

ORIGINAL RESEARCH

Open access

Why Lifelong Learning Fails for Materials Property Prediction: Catastrophic Interference in Sequential Fine-Tuning

Oliver Schmidt^{1*}, Lukas Weber¹, Jonas Richter²

Abstract

Lifelong learning, also known as continual or sequential fine-tuning, promises machine-learning models for materials science that continuously improve as new data streams arrive. In theory, a model trained on binary-alloy formation energies should seamlessly incorporate new knowledge about ternary systems, experimental band gaps, or updated DFT databases without losing prior accuracy. Yet in practice, lifelong learning systematically fails for materials property prediction. Sequential fine-tuning triggers catastrophic interference: the model rapidly forgets earlier tasks while attempting to master new ones. This failure mode analysis identifies four core failure modes that are especially destructive in the materials domain. Catastrophic forgetting occurs when shared parameters encoding old task knowledge are overwritten during fine-tuning on new tasks. Negative interference arises when conflicting signals from different data sources force the model into an impossible compromise. Representation overwriting destroys generic feature detectors in early layers that were useful across multiple materials tasks. Task imbalance allows later, larger, or easier datasets to dominate the optimization process, marginalizing earlier knowledge. The root causes are intrinsic to materials science: massive domain shifts between compositional spaces, overlapping but non-identical feature distributions, and the optimizer's inability to preserve multi-task optima in a shared parameter space. These problems are documented across recent studies of machine-learning interatomic potentials, continual learning in mechanical networks, and sequential task learning on evolving materials databases. This paper provides precise detection signatures for each failure mode and outlines mitigation principles that can be applied immediately by practitioners. By exposing why sequential fine-tuning collapses in materials property prediction, the analysis lays the foundation for genuinely lifelong materials models that accumulate knowledge rather than repeatedly erasing it.

Keywords Continual learning, Catastrophic forgetting, Materials property prediction, Lifelong learning, Sequential fine-tuning, Catastrophic interference

*Correspondence:

Oliver Schmidt
oliver.schmidt@gmail.com

¹ Department of Computational Materials Engineering, Faculty of Engineering, University of Stuttgart, Stuttgart, Germany

² Department of Intelligent Materials Analytics, Faculty of Technology, Technical University of Darmstadt, Darmstadt, Germany

Introduction

Materials science is fundamentally a lifelong learning problem. New materials are discovered every year, new properties are measured, and new databases are released with higher fidelity or broader coverage [1-4]. Ideally, predictive models should improve continuously as fresh data arrive, building on everything they have already learned instead of being retrained from scratch. This ideal is exactly what lifelong learning (also called continual learning or sequential fine-tuning) aims to deliver [5-8]. Recent work on forgetting-aware fine-tuning frameworks for universal machine-learning interatomic potentials and on lifelong machine learning potentials has highlighted the theoretical appeal of such approaches [5, 9]. Yet in practice, lifelong learning fails dramatically for materials property prediction.

Sequential fine-tuning leads to catastrophic interference: the model forgets earlier tasks while learning new ones [10, 11]. Performance on previously mastered properties collapses, sometimes by orders of magnitude, even though the new task appears to be learned successfully. This is not a minor degradation; it is a systematic collapse that renders the model unreliable for real-world materials design [12]. Studies of continual learning in mechanical networks and of continual learning of electrical conduction in resistive-switching-memory materials have shown that the same interference patterns appear across different materials domains [10, 11, 13-15].

The consequences are severe. Materials engineers who rely on these models for screening new alloys, predicting band gaps, or estimating elastic constants risk deploying predictions that are silently wrong for earlier material classes. Retraining from scratch every time a new database appears is computationally prohibitive and defeats the purpose of incremental improvement. The field therefore faces a paradox: the very data-rich, continuously evolving nature of materials science that should make lifelong learning ideal instead exposes its deepest weaknesses [1].

This failure mode analysis dissects exactly why sequential fine-tuning collapses in the materials setting. It identifies four interlocking failure modes—catastrophic forgetting, negative interference, representation overwriting, and task imbalance—and traces their root causes to domain shifts between tasks, the limitations of shared representations, and optimizer pathologies that are amplified by the physics and chemistry of materials data [16]. The analysis draws on the growing body of work on sequential task learning in materials machine learning and on forgetting in transfer learning for materials systems [17, 18].

Figure 1 organizes the manuscript's core argument by showing how sequential fine-tuning in materials property prediction moves from realistic update scenarios to shared-parameter disruption, four distinct failure modes, and observable reliability collapse.

