神经结构搜索的演变与效率：缩小专家设计与自动化优化之间的差距

Fanfei Meng, Chen-Ao Wang, Alexander Brown

fanfeimeng2023@u.northwestern.edu

Northwestern University, Evanston, IL, United States

Tianjin University, Weijin Road Campus: No. 92 Weijin Road, Nankai District, Tianjin, 300072

摘要. 本文全面概述了神经结构搜索（NAS），重点强调了它从手动设计到自动化、计算驱动的方法的演变。它介绍了 NAS 的起源和发展，强调了它在医学成像和自然语言处理等不同领域的应用。该文件详细说明了从专家驱动设计到算法驱动过程的转变，探索了强化学习和进化算法等初始方法。它还讨论了计算需求的挑战以及高效 NAS 方法（如可微分结构搜索和硬件感知 NAS）的出现。本文进一步阐述了 NAS 在计算机视觉、NLP 等领域的应用，展示了其多样性和优化不同任务中神经网络结构的潜力。文章还介绍了未来的方向和挑战，包括计算效率和与新兴 AI 领域的整合，展示了 NAS 的动态性质和持续向更复杂、更高效的架构搜索方法演变的趋势。

关键词: 神经结构搜索；自动化设计优化；计算效率；进化算法；可微分结构搜索

1. 介绍

神经结构搜索（NAS）的兴起和早期发展代表了人工智能，特别是深度学习领域的一次变革性阶段。自动化神经网络架构设计的追求已经取得了重要的里程碑，研究工作主要集中在克服手动架构设计的局限性，并利用计算策略发现最优的网络结构。NAS 领域的早期研究以理解和改进循环神经网络（如长短时记忆（LSTM）网络）为标志。[1] 对 LSTM 变体进行了最大规模的研究之一，通过随机搜索评估了它们的效用并优化了超参数，这突显了早期对自动化神经网络设计方面的兴趣。[2] 介绍了 Inception 架构，展示了精心设计的潜力，可以更高效地利用计算资源，这一原则后来影响了 NAS 方法。[3] 强调了医学成像领域对 NAS 的日益浓厚的兴趣，其中深度学习模型（包括通过 NAS 设计的模型）在某些识别任务中的能力显示出超越人类表现的潜力。[4] 进一步举例说明了 NAS 如何演化为自然语言处理等专门领域，并引入了 NAS-Bench-NLP，通过提供全面的基准测试来推动该领域的研究。.这些早期的研究工作不仅推动了 NAS 领域的发展，还为后续的研究奠定了坚实的基础。它们展示了 NAS 在改进神经网络设

计、提高模型性能以及扩展应用到不同领域方面的巨大潜力。随着技术的不断进步，我们有理由相信 NAS 将继续推动人工智能领域的创新和进步。

从 NAS 研究的早期阶段到其当前状态的发展轨迹，凸显了在不同领域自动化和优化神经网络设计的广泛而宏大的努力。从提升 LSTM 网络到开创卷积神经网络（CNN）架构，再到拓展到医学和语言处理应用，NAS 体现了从手动、专家驱动的设计向自动化、计算驱动的架构搜索过程的转变。这一演变标志着深度学习领域民主化和加速创新的重要转变，有望解锁 AI 系统的新能力和效率。NAS 不仅改变了神经网络设计的方式，还促进了深度学习领域的技术进步。通过自动化搜索和优化网络架构，NAS 能够快速发现和采用更有效的结构和配置，从而提高模型的性能和效率。这种自动化的设计优化减少了手动调整和网络设计的繁琐过程，使研究人员和开发者能够更专注于解决具体问题和改进模型性能。

随着 NAS 的不断发展，它在各个领域的应用也越来越广泛。在医学领域，NAS 已经应用于医学图像分析、疾病诊断和治疗辅助等方面，帮助医生更准确地诊断疾病和制定治疗方案。在自然语言处理领域，NAS 被用于改进语音识别、机器翻译和文本生成等任务，提高了模型的准确性和效率。此外，NAS 还促进了深度学习社区的创新和合作。通过共享资源和代码，研究人员可以相互合作、交流和验证 NAS 方法的有效性，从而加速了整个领域的发展。这种开放的合作精神为深度学习的发展注入了新的活力和动力。

综上所述，NAS 的演变和发展标志着深度学习领域的重大进步和创新。它不仅改变了神经网络设计的方式，还促进了技术进步、拓宽了应用领域，并推动了社区的合作和创新。随着 NAS 的不断发展和完善，我们有望见证 AI 系统的新能力和效率，推动人工智能的广泛应用和发展。

