International Journal of Scientific Research and Modern Technology (IJSRMT) ijsrmt.com
Volume 5, Issue 4, 2026
DOI: https://doi.org/10.38124/ijsrmt.v5i4.1399

Green AI: A Study of Optimization Approaches, Sustainability Metrics, Applications and Emerging Challenges

Dr. Hussein K. A. Kadhim[1]

[1]Ministry of Education Computer Technology Department

Publication Date: 2026/05/08

Abstract

The recent development of artificial intelligence (AI) has brought spectacular improvements in a great variety of spheres but has also caused significant issues related to the high energy usage, the cost of the calculations and the carbon emissions. In this respect, Green AI has become a significant research focus whereby efforts are undertaken to come up with AI systems that are not only accurate and powerful, but also energy-efficient, environmentally sustainable and economically viable. The study begins by defining what Green AI entails and why it is relevant in mitigating the environmental footprint of the current AI systems. It then introduces a taxonomy of Green AI techniques, which are model-level optimization, training-time optimization, system and infrastructure optimization and deployment/runtime optimization. Moreover, the survey examines the feasibility of deploying Green AI in various significant fields such as buildings, agriculture, manufacturing, data centers, and supply chains and how sustainability-oriented AI can be used to enhance the operational efficiency of facilities and minimize environmental impact. The paper also talks about the primary evaluation measures and instruments applied in Green AI such as energy usage, carbon footprint, training time and inference latency. Lastly, the survey also lists the main open challenges, including the compromise between accuracy and efficiency, the unstandardized benchmarks and the growing sustainability cost of large-scale AI models, as well as offers promising future directions, including sustainable scheduling, efficient foundation models and adaptive real-time Green AI systems.

Keywords: *Green AI, Sustainable Artificial Intelligence, Carbon-Aware Computing, Efficient Foundation Models.*

I. INTRODUCTION

The field AI has been transformed into a cornerstone of innovation in a wide range of industries, such as healthcare, transport, industry and digital services. Simultaneously, as the topic of deep learning and large-scale foundation models has grown, concerns about the computational cost, electricity usage and carbon footprint have increased. Green AI is a concept that has developed as a research paradigm in reaction to these fears and advocates not only high predictive performance, but also energy efficiency, environmental friendliness and increased accessibility. In contrast to the traditional AI methods, which focus on the accuracy factor as their key target, Green AI focuses on the significance of accounting computational cost and environmental cost in the process of evaluating and developing intelligent systems [1] [2].

Recent reports reveal that Green AI has become a wide and vibrant field of study that addresses several levels of technicality. Recent sources demonstrate that trying to be sustainable in AI can be achieved by using model optimization strategies via pruning, quantization and distillation, by using effective training strategies, hardware-efficient platforms and carbon-conscious deployment systems. Additionally, recent survey and review publications have reported the increasing significance of energy measurement, carbon reporting and benchmarking practices as conditions required to develop responsible AI. This is one of the key changes of the AI focused on accuracy to a more balanced approach, which takes into account performance, resource efficiency and environmental impact together [3] [4].

Thus, the purpose of this survey is to give a systematic picture of Green AI in the context of the investigation of its key methods, areas of its application, existing issues and prospects of research in this field. The article dwells on sustainability in the entire AI lifecycle such as in model design, training, infrastructure and deployment. It also talks about the necessity of standard metrics, transparent reporting systems and realistic frameworks that can help strike the balance between precision and energy consumption and carbon efficiency. The synthesis of recent developments in this field is aimed at clarifying the state of the art and finding some prospects in creating more sustainable and efficient AI systems through the survey [5][6][7].

II. BACKGROUND ON GREEN AI

Green AI has become a reaction to the rising computational and environmental cost of current artificial intelligence systems. The increasing size and complexity of AI models drive both their development and deployment to demand significant resources in terms of computing and energy usage and hardware support, which can pose a significant carbon footprint. That is why, Green AI is becoming more viewed as paradigm that is not limited to predictive accuracy and covers efficiency, sustainability and responsible use of resources as the basic design aims [1] [3] [4] [8] [9]. This change is indicative of a significant change in the research in AI towards less performance-driven development and more balanced approaches that take into account the technical ability and environmental consciousness [10] [11].

In a broader sense, Green AI spans several AI life cycle phases, such as model design, training processes, inference processes, software implementation and infrastructure operation. It is also not restricted to model size reduction, but also efficient training, energy-conscious inference, sustainable computing and improved evaluation of the environmental impact. Such considerations have gained particular significance as the proliferation of large scale AI services where long-term energy demand may be generated by combining continuous inference, cloud computing and edge deployment [7] [9] [10]. Most current researches, also reveal that Green AI must be considered as a multi-disciplinary area of research where machine learning, systems optimization, sustainable software creation and environmental reviewing are integrated so that efficiency is increased without lowering practical utility [11] [12]. As illustrated in Fig. 1, Green AI is also directly related to real-world adoption in various sectors, such as manufacturing, agriculture, energy and utility, data center and IT sector, building and construction, retail and supply chain. The areas show that Green AI does not simply focus on a limited technical challenge, but it is more of an inclusiveness sustainability based model of intelligent systems in a broad range of industrial and service settings. In this respect, the character assists to explain the fact that Green AI is able to assist in providing the efficient approach to decision-making, decreasing the waste of resources and making operations more sustainable within the wide scope of applications [5] [6] [7] [12].

