

# 基于神经网络结构搜索的轻量化网络构建

姚 潇<sup>1</sup> 史叶伟<sup>1</sup> 霍冠英<sup>1</sup> 徐 宁<sup>1</sup>

**摘 要** 轻量化网络可解决深度神经网络参数较多、计算量较高、难以部署在计算能力有限的边缘设备上等问题。针对轻量化网络中常用的分组卷积的分组结构问题,文中提出基于神经网络结构搜索的轻量化网络。将不同分组的卷积单元作为搜索空间,使用神经网络结构搜索,得到网络的分组结构和整体架构。同时为了兼顾准确率与计算量,提出循环退火搜索策略,用于解决神经网络结构搜索的多目标优化问题。在数据集上的实验表明,文中网络识别准确率较高,时间复杂度和空间复杂度较低。

**关键词** 轻量化网络,模型压缩,分组卷积,神经网络结构搜索,多目标优化

**引用格式** 姚 潇,史叶伟,霍冠英,徐 宁.基于神经网络结构搜索的轻量化网络构建.模式识别与人工智能,2021,34(11):1038-1048.

**DOI** 10.16451/j.cnki.issn1003-6059.202111007

中图法分类号 TP 18

## Lightweight Model Construction Based on Neural Architecture Search

YAO Xiao<sup>1</sup>, SHI Yewei<sup>1</sup>, HUO Guanying<sup>1</sup>, XU Ning<sup>1</sup>

**ABSTRACT** The traditional deep neural network cannot be deployed on the edge devices with limited computing capacity due to numerous parameters and high computation. In this paper, a lightweight network based on neural architecture search is specially designed to solve this problem. Convolution units of different groups are regarded as search space, and neural architecture search is utilized to obtain both the group structure and the overall architecture of the network. In the meanwhile, a cycle annealing search strategy is put forward to solve the multi-objective optimization problem of neural architecture search with the consideration of the accuracy and the computation cost of the model. Experiments on datasets show that the proposed network model achieves a better performance than the state-of-the-art methods.

**Key Words** Lightweight Network, Model Compression, Group Convolution, Neural Architecture Search, Multi-objective Optimization

**Citation** YAO X, SHI Y W, HUO G Y, XU N. Lightweight Model Construction Based on Neural Architecture Search. Pattern Recognition and Artificial Intelligence, 2021, 34(11): 1038 - 1048.

收稿日期: 2021-01-07; 录用日期: 2021-06-03

Manuscript received January 7, 2021;

accepted June 3, 2021

国家自然科学基金项目(No.61501170、41876097)、中央高校基本科研基金项目(No.B20020205)、江苏省重点研究开发项目(No.BK20192004、BE2018004-04)、东南大学生物电子学国家重点实验室开放研究基金项目(No.2019005)资助  
Supported by National Natural Science Foundation of China(No. 61501170、41876097), Fundamental Research Funds for the Central Universities(No.B20020205), Key Research and Development Program of Jiangsu Province(No.BK20192004、BE2018004-04), Open Research Fund of State Key Laboratory of Bio-

对于卷积神经网络(Convolutional Neural Network, CNN)如残差网络(Residual Network, ResNet)和视觉几何组(Visual Geometry Group, VGG)<sup>[1-2]</sup>精度的提高大多都是建立在网络深度和宽度的增加上,这显然会消耗大量的计算资源。对于机器人、自

electronics of Southeast University(No.2019005)

本文责任编辑 郝志峰

Recommended by Associate Editor HAO Zhifeng

1.河海大学 物联网工程学院 常州 213022

1.College of Internet of Things Engineering, Hohai University, Changzhou 213022

自动驾驶和增强现实这类需要在有限资源的计算平台上对本地数据在线处理的应用来说,需要更低开销的轻量级 CNN。

网络轻量化主要分为参数量化、模型剪枝、紧凑网络等方法。参数量化减少存储每个权重所需比特数,以此减少计算量和参数量,这一想法可以进一步扩展为以量化形式表示梯度和激活值。Young 等<sup>[3]</sup>训练后使用率失真框架对权值进行最优变换和量化,提高在任何给定量化比特数下的压缩性能。受结构化剪枝启发,Yu 等<sup>[4]</sup>提出以卷积核为量化单位的压缩方法。Gong 等<sup>[5]</sup>利用矢量量化损失作为网络权值,达到最小的量化损失和更优的模型精度。

模型剪枝舍弃对最终输出结果贡献较小的神经元,剪掉冗余参数,降低网络复杂度,解决过拟合问题。早期的剪枝方法多是非结构化的剪枝方案,精度相对较高,但依赖于特定算法库或硬件平台的支持。近期剪枝方法大都不需要特定算法库或硬件平台的支持,可直接在目前成熟的深度学习框架上运行结构化剪枝。He 等<sup>[6]</sup>提出基于几何中值的滤波器剪枝(Filter Pruning via Geometric Median, FPGM),基于几何中值修剪冗余卷积核而不是重要性较小的卷积核。He 等<sup>[7]</sup>提出软滤波器剪枝(Soft Filter Pruning, SFP),上一次迭代中修剪的卷积核依旧参与下一次迭代更新。Liu 等<sup>[8]</sup>提出多层滤波器剪枝(Multi-level Filter Pruning, MFP)。

紧凑网络是专在资源受限的边缘设备上设计的轻量级网络,与大型网络进行压缩的方法不同,紧凑网络直接训练一个可用的低参数量和计算复杂度较低的网络。这类方法大多使用深度可分离卷积等方法,将卷积操作分解为多个步骤,降低参数数量和计算复杂度。Sandler 等<sup>[9]</sup>提出使用深度可分离卷积和逆残差结构的 MobileNetV2。Zhang 等<sup>[10]</sup>提出 ShuffleNet,使用逐点群卷积,降低计算量,并通过通道混洗(Channel Shuffle)的方式解决分组间信息流通不畅的问题。

除上述方法以外,还有将复杂模型中的知识转移到简单模型中的知识蒸馏方法<sup>[11]</sup>、使用移位操作代替空间卷积的移位残差网络(ShiftResNet)<sup>[12]</sup>、将权重矩阵分解为多个小矩阵的低秩分解方法等<sup>[13]</sup>。

Shi 等<sup>[14]</sup>提出使用空洞卷积提取多尺度特征的空洞轻量网络(Dilated LightNet)。然而,这种基于经验和直觉的人工设计并不能保证最终的神经网络最优,而且通过穷举的方式列出所有的网络结构再从中找出最优架构并不现实。

