



神经结构搜索进展概述

何明捷^{1,2}, 张杰¹, 山世光^{1,2}

(1. 中国科学院计算技术研究所中国科学院智能信息处理重点实验室, 北京 100190;

2. 中国科学院大学, 北京 100049)

摘要: 近年来, 深度学习技术在大量的计算视觉任务上取得了巨大的成功, 深度神经结构是一个决定性能的关键要素, 全自动的神经结构搜索方法的研究近年来受到了越来越多的关注。全自动的神经结构搜索方法是指针对特定的任务, 通过算法自动地学习出适用的深度神经结构。各类神经结构搜索方法在探索高性能、高效率的神经结构方面已经展示出了巨大的潜力。从性能评估方法、搜索空间、结构搜索策略 3 个维度对神经结构搜索方法进行了分类概述: 重点介绍了 4 种降低计算开销的性能评估方法, 2 类典型的神经结构搜索空间以及基于离散空间和基于连续空间的 2 种搜索策略。基于连续空间的 NAS 算法正逐渐成为 NAS 算法的重要发展方向。

关键词: 神经结构搜索; 深度学习; 强化学习

中图分类号: TP183

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2019097

A survey of neural architecture search

HE Mingjie^{1,2}, ZHANG Jie¹, SHAN Shiguang^{1,2}

1. Key Lab of Intelligent Information Processing, Institute of Computing Technology
Chinese Academy of Sciences, Beijing 100190, China

2. University of Chinese Academy of Sciences, Beijing 100049, China

Abstract: Recently, deep learning has achieved impressive success on various computer vision tasks. The neural architecture is usually a key factor which directly determines the performance of the deep learning algorithm. The automated neural architecture search methods have attracted more and more attentions in recent years. The neural architecture search is the automated process of seeking the optimal neural architecture for specific tasks. Currently, the neural architecture search methods have shown great potential in exploring high-performance and high-efficiency neural architectures. In this paper, a survey in this research field and categorize existing methods based on their performance estimation methods, search spaces and architecture search strategies were presented. Specifically, there were four performance estimation methods for computation cost reduction, two typical neural architecture search spaces and two types of search strategies based on discrete and continuous spaces respectively. Neural architecture search methods based on continuous space are becoming the trend of researches on neural architecture search.

Key words: neural architecture search, deep learning, reinforcement learning

收稿日期: 2019-04-10; 修回日期: 2019-05-06

基金项目: 国家自然科学基金资助项目 (No.61806188)

Foundation Item: The National Natural Science Foundation of China (No.61806188)



1 引言

近年来,深度学习技术在计算机视觉领域取得了巨大的成功。以大规模图像分类任务 ImageNet^[1]为例,2012 年的卷积神经网络 AlexNet^[2]是深度学习领域最重要的突破之一,它取得了远好于传统方法的分类精度。在深度学习技术中,深度神经结构是一个决定性能的关键要素。近年来,很多优秀的深度神经结构被设计出来,如 VGG^[3]、GoogLeNet^[4]、ResNet^[5-6]、DenseNet^[7]、ResNeXt^[8]、DPN^[9]、SENet^[10]等。虽然这些人工设计的神经结构不断地刷新了 ImageNet 数据集上的分类精度,但人工设计严重依赖专家经验,且面对全新任务时,设计对其最适合的神经结构极其困难。为此,如何通过算法自动地搜索出优秀的神经结构是深度学习领域未来重要的研究方向之一。近年来,关于神经结构搜索(neural architecture search, NAS)方法的研究受到了越来越多的关注,各类 NAS 算法搜索出的神经结构已经逐渐在性能上超越了人工设计的神经结构。

神经结构搜索方法的流程如图 1 所示。首先依据结构搜索策略在预先定义的搜索空间中选择一个神经结构,然后用性能评估方法评估该神经结构,评估的结果将返回给结构搜索策略用于调整下一次的神经结构选择。流程中的步骤将循环地执行直到搜索出了性能符合要求的神经结构。在这个流程中,性能评估方法、搜索空间、结构搜索策略是 3 个关键要素,本文将从这 3 个角度对神经结构搜索方法进行分类概述。

