How Robust are Discriminatively Trained Zero-Shot Learning Models?

Supplementary Material

Mehmet Kerim Yucel[a], Ramazan Gokberk Cinbis[b], Pinar Duygulu[a]

[a] *Hacettepe University, Graduate School of Science and Engineering, 06800 Ankara, Turkey*
[b] *Department of Computer Engineering, Middle East Technical University, 06800 Ankara, Turkey*

---

1. Mean Corruption Errors

We provide mean corruption error (MCE) metrics for our corruption robustness experiments. These results are complementary to the relative, corruption-induced accuracy reduction ratios presented in the main manuscript and are given to establish a numerical baseline. We follow the calculation described in [1], but instead of AlexNet errors, we use the error of the original ALE model used in the paper. These average errors are given in Tables 1, 2, 3 for AWA2-C, CUB-C and SUN-C datasets, respectively. We provide MCE metrics based on ZSL top-1 scores in Tables 4, 8, 12, GZSL unseen scores in Tables 5, 9, 13, GZSL seen scores in Tables 6, 10, 14 and GZSL harmonic scores in Tables 7, 11, 15 for AWA2-C, CUB-C and SUN-C datasets, respectively.

In Tables 4 to 15, *OR* refers to the original ALE model, *SS* refers to spatial smoothing, *TVM* refers to total variance minimization and *LS* refers to label smoothing defenses. We note that since spatial smoothing and total variance minimization are preprocessing defenses, their *clean* (uncorrupted) errors are the same as the original model. Corruption types given in the Tables are (from left to right) Gaussian noise, shot noise, impulse noise, defocus blur, glass blur, motion blur, zoom blur, snow, frost, fog, brightness, contrast, elastic transformations, pixelate and JPEG compression corruptions.

Email addresses: mkerimyucel@hacettepe.edu.tr (Mehmet Kerim Yucel), gcinbis@ceng.metu.edu.tr (Ramazan Gokberk Cinbis), pinar@cs.hacettepe.edu.tr (Pinar Duygulu)

	Noise			Blur				Weather				Digital			
AWA2-C	Gaussian	Shot	Impulse	Defocus	Glass	Motion	Zoom	Snow	Frost	Fog	Bright	Contrast	Elastic	Pixelate	JPEG
ZSL	51.9	52.7	55.0	59.4	59.4	56.5	61.0	64.8	53.6	52.2	38.8	63.3	51.1	50.3	42.3
Unseen	93.3	94.1	94.8	92.9	94.0	91.8	96.3	95.4	92.8	92.3	85.8	95.0	91.7	90.7	89.2
Seen	46.9	50.3	53.1	65.1	61.7	55.6	71.6	65.5	52.0	44.8	24.5	63.3	39.7	42.1	29.8
H-Score	88.1	89.5	90.6	88.3	89.6	86.1	93.5	91.9	87.4	86.5	76.1	91.3	85.5	84.1	81.3

Table 1: AWA2-C average errors per corruption category.

	Noise			Blur				Weather				Digital			
CUB-C	Gaussian	Shot	Impulse	Defocus	Glass	Motion	Zoom	Snow	Frost	Fog	Bright	Contrast	Elastic	Pixelate	JPEG
ZSL	77.0	77.6	80.8	63.2	71.6	61.0	65.5	77.5	71.1	65.0	53.3	72.9	59.8	61.0	59.3
Unseen	90.8	91.4	93.0	86.8	89.3	84.9	86.2	90.8	88.9	87.7	77.9	90.6	82.4	83.5	83.3
Seen	76.9	78.8	82.0	64.5	72.6	60.6	67.0	81.2	73.4	61.7	47.3	70.6	57.2	62.3	57.1
H-Score	86.8	87.8	89.9	80.8	84.7	78.2	80.6	87.7	84.3	81.4	68.9	85.8	75.0	77.1	76.0

Table 2: CUB-C average errors per corruption category.