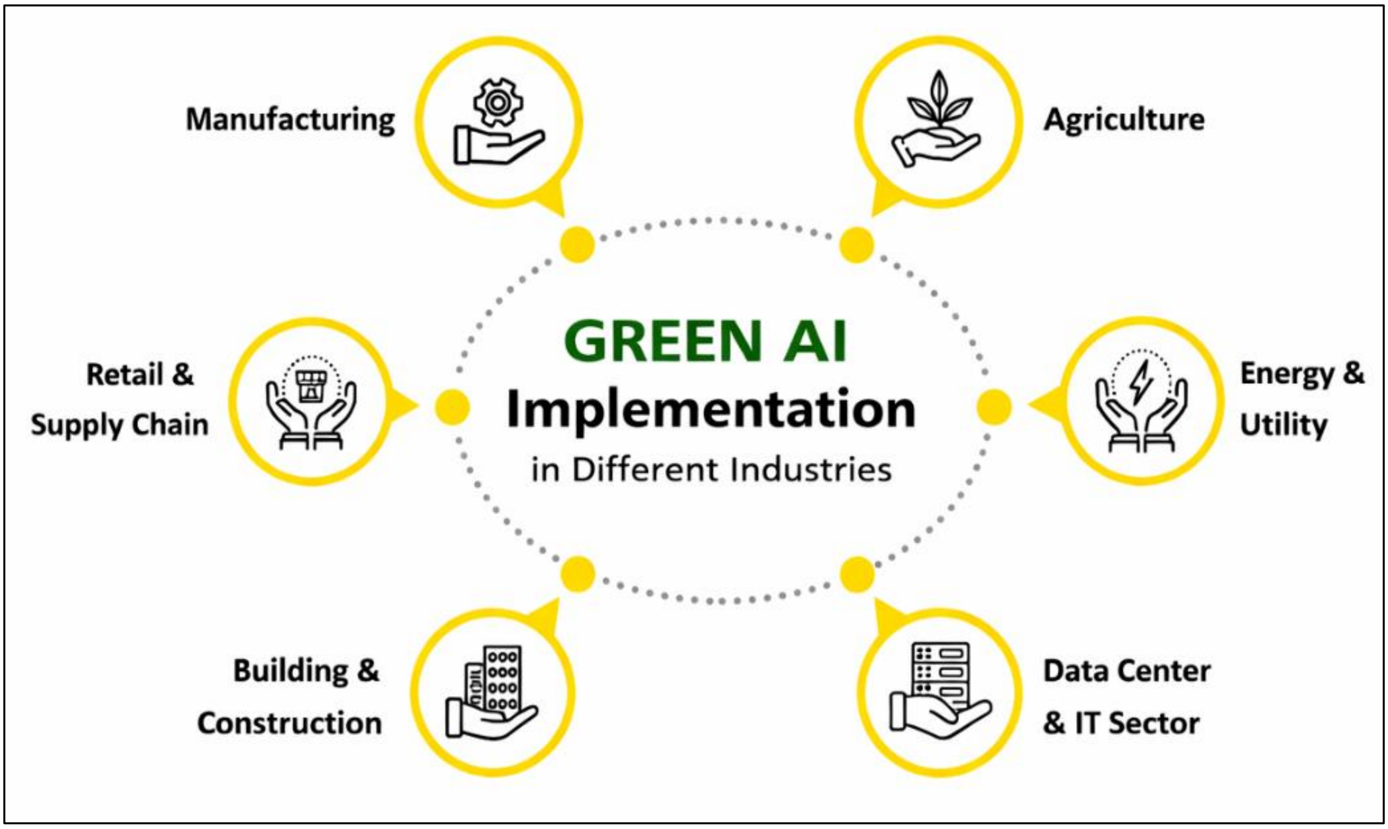

Fig 1 Overview of Green AI.

III. SURVEY METHODOLOGY

The type of chosen methodology is a systematic process of literature research, screening, selection based on the eligibility criteria of the participants, final selection and thematic analysis as shown in Fig. 2.

