EC-DARTS: Inducing Equalized and Consistent Optimization into DARTS Supplementary Material

In the following, we provide more details about the Kendall Tau [5] used in our manuscript. In addition, we show the samples of architectures that searched on CIFAR10/100, Tiny-ImageNet-200, and ImageNet by EC-DARTS.

1. Kendall Tau for Correlation Evaluations

To conduct the the correlation evaluation, we apply Kendall Tau [5] to measure correlations between different ranks. Specifically, Kendall Tau between two ranks R_1 and R_2 is denoted as $\tau(R_1, R_2)$, which belongs to $[-1, 1]$. $\tau(R_1, R_2)$ can be as high as 1 if R_1 and R_2 have a strong correlation, or as low as -1 if R_1 and R_2 are negatively correlated. Note that as $\tau(R_1, R_2)$ approaches 0, there tends to be no correlation between R_1 and R_2.

In our manuscript, we use Kendall Tau to evaluate the correlations of 3 paired ranks:

- The search accuracies and the corresponding retraining accuracies of 10 architectures are randomly selected from a single search. It should be noted that it is normal for different epochs to output the same intermediate architectures.
- The ranks of the operation weights and retraining accuracy of 10 architectures generated from one search result. Specifically, these 10 architectures are generated by replacing the 2 pairwise edges, that are connected with the last intermediate node in the searched architecture, with different possible combinations. Each possible edge connection only retains the operation corresponding to the largest operation weight.
- The search accuracies and the corresponding retraining accuracies of 10 independently searched architectures.

2. More Results in NAS-Bench-201 and Samples of Searched Results

To further fairly verify the effectiveness of EC-DARTS, we conduct experiments on NAS-Bench-201 [3]. NAS-Bench-201 is a large NAS benchmark, which includes 15,625 architectures in total. These architectures have been evaluated on CIFAR10, CIFAR100 and ImageNet-16-120 [1] by NAS-Bench-201. We compared our method with ResNet [4], RSPS [6], DARTS [7], GDAS [2]. As shown in Table 1, we run our method for 3 times on each dataset, and the architectures searched by our method achieves the state-of-the-art performance on the three datasets. Specifically, DARTS is affected by the optimization gaps, which is performed poorly on the three datasets. Different from DARTS, our method improves the correlations of different ranks from the levels of operation and structure. Besides, we provide samples of searched results on CIFAR10/100, Tiny-ImageNet-200, and ImageNet by EC-DARTS. The visualization results are shown in Figure 1.

References

[1] Patryk Chrabaszcz, Ilya Loshchilov, and Frank Hutter. A downsampled variant of imagenet as an alternative to the cifar datasets. *arXiv preprint arXiv:1707.08819*, 2017.

[2] Xuanyi Dong and Yi Yang. Searching for a robust neural architecture in four gpu hours. In *Proceedings of the IEEE Conference on computer vision and pattern recognition*, pages 1761–1770, 2019.

[3] Xuanyi Dong and Yi Yang. Nas-bench-201: Extending the scope of reproducible neural architecture search. *arXiv preprint arXiv:2001.00326*, 2020.

[4] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

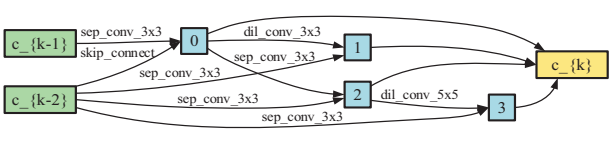
[5] Maurice G Kendall. A new measure of rank correlation. *Biometrika*, 30(1/2):81–93, 1938.

[6] Liam Li and Ameet Talwalkar. Random search and reproducibility for neural architecture search. In *Uncertainty in Artificial Intelligence*, pages 367–377. PMLR, 2020.

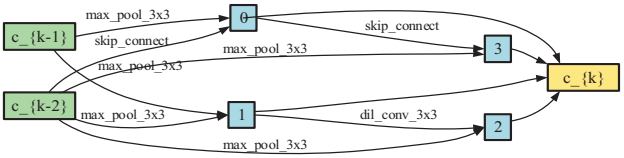
[7] Hanxiao Liu, Karen Simonyan, and Yiming Yang. Darts: Differentiable architecture search. *arXiv preprint arXiv:1806.09055*, 2018.

