

Physics-Informed Deep Learning for Automated Defect Localization and Material Characterization in Ultrasonic Non-Destructive Testing

Otis James

Department of Electrical Engineering and Computer Science, University of Missouri,
Columbia, MO, USA.

james908@missouri.edu

Diya C. Thakur

School of Electrical Engineering and Computer Science, Oregon State University, Corvallis,
OR, USA.

dthakur79@oregonstate.edu

Otis M. Hawkins

Department of Computer Science and Engineering, University of Nevada, Reno, Reno, NV,
USA.

otis.hawkins@unr.edu

Leif J. Warner

Department of Computer Science, University of Central Florida, Orlando, FL, USA.

leif.warner589@ucf.edu

Abstract

Ultrasonic non-destructive testing remains a cornerstone of structural integrity assurance across aerospace, energy, transportation, and civil infrastructure. Conventional interpretation of ultrasonic signals depends heavily on expert judgment, physics-based inversion, and controlled inspection protocols. Deep learning has created new possibilities for automated defect localization and material characterization, but purely data-driven approaches often suffer from limited labeled data, sensor variability, and physically inconsistent predictions. Physics-informed deep learning integrates elastic wave propagation principles, scattering constraints, and material behavior into neural network training and architecture design. This paper presents a system-level examination of physics-informed deep learning for ultrasonic non-destructive testing. It analyzes data acquisition infrastructure, model architecture, simulation-assisted training, deployment constraints, robustness, uncertainty quantification, governance, and sustainability. The discussion emphasizes structural trade-offs between physical fidelity and computational cost, centralized training and edge deployment, and automation and certification. By treating the inspection pipeline as a socio-technical system, the paper argues that industrial adoption requires modular design, standardized data provenance, continuous recalibration, and transparent decision support. Cross-domain comparisons with medical imaging and structural health monitoring further illustrate how governance mechanisms and uncertainty management can support trustworthy automation. The analysis concludes that physics-informed deep learning can improve defect localization and material characterization only when embedded within an institutional framework that supports validation, lifecycle maintenance, and regulatory alignment.

Keywords

physics-informed deep learning, ultrasonic non-destructive testing, defect localization, material characterization, structural health monitoring, uncertainty quantification, industrial deployment.

1. Introduction

Ultrasonic non-destructive testing is widely used to evaluate the internal condition of safety-critical components without causing damage. The technique relies on elastic waves transmitted through a material and measured after interaction with geometric boundaries, discontinuities, and microstructural features. Interpretations of these measurements have traditionally depended on trained inspectors who combine signal features with prior knowledge of component geometry and expected failure modes. This dependence on human expertise creates variability in inspection outcomes and limits the scalability of assessment programs across large fleets and distributed infrastructure [1]. Quantitative ultrasonic evaluation further requires the use of forward models and inversion strategies to estimate defect size, orientation, and material properties, yet such approaches are often constrained by simplifying assumptions about wave propagation and boundary conditions [2].

Recent advances in deep learning have demonstrated substantial potential for automating the interpretation of ultrasonic signals and images. Supervised convolutional networks, sequence models, and fully connected architectures can learn complex mappings from raw ultrasonic records to defect classes, locations, or material states. Reviews of the field indicate that deep learning can reduce operator workload and improve consistency in well-defined inspection tasks [3]. However, purely data-driven models treat physical knowledge as implicit and may produce predictions that violate known wave propagation behavior when test conditions shift. Physics-informed deep learning offers an alternative by embedding governing physics, such as elastic wave equations and scattering constraints, into the learning process [4]. This integration can reduce data requirements, improve generalization, and encourage predictions that remain consistent with known material behavior.

The problem of automated defect localization and material characterization must be addressed not only as a modeling challenge but also as a systems problem. Inspection data are generated by heterogeneous sensors, stored across organizational boundaries, and interpreted under regulatory and operational constraints. Algorithmic accuracy alone does not guarantee industrial acceptance. The architecture of the inspection system, data lineage, model lifecycle, uncertainty handling, and institutional governance all affect reliability [5]. This paper therefore examines physics-informed deep learning from a system-level perspective. It discusses how physical constraints can be embedded into data pipelines, model architectures, and deployment workflows, and how robustness, fairness, and sustainability concerns shape the design of trustworthy ultrasonic inspection systems.

