Progress in Natural Science: Materials International 36 (2026) 1–19

HOSTED BY

Contents lists available at ScienceDirect

Progress in Natural Science: Materials International

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

A comprehensive review of machine learning for superalloys: from data-driven prediction to intelligent design

Linlin Sun[a,b], Jie Xiong[c,d], Qingshuang Ma[a,b], Chenghao Pei[a,b], Huijun Li[e], Qiuzhi Gao[a,b,*]

[a] School of Materials Science and Engineering, Northeastern University, Shenyang, 110819, China
[b] School of Resources and Materials, Northeastern University at Qinhuangdao, Qinhuangdao, 066004, China
[c] Materials Genome Institute, Shanghai University, Shanghai, 200444, China
[d] State Key Laboratory of Materials for Advanced Nuclear Energy, Shanghai University, Shanghai, 200444, China
[e] Faculty of Engineering and Information Sciences, University of Wollongong, Wollongong, NSW, 2522, Australia

ARTICLE INFO

Keywords:
Machine learning (ML)
Superalloy
Forward prediction
Intelligent design
High temperature properties

ABSTRACT

Data-centric materials informatics has become a transformative paradigm for accelerating the discovery and design of superalloys, particularly by enabling efficient prediction of properties that are experimentally inaccessible or computationally intractable due to constraints in cost, time, or complexity. By harnessing the ability of machine learning (ML) to model complex, nonlinear, and high-dimensional relationships, this approach provides a compelling alternative to traditional trial-and-error and simulation-based strategies. This review presents a comprehensive and critical assessment of recent advances in ML for superalloys. We first delineate the essential workflow for ML-enabled superalloy design, encompassing foundational data resources, quantitative assessments of data quality, feature descriptors and feature-selection strategies, representative algorithms tailored to small and heterogeneous datasets, rigorous model-evaluation protocols, and model interpretation through explainable ML and symbolic regression. We then summarize state-of-the-art ML applications targeting specific high-temperature performance metrics, particularly γ' phase stability, creep behavior, fatigue life, and oxidation resistance, and highlight how approaches such as multi-fidelity learning, data augmentation, transfer learning, and optimization algorithms facilitate efficient exploration of vast composition-processing design spaces. Finally, we discuss persisting challenges and emerging opportunities, including data scarcity and reliability, model confidence and uncertainty quantification, cross-system generalizability across Co-, Ni-, and multi-principal superalloys, high-dimensional multi-objective optimization, and the integration of physics-informed models and large language models into materials-informatics workflows. By synthesizing these developments, this review outlines a strategic roadmap for harnessing ML to accelerate the discovery, performance optimization, and intelligent design of next-generation superalloys.

1. Introduction

Superalloys, renowned for their exceptional mechanical strength and resistance to thermal creep deformation, are inherently complex, high-dimensional, and multiscale-coupled systems. Despite their critical applications in aerospace, power generation, and other high-temperature environments, existing fundamental theories struggle to accurately and quantitatively describe the intricate relationships between composition, processing, microstructure, properties, and service behavior. Several key mechanisms remain insufficiently understood, necessitating a prolonged reliance on experimental trial-and-error approaches [1–3] and various computational simulations [4,5]. Experimental strategies, while invaluable for gaining insights, are often prohibitively costly, time-consuming, and labor-intensive, with outcomes heavily dependent on the researcher's expertise. Given the vast combinatorial possibilities in alloy design, the efficiency of experimental trial-and-error methods remains limited. Computational techniques like density functional theory (DFT) offer precise atomic-scale calculations [6], but their applicability is restricted by significant computational demands, especially for complex alloys [7]. Similarly, physical models based on in prior knowledge, such as molecular dynamics (MD) and Monte Carlo (MC) simulations, can handle large-scale systems over shorter timescales but fall short in predictive capability for alloy design [8]. Furthermore, while physical models grounded in prior knowledge have proven

* Corresponding author. School of Materials Science and Engineering, Northeastern University, Shenyang, 110819, China.
E-mail address: neuqgao@163.com (Q. Gao).

https://doi.org/10.1016/j.pnsc.2025.12.004
Received 6 August 2025; Received in revised form 6 December 2025; Accepted 11 December 2025
Available online 11 February 2026

successful for specific alloy systems, they often lose fidelity when applied to different alloy systems [9]. Thus, these limitations present significant challenges in exploring the extensive search space of superalloys.

Machine learning (ML) is rapidly bridging the gap between traditional experimental trial-and-error approaches and computational simulations by offering surrogate models that combine the strengths of both. Rather than relying on labor-intensive experimental processes or the specific functional forms and parameterizations characteristic of computational simulations, which limit their generalizability, ML methods leverage experimental data reported in the literature alongside simulation data to predict the specific properties of a given alloy composition, processing condition, microstructure, or crystal structure [10,11]. A key advantage of ML is its ability to provide significantly faster predictions, with speeds several orders of magnitude higher than those of traditional experimental and computational simulations. For instance, certain ML models can predict specific properties within milliseconds, compared to the hours or days required by DFT or MD simulations [12]. Moreover, by directly correlating alloy properties with composition and processing parameters, ML mitigates the cumulative errors often introduced by the complexity of multi-mechanism interactions, thereby enhancing both predictive accuracy and robustness [13–15].

Data is the cornerstone of ML. ML algorithms rely on extensive datasets to identify and learn underlying trends, correlations, and patterns [16–18]. In the realm of superalloys, sustained research and development efforts over decades have generated a wealth of data, encompassing alloy compositions, processing conditions, microstructural characterizations, and performance metrics. This has laid a solid foundation for the widespread application of ML in superalloy research. Through ML, it is feasible to establish mapping relationships between key feature factors, such as alloy composition, heat treatment processes, and microscopic electrophysical properties, and target performance. This capability not only facilitates the prediction of alloy properties but also accelerates the discovery of novel superalloy compositions [11]. Furthermore, ML can aid superalloy researchers in gaining deeper insights into the mechanisms governing superalloy behavior across various scales and dimensions, ultimately revealing their underlying scientific principles [19,20].

In recent years, ML methods have been increasingly applied to superalloy research, particularly in alloy composition design and performance prediction. While significant progress has been made, several challenges remain. Therefore, it is essential to comprehensively understand both the latest research advancements and the ongoing challenges in applying ML to superalloys. While several excellent reviews exist on the broader application of ML in materials science [7,8,21–28], there remains a noticeable gap in comprehensive reviews focusing specifically on superalloys. For instance, Liu et al. [7] reviewed ML applications for high-entropy alloys, focusing on state-of-the-art ML models that describe atomic interactions and atomistic simulations of thermodynamic and mechanical properties. Another review by Hu et al. [25] summarized the current state of ML-driven alloy research. Both review papers discuss ML applications for alloys in general; However, ML's unique challenges and opportunities for superalloys are not thoroughly discussed. Given the extensive attention that superalloys have attracted, a timely and focused review is warranted. Specifically, we present a comprehensive review on the interdisciplinary topic of ML for superalloys, discuss the critical aspects involved in implementing ML, summarize key applications to illustrate the advantages of ML for superalloys, and highlight the existing challenges and opportunities in this field.

2. Critical steps of ML methods for superalloy

Fig. 1 presents the overall framework for applying ML to superalloys. The following sections will focus on key components within this framework, including the fundamental data, data quality evaluation, feature descriptors and feature selection, representative ML algorithms, model evaluation, model interpretation, virtual space, optimization

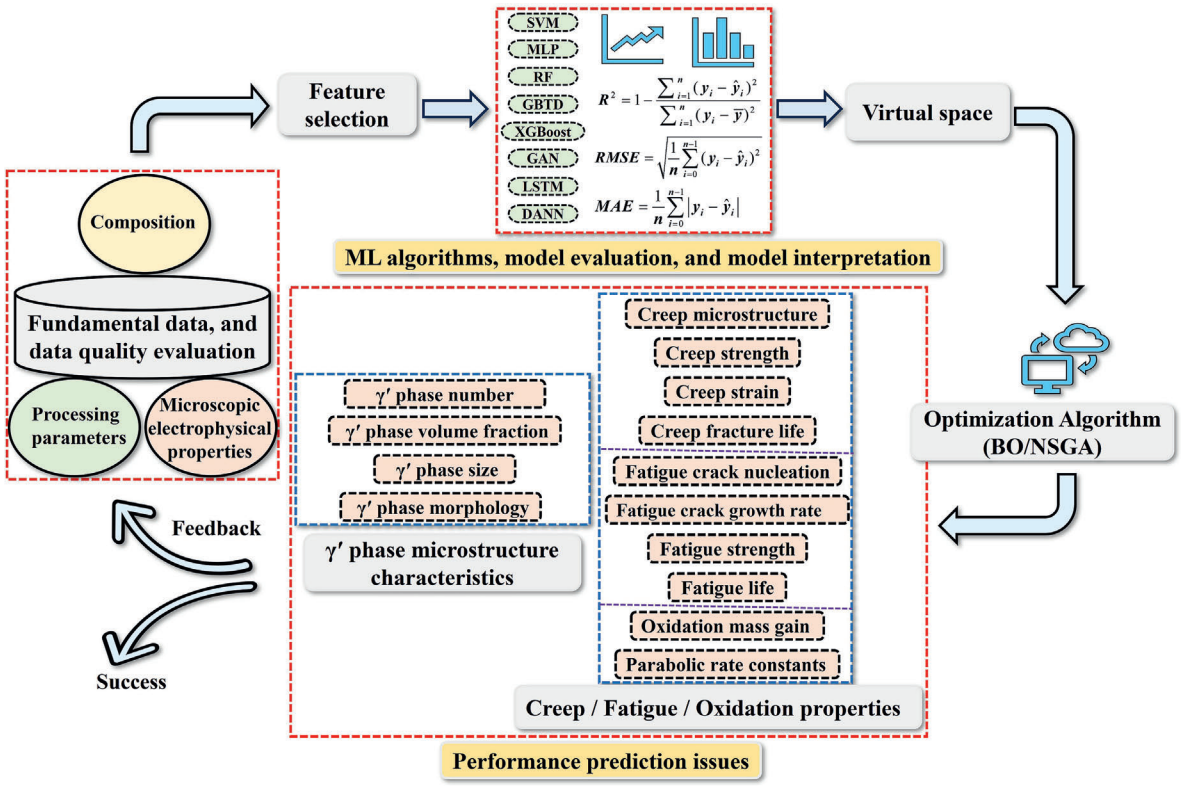

Fig. 1. Overall framework for applying ML to superalloys.

algorithms, and the performance prediction issues associated with superalloys using ML.

2.1. Fundamental data

The availability of reliable fundamental data is a prerequisite for accelerating research and development in superalloys. Existing databases containing structural, thermodynamic, and property information, derived from high-throughput experiments, computational thermodynamics, and first-principles calculations, provide essential resources for data-driven alloy design. Although a dedicated, comprehensive superalloy database is not yet available, several well-established repositories contain extensive data relevant to superalloys. Representative examples are summarized in Table 1. For instance, the MatNavi database developed by the National Institute for Materials Science (NIMS, Japan) offers a substantial collection of tensile, creep, and fatigue data for Fe-, Ni-, and Co-based superalloys.

Beyond these curated databases, vast quantities of high-quality, peer-reviewed experimental data are embedded in the scientific literature. However, manually mining such information is both time-consuming and knowledge-intensive. To address this challenge, Wang et al. [34] developed a rule-based named-entity-recognition framework coupled with a distance-based heuristic multi-relation extraction algorithm. As illustrated in Fig. 2, this workflow enables automated extraction of composition, processing, and performance data from literature with high accuracy, particularly for small and domain-specific corpora. Using similar strategies, large-scale datasets of superalloy compositions and properties can be constructed efficiently, providing a robust foundation for property prediction and rational alloy design.

Despite their value, literature-derived datasets typically remain small in size. To mitigate this intrinsic limitation, data-augmentation techniques have emerged as powerful tools [35]. Among them, generative adversarial networks (GANs) have proved particularly effective. In GANs, a generator is trained to produce synthetic samples that closely mimic the original data, while a discriminator learns to distinguish real from synthetic samples. Through this adversarial interplay, the model captures the underlying latent structure of the dataset, thereby enhancing both diversity and representativeness. For example, Lee et al. [36] successfully applied GAN-based augmentation to high-entropy alloys, improving the phase-prediction accuracy of a deep neural network from 84.75 % to 93.17 %. Such generative approaches substantially strengthen the data foundation of ML-driven superalloy research and accelerate the discovery of new alloy chemistries.

In addition to data scarcity, materials datasets often suffer from incomplete, ambiguous, or partially missing information, particularly regarding alloy compositions or processing histories. To address this pervasive issue, Lv et al. [37] innovatively integrated transfer learning with partial-label learning to construct an uncertainty-aware predictive framework (Fig. 3). This model leverages transferable knowledge from related material systems to intelligently infer, refine, and correct missing or ambiguous information in the target dataset. By reconstructing the "true" compositions from incomplete records, the framework significantly enhances both the accuracy and generalizability of fatigue-performance prediction models. This strategy highlights a promising avenue for harnessing machine learning to extract reliable knowledge from imperfect materials data, thereby advancing ML-assisted alloy design.

In summary, reliable fundamental data remain the cornerstone of ML-enabled superalloy research, yet current resources are fragmented across curated databases and dispersed literature. Recent advances have begun to close this gap through automated information extraction and data-augmentation frameworks, which together expand dataset coverage while preserving physical fidelity. Moreover, emerging strategies that integrate transfer learning with partial-label learning offer a pathway to correct incomplete or ambiguous records, thereby strengthening model robustness in realistic, imperfect data environments. Collectively, these developments transform data availability from a limiting bottleneck into an enabling capability, laying the groundwork for scalable, accurate, and generalizable ML-driven superalloy design.

2.2. Data quality evaluation

The reliability of ML applications in superalloy research is determined not only by dataset size, but increasingly by the quality, transparency, and internal coherence of the underlying data. Recent studies [38,39] have introduced quantitative evaluation frameworks centered on completeness, accuracy, and consistency, enabling systematic assessment of heterogeneous data sources.

(1) Completeness evaluation. Completeness reflects the availability of essential composition-processing-microstructure-property information and is commonly quantified using a completeness ratio:

$$C = \frac{N_{complete}}{N_{total}} \tag{1}$$

where N_{complete} denotes the number of records containing full chemical composition, heat-treatment descriptors, microstructural parameters, and corresponding property values (e.g., γ' phase coarsening rate, oxidation kinetics, tensile strength, creep life, fatigue lifetime, etc.), and N_{total} is total number of records in the dataset. Records missing critical target properties are typically excluded to avoid label uncertainty, whereas partially missing auxiliary fields may be reconstructed through literature tracing, CALPHAD-guided calculation, or statistically informed estimation when physically justified.

(2) Accuracy assessment. Accuracy describes the degree of agreement between reported data and more reliable references (e.g., curated databases, benchmark experiments, or well-established literature values). In superalloy datasets, accuracy is evaluated at both the single-entry and dataset level. At the single-entry level, the relative deviation between two independent sources can be expressed as:

Table 1
Online databases for superalloys.

