Supplementary material

Additional results on PASCAL VOC

In Tab. S1 and Tab. S2, we report the per-category detection APs and localization CorLocs on the *test* and *trainval* splits of PASCAL VOC 2007, respectively. Compared to other WSOD methods without external modules or data, we observe: (1) Our UWSOD outperforms all others on most categories (19 categories for APs and CorLocs). (2) The largest improvement in terms of AP is from category *cow*, which has gains of 37.9. (3) The performance of hard categories in our UWSOD, *i.e.*, *boat*, *chair* and *plant*, still surpass other methods by 10.3, 0.3 and 5.0 AP points, respectively. With class-agnostic ground-truth bounding-box known, we observe that the per-category performance of proposed UWSOD is competitive to that of Fast/Faster RCNN.

Table S1: Comparison with the state-of-the-art methods on PASCAL VOC 2007 in terms of AP (%) on *test*.

Method	Backbone	aero	bicy	bird	boa	bot	bus	car	cat	cha	cow	dtab	dog	hors	mbik	pers	plnt	she	sofa	trai	tv	Av.
WSOD with external object proposal modules or additional data																						
Multi-Fold MIL [6]	AlexNet	39.3	43.0	28.8	20.4	8.0	45.5	47.9	22.1	8.4	33.5	23.6	29.2	38.5	47.9	20.3	20.0	35.8	30.8	41.0	20.1	30.2
WSDDN [7]	VGG16	39.4	50.1	31.5	16.3	12.6	64.5	42.8	42.6	10.1	35.7	24.9	38.2	34.4	55.6	9.4	14.7	30.2	40.7	54.7	46.9	34.8
ContextLocNet [29]	VGG-F	57.1	52.0	31.5	7.6	11.5	55.0	53.1	34.1	1.7	33.1	49.2	42.0	47.3	56.6	15.3	12.8	24.8	48.9	44.4	47.8	36.3
WCCN [64]	VGG16	49.5	60.6	38.6	29.2	16.2	70.8	56.9	42.5	10.9	44.1	29.9	42.2	47.9	64.1	13.8	23.5	45.9	54.1	60.8	54.5	42.8
Jie *et al*. [34]	VGG16	52.2	47.1	35.0	26.7	15.4	61.3	66.0	54.3	3.0	53.6	24.7	43.6	48.4	65.8	6.6	18.8	51.9	43.6	53.6	62.4	41.7
TST [56]	AlexNet	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	33.8
SGWSOD [78]	VGG16	48.4	61.5	33.3	30.0	15.3	72.4	62.4	59.1	10.9	42.3	34.3	53.1	48.4	65.0	20.5	16.6	40.6	46.5	54.6	55.1	43.5
$TS^{C}2$ [32]	VGG16	59.3	57.5	43.7	27.3	13.5	63.9	61.7	59.9	24.1	46.9	36.7	45.6	39.9	62.6	10.3	23.6	41.7	52.4	58.7	56.6	44.3
CSC C5 [31]	VGG16	51.4	62.0	35.2	18.7	27.9	66.7	53.5	51.4	16.2	43.6	43.0	46.7	20.0	58.4	31.1	23.8	43.6	48.8	65.4	53.5	43.0