2. Background and Related Work

Ultrasonic inspection data are diverse, including A-scans, B-scans, C-scans, and full matrix captures from phased array transducers. The underlying physical process is governed by elastic wave propagation, mode conversion, attenuation, and scattering from boundaries and discontinuities. Traditional quantitative methods use beam models, ray tracing, and elastic wave inversion to infer defect parameters from measured signals. Although these methods provide principled uncertainty information and physical interpretability, they often require significant computational effort and are sensitive to modeling error. Data-driven approaches

have emerged as a complement by treating signal interpretation as pattern recognition. The use of synthetic ultrasonic data for training is particularly important because labeled real defect examples are scarce and often proprietary [6]. Deep learning architectures have been shown to learn hierarchical features from ultrasonic images and waveforms, reducing the need for handcrafted signal features [7].

Several studies have applied convolutional neural networks to ultrasonic weld inspection and defect classification. These models can achieve high classification accuracy when training and testing data share similar acquisition settings, but performance often degrades under transducer variability, coupling differences, or new material geometries [8]. Quantitative ultrasonic evaluation has also benefited from machine learning techniques that approximate inverse mappings between observed signals and defect parameters. The combination of physics-based forward modeling with learned inverse models has become a central theme in contemporary research [9]. Within this context, full-field ultrasonic characterization has gained attention because it can recover spatially distributed material properties rather than isolated flaw parameters. Deep learning has been used to process full-field ultrasonic measurements and improve the reconstruction of internal material maps [10]. These advances illustrate a broader trend toward hybrid models that retain physical consistency while exploiting the representational capacity of deep networks.

3. Conceptual Architecture of Physics-Informed Ultrasonic Deep Learning Systems

A physics-informed ultrasonic deep learning system can be conceptualized as a layered architecture in which physical knowledge is distributed across data generation, model structure, training objectives, and post-processing constraints. At the lowest layer, the acquisition environment produces raw waveform data whose characteristics depend on transducer frequency, beam angle, material anisotropy, and coupling conditions. Above this layer, a simulation environment or physics-based feature extractor supplies domain knowledge that may be used for data augmentation or as a reference for network predictions. The core model then maps measured signals or images to defect locations, types, and material properties. Physics can be introduced as a soft constraint in the loss function, as a hard constraint through specialized network layers, or as a generative prior that regularizes the learned inverse mapping [10]. This distributed integration avoids treating physics as a single monolithic constraint and allows the system to balance fidelity and flexibility.

The architecture should also separate concerns between sensing, inference, and decision support. A modular design supports the replacement of transducer hardware or inspection protocols without retraining the entire system from scratch. For example, a data normalization module can convert raw records into standardized representations that are less sensitive to equipment-specific characteristics. A physics-informed inference module then processes these representations, while an uncertainty module estimates confidence intervals for defect presence and sizing. Finally, a decision support layer presents results in a format that inspectors and engineers can interpret. Deep learning components within this architecture benefit from hierarchical feature extraction, but their outputs must be constrained by physically plausible ranges and validated against known material behavior [11]. Such modularity is essential for long-term maintenance and regulatory acceptance because it allows components to be certified independently.

4. Data Acquisition and Preprocessing Infrastructure

Data acquisition infrastructure is a critical determinant of model performance and generalization. Ultrasonic inspections are conducted in laboratories, manufacturing plants, and field environments with different transducer configurations, calibration routines, and material conditions. The resulting data are heterogeneous in sample rate, dynamic range, coordinate registration, and noise characteristics. Data-driven structural health monitoring research has repeatedly shown that inconsistent acquisition protocols introduce distribution shifts that undermine model reliability [12]. To address this, inspection systems must capture metadata describing transducer parameters, calibration settings, component geometry, environmental conditions, and operator actions. Without such metadata, physics-informed constraints cannot be correctly applied because the physical relationship between the measured signal and the material state depends on the acquisition context.

