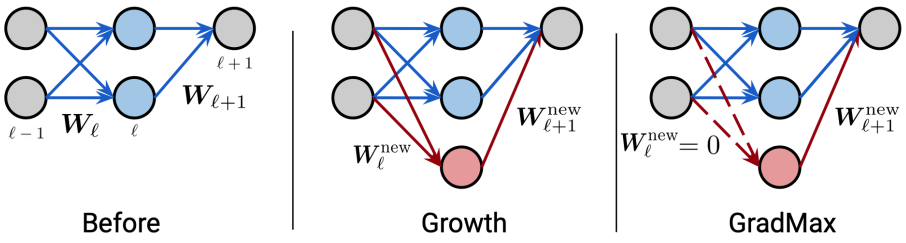
TECHNISCHE UNIVERSITÄT DARMSTADT

Thesis (B.Sc. / M.Sc.)
Growing Neural Networks



A great portion of the advances in machine learning in recent years can be attributed to a significant advance in the theory and realizations of artificial Neural Networks (NNs). The from the brain inspired Neural Networks consist of single neurons in at least two different layers. The neurons in neighbouring layers communicate which each other over weighted connections. Therefore "learning" in the context of NNs means tuning the weights of the connections between neurons such that the final result is the desired one.

Even though NNs improve their software partially by themselves, generally there needs to be a set underlying architecture possessing non-negligible influence on the Networks performance [1]. Depending on the task, said architecture can be complex or simple, however, there is no fixed rule on how to build it reliable and efficient. Therefore one cannot be certain if other (possibly unknown) architectures would result in a better performance for the specific problem at hand. Additionally, training the model and afterwards changing the architecture is often difficult since this would result in undesired changes of outcome due to newly introduced interactions with the new neurons. The optimal solution to this problem would be to allow the NN to reliably change its architecture by itself during the training process.

Searching for the best architecture is an active research area incorporating many different approaches like neural architecture search [2], evolutionary algorithms [3] or pruning [4]. However, these approaches are generally costly in terms of time and computational power. Other approaches rely on adding new neurons resulting in immediate improvements of the training objective. One recently proposed idea called **GradMax** tries to solve this problem from another angle by improving the training dynamics instead of the immediate response [5].

In this project we are investigating the novel GradMax approach, specifically the questions **When**, **Where** and **How** a layer of neurons should be added to Convolutional Neural Networks used in the regime of computer vision.
Alternatively one could investigate the question if and how such an approach can be combined with an degenerative approach like pruning, possibly resulting in novel effects of NNs. Ideally, the prospective student has already basic knowledge of (Convolutional) Neural Networks and python programming skills.

For further information, please contact Sascha H. Hauck.

Fachbereich 18
Elektrotechnik und Informationstechnik
Selbstorganisierende Systeme

Department 18
Electrical Engineering and Information Technology
Self-Organizing Systems

Prof. Dr. Heinz Koeppl
Head of lab

Sascha H. Hauck
Project supervisor

Rundeturmstraße 12
64283 Darmstadt

Phone: +49 6151 16 - 57 243
sascha.hauck@tu-darmstadt.de
https://www.bcs.tu-darmstadt.de

Date
March 22, 2022

References

[1] Behnam Neyshabur. Towards learning convolutions from scratch. https://arxiv.org/pdf/2007.13657.pdf, 2020.

[2] Thomas Elsken, Jan Hendrik Metzen, and Frank Hutter. Neural architecture search: A survey. *Journal of Machine Learning Research*, 20:1–21, 2019.

[3] Kenneth O. Stanley and Risto Miikkulainen. Evolving neural networks through augmenting topologies. *Evolutionary Computation*, 10, 2002.

[4] Thomas Liu, Mingjie Sun, Tinghui Zhou, Gao Huang, and Trevor Darrell. Rethinking the vlaue of network pruning. https://arxiv.org/pdf/1810.05270.pdf, 2018.

[5] Utku Evci, Bart van Merrienboer, Thomas Unterthiner, Max Vladymyrov, and Fabian Pedregosa. Gradmax: Growing neural networks using gradient information. https://arxiv.org/pdf/2201.05125.pdf, 2022.