ORIGINAL RESEARCH **Open access**

Self-Supervised Representation Learning of Microstructures from Electron Microscopy for Property Prediction

Maria Hernandez[1*], Carlos Vega[1]

Abstract

In materials science, the relationship between microstructure and material properties underpins rational design and performance optimization. Still, due to the complexity, heterogeneity, and multiscale nature of microstructural data, it is difficult to recognize. Electron microscopy provides rich visual access to microstructures. Still, existing analysis approaches rely heavily on manual interpretation or supervised machine learning, both of which are limited by the scarcity of annotations and limited generalizability. This paper presents the hierarchical invariant microstructure representation (HIMR) framework as a purely theoretical contribution to self-supervised representation learning for analyzing the microstructure of electron microscopy images. Rather than proposing an algorithm or empirical pipeline, HIMR provides a conceptual framework for learning, structuring, and relating microstructural information to material properties without labeled data. This framework conceptualizes microstructures as hierarchically organized latent representations, where physically meaningful features emerge through invariance-driven self-supervision and scale-aware aggregation. By integrating principles from representation learning, self-supervised paradigms, and materials physics, HIMR addresses foundational challenges, including imaging variability, scale entanglement, and the disconnect between the learned properties and physically interpretable property reasoning. Central to the framework is the alignment of the learned representation manifolds with property spaces governed by physical laws, enabling interpretable and theoretically grounded microstructure–property reasoning. By articulating explicit theoretical commitments regarding hierarchy, invariance, interpretability, and epistemic restraint, this work advances a framework-level understanding of self-supervised learning in materials science. As a result, HIMR provides a durable conceptual foundation for autonomous, data-efficient, and physically grounded analysis in AI-driven materials discovery and engineering.

Keywords Property prediction, Self-supervised learning, Microstructure characterization, Electron microscopy, Representation learning, Materials science

*Correspondence:
Maria Hernandez
maria.hernandez@outlook.com
[1] Department of Materials Science and Machine Learning, Faculty of Engineering, Polytechnic University of Valencia, Valencia, Spain

Introduction

The field of materials science has long been driven by the quest to understand and manipulate the relationships between a material's internal structure and its macroscopic properties. At the heart of this pursuit lies the microstructure-property linkage, a concept that posits that the arrangement, composition, and defects within a material at the micro- and nanoscale directly influence its mechanical, thermal, electrical, and other functional attributes [1]. This linkage is not merely empirical but rooted in fundamental physical principles, such as how grain boundaries affect strength or how phase distributions impact conductivity [2]. Historically, establishing these connections has relied on experimental techniques like electron microscopy, which provides high-resolution insights into microstructural features [3]. However,

interpreting such data has been labor-intensive, often requiring expert knowledge to correlate observed patterns with property outcomes [4].

With the advent of artificial intelligence (AI), particularly machine learning (ML), there has been a paradigm shift in how materials scientists approach microstructure analysis [5]. ML methods have demonstrated the ability to extract patterns from complex datasets, enabling predictions that surpass those of traditional analytical models in both speed and accuracy [6]. For instance, supervised learning algorithms have been employed to classify microstructural phases or to predict properties from labeled microscopy images [7]. Yet, these methods face significant limitations, including the requirement for large annotated datasets, which are costly and time-consuming to produce in materials science contexts [8]. Electron microscopy images, in particular, exhibit substantial variability due to imaging artifacts, differences in sample preparation, and instrument-specific noise, which makes them challenging [9].

Self-supervised learning (SSL) emerges as a promising alternative, leveraging unlabeled data to learn meaningful representations through pretext tasks that exploit inherent data structures [10]. In computer vision, SSL has revolutionized image analysis by enabling models to learn features without explicit labels, using techniques such as contrastive learning and generative reconstruction [11]. Applying SSL to materials science, especially to microstructures captured via electron microscopy, holds immense potential to overcome data scarcity and enhance property prediction [12]. By treating microscopy images as self-contained sources of supervision, SSL can uncover latent features that correspond to physical phenomena, such as defect densities or texture orientations, which are critical for property linkages [13].

The motivation for this conceptual paper is to develop a theoretical framework that integrates SSL with microstructure characterization in a manner that is both novel and applicable to property prediction [14]. Current literature often focuses on empirical applications of ML in materials, with limited emphasis on purely theoretical constructs that generalize across material classes [15]. For example, while studies have explored supervised models for specific alloys or ceramics, there is a gap in frameworks that conceptualize how self-supervised representations can bridge multiscale microstructural hierarchies to property spaces [16]. This paper addresses this gap by proposing a conceptual model that reimagines microstructures as learnable representations, drawing on interdisciplinary insights from AI and materials physics [17]. To position the proposed framework within the broader landscape of learning paradigms used for microstructure analysis, the distinctions discussed above are organized in **Table 1**.