Preprocessing for physics-informed deep learning must preserve physical information while reducing unwanted variability. Filtering, time-gain correction, and normalization are common operations, but they must be applied carefully to avoid removing diagnostically relevant features. Bayesian data analysis principles provide a useful framework for modeling measurement uncertainty at the preprocessing stage [13]. Rather than treating preprocessing as a deterministic transformation, a probabilistic representation can retain information about signal quality and ambiguity. Synthetic data generated from forward models can be integrated into the training pipeline, but the simulation parameters must reflect realistic material properties and inspection geometries. If the simulation is too idealized, the physics-informed model may learn features that do not transfer to field data. A robust infrastructure therefore combines real measurements, physics-based simulation, and uncertainty-aware preprocessing to produce training sets that cover both nominal and rare defect conditions.

5. Model Design and Training Strategies

Model design for physics-informed ultrasonic deep learning typically involves a hybrid architecture that combines data-driven representation learning with explicit physical constraints. One common strategy is to train a network using a loss function that includes both data fidelity and residual physics terms. The physics terms penalize predictions that violate wave propagation relations or known boundary conditions. Another strategy uses differentiable physics layers that transform latent defect representations into expected ultrasonic measurements. These layers enforce consistency between the inferred defect geometry and the observed signal. By embedding forward models within the network, the system can perform inversion in a differentiable manner while still using gradient-based optimization. This approach reduces the need for large labeled datasets and improves the physical plausibility of reconstructed defects [4].

Training strategies must also address calibration and reliability. Deep networks are often overconfident when tested outside their training distribution, which is particularly dangerous in safety-critical inspection. Calibration methods can adjust predicted probabilities so that confidence estimates match empirical accuracy on validation data [14]. In ultrasonic testing, calibration is especially important because defect probability thresholds often determine whether a component is accepted or rejected. Physics-informed training can improve calibration by constraining the network to produce outputs that are consistent with known physical relationships, thereby reducing the space of implausible predictions. However, this benefit depends on the quality of the embedded physics. If the forward model omits important phenomena such as attenuation anisotropy or mode conversion at complex interfaces, the resulting constraints may bias the model. System designers must therefore balance the

completeness of the physics model against the computational cost and complexity of the overall architecture.

6. Deployment and System Integration

Deployment of physics-informed deep learning in industrial ultrasonic inspection requires careful integration with existing non-destructive testing workflows. Inspection systems often operate at the edge, near production lines or field assets, where computational resources may be limited. Deep learning models can be compressed, quantized, or distilled to reduce latency, but these transformations must not compromise physical consistency. A compact model may lose the capacity to represent complex scattering behavior, thereby weakening the physics-informed constraints. System integrators must evaluate trade-offs between model size, inference speed, memory usage, and the fidelity of embedded physics. In many cases, a two-stage architecture is preferable: a lightweight edge model performs initial screening, while a more complete physics-informed model runs on centralized servers for detailed characterization.

Deployment also involves organizational and regulatory considerations. Inspection results are used by engineers, maintenance planners, and certification bodies who must trust the system. A physics-informed model can support trust by providing outputs that are consistent with known physical principles, but this alone is insufficient. The system should also produce uncertainty estimates that reflect measurement noise, model limitations, and distribution shift. Uncertainty quantification methods in deep learning, including ensemble approaches and Bayesian approximations, can be integrated into the decision pipeline to support risk-based inspection planning [15]. In practice, a deployed system should not merely output a defect label or size estimate. It should summarize evidence, confidence, and alternative interpretations so that human decision makers can evaluate the result in its full operational context.

7. Robustness, Uncertainty, and Generalization

Robustness is a central concern for physics-informed deep learning in ultrasonic non-destructive testing. Field conditions differ from laboratory training conditions in ways that are difficult to anticipate. Transducers age, couplant properties change, material microstructures vary across batches, and component geometries deviate from nominal drawings. Purely data-driven models may learn spurious correlations that fail under these shifts. Physics-informed constraints can improve robustness by limiting the model to physically plausible solutions, but they do not eliminate all vulnerabilities. If the embedded physics is inaccurate for a particular material or inspection mode, the constraint may actively mislead the model. Robustness therefore depends on a combination of diverse training data, physics-based regularization, and uncertainty awareness.

