoc				AP50				AP@[50:95]			
	C20K	C120K	Op50K	Op1.7M	C20K	C120K	Op50K	Op1.7M	C20K	C120K	Op50K	Op1.7M
EB [83]	28.8	29.1	32.7	32.8	4.86	4.91	5.46	5.49	1.41	1.43	1.53	1.53
Wei [71]	38.2	38.3	34.8	34.8	2.41	2.44	1.86	1.86	0.73	0.74	0.6	0.6
Kim [32]	35.1	34.8	37.0	-	3.93	3.93	4.13	-	0.96	0.96	0.98	-
Vo [67]	**48.5**	48.5	48.0	47.8	5.18	5.03	4.98	4.88	1.62	1.6	1.58	1.57
Ours (LOD+Self [18])	41.1	42.4	**49.5**	**49.4**	4.56	4.90	6.37	**6.28**	1.29	1.37	1.87	**1.86**
Ours (LOD)	**48.5**	**48.6**	48.1	47.7	**6.63**	**6.64**	**6.46**	**6.28**	**1.98**	**2.0**	**1.88**	1.83

[18] Gidaris et al., CVPR'21; [32] Kim et al., NIPS'09; [67] Vo et al., ECCV'20; [71] Wei et al., PR'19; [83] Zitnick et al., ECCV'14; Symonian et al., ICLR'15; Lin et al., ECCV'14; Krasin et al., 2017.

Running time



[32] Kim et al., NIPS’09; [67] Vo et al., ECCV’20; [71] Wei et al., PR’19; [83] Zitnick et al., ECCV’14.

Different design choices

Opt.	Proposal	Feature	Single-object CorLoc C20K	Single-object CorLoc Op50K	Multi-object AP50 C20K	Multi-object AP50 Op50K	Multi-object AP@[50:95] C20K	Multi-object AP@[50:95] Op50K
None	EB [83]	None	28.8	32.7	4.86	5.46	1.41	1.53
	[67]+Self		29.7	39.8	2.47	3.72	0.61	1.0
	[67]+Sup		23.6	38.1	4.07	4.81	1.03	1.39
Wei [71]	None	Self	37.9	42.4	2.53	3.13	0.69	0.9
		Sup	38.2	34.8	2.41	1.86	0.73	0.6
Kim [32]	EB [83]	Self	5.5	5.4	0.64	0.79	0.13	0.15
		Sup	15.6	20.2	1.96	2.56	0.36	0.47
	[67]+Self	Self	4.7	4.6	0.13	0.29	0.02	0.05
	[67]+Sup	Sup	35.1	37.0	3.93	4.13	0.96	0.98
Vo [67]	EB [83]	Self	35.6	43.6	3.34	4.43	0.99	1.39
		Sup	40.2	44.0	4.0	4.47	1.21	1.41
	[67]+Self	Self	37.8	48.1	2.65	4.19	0.82	1.45
	[67]+Sup	Sup	**48.5**	48.0	5.18	4.98	1.62	1.58
LOD	EB [83]	Self	35.5	39.7	5.87	6.73	1.57	1.76
		Sup	38.9	41.3	6.52	**7.01**	1.76	1.86
	[67]+Self	Self	41.1	**49.5**	4.56	6.37	1.29	1.87
	[67]+Sup	Sup	**48.5**	48.1	**6.63**	6.46	**1.98**	**1.88**

Opt.	Proposal	Feature	Single-object CorLoc C20K	Single-object CorLoc Op50K	Multi-object AP50 C20K	Multi-object AP50 Op50K	Multi-object AP@[50:95] C20K	Multi-object AP@[50:95] Op50K
Q	EB [83]	Self	32.8	40.3	4.15	6.43	1.07	1.67
		Sup	36.0	41.1	5.72	6.49	1.47	1.7
	[67]+Self	Self	38.7	48.9	4.38	6.39	1.17	1.84
	[67]+Sup	Sup	43.8	47.5	6.21	6.66	1.74	**1.88**
P	EB [83]	Self	35.5	39.7	4.91	6.73	1.34	1.75
		Sup	38.9	41.3	6.51	6.99	1.76	1.86
	[67]+Self	Self	41.2	**49.5**	4.38	6.13	1.24	1.81
	[67]+Sup	Sup	47.5	47.8	6.25	6.19	1.87	1.81
LOD	EB [83]	Self	35.5	39.7	5.87	6.73	1.57	1.76
		Sup	38.9	41.3	6.52	**7.01**	1.76	1.86
	[67]+Self	Self	41.1	**49.5**	4.56	6.37	1.29	1.87
	[67]+Sup	Sup	**48.5**	48.1	**6.63**	6.46	**1.98**	**1.88**

[32] Kim et al., NIPS'09; [67] Vo et al., ECCV'20; [71] Wei et al., PR'19; [83] Zitnick et al., ECCV'14; Gidaris et al., CVPR'21; Symonian et al., ICLR'15.