	Noise			Blur				Weather				Digital			
SUN-C	Gaussian	Shot	Impulse	Defocus	Glass	Motion	Zoom	Snow	Frost	Fog	Bright	Contrast	Elastic	Pixelate	JPEG
ZSL	76.0	79.1	81.0	63.1	79.0	67.2	63.5	79.5	75.5	57.2	50.7	65.1	70.2	61.2	54.9
Unseen	93.3	94.1	94.8	89.6	93.8	90.8	93.3	95.0	94.0	85.7	83.2	89.4	91.1	87.8	85.3
Seen	90.4	91.7	93.4	84.5	90.8	86.3	90.1	93.5	90.0	79.5	75.7	84.0	87.2	83.6	78.3
H-Score	92.2	93.1	94.2	87.6	92.7	89.0	92.0	94.4	92.5	83.1	80.1	87.3	89.5	86.0	82.5

Table 3: SUN-C average errors per corruption category.

			Noise			Blur				Weather				Digital			
Model	Clean	**mCE**	Gauss	Shot	Imp.	Defo.	Glass	Mot.	Zoom	Snow	Frost	Fog	Brigh.	Cont.	Elas.	Pixel	JPEG
OR	38.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
SS	38.0	106.6	121.0	118.7	90.9	104.2	102.7	105.0	106.8	108.3	108.9	109.2	106.2	104.7	102.8	103.6	106.5
TVM	38.0	102.7	99.8	100.0	103.3	100.5	100.1	101.7	101.9	106.3	103.3	106.3	105.2	104.3	103.1	101.0	103.4
LS	39.4	104.6	103.1	104.6	104.7	102.6	103.0	107.1	103.0	103.6	106.2	108.1	104.7	105.4	102.8	105.2	104.9
AM	55.1	107.0	109.6	108.6	106.4	101.8	103.0	104.9	106.9	101.8	104.0	104.8	119.0	102.5	107.4	107.5	116.5
ANT	54.9	104.9	100.5	101.0	98.1	101.8	102.4	105.8	106.7	102.3	108.0	105.4	118.7	104.2	104.9	103.3	110.7

Table 4: AWA2-C mCE values based on ZSL top-1 accuracy.

			Noise			Blur				Weather				Digital			
Model	Clean	**mCE**	Gauss	Shot	Imp.	Defo.	Glass	Mot.	Zoom	Snow	Frost	Fog	Brigh.	Cont.	Elas.	Pixel	JPEG
OR	84.7	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
SS	84.7	100.9	101.8	102.0	96.2	100.6	100.6	101.2	100.4	102.1	101.4	101.7	101.7	101.0	101.3	100.5	100.5
TVM	84.7	100.9	100.5	100.5	100.6	100.2	100.4	100.7	99.8	101.3	101.2	101.8	102.6	101.0	101.1	100.2	100.9
LS	83.7	98.2	98.5	98.7	98.1	98.1	98.9	98.6	98.3	98.1	98.5	97.6	97.6	98.6	98.6	97.2	97.9
AM	83.9	98.5	96.6	96.5	96.9	99.6	97.1	99.0	97.3	100.9	100.4	99.2	98.6	98.1	99.3	99.4	98.3
ANT	80.7	95.3	90.5	89.5	91.3	98.7	96.5	98.2	97.8	98.9	95.9	95.7	94.7	95.9	95.3	96.1	94.3

Table 5: AWA2-C mCE values based on unseen accuracy.

			Noise			Blur				Weather				Digital			
Model	Clean	**mCE**	Gauss	Shot	Imp.	Defo.	Glass	Mot.	Zoom	Snow	Frost	Fog	Brigh.	Cont.	Elas.	Pixel	JPEG
OR	21.2	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
SS	21.2	113.7	143.0	133.0	74.0	106.1	105.4	112.8	107.9	117.1	117.2	120.8	122.2	110.1	112.6	107.6	115.6
TVM	21.2	106.7	113.3	109.9	107.5	100.7	100.1	103.6	100.7	111.6	105.4	114.9	111.9	106.5	107.5	101.5	104.9
LS	25.8	106.1	107.4	105.8	105.6	103.1	103.2	104.6	101.6	102.3	104.7	105.9	117.1	103.2	107.3	107.0	113.1
AM	16.5	86.2	87.5	86.1	86.8	90.3	94.4	87.1	90.6	89.4	84.2	78.3	78.4	89.3	83.6	86.2	81.1
ANT	14.6	83.8	71.3	70.2	72.5	89.2	92.9	88.2	94.1	96.8	92.0	85.3	72.4	95.5	83.2	81.3	72.7