Interpretability is another factor in robust deployment. High-stakes inspection decisions require explanations that can be reviewed by domain experts. Black-box models that produce only a defect probability are difficult to validate. Research on interpretable machine learning has emphasized the value of models whose decision logic can be inspected or approximated in human-understandable terms [16]. In medical image analysis, explainable artificial intelligence techniques have been used to highlight image regions that contribute to diagnostic predictions [17]. Similar approaches can be applied to ultrasonic B-scans or C-scans to indicate which signal features support a defect call. When these highlighted regions align with

known scattering physics, the result is more credible. When they do not, the system may be detecting artifacts or spurious correlations, and the human reviewer can investigate further.

Generalization across materials, geometries, and inspection modes remains an open challenge. A model trained on steel welds may not perform well on composite aerospace structures because the physics of wave propagation differs substantially. Physics-informed models can be adapted through transfer learning, but the transfer process must preserve the physical constraints relevant to the new material. This requires a modular physics layer that can be updated without retraining the entire network. It also requires standardized data formats and simulation tools that can generate realistic training distributions for new applications. System designers should plan for continuous learning loops in which field data are periodically collected, labeled by experts or through physics-based inversion, and used to refine both the data-driven and physics-informed components of the model.

8. Governance, Fairness, and Policy Implications

Governance of physics-informed deep learning in ultrasonic inspection extends beyond technical validation. Organizations must define responsibilities for data quality, model updates, failure analysis, and human oversight. Without clear accountability, automated defect calls may be ignored by inspectors or accepted without critical review. Fairness in this context does not primarily refer to individual demographic fairness but to equitable treatment of different inspection contexts, facilities, and asset types. A model trained predominantly on data from one manufacturer or one class of components may provide less reliable assessments for other contexts. This creates systemic risk if the model is deployed uniformly across a heterogeneous fleet. Governance frameworks for artificial intelligence have proposed principles such as beneficence, non-maleficence, autonomy, justice, and explicability that can guide the design of inspection systems [18]. Applying these principles requires documenting training data coverage, physical assumptions, and known limitations.

Policy implications also include certification and standardization. Regulatory bodies typically require that non-destructive testing procedures be validated through round-robin trials and operator qualification programs. Automated systems based on deep learning introduce new challenges because their behavior can change after retraining. Current certification frameworks for personnel and procedures may not adequately address continuously updated models. Standards for non-destructive testing personnel qualification provide a useful baseline, but they were not designed for adaptive machine learning systems [19]. New policy instruments may be needed to define model versioning, validation data requirements, and the extent of human oversight. Physics-informed constraints can assist certification by providing a stable physical basis for model behavior, but certification must also evaluate the data-driven components and their interaction with physical priors. A governance system should therefore treat the model as a living artifact that requires periodic audits, incident reporting, and transparent documentation [20].

9. Sustainability and Long-Term Maintenance

Sustainability of physics-informed ultrasonic deep learning systems involves more than environmental considerations. Long-term technical sustainability requires that models remain accurate as inspection equipment, materials, and operational demands evolve. Data drift is inevitable because transducers are replaced, calibration procedures change, and new defect types emerge. A maintenance plan should include periodic performance monitoring against reference standards and archived field data. When performance degrades, the system should

support targeted retraining rather than full replacement. Physics-informed components may be more stable than purely data-driven components because they are anchored to known physical behavior, but they still require updates when the governing physics changes due to new materials or inspection modes.

Computational sustainability is also relevant. Training large physics-informed models can be energy-intensive, especially when simulation-based data generation is included. Organizations should evaluate whether model complexity is justified by the inspection task. In many cases, simpler architectures with strong physical priors may achieve sufficient accuracy with lower computational cost than large general-purpose models. Lifecycle assessment of inspection systems should include the energy and material costs of data acquisition, simulation, training, and edge hardware. Finally, knowledge sustainability requires maintaining the expertise needed to understand both the physics and the machine learning components. If organizations rely entirely on automated systems without retaining domain knowledge, they risk losing the ability to detect model failures. Sustainable deployment therefore requires investment in training, documentation, and interdisciplinary collaboration between non-destructive evaluation specialists, machine learning engineers, and reliability engineers.

