

Supplementary Material:

Few-shot Object Detection via Feature Reweighting

Implementation Details

All our models are trained using SGD with momentum 0.9, and L_2 weight-decay 0.0005 (on both feature extractor and reweighting module). The batch size is set to be 64. For base training we train for 80,000 iterations, a step-wise learning rate decay strategy is used, with learning rate being 10^{-4} , 10^{-3} , 10^{-4} , 10^{-5} , and changes happening in iteration 500, 40,000, 60,000. For few-shot fine-tuning, we use a constant learning rate of 0.001 and train for 1500 iterations. We use multi-scale training, and evaluate the model in 416×416 resolution, as with the original YOLOv2.

Additional Ablation Studies

Sampling of Examples for Testing During training, the reweighting module takes random input from the k -shot data each of the N classes. In testing, we take the k -shot example as reweighting module’s input and use the average of their predicted weights for detecting the corresponding class. If we replace the averaging process by randomly selecting reweighting module’s input (as during training), the performance on base/novel classes will drop significantly from 69.7%/47.2% to 63.9%/45.1%. This is similar to the ensembling effect, except that this averaging over reweighting coefficients do not need additional inference time as in normal ensembling.

Sharing Weights Between Feature Extractor and Reweighting Module The first few layers of the reweighting module and the backbone feature extractor share the same architecture. Thus some weights can be shared between them. We evaluate this alternative and found the performance on base/novel classes decrease from 69.7%/47.2% to 68.3%/44.8%. The reason could be it imposes more constraints in the optimization process.

Complete Results on PASCAL VOC

Here we present the complete results for each class and number of shot on PASCAL VOC dataset. The results for base/novel split 1/2/3 are shown in Table 1/2/3 respectively.