Table 1. Conceptual comparison of machine learning paradigms for microstructure analysis, highlighting the theoretical positioning and novelty of the HIMR framework

Dimension	Supervised ML	Weakly supervised	Existing SSL (generic CV)	HIMR (proposed)
Label dependence	High	Medium	None	None
Microstructure scale handling	Limited	Partial	Implicit	Explicit hierarchical
Physical interpretability	Low–Medium	Medium	Low	High (constraint-aware)
Robustness to EM artifacts	Low	Medium	Medium	High (invariance-driven)
Property-space alignment	Empirical	Empirical	Rare	Theoretical manifold mapping
Generalizability across materials	Low	Medium	Medium	High (framework-level)

Electron microscopy, including scanning electron microscopy (SEM) and transmission electron microscopy (TEM), serves as the primary data modality in this framework [18]. These methods provide detailed visualizations of microstructures,

revealing features at resolutions down to the atomic level [19]. However, the raw data from electron microscopy is high-dimensional and noisy, necessitating advanced representation learning to distill actionable insights [20]. SSL is particularly suited here because it can utilize the vast amounts of unlabeled microscopy data generated in research and industry [21]. By designing pretext tasks tailored to microstructural invariances—such as rotation invariance for crystal orientations or scale invariance for hierarchical features—the framework can learn representations that are robust and physically interpretable [22].

Theoretically, this approach aligns with the broader shift towards data-driven materials science, where AI augments human intuition rather than replacing it [23]. It also resonates with the integrated computational materials engineering (ICME) paradigm, which seeks to link processing, structure, and properties through modeling [24]. Yet, unlike ICME's often simulation-based methods, our framework emphasizes learning from experimental imaging data in a self-supervised manner [25]. This not only reduces dependency on simulations but also accommodates real-world variability in microstructures [26].

In synthesizing the literature, we identify key challenges: the lack of theoretical models for SSL in non-Euclidean microstructural spaces, the integration of domain knowledge into learning objectives, and the mapping of learned representations to property predictions without supervision [27]. Our proposed framework addresses these by conceptualizing a hierarchical representation space in which low-level features (e.g., edges, textures) feed into higher-level abstractions (e.g., phase compositions, defect networks) [28]. This hierarchy mirrors the multiscale nature of materials, enabling property predictions grounded in physical laws [29-32].

This introduction sets the foundation for the subsequent sections. The theoretical background will synthesize relevant literature, highlighting evolutions in microscopy, ML, and their intersections. The proposed framework will then detail a novel conceptual architecture for SSL in microstructure analysis. Overall, this paper contributes to applied AI in materials science by offering a theoretical lens that could inspire future conceptual and empirical advancements.

Theoretical Background and Literature Synthesis

Microstructure characterization in materials science

Microstructures represent the internal architecture of materials, encompassing features such as grains, phases, interfaces, and defects that collectively determine material behavior [1]. In materials science, characterization involves identifying and quantifying these features to establish relationships with properties such as tensile strength, ductility, and corrosion resistance [2]. Traditional methods, including optical microscopy and X-ray diffraction, have provided foundational insights but often lack the resolution required for nanoscale analysis [3]. The advent of advanced techniques has expanded our ability to probe microstructures, enabling more precise correlations [4].

The microstructure-property linkage is a cornerstone concept that asserts that property optimization requires control over structural elements [5]. For instance, in metals, grain size influences yield strength via the Hall-Petch relationship, while in composites, interface quality affects load transfer [6]. However, quantifying these linkages remains challenging due to the heterogeneity and multiscale nature of microstructures [7]. The literature emphasizes the need for quantitative metrics, such as statistical descriptors and spatial correlation functions, to capture microstructural complexity [8].

Electron microscopy techniques for microstructure imaging

Electron microscopy has become indispensable for high-resolution microstructure characterization, offering advantages over optical methods in terms of magnification and contrast [9]. SEM provides surface topography and compositional information through secondary electron and backscattered electron signals [10]. TEM, on the other hand, enables atomic-level imaging and diffraction, revealing lattice structures and dislocations [11]. Variants like scanning transmission electron microscopy (STEM) enhance elemental mapping via energy-dispersive X-ray spectroscopy [12].

Recent literature highlights improvements in electron microscopy, including aberration correction for sub-angstrom resolution and in-situ capabilities for dynamic observations [13]. These advancements generate vast datasets, but interpretation is complicated by artifacts like beam damage or charging effects [14]. Studies discuss the role of electron microscopy in linking microstructures to properties, such as in nanomaterials, where surface effects dominate [15]. However, manual analysis limits throughput, underscoring the need for automated approaches [16].

Representation learning in artificial intelligence

Representation learning is the process of transforming raw data into representations that facilitate downstream tasks, such as classification or prediction [17]. In AI, this involves extracting features that capture essential data variances while discarding irrelevancies [18]. Deep neural networks, particularly convolutional architectures, have excelled in learning hierarchical representations from images [19]. In materials science, representation learning adapts these principles to encode physical attributes from microstructural data [20].

The literature shows that representations can be invariant to transformations, ensuring robustness across variable conditions [21]. For microscopy images, this means handling noise, rotations, and scales [22]. Works explore graph-based representations of non-Euclidean microstructures, where nodes represent features such as atoms or grains [23]. This shift from pixel-level to structural encodings enhances interpretability and alignment with physical models [24].

