Automated machine learning for neural network verification

König, H.M.T.

Bibliography

[1] Michael Akintunde, Alessio Lomuscio, Lalit Maganti, and Edoardo Pirovano. Reachability Analysis for Neural Agent-Environment Systems. In *Proceedings of the 16th International Conference on Principles of Knowledge Representation and Reasoning (KR2018)*, pages 184–193, 2018.

[2] Syed Muhammad Anwar, Muhammad Majid, Adnan Qayyum, Muhammad Awais, Majdi Alnowami, and Muhammad Khurram Khan. Medical image analysis using convolutional neural networks: a review. *Journal of medical systems*, 42:1–13, 2018.

[3] Stanley Bak, Changliu Liu, and Taylor Johnson. The Second International Verification of Neural Networks Competition (VNN-COMP 2021): Summary and Results. *arXiv preprint arXiv:2109.00498*, 2021.

[4] Stanley Bak, Hoang-Dung Tran, Kerianne Hobbs, and Taylor T. Johnson. Improved Geometric Path Enumeration for Verifying ReLU Neural Networks. In *Proceedings of the 32nd International Conference on Computer Aided Verification (CAV 2020)*, pages 66–96, 2020.

[5] Thomas Bartz-Beielstein, Carola Doerr, Daan van den Berg, Jakob Bossek, Sowmya Chandrasekaran, Tome Eftimov, Andreas Fischbach, Pascal Kerschke, William La Cava, Manuel Lopez-Ibanez, et al. Benchmarking in Optimization: Best Practice and Open Issues. *arXiv preprint arXiv:2007.03488*, 2020.

[6] Osbert Bastani, Yani Ioannou, Leonidas Lampropoulos, Dimitrios Vytiniotis, Aditya Nori, and Antonio Criminisi. Measuring Neural Net Robustness with Constraints. In *Advances in Neural Information Processing Systems 29 (NeurIPS 2016)*, pages 2613–2621, 2016.

[7] Mauro Birattari, Zhi Yuan, Prasanna Balaprakash, and Thomas Stützle. F-Race and Iterated F-Race: An Overview. In Thomas Bartz-Beielstein, Marco Chiarandini, Luís Paquete, and Mike Preuss, editors, *Experimental Methods for the Analysis of Optimization Algorithms*, pages 311–336. Springer, 2010.

[8] Elena Botoeva, Panagiotis Kouvaros, Jan Kronqvist, Alessio Lomuscio, and Ruth Misener. Efficient Verification of ReLU-based Neural Networks via Dependency

Analysis. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI-20)*, pages 3291–3299, 2020.

[9] Leo Breiman. Random Forests. *Machine Learning*, 45(1):5–32, 2001.

[10] Christopher Brix, Stanley Bak, Changliu Liu, and Taylor T Johnson. The Fourth International Verification of Neural Networks Competition (VNN-COMP 2023): Summary and Results. *arXiv preprint arXiv:2312.16760*, 2023.

[11] Rudy Bunel, Jingyue Lu, Ilker Turkaslan, Philip H. S. Torr, Pushmeet Kohli, and M. Pawan Kumar. Branch and Bound for Piecewise Linear Neural Network Verification. *Journal of Machine Learning Research*, 21(42):4795–4804, 2020.

[12] Rudy Bunel, Ilker Turkaslan, Philip Torr, Pushmeet Kohli, and Pawan K Mudigonda. A Unified View of Piecewise Linear Neural Network Verification. In *Advances in Neural Information Processing Systems 31 (NeurIPS 2018)*, pages 1–10, 2018.

[13] Nicholas Carlini, Guy Katz, Clark Barrett, and David L Dill. Provably Minimally-Distorted Adversarial Examples. *arXiv preprint arXiv:1709.10207*, 2017.

[14] Nicholas Carlini and David Wagner. Towards Evaluating the Robustness of Neural Networks. In *Proceedings of the 38th IEEE Symposium on Security and Privacy (IEEE S&P 2017)*, pages 39–57, 2017.