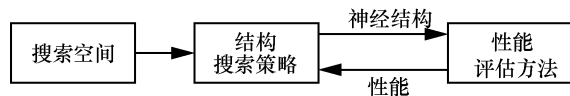


图 1 神经结构搜索方法的流程^[11]

2 性能评估方法

NAS 算法首先要设定搜索目标,由搜索目标

来引导算法按照相应的方向搜索最合适的神经结构。目前 NAS 算法主要以提高神经结构在图像分类任务上的分类精度为目标^[12-16]。除此之外,也可以将提高神经结构的计算速度作为搜索目标之一^[17]。

在神经结构的搜索过程中,NAS 算法需要不断地评估当前神经结构的性能是否达到了预定的搜索目标。最直接的性能评估方法是对当前的神经结构做一次完整训练并在测试集上测试其性能,但是过多的训练会带来巨大的计算开销并使搜索过程变得非常缓慢。因此,降低性能评估过程中的计算开销是性能评估方法的主要研究方向。目前 NAS 算法中的性能评估方法主要可以分成以下 4 种。

(1) 在小规模代理任务上估计神经结构的性能。如图 2 所示,这种做法是在一个小规模的代理(proxy)任务上进行神经结构搜索,在搜索过程中用神经结构在代理任务上的性能来估计它在实际任务上的性能,当搜索完成后再将神经结构迁移到实际任务上。

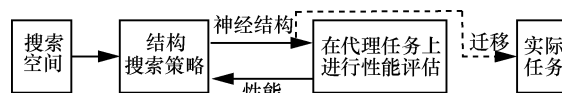


图 2 基于代理任务的 NAS 算法流程

构建代理任务的常用方式是使用与实际任务同类型但规模更小的数据集^[12-13,15-16,18-19]或分辨率更低的数据集^[20]。采用小规模代理任务的 NAS 算法降低了计算开销,但神经结构迁移成功的前提是搜索空间中的神经结构在代理任务上的性能排序和在实际任务上性能排序基本相似。当代理任务和实际任务差异较大时,代理任务有限的参考价值将降低性能评估的准确性。

(2) 利用神经结构在训练过程早期的性能估计最终的性能。这种做法是将神经结构只训练少量的迭代次数并在训练过程中记录神经结构的性能曲线,通过已有的性能曲线来插值未来的性能并提前停止训练预期性能不好的神经结构。例如,

参考文献[15]和参考文献[21]都只将搜索中的神经结构训练很少的迭代次数,然后参考文献[15]定义了一个性能估计公式来预测最终性能,参考文献[21]用一个长短期记忆网络(LSTM)来预测神经结构的最终性能。

(3)使用基于结构信息的性能预测器来预测神经结构的性能。参考文献[12]和参考文献[13]构建了一个性能预测器并使用神经结构的结构信息作为预测器的输入来预测神经结构的性能。性能预测器可以快速地排除搜索空间中性能较差的神经结构,大幅减少NAS算法的搜索次数。例如,使用了性能预测器的PNASNet^[12]的搜索速度比没有使用性能预测器的NASNet^[16]快5倍。

(4)通过权值共享来加速性能评估过程。这种方法让搜索过程中所有的神经结构之间共享权值,避免了重复地从随机初始化开始训练,大幅加快了神经结构的性能评估过程。表1中,使用了权值共享技术的ENAS^[14]只需要1块GPU运行0.45天就完成了神经结构搜索。

这4种性能评估方法既可以每种单独使用,也可以多种方法联合使用来进一步提高性能评估效率。例如参考文献[12-13,15]都将CIFAR-10^[22]分类任务作为代理任务,与此同时,参考文献[12-13,15]又分别使用了性能估计公式和基于结构信息的性能预测器来评估神经结构的性能。