最初，神经网络设计主要由专家驱动，需要深入了解不同的架构选择（如层数、层类型（例如卷积层、循环层）以及层之间的连接）如何影响网络的性能。这种专业知识是通过大量实验获得的。NAS 通过引入能够探索广阔架构空间的算法，改变了这一范式，识别出通常在手动过程中被忽略的最佳配置。NAS 中的第一个模型和算法是基础性的，为更复杂的技术奠定了基础。这些早期工作探索了简单的策略，如网格搜索和随机搜索，并迅速演变为更复杂的方法，如强化学习、进化算法和基于梯度的方法。NAS 的演变反映了人工智能中的一个更广泛趋势：即向能够自主学习和适应的系统转变，减少对人为输入的依赖。NAS 位于多个学科的交叉点：机器学习、优化、统计学和计算理论。这些领域的进步推动了 NAS 的发展，受益于计算资源的改进、对深度学习模型的理论理解以及大型数据集可用性的不断增长。这种多学科性质使 NAS 成为一个充满活力和快速发展的领域，吸引了学术界和工业界的关注。

2. 背景与相关工作

2.1 减少网络设计中的人力投入

神经结构搜索（NAS）的根本目标是减轻在神经网络架构开发过程中传统上所需的大量手动工作和专业知识。传统上，有效神经网络的设计是专家领域，需要漫长的试错过程来确定最佳配置。NAS 引入了一种革命性的方法，旨在自动化和改进这一设计过程。通过利用先进的算法，NAS 致力于使神经网络设计更加普及，更易于普遍访问，并减少对专业知识的依赖。这种方法不仅加速了高效神经架构的发现，而且通过探索可能超出人类直觉的大量设计可能性，扩大了创新的范围。NAS 通过自动化和优化设计过程，为神经网络架构的发展开辟了新的道路。它减

少了对手动调优和专家知识的依赖，使得神经网络的开发更加高效和普及。随着 NAS 技术的不断进步，我们有望看到更多创新、高效的神经网络架构的出现，推动人工智能领域的发展。

2.2 探索更广阔的架构可能性空间

传统设计神经网络的方法在很大程度上受到了对人类专业知识的依赖的限制，这本质上限制了对广阔架构空间的探索，仅限于先前的经验和直觉猜测的领域。这种方法经常错过发现创新和更高效的架构，这些架构超出了传统智慧的范围。神经结构搜索（NAS）作为解决这个问题的革命性方案而出现，利用强大的算法系统地探索更广泛的架构可能性。通过这样做，NAS 为识别新型神经网络配置打开了大门，这些配置虽然对人类设计师来说可能违反直觉，但可能提供卓越的性能和效率。NAS 促进的这种全面探索不仅拓宽了潜在设计的视野，而且使设计过程更加普及，使高性能神经网络的发展更加易于访问，并减少了专业知识瓶颈的约束。

因此，NAS 不仅是神经网络设计领域的重大突破，也为整个深度学习社区开辟了新的可能性。它使研究人员和开发者能够以前所未有的方式探索和优化神经网络架构，从而推动人工智能技术的进步。随着 NAS 技术的不断发展和完善，我们有望看到更多创新、高效的神经网络架构的出现，为各种应用带来更好的性能和效率。

2.3 最初的 NAS 策略与算力限制

神经网络架构搜索（NAS）的早期方法采用了强化学习（RL）和进化算法（EAs）等策略来探索和优化神经网络架构。基于 RL 的 NAS 方法利用策略网络根据性能反馈生成和优化架构，而 EA 则从生物进化中汲取灵感，通过选择、突变和重组多代架构来提高性能。尽管这些方法具有创新性，但这些初始策略需要大量的计算资源，从而限制了其实际应用。特别是对于图像分类中的深度卷积神经网络（CNN）的进化设计方法，强调了平衡分类性能和计算成本的多目标优化的重要性[4]。随后的研究为应对这些计算挑战引入了更有效的 NAS 方法[5]：