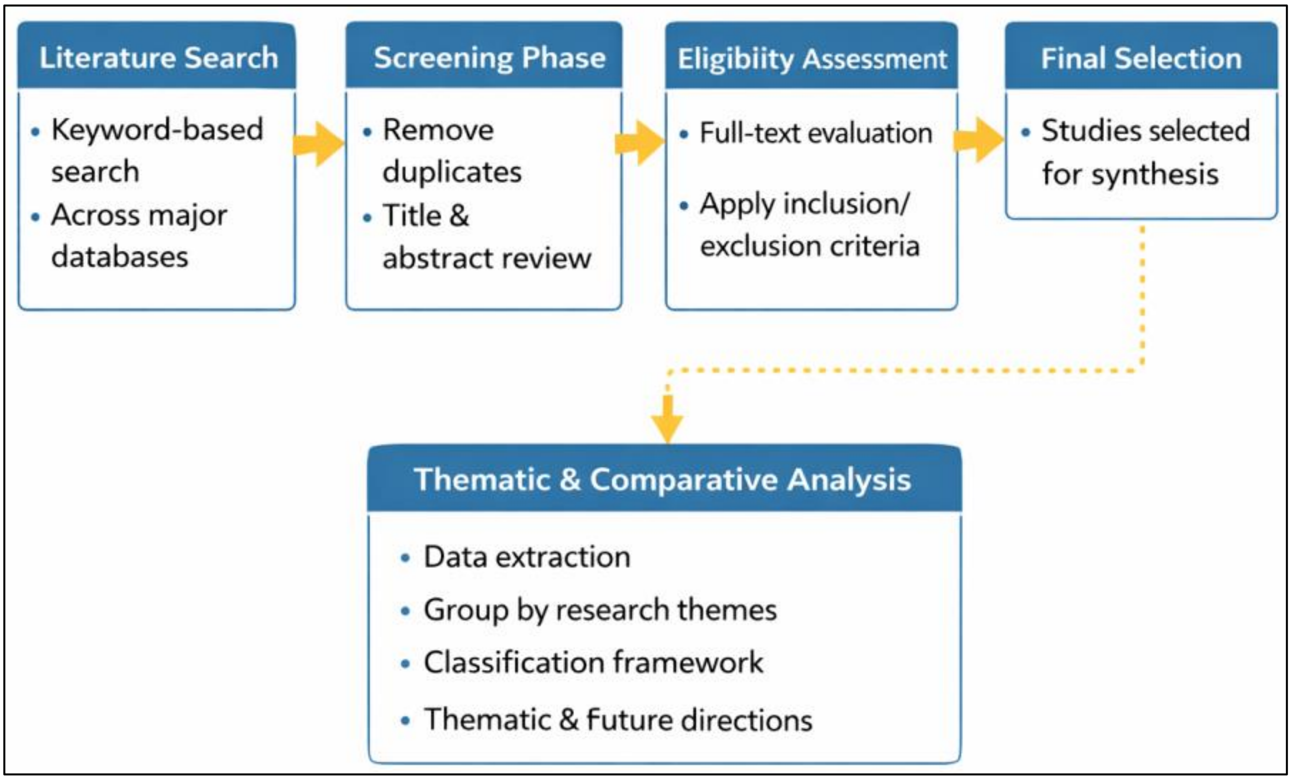

Fig 2 The Survey Methodology Workflow.

This survey aims at offering a systematic and comprehensive overview of Green AI studies, including sustainable methods, applications, significant issues or implications and future research opportunities. The literature review targets research addressing the minimization of computational cost, energy and environmental impact during the AI lifecycle, such as developing a model, training a model, deploying a model and operational use of models.

➢ *Literature Search*

The methodology was initially characterized by the first stage that was based on the systematic search of the scientific literature by using key words in the leading databases. The search was made comprehensive in order to capture the best academic sources that exist in the fields of artificial intelligence, sustainable computing and interdisciplinary research in engineering. Keywords used were terms like Green AI, Sustainable AI, Energy-Efficient AI, Low-Carbon AI, and Energy-Aware Machine Learning.

➢ *Screening Phase*

Having retrieved the first set of studies, this was followed by a screening phase to enhance relevancy and eliminate inappropriate records.

➢ *Eligibility Assessment*

The rest of the studies were then put through the full-text evaluation to help ascertain their appropriateness to be incorporated towards the survey. At this point, specific inclusion and exclusion criteria were used so that only the relevant, detailed enough, and scholarly reliable studies could be included in the further analysis.

➢ *Final Selection and Analysis*

Qualitative synthesis and comparative analysis were done on the last group of chosen studies. Each paper was searched and the relevant information was obtained: the year of publication, primary purpose, dimension of Green AI that the study addressed, methods proposed, domain of application, and major conclusions. These studies were then categorized into broad topics which included model level optimization, training efficiency, system and infrastructure optimization, deployment strategies, applications, challenge and future directions.

IV. TAXONOMY OF GREEN AI TECHNIQUES

Green AI methods are the collection of optimization methods aimed at decreasing the amount of energy, computational cost and environmental footprint of artificial intelligence systems without significantly affecting its performance. There are several techniques, namely, model-level optimization, training-time optimization, system and infrastructure optimization and deployment and runtime optimization, so as shown in Fig. 3.

