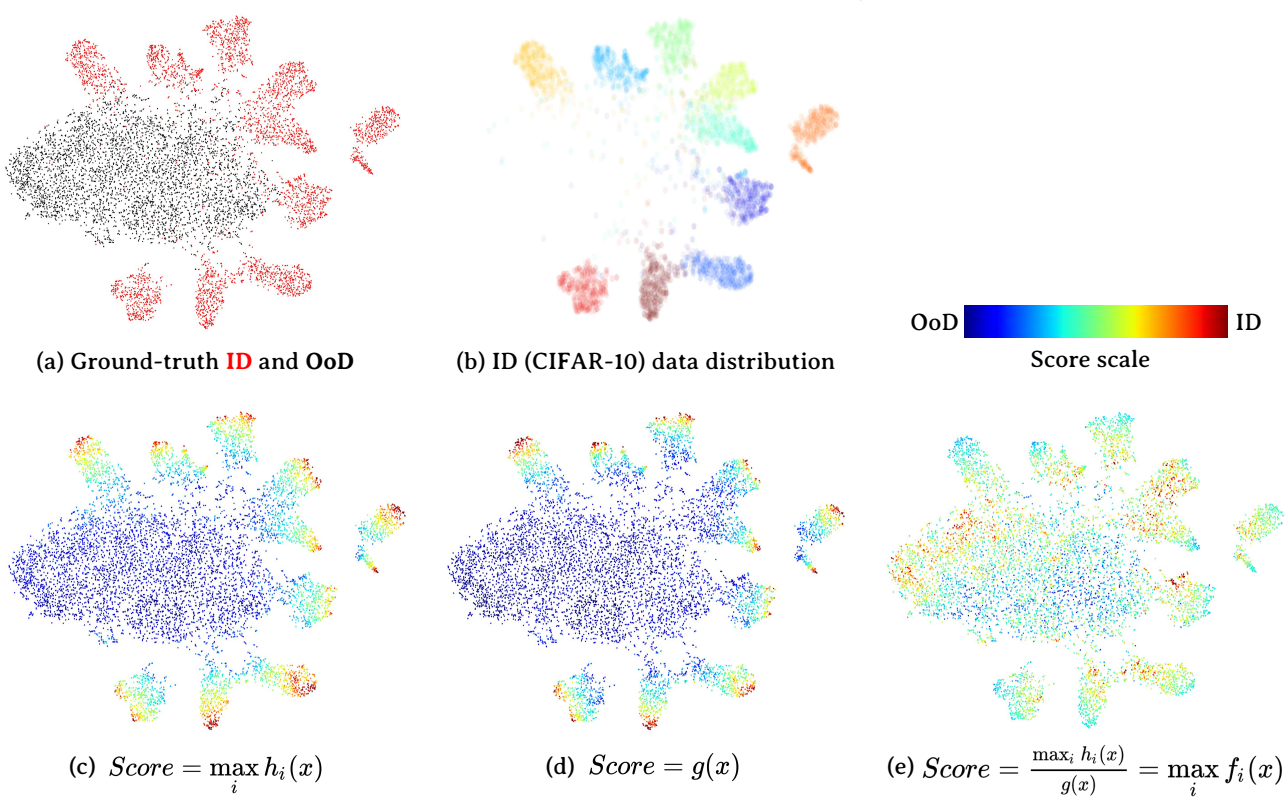
Generalized ODIN: Detecting Out-of-distribution Image without Learning from Out-of-distribution Data
- Supplementary Materials -

Yen-Chang Hsu[1], Yilin Shen[2], Hongxia Jin[2], Zsolt Kira[1]
[1]Georgia Institute of Technology, [2]Samsung Research America



(a) Ground-truth ID and OoD (b) ID (CIFAR-10) data distribution

(c) $Score = \max_i h_i(x)$ (d) $Score = g(x)$ (e) $Score = \frac{\max_i h_i(x)}{g(x)} = \max_i f_i(x)$

Figure 1: Visualization of the score distribution. The data are visualized with t-SNE using the features from the penultimate layer of the neural networks. The results are from DeConf-I with ResNet-34. The figure (a) visualizes the ground-truth in-distribution (ID, red, CIFAR-10) and out-of-distribution (OoD, black, Imagenet-resized) data. The colors in (b) represent different classes of CIFAR-10. The scores are obtained from (c) h function, (d) g function, or (e) the logits, and the scores are linearly re-scaled to between zero and one for visualization. The figure presents two phenomena. The first is that the OoD data in (e) have high scores. It is related to the overconfident effect discussed with equation 4. The second phenomenon is that high-score data in (c) and (d) are more significantly clustered in each class of CIFAR-10. It shows a tendency that the in-distribution data in high-density regions have higher scores than those in low-density regions (close to OoD data). This phenomenon is related to the discussion at the end of section 3.1.

