未命名稿件编号。（由编辑填写）

神经网络国际验证竞赛（VNN-COMP）前三年

Christopher Brix · Mark Niklas M¨uller · Stanley Bak · Taylor T. Johnson · ChangliuLiu

arXiv:2301.05815v1 [cs.LG] 14 Jan 2023

收稿日期：日期 / 接受日期：日期

摘要 本文总结并元分析了2020年、2021和2022年举办的年度神经网络国际验证竞赛（VNN-COMP）前三届赛事。在VNN-COMP中，参与者提交软件工具，用于分析给定神经网络是否满足描述其输入输出行为的规范。这些神经网络和规范涵盖多种问题类别和任务，对应于图像分类、神经控制、强化学习和自主系统中的安全性与鲁棒性属性。我们总结了关键流程、规则和结果，介绍了过去三年观察到的趋势，并展望了未来可能的发展。

关键词 认证鲁棒性、对抗鲁棒性、形式化验证、形式化方法、神经网络、机器学习、深度学习

Christopher Brix 亚琛工业大学，德国亚琛 电子邮箱：brix@cs.rwth-aachen.de

Mark M¨uller 苏黎世联邦理工学院，瑞士苏黎世 电子邮箱：mark.mueller@inf.ethz.ch

Stanley Bak 石溪大学，美国纽约州石溪 电子邮箱：stanley.bak@stonybrook.edu

Taylor T. Johnson 范德堡大学，美国田纳西州纳什维尔 电子邮箱：taylor.johnson@vanderbilt.edu

Changliu Liu 卡内基梅隆大学，美国宾夕法尼亚州匹兹堡 电子邮箱：cliu6@andrew.cmu.edu

1 引言

神经网络越来越多地应用于安全关键应用领域 [7,23]。然而，人们逐渐发现，神经网络极易受到对抗样本的影响[47], ，也就是说，微小且可能难以察觉的输入扰动就可能导致输出发生显著变化。由于这类扰动可能在现实世界中随机出现，也可能由恶意行为者造成，因此，在将基于深度学习的系统应用于安全关键领域之前，以数学上严谨的方式分析其鲁棒性至关重要。为此，人们开发了各种各样的方法及相应的软件工具 [12,16,22,24]。然而，随着工具的数量不断增加、专业化程度不断提高，从业者越来越难以决定该使用哪种工具。

2020年，首届 VNN-COMP 应运而生，旨在解决这一问题，让研究人员能够在一系列基准测试中比较各自的神经网络验证器。该竞赛最初被设想为一场标准化程度不高的友好竞赛，之后逐步实现了更高程度的标准化和自动化，以确保在成本相当的硬件上，使用标准化的性质和网络格式进行公平比较。

本文概述了这一发展历程，总结了主要规则和结果，介绍了过去三年观察到的总体趋势，并展望了未来可能的发展方向。

2 神经网络验证

我们考虑如下定义的神经网络验证问题：给定输入规约 $\phi \subseteq \mathrm{R}^{d_{\text{in}}}$（也称为前置条件），以及输出规约

$\psi \subseteq \mathbb{R}^{d_{\text{out}}}$（也称为后置条件）和一个神经网络 $N : \mathbb{R}^{d_{\text{in}}} \mapsto \mathbb{R}^{d_{\text{out}}}$，我们的目标是证明前置条件蕴含后置条件，即

$$\forall x : x \vDash \phi \Rightarrow N(x) \vDash \psi, \tag{2.1}$$

或提供一个反例。

受 SAT 求解器社区常用符号的启发，我们通过指定一个描述对抗样本的约束集合来编码这个问题，即

$$\exists x : x \vDash \phi \wedge N(x) \vDash \neg\psi. \tag{2.2}$$

因此，我们将满足方程 (2.2)，从而方程 (2.1) 所编码的性质不 成立的实例称为可满足，将方程 (2.2) 不可满足，且方程 (2.1) 所编码的性质已被证明成立的实例称为不可满足。请注意，虽然可以通过使用对抗攻击直接搜索反例来证明可满足性 [17,32], ，但这些方法并不完备；也就是说，如果它们未能找到反例，并不 意味着某个性质成立。

示例问题 一种尤为常见的性质是图像分类中对范数受限扰动的鲁棒性。对于每个输入 $x \in \mathbb{R}^{d_{\text{in}}}$，网络$N$ 都会计算一个数值分数 $y \in \mathbb{R}^{d_{\text{out}}}$，表示输入属于各个 d_{out} 类别的置信度。最终的分类结果 c则计算为 $c = \arg\max_i N(x)_i$。在这种情况下，攻击者可能希望扰动输入，使分类结果发生变化。因此，验证旨在证明

$$\arg\max_i N(x')_i = t,$$
$$\forall x' \in \{x' \in \mathbb{R}^{d_{\text{in}}} \mid \|x - x'\|_\infty \leq \epsilon\},$$

其中， t 是目标类别， x 是原始图像，ϵ 是允许的最大扰动幅度。在这里，前置条件 ϕ 描述攻击者可以选择的输入（$\phi = \{x' \in \mathbb{R}^{d_{\text{in}}} \mid \|x - x'\|_\infty \leq \epsilon\}$），即半径为 ϵ的 ℓ_∞球；后置条件ψ 描述对应于分类为目标类别 t 的输出空间（$\psi = \{y \in \mathbb{R}^{d_{\text{out}}} \mid y_t > y_i, \ \forall i \neq t\}$）。

当神经网络用作控制器时，可能涉及更复杂的性质。例如，在 ACAS Xu 场景中 [23]，神经控制器会根据受控飞机和入侵飞机的相对位置与航向给出动作建议。在这种情况下，我们希望确保，例如，对于 $\mathcal{D}$ 对应入侵飞机正前方且正朝我们飞来的输入，规避机动“强左转”（SL）或“强右转”（SR）都不会被视为最差选项。更正式地说，我们希望验证

$$\arg\min_i N(x')_i \notin \{\text{SL}, \text{SR}\}, \ \forall x' \in \mathcal{D}.$$

在这里，我们得到一个更复杂的非凸后置条件

$$\begin{aligned}\psi = &\mathbb{R}^{d_{\text{out}}} \setminus \\ &\big(\{y \in \mathbb{R}^{d_{\text{out}}} \mid y_{\text{SL}} < y_i, \ \forall i \notin \{\text{SL}, \text{SR}\}\} \\ &\cup \{y \in \mathbb{R}^{d_{\text{out}}} \mid y_{\text{SR}} < y_i, \ \forall i \notin \{\text{SL}, \text{SR}\}\}\big).\end{aligned}$$