# Shots		Novel						Base															
		bird	bus	cow	mbike	sofa	mean	aero	bike	boat	bottle	car	cat	chair	table	dog	horse	person	plant	sheep	train	tv	mean
1	YOLO-joint	0.0	0.0	0.0	0.0	0.0	0.0	78.4	76.9	61.5	48.7	79.8	84.5	51.0	72.7	79.0	77.6	74.9	48.2	62.8	84.8	73.1	70.2
	YOLO-ft	6.8	0.0	9.1	0.0	0.0	3.2	77.1	78.2	61.7	46.7	79.4	82.7	51.0	69.0	78.3	79.5	74.2	42.7	68.3	84.1	72.9	69.7
	YOLO-ft-full	11.4	17.6	3.8	0.0	0.0	6.6	75.8	77.3	63.1	45.9	78.7	84.1	52.3	66.5	79.3	77.2	73.7	44.0	66.0	84.2	72.2	69.4
	LSTD(YOLO)	12.0	17.8	4.6	0.0	0.1	6.9	75.5	76.9	63.2	46.2	78.9	84.1	52.5	66.8	79.2	79.4	74.1	44.7	66.4	84.6	73.6	69.7
	LSTD(YOLO)-full	13.4	21.4	6.3	0.0	0.0	8.2	73.4	73.5	61.8	44.7	78.4	83.9	50.8	68.3	79.3	80.5	72.3	41.0	64.5	83.2	72.5	68.5
	Ours	13.5	10.6	31.5	13.8	4.3	14.8	75.1	70.7	57.0	41.6	76.6	81.7	46.6	72.4	73.8	76.9	68.8	43.1	63.0	78.8	69.9	66.4
2	YOLO-joint	0.0	0.0	0.0	0.0	0.0	0.0	77.6	77.6	60.4	48.1	81.5	82.6	51.5	72.0	79.2	78.8	75.2	47.0	65.2	86.0	72.7	70.4
	YOLO-ft	11.5	5.8	7.6	0.1	7.5	6.5	77.9	75.0	58.5	45.7	77.6	84.0	50.4	68.5	79.2	79.7	73.8	44.0	66.0	77.5	72.9	68.7
	YOLO-ft-full	16.6	9.7	12.4	0.1	14.5	10.7	76.4	70.2	56.9	43.3	77.5	83.8	47.8	70.7	79.1	77.6	71.7	39.6	61.4	77.0	70.3	66.9
	LSTD(YOLO)	12.3	10.1	14.6	0.1	8.9	9.2	77.4	77.1	59.4	46.4	77.8	84.5	50.9	67.1	79.1	80.6	73.8	43.3	64.9	79.4	72.4	68.9
	LSTD(YOLO)-full	17.3	12.5	8.6	0.2	16.5	11.0	74.6	71.7	57.9	42.8	78.1	83.8	47.9	66.7	78.4	77.8	71.8	39.3	60.7	81.4	71.2	67.0
	Ours	21.2	12.0	16.8	17.9	9.6	15.5	74.6	74.9	56.3	38.5	75.5	68.0	43.2	69.3	66.2	42.4	68.1	41.8	59.4	76.4	70.3	61.7
3	YOLO-joint	0	0	0	0	9.1	1.8	78.0	77.2	61.2	45.6	81.6	83.7	51.7	73.4	80.7	79.6	75.0	45.5	65.6	83.1	72.7	70.3
	YOLO-ft	10.9	5.5	15.3	0.2	0.1	6.4	76.7	77.0	60.4	46.9	78.8	84.9	51.0	68.3	79.6	78.7	73.1	44.5	67.6	83.6	72.4	69.6
	YOLO-ft-full	21.0	22.0	19.1	0.5	0.0	12.5	73.4	67.5	56.8	41.2	77.1	81.6	45.5	62.1	74.6	78.9	67.9	37.8	54.1	76.4	71.9	64.4
	LSTD(YOLO)	12.3	7.1	17.7	0.1	0.0	7.5	75.9	76.2	59.7	46.6	78.3	84.4	49.4	64.5	78.7	79.7	72.6	42.5	63.8	80.5	73.9	68.4
	LSTD(YOLO)-full	23.1	22.6	15.9	0.4	0.0	12.4	74.8	68.7	57.1	44.1	78.0	83.4	46.9	64.0	78.7	79.1	70.1	39.2	58.1	79.8	71.9	66.3
	Ours	26.1	19.1	40.7	20.4	27.1	26.7	73.6	73.1	56.7	41.6	76.1	78.7	42.6	66.8	72.0	77.7	68.5	42.0	57.1	74.7	70.7	64.8
5	YOLO-joint	0.0	0.0	0.0	0.0	9.1	1.8	77.8	76.4	65.7	45.9	79.5	82.3	50.4	72.5	79.1	79.0	75.5	47.9	67.2	83.0	72.5	70.3
	YOLO-ft	11.6	7.1	10.7	2.1	6.0	7.5	76.5	76.4	61.0	45.5	78.7	84.5	49.2	68.7	78.5	78.1	73.7	45.4	66.8	85.3	70.0	69.2
	YOLO-ft-full	20.2	20.0	22.4	36.4	24.8	24.8	72.0	70.6	60.7	42.0	76.8	84.2	47.7	63.7	76.9	78.8	72.1	42.2	61.1	80.8	69.9	66.6
	LSTD(YOLO)	12.9	8.1	13.6	16.1	10.2	12.2	77.4	75.0	61.1	45.2	78.4	85.0	50.6	68.0	78.1	79.3	73.1	44.6	65.5	84.5	71.1	69.1
	LSTD(YOLO)-full	24.1	30.2	24.0	40.0	25.6	29.1	74.2	70.7	60.4	42.9	77.3	83.1	47.9	66.0	76.9	79.2	71.3	41.4	61.0	80.2	70.2	66.8
	Ours	31.5	21.1	39.8	40.0	37.0	33.9	69.3	57.5	56.8	37.8	74.8	82.8	41.2	67.3	74.0	77.4	70.9	40.9	57.3	73.5	69.3	63.4
10	YOLO-joint	0.0	0.0	0.0	0.0	9.1	1.8	76.9	77.1	62.2	47.3	79.4	85.1	51.3	70.1	78.6	78.0	75.2	47.4	63.9	85.0	72.3	70.0
	YOLO-ft	11.4	28.4	8.9	4.8	7.8	12.2	77.4	76.9	60.9	44.8	78.3	83.2	48.5	68.9	78.5	78.9	72.6	44.8	67.3	82.7	69.3	68.9
	YOLO-ft-full	22.3	53.9	32.9	40.8	43.2	38.6	71.9	69.8	57.1	41.0	76.9	81.7	43.6	65.3	77.3	79.2	70.1	41.5	63.7	76.9	69.1	65.7
	LSTD(YOLO)	11.3	32.2	5.6	1.3	7.7	11.6	77.1	75.2	62.0	44.5	78.2	84.2	49.9	68.6	78.8	78.8	72.6	45.0	66.9	82.6	69.5	68.9
	LSTD(YOLO)-full	22.8	52.5	31.3	45.6	40.3	38.5	70.9	71.3	59.8	41.1	77.1	81.9	45.1	67.2	78.0	78.9	70.7	41.6	63.8	79.7	66.8	66.3
	Ours	30.0	62.7	43.2	60.6	40.6	47.2	65.3	73.5	54.7	39.5	75.7	81.1	35.3	62.5	72.8	78.8	68.6	41.5	59.2	76.2	69.2	63.6