WS-JDS [48]	VGG16	52.0	64.5	45.5	26.7	27.9	60.5	47.8	59.7	13.0	50.4	46.4	56.3	49.6	60.7	25.4	28.2	50.0	51.4	66.5	29.7	45.6
OICR [8]	VGG16	58.0	62.4	31.1	19.4	13.0	65.1	62.2	28.4	24.8	44.7	30.6	25.3	37.8	65.5	15.7	24.1	41.7	46.9	64.3	62.6	41.2
K-EM [80]	VGG16	59.8	64.6	47.8	28.8	21.4	67.7	70.3	61.2	17.2	51.5	34.0	42.3	48.8	65.9	9.3	21.1	53.6	51.4	54.7	50.7	46.1
MELM [41]	VGG16	55.6	66.9	34.2	29.1	16.4	68.8	68.1	43.0	25.0	65.6	45.3	53.2	49.6	68.6	2.0	25.4	52.5	56.8	62.1	57.1	47.3
ZLDN [38]	VGG16	55.4	68.5	50.1	16.8	20.8	62.7	66.8	56.5	2.1	57.8	47.5	40.1	69.7	68.2	21.6	27.2	53.4	56.1	52.5	58.2	47.6
GAL-fWSD512 [47]	VGG16	58.4	63.8	45.8	24.0	22.7	67.7	65.7	58.9	15.0	58.1	47.0	53.7	23.8	64.3	36.2	22.3	46.7	50.3	70.8	55.1	47.5
ML-LocNet [42]	VGG16	59.3	68.9	45.7	29.0	24.5	64.8	68.4	59.3	18.6	49.1	50.2	43.1	65.8	70.2	19.9	24.3	48.1	54.2	62.8	41.8	48.4
WSRPN [11]	VGG16	57.9	70.5	37.8	5.7	21.0	66.1	69.2	59.4	3.4	57.1	57.3	35.2	64.2	68.6	32.8	28.6	50.8	49.5	41.1	30.0	45.3
PCL [36]	VGG16	54.4	69.0	39.3	19.2	15.7	62.9	64.4	30.0	25.1	52.5	44.4	19.6	39.3	67.7	17.8	22.9	46.6	57.5	58.6	63.0	43.5
Kosugi *et al*. [37]	VGG16	61.5	64.8	43.7	26.4	17.1	67.4	62.4	67.8	25.4	51.0	33.7	47.6	51.2	65.2	19.3	24.4	44.6	54.1	65.6	59.5	47.6
C-MIL [43]	VGG16	62.5	58.4	49.5	32.1	19.8	70.5	66.1	63.4	20.0	60.5	52.9	53.5	57.4	68.9	8.4	24.6	51.8	58.7	66.7	63.5	50.5
Pred Net [44]	VGG16	66.7	69.5	52.8	31.4	24.7	74.5	74.1	67.3	14.6	53.0	46.1	52.9	69.9	70.8	18.5	28.4	54.6	60.7	67.1	60.4	52.9
OICR W-RPN [12]	VGG16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	46.9
SDCN [49]	VGG16	59.8	67.1	32.0	34.7	22.8	67.1	63.8	67.9	22.5	48.9	47.8	60.5	51.7	65.2	11.8	20.6	42.1	54.7	60.8	64.3	48.3
Sona *et al*. [81]	VGG16	62.1	55.7	42.0	31.1	17.2	67.6	65.2	50.8	20.4	51.5	36.3	34.1	46.2	65.8	12.3	21.9	48.8	55.4	60.2	65.7	45.4
$WSOD^2$ [18]	VGG16	65.1	64.8	57.2	39.2	24.3	69.8	66.2	61.0	29.8	64.6	42.5	60.1	71.2	70.7	21.9	28.1	58.6	59.7	52.2	64.8	53.6
OICR+GAM+REG [35]	VGG16	55.2	66.5	40.1	31.1	16.9	69.8	64.3	67.8	27.8	52.9	47.0	33.0	60.8	64.4	13.8	26.0	44.0	55.7	68.9	65.5	48.6
