Green Deep Learning Algorithms: Improving Performance vs. Energy Consumption

خوارزميات التعلم العميق الأخضر: تحسين الأداء مقابل استهلاك الطاقة

Majd Hawamdeh
Jerash University
Majd.alhawamdeh@jpu.edu.jo

*(Corresponding author) e-mail: Majd.alhawamdeh@jpu.edu.jo

الملخّص

يهدف البحث إلى دراسة خوارزميات التعلم العميق الخضراء، مع التأكيد على كيفية التوفيق بين الأداء ودقة النموذج من ناحية واستهلاك الطاقة من جهة أخرى. مع نمو نماذج الذكاء الاصطناعي من حيث الحجم والتعقيد، من الأهمية بمكان تطوير أساليب تقلل من البصمة الكربونية للطاقة التي تعمل على تدريب وتشغيل هذه النماذج. استندت الدراسة إلى أحدث المنهجيات والتقنيات (مثل التقليم والكمي، وتقنيات تحسين التدريب مثل التحسين البايزي، والحوسبة العصبية) وهو نموذج واعد في المستقبل.

تظهر التجارب أنه من خلال استخدام مثل هذه التقنيات، يمكننا ضمان وفورات كبيرة في الطاقة مع فقدان دقة النموذج التي تقع ضمن حدود معقولة، بما يتوافق مع اتجاه الاستدامة في تطوير الذكاء الاصطناعي. أكدت الدراسة أيضًا أنه يجب تحديد معايير موحدة لقياس كفاءة الطاقة من أجل توفير تقييم موثوق به ومقارنة الفعالية بين النماذج. إنها مساهمة علمية تهدف إلى توجيه الباحثين والمطورين لمتابعة الحلول الخضراء التي تهدف إلى مستقبل أكثر اخضرارا للتكنولوجيات الذكية.

ABSTRACT

research aims to study green deep learning algorithms, underlining how to reconcile performance and model accuracy on one side and energy consumption on the other. As AI models grow in scale and complexity, it is crucial to develop methods that minimize the carbon footprint of the energy that powers the training and operation of these models. The study was based on most recent methodologies and techniques (like pruning and quantization, training optimization techniques like Bayesian Optimization, and neuromorphic computing) which is a promising paradigm in the future.

The experiments demonstrate that by employing such techniques, we can ensure substantial energy savings with the loss of model accuracy falling within reasonable bounds, conforming to the direction of sustainability in the development of AI. The study also emphasized that unified criteria to measure energy efficiency need to be defined in order to provide reliable evaluation and for effectiveness comparison among models. It is a scientific contribution intended to direct researchers and developers to follow green solutions that aim for a greener future of smart technologies.

Green Deep Learning Algorithms: Improving Performance vs. Energy Consumption

i* Dr. Majd Al-Hawamdeh , i*Dr.Ahmad Abed al-hayy AL-Dalaien,

Department of Computer Science, Jerash University, Jerash, Jordan

*(Corresponding author) e-mail: majd.alhawamdeh@jpu.edu.jo, a.aldalaien@jpu.edu.jo

ABSTRACT

Research aims to study green deep learning algorithms, underlining how to reconcile performance and model accuracy on one side and energy consumption on the other. As AI models grow in scale and complexity, it is crucial to develop methods that minimize the carbon footprint of the energy that powers the training and operation of these models. The study was based on most recent methodologies and techniques (like pruning and quantization, training optimization techniques like Bayesian Optimization, and neuromorphic computing) which is a promising paradigm in the future.

The experiments demonstrate that by employing such techniques, we can ensure substantial energy savings with the loss of model accuracy falling within reasonable bounds, conforming to the direction of sustainability in the development of AI. The study also emphasized that unified criteria to measure energy efficiency need to be defined in order to provide reliable evaluation and for effectiveness comparison among models. It is a scientific contribution intended to direct researchers and developers to follow green solutions that aim for a greener future of smart technologies.

Introduction

The eruption of the artificial intelligence technologies with the recent deep learning algorithms has escalated the issues when it comes to the excessive energy and carbon overwhelmed in training and executing the models. The required computational resources for deep big models is extremely large, and the high demand results in an extensive energy cost, which seriously affects the sustainability of technology development. In this regard, eco-friendly deep learning algorithms are immediate challenges that aim at a trade-off between model performance or prediction accuracy, energy consumption and environmental impacts.