3 竞赛目标

VNN-COMP 旨在进一步实现以下目标。

制定标准 为了让从业者能够轻松使用和评估多种不同的验证方法和工具，而无需承担大量额外工作，所有工具都能够处理采用标准化文件格式的网络和规范，这一点至关重要。为此，VNN-COMP 的第二届竞赛确立了这样一种标准。问题规范（前置条件和后置条件）采用 VNN-LIB [48] 格式定义，神经网络则采用 ONNX [3]标准定义。2022年还引入了反例的标准化格式。

促进验证工具比较每年都会发表数十篇关于神经网络验证的论文，其中许多不仅提出了新方法，还提出了新的基准测试。作者可能会投入更多时间，针对选定的基准测试调整自己的方法，因此要公平比较所有这些方法并不容易。VNN-COMP 使用成本相当的硬件，并采用参与者无法获取的测试实例，促进了在多样化基准测试集上对大量工具进行比较。让参与者和行业从业者提出各种有趣的基准测试，不仅能对该领域通常使用的问题进行排名，还能凸显哪些工具特别适用于更专业的问题。此外，通过确保安装和评估流程标准化，任何发表成果都可以与大量最先进的工具进行比较。

塑造未来研究方向VNN-COMP 对所考虑基准测试背后问题的关注，有望提升这些问题在社区中的影响力。由于基准测试由工业界和学术界共同制定，这为塑造未来研究、使其尽可能产生深远影响提供了绝佳机会。

近年来，基准测试中的网络规模不断扩大（见表 5.2），推动了可扩展性的发展；网络也变得更加复杂（例如，包含残差 [19] 和最大池化层[64]），从而促进了泛化能力；规范也更加复杂，使得能够分析更有趣的性质。

汇聚研究人员竞赛筹备期间的规则和基准测试讨论阶段，以及在机器学习赋能自主系统形式化方法研讨会（FoMLAS）上现场展示结果[1] ，都为参与者提供了与同行交流、探讨该领域未来发展的绝佳机会。此外，参与者为年度报告 [2,5,34] 撰写的工具和基准测试说明，是对最先进方法的精彩总结，能让刚进入该领域的人快速了解全貌。

4 VNN-COMP 三年概述

本节概述 VNN-COMP 从 2020年到 2022年 的发展历程，并在表5.1中列出所有参与者及最终排名。一般而言，性能是在一组权重相同的基准测试上衡量的，每个基准测试都由一组相关实例组成。每个实例包含一个训练好的神经网络、一个超时时间，以及输入和输出约束。下文将基准测试按类别分组，以便更快地比较各年份的情况。

4.1 VNN-COMP 2020年

首届 VNN-COMP[2] [2] 于 2020年 举办，形式是一次没有优胜者的“友好竞赛”。其主要目标是为后续赛事奠定基础，启动界定通用问题设置并确定可能的标准化途径这一进程。

4.1.1 基准测试

本次竞赛考虑了三类基准测试，但八支参赛队伍中只有一支参与了所有类别：

– 采用 ReLU 激活的全连接网络——两个基准测试，分别基于 ACAS Xu 和 MNIST。–采用 Sigmoid 和 Tanh 激活函数的全连接网络——一个基准测试，基于 MNIST。– 卷积网络——两个基准测试，分别基于 MNIST 和 CIFAR10。

4.1.2 评估

各参赛队伍使用自己的硬件评估工具。虽然这简化了评估流程，但由于硬件差异显著，所报告的结果无法进行比较。各队报告称，他们使用的 CPU 数量在 4 到 40 个之间，内存容量在 16 到 756 GB 之间。

4.2 VNN-COMP 2021

基于从2020年中获得的经验，VNN-COMP[3] 第二届赛事的组织更加注重参与工具[5]之间的可比性。

4.2.1 基准测试

各团队可以提出一个基准测试，其总超时时间最多为六小时，并分配给该基准测试包含的各个实例。网络采用 ONNX 格式 [3] 定义，问题规范采用 VNN-LIB 格式 [48]提供。为避免针对特定基准测试实例进行过度调优，鼓励基准测试提议者提供脚本，以便生成新的随机实例，用于最终的工具评估。不过，各团队可以使用基准测试实例的初始集合，针对每项基准测试调整其工具。

2021年，这些基准测试可分为以下类别，多个团队参与了所有类别：

– ReLU 激活的全连接网络——两个基准测试，分别基于 ACAS Xu 和 MNIST。– Sigmoid 激活的全连接网络——一个基准测试，基于 MNIST。– 卷积网络——三个基准测试，基于 CIFAR10。– 带最大池化层的网络——一个基准测试，基于 MNIST。–残差网络——一个基准测试，基于 CIFAR10。– 使用稀疏矩阵的大型网络——一个基准测试，基于数据库索引。

[1] https://fomlas2022.wixsite.com/fomlas2022
[2] https://sites.google.com/view/vnn20/vnncomp
[3] https://sites.google.com/view/vnn2021

表 4.1：可用的亚马逊云科技实例。

	2021	2022	虚拟 CPU	内存 [GB]	GPU
r5.12xlarge	✓	✗	48	384	✗
p3.2xlarge	✓	✓	8	61	16 GB 内存的 V100 图形处理器
m5.16xlarge	✗	✓	64	256	✗
g5.8xlarge	✗	✓	32	128	24 GB 内存的 A10G 图形处理器
t2.large	✗	✓	2	8	✗

4.2.2 评估

为确保结果具有可比性，所有工具均使用亚马逊云科技（AWS）提供的同等成本硬件进行评估。各团队可自行决定使用侧重 CPU 的 r5.12xlarge 实例，还是侧重图形处理器的 p3.2xlarge 实例来评估其工具（更多详情请参见表 4.1）。此外，团队不能自行提供结果和运行时间，而是必须准备好脚本，以自动完成工具的安装和运行。提交截止日期之后，组织者使用所提供的脚本安装并评估了每个工具。在许多情况下，这一过程需要组织者与团队反复沟通，进行一些调试。

表 4.2：2021 年每个实例的得分。根据简单对抗攻击是否成功，将可满足实例分为简单和复杂两类。

	返回结果		
真实值	SAT	UNSAT	其他
SAT，简单	+1	−100	0
SAT，复杂	+10	−100	0
UNSAT	−100	+10	0

评分 对于每个基准测试，正确判断实例为可满足/不可满足可得 10 分，结果错误则扣 100 分（见表 4.2）。我们使用一种简单的对抗攻击来识别“容易”的可满足实例，这类实例的可得分数从 10 分降至 1 分。如果工具对某个实例给出了相互矛盾的结果，则通过多数表决确定真实值。每个实例中运行最快的两个工具可获得奖励分（最快者得 2 分，第二快者得 1 分）。运行时间相差不到 0.2 秒或均低于 1 秒时，视为相同，因此多个团队可能同时获得 2 分奖励。为消除启动开销方面显著的差异（例如，获取图形处理器所需的开销），我们将其作为在简单实例上的运行时间进行测量，并从每次运行时间中扣除。基准测试得分根据上述获得的分数计算，并以可获得的最高分数进行归一化。因此，得分最高的工具被赋予 100% 的得分。竞赛总分则为各基准测试得分之和，相当于对每个基准测试赋予相同权重。