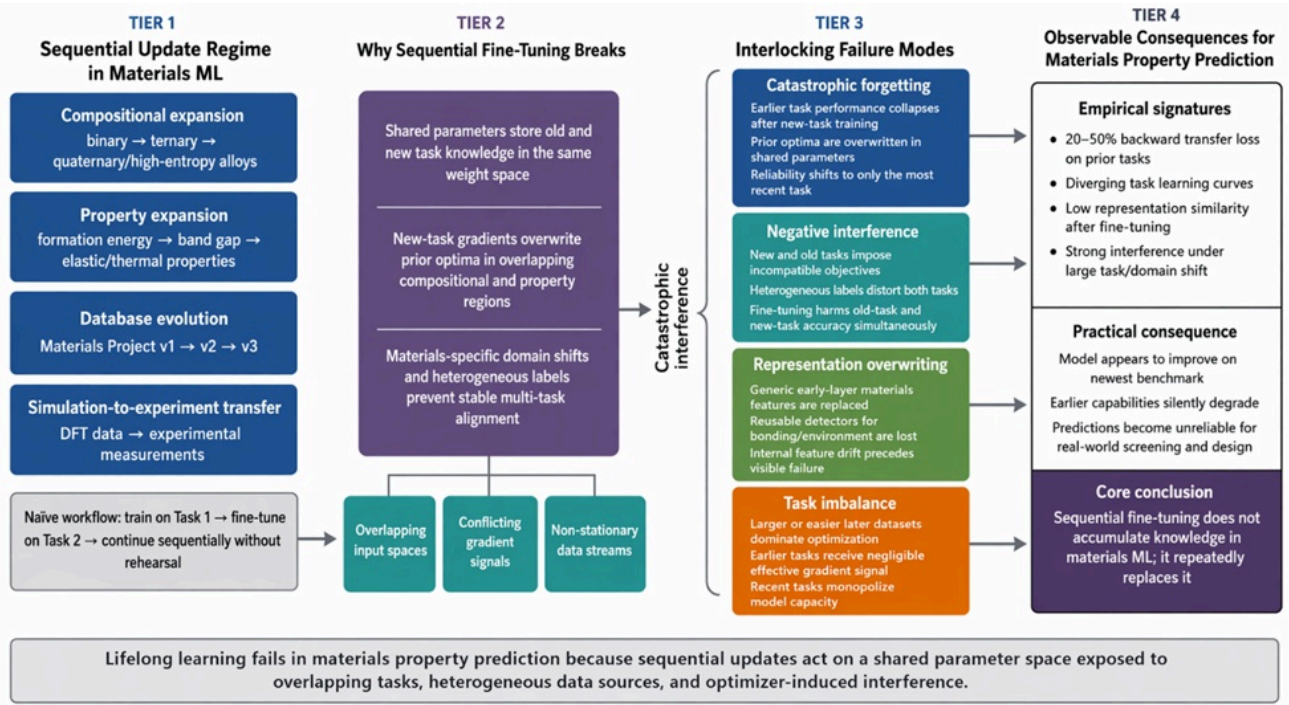


Figure 1. Hierarchical Failure Architecture of Lifelong Learning Collapse in Materials Property Prediction

By focusing exclusively on conceptual failure mechanisms rather than new experiments or performance tables, this paper provides a diagnostic framework that practitioners can apply today. Detection principles and mitigation strategies are

introduced so that future models can be designed or fine-tuned to resist catastrophic interference. Ultimately, the goal is to move materials property prediction from brittle sequential updates to genuine lifelong accumulation of knowledge.

What is Lifelong Learning for Materials?

Lifelong learning, also termed continual learning, is the ability to learn sequentially from a stream of tasks or datasets without forgetting previously acquired knowledge [5]. In materials science this capability is essential because the data landscape never stops evolving.

Four representative scenarios illustrate the problem. Scenario A involves compositional expansion: a model is first trained on binary alloys, then fine-tuned on ternary systems, then on quaternary and high-entropy alloys. Scenario B concerns property expansion: the same architecture is trained first on formation energy, then on band-gap prediction, then on elastic constants or thermal conductivity. Scenario C reflects database evolution: a model trained on Materials Project version 1 is later fine-tuned on version 2 and version 3 as new calculations and corrections are added [1]. Scenario D captures the shift from simulation to experiment: a model initially trained on DFT data is fine-tuned on experimental measurements that carry different uncertainties and systematic biases.

The standard approach to these scenarios is sequential fine-tuning. The workflow is straightforward: (1) train the model to convergence on Task 1 (for example, binary-alloy formation energies); (2) take the resulting weights and continue training on Task 2 (ternary alloys); (3) repeat for every subsequent task. No rehearsal of old data is performed, and the optimizer treats the new task as the sole objective [18].

