L^{2}NAS: Learning to Optimize Neural Architectures via Continuous-Action Reinforcement Learning

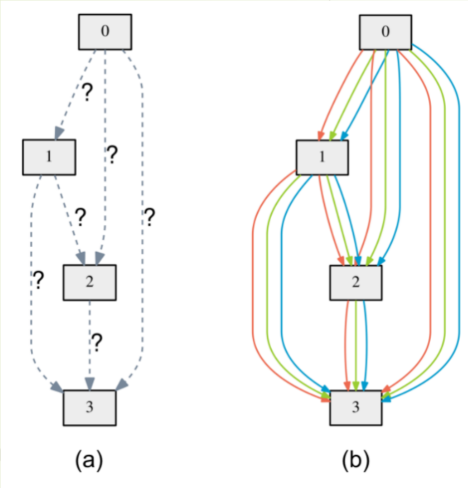
Keith G. Mills[1], Fred X. Han[2], Mohammad Salameh[2], Seyed Saeed Changiz Rezaei[2], Linglong Kong[1], Wei Lu[2], Shuo Lian[3], Shangling Jui[3] and Di Niu[1]

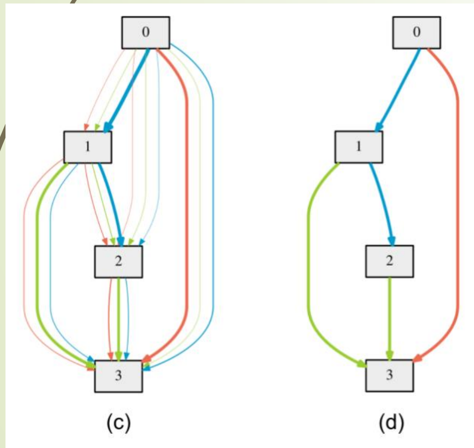
[1]University of Alberta

[2]Huawei Technologies Canada Co., Ltd.

[3]Huawei Kirin Solution, Shanghai, China

Motivation





Images from "DARTS: Differentiable Architecture Search" – ICLR 2019

- Neural Architecture Search (NAS) automates neural network design.
- Recently Differentiable Architecture Search (DARTS) has popularized continuous, gradient-based approaches.
  - Search using supernet which imposes all possible architectures.
  - Architecture described by continuous hyperparameters α.
- However, it has been shown that gradient descent has issues when applied to NAS.
  - Discretization error: Search in continuous domain on validation set, but final evaluation is performed in the discrete domain.
  - Gradient descent generally just seeks a local minima.

Contributions

- Learn an optimizer to generate the architecture hyperparameters α in a continuous domain while evaluating in the discrete domain.
- Specifically, L^2NAS is based on the actor-critic framework of Deep Deterministic Policy Gradient (DDPG), a continuous reinforcement learning (RL) algorithm.
  - Actor network learns to generate hyperparameters α.
  - Critic network learns a reward (accuracy) distribution.
- Depart from traditional RL by focus on learning from the best architectures.
  - Quantile check loss in critic learns the tail of the accuracy distribution.
- Transferability: Train an optimizer on a source dataset, then fine-tune on multiple different target dataset.
  - Outperforms training from scratch on each target dataset.
- State-of-the-art performance on NAS-Bench-201, DARTS and Once-for-All search spaces.

Our Solution

- Formulate NAS as a black-box optimization problem.
  - Let x represent a discrete architecture.
  - Let $S(.)$ be a real-valued function that evaluates an architecture, available through point queries.
  - Goal: Optimize $S(x)$.

- Continuous Relaxation
  - Let α be a continuous representation of x.
  - Let $Discretize(.)$ be a many-to-one mapping $x = Discretize(\alpha)$.
  - Result: Search on $S(x)$ is converted to search on $S(Discretize(\alpha))$, allowing gradient-based optimization of α.

