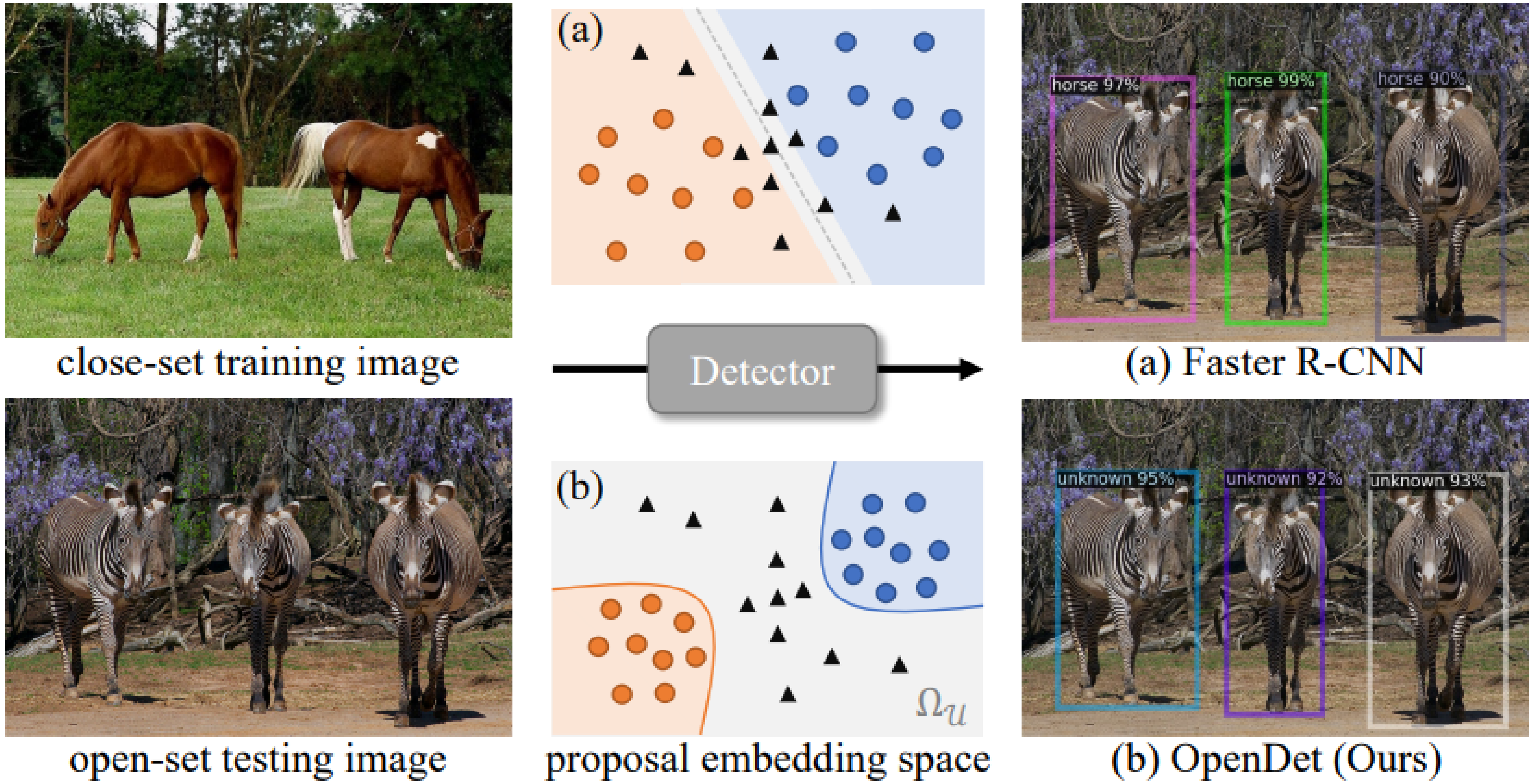
Expanding low-Density Latent Regions for Open-Set Object Detection

- Authors: Jiaming Han, Yuqiang Ren, Jian Ding, Xingjia Pan, Ke Yan, Gui-Song Xia

- Published in: CVPR 2022

- Presented by: Leonardo Rolandi (s301216)

Traditional Detector vs Open-Set Detector



a) threshold-based methods, like Faster R-CNN, usually misclassify unknown objects(black triangles, e.g. zebra), into known classes(colored dots, e.g. horse) due to limited low-density regions (in gray color). **b)** OpenDet instead identifies unknown objects by expanding low-density regions