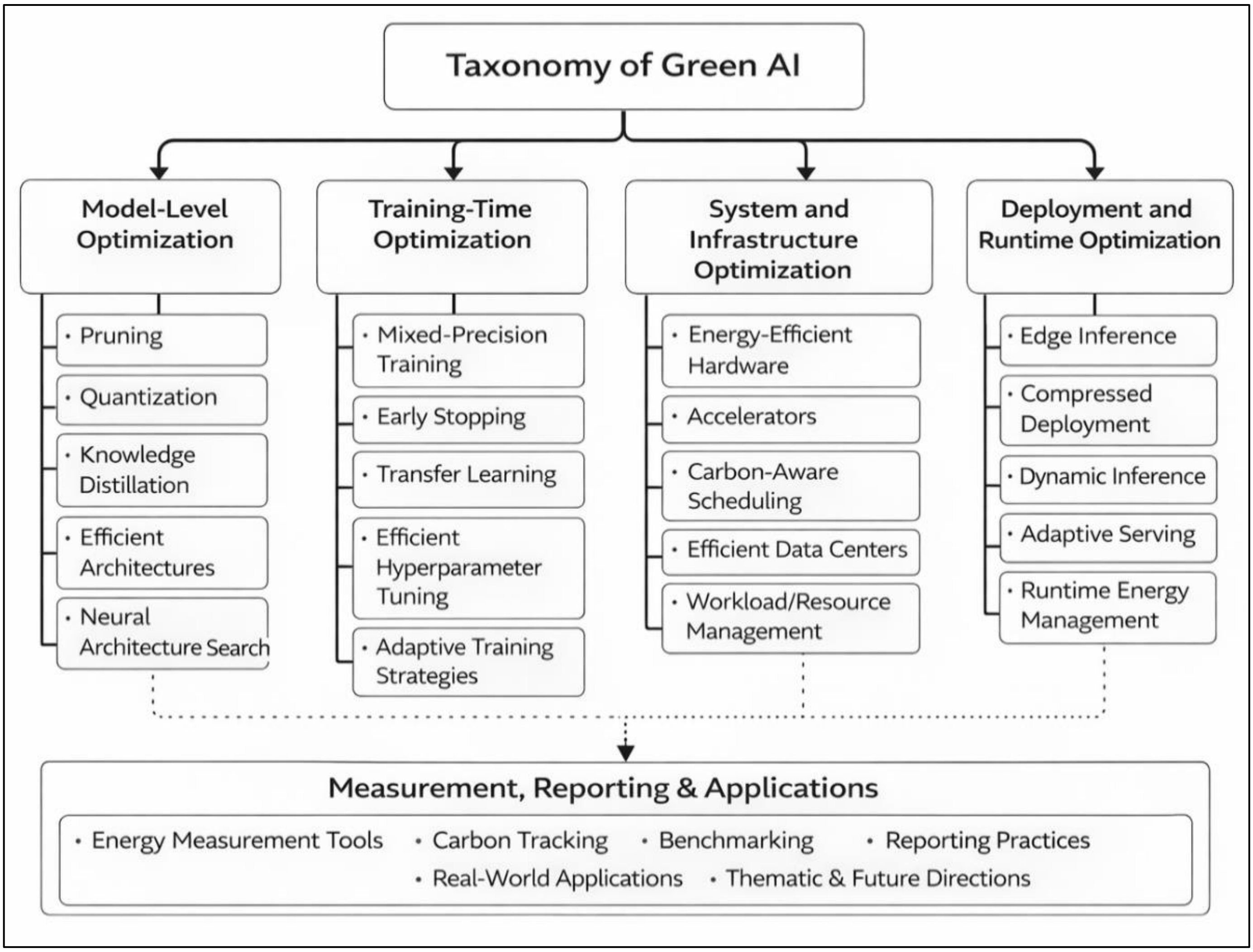

Fig 3 Taxonomy of Green AI

➢ *Model-Level Optimization*

Model-level optimization aims at enhancing the internal composition of AI models to minimize the computational intricacy, memory utilization and energy use without changing the satisfactory predictive outcomes. This type aims at the model itself either prior to or following the training and it is among the most popular directions of Green AI.

- *Pruning*

Pruning is the method of eliminating the superfluous or less significant parameters, neurons, channels or connection of a neural network. The rationale behind this is that a large number of trained models have extra weights, which do not add significant value to the final prediction. Pruning can decrease the size of models by removing redundant parts, reduce memory requirements and decrease the number of computations required at inference. Consequently, pruned models are able to operate quicker and use less power, particularly on constrained resources devices.

- *Quantization*

Quantization lowers the numerical accuracy of representation of model parameters and computations. As an example, a model can represent values as 16-bit, 8-bit, or lower-precision values instead of 32-bit floating-point values. This minimizes the consumption of memory and also speeds up arithmetic operations thus enhancing efficiency and energy saving. Quantization in particular can be applied during deployment as it enables large models to run on mobile, edge and embedded devices more efficiently.

- *Knowledge Distillation*

Knowledge distillation is a compression technique whereby a large and powerful model, which is referred to as the teacher model, transfers its learned knowledge to a smaller model, referred to as the student model. The student model is also designed to imitate the teacher behaviors but with fewer parameters and calculations. This allows it to save a lot of the predictive power of the original model but with a great deal of efficiency enhancement. In Green AI, distillation is relevant since it aids in developing lightweight models that can be used to deploy them sustainably.

- *Efficient Architectures*

Efficient architectures Neural network designs that are designed to be strong with reduced cost of computation are called efficient architectures. However, rather than reducing the size of a large model by training, this technique creates the model in a way that is efficient in the first place. Lightweight convolutional models and small variants of transformers can be used as examples. This aims at minimizing operations, memory access and parameter count without sacrificing the high precision. This is the focus of the Green AI approach since the architectural efficiency is a direct control of the training cost and inference energy.

- *Neural Architecture Search*

Neural Architecture Search (also known as NAS) is an automated way of searching network architectures that work. Accuracy is not the only criterion that can direct NAS in the context of Green AI, the efficiency-related criteria of latency, memory usage and energy consumption can also be employed. This facilitates the automatic creation of architectures that will be more sustainable AI systems. Even though not all NAS is computationally inexpensive, more efficient or constrained NAS techniques can aid in green model design discovery.

➢ *Training-Time Optimization*

The cost of model training can be lowered by training-time optimization. This category is critical in Green AI since the huge time, electricity and hardware demands that large AI models can consume, this segment is quite critical.

- *Mixed-Precision Training*

Mixed-precision training is a training method that involves using both high-precision and low-precision arithmetic. Some of the operations are not carried out in full precision, but in lower precision format that consumes less memory and less computation effort. This accelerates training, conserves energy and enables modern hardware accelerators to be used more productively. It is a viable method of ensuring that deep learning training is more sustainable.

- *Early Stopping*

Early stopping is a training method that ends the learning process when the model ceases to exhibit significant improvement on the validation data. The training is terminated when further computation may not enhance performance as opposed to running to a predetermined number of epochs. This will save unnecessary energy consumption, lessen the time of training and decrease the workload on hardware. Early stopping is also useful in Green AI since it eliminates overtraining expenses.

- *Transfer Learning*

Transfer learning uses the knowledge gained on one problem or dataset and transfers it to another but related problem. Rather than training a model, an existing model is fine-tuned to the application required. This significantly limits the size of the data, time and computational power needed to train. Transfer learning is a well-known application in Green AI since it enables the efficient creation of the model and reduces the use of resources in general.

- *Efficient Hyperparameter Tuning*

The art of finding optimal settings to a model, including learning rate, batch size or network depth, is known as hyperparameter tuning. The conventional tuning techniques are computationally complex in that they involve repeated training cycles. The goal of efficient hyperparameter tuning is to minimize this cost with more intelligent search algorithms, early stopping or adaptive optimization. This assists in reducing the environmental cost attributed to the repetitive experimentation.

- *Adaptive Training Strategies*

Adaptive training strategies are used to dynamically adjust training process based on the progress of the model, property of the data or the available resource. They include adaptive learning schedules, curriculum learning or dynamic batch adjustment. These techniques enhance efficiency by distributing the computation power where it is the most required. Adaptive training can be useful in Green AI since it facilitates smarter and less wasteful learning.

➢ *System and Infrastructure Optimization*

This group deals with the computer setting where AI models are trained and implemented. In the case of an efficient model, the total environment impact may be high due to the non-optimization of the infrastructure beneath them.

- *Energy-Efficient Hardware*

Hardware Energetic hardware consists of processors, accelerators and special purpose chips that are capable of running workloads of AI more efficiently. An example is GGPUs, TPUs and AI-specific low-power processors. Appropriate hardware can significantly minimise the amount of energy consumed and enhance the computational performance. Hardware selection is a significant consideration in Green AI as it is a direct cause of the energy footprint of training and inference.

