Journal of Systems Scalability
http://dx.doi.org/10.62762/JSS.2025.960646

PERSPECTIVE

Energy Scalability in the Training of AI Models for Image Processing: The Role of Hyperparameters

Blanca Atiénzar[1], **David Cortes**[1,*], **Belen Bermejo**[1] **and Carlos Juiz**[1]

[1] Department of Mathematics and Computer Science, University of the Balearic Islands, 07122 Palma, Spain

Abstract

This perspective article argues that hyperparameters such as learning rate, batch size, numerical precision, and training workers are key determinants of energy scalability in CNN training. These parameters directly influence convergence dynamics, hardware utilization, and training duration, leading to substantially different energy profiles even when comparable accuracy is achieved. Moreover, hyperparameter search itself introduces a significant cumulative energy cost, often exceeding that of the final selected model. By analyzing the interaction between convergence behavior and energy consumption, this work highlights the need to treat energy as an explicit scalability metric and to integrate energy-aware considerations into hyperparameter optimization. Adopting this perspective enables more efficient, sustainable, and reproducible training practices for large-scale image processing models.

Keywords: hyperparameters, energy, scalability, training of AI.

Submitted: 22 December 2025
Accepted: 11 February 2026
Published: 10 March 2026

Vol. 1, **No.** 1, 2026.
10.62762/JSS.2025.960646

***Corresponding author:**
David Cortes
d.cortes@uib.es

1 Introduction

The rapid growth of artificial intelligence has made large-scale model training a central focus of both academic research and industrial applications [1, 2]. In image processing, this trend is reflected in the extensive use of architectures based on tensor operations, particularly convolutional neural networks (CNNs) [3], which belong to machine learning (ML) and are part of deep learning (DL) models [4, 5]. CNNs constitute the core of most modern computer vision systems [4]. Traditionally, the scalability of these systems' performance has been evaluated using computational metrics such as the number of floating-point operations, total training time [6], and the number of processed images [7]. However, the sustained growth in model size, data resolution, and training duration has positioned energy consumption as a critical constraint on their viability and scalability [8].

In this context, energy ceases to be an indirect consequence of performance and becomes a first-order scalability metric [9], with a direct impact on operational costs [10], environmental sustainability, and the accessibility of large-scale research [11]. However, one of the most influential elements in training behavior-the selection of hyperparameters, understood as external configuration variables defined prior to training a CNN model-remains largely

Citation
Atiénzar, B., Cortes, D., Bermejo, B., & Juiz, C. (2026). Energy Scalability in the Training of AI Models for Image Processing: The Role of Hyperparameters. *Journal of Systems Scalability*, 1(1), 1–5.

decoupled from energy considerations [12, 13]. Common tuning practices prioritize final accuracy, convergence stability, or execution time, implicitly assuming that energy efficiency is automatically derived from these decisions [14].

In image processing models, this assumption is particularly problematic. Hyperparameters such as the learning rate determine not only the convergence dynamics of deep convolutional networks but also how hardware resources are utilized throughout training [15, 16]. These choices influence computational intensity, memory access patterns, and the overall duration of the process, leading to significantly different energy profiles even when similar accuracy is achieved. From a scalability perspective, hyperparameter search additionally introduces a combinatorial growth in energy cost, as each evaluated configuration entails a partial or complete training run. Since tuning procedures typically require the evaluation of multiple combinations of learning rates, batch sizes, or numerical precisions, the cumulative energy cost of the search phase can exceed that of the final training of the selected model [17].

This perspective article argues that, in the training of AI models for image processing, energy should be treated as a central scalability metric, and that hyperparameters constitute one of the primary mechanisms through which energy costs are amplified or mitigated as scaling increases.

2 Convolutional Neural Networks, Convergence, and Energy Consumption

CNNs excel at recognizing features and local patterns across an image using filters. Consequently, they are widely used in image recognition, analysis, and processing tasks [5]. They are composed of three main types of layers: convolutional layers, pooling layers, and fully connected layers [5]. As such, they constitute a paradigmatic case for studying the relationship between hyperparameters, convergence, and energy consumption. Their structure, based on intensive tensor operations such as convolutions and matrix multiplications, makes both the computational and energy costs of training highly sensitive to the chosen configuration.

