VNN-COMP2020 Report

Draft as of September 6, 2020

Authored by the participants and organizers

Abstract

This report summarizes the first International Verification of Neural Networks Competition (VNN-COMP) in 2020, held as a part of the 3rd International Workshop on Verification of Neural Networks (VNN) that was collocated with the 32nd International Conference on Computer-Aided Verification (CAV).

1 Introduction

The competition was organized using online discussion, primarily through a GitHub repository[1]. Based on discussion and feedback, three categories were selected, based primarily around the architecture of the neural networks, specifically in terms of the types of layers and activation functions used within the neural networks. The three categories are:

- CNN: convolutional neural networks, primarily restricted to convolutional layers,
- PWL: fully-connected neural networks with piecewise linear activation functions, specifically rectified linear units (ReLUs), and
- NLN: fully-connected neural networks with nonlinear activation functions, specifically tanh and sigmoid.

2 Participants

The following tools and teams participated in VNN-COMP.

2.1 nnenum (Neural Network Enumeration Tool)

Team: Stanley Bak

The nnenum tool uses multiple levels of abstraction to achieve high-performance verification of ReLU networks without sacrificing completeness [1]. Analysis combines three types of zonotopes with star set (triangle) overapproximations [24], and uses efficient parallelized ReLU case splitting [2]. The verification tree search can be augmented with adversarial example generation using multiple attacks from the foolbox library to quickly find property violations [15]. The tool is written in Python 3, uses GLPK for LP solving and directly accepts ONNX network files as input. The ImageStar method [22] allows sets to be quickly propagated through all layers supported by the ONNX runtime, such as convolutional layers with arbitrary parameters.

The code for the competition version is available at `https://github.com/stanleybak/nnenum` in the `vnn2020` branch, including scripts to run the tool in Docker. The measurements reported here were performed on an Amazon EC2 `m4.10xlarge` cloud instance, with a 40-core 2.4 GHz Intel Xeon E5-2676 v3 processor and 160 GB RAM, running Ubuntu 18.04. The nnenum tool participated in the CNN and PWL category.

[1] `https://github.com/verivital/vnn-comp/`

2.2 NNV (Neural Network Verification Tool)

Team: Neelanjana Pal, Hoang-Dung Tran, Taylor Johnson

The Neural Network Verification Tool (NNV) [22–25, 28] is written primarily with Matlab and implements reachability-analysis methods for neural network verification with a particular focus on applications of closed-loop neural network control systems in autonomous cyber-physical systems. NNV uses geometric representations such as star sets that allows for a layer-by-layer computation of the exact reachable set for feed-forward deep neural networks. In the event that a particular safety property is violated, NNV can construct and visualize the complete set of counterexample inputs for a neural network. NNV along with the relevant experiments and related publications are available at https://github.com/verivital/nnv.

NNV participated in the CNN (on GGN-Nets), PWL, and NLN categories. Most experiments were run on an 8-core Intel(R) Core(TM) i7-7700 3.60GHz CPU with 64 GB RAM and Ubuntu 16.04.6 LTS as the OS, with some on a 16 core Intel(R) Xeon(R) Platinum 8259CL CPU 2.50GHz Codeocean instance (shared cloud platform). The code for this competition is available at https://github.com/verivital/vnn-comp.

2.3 PeregriNN

Team: Haitham Khedr, Yasser Shoukry

PeregriNN [11] is a sound and complete Neural Network verification tool written in Python 3 focusing on ReLU activations. It uses search and optimization to verify the safety property. PeregriNN searches over neurons' phases in a way that emphasizes and exploits the geometry of their activation regions. In particular, PeregriNN splits neurons which result in small highly linear regions in the input space such that it would be faster for the optimizer to verify the property in these regions. Gurobi is used for LP solving and the tool supports **nnet** and **keras** network formats as input.

The code for the competition is available at https://github.com/rcpsl/PeregriNN/tree/vnn in **vnn** branch with scripts to run experiments on **ACASXU** and **MNIST** datasets. We ran our experiments on a 24-core 2.2 GHz Intel Xeon E5-2650 v4 machine with 256 GB of memory. The tool participated in PWL category.

2.4 VeriNet

Team: Patrick Henriksen, Alessio Lomuscio

VeriNet [8] is a complete symbolic-interval propagation-based toolkit for local robustness properties; the toolkit supports ReLU, Sigmoid and Tanh activation functions and Fully-Connected, Convolutional and Batch-normalisation layers. VeriNet uses several novel techniques to achieve a high level of performance, including a local gradient-based search around spurious counterexamples, an adaptive refinement strategy and optimal relaxations for the Sigmoid and Tanh activation functions.

We ran our toolkit on benchmarks in all three categories; however, we did not evaluate the ACAS Xu Benchmarks as VeriNet does not support some of the properties of ACAS Xu. We also could not verify the OVAL Cifar10 networks due to time constraints. We ran all benchmarks on a workstation with a Ryzen 3700X 3.6 GHz 8-core CPU, 64 GB ram running Ubuntu 20.04 LTS with Linux kernel 5.4.0. The VeriNet code is available at https://vas.doc.ic.ac.uk/software/neural/.

2.5 Venus

Team: Elena Botoeva, Panagiotis Kouvaros, Jan Kronqvist, Alessio Lomuscio, Ruth Misener

Venus [3] is a complete verification toolkit for Relu-based feed-forward neural networks. It can be used to check reachability and local adversarial robustness properties. Venus implements a MILP-based verification method whereby it leverages dependency relations between the ReLU nodes to reduce the search space that needs to be considered during branch-and-bound. It additionally implements methods based on symbolic interval propagation and input domain splitting.

Venus participated in the PWL category. The experiments were carried out on an Intel Core i7-7700K (4 cores) equipped with 16GB RAM, running Linux kernel 4.15. The code of Venus is available at `https://vas.doc.ic.ac.uk/software/neural/`.

2.6 OVAL Verification Methods

Team: Harkirat Singh Behl, Alessandro De Palma, Florian Jaeckle, Jingyue Lu, M. Pawan Kumar

The OVAL verification tools [4,5,12] are written in PyTorch [14] and focus on verifying piecewise linear neural networks. Verification is formulated as a non-convex minimisation problem, which is solved with a Branch-and-Bound framework [5]. As a bounding subroutine, OVAL employs two solvers operating on the Lagrangian Decomposition dual of the network: a *Proximal* method, and a *Supergradient* method [4]. The branching strategies are designed to split on the piecewise-linear activations' states: *SR*, a heuristic method [5], and *GNN*, a learnt branching strategy [12]. All methods in the framework allow for an efficient GPU implementation, resulting in a high degree of parallelism. OVAL participated in the CNN category, presenting the best performing configuration of the framework for each dataset.

The experiments were executed on Titan Xp GPUs, on machines with Intel Core i9-7900X (6 cores) and i7-6850K CPUs.

2.7 MIPVerify

Team: Vincent Tjeng, Kai Xiao, Russ Tedrake

MIPVerify [21] is a complete verification tool written in Julia, focusing on verification for piecewise linear neural networks. Users submits a query regarding the satisfiability of a property to this tool, which translates the query into a tightly specified mixed-integer linear programming problem.[2] The problem can then be solved by any of the open-source or commercial solvers supported by JuMP [6]. The tool supports the ReLU and maximum non-linearities, as well as fully-connected, convolutional, normalization layers and residual blocks [7].

MIPVerify participated in the PWL and CNN category. The experiments were carried out with MIPVerify `v0.2.3` and Gurobi `9.0.2` with 4 threads on a 4-core Intel Core i5-4300U 1.90GHz CPU, with 8 GB RAM (DDR3, 1600 MHz), running Ubuntu 18.04.4 LTS. Full details can be found in the accompanying code submission, and code for MIPVerify is available at `https://github.com/vtjeng/MIPVerify.jl`.

[2] The tool also supports *optimizing* the value of a user-specified objective, given constraints on the input and activation values in the network; however, this functionality was not used for any of the benchmarks.

2.8 ERAN

Team: Gagandeep Singh, Dimitar I. Dimitrov, Makarchuk Gleb, Markus Püschel, Martin Vechev

ERAN [16–19] is a neural network certifier based on abstract interpretation. ERAN supports both incomplete and complete verification and can handle fully-connected, convolutional, residual network architectures containing ReLU, Sigmoid, Tanh, or Maxpool non-linearities. The analyzer is written in Python and uses ELINA [20] for numerical domains, cdd [26] for computing convex hull for the k-ReLU [16] framework and Gurobi for solving MILP instances.

ERAN participated in all three categories. When run in complete mode, ERAN generates concrete counterexamples. In incomplete mode, ERAN produces abstract counter examples which are then checked for validity. If this fails, then ERAN runs PGD attack for generating counter examples. All steps are run in a way so that the runtime of ERAN never exceeds the specified timelimit. We note that we use a fix abstraction for all benchmarks corresponding to a network when run in incomplete mode, thus ERAN can fail to verify much earlier than the specified timeout even though it maybe possible to refine the abstraction for potentially verifying the property within the timelimit but we did not do so because of time constraints.

The code for ERAN is available at `https://github.com/eth-sri/eran`. We made changes to the front-end for handling the different types of networks and properties used in the competition. We used all cores of the following machines for running the experiments:

- 3.3 GHz 10 Core Intel i9-7900X Skylake CPU with a main memory of 64 GB: Certification with ERAN of all ACASXu, PAT-FCN, GGN-CNN (except mnist_0.3.onnx), and NEEL-NLN.
- 2.2 GHz 12 core Intel Xeon CPU E5-2650 with 756 GB of main memory: PGD attacks for generating concrete counterexamples for the PAT-FCN and NEEL-NLN networks.
- 2.6 GHz 14 core Intel Xeon CPU E5-2690 with 512 GB of main memory: Certification with ERAN of the networks from the OVAL group and mnist_0.3.onnx from GGN-CNN.

3 Benchmarks

Benchmarks were solicited in the different categories. After discussion among the participants in these categories, the following benchmarks were selected. Most benchmarks selected were to establish robustness to bounded perturbations for networks that perform image classification, with one additional benchmark selected as a form of input-output behavior.

3.1 PWL

3.1.1 ACASXU-ALL

ACASXU-ALL[3] consists of ten properties defined over 45 neural networks used to issue turn advisories to aircraft to avoid collisions. The neural networks have 300 neurons arranged in 6 layers, with ReLU activation functions. There are five inputs corresponding to the aircraft states, and five network outputs, where the minimum output is used as the turn advisory the system ultimately produces.

The inputs should be scaled using `means = [19791.091, 0.0, 0.0, 650.0, 600.0, 7.5188840201005975]` and `ranges = [60261.0, 6.28318530718, 6.28318530718, 1100.0,`

[3] `https://github.com/verivital/vnn-comp/tree/master/2020/PWL/benchmark/acasxu`

`1200.0, 373.94992]`, where `input[i] = (input[i] - means[i]) / ranges[i]`. The outputs are scaled using the last value, `output[i] = (output[i] * ranges[-1]) + means[-1]`. Output scaling only matters for property 1, as the other properties compare relative output values. We use the original 10 properties [9], where properties 1-4 are checked on all 45 networks as was done in later work by the original authors [10]. Properties 5-10 are checked on a single network. The total number of benchmarks is therefore 186. The original verification times ranged from seconds to days—including some benchmark instances that did not finish. VNN2020 used a five minute timeout for each property.

3.1.2 ACASXU-HARD

ACASXU-HARD consists of ten benchmark instances—specification-network pairs—run with a longer, six hour timeout. The input-output scaling is the same as in the ACAS Xu benchmarks in the previous section. The specific benchmarks considered are shown in Table 2.

3.1.3 PAT-FCN

These MNIST fully-connected benchmarks[4] use three fully-connected networks with 2, 4 and 6 layers, respectively, and 256 ReLU nodes in each layer. For each network a local robustness check is performed with a ϵ-perturbation of 0.02 and 0.05 on 25 images from the MNIST test set. Pixels are normalised to $[0, 1]$ by dividing each pixel by 255.

3.2 CNN

3.2.1 OVAL

The OVAL benchmark[5] consists of sets of adversarial robustness properties specifically designed for three adversarially trained CIFAR-10 convolutional neural networks with ReLU activations [12]. Two networks are composed of 2 convolutional layers followed by 2 fully connected layers: a "*Base*" model, and a wider "*Wide*" model. A "*Deep*" model has 2 additional convolutional layers, with a width analogous to the *Base* model. All three models are trained robustly using the method by Wong and Kolter [27]. The network input is normalised with: $\bar{x} = [0.485, 0.456, 0.406]$, $\sigma = [0.225, 0.225, 0.225]$.

The properties are of the following type: given a correctly classified image x with label y, verify that the trained network misclassifies a slightly perturbed image as y' ($y' \neq y$). The label y' is randomly selected, and the allowed perturbation is determined by an epsilon value under ℓ_∞, which is specific to each network and image (found via binary search) and applied to the normalised input space. All the properties are not satisfied (the network is robust). For each of the three models there are 100 (or a smaller subset of 20) properties, subsets of a previously presented dataset [12]. A timeout of one hour per property is suggested.

3.2.2 GGN-CNN

The 2 MNIST and 2 CIFAR10 neural networks[6] in the category are taken from COLT [13] available publicly at `https://github.com/eth-sri/colt`. The inputs for the networks are normalized normalized as follows:

[4]MNIST: `https://github.com/verivital/vnn-comp/tree/master/2020/PWL/benchmark/mnist/oval`
[5]CIFAR: `https://github.com/verivital/vnn-comp/tree/master/2020/CNN/benchmark/cifar/oval`
[6]MNIST: `https://github.com/verivital/vnn-comp/tree/master/2020/CNN/benchmark/mnist/eth`, CIFAR: `https://github.com/verivital/vnn-comp/tree/master/2020/CNN/benchmark/cifar/eth`

- MNIST: $\bar{x} = 0.1307$, $\sigma = 0.3081$.
- CIFAR10: $\bar{x} = [0.4914, 0.4822, 0.4465]$, $\sigma = [0.2023, 0.1994, 0.2010]$.

The goal is to verify that the classfication is robust within an adversarial region defined by an L_∞-ball of radius ϵ around an image in the test set. We consider the first 100 images of the test set discarding those that are not correctly classified. The ϵ values for the two MNIST networks are 0.1 and 0.3 respectively. The corresponding values for the CIFAR10 networks are 2/255 and 8/255 respectively.

3.3 NLN

3.3.1 NEEL-NLN

These nonlinear fully-connected MNIST benchmarks[7] use a total of 4 fully-connected networks (2 networks with tanh activation function and 2 with sigmoid) with 2 and 3 layers respectively. Each of the 2-layered networks have 200 and 50 ReLU and in each layer whereas the 3-layered ones have 200,100 and 50 nodes. Input images are flattened but not normalized (i.e. the pixel values are in $[0, 255]$). For each network a local robustness check is performed with ϵ-perturbation of 5 and 12 on first 16 images from the MNIST testset similar as Section 3.1.3. Results for this benchmark are shown in Table 6.