The problem manifests immediately after the first fine-tuning step. Performance on Task 1 drops dramatically—often by more than 50 %—while the model appears to make progress on Task 2. This pattern repeats with each new task, producing a model that is only reliable for the most recent data it has seen. Work on lifelong machine learning potentials and on agents for sequential learning using multiple-fidelity data has documented exactly this degradation in materials-relevant settings [5, 17].

Why does the simple sequential strategy fail so consistently? The core issue is that materials tasks are not independent. Compositional spaces overlap, local atomic environments recur across tasks, and the same input features can carry different target labels depending on the property or data source [16]. When the optimizer updates shared parameters to reduce loss on the new task, it inevitably disturbs the delicate balance that encoded earlier knowledge. Studies of sequential task learning in materials machine learning and of forgetting in transfer learning for materials confirm that this interference is not an edge case but the default outcome of naïve fine-tuning [17, 18].

In short, lifelong learning for materials property prediction is not merely an incremental upgrade; it is a fundamentally different regime that exposes the fragility of current architectures and training protocols. Understanding the precise failure modes is the first step toward building models that can truly accumulate knowledge across the expanding materials universe.

Catastrophic Forgetting

Catastrophic forgetting is the rapid and severe degradation of performance on previously learned tasks after training on a new task [9, 19]. In materials property prediction this failure mode is especially destructive because the model does not merely become slightly worse on old tasks; it often becomes completely unusable.

The mechanism is simple yet devastating. All knowledge about earlier tasks is stored in the same shared parameters that the optimizer modifies for the new task. When gradients from the new task push parameters away from the old-task optimum, the encoded representations for the old task are erased. The effect is not gradual; it is catastrophic. A model

that once predicted formation energies of binary alloys with high accuracy can, after only a few epochs of fine-tuning on ternary systems, produce errors five times larger than its original performance [5, 9].

This behavior has been observed in multiple materials contexts. Investigations of forgetting-aware fine-tuning for pretrained universal machine-learning interatomic potentials and of lifelong machine learning potentials both highlight how parameter updates during sequential fine-tuning destroy prior optima [5, 9]. The problem is exacerbated in materials because tasks are not cleanly separable. Binary and ternary alloys share the same local coordination environments and bonding motifs. The model therefore cannot “quarantine” knowledge; every update affects overlapping regions of the input space [10].

Detection is straightforward once the right metrics are tracked. The signature of catastrophic forgetting is a performance drop on Task 1 exceeding 50 % immediately after fine-tuning on Task 2, with no recovery even if training on Task 2 continues for many more epochs. The degradation is permanent unless explicit countermeasures are introduced.

Materials-specific factors make the problem worse than in other domains. The compositional space is continuous rather than discrete, so new tasks almost always overlap with old ones. Feature distributions shift subtly but meaningfully (for example, bond-length distributions change when moving from binary to ternary compositions). The optimizer therefore receives conflicting gradient signals that cannot be satisfied simultaneously without compromising earlier knowledge. Recent analyses of continual learning in mechanical networks and of sequential learning in materials databases confirm that this interference pattern is ubiquitous [10, 17].

Catastrophic forgetting is not a failure of the data or the model architecture alone; it is a direct consequence of attempting to solve a multi-task optimization problem sequentially with a single set of parameters. Until this structural limitation is acknowledged and addressed, sequential fine-tuning will continue to produce models that know only their most recent training task [20].

Negative Interference

Negative interference goes beyond simple forgetting. It occurs when learning a new task actively harms performance on previous tasks because the new knowledge directly conflicts with the old [21]. In materials property prediction this failure mode appears whenever tasks present contradictory signals for overlapping inputs.

The root cause is label inconsistency across tasks. Two different data sources may assign different target values to chemically similar structures. When the model tries to satisfy both sets of labels using the same parameters, it satisfies neither. The optimization landscape develops saddle points or flat regions where no single set of weights can minimize loss on all tasks simultaneously [16].

A concrete materials example illustrates the mechanism. Task 1 uses DFT formation energies calculated with the PBE functional; Task 2 uses experimental formation energies for the same compounds. During fine-tuning on the experimental data the model attempts to correct systematic biases, but the correction distorts the internal representations that were tuned to reproduce the entire consistent landscape. The result is worse performance on both tasks. Work on forgetting in transfer learning for materials and on continual learning of electrical conduction in resistive-switching materials has shown analogous conflicts when simulation and experiment are combined sequentially [11, 18].

Detection signatures are distinct from those of pure forgetting. The performance drop on Task 1 is larger than can be explained by simple parameter drift alone, and Task 2 itself shows unexpectedly poor accuracy. The model is caught between two incompatible objectives and cannot reconcile them.