A high learning rate can accelerate initial convergence but may also cause oscillations and instability, increasing the total number of iterations [18, 19]. Similarly, a large batch size can improve the per-step accuracy of a CNN model but may require additional adjustments to the learning rate or affect the model's generalization capability [16]. Numerical precision, in turn, reduces the energy cost per operation, although in some cases it may prolong the convergence process if training stability is compromised [20, 21].

Thus, it is not sufficient to analyze whether a convolutional network converges correctly; it is crucial to understand the energy cost at which it does so, particularly as model depth, image resolution, and data volume increase. Even small decisions in hyperparameter configuration can lead to substantial differences in energy consumption, demonstrating that energy scalability depends not only on the size of the model or the dataset, but also on how training is planned and computational workloads are managed [13, 22].

3 The Role of Hyperparameters in The Energy Scalability of AI Model Training

In large-scale training of image processing models, hyperparameters act as control variables that simultaneously determine algorithmic efficiency and energy efficiency. The learning rate defines the speed at which the model converges and the number of epochs required to reach acceptable accuracy [18], which directly translates into cumulative energy consumption [2]. Batch size and the number of workers influence the degree of effective parallelism, memory utilization, and the balance between computation and communication, thereby affecting total energy consumption in a non-trivial manner [23]. Experimental evidence shows that increasing the number of epochs leads to higher total energy consumption, although the growth is not strictly linear. In particular, intermediate configurations such as 80 epochs exhibit lower energy consumption than would be expected under a linear scaling assumption, while at 160 epochs the behavior approaches the linear trend again [25]. Similarly, numerical precision reduces the energy load per operation but may require more training steps if learning stability is reduced [24]. Finally, epoch scheduling strategies and early stopping determine whether training halts at an energetically reasonable point or continues to consume resources for marginal performance improvements.

These variables do not operate in isolation; their effects are amplified in long-running trainings and large-scale datasets; such that small configuration changes can produce significant differences in total energy consumption. This implies that ignoring this dimension can lead to systems that, while scalable in terms of accuracy, may be inefficient or even unviable from an energy perspective. By contrasting computational load with actual energy consumption, it becomes evident that hyperparameter optimization not only improves convergence but can also reduce total energy usage [1].

3.1 Implications and Future Research Directions

Methodologies are required that jointly characterize convergence, accuracy, and energy consumption, as well as hyperparameter search strategies that account for the cumulative energy cost of the training process. Likewise, it is relevant to investigate adaptive mechanisms that dynamically adjust hyperparameters such as learning rate, batch size, or numerical precision based on the energy behavior observed during training. Addressing these challenges would not only improve the efficiency of existing models but also advance toward more sustainable and reproducible training practices in large-scale scenarios.

4 Conclusion

Energy scalability has become a determining factor in the training of AI models for image processing. In this perspective article, we have argued that hyperparameters play a central role in how energy consumption scales with model size, data volume, and training duration, particularly in architectures based on tensor operations such as convolutional neural networks.

Adopting a perspective in which energy is treated as an explicit scalability metric has profound implications for the design and evaluation of AI systems for image processing. Hyperparameter selection ceases to be a purely algorithmic problem and becomes part of the overall system design, linking optimization decisions to cumulative energy costs.

This approach also enables a more transparent and reproducible evaluation of large-scale systems by revealing trade-offs that remain hidden when accuracy is the sole criterion for success. From an industrial standpoint, hyperparameter selection that is conscious of energy scalability can translate into significant cost reductions and more efficient use of infrastructure.

Data Availability Statement

Not applicable.

Funding

This work was supported without any funding.

Conflicts of Interest

The authors declare no conflicts of interest.