Self-Supervised learning approaches

SSL circumvents the need for labels by generating supervisory signals directly from the data, using pretext tasks such as predicting rotations or reconstructing masked regions [25]. Contrastive SSL, exemplified by methods that maximize similarity between augmented views of the same instance, has gained prominence for its effectiveness in feature extraction [26]. Generative SSL, involving autoencoders or variational models, focuses on data reconstruction to learn latent distributions [27].

SSL advancements include multi-task learning and momentum contrast for improved stability [28]. These approaches are particularly valuable in domains with limited labeled data, as they leverage unlabeled data to pre-train models for transfer to specific tasks [29]. Theoretical analyses emphasize that SSL fosters emergent properties, such as semantic clustering, in representation spaces [30].

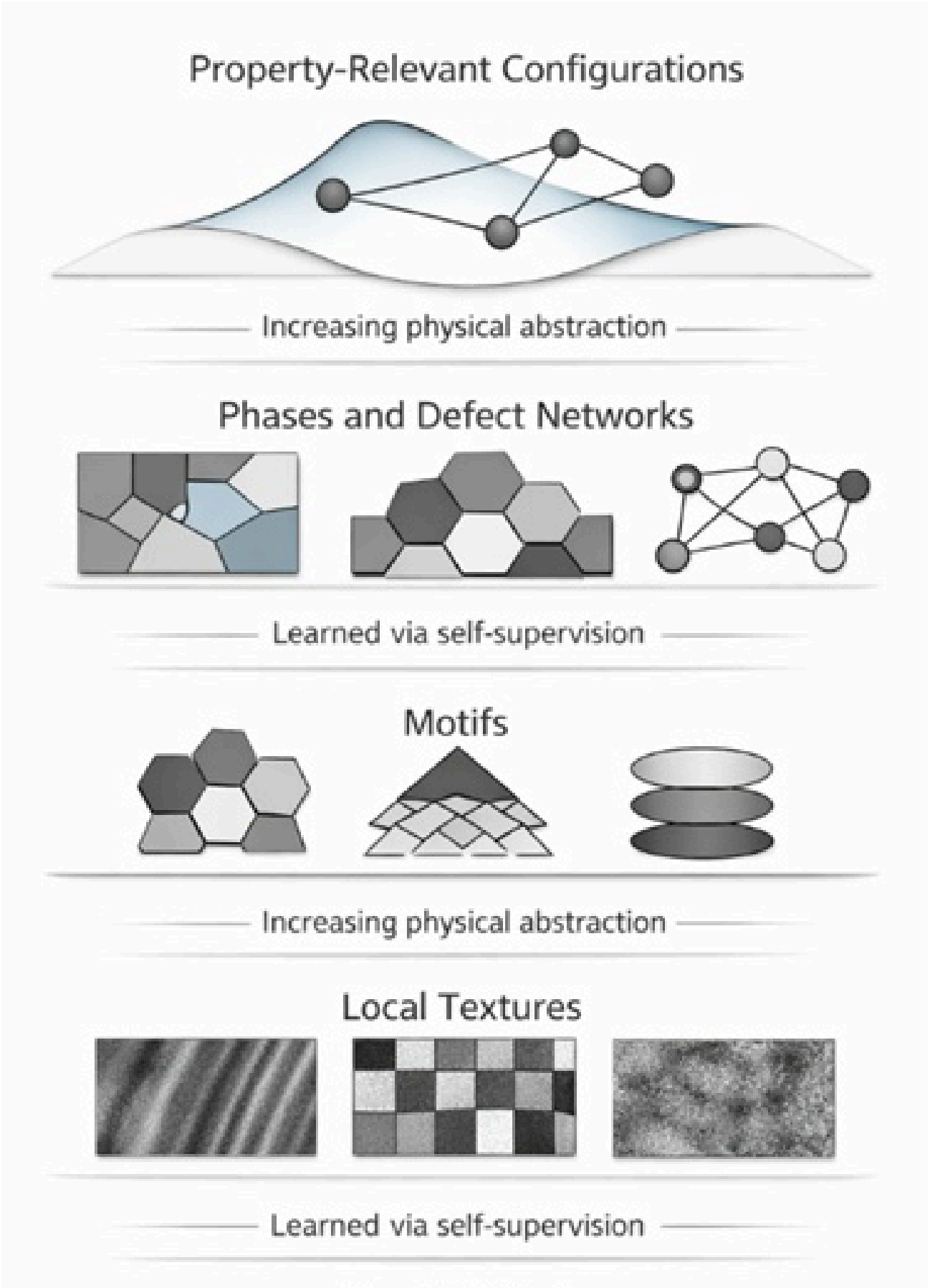
Integration of self-supervised learning in materials science