除了整体架构方面的网络结构以外,在网络的分组结构上,即网络每层的分组数保持相同的常量,这一点同样是基于经验和直觉的人工设计,同样存在性能次优的问题。对此,Zhang 等<sup>[15]</sup>提出使用二值矩阵表示卷积层输入输出通道之间的连接,通过一些约束条件让二值矩阵能表示分组的结构,再通过端到端的方式训练该矩阵,得到最终的分组结构。虽然此方法能解决网络分组结构的问题,但却不适用于网络的整体架构。

为了保证网络的轻量化,本文提出基于神经网络结构搜索的轻量化网络。考虑到资源有限,选择高效神经网络结构搜索(Efficient Neural Architecture Search, ENAS)<sup>[16]</sup>作为基础框架,在此基础上提出循环退火的神经网络结构搜索算法(Cycle Annealing Neural Architecture Search, CANAS),构建最佳网络结构,优化搜索空间和搜索策略。数据集上的实验表明本文网络的有效性和泛化性。

## 1 基于神经网络结构搜索的轻量化网络

### 1.1 多尺度空洞的轻量化网络

Dilated LightNet<sup>[14]</sup>使用不同扩张率空洞卷积多尺度地提取每个分组的信息,基本单元如图 1 所示。

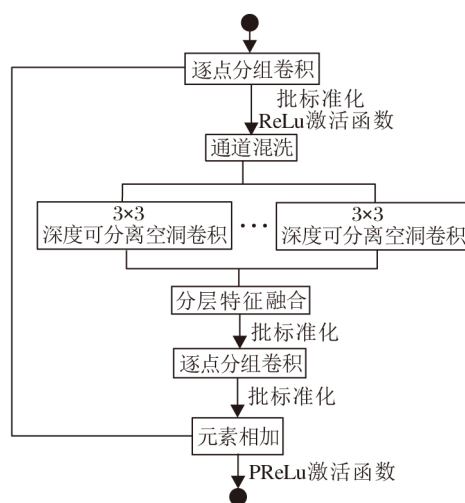


图 1 Dilated LightNet 基本单元<sup>[14]</sup>

Fig.1 Dilated LightNet unit<sup>[14]</sup>

但是,多层串联的 Dilated LightNet 单元会放大空洞卷积的网格效应。为了解决这一问题,在网络架

构上混合使用 ShuffleNet 单元与 Dilated LightNet 单元,保证在提取多尺度特征的同时减少网格效应的负面影响.具体网络架构如图 2 所示.

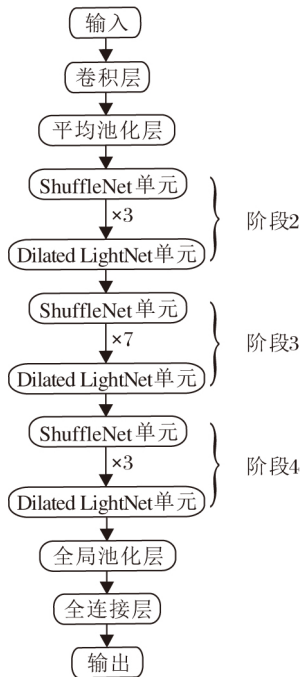


图 2 Dilated LightNet 网络架构<sup>[14]</sup>

Fig.2 Dilated LightNet architecture<sup>[14]</sup>

## 1.2 整体框架与搜索空间

图 2 中的 Dilated LightNet 网络架构由人工设计得到,为了避免次优问题,本文通过神经网络结构搜索的方式搜索整体网络架构,分组结构作为整体网络架构一部分可同时进行搜索.神经网络结构搜索主要分为控制器和生成器<sup>[16]</sup>.

控制器是一个长短期记忆网络 (Long Short-Term Memory, LSTM),以一定的概率对表示不同网络结构的子模型进行采样,生成一串表示网络结构的序列.例如, Dilated LightNet 网络需要搜索的部分共 16 层,每层可选的单元包括不同分组的 ShuffleNet 单元和 Dilated LightNet 单元共 8 种, LSTM 由 16 个节点组成,每个节点输出 0~7 之间的整数,16 个节点输出组成的序列表示采样的网络结构. LSTM 通过 Softmax 分类器以自回归的方式对决策进行采样,上一步的输出作为下一步的输入.

生成器部分是 CNN,本文使用 Dilated LightNet

网络.根据控制器采样的网络结构,即生成序列,构建同样结构的网络,在训练集上训练以收敛,并在验证集上测试得到准确率.控制器使用策略梯度方法,最大化生成器在验证集上获得准确率并更新控制器参数.随着迭代次数的增加,控制器生成更优的网络结构.

CANAS 最终迭代的所有结构都可视为一个更大结构的子结构,可使用一个有向无环图<sup>[16]</sup>表示 CANAS 的搜索空间.图 3 为 CANAS 的有向无环图示例,通过采样有向无环图的子图获得特定的网络结构,实线连接的子图就是一种网络结构.简单来说, CANAS 的有向无环图是所有可能的子模型在搜索空间中的叠加,其中,节点表示具体的单元,边表示信息流向. CANAS 保存每个节点单元的参数,每次只更新被激活节点的参数,保证每次采样完成后生成器不必重新训练,只在之前训练的参数基础上更新,因此所有子模型的参数共享.

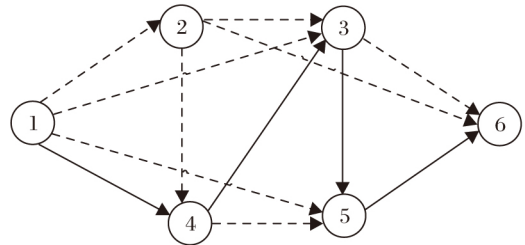


图 3 CANAS 的有向无环图

Fig.3 Direct acyclic graph of CANAS

根据搜索的目标不同,有向无环图中节点的单元也有所变化.从全局尺度上看,如果要搜索整体的网络结构,则每个节点就应该是网络中完整的一层.从局部的尺度上看,如果想要搜索用于搭建整体网络的具体单元,那么每个节点应该是不同的计算操作,如不同参数的卷积池化等操作.如果想要同时解决每层分组结构与卷积结构的问题,每个节点应该是不同分组的 ShuffleNet 单元和 Dilated LightNet 单元.具体的搜索空间如图 4 所示,图中  $g$  为分组数.根据图 2 中的网络架构,搭建成图 4 中的搜索空间,前两层和后两层保持不变,中间 16 层中每层具体的单元需要通过搜索得到,实线连接的结构表示控制器采样生成的网络架构.