[15] Marco Casadio, Ekaterina Komendantskaya, Matthew L. Daggitt, Wen Kokke, Guy Katz, Guy Amir, and Idan Refaeli. Neural Network Robustness as a Verification Property: A Principled Case Study. In *Proceedings of the 34rd International Conference on Computer Aided Verification (CAV 2022)*, pages 219–231, 2022.

[16] Pin-Yu Chen, Yash Sharma, Huan Zhang, Jinfeng Yi, and Cho-Jui Hsieh. EAD: Elastic-Net Attacks to Deep Neural Networks via Adversarial Examples. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI-18)*, pages 10–17, 2018.

[17] Chih-Hong Cheng, Georg Nührenberg, and Harald Ruess. Maximum Resilience of Artificial Neural Networks. In *Proceedings of the 15th International Symposium on Automated Technology for Verification and Analysis (ATVA2017)*, pages 251–268, 2017.

[18] Marco Chiarandini, Chris Fawcett, and Holger H Hoos. A Modular Multiphase Heuristic Solver for Post Enrolment Course Timetabling. In *Proceedings of the 7th International Conference on the Practice and Theory of Automated Timetabling (PATAT 2008)*, 2008.

[19] Jeremy Cohen, Elan Rosenfeld, and Zico Kolter. Certified Adversarial Robustness via Randomized Smoothing. In *Proceedings of the 36th International Conference on Machine Learning (ICML 2019)*, pages 1310–1320, 2019.

[20] Francesco Croce, Maksym Andriushchenko, and Matthias Hein. Provable Robustness of ReLU networks via Maximization of Linear Regions. In Kamalika Chaudhuri and Masashi Sugiyama, editors, *Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics (AISTATS 2019)*, pages 2057–2066, 2019.

[21] George B. Dantzig. Linear Programming. *Operations Research*, 50(1):42–47, 2002.

[22] Alessandro De Palma, Rudy Bunel, Alban Desmaison, Krishnamurthy Dvijotham, Pushmeet Kohli, Philip H. S. Torr, and M. Pawan Kumar. Improved Branch and Bound for Neural Network Verification via Lagrangian Decomposition. *arXiv preprint arXiv:2104.06718*, 2021.

[23] Gavin Weiguang Ding, Yash Sharma, Kry Yik Chau Lui, and Ruitong Huang. MMA training: Direct input space margin maximization through adversarial training. In *Proceedings of the 8th International Conference on Learning Representations (ICLR 2020)*, pages 2057–2066, 2020.

[24] Yinpeng Dong, Fangzhou Liao, Tianyu Pang, Hang Su, Jun Zhu, Xiaolin Hu, and Jianguo Li. Boosting Adversarial Attacks With Momentum. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2018)*, pages 9185–9193, 2018.

[25] Souradeep Dutta, Susmit Jha, Sriram Sankaranarayanan, and Ashish Tiwari. Output Range Analysis for Deep Neural Networks. In *Proceedings of the Tenth NASA Formal Methods Symposium (NFM 2018)*, pages 121–138, 2018.

[26] Krishnamurthy Dvijotham, Robert Stanforth, Sven Gowal, Timothy A Mann, and Pushmeet Kohli. A Dual Approach to Scalable Verification of Deep Networks. In *Proceedings of the 38th Conference on Uncertainty in Artificial Intelligence (UAI 2018)*, pages 550–559, 2018.

[27] Ruediger Ehlers. Formal Verification of Piece-Wise Linear Feed-Forward Neural Networks. In *Proceedings of the 15th International Symposium on Automated Technology for Verification and Analysis (ATVA 2017)*, pages 269–286, 2017.

[28] Thomas Elsken, Jan Hendrik Metzen, and Frank Hutter. Neural Architecture Search: A Survey. *The Journal of Machine Learning Research*, 20(1):1997–2017, 2019.