Table 1: Detection performance (AP) for the base and novel categories on the PASCAL VOC dataset for the 1st base/novel split. We evaluate the performance for different numbers of training examples for the novel categories.

# Shots		Novel						Base															
		aero	bottle	cow	horse	sofa	mean	bike	bird	boat	bus	car	cat	chair	table	dog	mbike	person	plant	sheep	train	tv	mean
1	YOLO-joint	0.0	0.0	0.0	0.0	0.0	0.0	78.8	73.2	63.6	79.0	79.7	87.2	51.5	71.2	81.1	78.1	75.4	47.7	65.9	84.0	73.7	72.7
	YOLO-ft	0.4	0.2	10.3	29.8	0.0	8.2	77.9	70.2	62.2	79.8	79.4	86.6	51.9	72.3	77.1	78.1	73.9	44.1	66.6	83.4	74.0	71.8
	YOLO-ft-full	0.6	9.1	11.2	41.6	0.0	12.5	74.9	67.2	60.1	78.8	79.0	83.8	50.6	72.7	75.5	74.8	71.7	43.9	62.5	81.8	72.6	70.0
	LSTD(YOLO)	0.5	0.1	11.1	37.7	0.0	9.9	76.9	69.8	61.5	78.2	81.0	85.7	51.9	73.7	79.6	76.7	73.4	43.8	66.0	82.2	74.1	71.6
	LSTD(YOLO)-full	0.1	1.5	10.4	44.9	0.0	11.4	76.1	68.0	58.7	78.1	79.0	85.0	50.7	72.2	76.2	75.2	71.8	43.3	62.7	82.8	72.2	70.1
	Ours	11.8	9.1	15.6	23.7	18.2	15.7	77.6	62.7	54.2	75.3	79.0	80.0	49.6	70.3	78.3	78.2	68.5	42.2	58.2	78.5	70.4	68.2
2	YOLO-joint	0.0	0.6	0.0	0.0	0.0	0.1	78.4	69.7	64.5	78.3	79.7	86.1	52.2	72.6	81.2	78.6	75.2	50.3	66.1	85.3	74.0	72.8
	YOLO-ft	0.2	0.2	17.2	1.2	0.0	3.8	78.1	70.0	60.6	79.8	79.4	87.1	49.7	70.3	80.4	78.8	73.7	44.2	62.2	82.4	74.9	71.4
	YOLO-ft-full	1.8	1.8	15.5	1.9	0.0	4.2	76.4	69.7	58.0	80.0	79.0	86.9	44.8	68.2	75.2	77.4	72.2	40.3	59.1	81.6	73.4	69.5
	LSTD(YOLO)	0.4	4.5	21.5	0.5	0.0	5.4	77.5	71.8	61.4	79.5	79.4	86.9	48.6	71.0	80.1	77.2	74.0	43.3	63.6	81.8	75.3	71.4
	LSTD(YOLO)-full	3.0	1.5	13.9	0.6	0.0	3.8	77.2	69.0	58.2	77.6	79.1	86.3	45.6	70.2	77.1	76.3	72.7	40.3	59.4	81.1	74.4	69.6
	Ours	28.6	0.9	27.6	0.0	19.5	15.3	75.8	67.4	52.4	74.8	76.6	82.5	44.5	66.0	79.4	76.2	68.2	42.3	53.8	76.6	71.0	67.2
3	YOLO-joint	0.0	0.0	0.0	0.0	0.0	0.0	77.6	72.2	61.2	77.9	79.8	85.8	49.9	73.2	80.0	77.9	75.3	50.8	64.3	84.2	72.6	72.2
