

MSINet: Twins Contrastive Search of Multi-Scale Interaction for Object ReID

摘要

神经网络结构搜索 (NAS), 是一门专门研究如何不借助于人工调试的方式就可以达到自动化设计高性能深度神经网络架构的技术, 这种针对特定任务的结构可以大大提高搜索性能。在重识别 (ReID) 任务中, NAS 方法倍受欢迎。先前的研究主要集中在如何使用 NAS 技术为 ReID 任务寻找新的优化目标和搜索空间, 但是这些研究忽略了图像分类任务和 ReID 任务之间的训练方法的不同之处。本文提出了一种新的 Twins Contrastive Mechanism(TCM) 方法, 更好的引导计算机学习如何构建适用于 ReID 任务的网络结构。其次, 本文提出了一个多尺度交互 (MSI) 搜索空间, 来搜索多尺度特征之间的合理交互操作。另外, 还引入了一个空间对齐模块 (SAM), 使得模型在面对来自不同来源的图像时能保持注意一致性。最后, 通过使用 NAS 方法, 论文找到了一种特定的神经网络架构, 称为 MSINet。

关键词: ReID; NAS; 多尺度交互

1 引言

目标重识别任务是在不同的摄像头视角或者不同的场景下, 找出是同一个目标, 比如在监控视频中追踪一个人或车辆。这在公共安全、零售和交通管理等方面都有很多实际应用。现有的目标重识别方法一般都是使用预先定义好的神经网络结构, 但是这些结构并不一定适用于 ReID 这个特殊的任务。因此, 通过 NAS 方法, 尝试找到更适合车辆重识别任务的网络结构是提高重识别性能的一种可能途径。这不仅有助于提高目标重识别系统的准确性, 也对公共安全、零售体验等有很多实际意义的应用产生积极影响。因此, 这个研究对于实际场景中应用目标重识别技术具有很重要的实际价值。

2 相关工作

2.1 神经网络结构搜索

神经网络结构搜索技术在计算机视觉领域越来越受欢迎, 这种针对特定任务的结构可以大大提高搜索性能。神经网络结构搜索方法主要分为四类: 强化学习 [1, 25]、进化算法 [14, 18]、梯度下降 [15, 16]、性能预测 [6, 13]。Liu 等人提出了一种可微分神经网络结构搜索的方法 [15], 提高了 NAS 的实用性。之后的研究者通过其他方法进一步改进了这个结构, 比如采

样策略 [22]、网络修剪 [2, 5]、渐进学习 [3]、协作竞争等 [4]。NAS 技术大部分使用在图像分类任务上，但是这种图像分类任务中的神经网络结构并不适用于目标重识别任务。因此本论文提出了一种新颖的搜索方案，使得搜索出来的神经网络结构更适合用于 ReID。

2.2 目标重识别网络的设计

当前的目标重识别任务主要使用为图像分类设计的网络结构 [19, 20]，比如基于常见的主干网络设计关注模块 [9, 10, 12, 23]，从而发现这些模块在区分局部差异方面的潜力。但是，这些方法通常需要较大的算力。还有的研究工作专门为 ReID 任务设计特定的结构，比如 Li 等人提出了一种 Filter Pairing 神经网络，它可以动态匹配特征图中的部分 [11]。Wang 等人通过 WConv 层分离和重新组合两个样本的特征 [21]。Guo 等人提取多尺度特征来直接评估样本之间的相似性 [8]。但是，如果在大型数据库上进行检索，使用孪生结构是非常不便的。Zhou et al. 将多尺度信息进行聚合，仅少量算力就有高精度 [24]。Quan 等人有部分感知模块引入 DARTS 搜索空间 [15]。Li 等人针对感受野尺度提出了一个新的搜索空间 [17]。这些方法的模型结构复杂性低，在参数量少的时候性能好，但性能比不上复杂性高的模型结构。与前面提到的方法不同，本论文设计了一个专门针对车辆重识别任务的轻量级的搜索结构，注重多尺度特征之间的合理交互。本论文通过这种方法搜索得到的 MSINet 在 ReID 任务上的性能表现较好。

