Table 4. Comparison of baseline method (Hendrycks & Gimpel, 2017) and confidence-based thresholding. All models are trained on SVHN, which is used as the in-distribution dataset (average of 5 runs). All values are shown in percentages. ↓ indicates that lower values are better, while ↑ indicates that higher scores are better.

	Out-of-distribution dataset	FPR (95% TPR) ↓	Detection Error ↓	AUROC ↑	AUPR In ↑	AUPR Out ↑
		Baseline (Hendrycks & Gimpel, 2017)/Confidence Thresholding				
	TinyImageNet (crop)	12.9/**3.1**	7.0/**4.0**	97.4/**99.2**	99.0/**99.5**	93.2/**98.0**
	TinyImageNet (resize)	7.2/**1.5**	5.3/**2.8**	98.4/**99.5**	99.4/**99.8**	95.6/**98.7**
	LSUN (crop)	28.3/**14.0**	11.7/**8.4**	93.7/**96.6**	96.8/**98.1**	85.6/**93.2**
DenseNet-BC	LSUN (resize)	6.0/**1.0**	4.9/**2.3**	98.6/**99.7**	99.5/**99.9**	96.0/**99.0**
	iSUN	6.0/**0.9**	4.9/**2.3**	98.6/**99.7**	99.5/**99.9**	95.7/**98.8**
	Uniform	11.9/**0.1**	5.2/**1.2**	97.9/**99.9**	99.3/**100.0**	93.4/**99.6**
	Gaussian	7.4/**0.0**	4.1/**0.9**	98.5/**99.9**	99.5/**100.0**	95.1/**99.7**
	All Images	12.2/**4.2**	7.2/**4.5**	97.3/**98.9**	95.1/**97.4**	98.4/**99.4**
	TinyImageNet (crop)	16.3/**4.1**	7.8/**4.4**	96.9/**99.1**	98.7/**99.5**	91.4/**98.1**
	TinyImageNet (resize)	10.6/**1.5**	6.1/**2.7**	97.8/**99.6**	99.2/**99.8**	93.6/**99.2**
	LSUN (crop)	31.9/**18.7**	12.5/**10.5**	93.0/**95.5**	96.4/**97.5**	84.0/**91.8**
WRN-16-8	LSUN (resize)	9.5/**0.6**	5.8/**1.8**	98.0/**99.8**	99.3/**99.9**	94.0/**99.5**
	iSUN	9.6/**0.8**	5.9/**2.1**	98.0/**99.8**	99.3/**99.9**	93.4/**99.4**
	Uniform	17.7/**0.3**	7.1/**1.2**	97.1/**99.9**	99.0/**100.0**	91.1/**99.6**
	Gaussian	11.0/**0.2**	5.8/**1.0**	97.9/**99.9**	99.3/**100.0**	93.7/**99.8**
	All Images	15.7/**5.3**	7.9/**5.0**	96.7/**98.7**	94.1/**96.8**	97.9/**99.4**
	TinyImageNet (crop)	17.3/**4.3**	7.7/**4.6**	96.9/**99.2**	98.8/**99.7**	91.3/**98.1**
	TinyImageNet (resize)	11.4/**1.8**	6.2/**3.1**	97.8/**99.6**	99.2/**99.8**	93.7/**99.1**
	LSUN (crop)	22.7/**13.0**	9.4/**7.8**	95.6/**97.6**	98.1/**99.0**	88.6/**94.7**
VGG13	LSUN (resize)	9.4/**0.8**	5.7/**2.0**	98.1/**99.8**	99.3/**99.9**	94.3/**99.6**
	iSUN	10.0/**1.0**	6.0/**2.2**	98.0/**99.8**	99.3/**99.9**	93.7/**99.5**
	Uniform	20.0/**0.5**	7.3/**1.4**	96.8/**99.9**	98.9/**100.0**	90.2/**99.7**
	Gaussian	12.9/**0.3**	6.0/**0.9**	97.8/**99.9**	99.2/**100.0**	93.1/**99.9**
	All Images	14.2/**4.3**	7.1/**4.6**	97.3/**99.2**	95.9/**98.5**	98.2/**99.6**