[29] Claudio Ferrari, Mark Niklas Mueller, Nikola Jovanović, and Martin Vechev. Complete Verification via Multi-Neuron Relaxation Guided Branch-and-Bound. In *Proceedings of the 10th International Conference on Learning Representations (ICLR 2022)*, pages 1–15, 2022.

[30] Matthias Feurer, Aaron Klein, Katharina Eggensperger, Jost Springenberg, Manuel Blum, and Frank Hutter. Efficient and Robust Automated Machine

Learning. In *Advances in Neural Information Processing Systems 28 (NeurIPS 2015)*, pages 2962–2970, 2015.

[31] Matthias Feurer, Jost Tobias Springenberg, and Frank Hutter. Initializing Bayesian Hyperparameter Optimization via Meta-Learning. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI–15)*, pages 1128–1135, 2015.

[32] Matteo Fischetti and Jason Jo. Deep neural networks and mixed integer linear optimization. *Constraints*, 23(3):296–309, 2018.

[33] Alexandre Fréchette, Lars Kotthoff, Tomasz Michalak, Talal Rahwan, Holger Hoos, and Kevin Leyton-Brown. Using the shapley value to analyze algorithm portfolios. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI-16)*, pages 3397–3403, 2016.

[34] Martin Gebser, Roland Kaminski, Benjamin Kaufmann, Torsten Schaub, Marius Thomas Schneider, and Stefan Ziller. A Portfolio Solver for Answer Set Programming: Preliminary Report. In *Proceedings of the 10th International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR2019)*, pages 1–6, 2011.

[35] Timon Gehr, Matthew Mirman, Dana Drachsler-Cohen, Petar Tsankov, Swarat Chaudhuri, and Martin Vechev. AI2: Safety and Robustness Certification of Neural Networks with Abstract Interpretation. In *Proceedings of the 39th IEEE Symposium on Security and Privacy (IEEE S&P 2018)*, pages 3–18, 2018.

[36] Carla P Gomes and Bart Selman. Algorithm portfolios. *Artificial Intelligence*, 126(1–2):43–62, 2001.

[37] Ian J. Goodfellow, Jonathon Shlens, and Christian Szegedy. Explaining and Harnessing Adversarial Examples. In *Proceedings of the 3rd International Conference on Learning Representations (ICLR 2015)*, pages 1–11, 2015.

[38] Warren He, Bo Li, and Dawn Song. Decision Boundary Analysis of Adversarial Examples. In *Proceedings of the 6th International Conference on Learning Representations (ICLR 2018)*, pages 1–15, 2018.

[39] Patrick Henriksen, Kerstin Hammernik, Daniel Rueckert, and Alessio Lomuscio. Bias Field Robustness Verification of Large Neural Image Classifiers. In *Proceedings of the 32nd British Machine Vision Conference 2021 (BMVC 2021)*, pages 202–2016, 2021.

[40] Patrick Henriksen and Alessio Lomuscio. Efficient Neural Network Verification via Adaptive Refinement and Adversarial Search. In *Proceedings of the 24th European Conference on Artificial Intelligence (ECAI 2020)*, pages 2513–2520, 2020.

[41] Holger H. Hoos and Thomas Stützle. *Stochastic Local Search: Foundations & Applications*. Elsevier / Morgan Kaufmann, 2004.

[42] Bernardo A. Huberman, Rajan M. Lukose, and Tad Hogg. An Economics Approach to Hard Computational Problems. *Science*, 275(5296):51–54, 1997.

[43] Frank Hutter, Domagoj Babic, Holger H Hoos, and Alan J Hu. Boosting Verification by Automatic Tuning of Decision Procedures. In *Proceedings of Formal Methods in Computer Aided Design (FMCAD'07)*, pages 27–34, 2007.

[44] Frank Hutter, Holger H Hoos, and Kevin Leyton-Brown. Automated Configuration of Mixed Integer Programming Solvers. In *Proceedings of the 7th International Conference on Integration of Artificial Intelligence (AI) and Operations Research (OR) Techniques in Constraint Programming (CPAIOR 2010)*, pages 186–202, 2010.

