

关于“MedNAS: Multiscale Training-Free Neural Architecture Search for Medical Image Analysis” [20] 一文的复现

摘要

深度神经网络在医学图像分类任务中已取得了显著成果，但为特定任务设计合适的架构既依赖专业知识又耗时。神经架构搜索 (NAS) 提供了一种有效的架构发现手段，在自然图像分类等领域取得了巨大成功。然而，医学图像具有独特的特点，如小区域和病变大小的多样性，这使它们与自然图像存在差异。该论文提出了一种新的 NAS 方法，称为多尺度无训练神经架构搜索 (MSTF-NAS)，并提出了一个新的缩减搜索空间，得到了一个适用于医学图像分类的神经架构搜索方法。

关键词：神经架构搜索；医学图像分类

1 引言

近年来，深度神经网络已经在多个图像任务中取得惊人的成果，如图像分类、目标检测等，深度神经网络在这些任务上的成功，大多归功于如 ResNet 等精心设计的神经架构。考虑到医学图像任务中，不同架构在不同数据集和疾病上的性能差异巨大，因此为特定的医学任务设计恰当的神经架构显得尤其重要。然而，为特定的医学任务单独设计神经架构是极其复杂、需要多次尝试和成本，并极度依赖专家经验的一个过程。为了克服这个缺点，神经架构搜索是一个有前景的解决方案。

而考虑到过去的神经架构搜索大多是基于验证的方法，这种方法在搜索的过程中需要对每个候选模型进行多次训练以对神经架构的性能进行评估，存在资源消耗过多的缺点；且常见的搜索空间大多为自然图像所设计，没有考虑到医学图像所具有的独特特点，如小区域和病变大小的多样性。因此，该论文引入了免训练的神经架构搜索方法，并使用三种免训练指标对神经架构的性能进行评估，以减少对设备资源的需求。此外，考虑到小区域和病变大小的多样性，该论文提出了一种使用了多种池化技术的搜索空间，以达到感知不同尺度的效果。最后使用 NSGA-II 算法构建多目标优化问题，来搜索到最佳的神经架构。该论文所提出的方法被称为多尺度无训练神经架构搜索 (MSTF-NAS)。

2 相关工作

该部分将对与论文论文相关的多尺度特征提取、神经架构搜索和免训练指标这三个相关方向进行回顾。

2.1 多尺度特征提取

多尺度特征提取方法的作用是从输入的数据中提取到不同尺度的特征，方法包括将输入图像变换到不同尺寸，再使用主干网络来提取不同尺度的特征 [3, 6]。该论文则是通过神经架构搜索技术，自动组合不同大小的池化核，从而达到多尺度特征提取的效果。

2.2 神经架构搜索

神经架构搜索是一种用于自动化确定神经网络架构的技术，其搜索的内容包括参数量、操作算子、网络连接、网络层数等一系列相关的内容。过去的神经架构搜索大多可以分为三类：基于强化学习的神经架构搜索 [2, 14–16, 21]、基于进化算法的神经架构搜索 [4, 5, 9, 11] 和基于梯度的神经架构搜索 [10]。这三类方法大多通过不断训练候选网络，将在验证集中的准确率等指标来评估候选网络的性能，这种手段存在消耗资源极多的缺点，为神经架构搜索这一方向带来了不可忽略的计算资源需求。为了解决这个问题，研究者们提出了 Zero-Shot 的方法 [1, 7, 13]。Zero-Shot 这种方法不需要对候选网络进行训练，而是通过免训练指标在一个小批次的前向传播中，直接评估候选网络的性能，从而极大地减少了计算花销。该论文便使用了 Zero-Shot 的方法来减少神经架构搜索过程中所需要的计算花销。

2.3 免训练指标

免训练指标是 Zero-Shot 方法的核心，它通过在一个小批次的前向传播中，直接评估候选网络的性能，以此来达到减少了计算花销的目标。不同的免训练指标会评估候选网络在不同方面的性能，比如表达性、渐进性、复杂性、可训练性等一系列的性能，再根据对各方面性能的偏好选择出最佳网络。常见的免训练指标 [1] 有：fisher [18]、grad_norm、grasp [19]、jacob_cov [12]、l2_norm、plain、snip [8]、synflow [17] 等。该文章采用了 jacob_cov、synflow、flops 这三个免训练指标，并将其构建为双目标优化问题。