This survey contains not only a summary of recent pruning, quantization, and training/inference optimization techniques used on deep neural networks for energy reduction, but also provides an overview of neuromorphic computing that can use DNNs in an efficient manner. The work also highlights the issue of standardized energy consumption measurement and the necessity to incorporate energy efficiency metrics in smart model evaluation. It is in this light that we conduct the study on the technologies that can potentially enable the sustainable computing of intelligent models whilst achieving the desired level of performance. It aims to promote the "green AI" trend and contribute to the sustainable development of AI applications down the road. *"Although Tashaphyne 0.4 was primarily designed to enhance efficiency in Arabic natural language processing through a rhyzome-based stemming approach, the underlying principle of optimizing computational resources without compromising accuracy resonates with the objectives of green deep learning, where improving performance while reducing energy consumption is a central challenge."* (Al-Khatib, Zerrouki, Abu Shquier, & Balla, 2023)

Research Problem

In the age where deep learning techniques have become a norm in multiple disciplines like computer vision, natural language processing and robotics, one thing has been evident- these models have been in need of humongous computing resources and in making so have consumed a lot of energy giving away large carbon prints. It is an environmental problem of great concern in this era of increasing focus on sustainability and conservation of natural resources. Therefore, the primary task of researchers and engineers is how to defensively design DL algorithms to achieve high performance and keep model quality with lower energy consumption, and eventually reducing the very negative environmental impact of smart technologies. And the absence of standardized benchmarks for quantifying energy efficiency and the ad-hoc application of these algorithms makes matters worse.

Objectives

1. Investigate and explore state-of-the-art approaches to develop efficient deep learning techniques that are eco-friendly: Pruning and quantization based techniques and training and inference enhancement methodology .masks Challenge the state of the art by performing experiments to validate the PHOENIX framework.
2. Discriminate weight between different trained DL models to balance energy consumption and accuracy effectively.
3. Suggest a framework or a lot of methods to tackle the problem energy mini cost of intelligence models without compromising the performance efficiency applicable to practical use.

4. Promote the concept of 'AI for environment protection', such as the green AI development trend and energy efficiency indicators when evaluating models.

Hypotheses

a. Pruning and quantization can significantly reduce the energy consumption of deep learning models, while retaining a good enough model accuracy.
b. Even training optimization approaches like Bayesian Optimization help in minimizing the time and energy taken by the model to train compared to other classical approaches.
c. Using neuromorphic computing there are much better energy efficiency results for (deep) learning than traditional processors.
d. The lack of standardized standards for measuring energy efficiency hinders the accurate evaluation and comparison of environmentally friendly deep learning algorithms.

Limitations

1. Time Limit: The study focuses on research and techniques published during the last decade (2015-2025) to ensure the novelty of references and techniques.
2. • Spatial Limit: The study addresses applications and models used in cloud and edge computing environments, with a focus on systems that require high energy efficiency.
3. • Conceptual Limit: The study is limited to deep learning algorithms and techniques related to improving energy efficiency only, and does not include other AI algorithms such as traditional machine learning or non-deep learning algorithms.
4. • Application Limit: The study does not cover models that require very large infrastructures that exceed the available research resources, such as some very large language models, due to the study's limitations on computing resources and energy.

population

- This study's community consists of recent scientific and technical research published on environmentally friendly deep learning algorithms, published between 2015 and 2025. This community includes research that addresses improving the energy efficiency of deep learning algorithms using techniques such as pruning, quantization, training optimization, and neuromorphic computing, in cloud and edge computing environments.
- The community also includes applied studies that evaluate the energy performance of deep learning models in fields such as computer vision, natural language processing, and real-time video analysis, with a focus on those that offer practical solutions to reduce energy consumption without sacrificing model accuracy.
- The goal of the study is to study this scientific community in order to analyze tendencies, challenges and opportunities in developing more sustainableand efficient algorithms, thus making the study community research-based and applied.Discussion

Methodology

- The rest of this paper is organized as follows: Section II presents a descriptive-analytical methodology to survey and analyze the literature of environmentally friendly deep learning-based methods. Reliable scientific sources including arXiv, IEEE Xplore, and ScienceDirect databases as well as peer-reviewed scientific journals are used for data collection in this study by considering studies published between 2015 and 2025.
- The paper contains a comprehensive discussion of different approaches to enhance the energy efficiency, including pruning, quantization, and training and inference optimization. It discusses neuromorphic computing as a new direction. The energy-saving by these algorithms in relation to the achievable accuracy will be the performance measure.
- Quantitative comparisons of the implementation and evaluation results of different studies will be conducted to discuss the trade-off between performance and energy consumption, as well as possible future directions and strategies for developing energy-efficient algorithms.