Table 6: AWA2-C mCE values based on seen accuracy.

			Noise			Blur				Weather				Digital			
Model	Clean	**mCE**	Gauss	Shot	Imp.	Defo.	Glass	Mot.	Zoom	Snow	Frost	Fog	Brigh.	Cont.	Elas.	Pixel	JPEG
OR	74.3	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
SS	74.3	101.7	103.6	103.7	93.4	101.2	101.1	102.3	100.8	103.7	102.5	103.1	103.0	101.8	102.3	100.9	101.1
TVM	74.3	101.5	101.1	100.9	101.2	100.3	100.7	101.3	99.8	102.4	102.1	103.3	104.4	101.8	102.0	100.4	101.6
LS	73.2	97.4	97.7	98.0	97.0	97.7	98.4	98.2	97.5	97.3	97.8	96.4	96.6	98.1	97.9	95.8	96.9
AM	73.0	97.4	94.4	94.2	94.8	98.9	95.5	98.1	95.6	101.5	100.5	98.4	97.5	96.9	98.7	98.9	97.0
ANT	68.4	92.4	84.9	83.5	86.2	97.4	94.6	96.9	96.4	98.3	93.6	93.0	91.2	94.1	92.2	93.5	90.5

Table 7: AWA2-C mCE values based on harmonic accuracy.

			Noise			Blur				Weather				Digital			
Model	Clean	**mCE**	Gauss	Shot	Imp.	Defo.	Glass	Mot.	Zoom	Snow	Frost	Fog	Brigh.	Cont.	Elas.	Pixel	JPEG
OR	45.5	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
SS	45.5	106.3	107.1	107.2	86.4	106.2	104.6	105.8	106.5	110.7	113.5	114.2	107.2	106.6	106.5	107.2	104.5
TVM	45.5	100.8	96.7	98.0	99.5	101.1	99.4	100.5	101.2	102.6	98.9	104.5	104.7	102.6	103.0	100.5	99.0
LS	47.8	101.3	104.1	105.2	102.8	103.0	99.6	98.5	98.4	95.4	99.7	106.2	98.2	101.6	99.8	102.5	104.7
AM	48.4	102.1	101.7	102.8	101.5	101.1	104.4	101.6	99.0	103.5	101.4	99.1	104.2	100.3	102.9	103.0	105.2
ANT	51.1	100.8	97.3	99.9	97.8	105.0	100.9	104.7	102.3	94.1	95.8	99.7	106.5	99.4	99.8	102.3	105.7

Table 8: CUB-C mCE values based on ZSL top-1 accuracy.

			Noise			Blur				Weather				Digital			
Model	Clean	**mCE**	Gauss	Shot	Imp.	Defo.	Glass	Mot.	Zoom	Snow	Frost	Fog	Brigh.	Cont.	Elas.	Pixel	JPEG
OR	74.4	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
SS	74.4	103.0	104.3	103.9	95.3	102.5	101.9	103.1	102.8	104.2	104.5	104.2	105.4	102.2	103.3	103.5	103.2
TVM	74.4	100.3	99.5	99.4	99.7	100.5	99.1	100.6	100.7	100.7	100.1	101.5	101.6	100.5	101.5	100.2	99.3
LS	77.3	101.8	102.6	102.2	102.0	101.8	99.7	101.9	102.1	99.3	101.3	102.8	101.5	102.7	101.4	102.4	103.1
AM	72.8	98.3	99.5	99.4	99.1	97.8	98.8	97.8	97.5	99.8	98.2	97.6	96.4	98.5	98.7	97.7	98.7
ANT	73.9	97.2	96.2	96.9	96.1	99.0	99.2	97.2	97.2	96.7	95.7	96.6	99.1	98.0	97.2	96.8	96.8

Table 9: CUB-C mCE values based on unseen accuracy.

