Array 27 (2025) 100448

Contents lists available at ScienceDirect

Array

journal homepage: www.elsevier.com/locate/array

Taxonomy, challenges, and future directions for AI-driven industrial cooling systems

Md Mohsin Kabir *, Shahina Begum, Shaibal Barua, Mobyen Uddin Ahmed

School of Innovation, Design and Engineering, Mälardalens University, Västerås, 722 20, Sweden

ARTICLE INFO

Keywords:
AI-driven diagnostics
Deep learning
Industrial cooling systems
Machine learning
Physics-informed neural networks

ABSTRACT

The efficiency and reliability of industrial cooling systems are critical for sectors such as energy systems, electronics manufacturing, and data centers. Traditional cooling systems rely on reactive maintenance, leading to increased downtime, energy consumption, and operating costs. Recent advances in artificial intelligence (AI), including machine learning (ML), deep learning (DL), and physics-informed neural networks (PINNs), have enabled proactive fault diagnosis and predictive maintenance in industrial cooling systems, significantly reducing energy use and improving operational reliability. However, current AI applications face challenges, such as limited access to quality datasets, computational complexity, integration with legacy systems, and model scalability. This paper systematically addresses these gaps by providing a detailed taxonomy of AI-driven cooling system diagnostics, categorizing state-of-the-art methods, and identifying critical research challenges. Our main contribution is a structured taxonomy that integrates ML, DL, and PINNs, offering a clear framework for analyzing current practices and potential improvements. The paper highlights critical insights across 138 reviewed studies, emphasizing the transformative role of hybrid AI frameworks in diagnostics, including use cases in HVAC, data centers, and thermal imaging. Notably, the integration of ML, DL, and PINNs has been shown to improve fault detection accuracy, energy efficiency, and model interpretability, paving the way for scalable, real-time deployments.

1. Introduction

Industrial cooling systems are integral to the functioning of modern industries, particularly in sectors such as energy systems [1], electronics manufacturing [2], and data centers [3,4]. These systems ensure that sensitive components, such as electronic circuits, transformers, and servers, remain within their permissible operating temperature ranges. This is vital to prevent overheating, which can cause system failures, decreased performance, and significant economic losses. However, the increasing complexity of industrial processes and the growing demand for higher energy efficiency have exposed the limitations of traditional cooling methods. For instance, conventional designs often rely on worst-case scenarios to ensure thermal stability, leading to over-dimensioned, energy-intensive, and costly systems [5].

Fig. 1 illustrates the workflow of traditional cooling systems, highlighting their reactive nature, where sensor data are manually monitored, and faults are detected only after they occur, leading to significant downtime and inefficiency due to delayed fault detection and reactive maintenance. In contrast, Fig. 2 showcases the AI-driven cooling system, highlighting proactive and automated capabilities. AI models leverage real-time sensor data for fault detection and predictive maintenance, supported by continuous learning and feedback mechanisms. This approach minimizes downtime, enhances reliability, and transitions cooling systems from reactive to proactive diagnostics, significantly improving efficiency and reducing energy consumption.

The global push for sustainability and carbon neutrality by 2050 further emphasizes the need for sustainable cooling solutions [7]. Cooling systems account for a significant portion of industrial energy consumption, and uncontrolled temperature levels can result in catastrophic failures, such as outages of the electric grid or data center shutdowns. Fig. 3 highlights the primary causes of electronic equipment failures, with temperature being the dominant factor (53%), followed by vibration (22%), humidity (19%), and dust (6%) [8]. These failures not only cause financial losses but also have environmental and social consequences. Industries are exploring advanced technologies to make cooling systems more intelligent, adaptive, and efficient. Artificial Intelligence (AI) [9], particularly Machine Learning (ML) [10], Deep Learning (DL) [11], Reinforcement Learning (RL) [12], and Physics-Informed Neural Networks (PINNs) [13], has emerged as a promising solution

* Corresponding author.
E-mail addresses: md.mohsin.kabir@mdu.se (M.M. Kabir), shahina.begum@mdu.se (S. Begum), shaibal.barua@mdu.se (S. Barua), mobyen.uddin.ahmed@mdu.se (M.U. Ahmed).

https://doi.org/10.1016/j.array.2025.100448
Received 3 April 2025; Received in revised form 15 June 2025; Accepted 2 July 2025
Available online 17 July 2025

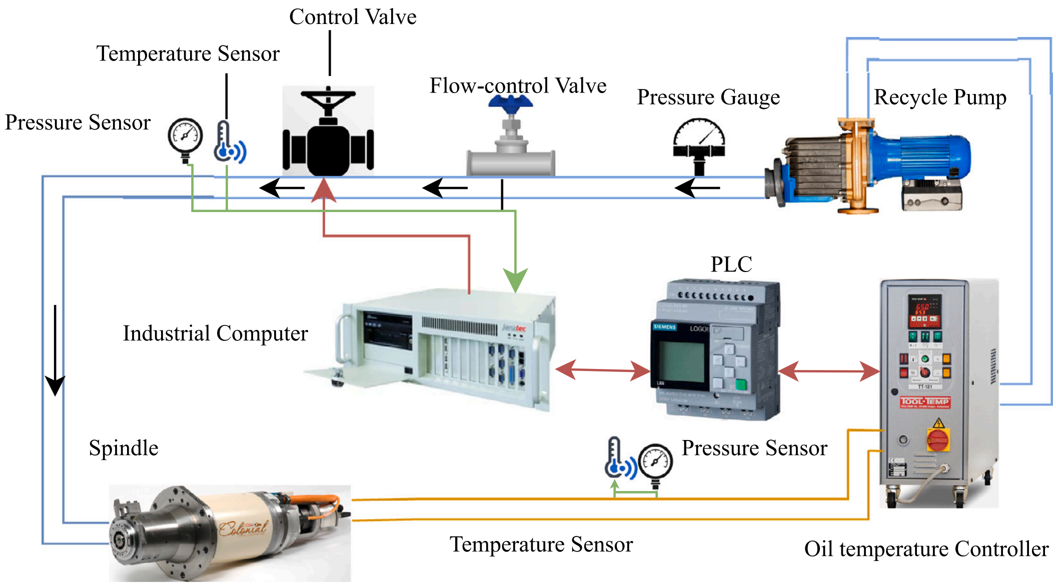


Fig. 1. Traditional cooling system framework [6].

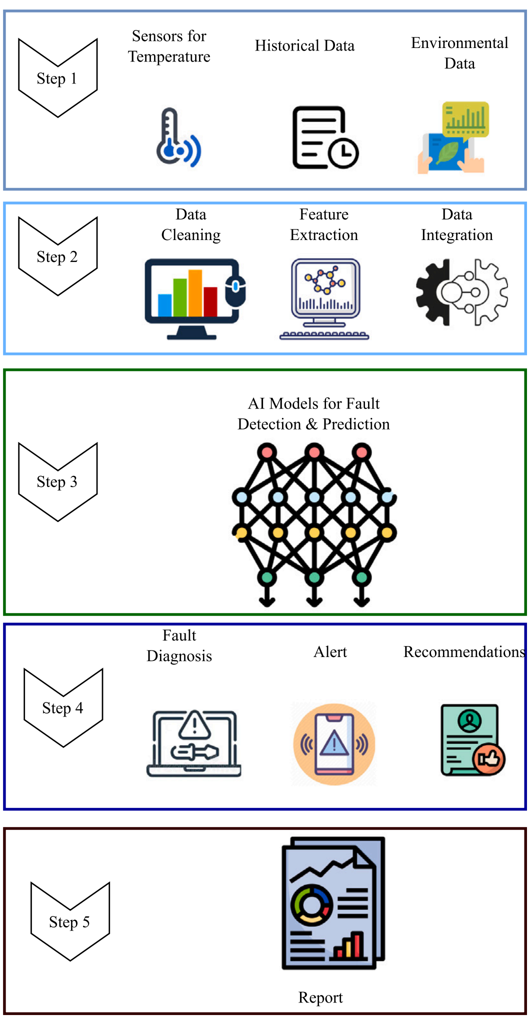


Fig. 2. AI-driven cooling system diagnostics framework.

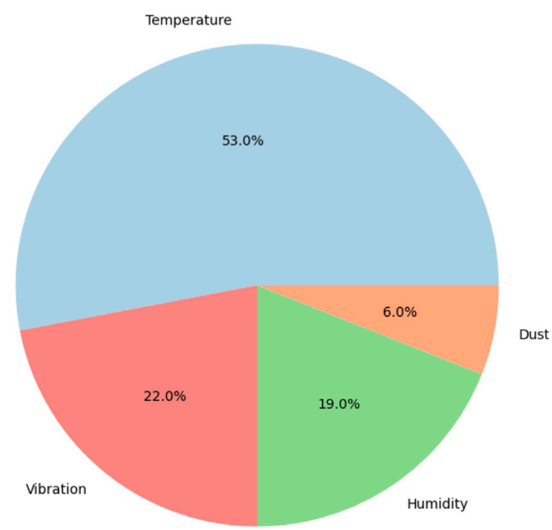


Fig. 3. Distribution of factors contributing to electronic equipment failures [8].

for AI-driven cooling systems. By leveraging these technologies, industries can shift from reactive maintenance to predictive and proactive diagnostics, ensuring system reliability and minimizing energy waste. Although AI technologies have shown potential in diagnostics and predictive maintenance, their application in industrial cooling systems and monitoring systems still faces several challenges. First, there is a lack of high-quality standardized datasets tailored to train AI models in the context of cooling systems. Existing data sets are often specific to particular industries or applications, limiting the generalizability of AI solutions. Second, understanding how cooling systems behave over time — especially as heat levels, outside conditions, and system settings change — requires advanced algorithms that combine real-world data with physical laws. Although physics-informed neural networks are a step forward [14,15], their implementation in real-world industrial settings is still in its infancy.

Furthermore, AI models face scalability challenges due to the high computational demands of processing large-scale industrial data, the need for seamless integration with legacy systems, and the variability in cooling system configurations across different industries. Real-time performance is hindered by the complexity of handling multi-source sensor

Table 1
Related surveys on AI-driven industrial cooling systems.

Focus area	Cite	Algorithms discussed	Key strengths	Limitations
AI for Industrial Failure Prediction	[17–23]	ML	Comprehensive overview of ML algorithms for failure prediction; Consistent and standardized reporting of studies to mitigate the impact of heterogeneity in knowledge accumulation.	Studies varied in methods, making comparisons difficult; Inconsistent reporting made data extraction hard; Study quality was judged only by publication
Predictive Maintenance for HVAC Systems	[9,24–26]	DL	Comprehensive review of predictive maintenance algorithms for HVAC systems; Not only highlights the advantages of predictive maintenance in HVAC systems but also critically assesses the constraints and challenges of different algorithmic approaches.	Limited Coverage of Recent Advances; Lacks real-world implementation or performance comparison; Issues like sensor failures, data sparsity, and noise may be overlooked; Limited discussion on feasibility in large-scale HVAC systems
Optimization of Cooling and Thermal Systems	[27–30]	ML, DL	Detailed overview of ML and DL techniques applied to optimize cooling and thermal systems; Comprehensive analysis of thermal and cooling methods, focusing on how ML and DL improve their efficiency and performance.	No empirical case studies to support AI applications; AI adaptability across different thermal systems is not fully addressed
Applications in Specialized Industrial Domains	[31–37,37,38]	ML, DL, IoT	Addresses Industry 4.0 practices; Highlights the application of ML, DL, and IoT in solving complex problems in specialized industries; Covers challenges and future opportunities for implementing smart technologies in industries.	AI is not yet fully integrated into Industry 4.0, and the study does not suggest clear solutions; Cybersecurity challenges are mentioned but not analyzed in detail
General AI Techniques for Diagnostics and Maintenance	[39–42]	PINNs, PIMLs	Comprehensive integration of PIMLs and PINNs; Analyze the benefits of integrating physical knowledge of AI models; Addresses real-time applications and sustainability.	A limited number of literary works analyzed; The findings may be skewed toward certain authors or methodologies, affecting the accuracy of the overall trend in Condition Monitoring (CM) applications

data, latency in decision-making processes, and the computational overhead of deploying deep learning models in resource-constrained environments [16]. Cooling systems are often integrated with other subsystems, creating interdependencies that complicate diagnostics. These gaps in research and practice highlight the urgent need for systematic studies to synthesize existing advancements, address limitations, and propose actionable solutions for improving the reliability and efficiency of industrial cooling systems. This review is timely and necessary as industries face increasing demands for energy-efficient and reliable cooling solutions amidst growing pressures to achieve carbon neutrality by 2050. Integrating AI-driven methodologies offers a pathway to improve system reliability and reduce energy consumption and CO2 emissions.

The analysis of recent surveys in the domain of AI-driven diagnostics and predictive maintenance for industrial cooling systems highlights several research gaps. As shown in Table 1, existing studies on industrial failure prediction offer a broad view of ML applications; however, these studies often present variations in methodology and inconsistent data reporting, making cross-study comparisons challenging. Surveys focused on predictive maintenance in HVAC systems provide insights into DL techniques. Additionally, research addressing optimization in cooling and thermal systems gives a comprehensive overview of ML and DL methods, but has limited practical applicability. Furthermore, surveys exploring AI applications in specialized industrial domains emphasize the potential of Industry 4.0 practices but do not provide clear guidance on overcoming integration challenges, particularly in terms of cybersecurity. Finally, studies examining general AI techniques, such as physics-informed neural networks, offer useful perspectives on the integration of physical principles into data-driven models but analyze a limited number of works, which may lead to biased conclusions. These gaps underscore the need for future research to adopt standardized methodologies, conduct more real-world validations, and explore hybrid AI approaches to improve the reliability and effectiveness of predictive maintenance systems in industrial cooling environments.

In contrast, this review addresses these gaps by offering a comprehensive integration of ML, DL, and PINNs, along with a detailed taxonomy and analysis of their applications in industrial cooling systems. Moreover, this review focuses on integrating AI-driven methodologies into industrial cooling systems' diagnostics, monitoring, and predictive maintenance. It examines how AI technologies can significantly improve fault detection, real-time monitoring, and system optimization when paired with advanced sensing techniques. The emphasis lies on three primary areas: First, the role of AI in diagnostics and predictive maintenance, highlighting its ability to identify faults and predict system failures before they occur, thus minimizing operational downtime and extending the lifespan of industrial systems. Second, AI-driven solutions that leverage advanced sensing technologies — including infrared (IR) imaging, IoT sensors, and real-time data fusion — to improve data quality, enable precise fault localization, and enhance decision-making in cooling system management. Finally, the review discusses industrial applications, such as cooling systems in energy-intensive environments and electronics-based industries, including electric grid cooling, HVAC systems in data centers, and manufacturing processes. The following research questions (RQs) guided the review process:

- **RQ1:** How do different AI techniques (machine learning, deep learning, and PINNs) impact the diagnostics and predictive maintenance in industrial cooling systems?
- **RQ2:** In what ways do AI-driven diagnostics and predictive maintenance approaches differ across various industrial cooling application domains, such as HVAC, data centers, and refrigeration systems?
- **RQ3:** Which critical gaps exist in current AI-driven cooling diagnostics research, and how can addressing these gaps guide future methodological advancements and practical deployments?

The research questions outlined above frame the key areas of investigation in this review. By addressing these questions, this study systematically explores the current landscape of AI-driven diagnostics

Table 2
Inclusion and exclusion criteria.

