NAS-Bench-201: Extending the Scope of Reproducible Neural Architecture Search

Xuanyi Dong, Yi Yang

INTRODUCTION

We propose NAS-Bench-201, a new NAS benchmark.

Motivation:

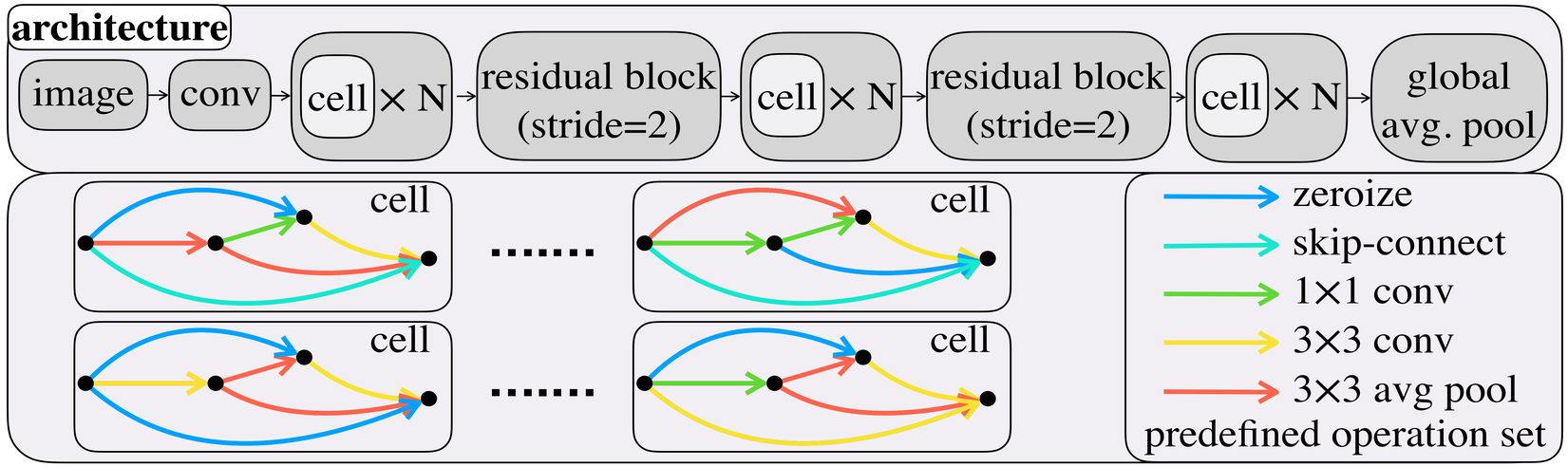
Different NAS methods have different setups, which raises a comparability problem when comparing their performance.

Highlighted features:

- Loss and accuracy on training / validation / test sets are provided ***for every epoch***.
- Results of 15625 architectures on ***three*** datasets are provided.
- Results trained with ***two*** different kinds of hyper-parameters are provided.
- The architecture space is ***agonistic*** to all NAS algorithms.
- The weights of all trained architectures are provided.
- 10 NAS algorithms are open sourced in one code base.

SCAN ME! FULL CODE PYPI API

ARCHITECTURE SPACE



Top: the macro skeleton of each architecture candidate. Bottom-left: examples of neural cell with 4 nodes. Each cell is a directed acyclic graph, where each edge is associated with an operation selected from a predefined operation set.

BENCHMARK 10 NAS METHODS

Method	Search (seconds)	CIFAR-10		CIFAR-100		ImageNet-16-120	
		validation	test	validation	test	validation	test
RSPS	8007.13	80.42±3.58	84.07±3.61	52.12±5.55	52.31±5.77	27.22±3.24	26.28±3.09
DARTS-V1	11625.77	39.77±0.00	54.30±0.00	15.03±0.00	15.61±0.00	16.43±0.00	16.32±0.00
DARTS-V2	35781.80	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	31609.80	89.89±0.08	93.61±0.09	71.34±0.04	70.70±0.30	41.59±1.33	41.71±0.98
SETN	34139.53	84.04±0.28	87.64±0.00	58.86±0.06	59.05±0.24	33.06±0.02	32.52±0.21
ENAS	14058.80	37.51±3.19	53.89±0.58	13.37±2.35	13.96±2.33	15.06±1.95	14.84±2.10
RSPS†	7587.12	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-V1†	10889.87	39.77±0.00	54.30±0.00	15.03±0.00	15.61±0.00	16.43±0.00	16.32±0.00
DARTS-V2†	29901.67	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†	28925.91	90.00±0.21	93.51±0.13	71.14±0.27	70.61±0.26	41.70±1.26	41.84±0.90
SETN†	31009.81	82.25±5.17	86.19±4.63	56.86±7.59	56.87±7.77	32.54±3.63	31.90±4.07
ENAS†	13314.51	39.77±0.00	54.30±0.00	15.03±0.00	15.61±0.00	16.43±0.00	16.32±0.00
REA	0.02	91.19±0.31	93.92±0.30	71.81±1.12	71.84±0.99	45.15±0.89	45.54±1.03
RS	0.01	90.93±0.36	93.70±0.36	70.93±1.09	71.04±1.07	44.45±1.10	44.57±1.25
REINFORCE	0.12	91.09±0.37	93.85±0.37	71.61±1.12	71.71±1.09	45.05±1.02	45.24±1.18
BOHB	3.59	90.82±0.53	93.61±0.52	70.74±1.29	70.85±1.28	44.26±1.36	44.42±1.49
ResNet	N/A	90.83	93.97	70.42	70.86	44.53	43.63
optimal		91.61	94.37	73.49	73.51	46.77	47.31

Searching results of 10 NAS methods on 3 datasets

TAKE AWAY

(1) NAS-Bench-201 helps you fairly and quickly compare your NAS method with others.

(2) The correlation of the model's performance between different datasets is not high.

(3) Using the batch mean and var of BN for searching algorithm instead of accumulated mean and var.