3 本文方法

3.1 本文方法概述

本文将要复现的工作，使用了与 Darts [10] 相似的搜索空间，包括 Normal Cell 和 Reduct Cell 这两个部分。该论文的整体算法流程是：首先生成一系列候选网络，再利用免训练指标对候选网络的性能进行衡量，利用 NSGA-II 算法根据性能筛选出亲本，通过交叉和变异生成新一代候选网络，直至满足终止条件。该算法的总体流程可以见图 1。

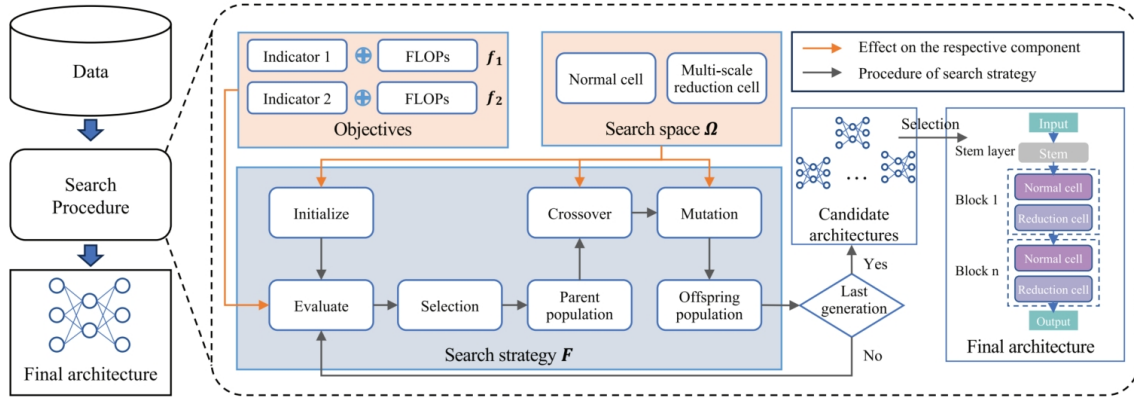


图 1. MSTF-NAS 整体流程

3.2 搜索空间

所复现论文的搜索空间由 Normal Cell 和 Reduction Cell 组成，其中 Normal Cell 与 Darts [10] 中的相似，在此不再复述。对于 Reduct Cell，该论文采用结合不同核大小的池化操作，在搜索过程中组合出 4 种最佳的池化操作组合，从而达到捕捉不同尺度特征的效果，其示意图可以见图 2。而整个搜索空间的编码则可以见图 3。

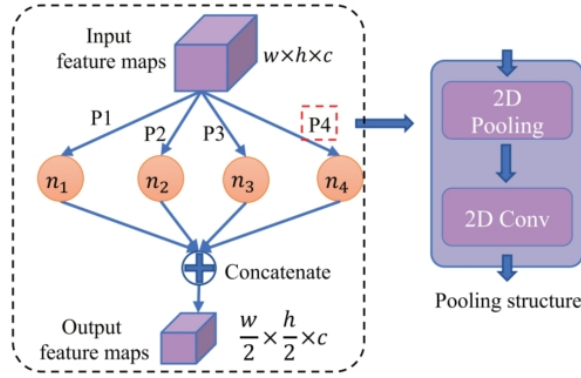


图 2. Reduction Cell 示意图

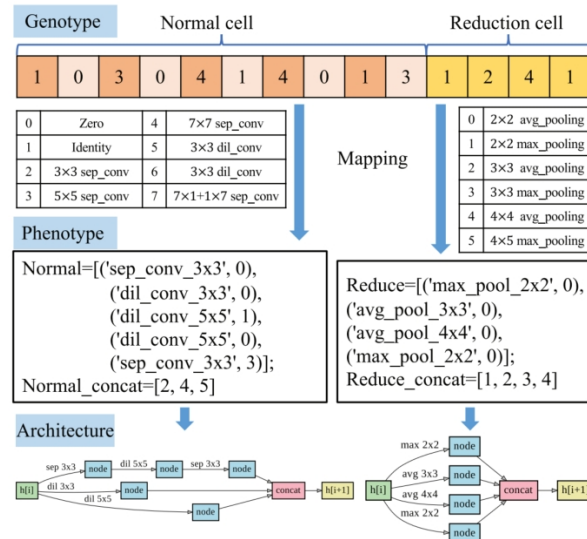


图 3. 搜索空间编码示意图

3.3 搜索算法

该论文的搜索算法使用了 NSGA-II 算法，对每一块的架构都分别进行搜索，而不采用通用的搜索。在优化问题的构建上，该论文将 $jacob_cov$ 、 $synflow$ 、 $flops$ 这三个免训练指标构建为双目标问题，以在搜索过程中平衡网络架构性能及其参数量。双目标问题的具体构建如下：