- *Accelerators*

The hardware used to accelerate the AI computation is referred to as accelerators. They enhance performance through optimization of the operations of matrices, parallel processing and memory access. AI tasks can be executed more quickly and using less energy by efficient accelerators than by general-purpose processors. Their purpose in Green AI is to offer increased computation per watt to offer sustainable AI workloads.

- *Carbon-Aware Scheduling*

Carbon-conscious scheduling, which is also known as carbon-awareness scheduling, is the process of assigning AI workloads in proportion to the carbon intensity of the available electricity or of the computing resources. As an illustration, the training of the model can be planned during the times or places when renewable energy is more accessible or where electricity has a lesser carbon footprint. This method does not necessarily decrease computation, but it decreases the environmental cost of the computation. It is one of the growing tendencies towards sustainable AI operations.

- *Efficient Data Centers*

Efficient data centers are computer centers that are tailored to minimize wastage of energy by efficient cooling mechanisms, integration of renewable energy, server usage and balancing of workloads. The sustainability of data centers has a significant impact on the sustainability of AI since most applications of AI are based on cloud-based infrastructures. The advantage of Green AI is that the high-quality data centers are efficient since large-scale computing lowers the cost of the environment.

- *Workload and Resource Management*

Workload and resource management refers to the allocation of computing, memory, storage, and network resources in an efficient manner among AI tasks. Management eliminates the idle hardware, redundant duplication and ineffective scheduling. This enhances the use and minimizes energy wastage. With Green AI, it is necessary to orchestrate resources to make computing infrastructure to run in a sustainable way.

➢ *Optimization of Deployment and Runtime*

Deployment and runtime optimization deal with efficiency of AI systems once they are used in training, when they are actively being used in the real-life situation. This is essential since the long-term deployment of the deployment can potentially use more energy than training.

- *Edge Inference*

Edge inference refers to the execution of AI models on local devices (smartphones, internet of things nodes, cameras and embedded systems) instead of fully depending on cloud servers. This is able to minimize the communication overhead, latency and bandwidth usage. Performed effectively edge inference can also help to achieve more ecological operation by reducing redundant data transfer and allowing local decisions to be made.

- *Compressed Deployment*

Compressed deployment is deployment of small versions of AI models following such techniques as pruning, quantization and distillation. This is aimed at ensuring that the deployed model uses less in terms of computational and memory resources. This enhances responsiveness, reduces energy consumption and renders AI services more sensible in limited hardware.

- *Dynamic Inference*

Dynamic inference is a plan whereby the level of computation to be applied during prediction varies with the complexity of the input. Simple inputs can be computed with fewer layers or components whereas challenging inputs can be computed with more computation. The adaptive behavior decreases the average energy consumption since the model does not necessarily need all of the inputs to achieve the entire model complexity.

- *Adaptive Serving*

Adaptive serving is the idea of runtime systems that vary the manner in which the AI services are provided according to conditions on workload, latency, or even resources. As an example, the system can select among the possible model versions or dynamically allocate its resources. This assists in ensuring efficiency and quality of services simultaneously. Green AI uses adaptive serving to minimize wastage in computing based on the real demand.

- *Runtime Energy Management*

Runtime energy management services consist of monitoring and controlling the energy utilization of AI systems during the runtime of the systems. This can involve changing the frequency of the processor, choosing efficient execution mode or redistribution of workloads between devices. This is to minimize the amount of energy used when Live is operating without impacting performance too hard. Specifically, it is significant when the AI service is large-scale or constantly running.

➢ *Measurement, Reporting, and Applications*

This type contributes to the entire Green AI concept since sustainability cannot be more efficiently enhanced unless it is assessed, communicated and implemented in the actual systems.

- *Energy Measurement Tools*

The electric power used in training and inference of AI is measured by energy measurement devices. These tools assist researchers and practitioners to have an idea of the amount of energy consumed by a model and where inefficiencies are present. Sustainability claims are not complete without this kind of measurement.

- *Carbon Tracking*

Carbon tracking estimates the greenhouse gas emissions with AI workloads. It generally integrates power consumption with the data of the source of electricity or grid carbon intensity. Carbon tracking is significant since energy consumption is not sufficient to measure environmental impact.

- *Benchmarking*

Benchmarking is a process that compares AI models and systems to standardized assessment locations. In Green AI, accuracy is not the only benchmarking parameter and others such as energy and latency, memory and carbon-related metrics should be addressed. This advocates a good and fair comparison of methods.

- *Reporting Practices*

Reporting practices are the practices through which researchers report the cost of their AI systems, including both computational and environmental costs. The data that can be reported transparently can be hardware details, time spent on training, energy consumption, estimated emissions and efficiency trade-offs. Reporting enhances the reproducibility and encourages the responsible development of AI.

- *Real-World Applications*

Practical applications illustrate the application of Green AI methods in practice in such areas as healthcare, transportation, agriculture, smart cities, industry, and cloud systems. These applications demonstrate that AI with a sustainability focus is not a far-fetched aspiration, but a real need.

- *Future Directions*

Future directions in Green AI involve creating more green foundation models, enhancing carbon-conscious scheduling, creating efficient hardware-software co-design and standardizing sustainability. Such guidelines are meant to help AI systems become more competent and responsible to the environment.

V. APPLICATIONS OF GREEN AI

Applications of Green AI refer to the practical use of sustainable artificial intelligence methods to reduce energy consumption, carbon emissions and resource waste across real-world domains such as buildings, industry, agriculture, data centers and supply chains. Representative application-oriented studies and their main contributions, benefits and limitations are summarized in Table 1.

