Operation and Topology Aware Differentiable Architecture Search

Shahid Siddiqui[1,2], Christos Kyrkou[1] and Theocharis Theocharides[1,2]
1. KIOS Research and Innovation Center of Excellence
2. Electrical & Computer Engineering Department, University of Cyprus

University of Cyprus
Imperial College London
KIOS Research and Innovation Center of Excellence
ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
ONLINE 10 | 15 January 2021

This work investigates DARTS search space and shows that:

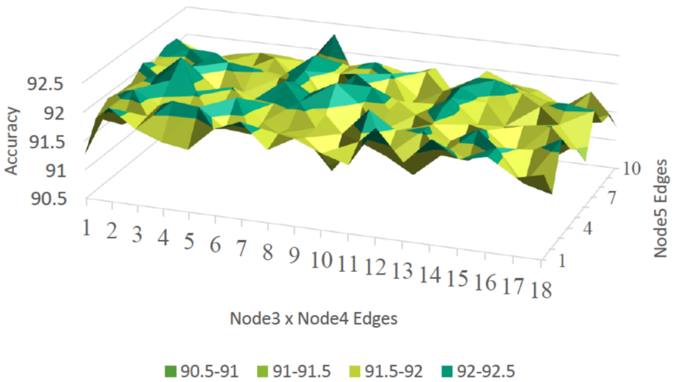
- Good cells can be discovered with *shallow search networks*.
- *Separable convolution* is the best operation in the search space.
- 3x3 kernel size is better than 5x5.
- Cell configuration also affects the accuracy.
- Skip connection bias problem can be solved faster than existing efforts.

This work is funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No 739551 (KIOS CoE) and from the Republic of Cyprus through the Directorate General for European Programmes, Coordination and Development.

Funded by:

Search Space	Operations
SS1	sep-3x3, dil-3x3
SS2	sep-5x5, dil-5x5
SS3	sep-3x3, sep-5x5
SS4	dil-3x3, dil-5x5
SS5	sep-3x3, dil-5x5
SS6	sep-5x5, dil-3x3
SS7	sep-3x3, sep-5x5, dil-3x3
SS8	sep-3x3, sep-5x5, dil-5x5

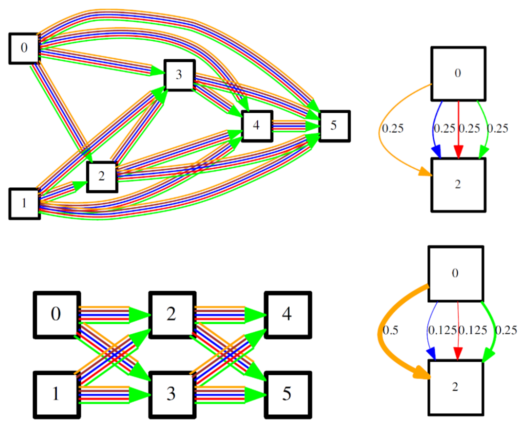
Searching in sub spaces reveals that searching deeper is only slightly better, separable convolution yields better cells than dilated and 3x3 kernel is stronger than 5x5 .



180 cell toplogies, each give different accuracy.

Method	Skip Connections: Normal	Skip Connections: Reduction	Shallow Network Accuracy	Parameters (M)	Search Cost (GPU Hours)
DARTS	14	19	90.25	0.11	36
DARTS + Data Subset	4	14	91.27	0.15	6
DARTS + Epoch Update	5	11	91.33	0.16	4.6
DARTS + Data Subset + Epoch Update	5	8	92.15	0.2	0.85

Simultaneous skip reduction and speed up mechanism.



DARTS (top) versus eDARTS search cell.

Architecture	Test Err. (%) C10	Test Err. (%) C100	Params (M)	Search Cost (GPU-days)
DenseNet-BC [24]	3.46	17.18	25.6	-
NASNet-A + cutout [7]	2.65	-	3.3	1800
AmoebaNet-A + cutout [8]	3.34	-	3.2	3150
AmoebaNet-B + cutout [8]	2.55	-	2.8	3150
Hierarchical evolution [9]	3.75	-	15.7	300
PNAS [37]	3.41	-	3.2	225
ENAS + cutout [27]	2.89	-	4.6	0.5
DARTS (first order) + cutout [10]	3.00	17.76†	3.3	1.5
DARTS (second order) + cutout [10]	2.76	17.54†	3.3	4
SNAS + mild constraint + cutout [38]	2.98	-	2.9	1.5
SNAS + moderate constraint + cutout [38]	2.85	-	2.8	1.5
SNAS + aggressive constraint + cutout [38]	3.10	-	2.3	1.5
ProxylessNAS + cutout [39]	2.08	-	5.7	4
PC-DARTS + cutout [12]	2.57	-	3.6	0.1
Fair DARTS [28]	2.54	-	2.8	0.3
P-DARTS CIFAR10 + cutout [11]	2.50	16.55	3.4	0.3
P-DARTS CIFAR100 + cutout [11]	2.62	15.92	3.6	0.3
eDARTS CIFAR10 + cutout	2.53	17.00	3.1	**0.015‡**
eDARTS CIFAR100 + cutout	2.72	16.83	3.5	0.016

eDARTS searches with fixed path and added weight to better operations. 6X faster than PC-DARTS.