Branching strategies for Zonotope-based Neural Network Verification

Technical University of Munich

Department of Informatics

Chair of Robotics, Artificial Intelligence and Real-time Systems

Background

Neural networks are developing in fast rates and are great at solving many complex tasks. However, the output of many neural networks is sensitive towards tiny input perturbations [4], thus making them unsuitable for high-safety domains. To enable their power in these domains, we need to focus on training neural networks that are robust against input perturbations and verifying their robustness formally.

Although these problems can be solved using SMT (Satisfiability Modulo Theories) solvers or MILP (Mixed-Integer Linear Programming), which are complete methods [5], they remain too expensive, as the problem is computationally hard Branch and Bound [2] framework, in order to control the exponentiality of the process and achieve completeness for piecewise linear neural networks.

By using this method we consider 2 important strategies; branching: where we partition our problem p into a set of easier subproblems $\{p_i\}$ and specify the order in which they are considered and bounding: where we bound the image of the input space through the network resulting in an overapproximation of the resulting output set, of which we check the safety.

In this framework, the verification problem is recursively split using the branching strategy into smaller subproblems that are evaluated using the bounding procedure when possible or pruned back awaiting further splits.

Description

The aim of this thesis is to extend the CORA neural network verification procedure [1]. In this approach, the bounding rule propagates zonotopes of the input set through the neural network using overapproximations of non-linear activations. Although branching rules can be agnostic to the bounding strategies, they can contribute to better overall verification time if optimized by controlling the number of subproblems treated for the complete verification. Therefore, we focus on the branching aspect to investigate different heuristics that leverage the structure of the network or symmetries of the problem. Several such heuristics are implemented in the literature using scores based on neuron sensitivity, unfixed ReLU activations, improvement in bound tightness estimations, cost estimation or optimized Lagrangian parameters [2, 3, 7, 8].

Tasks

- Familiarize with the toolbox CORA [1].
- Literature research of state of the art BaB verification methods.
- Implementation of different heuristics for problem branching.
- Evaluate implemented heuristics on zonotope-based training.
- Optional: optimality conditions for branching heuristics.
- Optional: integrate the branching heuristics in the training phase [6]

References

[1] Matthias Althoff. An introduction to cora 2015. In *ARCH@ CPSWeek*, pages 120–151, 2015.

[2] Rudy Bunel, Ilker Turkaslan, Philip H. S. Torr, M. Pawan Kumar, Jingyue Lu, and Pushmeet Kohli. Branch and bound for piecewise linear neural network verification. *JMLR*, 21(1), 2020.

Supervisor:
Prof. Dr.-Ing. Matthias Althoff

Advisor:
Lukas Koller, M.Sc.

Research project:
DFG–SPP2422

Type:
BT

Research area:
Formal verification, neural networks

Programming language:
MATLAB

Required skills:
Knowledge in formal methods and machine learning, good mathematical background

Language:
English

Date of submission:
6. März 2025

For more information please contact us:

Phone: +49 (89) 289 - 18140

E-Mail: lukas.koller@tum.de

Website: ce.cit.tum.de/cps/

[3] Claudio Ferrari, Mark Niklas Müller, Nikola Jovanović, and Martin Vechev. Complete verification via multi-neuron relaxation guided branch-and-bound. In *ICLR*, 2022.

[4] Ian Goodfellow, Jonathon Shlens, and Christian Szegedy. Explaining and harnessing adversarial examples. In *ICLR*, 2015.

[5] Guy Katz, Clark Barrett, David Dill, Kyle Julian, and Mykel Kochenderfer. Reluplex: An efficient smt solver for verifying deep neural networks. In *CAV*, 2017.

[6] Lukas Koller, Tobias Ladner, and Matthias Althoff. Set-based training for neural network verification. *arXiv preprint arXiv:2401.14961*, 2025.

[7] Tobias Ladner and Matthias Althoff. Automatic abstraction refinement in neural network verification using sensitivity analysis. In *HSCC*, 2023.

[8] Shiqi Wang, Huan Zhang, Kaidi Xu, Xue Lin, Suman Jana, Cho-Jui Hsieh, and Zico Kolter. Beta-crown: efficient bound propagation with per-neuron split constraints for neural network robustness verification. In *NeurIPS*, 2021.

Technical University of Munich

Department of Informatics

Chair of Robotics, Artificial Intelligence and Real-time Systems