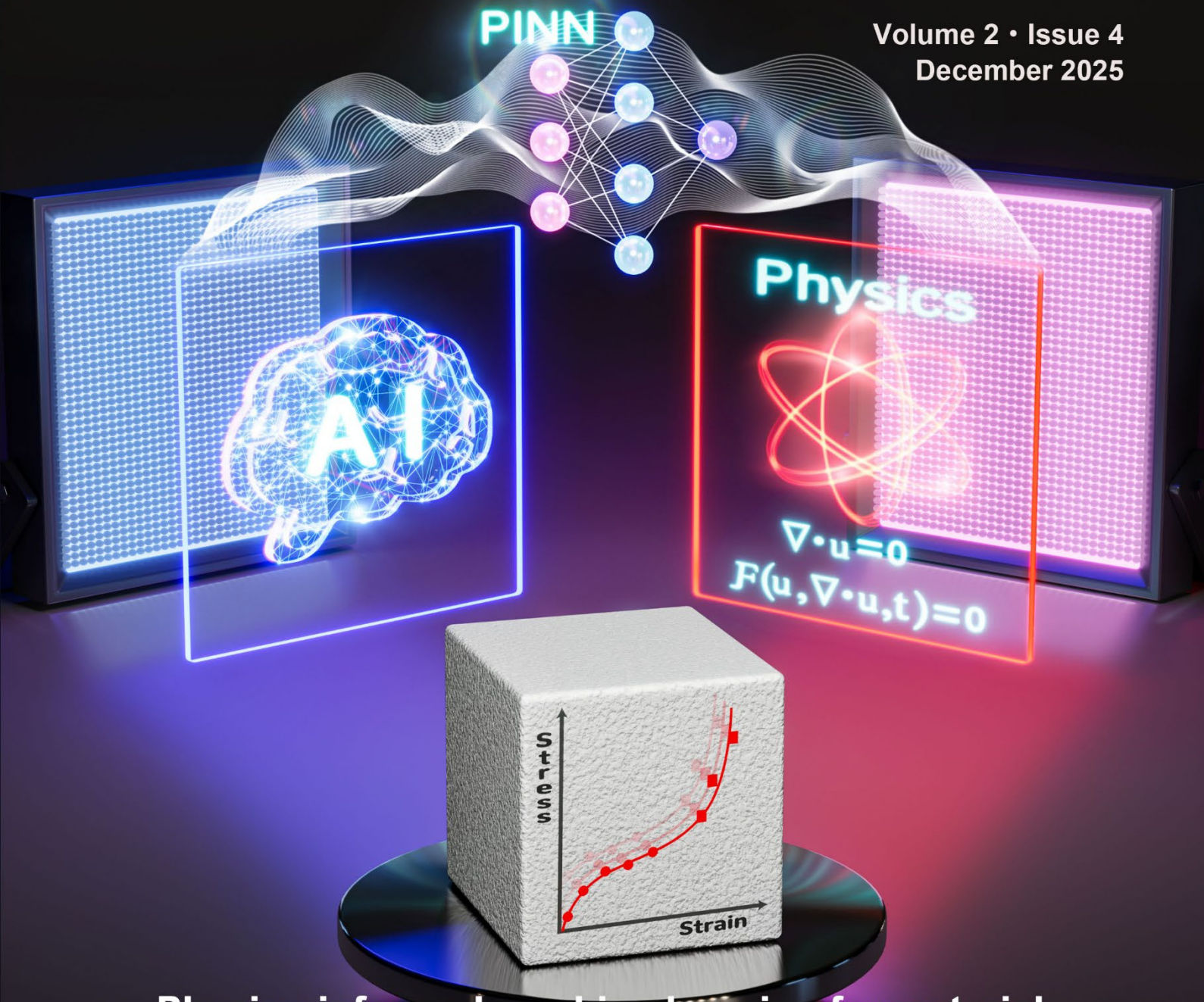
ISSN: 3041-0746 (Print)
ISSN: 3029-2573 (Online)

International Journal of AI for Materials and Design

Volume 2 • Issue 4
December 2025



Physics-informed machine learning for material characterization: A perspective on data-efficient discovery through physics-informed neural networks

Hyeonbin Moon,Junhyeong Lee,Jecheon Yu,Seunghwa Ryu

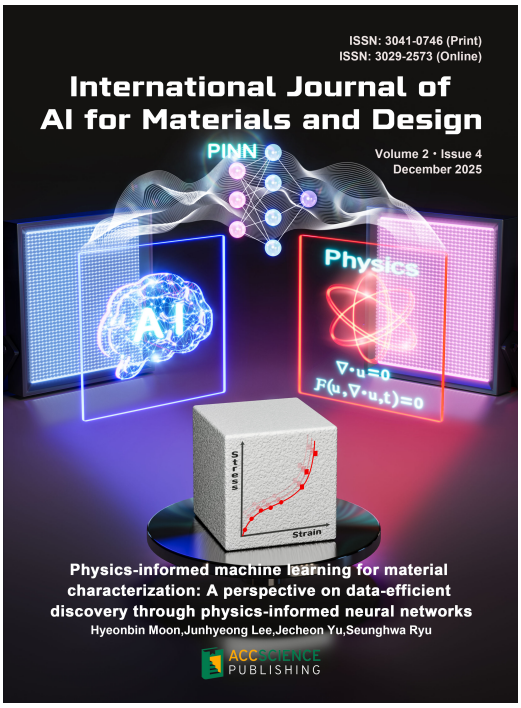
ACCSCIENCE PUBLISHING

ACCSCIENCE PUBLISHING

International Journal of AI for Materials and Design

Print ISSN: 3041-0746
Online ISSN: 3029-2573

The International Journal of AI for Material and Design (IJAMD) is a scholarly publication dedicated to advancing the intersection of artificial intelligence (AI), materials science and design. This peer-reviewed journal provides a platform for researchers, academics, and industry professionals to disseminate cutting-edge research, innovative methodologies, and practical applications that leverage AI techniques to enhance the understanding, development, and optimization of aspects related to materials and design processes. IJAMD seeks to contribute to the advancement of technology, innovation, and sustainability in materials design, engineering disciplines, product manufacturing and process technology.



About the Publisher

AccScience Publishing is a publishing company based in Singapore. We publish a range of high-quality, open-access, peer-reviewed journals and books from a broad spectrum of disciplines.

Contact Us

Managing Editor
ijamd.office@accscience.sg

AccScience Publishing
9 Raffles Place, Republic Plaza 1 #06-00 Singapore 048619.

Volume 2 • Issue 4 • December 2025
ISSN 3041-0746 (print) ISSN 3029-2573 (online)

INTERNATIONAL JOURNAL OF AI FOR MATERIALS AND DESIGN

ACCSCIENCE PUBLISHING

Access Science Without Barriers

International Journal of AI for Materials and Design

ISSN: 3041-0746 (print)
ISSN: 3029-2573 (online)

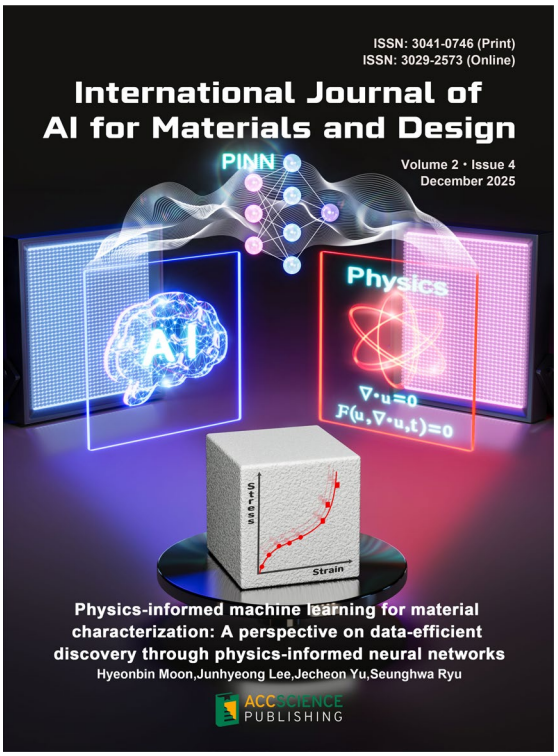
Editorial and Production Credits

Publisher: AccScience Publishing
Managing Editor: Shirley Lu
Production Editor: Sharmila Velapasamy
Article Layout and Typeset: Sinjore Technologies (India)

For all advertising queries, contact ijamd.office@accscience.sg.

Supplementary file

Supplementary files of articles can be obtained at https://accscience.com/journal/IJAMD/2/4.



Disclaimer

AccScience Publishing is not liable to the statements, perspectives, and opinions contained in the publications. The appearance of advertisements in the journal shall not be construed as a warranty, endorsement, or approval of the products or services advertised and/or the safety thereof. AccScience Publishing disclaims responsibility for any injury to persons or property resulting from any ideas or products referred to in the publications or advertisements. AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

International Journal of AI for Materials and Design

Editorial Board

*Editorial Board Members as of December 30, 2025

ACCSCIENCE PUBLISHING

Volume 2 | Issue 4 | December 2025 | Page 1 – 67 |

International Journal of AI for Materials and Design

https://accscience.com/journal/IJAMD

CONTENTS

ACCSCIENCE PUBLISHING

International Journal of AI for Materials and Design

PERSPECTIVE ARTICLE

Physics-informed machine learning for material characterization: A perspective on data-efficient discovery through physics-informed neural networks

Hyeonbin Moon, Junhyeong Lee, Jecheon Yu, and Seunghwa Ryu*

Department of Mechanical Engineering, Korea Advanced Institute of Science and Technology, 291 Daehak-ro, Yuseong-gu, Daejeon, Republic of Korea

Abstract

Accurate characterization of material properties is critical for modeling and optimizing advanced systems, yet conventional experimental and simulation-based approaches remain costly and data-intensive. As artificial intelligence evolves from data-driven modeling to physics-informed and knowledge-guided paradigms, this perspective article highlights the role of physics-informed machine learning (PIML), specifically physics-informed neural networks (PINNs), as a key enabler of data-efficient, physically consistent inference. PINNs embed governing equations into the learning process and have demonstrated strong capability in recovering constitutive and transport parameters from sparse or noisy data while preserving physical fidelity. This paper examines the fundamental structure, workflow integration, and recent advances of PINNs in the context of inverse material characterization. It also discusses open challenges in computational cost, training stability, and uncertainty quantification. Looking forward, integration with digital twins, generative modeling, and autonomous experimentation presents a pathway toward interpretable, adaptive, and automated characterization for next-generation intelligent manufacturing.

Keywords: Physics-informed neural network; Deep learning; Material property characterization; Physical AI

***Corresponding author:**
Seunghwa Ryu
(ryush@kaist.ac.kr)

Citation: Moon H, Lee J, Yu J, Ryu S. Physics-informed machine learning for material characterization: A perspective on data-efficient discovery through physics-informed neural networks. *Int J AI Mater Design*. 2025;2(4):1-9. doi: 10.36922/IJAMD025440043

Received: October 31, 2025

Revised: November 27, 2025

Accepted: December 3, 2025

Published online: December 12, 2025

1. Introduction

The rapid advancement of artificial intelligence (AI) and machine learning (ML) has profoundly transformed modern engineering. A clear trajectory has emerged toward knowledge-guided AI, where learning is informed not only by data but also by embedded physical principles and domain expertise.[1] The early phase of AI integration was largely driven by data-driven surrogate modeling, where large datasets were used to approximate complex physical responses without explicitly incorporating physical laws.[2-6] This paradigm has evolved toward generative modeling, where models approximate the data distribution and draw novel, distribution-consistent samples rather than memorized replicas of the training data.[7,8] It has further advanced toward physics-informed ML (PIML), which integrates governing equations and physical constraints directly into the learning process.[9,10] Most recently, the emergence of large

language models (LLMs) and agentic AI systems has accelerated this transformation, introducing autonomous reasoning, adaptive model construction, and closed-loop experimentation into engineering workflows,[11-13] as summarized in Table 1. Within this overarching framework of knowledge-guided AI, this perspective article represents the physics-knowledge-guided component, focusing on the physics-informed foundation through PIML and its core formulation, physics-informed neural networks (PINNs). As summarized in Figure 1, PIML encompasses several frameworks, including feature-augmented models, invariant and equivariant networks, PINNs, and operator learning methods.[9,14,15] Among these, this paper focuses on PINNs as a representative and widely adopted formulation that enables data-efficient, interpretable, and reliable characterization of material properties.

Such knowledge-guided transformation is particularly evident in material property characterization, where purely data-driven approaches often fail to capture the underlying physics that govern complex material behavior. Accurately identifying material properties, therefore, remains a persistent challenge in materials engineering and computational mechanics. Conventional experimental and numerical approaches—such as mechanical testing, curve fitting, or inverse finite element analysis—often require extensive data, significant computational effort, and substantial domain expertise.[16,17] These limitations are further exacerbated for non-linear, anisotropic, or coupled multi-physics systems, where replicating realistic boundary and environmental conditions is inherently difficult and measurement data are often sparse or noisy. Although ML-based surrogates can accelerate parameter estimation, they generally depend on large datasets and often lose physical interpretability, leading to limited generalization.[18-20]

To address these challenges, PINNs combine governing physical laws with data-driven learning, enabling reliable parameter identification under sparse or uncertain measurements while maintaining physical consistency.[21-23] Importantly, PINNs occupy a unique position in material characterization because they not only reduce data dependence but also restore the physical structure that governs constitutive behavior—something neither traditional inverse methods nor purely data-driven models can achieve. We argue that such physics-knowledge integration represents a necessary shift for material characterization problems in which data scarcity and physical fidelity must be jointly satisfied. A central theme of this perspective article is that PINNs should be viewed not merely as another modeling technique but as a foundational framework that unifies experimental observations, constitutive behavior, and computational modeling through explicit and implicit physics embedding. By providing structured inductive biases that guide inference toward physically admissible solutions, PINNs offer a principled route to overcome ill-posedness and enhance generalization across nonlinear, anisotropic, and emerging multi-physics systems. We highlight how PINNs complement existing experimental, numerical, and machine-learning approaches while also serving as a conceptual basis for next-generation hybrid, uncertainty-aware, and autonomous characterization pipelines. The remainder of this article outlines the PINN framework for material property identification and discusses representative applications, outstanding methodological challenges, and future opportunities for integrating PIML into digital-twin and intelligent manufacturing platforms.

2. Current status and challenges

This section supports our central argument by highlighting why traditional and data-driven approaches fall short and why physics-guided learning is needed. A critical challenge in material characterization is achieving reliable parameter identification under sparse, noisy, and often

Table 1. Summary of key AI paradigms, their implications, and limitations

Category	Representative models	Implications	Limitations
Data-driven surrogate modeling	GPR, CNN, LSTM	Accelerates surrogate modeling and design optimization	Strong data dependence; limited extrapolation
Generative AI	VAE, GAN, Diffusion models	Enables data augmentation and exploration of design spaces	Restricted by training distribution; no explicit physical constraints
PIML	PINNs, physics-informed neural operators, physics-constrained nets	Embeds governing equations to enhance data efficiency and physical consistency.	High computational cost
Agentic AI	LLM-based tool-using agents, multi-agent systems	Supports autonomous planning and knowledge-guided reasoning.	Complex implementation; limited real-time interpretability

Abbreviations: AI: Artificial intelligence; CNN: Convolutional neural network; GAN: Generative adversarial network; GPR: Gaussian process regression; LLM: Large language model; LSTM: Long short-term memory; PIML: Physics-informed machine learning; PINN: Physics-informed neural network; VAE: Variational autoencoder.

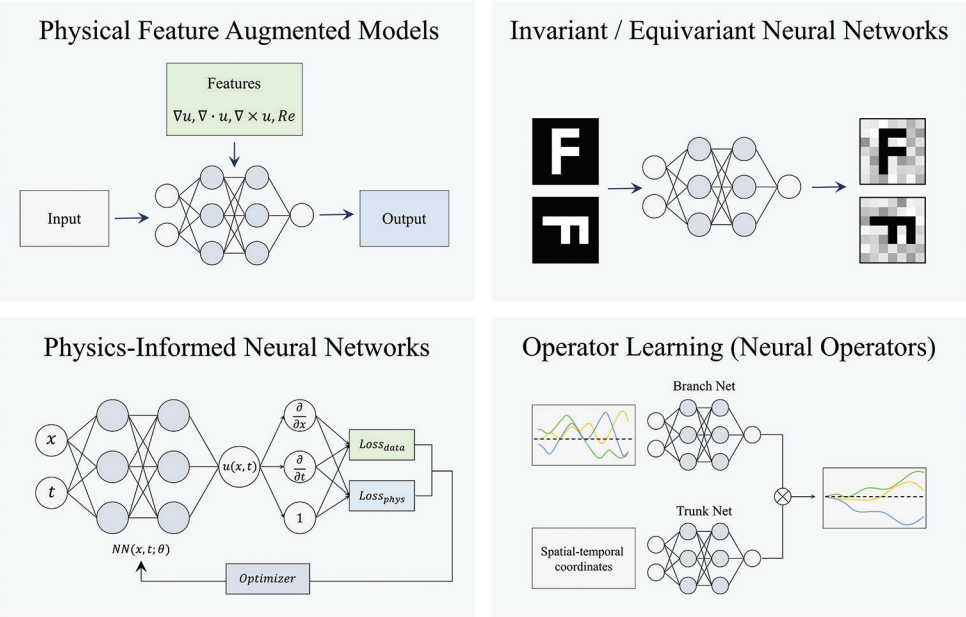


Figure 1. Four representative frameworks in physics-informed machine learning (PIML). Four representative frameworks of PIML are illustrated, including physical feature-augmented models that explicitly incorporate user-defined physical descriptors such as gradients or dimensionless numbers, invariant and equivariant networks that enforce physical symmetries through architectural design, physics-informed neural networks that embed governing equations into the loss function, and operator learning frameworks that learn mappings between function spaces. Image created by the authors.

costly measurements. Traditional inverse modeling approaches—such as curve fitting, finite element-based parameter optimization, and surrogate-assisted regression—have long supported the field, yet they remain limited by high computational cost, sensitivity to noise, and dependence on expert-driven trial-and-error procedures. These difficulties become more pronounced in nonlinear, anisotropic, or coupled multi-physics systems, where reproducing realistic loading and boundary conditions is inherently complex and experimental data are often sparse or uncertain.

In recent years, ML techniques have been explored as alternatives for accelerating material property identification. Methods such as neural networks, Gaussian process regression, and polynomial chaos expansion have been employed to correlate measured responses with underlying material parameters. These data-driven approaches can substantially reduce calibration time; however, their performance and generalization remain highly dependent on the quality and diversity of the available datasets. When data are limited, such models may suffer from overfitting and reduced generalization capability. Moreover, because physical constraints are often imposed externally rather than inherently embedded within the learning architecture, their interpretability and physical consistency can be difficult to maintain. On the other hand, physics-based inverse methods grounded in partial differential equation (PDE) formulations offer higher physical fidelity but remain computationally expensive and ill-posed, as multiple parameter sets can yield nearly identical responses. Even Bayesian inference, despite providing uncertainty quantification, can become computationally intractable for high-dimensional systems.

Building on these developments, PINNs have emerged as a promising hybrid paradigm that unifies data-driven learning and physics-based modeling. As illustrated in Figure 2, a PINN provides a unified framework applicable to both forward and inverse problems. In the forward setting, the network predicts physical field responses that satisfy governing equations for given material parameters and boundary conditions, whereas in the inverse setting, it identifies unknown material parameters from sparse or noisy measurements while preserving physical consistency. This bidirectional capability highlights the versatility of PINNs for integrating experimental data and physical modeling within a single learning architecture. Through this formulation, PINNs bridge the gap between conventional physics-based simulation and modern AI-driven discovery, laying the foundation for data-efficient and interpretable material characterization.

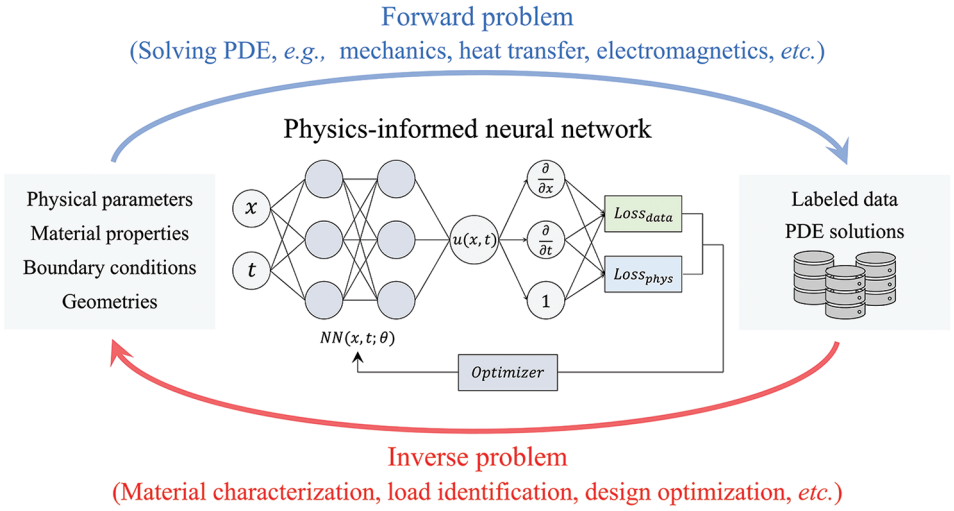


Figure 2. Schematic illustration of a physics-informed neural network (PINN) for forward and inverse problems. Image created by the authors. Abbreviation: PDE: Partial differential equation.

3. Opportunities of PINNs for material property characterization

A key claim of this perspective article is that PINNs uniquely bridge experimental data and governing physical laws, enabling material characterization approaches that remain reliable even under severe data scarcity. Rather than depending on large labeled datasets, PINNs embed the governing equations directly into the learning process, allowing the inference of material parameters and field quantities while maintaining physical consistency. By leveraging this physics-guided structure, PINNs mitigate overfitting and preserve admissible responses even when measurements are sparse or noisy, offering clear advantages over purely data-driven models. To better illustrate this concept, Figure 3 compares the conventional data-driven and physics-informed approaches for material property characterization. In the conventional approach, a neural network is trained on direct datasets such as stress–strain (σ, ε) pairs, which restricts its applicability to the training range and often results in non-physical extrapolation. In contrast, the physics-informed approach combines both direct and indirect data, such as displacement–force (u, F) measurements, with physical constraints and domain knowledge. By embedding this information into the training process, PINNs learn not only the mapping between inputs and outputs but also the underlying physical relationships. As shown in Figure 3, the conventional model fits the training data but diverges under unobserved conditions, whereas the physics-informed model maintains physically meaningful responses beyond the observed range. This capability makes PINNs particularly suitable for inverse material characterization, where experimental data are limited but the underlying physics is well established.

Beyond single-physics problems, the PINN framework can be readily extended to multi-physics systems such as thermo-mechanical, electro-mechanical, or flexoelectric materials by incorporating the corresponding governing equations into the physics-based loss term. This flexibility enables cross-domain inference, where different physical fields are coupled within a single learning framework. For instance, in thermoelectric systems, voltage measurements can be leveraged to infer thermal conductivity, enabling efficient estimation of transport properties without direct thermal measurements. Recent developments have further introduced Bayesian formulations of PINNs, which provide probabilistic uncertainty quantification and enhance robustness against sensor noise and incomplete data.[24] Such approaches extend the applicability of PINNs to real experimental settings while maintaining physical consistency and interpretability. Collectively, these capabilities position PINNs as a practical and robust tool for data-efficient, physics-consistent identification of material properties, bridging the gap between traditional inverse modeling and modern AI-driven materials design. The specific applications of this framework are discussed in Section 4.

4. PINN-based material characterization workflow

To illustrate how physics-informed learning can unify experiments, models, and simulations into a coherent characterization pipeline, this section outlines the integrated PINN-based workflow used in material

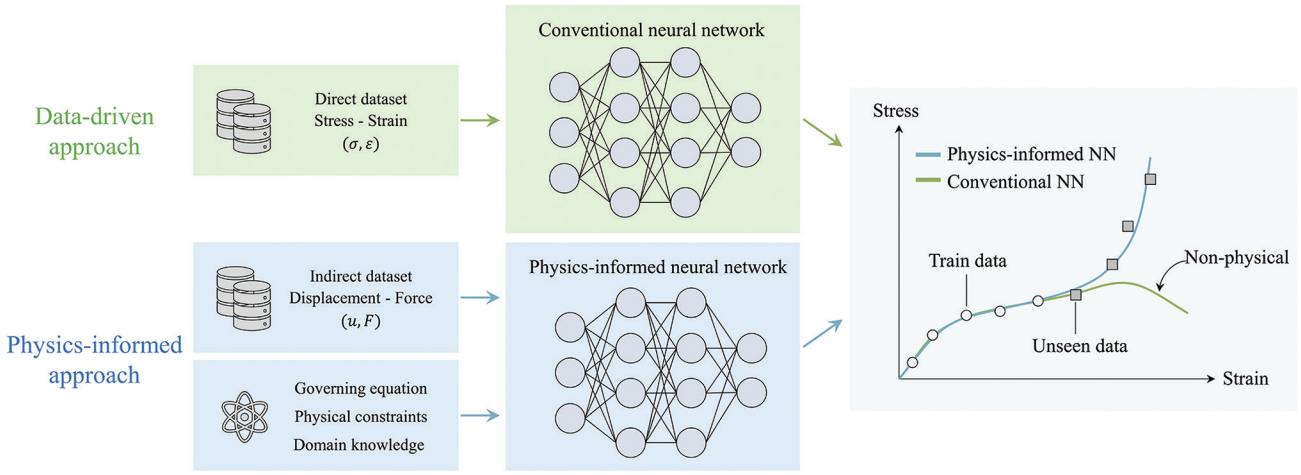


Figure 3. Comparison of conventional neural networks and physics-informed neural networks (PINNs) for material property characterization. Image created by the authors.
Abbreviations: NN: Neural network; PDE: Partial differential equation.

property identification. As illustrated in Figure 4, the workflow comprises five functional stages—data acquisition, dataset construction, constitutive modeling, PINN-based identification, and numerical validation—allowing physics-informed learning to bridge laboratory observations and computational mechanics in a coherent and data-efficient manner.

According to Figure 4, mechanical testing in stage (A) generates displacement, strain, or force data, which are organized in stage (B) into datasets that combine direct and indirect quantities for inference. Stage (C) introduces constitutive modeling, where analytical stress–strain relations may be replaced or augmented by neural representations whose parameters are treated as trainable variables. Stage (D) forms the core of the workflow: governing equations, boundary conditions, and measurement data are incorporated into a unified loss function that balances data fidelity with physical consistency. Different mathematical forms and training strategies may be used depending on the application, affecting computational cost without altering the underlying PINN formulation. Stage (E) then validates the inferred parameters or learned constitutive relations through finite element simulations or other high-fidelity solvers.

Although Figure 4 illustrates the workflow for mechanical characterization, the same structure extends naturally to thermal, electrical, thermoelectric, magneto-mechanical, and other coupled multi-physics systems through the incorporation of additional governing equations.[25-27] This generality underscores the role of PINNs as a versatile tool for physics-based characterization across diverse material classes. By unifying laboratory data, constitutive modeling, and computational verification in a single loop, PINNs provide the foundation for autonomous, closed-loop characterization pipelines central to digital-twin and AI-driven manufacturing systems.

5. Future directions and outlook

Looking ahead, establishing the foundations for reliable and widely adopted PINNs will require addressing several key challenges. The highest priority is establishing benchmark datasets that encompass diverse material systems, including elastic, thermal, electrical, and damage-sensitive materials, to ensure fair comparisons and reproducibility across studies. Equally important is integrating robust uncertainty quantification (UQ). Bayesian and probabilistic variants of PINNs provide a principled route to quantify predictive confidence and account for measurement noise, which is particularly important given the high sensitivity of inverse problems to sparse or biased data.[24,28-30]

Another major direction is improving the methodological scalability of PINNs for multi-scale and multi-physics systems. Although PINNs have shown strong potential in single-physics settings, extending them to coupled systems—such as thermo-mechanical or electro-chemical materials—remains challenging because multiple interacting field variables and complex constitutive couplings must be handled simultaneously. This gap highlights a notable opportunity for advancing data–physics fusion, where physics-informed learning could provide a unified representation across domains. Training such models, however, remains computationally demanding, particularly in high-dimensional or multi-physics settings where repeated differentiation of complex PDE operators amplifies the cost and causes instability

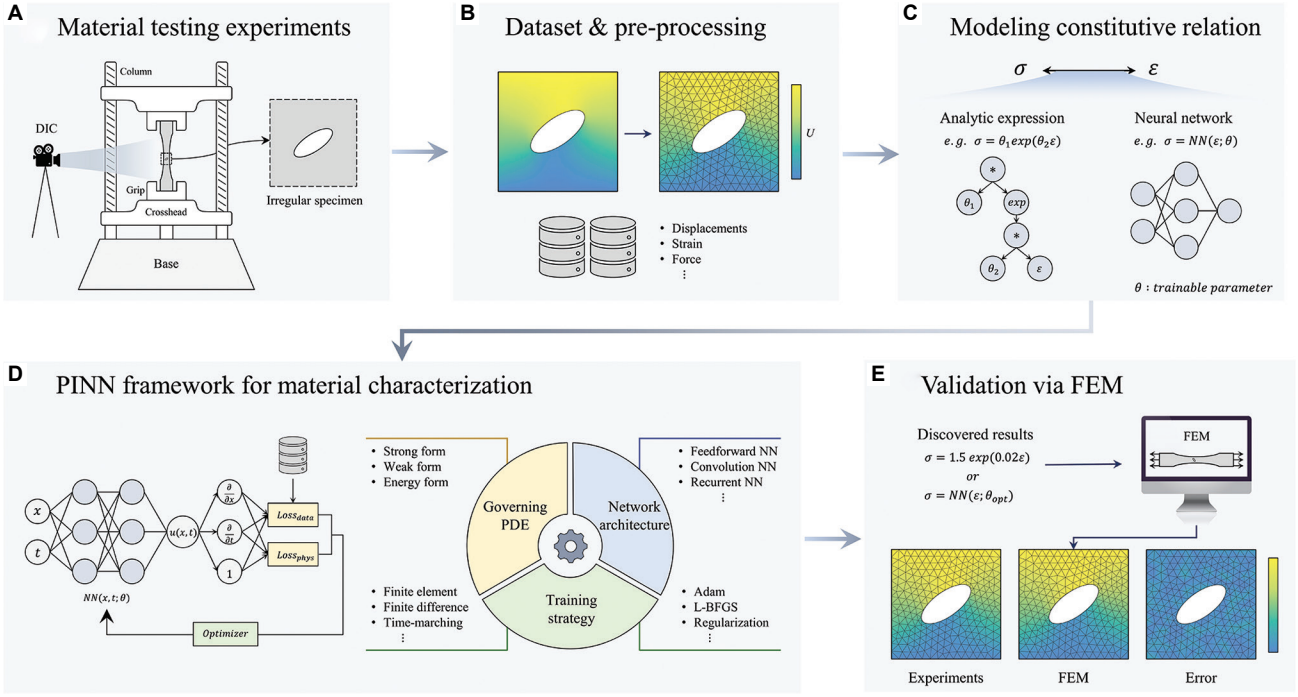


Figure 4. Representative workflow of physics-informed neural networks (PINNs) for material characterization. (A) Material testing and data acquisition; (B) Dataset construction and pre-processing; (C) Constitutive relation modeling; (D) PINN-based material characterization integrating governing equations, network architecture, and training strategy; (E) Validation via FEM simulations and experimental comparison. Image created by the authors. Abbreviations: FEM: Finite element method; L-BFGS: Limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm; NN: Neural network; PDE: Partial differential equation.

during optimization. Recent studies suggest that adaptive sampling, curriculum-based training, and operator-aware preconditioning can improve convergence stability, yet these techniques are not standardized and often require extensive problem-specific tuning.[31,32] In addition to computational cost, model-design decisions—including the choice of governing equations, network architectures, and loss formulations—remain highly problem-dependent and lack standardized best practices. A further challenge lies in extending PINNs to non-linear, history-dependent materials such as plasticity, viscoelasticity, and fracture, where internal state variables and path-dependent responses complicate learning.[33,34] Interpretability also remains limited, as the physical meaning encoded within learned latent representations is often unclear.