3 本文方法

本文的目标是使用一个高效的神经网络结构搜索方法，找到适合目标重识别任务的轻量级主干网络结构。针对目标重识别任务的训练方式的不同，本文提出了“Twins Contrastive Mechanism”（双对比）的机制，为搜索过程提供更合理的引导。为了让不同尺度的特征在网络中更好地传递信息，本文设计了一个专门用于多尺度交互的搜索空间。此外，本文还引入了一个空间对齐模块，可以在有限的参数增长下提高网络的泛化能力。

3.1 双对比机制

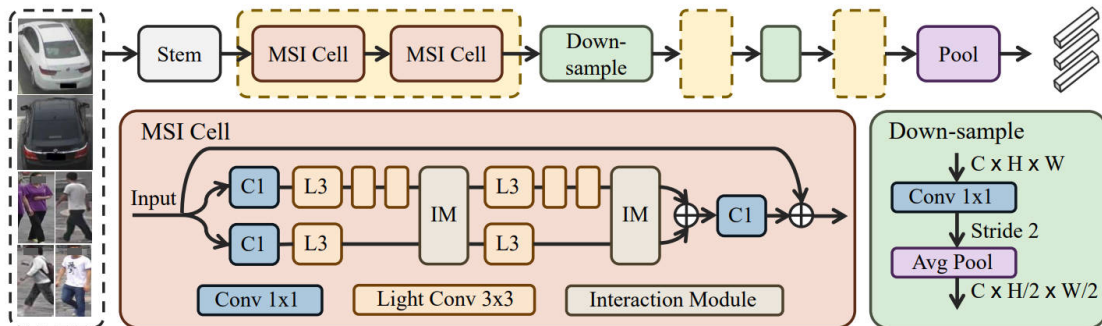


图 1. 论文提出的方法示意图

神经网络结构搜索技术的目标是自动搜索针对特定任务的最佳网络结构。受到 [15] 方法的启发，本论文建立了一个可微分的神经网络结构搜索方法。首先，将普通模型的参数定义

为 ω ，结构参数定义为 α 。对于搜索空间 \mathcal{O} 的网络层 i ， α_i 控制搜索空间中每个操作 o 的权重。输入的特征可以并行地通过所有操作进行传递，最终输出由交互操作模块输出的 softmax 函数进行加权：

$$f(\mathbf{x}_i) = \sum_{o \in \mathcal{O}} \frac{\exp\{\alpha_i^o\}}{\sum_{o' \in \mathcal{O}} \exp\{\alpha_i^{o'}\}} \cdot o(\mathbf{x}_i), \quad (1)$$

搜索空间进行搜索的过程会交替进行。利用训练集来更新模型的参数，然后利用验证集更新网络结构的参数。对于大多数为图像分类任务设计的神经网络结构搜索方法，训练集和验证集共用同一个类别的数据和损失计算也使用同一个线性分类层。

与标准的图像分类不同，作为一个开放集任务，目标重识别任务的训练集和测试集的类别不同。这种差异可能导致搜索到的网络结构在实际应用中表现不佳。因此，论文提出了一种新的双对比机制用于训练神经网络结构搜索到的 ReID 模型。具体方法如下：首先，使用两个独立的辅助内存，分别为 \mathcal{C}_{tr} 和 \mathcal{C}_{val} ，来存储训练集和验证集数据的嵌入特征。通过对每个类别的特征求平均值，得到质心特征，然后使用质心特征来对两个内存进行初始化。在每次训练迭代过程中，首先使用 \mathcal{C}_{tr} 计算训练的损失值，从而更新模型参数，使模型更好地适应训练数据。将特征 \mathbf{f} 嵌入到类别标签 j 中，通过在 \mathcal{C}_{tr} 中找到与类别 j 不同的样本并根据公式2计算对比分类损失函数，来引导模型学习更好的特征表示。

$$\mathcal{L}_{tr}^{cls} = -\log \frac{\exp(\mathbf{f} \cdot \mathbf{c}_{tr}^j / \tau)}{\sum_{n=0}^{N_{tr}^c} \exp(\mathbf{f} \cdot \mathbf{c}_{tr}^n / \tau)}, \quad (2)$$