Table 1 Applications of Green AI

Ref.	Research Paper	Main Idea of the Study	Benefit / Contribution	Limitation / Drawback
[2]	"The Carbon Footprint of Machine Learning Training Will Plateau, Then Shrink"	Studies how hardware choice, data-center efficiency and cleaner electricity can reduce the environmental impact of ML training.	Shows that better practices can reduce training energy by up to 100× and carbon emissions by up to 1000×, making it highly relevant for practical Green AI deployment pipelines.	Focuses mainly on large-scale training environments and operational best practices rather than a single domain-specific end-user application.
[5]	"Towards Sustainable AI: A Comprehensive Framework for Green AI"	Proposes a full framework for embedding sustainability across AI design, training, deployment and evaluation.	Useful as an implementation-oriented foundation for applying Green AI systematically across projects and sectors.	Framework-oriented rather than validated through many large real-world case studies.
[6]	"Green Artificial Intelligence Initiatives: Potentials and Challenges"	Examines practical Green AI initiatives and how organizations can adopt energy- and carbon-aware AI practices.	Helps map where Green AI can be practically implemented and what institutional value it can deliver.	More strategic and analytical than experimentally grounded in a single deployed application.
[8]	"Achieving Green AI with Energy-Efficient Deep Learning Using Neuromorphic Computing"	Explores neuromorphic computing as a practical pathway for lower-power deep learning execution.	Strong applied value for low-energy AI hardware design and future efficient inference platforms.	The approach depends on specialized hardware ecosystems that are not yet broadly available in all application settings.
[11]	"Green IN Artificial Intelligence from a Software Perspective: State-of-the-Art and Green Decalogue"	Addresses Green AI from the software-engineering side, emphasizing sustainable development practices and lifecycle thinking.	Valuable for turning Green AI into actionable software practices rather than only model-level optimization.	More guidance-oriented than experimentally benchmarked on a specific industrial deployment.
[13]	"Potential of Artificial Intelligence in Reducing Energy and Carbon Emissions of Commercial Buildings at Scale,"	Evaluates how AI can reduce energy use and carbon emissions in commercial buildings through equipment, occupancy, control/operation and design improvements.	Reports that AI adoption could reduce building energy use and carbon emissions by about 8%–19% by 2050, which is a strong real-sector Green AI application.	Scenario-based analysis for commercial buildings; transferability to other sectors is not automatic.
[14]	"AI-Enabled Energy Baselines for Verified Building Decarbonization"	Develops an AI-based framework for establishing dynamic energy baselines and verifying savings in real building operations.	Demonstrates practical decarbonization assessment in buildings, with reported RMSE below 5% in general and cumulative emissions reduction evidence in the case study.	Building-retrofit focused, its workflow is highly useful but still domain-specific.
[15]	"Complex Artificial Intelligence Models for Energy Consumption of School Buildings"	Uses ML and DL models to estimate school-building energy consumption from real building data.	Shows practical value for educational-building energy planning, the study reports strong results, with the decision-tree model reaching an average prediction error of about 3.58% and gradient boosting performing best overall.	Focused on school buildings and a limited building sample, so broader generalization needs further validation.

[16]	"Multi-Agent Attention-Based Deep Reinforcement Learning for Demand Response in Grid-Responsive Buildings"	Applies multi-agent DRL to coordinate building-level demand response and energy management.	Offers a practical AI control strategy for reducing electrical demand in building networks while respecting operating constraints.	Reinforcement-learning deployment may require careful tuning, simulation support, and reliable building-state data before real-world scaling.
[17]	"Multi-Agent Deep Reinforcement Learning Based Demand Response and Energy Management for Heavy Industries with Discrete Manufacturing Systems"	Proposes a MARL-based energy-management system for energy-intensive industrial settings under demand-response schemes.	Strong industrial Green AI application because it targets real heavy-industry energy optimization and provides a reusable framework for future implementations.	Industrial deployment complexity and data/integration requirements may limit immediate adoption across heterogeneous factories.
[18]	"A Data Center Energy Efficiency Optimization Method Based on Optimal Temperature Control of Designated Active Servers"	Optimizes joint IT–cooling operation in data centers using active-server temperature control.	Practical value for greener AI infrastructure; the proposed method reduced total data-center energy consumption by about 12.09% for homogeneous VMs and 29.1% for heterogeneous VMs in the reported tests.	Designed around a specific IT–cooling coordination strategy and tested in a particular data-center setting, so broader replication is still needed.
[19]	"Improving Efficiency and Sustainability via Supply Chain Optimization through CNNs and BiLSTM"	Uses CNNs and BiLSTM to improve demand forecasting, resource allocation and sustainable supply-chain optimization.	Connects AI directly to lower stock-outs, shorter lead times, route optimization and carbon-footprint minimization; the hybrid model reported 96.57% accuracy.	Mainly evaluated as a predictive/optimization model; operational deployment details and external validation remain limited.
[20]	"A Machine Learning-Based Irrigation Prediction Model for Cherry Tomatoes in Greenhouses: Leveraging Optimal Growth Data for Precision Irrigation"	Builds an ML-based smart irrigation system using greenhouse experiments and IoT deployment for precision irrigation.	Practical sustainable-agriculture application; field trials reported 15.1% higher yield and 17.2% greater water efficiency.	Crop- and greenhouse-specific design may require adaptation before transfer to other crops or open-field conditions.
[21]	"Smart Agriculture Using IoT for Automated Irrigation, Water and Energy Efficiency"	Integrates IoT sensing, predictive algorithms and automated control for irrigation and resource optimization.	Field trials reported about 30% lower water use while maintaining target soil moisture, with low average power demand and reduced labor burden.	The system's performance may depend on local sensing quality, environmental conditions and connectivity reliability.