Finally, future material characterization pipelines are expected to evolve beyond classical PINNs toward hybrid, interpretable, and generative frameworks. Integrating generative modeling principles could alleviate the ill-posedness inherent in inverse problems, providing probabilistic descriptions of admissible material parameters rather than single deterministic estimates. Meanwhile, LLMs and knowledge-augmented networks may further enrich physical priors and assist in model design through embedded domain expertise.[35-37] In our view, achieving reliable, data-efficient, and autonomous material characterization will require models that not only learn from data and physics but also reason about uncertainty and adapt across material classes. Realizing this vision will establish PIML as a cornerstone of next-generation materials discovery and manufacturing.

From a broader perspective, we believe that advancing these directions will require closer collaboration across disciplinary boundaries. Experimentalists, computational mechanicians, materials scientists, and machine-learning researchers each contribute essential expertise—ranging from high-quality measurements and constitutive insight to scalable algorithms and physics-grounded architectures. Strengthening these connections will be crucial for translating physics-informed AI from promising methodologies into reliable, interpretable, and deployable tools for complex material systems. We view such interdisciplinary convergence as a necessary step toward building the next generation of characterization frameworks that seamlessly integrate data, physics, and intelligent automation.

6. Conclusion

The evolution of AI in materials science has advanced from purely data-driven modeling to generative, physics-

informed, and ultimately knowledge-guided paradigms. Within this progression, PINNs represent a pivotal step toward unifying data and physics for reliable and data-efficient material characterization. By embedding governing equations and physical constraints directly into the learning process, PINNs enable the inference of constitutive and material parameters even when experimental data are sparse, incomplete, or noisy. This integration of physical reasoning and ML has fundamentally reshaped the way materials are modeled, characterized, and designed.

Realizing the full potential of PIML, however, requires more than algorithmic innovation. The field must address challenges in data standardization, reproducibility, and computational scalability, while also developing robust frameworks for uncertainty quantification and interpretability. Collaborative efforts toward open benchmark datasets, standardized training protocols, and transparent evaluation metrics will be essential to ensure reliability and comparability across studies. Strengthening the integration between experimental and computational domains will further accelerate the transition from proof-of-concept demonstrations to practical engineering applications.

Looking forward, PIML is expected to evolve into hybrid and autonomous paradigms that combine the interpretability of physics with the adaptability of data-driven learning. Generative modeling principles, probabilistic inference, and knowledge-augmented architectures—potentially guided by LLMs—will further enhance physical understanding and enable intelligent design exploration. Ultimately, this convergence of physics, data, and intelligence is poised to transform materials research into a self-adaptive, closed-loop ecosystem capable of reasoning, learning, and discovery.

Acknowledgments

None.

Funding

This work was supported by the InnoCORE program (N10250154) and the National Research Foundation of Korea (RS-2023-00247245) grant of the Ministry of Science and ICT.

Conflict of interest

Seunghwa Ryu serves as an Editorial Board Member of this journal, but was not in any way involved in the editorial and peer-review process conducted for this paper, directly or indirectly. Separately, other authors declared that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Author contributions

Conceptualization: Hyeonbin Moon, Seunghwa Ryu
Visualization: Hyeonbin Moon, Jecheon Yu
Writing–original draft: Hyeonbin Moon, Junhyeong Lee
Writing–review & editing: All authors

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

Not applicable.

References

1. Lee H, Moon H, Lee J, Ryu S. Toward knowledge-guided AI for inverse design in manufacturing: A perspective on domain, physics, and human-AI synergy. *Adv Intell Discov*. 2025:e202500107.

   doi: 10.1002/aidi.202500107.

2. Lee J, Park D, Lee M, *et al*. Machine learning-based inverse design methods considering data characteristics and design space size in materials design and manufacturing: A review. *Mater Horiz*. 2023;10(12):5436-5456.

   doi: 10.1039/d3mh00039g

3. Bommasani R, Hudson DA, Adeli A, *et al*. *On the Opportunities and Risks of Foundation Models*. arXiv [Preprint]; 2021.

   doi: 10.48550/arXiv.2108.07258

4. Karpatne A, Atluri G, Faghmous JH, *et al*. Theory-guided data science: A new paradigm for scientific discovery from data. *IEEE Trans Knowl Data Eng*. 2017;29(10):2318-2331.

   doi: 10.1109/tkde.2017.2720168

5. Fuhg JN, Padmanabha GA, Bouklas N, *et al*. A review on data-driven constitutive laws for solids. *Arch Comput Methods Eng*. 2025;32:1841-1883.

   doi: 10.1007/s11831-024-10196-2

6. Kirchdoerfer T, Ortiz M. Data-driven computational mechanics. *Comput Methods Appl Mech Eng*. 2016;304: 81-101.

   doi: 10.1016/j.cma.2016.02.001

7. Zheng X, Zhang X, Chen TT, Watanabe I. Deep learning in mechanical metamaterials: From prediction and generation to inverse design. *Adv Mater*. 2023;35(45):e2302530.

   doi: 10.1002/adma.202302530

8. Zeni C, Pinsler R, Zügner D, *et al*. A generative model for inorganic materials design. *Nature*. 2025;639(8055):624-632.

doi: 10.1038/s41586-025-08628-5

9. Karniadakis GE, Kevrekidis IG, Lu L, Perdikaris P, Wang S, Yang L. Physics-informed machine learning. *Nat Rev Phys*. 2021;3(6):422-440.

   doi: 10.1038/s42254-021-00314-5

10. Meng C, Griesemer S, Cao D, Seo S, Liu Y. When physics meets machine learning: A survey of physics-informed machine learning. *Mach Learn Comput Sci Eng*. 2025;1(1):1-23.

    doi: 10.1007/s44379-025-00016-0

11. Achiam J, Adler S, Agarwal S, *et al*. *GPT-4 Technical Report*. arXiv [Preprint]; 2023.

    doi: 10.48550/arxiv.2303.08774

12. Lee J, Kim JY, Kim H, Lee I, Ryu S. *IM-Chat: A Multi-Agent LLM Framework Integrating Tool-Calling and Diffusion Modeling for Knowledge Transfer in Injection Molding Industry*. arXiv [Preprint]; 2025.

    doi: 10.48550/arXiv.2507.15268

13. Park D, Moon H, Ryu S. *A Self-Correcting Multi-Agent Framework for Language-based Physics Simulation and Explanation*. engrXiv [Preprint]; 2025.

    doi: 10.31224/4723

14. Batzner S, Musaelian A, Sun L, *et al*. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat Commun*. 2022;13(1):2453.

    doi: 10.1038/s41467-022-29939-5

15. Lu L, Jin P, Pang G, Zhang Z, Karniadakis GE. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nat Mach Intell*. 2021;3(3):218-229.

    doi: 10.1038/s42256-021-00302-5

16. Carlsson LA, Adams DF, Pipes RB. *Experimental Characterization of Advanced Composite Materials*. Boca Raton: CRC Press; 2014.

    doi: 10.1201/b16618

17. Kefal A, Oterkus E. Displacement and stress monitoring of a Panamax containership using inverse finite element method. *Ocean Eng*. 2016;119:16-29.

    doi: 10.1016/j.oceaneng.2016.04.025

18. Nyshadham C, Rupp M, Bekker B, *et al*. Machine-learned multi-system surrogate models for materials prediction. *NPJ Comput Mater*. 2019;5(1):51.

    doi: 10.1038/s41524-019-0189-9

19. Willard J, Jia X, Xu S, Steinbach M, Kumar V. *Integrating Physics-based Modeling with Machine Learning: A Survey*. arXiv [Preprint]; 2020.

    doi: 10.48550/arXiv.2003.04919

20. Zhong X, Gallagher B, Liu S, Kailkhura B, Hiszpanski A, Han TYJ. Explainable machine learning in materials science. *NPJ Comput Mater*. 2022;8(1):204.

    doi: 10.1038/s41524-022-00884-7

21. Sun L, Gao H, Pan S, Wang JX. Surrogate modeling for fluid flows based on physics-constrained deep learning without simulation data. *Comput Methods Appl Mech Eng*. 2020;361:112732.

    doi: 10.1016/j.cma.2019.112732

22. Gültekin O, Moeineddin A, Cansız B, Sveric K, Linke A, Kaliske M. A physics-informed neural network model for the anisotropic hyperelasticity of the human passive myocardium. *Int J Numer Methods Eng*. 2025;126(14):e70067.

    doi: 10.1002/nme.70067

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

    doi: 10.1016/j.cma.2022.114823

24. Moon H, Cho H, Demeke W, Ryu B, Ryu S. *Thermal Conductivity Estimation of Thermoelectric Materials with Uncertainty Quantification using Bayesian Physics-Informed Neural Networks*. arXiv [Preprint]; 2025.

    doi: 10.48550/arxiv.2510.16723

25. Moon H, Park D, Yeo J, Ryu S. *Physics-Informed Neural Network Framework for Solving Forward and Inverse Flexoelectric Problems*. arXiv [Preprint]; 2025.

    doi: 10.48550/arXiv.2506.21810

26. Moon H, Park D, Cho H, Noh HK, Ryu S. Physics-informed neural network-based discovery of hyperelastic constitutive models from extremely scarce data. *Comput Methods Appl Mech Eng*. 2025;439:118258.

    doi: 10.1016/j.cma.2025.118258

27. Moon H, Lee S, Demeke W, Ryu B, Ryu S. Physics-informed neural operators for generalizable and label-free inference of temperature-dependent thermoelectric properties. *NPJ Comput Mater*. 2025;11(1):272.

    doi: 10.1038/s41524-025-01769-1

28. Li P, Grana D, Liu M. Bayesian neural network and Bayesian physics-informed neural network via variational inference for seismic petrophysical inversion. *Geophysics*. 2024;89(6):M185-M196.

    doi: 10.1190/geo2023-0737.1

29. Wang R, Kong W, Liu X, *et al*. A physics-informed Bayesian neural network model for probabilistic prediction of fatigue crack growth rate at different temperatures. *Int J Fatigue*. 2025;201:109184.

    doi: 10.1016/j.ijfatigue.2025.109184

30. Bastek JH, Sun W, Kochmann DM. *Physics-Informed Diffusion Models*. arXiv [Preprint]; 2024.

doi: 10.48550/arXiv.2403.14404

31. Wang F, Zhai W, Zhao S, Man J. A novel unsupervised PINN framework with dynamically self-adaptive strategy for solid mechanics. *J Comput Phys*. 2025;542:114373.

doi: 10.1016/j.jcp.2025.114373

32. Rathore P, Lei W, Frangella Z, Lu L, Udell M. *Challenges in Training PINNs: A Loss Landscape Perspective*. arXiv [Preprint]; 2024.

doi: 10.48550/arXiv.2402.01868

33. Lardy M, Tlili S, Gsell S. Inferring viscoplastic models from velocity fields: A physics-informed neural network approach. *J Non Newton Fluid Mech*. 2025;346:105512.

doi: 10.1016/j.jnnfm.2025.105512

34. Liu Y, Chen M, Zeng Q. Physics-informed identification of stress fields and thermo-viscoplastic model parameters for metals from full-field data under impact loading. *Comput Methods Appl Mech Eng*. 2026;449:118568.

doi: 10.1016/j.cma.2025.118568

35. He X, You L, Tian H, Han B, Tsang I, Ong YS. *Lang-PINN: From Language to Physics-Informed Neural Networks Via a Multi-Agent Framework*. arXiv [Preprint]; 2025.

doi: 10.48550/arXiv.2510.05158

36. Qi Y, Xu R, Chu X. *FeaGPT: An End-to-End Agentic-AI for Finite Element Analysis*. arXiv [Preprint]; 2025.

doi: 10.48550/arxiv.2510.21993

37. Jiang Q, Karniadakis G. *AgenticSciML: Collaborative Multi-Agent Systems for Emergent Discovery in Scientific Machine Learning*. arXiv [Preprint]; 2025.

doi: 10.48550/arXiv.2511.07262

ACCSCIENCE PUBLISHING

International Journal of AI for Materials and Design

ORIGINAL RESEARCH ARTICLE

A data-efficient machine learning approach for amorphous Fe-based bulk metallic glass fabrication in powder bed fusion

Jungyeon Kim[1], Sangjun Jeon[1,2], Seong Je Park[3], and Seung Ki Moon[1,4]*

[1]School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore

[2]Department of Future Convergence Engineering, Kongju National University, Cheonan, Republic of Korea

[3]School of Mechanical Engineering, Gyeongsang National University, Jinju, Republic of Korea

[4]Singapore Centre for 3D Printing, School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore

(This article belongs to the *Special Issue: AI Usage in the Analysis of the Additive Manufacturing Process*)

Abstract

The widespread adoption of bulk metallic glasses (BMGs) in aerospace and biomedical industries requires topology-optimized architectures that conventional manufacturing cannot achieve. In response, BMGs have been investigated for powder bed fusion (PBF), but the process remains challenging due to narrow thermal windows, expensive feedstock, and limited data. This study introduces a constrained multi-objective Bayesian optimization framework to optimize key PBF printing parameters, including laser power and scan speed, to maximize hardness while preserving the amorphous state of the printed BMG. Hardness is optimized as the primary objective with density incorporated in the scalarization to regularize the search space, and amorphous retention is enforced through a feasibility probability learned by a logistic classifier. Surrogate models are compared, including Gaussian process, Bayesian additive regression trees, Bayesian multivariate adaptive regression splines (BMARS), and a Bayesian attention neural network. Acquisition scores are computed with constrained expected improvement and are maximized on a uniform grid over power and velocity. Superior predictive accuracy is obtained with BMARS, and 95% credible intervals are calibrated to the measurements. A high-hardness region at high power and low velocity is localized by the surrogates. A fully amorphous sample at 60 W and 1300 mm/s is produced, and a hardness of 1010.4 HV is measured in agreement with the predicted high-hardness band. In conclusion, the study establishes a data-efficient process-window discovery method for BMG PBF, produces an interpretable process map, and demonstrates a screening framework suitable for constrained experimental budgets.

Keywords: Additive manufacturing; Bayesian optimization; Bulk metallic glass; Powder bed fusion

***Corresponding author:**
Seung Ki Moon
(skmoon@ntu.edu.sg)

Citation: Kim J, Jeon S, Park SJ, Moon SK. A data-efficient machine learning approach for amorphous Fe-based bulk metallic glass fabrication in powder bed fusion. *Int J AI Mater Design*. 2025;2(4): 10-23.
doi: 10.36922/IJAMD025390036

Received: September 24, 2025

Revised: October 25, 2025

Accepted: October 30, 2025

Published online: November 13, 2025

1. Introduction

Additive manufacturing allows efficient production of complex parts by directly converting digital designs into solid objects, which allows additive manufacturing to

be widely used in biomedical, automotive, construction, and aerospace industrial applications.[1-4] Among the seven additive manufacturing techniques based on ISO/ASTM 52900, powder bed fusion (PBF) is a printing method suitable for fabricating high-accuracy metal parts.[5,6] PBF achieves high-resolution fabrication by selectively melting fine layers of metal powder using a laser. PBF can fabricate complex geometries such as lattice structures and minimal surfaces that are difficult or impossible to produce with conventional technologies.[7,8] It also improves the usability of various metal materials. Bulk metallic glass (BMG) is an example of a unique material that benefits from PBF. Amorphous alloys of BMG exhibit high strength, hardness, corrosion resistance, and wear resistance.[9-11] However, traditional casting techniques limit the use of BMG to small parts with simple shapes, because BMG must solidify faster than the time it takes for crystals to form.[12-14] PBF provides extremely rapid cooling rates, up to 10^7 K/s, required to retain the amorphous structure, which makes it suitable for producing complex or larger BMG components.[15] However, applying PBF to BMGs remains challenging since high hardness, elastic limit, and corrosion resistance can be achieved only if the melt pool cools at 10^6–10^8 K/s.[16,17] Any local overheating, oxygen uptake, or insufficient enthalpy can trigger nanocrystallization, embrittlement, and cracking.[18,19] Each failed trial is a major loss of time and material and therefore makes data-efficient optimization necessary, since BMG powders are expensive and the amorphous processing window is extremely narrow.

Recent studies have been employing artificial intelligence (AI) approaches to reduce experimental efforts. Supervised convolutional and recurrent networks can classify defects or forecast thermal histories, but the models require thousands of labeled samples.[20-22] Bayesian optimization (BO) overcomes the high data demand by building an uncertainty-aware surrogate and choosing each new experiment where the expected improvement (EI) is highest.[23-26] In projection multiphoton lithography, Johnson *et al.*[27] reduced geometric error to the measurement limit after only four BO iterations by coupling a Gaussian process (GP) surrogate with an EI policy geometry. For laser directed-energy deposition (DED), Karkaria *et al.*[28] embedded a Bayesian long short-term memory surrogate in a digital-twin loop and introduced the Time Series Process Optimization (BOTSPO) algorithm, which achieved optimal ten-parameter laser-power profiles with far fewer function evaluations than Latin-hypercube or Markov chain Monte Carlo (MCMC) searches. Beyond additive manufacturing, Lei *et al.*[24] demonstrated that Bayesian additive regression trees (BART) and Bayesian multivariate adaptive regression splines (BMARS) can outperform GPs when the response surface is high-dimensional or exhibits abrupt transitions, arguing for more flexible surrogates in automated design frameworks. These previous studies confirm that BO delivers the sample efficiency needed for costly experiments and that surrogate choice critically affects convergence.

Despite recent advances, two critical gaps exist in the literature. First, applications of BO to metallic systems have focused almost exclusively on DED or wire-arc additive manufacturing, which leaves PBF unexplored.[29-32] Second, the complex amorphous-crystalline transition with the processability of BMGs has not been modeled in a systematic way.[33-36] Glass formation occurs only when the melt pool cools faster than a critical rate that prevents crystal nucleation. Cooling rates below the critical rate initiate nanocrystals to form and their growth during subsequent layers, which reduces hardness and toughness.[33,37] The resulting discontinuity in the process maps due to nanocrystals limits the development of a smooth surrogate model, but no published study has investigated such a phenomenon. The thermo-kinetic constraints in BMG PBF differ fundamentally from the gradual microstructural trends in Ti-6Al-4V or Inconel 718, so results obtained for conventional alloys cannot be transferred for analysis of BMG. Consequently, PBF still lacks a data-efficient strategy for identifying laser power-speed combinations that preserve glass structure while maximizing mechanical performance.

This study addresses the unmet need by integrating constrained, multi-objective BO with PBF of an FeCrBSiC BMG that has not previously been fabricated using PBF. The proposed framework starts with initial dataset formation through twelve randomly selected laser power and scan speed combinations. Four different surrogate models, which include GP, BART, BMARS, and a Bayesian-Attention Neural Network (BANN), are then trained to predict hardness, while density is included in the scalarization step. A probabilistic classifier is applied to enforce amorphous retention as a feasibility constraint. A constrained expected improvement (CEI) acquisition function is employed to evaluate the balance between hardness and density while excluding regions with high crystallization risk. Iterative retraining of the surrogates after each build is performed to refine the process map and direct the search toward the optimum power and speed region. A final validation build is conducted to confirm the predictive accuracy of the best-performing surrogate.

The four surrogates were chosen to capture smooth trends, sharp transitions, and predictive accuracy under limited data. A GP surrogate was used because GP-based BO showed fast convergence in the printing process and provided calibrated uncertainty for sample-efficient search.

The two non-parametric surrogates, including BART and BMARS, were added to capture local non-linearity and abrupt transitions in the power–speed domain, which were difficult to be well modeled by a smooth GP assumption. A compact BANN was included to provide neural function approximation with posterior uncertainty under a small dataset using approximate Bayesian inference. The selection of the four surrogates aligns with prior BO studies in 3D printing and with adaptive surrogate design in materials experimentation.

This paper is organized as follows: Section 2 discusses the experimental setup and dataset. Section 3 describes the surrogate modeling and optimization procedure. Section 4 reports the results of model comparison and optimization. Section 5 discusses the physical implications and novelty. Section 6 provides the conclusions and future research directions.

2. Data collection and methodology

2.1. Experimental setup and conditions

The powder and equipment employed were FeCrBSiC (Dura-Metal, Singapore) and EOS M100 (EOS GmbH, Germany), respectively. Laser energy densities were experimentally selected between 25.2 J/mm^3 and 35.7 J/mm^3 for sample fabrication, which were calculated using the volumetric energy density (E) equation:

$$E = \frac{P}{vht} \quad \text{(I)}$$

Where P represents the laser power, v is the scan speed, h is the hatch distance, and $t$is the layer thickness. E represents the average energy delivered per unit volume of material during one layer printing and serves as an integrated measure of process input. The hatch distance and layer thickness were fixed at 70 μm and 20 μm, respectively. Laser energy densities below 25.2 J/mm^3 were found to be insufficient to achieve full melting, leading to poor layer bonding and incomplete builds, which lowered both hardness and surface density. However, E values above 35.7 J/mm^3 resulted in excessive energy input, causing defects such as cracking due to thermal stress accumulation and promoting local crystallization. Crystallization led to embrittlement and a decrease in attainable hardness. Besides, three representative laser powers (60 W, 70 W, and 120 W) were randomly selected to explore the effects of laser power and laser scan speed on the resulting part quality. At each laser power, the scan speed was varied within the defined volumetric energy density range to fabricate 10 × 10 × 10 mm^3 cubic samples. After fabrication, the samples were separated from the substrate using wire cutting. X-ray diffraction (XRD) analysis was performed using a SmartLab diffractometer (Rigaku, Japan) with Cu Kα radiation (λ = 0.1541 nm). The XRD patterns were recorded in the 2θ range of 25°–90°. In contrast, for density and hardness measurements, the samples were mounted and polished to ensure smooth cross-sectional surfaces. The surface densities were evaluated by analyzing cross-sectional images obtained via an optical microscope, Olympus DP22 (Evident, Japan), using ImageJ software. Vickers hardness tests were performed on the same polished surfaces via a Microscan hardness tester (Omron, Japan), respectively. The resulting dataset served as the experimental foundation for the machine learning-based optimization process described in the subsequent sections.

2.2. Surrogate model construction and data pre-processing

In this study, a Bayesian surrogate modeling approach is employed to capture the relationships between PBF process parameters and the resulting material responses, including hardness, density, and amorphous retention. A surrogate that can learn from sparse input data and quantify predictive uncertainty is required because the high cost and lengthy duration of BMG PBF trials limit the number of feasible experiments. Therefore, GP regression is chosen as a primary surrogate model. The GP surrogate models the unknown objective function in a Bayesian manner, where GP assumes any set of function values has a joint Gaussian distribution defined by a mean function and a covariance kernel.[29] After observing data, the GP posterior provides a predictive mean μ and variance σ^2 for any candidate input x, which serve as the model's prediction and uncertainty estimate, respectively. An automatic relevance determination radial basis function kernel with a constant and white-noise term was used for GP. The kernel parameters were optimized by maximizing log-marginal likelihood with 15 random restarts. Uncertainty quantification is crucial in BO, since predictions with the lowest uncertainty are used inside the algorithm to identify the most favorable predicted outcome, and predictions with high uncertainties provide meaningful information for new data.

While GPs are suitable for smooth, continuous functions, the BMG PBF can exhibit non-stationary or discontinuous behaviors such as sudden changes in properties near processing limits, and heteroscedastic noise. A standard GP with a stationary kernel often fails to capture non-stationarity and input-dependent noise.[24,38,39] Therefore, more flexible nonparametric Bayesian surrogates are evaluated to address the non-stationarity and heteroscedastic noise: BMARS and BART. BMARS and BART can capture complex nonlinear relationships and

interactions that a simple GP cannot represent. BMARS constructs a flexible polynomial basis to model localized trends, whereas BART integrates many shallow trees to accommodate sudden changes and complex interactions. Both BMARS and BART perform automatic variable selection and include Bayesian regularization, which helps avoid overfitting even with limited data. Both BMARS and BART implement a fully Bayesian inference procedure, which produces a predictive distribution at each input. BMARS employed 40 maximum terms, a second-degree basis, and a penalty of 0.5. BART used 30 trees with 60–200 burn-in and posterior draws. The resulting uncertainty quantification aligns with the GP surrogate's mean-variance output and supports direct integration into the acquisition computations.

Besides, BANN is also studied, which is an additional nonparametric model alongside GP, BMARS, and BART. A compact BANN surrogate is developed, which is designed specifically for the limited dataset of this study. One hidden layer with a rectified linear unit (ReLU) is used. ReLU is chosen to avoid saturation and to keep a locally linear function class that controls capacity under small datasets, which is consistent with practical guidance for small-data BANN.[40] Dropout of 0.25 is kept active during inference to approximate Bayesian inference and to provide calibrated epistemic uncertainty.[41] Two output heads are used to predict hardness and surface density, and each head produced a mean $\mu(x)$ and a log variance $s(x)$ where x is defined as the two-dimensional input process vector (P, v). Training is performed with the Gaussian negative log likelihood with L2 weight, which is expressed as follows:[41,42]

$$\mathcal{L}(x,y)=\frac{1}{2}s(x)+\frac{1}{2}\frac{\left(y-\mu(x)\right)^2}{\exp\{s(x)\}}+\lambda\theta_2^2 \quad \text{(II)}$$

In Equation II, y represents the observed target, hardness, and surface density. λ is a weight-decay coefficient that sets the strength of L2 regularization, θ represents the set of all trainable weights in the neural network, and $\|\theta\|_2^2$ represents the squared L2 norm or sum of squares of all weights in θ. λ is set to 0.005, learning rate is 0.003, and the early stopping patience is fixed at 60 epochs.

At prediction, M stochastic forward passes with dropout were used. The predictive mean $\bar{\mu}(x)$ is expressed as follows:

$$\bar{\mu}(x)=\frac{1}{M}\sum_{m=1}^{M}\mu^{(m)}(x) \quad \text{(III)}$$

Where $\mu^{(m)}(x)$ is the predicted mean from the m-th stochastic pass with dropout, and M is the number of Monte Carlo passes.[40,41] This estimator was used as the MC-dropout approximation to Bayesian inference.[40,41]

Separate surrogate models predict hardness and density. Input-dependent noise is incorporated by allowing predictive uncertainty to vary across the process parameter space rather than imposing constant observation noise. A heteroscedastic GP formulation models the noise variance as a function of the input, for example, via an auxiliary GP over the noise term. Experimental observations indicate that regions of very high laser power or scanning speed yield greater outcome variability from unstable melt pool dynamics, whereas other regions behave more consistently. Local adaptation of the noise model improves uncertainty quantification because predictive variance rises where data is sparse or noisy and falls in well-characterized regions. The heteroscedastic modeling proves to be critical for BMG printing, where abrupt shifts in process conditions can drastically affect part quality variability.

Amorphous phase preservation is calculated as a probabilistic feasibility constraint in the optimization process rather than as a distinct objective. An ideal printer parameter combination must preserve the minimum critical amorphous fraction in the printed part. A probabilistic feasibility surrogate predicts whether a given set of process parameters produces an amorphous structure. Each experiment is labeled with a binary indicator , which describes the structural order between 0 and 1 as follows:

$$\Delta=\begin{cases}0\ (crystalline)\\ 1\ (amorphous)\end{cases} \quad \text{(IV)}$$

A logistic classifier learns a latent function $g(x)$ from the feasibility labels. The feasibility probability at x equals $\Phi[g(x)]$, with the standard-normal cumulative distribution evaluated at $g(x)$. The predicted probability $p(x)$ gives a continuous estimate from zero to one of meeting the amorphous preservation requirement. Training on feasible versus infeasible outcomes produces predictive uncertainty in regions with limited or conflicting data. The logistic classifier delivers both a feasibility prediction and a confidence level. Constrained BO multiplies the EI metric by $p(x)$ so that parameter settings predicted to violate the amorphous criterion receive a negligible acquisition score. Modeling the feasibility constraint via a probabilistic classifier follows a standard practice in constrained BO, and the acquisition function balances potential objective gains against the requirement to preserve the amorphous phase while accounting for uncertainty.

Four surrogate models, including GP, BART, BMARS, and BANN, are developed to model the physical responses under limited data and to provide predictive uncertainty

for BO. Data pre-processing is applied before model fitting. Raw measurements of laser power, scan speed, Vickers hardness, and surface density are used as recorded. GP targets are normalized using the model's internal normalization option. Vickers hardness and surface density were standardized for training and were inverse-scaled at prediction for BART and BANN. Targets were kept in original units for BMARS. The feasibility classifier for amorphous retention is trained on original power and scan speed with binary labels.

2.3. Acquisition and optimization procedure

Figure 1 illustrates the overall sequential workflow, which collects process data; trains surrogate models, applies Pareto efficient global optimization (ParEGO) scalarization, evaluates CEI, and selects the next build parameters. Surrogate models for hardness, density, and amorphous-phase feasibility are supplied to a constrained multi-objective BO that identifies optimal PBF process parameters. The optimization strategy balances two requirements: (i) concurrent improvement of hardness and density, and (ii) compliance with the amorphous-phase feasibility requirement. The ParEGO method is employed to reduce multiple objectives to a single-objective acquisition problem at each iteration.[42] Besides, scalarization is used to combine the surrogate predictions into a weighted sum, where a weight vector,, balances hardness and density. At each iteration, a random weight between 0 and 1 is drawn from the uniform distribution, and the composite objective is defined as:

$$f_{scalar}(x) = w \cdot \tilde{H}(x) + (1-w) \cdot \tilde{D}(x) \quad \text{(V)}$$

Where $\tilde{H}(x)$ and $\tilde{D}(x)$ are the predicted hardness and density, respectively. Variation of the weight w across iterations explores Pareto trade-offs between the two metrics: a larger w places more emphasis on predicted hardness at the expense of density, whereas a smaller w places more emphasis on predicted density at the expense of hardness. After multiple iterations with distinct random weights, ParEGO produces a solution set that approximates the Pareto-optimal frontier for both objectives. ParEGO extends single-objective EI using the same GP surrogates and acquisition framework and proves effective under the limited experimental feasibility, which is typical for BMG PBF research.

After scalarization, the multi-objective problem becomes a single-objective constrained optimization at each iteration, with the selected weight defining the hardness-density trade-off. Selection of the next experimental condition proceeds via maximization of the acquisition function, which evaluates each candidate according to the surrogate's predicted mean $\mu(x)$ and variance $\sigma^2(x)$. The EI criterion captures both exploitation and exploration by computing the expected gain over the current best value f^*. For a candidate x, the improvement random variable is defined as:

$$I(x) = \max\left(f_{scalar}(x) - f^*, 0\right) \quad \text{(VI)}$$

Where $f_{scalar}(x)$ represents the surrogate's predictive distribution. Under a Gaussian posterior, the $EI(x)$ is described in a closed form as:

$$EI(x) = \left(\mu(x) - f^*\right)\Phi(Z) + \sigma(x)\phi(Z) \quad \text{(VII)}$$

Where $\sigma(x)$ is the predicted standard deviation, and Z is the measure of the distance between the surrogate's predicted mean and f^*.[43] The functions Φ and ϕ represent the cumulative distribution function and probability density function of a standard normal distribution, respectively. The closed-form results from integrating the improvement variable under the surrogate's Gaussian posterior. The first term, $(\mu(x)\text{-}f^*)\,\Phi(Z)$, captures the expected gain if the predicted mean exceeds the current best, weighted by the corresponding probability. The

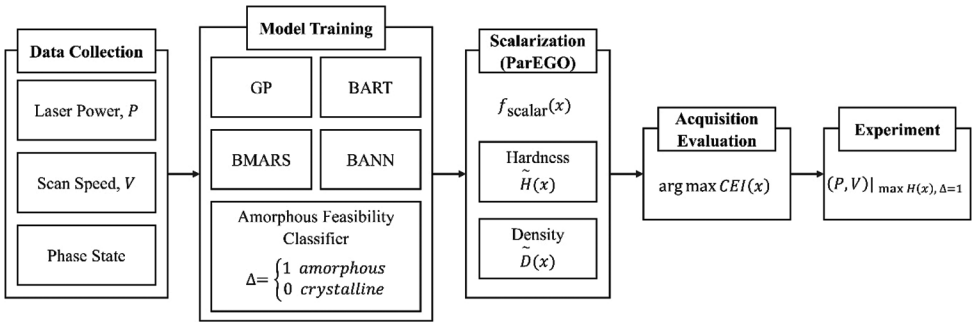


Figure 1. Constrained multi-objective Bayesian optimization workflow for BMG PBF parameter selection
Abbreviations: BANN: Bayesian attention neural network; BART: Bayesian additive regression trees; BMARS: Bayesian multivariate adaptive regression splines; BMG: Bulk metallic glass; CEI: Constrained expected improvement; GP: Gaussian process; ParEGO: Pareto efficient global optimization; PBF: Powder bed fusion.

second term, $\sigma(x)\phi(Z)$, represents the value of sampling in regions with high uncertainty, where unexpectedly large improvements remain possible. Consequently, $EI(x)$ attains large values in regions of high predicted performance or high uncertainty, which provides a balance between exploitation and exploration.