[45] Frank Hutter, Holger H Hoos, and Kevin Leyton-Brown. Sequential Model-Based Optimization for General Algorithm Configuration. In *Proceedings of the 5th International Conference on Learning and Intelligent Optimization (LION 5)*, pages 507–523, 2011.

[46] Frank Hutter, Holger H Hoos, Kevin Leyton-Brown, and Thomas Stützle. ParamILS: An Automatic Algorithm Configuration Framework. *Journal of Artificial Intelligence Research*, 36:267–306, 2009.

[47] Frank Hutter, Marius Lindauer, Adrian Balint, Sam Bayless, Holger Hoos, and Kevin Leyton-Brown. The Configurable SAT Solver Challenge (CSSC). *Artificial Intelligence*, 243:1–25, 2017.

[48] Frank Hutter, Lin Xu, Holger H Hoos, and Kevin Leyton-Brown. Algorithm Runtime Prediction: Methods & Evaluation. *Artificial Intelligence*, 206:79–111, 2014.

[49] Kai Jia and Martin C. Rinard. Efficient Exact Verification of Binarized Neural Networks. In Hugo Larochelle, Marc'Aurelio Ranzato, Raia Hadsell, Maria-Florina Balcan, and Hsuan-Tien Lin, editors, *Advances in Neural Information Processing Systems 33 (NeurIPS 2020)*, pages 1782–1795, 2020.

[50] Kyle D Julian, Mykel J Kochenderfer, and Michael P Owen. Deep Neural Network Compression for Aircraft Collision Avoidance Systems. *Journal of Guidance, Control, and Dynamics*, 42(3):598–608, 2019.

[51] Kyle D Julian, Jessica Lopez, Jeffrey S Brush, Michael P Owen, and Mykel J Kochenderfer. Policy compression for aircraft collision avoidance systems. In *Proceedings of the 35th Digital Avionics Systems Conference (DASC2016)*, pages 1–10, 2016.

[52] Serdar Kadioglu, Yuri Malitsky, Ashish Sabharwal, Horst Samulowitz, and Meinolf Sellmann. Algorithm Selection and Scheduling. In *Proceedings of the Seventeenth International Conference on Principles and Practice of Constraint Programming (CP2011)*, pages 454–469, 2011.

[53] Manisha M Kasar, Debnath Bhattacharyya, and TH Kim. Face recognition using neural network: a review. *International Journal of Security and Its Applications*, 10(3):81–100, 2016.

[54] Haniye Kashgarani and Lars Kotthoff. Is Algorithm Selection Worth It? Comparing Selecting Single Algorithms and Parallel Execution. In *AAAI Workshop on Meta-Learning and MetaDL Challenge*, pages 58–64, 2021.

[55] Guy Katz, Clark Barrett, David L Dill, Kyle Julian, and Mykel J Kochenderfer. Reluplex: An Efficient SMT Solver for Verifying Deep Neural Networks. In *Proceedings of the 29th International Conference on Computer Aided Verification (CAV 2017)*, pages 97–117, 2017.

[56] Guy Katz, Derek A. Huang, Duligur Ibeling, Kyle Julian, Christopher Lazarus, Rachel Lim, Parth Shah, Shantanu Thakoor, Haoze Wu, Aleksandar Zeljić, David L. Dill, Mykel J. Kochenderfer, and Clark Barrett. The Marabou Framework for Verification and Analysis of Deep Neural Networks. In *Proceedings of the 31st International Conference on Computer Aided Verification (CAV 2019)*, pages 443–452, 2019.

[57] Konstantin Kaulen, Matthias König, and Holger H Hoos. Dynamic algorithm termination for branch-and-bound-based neural network verification. In *To appear in Proceedings of the 39th AAAI Conference on Artificial Intelligence (AAAI-25)*, pages 1–9, 2025.

[58] Pascal Kerschke, Holger H Hoos, Frank Neumann, and Heike Trautmann. Automated Algorithm Selection: Survey and Perspectives. *Evolutionary Computation*, 27(1):3–45, 2019.