Database Name	Description	Availability
Materials Data Sharing (MSDSN)	13702 superalloy experimental data	http://www.materdata.cn/
MatWeb	1106 superalloy experimental data	https://matweb.com/
CINDAS LLC	32 superalloy experimental data	https://cindasdata.com/
NIMS Materials Database (MatNavi)	Experimental data	https://mits.nims.go.jp/
Inorganic Crystal Structure Database (ICSD) [29]	Crystal structure data	http://icsd.fiz-karlsruhe.de
Crystallography Open Database (COD) [30]		https://www.crystallography.net/cod/
The Open Quantum Materials Database (OQMD) [31]	DFT calculated thermodynamic and structural properties data	https://oqmd.org/
Materials Project (MP) [32]		https://www.materialsproject.org/
Automatic Flow (AFLOW) [33]		http://www.aflowlib.org/
Atomly		https://atomly.net/#/matdata

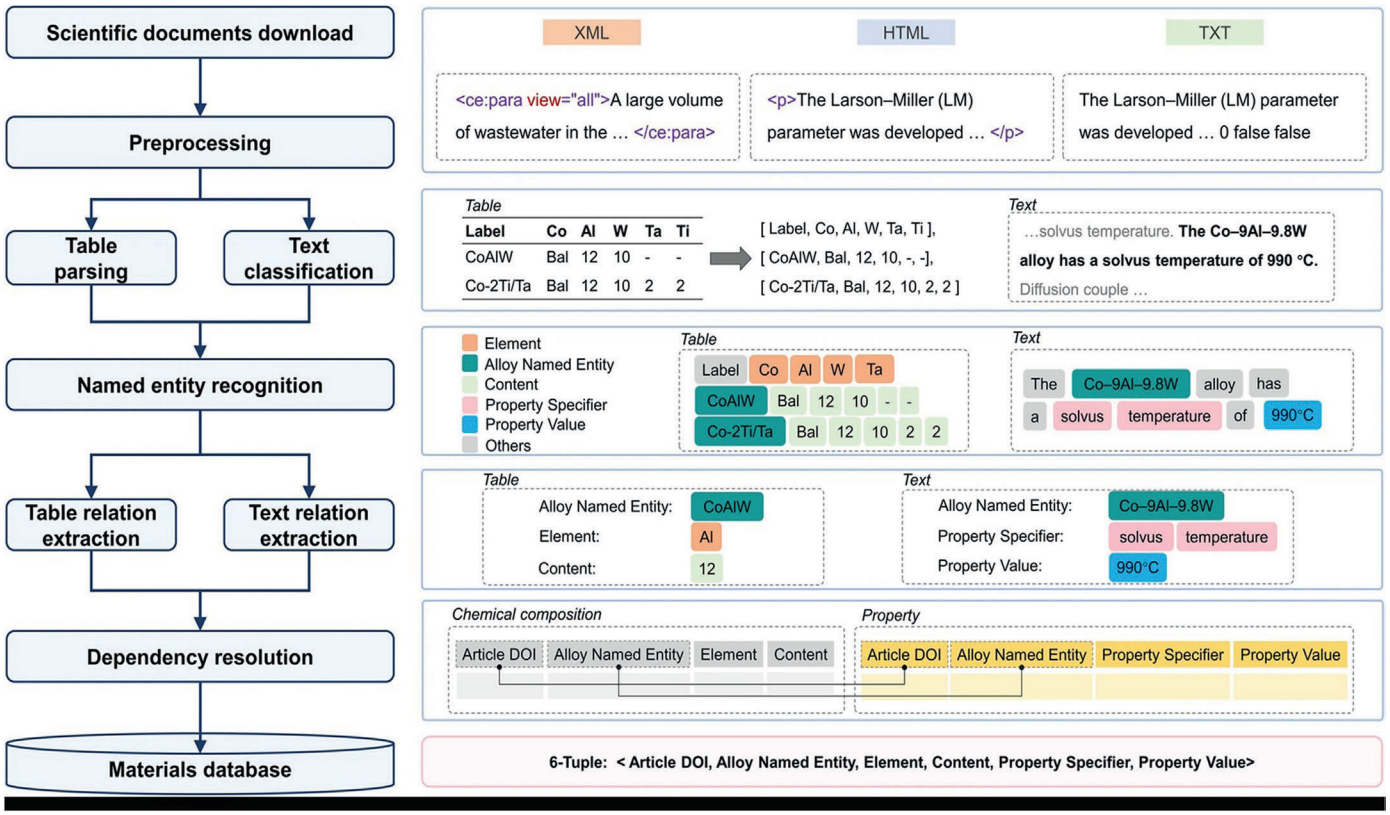

Fig. 2. Schematic workflow of the automated text mining pipeline [34]. The workflow involves several stages of scientific documents download, preprocessing, table parsing, text classification, named entity recognition, table and text relation extraction, and interdependency resolution. A corpus of scientific articles is scraped and the irrelevant information in raw corpus is then filtered during preprocessing. According to the table parsing and text classification, the tables and sentences with target information are determined for named entity recognition and relation extraction. The alloy named entity, property specifier and property value are recognized by named entity recognition, and relation extraction of text and table gives the specific tuple relations. Interdependency resolution resolves the linkage to chemical composition and property data fragments for one specific material, and finally outputs a complete record into materials database.

$$\delta_i = \frac{\left| x_i^{(source1)} - x_i^{(source2)} \right|}{x_i^{(source2)}} \tag{2}$$

where $x_i^{source1}$ is the value of the i-th data point reported in one source (e.g., a database or a specific paper), and $x_i^{source2}$ is the corresponding value reported in an independent reference source (e.g., another database or a different experimental study). A pointwise accuracy score is then defined as:

$$A_i = 1 - \delta_i \tag{3}$$

such that $A_i \to 1$ indicates excellent agreement, while lower A_i values reflect larger deviations. At the dataset level, an average accuracy indicator can be calculated over N_{pair} cross-comparable entries:

$$A_{avg} = \frac{1}{N_{pair}} \sum_{i=1}^{N_{pair}} A_i \tag{4}$$

where N_{pair} is the number of entries with cross-source correspondence, and A_{avg} represents the dataset-level average accuracy.

Beyond cross-source comparisons, statistical anomaly detection is widely employed to identify suspicious measurements. A commonly used approach is z-score analysis:

$$z_i = \frac{x_i - \mu}{\sigma} \tag{5}$$

where x_i is the value of the i-th data point and μ and σ are the sample mean and standard deviation, respectively. Values exceeding a threshold (typically $|z_i| > 3$) are flagged as statistical outliers and removed or subjected to manual inspection. Complementary methods, such as interquartile-range (IQR) filtering, serve a similar purpose. Physical plausibility checks further compare reported values against thermodynamic and phase-stability constraints (e.g., γ' volume fractions not exceeding equilibrium limits), providing an additional layer of reliability. Collectively, cross-source validation, statistical outlier detection, and physics-based screening form a multi-level accuracy evaluation framework that substantially reduces the impact of experimental noise, transcription errors, and inconsistent testing protocols.

(3) Consistency harmonization. Consistency harmonization ensures that data originating from heterogeneous sources can be meaningfully compared and integrated by eliminating discrepancies in representation, terminology, and measurement scales. A commonly cited quantitative indicator is the normalized consistency score:

$$S_f = \frac{\sigma_f}{\mu_f} \tag{6}$$

where σ_f and μ_f represent the standard deviation and mean of feature f, respectively. Features exhibiting extremely low variance (e.g., $S_f \approx 0$) are typically removed due to negligible information content and the risk of implicit weighting during model training. Consistency harmonization generally involves: (i) unit standardization, including converting of wt.% to at.% through atomic-weight normalization, unifying stress units (e.g., MPa vs GPa) and temperature scales, and aligning time formats (e.g., aging duration in hours vs minutes); and (ii) terminology and descriptor unification, such as standardizing phase labels (γ' vs L_{12}, FCC vs γ matrix), harmonizing heat-treatment notations (e.g., "solutionized at

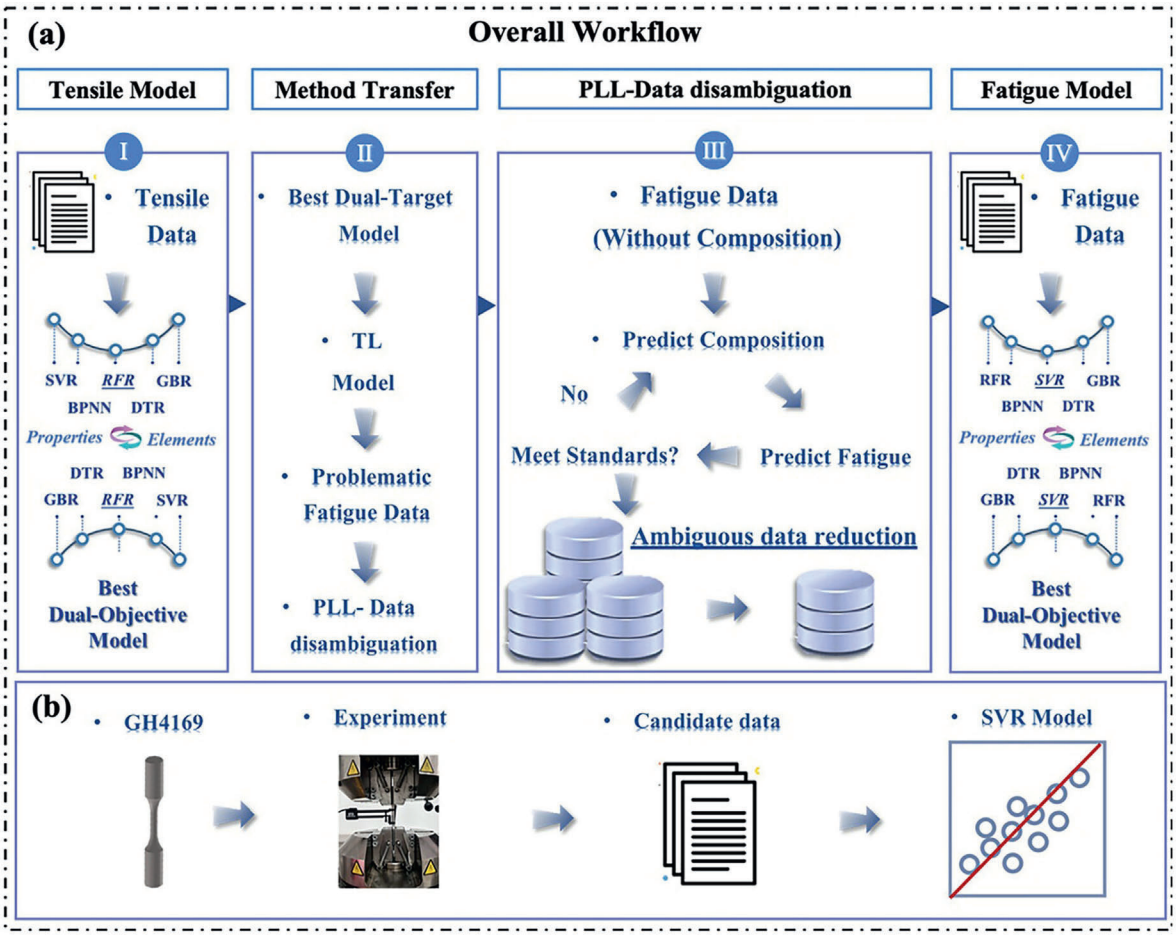

Fig. 3. Schematic overview of the proposed methodology [37]. (a) Integrated workflow combining partial label learning (PLL) and transfer learning (TL) for alloy composition mining and fatigue performance prediction. The framework consists of four key modules: (I) prediction of tensile properties and inferred composition; (II) knowledge transfer via a TL strategy; (III) ambiguity resolution using PLL to handle partially labeled compositional data; and (IV) final prediction of fatigue performance. (b) Experimental validation of the developed model, including alloy fabrication, fatigue testing, data acquisition, and quantitative assessment of prediction accuracy.

1250 °C for 1 h" vs "ST1250-1h"), and resolving ambiguous expressions lacking explicit conditions.

Overall, the adoption of quantitative data-quality evaluation signals a shift from data accumulation toward data reliability, establishing a more rigorous foundation for trustworthy, generalizable, and physics-consistent insights in superalloy informatics.

2.3. Feature descriptors, and feature selection

Feature descriptors are fundamental to establishing quantitative structure-property relationships that link alloy composition, processing conditions, microstructural attributes, and targeted performance metrics. However, excessively high-dimensional descriptor spaces can introduce the curse of dimensionality, data sparsity, and overfitting, ultimately degrading the generalization capability of ML models and increasing computational cost. Accordingly, feature selection constitutes a critical step in developing robust, efficient, and interpretable ML frameworks, particularly in superalloy research, where datasets are typically heterogeneous, strongly correlated, and limited in size. By identifying the most informative variables while eliminating redundant or weakly relevant descriptors, feature selection reduces model complexity, mitigates overfitting, and improves the physical interpretability of predictions. This step is especially important in Co- and Ni-based superalloys, where candidate descriptors span a wide spectrum, including alloying composition, heat-treatment schedules, γ' phase metrics, diffusion-controlled parameters, thermodynamic quantities, and electronic-structure features, many of which are interdependent and do not contribute equally to predictive performance. Existing feature-selection strategies are generally categorized into three groups.

(1) Filter methods evaluate descriptor relevance independently of the learning algorithm using statistical criteria such as Pearson correlation, mutual information, variance thresholds, or maximum-relevance-minimum-redundancy (mRMR). These approaches offer high computational efficiency and are well suited for the initial screening of large descriptor spaces.
(2) Wrapper methods, including recursive feature elimination (RFE), forward or backward stepwise selection, and cross-validated subset search, iteratively assess feature subsets by training ML models. Although computationally more demanding, they often deliver superior performance for small-sample materials datasets by capturing nonlinear interactions among descriptors.
(3) Embedded methods incorporate feature selection directly into model training, for example through LASSO or elastic-net regularization, sparsity-promoting Bayesian frameworks, or tree-based importance measures such as random forest and XGBoost.

These approaches provide a balanced compromise between efficiency, predictive accuracy, and reproducibility. Recent advances in interpretable and physics-aware ML have introduced hybrid strategies

that integrate statistical selection with domain knowledge. Examples include SHAP-based importance ranking, stability selection under bootstrapping, and metallurgically guided elimination informed by CALPHAD predictions or mechanistic constraints (e.g., Al-Ta governing γ' solvus temperature versus W-Mo controlling diffusion-limited coarsening). For superalloys, such combined strategies not only enhance predictive performance but also uncover meaningful composition-processing-microstructure-property linkages, prevent physically spurious correlations, and provide actionable guidance for alloy design.

Overall, rigorous and reproducible feature-selection pipelines are indispensable for enabling reliable model generalization, reducing computational overhead, and establishing interpretable and scientifically grounded ML frameworks for next-generation superalloy design. Fig. 4 summarizes representative feature descriptors associated with key target properties in Co- and Ni-based superalloys.

2.4. Representative ML algorithms

Different ML algorithms possess distinct strengths and limitations; therefore, selecting an appropriate model requires careful consideration of task characteristics, feature dimensionality, and, critically, the scale and fidelity of the available dataset. Although comprehensive introductions to classical ML algorithms are widely available [40,41], conventional approaches inherently rely on large, high-quality datasets containing hundreds to thousands of alloy samples [25,42]. While such data volumes are often attainable for properties that are inexpensive and straightforward to measure, they remain exceptionally difficult to obtain for high-cost, small-sample properties such as $L1_2$-γ' coarsening kinetics, creep lifetime, or fatigue resistance, which require prolonged testing, complex characterization, and are frequently constrained by industrial confidentiality.