	YOLO-ft	4.9	0.0	11.2	1.2	0.0	3.5	78.7	71.6	62.4	77.4	80.4	87.5	49.5	70.8	79.7	79.5	72.6	44.3	60.0	83.0	75.2	71.5
	YOLO-ft-full	10.7	4.6	12.9	29.7	0.0	11.6	74.9	69.2	60.4	79.4	79.1	87.3	43.4	69.7	75.8	75.2	70.5	39.4	52.9	80.8	73.4	68.8
	LSTD(YOLO)	4.5	0.1	10.8	0.8	0.0	3.2	78.4	71.5	60.9	78.5	80.2	87.7	47.8	70.4	79.5	77.8	73.1	42.9	58.9	81.6	75.4	71.0
	LSTD(YOLO)-full	12.6	0.7	11.3	0.4	0.0	5.0	75.5	69.7	61.0	79.5	79.1	87.8	43.2	68.5	76.0	75.7	71.0	41.2	61.2	80.9	73.3	69.6
	Ours	29.4	4.6	34.9	6.8	37.9	22.7	62.6	64.7	55.2	76.6	77.1	82.7	46.7	65.4	75.4	78.3	69.2	42.8	45.2	77.9	69.6	66.0
5	YOLO-joint	0.0	0.0	0.0	0.0	9.1	1.8	78.0	71.5	62.9	81.7	79.7	86.8	50.0	72.3	81.7	77.9	75.6	48.4	65.4	83.2	73.6	72.6
	YOLO-ft	0.8	0.2	11.3	5.2	0.0	3.5	78.6	72.4	61.5	79.4	81.0	87.8	48.6	72.1	81.0	79.6	73.6	44.9	61.4	83.9	74.7	72.0
	YOLO-ft-full	10.3	9.1	17.4	43.5	0.0	16.0	76.4	69.6	59.1	80.3	78.5	87.8	42.1	72.1	76.6	77.1	70.7	43.1	58.0	82.4	72.6	69.8
	LSTD(YOLO)	0.7	0.6	13.0	14.3	0.0	5.7	79.1	72.4	62.0	78.6	80.8	87.2	44.9	71.3	79.3	78.3	72.4	44.5	62.1	82.1	74.7	71.3
	LSTD(YOLO)-full	11.6	9.1	15.2	42.9	0.0	15.8	76.4	70.7	59.4	77.5	78.9	87.6	41.6	70.7	76.8	77.8	70.2	42.1	57.9	82.8	72.3	69.5
	Ours	33.1	9.4	38.4	25.4	44.0	30.1	73.2	65.6	52.9	75.9	77.5	80.0	43.7	65.0	73.8	78.4	68.9	39.2	56.4	78.0	70.8	66.6
10	YOLO-joint	0.0	0.0	0.0	0.0	0.0	0.0	77.4	71.5	61.1	78.8	82.7	87.1	52.5	74.6	80.8	79.3	75.4	46.1	64.2	85.2	73.6	72.7
	YOLO-ft	3.8	0.0	18.3	17.0	0.0	7.8	79.3	72.8	61.6	78.5	81.4	87.1	46.9	73.3	79.8	79.0	73.1	44.6	65.9	83.4	73.7	72.0
	YOLO-ft-full	41.7	9.5	34.5	45.1	38.4	33.9	75.5	69.4	60.0	78.3	78.8	86.8	44.9	68.4	75.8	76.9	70.7	44.0	64.1	81.6	71.1	69.8
	LSTD(YOLO)	31.2	9.1	22.3	25.6	7.8	19.2	78.8	72.5	62.3	78.5	80.9	86.8	47.4	70.8	79.6	78.6	72.7	44.2	66.5	83.7	73.3	71.8
	LSTD(YOLO)-full	41.5	9.3	29.2	38.9	36.1	31.0	74.6	70.2	59.6	77.3	78.6	86.5	45.1	68.1	77.6	75.2	70.6	44.5	59.8	79.7	71.2	69.2
	Ours	41.8	14.0	42.7	63.4	40.7	40.5	75.2	65.2	46.7	74.9	78.5	79.1	36.0	58.4	73.0	77.7	67.9	39.9	57.1	75.2	66.3	64.7