[59] Matthias König, Annelot W Bosman, Holger H Hoos, and Jan N van Rijn. Critically Assessing the State of the Art in CPU-based Local Robustness Verification. In *Proceedings of the Workshop on Artificial Intelligence Safety 2023 (SafeAI 2023) co-located with the Thirty-Seventh AAAI Conference on Artificial Intelligence (AAAI2023)*, pages 1–9, 2023.

[60] Matthias König, Annelot W Bosman, Holger H Hoos, and Jan N van Rijn. Critically Assessing the State of the Art in Neural Network Verification. *Journal of Machine Learning Research*, 25(12):1–53, 2024.

[61] Matthias König, Holger H Hoos, and Jan N van Rijn. Speeding up neural network robustness verification via algorithm configuration and an optimised mixed integer linear programming solver portfolio. *Machine Learning*, 111(12):4565–4584, 2022.

[62] Matthias König, Holger H Hoos, and Jan N van Rijn. Towards Algorithm-Agnostic Uncertainty Estimation: Predicting Classification Error in an Automated Machine Learning Setting. In *ICML Workshop on Automated Machine Learning*, pages 1–6, 2020.

[63] Matthias König, Holger H Hoos, and Jan N van Rijn. Speeding Up Neural Network Verification via Automated Algorithm Configuration. In *ICLR Workshop on Security and Safety in Machine Learning Systems*, pages 1–4, 2021.

[64] Matthias König, Holger H Hoos, and Jan N van Rijn. Accelerating Adversarially Robust Model Selection for Deep Neural Networks via Racing. In *Proceedings of the 38th AAAI Conference on Artificial Intelligence (AAAI-24)*, pages 21267–21275, 2024.

[65] Matthias König, Xiyue Zhang, Holger H Hoos, Marta Kwiatkowska, and Jan N van Rijn. Automated Design of Linear Bounding Functions for Sigmoidal Nonlinearities in Neural Networks. In *Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, 2024.

[66] Lars Kotthoff. Algorithm Selection for Combinatorial Search Problems: A Survey. In *Data Mining and Constraint Programming*, pages 149–190. Springer, 2016.

[67] Alexey Kurakin, Ian Goodfellow, and Samy Bengio. Adversarial examples in the physical world. *arXiv preprint arXiv:1607.02533*, 2016.

[68] Ailsa H. Land and Alison G. Doig. An Automatic Method for Solving Discrete Programming Problems. In *50 Years of Integer Programming 1958-2008 - From the Early Years to the State-of-the-Art*, pages 105–132. Springer, 2010.

[69] Mathias Lecuyer, Vaggelis Atlidakis, Roxana Geambasu, Daniel Hsu, and Suman Jana. Certified Robustness to Adversarial Examples with Differential Privacy. In *Proceedings of the 40th IEEE Symposium on Security and Privacy (SP2019)*, pages 656–672. IEEE, 2019.

[70] Kevin Leyton-Brown, Eugene Nudelman, Galen Andrew, Jim McFadden, and Yoav Shoham. A portfolio approach to algorithm selection. In *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence (IJCAI-03)*, pages 1542–1543, 2003.

[71] Linyi Li, Xiangyu Qi, Tao Xie, and Bo Li. Sok: Certified robustness for deep neural networks. *arXiv preprint arXiv:2009.04131*, 2020.

[72] Marius Lindauer, Holger H Hoos, Frank Hutter, and Torsten Schaub. AutoFolio: An Automatically Configured Algorithm Selector. *Journal of Artificial Intelligence Research*, 53:745–778, 2015.

[73] Changliu Liu, Tomer Arnon, Christopher Lazarus, Christopher A. Strong, Clark W. Barrett, and Mykel J. Kochenderfer. Algorithms for Verifying Deep Neural Networks. *Foundations and Trends in Optimization*, 4(3-4):244–404, 2021.