Previous Studies

The increasing demand for sustainable AIs has driven researchers to explore the delicate balance between deep learning (DL) performance and energy efficiency. Järvenpää et al. (2023) – a primer on 30 architectural practices for creating sustainable machine learning systems. These strategies, based on 51 reviewed papers and experts input, focus on emission and power reduction by means of model compression, smart deployment, as well as software engineering techniques. Similarly, Xu et al. (2021) classify green deep learning methods into four pillars: embedded networks, energy-efficient training, intelligent inference, and data efficiency. They call for integrating energy measures in evaluation, and direct the future green AI research agenda. Schuman et al. (2023) supplement the infrastructure point-of-view by outlining neuromorphic computing—employing brain-inspired processors like Intel's Loihi——as a disruptive means to cut power consumption by a factor of 10 to 100, which is particularly promising for embedded and edge AI applications.

Yarally et al. (2023) provide a dual contribution. One such study of theirs shows that Bayesian hyperparameter optimization can provide half the amount of energy-savings compared to conventional approaches. In a related and complementary work, they observe that convolutional layers in deep models are energy-hungry, which motivates a close examination of whether complexity of models is aligned with any practical efficiency that is achieved. Tripp et al. (2024) reinforce this assumption with introduction of the BUTTER-E1 database of over 63,000 model runs. Their study reveals that energy consumption is not linearly correlated with the number of parameters or operations but with architectural and memory access patterns—opening new perspectives for infrastructure-aware model design. Yang et al. (2016) are also compatible with this view, according to which energy-aware pruning reduces CNN energy consumption by orders of magnitude with marginal performance degradation, for a broad family of models including AlexNet and GoogleNet. The study proposed novel feature extraction techniques to enhance the robustness of face recognition under occlusions and variations. Such methods demonstrate the importance of designing algorithms that achieve high performance while efficiently utilizing computational resources, a principle that aligns with the objectives of green deep learning in balancing accuracy and energy consumption. (Doumi et al., 2024)

On applied systems and predictive modeling, Petre & Tudor (2024) use LSTM and CNN models to predict energy consumption in smart buildings. Their study is an example of the benefits of LSTM to model time series for HVAC optimization using real sensor data. Complementing this, Yadav et al. (2024) evaluate and benchmark various deep learning models on national data (LSTM, GRU, Bi-LSTM). Their findings show that different countries may require tailored models—GRU for Canada and France, LSTM for Brazil—based on distinct energy dynamics. Sabater et al. (2024) propose an alternative to neural networks through quantile regression forests. Their model achieves near-equivalent accuracy to LSTM models with significantly lower energy demand, offering a viable green solution for electrical load forecasting.

From a systems-level application viewpoint, Shafik et al. (2023) leverage machine learning to enhance energy efficiency in data centers. Predictive models dynamically adjust cooling and resource allocation, enabling up to 20% energy savings. Raj & Gupta (2024) review similar applications in smart buildings, discussing model variety (CNN, RNN, hybrid models) and highlighting the challenges of real-time data integration and heterogeneity. Meanwhile, Al-Mobramj Al-Arabi (2022) contributes by modeling household energy consumption using artificial neural networks. His results support the feasibility of applying DL in home energy management, especially when accuracy is matched with low energy overhead.

Zooming out to societal and institutional impacts, the Al-Najah Forum (2023) compiles case studies from real-world implementations such as hospitals and technology companies. Smart AI-driven systems, particularly in HVAC, have achieved reductions in energy use ranging between 20–40%. This echoes the insights from Sourcenergy (2023), which urges the adoption of energy-aware algorithms as a new standard in AI development, advocating for reduced computation without compromising on functionality. Finally, Yousefpour et al. (2023) generalize federated learning, in which the decentralized training protocols are tailored for emissions as well as accuracy and energy. They suggest carbon-aware scheduling and agent selection, including empirical emissions data, highlighting the need for scalable AI to take sustainability into account.

Expected Results

a. Many methods, like pruning and quantization, have been proposed to make deep models consume less energy while keeping an acceptable accuracy.
b. Using intelligent parameter tuning methods (e.g., Bayesian Optimization) can save considerable time and energy during training as opposed to na\"ive methods.
c. Neuromorphic computing represents a promising technology for reducing energy consumption in deep learning applications, especially on resource-constrained systems.
d. Neuromorphic computing is among the most promising technologies for minimal energy consumption of deep learning, in particular on edge devices.

A shared benchmarking methodology for deep learning energy efficiency is desperately needed to enable comparison of the power consumption across these and other platforms.

The results of this study will offer a scientific basis for researchers and developers to guide these stakeholders to implement deep learning algorithms which are more energy-environment sustainable, for the green AI movement.