Table 1: Performance of six OOD detection methods on 8 benchmark datasets. This is a full version of Table 1 in the main paper, which uses DenseNet for the backbone networks. All DeConf results are from the $h(x)$ branch. The value in parentheses is the standard deviation.

ID	OOD	AUROC
		Baseline / ODIN* / Maha* / DeConf-I* / DeConf-E* / DeConf-C*
CIFAR-100	Imagenet(c)	79.0(2.2) /90.5(1.1) /92.4(0.3) /84.4(2.3) /95.1(0.5) /97.6(0.2)
	Imagenet(r)	76.4(3.2) /91.1(1.3) /96.4(0.2) /81.2(3.6) /97.4(0.3) /98.6(0.2)
	LSUN(c)	78.6(1.1) /89.9(0.5) /81.2(0.6) /91.7(0.3) /90.1(0.3) /95.3(0.4)
	LSUN(r)	78.2(2.4) /93.0(0.8) /96.6(0.2) /84.1(2.1) /97.8(0.2) /98.7(0.0)
	iSUN	76.8(2.7) /91.6(1.1) /96.5(0.2) /82.1(2.9) /97.4(0.2) /98.4(0.0)
	SVHN	78.1(3.5) /85.6(0.0) /89.9(0.2) /89.7(0.4) /94.0(0.6) /95.9(0.7)
	Uniform	65.0(22.) /91.4(10.) /100.(0.0) /48.5(16.) /99.9(0.0) /99.9(0.0)
	Gaussian	48.0(28.) /62.0(38.) /100.(0.0) /6.79(4.9) /99.9(0.0) /99.9(0.0)
CIFAR-10	Imagenet(c)	92.1(1.0) /88.2(4.2) /96.3(0.1) /98.2(0.0) /98.0(0.2) /98.7(0.1)
	Imagenet(r)	91.5(1.4) /90.1(4.1) /98.2(0.0) /98.4(0.0) /98.2(0.2) /99.1(0.1)
	LSUN(c)	93.0(0.5) /91.3(2.0) /92.2(0.4) /98.4(0.0) /98.6(0.2) /98.3(0.2)
	LSUN(r)	93.9(0.4) /92.9(2.9) /98.2(0.0) /98.6(0.0) /98.8(0.0) /99.4(0.1)
	iSUN	93.0(0.7) /92.2(3.4) /98.2(0.0) /98.6(0.0) /98.8(0.0) /99.4(0.0)
	SVHN	88.1(4.8) /89.6(0.3) /98.0(0.3) /98.2(0.2) /98.4(0.6) /98.8(0.1)
	Uniform	95.4(0.7) /98.9(0.7) /99.9(0.0) /99.2(0.5) /99.9(0.0) /99.9(0.0)
	Gaussian	94.0(2.9) /98.6(1.7) /100.(0.0) /99.1(0.3) /99.9(0.0) /99.9(0.0)
ID	OOD	TNR@TPR95
		Baseline / ODIN* / Maha* / DeConf-I* / DeConf-E* / DeConf-C*
CIFAR-100	Imagenet(c)	25.3(2.8) /56.0(3.1) /63.5(2.1) /31.0(3.4) /74.6(2.8) /87.8(1.7)
	Imagenet(r)	22.3(3.1) /59.4(3.7) /82.0(1.6) /21.4(4.0) /87.6(1.7) /93.3(1.2)
	LSUN(c)	23.0(2.2) /53.0(1.0) /31.6(1.3) /59.6(1.9) /51.0(1.0) /75.0(1.9)
	LSUN(r)	23.7(2.5) /64.0(3.0) /82.6(1.8) /21.1(3.3) /89.8(1.5) /93.8(0.3)
	iSUN	21.5(2.8) /58.4(4.1) /81.2(1.4) /17.6(3.3) /87.3(1.2) /92.5(0.2)
	SVHN	18.9(4.9) /35.3(2.9) /43.3(2.7) /52.0(0.6) /67.1(3.4) /77.0(5.0)
	Uniform	2.95(4.1) /66.1(46.) /100.(0.0) /0.0(0.0) /100.(0.0) /100.(0.0)
	Gaussian	0.06(0.0) /33.3(47.) /100.(0.0) /0.0(0.0) /100.(0.0) /100.(0.0)
CIFAR-10	Imagenet(c)	50.0(2.8) /47.8(15.) /81.2(0.8) /92.0(0.2) /90.1(1.5) /93.4(1.2)
	Imagenet(r)	47.4(4.4) /51.9(16.) /90.9(0.5) /93.6(0.2) /91.7(1.6) /95.8(0.9)
	LSUN(c)	51.8(3.1) /63.5(7.8) /64.2(0.6) /92.5(0.4) /93.3(1.5) /91.5(1.2)
	LSUN(r)	56.3(3.6) /59.2(18.) /91.7(0.3) /94.9(0.2) /95.7(0.1) /97.6(0.5)
	iSUN	52.3(3.6) /57.2(18.) /90.6(0.7) /94.6(0.3) /95.4(0.2) /97.5(0.3)
	SVHN	40.5(6.9) /48.7(3.2) /90.6(1.7) /91.4(1.1) /92.1(3.4) /94.0(0.6)
	Uniform	59.9(12.) /98.1(2.6) /100.(0.0) /99.9(0.0) /100.(0.0) /100.(0.0)
	Gaussian	48.8(26.) /92.1(11.) /100.(0.0) /99.9(0.0) /100.(0.0) /100.(0.0)