[74] Weiyang Liu, Yandong Wen, Zhiding Yu, and Meng Yang. Large-Margin Softmax Loss for Convolutional Neural Networks. In *Proceedings of the 33nd International Conference on Machine Learning (ICML 2016)*, pages 507–516, 2016.

[75] Alessio Lomuscio and Lalit Maganti. An approach to reachability analysis for feed-forward ReLU neural networks. *arXiv preprint arXiv:1706.07351*, 2017.

[76] Manuel Lopez-Ibanez and Thomas Stützle. Automatically improving the anytime behaviour of optimisation algorithms. *European Journal of Operational Research*, 235(3):569–582, 2014.

[77] Aleksander Madry, Aleksandar Makelov, Ludwig Schmidt, Dimitris Tsipras, and Adrian Vladu. Towards Deep Learning Models Resistant to Adversarial Attacks. *arXiv preprint arXiv:1706.06083*, 2017.

[78] Yuri Malitsky, Ashish Sabharwal, Horst Samulowitz, and Meinolf Sellmann. Parallel SAT Solver Selection and Scheduling. In *Proceedings of the Eighteenth International Conference on Principles and Practice of Constraint Programming (CP2012)*, pages 1–15, 2012.

[79] Oded Maron and Andrew Moore. Hoeffding Races: Accelerating Model Selection Search for Classification and Function Approximation. In *Advances in Neural Information Processing Systems 6 (NeurIPS 1993)*, pages 59–66, 1993.

[80] Mark Huasong Meng, Guangdong Bai, Sin Gee Teo, Zhe Hou, Yan Xiao, Yun Lin, and Jin Song Dong. Adversarial robustness of deep neural networks: A survey from a formal verification perspective. *IEEE Transactions on Dependable and Secure Computing*, pages 1–19, 2022.

[81] Saeed Mian Qaisar. Baseline wander and power-line interference elimination of ecg signals using efficient signal-piloted filtering. *Healthcare Technology Letters*, 7(4):114–118, 2020.

[82] Jeet Mohapatra, Ching-Yun Ko, Lily Weng, Pin-Yu Chen, Sijia Liu, and Luca Daniel. Hidden Cost of Randomized Smoothing. In *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics (AISTATS 2021)*, pages 4033–4041, 2021.

[83] Jeet Mohapatra, Tsui-Wei Weng, Pin-Yu Chen, Sijia Liu, and Luca Daniel. Towards Verifying Robustness of Neural Networks Against A Family of Semantic Perturbations. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2020)*, pages 241–249, 2020.

[84] Leonardo de Moura and Nikolaj Bjørner. Satisfiability Modulo Theories: An Appetizer. In *Proceedings of the Brazilian Symposium on Formal Methods (SBMF 2018)*, pages 23–36, 2009.

[85] Christoph Müller, François Serre, Gagandeep Singh, Markus Püschel, and Martin Vechev. Scaling Polyhedral Neural Network Verification on gpus. In *Proceedings of Machine Learning and Systems 3 (MLSys 2021)*, pages 1–14, 2021.

[86] Mark Niklas Müller, Gleb Makarchuk, Gagandeep Singh, Markus Püschel, and Martin Vechev. PRIMA: General and Precise Neural Network Certification via Scalable Convex Hull Approximations. In *Proceedings of the 49th ACM SIGPLAN Symposium on Principles of Programming Languages (POPL 2022)*, pages 1–33, 2022.

[87] Mark Niklas Müller, Christopher Brix, Stanley Bak, Changliu Liu, and Taylor T. Johnson. The Third International Verification of Neural Networks Competition (VNN-COMP 2022): Summary and Results. *arXiv preprint arXiv:2212.10376*, 2023.

[88] Nina Narodytska, Shiva Prasad Kasiviswanathan, Leonid Ryzhyk, Mooly Sagiv, and Toby Walsh. Verifying Properties of Binarized Deep Neural Networks. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI-18)*, pages 6615–6624, 2018.