C-MIDN [50]	VGG16	53.3	71.5	49.8	26.1	20.3	70.3	69.9	68.3	28.7	65.3	45.1	64.6	58.0	71.2	20.0	27.5	54.9	54.9	69.4	63.5	52.6
OIM+IR [39]	VGG16	55.6	67.0	45.8	27.9	21.1	69.0	68.3	70.5	21.3	60.2	40.3	54.5	56.5	70.1	12.5	25.0	52.9	55.2	65.0	63.7	50.1
Ren *et al*. [20]	VGG16	68.8	77.7	57.0	27.7	28.9	69.1	74.5	67.0	32.1	73.2	48.1	45.2	54.4	73.7	35.0	29.3	64.1	53.8	65.3	65.2	54.9
Zeni *et al*. [82]	VGG16	68.6	62.4	55.5	27.2	21.4	71.1	71.6	56.7	24.7	60.3	47.4	56.1	46.4	69.2	2.7	22.9	41.5	47.7	71.1	69.8	49.7
PG-PS [13]	VGG16	63.0	64.4	50.1	27.5	17.1	70.6	66.0	71.1	25.8	55.9	43.2	62.7	65.9	64.1	10.2	22.5	48.1	53.8	72.2	67.4	51.1
WSOD without external object proposal modules or additional data																						
Beam Search [62]	VGG16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	25.7
OM+MIL [33]	AlexNet	37.2	35.7	25.8	13.8	12.7	36.2	42.4	22.3	14.3	24.2	9.4	13.1	27.9	38.9	3.7	18.7	20.1	16.3	36.1	18.4	23.4
OPG [84]	VGG16	48.9	49.9	25.6	14.0	6.1	47.1	22.5	52.7	3.4	19.6	33.2	33.3	55.3	30.2	9.9	9.1	14.8	25.8	50.8	24.7	28.8
SPAM-CAM [63]	VGG16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	27.5
UWSOD	VGG16	57.7	72.7	46.4	24.3	11.2	60.4	72.3	29.2	14.6	58.7	29.1	59.4	72.6	68.6	1.4	23.7	35.6	40.3	51.8	49.8	44.0
	WSR18	59.1	68.9	47.1	18.3	28.2	65.1	72.3	31.4	7.6	62.1	44.9	46.6	68.1	70.4	18.9	23.5	49.9	40.5	66.2	10.5	45.0
FSOD																						
Fast RCNN [85]	VGG16	74.5	78.3	69.2	53.2	36.6	77.3	78.2	82.0	40.7	72.7	67.9	79.6	79.2	73.0	69.0	30.1	65.4	70.2	75.8	65.8	66.9
Faster RCNN [2]	VGG16	70.0	80.6	70.1	57.3	49.9	78.2	80.4	82.0	52.2	75.3	67.2	80.3	79.8	75.0	76.3	39.1	68.3	67.3	81.1	67.6	69.9
WSOD with Cls-agnostic GT-bbox Known																						
OICR + GAM + REG [35]	VGG16	66.1	64.0	55.6	25.3	37.2	75.6	69.9	53.6	17.4	63.2	62.2	49.0	65.1	69.3	50.8	18.9	52.1	54.7	69.5	66.2	54.3
Ren *et al*. [20]	VGG16	70.2	76.1	57.7	46.2	42.3	73.8	72.0	78.5	35.3	69.4	46.0	72.1	76.5	72.4	64.7	30.5	65.1	59.2	69.3	66.0	62.2
UWSOD	VGG16	74.0	77.6	66.2	57.8	51.0	78.0	74.7	81.5	43.2	70.5	67.6	77.8	81.2	71.1	67.7	36.7	59.0	66.1	78.6	72.0	67.7
	WSR18	73.2	77.6	69.6	64.4	50.7	81.4	77.7	83.9	44.1	71.0	69.5	79.3	81.6	72.3	70.0	42.0	63.4	69.1	82.2	70.9	69.7