结果 2021 年，共有 12 支队伍参加了比赛。 α-β-CROWN 获得第一名，VeriNet 紧随其后获得第二名，ERAN/OVAL 获得第三名；具体排名取决于所使用的开销测量和投票机制，以确定结果是否正确。除 VeriNet 外，其他队伍均使用了图形处理器实例。

4.3 VNN-COMP 2022

在最近一届 VNN-COMP[4] [34], 中，评估已实现全自动化，因此能够增加基准测试的数量。

4.3.1 基准测试

2022年，每个参赛团队最多可提交或认可两项基准测试，因此工业界从业者无需参与工具竞赛，也能提出基准测试。今年，每项基准测试的超时时间为三至六小时，并且必须对实例进行随机化。仍然允许针对每项基准测试进行工具调优；实际上，也可以利用网络的统计数据，针对每个网络进行调优。

提交的基准测试可分为以下类别：

– 采用 ReLU 激活的全连接网络——三个基准测试，基于强化学习任务和 MNIST。– TLL 格式的全连接网络 [13] ——一个基准测试。– 使用稀疏矩阵的大型网络——一个基准测试，基于数据库索引和基数估计。

[4] https://sites.google.com/view/vnn2022

– 卷积网络——三个基准测试，基于 CIFAR10。–残差网络——两个基准测试，基于 CIFAR10、CIFAR100 和 TinyImageNet。– 采用平均池化和 softmax 的复杂 U-Net 网络——一个基于图像分割的基准测试。

4.3.2 评估

与上一年类似，参赛团队可以在多种亚马逊云科技实例类型之间进行选择（见表 4.1），这些实例分别侧重于 CPU、图形处理器或混合配置。除性能弱得多的 t2.large 实例外，所有实例的价格均约为每小时 3 美元。与 2021 年不同，当时组织者必须手动执行安装脚本并与参与者一起调试；本届赛事建立了自动化的提交和测试流程。团队可以通过网页界面提交其基准测试和工具，具体包括指定 Git 仓库、提交哈希值和安装后脚本（例如，用于获取许可证）。提交后，系统会启动一个新的亚马逊云科技实例，并在其中执行所有安装脚本。如果安装成功，系统会在预先选定的一组基准测试上自动评估该工具，随后终止该实例。为便于参与者进行调试，所有输出都会被记录，并通过提交网站实时提供，从而使他们能够监控进度。这种自动化使每个团队能够根据需要执行任意多次测试，而无需等待组织者的反馈。此外，团队还可以在最终评估所使用的同类亚马逊云科技实例上进行测试，且无需支付使用费用；相关成本由苏黎世联邦理工学院的 SRI 实验室慷慨承担。

评分 与 VNN-COMP 2021 期间不同，出于评分目的，可满足实例不再分为简单和复杂，因此所有正确结果均获得 10 分（见表 4.3）。此外，当工具之间存在异议时，不再依靠投票机制来确定真实值，而是将举证责任交给报告可满足结果的工具，要求其提供具体的反例。如果未提供有效反例，则判定相应工具的结果不正确，并处以 100 分的惩罚。

结果 在 11 支参赛队伍中， α-β- CROWN 名列第一，MN-BAB位居第二，VeriNet 位居第三。关于所有参与工具在所有基准测试中的比较，请参见图 4.1。

表 4.3：2022年每个实例的得分。

	返回结果		
真实值	SAT	UNSAT	其他
SAT	+10	−100	0
UNSAT	−100	+10	0

5 各年份比较

表5.1列出了参加过 VNN-COMP 任意一届比赛的所有工具；有兴趣的读者可以参阅相应的 VNN-COMP 报告，了解这些工具的简要介绍。表 5.2 对比了过去三年竞赛的范围。可以看出，基准测试的数量、种类、复杂度和规模在每届比赛中都有所增加。2020年，比赛从 5 个基准测试起步，涵盖简单的全连接（FC）网络和卷积（Conv）网络；到了 2022年，比赛的基准测试增至 12 个，其中包括多种复杂的残差和 U-Net 架构，参数量最高达 1.4 亿。此外，我们认为，注册工具数量不断增加，清楚地表明人们对该领域整体以及竞赛本身的兴趣逐年增长。然而，注册工具与提交工具之间的差距日益扩大，这或许表明许多团队认为，支持标准化的网络与规范格式以及种类繁多的基准测试需要投入大量精力，而他们无力承担。工具按总得分排名，每个基准测试最高可得 100% 的分数，因此最终排名偏向于支持所有基准测试的工具。我们认为，这能够有效激励工具开发者开发可轻松应用于新问题的方法，但要让新团队实现所有必要功能，可能会令人生畏，从而使他们望而却步，不愿参赛。

成功趋势 2020 年，所有团队起初都只使用 CPU；到了 2021 年，排名前四的团队中仅有一支完全依赖 CPU；而 2022 年，排名前三的团队都选择了 GPU 实例。这一转变既提高了对 DeepPoly [45], CROWN [62], 和 IBP [18] 等简单界传播方法的评估效率，也使验证过程中产生的线性规划（LP）问题得以近似求解 [14,55,58]。类似地，2021 年排名前两位的团队以及 2022 年排名前三的团队都采用了基于分支定界（BaB）的方法，递归地将验证问题拆解为更简单的子问题，直至问题可解，从而有效利用 GPU 求解更紧的混合