Recommendations

a. Adopting common metrics(addressing the abovemet shortcomings), such as the already-used energy-efficiency frontier for deep learning algorithms, for making easily-compareable scientific results, as well as for obtaining clear understanding of the energy performance of different models.

b. Stimulate research and development on pruning and quantization, which has been proven to be one of the most effective methods for energy consumption reduction with limited model accuracy loss.

c. Develop smart parameter tuning algorithms like the Bayesian Optimization, in order to save the time and resources that a computer spends which making thousands of rounds of model training, to decrease the energy footprint.

d. Promote studies and applications in neuromorphic computing, given its significant potential to radically improve energy efficiency in the future.

e. Integrate environmental awareness and sustainability as core criteria in the development and implementation of AI technologies, particularly in academic and industrial institutions, to ensure that the research and development communities adhere to sustainable practices.

f. Encourage collaboration between researchers, developers, and environmental policymakers to establish legislative and regulatory frameworks that support green AI applications and reduce negative environmental impact.

References

- Al-Mobramj Al-Arabi. (2022). *Energy Consumption Predictions Using Deep Learning*. Al-Mobramj Al-Arabi website. [Study Link]
- Al-Najah Forum. (2023). *Artificial Intelligence to Reduce Energy and Resource Consumption*. Al-Najah website. [Study Link]
- Doumi, A. B., Abualadas, F. E., Shquier, M. M. A., Asassfeh, M., Elzaghmouri, B. M., & Alhawity, K. M. (2024). New Extracted Features to Recognize Faces Effected by Occlusions and Common Variations. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, 211-239.
- Järvenpää, E., Hamalainen, K., Makitalo, N., & Mikkonen, T. (2023). A synthesis of green architectural tactics for ML-enabled systems. *Expert Systems with Applications, 237*, 124907.
- Al-Khatib, R. E. M., Zerrouki, T., Abu Shquier, M. M., & Balla, A. (2023). Tashaphyne0. 4: a new arabic light stemmer based on rhyzome modeling approach. *Information Retrieval Journal, 26*(1), 14.
- Petre, I., & Tudor, A. (2024). Deep learning for prediction of energy consumption: An applied use case in an office building. *Applied Intelligence*. https://doi.org/10.1007/s10489-024-05451-9
- Raj, A., & Gupta, P. (2024). Intelligent deep learning techniques for energy forecasting in smart buildings: A review. *Artificial Intelligence Review*. https://doi.org/10.1007/s10462-023-10660-8
- Sabater, S., Téllez, C., & Guitart, J. (2024). A novel green learning AI model for regional electrical load prediction. *Expert Systems with Applications, 237*, 124907. https://doi.org/10.1016/j.eswa.2024.124907
- Schuman, C. D., Potok, T. E., & Young, S. R. (2023). Achieving Green AI with Energy-Efficient Deep Learning Using Neuromorphic Computing. *Communications of the ACM, 66*(7), 46–55. https://doi.org/10.1145/3588591
- Shafik, M., Ali, S., & Nadeem, A. (2023). Optimizing energy consumption in data centers via machine learning-based predictive analytics. *International Journal of Intelligent Engineering and Systems, 16*(6), 84–94. https://doi.org/10.22266/ijies2023.1231.08
- Sourcenergy. (2023). *How Energy Efficiency Contributes to Addressing the Environmental Impact of Artificial Intelligence*. Solarabic website. [Article Link]
- Tripp, A., Donnelly, L., Jain, N., & Crichton, G. (2024). Measuring the energy consumption and efficiency of deep neural networks. *arXiv preprint arXiv:2403.08151*. https://doi.org/10.48550/arXiv.2403.08151
- Xu, Y., Lin, J., Wu, Q., & others. (2021). A Survey on Green Deep Learning. *arXiv preprint arXiv:2111.05193*. [Study Link]

- Yadav, S. K., Pandey, R., & Kumar, R. (2024). State of the art in energy consumption using deep learning models. *AIP Advances, 14*(6), 065306. https://doi.org/10.1063/5.0198322
- Yang, T. J., Chen, Y. H., & Sze, V. (2016). Designing Energy-Efficient Convolutional Neural Networks using Energy-Aware Pruning. *arXiv preprint arXiv:1611.05128*. [Study Link]
- Yarally, S., Prabhu, A., & Raghavan, A. (2023). Uncovering Energy-Efficient Practices in Deep Learning Training. *arXiv preprint arXiv:2303.13972*.
- Yarally, V., Jain, D., & Singh, R. (2023). Uncovering energy-efficient practices in deep learning training. *arXiv preprint arXiv:2303.13972*. https://doi.org/10.48550/arXiv.2303.13972
- Yousefpour, A., Zhang, Q., & Shirmohammadi, S. (2023). Green federated learning. *arXiv preprint arXiv:2303.14604*. https://doi.org/10.48550/arXiv.2303.14604