SIESTA: Efficient Online Continual Learning with Sleep

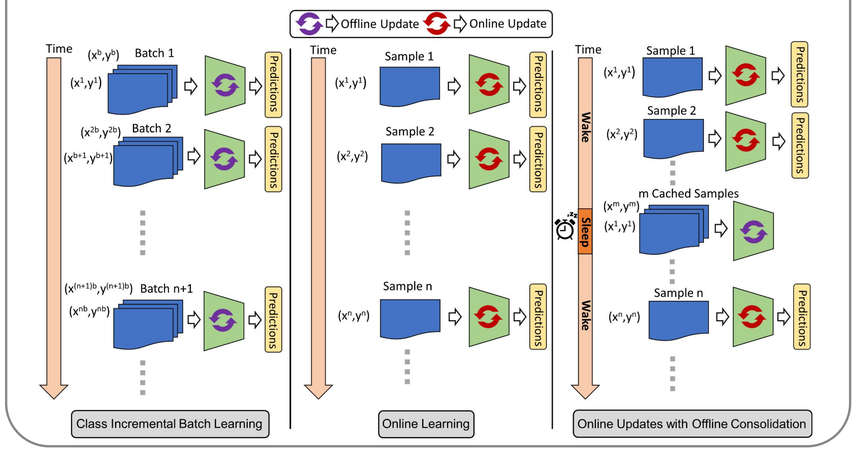
Md Yousuf Harun[1]*, Jhair Gallardo[1]*, Tyler L. Hayes[1], Ronald Kemker[2], Christopher Kanan[3]

[1]Rochester Institute of Technology, [2]United States Space Force, [3]University of Rochester, *Equal contributions

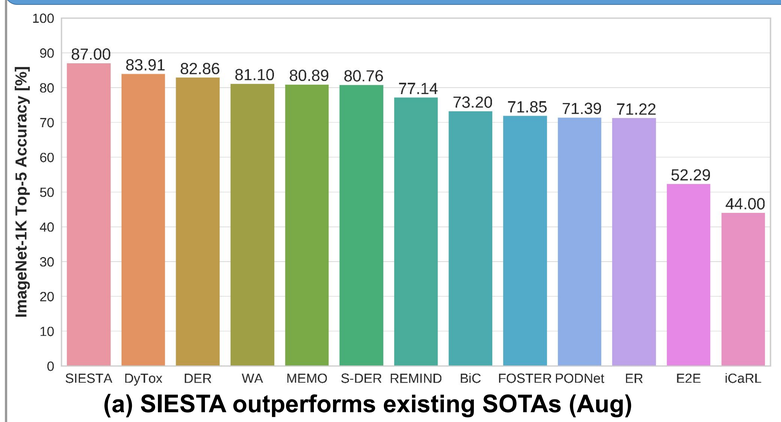
Code: https://yousuf907.github.io/siestasite

Overview

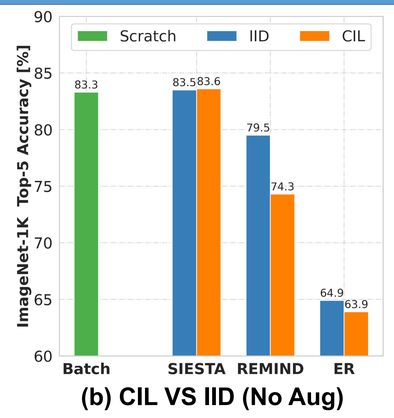
- Continual learning (CL) research solely focused on **catastrophic forgetting** overlooking **compute efficiency**.
- To make a real-world impact, CL must provide compute efficiency and **rival batch learners** trained from scratch.
- We propose a novel method, **SIESTA** based on a **wake/sleep** framework for rapid online CL. **SIESTA performs CL on ImageNet-1K in under 2 hours.**
- We propose a **new paradigm** combining offline and online CL, facilitating many applications e.g., on-device CL.



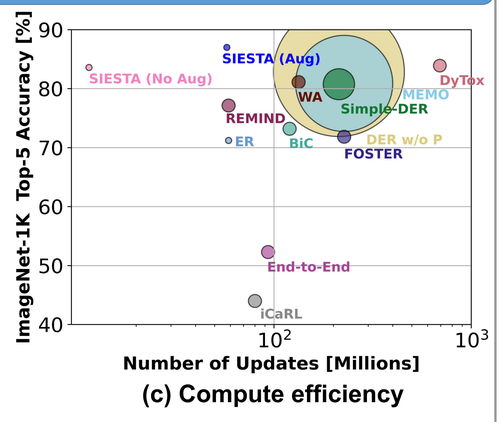
Results



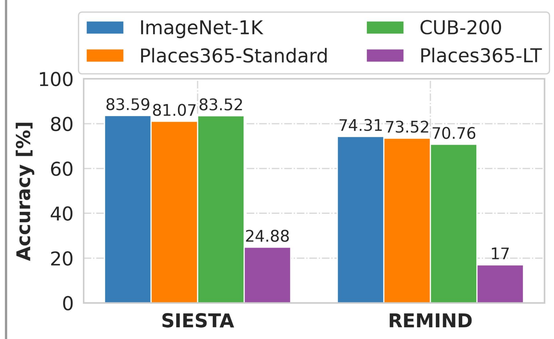
(a) SIESTA outperforms existing SOTAs (Aug)