Negative interference is particularly insidious in materials science because data sources are heterogeneous by nature. Sequential fine-tuning has no built-in mechanism to isolate or reconcile these discrepancies; it simply averages them in a

way that degrades everything. Studies of sequential task learning in materials machine learning repeatedly document this pattern when moving from one database version to the next or from computation to measurement [1, 17].

The failure mode reveals a deeper limitation: shared representations assume that all tasks are compatible. When they are not, negative interference becomes inevitable [22]. Recognizing this incompatibility early is essential for choosing appropriate mitigation strategies rather than hoping that more training data or longer fine-tuning will magically resolve the conflict.

Representation Overwriting, Task Imbalance

Failure Mode 3 is representation overwriting. Shared feature representations learned in the early layers of a model—features such as bond lengths, coordination numbers, or local chemical environments—are gradually overwritten to become optimal for the newest task [16]. The generic detectors that were useful across many materials tasks are replaced by task-specific detectors that no longer serve earlier properties.

This overwriting occurs because fine-tuning typically updates all layers. Even when only the final layers are intended to change, gradients propagate backward and alter generic representations. The result is that performance on Task 1 collapses even if the earliest layers are partially frozen. Analyses of representation drift in continual learning for physical systems and of transfer learning stability in materials contexts have identified exactly this loss of reusable structure [16, 23].

Failure Mode 4 is task imbalance. Later tasks often arrive with larger datasets or simpler prediction targets. The optimizer, driven by the volume and ease of the new data, devotes most of its capacity to the recent task. Earlier, smaller, or harder tasks receive negligible gradient signal and are effectively forgotten [18]. A model trained first on a modest set of DFT calculations (hard task) and then fine-tuned on CALPHAD-derived predictions (easy, abundant task) will overfit to the easy task while discarding the hard-earned knowledge from the first task. Benchmarking studies of sequential learning acceleration in materials discovery and of optimal pre-train/fine-tune strategies have shown that data volume and task difficulty strongly bias the optimization trajectory [18, 24].

Both failure modes compound each other. Overwritten representations make it impossible for the model to recover old tasks even if data replay were attempted later, while task imbalance ensures that the overwriting happens preferentially for the most recent and largest datasets [25]. The combined effect is a model that appears successful on the newest benchmark yet has silently lost all prior capability.

These four failure modes—catastrophic forgetting, negative interference, representation overwriting, and task imbalance—explain why sequential fine-tuning has not delivered the promised lifelong learning capability in materials property prediction. They are not implementation bugs; they are structural consequences of applying current training paradigms to the unique characteristics of materials data streams [5, 9].

Detection Principles

Detection principles equip practitioners with practical, low-cost mechanisms for identifying the onset of catastrophic interference in sequential fine-tuning. By enabling real-time monitoring after each step rather than awaiting final deployment to reveal knowledge loss [9, 10], they avert the gradual accumulation of errors that otherwise renders lifelong learning impractical for materials property prediction.

A core diagnostic arises from backward transfer measurement: following introduction of a new task, re-evaluation of performance across all prior tasks with the updated model reveals any degradation exceeding 20 percent as a direct marker of catastrophic forgetting [9]. In materials contexts, this metric proves particularly incisive, as it registers the swift

representational erosion observed when shifting from binary to ternary alloys or from DFT-derived to experimental data, with forgetting-aware fine-tuning of machine-learning interatomic potentials confirming that such backward transfer reliably signals the overwriting of prior optima in shared parameters [9].

Quantifying domain shift through task similarity analysis prior to fine-tuning further illuminates interference risks, whereby feature-distribution or embedding-distance metrics flag substantial dissimilarity as a precursor to negative transfer [16]. This becomes especially salient in materials science, where compositional spaces and property landscapes exhibit only partial overlap; a formation-energy model, for instance, encounters profound conflicts upon fine-tuning on band-gap data from divergent functionals or sources, as multifidelity continual-learning frameworks for physical systems demonstrate that preemptive shift quantification accurately forecasts interference magnitude [16].

Complementary insight emerges from representation similarity measurement, in which cosine similarity or centered kernel alignment of internal activations and feature embeddings before and after fine-tuning exposes sharp drops indicative of representation overwriting, even when new-task performance advances at the output layer [16]. Such overwriting assumes materials-specific forms, destabilizing stable generic features like coordination environments or bond-length distributions that should persist invariantly across tasks and thereby eroding utility for earlier properties.

Finally, learning-curve analysis traces performance trajectories for every task against training steps, exposing task imbalance through the characteristic signature of rapid convergence on later tasks coupled with steady degradation on earlier ones [18, 24]—a pattern recurrent when newer databases dominate in scale or when simpler properties disproportionately shape gradients. Sequential learning benchmarks in materials discovery affirm these diverging curves as unambiguous harbingers of instability [24].