To address this persistent data scarcity, recent studies have developed a suite of practical, domain-specific strategies tailored to the superalloy community. First, data-augmentation techniques, including generative adversarial networks (GANs) and Markov and Chain Monte Carlo (MCMC) sampling, have emerged as effective tools for synthesizing physically realistic microstructural or property data. These methods substantially alleviate the small-sample bottleneck by expanding the training domain while preserving thermodynamic and mechanistic consistency. Sun et al. [43] proposed an interpretable machine-learning (XML) framework combined with a multi-fidelity augmentation strategy specifically designed for small-sample scenarios. Their workflow integrates MCMC sampling and a Wasserstein GAN with gradient penalty (WGAN-GP) to generate thermodynamics-constrained medium-fidelity data, while low-fidelity expansions are produced using the SMOGN algorithm with Gaussian perturbations. The resulting hybrid dataset,

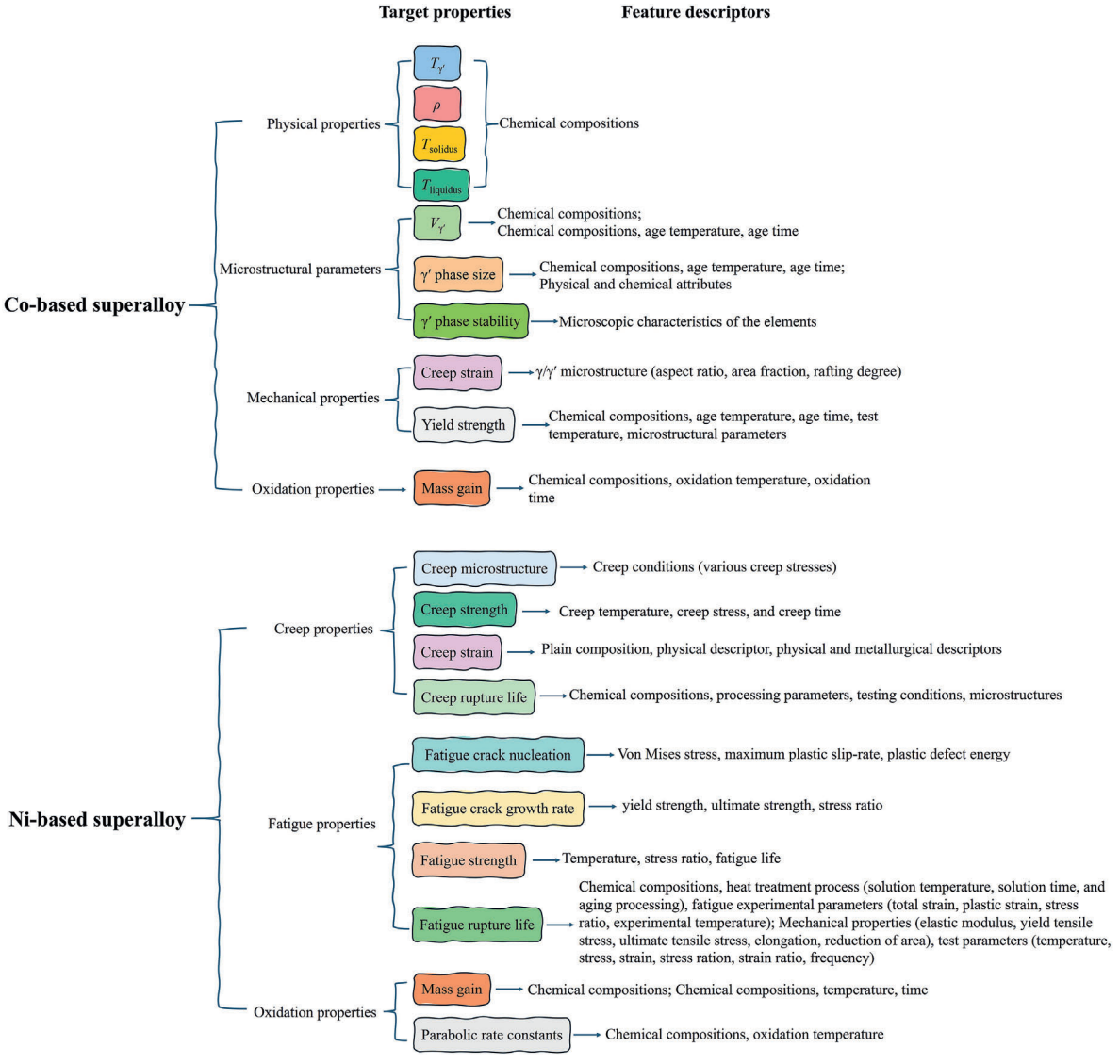

Fig. 4. The feature descriptors corresponding to target properties of Co/Ni-based superalloys.

comprising experimental measurements, thermodynamic simulations, and physically constrained generative data, enables robust multi-fidelity learning. Using this enriched dataset, the authors achieved high-accuracy prediction of the γ′ coarsening rate constant (K_r) with an XGBoost regression model, demonstrating the effectiveness of multi-fidelity augmentation in modeling superalloy properties otherwise limited by sample scarcity.

Importantly, this multi-fidelity paradigm is readily extendable to creep-property prediction. For instance, GAN-based generative models or MCMC sampling can be employed to approximate the underlying data distribution of creep-related features (e.g., composition, microstructure descriptors, temperature-stress conditions). These synthetic data provide a statistically enriched representation of the design space. Subsequently, physics-based computational tools, such as JMatPro, can be used to calculate creep life or steady-state creep rates for these generated compositions, thereby supplying medium-fidelity simulation data. Although such simulated creep data inevitably contain model-dependent uncertainties, they nevertheless offer valuable mechanistic insights and broaden the effective training set. When combined with a limited quantity of high-fidelity experimental creep measurements, the resulting multi-fidelity dataset enables the construction of significantly more reliable machine-learning models for predicting creep life or creep rate in advanced superalloys.

GANs represent one of the most influential generative frameworks in modern deep learning and are widely used to model complex material-property distributions through an adversarial game between a generator and a discriminator. In the classical formulation, the two networks optimize a minimax objective:

$$\min_G \max_D V(D,G) = E_{x\sim p_{data}}[\log D(x)] + E_{z\sim p_z}[\log(1-D(G(z)))] \tag{7}$$

where p_{data} denotes the real data distribution, p_z the latent prior, and $G(z)$ the synthesized sample with model distribution p_g. The generator learns to map latent variables onto the real data manifold, whereas the discriminator aims to distinguish real from synthesized samples; ideally, iterative optimization drives p_g toward p_{data}. However, classical GANs frequently suffer from mode collapse, gradient vanishing, and unstable convergence, limitations that are particularly pronounced in small, noisy materials datasets.

To mitigate these drawbacks, advanced variants such as the Wasserstein GAN (WGAN) replace the Jensen-Shannon divergence with the Wasserstein-1 (Earth mover) distance:

$$W(p_{data}, p_g) = \sup_{\|f\|_L \le 1} \left[E_{x\sim p_{data}}[f(x)] - E_{x\sim p_g}[f(x)] \right] \tag{8}$$

where f ranges over the set of 1-Lipschitz functions. The corresponding critical loss is expressed as:

$$L_{WGAN} = E_{x\sim p_g}[D(x)] - E_{x\sim p_{data}}[D(x)] \tag{9}$$

Further stabilization is achieved using the gradient-penalty variant (WGAN-GP), in which the Lipschitz constraint is enforced via an additional regularization term:

$$L_{WGAN-GP} = L_{WGAN} + \lambda E_{\hat{x}\sim p_{\hat{x}}}(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2 \tag{10}$$

where $\hat{x}$ is sampled along straight lines between real and generated samples and $\lambda > 0$ controls penalty strength. This formulation significantly improves training stability and increases the diversity of generated samples.

Performance evaluation of GAN-generated materials data typically considers both realism and diversity. Widely used quantitative metrics include the Fréchet Inception Distance (*FID*), Maximum Mean Discrepancy (*MMD*), and mode-count-based diversity scores. A common reproducing-kernel Hilbert-space (RKHS) definition of *MMD* between distributions P and Q is:

$$MMD^2(P,Q) = \left\| E_{x\sim P}[\phi(x)] - E_{y\sim Q}[\phi(y)] \right\|_{\mathrm{H}}^2 \tag{11}$$

where $\phi(\cdot)$ denotes the feature mapping into RKHS H. In materials science, however, the most practical criterion is often task-oriented evaluation, namely, whether the generated data enhance the predictive accuracy, robustness, or extrapolation capability of downstream ML models. Although GANs hold significant promise for augmenting high-cost datasets (e.g., γ′ phase coarsening, creep, fatigue, oxidation, and phase-transformation kinetics), their outputs may violate physicochemical constraints if trained without appropriate regularization. Consequently, physics-informed or constraint-embedded GAN frameworks have emerged as essential directions for ensuring both fidelity and physical validity. Overall, GAN-based data generation provides a powerful route to mitigating the small-sample bottleneck in advanced materials research and enabling more reliable exploration of composition-structure-property relationships in regimes where experimental data are limited or prohibitively costly to obtain.

Markov Chain Monte Carlo (MCMC) provides a complementary, fully probabilistic approach for synthesizing medium-fidelity materials data by sampling directly from a physically meaningful target distribution $p(x)$. MCMC constructs a Markov chain with transition kernel $T(x\to x')$ such that $p(x)$ is its stationary distribution, typically enforced through the detailed-balance condition:

$$p(x)P(x\to x') = p(x')P(x'\to x) \tag{12}$$

Representative algorithms include Metropolis-Hastings (MH), Gibbs sampling, and Hamiltonian Monte Carlo (HMC). In MH sampling, a candidate point x′ drawn from

$q(x'|x)$ is accepted with probability:

$$\alpha(x,x') = \min\left(1, \frac{p(x')q(x|x')}{p(x)q(x'|x)}\right) \tag{13}$$

Gibbs sampling sequentially updates each variable by drawing from its full conditional distribution, whereas HMC introduces auxiliary momentum variables and simulates Hamiltonian dynamics to perform long-range, low-autocorrelation moves that significantly reduce random-walk behavior. Convergence diagnostics, such as trace-plot inspection, the Gelman-Rubin statistic, effective sample size, and autocorrelation analysis, are routinely employed to ensure sampling reliability. Compared with GANs, MCMC offers stronger interpretability, explicit incorporation of physicochemical priors, and strict adherence to thermodynamic constraints. Thus, MCMC serves as a critical component of multi-fidelity data-augmentation workflows, complementing GAN-based generative models and substantially improving the physical consistency and generalization performance of ML models under small-sample regimes.

Second, transfer learning (TL) provides another highly effective pathway for small-sample materials modeling. By transferring knowledge learned from data-rich domains (e.g., tensile strength, oxidation kinetics) to data-poor domains (e.g., creep rupture life, fatigue performance), TL can substantially enhance predictive accuracy while reducing reliance on costly experimental datasets. Gao et al. [44] proposed a transfer learning Transformer-KAN (TLT-KAN) framework integrating Transformer encoders and Kolmogorov-Arnold Networks (KAN) to predict Ni-based superalloy residual fatigue damage under multi-level loading with limited samples. To address the dual limitations of physics-based models (inaccuracies caused by over simplification) and conventional data-driven methods (requiring sufficient data), TLT-KAN leverages simulation data from the validated Manson-Halford physics-based model as the source domain for pre-training, transferring knowledge by fine-tuning the model on scarce experimental data (target domain). Validation on a dataset comprising 14 materials demonstrates that TLT-KAN achieves competitive accuracy using only 8.3 % of the training data required by conventional ML models; when the training data increases to 16.7 % of that needed by conventional ML models, its accuracy surpasses all ML models. Ablation studies confirm the critical

roles of the Transformer encoder's feature interaction and KAN's adaptive nonlinear mapping, and the model exhibits exceptional data efficiency, enabling high-accuracy fatigue prediction under scenarios of limited experimental data.

The concept of TL is inspired by the human ability to apply previously acquired knowledge to solve new tasks with greater speed and accuracy [45], as illustrated in Fig. 5. TL is particularly advantageous when experimental data available for the target task are insufficient or the computational costs are prohibitively high [22]. By leveraging transferable knowledge from a source domain, TL enhances model performance in the target domain [46].

TL methodologies can be broadly categorized into model-based, sample-based, and feature-based strategies [47]. Model-based TL involves sharing components of a pre-trained architecture and fine-tuning it for the target task, which is the mainstream approach. In neural networks, shallow layers are typically retained for feature extraction, capturing fundamental patterns such as edges, textures, or spatial motifs. Freezing these early layers and retraining only the final one or a few layers using a small learning rate significantly reduces the need for labeled data and accelerates convergence, as shown in Fig. 6.

Sample-based TL requires selecting samples from the source domain that are highly similar to the target domain for utilization. This approach is straightforward and does not require additional model training or feature extraction. When the feature spaces of the source and target domains differ, feature-based TL, also known as marginal distribution adaptation, becomes necessary. This approach transfers knowledge by aligning feature representations between the source and target domains. The primary challenge in feature-based TL lies in addressing discrepancies in the marginal distribution of input data or features. By focusing on shared feature representations, feature-based TL provides an effective strategy for knowledge transfer across domains.

Fig. 7 presents the specific machine learning models employed for various target properties of Co- and Ni-based superalloys. To further promote broad engagement in the field of materials informatics and advance the research of high-temperature alloys, an array of integrated informatics platforms and software tools has emerged. These platforms are designed to assist researchers in efficiently developing models and accelerating the discovery of novel alloy systems, as summarized in Table 2.

2.5. Model evaluation

Generally, model performance in regression tasks is quantitatively evaluated using five standard evaluation metrics: coefficient of determination (R^2), mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE) and symmetric mean absolute percentage error (SMAPE), defined as follows [59]:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{14}$$

$$MAE = \frac{1}{n}\sum_{i=0}^{n-1}|y_i - \hat{y}_i| \tag{15}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{16}$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{17}$$

$$SMAPE = \frac{100\%}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{(|y_i| + |\hat{y}_i|)/2} \tag{18}$$

where y_i, $\hat{y}_i$, and $\bar{y}$ represent the actual value, predicted value, and mean of the actual values for the i-th sample, respectively, and n denotes the total number of samples.

R^2 measures the proportion of variance in the target variable explained by the model, with values approaching 1 indicating improved goodness of fit. MAE and RMSE metrics provide complementary perspectives on prediction accuracy. MAE, computed as the mean of absolute residuals, provides a robust estimate that is less affected by extreme values. RMSE applies quadratic penalization to residuals and is therefore more sensitive to large deviations, making it particularly useful for identifying substantial prediction errors. Both metrics decrease as predicted values more closely match experimental observations and are widely adopted for evaluating regression models.

MAPE quantifies the average percentage deviation between predicted and actual values, offering a scale-independent and interpretable measure of prediction accuracy. However, its formulation involving division by the ground truth renders it numerically unstable when actual

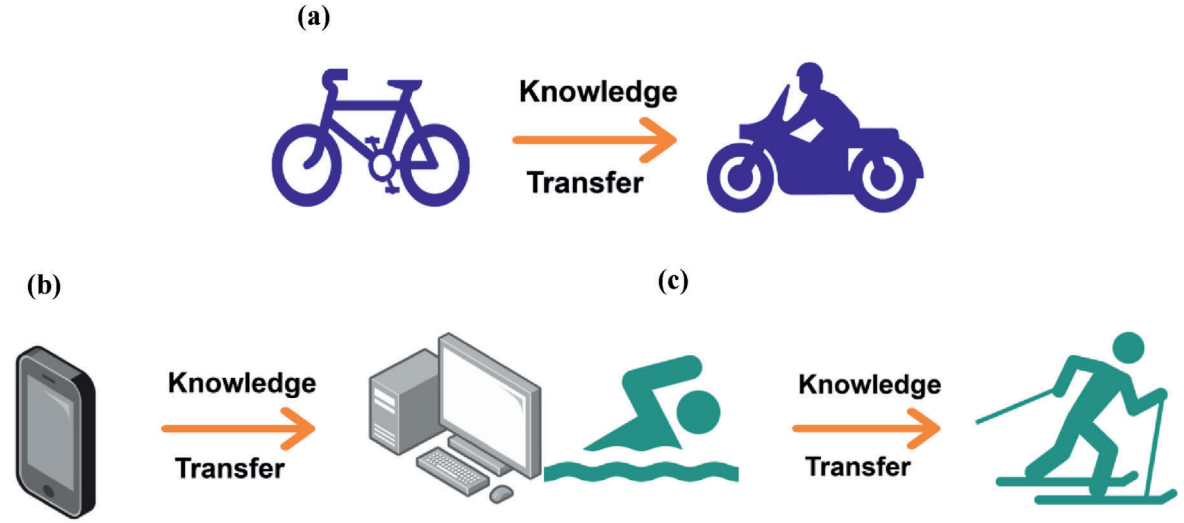

Fig. 5. Intuitive examples about TL. (a) Applying the knowledge of riding a bicycle to learning how to ride a motorcycle. (b) Utilizing the skills gained from using a smartphone to efficiently operate a computer. (c) Transferring swimming techniques to learning how to ski.

