COLA: Orchestrating Error COding and LeArning for Robust Neural Network Inference against Hardware Defects

Anlan Yu[1], Ning Lyu[1], Jieming Yin[2], Zhiyuan Yan[1], Wujie Wen[1]

[1]Department of Electrical and Computer Engineering, Lehigh University, Bethlehem, USA
[2]School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing, China

LEHIGH UNIVERSITY

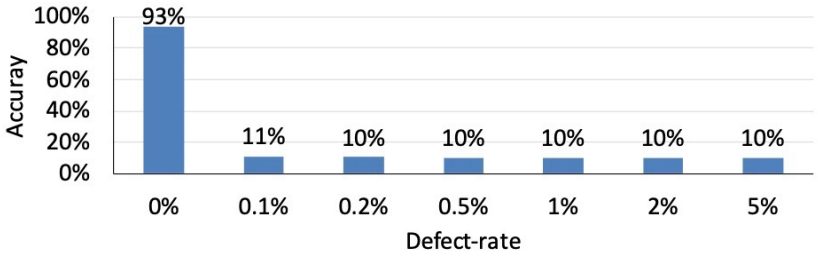
DNN Hardware Accelerator

- DNN hardware accelerators are essential
- DNN hardware accelerators suffer from hardware defects
- Hardware defects lead to DNN parameter deviation and performance degradation

Goal: improve robustness of DNN hardware accelerator

Existing solution: error correcting output codes (ECOC)

Encoding the sample labels to binary error correction codes other than one-hot code to **increase** the **inter-class distance**.



Performance of VGG-16/CIFAR10 under bit flip error

One-hot labeling

Label 0	1	0	0	0	0	0	0	0
Label 1	0	1	0	0	0	0	0	0
Label 2	0	0	1	0	0	0	0	0
…	…							
Label 6	0	0	0	0	0	0	1	0
Label 7	0	0	0	0	0	0	0	1

Minimum distance = 2

ECOC labeling

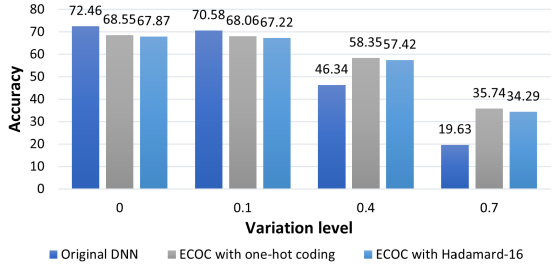
Label 0	1	1	1	1	1	1	1	1
Label 1	1	0	1	0	1	0	1	0
Label 2	1	1	0	0	1	1	0	0
…	…							
Label 6	1	1	0	0	0	0	1	1
Label 7	1	0	0	1	0	1	1	0

Minimum distance = 4

Why does existing solution perform undesirably?

Ideal ECOC solution should satisfy:

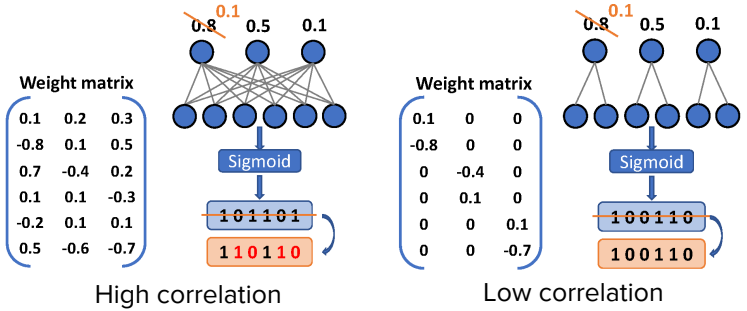
1. Keep high model accuracy with and without hardware defects
2. Improve model performance more prominently when adopting stronger ECOC codes (increased minimum Hamming distance)



1) Original DNN: using one-hot label and softmax output activation
2) ECOC with one-hot coding: using one-hot label and sigmoid output activation
3) ECOC with Hadamard-16: using Hadamard-16 label and sigmoid output activation

Root cause: ERROR CORRELATION

- Errors in DNNs are intrinsically correlated since DNN layers leverage shared information
- ECOC-enhanced DNNs require the independence of output classifiers