Criteria	Details
Inclusion Criteria	• Peer-reviewed review or survey articles published between 2018 and 2025. • Studies focusing on AI techniques (ML, DL, PINNs) in diagnostics and predictive maintenance for industrial cooling systems. • Papers discussing applications of AI in fault detection, energy optimization, and thermal management. • Articles providing detailed methodologies and frameworks for AI-driven solutions. • Studies addressing sustainability and energy efficiency in industrial systems.
Exclusion Criteria	• Articles not published in peer-reviewed journals or conferences. • Studies lacking methodological or experimental details. • Papers focusing on non-industrial applications, such as consumer electronics or automotive systems. • Articles unrelated to AI or predictive maintenance for cooling systems. • Duplicate studies or studies published as abstracts without full-text availability.

and predictive maintenance in industrial cooling systems, identifying existing methodologies, application domains, and associated challenges. Through this structured inquiry, the review not only synthesizes prior research but also highlights critical gaps that necessitate further exploration. Building on these research questions, this study makes several key contributions:

- Provides a comprehensive review and taxonomy of AI-driven methodologies, categorizing techniques into machine learning, deep learning, and physics-informed neural networks for diagnostics and predictive maintenance in industrial cooling systems.
- Identifies and analyzes critical challenges in deploying AI solutions, including data scarcity, model scalability, computational overhead, and the need for real-time integration in complex industrial environments.
- Explores diverse application domains of AI in industrial cooling systems, such as fault detection, thermal imaging, energy optimization, and anomaly detection, providing actionable insights for researchers and practitioners.
- Proposes future research directions, including the development of interpretable AI models, hybrid approaches combining data-driven and physics-based methods, and advancements in sustainable cooling solutions.

The remainder of this paper is organized as follows: Section 2 presents the review methodology, detailing the systematic review approach based on PRISMA guidelines. Section 4 provides an in-depth analysis of AI techniques for diagnostics, categorizing them into machine learning-based, deep learning-based, and physics-informed neural networks-based approaches. Section 3 explores the applications of AI-driven diagnostics and predictive maintenance in industrial cooling systems, highlighting use cases such as fault detection, energy efficiency optimization, and real-time monitoring. Section 5 discusses the research questions and outcomes of this survey. Finally, Section 6 concludes the paper by summarizing the findings and emphasizing the significance of integrating AI techniques for advancing industrial cooling systems.

2. Review methodology

This review adopts a systematic methodology to identify, analyze, and synthesize studies relevant to AI-driven diagnostics and predictive maintenance for industrial cooling systems. The approach follows the PRISMA (Preferred Reporting Items for Systematic Reviews) guidelines to ensure transparency and applicability.

2.1. Review methodology

2.1.1. Databases searched

The literature search was conducted across leading academic databases, including IEEE Xplore, ScienceDirect, SpringerLink, and Scopus. These databases were chosen for their extensive coverage of research in artificial intelligence, machine learning, deep learning, and industrial systems.

2.1.2. Search strategy

A structured search query was developed to identify relevant studies. The primary search string included combinations of keywords representing AI techniques, application domains, and review-specific terms. Such as:

> [“Machine Learning” OR “Deep Learning” OR “Physics Informed Neural Networks”] AND (“Industrial Cooling Systems” OR “Fault Diagnosis” OR “Predictive Maintenance”)

The search was refined with Boolean operators and tailored to the syntax of each database to maximize coverage.

2.1.3. Inclusion and exclusion criteria

Specific inclusion and exclusion criteria were defined to ensure the relevance and quality of the studies included in this review. Table 2 summarizes the criteria used for study selection. This table ensures clarity in the selection process by outlining the specific parameters for study inclusion and exclusion. The criteria were applied consistently during the review process to maintain rigor and relevance.

2.1.4. Study selection process

The study selection process involved three key steps:

1. **Initial Screening:** Titles and abstracts were screened for relevance to the research scope.
2. **Full-Text Review:** Full texts of potentially relevant studies were evaluated against the inclusion and exclusion criteria.
3. **Final Selection:** Studies meeting all criteria were included for detailed analysis and synthesis.

2.1.5. Analysis and synthesis

For each selected study, key data points were extracted, including the focus area, AI methodologies (e.g., ML, DL, PINNs), applications discussed, challenges addressed, and notable findings. This data was synthesized to provide a comprehensive understanding of the field, identify gaps in the literature, and propose directions for future research.

2.1.6. PRISMA flow diagram

The PRISMA flow diagram (Fig. 4) summarizes the review process, illustrating the stages of study identification, screening, eligibility assessment, and inclusion. The inclusion of clear research questions ensures a focused approach to addressing key gaps and advancements in AI-driven diagnostics and predictive maintenance. By adhering to PRISMA guidelines, this study maintains a transparent and reproducible framework, providing valuable insights into the existing body of knowledge and identifying future research opportunities.

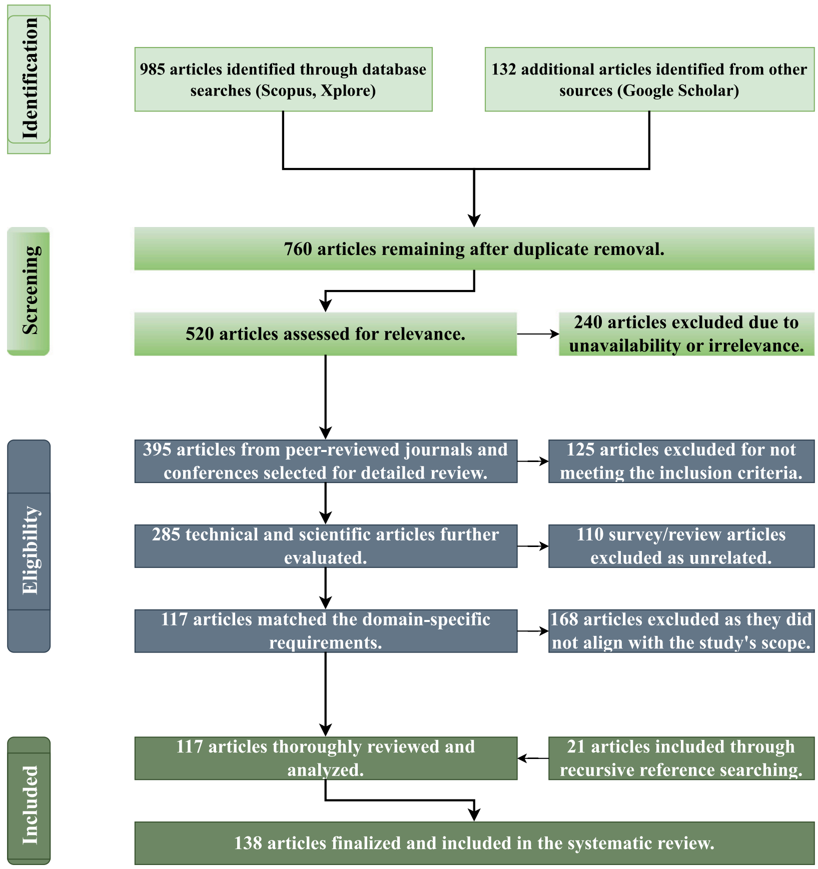


Fig. 4. PRISMA flow diagram for the review process.

2.2. Bibliometric analysis

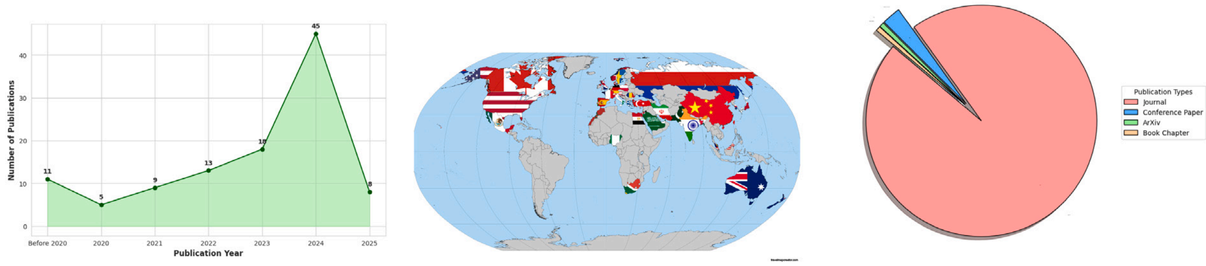
Fig. 5 provides a consolidated overview of key bibliometric insights into AI-driven diagnostics and predictive maintenance in industrial cooling systems. As shown in Fig. 5(a), the publication trend reveals a clear rise in research interest, with a sharp increase in contributions after 2020. While earlier years saw limited output, 2024 marks the peak with 45 published studies, reflecting a strong and growing academic focus on this topic. The geographic distribution in Fig. 5(b) further emphasizes the global scope of this research, with contributions from 42 countries spanning North America, Europe, Asia, Africa, and Australia. This international engagement highlights the widespread relevance of AI applications in industrial maintenance and suggests a collaborative global effort to address shared industrial challenges. Fig. 5(c) shows that journal articles comprise the majority (96%) of the selected sources, reinforcing the field's reliance on peer-reviewed, high-quality studies. Conference papers make up 2.6%, representing recent innovations and experimental approaches, while a small number of book chapters and preprints provide foundational context or emerging insights. Together, these figures indicate a maturing research domain marked by increasing output, global collaboration, and a preference for rigorously validated methodologies.

2.3. Evolution of AI methods in industrial cooling systems

Fig. 6 presents a structured timeline outlining the technological evolution of AI methodologies applied to fault detection and diagnostics (FDD) in industrial cooling systems. This progression reflects not only advancements in computational power and sensing infrastructure but also the increasing complexity and data-richness of modern industrial environments. Initial developments were grounded in rule-based systems and heuristic thresholds; however, the emergence of data-driven approaches marked a significant shift. Comprehensive reviews such as [26] have catalogued the transition toward statistical learning frameworks. Classical machine learning methods, including Support Vector Machines and ensemble classifiers, became prevalent in early applications due to their robustness and interpretability. Notable examples include the hybrid random forest and SVM approach by Tun et al. [43] and the LightGBM-based fault warning model proposed by Li et al. [44].

The subsequent rise of deep learning, enabled by increased availability of high-performance GPUs and thermal imaging sensors, facilitated automatic feature extraction and improved performance on high-dimensional data. Calderon-Uribe et al. [45] and Wiysobunri et al. [46] demonstrated the efficacy of convolutional neural networks (CNNs) for thermal fault detection in motors and server cooling systems, respectively. More recently, hybrid deep learning architectures have gained prominence to address limitations in labeled data availability and model generalizability. Transformer-based self-supervised learning for HVAC systems [47] and few-shot GAN approaches for imbalanced datasets [48] exemplify this trend.

The latest developments emphasize physics-informed learning paradigms, which integrate domain knowledge directly into the training process to enhance physical consistency and interpretability. Zhang et al. [49] employed PINNs for simulating conjugate heat transfer in microchannel heat sinks, while Wang et al. [50] introduced a physics-informed reinforcement learning framework for real-time control in data center cooling. Finally, emergent directions such as quantum-enhanced AI are beginning to show promise in complex energy systems, as illustrated by Sworna et al. [51]. Collectively, this timeline contextualizes the methodological trajectory reviewed in this paper and illustrates the domain-specific motivations behind each successive advancement.



(a) Publication trends over the years. (b) Geographic distribution. (c) Distribution of publication types.

Fig. 5. Overview of research trends in AI-driven diagnostics and predictive maintenance in industrial cooling systems: (a) publication trends, (b) geographic distribution, and (c) publication types.

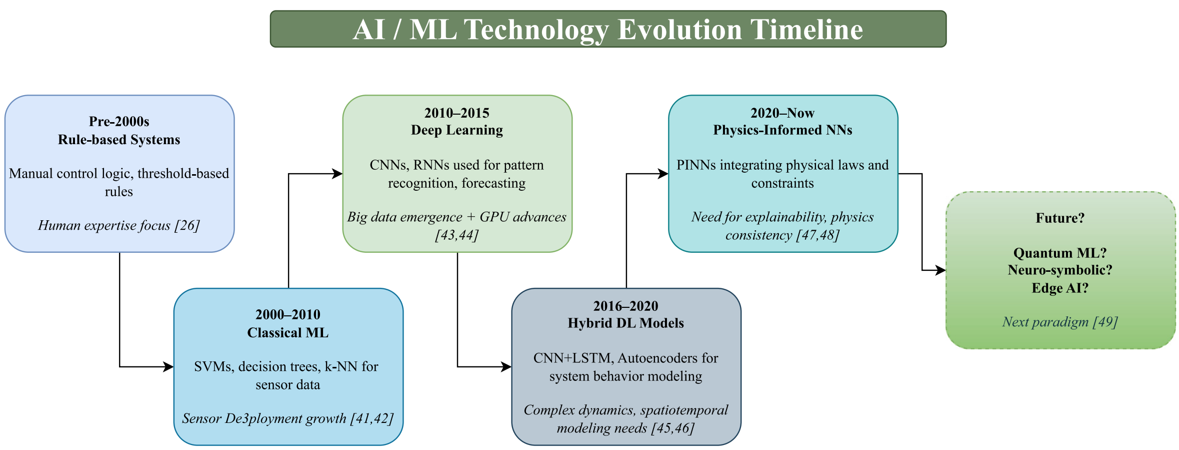


Fig. 6. Technology development pulse diagram showing the evolution of AI methods in industrial cooling systems and the key motivations for each transition.

3. Applications

The integration of AI technologies into industrial cooling systems has revolutionized diagnostics and predictive maintenance, enhancing system reliability, energy efficiency, and operational performance. This section categorizes the key application domains where AI-driven methods have demonstrated a significant impact. The taxonomy presented in Fig. 7 illustrates the key domains and techniques within AI-driven diagnostics and predictive maintenance for industrial cooling systems.

3.1. Fault detection and classification

AI algorithms are extensively used to identify and classify faults in industrial cooling systems. Fault detection aims to identify abnormal operating conditions, while classification determines the type and severity of the fault.

- **HVAC Systems:** Algorithms like Support Vector Machines and Random Forests classify faults such as refrigerant leaks, sensor failures, and compressor issues [52]. Chen et al. [26] provided a comprehensive review of data-driven fault detection and diagnostics (FDD) methods for HVAC systems, emphasizing the importance of supervised, semi-supervised, and unsupervised learning techniques. Their study outlined the steps in FDD processes, including data collection, preprocessing, baseline establishment, and fault detection. Support Vector Machine (SVM) and Random Forests were highlighted for their accuracy in fault classification, although challenges such as scalability, real-building deployment, and data privacy remain significant. The review stressed the need for improved interpretability and benchmarking to facilitate real-world adoption.
- **Thermal Imaging:** Deep learning models, particularly Convolutional Neural Networks, utilize thermal images to detect anomalies such as overheating or hotspots in cooling circuits. Javed et al. [53] proposed a novel methodology combining infrared thermography (IRT) with machine vision for fault detection in induction motors. The approach involved generating thermal image datasets under various operational conditions and using Local Octa Patterns (LOP) for feature extraction, followed by SVM classification. However, the reliance on pre-configured datasets and manual feature extraction limits scalability for broader applications.
- **Chiller Systems:** Physics-informed neural networks diagnose condenser fouling and other performance degradation issues by integrating operational data and physical laws. Shen et al. [54] proposed a data-driven self-attention-based deep learning method for diagnosing faults in HVAC chiller systems under imbalanced data scenarios. Their model combines the stable synthetic minority oversampling technique (SSMOTE) for data augmentation with a skip self-attention temporal convolutional network (STCN) for fault classification. Despite its high performance, the computational overhead and dependence on curated datasets remain limitations for real-time applications in complex industrial settings.