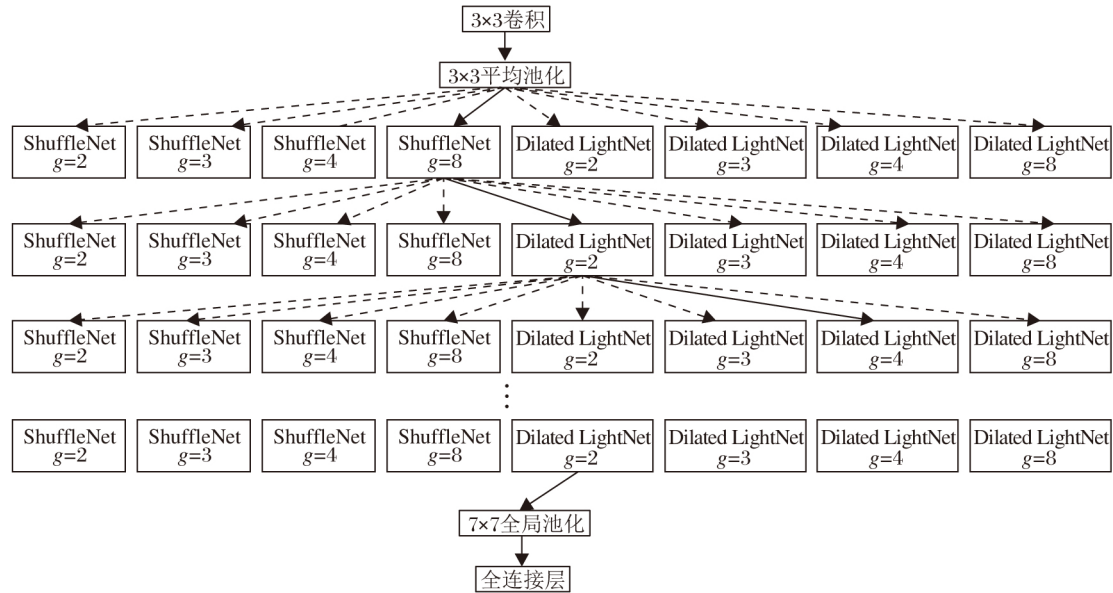


图 4 结合不同分组的搜索空间

Fig.4 Search space consisting of different groups

### 1.3 神经网络结构搜索的训练方法

在 CANAS 框架中,有 2 个可学习的参数:控制器 LSTM 的参数  $\theta$ 、子模型的共享参数  $\omega$ 。整体的训练过程由 2 个阶段交替进行,第一阶段通过数据集训练子模型的共享参数  $\omega$ ,梯度  $\nabla_{\omega}$  使用标准的反向传播计算。第二阶段是对  $\theta$  进行固定步数的训练,使用强化学习更新参数。

1) 训练子模型的共享参数  $\omega$ 。首先固定控制器的策略  $\pi(m; \theta)$ ,并对  $\omega$  使用随机梯度下降法(Stochastic Gradient Descent, SGD)最小化期望损失函数  $E_{m \sim \pi}[L(m; \omega)]$ 。为使用采样自  $\pi(m; \theta)$  的子模型  $m$  在训练集上以 minibatch 方式计算的标准交叉熵损失,可通过蒙特卡洛估计得到梯度,即

$$\nabla_{\omega} E_{m \sim \pi(m; \theta)}[L(m; \omega)] \approx \frac{1}{M} \sum_{i=1}^M \nabla_{\omega} L(m_i; \omega),$$

上式是梯度  $\nabla_{\omega} E_{m \sim \pi(m; \theta)}[L(m; \omega)]$  的一个无偏估计,这个估计在  $m$  固定时方差大于标准 SGD 梯度。然而在实验中发现,  $M=1$  时也可取得较优效果,即可使用从  $\pi(m; \theta)$  采样的任何单一模型  $m$  的梯度以更新  $\omega$ 。

2) 训练控制器参数  $\theta$ 。本文固定  $\omega$  更新  $\theta$ ,目标是最大化期望奖励  $E_{m \sim \pi(m; \theta)}[R(m; \omega)]$ 。奖励  $R(m; \omega)$  是在验证集上计算得到,目的是鼓励 CANAS 选择泛化性更优的模型而不是过拟合训练集的模型。在图像分类实验中,奖励函数是模型在验证集上的准确率。本文采用自适应矩估计(Adaptive Moment Estimation,

Adam) 优化器,由强化学习算法计算得到梯度,使用移动平均基线减少方差。

在强化学习计算梯度的过程中,本文将控制器产生的序列看作是一系列的动作,记做  $a_{1:G}$ ,这些序列表示生成的网络结构。将生成器在验证集上的准确率记作  $R$ ,为了找到最佳网络结构, CANAS 要求控制器最大化其期望奖励,重新表示为

$$J(\theta) = E_{P(a_{1:G}; \theta)}[R].$$

由于奖励  $R$  不可微,需使用策略梯度方法迭代更新  $\theta$ 。本文使用 REINFORCE 方法<sup>[17]</sup> 梯度

$$\nabla_{\theta} J(\theta) = \sum_{g=1}^G E_{P(a_{1:G}; \theta)}[\nabla_{\theta} \ln P(a_g | a_{(g-1):1}; \theta) R],$$

根据经验上式约等于

$$\frac{1}{N} \sum_{n=1}^N \sum_{g=1}^G \nabla_{\theta} \ln P(a_g | a_{(g-1):1}; \theta) R_n,$$

其中  $N$  为控制器在一个批次中采样的不同结构的数量,  $G$  为控制器设计网络结构时输出的序列长度。上式是对梯度的无偏估计,但有相对较大的方差,为了减少这一估计的方差,采用基线函数:

$$\frac{1}{N} \sum_{n=1}^N \sum_{g=1}^G \nabla_{\theta} \ln P(a_g | a_{(g-1):1}; \theta) (R_n - b). \quad (1)$$

只要基线函数  $b$  不依赖于当前动作,这仍是一个无偏梯度估计。此处,基线  $b$  是之前网络结构准确率的指数移动平均值。

### 1.4 循环退火搜索策略

神经网络结构搜索算法只是将最大化准确率作

为搜索目标,然而在实际应用中通常还要考虑其它条件.例如,本文的目标轻量化可使用浮点计算量进行量化表示.基于上述原因,讨论神经网络结构搜索(Neural Architecture Search, NAS)的多目标优化问题,为了在不同目标的权衡中选择理想的结果,需要计算该问题的帕累托前沿,即一组帕累托最优的网络结构.