3 搜索空间

神经结构搜索方法的另一个关键要素是神经结构的搜索空间(search space)。搜索空间定义了哪些神经结构将会被NAS算法搜索到。搜索空间的大小既要能够描述足够丰富的结构变化,也要尽可能避免使NAS算法的搜索时间过长。本节将讨论现有的NAS算法中经典的搜索空间。

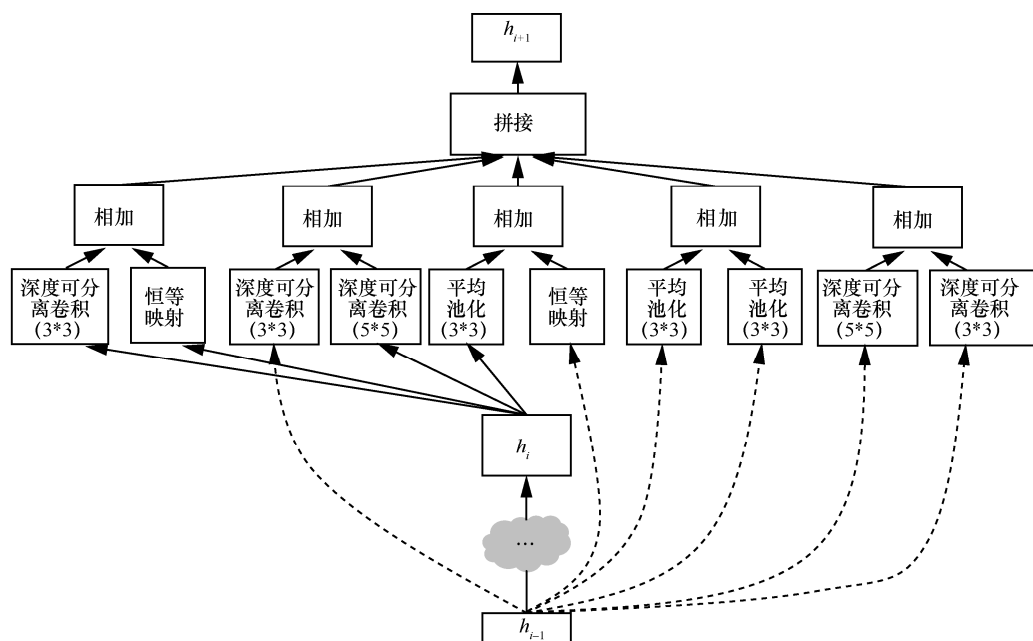
3.1 基于cell结构的搜索空间

很多人工设计的神经结构是由重复的结构单

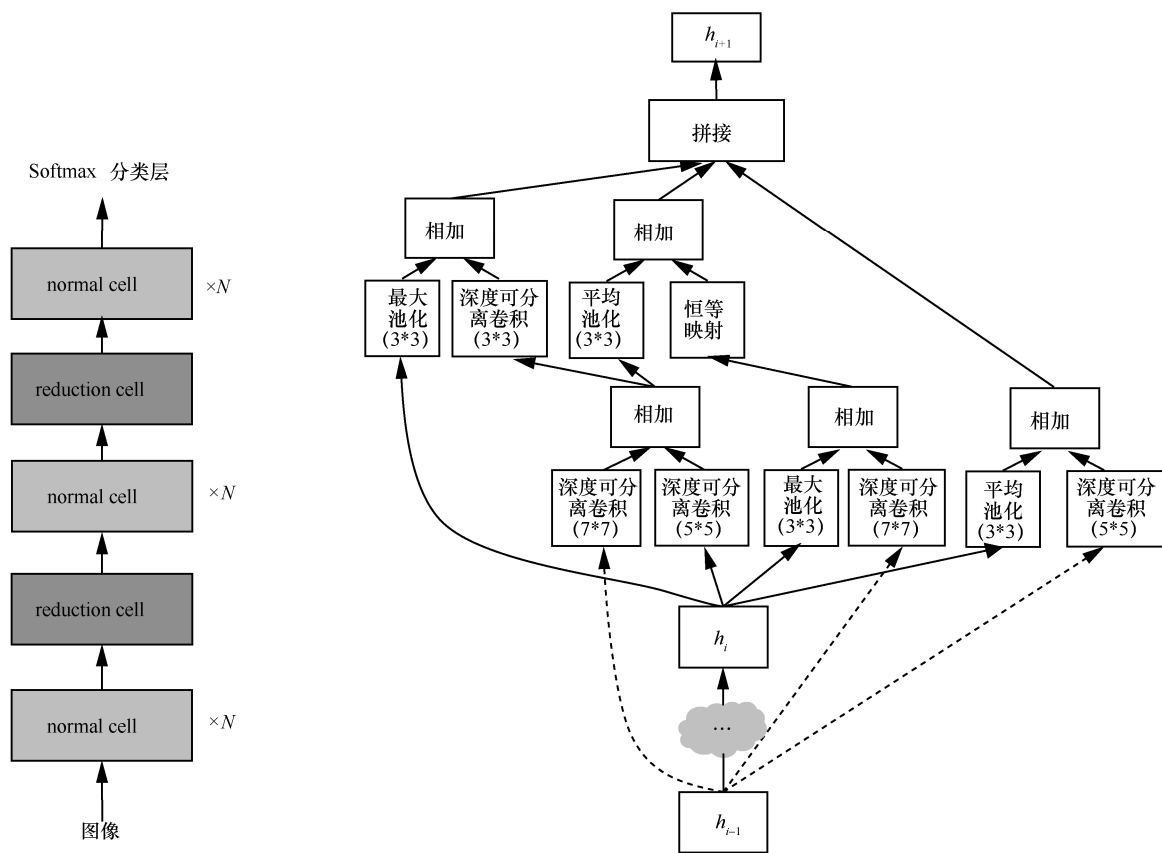
元组成的^[4-7,10],例如:ResNet^[5-6]结构由多个残差单元组成,GoogLeNet^[4]结构由Inception结构单元组成。由此启发,NAS算法^[12-13,15-16]通过搜索结构单元(cell)来替代搜索整个神经结构。NASNet^[16]提出了基于cell结构的搜索空间。在这种搜索空间中,NAS算法搜索2种类型的cell,一种为normal cell,另一种为reduction cell。normal cell的输出和输入尺寸保持一致,而reduction cell的输出尺寸是输入尺寸的一半。在搜索完cell之后,再按照预先设定的方式堆叠这2种cell来构建完整的神经结构。图3(a)展示了NASNet^[16]中的cell堆叠方式。在堆叠过程中normal cell重复堆叠 N 次, N 的大小由人工设定。NASNet^[16]搜索出了NASNet-A、NASNet-B、NASNet-C 3种神经结构。图3(b)展示了NASNet-A中的normal cell结构。图3(c)展示了NASNet-A中的reduction cell结构。

如图3所示,每个cell都有两个输入 h_i 和 h_{i-1} ,这2个输入可以来自网络中更浅的cell的输出,也可以来自图像数据。每个cell包含 B 个操作块(block),典型的 B 设置为5。每个block包含2个神经网络操作,在NASNet中,有13种操作可以被选择,例如卷积核长宽为3的卷积操作、卷积核长宽为1的卷积操作。在每个block中,这2个操作的输入可以从 h_i 、 h_{i-1} 以及前面的block的输出中选择,这2个操作的输出可以通过相加(add)或者拼接(concat)的方式来作为该block的输出。在搜索过程中,5个block按顺序逐个搜索,当5个block都搜索完成后,这5个block的输出中没有被使用过的输出会被拼接在一起作为cell的输出。虽然NASNet中有13种操作选项,但其中有5种在搜索出的cell中没有被选中过。因此,为了进一步缩减搜索空间,PNASNet^[12]将每个block中可选的操作选项由13种缩减到了8种。

基于cell结构的搜索空间有以下两点优势。



(b) NASNet-A的normal cell结构



(c) NASNet-A的reduction cell结构

(a) NASNet中cell的堆叠方式

图3 NASNet 中的神经结构^[16]

