Bernstein Conference 2019

An energy-based model of folded autoencoders for unsupervised learning in cortical hierarchies

Dominik Dold [1, 2], João Sacramento [3], Akos F. Kungl [1, 2], Walter Senn [2], Mihai A. Petrovici [1, 2]

1. Kirchhoff Institute for Physics, Heidelberg University, Im Neuenheimer Feld 227, 69120 Heidelberg, Germany
2. Department of Physiology, University of Bern, Bühlplatz 5, 3012 Bern, Switzerland
3. Institute of Neuroinformatics, University of Zurich / ETH Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

Recently, the problem of credit assignment in cortical networks has been addressed by several models suggesting a biologically plausible implementation of backprop [1], e.g., by drawing parallels to predictive coding [2] or proposing a circuit-level implementation using interneurons [3-5]. However, these models have so far been restricted to supervised learning.

Here, we propose an extension of these models to unsupervised learning by using a layer-wise recurrent network architecture with convex gating of the forward and backward information flow, controlled by λ. Similar to [2,4], the neurosynaptic dynamics are derived as gradient descent on an energy function composed of two squared error terms and a cost function, $E = \frac{\lambda}{2}\sum_i ||u_i - W_i r_{i-1}||^2 + \frac{1-\lambda}{2}\sum_i ||u_i - G_i r_{i+1}||^2 + \beta C$, where u_i and r_i are the membrane potentials and rates of neurons in layer i, W_i the discriminative weights (DW) projecting from layer i-1 to i, G_i the generative weights (GW) from layer i+1 to i and βC the cost function weighted by a scalar $\beta \geq 0$ (Fig. 1A). This way, we obtain standard leaky dynamics where forward and backward inputs are convexly combined at the soma (Fig. 1B). The resulting synaptic plasticity for W_i and G_i is driven by the dendritic prediction of somatic activity [6]. For small gating λ the plasticity rules for GW and DW, even though they are formally identical, perform different optimization tasks: the GW minimize a reconstruction error in the visible layer, whereas the DW learn to match the generative input entering the same layer.

Different from previous models [7-12], this network allows the simultaneous training of encoding (DW) and decoding (GW) weights in a deep folded autoencoder with a bottleneck in the highest layer (Fig. 1C,D). Both the encoding, decoding as well as the error propagation for the plasticity of the generative weights is done via the same neurons simultaneously. In addition, the visible layer is not clamped during training but only nudged towards the correct activity. The model can be directly connected to the microcircuits proposed in [3,4] by having the generative weights and errors project to apical compartments, and forward ones to basal compartments of pyramidal neurons (Fig. 1B). Thus, the presented model proposes a biologically plausible implementation of efficient simultaneous discriminative and generative learning in cortical hierarchies.

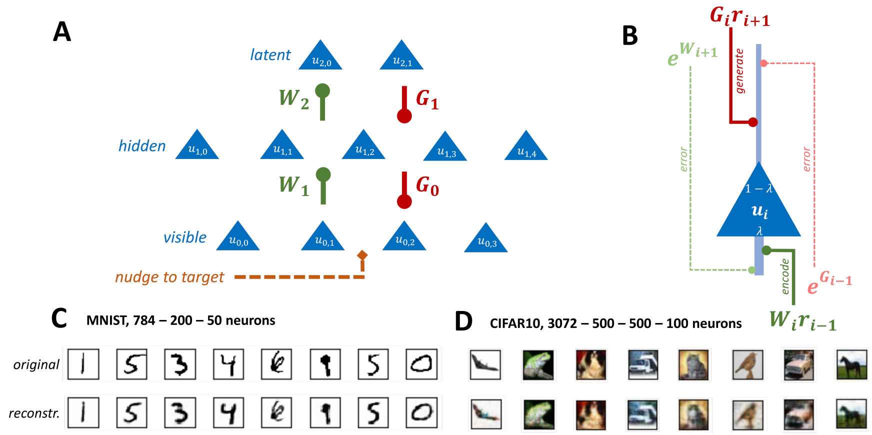

Figure 1: (A) Sketch of network architecture. (B) Physiological implementation of derived dynamics. (C) Encoding and decoding of MNIST images. We first encode the image with gating 0.9 and decode with gating 0.1. During training, the gating is kept constant at 0.1. (D) Same as (C) but for CIFAR10.

Acknowledgements

This work has received funding from the Manfred Stärk Foundation and the European Union Horizon 2020 Framework Programme (grant agreement 720270,785907). Calculations were performed on UBELIX (University of Bern HPC) and bwHPC (state BaWü HPC, funded by the DFG through grant no INST 39/963-1 FUGG).

References

1. Whittington, J. C., & Bogacz, R. (2019). Theories of error back-propagation in the brain. Trends in cognitive sciences.
2. Whittington, J. C., & Bogacz, R. (2017). An approximation of the error backpropagation algorithm in a predictive coding network with local Hebbian synaptic plasticity. Neural computation, 29(5), 1229-1262.
3. Sacramento, J., Costa, R. P., Bengio, Y., & Senn, W. (2018). Dendritic cortical microcircuits approximate the backpropagation algorithm. In Advances in Neural Information Processing Systems (pp. 8721-8732).
4. Dold, D., Kungl, A. F., Sacramento, J. , Petrovici, M. A., Schindler, K., Binas, J., Bengio, Y., & Senn, W. (2019). Lagrangian dynamics of dendritic microcircuits enables real-time backpropagation of errors. Cosyne Abstracts 2019, Lisbon, PT.
5. Guerguiev, J., Lillicrap, T. P., & Richards, B. A. (2017). Towards deep learning with segregated dendrites. ELife, 6, e22901.
6. Urbanczik, R., & Senn, W. (2014). Learning by the dendritic prediction of somatic spiking. Neuron, 81(3), 521-528.
7. Oh, J. H., & Seung, H. S. (1998). Learning generative models with the up propagation algorithm. In Advances in Neural Information Processing Systems (pp. 605-611).
8. Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nature neuroscience, 2(1), 79.
9. Seung, H. S. (1998). Learning continuous attractors in recurrent networks. In Advances in neural information processing systems (pp. 654-660).
10. Wang, J., He, H., & Prokhorov, D. V. (2012). A folded neural network autoencoder for dimensionality reduction. Procedia Computer Science, 13, 120-127.
11. Burbank, K. S. (2015). Mirrored STDP implements autoencoder learning in a network of spiking neurons. PLoS computational biology, 11(12), e1004566.

12. Pontes-Filho, S., & Liwicki, M. (2018). Bidirectional Learning for Robust Neural Networks. arXiv preprint arXiv:1805.08006.

Citation: Dold D, Sacramento J, Kungl AF, Senn W, Petrovici MA (2019) An energy-based model of folded autoencoders for unsupervised learning in cortical hierarchies. Bernstein Conference 2019.
doi: 10.12751/nncn.bc2019.0156 (http://doi.org/10.12751/nncn.bc2019.0156)