Table 1: Quantitative comparison results on NAS-Bench-201. “Optimal” indicates the best performing architecture in NAS-Bench-201 search space.

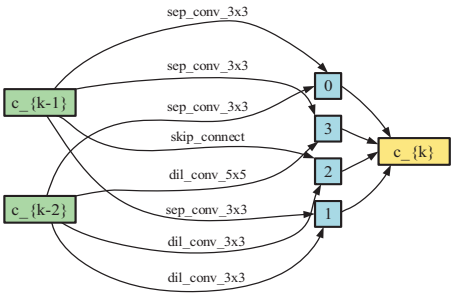
Method	CIFAR-10		CIFAR-100		ImageNet-16-120	
	Val	Test	Val	Test	Val	Test
ResNet [4]	90.83	93.97	70.42	70.86	44.53	43.63
RSPS [6]	84.16±1.69	87.66±1.69	59.00±4.60	58.33±4.34	31.56±3.28	31.14±3.88
DARTS [7]	39.77±0.00	54.30±0.00	15.03±0.00	15.61±0.00	16.43±0.00	16.32±0.00
GDAS [2]	90.0±00.21	93.51±0.13	71.14±0.27	70.61±0.26	41.70±1.26	41.84±0.90
Optimal	91.61	94.37	73.49	73.51	46.77	47.31
Ours	90.22±0.27	93.75±0.15	71.02±1.22	71.26±1.17	44.89±1.45	45.10±1.13



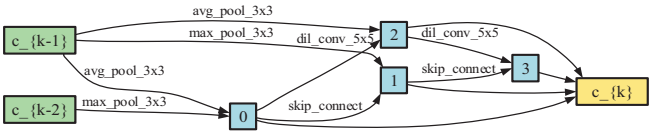
(a) the normal cell found by EC-DARTS on CIFAR10



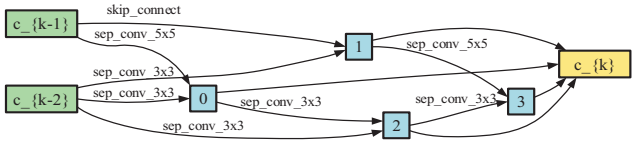
(b) the reduction cell found by EC-DARTS on CIFAR10



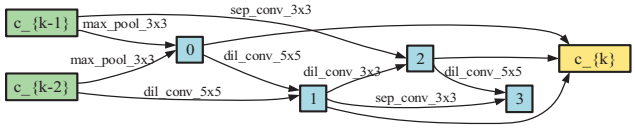
(c) the normal cell found by EC-DARTS on CIFAR100



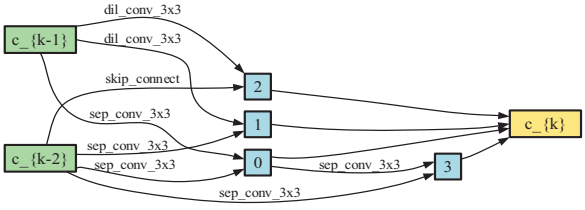
(d) the reduction cell found by EC-DARTS on CIFAR100



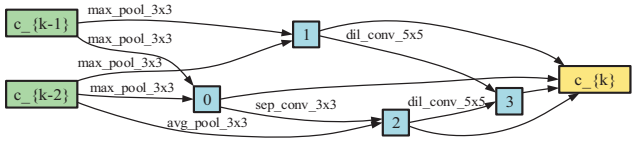
(e) the normal cell found by EC-DARTS on ImageNet



(f) the reduction cell found by EC-DARTS on ImageNet



(g) the normal cell found by EC-DARTS on tiny-ImageNet-200



(h) the reduction cell found by EC-DARTS on tiny-ImageNet-200

Figure 1: Normal cell and reduction cell searched by EC-DARTS on CIFAR10/100, Tiny-ImageNet-200, and ImageNet.