Table 2: Performance of six OOD detection methods on 8 benchmark datasets. The experiment here is the same as Table 1 but use Resnet-34 for the backbone network. All DeConf results are from the $h(x)$ branch. The value in parentheses is the standard deviation.

ID	OOD	AUROC
		Baseline / ODIN* / Maha* / DeConf-I* / DeConf-E* / DeConf-C*
CIFAR-100	Imagenet(c)	78.9(0.1) /84.8(0.6) /93.4(0.3) /88.2(0.6) /95.2(0.6) /95.3(0.6)
	Imagenet(r)	75.1(0.8) /85.7(0.2) /96.3(0.1) /84.6(1.0) /97.0(0.4) /95.9(0.7)
	LSUN(c)	78.8(0.6) /80.3(1.3) /79.8(0.3) /93.8(0.3) /92.6(0.2) /93.8(0.3)
	LSUN(r)	76.2(1.4) /86.6(0.8) /96.3(0.2) /85.9(1.8) /97.0(0.7) /96.1(0.5)
	iSUN	75.2(1.4) /85.9(0.8) /95.8(0.2) /84.7(1.4) /96.6(0.6) /95.7(0.5)
	SVHN	75.1(2.5) /80.2(2.0) /80.9(1.1) /89.2(2.6) /93.8(0.8) /93.2(1.1)
	Uniform	69.0(13.) /96.7(2.5) /100.(0.0) /79.3(8.3) /99.9(0.0) /99.9(0.0)
	Gaussian	51.5(1.8) /93.7(1.7) /99.9(0.0) /60.8(23.) /99.9(0.0) /99.9(0.0)
CIFAR-10	Imagenet(c)	90.0(0.9) /81.2(2.4) /94.2(0.1) /98.2(0.2) /98.2(0.1) /96.0(0.2)
	Imagenet(r)	87.3(1.3) /81.1(2.9) /96.5(0.1) /98.1(0.3) /98.1(0.3) /96.1(0.5)
	LSUN(c)	92.0(1.7) /77.9(4.6) /87.7(0.2) /98.8(0.1) /98.5(0.0) /97.2(0.1)
	LSUN(r)	91.6(1.2) /88.5(2.0) /97.2(0.1) /98.9(0.2) /99.0(0.1) /98.0(0.1)
	iSUN	90.1(1.4) /86.1(2.5) /96.5(0.2) /98.8(0.2) /98.9(0.1) /97.6(0.1)
	SVHN	87.7(2.4) /63.9(4.3) /87.8(1.6) /96.8(0.4) /96.1(1.4) /97.8(0.3)
	Uniform	85.9(10.) /93.3(4.5) /99.9(0.0) /99.6(0.1) /99.9(0.0) /99.9(0.0)
	Gaussian	89.9(10.) /97.1(2.0) /99.9(0.0) /99.7(0.0) /99.9(0.0) /99.9(0.0)
ID	OOD	TNR@TPR95
		Baseline / ODIN* / Maha* / DeConf-I* / DeConf-E* / DeConf-C*
CIFAR-100	Imagenet(c)	24.1(0.6) /44.0(2.2) /68.2(1.4) /42.6(2.7) /73.4(3.7) /72.6(3.7)
	Imagenet(r)	19.4(0.1) /45.5(1.4) /82.6(0.8) /30.4(3.0) /84.3(2.7) /76.5(3.8)
	LSUN(c)	21.9(0.4) /34.8(2.4) /27.7(1.4) /66.1(2.2) /59.7(0.7) /65.7(2.3)
	LSUN(r)	19.8(1.6) /48.2(3.0) /81.8(1.4) /29.4(5.2) /84.6(4.0) /76.8(3.3)
	iSUN	17.7(0.5) /45.3(2.8) /80.4(0.8) /27.1(4.3) /83.0(3.1) /75.3(3.3)
	SVHN	16.6(1.5) /27.5(5.0) /25.7(2.6) /43.7(10.) /60.8(5.3) /55.1(7.1)
	Uniform	5.63(7.0) /76.4(27.) /100.(0.0) /4.11(5.8) /100.(0.0) /100.(0.0)
	Gaussian	0.0(0.0) /46.6(20.) /100.(0.0) /0.06(0.0) /100.(0.0) /100.(0.0)
CIFAR-10	Imagenet(c)	54.6(2.6) /53.7(3.1) /74.6(0.6) /90.8(1.5) /91.1(0.9) /81.1(1.7)
	Imagenet(r)	48.3(3.2) /53.1(4.3) /85.1(0.6) /90.5(1.8) /90.8(1.8) /81.4(2.4)
	LSUN(c)	59.9(4.7) /50.9(6.1) /53.6(1.0) /93.9(0.5) /92.4(0.5) /87.3(1.0)
	LSUN(r)	57.5(4.4) /68.1(4.2) /87.4(0.8) /95.8(1.0) /96.0(0.7) /90.9(0.9)
	iSUN	53.7(3.8) /62.8(5.0) /84.6(0.9) /95.1(1.0) /95.3(0.5) /88.8(1.1)
	SVHN	44.5(8.1) /29.7(6.2) /46.2(4.8) /84.5(2.5) /78.8(7.6) /89.5(2.1)
	Uniform	27.9(20.) /74.5(20.) /100.(0.0) /100.(0.0) /100.(0.0) /100.(0.0)
	Gaussian	52.7(40.) /87.1(9.3) /100.(0.0) /100.(0.0) /100.(0.0) /100.(0.0)

Table 3: The AUROC of individual experimental setting in Figures 3 and 4 . The experiments do not use input preprocessing. All values are percentages averaged over three runs, and the value in parentheses is the standard deviation. The "+" means that the classifier is trained with extra regularization (dropout rate 0.7).