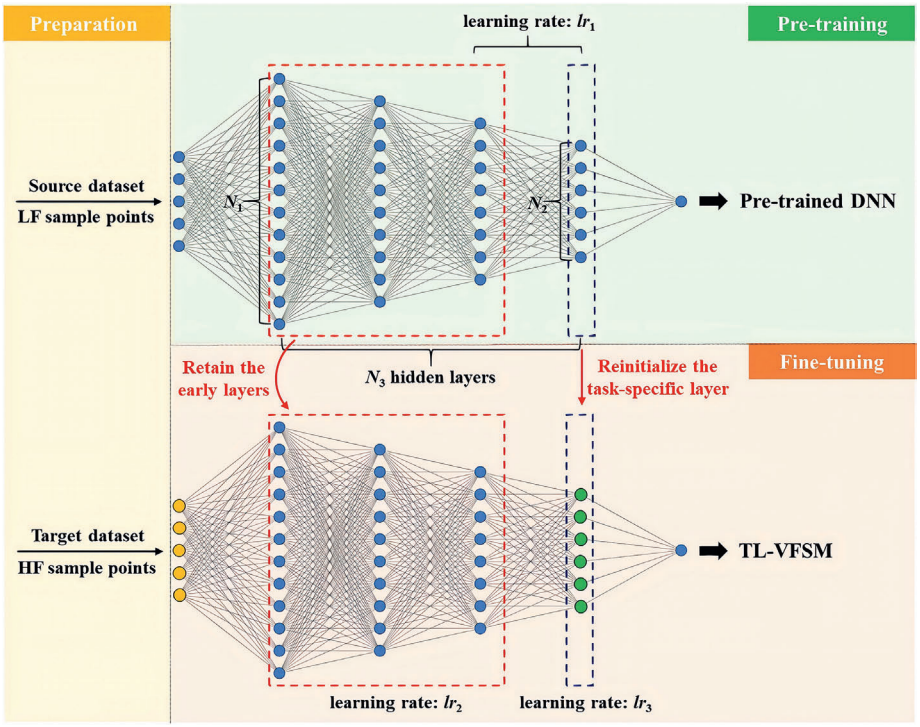

Fig. 6. Variable-fidelity surrogate model based on TL [48].

values approach zero. SMAPE mitigates this issue by replacing the denominator with the average of predicted and actual values, yielding a symmetric and more numerically stable estimate of relative error. This makes SMAPE especially advantageous for datasets containing heterogeneous magnitudes or near-zero targets.

In classification tasks, accuracy (ACC), recall (REC), precision (P), and the F1-score constitute the fundamental metrics for evaluating discriminative performance. Among these, the F1-score is particularly informative under class-imbalanced conditions because it balances precision and recall. The receiver operating characteristic (ROC) curve and its associated area under the curve (AUC) provide a threshold-independent measure of class separability and are applicable to both binary and multi-class classification problems. The definitions of the classification model evaluation metrics are provided as follows [60]:

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} \tag{19}$$

$$P = \frac{TP}{TP + FP} \tag{20}$$

$$REC = \frac{TP}{TP + FN} \tag{21}$$

$$F1score = \frac{2TP}{2TP + FP + FN} \tag{22}$$

where *TP* (True Positive) denotes the number of correctly predicted positive samples; FP (False Positive) represents the number of negative samples incorrectly predicted as positive; TN (True Negative) corresponds to the number of correctly predicted negative samples; and FN (False Negative) indicates the number of positive samples incorrectly predicted as negative.

Although many models achieve seemingly excellent fitting performance on the training dataset, such results provide only limited insight into their true predictive reliability. For superalloy applications, model evaluation should therefore be based on rigorously separated validation schemes, such as k-fold or nested cross-validation, repeated random subsampling, and, ideally, independent test sets, to obtain unbiased estimates on genuinely unseen data and to detect overfitting that is often hidden in small, noisy, or imbalanced datasets. In practice, robust generalization further requires careful control of data leakage (e.g., avoiding splitting data points originating from the same alloy or heat in different subsets), transparent reporting of evaluation protocols, and complementary uncertainty quantification to indicate the confidence level of each prediction. A distinct but equally important issue is generalization across alloy systems. Models trained exclusively on a single chemistry window, for example, Co-based γ′-strengthened superalloys, rarely transfer directly to Ni-based or multi-principal systems because of substantial shifts in composition space, phase-stability landscapes, diffusion kinetics, and deformation or damage mechanisms. Consequently, naïve extrapolation beyond the training domain can lead to deceptively confident yet physically unreliable predictions. Recent studies [61,62] have begun to mitigate this gap through strategies such as transfer learning (pretraining on broad, cross-system datasets followed by system-specific fine-tuning), domain adaptation and reweighting of training samples, multi-fidelity or physics-informed descriptors (e.g., CALPHAD-derived thermodynamic metrics, diffusion coefficients, and γ′ phase attributes), and multi-task frameworks that learn shared representations across related properties or alloy families. Nevertheless, systematic benchmarks on cross-system generalization remain scarce, underscoring the need for community-wide, standardized testbeds to evaluate how far current ML models can reliably extrapolate between Co-based, Ni-based, and emerging superalloy systems.

2.6. Model interpretation

Most conventional ML models operate as black boxes, providing limited transparency regarding how input features influence predicted outcomes. This lack of interpretability poses a fundamental challenge for data-driven superalloy design, where understanding mechanistic trends is essential for informed alloy development and scientifically grounded decision-making. The advent of explainable ML (XML) techniques has substantially alleviated this limitation by enabling quantitative and human-interpretable attribution of model behavior. Among available XML methods, SHapley Additive exPlanations (SHAP), local

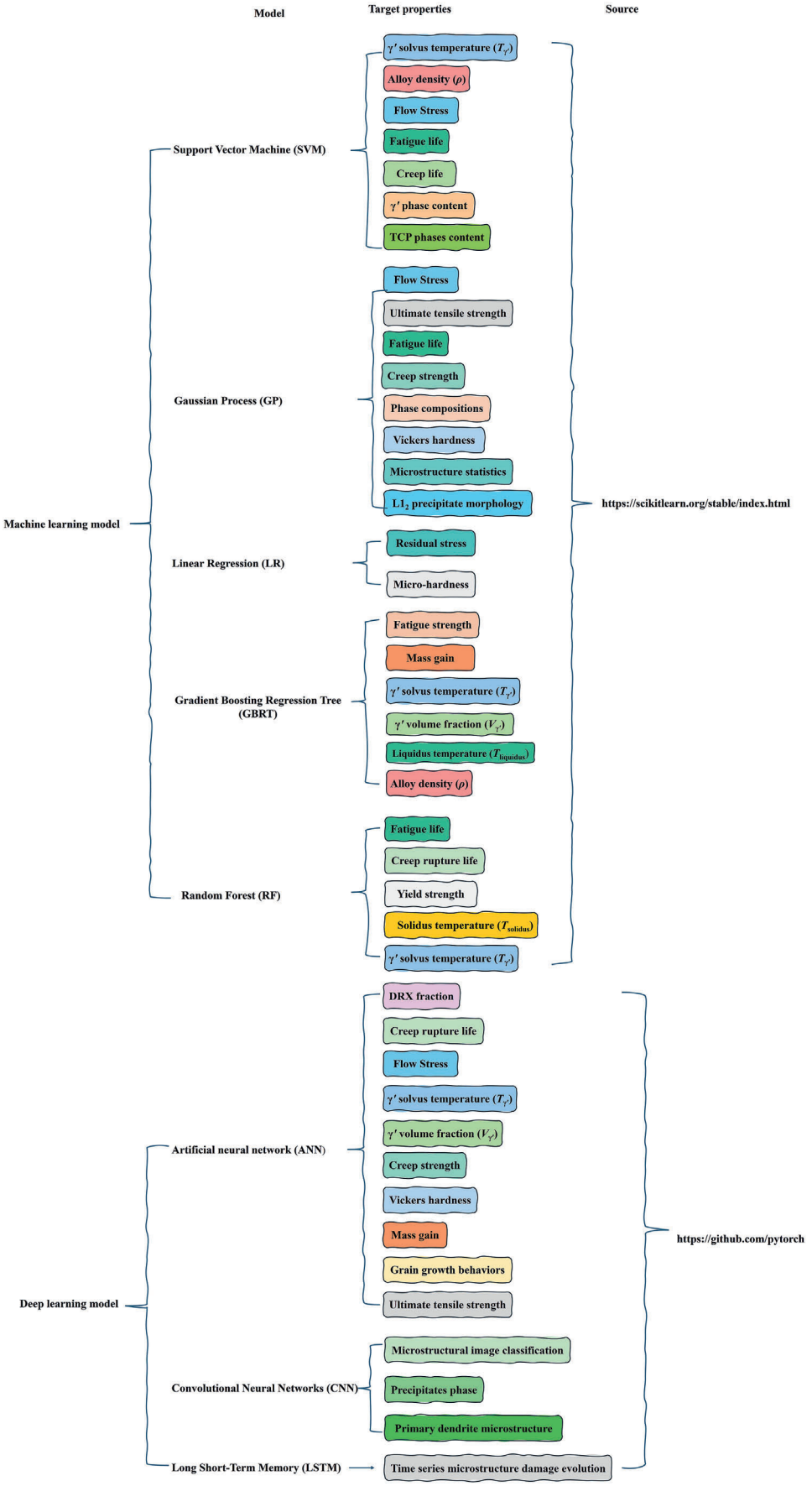

Fig. 7. Commonly used ML models to target properties in Co/Ni-based superalloys.

Table 2
All-in-one materials informatics platforms/software.

Platform/software	Function	Source
MLMD [49]	A programming-free AI platform to predict and design materials	http://matdesign.top/
MatCloud [50]	A high-throughput computational infrastructure for integrated management of materials simulation, data and resources	https://www.matcloudplus.com.cn/
AlphaMat [51]	A material informatics hub connecting data, features, models and applications	http://www.aimslab.cn/#/alphamat
Python Materials Genomics (Pymatgen) [52]	A robust, open-source python library for materials analysis	https://pymatgen.org/
MaterialsAtlas.org [53]	A materials informatics web app platform for materials discovery and survey of state-of-the-art	http://www.materialsatlas.org/
Joint automated repository for various integrated simulations (JARVIS) [54]	An integrated infrastructure to accelerate materials discovery and design using density functional theory (DFT), classical force-fields (FF), and machine learning (ML) techniques	https://jarvis.nist.gov/
Jilin Artificial-intelligence aided Materials-design Integrated Package (JAMIP) [55]	An open-source Python framework to meet the research requirements of computational materials informatics	http://www.jamip-code.com/
Materials Image Processing and Automated Reconstruction (MIPARTM) [56]	A novel software package for two- and three-dimensional microstructural characterization	https://www.mipar.us/
Artificial Learning and Knowledge Enhanced Materials Informatics Engineering (ALKEMIE) [57,58]	An intelligent computational platform for accelerating materials discovery and design	https://alkemine.org
Crystals.ai	An software frameworks for robust AI in materials science	https://crystals.ai/

interpretable model-agnostic explanations (LIME), and symbolic regression (SR) have been widely adopted in the superalloy community [59,63,64].

SHAP analysis, in particular, provides a unified, game-theoretic framework for decomposing model outputs into additive feature contributions while capturing nonlinear interactions. These capabilities make SHAP especially effective for elucidating the complex composition-processing-microstructure-property relationships characteristic of Co- and Ni-based superalloys.

For example, Sun et al. [59] reported that when age temperature (T_{age}) is below 1050 °C and the compositional window satisfies V > 0 at.%, Ti < 3.0 at.%, Cr > 3.0 at.%, and Mo between 1 and 5 at.%, the SHAP values of the γ' coarsening-rate model remain predominantly positive, indicating a significant reduction in the γ' coarsening rate constant (K_r) under these thermochemical conditions. Notably, T_{age} exerts the strongest influence on K_r. Conversely, when $T_{age} \leq 1050$ °C, Ti ≥ 3.0 at.%, W > 2.5 at.%, Co ≤ 55 at.%, Ta ≥1 at.%, Cr ≤ 2 at.%, V ≥ 1 at.% and Mo within 2–5 at.%, the γ' volume-fraction model exhibits consistently positive SHAP values, suggesting a pronounced enhancement in γ' phase volume fraction. These findings illustrate that XML tools not only identify dominant factors but also reveal intricate feature interactions that frequently elude traditional metallurgical intuition.

A concise formulation of SHAP is provided below. The SHAP value for the i-th feature is defined as:

$$\varphi_i = \sum_{s \subseteq F \setminus \{i\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} \left[f_{S \cup \{i\}}(x_{S \cup \{i\}}) - f_S(x_S) \right] \tag{23}$$

where F represents the set of all features, S is a subset of F, and $S \cup \{i\}$ denotes the union of subset S and the i-th feature. Here, $f_{S \cup \{i\}}(x_{S \cup \{i\}})$ is the model's prediction when the i-th feature is included, while $f_S(x_S)$ is the prediction of the model trained without the i-th feature.

The final prediction is obtained by aggregating the contributions of all features:

$$\hat{y} = \hat{y}_0 + \sum_{i=1}^{n} \varphi_i \tag{24}$$

where $\hat{y}$ is the predicted value and $\hat{y}_0$ represents the model's prediction without any features. Eqs. (23) and (24) are applied to each data instance, meaning that SHAP values vary across different samples. The mean absolute value of the SHAP values for each feature can be computed to assess its overall impact on the targeted property.

Beyond quantifying individual feature contributions, SHAP also captures feature interactions through SHAP interaction values, which measure the interaction effects between two features. The formula is defined as follows:

$$\varphi_{i,j} = \sum_{s \subseteq F \setminus \{i,j\}} \frac{|S|!(|F| - |S| - 2)!}{(|F| - 1)!} \left[f_{S \cup \{i,j\}}(x_{S \cup \{i,j\}}) - f_{S \cup \{i\}}(x_{S \cup \{i\}}) - f_{S \cup \{j\}}(x_{S \cup \{j\}}) + f_S(x_S) \right]. \tag{25}$$

where represents the interaction effect between features i and j. Here, F is the set of all features, and S is any subset that excludes i and j. The term $f_S(x_S)$ denotes the model's prediction when using only the feature subset S, while $f_{S \cup \{i\}}(x_{S \cup \{i\}})$ and $f_{S \cup \{j\}}(x_{S \cup \{j\}})$ represent the model's prediction when feature i or feature j is added, respectively. Finally, $f_{S \cup \{i,j\}}(x_{S \cup \{i,j\}})$ denotes the model's prediction when both features i and j are included. This formulation quantifies the incremental contribution of features i and j when considered jointly, relative to their individual effects. A positive $\varphi_{i,j}$ indicates a synergistic effect, suggesting that the two features reinforce each other's influence on the prediction. Conversely, a negative $\varphi_{i,j}$ suggests an antagonistic effect, where the two features counteract each other's impact.

Beyond model-specific explanations, XML techniques can also be leveraged to interrogate the data itself. Symbolic regression (SR), particularly when implemented via genetic programming, identifies explicit, closed-form equations that map alloy descriptors to target properties. Because the resulting expressions remain physically interpretable, SR offers a promising route to uncover mechanistic relationships and guide rational alloy design. For instance, Hansen [65] employed genetic-programming-based SR to predict fatigue-indicator parameters (FIPs) in additively manufactured IN625. The accurately predicted FIPs enabled reliable estimation of microstructurally short-crack initiation and growth. By coupling SR-derived FIPs with long-crack growth models, the authors established a unified framework for efficiently predicting total fatigue life of IN625 superalloys.

Collectively, these advances underscore that interpretability is not merely an auxiliary attribute but a central pillar of modern, trustworthy, data-driven superalloy design. XML techniques bridge the gap between predictive performance and scientific understanding, allowing ML models to serve not only as computational tools but also as engines for discovering mechanistic principles and accelerating the development of next-generation superalloys.