[7] MNIST: https://github.com/verivital/vnn-comp/tree/master/2020/NLN/benchmark/mnist

4 Results

4.1 PWL

4.1.1 ACASXU-ALL

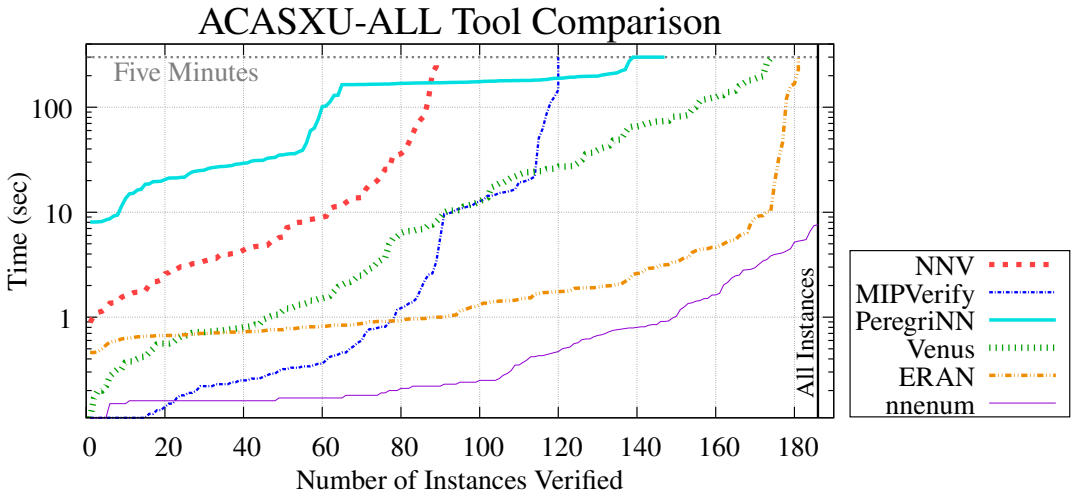

Figure 1: ACASXU-ALL Cactus Plot.

Table 1: Tool Runtime (sec) for ACASXU-ALL.

Prop	Net	Result	nnenum	NNV	PeregriNN	MIPVerify	Venus	ERAN
1	1-1	UNSAT	0.51	-	18.58	-	1.32	0.62
1	1-2	UNSAT	0.65	-	9.22	-	3.00	0.66
1	1-3	UNSAT	0.81	-	33.03	-	22.73	0.81
1	1-4	UNSAT	0.81	-	24.65	-	17.85	0.81
1	1-5	UNSAT	0.62	-	8.66	-	2.25	0.67
1	1-6	UNSAT	1.63	-	8.15	-	2.03	0.63
1	1-7	UNSAT	0.62	-	8.09	-	1.06	0.64
1	1-8	UNSAT	0.79	-	8.05	62.35	0.55	0.58
1	1-9	UNSAT	0.67	-	8.06	86.62	0.46	0.70
1	2-1	UNSAT	1.37	-	14.97	-	11.35	0.70
1	2-2	UNSAT	2.96	-	101.09	-	24.34	0.95
1	2-3	UNSAT	1.35	-	37.70	-	27.60	0.67
1	2-4	UNSAT	0.61	-	13.48	-	6.83	0.67
1	2-5	UNSAT	3.42	-	-	-	24.63	1.77
1	2-6	UNSAT	2.14	-	-	-	74.37	1.48
1	2-7	UNSAT	3.19	-	-	-	193.70	4.50
1	2-8	UNSAT	2.83	-	-	-	186.45	5.25
1	2-9	UNSAT	5.18	-	-	-	160.32	4.62
1	3-1	UNSAT	1.69	-	113.10	-	7.63	0.84

Table 1 Continued

Prop	Net	Result	nnenum	NNV	PeregriNN	MIPVerify	Venus	ERAN
1	3-2	UNSAT	0.78	-	59.74	-	10.43	0.84
1	3-3	UNSAT	1.25	-	26.87	-	6.93	0.78
1	3-4	UNSAT	0.75	-	16.59	-	6.54	0.66
1	3-5	UNSAT	4.10	-	191.42	-	12.20	1.15
1	3-6	UNSAT	5.48	-	-	-	263.55	3.15
1	3-7	UNSAT	3.84	-	-	-	124.21	3.17
1	3-8	UNSAT	3.54	-	-	-	122.72	2.92
1	3-9	UNSAT	2.47	-	-	-	197.10	5.47
1	4-1	UNSAT	2.03	-	-	-	82.56	1.53
1	4-2	UNSAT	2.86	-	19.62	-	21.36	0.76
1	4-3	UNSAT	0.92	-	63.32	-	13.22	0.67
1	4-4	UNSAT	0.98	-	26.58	-	7.58	0.72
1	4-5	UNSAT	1.62	-	129.67	-	80.94	2.57
1	4-6	UNSAT	5.30	-	-	-	179.19	5.37
1	4-7	UNSAT	4.28	-	-	-	-	-
1	4-8	UNSAT	3.96	-	-	-	-	3.69
1	4-9	UNSAT	6.59	-	-	-	-	159
1	5-1	UNSAT	0.84	-	11.43	-	5.66	0.54
1	5-2	UNSAT	1.18	-	35.20	-	10.42	0.59
1	5-3	UNSAT	0.73	-	9.39	-	3.51	0.47
1	5-4	UNSAT	0.79	-	15.14	-	8.86	0.50
1	5-5	UNSAT	1.51	-	77.21	-	29.46	1.04
1	5-6	UNSAT	4.03	-	-	-	-	1.91
1	5-7	UNSAT	2.35	-	-	-	130.38	1.71
1	5-8	UNSAT	2.81	-	-	-	-	7.92
1	5-9	UNSAT	4.48	-	-	-	120.19	3.15
2	1-1	UNSAT	0.77	-	300.11	-	27.19	2.60
2	1-2	SAT	0.35	-	28.83	-	2.89	4.40
2	1-3	SAT	0.91	-	300.15	-	1.16	10.2
2	1-4	SAT	0.35	-	27.20	-	27.21	1.91
2	1-5	SAT	0.37	-	103.18	-	81.99	4.1
2	1-6	SAT	0.76	-	35.71	125.49	1.44	3.69
2	1-7	UNSAT	0.91	-	300.11	-	72.14	1.88
2	1-8	UNSAT	1.42	-	300.09	-	180.88	2.98
2	1-9	UNSAT	0.85	-	300.16	-	73.94	1.72
2	2-1	SAT	0.17	-	21.19	21.86	13.78	1.95
2	2-2	SAT	0.16	-	26.08	145.67	0.78	1.51
2	2-3	SAT	0.16	-	21.53	10.00	1.44	1.52
2	2-4	SAT	0.17	-	21.19	20.02	0.56	4.12
2	2-5	SAT	0.17	-	27.84	11.32	0.77	1.93
2	2-6	SAT	0.17	-	28.60	12.28	0.95	1.91
2	2-7	SAT	0.16	-	29.31	19.51	0.57	4.47
2	2-8	SAT	0.16	-	31.22	15.44	0.76	2.94
2	2-9	SAT	1.44	-	30.70	-	0.98	6.18
2	3-1	SAT	0.16	-	19.53	9.59	0.56	1.72

Table 1 Continued

Prop	Net	Result	nnenum	NNV	PeregriNN	MIPVerify	Venus	ERAN
2	3-2	SAT	0.23	-	18.50	-	0.82	8.02
2	3-3	UNSAT	7.46	-	300.15	-	294.66	167
2	3-4	SAT	0.16	-	19.74	9.71	6.60	4.63
2	3-5	SAT	0.16	-	21.49	12.69	2.49	2.0
2	3-6	SAT	0.16	-	36.07	-	0.85	4.3
2	3-7	SAT	0.43	-	32.62	-	1.21	19.5
2	3-8	SAT	0.23	-	31.11	110.51	1.55	8.7
2	3-9	SAT	0.16	-	27.28	14.67	2.16	2.82
2	4-1	SAT	0.17	-	45.65	-	81.66	3.49
2	4-2	UNSAT	7.59	-	300.24	-	-	230
2	4-3	SAT	0.16	-	23.07	11.07	0.72	1.42
2	4-4	SAT	0.29	-	16.24	-	0.50	2.06
2	4-5	SAT	0.16	-	29.51	14.34	2.13	4.99
2	4-6	SAT	0.16	-	27.55	14.35	1.31	6.42
2	4-7	SAT	0.16	-	36.08	11.48	0.71	3.32
2	4-8	SAT	0.16	-	24.18	15.68	72.99	3.37
2	4-9	SAT	0.17	-	300.14	50.52	-	5.99
2	5-1	SAT	0.16	-	21.16	13.18	11.80	1.31
2	5-2	SAT	0.17	-	24.93	10.78	0.74	1.45
2	5-3	SAT	0.85	-	300.14	-	9.94	9.7
2	5-4	SAT	0.16	-	20.46	10.57	1.02	1.75
2	5-5	SAT	0.16	-	25.13	16.04	0.65	1.41
2	5-6	SAT	0.16	-	38.99	15.22	0.74	1.90
2	5-7	SAT	0.16	-	31.09	16.50	1.05	1.89
2	5-8	SAT	0.16	-	35.01	20.79	0.72	2.63
2	5-9	SAT	0.16	-	33.30	19.19	0.68	9.28
3	1-1	UNSAT	0.80	-	177.32	3.38	69.51	46.3
3	1-2	UNSAT	0.68	244.63	177.47	1.22	6.87	1.43
3	1-3	UNSAT	0.57	55.96	186.40	5.72	36.91	1.99
3	1-4	UNSAT	0.24	13.68	193.36	0.32	2.47	1.0
3	1-5	UNSAT	0.25	18.25	171.46	0.22	1.21	0.76
3	1-6	UNSAT	0.16	4.60	169.71	0.12	0.80	0.65
3	1-7	SAT	0.16	2.17	178.45	0.09	0.20	1.68
3	1-8	SAT	0.16	1.73	172.65	0.08	0.08	1.36
3	1-9	SAT	0.16	1.55	176.41	0.06	0.20	1.41
3	2-1	UNSAT	0.50	67.63	202.59	0.89	5.86	1.03
3	2-2	UNSAT	0.42	22.24	170.25	0.81	10.39	1.66
3	2-3	UNSAT	0.44	38.68	210.13	2.23	11.58	0.74
3	2-4	UNSAT	0.01	1.71	193.37	0.08	0.42	0.73
3	2-5	UNSAT	0.19	8.45	172.52	0.24	0.62	0.46
3	2-6	UNSAT	0.16	1.64	164.48	0.13	0.75	0.90
3	2-7	UNSAT	0.17	4.16	163.83	0.18	0.79	0.82
3	2-8	UNSAT	0.15	1.74	164.24	0.09	0.72	0.82
3	2-9	UNSAT	0.01	1.21	170.51	0.06	0.17	0.87
3	3-1	UNSAT	0.25	19.99	169.25	0.22	6.32	0.69

Table 1 Continued

Prop	Net	Result	nnenum	NNV	PeregriNN	MIPVerify	Venus	ERAN
3	3-2	UNSAT	0.52	210.03	175.24	2.54	4.92	0.78
3	3-3	UNSAT	0.39	35.06	168.00	0.50	2.69	0.84
3	3-4	UNSAT	0.25	8.64	170.87	1.17	1.51	0.67
3	3-5	UNSAT	0.18	3.96	176.65	0.23	0.38	0.71
3	3-6	UNSAT	0.22	7.90	178.48	0.77	3.34	0.72
3	3-7	UNSAT	0.16	1.11	164.21	0.06	0.22	0.72
3	3-8	UNSAT	0.18	2.87	170.57	0.22	0.66	0.63
3	3-9	UNSAT	0.17	4.83	170.25	0.12	0.40	0.80
3	4-1	UNSAT	0.26	8.77	174.71	0.57	3.70	0.86
3	4-2	UNSAT	0.47	83.21	181.34	1.27	24.07	1.36
3	4-3	UNSAT	0.56	121.59	173.23	1.35	8.14	0.69
3	4-4	UNSAT	0.18	2.19	179.86	0.31	0.38	0.46
3	4-5	UNSAT	0.16	1.40	164.54	0.10	0.37	0.67
3	4-6	UNSAT	0.22	11.27	166.16	0.79	1.94	0.68
3	4-7	UNSAT	0.16	3.50	165.80	0.15	1.26	0.65
3	4-8	UNSAT	0.17	2.66	168.72	0.19	1.49	0.92
3	4-9	UNSAT	0.20	4.03	165.40	0.29	1.60	0.71
3	5-1	UNSAT	0.43	34.87	181.44	1.39	27.17	1.78
3	5-2	UNSAT	0.21	7.28	179.12	0.26	0.56	2.32
3	5-3	UNSAT	0.25	8.53	196.05	0.44	1.36	0.72
3	5-4	UNSAT	0.17	3.02	197.42	0.22	0.73	0.73
3	5-5	UNSAT	0.18	4.36	196.68	0.34	0.85	0.70
3	5-6	UNSAT	0.23	4.75	167.22	0.27	1.76	0.91
3	5-7	UNSAT	0.01	0.89	165.77	0.06	0.32	0.91
3	5-8	UNSAT	0.24	9.14	165.30	0.20	0.81	0.71
3	5-9	UNSAT	0.01	1.05	164.27	0.09	0.32	0.92
4	1-1	UNSAT	0.45	76.15	-	0.81	118.26	4.79
4	1-2	UNSAT	0.46	46.07	-	0.71	95.36	1.98
4	1-3	UNSAT	0.42	36.03	-	1.52	140.03	2.19
4	1-4	UNSAT	0.18	3.23	-	0.26	70.28	0.97
4	1-5	UNSAT	0.31	25.71	-	0.33	57.38	0.88
4	1-6	UNSAT	0.23	12.37	-	0.25	24.60	0.97
4	1-7	SAT	0.16	2.03	164.72	0.07	0.2982	1.44
4	1-8	SAT	0.16	1.78	197.54	0.09	0.4168	1.43
4	1-9	SAT	0.16	1.61	172.78	0.06	0.2737	1.26
4	2-1	UNSAT	0.22	13.81	-	0.77	64.40	0.84
4	2-2	UNSAT	0.21	13.69	-	0.93	-	1.00
4	2-3	UNSAT	0.18	3.03	197.66	0.19	25.76	1.21
4	2-4	UNSAT	0.16	3.26	188.98	0.25	21.84	0.71
4	2-5	UNSAT	0.22	10.80	189.34	0.32	17.26	0.66
4	2-6	UNSAT	0.20	5.24	172.22	0.35	84.10	0.80
4	2-7	UNSAT	0.17	2.27	179.26	0.18	23.61	1.00
4	2-8	UNSAT	0.24	8.02	188.33	1.84	66.73	0.91
4	2-9	UNSAT	0.01	1.36	170.61	0.08	12.69	0.99
4	3-1	UNSAT	0.22	13.12	-	0.36	114.61	1.25

Table 1 Continued

Prop	Net	Result	nnenum	NNV	PeregriNN	MIPVerify	Venus	ERAN
4	3-2	UNSAT	0.24	25.68	179.99	0.24	65.56	0.96
4	3-3	UNSAT	0.15	3.30	169.07	0.14	19.04	0.83
4	3-4	UNSAT	0.15	4.07	180.10	0.23	6.61	0.68
4	3-5	UNSAT	0.22	16.45	183.60	0.53	35.79	0.74
4	3-6	UNSAT	0.17	5.81	-	0.42	143.95	1.0
4	3-7	UNSAT	0.22	3.92	179.41	0.15	23.98	0.73
4	3-8	UNSAT	0.24	2.88	-	0.46	20.13	0.75
4	3-9	UNSAT	0.21	11.37	222.50	0.36	25.69	0.81
4	4-1	UNSAT	0.15	2.59	182.54	0.11	49.35	1.75
4	4-2	UNSAT	0.18	4.61	-	0.46	139.31	0.78
4	4-3	UNSAT	0.25	8.95	-	0.43	48.96	0.96
4	4-4	UNSAT	0.20	4.73	-	2.61	27.50	0.76
4	4-5	UNSAT	0.16	3.95	204.67	0.29	43.19	0.93
4	4-6	UNSAT	0.16	7.13	171.17	0.34	36.77	0.96
4	4-7	UNSAT	0.15	1.37	174.02	0.17	48.39	0.79
4	4-8	UNSAT	0.17	5.72	179.51	0.33	100.79	0.84
4	4-9	UNSAT	0.16	8.01	273.96	0.39	41.78	0.90
4	5-1	UNSAT	0.18	18.68	171.78	0.34	167.89	1.17
4	5-2	UNSAT	0.28	11.32	-	0.25	18.83	1.76
4	5-3	UNSAT	0.17	4.72	227.95	0.60	40.87	0.76
4	5-4	UNSAT	0.17	3.63	166.48	0.26	10.37	0.60
4	5-5	UNSAT	0.18	5.33	195.84	0.28	31.43	0.68
4	5-6	UNSAT	0.17	3.10	208.34	0.23	25.99	0.92
4	5-7	UNSAT	0.16	1.14	-	0.13	73.34	0.85
4	5-8	UNSAT	0.19	3.42	-	0.33	65.65	0.72
4	5-9	UNSAT	0.16	3.63	212.76	0.23	39.06	1.05
5	1-1	UNSAT	1.02	-	130.10	-	-	32.3
6	1-1	UNSAT	5.38	-	-	-	-	10.20
7	1-9	SAT	1.50	-	-	-	-	91.4
8	2-9	SAT	0.22	-	8.45	-	0.73	122
9	3-3	UNSAT	2.52	-	-	-	-	9.20
10	4-5	UNSAT	0.70	-	-	-	130.63	2.07

4.1.2 ACASXU-HARD

Table 2: Tool Runtime (sec) for ACASXU-HARD.