对于 NAS 而言,可能存在的结构数量非常多,结构评估的成本很高,因此较难准确高效地获得完整的帕累托前沿.关于多目标 NAS 的大多工作都依靠标量化的方式以近似帕累托前沿,即通过加权和或乘积的方式将多个目标合为一个:

$$R(m) = Acc(m) \cdot \left( \frac{flops(m)}{F} \right)^\alpha, \quad (2)$$

$$\alpha = \begin{cases} -1, & flops(m) > F \\ 0, & flops(m) \leq F \end{cases}$$

此奖励函数是一种复杂的标量化方式,其中  $Acc$  表示模型  $m$  在验证集上的准确率,  $flops$  表示模型  $m$  的浮点计算量,  $F$  表示目标浮点计算量,只有在模型浮点计算量高于目标计算量时,函数右边才会变成一个小于 1 的惩罚项,用于控制奖励函数的整体值.

然而,为了获得多个帕累托最优解,需要在不同目标计算量下多次运行 NAS,这会消耗大量的计算资源.为了准确高效地搜索整个帕累托前沿,本文提出循环退火搜索策略(Cycle Annealing Search, CAS).不同于传统多目标 NAS,本文的奖励函数是非平稳的,可随搜索进程的推进不断调整策略,有效产生整个帕累托前沿,奖励函数

$$R(m) = Acc(m) \cdot \exp \left( - \frac{|flops(m) - F|}{\lambda \cdot \delta} \right). \quad (3)$$

式(3)重点在于奖励函数的右半部分, CANAS 将整个搜索过程分为加热和退火两个阶段,退火阶段又分为慢退火和快退火两个阶段,在不同阶段用于控制惩罚项参数  $\lambda \cdot \delta$  发生变化.以外,在退火阶段设  $F \in [F_{\min}, F_{\max}]$ .

由式(3)可见,想要获得更高的奖励取决于模型的准确率和每秒浮点计算量(Floating Point Operations per Second, FLOPs).如果  $\theta$  表示目标计算量为  $f$  时的最优控制器参数,可认为  $f$  附近的目标计算量  $f'$  的最优控制器参数  $\theta'$  也应接近  $\theta$ .因此,假设已知目标计算量  $f$  时的最佳控制器策略为  $\pi$ ,从已知的最优策略  $\pi$  开始,只需经过少量计算就可得到  $f$  附近目标计算量为  $f'$  的最佳控制器策略  $\pi'$ .

综上所述,本文的循环退火搜索算法步骤如下.

算法 CAS

输入 数据集  $D$ , 目标计算量  $[F_{\min}, F_{\max}]$ ,  
惩罚项参数  $\delta_{\text{warm}}, \delta_{\text{anneal}}$ , 循环周期  $T$ ,  
循环次数  $N_{\text{warm}}, N_{\text{slow\_anneal}}, N_{\text{fast\_anneal}}$

输出 子模型结构及性能的集合

$$\{m, flops(m), Acc_{\text{valid}}(m), Acc_{\text{test}}(m)\}$$

1. 随机初始化控制器参数  $\theta$

2. 随机初始化子模型的共享参数  $\omega$

// 加热阶段

3. 设  $F = F_{\min}$ ,  $\lambda = 1$ ,  $\delta = \delta_{\text{warm}}$

4. for  $step = 1$  to  $N_{\text{warm}}$  do:

5. 控制器采样子模型  $m$

6. 使用训练集  $D_{\text{train}}$  更新  $\omega$

7. 使用验证集  $D_{\text{valid}}$  的准确率, 根据式(3) 计算奖励  $r$

8. 使用奖励  $r$  更新参数  $\theta$

9. 在测试集  $D_{\text{test}}$  评估模型  $m$

10. end for

// 慢退火阶段

11. 设  $\lambda = 1$ ,  $\delta = \delta_{\text{anneal}}$

12. for  $step = 1$  to  $N_{\text{slow\_anneal}}$  do

13. 设

$$F = F_{\min} + \frac{F_{\max} - F_{\min}}{2} \cdot$$

$$\left( 1 - \cos \left( 2\pi \cdot \frac{step \cdot \text{mod}(T+1)}{T} \right) \right)$$

//mod 为取模函数

14. 重复 5 ~ 9

15. 更新  $m, flops(m), Acc_{\text{valid}}(m), Acc_{\text{test}}(m)$

16. end for

// 快退火阶段

17. 设  $0 < \lambda < 1$ ,  $\delta = \delta_{\text{anneal}}$

18. for  $step = 1$  to  $N_{\text{fast\_anneal}}$  do

19. 设  $\lambda = \lambda^{\lceil \text{Ceil}(step/T) \rceil}$

//Ceil 为向上取整函数

20. 重复 13 ~ 15

21. end for

CAS 可细分为三个阶段.第一阶段为加热阶段,在此阶段学习得到最小目标计算量  $F_{\min}$  时的最优策略,为后续的退火过程做准备,此时的  $\delta_{\text{warm}}$  为一个较大的值,对奖励函数惩罚也相对较小.第二阶段为慢退火阶段,此时引入参数循环周期  $T$ ,目标计算量  $F$  会在循环周期  $T$  内从  $F_{\min}$  遍历至  $F_{\max}$ .同时  $\delta_{\text{anneal}}$  为一个更小的值,意味着惩罚力度更大.第三阶段是快



退火阶段 相比慢退火阶段 此时只有参数  $\lambda$  发生变化 设为一个  $0 \sim 1$  之间的小数 随着循环次数的增加  $\lambda$  会越来越小  $\lambda \cdot \delta$  整体会变得 smaller 这意味着随着循环次数的增加 同时也是迭代次数的增加 模型浮点计算量偏离目标计算量付出的代价也越来越大。

综上所述 CAS 将在慢退火阶段和快退火阶段以不同的惩罚力度学习所有目标计算量时的最优策略 在退火阶段结束时 就可得到整个帕累托前沿。

## 2 实验及结果分析

### 2.1 实验设置

本文实验使用 MNIST<sup>[18]</sup>、CIFAR-10、CIFAR-100<sup>[19]</sup> 数据集 三者均为分类任务。MNIST 数据集由 0~9 的手写数字组成 图像大小为  $28 \times 28$  训练集包含 60 000 幅图像 测试集包含 10 000 幅图像。CIFAR-10 数据集包含 10 种类别 CIFAR-100 数据集包含 100 种类别 图像大小均为  $32 \times 32$  训练集包含 50 000 幅图像 测试集包含 10 000 幅图像。

实验使用 Pytorch 深度学习框架 全部实验均采用相同的软硬件和参数配置 CPU 为 Intel Broadwell E5-2620V4 GPU 为 GeForce RTX 2080 Ti 显存大小为 12 GB 内存大小为 128 GB 操作系统为 Ubuntu 16.04 LTS 开发语言为 Python3.6。