A feasibility constraint on amorphous-phase preservation multiplies the EI criterion. The logistic classifier estimates $p(x)$, which is the probability that parameter vector x yields an amorphous build. The CEI can be computed as:

$$CEI(x) = p(x)\,EI(x) \quad \text{(VIII)}$$

A candidate x with a large EI but low feasibility probability receives a near-zero CEI score. Only parameter settings with high $p(x)$ can retain the full EI value. As a result, the acquisition function balances gains in hardness and density against the requirement to preserve the amorphous structure. Regions where $p(x)$ is close to 1 reduce CEI to standard EI, whereas regions where $p(x)$ is close to 0 yield negligible CEI regardless of predicted performance.

Maximization of the CEI employs a uniform grid over the laser-power and scan speed ranges. The acquisition score $CEI(x)$ is computed at each grid point in milliseconds, and the point with the maximum score becomes the next experimental condition. A full-grid search identifies the maximum global CEI within the grid resolution and avoids the risk of settling in local maxima that can affect gradient-based or stochastic search methods.

2.4. Statistical analysis

All computations ran on a Windows 10 laptop equipped with an Intel Core i5 13420H (2.10 GHz) and 16 GB RAM. The Python 3.6 environment used scikit-learn 1.3 for Gaussian-process regression and logistic regression, bartpy for BART, pyEarth 0.1 for BMARS, and TensorFlow Probability 0.22.5 for the BANN. A global NumPy random seed fixed every stochastic step in the workflow.

Model accuracy was evaluated using leave-one-out cross-validation (LOOCV) for reliable estimation of predictive performance under the small sample condition. LOOCV provides reliable Bayesian predictive accuracy relative to information-criteria approaches for limited data.[44,45] In each iteration, one observation was used as the test case, and the remaining observations served as the training set. The procedure was repeated over all observations. The average prediction error across iterations provided a stable estimate of generalization.

Root mean square error (RMSE) and coefficient of determination (R^2) were computed from the LOOCV results for both hardness and surface density predictions. The logistic classifier that enforces the amorphous-phase constraint was also validated under the same LOOCV framework.

BO progress was tracked by recording the best hardness observed after each iteration. Matplotlib produces all figures, and pandas records the accuracy and optimization tables. No further hypothesis testing, calibration scoring, or predictive interval assessment is necessary for the objectives of this study.

3. Results

3.1. Model comparison

Figure 2 illustrates the predicted hardness response surfaces as functions of laser power and scan velocity for each surrogate model: GP, BMARS, BANN, and BART. The figure presents four distinct surfaces, each representing a different modeling approach applied to the same initial dataset, with warmer colors indicating higher hardness values.

Figure 2A displays the GP surrogate model results. The GP surface exhibited a smooth gradient, with hardness values rising from approximately 700 HV at low power and high velocity to over 1000 HV at high power and low velocity. Hardness increased with laser power and decreased with scan velocity in the GP predictions, which is consistent with the stationary kernel's ability to represent the dominant process-property relationship. The smoothness of the GP surface reflected its underlying assumption of continuous, gradual changes in the response, which proved suitable for the observed data. Figure 2B shows the BMARS surrogate model, which showed a similar overall gradient to the GP but with sharper transitions between regions of differing hardness. Hardness increased significantly with higher power and lower velocity, where the value reached its maximum in the same region. The spline-based formulation of BMARS allowed the model to fit both gradual and abrupt changes, resulting in a surface that captured the main trend while also highlighting localized variations. The flexibility of BMARS produced a clearer separation between high- and low-hardness regions compared with the GP model.

BANN surrogate model predictions are shown in Figure 2C. The BANN surface appeared as a broad plateau across the domain with multiple local peaks and valleys. A gentle rise was found toward high power and low velocity regions, but the model did not form a clear ridge. The response range was compressed, indicating limited sensitivity to the inputs compared with GP and BMARS. Figure 2D illustrates the BART surrogate model,

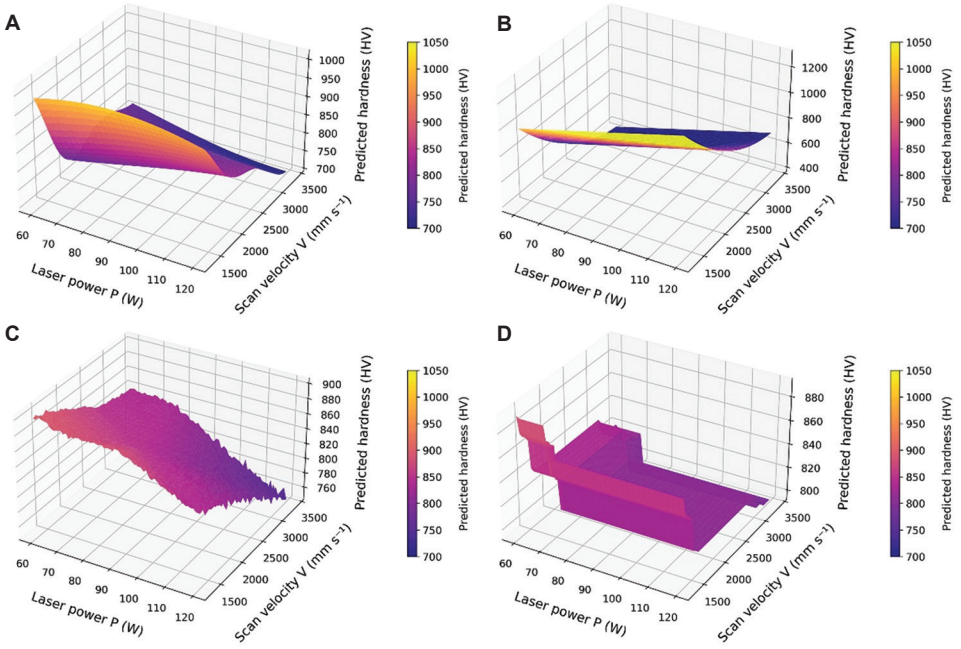


Figure 2. Predicted hardness response surfaces as functions of laser power (P) and scan velocity (V) for each surrogate model: (A) Gaussian process, (B) Bayesian multivariate adaptive regression splines, (C) Bayesian attention neural network, and (D) Bayesian additive regression trees

which partitions the design space into piecewise-constant regions of uniform hardness. A distinct high-hardness surface appeared at high power and low velocity, sharply separated from a lower-hardness regime elsewhere. The tree-ensemble structure of BART displayed abrupt changes in hardness and facilitated the identification of discrete processing regimes. The stepwise approximation effectively captured sharp transitions in the response surface.

All four models successfully identified the peak-hardness region at high power and low velocity, but the models differed in representation and smoothness. The GP and BMARS surrogates provided smooth interpolations of the underlying trend, where BMARS showed slightly sharper regional transitions. The BART model outperformed in isolating distinct hardness regimes through its stepwise approximation, whereas the BANN model exhibited excessive variability across the domain.

3.2. Surrogate-model accuracy

Figure 3 shows predicted hardness against laser power at 1500 mm/s. Blue curves are posterior means. Shaded regions mark the 95% credible intervals. Red symbols are the measured specimens.

The BMARS captured the monotonic rise in hardness as illustrated in Figure 3B. Hardness increased from about 880 HV at 60 W to roughly 1230 HV at 120 W, which matched the experimental data. The 95% confidence interval remained narrow, and every experimental data point lay inside the band. The tight span indicates precise uncertainty quantification. The GP model in Figure 3A followed the same trend but appears smoother. The GP model underestimated the peak by nearly 50 HV, with a broader confidence interval. The wider band indicated higher epistemic uncertainty.

The BANN model in Figure 3C showed weak dependence on power. The posterior mean remained near a flat level across the range. The 95% confidence interval was wide and enclosed the two high-hardness data points. The model did not recover the upward trend and reflected low confidence in the prediction rather than overconfidence. Besides, the BART model in Figure 3D predicted a flat segment beyond 90 W. The mean deviated from the highest experimental data point by more than 100 HV. The BART model's interval stayed narrow but excluded the data points. The model understated both the mean and the variance in the high-power regime.

The prediction performances of four different models are summarized in Table 1. BMARS yielded an RMSE of 36 HV and an R^2 of 0.89. The GP followed with RMSE of 46 HV and R^2 of 0.83. BART and BANN exceeded 90 HV in RMSE and fell below 0.35 in R^2 values.

3.3. Optimization outcomes

Figure 4 compares the measured and predicted hardness values for each surrogate under LOOCV. Points distributed near the dashed 45° lines represent accurate

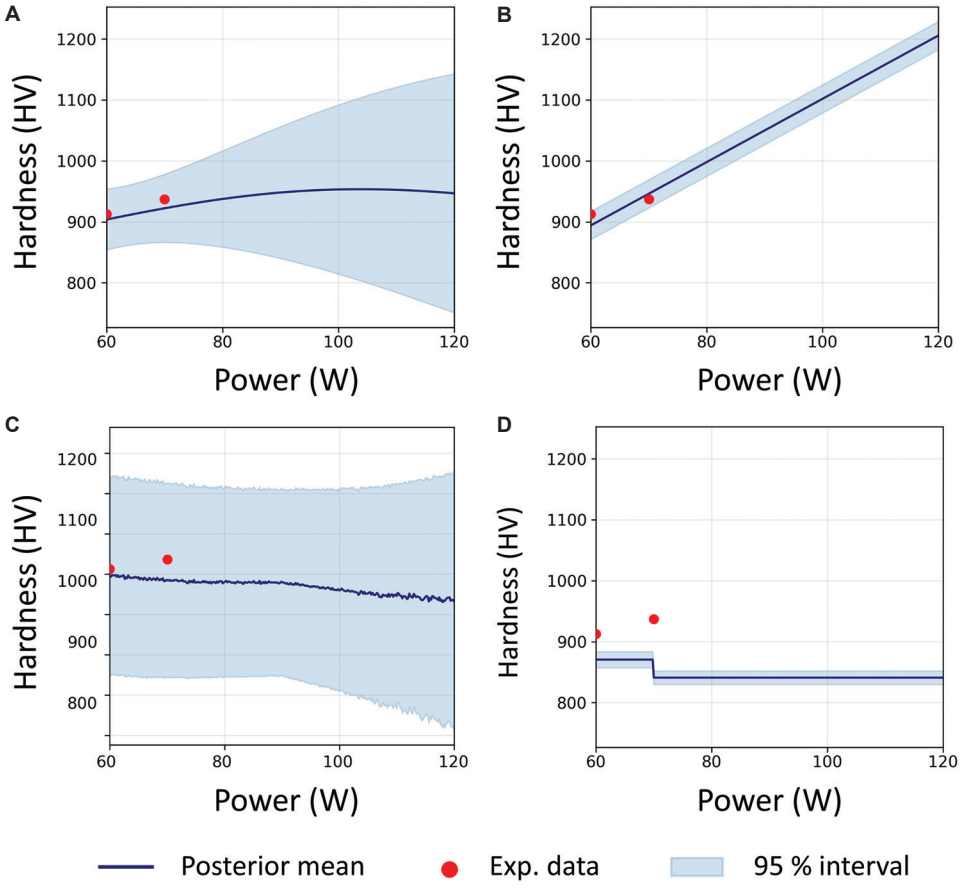


Figure 3. Hardness predictions of different models at scan velocity of 1500 mm/s: (A) Gaussian process, (B) Bayesian multivariate adaptive regression splines, (C) Bayesian attention neural network, and (D) Bayesian additive regression trees

Table 1. Surrogate accuracy metrics

Surrogate	RMSE (HV)	R^2
GP	45.59	0.832
BMARS	36.37	0.893
BANN	92.77	0.303
BART	96.59	0.244

Abbreviations: BANN: Bayesian attention neural network; BART: Bayesian additive regression trees; BMARS: Bayesian multivariate adaptive regression splines; GP: Gaussian process; RMSE: Root mean square error.

predictions. Both GP and BMARS models followed the ideal trend closely, whereas BANN and BART showed larger deviations. The GP points concentrated around the diagonal but slightly underestimated the highest hardness regions, whereas BMARS aligned most tightly with the measurements. The BANN model compressed the range of predicted values, and the BART model produced wider scatter that did not capture the full variation of the data. The plots confirm that BMARS and GP achieved the best agreement with experimental hardness, which is also consistent with the quantitative ranking in Table 1.

Figure 5 illustrates the running best hardness that each BO model recorded after every evaluation. BANN reached its final level by the third iteration at about 1080 HV. BMARS and GP were also stabilized by iteration 1 near 1027 HV and 1017 HV. BART reached its final level by iteration 2 near 1053 HV. Traces remained flat afterward, which indicates early convergence at different plateaus.

The rapid initial increase in the running best hardness during the first two to four iterations was consistent with the shapes of the surrogate response surfaces, as shown in Figure 2. A steep gradient near the high power and low velocity region was learned by BANN, and the region was prioritized by the acquisition. The same zone was reached by BART, but a lower peak was produced because the stepwise approximation underestimated the maximum. Curvature was preserved by BMARS, and the peak along the high hardness ridge was smoothed. The GP surface was smoothed more strongly, and the inflection near the crystalline boundary was blunted. Figures 2 and 3 show the same pattern for early entry into the high power and low velocity region and for peak height. GP and BMARS reached the region first, where BMARS captured sharper

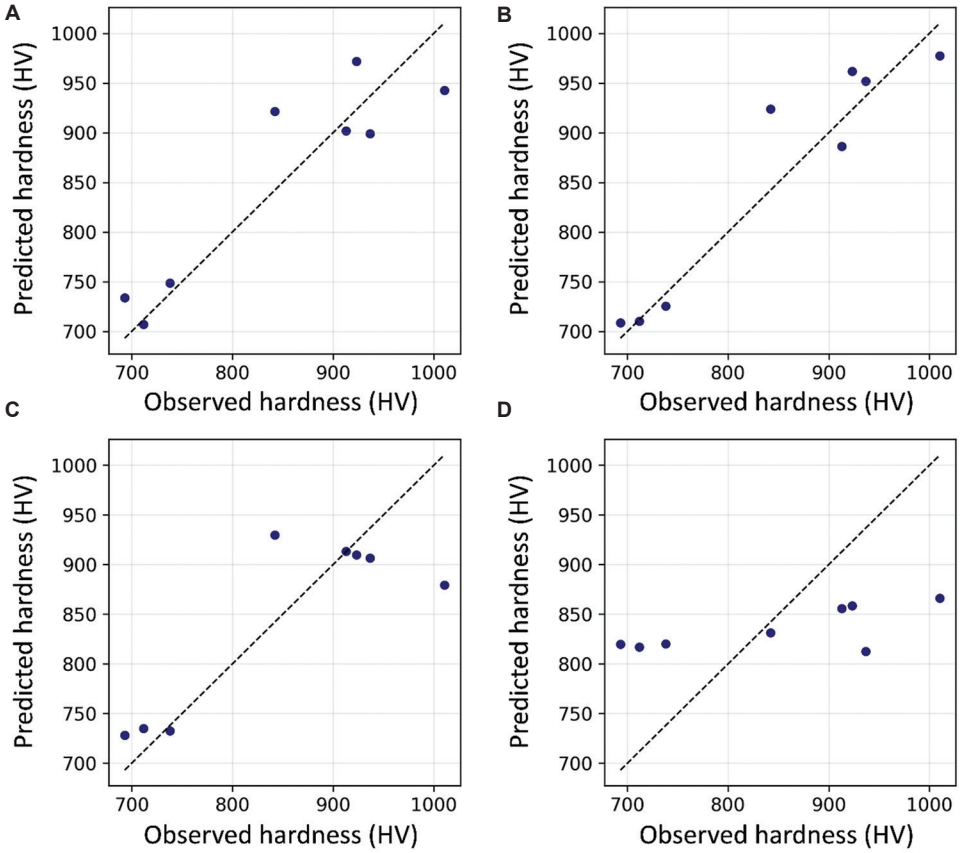


Figure 4. Observed versus predicted hardness for the four surrogate models under leave-one-out cross-validation: (A) Gaussian process, (B) Bayesian multivariate adaptive regression splines, (C) Bayesian attention neural network, and (D) Bayesian additive regression trees

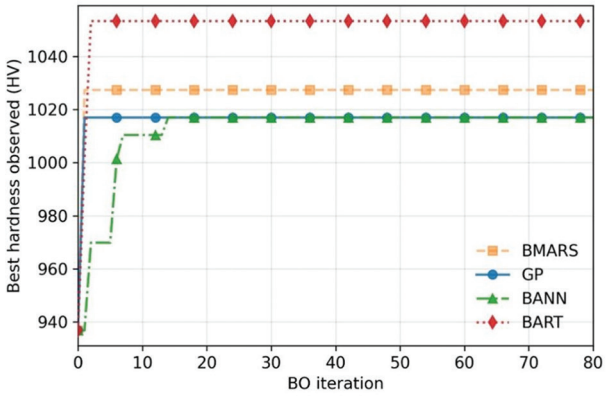


Figure 5. Optimization convergence trajectories for each surrogate model Abbreviations: BANN: Bayesian attention neural network; BART: Bayesian additive regression trees; BMARS: Bayesian multivariate adaptive regression splines; BO: Bayesian optimization; GP: Gaussian process.

local changes, GP smoothed the peak, BART reached the highest hardness peak, and BANN reached the optimum with the highest number of iterations.

Table 2 lists four quantities for each model, including maximum hardness, the first iteration at or above 95% of the maximum, and the recommended power and velocity. GP and BMARS reached the 95% criterion by iteration 1. BART met the criterion by iteration 2, and BANN reached the criterion by iteration 3.

Table 2. Optimization results

Surrogate	Maximum hardness (HV)	Iteration	Recommended power (W)	Recommended velocity (mm/s)
GP	1017	1	64	1300
BMARS	1027.4	1	66	1300
BANN	1017	3	60	1350
BART	1053.3	2	61	1300

Abbreviations: BANN: Bayesian attention neural network; BART: Bayesian additive regression trees; BMARS: Bayesian multivariate adaptive regression splines; GP: Gaussian process.

Near the crystalline boundary, a low probability of amorphous retention was predicted by the logistic classifier. EI was multiplied by the predicted probability, and the constrained score was near zero for power and velocity settings located in the boundary zone. Power and velocity settings in the boundary zone were ranked poorly

and were not selected. Sampling was concentrated along the ridge where high predicted hardness coincided with high feasibility. Sharp transitions in the BMARS and BART surfaces were mirrored in the selected evaluations.

A validation build was produced at a laser power of 60 W, a scan velocity of 1300 mm/s, and a hardness of 1010.4 HV was measured. A Vickers indentation from the build is shown in Figure 6. The value lay on the high hardness band in Figure 7 and confirmed movement toward a physically meaningful region.

Figure 7 maps predicted hardness across power and velocity using BMARS. Hardness values above 950 HV occupied the high-power and low-velocity region and appear as an elongated high-hardness band. Iso-hardness contours were oriented diagonally from higher power toward lower velocity and were nearly parallel to lines of constant volumetric energy density from Equation I. Hardness decreased when power fell below 80 W velocity exceeded 2600 mm/s. Beyond a volumetric density of about 60 J/mm^3, a saturation response was observed, and comparable hardness values were produced by many power and velocity settings. Trade-offs were present between hardness, build rate, and energy demand because regions with higher hardness often required lower velocity and higher power, which reduced throughput and increased energy input.

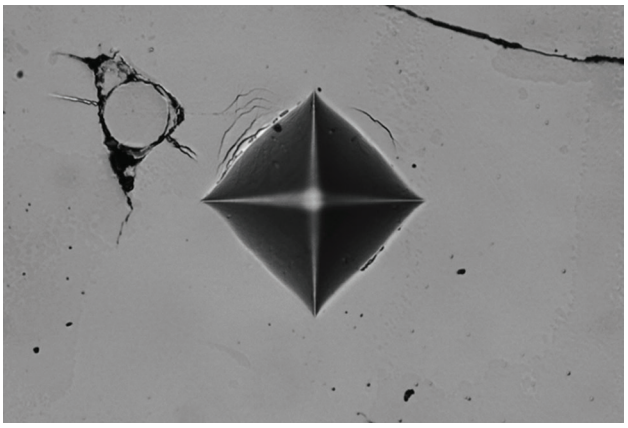


Figure 6. Vickers hardness indentation mark on the sample fabricated with a laser power of 60 W and a scanning speed of 1300 mm/s

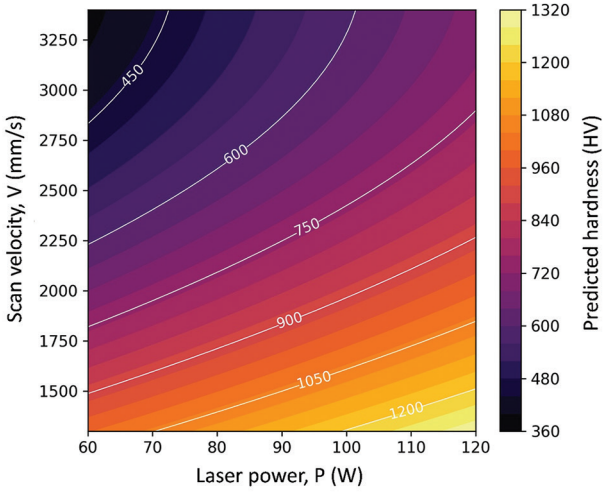


Figure 7. BMARS-predicted hardness across laser power and scan velocity
Abbreviation: BMARS: Bayesian multivariate adaptive regression splines.

4. Discussion

BART completed the search the fastest, and BMARS, GP, and BANN followed in that order. Although BMARS iteration finished second, the model produced the smallest error and the most reliable confidence intervals. The flexible spline bases in BMARS captured the broad rise in hardness as energy increased and preserved the sharp inflection point where crystallization began. In contrast, neither the stationary GP kernel nor BANN captured the sharp crystallization transition, as the models provided a smooth surface, a near-flat surface, respectively. Similar behavior appeared in high-dimensional MAX-phase screening, where Lei *et al.*[24] reported that BMARS and BART used fewer than 0.3% of the design space, whereas GP variants often ran out of budget BOTSPO. This research confirms that adaptive knots and tree splits give the surrogate the local freedom required when responses change abruptly.

The optimization results show that BMARS guided the search toward conditions that combined high predicted hardness with a high probability of amorphous retention. This focus avoided sampling power-velocity settings that thermokinetic theory identifies as infeasible and accelerated progress toward the best-performing region. Similar behavior has been reported in other additive manufacturing applications, such as projection multi-photon lithography and the BOTSPO digital-twin framework for DED, where BO reduced the number of required trials compared with conventional parameter sweeps.[28] In the present study, the method reached a setting that produced approximately 1030 HV hardness in physical builds while keeping the experimental effort low, demonstrating its potential for efficient process development in BMG PBF.

Hardness increased with volumetric energy density until about 60 J/mm^3 and remained nearly constant beyond that point. Above this volumetric energy density threshold, the melt pool temperature increased while the cooling rate stayed nearly unchanged, so the glass-forming driving force did not increase further. Below 80 W or

above 26 mm/s, hardness and density decreased because local cooling fell below the critical rate and nanocrystals formed. The logistic interaction term extracted from the classifier showed that raising scan speed compensates part of the excess power by shortening solidification time, which matches continuous-cooling-transformation theory for Fe-based glasses.

The optimum laser power and scan speed combination was located at the highest volumetric energy density input that maintained glass formation, and it corresponded to the first high-power pulse in the time-dependent power profiles selected by BOTSPO. Both static BMARS map and dynamic BOTSPO profiles converge on the same physical principle, which is to apply the highest energy input to close pores, and then cool rapidly to preserve the amorphous state. The consistency between the BMARS map and the BOTSPO profiles indicates that the BMARS map can provide a basis for initializing a process digital twin for PBF, which can then be refined layer by layer with *in situ* monitoring data, as demonstrated for DED control using the BOTSPO framework.[28]

The novelty of this study is defined by three main contributions. It is the first study to apply BO to PBF of BMG, which is a material system with an extremely narrow amorphous window for processability. The proposed framework introduces a constrained multi-objective approach that optimizes hardness and density while enforcing amorphous retention as a probabilistic feasibility condition. Furthermore, four surrogate models are systematically compared in terms of predictive accuracy, uncertainty calibration, and optimization efficiency, which establishes how surrogate selection influences convergence under limited experimental data. Overall, the contributions outlined above expand constrained BO's applicability to metastable alloys like BMG, which provides a data-efficient approach for process design in additive manufacturing under experimental constraints.

5. Conclusion

The objective of this study was to identify process parameters that preserved the amorphous structure of FeCrBSiC during PBF. The proposed approach was applied to a constrained multi-objective BO with surrogate modeling. This study established a data-efficient strategy for printing fully amorphous FeCrBSiC by PBF. A constrained multi-objective BO loop was applied to 80 PBF trials and identified a process condition that retained an amorphous FeCrBSiC structure with the maximum surface density and produced roughly 1030 HV Vickers hardness. Among the four surrogate models, BMARS exhibited the lowest RMSE, the highest R^2, and the only 95 % credible band that enclosed every measured value, which confirmed the superior fidelity compared to other models.

The experimental validation at 60 watts and 1300 mm/s further highlighted the predictive strength of the spline surrogate. The resulting contour map defined a narrow ridge of high hardness along the high-power, low-velocity border, which gave a clear processing window for engineering practice. A probabilistic classifier for amorphous-phase feasibility is combined with an uncertainty-aware acquisition function to determine experimental parameter selection. The combination of probabilistic feasibility modeling and uncertainty-aware acquisition makes the methodology applicable to other metastable alloys that require rapid solidification and precise energy control, and also supports the evaluation of new alloy systems for PBF processing.

Several limitations remain in this research. Hatch spacing and layer thickness were constant in this research, which makes keyhole porosity and track-shape effects unexplored. The surrogate was trained on one type of alloy, but transfer learning could reuse spline hyper-priors for alloys of lower glass-forming ability. Finally, hardness alone cannot guarantee structural reliability, so future work should also quantify fracture toughness and corrosion along the same energy range.

Future research will replace the squared exponential kernel with a spectral mixture or Gibbs kernel so that the length scale varies across the power and velocity domains. Reversible jump Markov chain Monte Carlo (RJMCMC) will detect change points that mark the transition from glass formation to crystallization and will switch kernel hyperparameters on each side. A hierarchical prediction derived from melt pool heat flow calculations will regularize predictions in regions with sparse data and will introduce parameters that correspond to physical quantities such as cooling rate. A multi-output GP will integrate hardness, density, and amorphous fraction, which will preserve coherent uncertainty when the optimizer chooses the next build. The acquisition function will account for estimated machine time and powder mass and will schedule batches that use fewer resources while still guiding the search toward the Pareto front. An embedded multi-sensor fusion-based pipeline will read optical and acoustic signals throughout the build, update the surrogate after each layer through Bayesian filtering, and adjust laser power and scan speed in real time to keep the melt pool within the glass-forming domain whenever drift emerges.

Acknowledgments

This research is supported by the Singapore Centre for 3D Printing (SC3DP) and the Ministry of Trade, Industry

and Energy (MOTIE, Korea) through the 2025 R&D HRD Program for Problem-Solving Based on Future Mobility Innovation Technology Development (Project No. P0023726).

Funding

This research is conducted by the Industrial Technology Innovation Program (KEIT project no. 20024344, Development of AI-based high carbon steel alloy design and sintering-based additive manufacturing technology for 7.0 L/Hr-level high-speed production of powertrain components with tensile strength over 1.0 GPa in the next-generation mobility) funded by the Ministry of Trade, Industry & Energy of the Republic of Korea.

Conflict of interest

Seung Ki Moon is an Editorial Board Member of this journal, but was not in any way involved in the editorial and peer-review process conducted for this paper, directly or indirectly. Separately, other authors declared that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Author contributions

Conceptualization: Jungyeon Kim
Formal analysis: Jungyeon Kim, Sangjun Jeon
Funding acquisition: Seung Ki Moon
Investigation: Jungyeon Kim, Seong Je Park
Methodology: Jungyeon Kim, Sangjun Jeon, Seong Je Park
Software: Jungyeon Kim
Writing–original draft: Jungyeon Kim, Sangjun Jeon, Seong Je Park
Writing–review & editing: Seong Je Park, Seung Ki Moon

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

The code and datasets supporting the findings of this study are available from the corresponding author upon reasonable request.

References

1. Tofail SA, Koumoulos EP, Bandyopadhyay A, Bose S, O'Donoghue L, Charitidis C. Additive manufacturing: Scientific and technological challenges, market uptake and opportunities. *Mat Today*. 2018;21(1):22-37.

   doi: 10.1016/j.mattod.2017.07.001

2. Kumar R, Kumar M, Chohan JS. The role of additive manufacturing for biomedical applications: A critical review. *J Manuf Process*. 2021;64:828-850.

   doi: 10.1016/j.jmapro.2021.02.022

3. Blakey-Milner B, Gradl P, Snedden G, *et al*. Metal additive manufacturing in aerospace: A review. *Mater Des*. 2021;209:110008.

   doi: 10.1016/j.matdes.2021.110008

4. Rehman M, Yanen W, Mushtaq RT, *et al*. Additive manufacturing for biomedical applications: A review on classification, energy consumption, and its appreciable role since COVID-19 pandemic. *Prog Addit Manuf*. 2023;8(5):1007-1041.

   doi: 10.1007/s40964-022-00373-9

5. *ISO/ASTM 52900:2021(en), Additive Manufacturing - General Principles — Fundamentals and vocabulary*. Available from: https://www.iso.org/obp/ui/#iso: std:iso.astm:52900:ed-2:v1:en [Last accessed on 2025 Aug 10].