Prop	Net	Result	nnenum	NNV	PeregriNN	MIPVerify	Venus	ERAN
1	4-6	UNSAT	5.30	-	3191.34	-	179.98	5.38
1	4-8	UNSAT	3.96	-	2568.02	-	372.11	3.69
2	3-3	UNSAT	7.46	-	-	-	294.53	167
2	4-2	UNSAT	7.59	-	-	-	648.57	230
2	4-9	SAT	0.17	-	-	37.42	446.13	6.0
2	5-3	SAT	0.85	9302	-	5390.41	9.63	9.7
3	3-6	UNSAT	0.22	7.38	178.48	0.64	3.36	0.72
3	5-1	UNSAT	0.43	35.07	181.43	1.30	27.13	1.78
7	1-9	SAT	1.50	-	-	-	8010.49	91.4
9	3-3	UNSAT	2.52	13326.19	1121.38	-	1795.17	9.21

4.1.3 Pat Fully Connected Networks

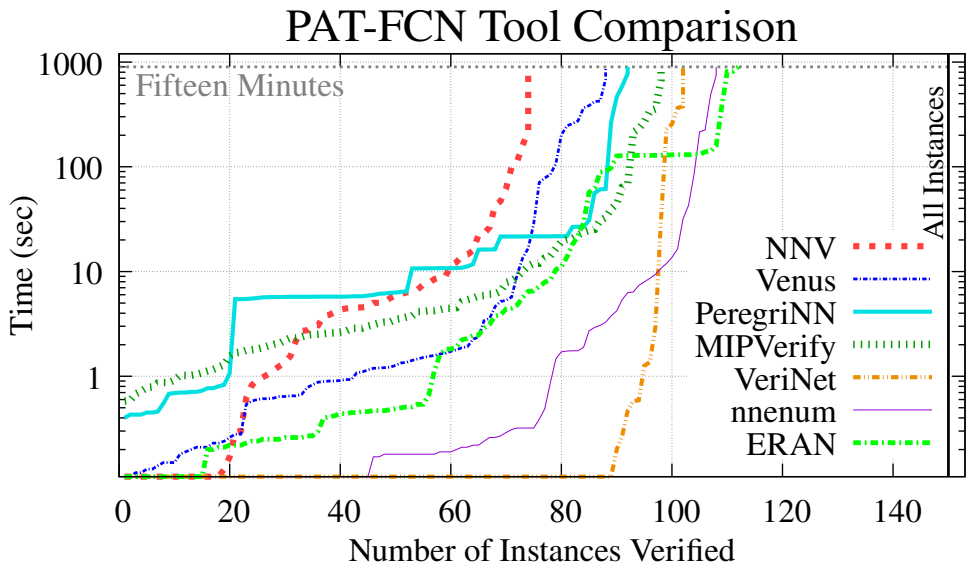

Figure 2: PAT-FCN Cactus Plot.

Table 3: Tool Runtime (sec) for PAT-FCN.

Net	Ep	Im	Result	nnenum	NNV	PeregriNN	VeriNet	MIPVerify	Venus	ERAN
MNIST2	0.02	1	SAT	0.19	-	5.81	0.09	4.31	0.67	0.66
MNIST2	0.02	2	UNSAT	0.02	1.15	5.47	0.01	0.58	0.09	0.04
MNIST2	0.02	3	SAT	2.71	-	30.75	0.02	2.55	0.89	4.39
MNIST2	0.02	4	UNSAT	0.02	1.25	0.45	0.01	0.59	0.61	0.04
MNIST2	0.02	5	SAT	3.78	-	60.80	0.02	0.78	0.64	2.3
MNIST2	0.02	6	UNSAT	2.92	4.43	5.58	0.20	1.02	0.61	0.05
MNIST2	0.02	7	UNSAT	10.72	-	5.76	378.33	3.29	0.91	2.45
MNIST2	0.02	8	UNSAT	3.06	-	5.74	2.94	4.53	1.27	3.13
MNIST2	0.02	9	SAT	0.21	-	5.76	0.02	1.56	0.64	1.11
MNIST2	0.02	10	UNSAT	0.02	0.29	0.45	0.01	0.79	0.11	0.03
MNIST2	0.02	11	UNSAT	0.02	0.10	0.43	0.01	0.68	0.14	0.05
MNIST2	0.02	12	UNSAT	0.02	0.62	0.43	0.01	0.92	0.15	0.03
MNIST2	0.02	13	UNSAT	0.02	0.92	0.43	0.01	1.29	0.12	0.04
MNIST2	0.02	14	UNSAT	0.02	1.05	5.45	0.01	0.87	0.15	0.04
MNIST2	0.02	15	UNSAT	0.02	0.98	5.67	0.01	1.02	0.12	0.05
MNIST2	0.02	16	UNSAT	0.02	0.30	0.40	0.01	1.04	0.58	0.04
MNIST2	0.02	17	UNSAT	0.02	0.82	5.45	0.01	1.08	0.54	0.04
MNIST2	0.02	18	UNSAT	1.74	4.55	5.90	30.92	1.77	0.62	0.05
MNIST2	0.02	19	SAT	0.18	0.00	5.97	0.03	1.16	0.65	1.81

Table 3 Continued

Net	Ep	Im	Result	nnenum	NNV	PeregriNN	VeriNet	MIPVerify	Venus	ERAN
MNIST2	0.02	20	UNSAT	1.47	2.71	5.76	0.59	1.12	0.65	2.47
MNIST2	0.02	21	UNSAT	0.02	1.35	5.52	0.02	1.02	0.60	0.04
MNIST2	0.02	22	UNSAT	7.37	5.57	10.87	1.36	3.57	1.15	4.45
MNIST2	0.02	23	UNSAT	0.02	0.17	0.55	0.02	0.72	0.13	0.04
MNIST2	0.02	24	UNSAT	0.02	0.13	0.46	0.01	0.86	0.13	0.04
MNIST2	0.02	25	SAT	0.20	-	5.79	0.02	7.86	0.83	2.89
MNIST2	0.05	1	SAT	0.19	0.00	10.75	0.02	20.68	1.73	3.75
MNIST2	0.05	2	SAT	0.18	-	5.74	0.02	13.47	0.78	3.86
MNIST2	0.05	3	SAT	0.17	0.00	10.70	0.02	68.44	28.59	900
MNIST2	0.05	4	SAT	0.18	-	10.72	0.02	39.47	5.31	7.21
MNIST2	0.05	5	SAT	0.17	0.01	5.69	0.02	11.25	1.42	7.89
MNIST2	0.05	6	SAT	0.18	-	10.72	0.02	63.70	1.22	140.76
MNIST2	0.05	7	SAT	0.18	0.00	5.73	0.02	5.16	1.46	21.77
MNIST2	0.05	8	SAT	0.18	0.01	10.79	0.02	4.45	1.63	62.5
MNIST2	0.05	9	SAT	0.18	0.00	5.71	0.02	2.50	1.34	10.41
MNIST2	0.05	10	SAT	1.71	-	5.72	0.02	2.52	1.10	1.9
MNIST2	0.05	11	UNSAT	7.96	8.73	-	-	2.86	1.21	4.77
MNIST2	0.05	12	SAT	0.21	-	5.76	0.03	9.95	0.90	0.46
MNIST2	0.05	13	SAT	1.75	0.00	10.78	0.02	1.88	1.58	1.59
MNIST2	0.05	14	SAT	0.29	-	56.08	0.02	30.83	5.71	6.46
MNIST2	0.05	15	SAT	3.28	-	10.78	0.47	341.35	4.62	159.11
MNIST2	0.05	16	UNSAT	490.25	-	-	-	20.20	13.49	11.16
MNIST2	0.05	17	SAT	0.22	-	10.76	0.03	14.08	2.26	8.36
MNIST2	0.05	18	SAT	0.18	0.00	5.75	0.02	2.33	1.51	5.66
MNIST2	0.05	19	SAT	0.18	0.00	5.70	0.02	0.87	0.90	1.82
MNIST2	0.05	20	SAT	1.75	0.00	5.74	0.02	8.84	1.79	88.58
MNIST2	0.05	21	SAT	1.89	0.00	5.76	0.02	1.22	1.55	2.23
MNIST2	0.05	22	SAT	0.19	0.00	5.74	0.02	1.79	1.66	13.76
MNIST2	0.05	23	UNSAT	31.96	-	-	-	6.07	3.57	6.54
MNIST2	0.05	24	SAT	0.22	-	5.81	0.03	32.71	1.19	56.89
MNIST2	0.05	25	SAT	0.18	0.00	10.75	0.02	19.39	1.91	18.12
MNIST4	0.02	1	UNSAT	0.02	15.90	0.77	0.10	2.69	0.18	0.24
MNIST4	0.02	2	UNSAT	0.02	5.91	1.07	0.02	1.67	0.15	0.26
MNIST4	0.02	3	UNSAT	215.64	-	-	-	23.73	382.91	372.43
MNIST4	0.02	4	UNSAT	0.02	3.07	0.83	0.03	1.70	0.21	0.22
MNIST4	0.02	5	UNSAT	0.02	8.44	0.70	0.02	2.64	0.21	0.25
MNIST4	0.02	6	UNSAT	6.28	219.99	16.15	0.27	8.14	1.87	0.27
MNIST4	0.02	7	UNSAT	0.02	14.07	6.40	0.03	2.58	0.21	0.21
MNIST4	0.02	8	UNKNOWN	-	-	-	-	-	-	811
MNIST4	0.02	9	UNSAT	4.21	62.61	11.70	1.26	3.68	1.36	0.26
MNIST4	0.02	10	UNSAT	0.03	6.33	0.80	0.03	1.95	0.19	0.20
MNIST4	0.02	11	UNSAT	0.02	12.24	6.26	0.02	2.60	0.93	0.26
MNIST4	0.02	12	UNSAT	0.02	4.32	0.77	0.02	2.07	0.23	0.20
MNIST4	0.02	13	UNSAT	0.03	11.50	0.69	0.02	2.38	0.28	0.22
MNIST4	0.02	14	UNSAT	0.02	6.36	0.68	0.02	2.02	0.23	0.21

Table 3 Continued

Net	Ep	Im	Result	nnenum	NNV	PeregriNN	VeriNet	MIPVerify	Venus	ERAN
MNIST4	0.02	15	UNSAT	0.03	42.87	6.17	0.02	3.50	0.88	0.30
MNIST4	0.02	16	UNSAT	0.02	14.51	0.71	0.02	2.48	0.24	0.23
MNIST4	0.02	17	UNSAT	5.39	99.55	16.24	-	5.90	2.36	0.25
MNIST4	0.02	18	UNSAT	0.02	4.52	0.70	0.03	1.82	0.20	0.25
MNIST4	0.02	19	SAT	0.26	-	61.39	0.02	22.33	9.35	29.3
MNIST4	0.02	20	UNSAT	0.03	112.49	6.28	0.02	4.14	0.93	0.25
MNIST4	0.02	21	UNSAT	42.75	-	16.28	-	24.65	84.41	97.5
MNIST4	0.02	22	UNSAT	0.03	22.97	6.14	0.03	3.01	1.08	0.27
MNIST4	0.02	23	UNSAT	0.03	8.94	0.73	0.02	2.33	0.28	0.22
MNIST4	0.02	24	UNSAT	0.02	7.73	0.71	0.02	2.21	0.26	0.26
MNIST4	0.02	25	UNSAT	6.37	160.83	16.20	-	11.98	3.39	0.24
MNIST4	0.05	1	UNKNOWN	-	-	-	-	-	-	-
MNIST4	0.05	2	UNSAT	-	-	-	-	214.22	-	811
MNIST4	0.05	3	SAT	-	-	-	0.58	-	-	129
MNIST4	0.05	4	UNSAT	743.72	-	-	-	231.16	-	-
MNIST4	0.05	5	UNKNOWN	-	-	-	-	-	-	-
MNIST4	0.05	6	UNKNOWN	-	-	-	-	-	-	-
MNIST4	0.05	7	SAT	0.27	-	-	0.04	-	-	130
MNIST4	0.05	8	SAT	0.39	-	267.41	0.03	-	273.51	128
MNIST4	0.05	9	SAT	-	-	-	0.03	-	-	139
MNIST4	0.05	10	UNKNOWN	-	-	-	-	-	-	-
MNIST4	0.05	11	UNKNOWN	-	-	-	-	-	-	-
MNIST4	0.05	12	UNSAT	-	-	-	-	671.22	-	-
MNIST4	0.05	13	UNKNOWN	-	-	-	-	-	-	-
MNIST4	0.05	14	UNSAT	-	-	-	-	357.02	-	-
MNIST4	0.05	15	SAT	-	-	588.97	0.04	-	-	130
MNIST4	0.05	16	UNKNOWN	-	-	-	-	-	-	-
MNIST4	0.05	17	SAT	-	-	-	0.06	-	-	-
MNIST4	0.05	18	UNKNOWN	-	-	-	-	-	-	-
MNIST4	0.05	19	SAT	0.26	0.00	11.19	0.04	-	70.97	90.7
MNIST4	0.05	20	SAT	0.49	-	-	0.03	-	-	127
MNIST4	0.05	21	SAT	0.24	-	774.89	0.03	-	-	128
MNIST4	0.05	22	UNKNOWN	-	-	-	-	-	-	-
MNIST4	0.05	23	UNKNOWN	-	-	-	-	-	-	-
MNIST4	0.05	24	UNKNOWN	-	-	-	-	-	-	-
MNIST4	0.05	25	SAT	11.93	-	-	0.04	-	-	-
MNIST6	0.02	1	UNSAT	0.03	5.09	21.57	0.11	5.85	-	0.5
MNIST6	0.02	2	UNSAT	0.03	1.56	6.46	0.03	3.23	1.20	0.45
MNIST6	0.02	3	UNKNOWN	-	-	-	-	-	-	-
MNIST6	0.02	4	UNSAT	0.03	4.51	21.61	0.04	3.30	78.12	0.44
MNIST6	0.02	5	UNSAT	0.03	3.86	21.58	0.03	3.76	419.65	0.51
MNIST6	0.02	6	UNSAT	13.63	46.96	-	-	25.24	-	0.51
MNIST6	0.02	7	UNSAT	225.18	-	-	-	-	-	-
MNIST6	0.02	8	UNSAT	8.63	22.35	26.72	-	12.00	-	0.46
MNIST6	0.02	9	SAT	0.32	0.00	463.74	0.05	268.27	5.18	-