ID	OoD	Plain-I	DeConf-I-$h(x)$	DeConf-I-$g(x)$	Plain-E	DeConf-E-$h(x)$	DeConf-E-$g(x)$	Plain-C	DeConf-C-$h(x)$	DeConf-C-$g(x)$
SVHN	Imagenet(c)	92.8(1.0)	98.7(0.1)	98.4(0.0)	92.9(0.8)	96.2(0.4)	97.2(0.7)	60.3(1.4)	93.9(0.4)	91.3(6.0)
	Imagenet(r)	92.4(0.8)	98.7(0.1)	98.4(0.0)	93.0(0.6)	96.1(0.4)	97.0(0.9)	63.9(1.5)	93.5(0.7)	90.8(6.0)
	LSUN(c)	91.0(0.6)	98.0(0.2)	97.5(0.5)	92.0(0.6)	95.0(0.1)	95.5(0.4)	51.9(0.3)	93.1(0.3)	92.4(5.8)
	LSUN(r)	91.3(1.0)	98.4(0.2)	98.3(0.3)	92.1(0.7)	95.5(0.7)	96.7(1.1)	61.4(1.6)	92.0(0.7)	90.6(5.8)
	iSUN	91.5(0.9)	98.6(0.1)	98.5(0.2)	92.4(0.8)	95.8(0.6)	96.7(1.1)	60.9(1.6)	93.1(0.8)	91.3(5.8)
	CIFAR10	91.4(0.7)	98.4(0.1)	98.0(0.0)	92.6(0.5)	95.6(0.4)	96.9(0.5)	63.1(1.7)	93.3(0.6)	89.5(5.8)
	CIFAR100	91.3(0.4)	98.1(0.1)	97.3(0.1)	92.5(0.5)	95.0(0.5)	96.3(0.6)	64.0(1.6)	93.0(0.5)	88.5(6.1)
	Uniform	93.7(1.6)	98.7(0.2)	98.6(0.3)	93.0(0.7)	94.8(0.9)	93.9(1.8)	64.6(2.4)	95.1(1.6)	90.1(6.4)
	Gaussian	94.4(1.3)	98.8(0.2)	98.7(0.3)	93.6(0.3)	95.6(0.6)	95.4(1.6)	66.5(3.4)	95.1(1.4)	90.3(6.5)
CIFAR-10	Imagenet(c)	90.0(0.9)	97.7(0.4)	96.6(0.7)	92.3(0.4)	97.4(0.2)	96.6(0.7)	74.3(2.9)	96.4(0.3)	87.3(10.)
	Imagenet(r)	87.3(1.3)	96.9(0.6)	95.5(1.0)	91.2(0.3)	96.8(0.4)	95.4(1.4)	71.7(3.1)	95.6(0.5)	86.0(12.)
	LSUN(c)	92.0(1.7)	99.0(0.0)	98.8(0.0)	93.7(0.5)	98.7(0.0)	98.7(0.1)	79.9(3.1)	98.4(0.0)	92.0(5.7)
	LSUN(r)	91.6(1.2)	98.2(0.4)	96.1(1.2)	93.5(0.3)	98.1(0.1)	95.8(1.3)	72.8(2.9)	97.6(0.2)	85.2(14.)
	iSUN	90.1(1.4)	98.0(0.4)	96.2(1.1)	92.9(0.4)	97.9(0.1)	96.0(1.2)	71.0(3.7)	97.3(0.3)	86.0(13.)
	SVHN	87.7(2.4)	98.3(0.4)	99.3(0.3)	91.7(0.7)	97.5(0.8)	99.0(0.3)	80.6(5.1)	98.6(0.5)	92.5(4.3)
	Uniform	85.9(10.)	93.5(1.0)	97.7(1.6)	88.6(2.5)	99.2(0.6)	93.3(9.0)	75.1(14.)	99.6(0.1)	87.2(5.2)