自适应可伸缩 NAS（AS-NAS）：该方法结合了简化的 RL 算法、增强的 I-Ching 占卜进化算法（IDEA）以及可变架构编码策略，以实现高效的操作符选择和可扩展的深度神经架构设计。与 L2 正则化的集成增强了架构的稀疏性，从而在保持或提高性能的同时显著降低了计算成本[6]。ENAS 的性能预测器：为进化 NAS 中的性能预测器引入了一种新颖的训练协议，通过提高架构性能预测的准确性，解决了高计算需求问题，而无需大量计算资源。该方法采用成对排名指标、逻辑回归拟合训练样本以及差异方法构建训练实例，显著提高了 NAS 的效率[7]。基于爬山算法的 NAS 框架：提出了一种使用爬山过程和形态学操作的新框架，通过关注神经网络层的老化来减少总体训练时间。该方法以显著较低的计算成本展示了具有竞争力的结果[8]。注意力卷积网络的高效进化搜索：该框架为卷积网络引入了一种计算高效的进化搜索方法，通过采用样本训练和节点继承来评估后代个体，而无需大量训练。在搜索空间中包含通道注意力机制可以进一步提高特征处理能力[9]。

2.4 克服算力限制

为了应对神经网络架构搜索（NAS）中固有的计算挑战，研究人员一直在追求更高效的方法，强调创新超越了早期的适应方法，如 DARTS（可微分架构搜索）。DARTS+算法通过融入早期停止来改进 DARTS，从而体现了这种进步，早期停止可以缩短优化过程以防止过拟合——

这是破坏模型泛化能力的常见陷阱。这种修改显著增强了架构搜索过程的鲁棒性，避免了在传统 DARTS 实现中经常观察到的有害性能崩溃[10]。

此外，还出现了替代策略来克服早期停止机制的局限性。值得注意的是，带权重共享的随机搜索方法表现出了非凡的有效性，挑战了 NAS 方法论的普遍复杂性。通过利用更简单、更随机的过程，该方法不仅与传统的 NAS 相关联的复杂算法相匹敌，而且在某些情况下还超过了它们，在 PTB 和 CIFAR-10 等基准测试中展示了其卓越的性能。带权重共享的随机搜索的成功强调了资源高效方法在实现具有竞争力甚至更出色的 NAS 结果方面的潜力，从而重塑了架构搜索范式的格局[11]。NAS 方法论的探索扩展到各种适应和细化，旨在克服传统方法最初提出的计算挑战。例如，Fair DARTS 通过允许每个操作的架构权重独立，引入了操作之间的协作而非排他性竞争。这种策略有效地缓解了与跳跃连接相关的性能崩溃问题，这是 DARTS 中的一个常见问题，从而促进了更公平、更平衡的搜索过程，在 CIFAR-10 和 ImageNet 上产生了最先进的结果[12]。另一方面，D-DARTS 通过在单元格级别嵌套神经网络来扩展搜索空间，偏离了传统的权重共享机制。这种创新允许生产多样化和专业化的架构，突出了提高各种计算机视觉任务性能和减少计算时间的潜力[13]。这些进步代表了向更高效、更有效、更公平的 NAS 方法论的重大转变。通过解决固有的计算密集和过拟合挑战，该领域在推动深度学习创新方面更接近实现自动架构搜索的全部潜力。

3. 网络检索机制

3.1 定义检索空间

神经网络架构搜索（NAS）的基础在于准确定义搜索空间，这是算法可能评估的所有潜在神经网络架构的综合目录。这个空间是多维的，包含了各种架构选择，如层的类型（例如，用于图像任务的卷积层，用于序列数据的循环层），这些层之间的连接模式（例如，顺序连接，跳跃连接），超参数（包括网络的深度，卷积层中滤波器的大小，密集层中的单元数），以及其他架构特征，如激活函数和正则化技术。搜索空间的广度和深度至关重要，因为它们直接影响 NAS 算法发现创新和高性能架构的能力。一个精心构建的搜索空间应在全面性和可行性之间取得平衡，确保可以探索各种各样的架构，而不会因计算限制而使搜索变得不切实际。

3.2 选择合适的策略

在 NAS 过程中，导航庞大的搜索空间的策略至关重要。各种方法在搜索效率、计算需求和所得架构的质量方面提供了不同的权衡。例如，强化学习方法使用策略网络按顺序选择架构组件，学习根据过去的性能提出更好的架构。进化算法模拟类似于自然选择的过程，其中架构发生突变和交叉，只有最适合的才能生存到下一代。基于梯度的方法允许通过梯度下降优化架构参数来采取更直接的方法，使它们高效但有时探索性较差。贝叶斯优化利用先验知识智能地探索搜索空间，在已知的良好区域和新区域的探索之间取得平衡。这些策略之间的选择取决于具体目标、可用的计算资源和手头问题的特点。