# Shots		Novel						Base															
		boat	cat	mbike	sheep	sofa	mean	aero	bike	bird	bottle	bus	car	chair	cow	table	dog	horse	person	plant	train	tv	mean
1	YOLO-joint	0.0	9.1	0.0	0.0	0.0	1.8	78.7	76.8	73.4	48.8	79.0	82.3	50.2	68.4	71.4	76.7	80.7	75.0	46.8	83.8	71.7	70.9
	YOLO-ft	0.1	25.8	10.7	3.6	0.1	8.1	77.2	74.9	69.1	47.4	78.7	79.7	47.9	68.3	69.6	74.7	79.4	74.2	42.2	82.7	71.1	69.1
	YOLO-ft-full	0.1	30.9	26.0	8.0	0.1	13.0	75.1	70.7	65.9	43.6	78.4	79.5	47.8	68.7	68.0	72.8	79.5	72.3	40.1	80.5	68.6	67.4
	LSTD(YOLO)	0.1	30.8	17.5	6.0	0.1	10.9	76.3	74.8	68.2	45.6	77.2	80.0	48.6	70.1	69.0	71.5	79.9	73.7	42.0	81.3	70.1	68.5
	LSTD(YOLO)-full	0.0	27.8	25.0	9.7	0.2	12.6	75.8	71.7	65.1	44.0	78.1	79.3	46.7	68.0	68.9	68.1	79.0	72.4	40.2	80.2	68.3	67.1
	Ours	10.3	41.4	29.1	16.2	9.4	21.3	77.6	72.6	65.7	39.6	77.0	78.2	49.7	53.9	64.6	67.4	79.3	67.2	41.0	82.5	72.5	65.9
2	YOLO-joint	0.0	9.1	0.0	0.0	0.0	1.8	77.6	77.1	74.0	49.4	79.8	79.9	50.5	71.0	72.7	76.3	81.0	75.0	48.4	84.9	72.7	71.4
	YOLO-ft	0.0	24.4	2.5	9.8	0.1	7.4	78.2	76.0	72.2	47.2	79.3	79.8	47.3	72.1	70.0	74.9	80.3	74.3	45.2	84.9	72.0	70.2
	YOLO-ft-full	0.0	35.2	28.7	15.4	0.1	15.9	75.3	72.0	69.8	44.0	79.1	78.8	42.1	70.0	64.9	73.8	81.7	71.4	40.9	80.9	69.4	67.6
	LSTD(YOLO)	0.0	25.4	0.0	12.6	0.1	7.6	78.1	76.4	71.9	46.8	78.8	79.6	45.3	70.6	66.9	75.3	81.7	73.9	43.4	84.1	71.8	69.7
	LSTD(YOLO)-full	0.2	27.3	0.1	15.0	0.2	8.5	77.4	73.3	69.5	44.8	78.5	79.2	43.0	69.2	66.4	71.9	82.0	72.3	39.8	84.5	69.3	68.1
	Ours	6.3	47.1	28.4	28.1	18.2	25.6	75.8	73.0	66.4	40.0	77.8	77.6	43.1	62.6	58.5	71.0	78.9	67.0	41.2	77.0	70.0	65.3
3	YOLO-joint	0.0	9.1	0.0	0.0	0.0	1.8	77.1	77.0	70.6	46.3	77.5	79.7	49.7	68.8	73.4	74.5	79.4	75.6	48.1	83.6	72.1	70.2
	YOLO-ft	0.0	27.0	1.8	9.1	0.1	7.6	77.7	76.6	71.4	47.5	78.0	79.9	47.6	70.0	70.5	74.4	80.0	73.7	44.1	83.0	70.9	69.7
	YOLO-ft-full	0.0	39.0	18.1	17.9	0.0	15.0	73.2	71.1	68.8	43.7	78.9	79.3	43.1	67.8	62.2	76.3	79.4	70.8	40.5	81.6	69.6	67.1