Table 3 Continued

Net	Ep	Im	Result	nnenum	NNV	PeregriNN	VeriNet	MIPVerify	Venus	ERAN
MNIST6	0.02	10	UNSAT	0.03	2.74	21.81	0.03	2.89	423.95	0.41
MNIST6	0.02	11	UNSAT	0.03	2.19	21.61	0.03	2.88	107.56	0.45
MNIST6	0.02	12	UNSAT	0.04	6.73	26.69	0.03	6.11	-	0.42
MNIST6	0.02	13	UNSAT	85.50	-	-	-	44.53	-	-
MNIST6	0.02	14	UNSAT	0.03	5.02	21.72	0.04	4.46	-	0.49
MNIST6	0.02	15	UNSAT	0.04	7.48	21.65	0.03	5.57	-	0.54
MNIST6	0.02	16	UNSAT	0.03	4.10	21.70	0.03	4.23	683.27	0.48
MNIST6	0.02	17	UNKNOWN	-	-	-	-	-	-	-
MNIST6	0.02	18	UNSAT	0.03	5.03	21.66	0.04	4.16	-	0.54
MNIST6	0.02	19	UNKNOWN	-	-	-	-	-	-	-
MNIST6	0.02	20	UNSAT	0.03	4.61	21.66	0.04	3.11	364.69	0.47
MNIST6	0.02	21	UNSAT	9.38	26.42	-	-	16.51	-	0.47
MNIST6	0.02	22	UNSAT	0.03	6.09	26.77	0.04	5.37	-	0.44
MNIST6	0.02	23	UNSAT	0.03	3.78	21.67	0.03	4.22	245.86	0.47
MNIST6	0.02	24	UNSAT	0.03	3.05	21.69	0.03	6.48	257.91	0.4
MNIST6	0.02	25	UNKNOWN	-	-	-	-	-	-	-
MNIST6	0.05	1	SAT	-	-	-	363.99	-	-	-
MNIST6	0.05	2	UNKNOWN	-	-	-	-	-	-	-
MNIST6	0.05	3	SAT	0.32	-	-	0.05	-	-	130.73
MNIST6	0.05	4	UNKNOWN	-	-	-	-	-	-	-
MNIST6	0.05	5	UNKNOWN	-	-	-	-	-	-	-
MNIST6	0.05	6	UNKNOWN	-	-	-	-	-	-	-
MNIST6	0.05	7	SAT	0.32	-	-	0.05	-	-	129.62
MNIST6	0.05	8	SAT	-	-	-	254.71	-	-	127.65
MNIST6	0.05	9	SAT	0.31	0.11	21.61	0.05	-	201.75	128.39
MNIST6	0.05	10	UNKNOWN	-	-	-	-	-	-	-
MNIST6	0.05	11	UNKNOWN	-	-	-	-	-	-	-
MNIST6	0.05	12	SAT	-	-	-	0.05	-	-	128.31
MNIST6	0.05	13	UNKNOWN	-	-	-	-	-	-	128.53
MNIST6	0.05	14	UNKNOWN	-	-	-	-	-	-	-
MNIST6	0.05	15	UNKNOWN	-	-	-	-	-	-	-
MNIST6	0.05	16	UNKNOWN	-	-	-	-	-	-	-
MNIST6	0.05	17	SAT	16.40	-	-	227.87	-	-	130.50
MNIST6	0.05	18	UNKNOWN	-	-	-	-	-	-	130.43
MNIST6	0.05	19	SAT	0.87	-	-	0.05	-	16.67	128.52
MNIST6	0.05	20	UNKNOWN	-	-	-	-	-	-	-
MNIST6	0.05	21	UNKNOWN	-	-	-	-	-	-	-
MNIST6	0.05	22	UNKNOWN	-	-	-	-	-	-	-
MNIST6	0.05	23	UNKNOWN	-	-	-	-	-	-	-
MNIST6	0.05	24	UNKNOWN	-	-	-	-	-	-	-
MNIST6	0.05	25	SAT	0.32	-	-	0.05	-	-	130.29

4.2 CNN

4.2.1 GGN - Convolutional Networks

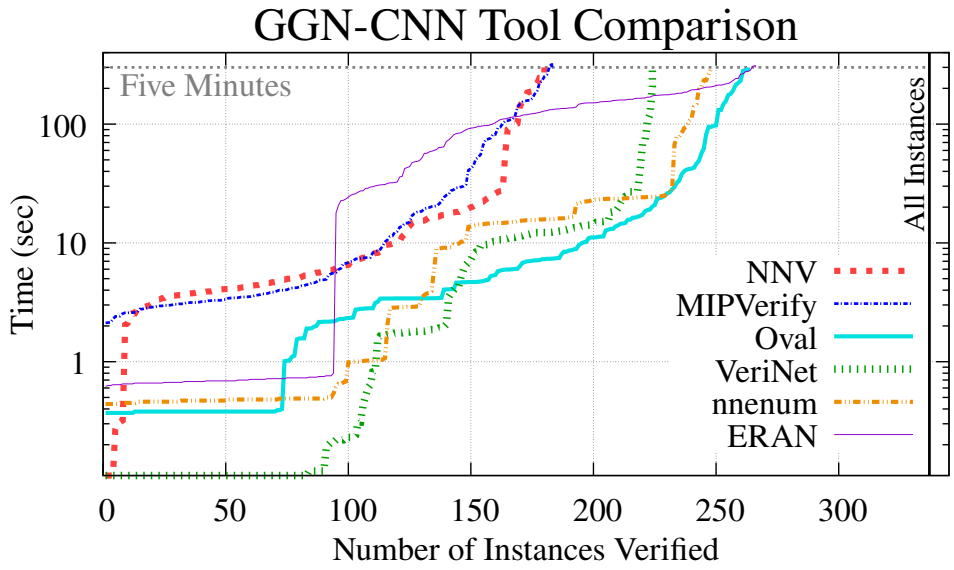

Figure 3: GGN-CNN Cactus Plot.

Table 4: Tool Runtime (sec) for GGN-CNN.

Net	Ep	Image	Result	nnenum	NNV	VeriNet	MIPVerify	Oval	ERAN
MNIST	0.1	1	UNSAT	0.44	5.90	0.19	4.45	0.37	0.64
MNIST	0.1	2	UNSAT	0.45	4.25	0.10	3.04	0.37	0.72
MNIST	0.1	3	UNSAT	0.44	3.64	0.10	2.86	0.37	0.66
MNIST	0.1	4	UNSAT	0.45	3.57	0.10	2.89	0.37	0.66
MNIST	0.1	5	UNSAT	0.48	4.79	0.10	3.72	0.37	0.66
MNIST	0.1	6	UNSAT	0.45	3.07	0.09	2.58	0.37	0.66
MNIST	0.1	7	UNSAT	-	-	-	6.73	24.83	63.67
MNIST	0.1	8	UNSAT	0.46	3.83	0.12	3.14	5.25	0.7
MNIST	0.1	9	UNSAT	-	-	-	6.82	17.26	20.7
MNIST	0.1	10	UNSAT	0.46	4.30	0.12	3.34	0.37	0.65
MNIST	0.1	11	UNSAT	0.47	7.09	0.09	4.61	0.37	0.69
MNIST	0.1	12	UNSAT	0.49	6.96	0.09	4.93	2.76	0.66
MNIST	0.1	13	UNSAT	0.48	5.69	0.09	4.35	2.74	0.69
MNIST	0.1	14	UNSAT	0.48	4.41	0.09	3.63	0.37	0.64
MNIST	0.1	15	UNSAT	0.45	2.11	0.09	2.13	0.37	0.71
MNIST	0.1	16	UNSAT	0.50	6.52	0.09	4.69	1.54	0.76
MNIST	0.1	17	UNSAT	0.48	4.19	0.09	3.15	0.37	0.67
MNIST	0.1	18	UNSAT	0.46	3.66	0.09	3.07	0.37	0.69

Table 4 Continued

Net	Ep	Image	Result	nnenum	NNV	VeriNet	MIPVerify	Oval	ERAN
MNIST	0.1	19	UNSAT	-	-	-	10.21	-	48.6
MNIST	0.1	20	UNSAT	0.47	4.52	0.12	3.51	0.39	0.67
MNIST	0.1	21	UNSAT	0.46	3.63	0.09	2.99	0.38	0.65
MNIST	0.1	22	UNSAT	0.49	4.40	0.09	3.42	0.40	0.73
MNIST	0.1	23	UNSAT	0.49	7.52	0.09	6.94	4.09	0.71
MNIST	0.1	24	UNSAT	0.47	4.89	0.10	3.84	0.39	0.63
MNIST	0.1	25	UNSAT	0.46	5.62	0.10	4.00	5.35	0.66
MNIST	0.1	26	UNSAT	0.49	5.43	0.09	4.03	0.38	0.75
MNIST	0.1	27	UNSAT	0.46	3.44	0.09	2.77	0.38	0.65
MNIST	0.1	28	UNSAT	0.47	3.82	0.09	2.92	0.38	0.69
MNIST	0.1	29	UNSAT	0.47	3.77	0.09	3.26	0.38	0.66
MNIST	0.1	30	UNSAT	0.44	2.55	0.09	2.28	0.38	0.71
MNIST	0.1	31	UNSAT	0.47	4.09	0.09	3.45	0.38	0.68
MNIST	0.1	32	UNSAT	0.44	2.98	0.09	2.60	0.38	0.67
MNIST	0.1	33	UNSAT	0.49	3.96	0.09	3.45	0.38	0.7
MNIST	0.1	34	UNSAT	0.49	5.31	0.10	4.36	0.38	0.68
MNIST	0.1	35	UNSAT	0.48	3.59	0.09	2.98	0.38	0.69
MNIST	0.1	36	UNSAT	0.48	5.03	0.10	4.00	0.38	0.68
MNIST	0.1	37	UNSAT	0.47	4.98	0.09	3.53	0.38	0.76
MNIST	0.1	38	UNSAT	0.44	2.94	0.09	2.40	0.38	0.64
MNIST	0.1	39	UNSAT	0.48	5.53	0.10	4.12	1.57	0.65
MNIST	0.1	40	UNSAT	0.46	3.33	0.09	2.80	0.38	0.72
MNIST	0.1	41	UNSAT	0.46	2.86	0.10	2.58	0.38	0.64
MNIST	0.1	42	UNSAT	0.47	4.14	0.09	3.29	0.38	0.62
MNIST	0.1	43	UNSAT	0.47	5.03	0.09	3.95	0.38	0.69
MNIST	0.1	44	UNSAT	0.47	4.15	0.09	3.29	4.06	0.74
MNIST	0.1	45	UNSAT	0.48	3.91	0.09	3.22	0.38	0.63
MNIST	0.1	46	UNSAT	0.49	6.91	0.09	5.57	4.09	0.66
MNIST	0.1	47	UNSAT	0.46	2.85	0.09	2.41	0.38	0.74
MNIST	0.1	48	UNSAT	0.49	5.84	0.10	4.89	4.04	0.68
MNIST	0.1	49	UNSAT	0.46	3.42	0.10	2.95	0.38	0.68
MNIST	0.1	50	UNSAT	0.49	4.48	0.10	3.49	0.38	0.71
MNIST	0.1	51	UNSAT	0.49	5.06	0.10	4.00	0.38	0.73
MNIST	0.1	52	UNSAT	0.49	5.92	0.09	4.25	0.38	0.69
MNIST	0.1	53	UNSAT	0.48	4.11	0.10	3.26	0.38	0.66
MNIST	0.1	54	UNSAT	0.49	5.91	0.09	4.48	2.82	0.72
MNIST	0.1	55	UNSAT	0.49	6.30	0.10	4.88	0.38	0.71
MNIST	0.1	56	UNSAT	0.47	4.60	0.10	3.66	0.38	0.65
MNIST	0.1	57	UNSAT	0.49	4.70	0.09	3.44	0.38	0.67
MNIST	0.1	58	UNSAT	0.44	2.04	0.10	2.14	0.38	0.7
MNIST	0.1	59	UNSAT	0.48	5.70	0.10	4.18	1.57	0.67
MNIST	0.1	60	UNSAT	9.83	4.99	0.43	3.84	8.03	0.66
MNIST	0.1	61	UNSAT	0.45	3.15	0.12	2.62	0.38	0.74
MNIST	0.1	62	UNSAT	0.48	5.17	0.09	3.72	0.38	0.73
MNIST	0.1	63	UNSAT	-	-	-	41.00	96.66	93.73

Table 4 Continued

Net	Ep	Image	Result	nnenum	NNV	VeriNet	MIPVerify	Oval	ERAN
MNIST	0.1	64	UNSAT	0.47	4.00	0.12	3.29	0.38	0.67
MNIST	0.1	65	UNSAT	0.47	4.67	0.10	3.75	0.38	0.76
MNIST	0.1	66	UNSAT	9.69	5.51	24.58	12.53	6.78	22.82
MNIST	0.1	67	UNSAT	0.49	6.42	0.12	4.59	2.82	0.67
MNIST	0.1	68	UNSAT	0.49	6.18	0.09	5.50	4.06	0.73
MNIST	0.1	69	UNSAT	0.48	4.10	0.10	3.28	0.38	0.69
MNIST	0.1	70	UNSAT	0.49	4.90	0.09	3.81	0.38	0.73
MNIST	0.1	71	UNSAT	0.46	3.35	0.10	2.77	0.38	0.67
MNIST	0.1	72	UNSAT	0.46	4.06	0.09	3.17	0.38	0.69
MNIST	0.1	73	UNSAT	0.49	5.01	0.10	3.52	0.38	0.72
MNIST	0.1	74	UNSAT	0.46	3.57	0.09	2.59	0.38	0.68
MNIST	0.1	75	UNSAT	0.45	2.81	0.09	2.59	0.38	0.68
MNIST	0.1	76	UNSAT	0.48	4.26	0.09	3.46	0.38	0.71
MNIST	0.1	77	UNSAT	0.46	3.75	0.09	3.15	0.38	0.72
MNIST	0.1	78	UNSAT	0.48	5.30	0.09	4.20	1.57	0.7
MNIST	0.1	79	UNSAT	0.46	4.09	0.10	3.33	2.79	0.66
MNIST	0.1	80	UNSAT	0.47	3.76	0.09	3.04	0.38	0.75
MNIST	0.1	81	UNSAT	0.46	2.89	0.10	2.54	0.38	0.69
MNIST	0.1	82	UNSAT	0.48	3.88	0.09	3.08	0.38	0.69
MNIST	0.1	83	UNSAT	0.48	4.07	0.09	2.98	0.38	0.73
MNIST	0.1	84	UNSAT	0.47	4.38	0.09	3.20	0.38	0.71
MNIST	0.1	85	UNSAT	0.50	6.46	0.10	4.88	2.81	0.73
MNIST	0.1	86	UNSAT	0.47	3.60	0.10	2.84	0.38	0.66
MNIST	0.1	87	UNSAT	0.47	3.73	0.10	3.07	0.38	0.72
MNIST	0.1	88	UNSAT	0.48	4.60	0.09	3.50	0.38	0.68
MNIST	0.1	89	UNSAT	0.48	4.67	0.10	3.61	0.38	0.72
MNIST	0.1	90	UNSAT	0.47	3.91	0.09	2.88	2.81	0.74
MNIST	0.1	91	UNSAT	0.46	3.62	0.10	2.95	0.38	0.72
MNIST	0.1	92	UNSAT	0.48	4.61	0.09	3.56	0.38	0.7
MNIST	0.1	93	UNSAT	-	-	-	12.34	-	63.67
MNIST	0.1	94	UNSAT	10.24	4.00	0.31	3.15	2.79	0.73
MNIST	0.1	95	UNSAT	0.45	2.37	0.11	2.12	0.38	0.69
MNIST	0.1	96	UNSAT	10.55	7.31	0.31	4.16	5.32	0.75
MNIST	0.1	97	UNSAT	0.46	2.55	0.12	2.42	0.38	0.73
MNIST	0.1	98	UNSAT	0.47	3.82	0.09	3.00	0.38	0.69
MNIST	0.1	99	UNSAT	0.48	4.23	0.10	3.21	0.38	0.68
MNIST	0.1	100	UNSAT	0.48	3.83	0.09	3.13	0.38	0.8
MNIST	0.3	1	UNSAT	14.72	18.23	12.23	20.21	10.93	31.73
MNIST	0.3	2	UNSAT	15.53	17.52	-	19.35	16.86	105.99
MNIST	0.3	3	UNSAT	13.80	10.25	-	9.41	11.11	17.89
MNIST	0.3	4	UNSAT	15.47	16.64	-	42.15	32.14	136.71
MNIST	0.3	5	UNSAT	15.57	17.35	-	-	41.81	103.06
MNIST	0.3	6	UNSAT	13.66	8.07	21.79	7.49	7.01	22.74
MNIST	0.3	7	UNKNOWN	-	-	-	-	-	-
MNIST	0.3	8	UNSAT	-	-	-	-	-	181.25