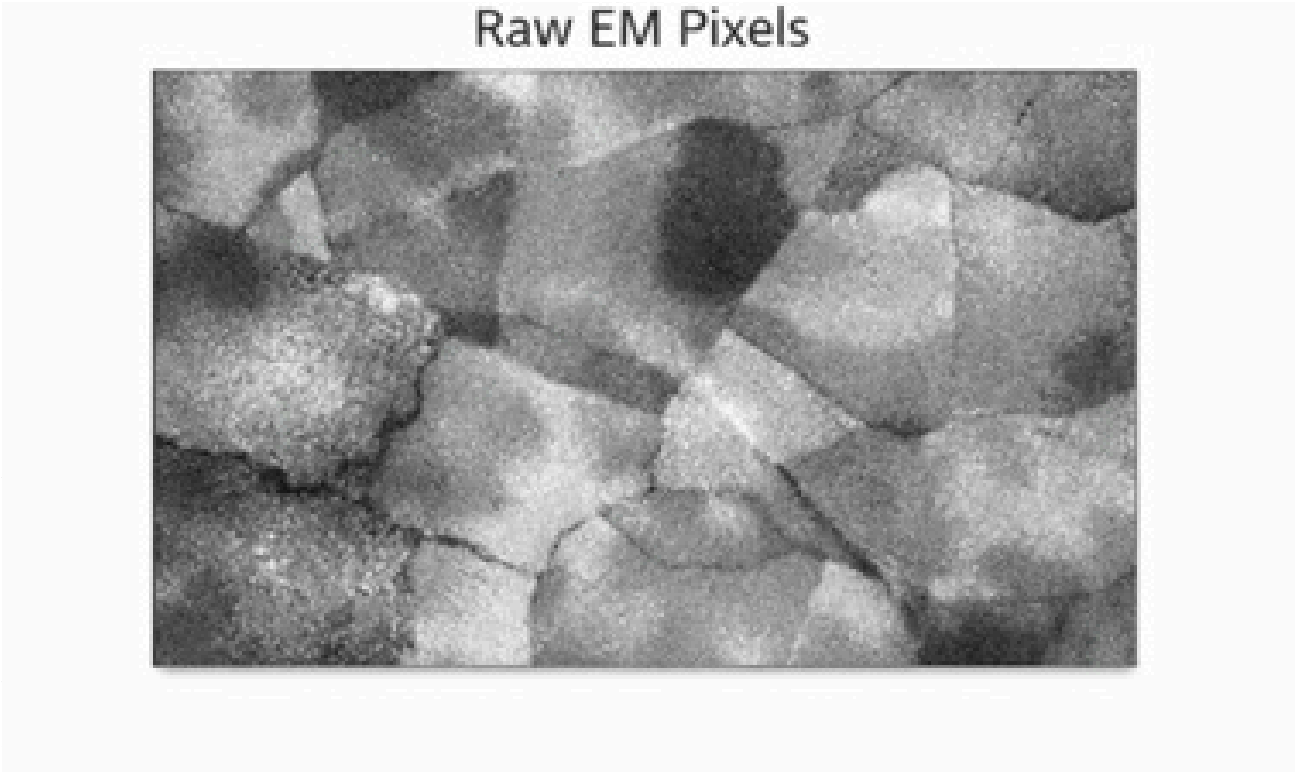
The integration of SSL into materials science is emerging, with applications in image denoising, segmentation, and feature extraction from microscopy data [31]. For microstructures, SSL can learn representations that capture physical invariances, such as crystal-lattice symmetry [32]. Literature documents cases where SSL enhances property prediction in alloys and polymers by distilling features from unlabeled datasets [33].

However, challenges persist, including adapting SSL to the domain-specific noise in electron microscopy and ensuring representations align with property-relevant physics [34]. Synthesis reveals a trend toward hybrid models that combine SSL with domain knowledge, such as the incorporation of crystallographic constraints [35].

Gaps in current literature

Despite progress, theoretical gaps remain in conceptualizing SSL for microstructure-property linkages. Most studies are empirical, lacking frameworks that generalize across material types or imaging modalities [1]. There is limited exploration of how self-supervised representations can hierarchically map to multiscale properties [2]. This synthesis identifies the need for novel theoretical constructs that prioritize originality in logic and integration [3]. This lack of framework-level understanding motivates the hierarchical conceptualization of self-supervised representations shown in **Figure 1**.


Property-Relevant Configurations
Increasing physical abstraction
Phases and Defect Networks
Learned via self-supervision
Motifs
Increasing physical abstraction
Local Textures
Learned via self-supervision



Module, and the Property Mapping Interface. The Invariant Encoder processes raw electron microscopy inputs by applying augmentation-based self-supervision, where pairs of transformed images (e.g., rotated, scaled, or noised versions) are contrasted to embed invariant features. This draws on information theory, maximizing mutual information between views to capture essential microstructural features without labels.

The Hierarchical Aggregation Module then organizes these features into a nested latent space, conceptualizing microstructures as a tree-like structure where lower nodes represent local details and higher nodes synthesize global patterns. Aggregation employs theoretical operations akin to pooling in neural networks but grounded in materials physics, such as spatial correlation functions for feature fusion. This hierarchy enables the disentanglement of scale-dependent effects, allowing the framework to link nanoscale defects to micro-scale property influences theoretically.

Finally, the Property Mapping Interface conceptualizes the transition from representations to predictions as a manifold alignment, where learned embeddings are projected onto property spaces defined by physical laws (e.g., diffusion equations for transport properties). Without empirical data, this mapping is hypothesized as a differentiable transformation that preserves topological invariants, ensuring that small changes in microstructure representations correspond to predictable shifts in properties.