Embedded as routine logging of prior-task metrics, task descriptors, representation snapshots, and training dynamics within any fine-tuning pipeline, this diagnostic suite renders lifelong learning a diagnosable process. It thereby enables timely intervention against failure modes in materials modeling pipelines [10, 16].

Mitigation Principles

Once failure modes are identified through real-time diagnostics, targeted mitigation strategies can reinstate genuine lifelong knowledge accumulation by directly addressing shared-parameter conflicts, representation drift, and optimizer bias, all without necessitating wholesale redesign of materials models [9, 21].

Elastic weight consolidation identifies critical parameters for prior tasks via Fisher information and applies penalties to subsequent updates, thereby anchoring established optima while permitting adaptation to new tasks; forgetting-aware frameworks for universal machine-learning interatomic potentials confirm its particular efficacy in overlapping compositional spaces [9]. Complementary rehearsal through memory replay maintains a compact, representative subset of historical data that is interleaved during new-task fine-tuning, compelling the model to sustain performance across the full task sequence rather than succumbing to recency bias, which has proven highly effective at curbing catastrophic forgetting in sequential materials modeling [21].

Where architectural flexibility exists, progressive networks introduce dedicated columns of neurons for each incoming task—linked to but not overwriting prior columns—while freezing earlier modules to insulate established knowledge entirely, delivering robust protection in physical systems characterized by ambiguous task boundaries despite elevated computational demands [10, 26]. Task-specific adapters achieve similar isolation more efficiently by freezing a pretrained backbone and updating only lightweight, low-rank layers with sparse masking, adding negligible overhead while preserving stability across incremental steps [25, 27].

Gradient episodic memory further constrains optimization by storing and projecting gradients from representative prior examples, explicitly nullifying updates that would exacerbate loss on historical data and thereby preventing representation

overwriting at its source [21]. These parameter- and gradient-level safeguards are augmented by deliberate task ordering from general to specific—broad compositional ranges preceding narrow subsets, simpler properties before complex ones—which cultivates reusable representations before they encounter specialized challenges, mitigating imbalance as evidenced in sequential learning benchmarks [18, 23].

Regularized fine-tuning provides an immediate, low-cost complement by lowering learning rates in early layers, strengthening weight decay on generic feature detectors, and triggering early stopping via backward transfer thresholds, integrating seamlessly with the above mechanisms to furnish practical defense without added memory or structural modifications [22, 28].

Applied synergistically, these interventions—parameter protection via consolidation or adapters, continuity through replay and gradient constraints, and balance via ordering and regularization—convert sequential fine-tuning into authentic lifelong learning. In materials property prediction, new datasets thereby reinforce rather than supplant accumulated knowledge [5, 6, 9].

Table 1 consolidates the paper’s diagnostic and design logic by linking each failure mode to its dominant warning signal, principal mitigation strategy, and expected implementation trade-off.

Table 1. Failure Mode–Detection–Mitigation Alignment Framework for Lifelong Materials Learning

Failure mode	Primary early-warning diagnostic	Most directly targeted mitigation principle(s)	Why the mitigation matches the failure mode	Implementation trade-off	Best-use condition in materials ML
Catastrophic forgetting	Backward transfer measurement showing substantial performance loss on earlier tasks after each fine-tuning step	Elastic weight consolidation (EWC); memory replay	EWC protects parameters critical to prior optima, while replay keeps earlier tasks active in the optimization objective	EWC can reduce plasticity; replay requires memory storage and sampling design	Best when new tasks partially overlap with prior compositional spaces and historical task accuracy must remain deployable
Negative interference	Task similarity analysis showing strong domain or label mismatch, combined with poor performance on both old and new tasks	Gradient episodic memory; task-specific adapters	Gradient constraints prevent harmful updates, while adapters isolate incompatible task-specific adjustments from the shared backbone	Gradient constraints increase optimization complexity; adapters add modular parameters and workflow overhead	Best when moving across heterogeneous data sources such as DFT, corrected databases, and experimental measurements
Representation overwriting	Representation similarity decline in early or intermediate layers after fine-tuning	Task-specific adapters; progressive networks;	Adapters and progressive structures preserve reusable	Progressive networks increase model size; adapters require task bookkeeping;	Best when early-layer materials descriptors are expected to remain useful

		regularized fine-tuning	backbone features, while regularized updates reduce drift in generic detectors	regularization may limit rapid adaptation	across many future tasks
Task imbalance	Learning-curve divergence showing rapid convergence on later tasks and monotonic decay on earlier ones	Careful task ordering; memory replay; regularized fine-tuning	Ordering builds broad reusable structure first, replay restores historical gradient visibility, and regularization tempers recency dominance	Requires planning of task sequence; replay adds storage cost; regularization may slow optimization	Best when task streams differ sharply in size, difficulty, or noise level
Cross-mode collapse	Multiple diagnostics deteriorate simultaneously across tasks, representations, and optimization traces	Combined strategy: EWC or adapters + replay or gradient memory + task ordering + regularized fine-tuning	The manuscript's argument is that failure modes are interlocking; robust lifelong learning therefore requires layered rather than single-point intervention	Higher implementation complexity and more hyperparameter tuning	Best for realistic materials pipelines involving evolving databases, multiple properties, and simulation-to-experiment transfer

