Sleep-Like Unsupervised Replay Improves Performance when Data are Limited or Unbalanced (Student Abstract)

Anthony Bazhenov[1,*], **Pahan Dewasurendra**[1,*], **Giri Krishnan**[2], **Jean Erik Delanois**[2,3]

[1] Del Norte High School, 16601 Nighthawk Ln, San Diego, CA 92127
[2] Department of Medicine, University of California, San Diego
[3] Department of Computer Science & Engineering, University of California, San Diego
9500 Gilman Drive, La Jolla, CA 92092
anthonyb11c234@gmail.com, pahan.dewa@gmail.com, gkrishnan@ucsd.edu, jdelanois@ucsd.edu

Abstract

The performance of artificial neural networks (ANNs) degrades when training data are limited or imbalanced. In contrast, the human brain can learn quickly from just a few examples. Here, we investigated the role of sleep in improving the performance of ANNs trained with limited data on the MNIST and Fashion MNIST datasets. Sleep was implemented as an unsupervised phase with local Hebbian type learning rules. We found a significant boost in accuracy after the sleep phase for models trained with limited data in the range of 0.5-10% of total MNIST or Fashion MNIST datasets. When more than 10% of the total data was used, sleep alone had a slight negative impact on performance, but this was remedied by fine-tuning on the original data. This study sheds light on a potential synaptic weight dynamics strategy employed by the brain during sleep to enhance memory performance when training data are limited or imbalanced.

Introduction

Deep learning methods have shown considerable performance when training datasets are large, however, existing techniques generally fail in low training data conditions. Additionally, training datasets are often imbalanced, with some categories occurring more frequently than others, resulting in reduced accuracy for ANNs. Several methods have been proposed to overcome these limitations. These include data augmentation (Shorten and Khoshgoftaar 2019), pretraining on other datasets (Zhuang et al. 2020) or alternative architectures such as neural tangent kernel (Arora et al. 2019). However, these approaches do not address the fundamental question of how to make overparameterized deep learning networks learn to generalize from small datasets without overfitting. In contrast, the human brain demonstrates the ability to learn quickly from just a few examples.

Sleep has been shown to play an important role in memory consolidation in biological systems (Stickgold 2005). Two critical components which are believed to underlie memory consolidation during sleep are spontaneous replay of memory traces and local unsupervised synaptic plasticity that restricts synaptic changes to relevant memories only. During sleep, replay of recently learned memories along with relevant old memories enables the network to form stable long-term memory representations (Rasch and Born 2013) and reduces competition between memories (González et al. 2020; Golden et al. 2022). The idea of replay has been explored in machine learning to enable continual learning. However, spontaneous unsupervised replay found in the biological brain and implemented here is significantly different compared to explicit replay of past inputs implemented in machine learning rehearsal methods (Hayes et al. 2021).

These results from neuroscience suggest that applying sleep replay principles to ANNs may enhance memory representations and, consequently, improve the performance of machine learning models trained on limited or unbalanced datasets, as tested in our study.

Algorithm

A fully-connected ANN with two hidden layers was first trained on a randomly selected subset of MNIST or Fashion MNIST (FMNIST) datasets using backpropagation. Subsequently, the sleep replay consolidation (SRC) algorithm was implemented as previously described in (Tadros et al. 2022). Briefly (see Supplementary Material for details), the ANN trained by limited data was mapped to a spiking neural network (SNN) with the same architecture. The SNN's activity was driven by randomly distributed Poisson spiking input that reflected average inputs observed in the training dataset. Local Hebbian-type plasticity was implemented to modify weights during the sleep phase, i.e., synaptic strength was increased if presynaptic activation was followed by postsynaptic activation and reduced if postsynaptic activation occurred without presynaptic activation. After the sleep phase, the SNN was remapped back to an ANN. In (Tadros et al. 2022) SRC was applied after each new task training to avoid catastrophic forgetting, here we applied it once after training with limited data.

Results

When the ANN was trained with the full dataset, it achieved an accuracy of over 90%. However, when less than 10% of the data was used during training, accuracy significantly declined (Figure 1, blue line). When 0.5% to 10% of the total data was used for ANN training, the subsequent application of SRC resulted in a substantial (20-30%) increase in accuracy for both MNIST and Fashion MNIST datasets (Figure

*These authors contributed equally.