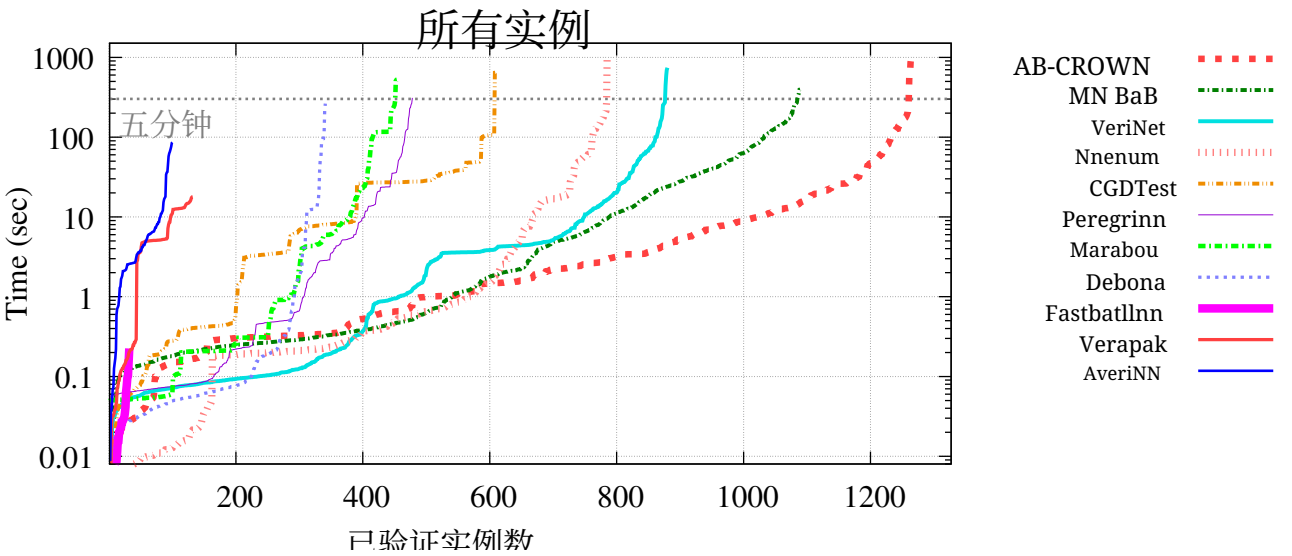


图 4.1：所有工具的仙人掌图

参加 VNN-COMP 2022 的所有基准测试。

整数线性规划（MILP）验证问题编码 [11,14,55,61]。最近一届排名前两位的团队都将这种方法与额外的多神经元约束 [14]以及求解器生成的割平面约束 [61]，相结合；这些约束最初由排名第三的 ERAN 在 2021 [36] 中引入。因此，我们得出结论：成功的工具会利用 GPU 等硬件加速器，高效处理验证问题的紧（混合整数）线性规划编码。

。将往届的所有基准测试都纳入计分基准测试的选择范围，会给参与者带来过重的负担；而每年挑选一个特别具有挑战性、代表性且有趣的基准测试，作为计分的额外基准测试纳入后续届次，或许是一个不错的折中方案。此外，对工具调优采取更严格的限制，或许能更具代表性地评估新工具在旧基准测试上的表现。

6 展望

下文将讨论一些考量因素，它们或可使 VNN-COMP 的未来迭代更好地实现其目标并服务相关社区（见第 3 节）。

6.1 跟踪逐年进展

虽然我们认为，VNN-COMP 已经提供了合理的机制来比较每一届提交的工具，但基准测试和工具不断变化，使得追踪整个领域逐年进展变得困难。由于有些工具针对当年竞赛的特定基准测试进行了大量优化，即使这些工具支持往届（或未来年份）的基准测试，直接在这些基准测试上评估它们也无法得出有意义的进展指标。首届竞赛中的一个基准测试曾作为不计分的额外基准测试纳入随后两届竞赛（cifar2020），但如今仍未解决的实例寥寥无几，因此它作为衡量进一步改进的指标非常不敏感

6.2 工具调优

许多最成功的工具并非只采用单一验证策略，而是提供一整套不同的模式，每种模式都有不同的超参数。工具性能会因选择而显著变化，因此，实践者在将这些工具用于新问题时，必须做出正确选择。然而，参数数量庞大且相互作用复杂，要做到这一点可能颇具挑战，尤其是在缺乏对工具深入了解的情况下。

对于 VNN-COMP，明确允许按基准测试使用调优工具，也默认允许按网络使用调优工具，从而让各团队能够展示其工具的最高性能。不过，在未来几届赛事中，可以考虑限制部分或全部基准测试的调优，以鼓励作者开发自动调优策略，让他们的工具更容易适应新问题。例如，

表5.1：参与工具。

Tool	机构	参与情况、地点			参考文献
		2020 [2]	2021 [5]	2022 [34]	
α-β-CROWN	卡内基梅隆大学 卡内基梅隆大学, 东北大学, 哥伦比亚大学，UCLA	✗	✓(1/12)	✓(1/11)	[63, 57, 58]
AveriNN	堪萨斯州立大学	✗	✗	✓(11/11)	N/A
CGDTest	滑铁卢大学	✗	✗	✓(5/11)	N/A
Debona	亚琛工业大学	✗	✓(6/12)	✓(8/11)	[9]
DNNF	弗吉尼亚大学	✗	✓(12/12)	✗	[42]
ERAN	苏黎世联邦理工学院、伊利诺伊大学厄巴纳-香槟分校	✓	✓(3/12)	✗	[46, 43, 45, 44, 36, 39]
FastBatLLNN	加利福尼亚大学	✗	✗	✓(9/11)	N/A
Marabou	希伯来 大学 耶路撒冷, 斯坦福大学，亚马逊云科技服务 ，NRI Secure	✗	✓(5/12)	✓(7/11)	[25]
MIPVerify	麻省理工学院-ogy	✓	✗	✗	[49]
MN-BAB	苏黎世联邦理工学院	✗	✗	✓(2/11)	[14]
Nnenum	石溪大学	✓	✓(8/12)	✓(4/11)	[4]
NNV	范德堡大学	✓	✓(9/12)	✗	[50, 53, 52, 51, 56]
NV.jl	卡内基梅隆大学、东北大学	✗	✓(10/12)	✗	[28]
Oval	牛津大学	✓	✓(3/12)	✗	[10, 11, 31]
PeregriNN	加利福尼亚大学	✓	✗	✓(6/11)	[26]
RPM	斯坦福大学	✗	✓(11/12)	✗	[54]
金星	伦敦帝国理工学院	✓	✓(7/12)	✗	[8, 27]
VeraPak	犹他州立大学	✗	✗	✓(10/11)	N/A
VeriNet	伦敦帝国理工学院	✓	✓(2/12)	✓(3/11)	[20, 21]

表 5.2：历年对比。

	2020	2021	2022
注册工具	N/A	15	18
提交工具	8	13	11
提交的基准测试	5	8（+1未评分）	12（+1未评分）
最大网络深度	8	18	27
最大网络参数	855,600	42,059,431（稀疏）	138,356,520
激活函数	ReLU、Tanh、Sigmoid	ReLU, Sigmoid, Max-池化、平均池化	ReLU、Sigmoid、最大池化
层类型	全连接、卷积	全连接, 卷积, 残差	全连接, 卷积, 残差 ，批归一化
应用	图像识别，Con-trol	图像识别，控-制，数据库索引	图像识别，控-制，数据库索引，基数估计
平均基准测试数/工具	3.0（最小值 2，最大值 5）	5.5（最小值 1，最大值 9）	7.3（最小值 1，最大值 13）

