Supplementary Materials of NLS

Xinyue Wei[1], Weichao Qiu[2], Yi Zhang[2], Zihao Xiao[2], and Alan Yuille[2]

[1]University of California San Diego
[2]Johns Hopkins University

In this supplementary material, we provide more details about data processing details, experiment setups and results, which are not included in our paper.

1. Data

1.1. Data Augmentation in MNIST [2]

We use GNT in [4] as data augmentation baseline in MNIST experiments. The Gaussian Noise standard deviation we use in GNT is randomly chosen from [0.4, 0.45, 0.5, 0.55, 0.6]. To explain the reason, according to [4], using Gaussian Noise with standard deviation around 0.5 as augmentation has the best performance on MNIST-C test sets.

1.2. UCF101 [6] Corruptions

We create a new test set with nine types of image corruptions. We use a python package *imgaug* [1] for implementation. We separately show the parameters of each corruption type for data augmentation in training and for testing in Fig. 1. These parameters have no specific meanings, we just increase corruption severity until there is obvious model performance reduction.

1.3. NTU RGB-D [5] CN Split

We show the detailed settings for creating Cross-Nuisance (CN) split. For background, we use setup id 1, 2, 3, 4, 5, 6, 8, 9, 12, 14, 15, 16, 17 for training and the rest for testing. For subject, we split the data following the same rule of Cross-Subject (CS) in NTU RGB-D. For viewpoint, we use 0 degrees for training and the rest (45 and 90 degrees) for testing instead of standard Cross-View (CV) split, making it more challenging. Thus we get a training set of 8760 videos and a testing set of 1680 videos.

2. Experiments

2.1. MNIST-C Results

We show the results in each type of corruption in MNIST-C test sets in Table 2.

References

[1] Alexander B. Jung, Kentaro Wada, Jon Crall, Satoshi Tanaka, Jake Graving, Christoph Reinders, Sarthak Yadav, Joy Banerjee, Gábor Vecsei, Adam Kraft, Zheng Rui, Jirka Borovec, Christian Vallentin, Semen Zhydenko, Kilian Pfeiffer, Ben Cook, Ismael Fernández, François-Michel De Rainville, Chi-Hung Weng, Abner Ayala-Acevedo, Raphael Meudec, Matias Laporte, et al. imgaug. `https://github.com/aleju/imgaug`, 2020. Online; accessed 01-Feb-2020.

[2] Yann LeCun and Corinna Cortes. Mnist handwritten digit database. 2010.

[3] Aleksander Madry, Aleksandar Makelov, Ludwig Schmidt, Dimitris Tsipras, and Adrian Vladu. Towards deep learning models resistant to adversarial attacks. *arXiv preprint arXiv:1706.06083*, 2017.

[4] Evgenia Rusak, Lukas Schott, Roland Zimmermann, Julian Bitterwolf, Oliver Bringmann, Matthias Bethge, and Wieland Brendel. Increasing the robustness of dnns against image corruptions by playing the game of noise. *arXiv preprint arXiv:2001.06057*, 2020.

[5] Amir Shahroudy, Jun Liu, Tian-Tsong Ng, and Gang Wang. Ntu rgb+ d: A large scale dataset for 3d human activity analysis. In *CVPR*, pages 1010–1019, 2016.

[6] Khurram Soomro, Amir Roshan Zamir, and Mubarak Shah. Ucf101: A dataset of 101 human actions classes from videos in the wild. *arXiv preprint arXiv:1212.0402*, 2012.

Corruption Type	Parameter Name	Training	Testing
Gaussian Noise	standard deviation	[0.03, 0.06, 0.09, 0.12, 0.15]	0.1
Salt & Pepper Noise	noise probability	-	0.1
Laplace Noise	standard deviation	-	0.1
Gray Scale	color invisibility	[0.2, 0.4, 0.6, 0.8, 1.0]	1.0
Multiply Hue	multiply value	-	0.5
Gamma Contrast	gamma	-	1.75
Gaussian Blur	sigma	[0.3, 0.6, 0.9, 1.2, 1.5, 1.8, 2.1]	1
Motion Blur	kernel size	-	7
Median Blur	kernel size	-	5

Table 1. Parameters of each corruption type for data augmentation in training and for testing. The parameter name is different for each corruption type. During data augmentation in training, the parameter values are randomly selected from the pool.

Model	Clean	Mean	Brightness	Canny Edges	Glass Blur	Motion Blur	Shear	Scale	Rotate	Brightness	Translate	Stripe	Fog	Splatter	Dotted Line	Zig Zag	Canny Edges
Baseline [3]	99.13	86.86	88	72	96	50	96	96	94	92	95	98	98	97	88	57	86
GNT [4]	99.40	92.39	98	76	99	93	98	99	97	93	95	98	99	99	90	55	96
ANT1×1 [4]	99.37	92.33	98	80	99	88	98	99	97	93	95	98	99	99	91	55	95
GNT+NLS	**99.44**	**92.51**	99	78	99	94	98	99	96	93	95	98	99	99	90	54	97

Table 2. Accuracy on MNIST and MNIST-C dataset of models trained with noise and our NLS.