Table 4 Continued

Net	Ep	Image	Result	nnenum	NNV	VeriNet	MIPVerify	Oval	ERAN
MNIST	0.3	9	SAT	2.08	-	0.22	29.93	2.00	150.31
MNIST	0.3	10	UNKNOWN	-	-	-	-	-	-
MNIST	0.3	11	UNSAT	15.53	20.89	23.35	20.50	16.35	23.47
MNIST	0.3	12	UNKNOWN	-	-	-	-	-	-
MNIST	0.3	13	UNSAT	-	-	-	-	283.06	155.56
MNIST	0.3	14	UNSAT	16.29	24.33	-	27.59	131.34	165.49
MNIST	0.3	15	UNSAT	3.94	5.62	0.20	5.37	7.13	31.03
MNIST	0.3	16	SAT	0.65	-	-	327.26	221.91	213.17
MNIST	0.3	17	UNSAT	-	-	-	148.17	97.55	110.81
MNIST	0.3	18	UNSAT	14.54	16.45	5.95	14.99	29.43	42.33
MNIST	0.3	19	SAT	0.54	-	0.24	74.91	1.91	161.25
MNIST	0.3	20	UNSAT	-	16.35	-	158.85	-	114.03
MNIST	0.3	21	UNSAT	15.53	15.85	-	-	42.13	135.29
MNIST	0.3	22	UNSAT	14.58	16.60	-	18.15	15.72	43.04
MNIST	0.3	23	UNKNOWN	-	-	-	-	-	-
MNIST	0.3	24	UNSAT	16.13	18.74	-	52.60	33.46	135.46
MNIST	0.3	25	UNKNOWN	-	-	-	-	-	-
MNIST	0.3	26	UNSAT	15.63	15.08	8.41	17.02	8.68	24.29
MNIST	0.3	27	UNSAT	14.67	14.74	-	19.85	14.45	95.8
MNIST	0.3	28	UNSAT	-	18.23	-	101.23	156.27	139.85
MNIST	0.3	29	UNSAT	16.01	19.44	-	25.22	42.97	-
MNIST	0.3	30	UNSAT	13.92	8.21	-	7.32	11.16	32.1
MNIST	0.3	31	UNSAT	-	-	-	-	-	150.79
MNIST	0.3	32	UNSAT	14.49	9.59	-	8.58	20.88	48.11
MNIST	0.3	33	UNSAT	14.74	15.31	4.25	13.80	16.31	26.7
MNIST	0.3	34	SAT	0.50	0.01	0.23	109.52	1.90	114.56
MNIST	0.3	35	UNSAT	14.88	13.87	-	14.68	11.29	30.52
MNIST	0.3	36	UNSAT	21.41	22.05	-	55.67	37.79	177.15
MNIST	0.3	37	UNSAT	-	-	-	179.30	213.53	125.33
MNIST	0.3	38	UNSAT	3.48	7.34	0.21	7.37	7.11	29.54
MNIST	0.3	39	UNKNOWN	-	-	-	-	-	-
MNIST	0.3	40	UNSAT	3.63	8.07	0.21	5.95	5.85	30.22
MNIST	0.3	41	UNSAT	6.11	6.27	0.51	6.12	5.77	29.72
MNIST	0.3	42	UNSAT	-	19.31	-	-	61.99	150.87
MNIST	0.3	43	UNSAT	15.79	27.37	-	29.24	46.38	-
MNIST	0.3	44	SAT	0.68	-	0.23	138.88	41.40	119.86
MNIST	0.3	45	UNSAT	-	-	-	116.54	-	-
MNIST	0.3	46	UNKNOWN	-	-	-	-	-	-
MNIST	0.3	47	UNSAT	14.31	8.23	0.64	7.53	8.62	28.83
MNIST	0.3	48	UNKNOWN	-	-	-	-	-	311
MNIST	0.3	49	UNSAT	14.35	10.09	-	8.28	18.49	56.73
MNIST	0.3	50	UNSAT	14.86	15.27	-	17.65	26.56	129.12
MNIST	0.3	51	UNSAT	16.61	17.20	-	230.23	82.46	124.42
MNIST	0.3	52	UNSAT	16.07	19.94	-	221.21	23.78	112.02
MNIST	0.3	53	UNSAT	296.55	24.30	-	-	280.84	-

Table 4 Continued

Net	Ep	Image	Result	nnenum	NNV	VeriNet	MIPVerify	Oval	ERAN
MNIST	0.3	54	UNKNOWN	-	-	-	-	-	-
MNIST	0.3	55	UNSAT	15.89	27.60	-	31.07	20.27	133.36
MNIST	0.3	56	UNSAT	15.43	13.37	15.22	14.62	13.12	32.24
MNIST	0.3	57	UNSAT	15.77	18.06	-	19.68	193.27	98.91
MNIST	0.3	58	UNSAT	14.27	7.73	-	6.70	10.14	51.8
MNIST	0.3	59	UNSAT	-	-	-	50.61	-	-
MNIST	0.3	60	UNKNOWN	-	-	-	-	-	-
MNIST	0.3	61	UNSAT	3.75	9.00	0.22	7.51	7.02	28.77
MNIST	0.3	62	UNKNOWN	-	-	-	-	-	-
MNIST	0.3	63	SAT	0.67	-	204.86	44.19	3.19	116.05
MNIST	0.3	64	SAT	0.65	-	0.23	76.05	1.90	185.65
MNIST	0.3	65	UNKNOWN	-	-	-	-	-	-
MNIST	0.3	67	UNKNOWN	-	-	-	-	-	-
MNIST	0.3	68	UNSAT	-	23.21	-	-	231.76	134.39
MNIST	0.3	69	UNSAT	-	-	-	155.62	-	-
MNIST	0.3	70	UNSAT	-	26.08	-	-	-	-
MNIST	0.3	71	UNSAT	3.78	9.95	0.22	8.45	7.08	27.45
MNIST	0.3	72	UNSAT	16.39	21.65	7.25	24.15	14.04	200.24
MNIST	0.3	73	UNKNOWN	-	-	-	-	-	194.04
MNIST	0.3	74	SAT	0.49	0.01	0.22	21.24	1.97	120.04
MNIST	0.3	75	UNSAT	14.66	9.86	0.60	8.78	8.43	30.92
MNIST	0.3	76	UNKNOWN	-	-	-	-	-	-
MNIST	0.3	77	UNSAT	-	15.37	-	25.63	94.92	59.07
MNIST	0.3	78	UNKNOWN	-	-	-	-	-	-
MNIST	0.3	79	UNSAT	-	-	-	158.51	-	-
MNIST	0.3	80	UNSAT	14.72	9.66	0.83	7.54	6.89	30.79
MNIST	0.3	81	UNSAT	-	-	-	18.28	-	-
MNIST	0.3	82	UNSAT	15.81	18.04	-	28.04	56.70	129.72
MNIST	0.3	83	UNSAT	16.08	15.42	0.82	13.58	8.44	25.93
MNIST	0.3	84	UNSAT	15.20	11.87	120.14	10.63	11.25	25.22
MNIST	0.3	85	UNKNOWN	-	-	-	-	-	-
MNIST	0.3	86	UNSAT	15.75	16.37	3.52	11.46	11.33	27.45
MNIST	0.3	87	UNSAT	14.78	11.34	-	11.34	13.74	35.83
MNIST	0.3	88	UNSAT	-	-	-	-	179.91	158.34
MNIST	0.3	89	UNSAT	15.77	22.26	-	22.50	26.06	93.01
MNIST	0.3	90	UNSAT	14.99	9.70	-	7.41	9.59	31.96
MNIST	0.3	91	UNSAT	16.51	19.06	-	260.92	18.90	174.56
MNIST	0.3	92	UNSAT	15.75	20.27	-	29.64	176.26	159.71
MNIST	0.3	93	SAT	0.61	-	0.23	66.66	3.13	137.32
MNIST	0.3	94	UNSAT	-	-	-	73.90	-	-
MNIST	0.3	95	UNSAT	14.13	7.51	67.79	5.63	11.11	25.66
MNIST	0.3	96	UNSAT	-	-	-	190.09	-	205.7
MNIST	0.3	97	UNSAT	14.98	8.45	-	7.06	19.05	43.92
MNIST	0.3	98	UNSAT	-	-	-	108.55	94.82	89.7
MNIST	0.3	99	UNKNOWN	-	-	-	-	-	204.6

Table 4 Continued

Net	Ep	Image	Result	nnenum	NNV	VeriNet	MIPVerify	Oval	ERAN
MNIST	0.3	100	UNSAT	15.28	17.10	-	18.49	14.72	73.68
CIFAR10	2/255	1	SAT	2.89	-	9.70	-	4.63	242.25
CIFAR10	2/255	2	UNSAT	24.19	-	7.96	-	3.45	155.93
CIFAR10	2/255	3	UNKNOWN	-	-	-	-	-	-
CIFAR10	2/255	4	UNSAT	207.14	-	18.21	-	8.89	155.06
CIFAR10	2/255	5	UNKNOWN	-	-	-	-	-	-
CIFAR10	2/255	6	UNSAT	-	-	71.23	-	-	284.65
CIFAR10	2/255	7	UNSAT	233.14	-	19.63	-	7.44	164.83
CIFAR10	2/255	9	UNKNOWN	-	-	-	-	-	-
CIFAR10	2/255	10	UNKNOWN	-	-	-	-	-	-
CIFAR10	2/255	11	UNSAT	196.77	-	15.32	-	16.57	165.11
CIFAR10	2/255	12	UNSAT	24.54	-	9.65	-	3.42	176.56
CIFAR10	2/255	13	SAT	2.85	-	13.84	-	4.67	241.88
CIFAR10	2/255	14	UNSAT	23.62	-	12.19	-	3.42	122.48
CIFAR10	2/255	15	UNSAT	24.46	-	13.40	-	3.42	204.69
CIFAR10	2/255	16	UNSAT	23.06	-	12.28	-	8.83	196.5
CIFAR10	2/255	17	UNKNOWN	-	-	-	-	-	-
CIFAR10	2/255	18	UNSAT	23.84	-	13.54	-	24.17	218.16
CIFAR10	2/255	19	UNSAT	23.38	-	11.43	-	7.39	147.51
CIFAR10	2/255	20	UNSAT	27.29	-	14.13	-	3.42	195.29
CIFAR10	2/255	21	UNSAT	233.86	-	25.12	-	13.12	150.84
CIFAR10	2/255	22	UNSAT	22.40	56.66	6.88	-	3.39	58.33
CIFAR10	2/255	24	UNSAT	24.21	-	12.76	-	3.41	178.68
CIFAR10	2/255	26	UNKNOWN	-	-	-	-	-	-
CIFAR10	2/255	27	SAT	2.85	-	15.14	-	4.69	-
CIFAR10	2/255	28	UNKNOWN	-	-	-	-	-	-
CIFAR10	2/255	29	SAT	2.87	-	12.58	-	4.70	213.14
CIFAR10	2/255	30	UNSAT	26.10	-	14.16	-	28.19	211.52
CIFAR10	2/255	31	UNKNOWN	-	-	-	-	-	-
CIFAR10	2/255	32	SAT	2.89	-	14.35	-	4.67	207.81
CIFAR10	2/255	33	SAT	2.87	-	14.85	-	4.68	276.66
CIFAR10	2/255	34	UNKNOWN	-	-	-	-	-	-
CIFAR10	2/255	35	UNSAT	24.24	-	12.41	-	3.44	175.09
CIFAR10	2/255	37	SAT	2.90	-	15.12	-	4.87	263.44
CIFAR10	2/255	38	SAT	2.89	-	11.47	-	4.77	190.11
CIFAR10	2/255	39	UNSAT	23.92	-	12.91	-	3.43	215.07
CIFAR10	2/255	40	UNSAT	-	-	41.37	-	-	277.17
CIFAR10	2/255	41	UNSAT	24.42	-	12.64	-	7.37	176.02
CIFAR10	2/255	42	UNSAT	24.33	-	12.31	-	3.42	157.77
CIFAR10	2/255	43	UNKNOWN	-	-	-	-	288.87	-
CIFAR10	2/255	44	UNSAT	26.89	-	14.87	-	6.03	182.36
CIFAR10	2/255	45	UNSAT	24.36	-	10.62	-	3.40	127.19
CIFAR10	2/255	46	UNSAT	23.45	-	11.16	-	3.41	176.48
CIFAR10	2/255	47	UNSAT	-	-	27.41	-	23.66	168.52
CIFAR10	2/255	49	UNKNOWN	-	-	-	-	-	-