Actor-Critic Framework

- A set of neural networks (MLPs) that parameterize the optimizer.
- The actor generates architectures, one per step *t*.
- The critic learns the reward distribution of found architectures.
  - Reward: $r_t = 100^{Acc(\alpha_t^d)}$; $Acc(\alpha_t{}^d)$ is the accuracy of the architecture at step *t*.
- For NAS, the critic focuses on immediate rewards, not maximizing the return like in traditional RL.
  - Immediate rewards corresponding to high-accuracy architectures.
- Quantile regression *check loss* learns tail of the accuracy distribution based on quantile $\tau \in [0, 1.0]$,

$$\rho_\tau(x) = x(\tau - \mathbf{1}(x \leq 0)), \quad \mathcal{L}_{\text{Critic}} = \frac{1}{|B_R|} \sum_{i \in B_R} \rho_\tau(r_i - Q(\alpha_i))$$

HUAWEI

Transferability

- Train an architecture optimizer on one dataset, then transfer to others.
- Modified reward function needed to normalize based on different accuracy distributions:

$$r_t = \frac{100^{Acc(\alpha_t^d)/Acc(Env)}}{100} - 1,$$

- Where *Acc(Env)* is a dataset-dependent hyperparameter.

Experimental Results
NAS-Bench-201

	CIFAR-10		CIFAR-100		ImageNet-16-120	
Method	**Valid [%]**	**Test [%]**	**Valid [%]**	**Test [%]**	**Valid [%]**	**Test [%]**
DARTS [22]	39.77 ± 0.00	54.30 ± 0.00	15.03 ± 0.00	15.61 ± 0.00	16.43 ± 0.00	16.32 ± 0.00
ENAS [26]	37.51 ± 3.19	53.89 ± 0.58	13.37 ± 2.35	13.96 ± 2.33	15.06 ± 1.95	14.84 ± 2.10
GDAS [11]	89.89 ± 0.08	93.61 ± 0.09	71.34 ± 0.04	70.70 ± 0.30	41.59 ± 1.33	41.71 ± 0.98
GAEA [19]	–	94.10 ± 0.29	–	**73.43 ± 0.13**	–	46.36 ± 0.00
RS [12]	90.93 ± 0.36	93.70 ± 0.36	70.93 ± 1.09	71.04 ± 1.07	44.45 ± 1.10	44.57 ± 1.25
REA [12]	91.19 ± 0.31	93.92 ± 0.30	71.81 ± 1.12	71.84 ± 0.99	45.15 ± 0.89	45.54 ± 1.03
REINFORCE [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 [12]	90.82 ± 0.53	93.61 ± 0.52	70.74 ± 1.29	70.85 ± 1.28	44.26 ± 1.36	44.42 ± 1.49
arch2vec-RL [39]	91.32 ± 0.42	94.12 ± 0.42	73.12 ± 0.72	73.15 ± 0.78	46.22 ± 0.30	46.16 ± 0.38
arch2vec-BO	91.41 ± 0.22	94.18 ± 0.24	**73.35 ± 0.32**	73.37 ± 0.30	46.34 ± 0.18	46.27 ± 0.37
L^{2}NAS-500	91.36 ± 0.19	94.11 ± 0.16	72.47 ± 0.74	72.69 ± 0.58	46.23 ± 0.28	46.74 ± 0.39
L^{2}NAS-1k	**91.47 ± 0.15**	**94.28 ± 0.08**	73.02 ± 0.52	73.09 ± 0.35	**46.58 ± 0.08**	**47.03 ± 0.27**
True Optimal	*91.61*	*94.37*	*73.49*	*73.51*	*46.77*	*47.31*

- NAS-Bench-201 contains 15,625 architectures evaluated on 3 datasets.
- L^{2}NAS obtains state-of-the-art results on NAS-Bench-201 CIFAR-10 and Imagenet-16-120.
- Only algorithm to achieve more than 47% average test accuracy on ImageNet-16-120.

Experimental Results
DARTS Weight-Sharing Search Space for CIFAR-10

Table 2: CIFAR-10 Results. Methods in the second and third categories search on P-DARTS and DARTS search spaces, respectively.