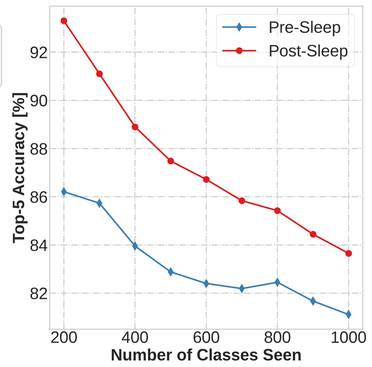
(b) CIL VS IID (No Aug)



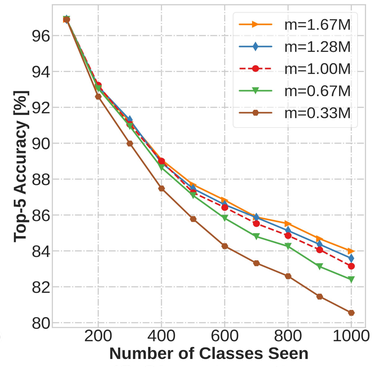
(c) Compute efficiency



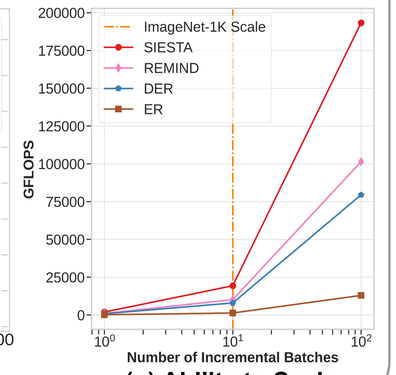
(d) CIL Performance (No Aug)



(e) Impact of Sleep



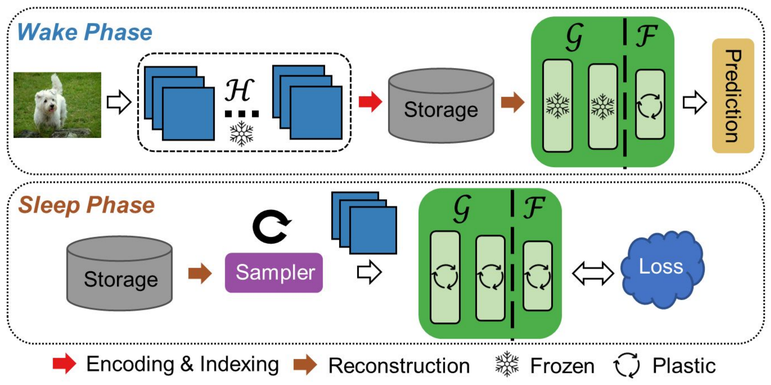
(f) Sleep Length



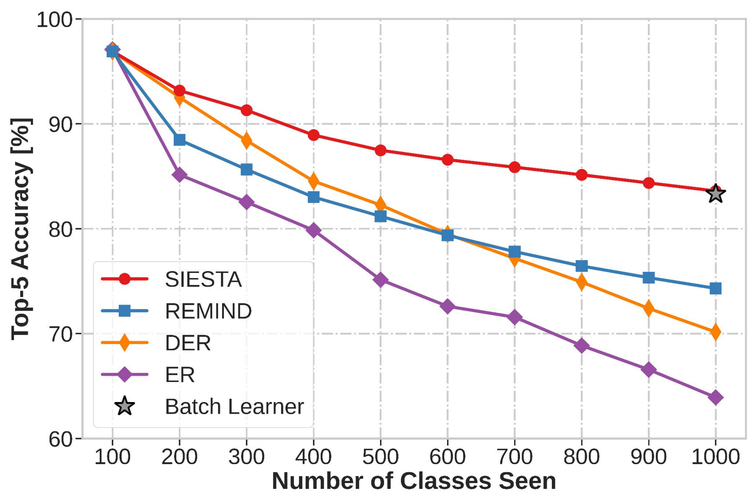
(g) Ability to Scale

How Does SIESTA Work?

During the **wake phase**, SIESTA transforms raw inputs into feature representations using H and stores them in a buffer in compressed form. F is updated in an online manner with running class means. During the **sleep phase**, a sampler retrieves examples from the buffer to reconstruct and update G & F with backpropagation.



SIESTA Achieves *Zero Forgetting*



SIESTA matches offline batch learner and outperforms ER, DER, and REMIND in CIL on ImageNet-1K.

Summary

- SIESTA outperforms SOTA CL methods while requiring **7x-60x fewer updates, 10x less memory, and 2x-20x fewer parameters**.
- Compared to REMIND, SIESTA is **3.4x** (ImageNet-1K) and **4.4x** (Places365-Standard) faster.
- Unlike others, SIESTA maintains similar performance across data distributions e.g., CIL and IID CL.
- SIESTA demonstrates the benefits of sleep for memory consolidation and learning efficiency.
- SIESTA rivals offline batch learners and achieves "**Zero Forgetting**" in large-scale CL.
- SIESTA unlocks opportunities for many real-world applications and has potential to advance Green AI.

Acknowledgements

We thank NSF for financially supporting this research.