10. Conclusion

Physics-informed deep learning holds significant promise for automated defect localization and material characterization in ultrasonic non-destructive testing. By embedding elastic wave physics, scattering constraints, and material behavior into data pipelines, model architectures, and training objectives, these systems can reduce data requirements, improve generalization, and produce physically plausible predictions. However, the path from laboratory demonstration to industrial deployment is not purely algorithmic. It requires a system-level approach that addresses data acquisition heterogeneity, preprocessing uncertainty, model design trade-offs, edge deployment constraints, robustness under distribution shift, governance mechanisms, and long-term maintenance. Modular architectures that separate sensing, inference, uncertainty estimation, and decision support can support incremental certification and adaptation. Uncertainty quantification and explainability are essential for human trust and regulatory acceptance. Cross-domain lessons from medical imaging and structural health monitoring highlight the importance of transparent validation and continuous monitoring. As the field matures, institutional frameworks must evolve to govern model updates, data provenance, and accountability. A sustainable and trustworthy ultrasonic inspection system will combine physics-informed learning with rigorous engineering practice, clear governance, and sustained investment in human expertise. Such a system can improve structural safety while respecting the operational realities of critical infrastructure.

References

1. Shull, P. J. (2002). *Nondestructive evaluation: Theory, techniques, and applications*. Marcel Dekker.
2. Felice, M. V., & Fan, Z. (2018). Sizing of flaws using ultrasonic bulk wave testing: A review. *Ultrasonics*, 88, 26–42.
3. Cantero-Chinchilla, S., Wilcox, P. D., & Croxford, A. J. (2022). Deep learning in automated ultrasonic NDE: A review. *Structural Health Monitoring*, 21(4), 1795–1819.

4. Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686–707.
5. Ye, J., Ito, S., & Toyama, N. (2020). Computerized ultrasonic imaging inspection: From shallow to deep learning. *Sensors*, 20(10), 2890.
6. Virkkunen, I., Koskinen, T., Jessen-Juhler, O., & Rinta-aho, J. (2021). Augmented ultrasonic data for machine learning. *Journal of Nondestructive Evaluation*, 40(1), 4.
7. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
8. Munir, N., Kim, H. J., Park, J., Song, S. J., & Kang, S. S. (2020). Convolutional neural network for ultrasonic weldment flaw classification. *Ultrasonics*, 105, 106128.
9. Thompson, R. B. (2002). Quantitative ultrasonic nondestructive evaluation methods. *Journal of Nondestructive Evaluation*, 21(3), 81–93.
10. Xu, Y., Pourahmadian, F., Song, J., & Wang, C. (2023). Deep learning for full-field ultrasonic characterization. *Mechanical Systems and Signal Processing*, 201, 110668.
11. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.
12. Azimi, M., Eslamlou, A. D., & Pekcan, G. (2020). Data-driven structural health monitoring and damage detection through deep learning: State-of-the-art review. *Sensors*, 20(10), 2778.
13. Sivia, D. S., & Skilling, J. (2006). *Data analysis: A Bayesian tutorial* (2nd ed.). Oxford University Press.
14. Guo, C., Pleiss, G., Sun, Y., & Weinberger, K. Q. (2017). On calibration of modern neural networks. In *Proceedings of the 34th International Conference on Machine Learning* (pp. 1321–1330). PMLR.
15. Abdar, M., Pourpanah, F., Hussain, S., Rezazadegan, D., Liu, L., Ghavamzadeh, M., Fieguth, P., Cao, X., Khosravi, A., Acharya, U. R., Makarenkov, V., & Nahavandi, S. (2021). A review of uncertainty quantification in deep learning: Techniques, applications and challenges. *Information Fusion*, 76, 243–297.
16. Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206–215.
17. Van der Velden, B. H., Kuijf, H. J., Gilhuijs, K. G., & Viergever, M. A. (2022). Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Medical Image Analysis*, 79, 102470.
18. Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1).
19. Dwivedi, S. K., Vishwakarma, M., & Soni, A. (2018). Advances and researches on non destructive testing: A review. *Materials Today: Proceedings*, 5(2), 3690–3698.
20. International Organization for Standardization. (2012). *Non-destructive testing: Qualification and certification of NDT personnel* (ISO 9712:2012). ISO.