- 缩小了搜索空间，加速了神经结构的搜索过程。和搜索整个神经结构相比，只搜索 2 种 cell 结构可以减小 NAS 算法的搜索时间。
- 神经结构的深度可以灵活地扩展。当搜索出了最优的 cell 结构后，通过调整 cell 的堆叠数量，可以灵活地控制神经结构的深度和复杂度以适应不同的应用场景。例如，NASNet^[16]和 BlockQNN^[15]通过增大堆叠次数 N 成功地利用在 CIFAR-10 上搜索出的 cell 结构构建出了在 ImageNet 数据集上性能优越的神经结构。

虽然基于 cell 的搜索空间被广泛应用于 NAS 算法，但通过 cell 堆叠整体神经结构的方式依然需要人工的设计和优化。这种堆叠方式可以认为是一种元结构（meta-architecture）^[11]。在未来，元结构的设计也应作为 NAS 算法的一部分。

3.2 基于 chain 结构的搜索空间

在基于 chain 结构的搜索空间中，神经结构由一连串(layer)的层组成。其中，第 $i-1$ 层的输出作为第 i 层的输入，第 i 层的输出再作为第 $i+1$ 层的输入。和基于 cell 结构的搜索空间相比，基于 chain 结构的搜索空间简化了神经结构的局部结构，搜索空间中的神经结构没有复杂的 cell 结构，但神经结构的所有部分都由 NAS 算法在搜索中确定。典型的基于 chain 结构的搜索空间是 MNASNet^[17]中的分解分层搜索空间（factorized hierarchical search space）。如图 4 所示，神经结构中的层被分为了若干 block，一个 block 内包含若干类型相同的层。每个块内层的类型、层的超参数以及层的数量都

由算法在搜索过程中确定。

基于 chain 结构的搜索空间平衡了局部结构的搜索复杂度和整体结构的搜索复杂度，适合搜索轻量化的神经结构，例如 MNASNet 搜索出了在 ImageNet 上精度一样，但速度比 MobileNetV2^[23]快 1.5 倍的神经结构。

4 结构搜索策略

4.1 基于离散空间的搜索策略

由于神经结构的搜索空间是由一系列离散的操作选项组成的，因此很多 NAS 算法直接在离散的搜索空间中设计搜索策略。这类基于离散空间的搜索策略主要有基于强化学习的策略^[14-16,19,24-26]和基于进化算法的策略^[18,27-29]。

基于强化学习的搜索策略包含 3 个关键的要素：状态（state）、动作（action）和奖励（reward）。状态表示当前搜索过程已经确定的部分神经结构，动作表示在当前状态时向结构中加入新的层，奖励是指每个动作带来的对预定搜索目标的增益。每个动作的奖励决定了是否在当前状态时选择该动作。目前常用的强化学习算法有 policy-based 的算法^[14,16,24]和 Q -learning 算法^[15]。使用 policy-based 算法的典型策略是 NASNet^[16]。在 NASNet 的搜索过程中，一个由循环神经网络组成的控制器负责采样神经结构。被采样的神经结构会被训练和测试，然后依据测试的结果使用 proximal policy optimization 来更新控制器的参数，使控制器的下一次采样能够获得更好的神经结构。使用 Q -learning 算法的典型策略是 BlockQNN^[15]，它将状态、动作、奖励组成一个 Q 值表（ Q -table），

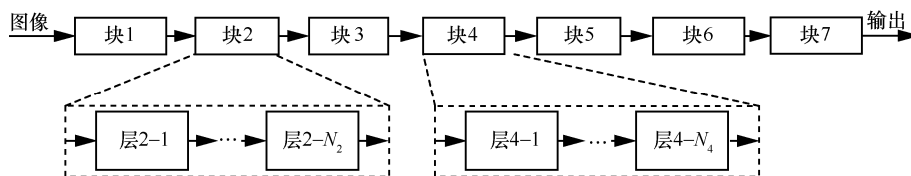


图 4 MNASNet 中的搜索空间^[17]