可以不仅生成随机规范，还生成随机网络来实现这一点。

6.3 批处理

每个 VNN-COMP 基准测试都由一组实例构成。这些实例通常彼此相关，但会被单独评估，且每次评估之间都会终止工具。遗憾的是，这意味着每个实例都会产生启动开销，例如获取图形处理器或预处理待评估网络。这与大多数实际场景不同，在实际场景中，通常会针对同一网络考虑大量输入输出规格。这种差异通过测量并从每次单独的运行时间中扣除这部分开销来弥补。然而，这种开销测量过程不仅存在缺陷并会引入噪声，还可能在简单实例的评估时间中占据主导。

在未来的迭代中，可以一次向工具提供一整批性质，使其更贴近实际应用场景。此外，目前超时时间是按实例设定的，因此始终尝试验证直至超时的策略最优。

然而，在实际场景中，识别出验证可能失败的实例并提前停止，可以显著提升方法的处理能力，进而提高其实用性。对于 VNN-COMP，改为按基准测试设定超时时间，将有助于推动开发实现这一目标的有效启发式方法。此外，工具还可以采用证明共享方法 [15], ，复用某个实例中已验证的子问题来处理后续实例。

6.4 持续竞赛

除了每年举办一次 VNN-COMP 之外，得益于今年引入的自动化提交和评估流程，最新基准测试集的工具提交还可以按滚动方式接收。这将使竞赛从对当前研究成果的年度快照，转变为一个集中式的最先进水平成果库，并随着各团队提交已发表的新方法而不断更新。不过，如果实施时不够谨慎，就可能让工具在提交前针对评估实例进行调优，从而导致比较失衡。此外，所需云计算的经费来源问题仍未解决。

6.5 可靠性评估

神经网络验证器的一项内在要求是必须可靠，即绝不能声称某项安全性属性成立，而实际上它并不成立。然而，评估可靠性并不容易，因为只有在以有效反例证明 VNN-COMP 问题实例可满足时，通常才能确定其真实值。当基准测试中没有任何实例可满足时，这一点尤其成问题：此时，即使对每个实例立即返回不可满足，也无法证明这种做法不可靠。要求每个基准测试中有一定比例的实例（预期）可满足，或许可以缓解这一问题。应对这一挑战的另一个有趣途径是证明生成 [37]。可以新增一个类别，要求或鼓励工具在声称某项属性不可满足时，额外提供可验证证明。

虽然严重的可靠性错误很少见，但提交的工具中几乎没有或完全没有能够保证浮点运算可靠性的工具。也就是说，即使某些工具在使用精确算术时能够保证可靠性，也可能因浮点算术引入的精度损失而变得不可靠。出于性能考虑而选择使用单精度计算的工具尤其如此。可以通过以下方式评估不同工具对这类问题的敏感程度：在专门设计用于揭示浮点运算可靠性问题的基准测试上进行评估。

6.6 其他竞赛模式

可以增设专门的证伪器类别，鼓励各团队开发并提交更强大的攻击方案，超越标准的对抗攻击。此外，还可以增设元求解器类别，以探究那些从其他领域行之有效的一系列方法中进行启发式选择的方案 [59], ，是否能显著优于单个工具。不过，需要确保这些工具相较于单独提交的工具确实具有足够价值，因为单独提交的工具已经结合了不同的验证策略。

6.7 推动通用工具开发

解析大型且复杂的 VNN-LIB 文件，或将 ONNX 文件转换为其他通用格式，可能需要耗费大量时间。尽管许多团队为此自行开发了工具，但仍应重点介绍用于解析 VNN-LIB 文件[1] 以及优化 ONNX 文件（ DNNV [41]）的现有开源工具，并鼓励持续开发。

6.8 尚存的挑战

我们大致可以将神经网络验证中的挑战分为四类：

– 验证规模相对较小但正则化较弱的网络，需要极其精确的分析；即使如此，使用当前方法仍可能难以处理。

– 将精确方法扩展到神经元数量庞大的中型网络（例如小型 ResNet）和数据集（例如 CIFAR100 或 TinyImageNet）颇具挑战，因为基于分支定界的算法，其成本会随所需分割深度呈指数级增长，这使得分支决策既更加困难，也更加重要。

–在稠密输入规范下，将验证扩展到大型网络（例如 VGG-Net 16）和数据集（例如 ImageNet）时，由于神经元数量庞大，需要采用特别节省内存的实现。

– 分类场景之外的验证研究仍不充分，因此缺乏成熟的方法，例如用于图像分割或目标检测的方法。

另一方面，可认证鲁棒网络的训练仍是一个尚未解决的问题。尽管近年来取得了显著进展 [6,18,33,35,38,40,60],，专门训练以展现可证明鲁棒性保证的网络，其标准准确率仍然严重下降。因此，VNN-COMP 中考虑的大多数基准测试，所基于的都是训练时未考虑后续认证的网络。更广泛地说，读者或许也会对人工智能与神经网络控制系统（AINNCS）类别下的连续与混合系统验证国际竞赛（ARCH-COMP）[5] 感兴趣；该竞赛自 2019 [30,29], 起每年举办一次，关注闭环系统中的神经网络验证。

7 给参与者的建议

本节为有意参加 VNN-COMP 但尚未参赛的团队提供一些指导。请注意，这些建议既不是规则，也不是要求。

7.1 给基准测试作者的建议

VNN-COMP 旨在展示神经网络验证能够成功应用的领域，并呈现参赛工具之间有趣的差异。因此，理想情况下，任务不应难到没有任何参与者能够解决任何实例，也不应简单到每种工具都能解决所有实例。对于与现实应用相关的基准测试，我们建议详细描述背景，以突出基准测试的相关性以及验证问题的特点，例如输入或某些网络层的稀疏性。在工具作者着手支持所提议的基准测试时，应预料到他们会提出进一步的问题，并请求进行修改。

7.2 给工具作者的建议

我们建议各团队在提交新的基准测试前，参考以往的基准测试来测试自己的工具。鉴于提交的基准测试种类日益增多，从一开始就支持所有基准测试可能并不可行。如果确实如此，我们建议优先支持全连接网络和卷积 ReLU 网络；过去，这些网络已覆盖广泛的基准测试，同时所需的实现工作量也较小。

[5] https://cps-vo.org/group/ARCH/FriendlyCompetition

有些操作（例如最大池化）也可以借助 DNNV[41] 等工具简化为多个 ReLU 层。此外，我们建议针对对抗攻击开展广泛测试，以尽可能减少正确性错误的发生概率。对于专为特定问题设计的工具，我们也鼓励作者提交相关基准测试，以突出其专业特长。最后，我们建议阅读表现优异的工具所关联的论文（见表5.1），以更好地了解成功团队采用的技术。