$$\begin{cases} f_1 = -jacob_cov - flops \\ f_2 = -synflow + flops \end{cases} \quad (1)$$

3.4 实验结果

所复现的论文对于提出的 MSTF-NAS 的有效性，进行了多个方面的实验来进行验证。本节将会抽取论文中部分实验进行介绍。

首先，论文在 MedMNIST 数据集上，与多个具有代表性的神经网络架构或神经架构搜索方法，通过准确率和 AUC 指标进行了对比。从图 4 和图 5 中可以看到，MSTF-NAS 所搜索得到的神经架构在多个子数据集上的准确率和 AUC 指标都超越了所对比的具有代表性的工作。

Methods	Path	Chest	Derma	OCT	Pneum	Retina	Breast	Blood	Tissue	OrganA	OrganC	OrganS	AVG.
ResNet-18 (28)	0.907	0.947	0.735	0.743	0.854	0.524	0.863	0.958	0.676	0.935	0.900	0.782	0.819
ResNet-18 (224)	0.909	0.947	0.754	0.763	0.864	0.493	0.833	0.963	0.681	0.951	0.920	0.778	0.821
ResNet-50 (28)	0.911	0.947	0.735	0.762	0.854	0.528	0.812	0.956	0.680	0.935	0.905	0.770	0.816
ResNet-50 (224)	0.892	0.948	0.731	0.776	0.884	0.511	0.842	0.950	0.680	0.947	0.911	0.785	0.821
auto-sklearn	0.716	0.779	0.719	0.601	0.855	0.515	0.803	0.878	0.532	0.762	0.829	0.672	0.722
AutoKeras	0.834	0.937	0.749	0.763	0.878	0.503	0.831	0.961	0.703	0.905	0.879	0.813	0.813
Google AutoML	0.728	0.948	0.768	0.771	0.946	0.531	0.861	0.966	0.673	0.886	0.877	0.749	0.809
DARTS	0.872	0.934	0.749	0.712	0.874	0.510	0.832	0.953	0.648	0.926	0.791	0.808	0.801
SNAS	0.850	0.938	0.737	0.708	0.871	0.515	0.811	0.946	0.708	0.918	0.891	0.778	0.806
HOPNAS	0.912	0.947	0.759	0.761	0.852	0.523	0.853	0.958	0.698	0.937	0.911	0.803	0.826
NSGA-NET	0.866	0.947	0.744	0.765	0.907	0.540	0.846	0.970	0.712	0.952	0.923	0.820	0.833
Random search	0.854	0.946	0.773	0.760	0.904	0.542	0.897	0.966	0.717	0.955	0.923	0.820	0.838
MSTF-NAS (Ours)	0.910	0.945	0.774	0.780	0.912	0.550	0.872	0.976	0.740	0.962	0.936	0.838	0.850