实验分为 2 类: 1) 神经网络结构搜索实验 用于搜索网络结构; 2) 常规的模型训练测试实验 用于验证模型的有效性。2 类实验在训练时均采用相同的数据增强方法 包括随机裁剪和几何翻转等。在神经网络结构搜索的训练中 将 CIFAR 训练集再次划分 其中 45 000 幅作为训练集 剩下 5 000 幅作为验证集 MNIST 训练集分为 50 000 幅和 10 000 幅。此外 为了与 ShuffleNet 网络进行对比分析 网络的输入图像分辨率大小为  $224 \times 224$  当与输入为  $224 \times 224$  的方法对比时 数据集的图像都使用双线性插值法将原图像变为  $224 \times 224$ 。当使用 CIFAR 数据集原图尺寸输入时 Dilated LightNet 网络去除开始的卷积层和池化层 此外 由于 MNIST 数据集较简单 Dilated LightNet 网络去除池化层 每个阶段只保留两层单元。

实验中参数涉及到 CANAS 的控制器和生成器参数及生成网络的实验参数 大部分参数设置参考 ENAS<sup>[16]</sup>。CANAS 的控制器是一个 16 层的 LSTM 即式 (1) 中  $G = 16$  采样数量  $N = 10$  使用学习率为 0.001 的 Adam 优化器进行训练。为了防止过早收敛 本文

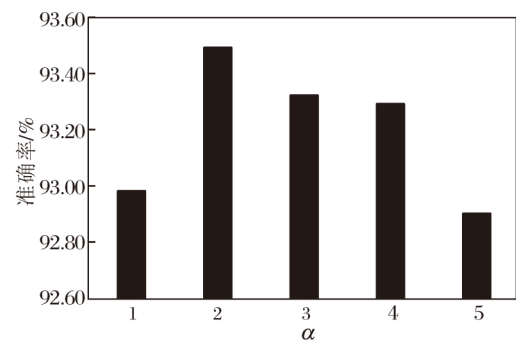
对控制器 logits 添加一个值为 1.5 的 tanh 常数 将控制器的样本熵添加到奖励中 权重为 0.1。生成器是根据控制器生成的 CNN 初始学习率设为 0.05。其它设置与生成网络的设置相同 迭代次数为 300。同时为了保证网络中每层的通道数能被不同的分组数整除 本文设置阶段 2、阶段 3、阶段 4 的输出通道数分别为 288、576、1 152。对于循环退火搜索算法中的参数 由于该通道数下的生成网络最低计算量为 83 M 所以设  $F_{\min} = 8 \times 10^7$  为了确保搜索到的网络计算量低于 ShuffleNet 网络的最低计算量 136 M 设  $F_{\max} = 1.2 \times 10^8$ 。由于搜索过程中的重点阶段是快退火阶段 在该阶段不断更新帕累托前沿 所以将总迭代次数中的较大部分分配给快退火阶段 前两个阶段只要训练至接近收敛即可 循环周期  $T = 50$  因此

$$N_{\text{warm}} = 50, N_{\text{slow\_anneal}} = 50, N_{\text{fast\_anneal}} = 200.$$

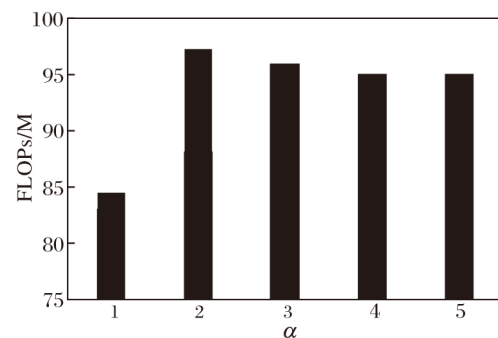
另外,

$$\delta_{\text{warm}} = \frac{1}{2}(F_{\max} - F_{\min}), \delta_{\text{anneal}} = \frac{1}{2\alpha}(F_{\max} - F_{\min}),$$

其中  $\alpha$  表示控制惩罚项的重要参数。 $\alpha$  不同时 CANAS 的准确率和计算量对比如图 5 所示。由图可知  $\alpha = 2$  时效果最优。



(a) 准确率  
(a) Accuracy



(b) 每秒浮点计算量  
(b) FLOPs

图 5  $\alpha$  不同时 CANAS 的准确率和计算量

Fig.5 Accuracy and FLOPs of CANAS with different  $\alpha$

对于生成网络的测试 使用带动量的随机梯度

下降法更新参数,损失函数使用交叉熵损失,动量设置为 0.9,权重衰减设置为 0.000 25,批大小设置为 128,迭代次数设置为 300,初始学习率设置为 0.02,使用余弦退火学习率调整方法<sup>[20]</sup>,循环周期设置为 10,最小学习率设置为 0.000 5。

本文在 3 个数据集上所有实验的参数设置都保持一致。

## 2.2 基于神经网络结构搜索的轻量化网络验证实验

本文将不同分组的 ShuffleNet 单元和 Dilated LightNet 单元作为搜索空间,使用 CANAS 搜索新网络结构的 Dilated LightNet (Dilated LightNet+CANAS)。对搜索得到的 Dilated LightNet 网络的性能进行实验评价,验证有效性。分别在 3 个数据集上进行神经网络结构搜索的实验,控制器的输出如图 6 所示,每个节点输出一层网络使用的单元。

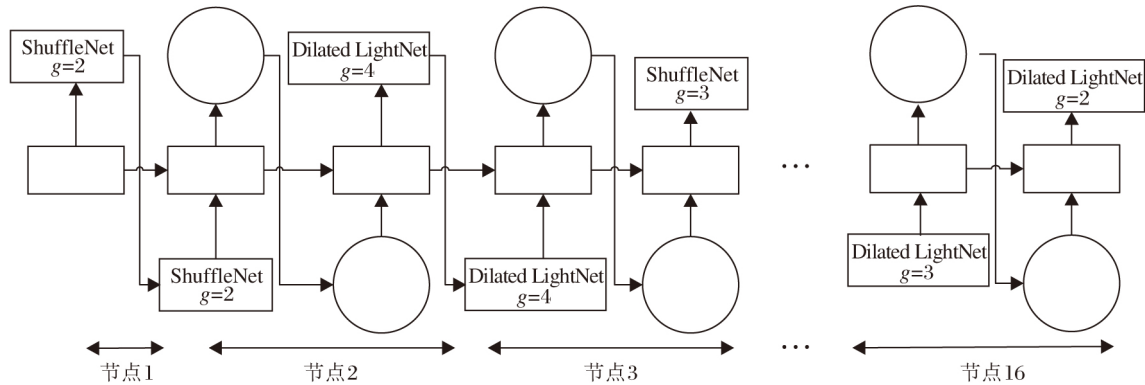
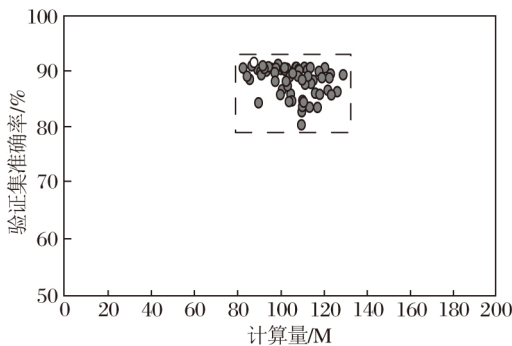