2.7. Virtual space

Machine learning models are often trained on small-sample datasets, typically ranging from a few hundred to several thousand data points, and subsequently employed to predict outcomes in substantially larger virtual datasets containing hundreds of thousands or even millions of samples. While models may exhibit exceptional performance on small training datasets, their predictive accuracy within the virtual sample

space remains uncertain, primarily due to potential discrepancies in the alignment and consistency of sample distributions between the training and virtual datasets. In the context of high-temperature alloys, virtual sample spaces are commonly constructed using domain knowledge or expertise, with predefined ranges for elemental compositions and processing parameters. This approach enables the systematic exploration of optimal compositions and processing conditions. However, the sample distribution in the virtual space may diverge significantly from that of the original experimental dataset. Such distributional disparities can compromise the reliability of model predictions, even when the model demonstrates strong performance on the training dataset. A practical strategy to assess the reliability of model predictions in the virtual space involves experimentally validating a subset of alloy compositions. If the experimental results closely correspond to the model predictions, the model's predictive capability in the virtual space can be considered credible. Nonetheless, this approach inevitably incurs additional experimental costs and effort. The primary challenge, therefore, lies in designing the virtual space to closely approximate the sample distribution of the training dataset. When the distributions are well-aligned, a model that performs effectively on the small training dataset is more likely to produce reliable predictions for the larger dataset. Through a review of the literature, it was found that numerous scholars have summarized various methods for virtual sample generation, including SMOTE, Monte Carlo, and Generative Adversarial Networks et al. [35, 66,67]. These methods have been described in detail in terms of their characteristics and the specific application scenarios for which they are most suitable. Therefore, a detailed introduction to these methods is not provided here.

2.8. Optimization algorithm

In general, the unexplored virtual space is prohibitively large, making it challenging for a trained ML model to efficiently identify an optimal solution. Consequently, optimization methods are typically integrated to facilitate the search process, wherein the established forward surrogate model serves as an evaluation solver. The genetic algorithm (GA) is a widely adopted optimization technique used as a global strategy for exploiting promising superalloys [68,69]. Inspired by Darwin's theory of natural selection, GA converges to an optimal solution by evolving successive generations through a series of evolutionary operations, including selection and genetic operators. Specifically, genetic operators such as crossover and mutation conduct stochastic searches within the solution space, while the selection operator directs the search trajectory. This combination enables GA to effectively balance exploration and exploitation [70]. In addition to GA, Bayesian optimization (BO) is extensively employed in the optimization design of superalloys [71–73]. As a robust global optimization framework, BO excels in tackling complex problems where function evaluations are prohibitively expensive or infeasible through direct computation or predictive modeling. Its core methodology involves constructing a probabilistic surrogate model, typically a Gaussian process, to quantify uncertainty in the objective function. This surrogate model facilitates informed exploration of the solution space by strategically balancing exploration and exploitation through acquisition functions such as Expected Improvement or Upper Confidence Bound. Consequently, BO efficiently navigates the vast and intricate parameter landscapes inherent in alloy design, enabling the identification of optimal compositions and processing conditions with reduced computational overhead.

3. Specific performance issues of ML applications in superalloys

ML can analyze and mine extensive experimental and computational data to establish complex mapping relationships between composition, processing conditions, microstructure, and properties. This approach allows for the prediction of novel alloys with specific properties in unexplored spaces, while simultaneously identifying the optimal compositions and processing conditions corresponding to superior properties. ML has thus become a widely adopted and effective strategy for advancing superalloy research and development. This section reviews recent progress in ML applications for superalloy performance prediction, mainly focusing on γ' phase high-temperature stability, high-temperature creep property, fatigue behavior, and oxidation resistance aspects.

3.1. High-temperature stability of the γ' phase

The high-temperature stability of the γ' phase, as the primary strengthening phase, is directly related to the structural and property stability of superalloys under elevated temperature service conditions. Structural stability under such conditions requires two critical factors: first, the γ/γ' two-phase microstructure must remain stable with minimal formation of detrimental secondary phases; second, the evolution of γ' phase in terms of volume fraction, size, and morphology occurs gradually to mitigate the degradation of high-temperature mechanical properties [63].

Traditional experimental characterization techniques face challenges such as long testing cycles, high costs, and complex influencing factors that introduce data uncertainties. With the ongoing shift toward the fourth paradigm of data-driven science and the fifth paradigm of AI for Science, the integration of materials informatics and ML has emerged as the optimal solution for designing and developing superalloys systematically and efficiently. In this context, researchers have extensively investigated the phase composition, stability, and coarsening behavior of precipitate phases. Yu et al. [74] employed 458 experimental data samples to develop a random forest classification model for predicting whether an alloy contains only the γ/γ' two-phase microstructure. The model achieved a prediction accuracy exceeding 95 %, enabling the design of a series of Co-Ti-V-based quaternary alloys with stable γ/γ' phases. Xi et al. [12] proposed a methodology integrating ML with first-principles calculations to predict the stability of the $L1_2$ phase in Co-V-Ta and Co-Al-V systems. Their results showed that the ML approach was more than four times as efficient as relying solely on first-principles calculations. Qin et al. [75] constructed a high-throughput experimental database comprising 33,484 Ni-based superalloy samples with detailed composition and microstructural data. Using the UNet3+ architecture for image recognition, they developed a high-precision backpropagation neural network (BPNN) model to predict the γ' phase volume fraction, average size, and size distribution. This model effectively identified γ', μ, σ, and η phases in Ni-based superalloys, and accurately predicted the morphology of the γ' phase. Additionally, it captured the distribution of alloying elements in various precipitate phases, providing insights into the effects of alloying elements on the microstructural morphology of Ni-based superalloys. Liu et al. [76] employed convolutional neural networks (CNNs) to develop an image segmentation model for the γ/γ' phase microstructure in superalloys. By utilizing the OpenCV ML vision library, they extracted detailed microstructural information related to the γ/γ' phases. A regression model was subsequently constructed to correlate microstructural electronic-physical properties with the γ' phase size, and the results were experimentally validated. Their findings revealed that elements with high Young's modulus, such as Re, Ru, W, and Mo, tend to segregate into the γ matrix or distribute relatively uniformly across the γ/γ' phases. This behavior increases the lattice constants a_γ and $a_{\gamma'}$, thereby enlarging the denominator of the lattice misfit δ without significantly altering the numerator. As a result, the misfit is reduced, lowering the elastic stress between the two phases and ultimately yielding a smaller γ' phase size. The observed coarsening behavior aligned closely with fitting from the Lifshitz-Slyozov-Wagner (LSW) theory [77]. The workflow of computer vision framework using deep learning-based CNNS to extract accurate microstructural information from γ/γ' microstructure images for the γ' coarsening analysis in kinetic theory and machine learning models is illustrated in Fig. 8. Addressing

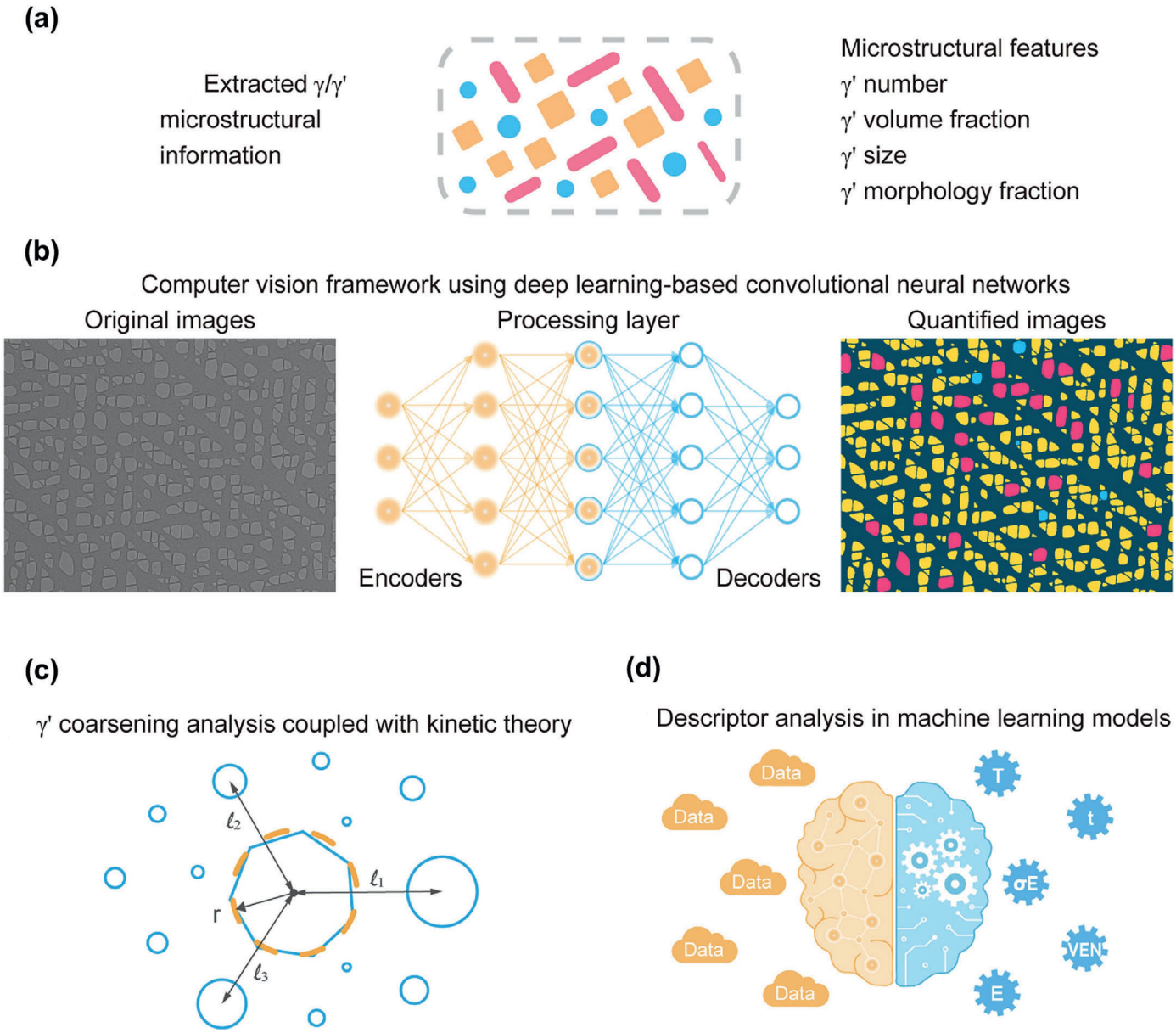

Fig. 8. The workflow of computer vision framework using deep learning-based convolutional neural networks to extract accurate microstructural information from γ/γ′ microstructure images for the γ′ coarsening analysis in kinetic theory and ML models (adapted from Ref. [76]). (a) The computer vision framework using deep learning-based convolutional neural networks to segment and quantify γ/γ′ microstructure images. (b) Extracted γ/γ′ microstructural features including γ′ number, γ′ volume fraction, γ′ size and γ′ morphology fraction. (c) γ′ coarsening analysis to compare with classical kinetic theory. (d) Descriptor analysis to construct ML models and explicit dimensionless relations using symbolic regression.

the stability of the γ/γ′ two-phase microstructure and the precipitation of harmful secondary phases during high-temperature service is essential for improving superalloy performance. By leveraging ML to optimize the regulation of alloying elements and uncover the underlying mechanisms, is of significant research value for enhancing the high-temperature service capabilities of superalloys.

3.2. High-temperature creep property prediction

Exceptional creep resistance is among the most critical metrics in the design, materials selection, and safety evaluation of high-temperature components. Under prolonged exposure to elevated temperatures and stresses, superalloys inevitably undergo creep deformation. From a microstructural perspective, this process manifests as γ′ phase dissolution, coarsening, and rafting, all of which contribute to microstructural degradation [78,79]. Such changes result in a marked deterioration in the mechanical properties of superalloys, thereby compromising their operational safety and service life. The relationship between microstructural evolution during creep and the corresponding macroscopic properties is highly complex and nonlinear, posing a significant challenge to traditional experimental methodologies and linear modeling approaches, which struggle to comprehensively and accurately capture these intrinsic mechanisms. Addressing this challenge and establishing a precise mapping relationship between microstructural evolution and creep performance remain pivotal issues in superalloy research. Furthermore, the creep rupture life of superalloys is a decisive factor in determining whether components can endure prolonged service in specific operational environments. However, the extended duration and high cost associated with creep testing have rendered traditional trial-and-error approaches increasingly unsuitable for novel materials design.

Recent advances in AI for Science have introduced powerful tools

and innovative methodologies for investigating the creep behavior of superalloys. By leveraging data-driven ML techniques, can effectively integrates multi-dimensional data, encompassing composition, processing, microstructure, and performance, to construct nonlinear predictive models of creep behavior. These models significantly enhance research efficiency and predictive accuracy.

Xu et al. [80] developed quantitative models to elucidate the creep-induced microstructural evolution of a single-crystal nickel-based superalloy containing Re and Ru, utilizing ML method that integrate physical and statistical microstructural features. High-temperature

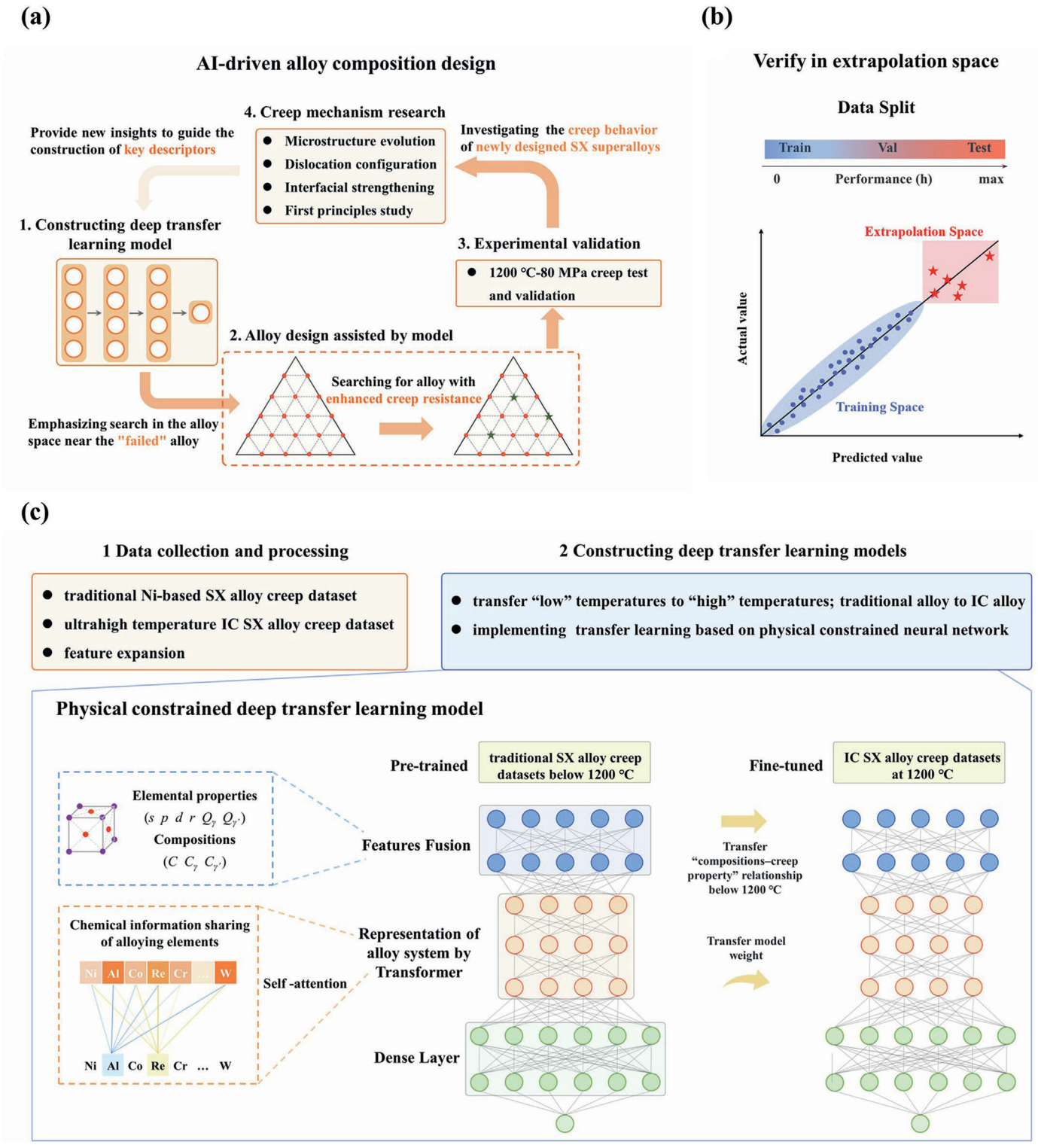

Fig. 9. Deep learning model construction strategy for improving extrapolation efficiency and alloy design flow (adapted from Ref. [82]). (a) The whole alloy design process including deep transfer learning model construction, composition screening, alloy design, experimental validation, and creep mechanism study at ultrahigh temperature. (b) The data split strategy, data with higher performance is utilized as a test set to evaluate the extrapolation efficiency of the model. (c) The deep learning model construction method based on transfer learning in A, the necessary dataset collection and processing steps are labeled, and the key feature fusion and element interaction representation modules involved in the model are shown in the illustrations.