3.2. Predictive maintenance

Predictive maintenance leverages AI to forecast potential failures, enabling timely interventions to prevent downtime.

- **Remaining Useful Life Prediction:** Recurrent Neural Networks and Long Short-Term Memory Networks (LSTMs) analyze time-series data to estimate the lifespan of cooling components. Nunes

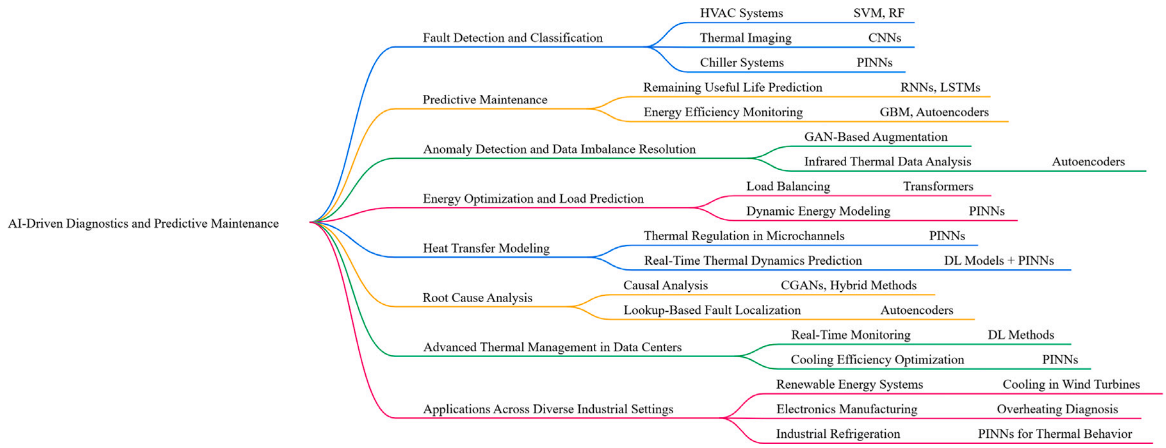


Fig. 7. Taxonomy of the applications of AI-driven diagnostics and predictive maintenance in industrial cooling systems (Application-driven approach).

et al. [55] reviewed various prognostic methods and their applications in predictive maintenance, with a focus on RUL prediction using data-driven approaches. The study emphasized the integration of anomaly detection with RNN and LSTM models to enhance the accuracy of RUL predictions in industrial systems. By addressing challenges such as noisy sensor data and data heterogeneity, the paper demonstrated how these models can effectively forecast equipment degradation and schedule timely maintenance actions. However, the computational demands and dependency on high-quality labeled datasets limit their scalability across diverse industrial scenarios.

- **Energy Efficiency Monitoring:** Algorithms like Gradient Boosting Machines and Autoencoders identify operational inefficiencies and predict energy-saving opportunities. Jeon et al. [56] proposed a predictive maintenance framework for compressed air systems, combining an LSTM-based motor power prediction model with ML-based radiator fault detection. However, the framework's reliance on curated datasets and high computational overhead for time-series analysis remains a limitation for scaling across diverse industrial applications.

3.3. Anomaly detection and data imbalance resolution

Anomaly detection identifies deviations from normal operational patterns, often in scenarios with imbalanced datasets.

- **GAN-Based Augmentation:** Generative Adversarial Networks generate synthetic data to improve fault detection accuracy and mitigate class imbalance. Ren et al. [48] introduced a Few-Shot GAN framework tailored for severe data imbalance in intelligent fault diagnosis. The method leverages a pre-training strategy on sample-rich classes followed by fine-tuning with anchor samples to generate diverse and realistic fault samples for sample-poor classes. The approach achieved very high accuracy in fault detection tasks by improving the diversity of generated samples while addressing GAN overfitting issues. Despite its advantages, the reliance on pre-trained generative models and computational complexity in high-dimensional sample spaces limits its scalability in real-world scenarios.
- **Infrared Thermal Data Analysis:** Autoencoders use reconstruction errors to detect anomalies in cooling systems based on thermal imaging. Sinap et al. [57] developed a CNN-based framework for detecting faults in photovoltaic solar modules using a dataset of 20,000 infrared images. The model demonstrated robust classification for various fault types, such as hotspots and shading, but relied heavily on curated datasets and required substantial computational resources for hyperparameter optimization, limiting scalability in broader industrial settings.

3.4. Energy optimization and load prediction

AI-driven methods improve the energy efficiency of industrial cooling systems by optimizing load distribution and minimizing wastage.

- **Load Balancing:** Transformer models analyze multivariate operational data to optimize cooling loads across components. Fan et al. [58] introduced the DTformer, a transformer-based model designed for multi-horizon, multi-energy load forecasting in integrated energy systems. The model incorporates a dual variable attention module and a temporal windowed attention (TWA) mechanism to capture both long-term temporal dependencies and variable interactions efficiently. Despite its high accuracy and reduced memory usage, the model's computational overhead and dependency on extensive hyperparameter tuning pose challenges for real-time industrial deployment.
- **Dynamic Energy Modeling:** PINNs predict thermal dynamics under varying loads, enabling real-time energy optimization. Xiao and You [59] proposed a Physically Consistent Deep Learning (PCDL) framework for building thermal modeling and energy optimization. The PCDL model integrates physics consistency into deep learning structures, leveraging RNN-LSTM hybrid architectures to predict indoor temperature and humidity dynamics. Despite these advancements, the model's computational overhead and dependence on extensive domain-specific constraints pose challenges for large-scale industrial deployment.

3.5. Heat transfer modeling

AI techniques are applied to simulate and optimize heat transfer processes critical to industrial cooling systems.

- **Thermal Regulation in Microchannels:** PINNs solve heat transfer equations to enhance cooling efficiency in compact systems. Zhang et al. [49] developed a Physics-Informed Neural Network framework to simulate conjugate heat transfer within manifold microchannel heat sinks (MMC) for high-power Insulated Gate Bipolar Transistor (IGBT) cooling. The study employed dual sub-PINNs to separately model flow dynamics and thermal behavior, integrating physical constraints from governing equations. Experimental results showed PINNs effectively predicted temperature distributions and pressure drops, achieving results comparable to Computational Fluid Dynamics (CFD) simulations while reducing computational cost. However, limitations included sensitivity to geometric complexities and numerical instabilities in scenarios with abrupt changes in flow patterns or gradients, which affected the accuracy near complex boundary conditions.

- **Real-Time Thermal Dynamics Prediction:** DL models combined with PINNs predict temperature distributions and identify inefficiencies in large-scale systems. Zhao et al. [60] proposed a Physics-Informed Convolutional Neural Network (PI-CNN) for real-time prediction of temperature fields in complex thermal systems without labeled data. By integrating heat conduction equations into the loss function and applying finite difference methods, the PI-CNN accurately mapped heat source layouts to steady-state temperature fields. Experimental evaluations demonstrated that the PI-CNN achieved competitive accuracy compared to finite difference methods (FDM), with a mean absolute error (MAE) below 0.03 K. While the framework significantly reduced computational costs, challenges included sensitivity to boundary conditions and difficulty in scaling to irregular geometries or dynamic thermal environments.

3.6. Root cause analysis

AI enables comprehensive fault diagnostics by identifying the underlying causes of performance issues.

- **Causal Analysis:** Algorithms like ML, DL, and hybrid approaches identify causal relationships between sensor anomalies and system faults. Oliveira et al. [61] provided a comprehensive overview of Automatic Root Cause Analysis (ARCA) techniques in manufacturing, highlighting the integration of machine learning and data mining for identifying causal pathways in system faults. The study emphasized the use of hybrid methodologies combining classification models and association rules to link operational data with root causes, achieving significant efficiency improvements in diagnosing complex anomalies. While promising, the approaches face limitations such as reliance on extensive labeled datasets and challenges in adapting to dynamic manufacturing environments, which restrict real-time scalability.
- **Lookup-Based Fault Localization:** Autoencoder-based methods provide interpretable diagnostics by associating reconstruction errors with specific faults. Qian et al. [62] presented a comprehensive review of autoencoder (AE) frameworks for fault detection and diagnosis in industrial processes. The study highlights how encoder–decoder structures and their variants, such as denoising AEs (DAEs) and sparse AEs, effectively identify fault locations by associating reconstruction errors with specific process variables. The findings underscore the capability of AE models in handling nonlinear, multimodal industrial data but also point out limitations such as sensitivity to data quality and challenges in interpreting abstract features in highly complex systems.

3.7. Advanced thermal management in data centers

AI technologies have been instrumental in optimizing cooling strategies for data centers, a critical application domain for industrial cooling systems.

- **Real-Time Monitoring:** DL-based methods analyze temperature distributions to detect thermal hotspots and suggest cooling adjustments. Wang et al. [63] proposed a multi-scale collaborative modeling framework combined with a CNN-BiLSTM-Attention network to predict thermal conditions in air-cooled data centers. The model incorporates boundary conditions derived from Computational Fluid Dynamics simulations and Bayesian optimization for hyperparameter tuning. This integration enabled precise real-time hotspot detection and cooling strategy optimization. However, the reliance on extensive simulation data and high computational overhead may limit scalability to diverse industrial scenarios.
- **Cooling Efficiency Optimization:** PINNs integrate physical laws with operational data to minimize energy consumption while maintaining thermal stability. Wang et al. [50] proposed Phyllis, a physics-informed lifelong reinforcement learning framework for data center cooling control. By embedding thermodynamic constraints into the learning process, Phyllis rapidly adapts to dynamic changes in data center configurations, such as the addition of IT devices or the installation of air containment. However, the reliance on extensive pre-training and the computational demands of lifelong adaptation present scalability challenges in larger and more diverse industrial cooling setups.

3.8. Applications across diverse industrial settings

AI-driven diagnostics and predictive maintenance extend beyond traditional cooling systems into diverse industrial contexts:

- **Renewable Energy Systems:** AI identifies cooling faults in wind turbine generators and photovoltaic panels, ensuring operational efficiency. Polymeropoulos et al. [64] reviewed vision-based monitoring techniques for fault detection in photovoltaic (PV) systems, highlighting advancements in unmanned aerial vehicles (UAVs) and AI-based methodologies, such as CNNs and YOLO frameworks. These approaches improved fault localization accuracy for anomalies like shading, cracks, and dirt accumulation, significantly enhancing PV performance and energy yield. However, challenges such as data overload, high-resolution processing requirements, and limited adaptability to diverse environmental conditions remain barriers to real-time industrial application [65].
- **Electronics Manufacturing:** DL models diagnose overheating and cooling failures in semiconductor fabrication processes. Moosavi et al. [66] proposed an Explainable Artificial Intelligence (XAI)-driven framework for diagnosing faults in induction furnaces used in semiconductor manufacturing. The system combines Deep Neural Networks (DNNs) with Shapley Additive Explanations and Local Interpretable Model-Agnostic Explanations (LIME) to interpret fault predictions based on electrical parameters such as voltage and current harmonics. The proposed model achieved an average F-measure of 0.9187, effectively identifying faults like phase-to-phase shorts and component overheating. However, challenges include reliance on curated datasets and the complexity of real-time XAI integration in dynamic industrial environments.
- **Industrial Refrigeration:** PINNs model dynamic thermal behavior to optimize energy use in refrigeration systems. Hussain et al. [67] developed a physics-informed, data-driven framework for estimating and optimizing two-phase pressure drops in mini- and macro channels for various refrigerants. The framework integrates deep neural networks with genetic algorithms (GAs) to optimize operational parameters such as hydraulic diameter, saturation temperature, and mass flux. However, challenges included sensitivity to irregular geometries and limited scalability to real-time applications across diverse operational conditions.

4. Materials and existing methods

This section provides a comprehensive overview of existing materials and methods relevant to AI-driven diagnostics and predictive maintenance in industrial cooling systems. It critically examines currently available datasets, categorizing them as either real-world or synthetic. Additionally, the section reviews established methods across machine learning, deep learning, and physics-informed neural networks, highlighting their strengths, weaknesses, and suitability for various cooling system scenarios.

Table 3
Dataset availability for AI-driven diagnostics and predictive maintenance in industrial cooling systems.

Dataset name	Reference	Dataset link	Dataset type
Simulation Data for Cooling Water System	[12]	Not available publicly	Synthetic
DAMADICS Benchmark	[14]	https://iair.mchtr.pw.edu.pl/Damadics	Both
ASHRAE RP-1312 Project	[43]	https://tinyurl.com/ASHRAE-RP-1312	Real data
AI4I 2020 Predictive Maintenance Dataset	[68]	https://archive.ics.uci.edu/dataset/601/ai4i+2020+predictive+maintenance+dataset	Synthetic
Infrared Thermographic Image Dataset	[69]	Not available publicly	Real data
Limited Thermal Image Dataset	[70]	Not available publicly	Real data
Infrared Thermal Image Dataset for Induction Motors	[45]	https://data.mendeley.com/datasets/m4sbt8hbvk/3	Real data
Industrial Fault Warning Dataset	[44]	Not available publicly	Real data
Cooling Load Estimation Dataset	[71]	Not available publicly	Real data
Industrial HVAC Energy Consumption Dataset	[72]	Not available publicly	Real data
Chandigarh UT Electricity Utility Data	[73]	Not available publicly	Real data
ITI/CERTH Smart House Dataset	[74]	https://tinyurl.com/ITI-CERTH-Dataset	Real data
Household Energy Consumption Dataset	[75]	https://archive.ics.uci.edu/dataset/235/individual+household+electric+power+consumption	Real data
Hydraulic Test Rig Dataset	[76]	https://archive.ics.uci.edu/dataset/447/condition+monitoring+of+hydraulic+systems	Real data
Real-Time Energy and Temperature Data	[77]	Not available publicly	Real data
VRF Refrigerant Charging Test Dataset	[78]	Not available publicly	Real data
Bus Voltage Dataset (24-Bus Network)	[31]	Not available publicly	Real data
Physics-Based Simulation Data	[60]	Not available publicly	Synthetic
Alibaba CPU Utilization Trace	[50]	Not available publicly	Real data
Temperature Data for Induction Heating Systems	[79]	Not available publicly	Real data
Libya Power System Fault Dataset	[80]	Generated using MATLAB/Simulink	Synthetic
ASHRAE RP-1043 Chiller Fault Dataset	[81]	https://tinyurl.com/ASHRAE-RP-1043	Real data

4.1. Dataset availability

The availability of high-quality datasets is crucial for developing and validating AI-driven diagnostics and predictive maintenance models in industrial cooling systems. Table 3 summarizes publicly available and proprietary datasets used in recent studies. These datasets cover various aspects, including fault detection, predictive maintenance, thermal analysis, and energy efficiency monitoring. Most datasets consist of real-world industrial data, such as ASHRAE RP-1312, Infrared Thermographic Image Dataset, and Industrial Cooling System Dataset, which provide valuable insights for machine learning models. However, several studies rely on synthetic data generated through simulations, such as Physics-Based Simulation Data and Libya Power System Fault Dataset, which may not fully capture real-world operational complexities. Additionally, some datasets remain proprietary, limiting the accessibility and reproducibility of research. Future efforts should focus on developing standardized, open-access datasets with diverse fault scenarios to enhance model generalization and benchmarking. Integrating real-world and physics-informed synthetic datasets can improve the robustness of AI models for predictive maintenance in industrial settings.