其中 \mathbf{c}_{tr}^n 指的是类别 n 的记忆特征， N_{tr}^c 指的是训练集中的类别总数， τ 是温度参数，一般设置为 0.05 [7]。在更新模型参数之后，具有类别标签 j 的嵌入特征 \mathbf{f} 被整合到相应的记忆特征 \mathbf{c}_{tr}^j 中，计算公式为：

$$\mathbf{c}_{tr}^j \leftarrow \beta \mathbf{c}_{tr}^j + (1 - \beta) \mathbf{f}, \quad (3)$$

其中， β 为 0.2 [7]。然后，在验证集上评估更新后的模型，用 \mathcal{C}_{val} 替换公式2中的 \mathcal{C}_{tr} ，来生成验证损失值。然后，用验证损失值来更新结构参数，接着完成一次迭代。

3.2 多尺度交互搜索空间

虽然之前的目标重识别任务已经提出了局部视角和多尺度特征的方法，但这种方法比较单一并且有限制的条件。本论文认为，沿着网络结构的浅层和深层来动态地调整多尺度特征的使用比例，并且使用交互模块可以实现特征信息的交换，并充分利用多尺度特征。因此，本论文提出了多尺度交互搜索空间，来建立适用于目标重识别的轻量级网络结构。

如图1所示，网络结构主要由 MSI 块和下采样块组成。在每个 MSI 块中，有两个分支，每个分支都负责处理输入的特征。为了在处理这些特征时减轻计算负担，每个分支内部采用了 1×1 卷积的堆栈和多个深度可分离的 3×3 卷积来实现特定尺度。两个分支的比例 ρ 为 3:1。另外，这两个分支除了交互模块共享模型参数之外，其余模块不共享模型参数。交互模块有 4 种操作选项。两个分支输入特征为 $(\mathbf{x}_1, \mathbf{x}_2)$ ，4 种操作可以表示为：

1. 无操作：不涉及参数，直接输出输入特征 $(\mathbf{x}_1, \mathbf{x}_2)$ 。
2. 交换：直接交换两个分支的特征并输出 $(\mathbf{x}_2, \mathbf{x}_1)$ 。
3. 通道注意力门：引入一个多层感知器（MLP）来生成通道注意力门，公式如下：

$$G(\mathbf{x}) = \sigma(MLP(\mathbf{x})), \quad (4)$$

然后返回 $(G(\mathbf{x}_1) \cdot \mathbf{x}_1, G(\mathbf{x}_2) \cdot \mathbf{x}_2)$ 。MLP 由 2 个全连接层组成，通过引入 sigmoid 函数，确保了生成的值在 $(0, 1)$ 范围内，使其可以用作特定通道的加权系数。

4. 交叉注意力模块：

传统的通道注意力模块计算单个特征图内的通道相关性。首先将原始特征图 $\mathbf{x} \in \mathcal{R}^{C \times H \times W}$ 转为查询特征 $\tilde{\mathbf{x}} \in \mathcal{R}^{C \times N}$ ，其中 $N = H \times W$ 。然后，对查询特征 $\tilde{\mathbf{x}}$ 和键特征 $\tilde{\mathbf{x}}^\top$ 用矩阵乘法来计算相关激活，将计算出的相关激活转换为一个掩码，并与原始特征进行相加。