8 结论

在本报告中，我们总结了迄今为止于2020年到2022年举办的三届 VNN-COMP 的主要流程和成果。我们着重指出，人们对这一领域的兴趣日益浓厚，这体现在报名参赛的团队数量不断增加，以及纳入考量的基准测试越来越多，其中还包括一些由工业界提交的基准测试。此外，我们观察到，每年纳入考量的网络和规范在规模与复杂性方面都不断增长，推动并体现了该领域的进步。最后，我们强调，标准化的输入和输出格式，以及参与 VNN-COMP 所需的自动化安装和评估流程，提高了验证方法的可访问性。

致谢 本材料基于以下项目的支持：空军科学研究办公室和海军研究办公室资助的项目，奖励编号为 FA9550-19- 1-0288、FA9550-21-1-0121、FA9550-22-1-0019、FA9550-22-1- 0450 和 N00014-22-1-2156；国防高级研究计划局（DARPA）通过合同编号 FA8750-18-C-0089 资助的 Assured Autonomy 计划；以及国家科学基金会（NSF）资助的项目，基金编号为 1911017、2028001、2220401、2220426 和 。本材料中表达的任何观点、发现、结论或建议均属于作者，不一定反映美国空军、美国海军、DARPA 或 NSF 的观点。

参考文献

1. 简单对抗生成器。 https://github.com/stanleybak/simple_adversarial_generator 。访问日期：2022-09-13 2. VNN-COMP 2020 https://www.overleaf.com/ 报告。project/5f0c85e8d15dc10001749fa9 2022 08 访问日期：- - 28 3. Bai, J., Lu, F., Zhang, K., 等：ONNX：开放神经网络交换格式。 https://github.com/onnx/onnx (2019) 4. Bak, S.：执行引导的过近似（EGO），用于提升神经网络验证的可扩展性（2020）

5. Bak, S., Liu, C., Johnson, T.：第二届国际神经网络验证竞赛（VNN-COMP 2021）：总结与结果（2021）。数字对象标识符 10.48550/ARXIV.2109.00498。网址 https://arxiv.org/abs/2109.00498 6。Balunovic, M., Vechev, M.T.：对抗训练与可证明防御：弥合差距。载于：第八届国际学习表征会议，ICLR 2020，埃塞俄比亚亚的斯亚贝巴，2020年4月26日至30日。OpenReview.net（2020）。网址 https://openreview.net/forum?id=SJxSDxrKDr7. Bojarski, M., Del Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., Jackel, L.D., Monfort, M., Muller, U., Zhang, J., Zhang, X., Zhao, J., Zieba, K.：面向自动驾驶汽车的端到端学习（2016）。数字对象标识符 10.48550/ARXIV.1604.07316。网址 https://arxiv.org/abs/1604.07316 8。Botoeva, E., Kouvaros, P., Kronqvist, J., Lomuscio, A., Misener, R.：通过依赖关系分析高效验证神经网络。载于：第34届AAAI人工智能会议（AAAI20）论文集。AAAI出版社（2020）9. Brix, C., Noll, T.：Debona：用于获得更紧界限并加速对抗鲁棒性证明的解耦边界网络分析。CoRR abs/2006.09040（2020）。网址 https://arxiv.org/abs/2006.09040 10。Bunel, R., De Palma, A., Desmaison, A., Dvijotham, K., Kohli, P., Torr, P.H., Kumar, M.P.：用于神经网络验证的拉格朗日分解。人工智能中的不确定性会议（2020）11. Bunel, R., Lu, J., Turkaslan, I., Kohli, P., Torr, P., Kumar, M.P.：用于分段线性神经网络验证的分支定界。《机器学习研究杂志》**21**（2020）（2020）12. Ehlers, R.：分段线性前馈神经网络的形式化验证。载于：自动化验证与分析技术国际研讨会，页码269–286（2017）。数字对象标识符 10/gh25vg 13. Ferlez, J., Shoukry, Y.：Aren：用于线性时不变系统模型预测控制的可靠 ReLU 神经网络架构。载于：第23届混合系统：计算与控制国际会议论文集，HSCC '20。计算机协会，纽约，美国（2020）。数字对象标识符 10.1145/3365365.3382213。网址 https://doi.org/10.1145/3365365.3382213 14。Ferrari, C., Muller, M.N., Jovanovic, N., Vechev, M.T.：通过多神经元松弛引导的分支定界实现完整验证。载于：第十届国际学习表征会议，ICLR 2022，线上会议，2022年4月25日至29日。OpenReview.net（2022）。网址 https://openreview.net/forum?id=l_amHf1oaK 15。Fischer, M., Sprecher, C., Dimitrov, D.I., Singh, G., Vechev, M.T.：用于神经网络验证的共享证书。载于：S. Shoham, Y. Vizel（编）《计算机辅助验证——第34届国际会议，CAV 2022》，以色列海法，8月 7 10, 2022, *Lec-* —论文 *ture Notes in Computer Science* 13371, 127 148 集，第一部分，卷. 页码 –。Springer（2022）。数字对象标识符 10.1007/978-3-031-13185-1_7。网址 https://doi.org/10.1007/978-3-031-13185-1_7 16。

Gehr, T., Mirman, M., Drachsler-Cohen, D., Tsankov, P., Chaudhuri, S., Vechev, M.T.：AI2：利用抽象解释对神经网络进行安全性与鲁棒性认证。载于：2018年IEEE安全与隐私研讨会，SP 2018，论文集，2018年5月21日至23日，美国加利福尼亚州旧金山，页码3–18。IEEE计算机学会（2018）。数字对象标识符 10.1109/SP.2018.00058。网址 https://doi.org/10.1109/SP.2018.00058