4.2. AI methodologies

This section provides a systematic overview of these techniques, categorized into three main groups: traditional machine learning algorithms, advanced deep learning models, and the emerging paradigm of physics-informed neural networks. Each category addresses the complexities of cooling system diagnostics in unique ways — from data-driven inference to hybrid physics-data integration — offering insights into fault detection, real-time monitoring, and system optimization. Fig. 8 presents the taxonomy of AI techniques in diagnostics for industrial cooling systems. This taxonomy categorizes the key AI methodologies into three primary domains: machine learning approaches, deep learning models, and physics-informed neural networks. Each domain is further divided into specific algorithms and frameworks used for fault detection, predictive maintenance, and optimization. The structure follows a hierarchical conceptual taxonomy, meaning it organizes concepts from general to specific in a tree-like format, making it easier to understand how broader AI categories branch into particular techniques. Simultaneously, it functions as a data-driven classification model, where the taxonomy is shaped by the analysis of reviewed literature, reflecting real-world usage patterns and research focus areas. This dual framework helps readers grasp both the theoretical organization of AI methods and their practical deployment in industrial cooling system diagnostics.

4.2.1. Machine learning approaches

Machine learning (ML) techniques have been extensively applied to fault diagnostics, anomaly detection, and predictive maintenance in industrial cooling systems [82]. Among these, Support Vector Machines [83] are widely employed for fault classification due to their ability to handle high-dimensional, nonlinear sensor data, making them particularly useful for identifying refrigerant leakage, compressor inefficiencies, and airflow disruptions in HVAC systems [84,85]. Random Forest (RF) [86] has emerged as a preferred ensemble learning method for anomaly detection and energy efficiency monitoring, excelling in situations where multivariate sensor data — such as temperature, pressure, and refrigerant flow rates — need to be analyzed for predictive diagnostics [76,87]. K-Nearest Neighbors (KNN) [88] has been applied in performance evaluation and fault detection through similarity-based classification, particularly in HVAC systems and energy efficiency studies [89,90]. Gradient Boosting Machines (GBM), including variants like XGBoost and LightGBM, are increasingly used for fault detection and energy load prediction due to their capability to optimize predictive accuracy through iterative ensemble learning [91,92]. Naïve Bayes (NB) [85] has proven useful for fault classification tasks where feature independence assumptions hold, though its effectiveness is limited in scenarios with correlated sensor data [80,93]. Logistic Regression (LR) remains a simple yet interpretable method for fault detection in cooling systems, though its reliance on linear separability constrains its applicability to more complex datasets [85]. While these ML algorithms each demonstrate strengths in different diagnostic tasks, challenges remain in handling noisy sensor data, adapting to dynamic operational conditions, and ensuring real-time efficiency. Future advancements should focus on hybrid approaches that combine ML models with physics-informed constraints and deep learning techniques to enhance diagnostic accuracy and robustness in industrial cooling applications. Table 4 summarizes the strengths, limitations, and applications of machine learning approaches in fault diagnostics for industrial cooling systems.

4.2.2. Deep learning models

Deep learning (DL) models have emerged as powerful tools for fault diagnostics, anomaly detection, and predictive maintenance in

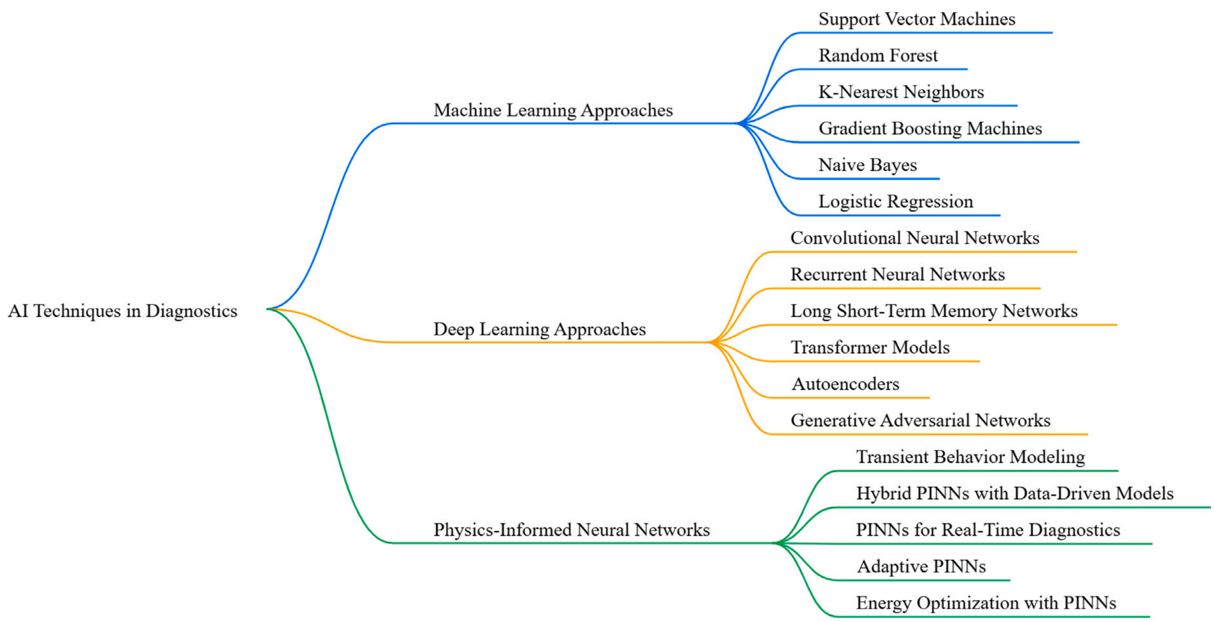


Fig. 8. Taxonomy of AI techniques in diagnostics for industrial cooling systems (hierarchical conceptual taxonomy | data-driven classification model).

Table 4
Comparison of machine learning Approaches for Fault Diagnostics in Industrial Cooling Systems.

Algorithm	Cite	Strengths	Limitations	Applications
SVM	[76,84,85,94–96]	High-dimensional data handling; robust against overfitting.	Computationally expensive; requires careful parameter tuning; struggles with noisy data.	Fault classification in HVAC and centrifugal chillers.
RF	[76,78,87,97,98]	Handles high-dimensional data; robust to overfitting; feature importance ranking.	Computationally intensive with large datasets; less interpretable.	Anomaly detection, refrigerant leakage diagnostics.
KNN	[77,98]	Simple and interpretable; effective for non-Gaussian data; works well with small datasets.	Sensitive to noisy data; computationally expensive for large datasets; requires careful selection of k.	Energy performance evaluation, fault detection in HVAC systems.
GBM	[44,73–75]	High accuracy; handles nonlinear relationships well; feature importance ranking.	Computationally intensive; prone to overfitting with small datasets; requires extensive hyperparameter tuning.	Fault detection, energy load prediction.
NB	[71,72,85,96]	Computationally efficient; effective with high-dimensional data; interpretable.	Assumes feature independence; struggles with noisy or correlated data; sensitive to data distribution.	Fault classification, HVAC energy prediction.
LR	[85,96]	Simple and interpretable; efficient for binary classification.	Assumes linear separability; struggles with multicollinearity; limited to simple relationships.	Fault detection in heat pumps, anomaly detection in HVAC systems.

industrial cooling systems. Their ability to automatically learn hierarchical representations from sensor data, infrared images, and time-series signals allows for enhanced fault detection, system monitoring, and optimization. Convolutional Neural Networks (CNNs) [99] have been widely applied in thermal imaging-based fault detection, identifying surface anomalies such as micro cracks, refrigerant leaks, and cooling inefficiencies in HVAC systems [46,100]. By leveraging feature extraction capabilities, CNNs process high-resolution infrared images to classify operational states and detect emerging faults in industrial cooling networks. Recurrent Neural Networks (RNNs) [101], particularly Gated Recurrent Units (GRUs) and Long Short-Term Memory, excel in predictive maintenance tasks by analyzing multivariate time-series data [102,103]. These models are effective for remaining useful life (RUL) estimation, predictive thermal modeling, and anomaly detection in cooling systems exposed to fluctuating environmental conditions. Transformer-based models [47,104] have recently gained attention for processing large-scale sensor data, offering superior scalability in multivariate time-series analysis and HVAC fault prediction through self-attention mechanisms. Autoencoders (AEs) [105,106] have been successfully utilized for unsupervised anomaly detection, where deviations between reconstructed and actual data points indicate potential system failures. Lastly, Generative Adversarial Networks (GANs) [107] play a key role in addressing class imbalance issues by generating synthetic fault data and improving the robustness of machine learning models for HVAC diagnostics [108,109]. While deep learning models have demonstrated remarkable success in various industrial applications, challenges remain in computational efficiency, real-time deployment, and the interpretability of complex neural network architectures. Future advancements should focus on hybrid AI frameworks that integrate physics-informed deep learning with domain knowledge to enhance diagnostic precision and predictive reliability. Table 5

Table 5
Comparison of deep learning models for fault diagnostics in industrial cooling systems.

Model	Cite	Strengths	Limitations	Applications
CNNs	[45,69,70, 110–112]	Excels in extracting spatial features; effective for image-based data such as thermal images.	Computationally expensive; requires large datasets for training; limited generalization to unseen industrial scenarios.	Fault detection using thermal imaging in HVAC systems and electronics.
RNNs	[68,113]	Captures sequential dependencies; effective for time-series data in dynamic systems.	High computational cost; challenges in scaling to larger datasets and diverse operational settings.	Predictive thermal modeling, RUL estimation for cooling systems.
LSTMs	[113]	Handles long-term dependencies in sequential data; effective for time-series fault prediction.	Computationally intensive; dependency on specific datasets limits generalization; challenges in hyperparameter optimization.	Early anomaly detection, predictive maintenance in heating and cooling.
Transformer	[47,114–116]	Efficient processing of multivariate and long sequential data; scalable to large datasets.	High computational overhead; requires careful hyperparameter tuning; training instability with small datasets.	Fault detection in HVAC systems, anomaly detection in data-scarce setups.
Autoencoders	[105,117, 118]	Effective for anomaly detection through reconstruction error; capable of learning complex data distributions.	High computational demands; dependency on expert domain knowledge for root cause analysis; limited scalability to diverse industrial applications.	Anomaly detection, explained fault localization in cooling systems.
GANs	[119]	Generates synthetic data for addressing data scarcity; robust against class imbalance in fault datasets.	Training instability; computationally intensive; requires extensive hyperparameter tuning and careful balancing of generator-discriminator networks.	Fault detection, synthetic data generation for HVAC systems.

summarizes the strengths, limitations, and applications of various deep learning models for fault diagnostics in industrial cooling systems.

4.2.3. Physics-informed neural networks (PINNs)

While conventional ML and DL methods rely heavily on data patterns, physics-informed neural networks incorporate domain knowledge via partial differential equations (PDEs) and other physical constraints into the training objective. PINNs are particularly valuable for industrial cooling, where thermal processes often exhibit complex transient behaviors governed by fundamental physical laws. PINNs can offer enhanced interpretability, reduced data requirements, and more robust extrapolation to unseen operating conditions by fusing data-driven insights with physics-based priors.

Diagram 9 illustrates the workflow of physics-informed neural networks for fault diagnostics in industrial cooling systems. Observational data points, such as thermal measurements, serve as inputs to a neural network that approximates the system behavior. The physics-guided residual loss ensures compliance with governing PDEs, while boundary and data loss terms enforce boundary conditions and data fidelity, respectively. These components are integrated into a composite loss function, which is optimized using gradient-based techniques and automatic differentiation. The resulting model provides accurate and physically consistent predictions for fault detection and system diagnostics.

This integration of physics with neural networks enables PINNs to handle sparse, noisy datasets effectively while preserving physical consistency, making them ideal for industrial cooling system applications. Jagtap et al. [120] introduced the CoolPINNs framework to model thermal regulation in microvasculatures, addressing challenges such as sharp thermal flux discontinuities and nonlinear radiative heat transfer. This meshless method demonstrated robust real-time monitoring and inverse modeling capabilities, offering significant advantages over traditional finite element methods (FEM). However, CoolPINNs required careful hyperparameter tuning to stabilize training when solving highly nonlinear problems, limiting its scalability. Pan et al. [121] applied PINNs for fault diagnostics in HVAC chillers, focusing on condenser fouling faults. By embedding physical inconsistencies into the loss function, their approach improved interpretability while achieving diagnostic accuracy comparable to purely data-driven methods. Nevertheless, the reliance on pre-defined fault-specific physical models constrained the framework's generalizability to other fault types. Chen et al. [122] developed an Adaptive PINNs (A-PCNNs) framework for thermal modeling in data centers, replacing static coefficients with adaptive ones to enhance flexibility and reduce computational costs. Their results showed a 79.2% reduction in long-term forecast errors compared to traditional PINNs, but the increased computational demand for adaptive coefficient estimation posed challenges for real-time applications. Zobeiry et al. [123] explored PINNs for heat transfer modeling in manufacturing processes. By incorporating convective boundary conditions directly into the loss function, their method achieved real-time thermal response predictions, outperforming FEM in speed. However, the model struggled to generalize across diverse manufacturing scenarios due to limitations in capturing complex transient thermal behaviors.

In addition, PINNs have gained significant attention for solving PDEs across various scientific and engineering domains. The foundational work by Raissi et al. introduced the original PINNs framework, enabling data-driven solutions that respect physical laws [124]. Subsequent advancements like VPINNs adopted a Petrov–Galerkin formulation to improve accuracy and reduce training costs by lowering differential order through integration by parts [125]. To address the limitations of PINNs in handling stiff PDEs, SA-PINNs introduced trainable adaptive weights that focus on regions with higher solution difficulty [126]. For high-dimensional or multi-scale problems, cPINNs and XPINNs employed spatial and space–time domain decomposition,

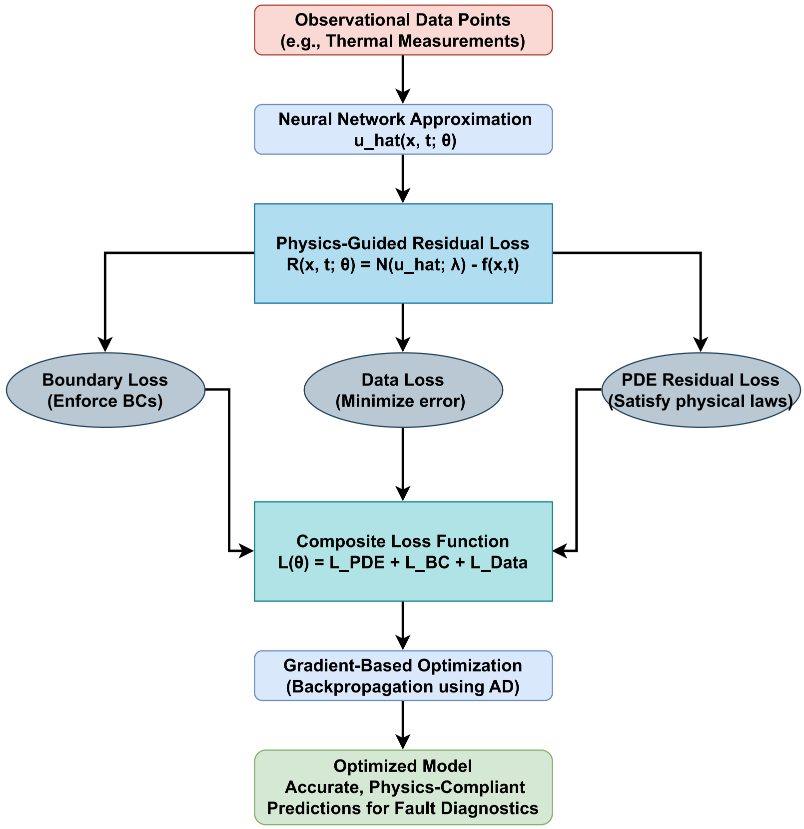


Fig. 9. Illustration of the PINNs framework for solving heat transfer problems, integrating physical laws into the neural network training process.

respectively, offering better scalability and parallelizability [127,128]. Further, APINNs enhanced domain decomposition through a gating network mechanism that learns soft partitions for improved generalization [129]. gPINNs introduced gradient-enhanced loss functions to improve convergence and training efficiency [130], while transfer learning approaches have been applied in PINNs for phase-field modeling of fracture, showing improved efficiency and accuracy in sequential loading scenarios [131]. Operator learning methods like PINO combined coarse data with high-resolution PDE constraints, outperforming standard PINNs in zero-shot super-resolution settings [132]. OL-PINNs hybridized operator learning and PINNs to solve PDEs with sharp solutions using fewer residual points [133]. B-PINNs introduced Bayesian inference for uncertainty quantification and noise robustness in both forward and inverse PDE problems [134]. Finally, multi-fidelity PINNs incorporated low- and high-fidelity models to balance accuracy and computational cost through structured surrogate modeling [135].