			Noise			Blur				Weather				Digital			
Model	Clean	**mCE**	Gauss	Shot	Imp.	Defo.	Glass	Mot.	Zoom	Snow	Frost	Fog	Brigh.	Cont.	Elas.	Pixel	JPEG
OR	35.4	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
SS	35.4	109.2	111.2	108.9	84.2	108.2	107.5	110.4	110.1	110.1	113.8	120.4	113.5	110.3	111.5	110.8	107.9
TVM	35.4	102.1	99.3	99.3	99.8	102.1	99.9	101.8	100.3	103.2	99.7	105.6	109.2	104.0	105.7	101.7	99.2
LS	43.8	105.7	105.6	106.0	104.4	108.4	101.1	105.7	104.8	96.3	102.9	110.9	112.6	107.3	104.7	104.5	109.7
AM	39.9	103.8	103.5	103.3	102.3	102.0	103.8	103.0	101.8	104.3	104.4	99.3	108.2	100.8	106.6	105.0	109.2
ANT	39.4	101.8	97.7	98.9	97.2	103.0	105.6	102.9	104.0	94.7	94.1	99.7	109.9	100.9	102.1	102.4	113.5

Table 10: CUB-C mCE values based on seen accuracy.

			Noise			Blur				Weather				Digital			
Model	Clean	**mCE**	Gauss	Shot	Imp.	Defo.	Glass	Mot.	Zoom	Snow	Frost	Fog	Brigh.	Cont.	Elas.	Pixel	JPEG
OR	63.3	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
SS	63.3	104.5	106.3	105.5	92.7	103.9	103.2	104.8	104.5	105.9	106.8	107.0	107.8	103.6	105.1	105.3	104.6
TVM	63.3	100.6	99.4	99.2	99.6	101.0	99.1	100.9	100.8	101.3	100.0	102.4	103.0	101.0	102.4	100.5	99.1
LS	67.6	102.7	103.7	103.5	102.8	103.0	99.8	102.8	103.0	98.6	101.8	104.5	103.3	104.0	102.2	103.3	104.6
AM	62.5	98.8	100.0	100.0	99.5	97.9	99.4	98.1	97.8	100.8	99.1	96.0	97.3	98.3	99.6	98.7	99.9
ANT	63.6	97.4	96.0	97.1	96.0	99.3	100.1	97.4	97.9	95.8	94.7	96.0	100.4	97.8	97.4	97.3	97.6

Table 11: CUB-C mCE values based on harmonic accuracy.

			Noise			Blur				Weather				Digital			
Model	Clean	**mCE**	Gauss	Shot	Imp.	Defo.	Glass	Mot.	Zoom	Snow	Frost	Fog	Brigh.	Cont.	Elas.	Pixel	JPEG
OR	42.6	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
SS	42.6	104.9	107.6	105.8	83.4	106.1	102.0	103.8	104.9	109.4	108.0	116.8	104.7	107.6	102.5	104.4	106.0
TVM	42.6	99.8	97.0	95.9	98.5	100.2	98.0	98.0	97.5	101.6	100.0	105.5	103.8	102.9	99.9	99.9	98.4
LS	44.8	101.2	100.6	99.6	100.5	102.0	96.9	102.0	99.8	98.8	99.0	106.5	101.3	105.1	99.7	103.4	102.7
AM	41.6	100.4	101.4	99.9	101.8	101.6	98.0	99.5	100.8	100.0	99.2	102.5	98.0	99.0	98.0	104.5	102.3
ANT	42.6	96.5	92.7	93.0	92.3	101.9	99.8	100.2	96.3	97.5	98.4	96.3	94.0	95.6	96.7	95.6	97.5

Table 12: SUN-C mCE values based on Top-1 ZSL accuracy.