然后依据 Q -table 来选择在当前状态下能获得最大奖励的动作。在神经结构的搜索过程中,性能评估方法会评估神经结构的性能并依据评估结果更新 Q -table。算法通过不断地更新 Q -table 使后续构建出的神经结构的性能不断提高,直到达到搜索目标。

基于进化算法的搜索策略的典型例子是 AmoebaNet^[18]。它首先随机生成一些神经结构作为初始的种群,并从中采样少量的神经结构进行训练和测试。然后将这些神经结构中测试结果较好的作为父代结构 (parent),并对这些父代结构做突变 (mutations) 来获得新的子结构 (child)。这些新的子结构将被训练和测试并加入种群中。随后,算法再继续从更新后的种群中采样一些神经结构持续地进行这种循环的进化过程,直到达到预设的循环次数。最终,种群中性能最好的神经结构将作为搜索算法的输出。

由于离散搜索空间的大小会随着结构中操作选项的增多呈指数增长,而且搜索空间的不连续性也限制了可以使用的搜索策略的种类。因此直接在离散空间中搜索神经结构的效率很低,见表1, NASNet 和 AmoebaNet 都需要几百块 GPU 来进行神经结构搜索。

4.2 基于连续空间的搜索策略

NAONet^[13]提出了在连续空间中进行神经结

构搜索的策略。该方法首先将离散搜索空间中的神经结构映射到一个连续的空间,然后在连续空间中通过连续优化的方法来更新神经结构。在 NAONet 算法中,实现连续空间搜索的关键是使用 LSTM 实现的编码器。编码器将表示 cell 结构的离散序列 x 映射到连续空间的 e_x 。然后,NAONet 使用一个性能预测器来预测 e_x 对应的神经结构的性能。随后,NAONet 通过连续优化方法依据性能预测器来更新 e_x 以获得更好的神经结构 $e_{x'}$ 。更新后的 e_x 依然在连续空间中,此时使用一个解码器来从 e_x 中解码出离散的神经结构序列。和编码器相对应,解码器也采用 LSTM 网络来实现。编码器、解码器以及性能预测器采用多任务的方式联合训练。NAONet 搜索出的神经结构在 CIFAR-10 上的精度超越了 PNASNet、AmoebaNet 和 BlockQNN 见表1。与此同时,这种基于连续空间的搜索策略大幅提高了算法的搜索效率,NAONet 只用 200 块 GPU 运行 1 天就搜索出了和 NASNet 相近性能的神经结构。

5 开放问题讨论

当前 NAS 算法虽然取了显著的进展,但目前还主要集中在针对图像分类任务进行神经结构搜索,针对其他领域的 NAS 算法相对较少。虽然图像分类领域大量优秀的人工设计的神经结构为设

表1 神经结构搜索算法的分类

方法名称	优化目标	搜索空间	搜索策略	参数数量 /MB	CIFAR-10 Top1 Error	GPU 数量	GPU 运行天数
NASNet ^[16]	精度	基于 cell 结构	离散, 强化学习	27.6	2.97%	500	4
PNASNet ^[12]	精度	基于 cell 结构	离散, 强化学习	3.2	3.63% ^[14]	100 ^[14]	1.5 ^[14]
MNASNet ^{[17]*}	精度, 速度	基于 chain 结构	离散, 强化学习	4.4	-	-	-
ENAS ^[14]	精度	基于 cell 结构	离散, 强化学习	4.6	3.54%	1	0.45
BlockQNN ^[15]	精度	基于 cell 结构	离散, 强化学习	39.8	3.54%	32	3
AmoebaNet ^[18]	精度	基于 cell 结构	离散, 进化算法	3.2	3.34%	450	7
NAONet ^[13]	精度	基于 cell 结构	连续优化	28.6	2.98%	200	1