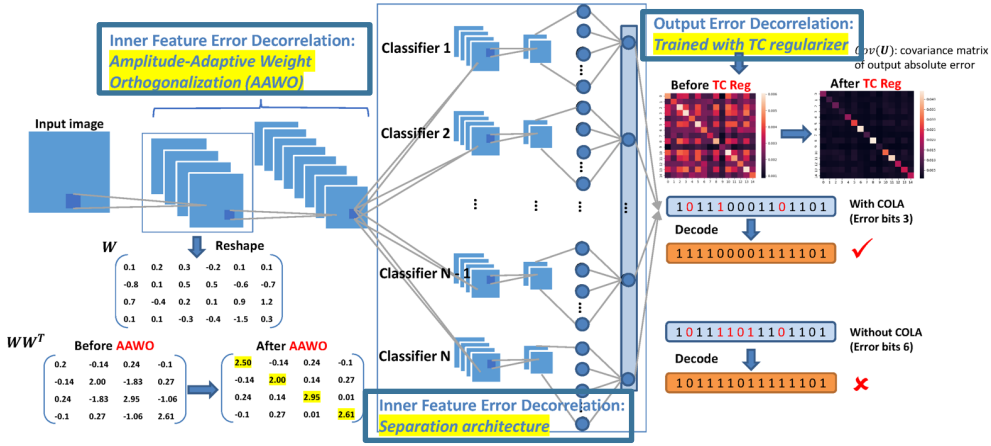


Overview of *COLA*

Goal: reduce error correlation

To achieve the goal, we propose *COLA*, a holistic framework for error decorrelation, which in turn consists of novel techniques:

1. **Amplitude-adaptive weight orthogonalization:** orthogonalizing feature errors on early layers to prevent error propagation and accumulation.
2. **Total correlation regularization:** reducing output error correlation rigorously.



Composition of *COLA*:

1) Inner feature error decorrelation: applying separation architecture; applying amplitude-adaptive weight orthogonalization
2) Output error decorrelation: training DNN with total correlation regularization

Evaluation Settings

- Datasets: MNIST, CIFAR10, CIFAR100, Tiny ImageNet
- Models: LeNet-5, AlexNet, VGG-16, ResNet-34, ResNet-50
- Codebook: Hadamard code
- Error model: 1) state error in analog DNN accelerator, 2) bit flip error in digital DNN accelerator

Please refer to the paper for more details.

Evaluation Results

- Experimental results on all the tasks show that the proposed *COLA* framework improves both **clean accuracy** and **robust accuracy** for up to **53%**

Table 1. Accuracy (% in the format of average ± standard deviation) with state error in analog accelerators. γ defines the variation level as introduced in Section 2.2. *Original*, *ECOC* and *ECOC+Sep* (Verma & Swami, 2019) are the benchmarks. *ECOC+Sep+orth*, *ECOC+TC* and *ECOC+TC+Sep+orth* are different combinations of techniques in *COLA*.

	γ	Original	ECOC	ECOC+Sep	*COLA* (Ours)		
					ECOC+Sep+orth	ECOC+TC	ECOC+TC+Sep+orth
LeNet-5 MNIST	0	98.87 ± 0.08	98.82 ± 0.04	98.86 ± 0.05	98.75 ± 0.08	**98.92** ± 0.09	98.79 ± 0.09
	0.1	98.43 ± 0.15	98.59 ± 0.07	98.66 ± 0.04	98.60 ± 0.08	98.68 ± 0.08	**98.71** ± 0.08
	0.4	70.10 ± 2.41	81.09 ± 1.92	82.85 ± 0.93	91.54 ± 0.56	85.98 ± 1.41	**93.12** ± 0.47
	0.7	23.63 ± 1.58	36.10 ± 1.94	36.94 ± 2.19	45.80 ± 1.26	40.90 ± 1.50	**49.50** ± 1.69
AlexNet CIFAR10	0	72.33 ± 0.19	68.04 ± 0.50	71.67 ± 0.71	77.08 ± 0.38	69.74 ± 0.39	**79.04** ± 0.20
	0.1	70.65 ± 0.11	67.45 ± 0.35	71.07 ± 0.63	76.89 ± 0.37	69.41 ± 0.36	**78.88** ± 0.15
	0.3	57.84 ± 0.43	62.61 ± 0.58	65.67 ± 0.33	75.04 ± 0.19	65.69 ± 0.23	**76.53** ± 0.17
	0.5	36.11 ± 0.77	51.99 ± 0.61	54.48 ± 0.83	68.49 ± 0.69	56.34 ± 0.54	**68.55** ± 0.20
VGG-16 CIFAR100	0	68.16 ± 0.52	49.06 ± 0.55	68.84 ± 0.41	68.82 ± 0.12	71.19 ± 0.15	**71.30** ± 0.29
	0.1	64.74 ± 0.95	48.07 ± 0.49	66.92 ± 0.45	67.74 ± 0.37	68.89 ± 0.39	**70.16** ± 0.35
	0.2	51.38 ± 1.56	44.00 ± 0.42	60.44 ± 0.34	66.12 ± 0.31	61.19 ± 0.58	**68.60** ± 0.37
	0.3	25.11 ± 1.18	33.06 ± 0.79	43.29 ± 0.89	62.93 ± 0.23	39.52 ± 0.55	**64.25** ± 0.13