creep tests were performed under multiple conditions on high-throughput specimens, enabling the continuous quantification of critical physical features of the microstructure, including the volume fraction, rafting degree, and raft thickness of γ' precipitates across eight distinct samples. To enhance the comprehensiveness of the microstructural representation, statistical features were incorporated through two-point correlation analysis and principal component analysis (PCA), thereby complementing the physical features with greater specificity. Two neural network-based ML models were subsequently constructed: one model predicted the microstructural features based on creep conditions, while the other inferred the creep conditions from specified microstructural features. Validation experiments confirmed the robust predictive performance of both models, enabling bidirectional prediction of the microstructure-performance relationship in single-crystal superalloys, thus providing a novel quantitative framework for understanding high-temperature creep behavior. Zhou et al. [81] proposed a simplified yet accurate and generalizable surrogate model for creep rupture life predictions for Ni-based single crystal superalloys using an automated ML algorithm implemented in AutoGluon. Remarkably, the model achieves high prediction accuracy on independent test data without integrating microstructural information or deformation mechanisms into the ML framework. The exclusion of microstructural and deformation mechanism details not only broadens the model's applicability across a vast, unexplored composition and processing space but also significantly facilitates the efficient inverse design of novel alloys. Yang et al. [82] introduced a physically constrained deep transfer learning framework, SaTNC_FT, illustrated in Fig. 9, to address two critical challenges in the predictive modeling of Ni-based single crystal (SX) superalloys at ultrahigh temperatures ($\geq$1200 °C): the scarcity of experimental data and the limited generalization capability of ML models in extrapolation spaces. Through the implementation of transfer learning, the framework effectively leveraged prior compositional sensitivity information from pre-trained models to accurately capture the complex, nonlinear relationships between alloy composition and creep rupture life at ultrahigh temperature, even when relying on a limited dataset. This innovative approach enabled the successful design of an advanced SX superalloy with a verified creep rupture life of approximately 170 h at 1200 °C and 80 MPa, achieving a 30 % improvement over the current state-of-the-art. Furthermore, the study elucidated the γ/γ' interface strengthening mechanism as the primary factor contributing to the enhanced creep performance, offering valuable insights into compositional optimization strategies. This proposed framework represents a significant advancement in the field, providing a robust, scalable, and cost-efficient methodology for the development of high-performance materials, while establishing a foundation for addressing broader challenges in alloy design. Ma et al. [83] proposed a back propagation neural networks (BPNN) model to predict the creep curves of MarM247LC superalloy under various conditions. Utilizing six creep curves for training, the model achieved prediction errors within $\pm$20 %, demonstrating its accuracy and reliability. Compared to the traditional θ projection model, the BPNN approach reduced the maximum error by 30 %, underscoring its superior predictive performance. Furthermore, the applicability of this method was validated on other superalloys, including DZ125 and CMSX-4, confirming its versatility and potential for widespread use in modeling the creep behavior of diverse superalloy systems.

Previous sections have primarily highlighted domestic advances in applying machine-learning techniques to predict and analyze the creep behavior of high-temperature superalloys. In recent years, several international researchers have also contributed to this field. For example, Hossain et al. [84] employed a multi-gene genetic programming algorithm to uncover an explicit mathematical relationship linking stress, temperature, intrinsic material properties, chemical composition, and creep-rupture life in Alloy 617. Brandon et al. [85] developed a machine-learning model capable of predicting the steady-state creep strain rate of γ'-strengthened Co-based superalloys as a function of temperature and applied stress, using alloy composition, heat-treatment parameters, and γ' precipitate volume fraction as model inputs. Overall, however, international studies remain limited, and most foreign research on creep behavior still relies predominantly on conventional experimental trial-and-error approaches rather than comprehensive data-driven frameworks.

In summary, ML provides a powerful tool for establishing precise mappings between input features and output performance metrics, enabling effective prediction and design of creep behavior in high-temperature alloys. By addressing risks associated with performance degradation under service conditions, ML-driven approaches offer a viable alternative to conventional methods. Nevertheless, the majority of ML applications in superalloy creep performance research, whether focused on microstructural evolution, rupture life, or creep curves, have predominantly concentrated on Ni-based superalloys due to the relative abundance of experimental and industrial data. In contrast, emerging alloys, such as Co-based superalloys, remain significantly underexplored. Leveraging creep data from Ni-based alloys as source data and experimental data from Co-based alloys as target data within a TL framework presents a promising pathway for advancing the understanding and predictive modeling of creep performance in Co-based superalloys, particularly in scenarios constrained by limited data availability.

3.3. Fatigue behavior

Assessing the fatigue life and reliability of high-temperature alloy structural materials is crucial, given their extensive use in demanding and extreme environments. The fatigue life of superalloys is governed by a diverse array of factors, including applied cyclic stress or strain, geometric design, cyclic strain rate, surface conditions (e.g., surface finish and treatment), environmental influences (e.g., high temperatures and corrosive atmospheres), manufacturing defects (e.g., keyholes, gas pores, and lack of fusion defects), microstructural characteristics, residual stresses, and heat treatment processes [86,87]. A key focus of fatigue research is the systematic investigation of these variables to ensure optimal performance in structural applications and to inform the selection of appropriate processing technologies tailored to specific operational requirements.

Traditional theoretical models for predicting the fatigue life of high-temperature alloys often emphasize the impact of a single influencing factor. This narrow focus poses challenges in capturing the intricate interactions among multiple variables that collectively drive fatigue failure. The development of comprehensive predictive approaches capable of addressing these multifactorial interactions is imperative for advancing the understanding and practical application of fatigue-resistant superalloys. The study of fatigue crack propagation is also a crucial aspect of high-temperature alloy fatigue performance, as crack growth often leads to catastrophic failure of the material. As such, understanding and predicting fatigue crack growth behavior is vital for assessing the durability of high-temperature alloys under cyclic loading conditions.

Recent literature has highlighted that, ML, as a data-driven solution, can overcome some of the limitations of existing model, by uncovering implicit relationships within data. Several studies have employed ML techniques to explore the fatigue properties of high-temperature alloys. For instance, Xu et al. [88] used chemical composition, heat treatment processes, and experimental parameters as input features, with the low-cycle fatigue life of nickel-based superalloys as the target variable. They applied random forest (RF), fully connected neural networks (FCNN), support vector machines (SVM), and a genetic algorithm-based random forest (GA-RF) model to predict fatigue life. The GA-RF model outperformed the others, delivering results comparable to the Coffin-Manson model. Menasche et al. [89] utilized convolutional neural network (CNN) image segmentation techniques to analyze fatigue crack images of Ni-based superalloys subjected to cyclic loading.

Using micro-computed tomography (μCT), they captured images of crack propagation following 26 in-situ fatigue tests. A dataset of crack images from a fatigue test conducted over 9500 cycles was selected for training deep learning models, including U-Net, FCN8, and FCN4s. The U-Net model achieved a Dice coefficient (IoU) of 0.995 ± 0.004, demonstrating excellent segmentation performance. Further analysis of crack images from 25 additional in-situ fatigue tests revealed that, as the number of fatigue cycles increased, the total crack area grew exponentially. This image processing approach provides valuable support to researchers studying high-temperature alloys, enabling efficient analysis of crack growth and propagation mechanisms.

In conclusion, ML holds substantial promise for uncovering complex relationships among multiple variables, thereby improving the prediction of fatigue properties. However, the availability of fatigue data for high-temperature alloys is often limited, and obtaining such data is both resource-intensive and costly. Overcoming the challenges associated with small sample sizes is crucial for advancing the application of ML in the study of fatigue properties.

3.4. High-temperature oxidation property

Superior oxidation resistance is a crucial performance criterion for high-temperature alloys during service. However, the regulation of multiple alloying elements introduces significant complexity to the high-temperature oxidation behavior of superalloys. The formation of continuous and stable oxide layers, such as Al_2O_3 and Cr_2O_3, on the alloy surface is essential for ensuring exceptional oxidation resistance at elevated temperatures [90–93]. Zhang et al. [91] observed that increasing Al content improves the oxidation resistance of Co-based superalloys at 800–900 °C while significantly reducing Ti diffusion. At 1000 °C, however, higher Al content promotes the diffusion of Ta, Ti, and W, resulting in the formation of Co(W, Mo)O_4 and $TaTiO_4$ oxides, which ultimately degrade the oxidation resistance of the alloy. The formation of Al_2O_3 oxide films is influenced by Al content and compositional changes within the γ' phase. Moreover, Al reduces the intrinsic stacking fault energy, thereby affecting the mechanical properties of the alloy [94]. The addition of Cr, meanwhile, tends to facilitate the formation of intermetallic compounds (TCP phases), which diminishes the stability of the γ' phase [95]. Interactions among alloying elements further influence their diffusion capacities and the formation energies of oxides, adding to the complexity of oxidation behavior and surface oxide layer formation in novel Co-based superalloys.

ML offers substantial advantages in modeling the intricate relationships between multiple variables and target properties. Kim et al. [96] utilized an artificial neural network (ANN) model to predict the high-temperature oxidation behavior of Ni-based superalloys. The researchers designed 62 multi-component alloys within the NiCoCrMoWAlTiTa system and conducted 20 cyclic oxidation experiments for each alloy at 850 °C, generating a total of 1240 experimental data points. Of these, 75 % were allocated for training and 25 % were reserved for validation. The input features for the ANN model included the cyclic oxidation cycles and the contents of alloying elements Co, Cr, Mo, W, Al, Ti, and Ta, while the output variable was the oxidation weight gain. The ANN model achieved a mean squared error of less than 0.01. Compared with the empirical response surface methodology (RSM), the ANN model demonstrated significantly higher predictive accuracy. For typical Ni-based alloys such as GTD-111, IN738LC, and CM247LC, whose experimental data were excluded from training, the ANN model predictions closely matched the experimental results, with deviations ranging from near 0 mg·cm-2 to below 0.15 mg·cm-2. In contrast, the RSM model exhibited substantial deviations, exceeding 0.9 mg·cm-2 for all three alloys, underscoring the superior predictive performance of the ANN model. Further analysis of the ANN model's predictions revealed that the addition of Mo and Ti increases oxidation weight gain, while the addition of Al and Ta reduces it, consistent with previous studies. By integrating gradient descent optimization algorithms, the model identified and recommended several new alloy compositions with high oxidation resistance. Nikhil et al. [97] constructed a dataset comprising over 2000 data points on the oxidation properties of Ni-based superalloys. Utilizing fully connected neural networks (FCNN) and extreme gradient boosting (XGBoost) algorithms, they developed high-accuracy ML models for predicting oxidation weight gain and oxidation rate constants. These models demonstrated exceptional precision and reliability in forecasting oxidation properties. Thus, employing ML to investigate the oxidation behavior of novel Co-based superalloys enables more scientifically informed regulation of Al and Cr contents while systematically uncovering the interactions between these elements and other alloying components. This approach offers a robust pathway to achieving precise control and reliable evaluation of the oxidation resistance of high-temperature alloys.

4. Bridging potential and practice: outlook and challenges in machine learning for superalloys

4.1. Strategic outlook

Multi-fidelity data sources. High-fidelity data derived from small-scale experimental studies, medium-fidelity data generated from larger-scale simulations, and low-fidelity data produced through virtual sampling technologies present substantial opportunities for advancing ML-based research. These multi-fidelity data sources offer a robust foundation for the discovery and optimization of high-temperature alloy systems exhibiting comprehensive and exceptional properties.

Interpretable ML models. ML models are often referred to as "black boxes" due to the difficulty in correlating their numerous complex parameters with established physical and chemical principles, rendering these models largely uninterpretable. In contrast, symbolic regression algorithms offer a tangible advantage by providing explicit mathematical expressions that elucidate the relationships between input and output variables within the model. Furthermore, Shapley Additive Explanations analysis can quantify the contributions of individual input features to the model's outputs, thereby offering a degree of interpretability. Together, these techniques enhance the transparency of ML models and facilitate a more robust understanding of their underlying mechanisms, which is particularly valuable in the study of superalloys.

Large language models (LLMs). The advent of LLMs has introduced transformative opportunities for the field of metallic material design, particularly in addressing the challenges posed by the multifaceted and interdisciplinary nature of this domain. Individual research teams often lack the comprehensive expertise required to navigate and integrate diverse aspects of material design effectively. Compounding this issue is the overwhelming volume of scientific publications, patents, conference abstracts, and other corpora, which are seldom organized to support holistic design approaches. Large language models, such as Deepseek-R1, ChatGPT-4o, Gemini, Llama, and Claude, have demonstrated exceptional potential in extracting and synthesizing information from extensive datasets. By leveraging these models, researchers can systematically process and distill insights from vast and fragmented corpora, thereby enabling a more integrated and efficient approach to material innovation. These capabilities position LLMs as valuable tools in the pursuit of advancing the design and development of high-performance metallic materials, including superalloys.

4.2. Challenges

Model confidence and reliability. Understanding the conditions under which a ML model provides accurate predictions and identifying scenarios where it may fail are critical challenges. Without such insights, reliance on ML predictions becomes precarious, potentially leading to erroneous conclusions or suboptimal decisions. This issue is particularly significant in the context of superalloy research, where the consequences of inaccurate predictions can profoundly affect material

performance and safety assessments.

A high-quality experimental dataset. The foundation of ML lies in the availability of high-quality experimental data. However, discrepancies in experimental conditions or gaps in data collection during testing can undermine the accuracy of subsequent ML models. Therefore, establishing comprehensive, high-quality and standardized experimental datasets for superalloys is crucial to advancing the application of ML in this field.

Highly interpretable ML/LLMs. The development of highly interpretable ML/LLMs to address the challenges of high-temperature alloy research and development has been a long-standing objective. Such models can provide deeper insights into the underlying mechanisms, fostering more effective and reliable advancements in the field.

High computational cost. The parameters of LLMs often range from billions to tens of billions, requiring immense computational resources for training.

Development of multi-objective optimization algorithms. In practice, a commercially viable superalloy must simultaneously meet numerous performance criteria, including yield strength, oxidation resistance, creep resistance, density, manufacturability, and cost. Conventional optimization approaches, such as Bayesian optimization or genetic algorithms, perform reasonably well for two or three objectives but become ineffective when scaling to seven, ten, or even twenty simultaneously optimized targets. High-dimensional multi-objective optimization for superalloys thus remains an open challenge. There is an urgent need to develop new optimization paradigms, potentially combining physics-based constraints, active learning, vector-valued surrogate models, and LLM-assisted search, that can efficiently navigate these extremely complex performance landscapes.