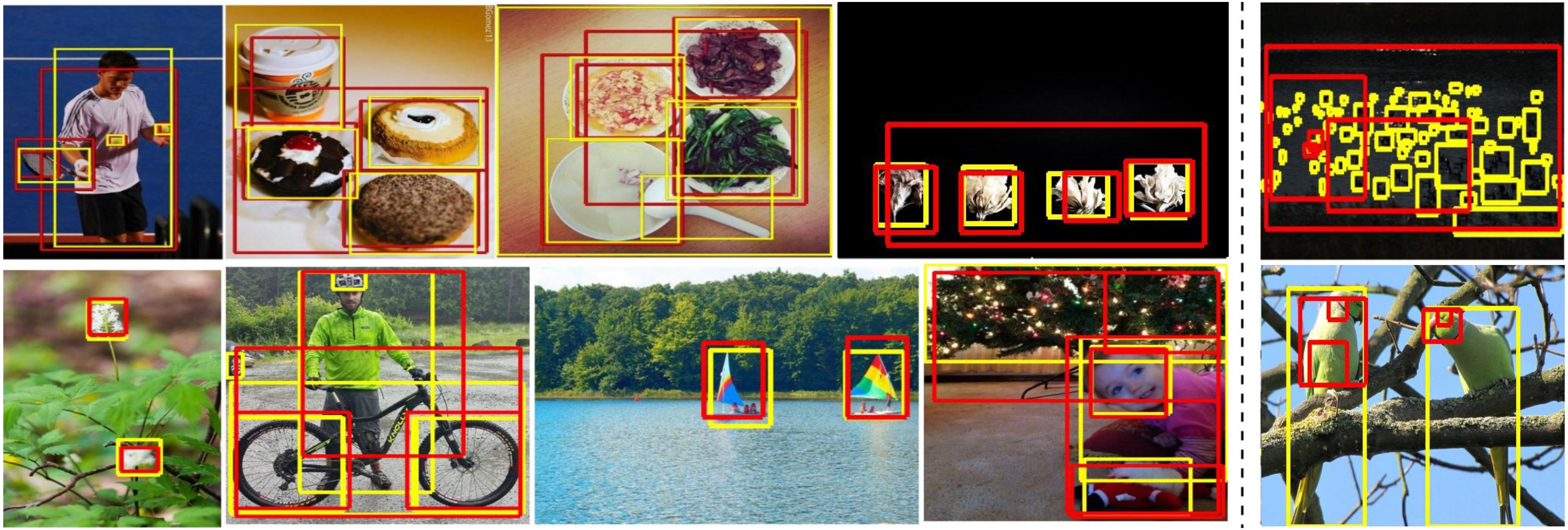
Different design choices

Opt.	Proposal	Feature	Single-object CorLoc		Multi-object AP50		Multi-object AP@[50:95]	
			C20K	Op50K	C20K	Op50K	C20K	Op50K
None	EB [83]	None	28.8	32.7	4.86	5.46	1.41	1.53
	[67]+Self		29.7	39.8	2.47	3.72	0.61	1.0
	[67]+Sup		23.6	38.1	4.07	4.81	1.03	1.39
Wei [71]	None	Self	37.9	42.4	2.53	3.13	0.69	0.9
		Sup	38.2	34.8	2.41	1.86	0.73	0.6
Kim [32]	EB [83]	Self	5.5	5.4	0.64	0.79	0.13	0.15
		Sup	15.6	20.2	1.96	2.56	0.36	0.47
	[67]+Self	Self	4.7	4.6	0.13	0.29	0.02	0.05
	[67]+Sup	Sup	35.1	37.0	3.93	4.13	0.96	0.98
Vo [67]	EB [83]	Self	35.6	43.6	3.34	4.43	0.99	1.39
		Sup	40.2	44.0	4.0	4.47	1.21	1.41
	[67]+Self	Self	37.8	48.1	2.65	4.19	0.82	1.45
	[67]+Sup	Sup	**48.5**	48.0	5.18	4.98	1.62	1.58
LOD	EB [83]	Self	35.5	39.7	5.87	6.73	1.57	1.76
		Sup	38.9	41.3	6.52	**7.01**	1.76	1.86
	[67]+Self	Self	41.1	**49.5**	4.56	6.37	1.29	1.87
	[67]+Sup	Sup	**48.5**	48.1	**6.63**	6.46	**1.98**	**1.88**