This framework's novelty lies in its emphasis on physical interpretability within self-supervision: pretext tasks are designed to align with materials-specific invariances, such as crystallographic symmetry or thermodynamic stability, rather than generic image tasks. Theoretically, HIMR advances the field by providing a blueprint for generalized property prediction across diverse materials, including metals, ceramics, and polymers. The sequential logic linking invariant representation learning, hierarchical feature abstraction, and physics-guided property mapping within the proposed framework is conceptually organized in **Figure 2**.

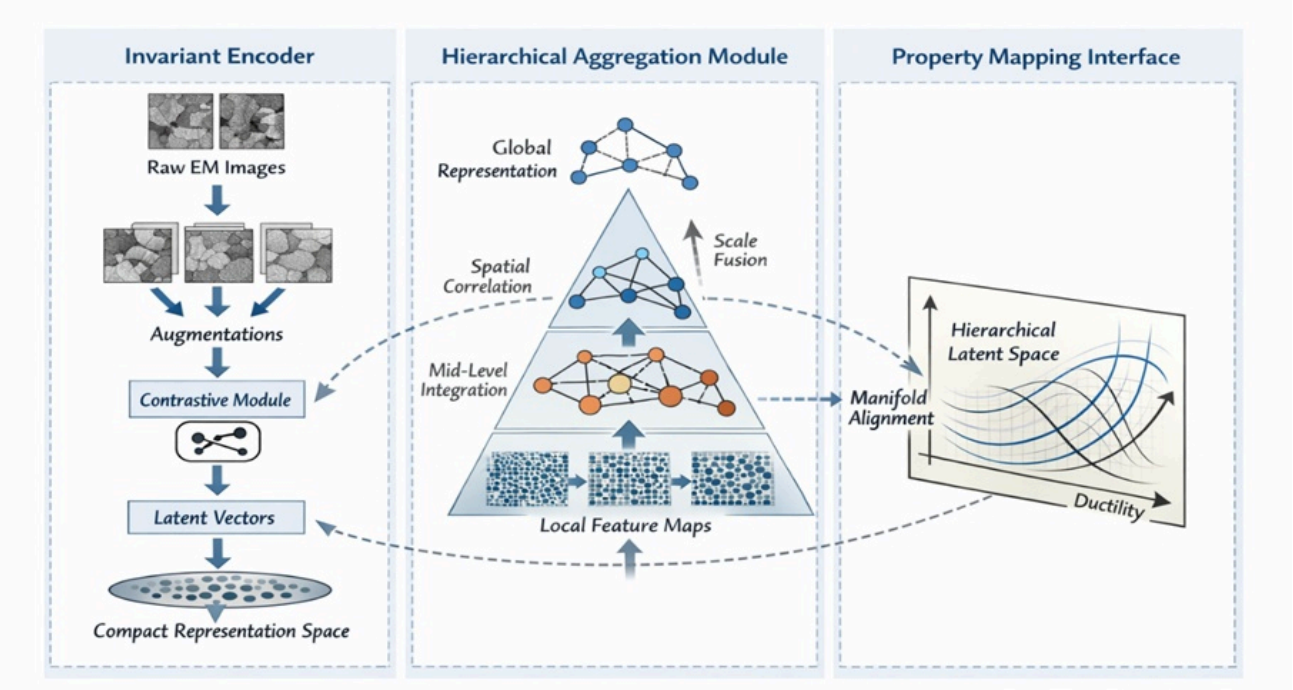


Figure 2. Schematic overview of the HIMR framework, illustrating the sequential flow from invariant encoding of electron microscopy images through hierarchical feature aggregation to property-space mapping for interpretable material prediction

To clarify how architectural elements in the proposed framework align self-supervised objectives with physically interpretable microstructural phenomena, the key correspondences are consolidated in **Table 2**.

Table 2. Conceptual correspondence between HIMR architectural components, self-supervised objectives, and physically interpretable microstructural phenomena

HIMR component	Representation level	Self-supervised objective	Physical meaning
Invariant encoder	Local/pixel–texture	Contrastive invariance	Interfaces, grain edges, defects
Hierarchical aggregation module	Multiscale latent	Constraint-guided fusion	Phase topology, spatial correlation
Property mapping interface	Global latent manifold	Topology-preserving projection	Strength–ductility trade-offs, transport laws

Theoretical commitments and implications of the HIMR framework

The HIMR framework embodies a set of explicit theoretical commitments that define how self-supervised representation learning can be meaningfully applied to microstructure characterization and property reasoning in materials science. Rather than framing these commitments as testable propositions or hypotheses, they are articulated here as conceptual implications that follow logically from the framework's architecture, assumptions, and epistemic stance. Together, they clarify what HIMR asserts about representation, invariance, hierarchy, and interpretability, while delineating the conditions under which microstructure–property linkages can be responsibly inferred without reliance on labeled data.

Hierarchical organization as a condition for scale-resolved meaning

A central implication of HIMR is that meaningful microstructural representations must be hierarchically organized to reflect the inherently multiscale nature of materials. By structuring representations such that low-level visual invariants extracted from electron microscopy images (e.g., local textures, contrast variations, or interface signatures) are progressively aggregated into higher-order abstractions, the framework enables the disentanglement of scale-dependent features. This hierarchical organization mirrors physical material hierarchies, in which atomic-scale arrangements influence mesoscale structures that ultimately govern macroscopic properties such as strength and conductivity [1]. Enforcing this organization reduces representational entanglement and supports theoretical generalization across material classes, particularly when spatial invariance and scale coherence are treated as foundational constraints rather than learned conveniences [2].