图 4. MedMNIST 数据集上，MSTF-NAS 与其他工作的准确率对比

Methods	Path	Chest	Derma	OCT	Pneum	Retina	Breast	Blood	Tissue	OrganA	OrganC	OrganS	AVG.
ResNet-18 (28)	0.983	0.768	0.917	0.943	0.944	0.717	0.901	0.998	0.930	0.997	0.992	0.972	0.922
ResNet-18 (224)	0.989	0.773	0.920	0.958	0.956	0.710	0.891	0.998	0.933	0.998	0.994	0.974	0.925
ResNet-50 (28)	0.990	0.769	0.913	0.952	0.948	0.726	0.857	0.997	0.931	0.997	0.992	0.972	0.920
ResNet-50 (224)	0.989	0.773	0.912	0.958	0.962	0.716	0.866	0.997	0.932	0.998	0.993	0.975	0.923
auto-sklearn	0.934	0.649	0.902	0.887	0.942	0.690	0.836	0.984	0.828	0.963	0.976	0.945	0.878
AutoKeras	0.959	0.742	0.915	0.955	0.947	0.719	0.871	0.998	0.941	0.994	0.990	0.974	0.917
Google AutoML	0.944	0.778	0.914	0.963	0.991	0.750	0.919	0.998	0.924	0.990	0.988	0.964	0.927
DARTS	0.975	0.732	0.913	0.953	0.965	0.742	0.912	0.994	0.901	0.987	0.969	0.910	0.913
SNAS	0.969	0.733	0.906	0.949	0.974	0.753	0.894	0.996	0.921	0.979	0.927	0.952	0.913
HOPNAS	0.987	0.763	0.899	0.948	0.971	0.770	0.907	0.996	0.913	0.995	0.998	0.975	0.927
NSGA-NET	0.979	0.779	0.915	0.958	0.965	0.759	0.857	0.999	0.942	0.999	0.993	0.978	0.927
Random search	0.980	0.774	0.923	0.956	0.963	0.750	0.921	0.999	0.944	0.999	0.994	0.982	0.932
MSTF-NAS (Ours)	0.990	0.791	0.934	0.968	0.963	0.755	0.930	0.999	0.951	0.999	0.996	0.983	0.938

图 5. MedMNIST 数据集上，MSTF-NAS 与其他工作的 AUC 对比

其次，论文还计算了 MSTF-NAS 算法搜索得到神经网络架构的用时，并与其他具有代表性的神经架构搜索算法所消耗的时间进行对比。通过图 6 中实验结果可以看到，相比于过去基于梯度、基于强化学习、基于进化算法等的神经架构搜索技术，采用 Zero-Shot 和免训练指标的 MSTF-NAS，在花销上有极为显著的优势。这证明了 MSTF-NAS 方法的高效性和低成本。

Methods	Path	Chest	Derma	OCT	Pneum	Retina	Breast	Blood	Tissue	OrganA	OrganC	OrganS	AVG.
Google AutoML	180	180	120	180	60	60	60	180	240	120	120	120	135
DARTS	1496	1184	121	1181	25	16	29	196	538	478	342	519	510
SNAS	691	585	45	443	11	7	13	474	1080	195	133	154	319
HOPNAS	465	390	38	372	8	5	8	55	985	152	92	112	224
NSGA-NET	2645	1067	223	3061	108	35	14	273	4147	799	280	281	1078
MSTF-NAS (Ours)	34	24	21	24	19	21	10	20	31	23	20	16	22

图 6. MSTF-NAS 与其他工作的搜索花销对比

然后，论文通过消融实验，探索了使用不同池化核大小对于不同尺度特征的影响，和不同免训练指标组合对 MSTF-NAS 算法的结果影响，为所提出的使用不同池化核大小来提取不同尺度特征的搜索空间和平衡性能与参数量的免训练指标组合，提供了实验结果作为依据。实验结果分别可见图 7 和图 8。

Pooling	AUC		
	Large	Medium	Small
2x2 average	74.69	74.18	77.05
2x2 max	76.36	74.32	80.13
3x3 average	80.29	73.58	<u>83.61</u>
3x3 max	<u>77.84</u>	76.89	<u>82.49</u>
4x4 average	<u>76.82</u>	71.28	83.78
4x4 max	73.25	<u>74.81</u>	77.72

图 7. 不同池化核对不同区域大小的 AUC 结果影响

Model	AUC	ACC	# Param (MB)
<i>fisher+synflow</i>	0.922	0.833	5.46
<i>jacob_cov+synflow</i>	0.938	0.850	4.47
<i>fisher+grad_norm</i>	0.929	0.828	5.91
<i>jacob_cov+grad_norm</i>	0.930	0.841	4.20
<i>fisher+jacob_cov</i>	0.936	0.846	5.08
<i>grad_norm+synflow</i>	0.920	0.834	5.25