6. Wang J, Jeong SG, Kim ES, Kim HS, Lee BJ. Material-agnostic machine learning approach enables high relative density in powder bed fusion products. *Nat Commun*. 2023;14(1):6557.

   doi: 10.1038/s41467-023-42319-x

7. Park SJ, Heogh W, Yang J, *et al*. Meta-structure of amorphous-inspired 65.1Co28.2Cr5.3Mo lattices augmented by artificial intelligence. *Adv Compos Hybrid Mater*. 2024;7(6):1-22.

   doi: 10.1007/s42114-024-01039-6

8. Park SJ, Lee JH, Yang J, *et al*. Lightweight injection mold using additively manufactured Ti-6Al-4V lattice structures. *J. Manuf Process*. 2022;79:759-766.

   doi: 10.1016/j.jmapro.2022.05.022

9. Howard J, Carlson K, Chidambaram D. High-temperature metallic glasses: Status, needs, and opportunities. *Phys Rev Mater*. 2021;5(4):040301.

   doi: 10.1103/physrevmaterials.5.040301

10. Scully JR, Gebert A, Payer JH. Corrosion and related mechanical properties of bulk metallic glasses. *J Mater Res*. 2007;22(2):302-313.

    doi: 10.1557/jmr.2007.0051

11. Xu T, Pang S, Li H, Zhang T. Corrosion resistant Cr-based bulk metallic glasses with high strength and hardness. *J Non Cryst Solids*. 2015;410:20-25.

    doi: 10.1016/j.jnoncrysol.2014.12.006

12. Chen M. A brief overview of bulk metallic glasses. *NPG Asia Mater*. 2011;3(9):82-90.

    doi: 10.1038/asiamat.2011.30

13. Mi XL, Hu L, Wan ZX, Wu BW, Wei B. Liquid state properties and amorphous solidification kinetics of multicomponent $Fe_{50-x}Co_xCr_{14}Mo_{14}C_9B_8Tm_5$ alloys investigated under containerless processing conditions. *Phys Rev E*. 2024;110(3):034612.

doi: 10.1103/physreve.110.034612

14. Mahbooba Z, Thorsson L, Unosson M, *et al*. Additive manufacturing of an iron-based bulk metallic glass larger than the critical casting thickness. *Appl Mater Today*. 2018;11:264-269.

doi: 10.1016/j.apmt.2018.02.011

15. Hooper PA. Melt pool temperature and cooling rates in laser powder bed fusion. *Addit Manuf*. 2018;22:548-559.

doi: 10.1016/j.addma.2018.05.032

16. Yang Z, Markl M, Körner C. Comprehensive numerical investigation of laser powder bed fusion process conditions for bulk metallic glasses. *Addit Manuf*. 2024;81:104026.

doi: 10.1016/j.addma.2024.104026

17. Li B, Yakubov V, Nomoto K, *et al*. Superior mechanical properties of a Zr-based bulk metallic glass via laser powder bed fusion process control. *Acta Mater*. 2024;266:119685.

doi: 10.1016/j.actamat.2024.119685

18. Chen C, Fan Y, Zhang W, *et al*. Tailoring Nano-crystallization in Zr50Ti4Y1Al10Cu25Ni7Co2Fe1 complex multicomponent bulk metallic glass by O doping. *J Non Cryst Solids*. 2021;553:120474.

doi: 10.1016/j.jnoncrysol.2020.120474

19. Chen Y, Zhang D, O'Toole P, *et al*. *In situ* observation and reduction of hot-cracks in laser additive manufacturing. *Commun Mater*. 2024;5(1):84.

doi: 10.1038/s43246-024-00522-3

20. Ansari MA, Crampton A, Garrard R, Cai B, Attallah M. A convolutional neural network (CNN) classification to identify the presence of pores in powder bed fusion images. *Int J Adv Manuf Technol*. 2022;120(7):5133-5150.

doi: 10.1007/s00170-022-08995-7

21. Westphal E, Seitz H. A machine learning method for defect detection and visualization in selective laser sintering based on convolutional neural networks. *Addit Manuf*. 2021;41:101965.

doi: 10.1016/j.addma.2021.101965

22. Mozaffar M, Paul A, Al-Bahrani R, *et al*. Data-driven prediction of the high-dimensional thermal history in directed energy deposition processes via recurrent neural networks. *Manuf Letters*. 2018;18:35-39.

doi: 10.1016/j.mfglet.2018.10.002

23. Chen Z, Mak S, Wu CF. A hierarchical expected improvement method for Bayesian optimization. *J Am Stat Assoc*. 2024;119(546):1619-1632.

doi: 10.1080/01621459.2023.2210803

24. Lei B, Kirk TQ, Bhattacharya A, *et al*. Bayesian optimization with adaptive surrogate models for automated experimental design. *NPJ Comput Mater*. 2021;7(1):194.

doi: 10.1038/s41524-021-00662-x

25. Zhang H, Chen WW, Iyer A, Apley DW, Chen W. Uncertainty-aware mixed-variable machine learning for materials design. *Sci Rep*. 2022;12(1):19760.

doi: 10.1038/s41598-022-23431-2

26. Ament S, Daulton S, Eriksson D, Balandat M, Bakshy E. *Unexpected Improvements to Expected Improvement for Bayesian Optimization*. [arXiv Preprint]; 2025.

doi: 10.48550/arXiv.2310.20708

27. Johnson JE, Jamil IR, Pan L, Lin G, Xu X. Bayesian optimization with Gaussian-process-based active machine learning for improvement of geometric accuracy in projection multi-photon 3D printing. *Light Sci Appl*. 2025;14(1):56.

doi: 10.1038/s41377-024-01707-8

28. Karkaria V, Goeckner A, Zha R, *et al*. Towards a digital twin framework in additive manufacturing: Machine learning and Bayesian optimization for time series process optimization. *J Manuf Syst*. 2024;75:322-332.

doi: 10.1016/j.jmsy.2024.04.023

29. Frazier PI. *A Tutorial on Bayesian Optimization*. [arXiv Preprint]; 2018.

doi: 10.48550/arXiv.1807.02811

30. Laghi V, Palermo M, Bruggi M, Gasparini G, Trombetti T. Blended structural optimization for wire-and-arc additively manufactured beams. *Prog Addit Manuf*. 2023;8(3):381-392.

doi: 10.1007/s40964-022-00335-1

31. Kavas B, Balta EC, Tucker MR, *et al*. *In-situ Controller Autotuning by Bayesian Optimization for Closed-loop Feedback Control of Laser Powder Bed Fusion Process*. Available from: https://arxiv.org/html/2406.19096v1 [Last accessed on 2025 Aug 11].

32. Squires L, Roberts E, Bandyopadhyay A. Radial bimetallic structures via wire arc directed energy deposition-based additive manufacturing. *Nat Commun*. 2023;14(1):3544.

doi: 10.1038/s41467-023-39230-w

33. Yang Z, Markl M, Körner C. Predictive simulation of bulk metallic glass crystallization during laser powder bed fusion. *Addit Manuf*. 2022;59:103121.

doi: 10.1016/j.addma.2022.103121

34. Laws KJ, Miracle DB, Ferry M. A predictive structural model for bulk metallic glasses. *Nat Commun*. 2015;6(1):8123.

doi: 10.1038/ncomms9123

35. Ding J, Ma E. Computational modeling sheds light on structural evolution in metallic glasses and supercooled liquids. *NPJ Comput Mater*. 2017;3(1):9.

    doi: 10.1038/s41524-017-0007-1

36. Xu D, Johnson WL. Crystallization kinetics and glass-forming ability of bulk metallic glasses $Pd_{40}Cu_{30}Ni_{10}P_{20}$ and $Zr_{41.2}Ti_{13.8}Cu_{12.5}Ni_{10}Be_{22.5}$ from classical theory. *Phys Rev B*. 2006;74(2):024207.

    doi: 10.1103/physrevb.74.024207

37. Bordeenithikasem P, Liu J, Kube SA, *et al*. Determination of critical cooling rates in metallic glass forming alloy libraries through laser spike annealing. *Sci Rep.* 2017;7(1):7155.

    doi: 10.1038/s41598-017-07719-2

38. Lázaro-Gredilla M, Titsias MK. Variational Heteroscedastic Gaussian process Regression. In: *Proceedings of the 28th International Conference on International Conference on Machine Learning*. *ICML'11*. Washington, DC: Omnipress; 2011. p. 841-848.

39. Heinonen M, Mannerström H, Rousu J, Kaski S, Lähdesmäki H. *Non-Stationary Gaussian Process Regression with Hamiltonian Monte Carlo*. [arXiv Preprint]; 2015.

    doi: 10.48550/arXiv.1508.04319

40. Jospin LV, Buntine W, Boussaid F, Laga H, Bennamoun M. Hands-on Bayesian Neural Networks -- a Tutorial for Deep Learning Users. *IEEE Comput Intell Mag*. 2022;17(2):29-48.

    doi: 10.1109/mci.2022.3155327

41. Gal Y, Ghahramani Z. *Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning*. [arXiv Preprint]; 2016.

    doi: 10.48550/arXiv.1506.02142

42. Knowles J. ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. *IEEE Trans Evol Comput*. 2006;10(1):50-66.

    doi: 10.1109/tevc.2005.851274

43. Hernández-Lobato JM, Gelbart MA, Adams RP, Hoffman MW, Ghahramani Z. *A General Framework for Constrained Bayesian Optimization using Information-Based Search*. [arXiv Preprint]; 2016.

    doi: 10.48550/arXiv.1511.09422

44. Wong TT. Performance evaluation of classification algorithms by *k*-fold and leave-one-out cross validation. *Pattern Recognit*. 2015;48(9):2839-2846.

    doi: 10.1016/j.patcog.2015.03.009

45. Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat Comput*. 2017;27(5):1413-1432.

    doi: 10.1007/s11222-016-9696-4

ACCSCIENCE PUBLISHING

International Journal of AI for Materials and Design

ORIGINAL RESEARCH ARTICLE

Computer vision and deep learning-based prediction for inkjet-printed electrodes

Gareth Quinn[1]*, Achu Titus[1,2,3]*, Anesu Nyabadza[1,2,3], Éanna McCarthy[1,2,3], Sithara Sreenilayam[1,2,3], and Dermot Brabazon[1,2,3]

[1]School of Mechanical and Manufacturing Engineering, Dublin City University, Dublin, Ireland

[2]I-Form Advanced Manufacturing Centre Research, Dublin City University, Dublin, Ireland

[3]DCU Institute for Advanced Processing Technology, Dublin City University, Dublin, Ireland

(This article belongs to the *Special Issue: Applications of Deep Learning in Advanced Materials Processing*)

Abstract

With the development of inkjet-printed electrodes, artificial intelligence-based quality control is essential for classifying inkjet-printed electrodes in a quality control environment. The quality of printed structures can be significantly affected by defects such as cracks, smudging, and misaligned deposits, which can degrade electrical performance and overall device reliability. Traditional quality control methods, including manual inspection and electrical testing, are time-consuming, subjective, and invasive, and they are unsuitable for high-throughput manufacturing environments. This work explores the application of computer vision and deep learning, specifically Convolutional Neural Networks (CNNs) and Feedforward Neural Networks, to automate defect detection and quality classification of inkjet-printed electrodes. To demonstrate the accessibility of deep learning techniques, Neural Architecture Search was implemented, showing the importance of automated model design in achieving high performance without extensive manual tuning or the need for expertise. The CNN models proved to be the most suitable approach for this image classification task, achieving a testing accuracy of 90.9% and a precision of 88.9% for a dataset of 2,406 electrode images containing both high-quality (1,020) and low-quality (1,386) prints.

Keywords: Inkjet printing; Electrodes; Defect detection; Deep learning; Computer vision; Convolutional Neural Networks; Feedforward neural networks; Neural architecture search

***Corresponding authors:**
Gareth Quinn (gareth.quinn39@mail.dcu.ie)
Achu Titus (achu.titus2@mail.dcu.ie)

Citation: Quinn G, Titus A, Nyabadza A, McCarthy E, Sreenilayam S, Brabazon D. Computer vision and deep learning-based prediction for inkjet-printed electrodes. *Int J AI Mater Design*. 2025;2(4):24-36. doi: 10.36922/IJAMD025430040

Received: October 22, 2025

Revised: November 28, 2025

Accepted: December 8, 2025

Published online: December 17, 2025

1. Introduction

Inkjet printing, also known as non-contact printing, has become a vital fabrication method in the production of flexible electronics, offering key advantages such as high customizability, minimal material waste, low temperature processing, fast deposition process, additive and digital patterning, and compatibility with a broad range of substrates[1] (*e.g.*, polymer, paper, textile, and fabric). It has wide applications in areas such as wearable and personalized healthcare devices,[2-4] electrochemical and biosensor devices,[5,6] flexible electronics (displays and solar cells), energy devices,[7,8] bioengineering,[9]

micro-electro-mechanical systems,[10] and microfluidic devices.[11] In an inkjet printer, inks stored in a reservoir are delivered through micrometer-sized nozzles in the print head, which operate either continuously or in a drop-on-demand mode to form electronically designed patterns or images. This process involves droplet generation, additive deposition of micro- and nanometer-scale materials, and ink spreading over the substrate, followed by drying without any physical impact on the substrate. However, this printing technique poses a significant issue known as the coffee ring effect (*i.e.*, formation of a ring-like thicker deposit at the print pattern edge due to the presence of solid particles in a drying ink droplet),[12] which hinders the formation of uniform patterns and negatively impacts the morphological and electrical properties of the electrodes.[13,14] Furthermore, print patterns experience geometric imperfections, such as cracking and discontinuities, due to the properties of the inkjet printable ink material, droplet ejection behavior, substrate surface morphology, and inkjet printing parameters, resulting in defective electrodes.[15] A prolonged sintering process also affects the microstructure of the printed electrode, especially porosity and particle order.[14] Defects are critical in inkjet-printed electrodes because they directly lower conductivity, degrade device performance, and compromise reliability.[16] Studies show that up to 66% of inkjet-printed thin-film transistors are defective due to process-related issues.[15] Issues, including particle agglomeration in the nanomaterials-based inkjet printable inks, air entrapment, nozzle obstruction, droplet misalignment, and environmental contamination, can significantly impair the precision and uniformity of prints. These factors result in unpredictable electrical properties and device unreliability.[17-19] Before the advent of digital microscopy, these defects were managed through quality control techniques, such as manual inspection and electrical testing, which were labor-intensive, subjective, intrusive, and inappropriate for large-scale manufacturing environments. At present, there are numerous methods to check the quality of a completed print, including inspection techniques, data analysis strategies, and quality assurance practices.[20] The quality assurance practices include optical photography, digital imaging, profilometry, and electrical resistance/conductivity measurements. The initial step for conductive inks involves assessing the success of the sintering process by examining the sample for any surface deformation and the presence of cracks. The inkjet-printed sample typically comprises particles at the nano- (10^{-9}) or micro- (10^{-6}) scale, which should be at least 50 times smaller than the nozzle diameter, which helps to avoid issues such as particle accumulation at the nozzle edge, droplet trajectory deviations, or nozzle blockage due to agglomeration.[21] In the absence of cracks, the sample undergoes a series of performance evaluations, which encompass electrical resistivity, adhesion, and mechanical deformation assessments.[22]

Anticipating defective electrodes before printing and modifying printing parameters is a crucial strategy for minimizing the incidence of defective printed electrodes, thereby enhancing the large-scale production of high-quality electrodes. The early identification of defects can be accomplished through a computer vision system, which represents an automated, non-destructive, and high-speed method utilizing advanced imaging and machine learning techniques.[20] Incorporating computer vision into the quality control of inkjet-printed electrodes allows manufacturers to identify and rectify problems, such as absent deposits, misalignment, and variations in line width, before compromising product performance.[23,24] It will enable immediate identification and correction of printing surface defects, minimizing material waste and improving yield rates. This technique offers several significant advantages, including the elimination of manual inspection, reduced human error, and increased efficiency.

Computer vision employs sophisticated imaging methods and machine learning algorithms to detect microscopic defects that may elude human observation. Several significant works have explored the intersection of inkjet printing and machine learning, focusing on improving monitoring, classification, prediction, and optimization of the printing processes.[23] Researchers have applied artificial intelligence (AI) to different aspects of the process, such as ink conductivity, line width, resistance, droplet velocity, jetting behavior, and overall print quality.[25-27] A wide range of methods have been explored, from basic regression models and ensemble techniques like Random Forest and Gradient Boosting to more advanced neural networks, as explained in Table 1. Deep learning models, including Convolutional Neural Networks (CNNs), Feedforward Neural Networks (FNNs), and Neural Architecture Searches (NAS), can analyze complex printed patterns and distinguish between acceptable and defective prints with high precision. In a recent work, a CNN-based algorithm was considered to identify four parameters of inkjet-printed functional structure, such as droplet spacing, line resistance, line quality, and the post-treatment method developed.[27] The model successfully classified printing features from the pictures, proving the use of CNNs for image classification and inkjet-printed structure classification.[28] A residual neural network was employed in another study to classify the three distinct stages of ejected droplets: none, non-spherical, and spherical.[27] The model attained an accuracy of approximately 96% after five epochs, after reducing the input size using a translation-invariant linear transformation, while the uncompressed model reached an accuracy of about 92%.[27] NAS was used

Table 1. Overview of the recent research in computer vision for the evaluation of inkjet-printed electronic structures

No	Predicted quality/defect of inkjet-printed structure	AI method used for prediction	References
1	Ink conductivity of the printed components	Regression models, including DT, XGBoost, LightGBM, and ResNet	30
2	Line width and resistance per unit length of printed traces	NLR, KNN regression, GPR	31
3	Drop pitch, line pitch, and average drop volume	RF, AdaBoost, and GB regressors	32,33
4	Predict drop velocity and formation	SVM, KNN, RFs, XGBoost, and MLP	33
5	Drop velocity and jetting morphology.	SVM, B-Tree, and MLP	34
6	Jetting status and predicting a high-resolution pattern	KNN, CART, RF, LR, gradient boost classifier, and bagging model	35
7	Jetting behavior of novel inks	GB and RF	36
8	Drop velocity and jetting morphology	MLP algorithm	37
9	Space into three regions based on predicted jetting probability: certain jetting, certain non-jetting, and doubt region	SVM, Multilayer Neural Network, and Gaussian Naïve Bayes	38
10	Printing quality optimization and electrical resistivity prediction	CNN architectures	39
11	Categorize the behavior of the droplet (normal, no-droplet, and satellite modes)	BPNN	25
12	Electrical conductivity of organic graphene-based electrodes	DTs, RFs, and KNN	40

Abbreviations: AdaBoost: Adaptive Boosting; BPNN: Backpropagation Neural Network; CART: Classification and Regression Tree; CNN: Convolutional Neural Network; DT: Decision Tree; GB: Gradient Boosting; GPR: Gaussian Process Regression; KNN: K-Nearest Neighbors; LightGBM: Light Gradient Boosting Machine; LR: Logistic Regression; MLP: Multilayer Perceptron; NLR: Non-linear regression; ResNet: Residual Neural Network; RF: Random Forest; SVM: Support Vector Machine; XGBoost: Extreme Gradient Boosting.

to detect and classify defects, such as the coffee ring effect and geometric defects (cracks, discontinuities, and surface impurities) in flexible inkjet-printed sensors. They used the computer vision model YOLOv5 as a baseline; along with NAS, the model achieved a mean average precision of 81.2%, 95.5% accuracy, and a detection time of 4.6 ms, making it suitable for real-time defect monitoring.[5] These two prediction models improved the print success rate to over 80%.[5] A separate study employed machine learning models, such as random forest, multilayer perceptron, and support vector machine, to forecast the printability and drug dosage of inkjet printing before formulation preparation, and it attained an accuracy of 97.22% in predicting the printability of formulations and 97.14% in assessing the quality of the prints.[29]

This work explores the application of deep learning, specifically CNNs and FNNs, to automate defect detection and quality classification of inkjet-printed electrodes. A custom dataset comprising over 400 images of inkjet-printed electrodes was developed, with each image manually classified as either exhibiting good- or poor-quality print based on visual inspection. These two classes were then examined through exploratory data analysis to identify distinguishing features and patterns. The images were subsequently pre-processed using a range of image processing techniques and used to train multiple neural network models. Model performance was assessed using several evaluation metrics, including accuracy, F1-score, and confusion matrices. To examine the accessibility of deep learning techniques, NAS was implemented, highlighting the importance of automated model design in achieving high performance without the need for extensive manual tuning or specialized expertise. This work demonstrates the effectiveness and accessibility of AI-powered visual inspection systems in improving production yield, reducing material waste, and enabling scalable, automated manufacturing of inkjet-printed electrodes.

2. Methods

2.1. Experimental setup

The manganese ink was produced using the pulsed laser ablation technique, and the electrodes were printed using a cartridge with 21-μm nozzle diameter, utilizing a FUJIFILM Dimatix Materials Printer DMP 2850 (FUJIFILM Dimatix, Inc., USA).[41] A 3 × 3 complete factorial Design of Experiments was conducted with the printing parameters to differentiate between good- and bad-quality prints. This study explored three different inkjet printing parameters, namely the jetting frequency (40, 60, 80 kHz), the print bed temperature (28, 30, 35°C), and the number of printed layers (10, 15, 30).[41] Each electrode printing parameter set permutation was printed 16 times with dimensions of 10 × 10 mm on a paper substrate, producing 432 electrodes. All electrode data, including printing and experimental parameters, are provided in Table S1 (in Supplementary File). Inkjet printing mechanism and representative printed electrodes are shown in Figure 1. Some samples did not print correctly for specific printer

parameters; therefore, only 401 electrodes were suitable for imaging.

Following the printing process, high-resolution images of individual electrodes and full electrode grids were captured using a Google Pixel 6 smartphone with a 50-megapixel camera and 7× zoom, positioned approximately 100 mm from the substrate. All images were taken parallel to the substrate under fluorescent diffuse white light illumination. This setup mirrors the typical capability of industrial line camera imaging systems. The inkjet-printed electrodes exhibited defects such as smudging, irregular deposition, contamination, and cracks. These conditions resulted in datasets labeled as either "Good Print Quality" or "Bad Print Quality," using binary classification based on visual inspection. Extensive exploratory data analysis techniques were used, such as pixel intensity distribution analysis, t-SNE visualization, Sobel edge detection, and various thresholding methods, in an attempt to highlight defects effectively. Figure 2 illustrates examples of electrodes classified as having "Good Print Quality" and "Bad Print Quality." Defective electrodes typically exhibited significant misprints or structural defects that impaired their ability to conduct electricity.

2.2. Exploring the datasets

Before proceeding with model training, an extensive exploratory data analysis was conducted to investigate the features of the good- and bad-quality printed electrodes. This involved statistical evaluations, intensity histograms, edge detection, thresholding techniques, and feature visualization methods such as t-SNE and Histogram of Oriented Gradients (HOG). The results of this section are described in Figure S1 (in Supplementary File). The pixel intensity distribution analysis showed that the distribution of good prints is more uniform. In contrast, bad prints exhibit greater randomness and a wider range of intensity values. Edge detection and thresholding techniques showed potential in highlighting defects but struggled with variations in the dataset, even when enhanced with automatic methods. Similarly, the HOG effectively captured large-scale defects by detecting sharp intensity changes but had difficulty identifying more minor, subtler anomalies.

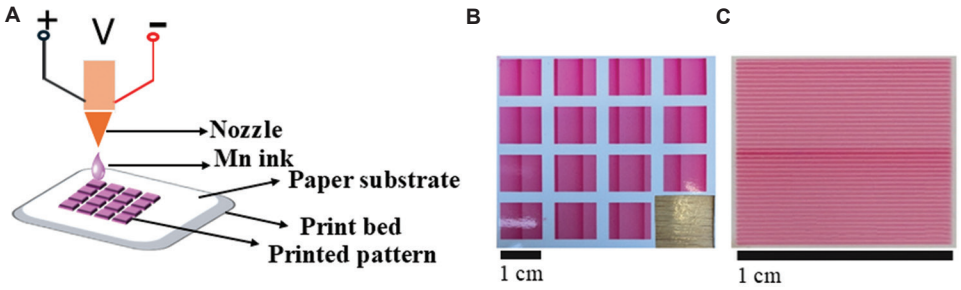


Figure 1. Inkjet printing process and printed electrodes. (A) Schematic view of the inkjet printing of manganese ink; (B) Printed grid (50 mm × 50 mm) with one sample cut out for various characterization; (C) An individual electrode (10 mm × 10 mm). The electrode grid (B) corresponds to sample 9, fabricated with 30 layers at 40 Hz and 30°C, while the electrode (C) was produced with 15 layers under the same frequency and temperature conditions.

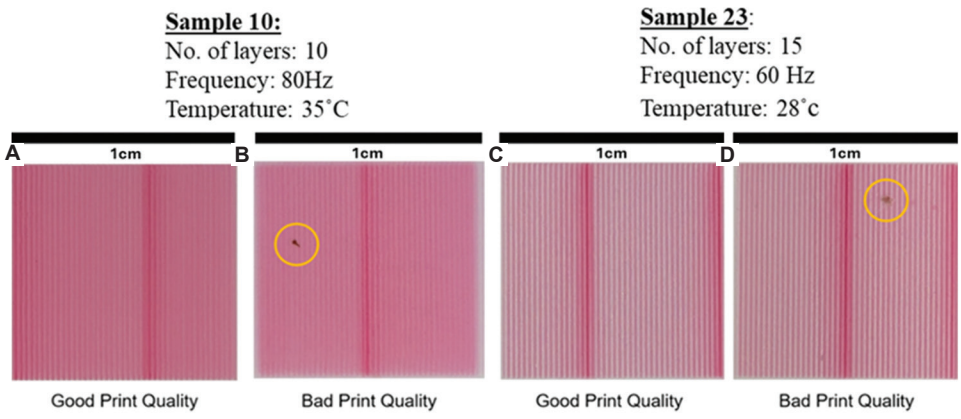


Figure 2. Pictures of electrodes exhibiting good (A and C) and bad (B and D) print quality from sample batches 10 and 23. Sample 10 was printed with 10 layers at 80 Hz and 35°C, while sample 23 was printed with 15 layers at 60 Hz and 28°C. Defects such as localized blotches are highlighted in yellow circles in the bad-quality prints.

2.3. Model preparation, training pipeline, and performance

To ensure optimal model performance and maintain consistency for fair comparison among models, a structured data preprocessing strategy and general pipeline were created. This included data augmentation to enlarge the dataset; a consistent train-test split; class balancing techniques to mitigate bias; callback functions for efficiency and effectiveness; normalization and standardization of the data; and, for evaluation, the performance metric values of accuracy, precision, recall, and F1-score were evaluated for each model. Along with these performance metric values, a confusion matrix was plotted to visualize the predictions made by the model. Finally, training graphs were plotted: loss over epochs and accuracy over epochs. The libraries used to evaluate the model were Tensorflow, Matplotlib, Scikit-Learn, and Pandas. The original datasets contained 401 images, representing a relatively small dataset size for neural networks, especially CNNs. To combat this, the images within the original dataset were augmented with rotation, brightness adjustments, zooming, and flipping to artificially enlarge the dataset while ensuring realistic variations. Each image underwent five augmentations, resulting in a new dataset size of 2,406, with 1,020 images of good print quality and 1,386 of bad print quality.

For most model iterations, except for a few during the optimization stages, all images were resized to a standard 128 × 128 resolution. The OpenCV's resize function was used for this transformation, which provided computational efficiency while preserving essential image details. Then, the dataset was divided into 70% training and 30% testing (15% testing and 15% validation). This train-test split ensured that the model had enough data to train on and that the training data represented the range of defects that could potentially occur. In addition, the testing and validation sets were large enough to provide reliable performance metrics, ensuring accurate and unbiased model assessment. The enlargement of the dataset through data augmentation resulted in a significant class imbalance. The bad print quality images outweighed the good print quality images by over 26%. This class imbalance could cause bias within the models, so class balancing was adopted to mitigate this. For each class, the class weights were calculated using the Scikit-Learn class weight function, which was applied during training by modifying the loss function to apply a higher weight to the minority class and lower weights for the majority class.

Learning rate schedulers were implemented to reduce the set learning rate when the validation loss stopped improving after five epochs, helping the model develop after reaching a plateau. For every model, the same metrics were measured. The metrics that were evaluated for each model were accuracy, precision, recall, and the F1-score. The performance of each neural network model was evaluated using four standard classification metrics: accuracy, precision, recall, and F1-score, each derived from the confusion matrix. In binary classification, predictions can be true positives (TP), true negatives (TN), false positives (FP), or false negatives (FN). Accuracy represents the proportion of correct predictions across all samples, computed using this formula: (TP+TN)/(TP+TN+FP+FN). Precision measures the proportion of correctly identified positive cases out of all positive predictions (TP/(TP+FP)), which is crucial in this project to minimize the number of defective electrodes incorrectly classified as good. Recall (TP/(TP+FN)) reflects the model's ability to correctly detect all good electrodes, while the F1-score provides a harmonic mean between precision and recall, balancing both FP and FN. These metrics were chosen because they provide valuable insights into the model's performance in real-world scenarios. Along with these performance metrics' values, a confusion matrix was plotted to visualize the predictions made by the CNN and FNN models, as shown in Figures S2 and S3, respectively. Finally, training graphs were plotted: loss over epochs and accuracy over epochs (Figures S2 and S3). Figure 3 shows examples of the two training graphs along with the confusion matrix

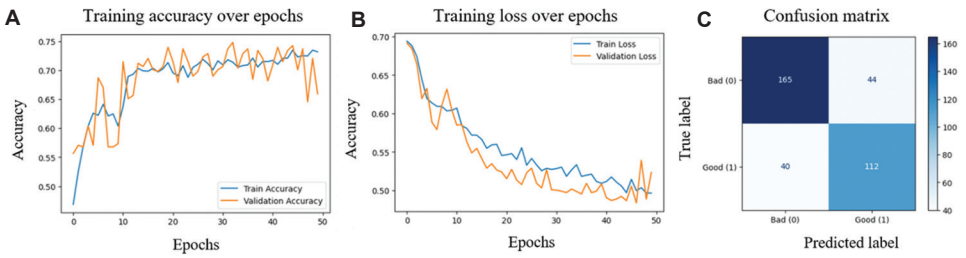


Figure 3. Performance evaluation of the basic CNN model. (A) Training and validation accuracy trends across epochs. (B) Training and validation loss trends across epochs, indicating model convergence. (C) Confusion matrix demonstrating the CNN's classification accuracy for good- and bad-print-quality electrodes.
Abbreviation: CNN: Convolutional Neural Network.

of the basic structured CNN training on the augmented dataset.

2.4. Model developments and optimization

This work modeled and compared CNNs, FNNs, and NAS to detect print quality. All model development and optimization were carried out using Python 3.10.0, primarily within the PyCharm development environment on a local laptop equipped with an Intel Core i7-1065G7 CPU (4 cores, 8 threads, 1.3–3.9 GHz), 8 GB RAM, and a 512 GB SSD. Additional training was performed on an NVIDIA Tesla T4 GPU through Google Colab for faster training and development times. The following section will explain how the previously mentioned neural network models are used to detect defective electrodes.

2.4.1. CNN model optimization

For the development of a fully optimized CNN model, basic and deep models were developed and fine-tuned using hyperparameter tuning and multiple optimization techniques to achieve the highest accuracy. After the initial iterations and improvements on the basic CNN architecture, the decision was made to restructure the network to enhance its ability by incorporating additional layers and nodes along with enhanced regularization techniques to detect more profound relationships within the image data and also to mitigate overfitting. Based on the hyperparameter tuning of the deep CNN model, a fully optimized model was developed to maximize performance and to ensure generalization to the data. The CNN architecture, from its basic structure to the concluding configuration, was modeled and optimized using the augmented datasets, as detailed in Tables S2 and S3 (in Supplementary File). The final architecture was a deep CNN with five convolutional layers, each using the Rectified Linear Unit (ReLU) activation function and L2 regularization ($\lambda = 0.001$) to ensure stable learning and help the model learn complex patterns. Batch normalization was applied after each convolutional layer. Max pooling was implemented after each layer for computational efficiency; however, it was left out in the final convolutional layer to preserve larger feature maps, ensuring the model had sufficient information for effective learning. The extracted features were flattened using the Flatten () function, due to the relatively small feature map size, and passed into a fully connected layer comprising three dense layers (512, 128, and 64 neurons). Each dense layer had a 30% dropout rate, which improved generalization by preventing overfitting and enabling the model to capture complex relationships. The output layer used the sigmoid activation function, producing a probability score between 0 and 1 to classify electrodes as defective or non-defective. The model was trained on RGB images, using the Adam optimizer and a binary cross-entropy loss function for robust optimization. The model had an initial learning rate of 0.0001. All other pipeline components, including callback functions, data splitting, class balancing, image resizing, and *ImageDataGenerator* setup, remained the same. The model was trained for 100 epochs, ensuring full convergence and optimal learning. In addition, early stopping techniques were employed to reduce unnecessary training and prevent overfitting. Figure 4 shows the optimized CNN model architecture.