Opt.	Proposal	Feature	Single-object CorLoc		Multi-object AP50		Multi-object AP@[50:95]	
			C20K	Op50K	C20K	Op50K	C20K	Op50K
Q	EB [83]	Self	32.8	40.3	4.15	6.43	1.07	1.67
		Sup	36.0	41.1	5.72	6.49	1.47	1.7
	[67]+Self	Self	38.7	48.9	4.38	6.39	1.17	1.84
	[67]+Sup	Sup	43.8	47.5	6.21	6.66	1.74	**1.88**
P	EB [83]	Self	35.5	39.7	4.91	6.73	1.34	1.75
		Sup	38.9	41.3	6.51	6.99	1.76	1.86
	[67]+Self	Self	41.2	**49.5**	4.38	6.13	1.24	1.81
	[67]+Sup	Sup	47.5	47.8	6.25	6.19	1.87	1.81
LOD	EB [83]	Self	35.5	39.7	5.87	6.73	1.57	1.76
		Sup	38.9	41.3	6.52	**7.01**	1.76	1.86
	[67]+Self	Self	41.1	**49.5**	4.56	6.37	1.29	1.87
	[67]+Sup	Sup	**48.5**	48.1	**6.63**	6.46	**1.98**	**1.88**

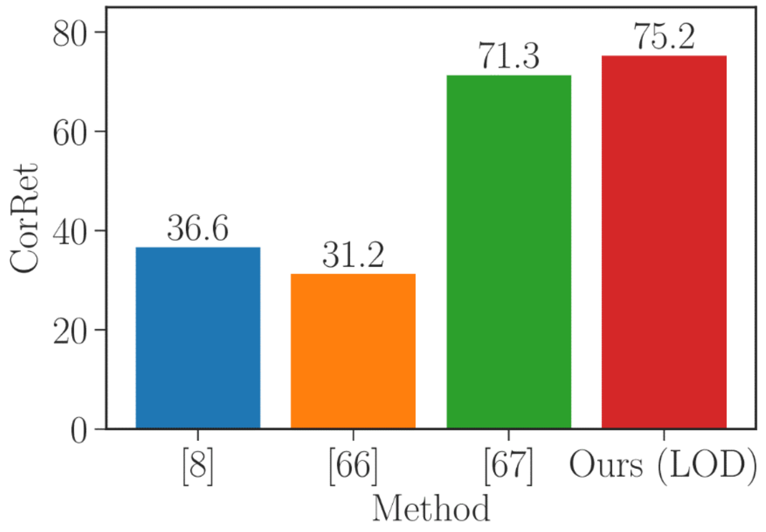
[32] Kim et al., NIPS’09; [67] Vo et al., ECCV’20; [71] Wei et al., PR’19; [83] Zitnick et al., ECCV’14; Gidaris et al., CVPR’21; Symonian et al., ICLR’15.

Qualitative results



Lin et al., ECCV'14; Krasin et al., 2017.