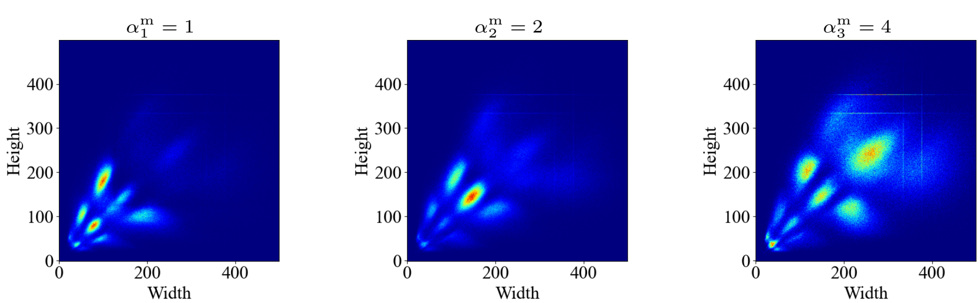

Figure S1: Proposal objectness score maps of different resampling rate α^{m} in MRRP on PASCAL VOC 2007 test.

Table S2: Comparison with the state-of-the-art methods on PASCAL VOC 2007 in terms of CorLoc (%) on *trainval*.

Method	Backbone	aero	bicy	bird	boa	bot	bus	car	cat	cha	cow	dtab	dog	hors	mbik	pers	plnt	she	sofa	trai	tv	Av.
WSOD with external object proposal modules or additional data																						
Multi-Fold MIL [6]	AlexNet	65.3	55.0	52.4	48.3	18.2	66.4	77.8	35.6	26.5	67.0	46.9	48.4	70.5	69.1	35.2	35.2	69.6	43.4	64.6	43.7	52.0
WSDDN [7]	VGG16	65.1	58.8	58.5	33.1	39.8	68.3	60.2	59.6	34.8	64.5	30.5	43.0	56.8	82.4	25.5	41.6	61.5	55.9	65.9	63.7	53.5
ContextLocNet [29]	VGG-F	83.3	68.6	54.7	23.4	18.3	73.6	74.1	54.1	8.6	65.1	47.1	59.5	67.0	83.5	35.3	39.9	67.0	49.7	63.5	65.2	55.1
WCCN [64]	VGG16	83.9	72.8	64.5	44.1	40.1	65.7	82.5	58.9	33.7	72.5	25.6	53.7	67.4	77.4	26.8	49.1	68.1	27.9	64.5	55.7	56.7
Jie *et al.*'17 [34]	VGG16	72.7	55.3	53.0	27.8	35.2	68.6	81.9	60.7	11.6	71.6	29.7	54.3	64.3	88.2	22.2	53.7	72.2	52.6	68.9	75.5	56.1
TST [56]	AlexNet	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	59.5
SGWSOD [78]	VGG16	71.0	76.5	54.9	49.7	54.1	78.0	87.4	68.8	32.4	75.2	29.5	58.0	67.3	84.5	41.5	49.0	78.1	60.3	62.8	78.9	62.9
TS^2C [32]	VGG16	84.2	74.1	61.3	52.1	32.1	76.7	82.9	66.6	42.3	70.6	39.5	57.0	61.2	88.4	9.3	54.6	72.2	60.0	65.0	70.3	61.0
CSC C5 [31]	VGG16	76.1	75.3	61.8	42.0	54.1	74.7	78.8	67.4	32.8	73.1	46.5	59.9	37.6	78.0	56.0	42.5	71.9	67.3	82.4	65.6	62.2
WS-JDS [48]	VGG16	82.9	74.0	73.4	47.1	60.9	80.4	77.5	78.8	18.6	70.0	56.7	67.0	64.5	84.0	47.0	50.1	71.9	57.6	83.3	43.5	64.5
OICR [8]	VGG16	81.7	80.4	48.7	49.5	32.8	81.7	85.4	40.1	40.6	79.5	35.7	33.7	60.5	88.8	21.8	57.9	76.3	59.9	75.3	81.4	60.6
K-EM [80]	VGG16	79.8	77.8	66.7	50.3	57.0	80.1	89.9	71.5	29.9	75.9	30.5	58.9	73.2	90.2	25.4	51.8	80.2	60.3	72.4	78.9	65.0
MELM [41]	VGG16	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	61.4
ZLDN [38]	VGG16	74.0	77.8	65.2	37.0	46.7	75.8	83.7	58.8	17.5	73.1	49.0	51.3	76.7	87.4	30.6	47.8	75.0	62.5	64.8	68.8	61.2
GAL-fWSD512 [47]	VGG16	78.6	81.9	63.6	40.3	48.8	80.7	85.3	76.3	30.3	78.0	54.5	65.3	48.4	86.5	56.3	46.9	76.0	68.1	83.9	73.1	66.1
ML-LocNet [42]	VGG16	78.6	82.3	68.2	42.0	53.3	78.5	88.5	70.3	36.4	70.2	60.5	58.0	80.5	88.2	38.8	59.2	75.0	69.0	78.2	64.5	67.0
WSRPN [11]	VGG16	77.5	81.2	55.3	19.7	44.3	80.2	86.6	69.5	10.1	87.7	68.4	52.1	84.4	91.6	57.4	63.4	77.3	58.1	57.0	53.8	63.8
Kosugi *et al.* [37]	VGG16	85.5	79.6	68.1	55.1	33.6	83.5	83.1	78.5	42.7	79.8	37.8	61.5	74.4	88.6	32.6	55.7	77.9	63.7	78.4	74.1	66.7
C-MIL [43]	VGG16	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	65.0
Pred Net [44]	VGG16	88.6	86.3	71.8	53.4	51.2	87.6	89.0	65.3	33.2	86.6	58.8	65.9	87.7	93.3	30.9	58.9	83.4	67.8	78.7	80.2	70.9
OICR W-RPN [12]	VGG16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	66.5