Relation to Other Failure Modes

The four failure modes identified here do not exist in isolation. They generalize and extend several related analyses of sequential learning difficulties reported in the recent literature.

Relation to catastrophic forgetting in mechanical networks [10]: The study focused on sequential learning in mechanical networks and demonstrated how shared parameters lead to rapid performance collapse when new physical tasks are introduced. This paper generalizes those observations to the full spectrum of materials property prediction tasks, showing that the same interference mechanisms operate whether the domain is mechanical metamaterials, alloy composition spaces, or electronic-structure databases. The materials setting adds an extra layer of severity because compositional and property overlaps are denser and more continuous than in purely mechanical systems [10].

Relation to the stability–plasticity dilemma in incremental learning [23, 29]: Earlier work examined the tension between preserving old knowledge (stability) and acquiring new knowledge (plasticity) in a single transfer step. The present analysis extends that dilemma to true sequential multi-step transfer across many evolving materials tasks. In materials property prediction the dilemma is amplified because each new database release or experimental campaign introduces not just one new target but an entire shift in data distribution, making stability even harder to maintain without explicit mitigation [23].

Relation to domain adaptation challenges [16, 18]: Domain shift between tasks has long been recognized as a source of poor transfer when moving from one dataset to another. Optimal pre-train/fine-tune strategies and multifidelity continual-learning methods have highlighted how even modest shifts in feature distributions degrade performance [16, 18]. This

paper shifts the perspective from single-step domain adaptation to repeated sequential adaptation, revealing how successive domain shifts accumulate into catastrophic interference, representation overwriting, and negative interference. The sequential view explains why naïve fine-tuning fails so consistently in materials science, where data sources evolve continuously rather than appearing as isolated source–target pairs.

By connecting these previously studied phenomena under the umbrella of lifelong learning failure modes, the analysis provides a unified diagnostic lens. Catastrophic forgetting, negative interference, representation overwriting, and task imbalance are not separate bugs but interlocking consequences of applying shared-parameter architectures to the non-stationary, overlapping data streams that define modern materials science [10, 16, 18, 23]. Recognizing these connections allows practitioners to borrow and adapt mitigation techniques across domains while tailoring them to the unique physics and chemistry of materials data.

Implications for Lifelong Learning in Materials

The failure modes and mitigation principles outlined above carry immediate consequences for three key audiences: practitioners, model developers, and benchmark designers.

For practitioners building day-to-day property-prediction pipelines the central lesson is simple: never assume sequential fine-tuning will work without safeguards. Default to at least one mitigation principle—EWC for parameter protection or memory replay for data continuity—whenever a new database version, experimental campaign, or property class is introduced [5, 9]. Before any model is deployed, backward transfer must be validated across the full task history. A model that looks excellent on the newest task but has silently lost accuracy on earlier ones cannot be trusted for high-stakes materials screening.

For model developers the implications are architectural. Future materials-specific networks should be designed with continual learning in mind from the outset. Progressive networks or task-specific adapters should become standard rather than experimental add-ons. Reporting forgetting metrics alongside final accuracy must become routine, just as uncertainty quantification is now expected. Lifelong machine-learning potentials already demonstrate that forgetting-aware training is feasible; the field should treat it as the baseline rather than an advanced feature [5, 6].

For benchmark designers the call is urgent: create and maintain lifelong learning benchmarks for materials. Current suites evaluate models on static splits or single transfers; new benchmarks must simulate realistic streams—evolving databases, shifting property sets, and mixed DFT–experimental data—while requiring explicit reporting of backward transfer, representation similarity, and task-imbalance curves [17, 24]. Only when forgetting is measured as rigorously as forward accuracy will the community move beyond brittle one-shot models toward genuinely cumulative intelligence.

Collectively these implications point toward a cultural shift in computational materials science. The era of retraining from scratch or accepting performance collapse as inevitable must end. With the detection and mitigation tools now available, lifelong learning can finally deliver on its promise: models that grow more knowledgeable and more reliable with every new piece of materials data rather than repeatedly erasing what they once knew [1, 16].