			Noise			Blur				Weather				Digital			
Model	Clean	**mCE**	Gauss	Shot	Imp.	Defo.	Glass	Mot.	Zoom	Snow	Frost	Fog	Brigh.	Cont.	Elas.	Pixel	JPEG
OR	79.5	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
SS	79.5	100.7	102.3	102.0	94.3	100.3	100.1	100.4	100.5	102.0	101.2	103.9	101.0	101.4	99.9	100.6	100.4
TVM	79.5	100.1	99.6	100.0	100.5	100.0	99.0	99.5	99.2	100.4	99.6	101.8	101.2	100.8	100.3	100.3	99.6
LS	81.6	101.4	100.9	101.0	101.6	103.3	101.7	101.3	100.4	99.5	100.4	102.5	102.5	101.8	101.1	101.8	101.8
AM	76.4	98.8	99.2	99.6	99.2	99.6	99.8	98.1	97.9	99.3	99.4	98.9	96.4	100.0	99.6	98.8	96.8
ANT	76.4	98.5	97.0	97.1	97.6	99.6	99.9	99.3	100.2	99.6	100.5	98.1	97.0	98.5	99.0	97.6	96.8

Table 13: SUN-C mCE values based on unseen accuracy.

			Noise			Blur				Weather				Digital			
Model	Clean	**mCE**	Gauss	Shot	Imp.	Defo.	Glass	Mot.	Zoom	Snow	Frost	Fog	Brigh.	Cont.	Elas.	Pixel	JPEG
OR	67.7	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
SS	67.7	102.5	104.6	103.4	93.0	102.8	101.1	101.2	101.9	102.9	103.4	108.0	103.3	104.2	102.1	102.3	103.1
TVM	67.7	100.7	100.3	100.1	100.4	100.0	100.0	99.3	99.6	100.7	100.8	103.6	102.4	102.0	100.9	100.3	100.7
LS	68.4	100.0	100.1	99.9	99.9	102.4	100.4	100.0	100.1	98.7	99.1	101.1	98.5	100.4	100.3	98.8	99.9
AM	64.3	98.7	99.6	99.5	1002	98.4	99.2	98.3	99.4	99.4	99.8	99.5	94.2	99.5	98.5	98.3	96.9
ANT	64.2	97.1	96.2	96.2	96.5	98.6	99.7	98.1	98.4	98.8	99.3	97.1	92.4	96.7	98.1	94.7	95.5

Table 14: SUN-C mCE values based on seen accuracy.

			Noise			Blur				Weather				Digital			
Model	Clean	**mCE**	Gauss	Shot	Imp.	Defo.	Glass	Mot.	Zoom	Snow	Frost	Fog	Brigh.	Cont.	Elas.	Pixel	JPEG
OR	74.9	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
SS	74.9	101.3	103.2	102.6	93.7	101.2	100.4	100.7	101.0	102.4	102.0	105.4	101.8	102.3	100.7	101.2	101.3
TVM	74.9	100.3	99.8	100.0	100.5	100.0	99.4	00.4	99.3	100.5	100.0	102.5	101.6	101.2	100.5	100.3	99.9
LS	76.7	101.1	100.7	100.7	101.1	103.3	101.5	101.0	100.4	99.2	100.1	102.2	101.5	101.6	101.0	100.9	101.4
AM	71.6	98.8	99.4	99.7	99.7	99.2	99.6	98.1	98.4	99.3	99.5	99.1	95.6	99.8	99.3	98.6	96.7
ANT	71.6	97.9	96.6	96.7	97.1	99.3	99.8	98.9	98.3	99.3	100.3	97.7	95.4	97.9	98.7	96.5	96.3

Table 15: SUN-C mCE values based on harmonic accuracy.

References

[1] D. Hendrycks, T. Dietterich, Benchmarking neural network robustness to common corruptions and perturbations, in: International Conference on Learning Representations, 2018.