2.4.2. FNN

The original FNN model featured a basic architecture consisting of three hidden layers and one output layer. The initial hidden layer comprised 12 neurons, succeeded by 64 neurons in the second layer and 16 neurons in the third. For binary classification, the output layer used the sigmoid activation function. To mitigate overfitting, a 20% dropout was introduced after the first hidden layer. The non-linear activation function, ReLU, was used across all the hidden layers to enhance feature learning and improve

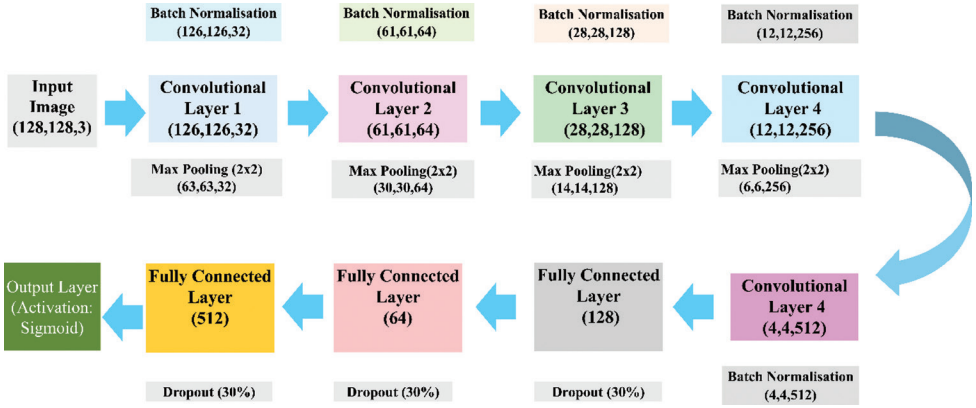


Figure 4. Flow chart illustrating the architecture of the optimized Convolutional Neural Network

generalization. The basic FNN architecture is shown in Table S4 (in Supplementary File). The dataset was further expanded and diversified using the *ImageDataGenerator* function for additional data augmentation. The model was optimized using the Adam optimizer, with binary cross-entropy as the loss function. The initial learning rate was set at 0.0001 during the training process. Following the first iteration's results, the decision was made to eliminate the *ImageDataGenerator* function from the pipeline. Dropouts and batch normalization were implemented following the initial two hidden layers to mitigate overfitting and enhance generalization within the model. In addition, L2-regularization, with $\lambda = 0.001$, was implemented on all three fully connected layers. The initial FNN with regularization is detailed in Table S5. Additional layers were incorporated to enhance the model's ability to learn complex patterns, thereby improving generalization and overall performance. Given that FNNs are not intrinsically suited for feature extraction, manual feature extraction was incorporated into the pipeline to assist the model in recognizing features within the images. The manual feature extraction methods implemented were Sauvola's thresholding technique and Sobel edge detection. Sauvola's thresholding was implemented using the libraries OpenCV and Scikit-image. After extensive hyperparameter tuning and iterative design, the final FNN model was composed of eight hidden layers arranged in a decreasing, funnel-shaped structure (256, 128, 64, 64, 32, 16, 1). Each layer employed the ReLU activation function, enabling the model to learn complex patterns and relationships. L2 regularization with a weight of 0.001 was implemented to prevent certain features from overshadowing the output and thereby enhancing generalization. Batch normalization and dropout were implemented after specific layers, as illustrated in the model architecture in Figure 5 and Table S6.

The output layer utilized the sigmoid activation function. The model was optimized using the Adam optimizer with a binary loss function and an initial learning rate of 0.0001 to support steady and stable learning. All other components of the pipeline, including callback functions, data splitting, class balancing, and image resizing, remained unchanged.

2.4.3. NAS

The utilization of NAS facilitated the efficient development of tailored, optimized neural networks for defect detection in inkjet-printed electrodes, demonstrating the accessibility of AI solutions for this issue. This was implemented using the AutoKeras library to automatically design an optimal neural network architecture tailored to the binary classification task. Pre-processed grayscale electrode images were standardized and split into training, validation, and testing sets (70/15/15). Class imbalance was addressed using the computed class weights. Due to computational limitations, only three trials were configured to run. The model was trained on Google Colab using the free T4 GPU, and the best-performing architecture was exported and evaluated.

3. Results and discussion

3.1. CNN results

The development, evaluation, and optimization of the CNN-based image classification model have demonstrated to effectively identify print defects in inkjet-printed electrodes for quality control purposes. The initial CNN model provided a foundational understanding of the network's capability to differentiate between good and bad-printed electrodes. While the base model achieved ~67% accuracy, it lacked robustness and stability, with a low precision score (~61%) indicating a high FP rate (bad-quality electrodes being classified as good), which is unsuitable for real-world quality control systems. However, the slight difference between training and testing performance suggested that the model was not overfitting and had potential for improvement.

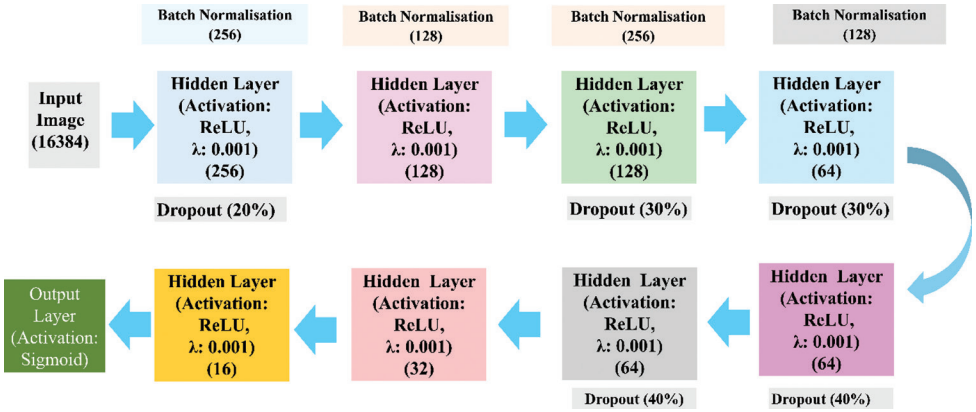


Figure 5. Flowchart illustrating the architecture of the optimized Feedforward Neural Network

After the implementation of data augmentation, there was a significant improvement in the model's performance. The validation accuracy increased from ~67% to ~77% and precision increased from ~61% to ~72%, indicating that there was roughly a 10% decrease in the number of FP (bad electrodes being misclassified). The average difference between training and validation also decreased, from ~5% to ~3%, highlighting the fact that the model generalized better. Standardizing the input data and adding additional augmentation using the *ImageDataGenerator* function, and transitioning to a deeper CNN architecture, significantly boosted performance and stability. Table 2 summarizes the model's accuracy before and after data augmentation. All metrics are derived from the performance on the testing set during training, which consists of the 20% of data withheld during training and not previously seen by the model.

The deeper model achieved an accuracy of up to approximately 75%, and the percentage differences between training and testing results showed improved generalization. While this iteration showed stronger performance, the misclassification of 27% of defective electrodes still posed a risk in a practical quality control implementation. In addition to that, the extensive hyperparameter tuning, which incorporated L2 regularization and dropout, proved effective in reducing overfitting by limiting dominant features and enforcing more general feature learning. *GlobalAveragePooling2D* was found to be a superior alternative to the *Flatten* function for smaller networks, acting as a natural regularizer and improving generalization. Adjustments to batch size and optimizer choice revealed Root Mean Square Propagation (RMSprop) as the most effective, and training on RGB images further increased classification accuracy by providing richer feature information. The summary of performance metrics from the hyperparameter tuning of the deeper CNN model is presented in Table 3.

The optimized CNN model achieved an impressive ~90% accuracy on the test data, along with an ~88% precision score, resulting in the lowest FP rate among all developed models. In binary classification, predictions can be TP, TN, FP, or FN. Accuracy represents the proportion of correct predictions across all samples, given by (TP+TN)/(TP+TN+FP+FN). Precision measures the proportion of correctly identified positive cases out of all positive predictions (TP/(TP+FP)), which is crucial in this project to minimize the number of defective electrodes incorrectly classified as good. Recall (TP/(TP+FN)) reflects the model's ability to correctly detect all good electrodes, while the F1-score provides a harmonic mean between precision and recall, balancing both FP and FN. Notably, the training and testing metrics were closely aligned, demonstrating strong generalizations on completely unseen data. The training graphs in Figure S2 show that validation accuracy tracked closely with training accuracy until around the 35th epoch, where a slight divergence emerged, indicating minimal overfitting; however, the model continued to learn effectively. The early stopping function stopped training at the 86th epoch, once validation loss had plateaued, ensuring optimal performance without unnecessary computing. Examining the confusion matrix, the model delivered the best performance across all iterations, misclassifying just 18 out of 203 electrodes. This translates to a 92% accuracy

Table 2. Comparison of the performance of a basic CNN model trained on the original dataset (401 images) versus an augmented dataset (2,406 images)

Dataset	Accuracy	Precision	Recall	F1-Score
Pre-augmentation	0.672131	0.615385	0.827586	0.705882
Post-augmentation	0.767313	0.717949	0.736842	0.727273
% Difference	+14	+17	−11	+3.03

Abbreviation: CNN: Convolutional Neural Network.

Table 3. Summary of performance metrics from the hyperparameter tuning of deeper CNN

CNN model development stages	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Average % difference to training	% increase from base CNN
Base deep CNN (Batch size: 32; optimizer: Adam)	74.38	68.72	76.88	72.57	2.18	#NA
Adding layers	82.55	76.33	84.87	80.37	5.30	9.74
Batch size 16	80.61	71.81	88.82	79.41	2.97	8.77
Batch size 64	80.33	72.13	86.84	78.81	2.91	8.04
L2 regularization	82.83	75.00	88.82	81.33	3.11	10.80
RGB input	82.37	77.27	85.00	80.95	5.68	10.15
RMSprop optimizer	76.86	73.17	75.00	74.07	3.04	2.20
M.SGD optimizer	72.45	67.65	71.88	69.70	0.88	−3.71
Global average pooling	77.14	74.21	73.75	73.98	2.97	2.19

Abbreviation: CNN: Convolutional Neural Network; M.SGD: Stochastic Gradient Descent with Momentum; RMSprop: Root Mean Square Propagation.

in identifying 121 defective electrodes, demonstrating the model's robustness and reliability for real-world defect detection applications. Overall, the results from this section demonstrate that CNNs, when adequately trained, tuned, and optimized, are highly effective for image-based defect classification, even with very minor defects. Figure 6 presents the results from the optimized CNN in training and testing. Training graphs along with the confusion matrix are shown in Figure S2 (in Supplementary File).

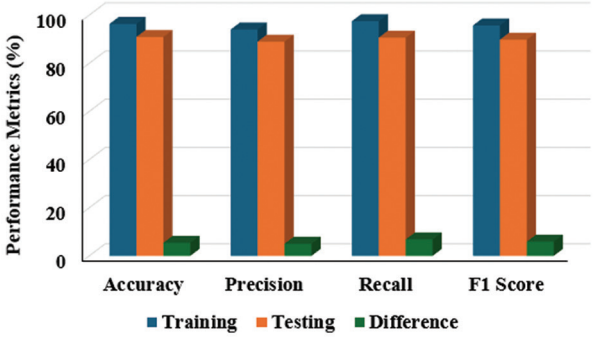


Figure 6. Graph showing the normalized performance metric of accuracy, precision, recall, and F1-score for the Convolutional Neural Network model

3.2. FNN results

Just like CNN, the FNN model's development also followed a structured, iterative approach to detect the defective electrode. The initial baseline model struggled to distinguish between the data due to its simple architecture, which struggled to handle the complexity of the data. Flattening the raw image data into one-dimensional (1D) arrays caused the loss of significant spatial information, making it very difficult for the FNN to distinguish between defective and non-defective electrode images. By removing the *ImageDataGenerator* function, the model's performance and ability to learn patterns in the data significantly increased. The model achieved a test accuracy of ~75% and a precision of ~71%, which are decent metrics. From the testing data, over 88/416 electrodes were classified as bad, giving the model an FP rate of 21%. Within the deeper FNN, employing Sauvola's thresholding yielded a testing accuracy of approximately 72%, while the Sobel edge detection method completely failed to distinguish between the classes. These findings underscore the FNN's sensitivity to noisy or overly complex inputs and its limited capacity to interpret spatial information effectively. The final model, which incorporated an optimized pipeline, regularization, and a deep architecture, achieved ~71% test accuracy and ~67% precision. The model misclassified 95 out of 416 bad electrodes, resulting in a 22% FP rate. Figure 7 presents the results of the optimized FNN. All training graphs, along with the confusion matrix, are illustrated in Figure S3 (in Supplementary File). In addition, receiver operating characteristic curves were plotted for both the CNN and FNN, which are shown in Figures S4 and S5 (in Supplementary File). The area under the curve was found to be 0.95 for optimized CNN and 0.76 for FNN, indicating excellent and acceptable discrimination performance, respectively.

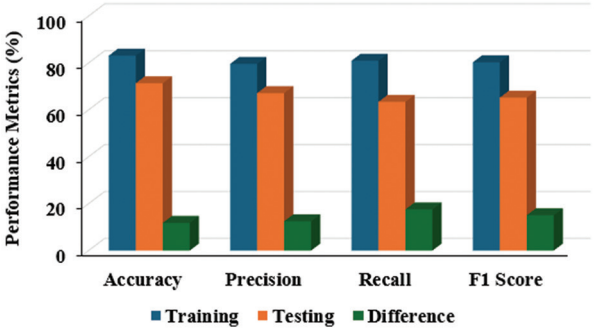


Figure 7. Graph showing the normalized performance metric of accuracy, precision, recall, and F1 score for the Feedforward Neural Network model

3.3. NAS results

From the model summary function, it was understood that the best-performing model obtained through NAS used a pre-trained EfficientNetB7 (a popular CNN architecture) as the backbone, with input shape being scaled grayscale images (224,224,3). The inputs were normalized and augmented (flip, translation). Global average pooling was used to translate the feature maps into a 1D array and the sigmoid activation function was used to classify. This NAS model demonstrated strong generalization, with a testing accuracy of ~91.4% and a training accuracy of ~94.8%. The ~3.67% accuracy difference between training and testing accuracy suggested that the model did not significantly overfit, indicating that it learned generalizable features rather than memorizing the training data. The results from using this automated method are presented in Figure 8 and Figure S6.

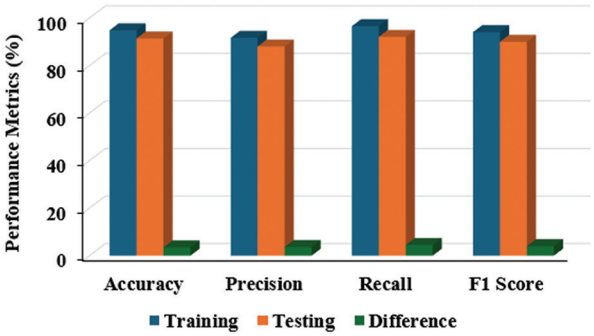


Figure 8. Graph showing the normalized performance metric of accuracy, precision, recall, and F1-score for the Neural Architecture Search-generated model

With a precision score of ~88% on testing, out of 209 printed electrodes, 19 were misclassified, but the model effectively classifies the majority of the remaining bad electrodes. With a minimal FP rate and effective classification, this NAS model effectively meets the strict quality control standards required in the manufacturing industry. These results demonstrate the use of automated deep learning methods to develop and optimize networks for classification tasks, without the need for expertise in machine learning or deep learning.

Overall, CNNs emerged as the most effective approach: the final optimized CNN achieved 90.9% test accuracy and 88.9% precision, aided by data augmentation, batch normalization, L2 regularization, and using RGB inputs. FNNs, by contrast, struggled due to the loss of spatial information when flattening images, with the best model reaching ~71% accuracy. Manual feature extraction (Sauvola thresholding, Sobel detection) offered limited benefit and often introduced noise. NAS with EfficientNetB7 achieved comparable results to the custom-built CNN (91.4% accuracy, 88% precision) while requiring minimal manual tuning. Overall, CNNs demonstrated strong generalization and industrial suitability, while FNNs proved less effective for image-based classification. Figure 9 shows a brief comparison of the three methods used to classify the test images. Figure 10 shows examples of the optimized CNN inference results.

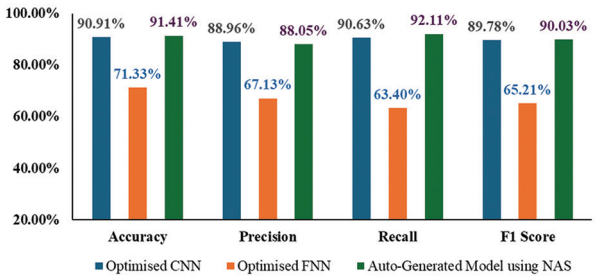


Figure 9. Graph showing accuracy, precision, recall, and F1-score as percentages, representing the proportion of correct predictions across all samples, for the CNN, FNN, and NAS models
Abbreviations: CNN: Convolutional Neural Network; FNN: Feedforward Neural Network; NAS: Neural Architecture Search.

Past literature has presented deep learning work to classify the jetting characteristics and conductivity measurements of inkjet-printed electrodes for assessing the quality of prints.[32,38] In this paper, the quality of inkjet-printed electrodes is classified on the basis of surface defects such as contamination and cracks, caused by irregularities from the deposition process. In particular, for the first time in this manuscript, the classification performance has been assessed and compared of deep learning models, including CNNs, FNNs, and NAS-generated models. CNN models were identified as one of the best classification models, resulting in a testing accuracy and precision of 90.9% and 88.9%, respectively, with improvements such as data augmentation, batch normalization, L2 regularization, global average pooling, and the use of RGB images. This prediction accuracy measured in this work for the CNN was in line with that found from previous work, which examined CNN architectures for extracting representative features of printed lines.[39]

4. Conclusion

This work demonstrates how deep learning, particularly CNNs, can be effectively leveraged to automate visual inspection tasks, offering a scalable, accurate, and efficient solution for industrial quality control. In particular, this work shows how neural networks can be leveraged for the classification of inkjet-printed electrodes in a quality control environment. A comparison between CNNs, FNNs, and NAS models highlights the strengths and limitations of each approach for the classification of inkjet-printed electrodes. The CNN models were the most suitable approach for this image classification task. After iterative improvements, the final optimized CNN model achieved 90.9% testing accuracy and 88.9% precision, demonstrating strong generalization with a low FP rate. Improvements such as data augmentation, batch normalization, L2 regularization, Global Average Pooling, and utilization of RGB images allowed the CNN to capture more complex patterns and stabilize training. These improvements also helped prevent overfitting, making the model reliable and suitable for industrial implementation. RGB input images

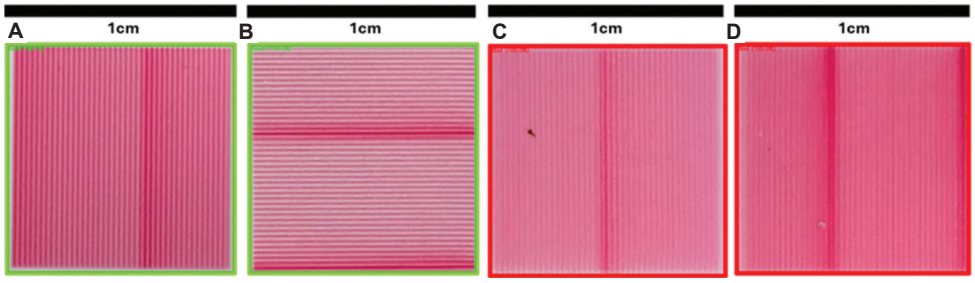


Figure 10. Pictures of electrodes correctly classified by the optimized Convolutional Neural Network. (A and B) Good-print-quality electrodes; (C and D) Bad-print-quality electrodes, which display noticeable blotches and contaminants.

contributed to a ~6% improvement in accuracy over grayscale images but at the cost of increased computational demand and longer training times. The base FNN using *ImageDataGenerator* showed random prediction behavior, with testing accuracy at ~58%. Removing this function improved the performance substantially (~75% testing accuracy and ~71% precision), showing that spatial transformations applied before flattening the data confused the model. Further improvements through applying dropout, batch normalization, and L2 regularization allowed for better generalization, although overfitting was still a recurring issue. The optimized final FNN model achieved ~71% test accuracy and ~67% precision. These results demonstrated the limitations of FNNs in extracting spatial features and reinforced that they are not ideal for this image classification task. It is important to note that, in terms of speed, FNNs are trained significantly faster than CNNs due to their simplistic architectures and reduced computational complexity. Despite the increased training time, CNNs remain the preferred choice because accuracy and reliability are critical for the required task. The NAS was employed to demonstrate automated model building. Using EfficientNetB7 (a popular CNN architecture) as a pre-trained backbone and with augmented grayscale input, this model achieved 91.4% testing accuracy and 88% precision, with only a 3.67% accuracy drop from training to testing, indicating excellent generalization. The model misclassified 19 out of 209 defective electrodes. The NAS AI methodology demonstrates the power of automated model building without requiring deep knowledge in machine learning or deep learning, reinforcing how accessible deep learning is becoming. The predictive capabilities of the established AI system, however, could be improved by collecting more data, including in the form of higher-resolution images. For more accurate defect capture, expanding these datasets with more labeled samples could also help in predicting a wider variety of defect types such as smudges, cracking, lack of connectivity, and contamination. The use of preprocessing methods such as feature extraction techniques could also be considered for creating AI models with more specific defect type predictability.

Acknowledgments

None.

Funding

This publication arises from research supported by Research Ireland under Grant Number 21/RC/10295_P2 and is co-funded by the European Regional Development Fund. This work is supported by I-Form.

Conflict of interest

Dermot Brabazon is one of the Associate Editors of the journal but was not involved in the editorial and peer-review process conducted for this paper, directly or indirectly. Separately, other authors declared that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Author contributions

Conceptualization: Gareth Quinn, Anesu Nyabadza, Dermot Brabazon
Formal analysis: Gareth Quinn
Funding acquisition: Dermot Brabazon
Investigation: Gareth Quinn, Achu Titus
Methodology: Gareth Quinn
Supervision: Dermot Brabazon
Writing–original draft: Achu Titus
Writing–review & editing: Gareth Quinn, Sithara Sreenilayam, Éanna McCarthy, Anesu Nyabadza, Dermot Brabazon

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

The data supporting this article have been included in the Supplementary File.

References

1. Vida J, Solak S, Shao Y, Homola T, List-Kratochvil E, Hermerschmidt F. Sintering of inkjet-printed silver nanoparticles by large-area atmospheric pressure nitrogen plasma. *Appl Phys A*. 2025;131(2):91.

   doi: 10.1007/s00339-024-08206-y

2. Du X, Wankhede SP, Prasad S, *et al*. A review of inkjet printing technology for personalized-healthcare wearable devices. *J Mater Chem C*. 2022;10(38):14091-14115.

   doi: 10.1039/D2TC02511F

3. Carey T, Cacovich S, Divitini G, *et al*. Fully inkjet-printed two-dimensional material field-effect heterojunctions for wearable and textile electronics. *Nat Commun*. 2017;8(1):1202.

   doi: 10.1038/s41467-017-01210-2

4. Rosati G, Cisotto G, Sili D, *et al*. Inkjet-printed fully customizable and low-cost electrodes matrix for gesture

recognition. *Sci Rep.* 2021;11(1):14938.

doi: 10.1038/s41598-021-94526-5

5. Yao C, Wang L, Wang Q, *et al*. Deep-learning-guided evaluation method for the high-volume preparation of flexible sensors based on inkjet printing. *ACS Appl Interfaces Mater*. 2024;16(10):13326-13334.

doi: 10.1021/acsami.4c00322

6. Rossetti M, Srisomwat C, Urban M, *et al*. Unleashing inkjet-printed nanostructured electrodes and battery-free potentiostat for the DNA-based multiplexed detection of *SARS-CoV-2* genes. *Biosens Bioelectron*. 2024;250:116079.

doi: 10.1016/j.bios.2024.116079

7. Mei Y, Cannizzaro C, Park H, *et al*. Cell-compatible, multi-component protein arrays with subcellular feature resolution. *Small*. 2008;4(10):1600.

doi: 10.1002/smll.200800363

8. Sarma Choudhury S, Katiyar N, Saha R, Bhattacharya S. Inkjet-printed flexible planar Zn-MnO_2 battery on paper substrate. *Sci Rep.* 2024;14(1):1597.

doi: 10.1038/s41598-024-51871-5

9. Kumar P, Ebbens S, Zhao X, *et al*. Inkjet printing of mammalian cells-theory and applications. *Bioprinting*. 2021;23:e00157.

doi: 10.1016/j.bprint.2021.e00157

10. Bernasconi R, Invernizzi GP, Gallo Stampino E, Gotti R, Gatti D, Magagnin L. Printing MEMS: Application of inkjet techniques to the manufacturing of inertial accelerometers. *Micromachines (Basel)*. 2023;14(11):2082.

doi: 10.3390/mi14112082

11. Chen S, He Z, Choi S, Novosselov IV. Characterization of inkjet-printed digital microfluidics devices. *Sensors (Basel)*. 2021;21(9):3064.

doi: 10.3390/s21093064

12. Tian L, Liu J, Chen X, Branicio PS, Lei Q. Mechanisms and strategies to achieve stability in inkjet printed 2D materials electronics. *Adv Electron Mater*. 2025;11(3):2400143.

doi: 10.1002/aelm.202400143

13. Yunker PJ, Still T, Lohr MA, Yodh AG. Suppression of the coffee-ring effect by shape-dependent capillary interactions. *Nature*. 2011;476(7360):308-311.

doi: 10.1038/nature10344

14. Soltman D, Subramanian V. Inkjet-printed line morphologies and temperature control of the coffee ring effect. *Langmuir*. 2008;24(5):2224-2231.

doi: 10.1021/la7026847

15. Sowade E, Ramon E, Mitra KY, *et al*. All-inkjet-printed thin-film transistors: Manufacturing process reliability by root cause analysis. *Sci Rep.* 2016;6(1):33490.

doi: 10.1038/srep33490

16. Chen Z, Gengenbach U, Koker L, *et al*. Systematic investigation of novel, controlled low-temperature sintering processes for inkjet printed silver nanoparticle ink. *Small*. 2024;20(21):2306865.

doi: 10.1002/smll.202306865

17. Kwon KS, Jo JY. Towards zero-defect inkjet printing via piezo self-sensing signals. *Sens Actuators A Phys*. 2025;393:116755.

doi: 10.1016/j.sna.2025.116755

18. Kwon KS. Methods for detecting air bubble in Piezo inkjet dispensers. *Sens Actuators A Phys*. 2009;153(1):50-56.

doi: 10.1016/j.sna.2009.04.024

19. Jeurissen R, De Jong J, Reinten H, *et al*. Effect of an entrained air bubble on the acoustics of an ink channel. *J Acoust Soc Am*. 2008;123(5):2496-2505.

doi: 10.1121/1.2835624

20. Law KNC, Yu M, Zhang L, *et al*. *Enhancing Printed Circuit Board Defect Detection through Ensemble Learning*. New Jersey: IEEE; 2024. p. 37-44.

doi: 10.1109/FITYR63263.2024.00013

21. Novikovs A, Tsebriienko T, Trausa A, *et al*. Novel terpineol-based silver nanoparticle ink with high stability for inkjet printing. *Nanomaterials (Basel)*. 2025;15(13):955.

doi: 10.3390/nano15130955

22. Beedasy V, Smith PJ. Printed electronics as prepared by inkjet printing. *Materials (Basel)*. 2020;13(3):704.

doi: 10.3390/ma13030704

23. Amini A, Gan TH. A computer vision-based quality assessment technique for R2R printed silver conductors on flexible plastic substrates. *Appl Sci*. 2023;13(2):1084.

doi: 10.3390/app13021084

24. Kim J, Jung Y, Parajuli S, *et al*. Quantifying printing quality for printed electrodes via deep learning and spatial association: Empowering process optimization. *Adv Intell Syst*. 2025;7:2500178.

doi: 10.1002/aisy.202500178

25. Ogunsanya M, Isichei J, Parupelli SK, Desai S, Cai Y. *In-situ* droplet monitoring of inkjet 3D printing process using image analysis and machine learning models. *Proced Manuf*. 2021;53:427-434.

doi: 10.1016/j.promfg.2021.06.045

26. Huang J, Segura LJ, Wang T, Zhao G, Sun H, Zhou C. Unsupervised learning for the droplet evolution prediction and process dynamics understanding in inkjet printing. *Addit Manuf*. 2020;35:101197.

doi: 10.1016/j.addma.2020.101197

27. Nandipati M, Ogunsanya M, Desai S, *et al*. Predictive

models for 3D inkjet material printer using automated image analysis and machine learning algorithms. *Manuf Lett*. 2024;41:810-821.

doi: 10.1016/j.mfglet.2024.09.101

28. Polomoshnov M, Reichert KM, Rettenberger L, *et al*. Image-based identification of optical quality and functional properties in inkjet-printed electronics using machine learning. *J Intell Manuf*. 2025;36(4):2709-2726.

doi: 10.1007/s10845-024-02385-4

29. Carou-Senra P, Ong JJ, Castro BM, *et al*. Predicting pharmaceutical inkjet printing outcomes using machine learning. *Int J Pharm X*. 2023;5:100181.

doi: 10.1016/j.ijpx.2023.100181

30. Sahu A, Aaen PH, Damacharla P, *et al*. *An Automated Machine Learning Approach to Inkjet Printed Component Analysis: A Step Toward Smart Additive Manufacturing*. New York: IEEE; 2024. p. 1-6.

doi: 10.1109/WMCS62019.2024.10618993

31. Lall P, Soni V, Kulkarni S, Miller S, *et al*. *Comparison of Machine Learning Approaches for Correlating Print Process Parameters to Realized Physical and Electrical Characteristics of Printed Electronics Using Inkjet Platform*. New York: American Society of Mechanical Engineers; 2023. p. V001T03A012.

doi: 10.1115/IPACK2023-112056

32. Queralto A, Pacheco A, Jimenez N, *et al*. Defining inkjet printing conditions of superconducting cuprate films through machine learning. *J Mater Chem C*. 2022;10(17):6885-6895.

doi: 10.1039/d1tc05913k

33. Kim S, Cho M, Jung S. The design of an inkjet drive waveform using machine learning. *Sci Rep.* 2022;12(1):4841.

doi: 10.1038/s41598-022-08784-y

34. Kim S, Wenger R, Bürgy O, *et al*. Predicting inkjet jetting behavior for viscoelastic inks using machine learning. *Flex Print Electron*. 2023;8(3):035007.

doi: 10.1088/2058-8585/acee94

35. Jiang L, Wolf R, Alharbi K, Qin H, *et al*. *In situ* monitoring and recognition of printing quality in electrohydrodynamic inkjet printing via machine learning. *J Manuf Sci Eng*. 2024;146(11):110901.

36. Brishty FP, Urner R, Grau G, *et al*. Machine learning based data driven inkjet printed electronics: Jetting prediction for novel inks. *Flexible and Printed Electronics*. 2022;7(1):015009.

doi: 10.1088/2058-8585/ac5a39

37. Kim S, Cho M, Jung S, *et al*. Reinforcement learning-based dynamic optimization of driving waveforms for inkjet printing of viscoelastic fluids. *Langmuir*. 2025;41(17):10831-10840.

doi: 10.1021/acs.langmuir.4c05141

38. Phung TH, Park SH, Kim I, Lee TM, Kwon KS. Machine learning approach to monitor inkjet jetting status based on the piezo self-sensing. *Sci Rep.* 2023;13(1):18089.

doi: 10.1038/s41598-023-45445-0

39. Li M, Yin S, Liu Z, Zhang H. Machine learning enables electrical resistivity modeling of printed lines in aerosol jet 3D printing. *Sci Rep.* 2024;14(1):14614.

doi: 10.1038/s41598-024-65693-y

40. Shirsavar MA, Taghavimehr M, Ouedraogo LJ, *et al*. Machine learning-assisted E-jet printing for manufacturing of organic flexible electronics. *Biosens Bioelectron*. 2022;212:114418.

doi: 10.1016/j.bios.2022.114418

41. Nyabadza A, Titus A, McCarthy É, *et al*. Fabrication and inkjet printing of manganese oxide electrodes for energy storage. *Chem Eng J Adv*. 2025;22:100761.

doi: 10.1016/j.ceja.2025.100761

ACCSCIENCE PUBLISHING

International Journal of AI for Materials and Design

ORIGINAL RESEARCH ARTICLE

Machine learning and exploratory data analysis for predicting tensile and thermal responses in friction stir spot welding

Sajad N. Alasadi and Raheem Al-Sabur*

Department of Mechanical Engineering College, University of Basrah, Basrah, Iraq

Abstract

Friction stir spot welding (FSSW) has gained increasing attention over the last decade due to its promising performance compared to conventional joining methods for similar metals. However, the thermal and tensile responses in this process are highly nonlinear. This study aims to explore the thermal and tensile performance of aluminum joints welded by FSSW using an innovative method based on exploratory data analysis (EDA), followed by several machine learning (ML) approaches. The welding parameters investigated in this study were tool rotational speed, dwelling time, and aluminum sheet thickness. The ML methods included linear and nonlinear regression models for welded joints at different welding parameters. We evaluated Bayesian ridge, elastic-net, support vector regression (SVR), random forest, polynomial regression (nonlinear), and robust regression. The random forest algorithm provided accurate predictions for lap-shear fracture load (R^2 = 0.96, mean squared error [MSE] = 0.01, and mean absolute error [MAE] = 0.07) in tensile performance, whereas the elastic net performed worst. Model-to-model differences were smaller for thermal performance, with the random forest model yielding the most accurate predictions (R^2 = 0.97, MSE = 26.51, and MAE = 3.86) while the SVR yielded the least accurate predictions. The study indicated that using EDA to address anomalies in welding conditions provides valuable insights into the best ML methods for predicting the thermal and mechanical performance of welding joints.