图 6 控制器的输出示意图

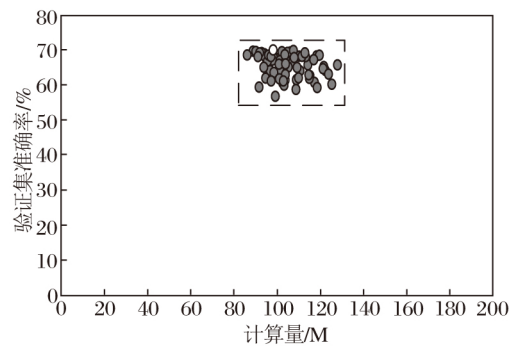
Fig.6 Illustration of controller output

在 CIFAR-10、CIFAR-100 数据集上搜索的网络结构分布如图 7 所示。由图可见,实验中有些网络结构在验证集上的准确率明显低于其它网络结构,这是由于在搜索过程中,神经网络结构搜索算法只能尽量靠近目标计算量,并不能准确得到符合目标计算量的绝对最佳网络结构,因此,由于网络结构上的细微差异,每次得到的网络计算量带有一定的随机

性。可能在之后的循环中,一些计算量下的网络结构不会被再次搜索到,也就不可能更新参数获得更高的准确率,所以在不丢弃这些网络结构的情况下,其在图中表现为准确率较低的点。搜索的网络结构收敛在最小目标计算量与最大目标计算量之间,选取空心点标注的准确率最高的网络结构用于之后的实验。



(a) CIFAR-10



(b) CIFAR-100

图 7 在 2 个数据集上搜索的网络结构

Fig.7 Network architecture searched on 2 datasets

搜索得到的网络结构在 3 个数据集上的指标值结果如表 1 所示,在表中,CIFAR-10 数据集上分组

数为 4 ,CIFAR-100 数据集上 ShuffleNet 网络的分组数为 8 ,Dilated LightNet 网络的分组数为 4.

表 1 各网络在 2 个数据集上的指标值结果  
Table 1 Index values of different networks on 2 datasets

网络	CIFAR-10			CIFAR-100		
	准确率/%	计算量/M	参数量/M	准确率/%	计算量/M	参数量/M
ShuffleNet	92.26	133	0.90	69.75	137	1.07
Dilated LightNet	93.63	134	0.90	72.15	134	1.00
Dilated LightNet+CANAS	93.49	88	0.55	72.36	99	0.71

ShuffleNet 网络和 Dilated LightNet 网络均选择 4 种分组中准确率最高的结果,在 CIFAR-10 数据集上 相比 ShuffleNet 网络,Dilated LightNet+CANAS 网络准确率增加 1.23%,计算量减少 33.8%,参数量减少 38.8%.相比 Dilated LightNet 网络,Dilated LightNet+CANAS 网络的准确率基本保持一致,但计算量只有原来的 65.6%,参数量只有原来的 61.1%.Dilated LightNet+CANAS 网络在 CIFAR-100 数据集上表现更佳.相比 ShuffleNet 网络,计算量减少 27.7%,参数量减少 33.6%,准确率增加 2.61%.相比 Dilated LightNet 网络,Dilated LightNet+CANAS 网络的准确率略有提升,计算量只有原来的 73.8%,参数量只有原来的 71%.由此说明本文结合循环退火搜索策略的神经网络结构搜索算法的有效性,也验证关于分组结构会影响模型表达的假设.

2.3 循环退火搜索策略的验证实验

本文的循环退火搜索策略是为了解决神经网络结构搜索中的多目标优化问题,当神经网络结构搜索只将准确率作为搜索目标时,得到的网络往往为了准确率倾向于搜索到计算量更大的网络结构,无法实现网络的轻量化.本节提出如下对比方法:直接

使用 ENAS、将离散的搜索空间松弛为可导的连续空间的微架构搜索 (Differentiable Architecture Search ,DARTS)<sup>[21]</sup>、在 ENAS 基础上使用标量化多目标优化的标量化多目标搜索 (Scalar Multi-objective Search ,SMS)、使用循环退火搜索策略的 CANAS.对比分析各方法性能,验证 CANAS 中循环退火搜索策略的有效性.

SMS 使用的奖励函数如式 (2) 所示,目标浮点计算量  $F=1.4\times10^8$ ,训练方法和参数设置与 2.2 节中本文方法相同.其它方法的参数设置与原文档完全一致,实验中本文方法的参数设置也与 2.2 节保持一致,实验结果如表 2 所示.

从 CIFAR-10 数据集上的实验结果可看出,Dilated LightNet+CANAS 网络的计算量只有 Dilated LightNet+ENAS 网络的 44.8%,参数量也只有 Dilated LightNet+ENAS 网络的 41%,准确率比 Dilated LightNet+ENAS 网络下降 0.38%.相比 Dilated LightNet+DARTS 网络,Dilated LightNet+CANAS 网络结果也相似.进一步与 Dilated LightNet+SMS 网络对比,Dilated LightNet+CANAS 网络在准确率、计算量和参数量上都有更优表现.

表 2 不同搜索算法得到的 Dilated LightNet 网络在 2 个数据集上的实验结果  
Table 2 Experimental results of Dilated LightNet obtained by different search algorithms on 2 datasets

网络	CIFAR-10			CIFAR-100		
	准确率/%	计算量/M	参数量/M	准确率/%	计算量/M	参数量/M
Dilated LightNet+ENAS	93.87	196	1.34	72.71	197	1.66
Dilated LightNet+DARTS	93.76	191	1.04	72.69	191	1.15
Dilated LightNet+SMS	93.41	102	0.61	72.46	133	1.03
Dilated LightNet+CANAS	93.49	88	0.55	72.36	99	0.71

在 CIFAR-100 数据集上,Dilated LightNet +CANAS 网络在准确率基本保持不变的基础上:相比 Dilated LightNet+ENAS 网络,计算量减少 49.7%,参

数量减少 57.2%;相比 Dilated LightNet+DARTS 网络,计算量减少 48.1%,参数量减少 38.2%;相比 Dilated LightNet+SMS 网络,计算量减少 25.5%,参



量减少 31%.