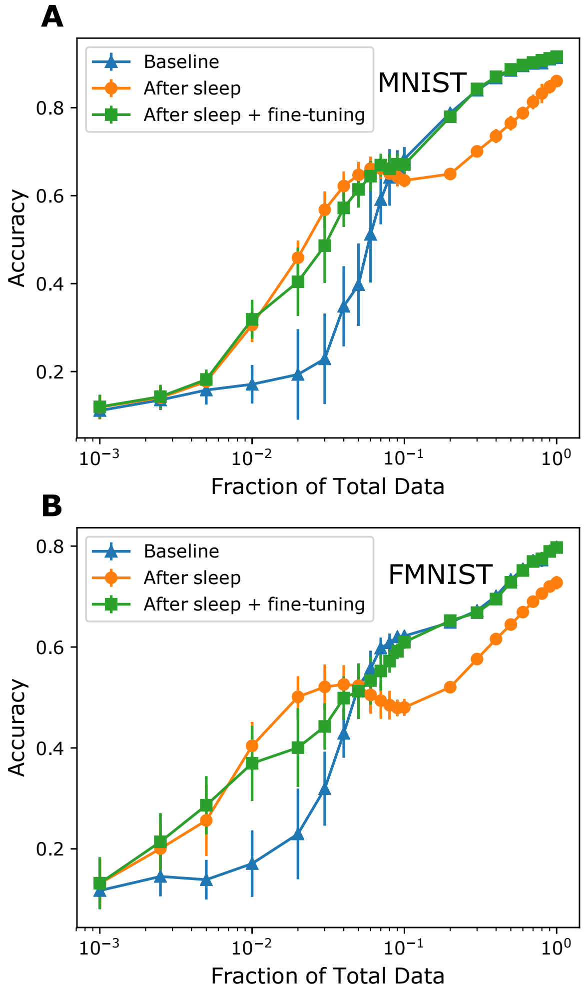

Figure 1: Accuracy on MNIST **(A)** and FMNIST **(B)** with mean (lines) and standard deviation (error bars) across 10 trials. X-axis - log of the relative amount of data used for training (e.g., 0.01=1% of data). Blue - baseline (after ANN training); Orange - baseline + sleep; Green - baseline + sleep + fine-tuning. Note significant gain in accuracy after sleep phase on low data. The sleep phase reduced performance on high data but was largely recovered by fine-tuning.

1, orange line). Increasing the training duration (number of epochs), increased performance before sleep but a significant performance gain after sleep remained.

Analysis of the confusion matrix (see Supplementary Material) revealed that networks trained with limited data can exhibit biases towards a few classes. For example, when 3% of the MNIST data was used in training, classes 0, 2, 5, and 6 were all classified as 0. However, after sleep, classes 0, 2, and 6 were classified correctly. Succinctly, the model exhibited a more balanced response after the application of SRC.

While performance improved when there was limited training data, we also observed a slight (10-15%) decrease in performance when more than approximately 10% of the data was employed for ANN training. We found that this decrease in performance could be mitigated by fine-tuning the ANN after sleep using the original (limited) training data (Figure 1, green line). Thus, by incorporating both sleep and fine-tuning, we were able to maintain performance on models trained with the full dataset while still achieving performance gains on models trained with limited data.

Next, we examined accuracy when a significant class imbalance was introduced to the training set by selectively reducing the number of training examples used for certain classes. We found that class-wise model performance was more robust to data reduction for some classes when compared to others. After SRC, most classes showed a positive improvement in class-wise accuracy (see Supplementary Material). Thus, the sleep phase proved effective in increasing model accuracy on underrepresented classes while preserving accuracy on well-trained classes.

Analysis of synaptic weights revealed that SRC increased strength for a small fraction of critical synapses, while many other synapses were weakened (see Supplementary Material). This suggests that the overall accuracy increase after SRC was a result of increasing the sparsity of responses.

Our study sheds light on a potential synaptic weight dynamics strategy employed by the brain during sleep to enhance memory performance when training data are limited or imbalanced. Applied to ANNs, sleep-like replay improves performance in a completely unsupervised manner, requiring no additional data, and can be applied to already trained models.

Acknowledgments

Supported by NSF (grants 2209874 and 2223839).

References

Arora, S.; Du, S. S.; Li, Z.; Salakhutdinov, R.; Wang, R.; and Yu, D. 2019. Harnessing the power of infinitely wide deep nets on small-data tasks. *arXiv preprint arXiv:1910.01663*.

Golden, R.; Delanois, J. E.; Sanda, P.; and Bazhenov, M. 2022. Sleep prevents catastrophic forgetting in spiking neural networks by forming joint synaptic weight representations. *PLoS Computational Biology*, 18(11): e1010628.

González, O. C.; Sokolov, Y.; Krishnan, G. P.; Delanois, J. E.; and Bazhenov, M. 2020. Can sleep protect memories from catastrophic forgetting? *Elife*, 9: e51005.

Hayes, T. L.; Krishnan, G. P.; Bazhenov, M.; Siegelmann, H. T.; Sejnowski, T. J.; and Kanan, C. 2021. Replay in deep learning: Current approaches and missing biological elements. *Neural Computation*, 33(11): 2908–2950.

Rasch, B.; and Born, J. 2013. About sleep's role in memory. *Physiological reviews*, 93(2): 681–766.

Shorten, C.; and Khoshgoftaar, T. M. 2019. A survey on image data augmentation for deep learning. *Journal of big data*, 6(1): 1–48.

Stickgold, R. 2005. Sleep-dependent memory consolidation. *Nature*, 437(7063): 1272–1278.

Tadros, T.; Krishnan, G.; Ramyaa, R.; and Bazhenov, M. 2022. Sleep-like unsupervised replay reduces catastrophic forgetting in artificial neural networks. *Nature Communications*, 13(1): 7742.

Zhuang, F.; Qi, Z.; Duan, K.; Xi, D.; Zhu, Y.; Zhu, H.; Xiong, H.; and He, Q. 2020. A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109(1): 43–76.