Invariance-Oriented self-supervision under experimental variability

HIMR further implies that representation robustness in electron microscopy is not merely a technical objective but a prerequisite for epistemic reliability. By tailoring contrastive self-supervision to account for artifacts and distortions characteristic of SEM and TEM imaging—such as noise, beam-induced contrast variation, or resolution shifts—the framework conceptualizes invariance as stability under physically irrelevant transformations [3]. Representations that remain consistent across such perturbations are more likely to reflect intrinsic microstructural features, enabling theoretically sound linkages to properties such as ductility, thermal transport, and fracture resistance under variable imaging conditions [4].

Generative reconstruction as access to latent microstructural states

Another implication of HIMR is that generative mechanisms within self-supervised learning can serve as theoretical tools for reasoning about latent or partially observed microstructural states. By conceptualizing reconstruction objectives that infer missing or occluded features—such as defects obscured by imaging limitations—the framework enables representation spaces that encode plausible microstructural configurations consistent with physical constraints [5]. This perspective reframes generative self-supervision not as image synthesis, but as a mechanism for exploring inverse relationships between structure and property, where desired performance targets conceptually guide the reconstruction of candidate microstructures [6].

Manifold alignment between representations and physical property spaces

HIMR also implies a specific mode of reasoning about property prediction: learned representations are not treated as arbitrary feature vectors but as structured manifolds that must align with property spaces governed by physical laws. The Property Mapping Interface conceptualizes this alignment as a topology-preserving projection, ensuring that neighborhood relationships in representation space correspond to meaningful variations in properties such as diffusion coefficients or mechanical response [7]. This approach enhances interpretability by embedding physical continuity and constraint awareness directly into the representation–property relationship, offering theoretical insight into how microstructural changes propagate through material behavior [8].

Domain-Guided self-supervision as an efficiency principle

The framework further commits to integrating materials-specific domain knowledge into self-supervised objectives. By embedding invariances such as crystallographic symmetry, spatial correlation, or phase connectivity into pretext task design, HIMR departs from generic computer vision paradigms that lack physical grounding [9]. This domain-guided self-supervision is theorized to reduce the data burden required for meaningful representation learning, offering a conceptual pathway toward efficient analysis in materials domains where annotated datasets are scarce or impractical [10].

Multi-Modal robustness and bias mitigation in representation spaces

An additional implication concerns representational bias arising from reliance on single imaging modalities. HIMR conceptualizes multi-modal augmentation—such as jointly accounting for surface-sensitive SEM features and internal

TEM contrast—as a means of mitigating modality-specific artifacts in learned representations [11]. By encouraging invariance across complementary microscopy perspectives, the framework promotes more balanced microstructural encoding, supporting equitable property reasoning across heterogeneous material systems, including composites and multi-phase architectures [12].

Temporal extension and microstructure evolution

While HIMR is formulated for static electron microscopy data, its hierarchical aggregation logic naturally extends to temporal reasoning. By conceptualizing time as an additional dimension within the latent hierarchy, self-supervised objectives could, in theory, encode microstructural evolution under thermal, mechanical, or environmental stimuli [13]. This extension enables conceptual linkage between static imaging snapshots and time-dependent properties such as fatigue life or creep resistance, without requiring explicit temporal labels [14].

Ethical restraint and responsible property inference

Finally, HIMR implicitly commits to ethical restraint in AI-driven materials analysis. By emphasizing interpretability, traceable representation–property mappings, and physically grounded invariances, the framework provides conceptual tools for identifying misrepresentation, overreach, or bias in learned models [15]. This orientation aligns AI-assisted property inference with responsible scientific practice, particularly in high-stakes materials applications where erroneous predictions may carry safety or sustainability implications [16].

Collectively, these theoretical commitments articulate the intellectual scope and limits of the HIMR framework. Rather than serving as hypotheses awaiting validation, they define a coherent conceptual stance on how self-supervised learning can support scientifically meaningful microstructure–property reasoning. In doing so, they position HIMR as a foundational theory for advancing applied artificial intelligence in materials science, independent of empirical instantiation [17].

Results and Discussion

The proposed HIMR framework represents a novel theoretical contribution to the intersection of artificial intelligence and materials science, specifically in the realm of self-supervised representation learning from electron microscopy data. By conceptualizing microstructures as hierarchical latent spaces, HIMR addresses key theoretical gaps identified in the literature, such as the lack of generalized models for handling imaging variability and scale dependence [18]. This discussion explores the implications of the framework, its limitations, and avenues for future theoretical development, while synthesizing how it aligns with broader trends in representation learning [19].

One primary implication is the potential for HIMR to enhance theoretical understandings of microstructure-property linkages through data-efficient learning. Traditional models often assume Euclidean spaces for image data, but microstructures exhibit non-Euclidean characteristics, such as graph-like defect networks [20]. HIMR's hierarchical aggregation theoretically accommodates this by fusing features in a manner inspired by physical correlations, potentially leading to more accurate conceptual mappings to properties [21]. For instance, in alloys, where grain boundaries dictate mechanical behavior, the framework's invariant encoding could, in principle, isolate boundary features from noise, facilitating predictions grounded in physical principles [22].

