Lipschitz-bounded convolutional neural networks

Patricia Pauli, Frank Allgöwer

University of Stuttgart, Institute for Systems Theory and Automatic Control,
Pfaffenwaldring 9, 70569 Stuttgart, Germany
patricia.pauli@ist.uni-stuttgart.de

Convolutional neural networks (CNNs) are successfully applied in many fields, e.g., they are the state of the art in image and audio processing tasks. However, as CNNs are black box models, their behavior is not fully understood which is especially problematic in safety-critical applications such as autonomous driving and medical devices. Consequently, there is a need to augment neural networks by including safety and robustness guarantees. To quantify robustness for neural networks, we use the Lipschitz constant of the input-output mapping characterized by the CNN which has become common practice [1]. In this context, we address two problem setups. On the one hand, we aim to determine an accurate upper bound on the Lipschitz constant for given CNNs [2, 3] and on the other hand, we propose a parameterization to design Lipschitz-bounded CNNs, i.e., we find a description of a CNN with built-in guarantees on the Lipschitz constant [4].

The proposed methods for both analysis and training of CNNs utilize well-established techniques from control theory, including semidefinite programming. We particularly take the perspective of a control engineer onto neural networks using tools and property definitions that have proven to be expedient over decades to now address problems in the field of deep learning.

The calculation of the Lipschitz constant for neural networks is an NP-hard problem and it is hence not feasible for large and deep NNs. Instead, we find an accurate upper bound on the Lipschitz constant. To do this, we overapproximate common nonlinear activation functions, such as ReLU, tanh, sigmoid, and MaxMin via incremental quadratic constraints. They satisfy certain properties, e.g., they are slope-restricted, which we exploit to formulate linear matrix inequalities (LMIs) that are sufficient for Lipschitz continuity of the CNNs. The Lipschitz constant estimation problem then boils down to a semidefinite program, which is a convex optimizazion problem [5, 6].

To further reduce the computational effort and increase the scalability of the method, we exploit the structure of the LMIs which form the constraints in the semidefinite program. Rather than formulating one big and sparse LMI constraint, we break the LMI down into multiple LMIs, yielding exactly one LMI per layer. Hereby, the underlying idea once again is inspired by control theory. We enforce disspativity onto all individual layers in the CNN, that typically consists of convolutional layers, pooling layers, and fully-connected layers as shown in Figure 1. Finally, we connect the layers through their dissipativity properties to determine an upper bound on the Lipschitz constant of the input-output mapping [2].

In some applications, it might not only be useful to verify robustness for given CNNs, but it may also be desirable to design robust neural networks with guarantees on the Lipschitz constant. Thus we establish a parameterization of

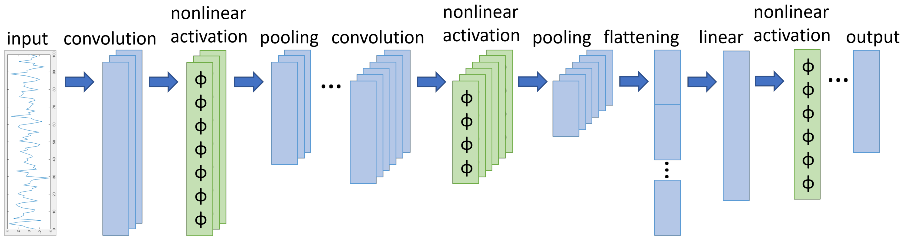

Figure 1: Illustration of a 1D convolutional neural network.

CNNs with built-in guarantees on the Lipschitz constant. The way the parameters of the CNNs are designed ensures that the LMI conditions that ensure Lipschitz continuity for the CNN are satisfied. Our parameterization uses the Cayley transform and the controllability gramian, mathematical tools that are known from Riemannian optimization and control theory, respectively.

References

[1] C. Szegedy, W. Zaremba, I. Sutskever, J. Bruna, D. Erhan, I. Goodfellow, and R. Fergus, "Intriguing properties of neural networks," *arXiv preprint arXiv:1312.6199*, 2013.

[2] P. Pauli, D. Gramlich, and F. Allgöwer, "Lipschitz constant estimation for 1D convolutional neural networks," *arXiv preprint arXiv:2211.15253*, 2022.

[3] D. Gramlich, P. Pauli, C. W. Scherer, F. Allgöwer, and C. Ebenbauer, "Convolutional neural networks as 2-D systems," 2023, to be published.

[4] P. Pauli, R. Wang, I. R. Manchester, and F. Allgöwer, "Lipschitz-bounded 1d convolutional neural networks using the cayley transform and the controllability gramian," 2023, to be published.

[5] M. Fazlyab, A. Robey, H. Hassani, M. Morari, and G. Pappas, "Efficient and accurate estimation of Lipschitz constants for deep neural networks," *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[6] P. Pauli, A. Koch, J. Berberich, P. Kohler, and F. Allgöwer, "Training robust neural networks using Lipschitz bounds," *IEEE Control Systems Letters*, vol. 6, pp. 121–126, 2022.