在交互之后，多尺度分支通过求和操作进行特征融合。在网络结构的开始部分，该模块使用与 OSNet [24] 中相同的干扰模块，包含一个 7×7 的卷积层和一个步幅为 2 的 3×3 最大池化。在搜索结束后，保留每个层中有最大权重 α_i^o 的交互操作 o ，来形成最终搜索到的网络结构。

在搜索完网络结构之后，模型在各种 ReID 任务上对网络结构进行验证。训练受到分类 ID 损失和三元损失的约束，其公式分别如公式5和公式6所示。

$$\mathcal{L}_{id} = \frac{1}{N} \sum_{i=1}^N -\log \left(\frac{\exp \mathbf{W}_i^\top \mathbf{f}_i}{\sum_j \exp \mathbf{W}_j^\top \mathbf{f}_i} \right), \quad (5)$$

其中 \mathbf{f}_i 是特征向量，其对应的分类器权重是 \mathbf{W}_i 。

$$\mathcal{L}_{tri} = [\mathcal{D}(\mathbf{f}_a, \mathbf{f}_p) - \mathcal{D}(\mathbf{f}_a, \mathbf{f}_n) + m]_+, \quad (6)$$

其中 \mathbf{f}_a , \mathbf{f}_p , \mathbf{f}_n 分别是锚点、最难的正样本和负样本的嵌入特征， $\mathcal{D}(\cdot, \cdot)$ 是欧氏距离， \mathbf{m} 是间隔参数， $[\cdot]_+$ 是 $\max(\cdot, 0)$ 函数。

3.3 空间对齐模块

在重识别的任务中，如果车辆受到姿势、光照或遮挡的影响，汽车的外观会发生变化。为了让模型能够关注到用来区分不同车辆的细节，该论文提出了空间对齐模块。

将该模块添加进去后，模型能够在处理不同的图像时，能够识别并对齐图像中的关键点，这有助于比较、分析和理解图像。

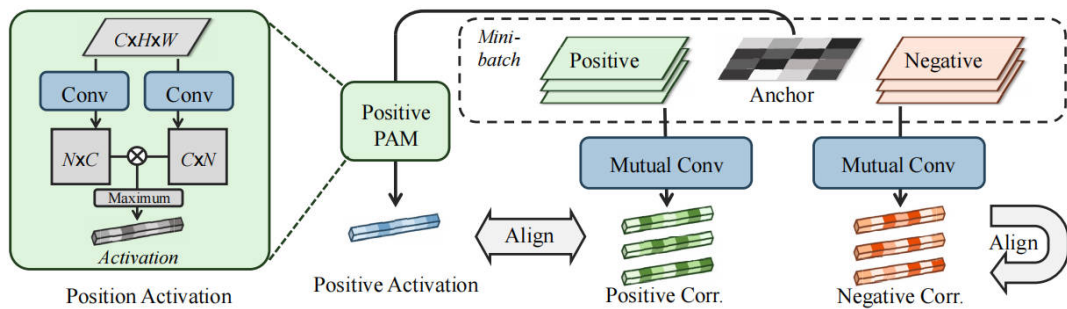


图 2. 空间对齐模块的网络结构图，左侧为位置激活模块

在图2的网络结构图中，首先，计算锚点特征与小批量中所有特征之间的相关激活向量。正向量与可学习的自激活进行对齐，而负向量与彼此进行对齐。左侧为位置激活模块的结构。

在“Mutual Conv”中，执行以下操作：

1. 计算小批量中特征图之间的位置相关图 \mathbf{A}

2. 数据 i 和 j 之间的激活表示为: $\mathbf{A}(i, j) = \tilde{\mathbf{x}}_j^\top \times \tilde{\mathbf{x}}_i, \tilde{\mathbf{x}} \in \mathcal{R}^{C \times N}$
3. 对数据 i 的每个位置取最大激活值, 公式为:

$$\mathbf{a}(i, j) = \max_{dim=1} \mathbf{A}(i, j), \quad (7)$$

对于负样本, 网络可能会将其识别成其背景。对齐样本 i 的所有相关性, 可以纠正网络存在的一些关注方向的偏差, 并关注更能区分不同车辆的细节。但是, 对齐负样本对不能突出与身份相关的特征, 因为负样本之间可能有很多差异, 比如不同的图像背景、车辆位置等。因此对于上面的问题, 该论文提出了一个位置激活模块来生成正样本对之间对齐的监督。空间对齐损失的公式为:

$$\mathcal{L}_{sa}(i) = \frac{1}{N_+} \sum_{p \in \mathcal{I}_+} (1 - S(\hat{\mathbf{a}}(i), \mathbf{a}(i, p))) + \frac{1}{N_-} \sum_{n_1, n_2 \in \mathcal{I}_-} (1 - S(\mathbf{a}(i, n_1), \mathbf{a}(i, n_2))), \quad (8)$$

其中, \mathcal{I}_+ 样本 i 的正指标的集合, 样本 i 的总数为 N_+ , 反之亦然。 $\hat{\mathbf{a}}(i)$ 表示生成的样本正对齐激活向量, $S(\cdot, \cdot)$ 表示余弦相似度。

4 复现细节

4.1 与已有开源代码对比

原论文代码使用的搜索空间是更加符合行人重识别任务的, 本次复现主要的改进是重新使用神经网络结构搜索出符合车辆重识别任务的神经网络结构框架, 使用 search 网络训练 VeRi-776 数据集, 每个 cell 有 2 个交互模块, 用了 6 个 cell, 然后得到 12 个交互模块的选择, 并将原来搜索出的操作顺序进行更新, 搜索出来的结构将更符合车辆重识别任务。对比浮现出来原论文的结果, 新搜索出来的神经网络结构的搜索效果有所提升。

4.2 实验环境搭建

- 1、硬件要求: Intel(R) Xeon(R) CPU E5-2643 v4 @ 3.40GHz 处理器的服务器, 总内存为 62GB, 以及装有 Windows 11 系统的计算机
- 2、软件要求: Pycharm、MobaXterm

4.3 创新点

本次复现的创新点如下:

1、神经网络结构的重新设计

本次复现通过神经网络结构搜索重新设计了网络结构, 使其更符合车辆重识别任务, 根据任务特点重新设计了神经网络结构的搜索空间。

2、操作顺序的更新

本次复现对搜索出的操作顺序进行了更新, 来优化网络的性能。

3、搜索效果的提升

本次复现新搜索出来的神经网络结构相比原论文在车辆重识别任务上有所提升。

5 实验结果分析

5.1 数据集

本次实验所用的数据集是 VeRi-776，包含超过 50000 张 776 辆车的图像，这些图像由 20 台摄像机拍摄，在 24 小时内覆盖 1.0 平方公里的面积，这使得该数据集可扩展到足以用于车辆 ReID 和其他相关研究。图像是在真实世界的无约束监视场景中捕获的，并标有不同的属性，例如：BBBox，类型，颜色和品牌。因此可以学习和评估车辆重识别的复杂模型。每辆车在不同的视点，照明，分辨率和遮挡下由 2 18 台摄像机拍摄，在实际监控环境中为车辆重识别提供高复发率。它还标有足够的牌照和时空信息，例如板块的 BBBox，板条，车辆的时间戳以及相邻相机之间的距离。

5.2 结果分析

下面表1是原论文以及对原论文代码进行复现的结果：

表 1. 原论文及原论文的复现结果

方法	轮数	R-1	R-5	R-10	mAP
MSINet(原)	350	96.8	-	-	78.8
MSINet(复现)	350	95.9	97.6	98.5	78.0