注: “*” 表示该工作发布于 arXiv, 尚未正式发表。

计 NAS 算法的结构搜索空间提供了参考,但也使目前的搜索空间存在局限性,例如通过同构的 cell 堆叠整体神经结构的方式依然包含大量的人工设定,支持神经结构不同深度时 cell 异构的搜索空间依然很少。此外,目前搜索空间的部分超参数(如卷积核的数量)是固定的,没有通过搜索的方法来得到。受限的搜索空间使得各类搜索策略难以充分发挥搜索能力,使目前由 NAS 算法搜索出的结构难以在更多的任务上取得本质的性能进步。因此,如何设计更灵活同时也能适应更多领域的搜索空间是未来 NAS 算法需要研究的关键问题之一。在性能评估方法方面,目前大部分 NAS 算法依然在小规模的代理任务上进行神经结构的搜索,虽然降低了计算开销,但是当代理任务和实际任务差异较大时,这种搜索方式将难以保证神经结构迁移到实际任务后的性能。如何不采用代理任务,直接针对大规模的实际任务进行神经结构搜索依然是一个需要研究的问题。在神经结构搜索策略方面,目前巨大的计算开销依然是当前 NAS 算法的主要问题,降低计算复杂度,提高搜索速度是未来 NAS 搜索策略的探索方向。

6 结束语

本文从性能评估方法、搜索空间、结构搜索策略 3 个维度对神经结构搜索方法进行了分类概述。在性能评估方法方面,本文重点概述了 4 种可以降低计算开销的性能评估方法;在搜索空间方面,本文概述了 2 类典型的神经结构搜索空间;在神经结构搜索策略方面,本文概述了基于离散空间的搜索策略和基于连续空间的搜索策略。由于在连续空间中可以使用连续优化方法高效率地更新神经结构,因此基于连续空间的 NAS 算法正逐渐成为 NAS 算法的重要发展方向。此外,在搜索过程中的权值共享技术由于可以大幅度降低 NAS 算法对计算资源的需求,它在 NAS 算法中的应用将越来越广泛。

参考文献:

- [1] DENG J, DONG W, SOCHER R, et al. ImageNet: a large-scale hierarchical image database[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2009, Miami, Florida, USA. Piscataway: IEEE Press, 2009.
- [2] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C]//Advances in Neural Information Processing Systems (NIPS), Dec 3-8, 2012, Lake Tahoe, Nevada, USA. New York: Curran Associates Inc, 2012.
- [3] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. Computer Science, 2004.
- [4] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. Piscataway: IEEE Press, 2015.
- [5] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. Piscataway: IEEE Press, 2016.
- [6] HE K, ZHANG X, REN S, et al. Identity mappings in deep residual networks[C]//European Conference on Computer Vision (ECCV), Oct 11-14, 2016, Amsterdam, Netherlands. Cham: Springer, 2016.
- [7] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. Piscataway: IEEE Press, 2017.
- [8] XIE S, GIRSHICK R, DOLL R P, et al. Aggregated residual transformations for deep neural networks[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. Piscataway: IEEE Press, 2017.
- [9] CHEN Y, LI J, XIAO H, et al. Dual path networks[C]//Advances in Neural Information Processing Systems (NIPS), Dec 4-9, 2017, Long Beach, CA, USA. New York: Curran Associates Inc, 2017.
- [10] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 18-22, 2018, Salt Lake, UT, USA. Piscataway: IEEE Press, 2018.
- [11] ELSKEN T, METZEN J H, HUTTER F. Neural architecture search: A survey[J]. arXiv: 1808.05377, 2018.
- [12] LIU C, ZOPH B, NEUMANN M, et al. Progressive neural architecture search[C]//European Conference on Computer Vision (ECCV), Sep 8-14, 2018, Munich, Germany. Cham: Springer, 2018.
- [13] LUO R, TIAN F, QIN T, et al. Neural architecture optimization[C]//Advances in Neural Information Processing Systems