Please refer to the paper for more results and discussions.

Evaluation Results

- Experimental results on all the tasks show that the proposed *COLA* framework improves both **clean accuracy** and **robust accuracy** for up to **53%**

Table 4. Accuracy (% in the format of average ± standard deviation) with bit-flip error in digital accelerators. α defines the bit-flip rate as introduced in Section 2.2. *Original*, *ECOC* and *ECOC+Sep* (Verma & Swami, 2019) are the benchmarks. *ECOC+Sep+orth*, *ECOC+TC* and *ECOC+TC+Sep+orth* are different combinations of techniques in *COLA*.

	α	Original	ECOC	ECOC+Sep	*COLA* (Ours)		
					ECOC+Sep+orth	ECOC+TC	ECOC+TC+Sep+orth
LeNet-5 MNIST	0	98.82 ± 0.08	98.80 ± 0.10	98.73 ± 0.08	98.61 ± 0.05	**98.86** ± 0.10	98.69 ± 0.10
	0.01	86.45 ± 2.13	89.68 ± 1.83	93.96 ± 0.58	97.29 ± 0.19	95.29 ± 1.51	**97.58** ± 0.25
	0.05	26.48 ± 1.96	33.20 ± 2.40	48.60 ± 2.24	72.31 ± 1.58	57.62 ± 2.42	**73.54** ± 1.00
	0.10	13.95 ± 0.51	14.55 ± 0.68	31.02 ± 1.38	35.05 ± 0.99	26.58 ± 0.78	**38.53** ± 1.17
AlexNet CIFAR10	0	71.69 ± 0.47	67.25 ± 0.39	71.04 ± 0.76	76.91 ± 0.52	69.51 ± 0.51	**78.35** ± 0.45
	0.01	58.56 ± 0.50	60.83 ± 0.67	62.59 ± 0.33	74.53 ± 0.15	64.54 ± 0.37	**76.02** ± 0.26
	0.05	25.64 ± 0.13	31.58 ± 1.31	33.51 ± 1.26	59.02 ± 0.92	38.74 ± 1.66	**62.12** ± 0.58
	0.10	13.72 ± 0.25	15.68 ± 1.52	16.35 ± 1.49	37.66 ± 0.73	16.78 ± 1.17	**41.50** ± 0.79
VGG-16 CIFAR100	0	66.12 ± 0.56	47.97 ± 1.08	68.55 ± 0.48	68.68 ± 0.28	70.55 ± 0.32	**71.13** ± 0.21
	0.001	59.01 ± 0.93	46.15 ± 1.19	68.13 ± 0.63	68.19 ± 0.35	70.32 ± 0.35	**70.57** ± 0.54
	0.010	09.67 ± 1.83	31.62 ± 1.70	60.29 ± 1.45	61.85 ± 1.52	57.99 ± 1.75	**63.24** ± 1.21
	0.050	01.03 ± 0.05	01.47 ± 0.16	04.96 ± 0.25	34.67 ± 1.70	02.58 ± 0.46	**41.51** ± 1.58

Please refer to the paper for more results and discussions.

Evaluation Results

- *COLA* is proposed for error decorrelation, so its applicability is not limited to the ECOC framework. It also enhances the performance of vanilla DNN

Table 5. Performance Comparison of ECOC and the original DNNs on AlexNet/CIFAR10 with *COLA*.

γ	Original	Original-*COLA*	ECOC-*COLA*
0	72.33	76.68	79.04
0.1	70.65	76.29	78.88
0.3	57.84	73.37	76.53
0.5	36.11	66.74	68.55

Please refer to the paper for more results and discussions.


LEHIGH
UNIVERSITY