Table 4 Continued

Net	Ep	Image	Result	nnenum	NNV	VeriNet	MIPVerify	Oval	ERAN
CIFAR10	2/255	51	UNKNOWN	-	-	-	-	-	-
CIFAR10	2/255	52	UNKNOWN	-	-	-	-	-	-
CIFAR10	2/255	54	UNKNOWN	-	-	-	-	-	-
CIFAR10	2/255	55	UNSAT	23.82	-	10.80	-	3.41	154.69
CIFAR10	2/255	56	UNSAT	22.35	-	10.65	-	3.40	148.47
CIFAR10	2/255	61	UNSAT	23.78	-	13.06	-	3.42	165.92
CIFAR10	2/255	63	UNSAT	24.63	-	11.31	-	6.05	152.65
CIFAR10	2/255	65	UNKNOWN	-	-	-	-	-	-
CIFAR10	2/255	66	SAT	2.81	-	13.41	-	4.69	265.71
CIFAR10	2/255	67	UNKNOWN	-	-	-	-	-	-
CIFAR10	2/255	68	UNSAT	22.23	-	10.12	-	3.41	137.36
CIFAR10	2/255	69	UNKNOWN	-	-	-	-	-	-
CIFAR10	2/255	72	UNKNOWN	-	-	-	-	-	-
CIFAR10	2/255	73	UNSAT	23.56	-	9.58	-	3.40	123.07
CIFAR10	2/255	74	UNSAT	23.49	-	12.03	-	3.41	159.44
CIFAR10	2/255	75	SAT	2.93	0.31	12.18	-	4.68	201.93
CIFAR10	2/255	76	UNSAT	23.62	-	11.16	-	3.41	151.24
CIFAR10	2/255	77	SAT	2.86	-	12.46	-	4.74	-
CIFAR10	2/255	78	UNKNOWN	-	-	-	-	-	-
CIFAR10	2/255	79	UNSAT	263.65	-	27.22	-	10.15	159.36
CIFAR10	2/255	80	UNSAT	23.20	-	9.53	-	3.40	110.23
CIFAR10	2/255	81	UNSAT	23.26	-	11.73	-	5.97	168.94
CIFAR10	2/255	82	UNSAT	21.28	-	10.26	-	3.41	115.31
CIFAR10	2/255	83	UNSAT	24.87	-	12.22	-	3.41	161.86
CIFAR10	2/255	84	UNSAT	22.72	-	10.25	-	3.41	134.15
CIFAR10	2/255	85	UNSAT	22.09	-	10.59	-	6.01	132.31
CIFAR10	2/255	87	SAT	2.92	-	11.85	-	7.35	-
CIFAR10	2/255	89	UNSAT	23.55	-	11.72	-	7.29	178.98
CIFAR10	2/255	90	UNSAT	23.54	-	12.22	-	7.28	219.26
CIFAR10	2/255	91	UNSAT	24.11	-	11.15	-	3.40	159.58
CIFAR10	2/255	93	UNSAT	23.47	-	10.91	-	3.41	135.76
CIFAR10	2/255	94	UNSAT	-	-	26.90	-	-	233.26
CIFAR10	2/255	95	UNSAT	-	-	26.94	-	12.92	194.57
CIFAR10	2/255	96	SAT	2.87	-	13.20	-	4.65	-
CIFAR10	2/255	97	UNSAT	25.00	-	13.98	-	5.96	225.1
CIFAR10	2/255	98	UNSAT	140.53	204.61	18.30	-	5.93	92.31
CIFAR10	2/255	99	UNSAT	24.15	189.51	8.34	-	3.41	83.18
CIFAR10	2/255	100	UNSAT	23.58	-	12.21	-	5.97	172.84
CIFAR10	8/255	2	UNSAT	78.24	141.58	6.53	299.96	49.32	95.78
CIFAR10	8/255	3	UNKNOWN	-	-	-	-	-	-
CIFAR10	8/255	5	SAT	-	-	1.85	-	-	-
CIFAR10	8/255	6	UNKNOWN	-	-	-	-	-	-
CIFAR10	8/255	10	SAT	0.99	-	1.77	-	2.29	150.54
CIFAR10	8/255	11	UNKNOWN	-	-	-	-	-	-
CIFAR10	8/255	12	UNSAT	-	247.23	-	-	40.10	185.85

Table 4 Continued

Net	Ep	Image	Result	nnenum	NNV	VeriNet	MIPVerify	Oval	ERAN
CIFAR10	8/255	14	UNSAT	67.14	90.56	5.19	81.92	9.89	49.33
CIFAR10	8/255	15	UNKNOWN	-	-	-	-	-	-
CIFAR10	8/255	16	UNKNOWN	-	-	-	-	-	-
CIFAR10	8/255	17	UNSAT	9.03	153.17	1.80	-	12.74	61.13
CIFAR10	8/255	19	UNSAT	9.19	94.27	1.74	94.44	1.02	60.02
CIFAR10	8/255	20	SAT	1.02	-	1.88	-	2.17	306.36
CIFAR10	8/255	22	UNSAT	8.97	23.90	1.53	20.32	1.01	32.58
CIFAR10	8/255	24	SAT	1.03	-	1.81	-	2.17	116.74
CIFAR10	8/255	29	UNKNOWN	-	-	-	-	-	-
CIFAR10	8/255	30	UNSAT	84.17	187.26	3.80	-	14.56	154.32
CIFAR10	8/255	31	SAT	1.07	-	2.06	-	2.16	100.2
CIFAR10	8/255	33	SAT	-	-	1.75	-	131.60	-
CIFAR10	8/255	34	SAT	1.02	-	1.80	-	2.27	91.46
CIFAR10	8/255	35	UNSAT	95.32	244.93	4.30	-	15.66	121.39
CIFAR10	8/255	39	SAT	106.72	-	1.82	-	-	-
CIFAR10	8/255	40	SAT	1.00	-	1.78	-	2.18	181.87
CIFAR10	8/255	42	UNSAT	9.12	98.35	1.92	92.68	1.10	47.41
CIFAR10	8/255	45	SAT	1.04	-	1.69	164.30	7.34	-
CIFAR10	8/255	46	UNKNOWN	-	-	-	-	-	-
CIFAR10	8/255	50	UNSAT	92.53	184.63	18.10	-	-	84.62
CIFAR10	8/255	51	UNSAT	10.27	165.69	1.96	153.49	7.44	71.34
CIFAR10	8/255	52	UNKNOWN	-	-	-	-	-	-
CIFAR10	8/255	55	UNSAT	78.35	220.51	4.01	-	6.37	81.19
CIFAR10	8/255	56	UNSAT	148.05	-	105.64	-	-	132.61
CIFAR10	8/255	61	SAT	1.02	0.27	1.81	-	2.34	161.75
CIFAR10	8/255	65	UNKNOWN	-	-	-	-	-	-
CIFAR10	8/255	67	UNKNOWN	-	-	-	-	-	-
CIFAR10	8/255	68	SAT	0.99	0.31	1.78	-	2.32	109.07
CIFAR10	8/255	72	UNKNOWN	-	-	-	-	-	-
CIFAR10	8/255	73	UNKNOWN	-	-	-	-	-	-
CIFAR10	8/255	74	UNKNOWN	-	-	-	-	-	-
CIFAR10	8/255	75	SAT	1.04	-	1.92	251.68	2.35	87.26
CIFAR10	8/255	77	UNKNOWN	-	-	-	-	-	-
CIFAR10	8/255	80	SAT	1.01	0.01	1.64	105.95	2.33	77.08
CIFAR10	8/255	81	UNKNOWN	-	-	-	-	-	-
CIFAR10	8/255	82	UNSAT	9.72	97.45	1.86	106.86	4.63	64.46
CIFAR10	8/255	83	UNSAT	9.15	-	1.81	-	7.16	84.02
CIFAR10	8/255	84	SAT	1.00	0.32	1.66	-	2.16	96.95
CIFAR10	8/255	85	UNKNOWN	-	-	-	-	-	-
CIFAR10	8/255	87	SAT	1.02	-	1.86	-	2.29	164.94
CIFAR10	8/255	89	UNSAT	-	244.84	-	-	147.67	97.42
CIFAR10	8/255	90	SAT	111.75	-	1.80	-	4.86	118.8
CIFAR10	8/255	91	SAT	1.00	-	1.75	-	2.18	174.29
CIFAR10	8/255	93	UNSAT	9.11	100.68	1.70	82.04	1.02	56.57
CIFAR10	8/255	96	SAT	1.00	0.01	1.75	-	2.17	-

Table 4 Continued

Net	Ep	Image	Result	nnenum	NNV	VeriNet	MIPVerify	Oval	ERAN
CIFAR10	8/255	97	SAT	-	-	14.08	-	-	-
CIFAR10	8/255	98	UNKNOWN	-	-	-	-	-	-
CIFAR10	8/255	99	UNSAT	8.91	26.79	1.64	29.06	1.12	36.38
CIFAR10	8/255	100	SAT	1.02	-	1.71	-	2.22	97.06

4.2.2 OVAL - Convolutional Networks

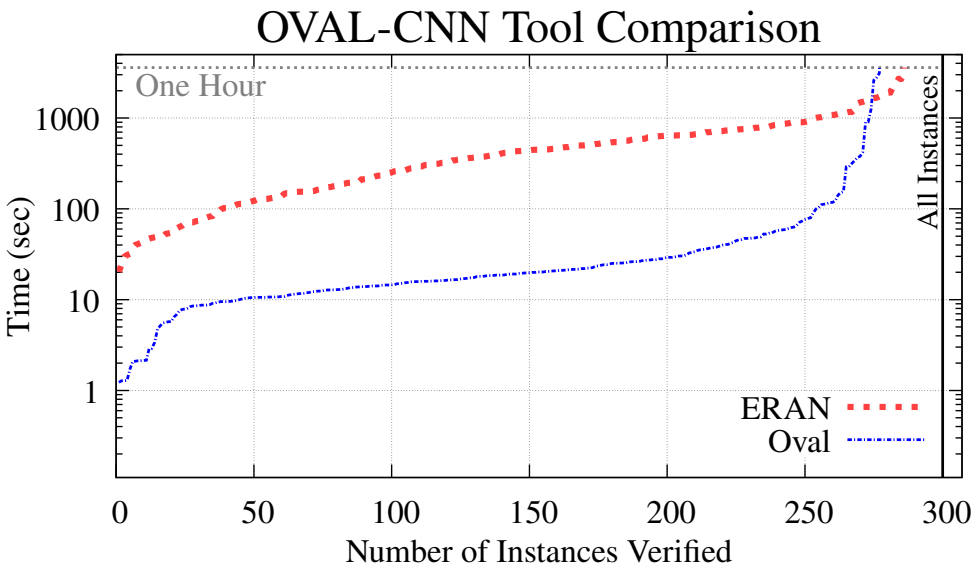

Figure 4: OVAL-CNN Cactus Plot.

Table 5: Tool Runtime (sec) for OVAL-CNN.

Net	Image	Ep	Result	Oval	ERAN
base	7400	0.147	UNSAT	37.94	661.37
base	4522	0.122	UNSAT	-	54.37
base	4425	0.241	UNSAT	289.09	893.12
base	362	0.235	UNSAT	19.95	447.55
base	4461	0.210	UNSAT	19.74	-
base	6546	0.172	UNSAT	14.43	120.18
base	349	0.197	UNSAT	99.79	71.28
base	9441	0.185	UNSAT	13.59	564.85
base	5402	0.228	UNSAT	22.57	616.37
base	5569	0.090	UNSAT	21.52	-
base	6548	0.203	UNSAT	-	211.7
base	3510	0.160	UNSAT	25.22	920.7
base	6524	0.122	UNSAT	2.15	1111.64
base	559	0.172	UNSAT	-	1039.06
base	6037	0.210	UNSAT	10.74	-
base	5501	0.122	UNSAT	25.08	180.31
base	3410	0.122	UNSAT	-	98.32
base	2510	0.178	UNSAT	-	732.31
base	3374	0.122	UNSAT	18.07	1538.78

Table 5 Continued

Net	Image	Ep	Result	Oval	ERAN
base	3057	0.185	UNSAT	6.44	52.89
base	6527	0.225	UNSAT	12.23	873.27
base	7166	0.164	UNSAT	9.10	649.78
base	409	0.185	UNSAT	9.21	352.99
base	2503	0.185	UNSAT	-	174.13
base	4018	0.220	UNSAT	421.14	153.8
base	7467	0.147	UNSAT	10.68	252.36
base	8314	0.072	UNSAT	23.51	756.01
base	350	0.178	UNSAT	366.97	127.03
base	9091	0.177	UNSAT	16.46	213.49
base	6519	0.197	UNSAT	13.77	1110.12
base	4478	0.210	UNSAT	13.93	-
base	6324	0.235	UNSAT	15.93	842.02
base	5021	0.205	UNSAT	11.98	1660.64
base	4567	0.128	UNSAT	10.81	445.21
base	4501	0.253	UNSAT	6.87	1783.64
base	4545	0.172	UNSAT	15.92	642.72
base	3556	0.141	UNSAT	377.81	1448.42
base	8348	0.197	UNSAT	25.30	867.98
base	7476	0.172	UNSAT	1.28	2693.77
base	146	0.138	UNSAT	27.65	791.22
base	5304	0.090	UNSAT	7.81	128.29
base	6313	0.210	UNSAT	-	122.47
base	2335	0.228	UNSAT	13.34	503.77
base	3370	0.253	UNSAT	16.91	750.35
base	4440	0.061	UNSAT	7.89	528.8
base	3436	0.128	UNSAT	-	39.99
base	345	0.250	UNSAT	-	132.21
base	5352	0.172	UNSAT	22.19	-
base	3552	0.172	UNSAT	47.29	1234.8
base	5526	0.210	UNSAT	18.06	2706.16
base	5463	0.100	UNSAT	14.14	1079.72
base	3141	0.210	UNSAT	908.04	73.74
base	1030	0.164	UNSAT	10.84	397.17
base	4455	0.200	UNSAT	10.07	1112.86
base	4451	0.235	UNSAT	13.79	273.88
base	4505	0.197	UNSAT	15.50	652.46
base	3529	0.122	UNSAT	17.56	2402.41
base	4056	0.200	UNSAT	-	1167.25
base	7414	0.141	UNSAT	18.42	1729.51
base	3536	0.219	UNSAT	60.65	45.97
base	3435	0.097	UNSAT	28.64	645.05
base	6503	0.103	UNSAT	-	581.38
base	9326	0.128	UNSAT	3117.31	264.82
base	5556	0.122	UNSAT	340.35	48.69

Table 5 Continued

Net	Image	Ep	Result	Oval	ERAN
base	5334	0.093	UNSAT	2.08	1672.26
base	1409	0.110	UNSAT	154.81	47.08
base	8431	0.178	UNSAT	-	193.67
base	371	0.247	UNSAT	29.25	310.65
base	9416	0.185	UNSAT	20.14	2212.59
base	4502	0.147	UNSAT	16.00	106.29
base	160	0.234	UNSAT	28.15	430.89
base	2340	0.197	UNSAT	19.49	154.32
base	6047	0.200	UNSAT	5.53	91.16
base	9419	0.210	UNSAT	12.81	197.71
base	2088	0.210	UNSAT	33.59	696.87
base	2349	0.172	UNSAT	10.57	1836.57
base	7438	0.238	UNSAT	1244.19	797.35
base	2310	0.090	UNSAT	-	366.2
base	5307	0.147	UNSAT	-	149.11
base	2043	0.140	UNSAT	48.92	382.64
base	6303	0.135	UNSAT	2.12	312.1
base	4185	0.215	UNSAT	1.29	476.92
base	2553	0.181	UNSAT	1.23	377.29
base	9409	0.244	UNSAT	22.03	759.53
base	7170	0.229	UNSAT	15.83	697.38
base	1316	0.172	UNSAT	13.14	637.95
base	4555	0.191	UNSAT	13.91	1888.69
base	2314	0.141	UNSAT	62.04	741.78
base	510	0.160	UNSAT	-	31.03
base	1545	0.115	UNSAT	14.50	281.7
base	6457	0.197	UNSAT	2.13	53.75
base	6480	0.216	UNSAT	12.79	944.95
base	1447	0.144	UNSAT	19.12	586
base	7345	0.178	UNSAT	39.89	55.05
base	8365	0.222	UNSAT	-	231.89
base	2012	0.224	UNSAT	31.28	189.82
base	4047	0.160	UNSAT	2.14	715.36
base	3062	0.205	UNSAT	2.10	854.62
base	6516	0.185	UNSAT	1.28	245.59
wide	2013	0.090	UNSAT	13.60	1087.27
wide	4025	0.088	UNSAT	32.72	193.39
wide	5022	0.063	UNSAT	65.65	220.16
wide	3021	0.090	UNSAT	-	65.36
wide	5021	0.172	UNSAT	17.46	82.33
wide	8008	0.122	UNSAT	27.07	1642.91
wide	1060	0.128	UNSAT	2615.61	614.51
wide	9058	0.082	UNSAT	40.47	497.02
wide	8001	0.135	UNSAT	150.86	91.24
wide	8037	0.141	UNSAT	14.56	101.76