	LSTD(YOLO)	0.0	29.0	9.5	9.1	0.1	9.5	77.7	76.2	69.8	48.1	77.9	79.9	46.9	69.7	69.0	75.0	79.9	73.8	43.9	83.8	70.9	69.5
	LSTD(YOLO)-full	0.0	36.6	21.4	16.9	0.0	15.0	74.7	73.2	67.7	44.7	78.1	79.5	40.5	69.0	60.9	76.0	79.1	71.0	40.1	83.0	69.6	67.1
	Ours	11.7	48.2	17.4	34.7	30.1	28.4	73.2	69.3	66.5	41.8	77.6	76.3	42.8	61.1	63.7	67.3	77.4	68.2	39.7	78.6	70.7	65.0
5	YOLO-joint	0.0	9.1	0.0	0.0	9.1	3.6	78.2	78.5	72.1	47.8	76.6	82.1	50.7	70.1	71.8	77.6	80.4	75.4	46.0	84.8	72.5	71.0
	YOLO-ft	0.0	33.8	2.6	7.8	3.2	9.5	77.2	77.1	71.9	47.3	78.8	79.8	47.1	69.8	71.8	77.0	80.2	74.3	44.2	82.5	70.6	70.0
	YOLO-ft-full	7.9	48.0	39.1	29.4	36.6	32.2	75.5	73.6	69.1	43.3	78.4	78.9	42.3	70.2	66.1	77.4	79.8	72.2	41.9	82.8	69.3	68.1
	LSTD(YOLO)	0.0	39.1	12.4	15.8	9.2	15.3	77.6	76.8	71.0	46.3	78.2	79.9	46.2	71.3	69.8	77.7	80.0	74.3	45.1	83.1	71.2	69.9
	LSTD(YOLO)-full	0.2	51.5	37.2	26.9	20.7	27.3	74.5	73.5	69.1	42.9	78.4	79.2	42.3	69.3	66.0	77.7	79.6	71.7	41.6	82.8	69.2	67.8
	Ours	14.8	59.1	49.6	45.0	45.6	42.8	70.4	69.3	65.9	40.7	76.6	77.4	43.0	63.5	63.8	68.9	79.6	71.5	44.1	80.9	69.7	65.7
10	YOLO-joint	0.0	9.1	1.5	0.0	9.1	3.9	78.7	77.1	73.3	48.0	79.4	79.8	51.6	71.7	71.1	77.6	79.9	74.4	47.8	83.2	73.4	71.1
	YOLO-ft	0.0	35.6	1.0	14.2	1.5	10.5	78.0	77.8	69.0	46.0	78.5	79.6	45.3	69.9	70.9	77.0	80.8	74.2	45.2	83.3	70.7	69.8
	YOLO-ft-full	12.0	59.6	42.5	39.1	38.9	38.4	73.2	74.0	66.5	44.0	78.1	78.5	43.6	68.0	66.9	76.9	81.4	72.1	43.8	82.1	68.2	67.8
	LSTD(YOLO)	0.0	37.0	18.6	22.7	6.5	16.9	77.2	77.2	68.5	44.4	78.7	79.3	44.4	69.9	70.2	78.6	80.1	73.3	43.6	81.4	69.7	69.1
	LSTD(YOLO)-full	10.5	53.3	41.9	36.2	39.5	36.3	75.5	74.0	66.3	43.6	77.6	78.9	41.9	65.8	66.5	77.5	81.7	71.7	42.5	80.6	67.5	67.4
	Ours	22.4	59.9	59.8	46.1	41.4	45.9	69.4	67.6	67.1	39.3	69.5	76.6	33.4	60.1	58.5	70.1	79.2	67.8	43.5	78.8	65.0	63.1

Table 3: Detection performance (AP) for the base and novel categories on the PASCAL VOC dataset for the 3rd base/novel split. We evaluate the performance for different numbers of training examples for the novel categories.