本次复现与原论文使用了相同的方法（MSINet），并且在轮数上也一致（350 轮）。在评价性能的指标上，使用了 R-1、R-5、R-10 和 mAP（平均准确率）。复现结果相比原论文在 R-1 上略有下降，在前几个匹配中的性能稍有下降，由于原论文没给出 R-5 和 R-10 的数据，所以此处 R-5 和 R-10 的数据对比意义不大。复现结果在 mAP 上相比原论文略有下降，在平均准确率方面，复现的模型性能稍逊于原论文。最后，复现的结果在一些指标上与原论文相比有轻微的波动，表现出略微的性能差异。这种差异可能是由于复现过程中的一些微小差异、超参数调整等原因导致的。

下面表2是原论文的复现结果与改进的复现结果对比：

表 2. 原论文的复现结果与改进的复现结果对比

方法	轮数	R-1	R-5	R-10	mAP
MSINet(复现)	350	95.9	97.6	98.5	78.0
MSINet(改)	350	96.2	98.0	98.6	78.2

本次对原论文的改进与复现的原论文使用了相同的方法（MSINet），并且在轮数上也一致（350 轮）。在评价性能的指标上，使用了 R-1、R-5、R-10 和 mAP，改进的复现结果在 R-1、R-5 和 R-10 上分别提高了 0.3、0.4、0.1，表明在前几个匹配中以及整体匹配性能上都有一些改善。改进的复现结果在 mAP 上略微提高了 0.2，说明改进的模型在平均准确率方面稍优于原复现结果。改进的复现结果在所有评价指标上都略有提升，说明通过重新设计神经网络结构，调整操作顺序，以符合车辆重识别任务的需求，改进的复现结果相对于原论文的复现结果在车辆重识别任务上表现更好，这也验证了改进方法的有效性。

6 总结与展望

本论文的研究重点在于对目标重识别任务，利用神经网络结构搜索技术，找到更适合重识别任务的神经网络结构。在本文中，我主要对原论文进行了复现，并在复现基础上进行了一定的改进，以适应车辆重识别任务。主要的改进重点在于神经网络结构的重新设计和操作顺序的优化，来提升模型在该任务上的性能。

在决定对神经网络结构重新设计时，也尝试了不同的改进方法，但是效果都不是很好。比如尝试增加模型的精度，去加神经网络的层数，但是出现了过拟合的情况，所以加层数是不能改善模型效果的。另外，修改激活函数、学习率、优化器、sam mode、sam radio 等参数，效果比复现的论文结果更差。最后，发现原论文的神经网络结构是更针对行人重识别任务的，然后就重新调整了神经网络结构搜索空间，使其更符合车辆重识别的特点。通过搜索优化，得到了新的神经网络结构，并对操作顺序进行了调整。最终，对比了原论文、复现结果以及改进的复现结果，发现改进的模型在车辆重识别任务上取得了一些性能提升。

然而，实现过程中也存在一些不足。首先，由于深度学习模型的训练过程复杂，复现和改进的过程中可能受到超参数、数据集划分等因素的影响，这些因素可能会引入一些随机性。其次，改进的性能提升相对较小，可能需要更深入的分析 and 实验，包括更多的超参数调优、数据增强策略等。

未来可以进一步深入分析模型性能的提升原因，包括每个改进点的贡献程度，并进行更精细的调优。进行更系统的超参数搜索，寻找最优的超参数配置，来提高模型性能。

参考文献

- [1] Irwan Bello, Barret Zoph, Vijay Vasudevan, and Quoc V Le. Neural optimizer search with reinforcement learning. In *International Conference on Machine Learning*, pages 459–468. PMLR, 2017.
- [2] Han Cai, Ligeng Zhu, and Song Han. Proxylessnas: Direct neural architecture search on target task and hardware. *arXiv preprint arXiv:1812.00332*, 2018.
- [3] Xin Chen, Lingxi Xie, Jun Wu, and Qi Tian. Progressive differentiable architecture search: Bridging the depth gap between search and evaluation. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 1294–1303, 2019.
- [4] Xiangxiang Chu, Tianbao Zhou, Bo Zhang, and Jixiang Li. Fair darts: Eliminating unfair advantages in differentiable architecture search. In *European conference on computer vision*, pages 465–480. Springer, 2020.
- [5] Xiyang Dai, Dongdong Chen, Mengchen Liu, Yinpeng Chen, and Lu Yuan. Da-nas: Data adapted pruning for efficient neural architecture search. In *Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXVII 16*, pages 584–600. Springer, 2020.