图 8. 不同免训练指标组合对结果的影响

最后，论文还展示了一个在 PneumoniaMNIST 上搜索得到的神经架构的示意图，可以见图 9。

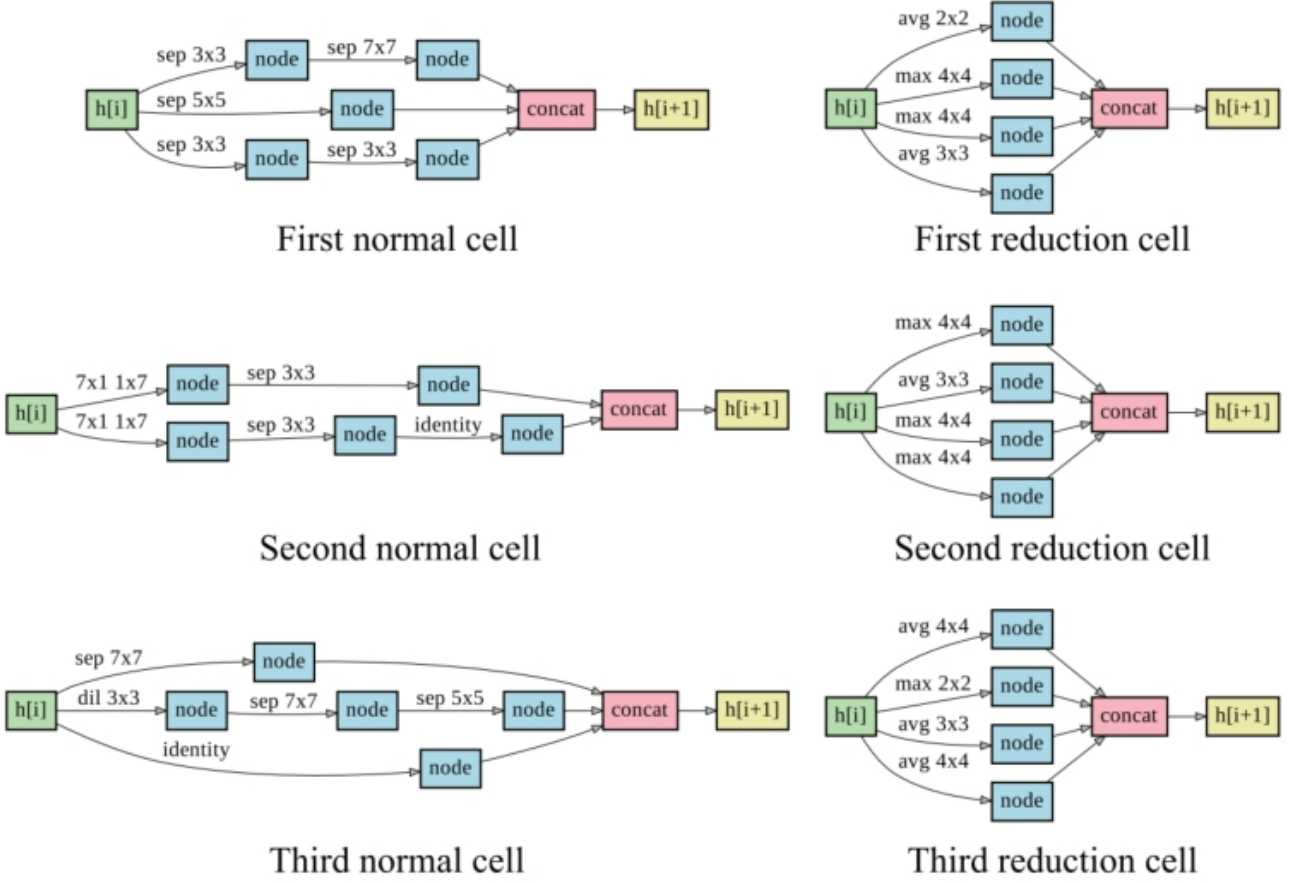


图 9. 在 PneumoniaMNIST 上搜索得到的神经架构

4 复现细节

4.1 与已有开源代码对比

由于 [20] 一文并没有公开其源代码和学习率、预处理、批大小等相关的实验参数设置，此文章的工作量主要在基于对该论文理解进行的原论文的复现。在复现的过程中，对于模型的基础操作，如深度分离卷积、卷积等一系列基本操作，引用了 [10] 中所实现的基础代码；对于免训练指标 jacob_cov 和 synflow 的计算，则引用了 [1] 中所实现的免训练指标代码。

根据论文的思路，该文章对论文的搜索空间、搜索方法、多目标问题的构建等进行了复现，并在医学数据集上进行了测试。此外，还将原论文中的 NSGA-II 算法使用 MOEA/D 算法进行代替，以图改进。