To illustrate a practical implementation of PINNs in real-world cooling systems, Liang et al. [136] proposed a comprehensive framework for applying physics-informed neural networks in the optimal control of commercial chiller plants. The study introduces a dual-knowledge embedding strategy by integrating both structure-type and trend-type prior knowledge into the network architecture (S-PINN) and loss function (T-PINN), respectively. The approach was applied to develop energy models for all critical chiller plant components — chillers, cooling towers, and pumps — using actual operational data from a commercial building in Shanghai. The results demonstrated substantial improvements in extrapolation robustness and energy efficiency. In a field deployment, the proposed PINN-based optimal control system achieved a 23.2% increase in energy efficiency compared to traditional fixed-setpoint strategies.

These applications demonstrate PINNs' versatility in solving a range of heat transfer and diagnostic problems in industrial cooling systems. While PINNs offer significant advantages in terms of accuracy, interpretability, and computational efficiency, challenges such as scalability, training stability, and generalizability to diverse conditions remain active areas of research. Lastly, the integration of machine learning, deep learning, and physics-informed neural networks has revolutionized diagnostics and predictive maintenance in industrial cooling systems. Each approach offers unique strengths and limitations tailored to specific application domains. Table 6 compares ML-based, DL-based, and PINN-based models for diagnostics and predictive maintenance in industrial cooling systems.

4.3. Method–application mapping

To strengthen the connection between the technical methods reviewed in Section 4.2 and the industrial use-cases presented in Chapter 3, Fig. 7 illustrates a high-level mapping between different techniques and their primary cooling-related applications. Traditional machine learning models such as Support Vector Machines, Decision Trees, and Random Forests are widely applied in HVAC fault classification tasks. Their strength lies in handling structured sensor data with relatively small datasets while offering model transparency. Deep learning methods show broader applicability across spatial and temporal domains. Convolutional and recurrent networks (e.g., CNNs and LSTMs) are mainly used for detecting data-center hotspots and for processing thermal imaging tasks. Transformer models, known for their ability to capture long-term dependencies, are primarily deployed in multi-energy load forecasting problems where accurate prediction over extended horizons is necessary for energy-efficient cooling strategies.

Unsupervised generative models such as autoencoders and GANs are particularly effective in PV and chiller anomaly detection, where labeled fault data are rare. These models learn compact latent representations of normal operating states and flag significant deviations, making them ideal for fault detection in imbalanced or unlabeled settings. Physics-Informed Neural Networks extend deep learning by embedding governing heat transfer equations into the learning process.

Table 6
Comparative advantages and limitations of ML, DL, and PINN approaches for industrial-cooling diagnostics and predictive-maintenance tasks.

Approach	Key advantages	Key limitations
Traditional Machine-Learning (ML) algorithms (SVM, RF, GBM, k-NN, NB, LR)	• Lightweight; fast training and inference on standard CPUs/edge devices • Higher interpretability (feature importance, decision paths) • Perform well on structured/tabular sensor data with limited samples • Tree ensembles provide built-in variable-ranking for root-cause analysis	• Depend on manual feature engineering; weak on raw image/signal data • Struggle with highly non-linear or high-dimensional relationships • Generally lower accuracy than DL on complex patterns • Often requires separate models per fault class, limiting scalability
Deep-Learning (DL) models (CNN, LSTM, Transformer, AE, GAN)	• Automatically learn hierarchical features from high-dimensional/ unstructured inputs (e.g. IR images, raw time-series) • Achieve state-of-the-art accuracy on complex non-linear tasks (fault detection, RUL prediction, anomaly segmentation) • GANs generate synthetic data to mitigate class imbalance	• Need large labeled datasets; annotation is costly and time-consuming • Substantial memory/compute footprint hinders real-time edge deployment • Often behave as "black boxes", reducing explainability and user trust • Hyper-parameter tuning is labour-intensive
Physics-Informed Neural Networks (PINNs)	• Embed governing physics in the loss; predictions remain *physically consistent* even with sparse/noisy data • Require smaller labeled datasets than purely data-driven models • Better extrapolation outside the training domain (out-of-distribution regimes) • Remove mesh/discretization errors typical of CFD/FEM solvers	• Training can be unstable; balancing data vs. physics losses is non-trivial • High computational overhead (automatic differentiation over PDE residuals) • Limited scalability to large 3-D, transient or strongly coupled multi-physics problems • Depend on domain expertise to specify correct physical constraints

They are suited to data-center thermal control and micro-channel heat sink design problems, where physical constraints and sparse sensor data are prevalent. Lastly, reinforcement learning and hybrid PINNs-RL models are used in real-time cooling optimization, where agents learn control policies that adapt to dynamic environments while satisfying thermal and energy efficiency constraints.

Fig. 10 illustrates the mapping between AI methodologies and their application domains in industrial cooling systems, alongside the frequency of articles associated with each category. Traditional machine learning methods such as SVM, RF, GBM, k-NN, NB, and LR were the most frequently studied, with 8 articles primarily focused on HVAC fault classification. Deep learning approaches using CNNs and LSTMs were discussed in 5 studies, particularly for hotspot detection in data centers and thermal imaging-based diagnostics. Autoencoders and GANs appeared in 6 papers, commonly applied to anomaly detection in photovoltaic systems and chiller units. Transformer models, used for multi-energy load forecasting, were referenced in only 1 article, suggesting this area is still emerging. Physics-informed neural networks (PINNs) and their reinforcement learning hybrids were each mentioned in 2 to 3 articles, indicating growing interest in real-time cooling optimization and microchannel heat sink design. This frequency analysis highlights the current research emphasis on traditional ML and anomaly detection techniques, while also pointing to underexplored opportunities in transformer-based forecasting and hybrid PINNs for dynamic thermal control.

5. Discussion

The integration of AI-driven diagnostics and predictive maintenance in industrial cooling systems has shown considerable progress, enhancing fault detection, energy efficiency, and operational reliability. However, despite these advancements, several challenges remain, limiting broader adoption and real-world implementation. This section critically analyzes the key findings of this review by addressing the research questions, discussing the major limitations identified in existing AI applications, and outlining future research directions. Instead of reiterating the results, this discussion examines why these findings are significant, how they compare across different studies, and what gaps remain to be addressed.

5.1. Application-specific analysis

The effectiveness of AI-driven diagnostics and predictive maintenance methods depends heavily on the application context of industrial cooling systems. These systems vary in scale, function, and operating conditions across different industrial sectors. As a result, the selection and performance of AI algorithms must align with application-specific requirements such as reliability, energy efficiency, data availability, and system complexity.

In commercial buildings and HVAC systems, energy efficiency and occupant comfort are primary concerns. These systems typically operate under relatively stable conditions and produce structured sensor data. Traditional ML algorithms such as SVM and Random Forests are well-suited for these environments. They offer interpretable results and are efficient to deploy, especially in scenarios where computational resources are limited. Data centers, on the other hand, require high operational reliability and rapid fault detection to avoid service disruptions. Deep learning models, particularly CNNs, have shown strong performance in analyzing thermal imaging data for identifying hot spots and hardware faults. Recurrent models such as LSTM networks are used for temperature trend forecasting and predictive maintenance based on time-series data. In addition, Transformer-based models offer scalable solutions for handling high-dimensional sensor inputs in real-time monitoring environments.

In manufacturing and heavy industry, cooling systems often operate under dynamic and harsh conditions with fluctuating thermal loads. These settings pose challenges for data collection and model generalization. Physics-Informed Neural Networks are useful in such cases, as they integrate physical laws into the learning process. PINNs can improve accuracy and reduce dependency on large datasets. However, their practical implementation may require careful tuning and substantial computational resources. Cryogenic systems and high-performance industrial

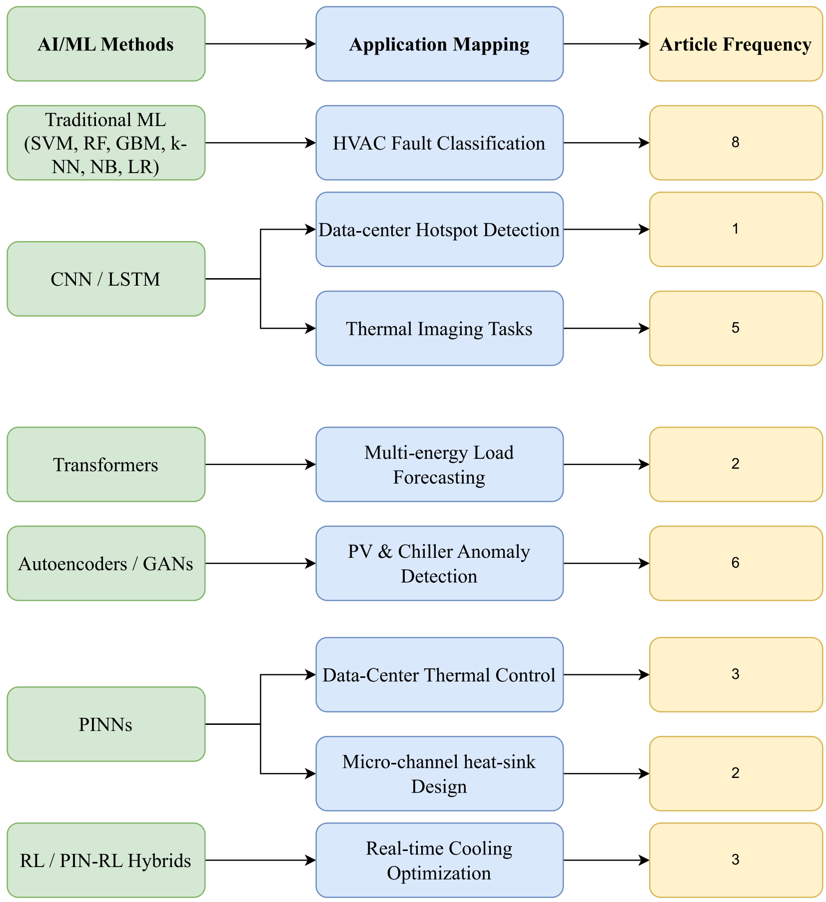


Fig. 10. Mapping of AI/ML methods to industrial cooling applications.

cooling applications often involve operating regimes that are underrepresented in available datasets. Generative models, such as GANs, can be employed to generate synthetic fault scenarios and enhancing the training process of diagnostic models. Reinforcement Learning methods are also relevant in adaptive cooling control, where the system dynamically adjusts cooling parameters in response to changing operational demands.

Across all application domains, model explainability is a key factor influencing adoption. Black-box models are often met with skepticism in industrial settings, particularly where compliance or operator trust is required. Explainable AI techniques, including feature attribution methods and attention mechanisms, can improve transparency and facilitate acceptance by maintenance personnel and decision-makers.

5.2. Algorithm-specific analysis

The comparative analysis of AI methodologies reveals distinct strengths and limitations when applied to industrial cooling systems. Deep learning methods, particularly Convolutional Neural Networks, excel in analyzing thermal imaging data. CNNs effectively extract spatial patterns from high-resolution infrared images, providing robust anomaly detection capabilities such as identifying overheating or refrigerant leakage. However, their performance heavily depends on extensive labeled datasets and computational resources, limiting their real-time deployment. In contrast, Long Short-Term Memory networks demonstrate exceptional performance in handling sequential time-series data, essential for predictive maintenance tasks such as forecasting remaining useful life (RUL) and detecting subtle patterns indicating impending failures. Their ability to retain long-term temporal dependencies allows accurate fault predictions even in dynamic operational conditions. Yet, LSTMs face scalability challenges, particularly when modeling multivariate sensor data streams simultaneously, where Transformer-based models may offer superior performance through their self-attention mechanisms and ability to scale efficiently with increasing data complexity.

Physics-informed neural networks play a crucial role by integrating governing physical laws directly into their architecture. This integration allows PINNs to leverage limited data more effectively, enhancing the accuracy and interpretability of thermal diagnostics. PINNs effectively handle sparse datasets and reduce the need for extensive labeled data, making them highly valuable in industrial scenarios where collecting comprehensive, high-quality data is challenging. Nonetheless, PINNs' computational complexity, sensitivity to the precise formulation of physical constraints, and potential instability during training remain critical challenges that necessitate careful tuning and methodological improvements.

Despite the dominance of sophisticated deep learning and physics-informed models, traditional machine learning algorithms, such as Support Vector Machines and Random Forests, remain relevant due to their interpretability, simplicity, and lower computational demands. ML models provide transparent decision-making paths, making them more trusted by practitioners, especially in resource-constrained industrial environments. Their lower computational requirements allow deployment in real-time scenarios, offering practical benefits where immediate diagnostics are essential. Although their predictive performance may not always match deep learning models, their straightforward implementation, ease of deployment, and ability to handle structured data effectively justify their continued use.

Generative models, including Generative Adversarial Networks, further extend the capabilities of AI-driven diagnostics by addressing data scarcity through synthetic data generation. GANs are instrumental in mitigating dataset imbalance by creating realistic fault scenarios,

Table 7
Comparison of data-driven and physics-based modeling approaches.

Aspect	Data-driven modeling	Physics-based modeling
Core Principle	Learns patterns directly from empirical data using statistical or machine learning methods	Uses governing physical laws (e.g., PDEs) derived from first principles
Data Requirements	Requires large volumes of high-quality labeled data	Can work with limited data if physical laws are well-defined
Interpretability	Often a black-box; difficult to interpret	Physically interpretable and consistent with domain knowledge
Generalization	May struggle outside training distribution; prone to overfitting	Generalizes better in unseen scenarios if physics is applicable
Robustness to Noise	Sensitive to noise and outliers in data	Robust to noise if equations are well-posed
Model Flexibility	Can approximate unknown or complex phenomena	Limited by the accuracy and completeness of physical equations
Computational Complexity	High for deep learning models; requires extensive training data and tuning	Can be computationally expensive for PDE solvers, especially in complex domains
Challenges in Integration (PINNs)	Conflicts with physical laws when data is noisy or inconsistent	Difficulty in incorporating real-world imperfections or data-driven residuals
References	[15,26,84,122,137,138]	[14,15,39,49,120–122]

enhancing the training robustness of other AI models. However, generative models carry inherent risks, such as mode collapse and generation of unrealistic samples, potentially introducing biases and inaccuracies. Addressing these limitations through rigorous training, validation, and hybrid modeling strategies is crucial for their successful application in industrial cooling diagnostics.