Open-Set Object Detection(OSOD) vs Open-Set Recognition(OSR)

$C_{\mathcal{K}} = \{1, \ldots, K\}$ $\longrightarrow$ set of K know classes

$C_{\mathcal{U}}$ $\longrightarrow$ the unknown classes

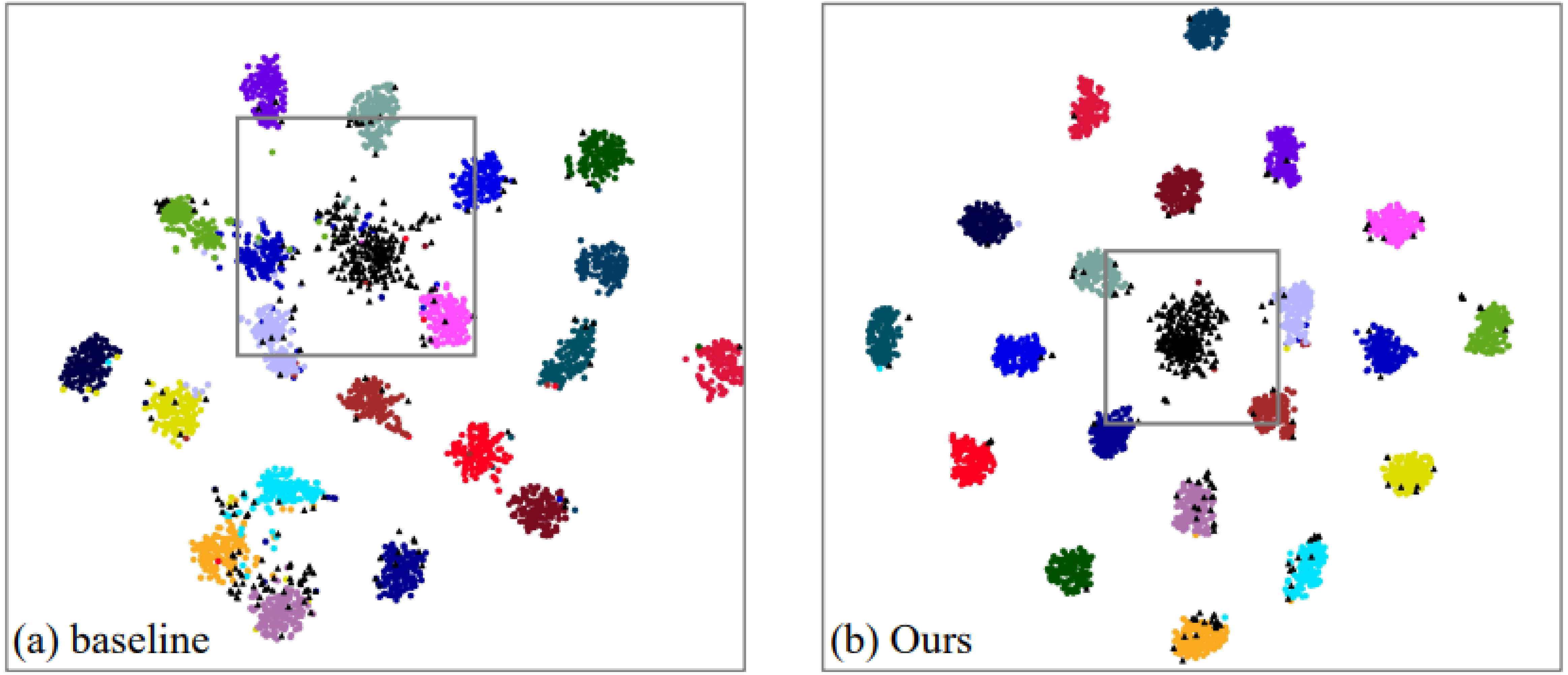
In **OSR**, an image only belongs to $C_{\mathcal{K}}$ or $C_{\mathcal{U}}$

In **OSOD**, an image may contains objects from both $C_{\mathcal{K}}$ and $C_{\mathcal{U}}$

