



Stochastic Weight Matrix-Based Regularization Methods for Deep Neural Networks

International Conference on Machine Learning, Optimization, and Data Science

LOD 2019: Machine Learning, Optimization, and Data Science pp 45-57 | Cite as

- Patrik Reizinger (1) Email author (rpatrik96@sch.bme.hu) View author's OrcID profile (View OrcID profile)
- Bálint Gyires-Tóth (1)

1. Budapest University of Technology and Economics, , Budapest, Hungary

Conference paper

First Online: 03 January 2020

- 790 Downloads

Part of the [Lecture Notes in Computer Science](#) book series (LNCS, volume 11943)

Abstract

The aim of this paper is to introduce two widely applicable regularization methods based on the direct modification of weight matrices. The first method, Weight Reinitialization, utilizes a simplified Bayesian assumption with partially resetting a sparse subset of the parameters. The second one, Weight Shuffling, introduces an entropy- and weight distribution-invariant non-white noise to the parameters. The latter can also be interpreted as an ensemble approach. The proposed methods are evaluated on benchmark datasets, such as MNIST, CIFAR-10 or the JSB Chorales database, and also on time series modeling tasks. We report gains both regarding performance and entropy of the analyzed networks. We also made our code available as a GitHub repository (<https://github.com/rpatrik96/lod-wmm-2019>)

Your privacy

We use cookies to make sure that our website works properly, as well as some "optional" cookies to personalise content and advertising, provide social media features and analyse how people use our site. By accepting some or all optional cookies you give consent to the processing of your personal data, including transfer to third parties, some in countries outside of the European Economic Area that do not offer the same data protection standards as the country where you live. You can decide which optional cookies to accept by clicking on "Manage Settings", where you can also find more information about how your personal data is processed. **View our privacy policy**

[Manage Settings](#)

[Accept All Cookies](#)

Sciences. We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research. Patrik Reizinger expresses his gratitude for the financial support of the Nokia Bell Labs Hungary.

This is a preview of subscription content, [log in](#) to check access.

References

1. Ignite (2019). <https://github.com/pytorch/ignite> (<https://github.com/pytorch/ignite>)
2. Bergstra, J., Bardenet, R., Bengio, Y., Kégl, B.: Algorithms for hyper-parameter optimization. In: Advances in Neural Information Processing Systems (NIPS), pp. 2546–2554 (2011). <https://doi.org/2012arXiv1206.2944S> (<https://doi.org/2012arXiv1206.2944S>)
3. Cho, K., et al.: Learning Phrase Representations Using RNN Encoder-Decoder for Statistical Machine Translation. arXiv (2014). <https://doi.org/10.1074/jbc.M608066200> (<https://doi.org/10.1074/jbc.M608066200>)
4. Glorot, X., Bengio, Y.: Understanding the difficulty of training deep feedforward neural networks. PMLR **9**, 249–256 (2010). <https://doi.org/10.1.1.207.2059> (<https://doi.org/10.1.1.207.2059>)
5. Guennec, A.L., et al.: Data augmentation for time series classification using convolutional neural networks. In: ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data (2016) [Google Scholar](#) (<https://scholar.google.com/scholar?q=Guenne%2C%20A.L.%2C%20et%20al.%3A%20Data%20augmentation%20for%20time%20series%20classification%20using%20convolutional%20neural%20networks.%20In%3A%20ECML%20PKDD%20Workshop%20on%20Advanced%20Analytics%20and%20Learning%20on%20Temporal%20Data%20%282016%29>)
6. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**(8), 1735–1780 (1997). <https://doi.org/10.1162/neco.1997.9.8.1735> (<https://doi.org/10.1162/neco.1997.9.8.1735>) [CrossRef](#) (<https://doi.org/10.1162/neco.1997.9.8.1735>) [Google Scholar](#) (http://scholar.google.com/scholar_lookup?title=Long%20short-

Your privacy

We use cookies to make sure that our website works properly, as well as some "optional" cookies to personalise content and advertising, provide social media features and analyse how people use our site. By accepting some or all optional cookies you give consent to the processing of your personal data, including transfer to third parties, some in countries outside of the European Economic Area that do not offer the same data protection standards as the country where you live. You can decide which optional cookies to accept by clicking on "Manage Settings", where you can also find more information about how your personal data is processed. **View our privacy policy**

Manage Settings

Accept All Cookies

<https://doi.org/10.1227/01.NEU.0000210260.55124.A4>

(<https://doi.org/10.1227/01.NEU.0000210260.55124.A4>).

<http://arxiv.org/abs/1606.01305> (<http://arxiv.org/abs/1606.01305>)

10. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. *Proc. IEEE* **86**(11), 2278–2323 (1998).
<https://doi.org/10.1109/5.726791> (<https://doi.org/10.1109/5.726791>)
[CrossRef](#) (<https://doi.org/10.1109/5.726791>)
[Google Scholar](#) (http://scholar.google.com/scholar_lookup?title=Gradient-based%20learning%20applied%20to%20document%20recognition&author=Y.%20LeCun&author=L.%20Bottou&author=Y.%20Bengio&author=P.%20Haffner&journal=Proc.%20IEEE&volume=86&issue=11&pages=2278-2323&publication_year=1998&doi=10.1109%2F5.726791)
11. Masters, D., Luschi, C.: Revisiting Small Batch Training for Deep Neural Networks. *arXiv*, pp. 1–18 (2018).
<https://doi.org/10.1016/j.biortech.2007.04.007>
(<https://doi.org/10.1016/j.biortech.2007.04.007>).
<http://arxiv.org/abs/1804.07612> (<http://arxiv.org/abs/1804.07612>)
12. Neil, D., Pfeiffer, M., Liu, S.C.: Phased LSTM: accelerating recurrent network training for long or event-based sequences. In: Lee, D.D., Sugiyama, M., Luxburg, U.V., Guyon, I., Garnett, R. (eds.) *Advances in Neural Information Processing Systems* 29, pp. 3882–3890. Curran Associates, Inc. (2016)
[Google Scholar](#) (<https://scholar.google.com/scholar?q=Neil%2C%20D.%2C%20Pfeiffer%2C%20M.%2C%20Liu%2C%20S.C.%3A%20Phased%20LSTM%3A%20accelerating%20recurrent%20network%20training%20for%20long%20or%20event-based%20sequences.%20In%3A%20Lee%2C%20D.D.%2C%20Sugiyama%2C%20M.%2C%20Luxburg%2C%20U.V.%2C%20Guyon%2C%20I.%2C%20Garnett%2C%20R.%20%28eds.%29%20Advances%20in%20Neural%20Information%20Processing%20Systems%2029%2C%20pp.%203882%E2%80%933890.%20Curran%20Associates%2C%20Inc.%20%282016%29>)
13. Paszke, A., et al.: Automatic differentiation in PyTorch. In: *Advances in Neural Information Processing Systems* 30 (NIPS), pp. 1–4 (2017)
[Google Scholar](#) (<https://scholar.google.com/scholar?q=Paszke%2C%20A.%2C%20et%20al.%3A%20Automatic%20differentiation%20in%20PyTorch.%20In%3A%20Advances%20in%20Neural%20Information%20Processing%20Systems%2030%20%28NIPS%29%2C%20pp.%201%E2%80%934.%20%282017%29>)

Your privacy

We use cookies to make sure that our website works properly, as well as some "optional" cookies to personalise content and advertising, provide social media features and analyse how people use our site. By accepting some or all optional cookies you give consent to the processing of your personal data, including transfer to third parties, some in countries outside of the European Economic Area that do not offer the same data protection standards as the country where you live. You can decide which optional cookies to accept by clicking on "Manage Settings", where you can also find more information about how your personal data is processed. **View our privacy policy**

Manage Settings

Accept All Cookies

[CrossRef](https://doi.org/10.1214/12-AOS1000) (<https://doi.org/10.1214/12-AOS1000>)

[MathSciNet](http://www.ams.org/mathscinet-getitem?mr=3231592) (<http://www.ams.org/mathscinet-getitem?mr=3231592>)

[zbMATH](http://www.emis.de/MATH-item?1318.68153) (<http://www.emis.de/MATH-item?1318.68153>)

[Google Scholar](http://scholar.google.com/scholar_lookup?title=Dropout%3A%20a%20simple%20way%20to%20prevent%20neural%20networks%20from%20overfitting&author=N.%20Srivastava&author=G.%20Hinton&author=A.%20Krizhevsky&author=I.%20Sutskever&author=R.%20Salakhutdinov&journal=J.%20Mach.%20Learn.%20Res.&volume=15&pages=1929-1958&publication_year=2014&doi=10.1214%2F12-AOS1000) (http://scholar.google.com/scholar_lookup?title=Dropout%3A%20a%20simple%20way%20to%20prevent%20neural%20networks%20from%20overfitting&author=N.%20Srivastava&author=G.%20Hinton&author=A.%20Krizhevsky&author=I.%20Sutskever&author=R.%20Salakhutdinov&journal=J.%20Mach.%20Learn.%20Res.&volume=15&pages=1929-1958&publication_year=2014&doi=10.1214%2F12-AOS1000)

17. Wan, L., Zeiler, M., Zhang, S., LeCun, Y., Fergus, R.: Regularization of neural networks using DropConnect. In: ICML, pp. 109–111 (2013).
<https://doi.org/10.1109/TPAMI.2017.2703082>
(<https://doi.org/10.1109/TPAMI.2017.2703082>)

Copyright information

© Springer Nature Switzerland AG 2019

About this paper

Cite this paper as:

Reizinger P., Gyires-Tóth B. (2019) Stochastic Weight Matrix-Based Regularization Methods for Deep Neural Networks. In: Nicosia G., Pardalos P., Umeton R., Giuffrida G., Sciacca V. (eds) Machine Learning, Optimization, and Data Science. LOD 2019. Lecture Notes in Computer Science, vol 11943. Springer, Cham.
https://doi.org/10.1007/978-3-030-37599-7_5

- First Online 03 January 2020
- DOI https://doi.org/10.1007/978-3-030-37599-7_5
- Publisher Name Springer, Cham
- Print ISBN 978-3-030-37598-0
- Online ISBN 978-3-030-37599-7
- eBook Packages [Computer Science](#) [Computer Science \(Ro\)](#)
- [Buy this book on publisher's site](#)
- [Reprints and Permissions](#)

Your privacy

We use cookies to make sure that our website works properly, as well as some "optional" cookies to personalise content and advertising, provide social media features and analyse how people use our site. By accepting some or all optional cookies you give consent to the processing of your personal data, including transfer to third parties, some in countries outside of the European Economic Area that do not offer the same data protection standards as the country where you live. You can decide which optional cookies to accept by clicking on "Manage Settings", where you can also find more information about how your personal data is processed. **View our privacy policy**

Manage Settings

Accept All Cookies