综合 2 个数据集的结果可见,只使用准确率作为搜索目标时,得到的是计算代价较高的网络结构,无法兼顾轻量化,即使使用标量化的多目标优化方法,本文的循环退火搜索策略在神经网络结构搜索的多目标优化问题上也表现更优.

#### 2.4 对比实验

在网络轻量化的问题上,有诸如紧凑网络、网络剪枝、低秩分解、参数量化等多种方法,为了能全面客观地评价本文方法,选择当前用于网络轻量化的一些先进方法进行对比分析:对 ResNet-56 使用基于几何中值剪枝的 FPGM<sup>[6]</sup>、将 ResNet-56 中空间卷积替换为移位操作且膨胀系数为 6 时的 ShiftResNet<sup>[12]</sup>、对 ResNet-32 进行软剪枝的 SFP<sup>[7]</sup>、对 VGG-16 剪枝的 MFP<sup>[8]</sup>、使用深度可分离卷积和逆残差结构的紧凑网络 MobileNetV2<sup>[9]</sup>、根据最终响应的重要性分数指导 LeNet 和 ResNet-56 剪枝的神经元重要性分数传播剪枝( Neuron Importance Score Propagation, NISP)<sup>[22]</sup>、使用生成对抗网络对 LeNet 和 ResNet-56 剪枝的生成对抗剪枝( Generative Adversarial Learning, GAL)<sup>[23]</sup>、对 LeNet 和 ResNet-20 进行结构稀疏化剪枝的结构稀疏化剪枝( Structured Sparsity Learning, SSL)<sup>[24]</sup>.NISP、GAL 和 SSL 在 MNIST 数据集上对 LeNet 剪枝,在 CIFAR 数据集上对 ResNet 剪枝.对比方法的参数设置与原文献完全一致,本文方法的实验设置与 2.1 节相同,实验结果如表 3~表 5 所示,表中黑体数字表示最优值.

表 3 各方法在 CIFAR-10 数据集上的指标值

Table 3 Index values of different methods on CIFAR-10

输入	方法	CIFAR-10		
		准确率/%	计算量/M	参数量/M
32×32	FPGM	92.93	59	—
	ShiftResNet	92.69	102	0.58
	SFP	92.08	40	—
	MFP	93.14	154	7.20
	SSL	92.60	36	0.26
	NISP	93.01	81	0.49
	GAL	91.58	50	0.29
	Dilated LightNet+CANAS	<b>93.21</b>	45	0.95
	MobileNetV2	93.25	126	0.70
224×224	Dilated LightNet+CANAS	<b>93.49</b>	<b>88</b>	<b>0.55</b>

表 4 各方法在 CIFAR-100 数据集上的指标值

Table 4 Index values of different methods on CIFAR-100

输入	方法	CIFAR-100		
		准确率/%	计算量/M	参数量/M
32×32	FPGM	69.66	59	—
	ShiftResNet	72.13	102	0.58
	SFP	68.37	40	—
	SSL	67.47	31	0.25
	GAL	70.13	73	0.73
	Dilated LightNet+CANAS	<b>72.72</b>	52	1.19
224×224	MobileNetV2	71.93	126	0.81
	Dilated LightNet+CANAS	<b>72.36</b>	<b>99</b>	<b>0.71</b>

表 5 各方法在 MNIST 数据集上的指标值

Table 5 Index values of different methods on MNIST

方法	MNIST		
	准确率/%	计算量/M	参数量/M
SSL	99.00	0.20	0.10
NISP	99.18	0.65	0.11
GAL	99.14	0.43	0.05
SOTA	99.84	24	1.5
Dilated LightNet+CANAS	<b>99.21</b>	0.30	<b>0.04</b>

为了在相近的计算量下对比,实验中使用缩放因子为 0.5 的 MobileNetV2 进行对比,输入为 32×32 时的 Dilated LightNet 网络宽度扩大为原来的 4/3 倍,在 MNIST 数据集上宽度缩小为原来的 1/6 倍.在 CIFAR-10 数据集上,本文方法准确率最佳时,计算量少于其中大部分方法.输入为 32×32 时,相比准确率最高的 MFP,本文方法准确率只高出 0.07%,计算量却减少 70.7%.相比计算量最低的 SSL,本文方法在计算量上略有增加,但准确率提高 0.61%.输入为 224×224 时,相比 MobileNetV2,本文方法在准确率、计算量、参数量上都表现更优.

CIFAR-100、CIFAR-10 数据集上的图像均来自 Tiny Images 数据集,但这 2 个数据集的类别互斥,因此 CIFAR-100 数据集可作为 CIFAR-10 数据集的反例.CIFAR-100 数据集上有 100 类图像,每类包含 600 幅图像.CIFAR-10 数据集上有 10 类图像,每类包含 6 000 幅图像.因此 CIFAR-100 数据集上的训练图像少于 CIFAR-10 数据集,任务难度更大,分类效果下降更明显.由 CIFAR-100 数据集上的实验结果看出,本文方法仍得到最高的识别准确率,而计算

量仍低于平均计算量。SSL 虽然时间复杂度相对较低,但识别准确率却下降 5.25%。此外,本文方法在输入为  $32 \times 32$  时表现优于  $224 \times 224$ ,计算量更少。这是因为  $224 \times 224$  的图像是通过插值得到,信息量没有增加,计算量却翻了数倍。

MNIST 数据集上的结果也与 CIFAR-10、CIFAR-100 数据集基本保持一致,本文方法在综合考虑性能与复杂度的情况下均最优。

从 3 个数据集的综合对比实验结果可看出,本文方法在平衡识别性能和轻量化方面表现更优,这表明方法的有效性。本文同时与 MNIST 的最高水平/最佳结果(State of the Art, SOTA)<sup>[25]</sup>进行对比分析,实验表明,本文方法在识别准确率上已经接近胶囊网络方法<sup>[25]</sup>,但在时间复杂度和空间复杂度上更优,在平衡性能与时间空间复杂度方面,本文方法更有优越性。

### 3 结 束 语

本文针对传统深度神经网络参数较多、计算量较高的问题,通过神经网络架构搜索构建轻量化网络,提出使用不同分组的 ShuffleNet 单元和 Dilated LightNet 单元作为搜索空间,基于神经网络结构搜索得到网络的分组结构和整体架构。同时,提出循环退火搜索策略,解决神经网络结构搜索的多目标优化问题。在 CIFAR-10、CIFAR-100、MNIST 数据集上的对比实验表明,本文方法无论是在准确率还是在计算量上都占有优势,较好地平衡识别性能和复杂度,证实轻量化网络构建方法的有效性。本文通过神经网络结构搜索优化分组结构和整体架构,但通道结构依旧是人工设置,今后可使用剪枝技术对轻量化网络进行通道冗余消除,实现紧凑模型的进一步压缩。