Aims

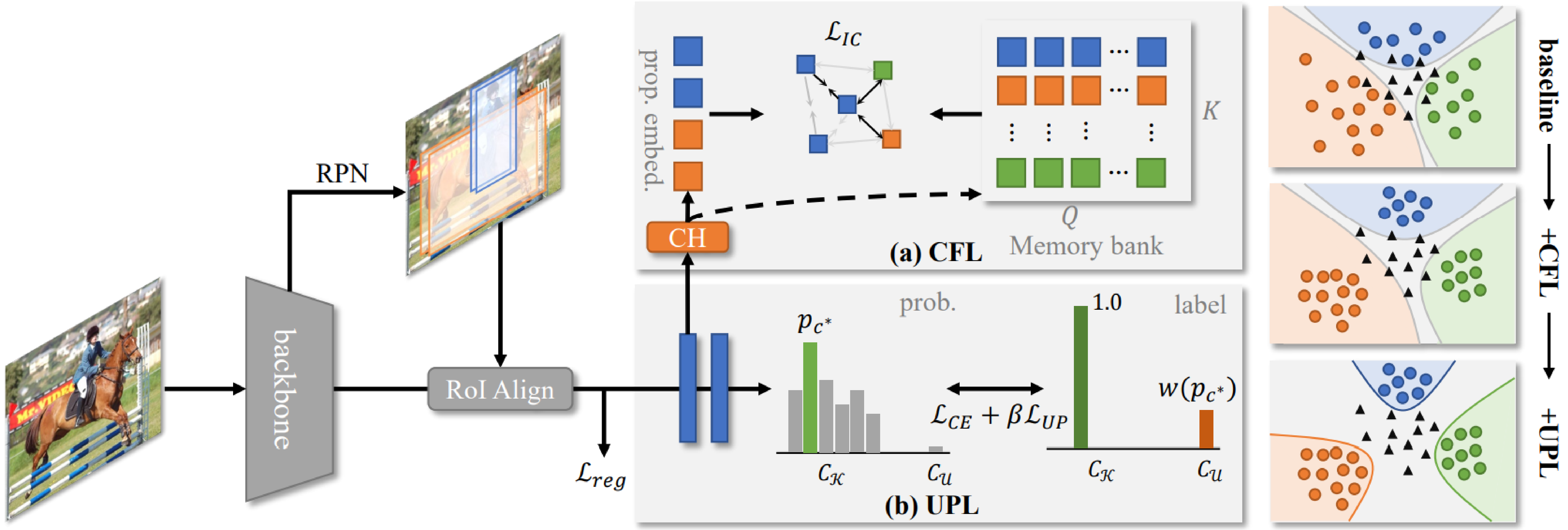
- Identify unknown objects
- No additional training step
- No post-processing

Separating high-low density regions



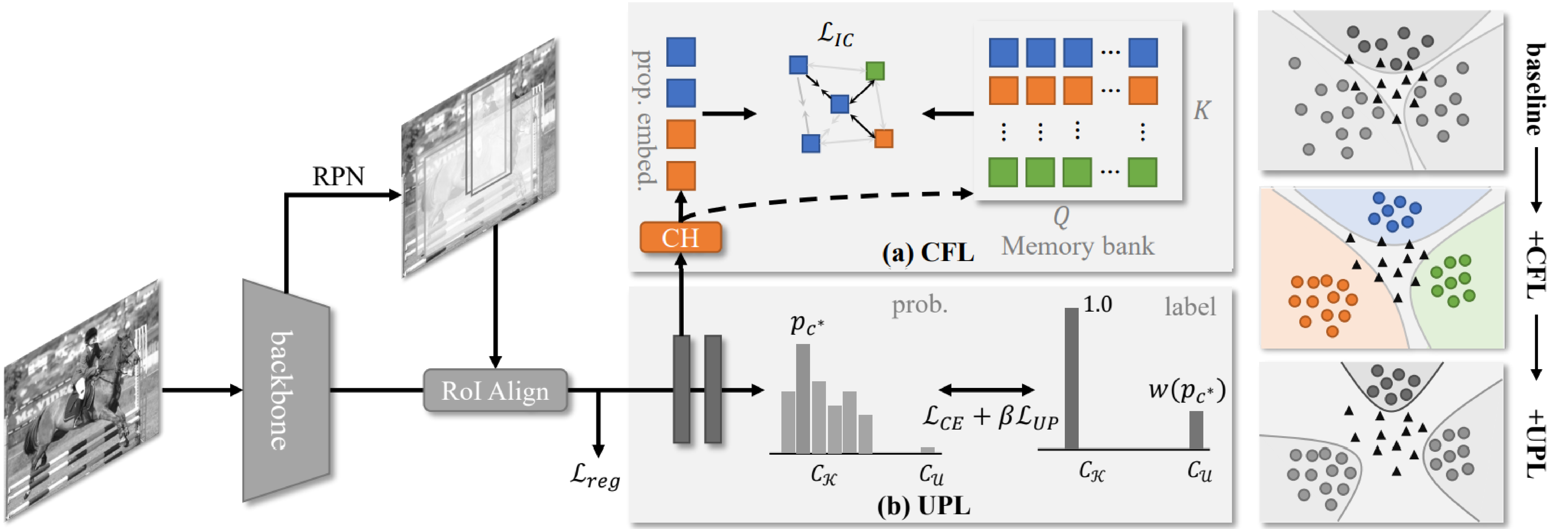
separation between known classes (colored dots) and unknown classes (black triangles) from **a)** a baseline detector and **b)** OpenDet