[89] Nicolas Papernot, Patrick McDaniel, Xi Wu, Somesh Jha, and Ananthram Swami. Distillation as a Defense to Adversarial Perturbations Against Deep Neural Networks. In *Proceedings of the 37th IEEE Symposium on Security and Privacy (IEEE S&P 2016)*, pages 582–597, 2016.

[90] Luca Pulina and A. Tacchella. Checking Safety of Neural Networks with SMT Solvers: A Comparative Evaluation. In *AI*IA*, pages 127–138, 2011.

[91] Luca Pulina and A. Tacchella. NeVer: A Tool for Artificial Neural Networks Verification. *Annals of Mathematics and Artificial Intelligence*, pages 403–425, 2011.

[92] Luca Pulina and Armando Tacchella. Challenging SMT Solvers to Verify Neural Networks. *AI Communications*, pages 117–135, 2012.

[93] Aditi Raghunathan, Jacob Steinhardt, and Percy Liang. Certified Defenses against Adversarial Examples. *arXiv preprint arXiv:1801.09344*, 2018.

[94] Karsten Scheibler, Leonore Winterer, Ralf Wimmer, and Bernd Becker. Towards Verification of Artificial Neural Networks. In *Proceedings of the 18th Workshop on Methoden und Beschreibungssprachen zur Modellierung und Verifikation von Schaltungen und Systemen (MBMV 2015)*, pages 30–40, 2015.

[95] David Shriver, Sebastian Elbaum, and Matthew B. Dwyer. DNNV: A Framework for Deep Neural Network Verification. In *Proceedings of the 33rd International Conference on Computer Aided Verification (CAV 2021)*, pages 137–150, 2021.

[96] 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 (NeurIPS 2019)*, pages 15072–15083, 2019.

[97] Gagandeep Singh and Timon Gehr. Boosting Robustness Certification of Neural networks. In *Proceedings of the 7th International Conference on Learning Representations (ICLR 2019)*, pages 1–12, 2019.

[98] Gagandeep Singh, Timon Gehr, Matthew Mirman, Markus Püschel, and Martin Vechev. Fast and Effective Robustness Certification. In *Advances in Neural Information Processing Systems 31 (NeurIPS 2018)*, pages 10825–10836, 2018.

[99] Gagandeep Singh, Timon Gehr, Markus Püschel, and Martin Vechev. An Abstract Domain for Certifying Neural Networks. In *Proceedings of the 46th ACM SIGPLAN Symposium on Principles of Programming Languages (POPL 2019)*, pages 1–30, 2019.

[100] Weidi Sun, Yuteng Lu, Xiyue Zhang, and Meng Sun. DeepGlobal: A framework for global robustness verification of feedforward neural networks. *Journal of Systems Architecture*, 128:102582, 2022.

[101] Christian Szegedy, Wojciech Zaremba, Ilya Sutskever, Joan Bruna, Dumitru Erhan, Ian J. Goodfellow, and Rob Fergus. Intriguing properties of neural networks. In *Proceedings of the 2nd International Conference on Learning Representations (ICLR 2014)*, pages 1–10, 2014.

[102] Merle W Tate and Sara M Brown. Note on the Cochran Q test. *Journal of the American Statistical Association*, 65(329):155–160, 1970.

[103] Chris Thornton, Frank Hutter, Holger H Hoos, and Kevin Leyton-Brown. Auto-WEKA: Combined Selection and Hyperparameter Optimization of Classification Algorithms. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD2013)*, pages 847–855, 2013.

[104] Vincent Tjeng, Kai Xiao, and Russ Tedrake. Evaluating Robustness of Neural Networks with Mixed Integer Programming. In *Proceedings of the 7th International Conference on Learning Representations (ICLR 2019)*, pages 1–21, 2019.

[105] Alan Turing. Intelligent machinery (1948). *B. Jack Copeland*, pages 395–432, 2004.

[106] European Union. Regulation (eu) 2024/1689 of the european parliament and of the council of 13 june 2024 laying down harmonised rules on artificial intelligence and amending certain union legislative acts (artificial intelligence act), 2024.