17. Goodfellow, I.J., Shlens, J., Szegedy, C.：解释并利用对抗样本。载于：Y. Bengio, Y. LeCun（编）《第三届国际学习表征会议》，ICLR 2015，美国加利福尼亚州圣迭戈，2015年5月7日至9日，会议论文集（2015）。URL http://arxiv.org/abs/1412.6572 18。Gowal, S., Dvijotham, K., Stanforth, R., Bunel, R., Qin, C., Uesato, J., Arandjelovic, R., Mann, T.A., Kohli, P.：关于区间界传播在训练可验证鲁棒模型中的有效性。CoRR **abs/1810.12715**（2018）。URL http://arxiv.org/abs/1810.12715 19. He, K., Zhang, X., Ren, S., Sun, J.：用于图像识别的深度残差学习。载于：2016年IEEE计算机视觉与模式识别会议（CVPR），第770–778页（2016）。数字对象标识符 10.1109/CVPR.2016.90。Henriksen, P., Lomuscio, A.：通过自适应细化和对抗搜索高效验证神经网络。载于：第二十四届欧洲人工智能会议论文集（ECAI20）（2020）。21. Henriksen, P., Lomuscio, A.：Deepsplit：一种通过间接效应分析验证神经网络的高效分割方法。载于：第三十届国际人工智能联合会议论文集（IJCAI21）（即将发表，2021年8月）。22. Huang, X., Kwiatkowska, M., Wang, S., Wu, M.：深度神经网络的安全性验证。载于：R. Majumdar, V. Kunˇcak（编）《计算机辅助验证》，第3–29页。Springer International Publishing，Cham（2017）。23. Julian, K.D., Lopez, J., Brush, J.S., Owen, M.P., Kochenderfer, M.J.：飞机防撞系统的策略压缩。载于：2016年IEEE/AIAA第三十五届数字航空电子系统会议（DASC），第1–10页（2016）。数字对象标识符 10.1109/DASC.2016.7778091。24. Katz, G., Barrett, C., Dill, D.L., Julian, K., Kochenderfer, M.J.：Reluplex：一种用于验证深度神经网络的高效SMT求解器。载于：R. Majumdar, V. Kunˇcak（编）《计算机辅助验证》，第97–117页。Springer International Publishing，Cham（2017）。25. Katz, G., Huang, D.A., Ibeling, D., Julian, K., Lazarus, C., Lim, R., Shah, P., Thakoor, S., Wu, H., Zelji´c, A. 等：Marabou框架：深度神经网络验证与分析框架。载于：国际计算机辅助验证会议，第443–452页。Springer（2019）。26. Khedr, H., Ferlez, J., Shoukry, Y.：利用线性区域几何有效地形式化验证神经网络。arXiv预印本 arXiv:2006.10864（2020）。27. Kouvaros, P., Lomuscio, A.：通过基于依赖关系的分支实现可扩展的完备ReLU神经网络验证。载于：第三十届国际人工智能联合会议论文集（IJCAI21）（即将发表，2021年）。28. Liu, C., Arnon, T., Lazarus, C., Kochenderfer, M.J.：Neuralverification.jl：验证深度神经网络的算法。载于：ICLR 2019机器学习模型调试研讨会（2019）。URL https://debug-ml-iclr2019.github.io/cameraready/DebugML-19_paper_22.pdf 29。Lopez, D.M., Althoff, M., Benet, L., Chen, X., Fan, J., Forets, M., Huang, C., Johnson, T.T., Ladner, T., Li, W., Schilling, C., Zhu, Q.：ARCH-COMP22类别报告：面向连续与混合系统工厂的人工智能与神经网络控制系统（AINNCS）。载于：G. Frehse, M. Althoff, E. Schoitsch, J. Guiochet（编）《第九届连续与混合系统应用验证国际研讨会论文集

（ARCH22），*EPiC Series in Computing*，第90卷，第142–184页。EasyChair（2022）。数字对象标识符 10.29007/wfgr。URL https://easychair.org/publications/paper/C1S8。

Lopez, D.M., Musau, P., Tran, H.D., Dutta, S., Carpenter, T.J., Ivanov, R., Johnson, T.T.：Arch-comp19类别报告：面向连续与混合系统工厂的人工智能与神经网络控制系统（AINNCS）。载于：G. Frehse, M. Althoff（编）ARCH19。《第六届连续与混合系统应用验证国际研讨会》，*EPiC Series in Computing* 61，103–119，第卷，第–页。EasyChair（2019）。数字对象标识符 10.29007/rgv8。URL https://easychair.org/publications/paper/8Krs。

Lu, J., Kumar, M.P.：用于神经网络验证的神经网络分支。载于：国际学习表征会议（2020）。32. Madry, A., Makelov, A., Schmidt, L., Tsipras, D., Vladu, A.：迈向能够抵御对抗攻击的深度学习模型。载于：第六届国际学习表征会议，ICLR 2018，加拿大不列颠哥伦比亚省温哥华，2018年4月30日至5月3日，会议论文集。OpenReview.net（2018）。URL https://openreview.net/forum?id=rJzIBfZAb33. Mirman, M., Gehr, T., Vechev, M.T.：用于可证明鲁棒神经网络的可微抽象解释。载于：J.G. Dy, A. Krause（编）《第三十五届国际机器学习会议论文集》，ICML 2018，瑞典斯德哥尔摩，2018年7月10日至15日，*Proceedings of Machine Learning Research*，第80卷，第3575–3583页。PMLR（2018）。URL http://proceedings.mlr.press/v80/mirman18b.html。Müller, M.N., Brix, C., Bak, S., Liu, C., Johnson, T.T.：第三届神经网络验证国际竞赛（VNN-COMP 2022）：总结与结果（2022）。数字对象标识符 10.48550/arXiv.2212.10376。URL https://arxiv.org/abs/2212.10376。Müller, M.N., Eckert, F., Fischer, M., Vechev, M.T.：认证训练：小框足矣。CoRR **abs/2210.04871**（2022）。数字对象标识符 10.48550/arXiv.2210.04871。URL https://doi.org/10.48550/arXiv.2210.04871 36. Müller, M.N., Makarchuk, G., Singh, G., Püschel, M., Vechev, M.：Prima：通过多神经元凸松弛实现精确且通用的神经网络认证。arXiv预印本 arXiv:2103.03638（2021）。37. O. Isac C. Barrett, M.Z., Katz, G.：通过证明生成进行神经网络验证。载于：第二十二届计算机辅助设计形式化方法国际会议（FMCAD）（2022）。38. Palma, A.D., Bunel, R., Dvijotham, K., Kumar, M.P., Stanforth, R.：通过分支定界实现经验证的对抗鲁棒性的IBP正则化。CoRR**abs/2206.14772**（2022）。数字对象标识符 10.48550/arXiv.2206.14772。URL https://doi.org/10.48550/arXiv.2206.14772 39. Serre, F., Müller, C., Singh, G., Püschel, M., Vechev, M.：在图形处理器上扩展多面体神经网络验证。载于：《机器学习与系统论文集》（MLSys）（2021）。40. Shi, Z., Wang, Y., Zhang, H., Yi, J., Hsieh, C.：通过短预热实现快速认证鲁棒训练。载于：M. Ranzato, A. Beygelzimer, Y.N. Dauphin, P. Liang, J.W. Vaughan（编）《神经信息处理系统进展34：神经信息处理系统年度会议2021》，NeurIPS 2021，2021年12月6日至14日，线上，第18335–18349页（2021）。URL https://proceedings.neurips.cc/paper/2021/hash/988f9153ac4fd966ea302dd9ab9bae15-Abstract.html41. Shriver, D., Elbaum, S., Dwyer, M.B.：Dnnv：深度神经网络验证框架。载于：A. Silva、K.R.M. Leino（编）：计算机辅助验证，第137–150页。Springer International Publishing，Cham（2021）42. Shriver, D., Elbaum, S.G., Dwyer, M.B.：通过对抗攻击简化深度神经网络属性以实现证伪。载于：第43届IEEE/ACM国际软件工程会议，ICSE 2021，马德里，西班牙，2021年5月22日至30日，第275–287页。IEEE（2021）。数字对象标识符 10.1109/ICSE43902.2021.00036。URL https://doi.org/10.1109/ICSE43902.2021.0003643. Singh, G., Ganvir, R., Püschel, M., Vechev, M.：突破神经网络认证的单神经元凸屏障。载于：神经信息处理系统进展32，第15098–15109页。Curran Associates, Inc.（2019）44. Singh, G., Gehr, T., Mirman, M., Püschel, M., Vechev, M.：快速且有效的鲁棒性认证。载于：S. Bengio、H. Wallach、H. Larochelle、K. Grauman、N. Cesa-Bianchi、R. Garnett（编）：神经信息处理系统进展31，第10802–10813页。Curran Associates, Inc.（2018）。URL http://papers.nips.cc/paper/8278-fast-and-effective-robustness-certification.pdf。Singh, G., Gehr, T., Püschel, M., Vechev, M.：用于认证神经网络的抽象域。ACM程序设计语言汇刊。**3**（POPL），41:1–41:30（2019）46. Singh, G., Gehr, T., Püschel, M., Vechev, M.：提升神经网络的鲁棒性认证。载于：国际学习表征会议论文集（ICLR）（2019）47. Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I.J., Fergus, R.：神经网络的奇特属性。载于：Y. Bengio、Y. LeCun（编）：第2届国际学习表征会议，ICLR 2014，加拿大艾伯塔省班夫，2014年4月14日至16日，会议论文集（2014）。URL http://arxiv.org/abs/1312.6199。Tacchella, A., Pulina, L., Guidotti, D., Demarchi, S.：神经网络验证库（vnn-lib）。https://www.vnnlib.org（2019）49. Tjeng, V., Xiao, K.Y., Tedrake, R.：使用混合整数规划评估神经网络的鲁棒性。载于：ICLR（2019）50. Tran, H.D., Bak, S., Xiang, W., Johnson, T.T.：使用图像星集验证深度卷积神经网络。载于：第32届计算机辅助验证国际会议（CAV）。Springer（2020）51. Tran, H.D., Musau, P., Lopez, D.M., Yang, X., Nguyen, L.V., Xiang, W., Johnson, T.T.：面向前馈神经网络的可并行化可达性分析算法。载于：第7届软件工程形式化方法国际研讨会（FormaliSE'19），FormaliSE '19，第31–40页。IEEE Press，美国新泽西州皮斯卡塔韦（2019）。数字对象标识符 10.1109/FormaliSE.2019.00012 52. Tran, H.D., Musau, P., Lopez, D.M., Yang, X., Nguyen, L.V., Xiang, W., Johnson, T.T.：面向深度神经网络的基于星集的可达性分析。载于：第23届形式化方法国际研讨会（FM'19）。Springer International Publishing（2019）53. Tran, H.D., Yang, X., Lopez, D.M., Musau, P., Nguyen, L.V., Xiang, W., Bak, S., Johnson, T.T.：NNV：用于深度神经网络和