- [6] Tobias Domhan, Jost Tobias Springenberg, and Frank Hutter. Speeding up automatic hyperparameter optimization of deep neural networks by extrapolation of learning curves. In *Twenty-fourth international joint conference on artificial intelligence*, 2015.
- [7] Yixiao Ge, Feng Zhu, Dapeng Chen, Rui Zhao, et al. Self-paced contrastive learning with hybrid memory for domain adaptive object re-id. *Advances in neural information processing systems*, 33:11309–11321, 2020.
- [8] Yiluan Guo and Ngai-Man Cheung. Efficient and deep person re-identification using multi-level similarity. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2335–2344, 2018.
- [9] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [10] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.
- [11] Wei Li, Rui Zhao, Tong Xiao, and Xiaogang Wang. Deepreid: Deep filter pairing neural network for person re-identification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 152–159, 2014.
- [12] Wei Li, Xiatian Zhu, and Shaogang Gong. Harmonious attention network for person re-identification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2285–2294, 2018.
- [13] Chenxi Liu, Barret Zoph, Maxim Neumann, Jonathon Shlens, Wei Hua, Li-Jia Li, Li Fei-Fei, Alan Yuille, Jonathan Huang, and Kevin Murphy. Progressive neural architecture search. In *Proceedings of the European conference on computer vision (ECCV)*, pages 19–34, 2018.
- [14] Hanxiao Liu, Karen Simonyan, Oriol Vinyals, Chrisantha Fernando, and Koray Kavukcuoglu. Hierarchical representations for efficient architecture search. *arXiv preprint arXiv:1711.00436*, 2017.
- [15] Hanxiao Liu, Karen Simonyan, and Yiming Yang. Darts: Differentiable architecture search. *arXiv preprint arXiv:1806.09055*, 2018.
- [16] Renqian Luo, Fei Tian, Tao Qin, Enhong Chen, and Tie-Yan Liu. Neural architecture optimization. *Advances in neural information processing systems*, 31, 2018.
- [17] Ruijie Quan, Xuanyi Dong, Yu Wu, Linchao Zhu, and Yi Yang. Auto-reid: Searching for a part-aware convnet for person re-identification. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 3750–3759, 2019.

- [18] Esteban Real, Alok Aggarwal, Yanping Huang, and Quoc V Le. Regularized evolution for image classifier architecture search. In *Proceedings of the aaai conference on artificial intelligence*, volume 33, pages 4780–4789, 2019.
- [19] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4510–4520, 2018.
- [20] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826, 2016.
- [21] Yicheng Wang, Zhenzhong Chen, Feng Wu, and Gang Wang. Person re-identification with cascaded pairwise convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1470–1478, 2018.
- [22] Yuhui Xu, Lingxi Xie, Xiaopeng Zhang, Xin Chen, Guo-Jun Qi, Qi Tian, and Hongkai Xiong. Pc-darts: Partial channel connections for memory-efficient architecture search. *arXiv preprint arXiv:1907.05737*, 2019.
- [23] Zhizheng Zhang, Cuiling Lan, Wenjun Zeng, Xin Jin, and Zhibo Chen. Relation-aware global attention for person re-identification. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 3186–3195, 2020.
- [24] Kaiyang Zhou, Yongxin Yang, Andrea Cavallaro, and Tao Xiang. Omni-scale feature learning for person re-identification. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 3702–3712, 2019.
- [25] Barret Zoph, Vijay Vasudevan, Jonathon Shlens, and Quoc V Le. Learning transferable architectures for scalable image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 8697–8710, 2018.