CRediT authorship contribution statement

Linlin Sun: Writing – original draft, Visualization, Supervision, Methodology, Investigation, Conceptualization. **Jie Xiong:** Writing – review & editing, Writing – original draft, Visualization, Resources, Methodology, Investigation, Conceptualization. **Qingshuang Ma:** Visualization, Data curation. **Chenghao Pei:** Data curation. **Huijun Li:** Resources, Methodology. **Qiuzhi Gao:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

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 grants and financial supports from the National Natural Science Foundation of China (Grant Nos. 52471004, 52201203, 52401015), and the Industry-University-Research-Cooperation Project of Hebei Based Universities and Shijiazhuang City (Grant No. 241791237A) are gratefully acknowledged. The authors also very much appreciate Prof. Xiao-Feng Zhao and Prof. Jie Lu from Shanghai Jiao Tong University, and Dr. Zhi-Min Pan from University of Science and Technology Beijing for offering insightful proposition and useful guidance.

References

[1] L. Wang, Y.H. Wang, F. Pyczak, M. Oehring, M. Song, Y. Liu, Revealing complex planar defects and their interactions at atomic resolution in a laves phase containing Co-base superalloy, Acta Mater. 264 (2024) 119568, https://doi.org/10.1016/j.actamat.2023.119568.

[2] X.T. Li, Q.S. Ma, E.Y. Liu, Z.Z. Li, J. Bai, H.j. Li, Q.Z. Gao, Order phase transition of HIP nickel-based powder superalloy during isothermal aging, J. Alloys Compd. 1010 (2025) 177269, https://doi.org/10.1016/j.jallcom.2024.177269.

[3] A.B. Parsa, D. Bürger, T.M. Pollock, G. Eggeler, Misfit and the mechanism of high temperature and low stress creep of Ni-base single crystal superalloys, Acta Mater. 264 (2024) 119576, https://doi.org/10.1016/j.actamat.2023.119576.

[4] M. Motamedi, M. Nikzad, M. Nasri, Molecular dynamics simulation of superalloys: a review, Arch. Comput. Methods Eng. 31 (2024) 2417–2429, https://doi.org/10.1007/s11831-023-10051-w.

[5] Q.Z. Gao, X.M. Zhang, Q.S. Ma, H.T. Zhu, H.L. Zhang, L.L. Sun, H.J. Li, Accelerating design of novel cobalt-based superalloys based on first-principles calculations, J. Alloys Compd. 927 (2022) 167012, https://doi.org/10.1016/j.jallcom.2022.167012.

[6] R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, C. Kim, Machine learning in materials informatics: recent applications and prospects, npj Comput. Mater. 3 (2017) 54, https://doi.org/10.1038/s41524-017-0056-5.

[7] X.L. Liu, J.X. Zhang, Z.R. Pei, Machine learning for high-entropy alloys: progress, challenges and opportunities, Prog. Mater. Sci. 131 (2023) 101018, https://doi.org/10.1016/j.pmatsci.2022.101018.

[8] D. Raabe, J.R. Mianroodi, J. Neugebauer, Accelerating the design of compositionally complex materials via physics-informed artificial intelligence, Nat. Comput. Sci. 3 (2023) 198–209, https://doi.org/10.1038/s43588-023-00412-7.

[9] F.Y. Wang, H.-H. Wu, L.S. Dong, G.F. Pan, X.Y. Zhou, S.Z. Wang, R.Q. Guo, G. L. Wu, J.H. Gao, F.-Z. Dai, X.P. Mao, Atomic-scale simulations in multi-component alloys and compounds: a review on advances in interatomic potential, J. Mater. Sci. Technol. 165 (2023) 49–65, https://doi.org/10.1016/j.jmst.2023.05.010.

[10] Y.Y. Yu, J. Xiong, X. Wu, Q. Qian, From small data modeling to large language model screening: a dual-strategy framework for materials intelligent design, Adv. Sci. 11 (2024) 2403548, https://doi.org/10.1002/advs.202403548.

[11] L.L. Sun, B. Cao, Q.S. Ma, Q.Z. Gao, J.H. Luo, M.L. Gong, J. Bai, H.J. Li, Machine learning-assisted composition design of W-free Co-based superalloys with high γ′-solvus temperature and low density, J. Mater. Res. Technol. 29 (2024) 656–667, https://doi.org/10.1016/j.jmrt.2024.01.040.

[12] S.K. Xi, J.X. Yu, L.K. Bao, L.P. Chen, Z. Li, R.P. Shi, C.P. Wang, X.J. Liu, Machine learning-accelerated first-principles predictions of the stability and mechanical properties of $L1_2$-strengthened cobalt-based superalloys, J. Mater. Inform. 2 (2022) 15, https://doi.org/10.20517/jmi.2022.22.

[13] J. Xiong, B.W. Bai, H.R. Jiang, A. Faus-Golfe, Determinants of saturation magnetic flux density in Fe-based metallic glasses: insights from machine-learning models, Rare Met. 43 (2024) 5256–5267, https://doi.org/10.1007/s12598-024-02805-7.

[14] J. Xiong, J.C. He, X.S. Leng, T.Y. Zhang, Gaussian process regressions on hot deformation behaviors of FGH98 nickel-based powder superalloy, J. Mater. Sci. Technol. 146 (2023) 177–185, https://doi.org/10.1016/j.jmst.2022.10.063.

[15] J. Xiong, T.-Y. Zhang, Data-driven glass-forming ability criterion for bulk amorphous metals with data augmentation, J. Mater. Sci. Technol. 121 (2022) 99–104, https://doi.org/10.1016/j.jmst.2021.12.056.

[16] J. Xiong, S.-Q. Shi, T.-Y. Zhang, Machine learning of phases and mechanical properties in complex concentrated alloys, J. Mater. Sci. Technol. 87 (2021) 133–142, https://doi.org/10.1016/j.jmst.2021.01.054.

[17] J. Xiong, T.-Y. Zhang, S.-Q. Shi, Machine learning of mechanical properties of steels, Sci. China Technol. Sci. 63 (2020) 1247–1255, https://doi.org/10.1007/s11431-020-1599-5.

[18] Q.S. Ma, X.T. Li, R.F. Xin, E.Y. Liu, Q.Z. Gao, L.L. Sun, X.M. Zhang, C.X. Zhang, Thermodynamic calculation and machine learning aided composition design of new nickel-based superalloys, J. Mater. Res. Technol. 26 (2023) 4168–4178, https://doi.org/10.1016/j.jmrt.2023.08.139.

[19] J.W. Yin, Z.Y. Rao, D.Y. Wu, H.P. Lv, H.K. Ma, T. Long, J. Kang, Q. Wang, Y. D. Wang, R. Su, Interpretable predicting creep rupture life of superalloys: enhanced by domain-specific knowledge, Adv. Sci. 11 (2024) 2307982, https://doi.org/10.1002/advs.202307982.

[20] F. Yang, W.Y. Zhao, Y. Ru, Y.L. Pei, S.S. Li, S.K. Gong, H.B. Xu, Predicting the oxidation kinetic rate and near-surface microstructural evolution of alumina-forming Ni-based single crystal superalloy based on machine learning, Acta Mater. 266 (2024) 119703, https://doi.org/10.1016/j.actamat.2024.119703.

[21] B.T. Zong, J.S. Li, T.H. Yuan, J. Wang, R.H. Yuan, Recent progress on machine learning with limited materials data: using tools from data science and domain knowledge, J. Mater. Inform. 11 (2024) 100916, https://doi.org/10.1016/j.jmat.2024.07.002.

[22] P.C. Xu, X.B. Ji, M.J. Li, W.C. Lu, Small data machine learning in materials science, npj Comput. Mater. 9 (2023) 42, https://doi.org/10.1038/s41524-023-01000-z.

[23] A. Valizadeh, R. Sahara, M. Souissi, Alloys innovation through machine learning: a statistical literature review, Sci. Technol. Adv. Mater.: Methods 4 (2024) 2326305, https://doi.org/10.1080/27660400.2024.2326305.

[24] Y. Jun Hao, L. Zhen, Y. Yu Jia, Y. An Yi, L. Wen Jie, S. Dan, W. Qing, Applications of machine learning method in high-performance materials design: a review, J. Mater. Inform. 4 (2024) 14, https://doi.org/10.20517/jmi.2024.15.

[25] M.W. Hu, Q.Y. Tan, R. Knibbe, M. Xu, B. Jiang, S. Wang, X. Li, M.-X. Zhang, Recent applications of machine learning in alloy design: a review, Mater. Sci. Eng., R 155 (2023) 100746, https://doi.org/10.1016/j.mser.2023.100746.

[26] X.Y. Gao, H.Y. Wang, H.J. Tan, L. Xing, Z.Y. Hu, Data-driven machine learning for alloy research: recent applications and prospects, Mater. Today Commun. 36 (2023) 106697, https://doi.org/10.1016/j.mtcomm.2023.106697.

[27] X.J. Liu, P.C. Xu, J.J. Zhao, W.C. Lu, M.J. Li, G. Wang, Material machine learning for alloys: applications, challenges and perspectives, J. Alloys Compd. 921 (2022) 165984, https://doi.org/10.1016/j.jallcom.2022.165984.

[28] J.F. Durodola, Machine learning for design, phase transformation and mechanical properties of alloys, Prog. Mater. Sci. 123 (2022) 100797, https://doi.org/10.1016/j.pmatsci.2021.100797.