Across the previous works, Green AI applications generally concentrate on improving sustainability in a small number of high-impact domains, especially smart buildings, data centers, manufacturing, agriculture, and supply chains. Majority of the studies employ AI to optimize energy usage, suppress carbon emission, enhance resource allocation or assist decision-making regarding operations within real world conditions. One general trend is that a certain number of works are predictive in nature like energy forecasting, irrigation prediction or demand estimation, whereas others are control-oriented like demand response, energy management and infrastructure scheduling. On the whole, these papers demonstrate that Green AI can be used as a measure to reduce the waste of operations and to enhance the sustainability performance, but most of them are domain-specific, based on the single-objective optimization and usually measure the success in terms of energy savings or their prediction accuracy instead of a coherent framework of energy, carbon, cost and real-time adaptive intelligence.

We think should be to come up with a multi-objective, dynamically-real-time Green AI architecture that simultaneously optimizes energy use, carbon footprint, system performance and operational cost via adaptive intelligence as opposed to solely using a static prediction. As an illustration, you might create a practical scholarly paper founded on reinforcement learning or multi-agent reinforcement learning of carbon-conscious energy management in intelligent structures or information facilities, with the agent selecting control measures in real-time in accordance with workload, renewable energy accessibility, the intensity of electricity carbon and service demands. This concept can do better than most of the past works since it would shift out of individual-sector prediction models and implement an enclosed loop, decision-driven Green AI system that not only can be measured, deployed and comprehensive in assessing sustainability.

VI. EVALUATION METRICS AND TOOLS

Sustainability-oriented and computational metrics are often used together to evaluate green AI systems as accuracy on its own cannot be deemed enough to determine whether an AI model is environmentally efficient. The indicators most commonly employed in this case are energy consumption, carbon footprint, training time, inference latency, memory footprint and energy-efficiency ratio and practical evaluation is commonly aided by experiments impact tracker, carbon tracker, green algorithms, code carbon, eco2 AI [6] [7] [22] [23] [24] [25] [26]. These measures and indicators enable researchers to measure environmental and operational impact cost of AI models when training and deploying them as presented in Table 2 and Table 3.

Table 2 Equations and Metrics of Green AI Evaluation

Metric	Equation	Symbols	Description
Energy Consumption	$E \approx \sum i = 1nPi\Delta t$	(E): total energy consumption; (P(t)): instantaneous power at time (t); (T): total execution time	Measures the total electrical energy consumed during execution.
Discrete Energy Estimate	$E \approx \sum i = 1nPi\Delta t$	(E): total energy; (Pi): power measured at interval (i); (Δt): sampling interval; (n): number of intervals	Provides a practical approximation of total energy based on sampled power values.
Carbon Footprint	$C = E \times \gamma$	(C): carbon footprint; (E): energy consumption; (γ): carbon intensity factor of electricity	Estimates carbon emissions based on energy use and electricity source.
Training Time	Ttrain=tend - tstart	(Ttrain): total training time; (tstart): training start time; (tend): training end time	Represents the total duration required for model training.
Inference Latency	$Linf = N \sum i = 1Nti$	(Linf): average inference latency; (ti): inference time for sample (i); (N): number of inference samples	Measures the average response time per inference.
Memory Footprint	$M = Np \times b$	(M): memory footprint; (Np): number of model parameters; (b): bytes per parameter	Estimates the memory required to store model parameters.
Energy Efficiency Ratio	$\eta = EA$	(η): energy-efficiency ratio; (A): accuracy or utility measure; (E): energy consumption	Evaluates how much useful performance is achieved per unit of energy.

Table 3 Evaluation Tools

Tool	Main Purpose	Main Output
Experiment Impact Tracker	Tracks environmental impact during ML experiments	Energy use, carbon emissions
Carbontracker	Tracks and predicts carbon cost during training	Energy and CO2 estimates
Green Algorithms	Estimates carbon footprint of computational workloads	Carbon footprint estimate
CodeCarbon	Lightweight emissions tracking for AI experiments	Energy and CO2 estimates
eco2AI	Automated carbon accounting for AI workflows	Emissions and resource statistics

VII. CHALLENGES AND FUTURE DIRECTIONS

Although Green AI is increasingly becoming more important, there are various obstacles that curb its consumption. Some of the major concerns are the trade-off between model accuracy and energy efficiency, lack of standardized evaluation metrics, inconsistent carbon and energy costs reporting, hardware dependency and high computational costs of large-scale models, including foundation models and LLMs. Moreover, a number of Green AI studies remain hard to compare due to the fact that they involve various experimental conditions, instruments and sustainability measurements.

Future studies ought to work on the creation of uniform standards of Green AI assessment, the enhancement of carbon-conscious scheduling and energy-efficient training techniques, and designs of lighter but strong models to be deployed into the real world. There is also the necessity to pay more attention to sustainable foundation models, efficient edge AI, hardware-software co-design and transparent reporting structures that would make Green AI more feasible, quantifiable, and universally applicable.

VIII. CONCLUSION

Green AI has become an important research topic that aims to achieve sustainability, efficiency and environmental responsibility of artificial intelligence systems. This survey showed that the phrase Green AI goes beyond model compression and energy efficiency and covers optimization throughout the entire AI lifecycle, both in model design and training and in deployment, evaluation and application. According to the reviewed literature, the principal results of this survey could be concluded as follows:

- Green AI offers a generalized conceptual framework, which minimizes energy usage, carbon emissions and computational overhead in the state-of-the-art AI systems.
- The primary research trends in Green AI are model level optimization, training time optimization, system and infrastructure optimization and deployment/runtime optimization.
- Existing ways of using Green AI have demonstrated positive outcomes in various practical fields such as buildings, agriculture, manufacturing, data centers and supply chains.
- Even with this development, Green AI is not free of significant challenges, especially regarding standardized assessment, open disclosure and the sustainability requirements of AI models of large scale.

In general, it can be stated that Green AI will become the core of the future of artificial intelligence when it comes to balancing technical development with environmental friendliness. Continued research in this area will be essential for developing AI systems that are not only powerful, but also efficient, responsible and sustainable.

REFERENCES

[1]. Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. Communications of the ACM, 63(12), 54–63. https://doi.org/10.1145/3381831.