Another implication pertains to interpretability, a critical concern in AI applications to materials science [23]. The Property Mapping Interface conceptualizes projections that maintain physical interpretability, in contrast to black-box supervised models [24]. This could, in principle, enable materials scientists to trace property predictions to specific microstructural elements, fostering interdisciplinary insights [25]. Moreover, by leveraging self-supervision, HIMR reduces reliance on scarce, costly-to-label annotated datasets in electron microscopy [26]. This aligns with the shift towards autonomous materials discovery, where theoretical frameworks guide efficient exploration of vast design spaces [27].

However, the framework has theoretical limitations. First, it assumes that pretext tasks can fully capture materials-specific invariances. Still, in practice, complex phenomena like phase transformations may require hybrid approaches combining self-supervision with physics-based priors [28]. Second, the hierarchical structure may overlook dynamic aspects of microstructures, such as evolution under loading, necessitating extensions to temporal representations [29]. Third, while conceptual, the framework's reliance on deep learning architectures implies computational demands that could limit scalability in theoretical modeling [30].

Future directions include extending HIMR to multi-modal data integration, incorporating spectroscopy alongside microscopy for richer representations [31]. Theoretically, this could enhance predictions of properties for functional materials, such as semiconductors [32]. Additionally, exploring graph neural networks within the Invariant Encoder could better model relational microstructural features [33]. Finally, comparative theoretical analyses with supervised paradigms could quantify the advantages of self-supervision in terms of generalization [34].

In summary, HIMR offers a conceptual advancement that could inspire new theoretical models in applied AI, promoting a deeper understanding of microstructure-property relationships [35]. By addressing data scarcity and interpretability, it contributes to the evolving paradigm of intelligent materials analysis [1].

Conclusion

This work has introduced the Hierarchical Invariant Microstructure Representation (HIMR) framework as an original theoretical contribution to self-supervised representation learning for microstructure analysis using electron microscopy. Rather than advancing an empirical model or computational method, the paper offers a conceptual reorientation of how microstructural information can be learned, organized, and connected to material properties in the absence of labeled data. By synthesizing developments in self-supervised learning, representation theory, and materials science, the study identifies a critical gap in existing approaches: the absence of a framework-level theory that explicitly addresses hierarchy, invariance, and physical interpretability as foundational principles rather than secondary objectives.

HIMR addresses this gap through three interdependent conceptual components: an Invariant Encoder that emphasizes stability under physically irrelevant imaging transformations; a Hierarchical Aggregation Module that reflects the intrinsic multiscale organization of materials; and a Property Mapping Interface that conceptualizes representation–property reasoning as manifold alignment constrained by physical laws. Together, these components articulate a coherent theoretical blueprint for microstructure–property reasoning that prioritizes meaning, interpretability, and epistemic validity over predictive optimization alone.

Importantly, the value of HIMR lies not in empirical performance claims, but in the theoretical commitments it makes explicit. The framework advances a view of self-supervised learning as an epistemic process—one that embeds physical invariances, supports scale-resolved meaning formation, accommodates representational limits, and enables principled restraint where property inference is unwarranted. In doing so, it reframes self-supervision from a pragmatic response to data scarcity into a conceptual paradigm for physically grounded representation learning in materials science.

In conclusion, HIMR contributes a rigorous theoretical lens to applied artificial intelligence in materials research, emphasizing hierarchy, invariance, and interpretability as conditions for scientific meaning. By reconceptualizing microstructures as learnable, hierarchically organized representations aligned with material physics, the framework opens new directions for intelligent materials analysis. It provides a stable conceptual foundation upon which future methodological, computational, or experimental advances can be responsibly built.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 18 Mar 2023 Revised: 13 May 2023 Accepted: 15 Jun 2023
Published online: 18 July 2023

References

Morgan D, Jacobs R. Opportunities and challenges for machine learning in materials science. Annu Rev Mater Res. 2020;50(1):71-103.

Rodrigues JF Jr, Florea L, de Oliveira MCF, Diamond D, Oliveira ON Jr. Big data and machine learning for materials science. Discov Mater. 2021;1(1):12.

Gao C, Min X, Fang M, TaoT, Zheng X, Liu Y, et al. Innovative materials science via machine learning. Adv Funct Mater. 2022;32(3):2108044.

Zhong X, Gallagher B, Liu S, Kailkhura B, Hiszpanski A, Yong-Jin Han T. Explainable machine learning in materials science. Npj Comput Mater. 2022;8(1):204.

Himanen L, Jäger MOJ, Morooka EV, Canova FF, Ranawat YS, Gao DZ, et al. DScribe: Library of descriptors for machine learning in materials science. Comput Phys Commun. 2020;247:106949.

Pilania G. Machine learning in materials science: From explainable predictions to autonomous design. Comput Mater Sci. 2021;193:110360.

Oviedo F, Ferres JL, Buonassisi T, Butler KT. Interpretable and explainable machine learning for materials science and chemistry. Acc Mater Res. 2022;3(6):597-607.