[29] D. Zagorac, H. Müller, S. Ruehl, J. Zagorac, S. Rehme, Recent developments in the inorganic crystal structure database: theoretical crystal structure data and related features, J. Appl. Crystallogr. 52 (2019) 918–925, https://doi.org/10.1107/S160057671900997X.
[30] A. Merkys, A. Vaitkus, A. Grybauskas, A. Konovalovas, M. Quirós, S. Gražulis, Graph isomorphism-based algorithm for cross-checking chemical and crystallographic descriptions, J. Cheminf. 15 (2023) 25, https://doi.org/10.1186/s13321-023-00692-1.
[31] S. Kirklin, J.E. Saal, B. Meredig, A. Thompson, J.W. Doak, M. Aykol, S. Rühl, C. Wolverton, The open quantum materials database (OQMD): assessing the accuracy of DFT formation energies, npj Comput. Mater. 1 (2015) 15010, https://doi.org/10.1038/npjcompumats.2015.10.
[32] A. Jain, S.P. Ong, G. Hautier, W. Chen, W.D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, K.A. Persson, Commentary: the materials project: a materials genome approach to accelerating materials innovation, APL Mater. 1 (2013) 011002, https://doi.org/10.1063/1.4812323.
[33] S. Curtarolo, W. Setyawan, G.L.W. Hart, M. Jahnatek, R.V. Chepulskii, R.H. Taylor, S. Wang, J. Xue, K. Yang, O. Levy, M.J. Mehl, H.T. Stokes, D.O. Demchenko, D. Morgan, AFLOW: an automatic framework for high-throughput materials discovery, Comput. Mater. Sci. 58 (2012) 218–226, https://doi.org/10.1016/j.commatsci.2012.02.005.
[34] W.R. Wang, X. Jiang, S.H. Tian, P. Liu, D.P. Dang, Y.J. Su, T. Lookman, J.X. Xie, Automated pipeline for superalloy data by text mining, npj Comput. Mater. 8 (2022) 9, https://doi.org/10.1038/s41524-021-00687-2.
[35] P.C. Xu, X.B. Ji, M.J. Li, W.C. Lu, Virtual sample generation in machine learning assisted materials design and discovery, J. Mater. Inform. 3 (2023) 16, https://doi.org/10.20517/jmi.2023.18.
[36] S.Y. Lee, S. Byeon, H.S. Kim, H. Jin, S. Lee, Deep learning-based phase prediction of high-entropy alloys: optimization, generation, and explanation, Mater. Des. 197 (2021) 109260, https://doi.org/10.1016/j.matdes.2020.109260.
[37] H. Lv, J. Yin, D. Wu, Z. Rao, C. Su, J. Kang, Q. Wang, H. Ma, H. Dong, Y. Wang, R. Su, Data-driven fatigue prediction of superalloys: a novel strategy integrating transfer learning and partial label learning for addressing ambiguous data, Adv. Sci. (2025) e07362, https://doi.org/10.1002/advs.202507362.
[38] R.L. Marchese Robinson, I. Lynch, W. Peijnenburg, J. Rumble, F. Klaessig, C. Marquardt, H. Rauscher, T. Puzyn, R. Purian, C. Aberg, S. Karcher, H. Vriens, P. Hoet, M.D. Hoover, C.O. Hendren, S.L. Harper, How should the completeness and quality of curated nanomaterial data be evaluated? Nanoscale 8 (2016) 9919–9943, https://doi.org/10.1039/c5nr08944a.
[39] T. Peixoto, B. Oliveira, Ó. Oliveira, F. Ribeiro, Data quality assessment in smart manufacturing: a review, Systems 13 (2025) 243, https://doi.org/10.3390/systems13040243.
[40] L. Yang, H. Wang, D. Leng, S. Fang, Y. Yang, Y. Du, Machine learning applications in nanomaterials: recent advances and future perspectives, Chem. Eng. J. 500 (2024) 156687, https://doi.org/10.1016/j.cej.2024.156687.
[41] H. Wang, H. Cao, L. Yang, Machine learning-driven multidomain nanomaterial design: from bibliometric analysis to applications, ACS Appl. Nano Mater. 7 (2024) 26579–26600, https://doi.org/10.1021/acsanm.4c04940.
[42] Z.L. Wang, A. Chen, K.H. Tao, Y.Q. Han, J.J. Li, MatGPT: a vane of materials informatics from past present to future, Adv. Mater. 36 (2023) 2306733, https://doi.org/10.1002/adma.202306733.
[43] L.L. Sun, Q.S. Ma, C.H. Pei, H.W. Yao, X.L. Liu, J. Xiong, C.X. Liu, H.J. Li, Q.Z. Gao, Explainable machine learning-enabled dual-objective design of γ' phase characteristic parameters in γ'-strengthened Co-based superalloys, npj Comput. Mater. 11 (2025) 316, https://doi.org/10.1038/s41524-025-01797-x.
[44] Z. Y. Gao, X. M. Jiang, A transfer learning Transformer-KAN fusion paradigm for fatigue damage prediction with data scarcity, Available at SSRN: https://ssrn.com/abstract=5430795orhttps://doi.org/10.2139/ssrn.5430795.
[45] F.Z. Zhuang, Z.Y. Qi, K.Y. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, Q. He, A comprehensive survey on transfer learning, Proc. IEEE 109 (2021) 43–76, https://doi.org/10.1109/JPROC.2020.3004555.
[46] A. Hosna, E. Merry, J. Gyalmo, Z. Alom, Z. Aung, M.A. Azim, Transfer learning: a friendly introduction, J. Big Data 9 (2022) 102, https://doi.org/10.1186/s40537-022-00652-w.
[47] S.J. Pan, Q. Yang, A survey on transfer learning, IEEE Trans. Knowl. Data Eng. 22 (2010) 1345–1359, https://doi.org/10.1109/TKDE.2009.191.
[48] B. Wang, Z.C. Li, Z.Y. Xu, Z.Y. Sun, K. Tian, Digital twin modeling for structural strength monitoring via transfer learning-based multi-source data fusion, Mech. Syst. Signal Process. 200 (2023) 110625, https://doi.org/10.1016/j.ymssp.2023.110625.
[49] J.X. Ma, B. Cao, S.Y. Dong, Y. Tian, M.H. Wang, J. Xiong, S. Sun, MLMD: a programming-free AI platform to predict and design materials, npj Comput. Mater. 10 (2024) 59, https://doi.org/10.1038/s41524-024-01243-4.
[50] X.Y. Yang, Z.G. Wang, X.S. Zhao, J.L. Song, M.M. Zhang, H.D. Liu, MatCloud: a high-throughput computational infrastructure for integrated management of materials simulation, data and resources, Comput. Mater. Sci. 146 (2018) 319–333, https://doi.org/10.1016/j.commatsci.2018.01.039.
[51] Z.L. Wang, A. Chen, K.H. Tao, J.F. Cai, Y.Q. Han, J. Gao, S.M. Ye, S.W. Wang, I. Ali, J.J. Li, AlphaMat: a material informatics hub connecting data, features, models and applications, npj Comput. Mater. 9 (2023) 130, https://doi.org/10.1038/s41524-023-01086-5.
[52] S.P. Ong, W.D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K.A. Persson, G. Ceder, Python materials genomics (pymatgen): a robust, open-source python library for materials analysis, Comput. Mater. Sci. 68 (2013) 314–319, https://doi.org/10.1016/j.commatsci.2012.10.028.
[53] J.J. Hu, S. Stefanov, Y.Q. Song, S.S. Omee, S.-Y. Louis, E.M.D. Siriwardane, Y. Zhao, L. Wei, MaterialsAtlas.org: a materials informatics web app platform for materials discovery and survey of state-of-the-art, npj Comput. Mater. 8 (2022) 65, https://doi.org/10.1038/s41524-022-00750-6.
[54] K. Choudhary, K.F. Garrity, A.C.E. Reid, B. DeCost, A.J. Biacchi, A.R. Hight Walker, Z. Trautt, J. Hattrick-Simpers, A.G. Kusne, A. Centrone, A. Davydov, J. Jiang, R. Pachter, G. Cheon, E. Reed, A. Agrawal, X. Qian, V. Sharma, H. Zhuang, S. V. Kalinin, B.G. Sumpter, G. Pilania, P. Acar, S. Mandal, K. Haule, D. Vanderbilt, K. Rabe, F. Tavazza, The joint automated repository for various integrated simulations (JARVIS) for data-driven materials design, npj Comput. Mater. 6 (2020) 173, https://doi.org/10.1038/s41524-020-00440-1.
[55] X.G. Zhao, K. Zhou, B.Y. Xing, R.T. Zhao, S.L. Luo, T.S. Li, Y.H. Sun, G.R. Na, J. H. Xie, X.Y. Yang, X.J. Wang, X.Y. Wang, X. He, J. Lv, Y.H. Fu, L.J. Zhang, JAMIP: an artificial-intelligence aided data-driven infrastructure for computational materials informatics, Sci. Bull. 66 (2021) 1973–1985, https://doi.org/10.1016/j.scib.2021.06.011.
[56] J.M. Sosa, D.E. Huber, B. Welk, H.L. Fraser, Development and application of MIPAR™: a novel software package for two- and three-dimensional microstructural characterization, Integr. Mater. Manuf. Innov. 3 (2014) 123–140, https://doi.org/10.1186/2193-9772-3-10.
[57] G.J. Wang, L.Y. Peng, K.Q. Li, L.G. Zhu, J. Zhou, N.H. Miao, Z.M. Sun, ALKEMIE: an intelligent computational platform for accelerating materials discovery and design, Comput. Mater. Sci. 186 (2021) 110064, https://doi.org/10.1016/j.commatsci.2020.110064.
[58] G.J. Wang, K.Q. Li, L.Y. Peng, Y.M. Zhang, J. Zhou, Z.M. Sun, High-throughput automatic integrated material calculations and data management intelligent platform and the application in novel alloys, Acta Metall. Sin. 58 (2021) 75–88, https://doi.org/10.11900/0412.1961.2021.00041.
[59] L.L. Sun, Q.S. Ma, C.H. Pei, H.W. Yao, X.L. Liu, J. Xiong, C.X. Liu, H.J. Li, Q.Z. Gao, Explainable machine learning-enabled dual-objective design of γ' phase characteristic parameters in γ'-strengthened Co-based superalloys, npj Comput. Mater. 11 (2025) 316, https://doi.org/10.1038/s41524-025-01797-x.
[60] H. Wang, L. Yang, D. Leng, Y. Du, H. Ning, Accelerating the discovery and optimization of metal-organic framework materials via machine learning, Adv. Colloid Interfac. 346 (2025) 103671, https://doi.org/10.1016/j.cis.2025.103671.
[61] C.L. Zhou, R.H. Yuan, B.L. Su, J.K. Fan, B. Tang, P.X. Zhang, J.S. Li, Creep rupture life prediction of high-temperature titanium alloy using cross-material transfer learning, J. Mater. Sci. Technol. 178 (2024) 39–47, https://doi.org/10.1016/j.jmst.2023.08.046.
[62] B. Chen, J. Zhang, S. Zhou, G. Zhang, F. Xu, Physics-informed transfer learning model for fatigue life prediction of IN718 alloy, J. Mater. Res. Technol. 32 (2024) 2767–2779, https://doi.org/10.1016/j.jmrt.2024.08.075.
[63] L.L. Sun, Q.S. Ma, J.W. Zhang, L.M. Yu, J. Xiong, H.J. Li, Q.Z. Gao, Identifying determinants of γ' phase coarsening behavior in Co/CoNi-based superalloys with explainable artificial intelligence (XAI), J. Mater. Inform. 4 (2024) 30, https://doi.org/10.20517/jmi.2024.57.
[64] Y.Y. Huang, J.D. Liu, C.W. Zhu, X.G. Wang, Y.Z. Zhou, X.F. Sun, J.G. Li, An explainable machine learning model for superalloys creep life prediction coupling with physical metallurgy models and CALPHAD, Comput. Mater. Sci. (2023), https://doi.org/10.1016/j.commatsci.2023.112283.
[65] C.K. Hansen, Predicting Fatigue Indicator Parameters in Additively Manufactured IN625 Using Genetic Programming for Symbolic Regression, The Unversity of Utah, USA, 2023.
[66] J.M. Wen, A. Su, X.L. Wang, H. Xu, J.J. Ma, K. Chen, X.Y. Ge, Z. Xu, Z. Lv, Virtual sample generation for small sample learning: a survey, recent developments and future prospects, Neurocomputing 615 (2025) 128934, https://doi.org/10.1016/j.neucom.2024.128934.
[67] Y.M. Jiang, X.Y. Ma, X. Li, Towards virtual sample generation with various data conditions: a comprehensive review, Inf. Fusion 117 (2025) 102874, https://doi.org/10.1016/j.inffus.2024.102874.
[68] H. Yu, Q. Zhao, J.B. Fu, Y.Z. Hu, J.J. Liang, J.G. Li, W. Xu, The design of oxidation resistant Ni superalloys for additive manufacturing, Addit. Manuf. 97 (2025) 104616, https://doi.org/10.1016/j.addma.2024.104616.
[69] Y.X. Liu, B. Xu, W. Huang Fu, H.Q. Yin, Nickel-based polycrystalline superalloy composition design framework based on non-dominated sorting genetic algorithm II, Comput. Mater. Sci. 220 (2023) 112065, https://doi.org/10.1016/j.commatsci.2023.112065.
[70] B. Alhijawi, A. Awajan, Genetic algorithms: theory, genetic operators, solutions, and applications, Evol. Intell. 17 (2024) 1245–1256, https://doi.org/10.1007/s12065-023-00822-6.
[71] D. Khatamsaz, B. Vela, P. Singh, D.D. Johnson, D. Allaire, R. Arróyave, Bayesian optimization with active learning of design constraints using an entropy-based approach, npj Comput. Mater. 9 (2023) 49, https://doi.org/10.1038/s41524-023-01006-7.
[72] P. Liu, H.Y. Huang, C. Wen, T. Lookman, Y.J. Su, The γ/γ' microstructure in CoNiAlCr-based superalloys using triple-objective optimization, npj Comput. Mater. 9 (2023) 140, https://doi.org/10.1038/s41524-023-01090-9.
[73] P. Liu, H.Y. Huang, S. Antonov, C. Wen, D.Z. Xue, H.W. Chen, L.F. Li, Q. Feng, T. Omori, Y.J. Su, Machine learning assisted design of γ'-strengthened Co-base superalloys with multi-performance optimization, npj Comput. Mater. 6 (2020) 62, https://doi.org/10.1038/s41524-020-0334-5.
[74] J.X. Yu, S. Guo, Y.C. Chen, J.J. Han, Y. Lu, Q.S. Jiang, C.P. Wang, X.J. Liu, A two-stage predicting model for γ' solvus temperature of $L1_2$-strengthened Co-base superalloys based on machine learning, Intermetallics 110 (2019) 106466, https://doi.org/10.1016/j.intermet.2019.04.009.

[75] Z.J. Qin, W.F. Li, Z. Wang, J.L. Pan, Z.X. Wang, Z.H. Li, G.W. Wang, J. Pan, F. Liu, L. Huang, L.M. Tan, L.N. Zhang, H. Han, H. Chen, L. Jiang, High-throughput characterization methods for Ni-based superalloys and phase prediction via deep learning, J. Mater. Res. Technol. 21 (2022) 1984–1997, https://doi.org/10.1016/j.jmrt.2022.10.032.

[76] P. Liu, H.Y. Huang, X. Jiang, Y. Zhang, T. Omori, T. Lookman, Y.J. Su, Evolution analysis of γ′ precipitate coarsening in Co-based superalloys using kinetic theory and machine learning, Acta Mater. 235 (2022) 118101, https://doi.org/10.1016/j.actamat.2022.118101.

[77] I.M. Lifshitz, V.V. Slyozov, The kinetics of precipitation from supersaturated solid solutions, J. Phys. Chem. Solid. 19 (1961) 35–50, https://doi.org/10.1016/0022-3697(61)90054-3.

[78] Y. Cheng, F.G. Xu, X.B. Zhao, Q.Z. Yue, B. Yu, W.S. Xia, Y.F. Gu, Z. Zhang, Degradation effects of the rafts and dislocation network on creep property of single crystal superalloy at medium temperature, Vacuum 232 (2025) 113904, https://doi.org/10.1016/j.vacuum.2024.113904.

[79] L.T. Tang, C. Li, Q.Y. Guo, R. Ding, L.M. Yu, Y.C. Liu, Microstructural degradation and creep life reduction in allvac 718Plus superalloy after long-term thermal exposures and its creep mechanisms, Mater. Sci. Eng., A 922 (2025) 147656, https://doi.org/10.1016/j.msea.2024.147656.

[80] J.H. Xu, L.F. Li, X.G. Liu, H. Li, Q. Feng, Quantitative models of high temperature creep microstructure-property correlation of a nickel-based single crystal superalloy with physical and statistical features, J. Mater. Res. Technol. 19 (2022) 2301–2313, https://doi.org/10.1016/j.jmrt.2022.06.011.

[81] C.L. Zhou, R.H. Yuan, W.J. Liao, T.H. Yuan, J.K. Fan, B. Tang, P.X. Zhang, J.S. Li, T. Lookman, Creep rupture life predictions for Ni-based single crystal superalloys with automated machine learning, Rare Met. 43 (2024) 2884–2890, https://doi.org/10.1007/s12598-023-02559-8.

[82] F. Yang, W.Y. Zhao, Y. Ru, S.Y. Lin, J. Huang, B.X. Du, Y.L. Pei, S.S. Li, S.K. Gong, H.B. Xu, Transfer learning enables the rapid design of single crystal superalloys with superior creep resistances at ultrahigh temperature, npj Comput. Mater. 10 (2024) 149, https://doi.org/10.1038/s41524-024-01349-9.

[83] B.H. Ma, X.T. Wang, G. Xu, J.W. Xu, J.S. He, Prediction of creep curves based on back propagation neural networks for superalloys, Mater 15 (2022) 6523, https://doi.org/10.3390/ma15196523.

[84] M.A. Hossain, L. Hao, W. Xiong, C.M. Stewart, Discovering chemistry to creep rupture equations in alloy 617 with machine learning, Sci. Rep. 15 (2025) 6051, https://doi.org/10.1038/s41598-025-89743-1.

[85] B. Ohl, C. Campbell, D.C. Dunand, Machine-learning prediction of creep strain rate in γ/γ′ cobalt-based superalloys, Mater. Sci. Eng., A 934 (2025), https://doi.org/10.1016/j.msea.2025.148304.

[86] M. Liu, Q.Y. Wang, Y.Q. Jiang, T.F. Zou, H. Wu, Z.H. Gao, Y.B. Pei, H. Zhang, Y. J. Liu, Q.Y. Wang, Low cycle fatigue behavior of MAR-M247 nickel-based superalloy from 500 to 900°C: analysis of cyclic response, microstructure evolution and failure mechanism, Int. J. Fatig. 189 (2024) 108564, https://doi.org/10.1016/j.ijfatigue.2024.108564.

[87] E. Sadeghi, P. Karimi, R. Esmaeilizadeh, F. Berto, S. Shao, J. Moverare, E. Toyserkani, N. Shamsaei, A state-of-the-art review on fatigue performance of powder bed fusion-built alloy 718, Prog. Mater. Sci. 133 (2023) 101066, https://doi.org/10.1016/j.pmatsci.2022.101066.

[88] L.P. Xu, R.L. Zhang, M.Q. Hao, L. Xiong, Q. Jiang, Z.X. Li, Q.Y. Wang, X.P. Wang, A data-driven low-cycle fatigue life prediction model for nickel-based superalloys, Comput. Mater. Sci. 229 (2023) 112434, https://doi.org/10.1016/j.commatsci.2023.112434.

[89] D.B. Menasche, P.A. Shade, S. Safriet, P. Kenesei, J.-S. Park, W.D. Musinski, Deep learning approaches to semantic segmentation of fatigue cracking within cyclically loaded nickel superalloy, Comput. Mater. Sci. 198 (2021) 110683, https://doi.org/10.1016/j.commatsci.2021.110683.

[90] L. Klein, A. Bauer, S. Neumeier, M. Göken, S. Virtanen, High temperature oxidation of γ/γ′-strengthened Co-base superalloys, Corros. Sci. 53 (2011) 2027–2034, https://doi.org/10.1016/j.corsci.2011.02.033.

[91] Y. Zhang, H.D. Fu, F.J. Zhou, J.X. Xie, Revealing the effect of Al content on the oxidation of γ′-strengthened cobalt-based superalloys, Corros. Sci. 198 (2022) 110122, https://doi.org/10.1016/j.corsci.2022.110122.

[92] C.H. Pei, Q.S. Ma, Q.Z. Gao, Y. Yang, Y.H. Du, H.L. Zhang, H.J. Li, A critical review on oxidation behavior of Co-based superalloys, Chin. J. Aeronaut. 38 (2025) 103380, https://doi.org/10.1016/j.cja.2024.103380.

[93] C.H. Pei, Q.S. Ma, J.W. Zhang, L.M. Yu, H.J. Li, Q.Z. Gao, J. Xiong, A novel model to predict oxidation behavior of superalloys based on machine learning, J. Mater. Sci. Technol. 235 (2025) 232–243, https://doi.org/10.1016/j.jmst.2025.01.071.

[94] D. Kubacka, M. Weiser, E. Spiecker, Early stages of high-temperature oxidation of Ni- and Co-base model superalloys: a comparative study using rapid thermal annealing and advanced electron microscopy, Corros. Sci. 191 (2021) 109744, https://doi.org/10.1016/j.corsci.2021.109744.

[95] Y.H. Chen, F. Xue, C.H. Wang, X.Q. Li, Q. Deng, X. Yang, H. Long, W. Li, L. Yang, A. Li, Effect of Cr on the microstructure and oxidation properties of Co-Al-W superalloys studied by in situ environmental TEM, Corros. Sci. 161 (2019) 108179, https://doi.org/10.1016/j.corsci.2019.108179.

[96] H.-S. Kim, S.-J. Park, S.-M. Seo, Y.-S. Yoo, H.-W. Jeong, H. Jang, Regression analysis of high-temperature oxidation of Ni-based superalloys using artificial neural network, Corros. Sci. 180 (2021) 109207, https://doi.org/10.1016/j.corsci.2020.109207.

[97] N. Khatavkar, A.K. Singh, Combined approach to capture the evolution of oxidation of nickel based superalloys using data driven approaches, Phys. Rev. Mater. 8 (2024) 053601, https://doi.org/10.1103/PhysRevMaterials.8.053601.