3.3 表现评估

在 NAS 中，评估候选架构是一个关键步骤。这一步确定了每个提出的架构在给定任务上的表现如何，从而告知搜索算法未来的决策。通常，评估涉及在数据集上训练架构，并使用诸如准

确性、损失或更具体的任务度量等指标来衡量其性能。然而，为每个评估从头开始训练神经网络可能是非常昂贵的。像权重共享这样的技术，其中不同的架构共享权重，以及代理任务，其中模型在任务的简化版本或简化数据上进行训练和评估，可以显著减少计算需求。这一步不仅确定了有前途的架构，还有助于理解不同的架构选择对性能的影响。

3.4 优化过程

通过评估机制的性能反馈，NAS 算法优化了搜索，以寻找显示高性能潜力的架构。这种优化可以根据搜索策略采取各种形式。例如，强化学习算法调整策略网络以增加选择高性能架构的概率。相比之下，进化算法可能会调整种群以朝向具有更高适应度分数的架构。这个优化过程是迭代的，每个周期都旨在细化搜索方向并收敛到最优的架构配置。最终目标是发现不仅在评估指标上表现良好，而且还满足其他标准（如计算资源和模型大小方面的效率）的架构。

3.5 提高和最终策略

在通过搜索和优化过程确定了一组有前途的架构子集后，可能会进行精炼阶段。这个阶段涉及进一步微调架构，可能通过额外的训练轮次、超参数优化或架构调整，以挤出额外的性能提升。然后，根据性能指标的全面评估以及其他考虑因素，如计算效率、模型复杂性和对不同任务或数据集的适应性，最终选择最优架构。这一步确保了所选架构不仅在受控评估设置中表现良好，而且对于实际应用也是实用的。

3.6 验证与测试

所选架构经历了严格的测试，以验证其在不同数据集或任务上的性能、泛化能力以及对输入或条件变化的鲁棒性。这一步对于确保模型的性能不过于依赖于训练集或在 NAS 过程中使用的评估条件至关重要。广泛的测试提供了信心，即所选架构将在实际应用中可靠地执行，实现了 NAS 自动化设计和有效高效神经网络的承诺。

3.7 部署

一旦经过验证，最终架构就可以部署到实际应用中了。这可以涵盖从消费电子产品的图像和语音识别系统到自动驾驶车辆中的复杂决策系统或电子商务中的个性化推荐系统。部署阶段标志着 NAS 过程的顶点，将计算和算法成就转化为各种应用中的实际利益。NAS 的迭代性质，包括其探索、评估、优化和精炼的周期，封装了一种全面的自动化神经网络设计方法。虽然计算量大，但 NAS 有望发现创新的架构，推动人工智能领域的可能性，并在效率和性能上超越传统的手动设计模型。

4. 神经网络检索的主流学习形态

1. 基于强化学习的 NAS（RL-NAS）：RL-NAS 使用策略网络作为决策者，按顺序提出神经网络架构。这个策略网络通过强化学习技术（如 Q 学习或策略梯度）进行训练，学会通过接收其建议的性能（例如，准确性、效率）反馈来导航架构搜索空间。提出架构、在验证集上评估其性能以及根据此反馈更新策略网络的迭代过程允许 RL-NAS 随时间优

化其搜索策略。尽管 RL-NAS 有可能发现高性能架构，但它因高计算需求而闻名，因为每个提出的架构都需要单独的训练和评估，使其资源密集。

2. 基于进化算法的 NAS（EA-NAS）：EA-NAS 将进化计算的原则应用于神经架构搜索。它初始化一组架构，并对它们进行受自然进化启发的选择、交叉和变异操作。在验证集上表现良好的架构更有可能存活下来，并将其遗传特征（即结构特征）传递给后代。这个过程是迭代的，逐渐使种群向高性能架构进化。EA-NAS 提供了一种可并行的搜索策略，可以有效地探索一组多样化的架构。然而，它可能是计算密集型的，尤其是当处理大型搜索空间或复杂架构时。