[107] Mauro Vallati, Chris Fawcett, Alfonso Emilio Gerevini, Holger Hoos, and Alessandro Saetti. Automatic Generation of Efficient Domain-Specific Planners from Generic Parametrized Planners. In *Proceedings of the 6th Annual Symposium on Combinatorial Search (SOCS)*, pages 184–192, 2013.

[108] Bruno Veloso, Luciano Caroprese, Matthias König, Sónia Teixeira, Giuseppe Manco, Holger H Hoos, and João Gama. Hyper-Parameter Optimization for Latent Spaces in Dynamic Recommender Systems. In *Proceedings of the European*

Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, pages 249–264, 2021.

[109] Shiqi Wang, Kexin Pei, Justin Whitehouse, Junfeng Yang, and Suman Jana. Efficient Formal Safety Analysis of Neural Networks. In *Advances in Neural Information Processing Systems 31 (NeurIPS 2018)*, pages 6369–6379, 2018.

[110] Shiqi Wang, Kexin Pei, Justin Whitehouse, Junfeng Yang, and Suman Jana. Formal Security Analysis of Neural Networks using Symbolic Intervals. In *Proceedings of the 27th USENIX Security Symposium (USENIX Security 18)*, pages 1599–1614, 2018.

[111] 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 *Advances in Neural Information Processing Systems 34 (NeurIPS 2021)*, pages 29909–29921, 2021.

[112] Lily Weng, Huan Zhang, Hongge Chen, Zhao Song, Cho-Jui Hsieh, Luca Daniel, Duane Boning, and Inderjit Dhillon. Towards Fast Computation of Certified Robustness for ReLU Networks. In *Proceedings of the 35th International Conference on Machine Learning (ICML 2018)*, pages 5276–5285, 2018.

[113] Eric Wong and Zico Kolter. Provable Defenses against Adversarial Examples via the Convex Outer Adversarial Polytope. In *Proceedings of the 35th International Conference on Machine Learning (ICML 2018)*, pages 5286–5295, 2018.

[114] Haoze Wu, Aleksandar Zeljic, Guy Katz, and Clark W. Barrett. Efficient Neural Network Analysis with Sum-of-Infeasibilities. In Dana Fisman and Grigore Rosu, editors, *Proceedings of the 28th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS 2022)*, volume 13243, pages 143–163, 2022.

[115] Weiming Xiang, Hoang-Dung Tran, and Taylor 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.

[116] Lin Xu, Holger Hoos, and Kevin Leyton-Brown. Hydra: Automatically Configuring Algorithms for Portfolio–Based Selection. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI–10)*, pages 210–216, 2010.

[117] Lin Xu, Frank Hutter, Holger H Hoos, and Kevin Leyton-Brown. SATzilla: Portfolio-based Algorithm Selection for SAT. *Journal of Artificial Intelligence Research*, 32:565–606, 2008.

[118] Lin Xu, Frank Hutter, Holger H Hoos, and Kevin Leyton-Brown. Hydra-MIP: Automated Algorithm Configuration and Selection for Mixed Integer Programming. In *RCRA Workshop on Experimental evaluation of Algorithms for Solving Problems with Combinatorial Explosion*, pages 16–30, 2011.

[119] Huan Zhang, Tsui-Wei Weng, Pin-Yu Chen, Cho-Jui Hsieh, and Luca Daniel. Efficient Neural Network Robustness Certification with General Activation Functions. In *Advances in Neural Information Processing Systems 31 (NeurIPS 2018)*, volume 31, pages 4944–4953, 2018.

[120] Jingfeng Zhang, Xilie Xu, Bo Han, Gang Niu, Lizhen Cui, Masashi Sugiyama, and Mohan S. Kankanhalli. Attacks Which Do Not Kill Training Make Adversarial Learning Stronger. In *Proceedings of the 37th International Conference on Machine Learning (ICML 2020)*, pages 11278–11287, 2020.