Open-Det



Left: Opendet with baseline and **a)** Contrastive Feature Learner (CFL) and **b)** Unknown Probability Learner (UPL)
Right: illustration of how different components work

Contrastive Feature Learner (CFL)



Contrastive Head (CH)

Multilayer Perceptron to map high-dimensional proposal feature to low-dimensional feature embedding

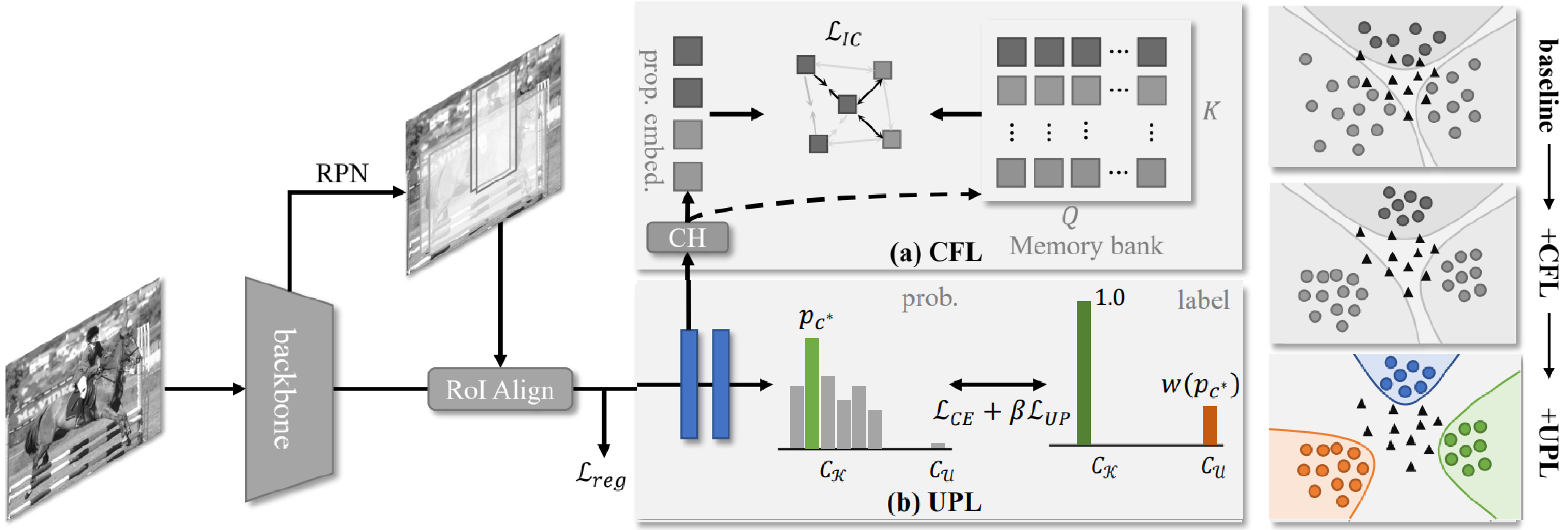
Class-Balanced Memory Bank (M(c))

For each class, first it screams proposals taking the most relevant, then samples the least similar

Instance Contrastive Loss (Lic)

it pushes the cluster of known classes away from low-density regions

Unknown Probability Learner (UPL)



$$\mathcal{L}_{CE} = -\sum_{c \in C} y_c \log(p_c), \quad y_c = \begin{cases} 1, c = c^* \\ 0, c \neq c^* \end{cases}$$

- $\mathcal{L}_{CE} = -\log(p_{c^*})$
- $\mathcal{L}_{UP} = -\log(p_u^{'}) \longrightarrow \mathcal{L}_{UP} = -w(p_{c^*})\log(p_u^{'})$

Total End-to-End Loss (3 default Faster R-CNN losses+2 added)

$$\mathcal{L} = \mathcal{L}_{rpn} + \mathcal{L}_{reg} + \mathcal{L}_{CE} + \beta\mathcal{L}_{UP} + \gamma_t\mathcal{L}_{IC}$$