4.2 实验环境搭建

本文所使用的深度学习框架为 Pytorch，进化算法所使用的库为 pymoo，所使用的数据集为 MedMNIST 中的 MedMNIST2D 数据集。MeDMNIST2D 是一个公开的医学图像数据集，它包括多种子数据集，包括结肠病理学、皮肤镜、视网膜光学相干断层扫描、肺炎胸部 X 射线、视网膜眼底摄像、乳腺超声、血细胞显微镜、A 组腹部 CT、C 组腹部 CT、S 组腹部 CT。

对于实验结果性能的评价，本实验采取在 MedMNIST 数据集上的准确率作为性能指标进行衡量。对于数据预处理，主要使用 AugMix 进行处理，有部分数据集没有使用 AugMix 进行

处理；优化器使用 SGD；损失函数使用 CrossEntropyLoss；对于学习率的调节使用 CosineAnnealingLR 或 ExponentialLR。部分实验参数可见表 1。

数据集	学习率	批大小	severity	mixture_width	alpha
PathMNIST	0.05	512	10	1	1.5
DermaMNIST	0.025	128	10	1	1.0
OCTMNIST	0.05	512	10	1	1.5
PneumoniaMNIST	0.0005	64/128	10	1	1.5
RetinaMNIST	0.001	128	3	3	1.0
BreastMNIST	0.001	128	3	3	1.0
BloodMNIST	0.025	128	-	-	-
OrganAMNIST	0.025	256	10	1	1.5
OrganCMNIST	0.025	128	-	-	-
OrganSMNIST	0.025	128	10	1	1.5

表 1. 部分实验参数

4.3 使用说明

Search.py 实现了搜索功能，运行该文件，可以为选定的数据集搜索出一个恰当的神经网络架构；Train.py 实现了训练功能，运行该文件，可以在指定数据集上训练当前的神经网络架构，并输出性能指标。

4.4 创新点

本文将原论文中的 NSGA-II 算法使用 MOEA/D 算法进行代替，以图改进。

5 实验结果分析

通过表 2 中的数据可以知道，本实验中复现的 MSTF-NAS 在 MedMNIST 数据集上的效果，对比原论文中给出的数据，可以说基本达到了原文的水平。

而对比使用 NSGA-II 算法和 MOEA/D 算法的实验结果可以看到，两种不同的搜索算法之间最终的结果，并无较大差异，两者在各个子数据集上的实验结果基本一致，可以知道对于神经结构搜索这一学科方向，限制其搜索到更佳的神神经架构的瓶颈，不在于搜索算法的好坏、差异，更多的在于免训练指标上。免训练指标对神经架构性能预测的准确与否、鲁棒性如何、泛用性如何是神经架构搜索技术得到更进一步提升的关键所在。因此，在目前阶段下，应该更注重对免训练指标的改善。

Methods	Path	Derma	OCT	Pneum	Retina	Breast	Blood	OrganA	OrganC	OrganS	AVG.
MSTF-NAS	0.910	0.774	0.780	0.912	0.550	0.872	0.976	0.962	0.936	0.838	0.851
MSTF-NAS(复现)	0.938	0.757	0.776	0.897	0.547	0.883	0.969	0.957	0.922	0.825	0.847
MSTF-NAS(改进)	0.939	0.758	0.771	0.902	0.535	0.879	0.967	0.954	0.928	0.823	0.846

表 2. MSTF-NAS、MSTF-NAS(复现)、MSTF-NAS(修改) 在 MedMNIST 上的准确率

6 总结与展望

本报告对 [20] 进行了复现，且基本在 MedMNIST 数据集上的准确率指标达到了原文的水平。在所复现的论文中，其使用的免训练指标均为普通任务中所使用的指标，缺少对医学任务的定制。在未来的工作中，可以考虑对使用的免训练指标进行更进一步的挖掘，尝试能否提出、寻找到专用于医学任务的免训练指标。