	Gaussian	89.9(10.)	94.6(1.4)	98.3(0.9)	89.6(5.1)	99.2(0.5)	88.6(15.)	77.8(3.4)	99.6(0.2)	85.1(1.5)
CIFAR-100	Imagenet(c)	78.9(0.1)	83.2(0.6)	63.7(6.6)	76.0(1.5)	93.4(0.7)	57.0(11.)	64.6(0.3)	92.6(0.8)	46.3(10.)
	Imagenet(r)	75.1(0.8)	76.6(1.4)	50.1(9.1)	72.0(1.8)	95.5(0.6)	46.6(15.)	61.8(1.5)	91.8(1.1)	51.3(14.)
	LSUN(c)	78.8(0.6)	91.3(0.5)	85.7(1.4)	77.5(0.3)	90.1(0.6)	78.1(3.0)	60.5(1.0)	93.3(0.7)	35.6(1.4)
	LSUN(r)	76.2(1.4)	78.4(2.5)	46.0(10.)	71.3(0.7)	95.5(0.8)	43.0(14.)	64.1(1.8)	92.0(0.7)	44.6(12.)
	iSUN	75.2(1.4)	76.6(2.0)	45.7(10.)	71.4(1.1)	95.2(0.7)	40.0(15.)	61.9(1.9)	91.6(0.8)	45.0(14.)
	SVHN	75.1(2.5)	89.6(2.0)	87.9(3.3)	77.5(2.2)	91.5(1.7)	75.3(7.8)	60.0(5.2)	93.6(1.3)	54.3(4.8)
	Uniform	69.0(13.)	50.4(11.)	46.8(26.)	84.0(10.)	99.8(0.1)	25.0(17.)	59.2(36.)	99.6(0.1)	97.7(0.8)
	Gaussian	51.5(1.8)	31.9(17.)	27.0(19.)	84.8(5.7)	99.9(0.0)	7.75(4.2)	23.5(14.)	99.3(0.3)	99.4(0.4)
CIFAR-100+	Imagenet(c)	77.0(1.4)	87.0(0.1)	82.1(1.2)	78.2(0.4)	86.8(1.2)	83.1(1.8)	69.4(3.4)	88.3(1.1)	81.6(0.7)
	Imagenet(r)	73.7(1.4)	83.8(0.6)	76.5(2.6)	76.3(0.5)	84.3(1.5)	78.0(2.7)	72.4(2.8)	87.0(1.1)	75.4(1.2)
	LSUN(c)	77.6(0.5)	89.3(0.5)	89.8(0.8)	76.5(1.1)	90.0(0.3)	90.7(0.7)	55.0(2.3)	86.2(1.3)	89.6(1.1)
	LSUN(r)	75.4(2.6)	84.6(1.3)	76.8(1.9)	75.8(2.0)	84.5(0.4)	78.0(2.8)	70.1(3.8)	87.0(2.2)	74.8(2.2)
	iSUN	74.5(1.6)	83.6(0.8)	76.2(2.4)	74.9(1.1)	84.0(1.1)	77.7(2.0)	67.9(3.8)	85.7(1.7)	74.5(1.8)
	SVHN	72.2(5.9)	83.2(2.9)	81.4(5.2)	74.5(3.4)	86.0(2.6)	83.8(3.4)	67.8(2.4)	86.3(2.7)	84.9(2.2)
	Uniform	87.0(1.5)	85.1(6.1)	75.6(14.)	88.7(3.2)	95.0(2.1)	83.6(11.)	60.0(15.)	81.7(14.)	57.3(2.4)
	Gaussian	87.1(4.7)	84.2(10.)	81.6(14.)	77.7(4.3)	95.4(4.7)	85.1(10.)	70.0(2.2)	75.9(4.6)	34.9(4.6)