Keywords: Friction stir spot welding; Machine learning; Exploratory data analysis; Tensile performance; Thermal performance

***Corresponding author:**
Raheem Al-Sabur
(raheem.musawel@uobasrah.edu.iq)

Citation: Alasadi SN, Al-Sabur R. Machine learning and exploratory data analysis for predicting tensile and thermal responses in friction stir spot welding. *Int J AI Mater Design*. 2025;2(4):37-51.
doi: 10.36922/IJAMD025360032

Received: September 1, 2025

Revised: December 15, 2025

Accepted: December 16, 2025

Published online: December 30, 2025

1. Introduction

Over the past decade, the use of solid-state welding has grown significantly, primarily due to its ability to generate heat below the melting point of the base metals, thereby minimizing defects in the weld zone and the surrounding material.[1,2] A notable trend during this period is the dominance of friction stir welding (FSW) in academic research compared to other solid-state welding techniques.[3] In addition, FSW is classified as an eco-friendly welding method; it does not produce any fumes or spatter, is less energy-intensive, and does not cause significant changes in the microstructure.[4,5] Initially, FSW was limited to aluminum alloys and gradually expanded to include other alloys, such as magnesium, copper, and even steel. Further developments have given FSW a significant advantage over conventional

welding, particularly in its ability to weld dissimilar metals, including those with substantial differences in melting points.[6] Furthermore, FSW has expanded to include the welding of polymeric materials and composites.[7]

The FSW mechanism differs significantly from most welding processes, as the welding tool is reusable. The welding tool is typically made of tool steel or cemented carbide, selected for high hardness and wear resistance.[8,9] A lathe process typically forms the tool, consisting of two interconnected parts. The first part is cylindrical, has a suitable length and diameter, and is referred to as the shoulder. The other part is shorter and narrower, can be tapered, and is called the pin.[10] The welding tool is securely attached to the FSW machine, typically a specialized unit designed for this process. Alternatively, computer numerical control machines can be adapted for FSW. For friction stir spot welding (FSSW), milling machines or even punching machines may also be used.[11,12] The two plates to be welded are firmly clamped together, usually in a butt joint configuration, though lap joints are also possible.[13] The process begins when the welding tool is lowered vertically onto the area to be welded at a specific rotational speed, depending on the type of metal to be welded.[14,15] The tool penetrates the right plate and a portion of the left plate due to frictional heat. Once the material beneath the shoulder becomes sufficiently plasticized, the welding tool begins a transverse movement along the weld line in addition to its rotational movement. This process yields a satisfactory weld between the two plates, with fewer defects, while generating flash metal on both ends of the weld line.[16] Upon reaching the endpoint of the weld, the tool rises, leaving a keyhole. FSSW is a special case of FSW in which spot welding is performed using the heat generated from the tool's rotational motion, and the transverse velocity is reduced, as shown in Figure 1.

Since John laid the foundation for exploratory data analysis (EDA),[17] the EDA approach has been steadily developed across numerous engineering applications,[18] although its use has also been limited to some welding problems.[19] The EDA approach begins by examining how the variables in the problem are distributed. EDA typically begins by examining the distributions of the variables of interest. In FSSW studies, EDA is primarily used to identify outliers and to explore relationships among welding parameters.[20] EDA therefore provides an initial basis for selecting appropriate models and interpreting results. The histograms visualize how key process parameters—such as material thickness, tool rotational speed, and dwell time—are distributed across the data.

Machine learning (ML) is a branch of artificial intelligence that has gained significant traction recently. ML focuses on developing algorithms and statistical models that enable computer systems to learn patterns and make decisions based on experimental data without requiring explicit programming for each task.[21,22] Welding has similarly benefited from those advances and is now widely applied in techniques such as arc welding,[23] laser welding,[24] ultrasonic welding,[25] and resistance spot welding.[26]

In the field of FSW and related welding methods, particularly FSSW, a substantial part of beneficial research has focused on the application of ML techniques. Studies have shown that ML can improve FSW process parameters by analyzing the relationship between input parameters and weld quality.[27] It can also accurately detect and classify defects by analyzing microstructure and classifying image properties.[28] It can also provide an in-depth description and better classification of the resulting weld joint.[29] Furthermore, it can predict weld quality by optimizing welding parameters and reducing defects, with greater potential to analyze the relationship between tool geometry and material properties and their association with weld joint quality.[30] In a previous study, artificial neural networks (ANNs) were used to predict stir zone hardness and determine optimal FSW parameters, particularly tool rotational speed and weld line velocity,

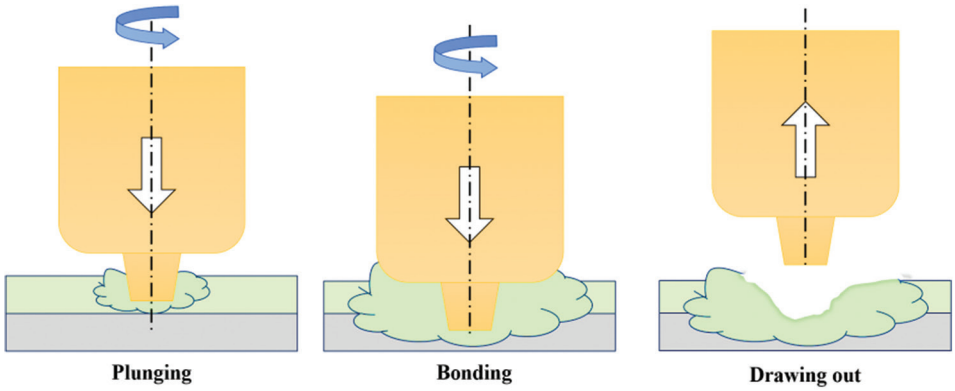


Figure 1. Process flow diagram of friction stir spot welding

in submerged 6061-T6 aluminum alloy.[31] ANNs have also shown remarkable performance in predicting wear rates during friction stir processing,[32] predicting tensile failure properties and fracture location of aluminum alloy 7075-T6 and duplex stainless steel,[33] and have been used to explore the thermal behavior and peak temperature in aluminum alloy.[34] In addition to ANNs, support vector machines (SVMs) have been utilized as ML tools in FSW and FSSW applications. For example, SVMs have been used for tensile-strength prediction[35] and for welding-process monitoring.[36] Several studies in FSSW and FSW focused on utilizing other ML tools, such as random forest.[37,38] Moreover, elastic-net,[39] ANN, and response surface methodology[40] were used in few FSW studies. Table 1 summarizes the most related studies.

FSSW is a solid-state welding process with a complex nonlinear nature. Therefore, understanding the tensile (lap-shear load) and thermal (maximum temperature) responses is complex. EDA and ML can help identify hidden correlations, while ML models effectively learn these nonlinear dependencies. ML, such as random forest, provides measures of feature importance that help understand parameter sensitivity. This study combines two approaches for an in-depth understanding of the prediction of the thermal and mechanical behavior of FSSW-welded aluminum alloys. EDA was employed alongside ML algorithms, marking a fundamental departure from traditional methods that rely solely on experimental approaches, simulations, or ML alone. In this study, we used EDA—including histograms, correlation matrices, and feature ranking—to uncover key patterns regarding shear stress and thermal behavior in the weld zone. We also conducted an in-depth comparison of several ML techniques, including random forest, elastic net, and Bayesian ridge, to determine the optimal welding conditions. Ultimately, this study offers a smart, scalable approach to optimizing the FSSW process, enabling easy prediction of the mechanical and thermal quality of the final joint.

2. Methodology

2.1. ML algorithms

ML applications have recently become evident across all sectors of industry. In this context, numerous studies

Table 1. Overall comparison between related previous studies

Process	Materials	ML algorithms	Key findings	References
FSW	AA2195	BPNN	Maximum tensile strength of 415 MPa with 92% prediction accuracy.	27
FSSW	Stainless steel and AA7075	SVM and ANN	The predicted tensile shear failure load is 4,967 N at 1,900 rpm.	33
FSP	AA6083 and AA8011	SVM, ANN, and RF	A 5.06 μm grain size is necessary to achieve the ultimate tensile strength.	28
FSW	Aluminum alloys	Multi-linear regression, K-nearest neighbor, RF, ANN, and SVM	ML is used to capture the relationship between the input and output FSW welding process	29
FSW	Aluminum alloys	Decision tree model, SVM, ANN, and CNN	An ML machine is used to achieve two key outcomes: enhanced quality assurance and improved automation	41
FSP	AA6061	ANN	Stir zone hardness prediction	31
FSW	Copper	ANN	Wear rate prediction in surface composites	32
FSW	AA7075	ANN	Peak temperature, torque, traverse force, bending stress, and maximum shear stress	34
FSW	AA5083 and AA5061	GPR and SVM	ML achieves 10.32% higher accuracy in tensile strength than regression methods	35
FSW	AA1100	discrete wavelet analysis and SVM	Using ML approaches to monitor and analyze FSW defect image classification	36
FSW	AA6061	SVM	FSW process parameter optimization	37
FSSW	AA7075	DTR, RF, and linear regression	The DTR can achieve better performance in the prediction of cross-tensile strength	38
FSW	Aluminum alloys	Bayesian ridge, k-nearest neighbor, RF, elastic-net, and SVM	The study highlights the potential of ML to forecast weld joint quality	39
	AA6061 and AZ31B	ANN and RSM	The accuracy of ANN's tensile strength prediction is better than RSM's	40

Abbreviations: ANNs: Artificial neural network; BPNN: Backpropagation neural network; CNN: Convolutional neural network; DTR: Decision tree regression; FSP: Friction stir processing; FSSW: Friction stir spot welding; FSW: Friction stir welding; ML: Machine learning; RF: Random forest; RSM: Response surface methodology; SVMs: Support vector machines.

have recommended adopting ML techniques in welding applications, specifically in two main categories: fusion welding and solid-state welding. Within the framework of fusion welding, Ramesh *et al.*[42] concluded that ML can be appropriate for detecting surface weld defects in electric arc welding, particularly shield metal arc welding. Tsai *et al.*[43] recommended ML as a suitable tool for optimizing laser welding parameters for lap joints. Based on these and other similar studies, this study aims to utilize and compare a range of ML tools to select optimal input parameters for achieving superior mechanical and thermal performance. Random forest, elastic net, Bayesian ridge, polynomial regression, and robust regression methods were used.

The random forest algorithm is considered an ensemble learning method and a supervised ML tool. It is a tree-based, non-parametric algorithm, with a primary purpose being regression or classification.[44,45] In welding, this technique has been used for several applications, including real-time defect detection in arc welding.[46,47] The random forest algorithm works by creating numerous decision trees, each trained on a random subset of the data (known as bootstrapping) and a random selection of features (referred to as feature bagging).[48] Each tree makes its prediction, and for regression tasks, the result is the average of all individual tree outputs. Elastic net regression is a regularized linear regression and a supervised ML approach. It is a parametric, linear model with Lasso and ridge regularizations. The primary purpose of elastic net regression is to enhance model generalization and address multicollinearity.[49] Several studies used elastic net regression in welding applications to assess quality in resistance spot welding.[50] The main purpose of using the Bayesian ridge regression is to incorporate prior distributions and provide uncertainty estimates. It is categorized as a probabilistic, linear regression, supervised ML tool.[51] Welding studies, such as predictive modeling of weld-bead geometry in wire-arc additive manufacturing, also employ Bayesian ridge regression.[52] Polynomial regression and robust regression are two supervised ML techniques used for predictive modeling, but they serve different purposes. Polynomial regression is a parametric approach that builds on linear regression by adding polynomial terms, such as squared or cubed predictors.[53] This enables it to capture curved, nonlinear relationships in the data. On the other hand, robust regression is a type of linear model built to handle messy datasets—especially those with outliers or heavy noise.[54] Unlike standard linear regression, it minimizes the impact of irregularities, making its predictions stable and reliable.

The Bayesian ridge was selected for its ability to produce probabilistic predictions, thereby enhancing the interpretability of uncertainty in tensile and thermal responses. Bayesian ridge provides stable estimates in the presence of multicollinearity, which can occur when process parameters (e.g., rotational speed and dwell time) influence responses in correlated ways. Elastic-net regression is suitable when only a subset of input features may strongly influence the output. It can lead to a better understanding of the influence of FSSW process parameters on the thermal and tensile behavior output. Random forest was selected for its ability to model complex nonlinear interactions among the FSSW parameters. The statistical basis of polynomial regression makes it an easy and suitable tool for interpretable nonlinear alternatives during the FSSW. The kernel functions of support vector regression (SVR) make it a powerful technique for learning nonlinear patterns without overfitting during the study of the FSSW process parameters and output responses. Robust regression reduces the influence of outliers and provides reliable predictions during the welding processes. In this study, Python (v3.15, Python Software Foundation, Netherlands) with Google Colab was used in EDA and ML prediction.

2.2. Experimental work

Aluminum alloy AA 6061 specimens of different thicknesses (1, 2, and 3 mm) and constant length and width (100 mm × 25 mm) were prepared in lap joint configurations and joined by FSSW in this study. A cylindrical welding tool with a diameter of 10 mm and a length of 50 mm was prepared. The pin was ready as a non-threaded type with lengths of 1, 2, and 3.5 mm, depending on the selected aluminum sheet thickness, while the diameter was maintained as a constant (3 mm). The welding tool was made from carbon steel with a carbon content of up to 1.5%. It was characterized by high hardness, wear resistance, and thermal stability, making it efficient at generating heat through friction.

Figure 2 shows the lap joint configuration of the aluminum alloy specimens. A load cell was used to measure the normal force exerted by the welding tool on the specimen surface, with the signal converted to digital using an analog-to-digital converter (U3-LV, LabJack, USA). A welding platform was also prepared to securely hold the specimens. A thermal imaging camera (TC002, TOPDON, USA), compatible with smartphones, was used to monitor temperature evolution during and after welding by tracking temperature loss and decline over time. The thermal camera has a resolution of 256 × 192 pixels and a high heat sensitivity of 40 mK. The chemical composition of the AA 6061 alloys, as measured using a metal analyzer (SPECTROPORT PXC01, AMETEK, USA), is presented in Table 2. Table 2 also shows the mechanical properties of AA 6061 aluminum.

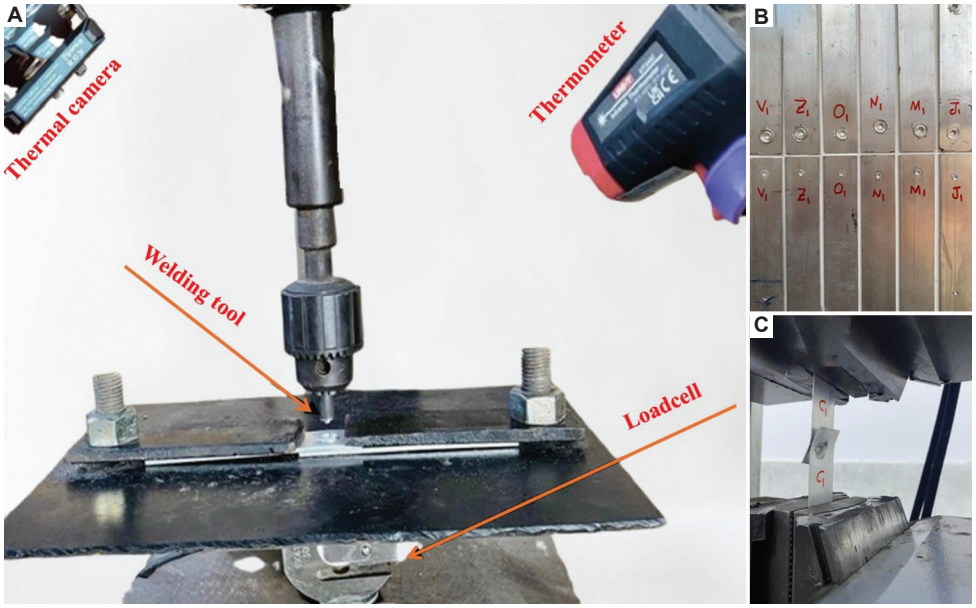


Figure 2. Friction stir spot welding process. (A) Welding arrangements; (B) Welded specimens; (C) Lap-shear tensile test.

Table 2. Characteristics of AA 6061 aluminum

Characteristics	Value
Chemical composition (element, %)	
Chromium	0.260
Copper	0.340
Magnesium	1.080
Iron	0.660
Silicon	0.550
Manganese	0.110
Titanium	0.160
Zinc	0.223
Aluminum	96.617
Mechanical properties	
Hardness (HV)	40
Yield strength (MPa)	276
Tensile strength (MPa)	310

Due to the nonlinearity in predicting the maximum thermal and tensile responses of the FSSW joint, this topic remains actively studied. This study was designed to predict the maximum temperature (°C) and the maximum lap-shear fracture load (kN) from FSSW process parameters. The total number of specimens used in this study was 81; every three specimens were examined with the same input parameters. The parameters involved were sheet thickness (mm), tool rotational speed (rpm), and dwelling time (s). The selected input parameter values were based on previous studies. The average of each of the three specimens was calculated. Table 3 shows the values of the welding parameters used, the corresponding maximum temperatures, and lap-shear fracture loads. The EDA and ML techniques were then used to maximize output response for temperature and lap-shear fracture load.

3. Results and discussion

3.1. Evaluation of EDA

This study employed EDA to analyze sets of experimental results obtained during the FSSW process. Several specific procedures were followed to ensure the accuracy of the results. The input parameters were aluminum sheet thickness, dwelling time, and tool rotational speed. The primary purpose of using EDA is to provide a deep, comprehensive understanding of the dataset characteristics and their relationship with the output variables—maximum temperature (°C) and lap-shear fracture load (kN). Histograms were plotted to visualize the distribution of key input parameters: thickness (mm) (Figure 3), rotational speed (rpm) (Figure 4), and dwell time (s) (Figure 5). These visualizations revealed variations in parameter distributions, highlighting potential areas of skewness or concentration.

Binning is a practical preprocessing method that simplifies complex data during exploratory analysis. By grouping raw continuous values into manageable intervals—or "bins"—this technique reduces noise and identifies underlying trends.[55] For the x-axis in Figures 3-5, binning was used to group the 27 original datasets into eight intervals, yielding clean, representative bar graphs that highlighted meaningful patterns.

Table 3. Friction stir spot welding process parameters and corresponding output parameters

No.	Thickness (mm)	Rotation speed (rpm)	Dwelling time (s)	Max temperature (°C)	Lap-shear fracture load (kN)
1	1	1,000	3	74	1.23
2	1	1,500	3	102	1.42
3	1	2,000	3	112	1.61
4	1	1,000	6	93	1.33
5	1	1,500	6	139	1.52
6	1	2,000	6	166	1.73
7	1	1,000	9	125	1.55
8	1	1,500	9	147	2.00
9	1	2,000	9	176	2.20
10	2	1,000	3	85	1.20
11	2	1,500	3	115	1.34
12	2	2,000	3	128	1.53
13	2	1,000	6	96	1.65
14	2	1,500	6	133	1.84
15	2	2,000	6	165	2.19
16	2	1,000	9	132	2.26
17	2	1,500	9	153	2.33
18	2	2,000	9	179	2.59
19	3	1,000	3	97	1.51
20	3	1,500	3	126	1.82
21	3	2,000	3	138	2.11
22	3	1,000	6	118	1.84
23	3	1,500	6	148	2.14
24	3	2,000	6	187	2.35
25	3	1,000	9	136	2.13
26	3	1,500	9	180	2.36
27	3	2,000	9	215	2.54

Figure 3 presents the histogram for aluminum sheet thickness during the FSSW process. The statistics reveal a standard deviation of approximately 0.832 over the 1–3 mm range, with a median value of 2 mm. The 25th, 50th, and 75th percentiles are 1, 2, and 3 mm, respectively, indicating a reasonably symmetric distribution around the median. Such a balanced distribution is critical, as it indicates the dataset is free from significant skewness or extreme outliers that could otherwise distort the analysis.[55]

The histogram of aluminum sheet thickness did not differ significantly from that of tool rotation speed (Figure 4) and dwelling time (Figure 5), with a symmetric distribution centered around the mean. The standard deviations were 392.558 and 2.496 for the tool rotational speed and the dwelling time, respectively. The 25th, 50th, and 75th percentiles of the rotational speeds were 1,000 rpm, 1,500 rpm, and 2,000 rpm, respectively, while those of the dwelling times were 3 s, 6 s, and 9 s, respectively.

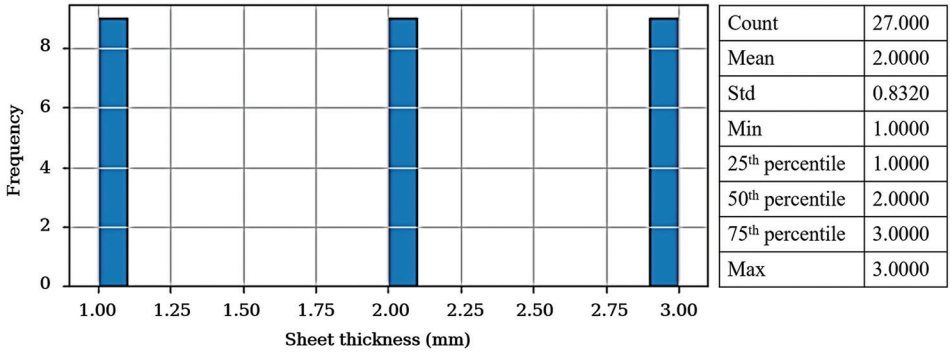
EDA, a crucial component of the FSSW process analysis in this study, involved generating correlations between thermal behavior and mechanical behavior maps to visualize linear relationships between pairs of variables during the FSSW of the aluminum alloy.

A correlation heat map is typically used to display the strength and direction of linear relationships between numerical variables in a dataset. The correlation values ranged from −1 to +1, while 0 indicates no linear correlation, −1 indicates a strong negative correlation (as one variable increases, the other decreases), and +1 indicates a strong positive correlation (both variables increase together). Moreover, the heat map shows a color-coded matrix, in which darker colors indicate stronger correlations, while lighter or neutral colors indicate weaker or no correlation.

In FSSW studies, the correlation heat map reveals how welding parameters, such as thickness, rotational speed, and dwell time, correlate with output responses, including maximum temperature and lap-shear fracture load. Figure 6 shows the heat map of correlations among the variables in this study. A value of 1.00 indicates perfect correlation, whereas values close to 0 indicate weak or no correlation. There are strong links between maximum temperature and lap-shear fracture load (r = 0.86) and between dwelling time and lap-shear fracture load (r = 0.96). At the same time, there is a weak correlation between thickness and maximum temperature (r = 0.28).

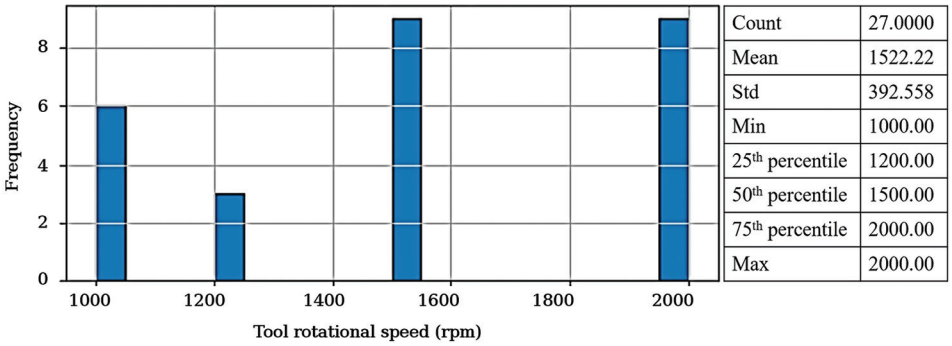
It is crucial to verify the significance of features in ML algorithms after model training, as it measures the impact of each variable (feature) on the model's results, reveals the contribution of each input to the output values, and identifies the most influential variables. This study used feature significance as a linear regression model to illustrate the impact of friction welding process parameters (dwell time, sheet thickness, and tool rotation speed) on the process outputs (maximum temperature and lap-shear fracture load).

Dwell time demonstrated the highest feature importance for lap-shear strength, indicating it most strongly influences weld quality. The feature importance of dwelling time was 0.28, corresponding to lower values for sheet thickness (0.19) and tool rotational speed (0.18), as shown in Figure 7. The most significant impact of dwell time, as the highest feature importance for lap-shear strength, might be attributed to greater frictional heating between the tool and the workpieces. The increase in



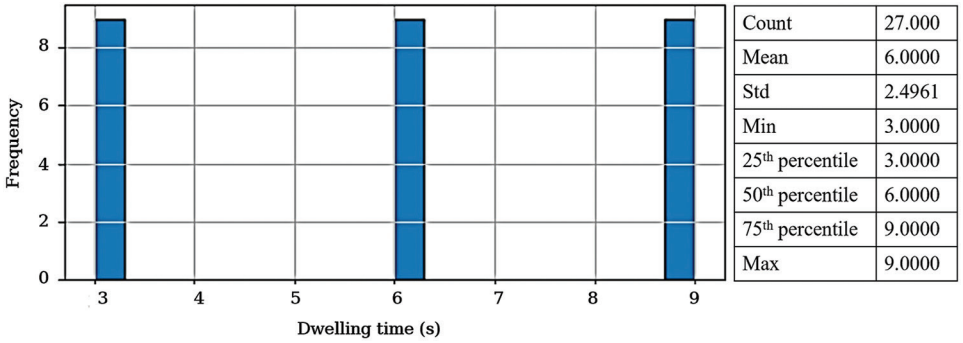
Count	27.000
Mean	2.0000
Std	0.8320
Min	1.0000
25th percentile	1.0000
50th percentile	2.0000
75th percentile	3.0000
Max	3.0000

Figure 3. Histogram analysis of the aluminum sheet thickness (mm) during the friction stir spot welding



Count	27.0000
Mean	1522.22
Std	392.558
Min	1000.00
25th percentile	1200.00
50th percentile	1500.00
75th percentile	2000.00
Max	2000.00

Figure 4. Histogram analysis of the tool rotational speed (rpm) during the friction stir spot welding



Count	27.000
Mean	6.0000
Std	2.4961
Min	3.0000
25th percentile	3.0000
50th percentile	6.0000
75th percentile	9.0000
Max	9.0000

Figure 5. Histogram analysis of the dwelling time (s) during the friction stir spot welding

dwelling time extended the time the tool spent rotating in the weld pool, resulting in greater homogeneity and mixing between the two aluminum plates during the FSSW process. This improved mixing and homogeneity ultimately led to increased shear forces and enhanced mechanical properties.

Figure 7 reflects the nature of the specimens in FSSW. The lower the dwell time, the poorer the joint quality, and thereby the weaker the tensile strength. Conversely, a longer dwell time resulted in a rigid joint, requiring a greater force to cause failure and lap fracture, as shown in Table 4 (tool rotational speed of 1,500 rpm and aluminum sheet thickness of 3 mm).

Tool rotational speed had the largest influence on maximum temperature, with a feature importance of 22.72. However, dwell time did not significantly affect the maximum temperature, with a feature importance of 21.13, slightly lower than that of tool rotation speed. The effect of aluminum sheet thickness was small, accounting for approximately 48% of the total effect, compared to the effects of rotation speed and dwell time, with a feature importance of 9.57, as shown in Figure 8.

The thermal imaging temperature readings confirm the EDA results. Under constant-sheet-thickness aluminum samples (2 mm) and a constant dwelling time (6 s), the controlled increase in rotational speed from 1,000 to

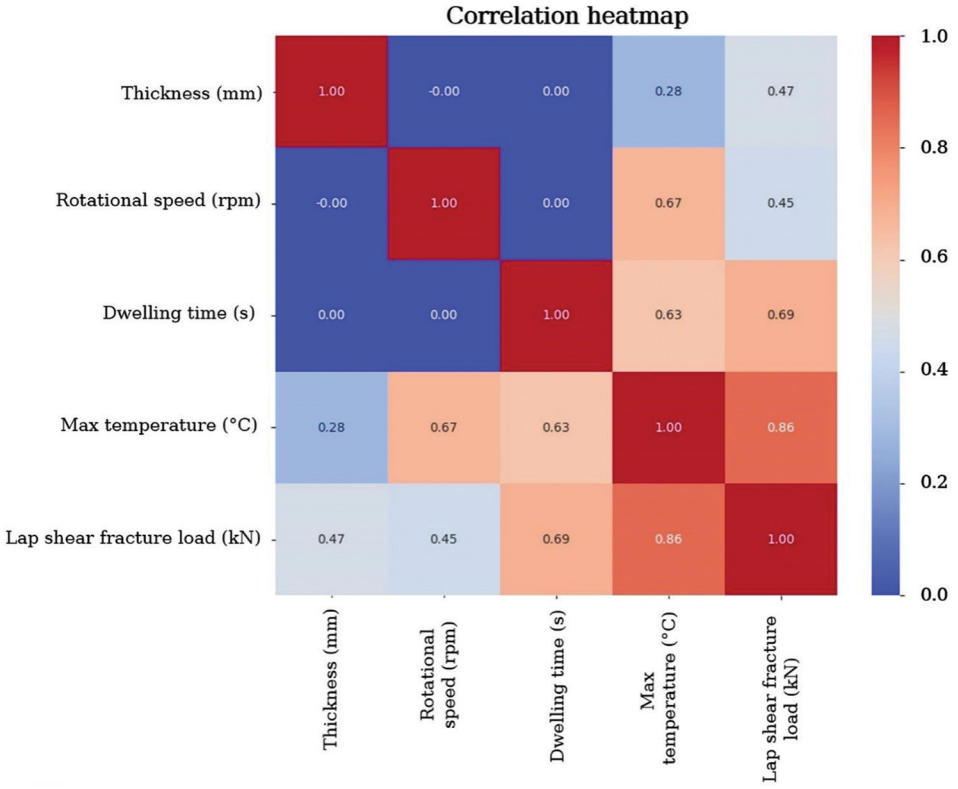


Figure 6. Correlation heatmap for the input and output parameters during the friction stir spot welding

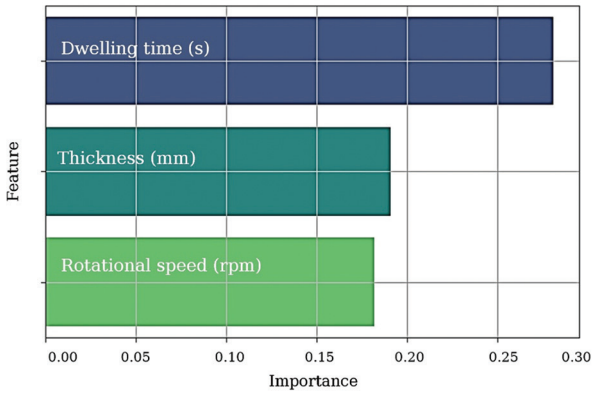


Figure 7. Feature importance analysis of the friction stir spot welding process parameters on the lap-shear fracture load

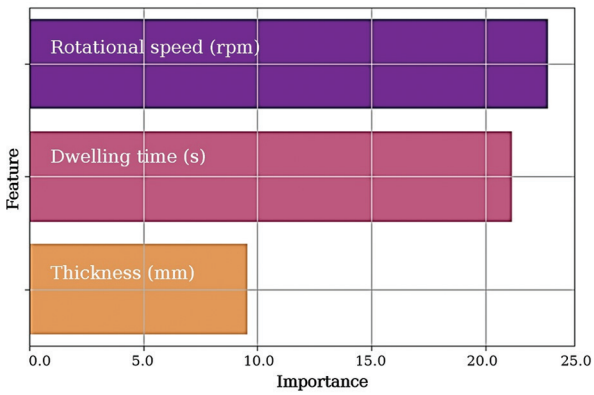


Figure 8. Feature importance analysis of the friction stir spot welding process parameters on maximum temperature

2,000 rpm, with a 500 rpm step, led to nonlinear behavior in the generated temperature of 96°C, 133°C, and 165°C, respectively, as shown in Figure 9. The increase in temperature can be attributed to increased frictional heat generation and the intensification of plastic deformation as the rotational speed increases.