3. DARTS（Differentiable Architecture Search）是 D-NAS 的一种实现方式，它通过引入可优化的架构参数，使得网络架构和网络权重可以同时进行学习。这种同时学习的方式极大地加速了搜索过程，与传统的非可微方法相比，效率显著提高。D-NAS 因其高效性而受到广泛赞誉。它不仅降低了进行 NAS 所需的计算资源，使得更多的研究者和实践者能够接触和使用 NAS，而且为 NAS 在实际应用中的普及和推广奠定了基础。总的来说，D-NAS 通过引入可微分的架构搜索方法，使得神经网络架构搜索变得更加高效和实用。它的出现不仅推动了 NAS 领域的发展，也为深度学习在实际应用中的进步提供了新的动力。
4. 基于贝叶斯优化的 NAS（Neural Architecture Search）：贝叶斯优化（Bayesian Optimization，BO）是一种针对黑盒函数全局优化的策略，尤其适用于函数评估（在 NAS 的上下文中为架构评估）成本高昂的情况。基于 BO 的 NAS 使用概率模型来预测架构的性能，并应用采集函数来平衡新架构的探索和已知优秀架构的利用。这种方法通过优先评估最有可能带来改进的架构来有效地导航搜索空间，因此适合计算资源有限的场景。
5. 基于图的 NAS（Neural Architecture Search）：利用图论，基于图的 NAS 将架构搜索空间表示为图，其中节点代表架构组件（例如，层、操作），边代表这些组件之间的连接。这种表示允许对具有不同深度和连接模式的复杂架构进行建模。基于图的方法可以使用操纵图结构的算法来有效地探索这个空间，从而提供一种灵活的方法来发现能够捕获复杂数据模式的架构。.
6. 分层 NAS（Hierarchical NAS）：分层 NAS 在多个粒度级别上处理架构搜索，区分宏观架构（网络的整体结构和连接性）和微观架构（网络内各个层或块的设计）。通过优化这两个级别，分层 NAS 能够更详细、更细致地探索架构空间，从而可能导致更优化、更针对任务的设计，同时考虑网络的全局结构和局部操作。
7. 一次性 NAS（One-shot NAS）：一次性 NAS 方法通过构建一个单一的、过度参数化的网络（通常称为超网）来简化搜索过程，该网络包含搜索空间内的所有可能架构。通过训练这个超网，然后评估子网络（候选架构）而无需从头开始重新训练，一次性 NAS 显著降低了 NAS 的计算成本。这种方法受益于子网络之间的权重共享，可以快速评估大量架构。
8. 基于单元的 NAS（Cell-based NAS）将重点放在寻找最优的构建块（单元）上，这些构建块随后被复制以构建整个网络。这种方法简化了搜索空间，并专注于发现通用且可重用的架构模式。由于发现的单元可以适应不同规模和复杂性的问题，这种策略在识别在不同数据集和任务上表现良好的架构方面已被证明是有效的。神经进化（Neuroevolution）是一种将进化算法的原理与神经网络相结合的方法，用于共同进化网络的架构和权重。这种方法将架构和权重视为一个统一的进化过程的一部分，通过变异、重组和基于性能选择网络来探索大量的潜在解决方案。神经进化因其能够发现创新架构和最优权重配置的能力而特别吸引人，为网络设计和优化提供了一种整体方

法。这两种方法各有优势。基于单元的 NAS 通过专注于构建块的优化来简化搜索空间，而神经进化则通过其全局搜索能力来发现创新的架构和权重配置。它们可以相互补充，结合使用可能会进一步提高网络设计的效率和性能。随着 NAS 的不断发展，这些方法之间的融合和创新将为自动化架构搜索带来更大的突破。

这些 NAS 方法各自具有独特的优势和权衡，反映了神经网络设计中存在的各种挑战和目标。随着 NAS 的不断发展，这些方法之间的相互作用可能会产生更加复杂和高效的自动化架构搜索方法。