学习型信息物理系统的神经网络验证工具。载于：第32届计算机辅助验证国际会议（CAV）（2020）54. Vincent, J.A., Schwager, M.：可达多面体推进（rpm）：一种用于配备深度神经网络组件的机器人系统的安全验证算法（2021）55. Wang, S., Zhang, H., Xu, K., Lin, X., Jana, S., Hsieh, C.J., Kolter, Z.：Beta-CROWN：通过逐神经元分裂约束进行高效界传播，以实现完备和不完备的神经网络验证。arXiv预印本 arXiv:2103.06624（2021）56. Xiang, W., Tran, H., Johnson, T.T.：多层神经网络的输出可达集估计与验证。IEEE神经网络与学习系统汇刊 **29**（11），5777–5783（2018）57. Xu, K., Shi, Z., Zhang, H., Wang, Y., Chang, K.W., Huang, M., Kailkhura, B., Lin, X., Hsieh, C.J.：面向可扩展认证鲁棒性及其他任务的自动扰动分析。神经信息处理系统进展 **33** （2020）58. Xu, K., Zhang, H., Wang, S., Wang, Y., Jana, S., Lin, X., Hsieh, C.J.：快速且完备：利用快速且大规模并行的不完备验证器实现完备神经网络验证。载于：国际学习表征会议（2021）。URL https://openreview.net/forum?id=nVZtXBI6LNn 59。Xu, L., Hutter, F., Hoos, H.H., Leyton-Brown, K.：Satzilla：面向可满足问题的基于算法组合的算法选择。人工智能研究期刊。**32**（1），565–606（2008）60. Zhang, H., Chen, H., Xiao, C., Gowal, S., Stanforth, R., Li, B., Boning, D.S., Hsieh, C.：迈向可验证鲁棒神经网络的稳定高效训练。载于：第8届国际学习表征会议，ICLR 2020，亚的斯亚贝巴，埃塞俄比亚，2020年4月26日至30日。OpenReview.net（2020）。URL https://openreview.net/forum?id=Skxuk1rFwB61. Zhang, H., Wang, S., Xu, K., Li, L., Li, B., Jana, S., Hsieh, C., Kolter, J.Z.：用于基于界传播的神经网络验证的通用割平面。CoRR **abs/2208.05740** （2022）。数字对象标识符 10.48550/arXiv.2208.05740。URL https://doi.org/10.48550/arXiv.2208.05740 62. Zhang, H., Weng, T., Chen, P., Hsieh, C., Daniel, L.：使用通用激活函数进行高效的神经网络鲁棒性认证。载于：S. Bengio、H.M. Wallach、H. Larochelle、K. Grauman、N. Cesa-Bianchi、R. Garnett（编）：神经信息处理系统进展31：神经信息处理系统年度会议 2018, 2018, 3 8, 2018, ´ 4944 4953 2018 NeurIPS 12月—加拿大蒙特利尔，第–页（）。URL https://proceedings.neurips.cc/paper/2018/hash/d04863f100d59b3eb688a11f95b0ae60-Abstract.html63. Zhang, H., Weng, T.W., Chen, P.Y., Hsieh, C.J., Daniel, L.：使用通用激活函数进行高效的神经网络鲁棒性认证。神经信息处理系统进展 **31**，4939–4948（2018）。URL https://arxiv.org/pdf/1811.00866.pdf 64。Zhou, Chellappa：使用神经网络计算光流。载于：IEEE 1988年国际神经网络会议，第71–78页，第2卷（1988）。数字对象标识符 10.1109/ICNN.1988.23914