In recent years, efforts to bridge data-driven methods with physics-based modeling have gained traction, notably through frameworks such as PINNs. One key difficulty is the tuning of loss function components. In PINNs, the total loss typically includes terms for both data fidelity and physical consistency. Determining appropriate weights for these components is non-trivial and often domain-specific. Overweighting the physical loss may reduce sensitivity to data patterns, especially in noisy industrial environments, while underweighting it can violate key physical laws. Another issue is the potential conflict between empirical data and governing equations. Noisy sensor data may contradict idealized PDEs, leading to optimization instability or biased learning. Moreover, expressing complex domain-specific phenomena — like turbulent cooling behavior or coupled heat-mass transfer — as mathematical constraints may not be feasible, limiting the practical applicability of PINNs. These challenges underline the need for adaptive weighting strategies, robust hybrid architectures, and better methods to encode domain knowledge without rigid formulations. Table 7 provides a comparative overview of data-driven and physics-based modeling approaches, highlighting their respective strengths, limitations, and integration challenges.

Overall, selecting an optimal AI model for industrial cooling diagnostics requires careful consideration of data availability, computational resources, interpretability, and domain-specific constraints. Balancing data-driven methodologies with physics-based insights provides the best avenue for achieving reliable, efficient, and interpretable AI-driven predictive maintenance solutions in industrial cooling systems.

5.3. Data challenges and implications

Data availability plays a crucial role in developing AI-driven diagnostics and predictive maintenance for industrial cooling systems. This study reveals a notable reliance on synthetic data compared to real-world datasets. The primary reason for this trend is the practical difficulties industries face in collecting and sharing real data. Industrial data collection encounters numerous barriers, including privacy concerns, cybersecurity threats, and operational confidentiality. The cost of data acquisition and maintenance further discourages companies from openly distributing their data, driving researchers toward synthetic data generation methods such as physics-based simulations and Generative Adversarial Networks. However, synthetic data introduces potential biases, as simulated scenarios may not accurately reflect the complex, nonlinear, and noisy conditions found in real-world cooling systems. Also, real-time data collection in industrial cooling systems poses distinct challenges, including sensor failures, latency issues, environmental disturbances, and integration complexities. Sensor failures and environmental noise create gaps and inaccuracies in continuous monitoring, significantly impacting the quality and reliability of data used for diagnostics. Industrial environments frequently involve harsh conditions and complex operational dynamics, complicating consistent sensor performance and data reliability. Addressing these challenges requires robust sensor networks, advanced sensor fusion techniques, and resilience to operational disturbances.

An essential consideration is whether limited data can effectively support AI-driven diagnostics. Techniques such as transfer learning, few-shot learning, and synthetic data augmentation demonstrate significant promise in compensating for limited real data. Transfer learning enables models to leverage knowledge acquired from related domains, improving performance with fewer data points. Few-shot learning methodologies further facilitate model development by requiring minimal labeled data and rapidly adapting AI models to new fault scenarios. Synthetic data augmentation, especially through GAN-based approaches, can artificially expand datasets, addressing class imbalance and providing representative examples of rare failure conditions. Nevertheless, reliance on synthetic data demands caution due to potential biases and discrepancies from real-world conditions. The debate between data quality and quantity is also critical. High-quality data, characterized by accuracy, consistency, and relevance, typically leads to superior diagnostic performance compared to large but noisy datasets. High-quality datasets, even if smaller, contribute to better model accuracy and lower false detection rates, highlighting the need for precise data collection and rigorous preprocessing techniques. In contrast, large, noisy datasets can obscure critical fault indicators, reduce the reliability of predictions, and demand increased computational resources.

The generalizability of existing benchmark datasets, such as the ASHRAE RP-1312 dataset, to diverse industrial environments is limited. Many publicly available datasets reflect highly controlled conditions, potentially unrepresentative of dynamic industrial scenarios characterized by fluctuating loads and varying cooling configurations. The limited applicability of these datasets underscores the necessity for standardized datasets that capture diverse fault scenarios across different operational settings, thereby improving model generalization. Lastly, concerns regarding data privacy and security significantly influence the industry's willingness to share cooling system operational data. The interconnected nature of modern cooling systems, often integrated with IoT networks, exposes them to cybersecurity vulnerabilities. Companies remain cautious about data sharing, fearing potential exploitation or unauthorized access. Developing secure, anonymized data-sharing platforms, coupled with robust cybersecurity frameworks, is essential to facilitate data exchange, enhance collaboration, and accelerate the development of effective AI-driven diagnostic solutions.

5.4. Practical deployment and performance

A critical consideration in deploying AI-driven diagnostic systems is the inherent trade-off between accuracy and computational cost. High-performing deep learning models, including Transformer architectures and deep CNNs, achieve exceptional accuracy but demand substantial computational resources. This requirement poses challenges in real-time industrial environments, where processing power, memory constraints, and latency restrictions significantly impact the feasibility of deployment. Although pruning, quantization, and model compression techniques can reduce computational burdens, these methods often compromise predictive accuracy, highlighting the need for careful consideration of performance trade-offs. Furthermore, deploying AI models effectively in real-time scenarios involves a balance between diagnostic accuracy and inference speed. Lightweight models, optimized through edge computing techniques, provide practical solutions by enabling rapid, resource-efficient processing directly at the industrial endpoint. However, the reduction in model complexity can result in lower accuracy, making it imperative to carefully evaluate the acceptable trade-offs between precision and operational efficiency based on application-specific requirements.

Explainability remains another critical factor influencing the adoption of AI-driven diagnostics in industrial cooling systems. The black-box nature of complex neural networks raises concerns among industry practitioners regarding transparency and accountability. Integration of Explainable AI techniques, such as attention mechanisms, SHAP values, and Local Interpretable Model-Agnostic Explanations, can significantly enhance model transparency, trustworthiness, and user acceptance. Transparent models not only improve stakeholder confidence but also enable technicians to interpret diagnostic decisions clearly, facilitating more informed maintenance actions. Scalability across different industrial settings represents another substantial challenge. Models trained on specific HVAC datasets may struggle to generalize effectively to different domains, such as data center cooling or industrial refrigeration, due to variations in operating conditions, cooling system configurations, and fault characteristics. Therefore, ensuring broader model generalization demands hybrid approaches combining data-driven and physics-informed methods, transfer learning strategies, and standardized cross-domain datasets. Models designed with modular or adaptive structures could help maintain accuracy while facilitating adaptation to diverse cooling environments.

Integration of AI-driven diagnostic solutions into existing legacy systems is particularly challenging, primarily due to compatibility issues, communication protocol mismatches, and outdated system architectures. Legacy systems often lack the necessary sensor infrastructure and computing capabilities required by advanced AI algorithms. Overcoming these integration barriers demands significant infrastructure updates, middleware solutions that enable seamless communication between legacy and modern AI systems, and incremental deployment strategies. Future research should explore interoperability frameworks and lightweight AI approaches specifically tailored for integration with existing industrial cooling infrastructures.

In addition, Table 8 provides a comparative summary of how different categories of AI methods — traditional machine learning, deep learning, and physics-informed neural networks — address key challenges in AI-driven diagnostics for industrial cooling systems. Traditional ML algorithms exhibit high computational efficiency, interpretability, and suitability for real-time deployment, making them ideal for structured sensor data in resource-constrained settings. However, they moderately address data scarcity and scalability. Deep learning techniques offer superior modeling capabilities for complex and unstructured data such as thermal images but face limitations in data requirements, computational cost, and real-time deployment. PINNs effectively handle sparse and noisy data while maintaining physical consistency, making them well-suited for scenarios with limited labeled data. Despite their interpretability and integration potential, PINNs currently struggle with scalability and training stability. This comparative analysis helps align methodological choices with domain-specific requirements in industrial cooling diagnostics.

5.5. Future research directions and recommendations

Based on the insights gained from this study, several important directions for future research can further enhance AI-driven diagnostics and predictive maintenance in industrial cooling systems:

- Developing standardized, publicly accessible benchmark datasets covering various fault scenarios, operational conditions, and cooling configurations to facilitate comparative analyses and model generalization.
- Analyzing and comparing different variants of physics-informed neural networks in the context of industrial cooling systems. This includes evaluating their effectiveness in modeling, fault detection, and optimization tasks.
- Investigate the integration of GAN-based loss functions within PINNs architectures to improve learning stability and enhance solution accuracy under limited or imbalanced data scenarios.
- Integrating Explainable AI methodologies, including hybrid deep learning approaches that combine physics-informed constraints and attention mechanisms, to enhance model interpretability and trustworthiness.
- Investigating lightweight AI architectures and deploying model compression techniques such as pruning, quantization, and knowledge distillation to enable efficient real-time deployment in resource-constrained industrial environments.
- Exploring reinforcement learning methods for adaptive cooling optimization, enabling AI models to dynamically adjust cooling parameters in response to real-time operational and environmental changes.
- Promoting self-adaptive AI systems that employ reinforcement learning and continual learning paradigms, allowing real-time optimization and adaptation to evolving cooling system conditions.
- Advancing hybrid AI approaches that combine physics-based models with data-driven methods, enabling efficient diagnostics even with limited real-world data, thus improving robustness and reducing training data requirements.
- Expanding the use of generative models such as GANs and variational autoencoders (VAEs) to generate realistic synthetic datasets for fault scenarios, addressing class imbalance and data scarcity issues.
- Conducting detailed studies on edge computing implementations to facilitate real-time analytics, reduce latency, and minimize dependence on centralized computational resources.
- Exploring federated learning strategies to overcome data privacy barriers, enabling collaborative AI model training without compromising sensitive industrial data.

Table 8
Mapping of AI Methods to Key Challenges in Industrial Cooling Diagnostics. (High = Effectively addresses the challenge; Medium = Partially addresses it; Low = Struggles to address it).

AI technique	Data scarcity	Computational efficiency	Real-time deployment	Legacy integration	Interpretability	Scalability
Traditional ML	Medium	High	High	Medium	High	Medium
Deep Learning	Low	Low	Low	Low	Low to Medium	Medium to High
PINNs	High	Medium	Medium	Medium	High	Low

- Examining the role of multi-modal data fusion techniques in integrating diverse data sources, including thermal images, sensor readings, and operational data, to enhance diagnostic accuracy and robustness.
- Optimizing energy efficiency and sustainability through AI-driven thermal management, utilizing AI techniques to dynamically control cooling parameters, minimize energy consumption, and reduce carbon footprints.

Overall, this study underscores the significant potential of AI-driven diagnostics and predictive maintenance for transforming industrial cooling systems. Continued research along these paths promises to substantially improve reliability, efficiency, and sustainability in industrial cooling, thereby aligning with broader goals of operational excellence and environmental responsibility.

6. Conclusions

This paper presents a comprehensive review of AI-driven diagnostics and predictive maintenance techniques for industrial cooling systems, focusing on the integration of machine learning, deep learning, and physics-informed neural networks. The study explores the applications of these methodologies across various domains, including fault detection, energy optimization, thermal imaging, and real-time monitoring. By synthesizing advancements and analyzing key challenges such as data scarcity, scalability, model interpretability, and cybersecurity, this review highlights the transformative potential of AI technologies in improving system reliability, energy efficiency, and operational sustainability. Despite significant progress, several gaps remain in the development and deployment of AI solutions for industrial cooling systems. Challenges such as the availability of labeled data, computational demands of advanced models, and the need for real-time scalability hinder widespread adoption. Furthermore, ensuring the integration of domain-specific knowledge with data-driven approaches remains a critical area for future research. To address these challenges, this paper outlines actionable research directions, including the development of standardized datasets, explainable AI frameworks, scalable real-time algorithms, and secure AI deployment protocols. By addressing these issues, future advancements can enable the design of robust, interpretable, and sustainable cooling systems.

CRediT authorship contribution statement

Md Mohsin Kabir: Writing – original draft, Validation, Software, Resources, Methodology, Formal analysis, Data curation, Conceptualization. **Shahina Begum:** Writing – review & editing, Validation, Supervision, Project administration, Funding acquisition, Formal analysis. **Shaibal Barua:** Writing – review & editing, Supervision, Investigation. **Mobyen Uddin Ahmed:** Writing – review & editing, Supervision, Investigation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to extend their sincere gratitude to the partners of the HeatTrack: Enhanced Reliability, Monitoring, and Diagnostics of Complex Cooling Systems through Advanced Thermal Management Grant Number: 559441–9839 project at the RESILIENT Centre, Mälardalen University, Sweden, funded by the Swedish Energy Agency and Hitachi Energy for their support and contributions in this research. We would also like to thank CPMXai: Cognitive Predictive Maintenance and Quality Assurance using Explainable AI and Machine Learning (Grant Number :2021-03679) for their support.

Data availability

No data was used for the research described in the article.

References

[1] Chen H, Ruan X-h, Peng Y-h, Wang Y-l, Yu C-k. Application status and prospect of spray cooling in electronics and energy conversion industries. Sustain Energy Technol Assess 2022;52:102181. http://dx.doi.org/10.1016/j.seta.2022.102181.

[2] Szabó L. Additive manufacturing of cooling systems used in power electronics. a brief survey. In: 2022 29th international workshop on electric drives: advances in power electronics for electric drives. IEEE; 2022, p. 1–8. http://dx.doi.org/10.1109/IWED54598.2022.9722580.

[3] Zhang Q, Meng Z, Hong X, Zhan Y, Liu J, Dong J, et al. A survey on data center cooling systems: Technology, power consumption modeling and control strategy optimization. J Syst Archit 2021;119:102253. http://dx.doi.org/10.1016/j.sysarc.2021.102253.

[4] Koronen C, Åhman M, Nilsson LJ. Data centres in future European energy systems—energy efficiency, integration and policy. Energy Effic 2020;13(1):129–44. http://dx.doi.org/10.1007/s12053-019-09833-8.

[5] Jones D, Eckert C. Hidden overdesign in building services: insights from two UK hospital case studies. J Eng Des 2023;34(7):437–61.

[6] Lei M, Jiang G, Zhao L, Wang J, Li BQ, Xia P, et al. Thermal error controlling for the spindle in a precision boring machine with external cooling across coated joints. Proc Inst Mech Eng Part C: J Mech Eng Sci 2020;234(2):658–75. http://dx.doi.org/10.1177/0954406219879720.

[7] Chen JM. Carbon neutrality: Toward a sustainable future. Innov 2021;2(3). https://www.cell.com/the-innovation/fulltext/S2666-6758(21)00052-7.

[8] Rahman SI, Moghassemi A, Arsalan A, Timilsina L, Chamarthi PK, Papari B, et al. Emerging trends and challenges in thermal management of power electronic converters: A state of the art review. IEEE Access 2024. http://dx.doi.org/10.1109/ACCESS.2024.3385429.

[9] Ucar A, Karakose M, Kırımça N. Artificial intelligence for predictive maintenance applications: key components, trustworthiness, and future trends. Appl Sci 2024;14(2):898. http://dx.doi.org/10.3390/app14020898.