[2]. Patterson, D., Gonzalez, J., Holzle, U., Le, Q., Liang, C., Munguia, L., Rothchild, D., So, D. R., Texier, M., & Dean, J. (2022). The carbon footprint of machine learning training will plateau, then shrink. Computer, 55(7), 18–28. https://doi.org/10.1109/mc.2022.3148714.

[3]. Verdecchia, R., Sallou, J., & Cruz, L. (2023). A systematic review of Green AI. Wiley Interdisciplinary Reviews Data Mining and Knowledge Discovery, 13(4). https://doi.org/10.1002/widm.1507.

[4]. Bolón-Canedo, V., Morán-Fernández, L., Cancela, B., & Alonso-Betanzos, A. (2024). A review of green artificial intelligence: Towards a more sustainable future. Neurocomputing, 599, 128096. https://doi.org/10.1016/j.neucom.2024.128096.

[5]. Tabbakh, A., Amin, L. A., Islam, M., Mahmud, G. M. I., Chowdhury, I. K., & Mukta, M. S. H. (2024). Towards sustainable AI: a comprehensive framework for Green AI. Discover Sustainability, 5(1). https://doi.org/10.1007/s43621-024-00641-4.

[6]. Alzoubi, Y. I., & Mishra, A. (2024). Green artificial intelligence initiatives: Potentials and challenges. Journal of Cleaner Production, 468, 143090. https://doi.org/10.1016/j.jclepro.2024.143090.

[7]. Khan, S., Naz, N. S., Mazhar, T., Tariq, M. U., Shahzad, T., Guizani, S., & Hamam, H. (2025). Green AI techniques for reducing energy consumption in AI systems. Array, 29, 100652. https://doi.org/10.1016/j.array.2025.10065.

[8]. Luo, T., Wong, W., Goh, R. S. M., Tuan, A., DO, Chen, Z., Li, H., Jiang, W., & Yau, W. (2023). Achieving Green AI with Energy-Efficient Deep Learning Using Neuromorphic Computing. Communications of the ACM, 66(7), 52–57. https://doi.org/10.1145/3588591.

[9]. Desislavov, R., Martínez-Plumed, F., & Hernández-Orallo, J. (2023). Trends in AI inference energy consumption: Beyond the performance-vs-parameter laws of deep learning. Sustainable Computing Informatics and Systems, 38, 100857. https://doi.org/10.1016/j.suscom.2023.100857

[10]. De Vries, A. (2023). The growing energy footprint of artificial intelligence. Joule, 7(10), 2191–2194. https://doi.org/10.1016/j.joule.2023.09.004.

[11]. Gutiérrez, M., Moraga, M. Á., García, F., & Calero, C. (2024). Green IN Artificial Intelligence from a Software Perspective: State-of-the-Art and Green Decalogue. ACM Computing Surveys, 57(3), 1–30. https://doi.org/10.1145/3698111.

[12]. Raman, R., Pattnaik, D., Lathabai, H. H., Kumar, C., Govindan, K., & Nedungadi, P. (2024). Green and sustainable AI research: an integrated thematic and topic modeling analysis. Journal of Big Data, 11(1). https://doi.org/10.1186/s40537-024-00920-x.

[13]. Ding, C., Ke, J., Levine, M., Granderson, J., & Zhou, N. (2024). Potential of artificial intelligence in reducing energy and carbon emissions of commercial buildings at scale. Nature Communications, 15(1), 5916. https://doi.org/10.1038/s41467-024-50088-4.

[14]. Li, J., Hao, Y., & Li, Y. (2026). AI-enabled energy baselines for verified building decarbonization. Scientific Reports, 16(1), 5815. https://doi.org/10.1038/s41598-026-36284-w.

[15]. Tariq, R., Mohammed, A., Alshibani, A., & Ramírez-Montoya, M. S. (2024). Complex artificial intelligence models for energy sustainability in educational buildings. Scientific Reports, 14(1), 15020. https://doi.org/10.1038/s41598-024-65727-5.

[16]. Xie, J., Ajagekar, A., & You, F. (2023). Multi-Agent attention-based deep reinforcement learning for demand response in grid-responsive buildings. Applied Energy, 342, 121162. https://doi.org/10.1016/j.apenergy.2023.121162.

[17]. Bashyal, A., Boroukhian, T., Veerachanchai, P., Naransukh, M., & Wicaksono, H. (2025). Multi-agent deep reinforcement learning based demand response and energy management for heavy industries with discrete manufacturing systems. Applied Energy, 392, 125990. https://doi.org/10.1016/j.apenergy.2025.125990.

[18]. Xu, Y., Tian, M., Yan, C., Li, D., Deng, T., Guo, P., & Hu, K. (2025). A data center energy efficiency optimization method based on optimal temperature control of designated active servers. Energy and Buildings, 345, 116126. https://doi.org/10.1016/j.enbuild.2025.116126.

[19]. Dalal, S., Lilhore, U. K., Simaiya, S., Radulescu, M., & Belascu, L. (2024). Improving efficiency and sustainability via supply chain optimization through CNNs and BiLSTM. Technological Forecasting and Social Change, 209, 123841. https://doi.org/10.1016/j.techfore.2024.123841.

[20]. Wang, Z., Liu, Z., Yuan, M., Yin, W., Zhang, C., Zhang, Z., & Hu, X. (2025). A machine learning-based irrigation prediction model for cherry tomatoes in greenhouses: Leveraging optimal growth data for precision irrigation. Computers and Electronics in Agriculture, 237, 110558. https://doi.org/10.1016/j.compag.2025.110558.

[21]. Gupta, S., Chowdhury, S., Govindaraj, R., Amesho, K. T., Shangdiar, S., Kadhila, T., & Iikela, S. (2025). Smart agriculture using IoT for automated irrigation, water and energy efficiency. Smart Agricultural Technology, 12, 101081. https://doi.org/10.1016/j.atech.2025.101081