Architecture	Test Acc. [%]	Params [M]
ENAS [26]	97.11	4.6
GDAS [11]	97.18	2.5
AlphaX [32]	97.46 ± 0.06	2.8
P-DARTS [5]	97.50	3.4
P-SDARTS [4]	97.52 ± 0.02	3.4
DARTS 1st [22]	97.00 ± 0.14	3.3
DARTS 2nd	97.24 ± 0.09	3.3
SNAS [37]	97.15 ± 0.02	2.8
EcoNAS [44]	97.38 ± 0.02	2.9
ISTA-NAS 2S [40]	97.46 ± 0.05	3.3
arch2vec-RL [39]	97.35 ± 0.05	3.3
arch2vec-BO	97.44 ± 0.05	3.6
MdeNAS [43]	97.45	3.6
MiLeNAS [14]	97.49 ± 0.04	3.9
PC-DARTS [38]	97.43 ± 0.07	3.6
PC-SDARTS	97.51 ± 0.04	3.5
PC-GAEA [19]	97.50 ± 0.06	3.7
L^{2}NAS	**97.51 ± 0.12**	**3.8**

- DARTS on CIFAR-10 is a popular benchmark.
  - NAS-Bench-201 search space is based on DARTS
- L^{2}NAS finds architectures that achieve state-of-the-art performance.

Experimental Results
Once-For-All Search Space

- L^2NAS can generalize to different search spaces.
  - Once-for-All (OFA) based on MobileNetV3.
  - Not like DARTS.
- L^2NAS is a competent search algorithm as it can be applied to different search spaces and achieve high performance.

Table 3: Comparison of L^2NAS with other state-of-the-art architectures on ImageNet.

Architecture	Top-1 Acc. [%]	Top-5 Acc. [%]	MACs [M]
PC-SDARTS	75.7	92.6	–
P-SDARTS	75.8	92.8	–
PC-GAEA	76.0	92.7	–
MBv2	74.7	–	585
ProxylessNAS	75.1	92.5	320
MBv3-L 0.75	73.3	–	155
MBv3-L 1.0	75.2	–	219
EfficientNet-B0	77.1	93.9	390
EfficientNet-B1	79.1	94.4	700
EfficientNet-B2	80.1	94.9	1000
Cream-S	77.6	93.3	287
Cream-M	79.2	94.2	481
Cream-L	80.0	94.7	604
OFA	76.0	–	230
OFA_{Large}	79.0	94.5	595
L^2NAS	**77.4**	**93.4**	**467**
L^2NAS_{Large}	**79.3**	**94.6**	**618**

Transferability of Search Agent Across Datasets

- Train an optimizer on CIFAR-10, the source dataset.
- Transfer to CIFAR-100 or ImageNet, target datasets, and fine-tune.
- Outperform the original DARTS and direct search on target.
- Better performance, faster search.
- Can re-use optimizer many times, saving search cost.

Table 5: Transferability results for L^2NAS for CIFAR-100 and ImageNet. 'Direct' means directly searching on CIFAR-100 and ImageNet using the procedure in Section 4.2.

CIFAR-100	Test Acc. [%]	Params [M]	GPU days
DARTS 1st	82.37	3.3	1.5
DARTS 2nd	82.65	3.3	4.0
L^2NAS-Direct	82.24 ± 0.19	3.5	1.0
L^2NAS-Transfer	82.97 ± 0.29	4.0	0.1
ImageNet	**Top-1/Top-5 [%]**	**Params [M]**	**Search Cost**
DARTS 2nd	73.1/91.3	4.7	4.0
L^2NAS-Direct	74.8/92.2	4.9	1.0
L^2NAS-Transfer	75.4/92.5	5.4	0.1

Conclusion

- We propose L^2NAS, a continuous-action scheme for NAS.
- L^2NAS is based in Reinforcement Learning and using an actor-critic design based on DDPG, while incorporating features like the quantile regression check loss to meet the needs of NAS.
- Train an optimizer that learns from the best architectures.
- Achieves state-of-the-art results on NAS-Bench-201, DARTS and OFA search spaces.
- Transferability shows that a pre-trained agent from CIFAR-10 can be moved to CIFAR-100 or ImageNet and achieve higher performance at lower search cost.