5. 发展和改进

NAS 近年来取得了显著的进展，解决了其初始的短板，并引入了更先进和高效的方法：

1. **硬件感知的 NAS**：像硬件感知框架这样的最新方法不仅关注模型的准确性，还关注硬件效率。这些方法使用进化算法与目标预测器相结合，以高效地为各种性能指标和硬件配置找到优化的架构[14]。
2. **NAS-Bench-101 用于可重复性**：NAS 研究的首个公共架构数据集 NAS-Bench-101 的引入，旨在解决 NAS 的高计算需求并使实验更具可重复性。它编译了一个包含超过 500 万个训练模型的大型数据集，可以迅速评估各种架构[15]。
3. **嵌入式设备的优化**：像 Efficient Neural Architecture Search (ENAS)这样的方法的改编考虑了在嵌入式设备上部署网络的约束，展示了 NAS 在资源受限环境中的灵活性和适用性[16]。
4. **Pareto 最优方法**：像 MONAS 和 DPP-Net 这样的多目标框架将 NAS 扩展到优化精度和其他由设备施加的目标，搜索可以在广泛设备上部署的神经架构[17]。
5. **图像去噪的高效 NAS**：NAS 的发展也应用于图像去噪，像超核实现这样的技术使模型能够快速训练进行密集预测，展示了 NAS 在不同应用领域的多功能性[18]。
6. **图像分类的进化设计**：进化算法已用于设计用于图像分类的深度卷积神经网络，解决了分类性能和计算效率等多个目标[19]。

这些进展展示了 NAS 的动态性质，它不断发展和适应以满足新的挑战和需求。随着技术的进步和算法的优化，NAS 有望在未来继续推动神经网络设计和性能的新边界。

6. 神经网络检索的应用

在计算机视觉及其以外的领域[20-22]，神经架构搜索（NAS）已成为一股变革性的力量，推动了图像分类、目标检测、语义分割和自然语言处理（NLP）等领域的进步。像 DQNAS 这样的框架的集成，将强化学习与一次性训练相结合，展示了 NAS 在各类图像相关应用中超越手工设计模型的能力。这种方法不仅简化了为语言建模和翻译等复杂任务创建高效模型的过程，而且还将其用途扩展到优化用于多分辨率特征提取的卷积网络[23-26]。NAS 的适应性进一步体现在其应用于人类活动识别（HAR）中，它使用贝叶斯优化等技术来优化神经架构，以分析与人类移动性相关的活动。这展示了 NAS 在专门领域中的有效性，并强调了其在提高针对特定任务和数据类型定制模型的准确性和效率方面的作用[27]。

从扩大其影响力角度来说，NAS 在智能城市应用中的时空预测任务中扮演着关键角色，例如 AutoST 等用于人群流量预测的方法，突显了 NAS 在城市规划和智能交通方面的贡献。NAS 在不同领域中的广泛应用——从提高机器对视觉信息的解释能力，到改善与人类语言的交互，以及通过技术推进我们对人类运动的理解——强调了它在自动化神经网络架构设计中的核心作用。随着 NAS 的不断发展[29-30]，它有望进一步推动神经网络架构的创新和优化，巩固其在

人工智能和机器学习各个领域的影响。正在进行的进步突显了 NAS 革命化研究和应用各个方面的潜力，使其成为未来技术发展中的基石。

7. 参考文献

[1] Klaus Greff et al. "LSTM: A Search Space Odyssey." *IEEE Transactions on Neural Networks and Learning Systems*, 28 (2015): 2222-2232. https://doi.org/10.1109/tnnls.2016.2582924.

[2] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2014). Going deeper with convolutions. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1-9. https://doi.org/10.1109/CVPR.2015.7298594.

[3] Lee, J., Jun, S., Cho, Y., Lee, H., Kim, G., Seo, J., & Kim, N. (2017). Deep Learning in Medical Imaging: General Overview. *Korean Journal of Radiology*, 18, 570 - 584. https://doi.org/10.3348/kjr.2017.18.4.570.

[4] Klyuchnikov, N., Trofimov, I., Artemova, E., Salnikov, M., Fedorov, M., & Burnaev, E. (2020). NAS-Bench-NLP: Neural Architecture Search Benchmark for Natural Language Processing. *IEEE Access*, PP, 1-1. https://doi.org/10.1109/access.2022.3169897.

[5] Lu, Z., Whalen, I., Dhebar, Y., Deb, K., Goodman, E., Banzhaf, W., & Boddeti, V. (2019). Multiobjective Evolutionary Design of Deep Convolutional Neural Networks for Image Classification. *IEEE Transactions on Evolutionary Computation*, 25, 277-291. https://doi.org/10.1109/TEVC.2020.3024708.

[6] Zhang, T., Lei, C., Zhang, Z., Meng, X., & Chen, C. (2021). AS-NAS: Adaptive Scalable Neural Architecture Search With Reinforced Evolutionary Algorithm for Deep Learning. *IEEE Transactions on Evolutionary Computation*, 25, 830-841. https://doi.org/10.1109/TEVC.2021.3061466.