Conclusion

Lifelong learning through sequential fine-tuning fails systematically for materials property prediction. The core mechanism is catastrophic interference, driven by four interlocking failure modes: catastrophic forgetting, negative interference, representation overwriting, and task imbalance. These modes arise because materials tasks are strongly overlapping, data sources are heterogeneous, and shared parameter spaces cannot simultaneously satisfy conflicting optima. The result is models that appear to learn new tasks while silently destroying earlier knowledge, rendering them unreliable for real-world materials design.

Detection is straightforward once the right principles are applied: backward transfer measurement, task similarity analysis, representation similarity, and learning-curve monitoring. These signatures allow practitioners to catch failure modes in real time rather than after deployment. Mitigation is equally practical: elastic weight consolidation, memory replay, progressive networks, task-specific adapters, gradient episodic memory, intelligent task ordering, and regularized fine-tuning together restore the possibility of genuine knowledge accumulation.

The analysis generalizes earlier observations from mechanical networks, incremental transfer studies, and domain-adaptation work, showing that the materials setting exposes and amplifies these challenges more severely than other domains. The implications are clear for everyone involved in computational materials engineering. Practitioners must stop treating sequential fine-tuning as a safe default. Model developers must build architectures that are lifelong by design. Benchmark creators must demand forgetting-aware evaluation.

Only by confronting these failure modes head-on can the field move from brittle, task-specific models to truly lifelong materials intelligence—systems that improve continuously as new elements, new properties, and new data arrive. The path forward is no longer theoretical; the detection principles and mitigation strategies are ready to be implemented today. The next generation of materials property prediction must be evaluated not just on how well it learns the latest task, but on how faithfully it remembers everything that came before.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 23 May 2025

Revised: 18 Aug 2025

Accepted: 06 Dec 2025

Published online: 18 January 2026

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

References

Chang R, Wang YX, Ertekin E. Towards overcoming data scarcity in materials science: Unifying models and datasets with a mixture of experts framework. *npj Comput Mater.* 2022;8(1):242.

<https://doi.org/10.1038/s41524-022-00929-x>.

Li X, Liu Z, Liu Y, Karki S, Li X, Akinwande D, et al. All-electrical control and temperature dependence of the spin and valley Hall effect in monolayer WSe₂ transistors. *ACS Appl Electron Mater.* 2022;4(8):3930-7.

<https://doi.org/10.1021/acsaelm.2c00599>.

Malakoutian M, Field DE, Hines NJ, Pasayat S, Graham S, Kuball M, et al. Record-low thermal boundary resistance between diamond and GaN-on-SiC for enabling radiofrequency device cooling. *ACS Appl Mater Interfaces.* 2021;13(50):60553-60.

<https://doi.org/10.1021/acsaami.1c13833>.

Narita T, Kanechika M, Kojima J, Watanabe H, Kondo T, Uesugi T, et al. Identification of type of threading dislocation causing reverse leakage in GaN p-n junctions after continuous forward current stress. *Sci Rep.* 2022;12(1):1458.

<https://doi.org/10.1038/s41598-022-05416-3>.

Eckhoff M, Reiher M. Lifelong machine learning potentials. *J Chem Theory Comput.* 2023;19(12):3509-25.

<https://doi.org/10.1021/acs.jctc.3c00279>.

Meuwly M. Machine learning for chemical reactions. *Chem Rev.* 2021;121(16):10218-39.

<https://doi.org/10.1021/acs.chemrev.1c00033>.

Wang Z, Hou F, Wang R. CLRL-tuning: A novel continual learning approach for automatic speech recognition. In: *Interspeech 2023*. 2023. p. 1279-83.

<https://doi.org/10.21437/Interspeech.2023-503>.

Han B, Zhao F, Sun Y, Pan W, Zeng Y. Continual learning of multiple cognitive functions with a brain-inspired temporal development mechanism. *Natl Sci Rev.* 2026;13(7):nwag066.

<https://doi.org/10.1093/nsr/nwag066>.

Kim J, Lee J, Oh S, Park Y, Hwang S, Han S, et al. An efficient forgetting-aware fine-tuning framework for pretrained universal machine-learning interatomic potentials. *npj Comput Mater.* 2026;12:26.

<https://doi.org/10.1038/s41524-025-01895-w>.

Stern M, Pinson MB, Murugan A. Continual learning of multiple memories in mechanical networks. *Phys Rev X.* 2020;10(3):031044.

<https://doi.org/10.1103/PhysRevX.10.031044>.

Go SX, Wang Q, Wang B, Jiang Y, Bajalovic N, Loke DK. Continual learning electrical conduction in resistive-switching-memory materials. *Adv Theory Simul.* 2022;5(8):2200226.