Object class discovery

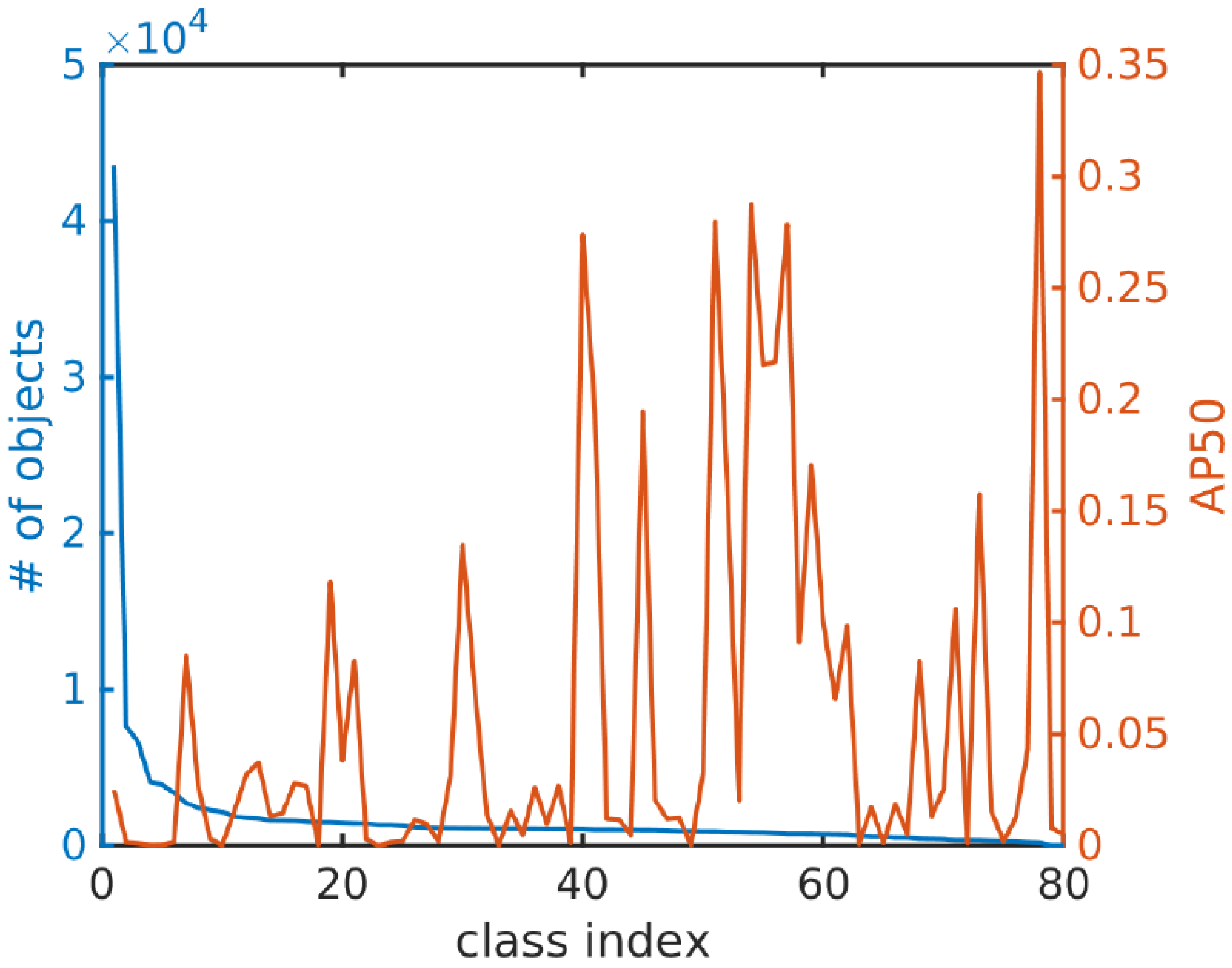


Dataset/Method	Ours (LOD)	[75]	[82]	[14]	[73]	[74]	[31]	[33]
SIVAL1	**97.4**	89.0	95.3	80.4	39.3	38.0	27.0	45.0
SIVAL2	**99.0**	93.2	84.0	71.7	40.0	33.3	35.3	33.3
SIVAL3	88.3	**88.4**	74.7	62.7	37.3	38.7	26.7	41.3
SIVAL4	**97.7**	87.8	94.0	86.0	33.0	37.7	27.3	53.0
SIVAL5	**94.3**	92.7	75.3	70.3	35.3	37.7	25.0	48.3
Average	**95.3**	90.2	84.7	74.2	37.0	37.1	28.3	44.2

[8] Cho et al., CVPR'15; [14] Feng et al., ICCV'11; [31] Kim et al., CVPR'08; [33] Kim et al., CVPR'12; [66] Vo et al., CVPR'19; [67] Vo et al., ECCV'20; [73] Zhang et al., Transactions on Neural Networks 2011; [74] Zhang et al., Applied Intelligence 2009; [75] Zhang et al., ICCV'15; [82] Zhu et al., CVPR'12; Rahmani et al., TPAMI'08.

Discussions

Finding frequently appearing visual patterns could favorize frequent object classes than rare ones. But is practice, it is not the case.



Conclusions

Ranking formulations are effective in scaling up unsupervised object discovery.

LOD combined with self-supervised features for UOD offers a viable UOD pipeline without any supervision whatsoever.

Code: https://github.com/huyvvo