3.2. Effectiveness of ML predictions

This study examined and compared several ML algorithms. Random forest, elastic net, Bayesian ridge, polynomial regression, and robust regression were used to predict thermal behavior (maximum temperature) and the mechanical behavior (lap-shear fracture load) during the FSSW process. Mean squared error (MSE), mean absolute error (MAE), and coefficient of determination (R^2) were used to evaluate the ML models. MSE measures the average squared prediction error; lower values indicate that the predictions are closer to the actual values. Moreover, it penalizes larger errors heavily, making it useful when large deviations in the FSSW process are undesirable. MAE is

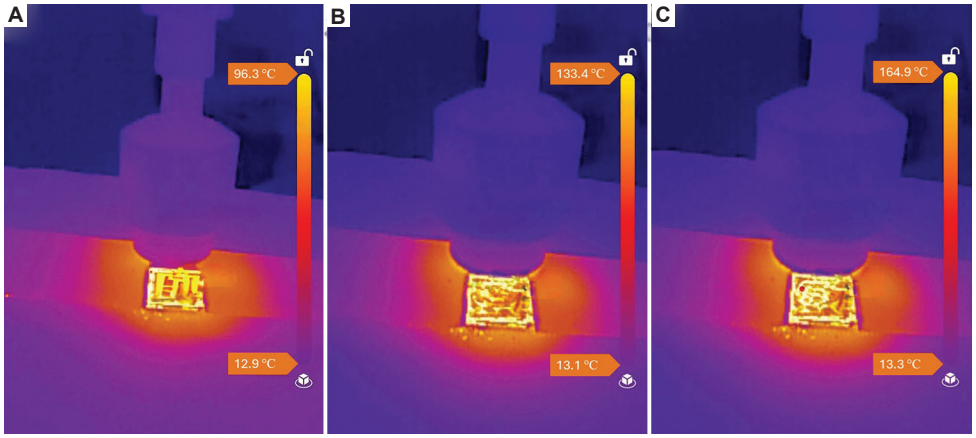
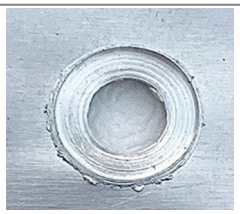
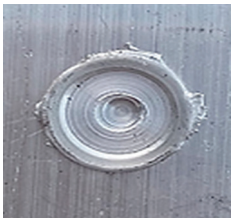


Figure 9. Effect of rotational speed on the maximum temperature in the friction stir spot welding process at a dwelling time of 6 sec and a sheet thickness of 2 mm. (A) 1,000 rpm; (B) 1,500 rpm; (C) 2,000 rpm.

Table 4. Effect of dwelling time of lap-shear fracture in friction stir spot welding at 1,500 rpm and 3 mm thickness

Dwelling time (s)	Lap-shear fracture load (kN)	Joint after FSSW	Joint after fracture
3	1.82		
6	2.14		
9	2.36		

Abbreviations: FSSW: Friction stir spot welding.

the average of the absolute differences between predictions and actual values during the FSSW process; lower MAE values indicate higher average prediction accuracy. The coefficient of determination (R^2) is an indicator of how well the independent variables explain the variance in the dependent variable during the experimental process. Their values ranged from 0 to 1, depending on the model that provided the better fit. The following subsections will examine the accuracy of the ML algorithms based on the obtained values of MSE, MAE, and R^2.

3.2.1. Evaluation of ML models in predicting thermal behavior

The total number of FSSW parameter sets examined in this section is 27. All selected ML models (random forest algorithm, elastic net regression, Bayesian ridge regression,

SVR, polynomial regression, and robust regression) were evaluated using MSE, MAE, and R^2. This step is essential for both linear and nonlinear regression algorithms to compare the efficacy of ML models. Figure 10 shows significantly lower MSE values for all ML models, except SVR, indicating suitable performance across all algorithms. The high MSE in SVR compared to other models suggests poor performance in predicting maximum temperatures. This ineffective behavior of the SVR may be due to the applied SVR kernel struggling to capture complex, nonlinear relationships between input features (like thickness, speed, and dwell time) and the corresponding temperature. Moreover, in numerous engineering applications, SVR can be data-sensitive, leading to poor generalization and higher errors.[56]

The random forest algorithm demonstrated superior performance (MSE = 26.51) in predicting maximum temperature due to its ability to model complex relationships, such as those in the FSSW process, between FSSW process parameters and output responses. Moreover, random forest is robust to outliers and noise, which are often present in experimental welding datasets.

Figure 11 shows the MAE values of various ML methods applied to temperature prediction during the FSSW process. A notable observation is the strong alignment between the MSE and MAE results. Although these two metrics differ in how they mathematically penalize errors and handle outliers, their parallel behavior here confirms that they provide a consistent assessment of the models' prediction accuracy. Notably, the experimental data were collected under the supervision of highly experienced FSSW personnel, resulting in very few, if any, outliers. This approach led to greater prediction stability for all ML approaches used in this study.

Figure 12 displays the R^2 values of various regression models. Models such as the random forest algorithm, elastic net regression, Bayesian ridge regression, polynomial regression, and robust regression algorithm achieved high R^2 values, indicating that they explained a large portion of the variance in the data and thus demonstrated better predictive power. In contrast, the SVR exhibited low R^2 values, suggesting poor performance in explaining the data's variance. The comparative analysis of the models using MSE, MAE, and R^2 scores highlights several key findings. The random forest algorithm model stands out with the lowest MSE (26.51) and MAE (3.86), along with the highest R^2 score (0.97), demonstrating exceptional prediction accuracy and a strong fit to the data.

The polynomial regression model also exhibited robust performance with low MSE (51.63) and MAE (6.26), and a high R^2 score (0.95). Similarly, Bayesian ridge regression

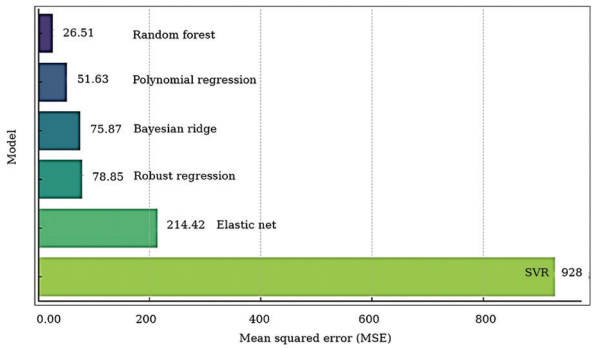


Figure 10. Evaluation of machine learning algorithms in maximum temperature prediction using mean squared error during the friction stir spot welding process
Abbreviation: SVR: Support vector regression.

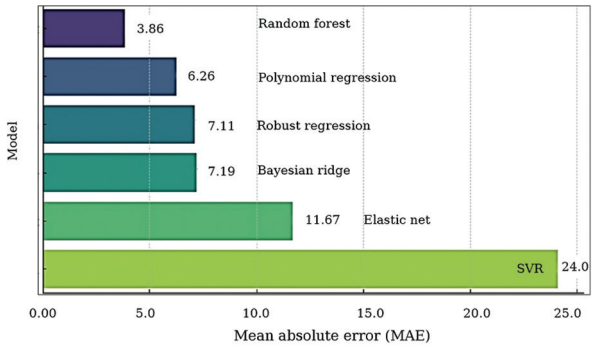


Figure 11. Evaluation of machine learning algorithms in maximum temperature prediction using mean absolute error during the friction stir spot welding process
Abbreviation: SVR: Support vector regression.

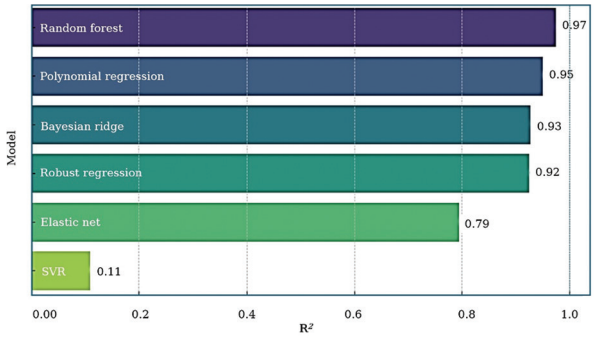


Figure 12. Evaluation of machine learning algorithms in maximum temperature prediction using R^2 during the friction stir spot welding process
Abbreviation: SVR: Support vector regression.

demonstrated remarkably low MSE (75.87), low MAE (7.19), and a high R^2 score (0.93), indicating its effectiveness.

On the other hand, with high MSE (928.44), MAE (24.09), and a low R^2 score (0.11), SVR performs poorly,

indicating that it is the least appropriate model without additional adjustment. Except for the elastic net, which yielded an R^2 of 0.7, other models exhibited strong R^2 values, making them suitable for temperature prediction. For applications requiring high predictive precision, the random forest, robust regression, polynomial regression, and Bayesian ridge regression models are recommended.

This study reports poor/good SVR responses in thermal and tensile performance, respectively. Temperature is monitored in real time using an infrared thermometer and thermal camera throughout the welding process. Due to the high nonlinearity of the friction temperature at the interface between the pin and the shoulder bottom, which are in contact with the aluminum alloy specimens, it is difficult to control the temperature and to tune the model or use SVR kernels to capture it effectively. The poor performance of SVR relative to other ML models in thermal applications, such as welding, has been well documented in previous studies.[57] Careful specimen preparation and controlled testing conditions likely reduced experimental variability, which may have improved model performance for tensile outcomes. The mechanical properties exhibited consistent, stable relationships with the input features, making them amenable to SVR modeling with the chosen parameters. Therefore, the SVR performed better in predicting mechanical properties than thermal behavior. This better SVR performance related to other ML techniques was also noticed in previous studies.[39]

3.2.2. Evaluation of ML models in predicting mechanical behavior

The total number of FSSW parameter sets examined in this section is 27. The previously mentioned models were also used to study and predict the mechanical properties of the weld joint and to compare them based on MSE, MAE, and R^2. The elastic net model yielded an MSE of 0.16, indicating poor model performance in this process. In contrast, the random forest, SVR, Bayesian ridge, and polynomial regression models exhibited favorable performance, with MSE values below 0.02, as shown in Figure 13. In contrast, Figure 14 compares the ML algorithms based on the MAE for predicting the maximum lap-shear fracture load. The lowest MAE values were observed with the random forest, SVR, and robust regression algorithms, at 0.07, 0.08, and 0.09, respectively. On the other hand, the elastic net model yielded a high value of 0.3, indicating poor performance in predicting the tensile performance of the FSSW weld joint, as indicated by the MAE.

Figure 15 displays the R^2 values as a bar chart of various regression models for the prediction of maximum lap-shear fracture load. Models such as the random forest

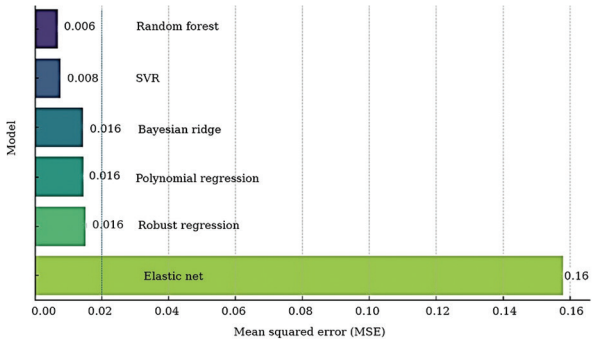


Figure 13. Evaluation of machine learning algorithms in predicting maximum lap-shear fracture load of weld joint using mean squared error during the friction stir spot welding process
Abbreviation: SVR: Support vector regression.

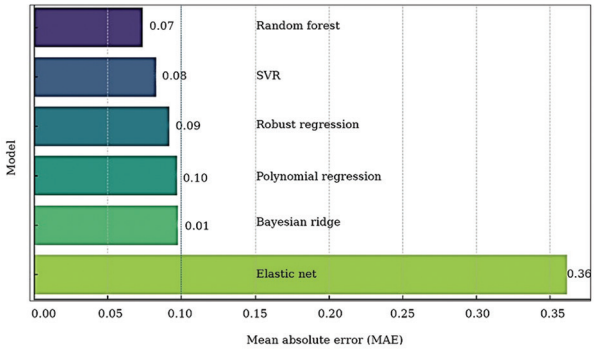


Figure 14. Evaluation of machine learning algorithms in predicting the maximum lap-shear fracture load of the weld joint using the mean absolute value during the friction stir spot welding process
Abbreviation: SVR: Support vector regression.

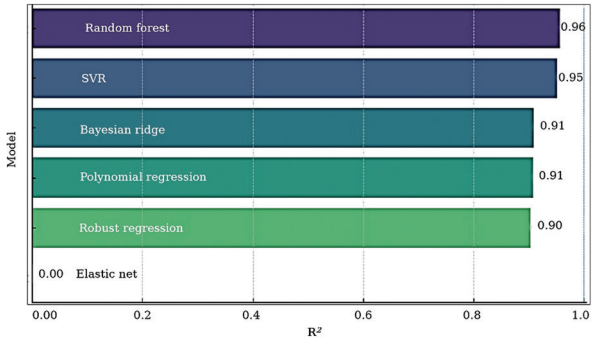


Figure 15. Evaluation of machine learning algorithms in predicting the maximum lap-shear fracture load of the weld joint using the coefficient of determination (R^2) during the friction stir spot welding process
Abbreviation: SVR: Support vector regression.

algorithm, elastic net regression, Bayesian ridge regression, polynomial regression, and robust regression algorithm achieve high R^2 values. All models recorded excellent performance in predicting the mechanical properties of the weld joint with R^2 values approaching 0.9, except for

the elastic net regression model, which showed inaccuracy and poor performance of $R^2 = 0$.

Based on the evaluation and comparison of models, R^2, MAE, and MSE, the elastic net regression model showed the weakest performance in terms of predicting the mechanical performance of the weld joint, as it recorded the lowest value for $R^2 = 0$, MAE = 0.36, and MSE = 0.16.

4. Conclusion

The major research question in this study concerns the feasibility of using EDA and ML approaches to determine the optimal thermal and tensile performances of aluminum joints welded using FSSW. The main challenge focuses on modeling the nonlinear behavior of the FSSW process outputs. The key findings that can be concluded from this study are as follows:

(i) In the FSSW process, there is no single model that can be relied upon for evaluation and prediction to suit all objectives, as the outcome depends on the variables of the welding process. In general, high-flexibility algorithms, such as the random forest algorithm and polynomial regression, have proven effective for predicting welding process outputs because they handle nonlinear relationships. The SVR model did not yield satisfactory results in determining the optimal thermal or mechanical performance of FSSW within the specified welding parameters. In contrast, Bayesian ridge regression, while not performing as well as the random forest algorithm, is considered satisfactory and can be recommended.

(ii) ML models are suitable tools to determine input variables. The dwelling time may have the highest significance (0.28) for the lap-shear fracture load, indicating that it significantly affects weld quality. In contrast, the tool rotation speed had the greatest impact on the maximum temperature, with a feature importance of 22.72.

(iii) The evaluation of ML models using MSE, MAE, and R^2 is essential for comparing the ML algorithms that can be used to predict thermal and tensile performance of the FSSW process.

The application of EDA—limited in this study to determining relationships among welding process variables for algorithm selection—together with ML techniques provided valuable insights into the thermal and mechanical behavior of the resulting joints. These results suggest that the proposed approach may serve as an effective analytical framework for complex solid-state welding processes, including repetitive friction-flip welding and double friction-flip welding. For future studies, ML techniques may be combined with X-ray diffraction images to classify joint defects and to accurately investigate metallic effects in friction-flip weld joints.

Acknowledgments

None.

Funding

None.

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: Raheem Al-Sabur
Formal analysis: Sajad N. Alasadi
Investigation: All authors
Methodology: Sajad N. Alasadi
Writing–original draft: Sajad N. Alasadi
Writing–review & editing: Raheem Al-Sabur

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

The data presented in this study are available upon request from the corresponding author.

References

1. Chen H, Zhu Z, Zhu Y, Sun L, Guo Y. Solid-state welding of aluminum to magnesium alloys: A review. *Metals (Basel)*. 2023;13(8):1410.

   doi: 10.3390/met13081410

2. Khedr M, Hamada A, Järvenpää A, Elkatatny S, Abd-Elaziem W. Review on the solid-state welding of steels: Diffusion bonding and friction stir welding processes. *Metals (Basel)*. 2023;13(1):54.

   doi: 10.3390/met13010054

3. Badavath HJ, Chattopadhyay S, Shankar S. Solid-State Welding and its Applications: A Methodological Review. In: *AIP Conference Proceedings*. Vol. 2681; 2022.

   doi: 10.1063/5.0116893

4. Gibson BT, Lammlein DH, Prater TJ, *et al*. Friction stir welding: Process, automation, and control. *J Manuf Process*. 2014;16(1):56-73.

   doi: 10.1016/j.jmapro.2013.04.002

5. Majeed T, Wahid MA, Alam MN, Mehta Y, Siddiquee AN. Friction stir welding: A sustainable manufacturing process. In: *Materials Today: Proceedings*. Vol. 46; 2020.

   doi: 10.1016/j.matpr.2021.04.025

6. Khalaf HI, Al-Sabur R. Friction stir techniques exploring fracture and fatigue behavior in hybrid composite joints. In: *Utilizing Friction Stir Techniques for Composite Hybridization*. United States: IGI Global; 2024. p. 108-134.

   doi: 10.4018/979-8-3693-3993-0.ch007

7. Avettand-Fènoël MN, Simar A. A review about friction stir welding of metal matrix composites. *Mater Charact*. 2016;120:1-17.

   doi: 10.1016/j.matchar.2016.07.010

8. Belalia SE, Serier M, Al-Sabur R. Parametric analysis for torque prediction in friction stir welding using machine learning and shapley additive explanations. *J Computat Appl Mech*. 2024;55(1):113-124.

   doi: 10.22059/JCAMECH.2024.370055.924

9. Tiwari A, Singh P, Pankaj P, Biswas P, Kore SD. FSW of low carbon steel using tungsten carbide (WC-10wt.%Co) based tool material. *J Mech Sci Technol*. 2019;33(10):4931-4938.

   doi: 10.1007/s12206-019-0932-7

10. Serier M, Jassim RJ, Al-Sabur R, Siddiquee AN. Thermal Diffusivity Modeling for Aluminum AA6060 Plates During Friction Stir Welding. In: *AIP Conference Proceedings*. Vol. 3051; 2024.

    doi: 10.1063/5.0191742

11. Jassim AK, Al-Subar RK. Studying the possibility to Weld AA1100 aluminum alloy by friction stir spot welding. *Int J Mater Metall Eng*. 2017;11(9):660.

12. Borah MJ, Saha N. Effect of Tool Geometry and Process Parameters on Strength of Various Friction Stir Spot Welded Lap Joints. In: *Proceedings of the World Congress on Mechanical, Chemical, and Material Engineering*; 2019.

    doi: 10.11159/icmie19.124

13. Al-Sabur R, Serier M. Material sustainability during friction stir joining. In: *Sustainable Machining and Green Manufacturing*. United States: Wiley; 2024.

    doi: 10.1002/9781394197866.ch7

14. Sree Sabari S, Malarvizhi S, Balasubramanian V. Characteristics of FSW and UWFSW joints of AA2519-T87 aluminium alloy: Effect of tool rotation speed. *J Manuf Process*. 2016;22:278-289.

    doi: 10.1016/j.jmapro.2016.03.014

15. Alasdi SN, Al-Sabur R. In-depth thermal analysis of different pin configurations in friction stir spot welding of similar and dissimilar alloys. *J Manuf Mater Process*. 2025;9(6):184.

    doi: 10.3390/jmmp9060184

16. Marazani T, Jeje SO, Shongwe MB, Malatji N. Mass flash reduction strategies in friction stir processing of aluminum alloys: A review. *Eng Rep*. 2024;6(10):e12981.

    doi: 10.1002/eng2.12981

17. Morgenthaler S. Exploratory data analysis. *Wiley Interdiscip Rev Comput Stat*. 2009;1(1):33-44.

    doi: 10.1002/wics.2

18. Bezerra A, Silva I, Guedes LA, Silva D, Leitão G, Saito K. Extracting value from industrial alarms and events: A data-driven approach based on exploratory data analysis. *Sensors (Basel)*. 2019;19(12):2772.

    doi: 10.3390/s19122772

19. Mathew BK, Velmurugan C, Vanitha V. Estimating Most Influencing Parameter in TIG Welding of SS304 Using Exploratory Data Analysis. In: *2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation, ICAECA 2021*; 2021.

    doi: 10.1109/ICAECA52838.2021.9675521

20. Wang LC. Experience of data analytics in EDA and test - Principles, promises, and challenges. *IEEE Trans Comput Aided Des Integr Circ Syst*. 2017;36(6):885-898.

    doi: 10.1109/TCAD.2016.2621883

21. Alzubi J, Nayyar A, Kumar A. Machine learning from theory to algorithms: An overview. *J Phys Conf Ser*. 2018;1142:012012.

    doi: 10.1088/1742-6596/1142/1/012012

22. Tufail S, Riggs H, Tariq M, Sarwat AI. Advancements and challenges in machine learning: A comprehensive review of models, libraries, applications, and algorithms. *Electronics (Switzerland)*. 2023;12(8):1789.

    doi: 10.3390/electronics12081789

23. Wang Z, Li L, Chen H, *et al*. Penetration recognition based on machine learning in arc welding: A review. *Int J Adv Manuf Technol*. 2023;125(9-10):3899-3923.

    doi: 10.1007/s00170-023-11035-7

24. Maculotti G, Genta G, Galetto M. Optimisation of laser welding of deep drawing steel for automotive applications by Machine Learning: A comparison of different techniques. *Qual Reliab Eng Int*. 2024;40(1):202-219.

    doi: 10.1002/qre.3377

25. Sun H, Ramuhalli P, Jacob RE. Machine learning for ultrasonic nondestructive examination of welding defects: A systematic review. *Ultrasonics*. 2023;127:106854.

    doi: 10.1016/j.ultras.2022.106854

26. Zhou B, Pychynski T, Reischl M, Kharlamov E, Mikut R. Machine learning with domain knowledge for predictive quality monitoring in resistance spot welding. *J Intell Manuf*. 2022;33(4):1139-1163.

doi: 10.1007/s10845-021-01892-y

27. Yu F, Zhao Y, Lin Z, Miao Y, Zhao F, Xie Y. Prediction of mechanical properties and optimization of friction stir welded 2195 aluminum alloy based on BP neural network. *Metals (Basel)*. 2023;13(2):267.

    doi: 10.3390/met13020267

28. Verma S, Msomi V, Mabuwa S, *et al*. Machine learning application for evaluating the friction stir processing behavior of dissimilar aluminium alloys joint. *Proc Instit Mech Eng L J Mater Des Appl*. 2022;236(3):633-646.

    doi: 10.1177/14644207211053123

29. Elsheikh AH. Applications of machine learning in friction stir welding: Prediction of joint properties, real-time control and tool failure diagnosis. *Eng Appl Artif Intell*. 2023;121:105961.

    doi: 10.1016/j.engappai.2023.105961

30. Chadha U, Selvaraj SK, Gunreddy N, *et al*. A Survey of machine learning in friction stir welding, including unresolved issues and future research directions. *Mater Des Process Commun*. 2022;2022:2568347.

    doi: 10.1155/2022/2568347

31. Ebnonnasir A, Karimzadeh F, Enayati MH. Novel artificial neural network model for evaluating hardness of stir zone of submerge friction stir processed Al 6061-T6 plate. *Mater Sci Technol*. 2011;27(6):990-995.

    doi: 10.1179/174328409X425290

32. Dinaharan I, Palanivel R, Murugan N, Laubscher RF. Application of artificial neural network in predicting the wear rate of copper surface composites produced using friction stir processing. *Aust J Mech Eng*. 2022;20(4):1079-1090.

    doi: 10.1080/14484846.2020.1769803

33. Pattanaik AK, Pal K, Mishra D. Tribiological investigation and optimization of friction stir spot welding of dissimilar metals by LSSM-ANN method. *Mech Based Des Struct Mach*. 2022;50(5):1595-1613.

    doi: 10.1080/15397734.2020.1759429

34. Manvatkar VD, Arora A, De A, DebRoy T. Neural network models of peak temperature, torque, traverse force, bending stress and maximum shear stress during friction stir welding. *Sci Technol Weld Join*. 2012;17(6):460-466.

    doi: 10.1179/1362171812Y.0000000035

35. Matitopanum S, Pitakaso R, Sethanan K, Srichok T, Chokanat P. Prediction of the ultimate tensile strength (UTS) of asymmetric friction stir welding using ensemble machine learning methods. *Processes*. 2023;11(2):460-466.

    doi: 10.3390/pr11020391

36. Bhat NN, Kumari K, Dutta S, Pal SK, Pal S. Friction stir weld classification by applying wavelet analysis and support vector machine on weld surface images. *J Manuf Process*. 2015;20:274-281.

    doi: 10.1016/j.jmapro.2015.07.002

37. Sefene EM, Tsegaw AA, Mishra A. Process parameter optimization of 6061AA friction stir welded joints using supervised machine learning regression-based algorithms. *J Soft Comput Civil Eng*. 2022;6(1):274-281.

    doi: 10.22115/SCCE.2022.299913.1350

38. Bock FE, Paulsen T, Brkovic N, *et al*. Evaluation of Mechanical Property Predictions of refill Friction Stir Spot Welding Joints via Machine Learning Regression Analyses on DoE Data. In: *ESAFORM 2021 - 24th International Conference on Material Forming*; 2021.

    doi: 10.25518/esaform21.2589

39. Mothilal M, Kumar A. Predictive modeling of ultimate tensile strength in dissimilar friction stir welded aluminum alloys via machine learning approach. *Philos Mag Lett*. 2025;105(1):2472669.

    doi: 10.1080/09500839.2025.2472669

40. Sambath Y, Natarajan R, Babu PK, *et al*. Comparative analysis of predictive modeling techniques for mechanical properties in dissimilar friction stir welding of AA6061 and AZ31B. *J Mater Eng Perform*. 2025;34:15597-15613.

    doi: 10.1007/s11665-024-10317-9

41. El-Said EMS, Abd Elaziz M, Elsheikh AH. Machine learning algorithms for improving the prediction of air injection effect on the thermohydraulic performance of shell and tube heat exchanger. *Appl Therm Eng*. 2021;185:116471.

    doi: 10.1016/j.applthermaleng.2020.116471

42. Ramesh K, Ramana EV, Srikanth L, Sri Harsha C, Kiran Kumar N. Identification of SMAW surface weld defects using machine learning. In: *Lecture Notes in Mechanical Engineering*. Berlin: Springer; 2023.

    doi: 10.1007/978-981-19-5347-7_28

43. Tsai YA, Lo YL, Raza MM, *et al*. Optimization of lap-joint laser welding parameters using high-fidelity simulations and machine learning mode. *J Mater Res Technol*. 2023;24:6876-6892.

    doi: 10.1016/j.jmrt.2023.04.256

44. España VJ, Aparicio J, Barber X. An adaptation of Random Forest to estimate convex non-parametric production technologies: An empirical illustration of efficiency measurement in education. *Int Trans Oper Res*. 2025;32(5):2523-2546.

    doi: 10.1111/itor.13561

45. Tabib S, Larocque D. Non-parametric individual treatment effect estimation for survival data with random forests. *Bioinformatics*. 2020;36(2):629-636.

    doi: 10.1093/bioinformatics/btz602

46. Zhang Z, Yang Z, Ren W, Wen G. Random forest-based real-time defect detection of Al alloy in robotic arc welding using optical spectrum. *J Manuf Process*. 2019;42:51-59.

doi: 10.1016/j.jmapro.2019.04.023

47. Zhu C, Yuan H, Ma G. An active visual monitoring method for GMAW weld surface defects based on random forest model. *Mater Res Express*. 2022;9(3):036503.

doi: 10.1088/2053-1591/ac5a38

48. Strobl C, Malley J, Tutz G. An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychol Methods*. 2009;14(4):323-348.

doi: 10.1037/a0016973

49. Zhang Z, Lai Z, Xu Y, Shao L, Wu J, Xie GS. Discriminative elastic-net regularized linear regression. *IEEE Trans Image Process*. 2017;26(3):1466-1481.

doi: 10.1109/TIP.2017.2651396

50. Martín Ó, Ahedo V, Santos JI, De Tiedra P, Galán JM. Quality assessment of resistance spot welding joints of AISI 304 stainless steel based on elastic nets. *Mater Sci Eng A*. 2016;676:173-181.

doi: 10.1016/j.msea.2016.08.112

51. Michimae H, Emura T. Bayesian ridge estimators based on copula-based joint prior distributions for regression coefficients. *Comput Stat*. 2022;37(5):2741-2769.

doi: 10.1007/s00180-022-01213-8

52. Šket K, Brezočnik M, Karner T, *et al.* Predictive modelling of weld bead geometry in wire arc additive manufacturing. *J Manuf Mater Process*. 2025;9(2):67.

doi: 10.3390/jmmp9020067

53. Hinkle J, Muralidharan P, Fletcher PT, Joshi S. Polynomial regression on Riemannian manifolds. In: *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol 7574. Berlin: Springer; 2012.

doi: 10.1007/978-3-642-33712-3_1

54. Liu C, Li B, Vorobeychik Y, Oprea A. Robust linear regression against training data poisoning. In: *AISec 2017 - Proceedings of the 10th ACM Workshop on Artificial Intelligence and Security, Co-Located with CCS 2017*; 2017.

doi: 10.1145/3128572.3140447

55. Sandfeld S. Exploratory Data Analysis. In: *Materials Data Science. The Materials Research Society Series*. Cham: Springer; 2024. p. 179-206.

doi: 10.1007/978-3-031-46565-9_9

56. Li S, Li M, Liu Z, Li M. A Data-Driven Residual Life Prediction Method for Rolling Bearings. In: *Proceedings of 2023 IEEE 12th Data Driven Control and Learning Systems Conference, DDCLS 2023*; 2023.

doi: 10.1109/DDCLS58216.2023.10166951

57. Anandan B, Manikandan M. Machine learning approach for predicting the peak temperature of dissimilar AA7050-AA2014A friction stir welding butt joint using various regression models. *Mater Lett*. 2022;325:132879.

doi: 10.1016/j.matlet.2022.132879

ACCSCIENCE PUBLISHING

International Journal of AI for Materials and Design

ORIGINAL RESEARCH ARTICLE

Design of speculative artifacts: Integrating generative artificial intelligence, biomaterials, and digital fabrication in co-creative and participatory design

Guilherme Giantini[1]*, Lígia Lopes[1,2], and Jorge Lino Alves[2,3,4]

[1]Department of Design, Faculty of Fine Arts, University of Porto, Porto, Portugal

[2]Course Director of the Master's in Product and Industrial Design, University of Porto, Porto, Portugal

[3]Department of Mechanical Engineering, Faculty of Engineering, University of Porto, Porto, Portugal

[4]Department of Mechanical Engineering, Institute of Science and Innovation in Mechanical and Industrial Engineering, Porto, Portugal

Abstract

This study explores artificial intelligence (AI)-mediated participatory design integrating biomaterials and digital fabrication to co-create speculative artifacts grounded in lived experiences. The present study involves experimentation with biomaterials, exploring the intersection of image-based generative AI, participatory, and co-creative methodologies within a design framework that reimagines lived experiences shaped by identity-based exclusionary processes. Rather than pursuing AI-driven discovery of new materials, this study positions design as a mediating process among human experience, critical reflection, biomaterial exploration, and digital fabrication. The research introduces a three-stage workflow (co-creation, fabrication, and materialization) that employs AI as a mediating tool between subjective narratives and tangible speculative artifacts. During the co-creation stage, participants shared their personal experiences through open-ended surveys, text-to-image generative AI visualization, and algorithmic three-dimensional (3D) modeling. This process enabled participants to speculatively reimagine lived experiences of social exclusion, demonstrating how AI can support new modes of participatory and social engagement. During the fabrication stage, digital models were translated into physical counter-molds through 3D printing and subsequently cast in silicon, reaffirming the reciprocal relationship between digital and craft-based production. The materialization stage explored biomaterial compositions informed by participants' narratives and materialities, incorporating hair, wood ash, and plastic waste into biomaterial compositions grounded in circular economy principles. The resulting artifacts function as speculative objects that incite interpretation beyond fixed symbolic representation. This study contributes to ongoing discussions in digital fabrication, material design, and critical craft by demonstrating how AI-mediated participatory co-creation can foster ethically conscious, socially engaged, and materially grounded design practices. Future work may extend this approach to larger collective settings and further explore the integration of biomaterials and AI within ecological and inclusive design frameworks.