- (NIPS), Dec 3-8, 2018, Montreal, Canada. New York: Curran Associates, Inc, 2018.
- [14] PHAM H, GUAN M Y, ZOPH B, et al. Efficient neural architecture search via parameter sharing[C]//International Conference on Machine Learning (ICML), July 10-15, 2018, Stockholmsmässan, Stockholm, Sweden. [S.l.: s.n], 2018.
- [15] ZHONG Z, YAN J, WU W, et al. Practical block-wise neural network architecture generation[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 18-22, 2018, Salt Lake City, UT, USA. Piscataway: IEEE Press, 2018.
- [16] ZOPH B, VASUDEVAN V, SHLENS J, et al. Learning transferable architectures for scalable image recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 18-22, 2018, Salt Lake City, UT, USA. Piscataway: IEEE Press, 2018.
- [17] TAN M, CHEN B, PANG R, et al. Mnasnet: platform-aware neural architecture search for mobile[J]. arXiv: 1807.11626, 2018.
- [18] REAL E, AGGARWAL A, HUANG Y, et al. Regularized evolution for image classifier architecture search[C]//Thirty-Third AAAI Conference on Artificial Intelligence, Jan 27-Feb 1, 2019, Honolulu, Hawaii, USA. [S.l.: s.n], 2019.
- [19] CAI H, YANG J, ZHANG W, et al. Path-level network transformation for efficient architecture search[C]//International Conference on Machine Learning (ICML), July 10-15, 2018, Stockholmsmässan, Stockholm, Sweden. [S.l.: s.n], 2018.
- [20] CHRABASZCZ P, LOSCHILOV I, HUTTER F. A downsampled variant of imagenet as an alternative to the cifar datasets[J]. arXiv:1707.08819, 2017.
- [21] RAWAL A, MIIKKULAINEN R. From nodes to networks: evolving recurrent neural networks[J]. arXiv: 1803.04439, 2018.
- [22] KRIZHEVSKY A, HINTON G. Learning multiple layers of features from tiny images[R]. 2009.
- [23] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: inverted residuals and linear bottlenecks[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 18-22, 2018, Salt Lake City, UT, USA. Piscataway: IEEE Press, 2018.
- [24] ZOPH B, LE Q V. Neural architecture search with reinforcement learning[C]//International Conference on Learning Representations (ICLR), April 24-26, 2017, Toulon, France. [S.l.: s.n], 2017.
- [25] BAKER B, GUPTA O, NAIK N, et al. Designing neural network architectures using reinforcement learning[C]//International Conference on Learning Representations (ICLR), April 24-26, 2017, Toulon, France. [S.l.: s.n], 2017.
- [26] CAI H, CHEN T, ZHANG W, et al. Reinforcement learning for architecture search by network transformation [J]. arXiv: 1707.04873, 2017.
- [27] LIU H, SIMONYAN K, VINYALS O, et al. Hierarchical representations for efficient architecture search[C]//International Conference on Learning Representations (ICLR), April 30-May 3, 2018, Vancouver, BC, Canada. [S.l.: s.n], 2018.
- [28] REAL E, MOORE S, SELLE A, et al. Large-scale evolution of image classifiers[C]//International Conference on Machine Learning (ICML), Aug 6-11, 2017, Sydney, NSW, Australia. [S.l.: s.n], 2017.
- [29] XIE L, YUILLE A. Genetic CNN[C]//IEEE International Conference on Computer Vision (ICCV), Oct 22-29, 2017, Venice, Italy. Piscataway: IEEE Press, 2017.

[作者简介]



何明捷（1990—），男，中国科学院计算技术研究所工程师，中国科学院大学博士生，主要研究方向为计算机视觉与机器学习。



张杰（1988—），男，中国科学院计算技术研究所助理研究员，主要研究方向为计算机视觉与机器学习。



山世光（1975—），男，中国科学院计算技术研究所研究员、博士生导师，中国科学院智能信息处理重点实验室常务副主任。国家自然科学基金委员会“优秀青年科学基金项目”入选者，第三批国家“万人计划”入选者，科技部“创新人才推进计划”中青年科技创新领军人才，CCF 青年科学家奖获得者。主要研究方向为计算机视觉和机器学习。