Table 5 Continued

Net	Image	Ep	Result	Oval	ERAN
wide	6056	0.122	UNSAT	42.72	711.32
wide	105	0.153	UNSAT	-	-
wide	3034	0.078	UNSAT	47.29	130.48
wide	26	0.090	UNSAT	12.52	83.62
wide	9024	0.191	UNSAT	44.51	116.06
wide	3005	0.078	UNSAT	23.87	497.89
wide	9060	0.133	UNSAT	38.92	47.83
wide	1048	0.133	UNSAT	59.15	154.72
wide	2041	0.197	UNSAT	25.91	211.02
wide	0	0.063	UNSAT	16.62	824.13
wide	5000	0.210	UNSAT	-	232.99
wide	7011	0.090	UNSAT	18.61	743.44
wide	2045	0.122	UNSAT	141.46	101.89
wide	5043	0.145	UNSAT	37.51	157.4
wide	7046	0.103	UNSAT	25.37	261.88
wide	8002	0.160	UNSAT	19.12	112.3
wide	2028	0.122	UNSAT	12.43	2959.68
wide	5045	0.108	UNSAT	73.41	117.73
wide	6017	0.069	UNSAT	24.23	-
wide	6000	0.103	UNSAT	1.72	49.67
wide	3030	0.063	UNSAT	122.32	166
wide	1021	0.072	UNSAT	-	449.94
wide	4021	0.122	UNSAT	20.76	60.82
wide	19	0.095	UNSAT	27.25	1146.84
wide	5020	0.066	UNSAT	52.78	41.09
wide	5010	0.122	UNSAT	-	29.61
wide	9	0.076	UNSAT	30.26	-
wide	8052	0.147	UNSAT	20.82	221.64
wide	4059	0.197	UNSAT	26.27	172.98
wide	4057	0.077	UNSAT	33.19	632.05
wide	8039	0.158	UNSAT	49.10	341.48
wide	8046	0.072	UNSAT	10.22	194.83
wide	8013	0.210	UNSAT	26.46	264.3
wide	3001	0.210	UNSAT	15.02	894.92
wide	6007	0.141	UNSAT	16.21	641.31
wide	4015	0.128	UNSAT	12.99	-
wide	3002	0.191	UNSAT	63.09	248.63
wide	8021	0.122	UNSAT	22.33	897.58
wide	2052	0.122	UNSAT	117.66	1019.46
wide	2016	0.084	UNSAT	8.41	162.32
wide	2043	0.141	UNSAT	40.50	145.97
wide	1002	0.072	UNSAT	12.94	508.88
wide	100	0.084	UNSAT	315.36	113.57
wide	1005	0.122	UNSAT	8.72	63.28
wide	7033	0.147	UNSAT	15.93	156.52

Table 5 Continued

Net	Image	Ep	Result	Oval	ERAN
wide	1054	0.079	UNSAT	10.61	265.96
wide	1025	0.104	UNSAT	112.50	226.69
wide	4048	0.095	UNSAT	10.70	73.1
wide	4041	0.072	UNSAT	76.33	184.62
wide	34	0.172	UNSAT	58.70	49.44
wide	1013	0.178	UNSAT	15.38	429.41
wide	6052	0.172	UNSAT	21.29	20.23
wide	8049	0.128	UNSAT	35.71	593.14
wide	8035	0.072	UNSAT	47.83	164.1
wide	7020	0.104	UNSAT	14.47	36.94
wide	6034	0.122	UNSAT	15.72	71.25
wide	4009	0.135	UNSAT	27.34	137.73
wide	7047	0.160	UNSAT	5.14	279.51
wide	2014	0.150	UNSAT	11.14	472.11
wide	25	0.133	UNSAT	24.22	1559.95
wide	2040	0.106	UNSAT	117.64	126.63
wide	104	0.185	UNSAT	47.28	24.57
wide	37	0.191	UNSAT	17.33	414.71
wide	3020	0.133	UNSAT	20.13	-
wide	7008	0.106	UNSAT	37.25	148.98
wide	7031	0.106	UNSAT	89.94	126.16
wide	3017	0.216	UNSAT	53.90	435.54
wide	2024	0.097	UNSAT	15.83	725.89
wide	6019	0.141	UNSAT	8.05	79.34
wide	1044	0.191	UNSAT	57.85	-
wide	6043	0.063	UNSAT	25.06	1002.46
wide	3042	0.135	UNSAT	79.62	42.38
wide	2009	0.131	UNSAT	18.61	-
wide	6031	0.093	UNSAT	54.26	154.28
wide	1012	0.101	UNSAT	19.02	75.59
wide	6039	0.159	UNSAT	25.40	70.89
wide	9004	0.083	UNSAT	47.76	542.69
wide	3049	0.147	UNSAT	9.52	43.54
wide	1033	0.097	UNSAT	12.76	-
wide	9016	0.061	UNSAT	11.29	258.87
wide	15	0.096	UNSAT	-	167.79
wide	6051	0.084	UNSAT	28.16	69.15
wide	4	0.097	UNSAT	58.10	108.76
wide	6025	0.069	UNSAT	3.39	500.05
wide	9018	0.084	UNSAT	34.83	49.72
wide	8056	0.106	UNSAT	8.54	31.98
wide	6020	0.078	UNSAT	78.48	433.11
wide	9037	0.135	UNSAT	111.56	456.04
wide	8040	0.108	UNSAT	17.10	432.99
wide	18	0.178	UNSAT	7.58	128.83

Table 5 Continued

Net	Image	Ep	Result	Oval	ERAN
deep	1033.0	0.103	UNSAT	114.08	-
deep	8016.0	0.141	UNSAT	27.67	366.58
deep	43.0	0.128	UNSAT	18.43	519.13
deep	9001.0	0.106	UNSAT	23.38	612.62
deep	6039.0	0.135	UNSAT	19.16	313.83
deep	2104.0	0.135	UNSAT	10.59	904.55
deep	5036.0	0.181	UNSAT	12.39	390.66
deep	5009.0	0.100	UNSAT	10.45	422.84
deep	8021.0	0.141	UNSAT	18.22	304.16
deep	6033.0	0.151	UNSAT	16.35	363.93
deep	2045.0	0.110	UNSAT	21.47	740.78
deep	7026.0	0.084	UNSAT	19.53	451.6
deep	9010.0	0.122	UNSAT	11.59	328.64
deep	104.0	0.178	UNSAT	56.30	583.05
deep	4055.0	0.135	UNSAT	5.71	678.75
deep	6058.0	0.147	UNSAT	26.86	575.47
deep	1001.0	0.147	UNSAT	16.23	1507.29
deep	9037.0	0.160	UNSAT	16.03	356.68
deep	5030.0	0.078	UNSAT	8.51	497.79
deep	5040.0	0.172	UNSAT	10.66	80.39
deep	51.0	0.063	UNSAT	36.66	412.87
deep	7032.0	0.122	UNSAT	2801.97	636.16
deep	3007.0	0.122	UNSAT	11.53	783.28
deep	5041.0	0.147	UNSAT	16.67	326.3
deep	6048.0	0.160	UNSAT	18.62	540.41
deep	3021.0	0.147	UNSAT	15.78	458.43
deep	1041.0	0.122	UNSAT	36.00	516.45
deep	41.0	0.178	UNSAT	35.44	440.51
deep	8018.0	0.063	UNSAT	8.70	766.76
deep	3102.0	0.084	UNSAT	52.86	156.76
deep	7031.0	0.100	UNSAT	15.18	302.95
deep	2.0	0.122	UNSAT	103.39	966.81
deep	6019.0	0.178	UNSAT	9.79	452.53
deep	8023.0	0.097	UNSAT	41.70	1022.03
deep	8043.0	0.103	UNSAT	14.16	371.02
deep	8000.0	0.160	UNSAT	10.39	408.01
deep	2008.0	0.120	UNSAT	14.65	526.21
deep	4031.0	0.110	UNSAT	-	113.18
deep	9023.0	0.141	UNSAT	11.73	510.42
deep	9022.0	0.128	UNSAT	18.77	108.82
deep	2002.0	0.122	UNSAT	21.59	543.19
deep	9016.0	0.084	UNSAT	8.62	388.61
deep	3002.0	0.110	UNSAT	9.56	376.05
deep	1104.0	0.110	UNSAT	290.28	152.51
deep	2112.0	0.135	UNSAT	11.67	478.5

Table 5 Continued

Net	Image	Ep	Result	Oval	ERAN
deep	6056.0	0.095	UNSAT	21.18	534.15
deep	3022.0	0.160	UNSAT	14.37	625.31
deep	2043.0	0.097	UNSAT	16.54	834.94
deep	6010.0	0.061	UNSAT	11.95	533.59
deep	2024.0	0.160	UNSAT	12.83	137.15
deep	15.0	0.122	UNSAT	9.61	641.74
deep	5012.0	0.122	UNSAT	9.70	753.82
deep	7005.0	0.110	UNSAT	29.26	553.31
deep	1102.0	0.090	UNSAT	14.07	321.48
deep	7013.0	0.107	UNSAT	9.53	237.83
deep	5027.0	0.165	UNSAT	18.50	306.66
deep	3030.0	0.063	UNSAT	9.52	649.61
deep	1009.0	0.122	UNSAT	19.96	192.59
deep	7043.0	0.153	UNSAT	17.95	348.3
deep	2012.0	0.160	UNSAT	20.53	646.02
deep	4104.0	0.147	UNSAT	19.64	896.67
deep	9028.0	0.084	UNSAT	12.90	563.38
deep	6049.0	0.128	UNSAT	8.83	293.96
deep	50.0	0.185	UNSAT	10.61	439.81
deep	8035.0	0.108	UNSAT	20.66	646.87
deep	2028.0	0.210	UNSAT	10.59	591.3
deep	3028.0	0.110	UNSAT	21.04	631.45
deep	2001.0	0.122	UNSAT	11.39	459.31
deep	3012.0	0.100	UNSAT	29.29	780.81
deep	9005.0	0.210	UNSAT	8.72	177.76
deep	1006.0	0.160	UNSAT	5.62	153.08
deep	7011.0	0.122	UNSAT	10.69	800.43
deep	9014.0	0.090	UNSAT	30.31	133.5
deep	2101.0	0.147	UNSAT	10.52	282.17
deep	4022.0	0.122	UNSAT	19.83	388.12
deep	4015.0	0.185	UNSAT	9.38	460.96
deep	2107.0	0.178	UNSAT	5.73	666.75
deep	2050.0	0.100	UNSAT	4.76	718.02
deep	7025.0	0.110	UNSAT	21.05	361.81
deep	0.0	0.110	UNSAT	20.17	487.95
deep	2005.0	0.153	UNSAT	14.95	293.38
deep	8002.0	0.128	UNSAT	14.17	564.07
deep	8004.0	0.178	UNSAT	45.55	350.78
deep	4008.0	0.084	UNSAT	901.76	1907.72
deep	29.0	0.135	UNSAT	16.07	320.31
deep	2119.0	0.122	UNSAT	29.68	861.21
deep	9030.0	0.133	UNSAT	70.85	499.71
deep	39.0	0.084	UNSAT	26.19	897.68
deep	9.0	0.153	UNSAT	12.26	447.03
deep	45.0	0.128	UNSAT	24.67	802.76

Table 5 Continued

Net	Image	Ep	Result	Oval	ERAN
deep	1029.0	0.160	UNSAT	16.11	342.13
deep	7042.0	0.100	UNSAT	20.46	701.57
deep	2011.0	0.135	UNSAT	46.00	450.07
deep	4100.0	0.206	UNSAT	17.00	1072.09
deep	7001.0	0.210	UNSAT	2.91	668.23
deep	9026.0	0.228	UNSAT	26.26	1167.4
deep	37.0	0.256	UNSAT	22.02	188.34
deep	1011.0	0.172	UNSAT	2.89	765.68
deep	6002.0	0.259	UNSAT	21.67	1063.11
deep	1013.0	0.228	UNSAT	13.94	1489.22

4.3 NLN

4.3.1 NEEL - Nonlinear Networks

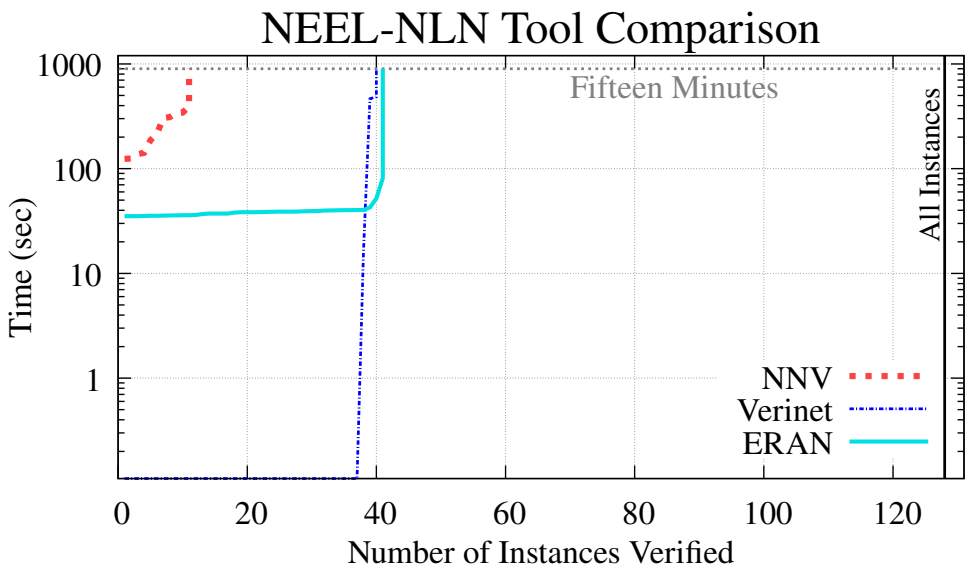

Figure 5: NEEL-NLN Cactus Plot.

Table 6: Tool Runtime (sec) for NEEL-NLN.