[10] Çınar ZM, Abdussalam Nuhu A, Zeeshan Q, Korhan O, Asmael M, Safaei B. Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0. Sustainability 2020;12(19):8211. http://dx.doi.org/10.3390/su12198211.

[11] Rafati A, Shaker HR. Predictive maintenance of district heating networks: A comprehensive review of methods and challenges. Therm Sci Eng Prog 2024;102722. http://dx.doi.org/10.1016/j.tsep.2024.102722.

[12] Xu J, Li H, Zhang Q. Adaptive control for circulating cooling water system using deep reinforcement learning. PLoS One 2024;19(7):e0307767. http://dx.doi.org/10.1371/journal.pone.0307767.

[13] Fassi Y, Heiries V, Boutet J, Boisseau S. Towards physics-informed machine learning-based predictive maintenance for power converters–A review. IEEE Trans Power Electron 2023. http://dx.doi.org/10.1109/TPEL.2023.3328438.

[14] Lai C, Ahmed I, Zio E, Li W, Zhang Y, Yao W, et al. A multistage physics-informed neural network for fault detection in regulating valves of nuclear power plants. Energies 2024;17(11):2647. http://dx.doi.org/10.3390/en17112647.

[15] Li H, Zhang Z, Li T, Si X. A review on physics-informed data-driven remaining useful life prediction: Challenges and opportunities. Mech Syst Signal Process 2024;209:111120. http://dx.doi.org/10.1016/j.ymssp.2024.111120.

[16] Sinha S, Lee YM. Challenges with developing and deploying AI models and applications in industrial systems. Discov Artif Intell 2024;4(1):55. http://dx.doi.org/10.1007/s44163-024-00151-2.

[17] Leukel J, González J, Riekert M. Adoption of machine learning technology for failure prediction in industrial maintenance: A systematic review. J Manuf Syst 2021;61:87–96. http://dx.doi.org/10.1016/j.jmsy.2021.08.012.

[18] Sahoo AK, Tripathy DP. Applications of AI and machine learning in mining: digitization and future directions. Saf Extrem Environ 2025;7(1):4. http://dx.doi.org/10.1007/s42797-025-00118-1.

[19] Ogidan ET, Olawale OP, Dimililer K. Machine learning applications in industry 4.0: opportunities and challenges. Handb Intell Sustain Manuf 2025;284–304, https://shorturl.at/dj7ht.

[20] Andrianandrianina Johanesa TV, Equeter L, Mahmoudi SA. Survey on AI applications for product quality control and predictive maintenance in industry 4.0. Electronics 2024;13(5):976. http://dx.doi.org/10.3390/electronics13050976.

[21] Bunian S, Al-Ebrahim MA, Nour AA. Role and applications of artificial intelligence and machine learning in manufacturing engineering: A review. Eng Sci 2024;29:1088. http://dx.doi.org/10.30919/es1088.

[22] Sahu AR, Palei SK, Mishra A. Data-driven fault diagnosis approaches for industrial equipment: A review. Expert Syst 2024;41(2):e13360. http://dx.doi.org/10.1111/exsy.13360.

[23] Kamm S, Veekati SS, Müller T, Jazdi N, Weyrich M. A survey on machine learning based analysis of heterogeneous data in industrial automation. Comput Ind 2023;149:103930. http://dx.doi.org/10.1016/j.compind.2023.103930.

[24] Es-Sakali N, Cherkaoui M, Mghazli MO, Naimi Z. Review of predictive maintenance algorithms applied to HVAC systems. Energy Rep 2022;8:1003–12. http://dx.doi.org/10.1016/j.egyr.2022.07.130.

[25] Mckoy DR, Tesiero RC, Acquaah YT, Gokaraju B. Review of HVAC systems history and future applications. Energies 2023;16(17):6109. http://dx.doi.org/10.3390/en16176109.

[26] Chen Z, O'Neill Z, Wen J, Pradhan O, Yang T, Lu X, et al. A review of data-driven fault detection and diagnostics for building HVAC systems. Appl Energy 2023;339:121030.

[27] Devasenan M, Madhavan S. Thermal intelligence: exploring AI's role in optimizing thermal systems–a review. Interactions 2024;245(1):282. http://dx.doi.org/10.1007/s10751-024-02122-6.

[28] Alamu O, Olwal TO, Migabo EM. Machine learning applications in energy harvesting internet of things networks: A review. IEEE Access 2025. http://dx.doi.org/10.1109/ACCESS.2024.3525263.

[29] Villano F, Mauro GM, Pedace A. A review on machine/deep learning techniques applied to building energy simulation, optimization and management. Thermo 2024;4(1):100–39. http://dx.doi.org/10.3390/thermo4010008.

[30] Zhou X, Du H, Xue S, Ma Z. Recent advances in data mining and machine learning for enhanced building energy management. Energy 2024;132636. http://dx.doi.org/10.1016/j.energy.2024.132636.

[31] Kotsiopoulos T, Sarigiannidis P, Ioannidis D, Tzovaras D. Machine learning and deep learning in smart manufacturing: The smart grid paradigm. Comput Sci Rev 2021;40:100341. http://dx.doi.org/10.1016/j.cosrev.2020.100341.

[32] Forootan MM, Larki I, Zahedi R, Ahmadi A. Machine learning and deep learning in energy systems: A review. Sustainability 2022;14(8):4832. http://dx.doi.org/10.3390/su14084832.

[33] Ntakolia C, Anagnostis A, Moustakidis S, Karcanias N. Machine learning applied on the district heating and cooling sector: A review. Energy Syst 2022;13(1):1–30. http://dx.doi.org/10.1007/s12667-020-00405-9.

[34] Khalil RA, Saeed N, Masood M, Fard YM, Alouini M-S, Al-Naffouri TY. Deep learning in the industrial internet of things: Potentials, challenges, and emerging applications. IEEE Internet Things J 2021;8(14):11016–40. http://dx.doi.org/10.1109/JIOT.2021.3051414.

[35] Fernandes M, Corchado JM, Marreiros G. Machine learning techniques applied to mechanical fault diagnosis and fault prognosis in the context of real industrial manufacturing use-cases: a systematic literature review. Appl Intell 2022;52(12):14246–80. http://dx.doi.org/10.1007/s10489-022-03344-3.

[36] Rodriguez-Fernandez V, Camacho D. Recent trends and advances in machine learning challenges and applications for industry 4.0. Expert Syst 2024;41(2):e13506. http://dx.doi.org/10.1111/exsy.13506.

[37] Ullah H, Uzair M, Jan Z, Ullah M. Integrating industry 4.0 technologies in defense manufacturing: Challenges, solutions, and potential opportunities. Array 2024;100358.

[38] Zambrano V, Mueller-Roemer J, Sandberg M, Talasila P, Zanin D, Larsen PG, et al. Industrial digitalization in the industry 4.0 era: Classification, reuse and authoring of digital models on digital twin platforms. Array 2022;14:100176.

[39] Wu Y, Sicard B, Gadsden SA. Physics-informed machine learning: a comprehensive review on applications in anomaly detection and condition monitoring. Expert Syst Appl 2024;124678. http://dx.doi.org/10.1016/j.eswa.2024.124678.

[40] Lee C-Y, Maceren EDC. Physics-informed anomaly and fault detection for wind energy systems using deep CNN and adaptive elite PSO-xgboost. IET Gener Transm Distrib 2025;19(1):e13289.

[41] Yan R, Zhou Z, Shang Z, Wang Z, Hu C, Li Y, et al. Knowledge driven machine learning towards interpretable intelligent prognostics and health management: Review and case study. Chin J Mech Eng 2025;38(1):5. http://dx.doi.org/10.1186/s10033-024-01173-8.

[42] Vadyala SR, Betgeri SN, Betgeri NP. Physics-informed neural network method for solving one-dimensional advection equation using PyTorch. Array 2022;13:100110.

[43] Tun W, Wong JK-W, Ling S-H. Hybrid random forest and support vector machine modeling for HVAC fault detection and diagnosis. Sensors 2021;21(24):8163.

[44] Li S, Jin N, Dogani A, Yang Y, Zhang M, Gu X. Enhancing LightGBM for industrial fault warning: An innovative hybrid algorithm. Processes 2024;12(1):221.

[45] Calderon-Uribe U, Lizarraga-Morales RA, Guryev IV. Fault diagnosis in induction motors through infrared thermal images using convolutional neural network feature extraction. Machines 2024;12(8):497.

[46] Wiysobunri BN, Erden HS, Toreyin BU. CNN-based server state monitoring and fault diagnosis using infrared thermal images. Soft Comput 2024;1–22.

[47] Abdollah MAF, Scoccia R, Aprile M. Transformer encoder based self-supervised learning for HVAC fault detection with unlabeled data. Build Environ 2024;258:111568.

[48] Ren Z, Zhu Y, Liu Z, Feng K. Few-shot GAN: Improving the performance of intelligent fault diagnosis in severe data imbalance. IEEE Trans Instrum Meas 2023;72:1–14.

[49] Zhang X, Tu C, Yan Y. Physics-informed neural network simulation of conjugate heat transfer in manifold microchannel heat sinks for high-power IGBT cooling. Int Commun Heat Mass Transfer 2024;159:108036.

[50] Wang R, Cao Z, Zhou X, Wen Y, Tan R. Phyllis: Physics-informed lifelong reinforcement learning for data center cooling control. In: Proceedings of the 14th ACM international conference on future energy systems. 2023, p. 114–26.

[51] Sworna Kokila M, Bibin Christopher V, Ramya G. Enhanced power system fault detection using quantum-AI and herd immunity quantum-AI fault detection with herd immunity optimisation in power systems. IET Quantum Commun 2024;5(4):340–8.

[52] Liu Z, Fang L, Jiang D, Qu R. A machine-learning-based fault diagnosis method with adaptive secondary sampling for multiphase drive systems. IEEE Trans Power Electron 2022;37(8):8767–72.

[53] Javed MR, Shabbir Z, Asghar F, Amjad W, Mahmood F, Khan MO, et al. An efficient fault detection method for induction motors using thermal imaging and machine vision. Sustainability 2022;14(15):9060. http://dx.doi.org/10.3390/su14159060.

[54] Shen C, Zhang H, Meng S, Li C. Augmented data driven self-attention deep learning method for imbalanced fault diagnosis of the HVAC chiller. Eng Appl Artif Intell 2023;117:105540. http://dx.doi.org/10.1016/j.engappai.2022.105540.

[55] Nunes P, Santos J, Rocha E. Challenges in predictive maintenance–a review. CIRP J Manuf Sci Technol 2023;40:53–67. http://dx.doi.org/10.1016/j.cirpj.2022.11.004.

[56] Jeon SH, Yoo S, Yoo Y-S, Lee I-W. ML-and LSTM-based radiator predictive maintenance for energy saving in compressed air systems. Energies 2024;17(6):1428. http://dx.doi.org/10.3390/en17061428.

[57] Sinap V, Kumtepe A. CNN-based automatic detection of photovoltaic solar module anomalies in infrared images: a comparative study. Neural Comput Appl 2024;36(28):17715–36. http://dx.doi.org/10.1007/s00521-024-10322-y.

[58] Fan J, Zhuang W, Xia M, Fang W, Liu J. Optimizing attention in a transformer for multihorizon, multienergy load forecasting in integrated energy systems. IEEE Trans Ind Inform 2024. http://dx.doi.org/10.1109/TII.2024.3392278.

[59] Xiao T, You F. Building thermal modeling and model predictive control with physically consistent deep learning for decarbonization and energy optimization. Appl Energy 2023;342:121165. http://dx.doi.org/10.1016/j.apenergy.2023.121165.

[60] Zhao X, Gong Z, Zhang Y, Yao W, Chen X. Physics-informed convolutional neural networks for temperature field prediction of heat source layout without labeled data. Eng Appl Artif Intell 2023;117:105516. http://dx.doi.org/10.1016/j.engappai.2022.105516.

[61] e Oliveira E, Miguéis VL, Borges JL. Automatic root cause analysis in manufacturing: an overview & conceptualization. J Intell Manuf 2023;34(5):2061–78. http://dx.doi.org/10.1007/s10845-022-01914-3.

[62] Qian J, Song Z, Yao Y, Zhu Z, Zhang X. A review on autoencoder based representation learning for fault detection and diagnosis in industrial processes. Chemometr Intell Lab Syst 2022;231:104711. http://dx.doi.org/10.1016/j.chemolab.2022.104711.

[63] Wang N, Guo Y, Huang C, Tian B, Shao S. Multi-scale collaborative modeling and deep learning-based thermal prediction for air-cooled data centers: An innovative insight for thermal management. Appl Energy 2025;377:124568. http://dx.doi.org/10.1016/j.apenergy.2024.124568.

[64] Polymeropoulos I, Bezyrgiannidis S, Vrochidou E, Papakostas GA. Enhancing solar plant efficiency: A review of vision-based monitoring and fault detection techniques. Technologies 2024;12(10):175. http://dx.doi.org/10.3390/technologies12100175.

[65] Katkuri AVR, Madan H, Khatri N, Abdul-Qawy ASH, Patnaik KS. Autonomous UAV navigation using deep learning-based computer vision frameworks: A systematic literature review. Array 2024;23:100361.

[66] Moosavi S, Razavi-Far R, Palade V, Saif M. Explainable artificial intelligence approach for diagnosing faults in an induction furnace. Electronics 2024;13(9):1721. http://dx.doi.org/10.3390/electronics13091721.

[67] Hussain I, Raza W, Sajjad U, Abbas N, Ali HM, Hamid K, et al. A physics-informed, data-driven framework for estimation and optimization of two-phase pressure drop of refrigerants in mini-and macro channels. Results Eng 2024;23:102538. http://dx.doi.org/10.1016/j.rineng.2024.102538.

[68] Eang C, Lee S. Predictive maintenance and fault detection for motor drive control systems in industrial robots using CNN-RNN-based observers. Sensors 2024;25(1):25. http://dx.doi.org/10.3390/s25010025.

[69] Ukiwe EK, Adeshina SA, Jacob T, Adetokun BB. Deep learning model for detection of hotspots using infrared thermographic images of electrical installations. J Electr Syst Inf Technol 2024;11(1):24. http://dx.doi.org/10.1186/s43067-024-00148-y.

[70] Siraj FM, Ayon STK, Samad MA, Uddin J, Choi K. Few-shot lightweight SqueezeNet architecture for induction motor fault diagnosis using limited thermal image dataset. IEEE Access 2024. http://dx.doi.org/10.1109/ACCESS.2024.3385430.

[71] Xu Y. Research on cooling load estimation through optimal hybrid models based on Naive Bayes. J Eng Appl Sci 2024;71(1):75. http://dx.doi.org/10.1186/s44147-024-00396-9.

[72] Ciampi FG, Rega A, Diallo TM, Pelella F, Choley J-Y, Patalano S. Energy consumption prediction of industrial HVAC systems using Bayesian networks. Energy Build 2024;309:114039. http://dx.doi.org/10.1016/j.enbuild.2024.114039.

[73] Jain A, Gupta S. Evaluation of electrical load demand forecasting using various machine learning algorithms. Front Energy Res 2024;12:1408119. http://dx.doi.org/10.3389/fenrg.2024.1408119.

[74] Koukaras P, Mustapha A, Mystakidis A, Tjortjis C. Optimizing building short-term load forecasting: A comparative analysis of machine learning models. Energies 2024;17(6):1450. http://dx.doi.org/10.3390/en17061450.