[7] Sun, Y., Sun, X., Fang, Y., Yen, G., & Liu, Y. (2020). A Novel Training Protocol for Performance Predictors of Evolutionary Neural Architecture Search Algorithms. *IEEE Transactions on Evolutionary Computation*, 25, 524-536. https://doi.org/10.1109/TEVC.2021.3055076..

[8] Verma, M., Sinha, P., Goyal, K., Verma, A., & Susan, S. (2019). A Novel Framework for Neural Architecture Search in the Hill Climbing Domain. *2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)*, 1-8. https://doi.org/10.1109/AIKE.2019.00009.

[9] Zhang, H., Jin, Y., Cheng, R., & Hao, K. (2020). Efficient Evolutionary Search of Attention Convolutional Networks via Sampled Training and Node Inheritance. *IEEE Transactions on Evolutionary Computation*, 25, 371-385. https://doi.org/10.1109/TEVC.2020.3040272.

[10] Liang, H., Zhang, S., Sun, J., He, X., Huang, W., Zhuang, K., & Li, Z. (2019). DARTS+: Improved Differentiable Architecture Search with Early Stopping. *ArXiv*, abs/1909.06035.

[11] Li, L., & Talwalkar, A. (2019). Random Search and Reproducibility for Neural Architecture Search. *ArXiv*, abs/1902.07638.

[12] Chu, X., Zhou, T., Zhang, B., & Li, J. (2019). Fair DARTS: Eliminating Unfair Advantages in Differentiable Architecture Search. *ArXiv*, abs/1911.12126. https://doi.org/10.1007/978-3-030-58555-6_28.

[13] Heuillet, A., Tabia, H., Arioui, H., & Youcef-Toumi, K. (2021). D-DARTS: Distributed Differentiable Architecture Search. *ArXiv*, abs/2108.09306.

[14] Cummings, D., Sarah, A., Sridhar, S., Szankin, M., Muñoz, J., & Sundaresan, S. (2022). A Hardware-Aware Framework for Accelerating Neural Architecture Search Across Modalities. *ArXiv*, abs/2205.10358.https://doi.org/10.48550/arXiv.2205.10358.

[15] Ying, C., Klein, A., Real, E., Christiansen, E., Murphy, K., & Hutter, F. (2019). NAS-Bench-101: Towards Reproducible Neural Architecture Search. *ArXiv*, abs/1902.09635.