Net	Ep	Image	Result	NNV	VeriNet	ERAN
sig_2L	5	1	UNSAT	339.62	-	-
sig_2L	5	2	UNSAT	137.80	-	-
sig_2L	5	3	UNSAT	304.81	-	-
sig_2L	5	4	UNSAT	-	464.32	-
sig_2L	5	5	UNSAT	186.35	-	-
sig_2L	5	6	UNSAT	391.73	-	-
sig_2L	5	7	UNSAT	310.29	-	-
sig_2L	5	8	UNSAT	342.61	-	-
sig_2L	5	9	MISMATCH	(U)275.52	(S)0.02	(S)42.5
sig_2L	5	10	UNSAT	141.63	-	-
sig_2L	5	11	UNSAT	215.18	-	-
sig_2L	5	12	UNKNOWN	-	-	-
sig_2L	5	13	UNKNOWN	-	-	-
sig_2L	5	14	UNSAT	124.93	-	-
sig_2L	5	15	UNKNOWN	-	-	-
sig_2L	5	16	UNSAT	123.99	-	-
sig_2L	12	1	UNKNOWN	-	-	-
sig_2L	12	2	SAT	-	-	42.6

Table 6 Continued

Net	Ep	Image	Result	NNV	VeriNet	ERAN
sig_2L	12	3	SAT	-	0.03	39.2
sig_2L	12	4	UNKNOWN	-	-	-
sig_2L	12	5	SAT	-	0.03	37.3
sig_2L	12	6	UNKNOWN	-	-	-
sig_2L	12	7	SAT	-	0.02	38.4
sig_2L	12	8	UNKNOWN	-	-	-
sig_2L	12	9	SAT	-	0.02	38.9
sig_2L	12	10	SAT	-	0.01	38.8
sig_2L	12	11	UNKNOWN	-	-	-
sig_2L	12	12	SAT	-	0.04	37.3
sig_2L	12	13	UNKNOWN	-	-	-
sig_2L	12	14	UNKNOWN	-	-	-
sig_2L	12	15	UNKNOWN	-	-	-
sig_2L	12	16	SAT	-	0.02	39.8
sig_3L	5	1	UNKNOWN	-	-	-
sig_3L	5	2	UNKNOWN	-	-	-
sig_3L	5	3	UNKNOWN	-	-	-
sig_3L	5	4	UNKNOWN	-	-	-
sig_3L	5	5	UNKNOWN	-	-	-
sig_3L	5	6	UNKNOWN	-	-	-
sig_3L	5	7	UNKNOWN	-	-	-
sig_3L	5	8	UNKNOWN	-	-	-
sig_3L	5	9	SAT	-	0.03	38.6
sig_3L	5	10	UNKNOWN	-	-	-
sig_3L	5	11	UNKNOWN	-	-	-
sig_3L	5	12	UNKNOWN	-	-	-
sig_3L	5	13	UNKNOWN	-	-	-
sig_3L	5	14	UNKNOWN	-	-	-
sig_3L	5	15	UNKNOWN	-	-	-
sig_3L	5	16	UNKNOWN	-	-	-
sig_3L	12	1	UNKNOWN	-	-	-
sig_3L	12	2	UNKNOWN	-	-	-
sig_3L	12	3	SAT	-	0.03	35.3
sig_3L	12	4	UNKNOWN	-	-	-
sig_3L	12	5	SAT	-	0.02	35.7
sig_3L	12	6	SAT	-	0.02	35.5
sig_3L	12	7	SAT	-	0.01	35.6
sig_3L	12	8	SAT	-	0.02	35.2
sig_3L	12	9	SAT	-	0.01	35.2
sig_3L	12	10	SAT	-	484.32	51.8
sig_3L	12	11	SAT	-	16.27	35.5
sig_3L	12	12	SAT	-	0.03	35.2
sig_3L	12	13	SAT	-	0.01	35.9
sig_3L	12	14	UNKNOWN	-	-	-
sig_3L	12	15	SAT	-	-	81.3

Table 6 Continued

Net	Ep	Image	Result	NNV	VeriNet	ERAN
sig_3L	12	16	SAT	-	0.03	35.8
tanh_2L	5	1	UNKNOWN	-	-	-
tanh_2L	5	2	UNKNOWN	-	-	-
tanh_2L	5	3	UNKNOWN	-	-	-
tanh_2L	5	4	UNKNOWN	-	-	-
tanh_2L	5	5	UNKNOWN	-	-	-
tanh_2L	5	6	UNKNOWN	-	-	-
tanh_2L	5	7	UNKNOWN	-	-	-
tanh_2L	5	8	UNKNOWN	-	-	-
tanh_2L	5	9	SAT	-	0.02	38.5
tanh_2L	5	10	UNKNOWN	-	-	-
tanh_2L	5	11	UNKNOWN	-	-	-
tanh_2L	5	12	UNKNOWN	-	-	-
tanh_2L	5	13	UNKNOWN	-	-	-
tanh_2L	5	14	UNKNOWN	-	-	-
tanh_2L	5	15	UNKNOWN	-	-	-
tanh_2L	5	16	SAT	-	0.02	40.2
tanh_2L	12	1	UNKNOWN	-	-	-
tanh_2L	12	2	UNKNOWN	-	-	-
tanh_2L	12	3	SAT	-	0.02	38.4
tanh_2L	12	4	UNKNOWN	-	-	-
tanh_2L	12	5	SAT	-	0.03	37.2
tanh_2L	12	6	SAT	-	0.04	40.1
tanh_2L	12	7	SAT	-	0.02	40.3
tanh_2L	12	8	UNKNOWN	-	-	-
tanh_2L	12	9	SAT	-	0.02	39.4
tanh_2L	12	10	SAT	-	0.05	36.8
tanh_2L	12	11	SAT	-	0.01	36.1
tanh_2L	12	12	SAT	-	0.03	37.3
tanh_2L	12	13	UNKNOWN	-	-	-
tanh_2L	12	14	UNKNOWN	-	-	-
tanh_2L	12	15	UNKNOWN	-	-	-
tanh_2L	12	16	SAT	-	0.02	38.7
tanh_3L	5	1	UNKNOWN	-	-	-
tanh_3L	5	2	UNKNOWN	-	-	-
tanh_3L	5	3	UNKNOWN	-	-	-
tanh_3L	5	4	UNKNOWN	-	-	-
tanh_3L	5	5	UNKNOWN	-	-	-
tanh_3L	5	6	UNKNOWN	-	-	-
tanh_3L	5	7	UNKNOWN	-	-	-
tanh_3L	5	8	UNKNOWN	-	-	-
tanh_3L	5	9	SAT	-	0.02	40
tanh_3L	5	10	UNKNOWN	-	-	-
tanh_3L	5	11	UNKNOWN	-	-	-
tanh_3L	5	12	UNKNOWN	-	-	-

Table 6 Continued

Net	Ep	Image	Result	NNV	VeriNet	ERAN
tanh_3L	5	13	UNKNOWN	-	-	-
tanh_3L	5	14	UNKNOWN	-	-	-
tanh_3L	5	15	UNKNOWN	-	-	-
tanh_3L	5	16	UNKNOWN	-	-	-
tanh_3L	12	1	UNKNOWN	-	-	-
tanh_3L	12	2	UNKNOWN	-	-	-
tanh_3L	12	3	SAT	-	0.02	38.8
tanh_3L	12	4	UNKNOWN	-	-	-
tanh_3L	12	5	SAT	-	0.04	40.3
tanh_3L	12	6	SAT	-	0.02	35.9
tanh_3L	12	7	SAT	-	0.02	39.3
tanh_3L	12	8	SAT	-	0.02	38.0
tanh_3L	12	9	SAT	-	0.01	38.4
tanh_3L	12	10	SAT	-	0.04	39.9
tanh_3L	12	11	UNKNOWN	-	-	-
tanh_3L	12	12	UNKNOWN	-	-	-
tanh_3L	12	13	UNKNOWN	-	-	-
tanh_3L	12	14	UNKNOWN	-	-	-
tanh_3L	12	15	UNKNOWN	-	-	-
tanh_3L	12	16	SAT	-	0.02	39.0

5 Discussion

The VNN-COMP2020 competition severed as a useful case-study to understand the performance and capabilities of the individual verification tools. However, the results particularly on performance were vastly varied as they were highly dependent on some external and internal factors as well, such as differences in hardware used for execution, the tool implementation platforms, optimization techniques. In the future, having bootstrapped this process, we plan to utilize a common infrastructure to improve conclusions that can be drawn about the underlying verification methods. We also plan to require usage of the draft VNN-LIB format[8] for future iterations that allows a restricted subset of ONNX and an SMT-based specification language, which was released around the time of the presentation of the VNN-COMP results.

Related to this formatting and interchange, several participants encountered difficulties related to the model formats. Several of the model format issues encountered include the following.

1. Some underlying libraries represent certain operations or layers in different ways, for instance, PyTorch utilizing row major for 'reshape' with MATLAB utilizing column major, which can result in differing outputs.

2. Some ONNX conversion tools (e.g., PyTorch and MATLAB) add extra and unnecessary layers to the original network after conversion.

[8] http://www.vnnlib.org/

3. Some ONNX conversion tools converted fully-connected layers to convolutional layers. While there is an equivalence between several types of layers, this can lead to different analysis approaches being used in the verification tool solvers, and can impact performance.

Working with the restricted set of ONNX operations allowed, plus utilizing some community-driven conversion tools, such as to take a (somewhat) arbitrary ONNX model and convert it to the VNN-LIB allowed subset of ONNX may help with some of these interchange and formatting issues.

6 Conclusion

This report summarizes the 1st Verification of Neural Networks Competition (VNN-COMP) held in 2020.

Acknowledgements

The material presented in this report is based in part upon work supported by the Defense Advanced Research Projects Agency (DARPA) through contract number FA8750-18-C-0089, the National Science Foundation (NSF) under grant numbers SHF 1910017 and FMitF 1918450, and the Air Force Office of Scientific Research (AFOSR) through award numbers FA9550-18-1-0122 and FA9550-19-1-0288. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon. Any opinions, finding, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of Air Force, AFOSR, DARPA, or NSF.

References

[1] Stanley Bak. Execution-guided overapproximation (ego) for improving scalability of neural network verification, 2020.

[2] Stanley Bak, Hoang-Dung Tran, Kerianne Hobbs, and Taylor T. Johnson. Improved geometric path enumeration for verifying ReLU neural networks. In *32nd International Conference on Computer-Aided Verification (CAV)*, July 2020.

[3] E. Botoeva, P. Kouvaros, J. Kronqvist, A. Lomuscio, and R. Misener. Efficient verification of neural networks via dependency analysis. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI20)*. AAAI Press, 2020.

[4] Rudy Bunel, Alessandro De Palma, Alban Desmaison, Krishnamurthy Dvijotham, Pushmeet Kohli, Philip HS Torr, and M Pawan Kumar. Lagrangian decomposition for neural network verification. *Conference on Uncertainty in Artificial Intelligence*, 2020.

[5] Rudy Bunel, Jingyue Lu, Ilker Turkaslan, P Kohli, P Torr, and M Pawan Kumar. Branch and bound for piecewise linear neural network verification. *Journal of Machine Learning Research*, 21(2020), 2020.

[6] Iain Dunning, Joey Huchette, and Miles Lubin. Jump: A modeling language for mathematical optimization. *SIAM Review*, 59(2):295–320, 2017.

[7] 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.

[8] P. Henriksen and A. Lomuscio. Efficient neural network verification via adaptive refinement and adversarial search. In *Proceedings of the 24th European Conference on Artificial Intelligence (ECAI20)*, 2020.

[9] Guy Katz, Clark Barrett, David L Dill, Kyle Julian, and Mykel J Kochenderfer. Reluplex: An efficient smt solver for verifying deep neural networks. In *International Conference on Computer Aided Verification*, pages 97–117. Springer, 2017.

[10] Guy Katz, Derek A Huang, Duligur Ibeling, Kyle Julian, Christopher Lazarus, Rachel Lim, Parth Shah, Shantanu Thakoor, Haoze Wu, Aleksandar Zeljić, et al. The marabou framework for verification and analysis of deep neural networks. In *International Conference on Computer Aided Verification*, pages 443–452. Springer, 2019.

[11] Haitham Khedr, James Ferlez, and Yasser Shoukry. Effective formal verification of neural networks using the geometry of linear regions. *arXiv preprint arXiv:2006.10864*, 2020.

[12] Jingyue Lu and M Pawan Kumar. Neural network branching for neural network verification. In *International Conference on Learning Representations*, 2020.

[13] Martin Vechev Mislav Balunovic. Adversarial training and provable defenses: Bridging the gap. In *Proc. International Conference on Learning Representations (ICLR)*, 2020.

[14] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in pytorch. *NIPS Autodiff Workshop*, 2017.

[15] Jonas Rauber, Wieland Brendel, and Matthias Bethge. Foolbox: A python toolbox to benchmark the robustness of machine learning models. *arXiv preprint arXiv:1707.04131*, 2017.

[16] Gagandeep Singh, Rupanshu Ganvir, Markus Püschel, and Martin Vechev. Beyond the single neuron convex barrier for neural network certification. In *Advances in Neural Information Processing Systems 32*, pages 15098–15109. Curran Associates, Inc., 2019.

[17] Gagandeep Singh, Timon Gehr, Matthew Mirman, Markus Püschel, and Martin Vechev. Fast and effective robustness certification. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems 31*, pages 10802–10813. Curran Associates, Inc., 2018.

[18] Gagandeep Singh, Timon Gehr, Markus Püschel, and Martin Vechev. An abstract domain for certifying neural networks. *Proc. ACM Program. Lang.*, 3(POPL):41:1–41:30, 2019.

[19] Gagandeep Singh, Timon Gehr, Markus Püschel, and Martin Vechev. Boosting robustness certification of neural networks. In *Proc. International Conference on Learning Representations (ICLR)*, 2019.

[20] Gagandeep Singh, Markus Püschel, and Martin Vechev. Fast polyhedra abstract domain. In *Proc. Principles of Programming Languages (POPL)*, pages 46–59, 2017.

[21] Vincent Tjeng, Kai Y. Xiao, and Russ Tedrake. Evaluating robustness of neural networks with mixed integer programming. In *ICLR*, 2019.

[22] Hoang-Dung Tran, Stanley Bak, Weiming Xiang, and Taylor T. Johnson. Verification of deep convolutional neural networks using imagestars. In *32nd International Conference on Computer-Aided Verification (CAV)*. Springer, July 2020.

[23] Hoang-Dung Tran, Patrick Musau, Diego Manzanas Lopez, Xiaodong Yang, Luan Viet Nguyen, Weiming Xiang, and Taylor T. Johnson. Parallelizable reachability analysis algorithms for feed-forward neural networks. In *Proceedings of the 7th International Workshop on Formal Methods in Software Engineering (FormaliSE'19)*, FormaliSE '19, pages 31–40, Piscataway, NJ, USA, May 2019. IEEE Press.

[24] Hoang-Dung Tran, Patrick Musau, Diego Manzanas Lopez, Xiaodong Yang, Luan Viet Nguyen, Weiming Xiang, and Taylor T. Johnson. Star-based reachability analysis for deep neural networks. In *23rd International Symposium on Formal Methods (FM'19)*. Springer International Publishing, October 2019.

[25] Hoang-Dung Tran, Xiaodong Yang, Diego Manzanas Lopez, Patrick Musau, Luan Viet Nguyen,

Weiming Xiang, Stanley Bak, and Taylor T. Johnson. NNV: The neural network verification tool for deep neural networks and learning-enabled cyber-physical systems. In *32nd International Conference on Computer-Aided Verification (CAV)*, July 2020.

[26] Matthias C. M. Troffaes. pycddlib, 2018.

[27] Eric Wong and Zico Kolter. Provable defenses against adversarial examples via the convex outer adversarial polytope. In *International Conference on Machine Learning*, pages 5286–5295, 2018.

[28] W. Xiang, H. Tran, and T. T. Johnson. Output reachable set estimation and verification for multilayer neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 29(11):5777–5783, 2018.