[75] Munir S, Pradhan MR, Abbas S, Khan MA. Energy consumption prediction based on lightgbm empowered with explainable artificial intelligence. IEEE Access 2024. http://dx.doi.org/10.1109/ACCESS.2024.3418967.

[76] Khan HA, Bhatti U, Kamal K, Alkahtani M, Abidi MH, Mathavan S. Fault classification for cooling system of hydraulic machinery using AI. Sensors 2023;23(16):7152. http://dx.doi.org/10.3390/s23167152.

[77] Kaushik K, Naik V. Detecting faults in the cooling systems by monitoring temperature and energy. Energy Inform 2024;7(1):46. http://dx.doi.org/10.1186/s42162-024-00351-1.

[78] Zhao T, Yang J, Zhu J, Peng M, Lu C, Shi Z. Quantitative detection of refrigerant charge faults in multi-unit air conditioning systems based on machine learning algorithms. Int J Refrig 2025;169:184–93. http://dx.doi.org/10.1016/j.ijrefrig.2024.10.026.

[79] Qais MH, Kewat S, Loo KH, Lai C-M, Leung A. LSTM-based stacked autoencoders for early anomaly detection in induction heating systems. Mathematics 2023;11(15):3319. http://dx.doi.org/10.3390/math11153319.

[80] Aker E, Othman ML, Veerasamy V, Aris Ib, Wahab NIA, Hizam H. Fault detection and classification of shunt compensated transmission line using discrete wavelet transform and naive bayes classifier. Energies 2020;13(1):243. http://dx.doi.org/10.3390/en13010243.

[81] Xia Y, Zhao J, Ding Q, Jiang A. Incipient chiller fault diagnosis using an optimized least squares support vector machine with gravitational search algorithm. Front Energy Res 2021;9:755649. http://dx.doi.org/10.3389/fenrg.2021.755649.

[82] Guo C, Dong M, Wu Z. Fault diagnosis of power transformers based on comprehensive machine learning of dissolved gas analysis. In: 2019 IEEE 20th international conference on dielectric liquids. IEEE; 2019, p. 1–4.

[83] Hearst M, Dumais S, Osuna E, Platt J, Scholkopf B. Support vector machines. IEEE Intell Syst Appl 1998;13(4):18–28. http://dx.doi.org/10.1109/5254.708428.

[84] Bi J, Wang H, Yan E, Wang C, Yan K, Jiang L, et al. AI in HVAC fault detection and diagnosis: A systematic review. Energy Rev 2024;100071. http://dx.doi.org/10.1016/j.enrev.2024.100071.

[85] Barandier P, Mendes M, Cardoso AJM. Comparative analysis of four classification algorithms for fault detection of heat pumps. Energy Build 2024;114342. http://dx.doi.org/10.1016/j.enbuild.2024.114342.

[86] Amit Y, Geman D. Shape quantization and recognition with randomized trees. Neural Comput 1997;9(7):1545–88. http://dx.doi.org/10.1162/neco.1997.9.7.1545.

[87] Li M, Tao H, Liu M, He T. Study on enhanced fault diagnosis of chiller units in HVAC systems under the imbalanced data environment using GA-optimized lightgbm. Energy Build 2025;115360. http://dx.doi.org/10.1016/j.enbuild.2025.115360.

[88] Guo G, Wang H, Bell D, Bi Y, Greer K. KNN model-based approach in classification. In: On the move to meaningful internet systems 2003: coopIS, DOA, and ODBASE: OTM confederated international conferences, coopIS, DOA, and oDBASE 2003. Springer; 2003, p. 986–96.

[89] Ho W, Yu F. Measurement and verification of energy performance for chiller system retrofit with k nearest neighbour regression. J Build Eng 2022;46:103845. http://dx.doi.org/10.1016/j.jobe.2021.103845.

[90] Gao L, Li D, Liu X, Liu G. Enhanced chiller faults detection and isolation method based on independent component analysis and k-nearest neighbors classifier. Build Environ 2022;216:109010. http://dx.doi.org/10.1016/j.buildenv.2022.109010.

[91] Qin K, Chen L, Shi J, Li Z, Hao K. Root cause analysis of industrial faults based on binary extreme gradient boosting and temporal causal discovery network. Chemometr Intell Lab Syst 2022;225:104559. http://dx.doi.org/10.1016/j.chemolab.2022.104559.

[92] Guo J, Yun S, Meng Y, He N, Ye D, Zhao Z, et al. Prediction of heating and cooling loads based on light gradient boosting machine algorithms. Build Environ 2023;236:110252. http://dx.doi.org/10.1016/j.buildenv.2023.110252.

[93] Lin C-M, Lin S-F, Liu H-Y, Tseng K-Y. Applying the naïve Bayes classifier to HVAC energy prediction using hourly data. Microsyst Technol 2019;1–15. http://dx.doi.org/10.1007/s00542-019-04479-z.

[94] Soltani Z, Sørensen KKr, Leth J, Bendtsen JD. Fault detection and diagnosis in refrigeration systems using machine learning algorithms. Int J Refrig 2022;144:34–45. http://dx.doi.org/10.1016/j.ijrefrig.2022.08.008.

[95] Han H, Cui X, Fan Y, Qing H. Least squares support vector machine (LS-SVM)-based chiller fault diagnosis using fault indicative features. Appl Therm Eng 2019;154:540–7. http://dx.doi.org/10.1016/j.applthermaleng.2019.03.111.

[96] Havaeji S, Ghanizadeh Anganeh P, Torbat Esfahani M, Rezaeihezaveh R, Rezaei Moghadam A. A comparative analysis of machine learning techniques for building cooling load prediction. J Build Pathol Rehabil 2024;9(2):119. http://dx.doi.org/10.1007/s41024-024-00466-8.

[97] Han H, Gao J, Gu B, Ren Z. Comprehensive study on building chiller fault feature and diagnosis using hybrid CNN. Sci Technol Built Environ 2025;31(1):47–65. http://dx.doi.org/10.1080/23744731.2024.2402200.

[98] Khan PW, Byun YC. Fault prognosis of wind turbines using multimodal machine learning. Energy Rep 2024;12:4938–52. http://dx.doi.org/10.1016/j.egyr.2024.10.043.

[99] Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, et al. Recent advances in convolutional neural networks. Pattern Recognit 2018;77:354–77. http://dx.doi.org/10.1016/j.patcog.2017.10.013.

[100] Zhao X, Zhao Y, Hu S, Wang H, Zhang Y, Ming W. Progress in active infrared imaging for defect detection in the renewable and electronic industries. Sensors 2023;23(21):8780. http://dx.doi.org/10.3390/s23218780.

[101] Medsker LR, Jain L, et al. Recurrent neural networks. Des Appl 2001;5(64–67):2.

[102] Pérez J, Pérez S, Moya JM, Arroba P. Thermal prediction for immersion cooling data centers based on recurrent neural networks. In: International conference on intelligent data engineering and automated learning. Springer; 2018, p. 491–8.

[103] Wu S, Jiang Y, Luo H, Yin S. Remaining useful life prediction for ion etching machine cooling system using deep recurrent neural network-based approaches. Control Eng Pract 2021;109:104748. http://dx.doi.org/10.1016/j.conengprac.2021.104748.

[104] Fan C, Lei Y, Sun Y, Mo L. Novel transformer-based self-supervised learning methods for improved HVAC fault diagnosis performance with limited labeled data. Energy 2023;278:127972. http://dx.doi.org/10.1016/j.energy.2023.127972.

[105] El Mokhtari K, McArthur J. Autoencoder-based fault detection using building automation system data. Adv Eng Inform 2024;62:102810. http://dx.doi.org/10.1016/j.aei.2024.102810.

[106] Shahid ZK, Saguna S, Ahlund C. Variational autoencoders for anomaly detection and transfer knowledge in electricity and district heating consumption. IEEE Trans Ind Appl 2024. http://dx.doi.org/10.1109/TIA.2024.3425805.

[107] Creswell A, White T, Dumoulin V, Arulkumaran K, Sengupta B, Bharath AA. Generative adversarial networks: An overview. IEEE Signal Process Mag 2018;35(1):53–65. http://dx.doi.org/10.1109/MSP.2017.2765202.

[108] Yan K, Chong A, Mo Y. Generative adversarial network for fault detection diagnosis of chillers. Build Environ 2020;172:106698. http://dx.doi.org/10.1016/j.buildenv.2020.106698.

[109] Yan K. Chiller fault detection and diagnosis with anomaly detective generative adversarial network. Build Environ 2021;201:107982. http://dx.doi.org/10.1016/j.buildenv.2021.107982.

[110] Aslan B, Balci S, Kayabasi A. Fault diagnosis in thermal images of transformer and asynchronous motor through semantic segmentation and different CNN models. Appl Therm Eng 2025;125599. http://dx.doi.org/10.1016/j.applthermaleng.2025.125599.

[111] Shiney SA, Seetharaman R, Sharmila V, Prathiba S. Deep learning based gasket fault detection: a CNN approach. Sci Rep 2025;15(1):4776. http://dx.doi.org/10.1038/s41598-025-85223-8.

[112] Mahmoud KA, Badr MM, Elmalhy NA, Hamdy RA, Ahmed S, Mordi AA. Transfer learning by fine-tuning pre-trained convolutional neural network architectures for switchgear fault detection using thermal imaging. Alex Eng J 2024;103:327–42. http://dx.doi.org/10.1016/j.aej.2024.05.102.

[113] Zheng W, Zhang S, Omer AM, Wu Z, Tao N, Zhang C, et al. A novel approach for one-step defect detection and depth estimation using sequenced thermal signal encoding. Nondestruct Test Eval 2024;1–16. http://dx.doi.org/10.1080/10589759.2024.2304719.

[114] Memari M, Shekaramiz M, Masoum MA, Seibi AC. Data fusion and ensemble learning for advanced anomaly detection using multi-spectral RGB and thermal imaging of small wind turbine blades. Energies 2024;17(3):673. http://dx.doi.org/10.3390/en17030673.

[115] Liu J, Long X, Jiang C, Liao W. Multi-feature vision transformer for automatic defect detection and quantification in composites using thermography. NDT E Int 2024;143:103033. http://dx.doi.org/10.1016/j.ndteint.2023.103033.

[116] Wang C, Fu Z, Zhang Z, Wang W, Chen H, Xu D. Fault diagnosis of power transformer in one-key sequential control system of intelligent substation based on a transformer neural network model. Processes 2024;12(4):824. http://dx.doi.org/10.3390/pr12040824.

[117] John LS, Yoon S, Li J, Wang P. Anomaly detection using convolutional autoencoder with residual gated recurrent unit and weak supervision for photovoltaic thermal heat pump system. J Build Eng 2025;100:111694. http://dx.doi.org/10.1016/j.jobe.2024.111694.

[118] Zhang Y, Xu C, Liu P, Xie J, Han Y, Liu R, et al. One-dimensional deep convolutional autoencoder active infrared thermography: Enhanced visualization of internal defects in FRP composites. Compos Part B: Eng 2024;272:111216. http://dx.doi.org/10.1016/j.compositesb.2024.111216.

[119] Yang W, Zhang H, Lim JB, Zhang Y, Meng H. A new chiller fault diagnosis method under the imbalanced data environment via combining an improved generative adversarial network with an enhanced deep extreme learning machine. Eng Appl Artif Intell 2024;137:109218. http://dx.doi.org/10.1016/j.engappai.2024.109218.

[120] Jagtap NV, Mudunuru MK, Nakshatrala KB. Coolpinns: A physics-informed neural network modeling of active cooling in vascular systems. Appl Math Model 2023;122:265–87. http://dx.doi.org/10.1016/j.apm.2023.04.020.

[121] Pan X, Liang K, Jiao Y, Chen Z, Wang Y. Interpretable chiller fault diagnosis based on physics-guided neural networks. J Build Eng 2024;86:108991. http://dx.doi.org/10.1016/j.jobe.2024.108991.

[122] Chen D, Chui C-K, Lee PS. Adaptive physically consistent neural networks for data center thermal dynamics modeling. Appl Energy 2025;377:124637. http://dx.doi.org/10.1016/j.apenergy.2024.124637.

[123] Zobeiry N, Humfeld KD. A physics-informed machine learning approach for solving heat transfer equation in advanced manufacturing and engineering applications. Eng Appl Artif Intell 2021;101:104232. http://dx.doi.org/10.1016/j.engappai.2021.104232.

[124] Raissi M, Perdikaris P, Karniadakis GE. Physics informed deep learning (part i): Data-driven solutions of nonlinear partial differential equations. 2017, arXiv preprint arXiv:1711.10561.

[125] Kharazmi E, Zhang Z, Karniadakis GE. Variational physics-informed neural networks for solving partial differential equations. 2019, arXiv preprint arXiv:1912.00873.

[126] McClenny LD, Braga-Neto UM. Self-adaptive physics-informed neural networks. J Comput Phys 2023;474:111722.

[127] Shukla K, Jagtap AD, Karniadakis GE. Parallel physics-informed neural networks via domain decomposition. J Comput Phys 2021;447:110683.

[128] Jagtap AD, Karniadakis GE. Extended physics-informed neural networks (XPINNs): A generalized space-time domain decomposition based deep learning framework for nonlinear partial differential equations. Commun Comput Phys 2020;28(5).

[129] Hu Z, Jagtap AD, Karniadakis GE, Kawaguchi K. Augmented physics-informed neural networks (APINNs): A gating network-based soft domain decomposition methodology. Eng Appl Artif Intell 2023;126:107183.

[130] Yu J, Lu L, Meng X, Karniadakis GE. Gradient-enhanced physics-informed neural networks for forward and inverse PDE problems. Comput Methods Appl Mech Engrg 2022;393:114823.

[131] Goswami S, Anitescu C, Chakraborty S, Rabczuk T. Transfer learning enhanced physics informed neural network for phase-field modeling of fracture. Theor Appl Fract Mech 2020;106:102447.

[132] Li Z, Zheng H, Kovachki N, Jin D, Chen H, Liu B, et al. Physics-informed neural operator for learning partial differential equations. ACM/ JMS J Data Sci 2024;1(3):1–27.

[133] Lin B, Mao Z, Wang Z, Karniadakis GE. Operator learning enhanced physics-informed neural networks for solving partial differential equations characterized by sharp solutions. 2023, arXiv preprint arXiv:2310.19590.

[134] Yang L, Meng X, Karniadakis GE. B-PINNs: Bayesian physics-informed neural networks for forward and inverse PDE problems with noisy data. J Comput Phys 2021;425:109913.

[135] Penwarden M, Zhe S, Narayan A, Kirby RM. Multifidelity modeling for physics-informed neural networks (PINNs). J Comput Phys 2022;451:110844.

[136] Liang X, Liu Y, Chen S, Li X, Jin X, Du Z. Physics-informed neural network for chiller plant optimal control with structure-type and trend-type prior knowledge. Appl Energy 2025;390:125857.

[137] Matetić I, Štajduhar I, Wolf I, Ljubic S. A review of data-driven approaches and techniques for fault detection and diagnosis in HVAC systems. Sensors 2022;23(1):1. http://dx.doi.org/10.3390/s23010001.

[138] Ala'raj M, Radi M, Abbod MF, Majdalawieh M, Parodi M. Data-driven based HVAC optimisation approaches: A systematic literature review. J Build Eng 2022;46:103678. http://dx.doi.org/10.1016/j.jobe.2021.103678.