[16] Cassimon, T., Vanneste, S., Bosmans, S., Mercelis, S., & Hellinckx, P. (2019). Using Neural Architecture Search to Optimize Neural Networks for Embedded Devices. , 684-693. https://doi.org/10.1007/978-3-030-33509-0_64.
[17] Cheng, A., Dong, J., Hsu, C., Chang, S., Sun, M., Chang, S., Pan, J., Chen, Y., Wei, W., & Juan, D. (2018). Searching Toward Pareto-Optimal Device-Aware Neural Architectures. *2018 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 1-7. https://doi.org/10.1145/3240765.3243494.
[18] Mo.zejko, M., Latkowski, T., Treszczotko, L., Szafraniuk, M., & Trojanowski, K. (2020). Superkernel Neural Architecture Search for Image Denoising. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2002-2011. https://doi.org/10.1109/cvprw50498.2020.00250.
[19] Lu, Z., Whalen, I., Dhebar, Y., Deb, K., Goodman, E., Banzhaf, W., & Boddeti, V. (2019). Multiobjective Evolutionary Design of Deep Convolutional Neural Networks for Image Classification. *IEEE Transactions on Evolutionary Computation*, 25, 277-291. https://doi.org/10.1109/TEVC.2020.3024708.
[20] Meng, F., & Wang, Y. (2023). Transformers: Statistical interpretation, architectures and applications. *Authorea Preprints*.
[21] Fanfei Meng, Branden Ghena. (2023) Research on Text Recognition Methods Based on Artificial In-telligence and Machine Learning. Advances in Computer and Communication, 4(5), 340-344.
[22] Meng, F., & Demeter, D. (2023). Sentiment analysis with adaptive multi-head attention in Transformer. *arXiv preprint arXiv:2310.14505*.
[23] Manijeh Razeghi, Arash Dehzangi, Donghai Wu, Ryan McClintock, Yiyun Zhang, Quentin Durlin, Jiakai Li, and Fanfei Meng. Antimonite-based gap-engineered type-ii superlattice materials grown by mbe and mocvd for the third generation of infrared imagers. In Infrared Technology and Applications XLV, volume 11002, pages 108–125. SPIE, 2019.
[24] Meng, F., Zhang, L., Chen, Y., & Wang, Y. (2023). FedEmb: A Vertical and Hybrid Federated Learning Algorithm using Network And Feature Embedding Aggregation. *Authorea Preprints*..
[25] Meng, F., Zhang, L., Chen, Y., & Wang, Y. (2023). Sample-based Dynamic Hierarchical Transformer with Layer and Head Flexibility via Contextual Bandit. *Authorea Preprints*.
[26] Meng, F., & Wang, C. A. (2023). A Dynamic Interactive Learning Interface for Computer Science Education: Programming Decomposition Tool. *Authorea Preprints*.
[27] Chang Ling, Chonglei Zhang, Mingqun Wang, Fanfei Meng, Luping Du, and Xiaocong Yuan, "Fast structured illumination microscopy via deep learning," Photon. Res. 8, 1350-1359 (2020)
[28] Meng, F., Jagadeesan, L., & Thottan, M. (2021). Model-based reinforcement learning for service mesh fault resiliency in a web application-level. *arXiv preprint arXiv:2110.13621*.
[29] Wang, Y., Meng, F., Wang, X., & Xie, C. (2023). Optimizing the Passenger Flow for Airport Security Check. *arXiv preprint arXiv:2312.05259*.
[30] Chen, Jin-Jin, et al. "A dataset of diversity and distribution of rodents and shrews in China." *Scientific Data* 9.1 (2022): 304
[30] Meng, F., Zhang, L., Wang, Y., & Zhao, Y. (2023). Joint detection algorithm for multiple cognitive users in spectrum sensing. *Authorea Preprints*.
[31] Fanfei Meng, Branden Ghena. (2023) Research on Text Recognition Methods Based on Artificial In-telligence and Machine Learning. Advances in Computer and Communication, 4(5), 340-344.
[32] Meng, F., & Demeter, D. (2023). Sentiment analysis with adaptive multi-head attention in Transformer. *arXiv preprint arXiv:2310.14505*.
[33] Manijeh Razeghi, Arash Dehzangi, Donghai Wu, Ryan McClintock, Yiyun Zhang, Quentin Durlin, Jiakai Li, and Fanfei Meng. Antimonite-based gap-engineered type-ii superlattice materials grown by mbe and mocvd for the third generation of infrared imagers. In Infrared Technology and Applications XLV, volume 11002, pages 108–125. SPIE, 2019.
[34] Meng, F., Zhang, L., Chen, Y., & Wang, Y. (2023). FedEmb: A Vertical and Hybrid Federated Learning Algorithm using Network And Feature Embedding Aggregation. *Authorea Preprints*..

[35] Meng, F., Zhang, L., Chen, Y., & Wang, Y. (2023). Sample-based Dynamic Hierarchical Transformer with Layer and Head Flexibility via Contextual Bandit. *Authorea Preprints*.
[36] Meng, F., & Wang, C. A. (2023). A Dynamic Interactive Learning Interface for Computer Science Education: Programming Decomposition Tool. *Authorea Preprints*.
[37] Chang Ling, Chonglei Zhang, Mingqun Wang, Fanfei Meng, Luping Du, and Xiaocong Yuan, "Fast structured illumination microscopy via deep learning," Photon. Res. 8, 1350-1359 (2020)
[38] Meng, F., Jagadeesan, L., & Thottan, M. (2021). Model-based reinforcement learning for service mesh fault resiliency in a web application-level. *arXiv preprint arXiv:2110.13621*.
[39] Wang, Y., Meng, F., Wang, X., & Xie, C. (2023). Optimizing the Passenger Flow for Airport Security Check. *arXiv preprint arXiv:2312.05259*.
[40] Chen, Jin-Jin, et al. "A dataset of diversity and distribution of rodents and shrews in China." *Scientific Data* 9.1 (2022): 304
[41] Meng, F., Zhang, L., Wang, Y., & Zhao, Y. (2023). Joint detection algorithm for multiple cognitive users in spectrum sensing. *Authorea Preprints*.
[42]Meng, F., & Wang, Y. (2023). Transformers: Statistical interpretation, architectures and applications. *Authorea Preprints*.
[43]Fanfei Meng, Chen-Ao Wang, Alexander Brown. Evolution and Efficiency in Neural Architecture Search: Bridging the Gap Between Expert Design and Automated Optimization. *TechRxiv.* February 14, 2024.