AI Use Statement

The authors declare that no generative AI was used in the preparation of this manuscript.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 3645-3650). [CrossRef]

[2] Patterson, D., Gonzalez, J., Le, Q., Liang, C., Munguia, L. M., Rothchild, D., ... & Dean, J. (2021). Carbon emissions and large neural network training. *arXiv preprint arXiv:2104.10350.*

[3] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444. [CrossRef]

[4] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012, December). ImageNet classification with deep convolutional neural networks. In *Proceedings of the 26th International Conference on Neural Information Processing Systems-Volume 1* (pp. 1097-1105).

[5] Cortes, D., Bermejo, B., & Juiz, C. (2024). The use of CNNs in VR/AR/MR/XR: a systematic literature review. *Virtual Reality, 28*(3), 154. [CrossRef]

[6] Tan, M., & Le, Q. (2021, July). Efficientnetv2: Smaller models and faster training. In *International conference on machine learning* (pp. 10096-10106). PMLR.

[7] Cortes, D., Juiz, C., & Bermejo, B. (2025). Estudio de la eficiencia en la escalabilidad de GPUs para el entrenamiento de Inteligencia Artificial. *arXiv preprint arXiv:2509.03263.*

[8] García-Martín, E., Rodrigues, C. F., Riley, G., & Grahn, H. (2019). Estimation of energy consumption in machine learning. *Journal of Parallel and Distributed Computing, 134*, 75-88. [CrossRef]

[9] Aquino-Brítez, S., García-Sánchez, P., Ortiz, A., & Aquino-Brítez, D. (2025). Towards an energy consumption index for deep learning models: A comparative analysis of architectures, GPUs, and measurement tools. *Sensors, 25*(3), 846. [CrossRef]

[10] Winsta, J. (2025). The hidden costs of ai: A review of

energy, e-waste, and inequality in model development. *arXiv preprint arXiv:2507.09611*.

[11] Tripp, C. E., Perr-Sauer, J., Gafur, J., Nag, A., Purkayastha, A., Zisman, S., & Bensen, E. A. (2024). Measuring the energy consumption and efficiency of deep neural networks: An empirical analysis and design recommendations. arXiv preprint *arXiv:2403.08151*.

[12] Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green ai. *Communications of the ACM, 63*(12), 54-63. [CrossRef]

[13] Henderson, P., Hu, J., Romoff, J., Brunskill, E., Jurafsky, D., & Pineau, J. (2020). Towards the systematic reporting of the energy and carbon footprints of machine learning. *Journal of machine learning research, 21*(248), 1-43.

[14] Geissler, D., Zhou, B., Suh, S., & Lukowicz, P. (2024). Spend more to save more (sm2): An energy-aware implementation of successive halving for sustainable hyperparameter optimization. *arXiv preprint arXiv:2412.08526*.

[15] Geißler, D., Zhou, B., Liu, M., Suh, S., & Lukowicz, P. (2024, May). The power of training: How different neural network setups influence the energy demand. In *International Conference on Architecture of Computing Systems* (pp. 33-47). Cham: Springer Nature Switzerland. [CrossRef]

[16] Smith, S. L., Kindermans, P. J., Ying, C., & Le, Q. V. (2017). Don't decay the learning rate, increase the batch size. *arXiv preprint arXiv:1711.00489*.

[17] Li, L., & Talwalkar, A. (2020, August). Random search and reproducibility for neural architecture search. In *Uncertainty in artificial intelligence* (pp. 367-377). PMLR.

[18] Smith, L. N. (2017, March). Cyclical learning rates for training neural networks. In *2017 IEEE winter conference on applications of computer vision (WACV)* (pp. 464-472). IEEE. [CrossRef]

[19] Bottou, L. (2010, September). Large-scale machine learning with stochastic gradient descent. In *Proceedings of COMPSTAT'2010: 19th International Conference on Computational StatisticsParis France, August 22-27, 2010 Keynote, Invited and Contributed Papers* (pp. 177-186). Heidelberg: Physica-Verlag HD. [CrossRef]

[20] Ortiz, M., Cristal, A., Ayguadé, E., & Casas, M. (2018). Low-precision floating-point schemes for neural network training. *arXiv preprint arXiv:1804.05267*.