<https://doi.org/10.1002/adts.202200226>.

Borg CK, Muckley ES, Nyby C, Saal JE, Ward L, Mehta A, et al. Quantifying the performance of machine learning models in materials discovery. *Digit Discov.* 2023;2(2):327-38.

<https://doi.org/10.1039/d2dd00113f>.

Pandey S, Nguyen XB, Borys N, Churchill H, Luu K. CLIFF: Continual learning for incremental flake features in 2D material identification. *arXiv Preprint.* 2025. arXiv:2508.17261.

Dutta S, Khanna A, Ye H, Sharifi MM, Kazemi A, San Jose M, et al. Lifelong learning with monolithic 3D ferroelectric ternary content-addressable memory. In: *2021 IEEE International Electron Devices Meeting (IEDM); 2021 Dec 11-16; San Francisco, CA. IEEE; 2021.* p. 1-4.

<https://doi.org/10.1109/IEDM19574.2021.9720495>.

Leonard T, Liu S, Alamdar M, Jin H, Cui C, Akinola OG, et al. Shape-dependent multi-weight magnetic artificial synapses for neuromorphic computing. *Adv Electron Mater.* 2022;8(12):2200563.

<https://doi.org/10.1002/aelm.202200563>.

Howard A, Fu Y, Stinis P. A multifidelity approach to continual learning for physical systems. *Mach Learn Sci Technol.* 2024;5(2):025042.

Palizhati A, Torrisi SB, Aykol M, Suram SK, Hummelshøj JS, Montoya JH. Agents for sequential learning using multiple-fidelity data. *Sci Rep.* 2022;12(1):4694.
<https://doi.org/10.1038/s41598-022-08413-8>.

Devi R, Butler KT, Sai Gautam G. Optimal pre-train/fine-tune strategies for accurate material property predictions. *npj Comput Mater.* 2024;10(1):300.
<https://doi.org/10.1038/s41524-024-01486-1>.

Subramanian M, Sathishkumar VE, Cho J, Shanmugavadeivel K. Learning without forgetting by leveraging transfer learning for detecting COVID-19 infection from CT images. *Sci Rep.* 2023;13(1):8516.
<https://doi.org/10.1038/s41598-023-34908-z>.

Serra J, Suris D, Miron M, Karatzoglou A. Overcoming catastrophic forgetting with hard attention to the task. In: *Proceedings of the 35th International Conference on Machine Learning*. PMLR; 2018. p. 4548-57.

Ke Z, Liu B, Ma N, Xu H, Shu L. Achieving forgetting prevention and knowledge transfer in continual learning. *Adv Neural Inf Process Syst.* 2021;34:22443-56.

Boschini M, Bonicelli L, Porrello A, Bellitto G, Pennisi M, Palazzo S, et al. Transfer without forgetting. In: *Computer Vision – ECCV 2022*. Cham: Springer Nature Switzerland; 2022. p. 692-709.
https://doi.org/10.1007/978-3-031-20050-2_40.

Kim K, Lee B, Kim H, Choi YS. Transfer is all you need: Revisiting the stability-plasticity dilemma through backward and forward transfer in PLMs. *OpenReview* [Internet]. 2026 [cited 2026 Jun 23]. Available from: <https://openreview.net/forum?id=ytiKTxE5up>

Rohr B, Stein HS, Guevarra D, Wang Y, Haber JA, Aykol M, et al. Benchmarking the acceleration of materials discovery by sequential learning. *Chem Sci.* 2020;11(10):2696-706.
<https://doi.org/10.1039/c9sc05999g>.

Zhang J, You J, Panda A, Goldstein T. LoRA without forgetting: Freezing and sparse masking for low-rank adaptation. In: *ICLR 2025 Workshop on Sparsity in LLMs (SLLM)*. 2025.

Iman M. *Expanse: A deep continual/progressive learning system for deep transfer learning* [dissertation]. Athens (GA): University of Georgia; 2022.

Pandey A. Low-rank adaptation reduces catastrophic forgetting in sequential transformer encoder fine-tuning: Controlled empirical evidence and frozen-backbone representation probes. *arXiv Preprint.* 2026. arXiv:2603.27707.

Chronopoulou A, Baziotis C, Potamianos A. An embarrassingly simple approach for transfer learning from pretrained language models. In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Minneapolis (MN): Association for Computational Linguistics; 2019. p. 2089-95.
<https://doi.org/10.18653/v1/N19-1213>.

Laborieux A, Ernoult M, Hirtzlin T, Querlioz D. Synaptic metaplasticity in binarized neural networks. *Nat Commun.* 2021;12(1):2549.
<https://doi.org/10.1038/s41467-021-22768-y>.