$\mathcal{L}_{rpn}$ -> total loss of RPN, $\mathcal{L}_{reg}$ -> smooth L1 loss for box regression,

β and γ_t are weighting coefficient, with γ_t that gradually decreases for every iteration t

Comparison with other methods(1)

Training dataset: Pascal VOC
Testing dataset: Pascal VOC, VOC-COCO-20, VOC-COCO-40,VOC-COCO-60

Open-set metrics:

- Wilderness Impact(WI): it measures the degree of unknown objects missclassified to known classes: $WI = (\frac{P_{\mathcal{K}}}{P_{\mathcal{K}\cup\mathcal{U}}} - 1) \times 100$,where $P_{\mathcal{K}}$ and $P_{\mathcal{K}\cup\mathcal{U}}$ denote the precision of close-set and open-set classes, respectively
- Absolute Open-Set Error(AOSE): count the number of misclassified unknown objects
- Average precision of u (APu): average precision of unknown class

Close-set metric:

- mean Average Precision (mAP)

Comparison with other methods(2)

Method	VOC	VOC-COCO-20				VOC-COCO-40				VOC-COCO-60			
	$\mathrm{mAP}_{\mathcal{K}\uparrow}$	$\mathrm{WI}_{\downarrow}$	$\mathrm{AOSE}_{\downarrow}$	$\mathrm{mAP}_{\mathcal{K}\uparrow}$	$\mathrm{AP}_{\mathcal{U}\uparrow}$	$\mathrm{WI}_{\downarrow}$	$\mathrm{AOSE}_{\downarrow}$	$\mathrm{mAP}_{\mathcal{K}\uparrow}$	$\mathrm{AP}_{\mathcal{U}\uparrow}$	$\mathrm{WI}_{\downarrow}$	$\mathrm{AOSE}_{\downarrow}$	$\mathrm{mAP}_{\mathcal{K}\uparrow}$	$\mathrm{AP}_{\mathcal{U}\uparrow}$
FR-CNN [39]	80.10	18.39	15118	58.45	0	22.74	23391	55.26	0	18.49	25472	55.83	0
FR-CNN* [39]	80.01	18.83	11941	57.91	0	23.24	18257	54.77	0	18.72	19566	55.34	0
PROSER [53]	79.68	19.16	13035	57.66	10.92	24.15	19831	54.66	7.62	19.64	21322	55.20	3.25
ORE [23]	79.80	18.18	12811	58.25	2.60	22.40	19752	55.30	1.70	18.35	21415	55.47	0.53
DS [33]	80.04	16.98	12868	58.35	5.13	20.86	19775	55.31	3.39	17.22	21921	55.77	1.25
OpenDet	80.02	**14.95**	**11286**	**58.75**	**14.93**	**18.23**	**16800**	**55.83**	**10.58**	**14.24**	**18250**	**56.37**	**4.36**

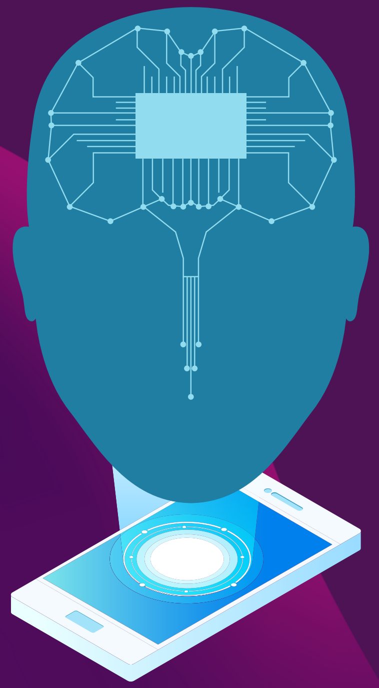
Table 1. **Comparisons with other methods on VOC and VOC-COCO-T_1.** We report close-set performance ($\mathrm{mAP}_{\mathcal{K}}$) on VOC, and both close-set ($\mathrm{mAP}_{\mathcal{K}}$) and open-set (WI, AOSE, $\mathrm{AP}_{\mathcal{U}}$) performance of different methods on VOC-COCO-{20, 40, 60}. * means a higher score threshold (*i.e.* 0.1) for testing.

Open-set Object detection methods:

- DS
- ORE
- PROSER

Traditional methods:

- Faster R-CNN (threshold of 0.05 for testing)
- Faster R-CNN* (higher score threshold for testing(0.1)



Thanks for your attention!

Leonardo Rolandi (s301216)