Table 4: The summary of classifiers analyzed in the experiment section. Their in-domain classification accuracy is provided in the right four columns. The "+" means that the classifier is trained with extra regularization (dropout rate 0.7).

Classifier	Image size	#class	Model	Experiment	Baseline	DeConf-I	DeConf-E	DeConf-C
CIFAR-10	32x32	10	DenseNet	Table 1,2	95.2±0.1	94.9±0.1	95.0±0.1	95.0±0.1
CIFAR-10	32x32	10	ResNet34	Figure 3	95.2±0.1	95.0±0.1	94.9±0.1	95.1±0.1
SVHN	32x32	10	ResNet34	Figure 3	96.9±0.1	96.8±0.1	96.5±0.1	96.7±0.1
CIFAR-100	32x32	100	DenseNet	Table 1,2; Figure 7	77.0±0.2	75.8±0.4	76.4±0.1	75.9±0.1
CIFAR-100	32x32	100	WRN	Figure 7	80.8±0.1	78.3±0.1	78.4±0.1	78.4±0.1
CIFAR-100	32x32	100	ResNet50	Figure 7	78.8±0.3	76.4±0.1	76.5±0.3	76.2±0.2
CIFAR-100	32x32	100	ResNet34	Figure 4,5,7	78.5±0.2	76.0±0.1	76.2±0.1	75.8±0.2
CIFAR-100	32x32	100	ResNet18	Figure 7	77.3±0.1	75.2±0.2	75.8±0.1	75.1±0.1
CIFAR-100	32x32	100	ResNet10	Figure 7	75.0±0.1	73.4±0.1	74.2±0.1	73.5±0.1
CIFAR-100$^+$	32x32	100	ResNet34	Figure 4	78.2±0.1	77.4±0.3	77.2±0.3	77.2±0.1
DomainNet (Real-A)	180x180 to 640x880	173	ResNet34	Table 3	73.6±0.1	73.0±0.1	73.4±1.5	72.2±0.5