参考文献

- [1] Mohamed S Abdelfattah, Abhinav Mehrotra, Łukasz Dudziak, and Nicholas D Lane. Zero-cost proxies for lightweight nas. *arXiv preprint arXiv:2101.08134*, 2021.
- [2] Zoph Barret, V Le Quoc, et al. Neural architecture search with reinforcement learning. In *International conference on learning representatoinis*, volume 1, 2017.
- [3] Liang-Chieh Chen, Yi Yang, Jiang Wang, Wei Xu, and Alan L Yuille. Attention to scale: Scale-aware semantic image segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3640–3649, 2016.
- [4] Xiangxiang Chu, Bo Zhang, and Ruijun Xu. Multi-objective reinforced evolution in mobile neural architecture search. In *europcan conference on computer vision*, pages 99–113. Springer, 2020.
- [5] Emmanuel Dufourq and Bruce A Bassett. Eden: Evolutionary deep networks for efficient machine learning. In *2017 Pattern Recognition Association of South Africa and Robotics and Mechatronics (PRASA-RobMech)*, pages 110–115. IEEE, 2017.
- [6] Xi Fang and Pingkun Yan. Multi-organ segmentation over partially labeled datasets with multi-scale feature abstraction. *IEEE Transactions on Medical Imaging*, 39(11):3619–3629, 2020.
- [7] Junghyup Lee and Bumsu Ham. Az-nas: Assembling zero-cost proxies for network architecture search. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5893–5903, 2024.
- [8] Namhoon Lee, Thalaiyasingam Ajanthan, and Philip HS Torr. Snip: Single-shot network pruning based on connection sensitivity. *arXiv preprint arXiv:1810.02340*, 2018.
- [9] Hanxiao Liu, Karen Simonyan, Oriol Vinyals, Chrisantha Fernando, and Koray Kavukcuoglu. Hierarchical representations for efficient architecture search. *arXiv preprint arXiv:1711.00436*, 2017.
- [10] Hanxiao Liu, Karen Simonyan, and Yiming Yang. Darts: Differentiable architecture search. *arXiv preprint arXiv:1806.09055*, 2018.

- [11] Zhichao Lu, Ian Whalen, Vishnu Boddeti, Yashesh Dhebar, Kalyanmoy Deb, Erik Goodman, and Wolfgang Banzhaf. Nsga-net: neural architecture search using multi-objective genetic algorithm. In *Proceedings of the genetic and evolutionary computation conference*, pages 419–427, 2019.
- [12] Joe Mellor, Jack Turner, Amos Storkey, and Elliot J Crowley. Neural architecture search without training. In *International conference on machine learning*, pages 7588–7598. PMLR, 2021.
- [13] Yameng Peng, Andy Song, Haytham M Fayek, Vic Ciesielski, and Xiaojun Chang. Swapnas: Sample-wise activation patterns for ultra-fast nas. *arXiv preprint arXiv:2403.04161*, 2024.
- [14] Hieu Pham, Melody Guan, Barret Zoph, Quoc Le, and Jeff Dean. Efficient neural architecture search via parameters sharing. In *International conference on machine learning*, pages 4095–4104. PMLR, 2018.
- [15] Ilija Radosavovic, Raj Prateek Kosaraju, Ross Girshick, Kaiming He, and Piotr Dollár. Designing network design spaces. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10428–10436, 2020.
- [16] Mingxing Tan, Bo Chen, Ruoming Pang, Vijay Vasudevan, Mark Sandler, Andrew Howard, and Quoc V Le. Mnasnet: Platform-aware neural architecture search for mobile. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2820–2828, 2019.
- [17] Hidenori Tanaka, Daniel Kunin, Daniel L Yamins, and Surya Ganguli. Pruning neural networks without any data by iteratively conserving synaptic flow. *Advances in neural information processing systems*, 33:6377–6389, 2020.
- [18] Jack Turner, Elliot J Crowley, Michael O’Boyle, Amos Storkey, and Gavin Gray. Block-swap: Fisher-guided block substitution for network compression on a budget. *arXiv preprint arXiv:1906.04113*, 2019.
- [19] Chaoqi Wang, Guodong Zhang, and Roger Grosse. Picking winning tickets before training by preserving gradient flow. *arXiv preprint arXiv:2002.07376*, 2020.
- [20] Yan Wang, Liangli Zhen, Jianwei Zhang, Miqing Li, Lei Zhang, Zizhou Wang, Yangqin Feng, Yu Xue, Xiao Wang, Zheng Chen, et al. Mednas: Multi-scale training-free neural architecture search for medical image analysis. *IEEE Transactions on Evolutionary Computation*, 2024.
- [21] 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.