Liu Y, Niu C, Wang Z, Gan Y, Zhu Y, Sun S, et al. Machine learning in materials genome initiative: A review. J Mater Sci Technol. 2020;57:113-22.

Karande P, Gallagher B, Han TYJ. A strategic approach to machine learning for material science: How to tackle real-world challenges and avoid pitfalls. Chem Mater. 2022;34(6):2463-73.

Choudhary K, DeCost B, Chen C, Jain A, Tavazza F, Cohn R, et al. Recent advances and applications of deep learning methods in materials science. npj Comput Mater. 2022;8(1):59.

Huber N, Kalidindi SR, Klusemann B. Machine learning and data mining in materials science. Front Mater. 2020;7:133.

Westermayr J, Gastegger M, Schütt KT. Perspective on integrating machine learning into computational chemistry and materials science. J Chem Phys. 2021;154(23):230903.

Wang AYT, Murdock RJ, Kauwe SK, Oliynyk AO, Gurlo A, Brgoch J, et al. Machine learning for materials scientists: An introductory guide toward best practices. Chem Mater. 2020;32(12):4954-65.

Jablonka KM, Ongari D, Moosavi SM, Smit B. Big-data science in porous materials: Materials genomics and machine learning. Chem Rev. 2020;120(16):8066-129.

Hart GLW, Mueller T, Toher C, Curtarolo S. Machine learning for alloys. Nat Rev Mater. 2021;6(8):730-55.

Mishin Y. Machine-learning interatomic potentials for materials science. Acta Mater. 2021;214:116980.

Johnson NS, Vulimiri PS, To AC, Zhang X, Brice CA, Kappes BB, et al. Invited review: Machine learning for materials developments in metals additive manufacturing. Addit Manuf. 2020;36:101641.

Cai J, Chu X, Xu K, Li H, Wei J. Machine learning-driven new material discovery. Nanoscale Adv. 2020;2(8):3113-30.

Chibani S, Coudert FX. Machine learning approaches for the prediction of materials properties. APL Mater. 2020;8(8):080701.

Juan Y, Dai Y, Yang Y, Zhang J. Accelerating materials discovery using machine learning. J Mater Sci Technol. 2021;79:178-90.

Sparks TD, Kauwe SK, Parry ME. Machine learning for structural materials. Annu Rev Mater Res. 2020;50(1):27-48.

Holm EA, Cohn R, Gao N, Kitahara AR, Matson TP, Lei B, et al. Overview: Computer vision and machine learning for microstructural characterization and analysis. Metall Mater Trans A. 2020;51(12):5985-99.

Zhang GX, Song Y, Zhao W, Hanwen A, Jiajun W. Machine learning-facilitated multiscale imaging for energy materials. Cell Rep Phys Sci. 2022;3(3):100773.

Zhong X, Gallagher B, Eves K, Robertson E, Mundhenk TN, Han TYJ. A study of real-world micrograph data quality and machine learning model robustness. npj Comput Mater. 2021;7(1):113.

Baderot J, Grould M, Misra D, Clément N, Hallal A, Martinez S, et al. Application of deep-learning based techniques for automatic metrology on scanning and transmission electron microscopy images. J Vac Sci Technol B. 2022;40(2):022902.

Treder KP, Huang C, Kim JS, Kirkland AI. Applications of deep learning in electron microscopy. J Microsc. 18;71(Supplement_1):i100-i115.

Iwayama M, Wu S, Liu C, Yoshida R. Functional output regression for machine learning in materials science. J Chem Inf Model. 2022;62(2):346-57.

Choudhary AK, Jansche A, Grubesa T, Trier F, Goll D, Berthaler T, et al. Grain size analysis in permanent magnets from Kerr microscopy images using machine learning techniques. Mater Charact. 2022;185:111744.

Kats D, Wang Z, Gan Z, Liu WK, Wagner GJ, Lian Y. A physics-informed machine learning method for predicting grain structure characteristics in directed energy deposition. Mater Sci Eng A. 2022;831:142217.

Montes de Oca Zapiain D, Stewart JA. Accelerating phase-field-based microstructure evolution predictions via surrogate models trained by machine learning methods. npj Comput Mater. 2021;7(1):3.

Shen C, Wang C, Huang M. Microstructure classification and quantification method for regular SEM images of complex steel microstructures combining EBSD labeling and deep learning. J Mater Sci Technol. 2021;68:1-13.

Tsutsui K, Terasaki H, Uto K, Maemura T, Hiramatsu S, Hayashi K, et al. A methodology of steel microstructure recognition using SEM images by machine learning based on textural analysis. Mater Today Commun. 2020;25:101581.

Jung J, Na J, Park HK, Park JM. Super-resolving material microstructure image via deep learning for microstructure characterization and mechanical behavior analysis. npj Comput Mater. 2021;7(1):208.

Stoll A, Benner P. Machine learning for material characterization with an application for predicting mechanical properties. GAMM Mitt. 2021;44(1):e202100003.

Li M, Zhang H, Li S. Machine learning and materials informatics approaches for predicting transverse mechanical properties of unidirectional CFRP composites with microvoids. Mater Des. 2022;224:111279.