[21] Narang, S., Diamos, G., Elsen, E., Micikevicius, P., Alben, J., Garcia, D., ... & Wu, H. (2017, October). Mixed precision training. In *Int. Conf. on Learning Representation*.

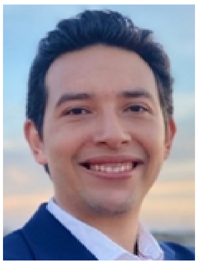
[22] You, J., Chung, J. W., & Chowdhury, M. (2023). Zeus: Understanding and optimizing {GPU} energy consumption of {DNN} training. In *20th USENIX Symposium on Networked Systems Design and Implementation (NSDI 23)* (pp. 119-139).

[23] Keskar, N. S., Mudigere, D., Nocedal, J., Smelyanskiy, M., & Tang, P. T. P. (2016). On large-batch training for deep learning: Generalization gap and sharp minima. *arXiv preprint arXiv:1609.04836*.

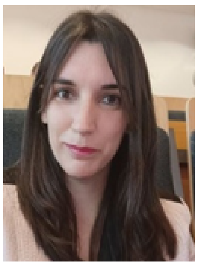
[24] Frey, N. C., Zhao, D., Axelrod, S., Jones, M., Bestor, D., Gadepally, V., ... & Samsi, S. (2022, May). Energy-aware neural architecture selection and hyperparameter optimization. In *2022 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)* (pp. 732-741). IEEE. [CrossRef]

[25] dos Santos, G. C., Araújo, A. L. D., Nakamura, T. C. R., Sousa-Neto, S. S., Giraldo-Roldán, D., Lopes, M. A., ... & Moraes, M. C. (2024, September). Assessment of the Influence of Batch Size on ResNet-50 Processing Applied to Histopathological Datasets. In *Brazilian Congress on Biomedical Engineering* (pp. 548-556). Cham: Springer Nature Switzerland. [CrossRef]

Blanca Atienzar is a student in Computer Engineering, with research interests related to database monitoring and management. She is beginning to study deep learning techniques with the aim of expanding her background in Artificial Intelligence. Her academic interests include exploring different technological areas for the analysis and understanding of complex systems. She aims to continue developing research-oriented skills in digital and technological environments. (Email: blanca.atienzar1@estudiant.uib.cat)



David Cortes received his undergraduate degree from the National University of Colombia and an M.Sc. in Intelligent Systems from the University of the Balearic Islands. He is a Ph.D. candidate in Information Technologies at the University of the Balearic Islands. His research focuses on Artificial Intelligence, with emphasis on convolutional neural networks for image and medical image processing. His interests include AR/VR, efficient training, scalability, and energy-efficient computing. (Email: d.cortes@uib.es)



Belen Bermejo received her Ph.D. in Computer Science in 2020 and is a member of the ACSIC research group. Her research focuses on improving the trade-off between energy consumption and performance in virtualized servers and cloud computing systems, as well as on the energy impact of surveillance capitalism. She has carried out pre-doctoral and post-doctoral research stays at Newcastle University and the University of Seville. She has authored several publications in JCR-indexed journals, international conferences, and book chapters, and currently serves as Chair of IEEE Women in Engineering Spain. (Email: belen.bermejo@uib.es)

Carlos Juiz is a Full Professor of Computer Science and holds a Ph.D. in Computer Science from the University of the Balearic Islands, Spain. His research focuses on performance engineering, Green IT, and IT governance, leading the ACSIC research group. He has served as a visiting researcher at the University of Vienna (2003) and a visiting associate professor at Stanford University (2011). He has supervised nine Ph.D. dissertations and authored over 250 international scientific publications and books, and is a Senior Member of IEEE and ACM. Currently, is a visiting professor at Keio University. (Email: cjuiz@uib.es)