### 参 考 文 献

- [1] HE K M, ZHANG X Y, REN S Q, *et al.* Deep Residual Learning for Image Recognition // Proc of the IEEE Conference on Computer Vision and Pattern Recognition. Washington, USA: IEEE, 2016: 770-778.
- [2] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large-Scale Image Recognition [C/OL]. [2020-12-10]. <https://arxiv.org/pdf/1409.1556.pdf>.
- [3] YOUNG S L, ZHE W, TANBMAN D, *et al.* Transform Quantization for CNN Compression [C/OL]. [2020-12-10]. <https://arxiv.org/ftp/arxiv/papers/2009/2009.01174.pdf>.
- [4] YU Z Z, SHI Y M, HUANG T J, *et al.* Kernel Quantization for Efficient Network Compression [C/OL]. [2020-12-10]. <https://arxiv.org/pdf/2003.05148.pdf>.
- [5] GONG C, CHEN Y, LU Y, *et al.* VecQ: Minimal Loss DNN Model Compression with Vectorized Weight Quantization [C/OL]. [2020-12-10]. <https://arxiv.org/pdf/2005.08501.pdf>.
- [6] HE Y, LIU P, WANG Z W, *et al.* Pruning Filter via Geometric Median for Deep Convolutional Neural Networks Acceleration // Proc of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington, USA: IEEE, 2019: 4335-4354.
- [7] HE Y, KANG G L, DONG X Y, *et al.* Soft Filter Pruning for Accelerating Deep Convolutional Neural Networks // Proc of the 27th International Joint Conference on Artificial Intelligence. New York, USA: ACM, 2018: 2234-2240.
- [8] LIU X G, WU L S, DAI C, *et al.* Compressing CNNs Using Multi-level Filter Pruning for the Edge Nodes of Multimedia Internet of Things. IEEE Internet of Things Journal, 2021. DOI: 10.1109/JIOT.2021.3052016.
- [9] SANDLER M, HOWARD A, ZHU M L, *et al.* Mobilenetv2: Inverted Residuals and Linear Bottlenecks // Proc of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington, USA: IEEE, 2018: 4510-4520.
- [10] ZHANG X Y, ZHOU X Y, LIN M X, *et al.* ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices // Proc of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington, USA: IEEE, 2018: 6848-6856.
- [11] HINTON G, VINYALS O, DEAN J. Distilling the Knowledge in a Neural Network [C/OL]. [2020-12-10]. <https://arxiv.org/pdf/1503.02531.pdf>.
- [12] WU B C, WAN A, YUE X Y, *et al.* Shift: A Zero Flop, Zero Parameter Alternative to Spatial Convolutions // Proc of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington, USA: IEEE, 2018: 9127-9135.
- [13] DENTON E, ZAREMBA W, BRUNA J, *et al.* Exploiting Linear Structure within Convolutional Networks for Efficient Evaluation // Proc of the 28th International Conference on Neural Information Processing Systems. Cambridge, USA: The MIT Press, 2014: 1269-1277.
- [14] SHI Y W, YAO X, CHEN R X, *et al.* Image Recognition Based on Multi-scale Dilated Lightweight Network Model // Proc of the 5th International Conference on Multimedia and Image Processing. New York, USA: ACM, 2020: 43-48.
- [15] ZHANG Z Y, LI J Y, SHAO W Q, *et al.* Differentiable Learning-to-Group Channels via Groupable Convolutional Neural Networks // Proc of the IEEE/CVF International Conference on Computer

- Vision. Washington , USA: IEEE , 2019: 3542–3551.
- [16] PHAM H , CUAN M , ZOPH B , *et al.* Efficient Neural Architecture Search via Parameters Sharing // Proc of the 35th International Conference on Machine Learning. New York , USA: ACM , 2018: 4095–4104.
- [17] WILLIAMS R J. Simple Statistical Gradient-Following Algorithms for Connectionist Reinforcement Learning. Machine Learning , 1992 , 8( 3/4 ) : 229–256.
- [18] LECUN Y , BOTTOU L , BENGIO Y , *et al.* Gradient-Based Learning Applied to Document Recognition. Proceedings of the IEEE , 1998 , 86( 11 ) : 2278–2324.
- [19] KRIZHEVSKY A. Learning Multiple Layers of Features from Tiny Images[R/OL]. [2020–12–10]. <https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf>.
- [20] LOSHCHELOV I , HUTTER F. SGDR: Stochastic Gradient Descent with Warm Restarts [C/OL]. [2020–12–10]. <https://arxiv.org/pdf/1608.03983v5.pdf>.
- [21] LIU H X , SIMONYAN K , YANG Y M. DARTS: Differentiable Architecture Search [C/OL]. [2020–12–10]. <https://openreview.net/pdf?id=S1eYHoC5FX>.
- [22] YU R C , LI A , CHEN C F , *et al.* NISP: Pruning Networks Using Neuron Importance Score Propagation // Proc of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington , USA: IEEE , 2018: 9194–9203.
- [23] LIN S H , JI R R , YAN C Q , *et al.* Towards Optimal Structured CNN Pruning via Generative Adversarial Learning // Proc of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington , USA: IEEE , 2019: 2785–2794.
- [24] WEN W , WU C P , WANG Y D , *et al.* Learning Structured Sparsity in Deep Neural Networks[C/OL]. [2020–12–10]. <https://arxiv.org/pdf/1608.03665v4.pdf>.

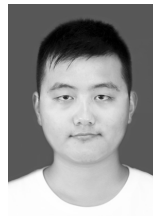
- [25] BYERLY A , KALGANOVA T , DEAR I. A Branching and Merging Convolutional Network with Homogeneous Filter Capsules[C/OL]. [2020–12–10]. <https://arxiv.org/pdf/2001.09136.pdf>.

### 作者简介



姚 潇 博士 副教授 ,主要研究方向为模式识别、计算机视觉.E-mail: yxkelvin@126.com.

( YAO Xiao , Ph.D. , associate professor. His research interests include pattern recognition and computer vision.)



史叶伟 硕士研究生 ,主要研究方向为模式识别、神经网络.E-mail: shiyewei@hhu.edu.cn.

( SHI Yewei , master student. His research interests include pattern recognition and neural network.)



霍冠英 博士 ,教授 ,主要研究方向为图像处理.E-mail: 20041686@hhu.edu.cn.

( HUO Guanying , Ph.D. , professor. His research interests include image processing.)



徐 宁( 通信作者) 博士 副教授 ,主要研究方向为模式识别.E-mail: 20101832@hhu.edu.cn.

( XU Ning ( Corresponding author ) , Ph.D. , associate professor. His research interests include pattern recognition.)