SDCN [49]	VGG16	85.8	83.1	56.2	58.5	44.7	80.2	85.0	77.9	29.6	78.8	53.6	74.2	73.1	88.4	18.2	57.5	74.2	60.8	76.1	79.2	66.8
$WSOD^2$ [18]	VGG16	87.1	80.0	74.8	60.1	36.6	79.2	83.8	70.6	43.5	88.4	46.0	74.7	87.4	90.8	44.2	52.4	81.4	61.8	67.7	79.9	69.5
OICR+GAM+REG [35]	VGG16	81.7	81.2	58.9	54.3	37.8	83.2	86.2	77.0	42.1	83.6	51.3	44.9	78.2	90.8	20.5	56.8	74.2	66.1	81.0	86.0	66.8
C-MIDN [50]	VGG16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	68.7
OIM+IR [39]	VGG16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	67.2
Ren *et al.* [20]	VGG16	87.5	82.4	76.0	58.0	44.7	82.2	87.5	71.2	49.1	81.5	51.7	53.3	71.4	92.8	38.2	52.8	79.4	61.0	78.3	76.0	68.8
Zeni *et al.* [82]	VGG16	86.7	73.3	72.4	55.3	46.9	83.2	87.5	64.5	44.6	76.7	46.4	70.9	67.0	88.0	9.6	56.4	69.1	52.4	79.8	82.8	65.7
PG-PS [13]	VGG16	85.4	80.4	69.1	58.0	35.9	82.7	86.7	82.6	45.5	84.9	44.1	80.2	84.0	89.2	12.3	55.7	79.4	63.4	82.1	82.1	69.2
WSOD without external object proposal modules or additional data																						
Shi *et al.* [83]	–	67.3	54.4	34.3	17.8	1.3	46.6	60.7	68.9	2.5	32.4	16.2	58.9	51.5	64.6	18.2	3.1	20.9	34.7	63.4	5.9	36.2
OM + MIL [33]	AlexNet	64.3	54.3	42.7	22.7	34.4	58.1	74.3	36.2	24.3	50.4	11.0	29.2	50.5	66.1	11.3	42.9	39.6	18.3	54.0	39.8	41.2
OPG [84]	VGG16	57.1	43.2	53.9	23.8	12.3	47.9	48.8	69.1	16.6	47.5	39.0	61.3	54.7	60.8	32.1	22.0	49.0	44.1	59.4	27.7	43.5
UWSOD	VGG16	77.8	85.8	66.0	56.0	39.1	74.2	91.4	41.4	30.3	81.9	33.0	78.9	90.5	85.6	7.6	46.4	68.8	67.0	76.1	61.7	63.0
	WSR18	80.4	85.3	79.4	42.0	65.5	78.4	90.7	49.7	18.8	73.9	48.5	63.1	87.8	90.8	37.4	47.4	77.1	54.1	81.9	23.4	63.8
WSOD with Cls-agnostic GT-bbox Known																						
OICR + GAM + REG [35]	VGG16	91.9	82.7	87.2	65.4	69.7	92.7	92.2	82.8	45.9	88.4	80.4	84.7	89.9	90.8	79.8	59.4	89.5	79.2	90.5	83.5	81.3
Ren *et al.* [20]	VGG16	95.3	94.5	93.0	83.1	68.5	92.8	91.8	93.0	69.8	86.4	71.2	89.1	92.9	96.3	88.9	75.4	93.8	82.6	92.4	89.4	87.1
UWSOD	VGG16	98.4	96.8	96.6	98.0	81.9	96.6	96.7	97.1	78.2	98.6	82.5	96.3	97.3	95.2	95.1	80.3	97.9	91.5	98.5	92.1	93.3
	WSR18	99.1	96.6	97.8	91.4	81.3	93.8	95.5	94.5	70.1	95.7	90.7	97.0	98.6	97.5	94.7	81.2	100.0	91.0	97.6	86.0	92.5

Additional analysis of MRRP

In Fig. S1, we show the proposal objectness score maps of different resampling rates in MRRP with $n^{\mathrm{m}} = 3$ and $\alpha^{\mathrm{m}} = \{1, 2, 4\}$. We accumulate the objectness scores of proposals with the same height and width from PASCAL VOC 2007 *test*. Then we map the accumulated objectness scores to the jet colourmap. We find that different resampling rates tend to favour different sizes of proposals. With large resampling rates, the number of the large proposals with high objectness scores is much more than that of small ones. And small resampling rates, *e.g.*, $\alpha^{\mathrm{m}} = 1$, predict more small proposals with high objectness scores. It demonstrates that our MRRP uses the in-network feature hierarchy to handle large scale variation.