Keywords: Artificial intelligence; Biomaterials; Co-creative design; Participatory design; Three-dimensional printing; Artifact-based design research

***Corresponding author:**
Guilherme Giantini
(up202002299@edu.fba.up.pt)

Citation: Giantini G, Lopes L, Alves JL. Design of speculative artifacts: Integrating generative artificial intelligence, biomaterials, and digital fabrication in co-creative and participatory design. *Int J AI Mater Design*. 2025;2(4):52-67. doi: 10.36922/IJAMD025480048

Received: November 25, 2025

Revised: December 14, 2025

Accepted: December 19, 2025

Published online: December 30, 2025

1. Introduction

The increasing role of artificial intelligence (AI) in data management and community interaction is shaping practice-based applications in material exploration and design, broadening opportunities for developing artifacts at the intersection of craft, digital fabrication, and material agency. While materials science research frequently focuses on AI-powered discovery, optimization, and prediction, design may extend that to esthetic exploration, analysis of experience, and contextual meaning.

AI has increasingly been used to support both the prediction of material behavior and the optimization of processing methods in material development. Yet, these advances often prioritize functional performance while overlooking the exploratory and embodied engagement with materials, which is central to craft and design practices and addresses issues of meaning-making and social context. When undertaken with participatory and co-creative design methods, generative AI can unlock new opportunities for agency, transparency, and meaning in human–machine collaboration.

Emerging research indicates that generative systems enhance craft by mediating gesture-driven fabrication, supporting speculative practice, and enabling participants to actively co-create. Within this setting, generative AI tools help mediate relationships among participants, tools, materials, and artifacts, transforming the digital–craft interface into a speculative space where lived experience, biomaterial engagement, and digital fabrication blend to create new material meanings.

This study uses a three-stage workflow of co-creation, fabrication, and materialization to bridge design gaps involving participatory/co-creative methods, generative AI tools, biomaterials or craft processes, and meaning derived from lived experience. By engaging participants in a co-creative process mediated by generative AI, which explores biomaterials and digital fabrication, this study introduces, as a novelty, the creation of three artifacts that reflect participants' agency over their lived experiences as a unique way of expressing subjective meaning. This study contributes to interdisciplinary debates on digital fabrication, craft, AI, and design by foregrounding participation, material agency, and subjectivity at a time when technological progress often outpaces attention to human experience and social justice. It argues for reconnecting design technologies with lived narratives, critical reflection, and speculative imagination, underscoring the continued value of these dimensions amid rapid technological change.

1.1. State of the art

The intersection of AI, data management, and community interaction has increasingly converged into practice-based applications,[1] such as in material exploration and design practice, thereby opening new possibilities for the conception and production of artifacts at the threshold of craft, digital fabrication, and material agency.

While most materials science literature highlights AI-driven discovery of new materials, task optimization, and faster prediction, many design-focused approaches emphasize different aims (esthetic exploration, experience, and situated meaning) than purely functional optimization.[2] This study contributes to the latter aspect by exploring the mediation of subjective lived experience and speculative materialization through a process combining biomaterial exploration, participatory/co-creative design, and generative AI tools.

In materials science, AI and machine learning are widely used to predict relationships between structure and properties and to optimize processing methods, thereby enhancing the development of new materials. These technological advances indicate significant changes in how material systems are understood and conceived,[3,4] shifting from trial-and-error methods to data-driven, algorithm-supported processes.[5] Although these advances are important, they often focus on functional improvement and tend to overlook the more contextual and exploratory interactions with materials through craft and design-led methods. In research related to design and craft, digital technology has been demonstrated to change how creators interact with materials and tools, where digital tools risk disconnecting the hand from the craft process, thereby challenging traditional human–material–tool dynamics.[6]

At the same time, participatory and co-creative design approaches emphasize involving users, especially marginalized ones in the context of this study, in the design process, granting them agency and voice in shaping outcomes.[7] When these approaches are combined with generative AI tools, new possibilities and tensions emerge: participants may engage with generative systems, interact with AI-driven suggestions or visualizations, and thereby become co-creators and active agents, rather than passive recipients of design outcomes. Research on AI and participatory design highlights the need to attend to power, agency, transparency, and meaning in the human–machine collaboration.[8] Within the fields of craft and digital fabrication, recent work has examined how AI and generative tools can support craft practices rather than replace them, for example, by enabling new forms of gesture-driven digital fabrication or by integrating AI-augmented

workflows into human–machine collaborative processes.[9,10] This highlights AI as a tool of mediation and collaboration for technology-driven processes within craft practice[11] rather than an autonomous creator. This also emphasizes AI as a potential tool for encouraging reflection, critical thinking, and speculative imaginaries, especially when technology is employed as a generative tool for learning or co-creative and participatory processes.[12,13]

While much of the AI–materials literature focuses on high performance, particularly for functional biomaterials used in medicine,[14] craft- and design-led practices often take a more exploratory approach, engaging with material affordances, identity, subjectivity, and the politics of material form. By working with biomaterials, the designer or maker may encourage a more contextually grounded exploration of materials that resist purely techno-functional instrumentalization. In this context, the use of generative AI tools in design aids to mediate the relationships among participants, tools, materials, and artifacts. The digital–craft interface becomes a site of speculative meaning-making, where generative AI supports the exploration of form, structure, and fabrication possibilities. Participants bring their lived experiential knowledge (especially within identity-based exclusionary contexts), and the materialization through digital fabrication and biomaterial engagement anchors the artifact in the tangible world.

Although there is a growing body of work on AI-driven materials discovery and separate research on participatory or co-creative design with digital tools, this study advances relatively few studies that combine: (i) participatory/co-creative design methods, (ii) generative AI tools as mediators of form exploration, (iii) biomaterial or craft-oriented fabrication processes, and (iv) the agency of participants bringing identity-based lived experience. This research gap is addressed in this study by engaging participants in a co-creative process mediated by generative AI, which explores biomaterials, digital fabrication, and craft-inspired engagement to materialize artifacts that reflect their subjective ownership and lived experience. Rather than generating novel material for performance-driven design, the participatory and co-creative outcomes explore esthetic experimentation with biomaterials as a unique way of expressing subjective meaning. In particular, this paper emphasizes how generative AI can enable non-designer participants to engage materially and experientially with design processes and fabrication, rather than leaving it exclusively to design or engineering experts.

1.2. Project-based academic context

This approach was undertaken as part of a practical design project within a master's thesis in the Product and Industrial Design course at the University of Porto, Portugal, between 2023 and 2024. This Master's program fosters a project-based learning environment,[15-20] in which students are encouraged to work on real design problems, prototype their ideas, and critically analyze their creative processes. The program is primarily grounded in multidisciplinary collaboration across academic institutions, jointly managed by the Faculty of Fine Arts and the Faculty of Engineering at the University of Porto. This convergence between design and engineering education[19] fosters a strong connection with industry, leading to partnerships between the Master's course program and local industrial firms, enabling students to work on real briefs, constraints, and opportunities. The program also features a research-oriented approach, including a dedicated teaching unit on research methodology and the option for a scientific research dissertation or project.

2. Methodology

Within this background and project-based academic context, the practical development of this project, described in the following sections, adopted a methodological approach comprising human–AI collaboration[21] that undertakes participation and co-creation methods to actively engage participants in the design process[22-24] and grant them interactive agency over design decisions,[25-27] respectively. Contributing to the expansion of the intersection between participatory and co-creative design approaches,[28] this methodology was undertaken with three individuals whose subject experiences were consistently shaped by identity-based exclusionary processes (i.e., racialization, LGBT-phobia and transphobia, misogyny, xenophobia, and ableism, among others). For ethical reasons, the identities of the participants, who willingly agreed to respond to the survey, were anonymized. The objective of this study is to enable participants to express creative and subjective agency of their lived experiences by reimagining them through the materialization of a design artifact conceptualized with generative AI tools and produced through digital fabrication techniques.

During the co-creation phase, participants had full decision-making agency regarding the narrative elements included in their narratives, as well as the selection of images derived from them. In the participatory phases (including material base definition, material experiments, and artifact production), they actively engaged in design decisions that ultimately influenced the final outcome. This blend of co-creative and participatory phases enabled the focus on the participants' creative and decision-making agency while maintaining the realization of technical aspects of the design process (such as AI prompts and image generation, merging of images, conversion of two-dimensional files to 3D, and the 3D printing of both

counter-molds and final molds) for the designer. This design process was undertaken in three stages: co-creation, digital fabrication, and materialization (Figure 1).

2.1. Co-creation

The objective of this stage was to design a 3D digital object based on each participant's interpretation of a lived experience relating to social exclusion. To this end, a co-creation methodology was developed (Figure 2), comprising three subsequent processes.

2.1.1. Memory to text: Data collection and surveys

The participants were asked to complete an open-ended survey centered on their social context and a significant object from their lived experience of discrimination, which they were invited to share. The aim was to collect words, expressions, and nuances of meaning that could help qualify the participants' experiential accounts in response to three questions. Question one referred to the object that was most relevant to the shared memory, requiring participants to disclose its main features or qualities and explain how that object could be used to address an experience of discrimination. Question two explored the context in which the discrimination experience took place, asking participants to describe the physical environment, the inner feelings caused by the aggressor, and their feelings toward the aggressor. The final question focused on the materiality of the experience, asking participants to identify the type of waste that was most relevant in such a context. The survey was also structured in a way that the output responses' syntactic morphology would relate to the typical syntax of a text-to-image generative AI art prompt.[29] Following the conventional order of subject, verb, and predicate, this prompt structure loosely framed the data collected from each participant's narrative. The data provided by participants are summarized in Table 1.

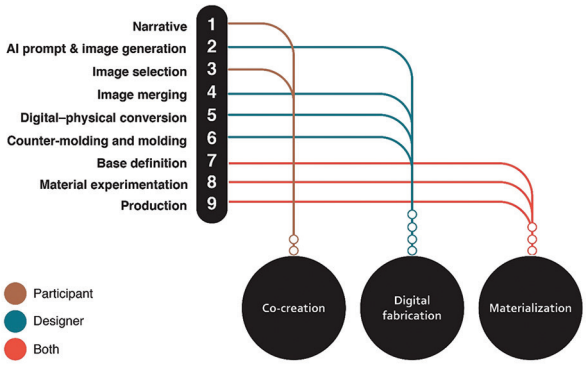


Figure 1. Map of the design process comprising methodological stages of co-creation, digital fabrication, and materialization
Abbreviation: AI: Artificial intelligence.

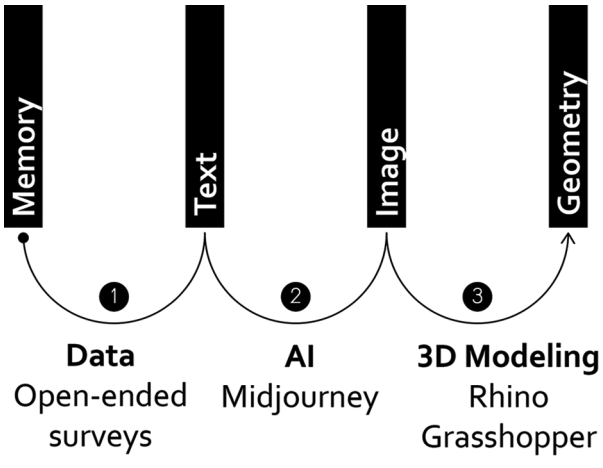


Figure 2. Methodology of the co-creation process
Abbreviations: 3D: Three-dimensional; AI: Artificial intelligence.

2.1.2. Text to image: Meaning beyond representation

After collecting and organizing data on each participant's experience, a series of prompts was created to visually translate these experiential narratives. This was undertaken in the Midjourney text-to-image generative AI model (version 5), which also allows for specifying a representational style as input data. In this case, however, such stylistic specification was intentionally excluded to avoid visually biasing the graphical results. In addition, Midjourney also allows for the insertion of descriptive prompt modifiers, such as "4K" or "16K," to simulate high fidelity or fine detail, as well as the customization of parameters, such as "--v" for version control, "--ar" for aspect ratio, and "—s," where "s" stands for lower and higher values to control how faithful to the prompt the AI tool will process the information. For each participant's narrative, the prompt was iteratively modified. The following example shows one of these prompt iterations: "A black transgender sentient, delicate being feeling rage, anger, shame, repulsion, and abjection when destroying a burning wooden pulpit, 8K --v 5 --ar 9:16 --s 100."

The objective of using AI in this creative exploration was to leverage the generative potential of AI techniques and tools, understood here in the Heideggerian sense of technology.[30,31] This concept originates from the field of philosophy, critiquing the common and current use of technology as framed by a modern mindset of resource optimization and control. When not enframed, technology may act as a form of poetic bringing-forth or an art-like revelation that cooperates with the world by uncovering processes and meanings beyond utility.

(a) Image creation and selection

A comprehensive set of images was produced for each experience, after which participants were requested

to select the images that most accurately represented their own narrative, as illustrated by Figures 3-5. The number of images was a variable parameter, and its resulting numeric value depended on each participant's personal decision.

(b) Image merging

The image merging stage was designed to, on one hand, integrate each set into a single composite abstract image, and on the other hand, to minimize any representational characteristics that might suggest a fixed meaning, thereby moving away from symbolic signification. To achieve this, the Photomerge technique in Adobe Photoshop version 14 was applied to integrate overlapping regions across the images. Figure 6 displays the process of merging the images for each participant, where similar features among images within each set are identified and seamlessly joined. This process resulted in three composite images, as shown in Figure 7.

2.1.3. Image to geometry: Digital conversion

Once the images from each set were merged, the next step involved assigning them a third dimension. This two-dimensional-to-3D conversion was performed in Rhinoceros version 7 using algorithmic modeling with the Grasshopper 3D plug-in (version 1.0.007), employing a simple algorithmic definition, that is, a visual script made from connecting parameters and components (patches of code) to generate and manipulate geometry parametrically without traditional code scripting (Figure 8).First, the algorithmic definition was used to define an 800 × 450 mm^2 mesh, of which normal vectors were extracted for each patch. The normal vector's amplitude of each patch was equated to the merged images from the participants using the "Image Sampler" component. This component is capable of identifying color or brightness values from any given image and applying them to control design parameters. Given the mesh size and subdivision, the vector's amplitude was mapped to the corresponding brightness-scale control values of each image to define the final mesh topology across different displacement heights. With displacement intervals of 50 mm within a domain of 50 to 350 mm, and a maximum height of 30 mm, brightness values of 0 produced minimal Z-displacement (50 mm),

Table 1. Summary of participants' data collected in surveys

Parameters	Predicate				
	Verb	**Object**	**Object's features**	**Social context**	**Emotional response**
Participant A	Release/loosen	Hair	Curly and coiled	Racism on a school bus route	Confusion, rage, hatred, revolt
Participant B	Burn	Pulpit	Tall, aggressive, dark, and slender	Transphobia at a Pentecostal Evangelical Church	Shame, humiliation, dishonor, smallness, anger, rage, disgust, repulsion
Participant C	Crush and discard	Plastic bag	Rigid	Xenophobia in a supermarket	Inferiority, rejection, exclusion, injustice, displeasure, sadness

Figure 3. Participant A. Each panel of the figure shows a generated iteration output that was selected by the participant as the most representative of their experience.

Figure 4. Participant B. Each panel of the figure shows a generated iteration output that was selected by the participant as the most representative of their experience.

Figure 5. Participant C. Each panel of the figure shows a generated iteration output that was selected by the participant as the most representative of their experience.

whereas values of 100 produced maximum displacement (350 mm). This process generated a mesh morphology whose geometric topography directly corresponded to each input image (Figure 9). From this process, three digital artifacts were obtained, each corresponding to the narrative of a participant, as shown in Figure 10.

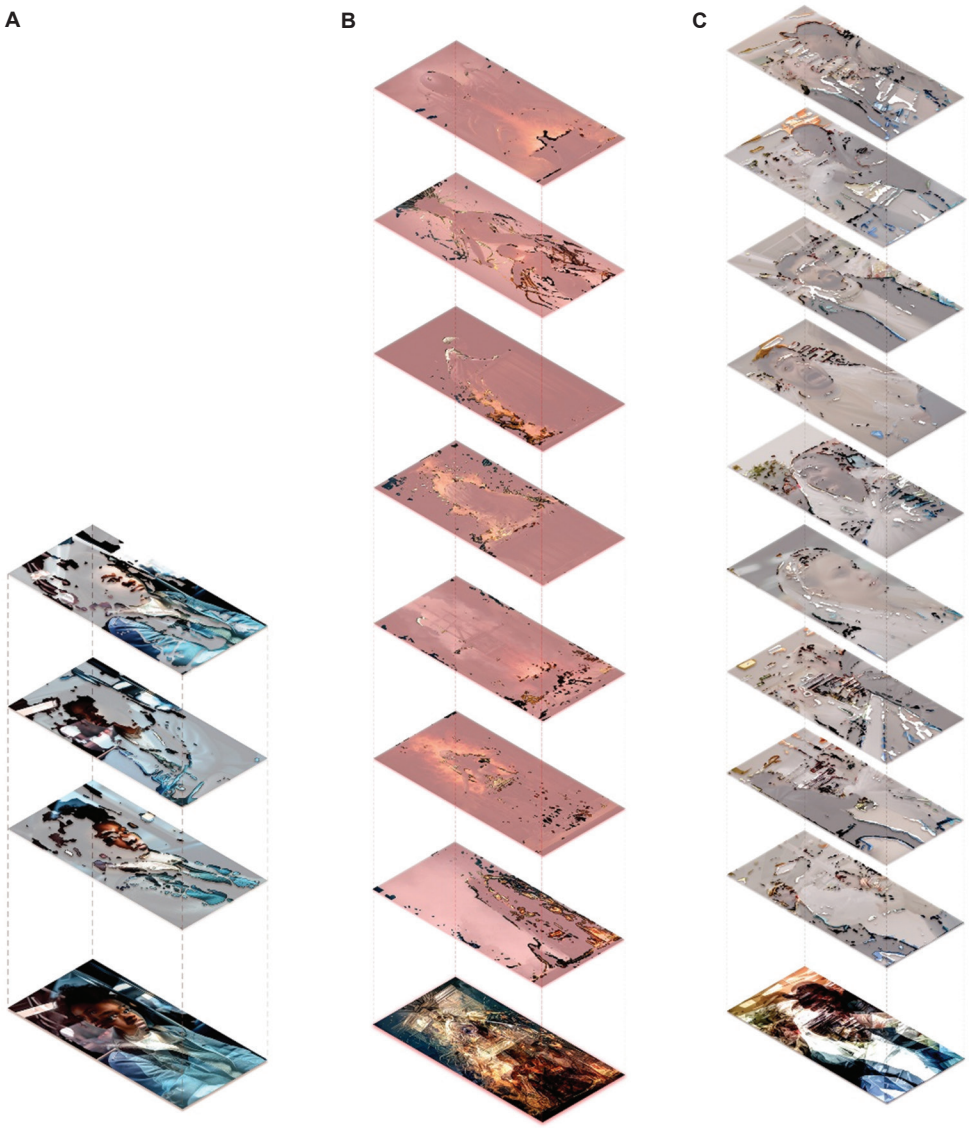


Figure 6. Merging images for each participant's selected images, accounting for overlapping regions. Images of (A) participant A, (B) participant B, and (C) participant C.

2.2. Digital fabrication

The production workflow comprised two main stages. In the fabrication stage, counter-molds were produced through 3D printing, from which silicon molds were created to enable an easier extraction of the final pieces. In the materialization stage, materials were selected, binding compositions were tested and refined, and the three narratives were ultimately materialized as artifacts. The production strategy is illustrated in Figure 11.

2.2.1. Mold fabrication

Once the 3D files of the artifacts were acquired, the next step involved creating molds to facilitate materialization.

The counter-molds were produced considering the maximum build volume of the Prusa i3 MK3S+ printer (Czech Republic) (250 × 210 × 210 mm) at the Product and Service Development Laboratory, Faculty of Engineering, University of Porto. Each digital artifact was segmented into 12 parts (200 × 150 × 30 mm each) (Figure 12) and converted into.STL files for 3D printing using the Prusa Slicer software (version 2.8.1) (Figure 13). The 3D printing process was conducted in a polylactic acid thermoplastic polymer widely used in 3D printing.[32] A 1.75 mm red filament (Tucab, Portugal) was used; the filament color was selected for practical reasons and did not influence the results. The other parameters were a layer height of 0.15 mm, an extrusion temperature of

Figure 7. Result of image merging. Images of (A) participant A, (B) participant B, and (C) participant C.

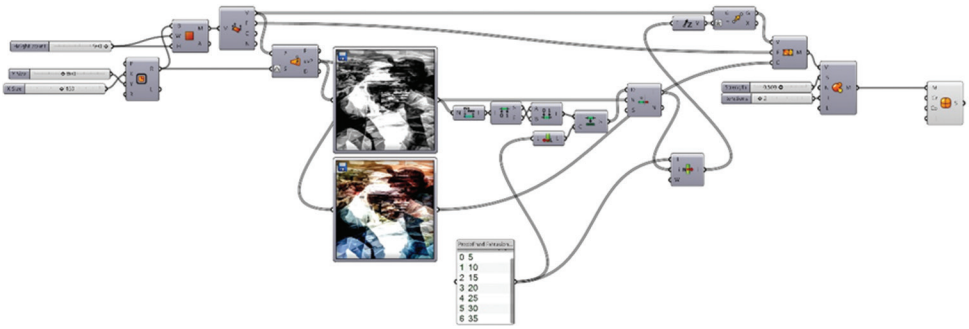


Figure 8. Algorithmic definition for three-dimensional image conversion

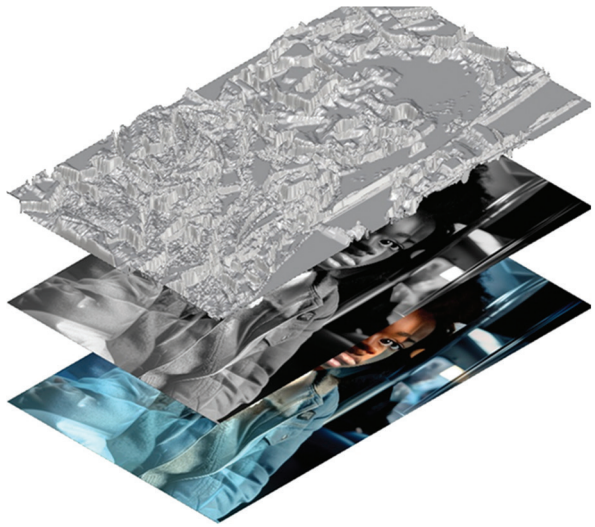


Figure 9. Outcome of the Image Sampler component applied to mesh deformation

approximately 210°C, a print speed of 60 mm/s, and an infill of 20%. These parameters were selected according to a performance criterion that balanced low weight, adequate resolution, and reduced fabrication time. Figure 14 shows the 3D printing of each of the 12 parts comprising one artifact.

For the production of the silicon molds, each set of 12 counter-mold parts was arranged on a wooden base measuring 1,030 × 680 × 25 mm, fitted within a frame of 1,080 × 730 × 45 mm to contain the silicon. Following the sequence defined in the design, each module was bonded to the base and to adjacent parts using a drop of cyanoacrylate, a fast-acting adhesive. To cover the counter-molds, an adequate amount of XIAMETER® RTV-4234-T4 (Dow Silicones Corporation, USA) industrial silicon was mechanically mixed and degassed in a vacuum chamber (Protoclick, Portugal)[33] with a catalyst before being poured over the assembly (Figure 15).[34]

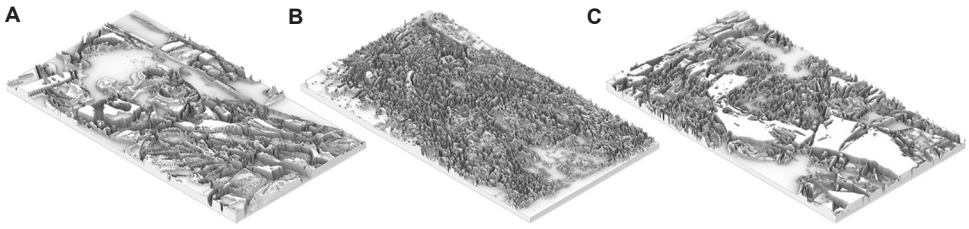


Figure 10. Outcome of three-dimensional mesh deformation. Outcomes of (A) participant A, (B) participant B, and (C) participant C

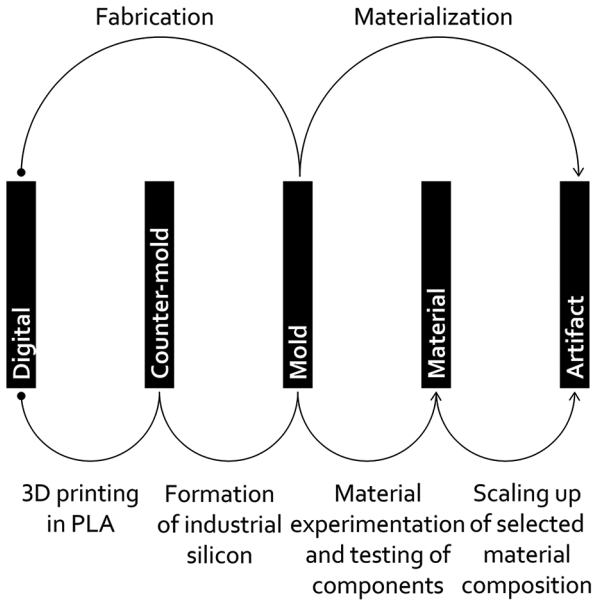


Figure 11. Methodology of the fabrication process
Abbreviations: 3D: Three-dimensional; PLA: Polylactic acid.

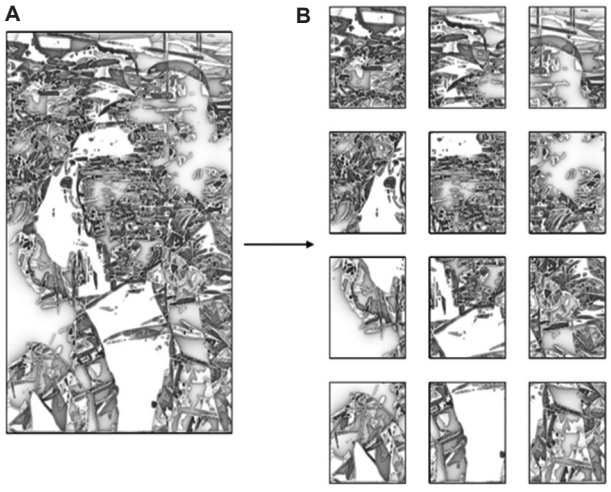


Figure 12. Segmentation of the digital model. (A) The full digital model before segmentation. (B) The same digital model segmented into 12 parts.

Figure 16 illustrates the silicon mold after it has dried and been removed from the wooden base. Each mold had an approximate weight of 13 kg.

2.3. Materialization

The materialization of the artifacts involved the preliminary development of biomaterials whose primary constituents were directly related to the materials of each object identified in the participant interviews. Following the co-creative agency of the methodology undertaken, the criteria for selecting these materials were left to the participants, guided by the final survey question about materiality. They were hair strands, wood ash, and plastic bags, as summarized in Table 1. Drawing on principles from emerging material disciplines, such as biodesign, biomaterials, and design for a circular economy, preliminary studies were conducted to develop biomaterial pastes with varying compositions, building on prior work. Following the material design guidelines explored in the Master's program on Industrial and Product Design at the University of Porto, which focused on waste materials for the circular economy,[18] the materialization process of the artifacts discussed in this paper was inspired by a preceding biomaterial experimentation. The methodological development of such experimentation assessed a wide range of base materials (such as used coffee grounds, vegetable peels, and eggshells), leading to the development of a circular tote bag made from garlic husks.[35] Characterization and evaluation of the biomaterial pastes enabled the development of a versatile composition capable of incorporating the three specified materials while meeting criteria such as dimensional stability, resistance to mildew, and resistance to mechanical impact. Each composition was poured into the corresponding mold, thereby assuming the 3D form equivalent to the previously conceived virtual image.

2.3.1. Biomaterial experiments

To develop a single composition capable of binding all the materials used, experimentation began with wood ashes, which were readily available during the development of the experiences. Drawing on compositions sourced from platforms such as Materiom and the Rhode Island School of Design Nature Lab, and scientific dissemination platforms including ResearchGate and MDPI, Composition 1 (Table 2) was prepared using plastic molds measuring 230 × 140 × 30 mm, greased with petroleum jelly to facilitate demolding.

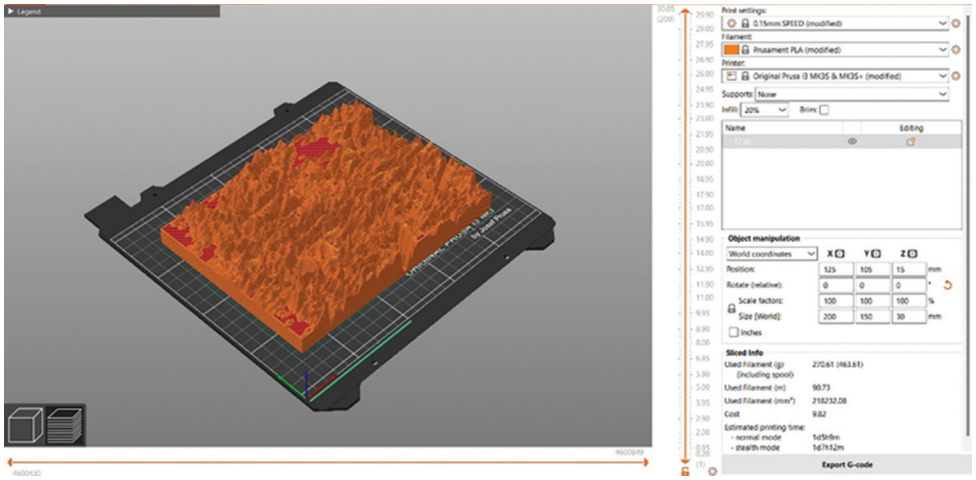


Figure 13. Preparation of the three-dimensional printing process of each counter-mold using the Prusa Slicer software

Table 2. Substances used in making the compositions

Composition	Substance	Amount
Composition 1	Water	300 mL
	Sodium alginate	45 g
	Wood ashes	60 g
	Glycerin	6 mL
	Gelatin	15 g
Composition 2	Water	150 mL
	Sodium bicarbonate	4 g
	Sugar	8 g
	Wood ashes	60 g
	Gelatin	60 g
Composition 3	Water	100 mL
	Glycerin	12 mL
	Sodium bicarbonate	4 g
	Wood ashes	120 g
	Gelatin	30 g

The liquid substances were combined in a container, while the solid substances were aggregated in a separate one. Both were then combined, without fusing, using heat, and poured into a single container to be air-dried. Figure 17 displays the result of the experiment using Composition 1, where the material's fragility and prolonged drying time rendered it unsuitable for the project.

Composition 2 was also inspired by the sources cited above and developed accordingly (Table 2). The preparation process followed the same procedure as before, in which liquid and solid substances were combined separately before mixing. However, in this and the subsequent compositions, heat fusion was applied using a FLAMA (Portugal) 8170FL gas stove for 270 s at an approximate power of 6,800 W. Figure 18 presents the result of the experiment using Composition 2. The material was considered unsuitable for large-scale applications due to its dimensional and shape instability.

The third test composition differs from the previous one by omitting sugar and by adjusting the quantities of the other binding agents (Table 2). Figure 19 shows the result of Composition 3. The resulting material demonstrated relative dimensional stability, mechanical and mold resistance, reduced drying time, and ease of demolding, making it suitable for replication with the other substances identified in Table 1, "object" column. It is also relevant to highlight the role of each component: gelatin as a thickening agent, cornstarch as a material commonly used in the production of gums and bioplastics, glycerin as a plasticizer, and sodium bicarbonate as a fungicidal agent. Accordingly, Composition 3 was used in the following experiments, considering substance replacement. Therefore, the following experiment replaced wood ashes with an equivalent amount of shredded plastic bags, while the subsequent experiment utilized hair as the primary substance. Respectively, Figures 20 and 21 present the results of an experiment that exhibited consistent characteristics when compared to the preceding experiment, undertaken with wood ashes.

This multivalent composition, adaptable across different materials, was therefore deemed suitable for large-scale application in the materialization of the artifacts. Table 3 offers a summary of the experiments conducted until arriving at the multivalent formula.

2.3.2. Biomaterial to artifact

Scaling the composition was necessary to achieve an adequate volume for producing the final artifacts. This

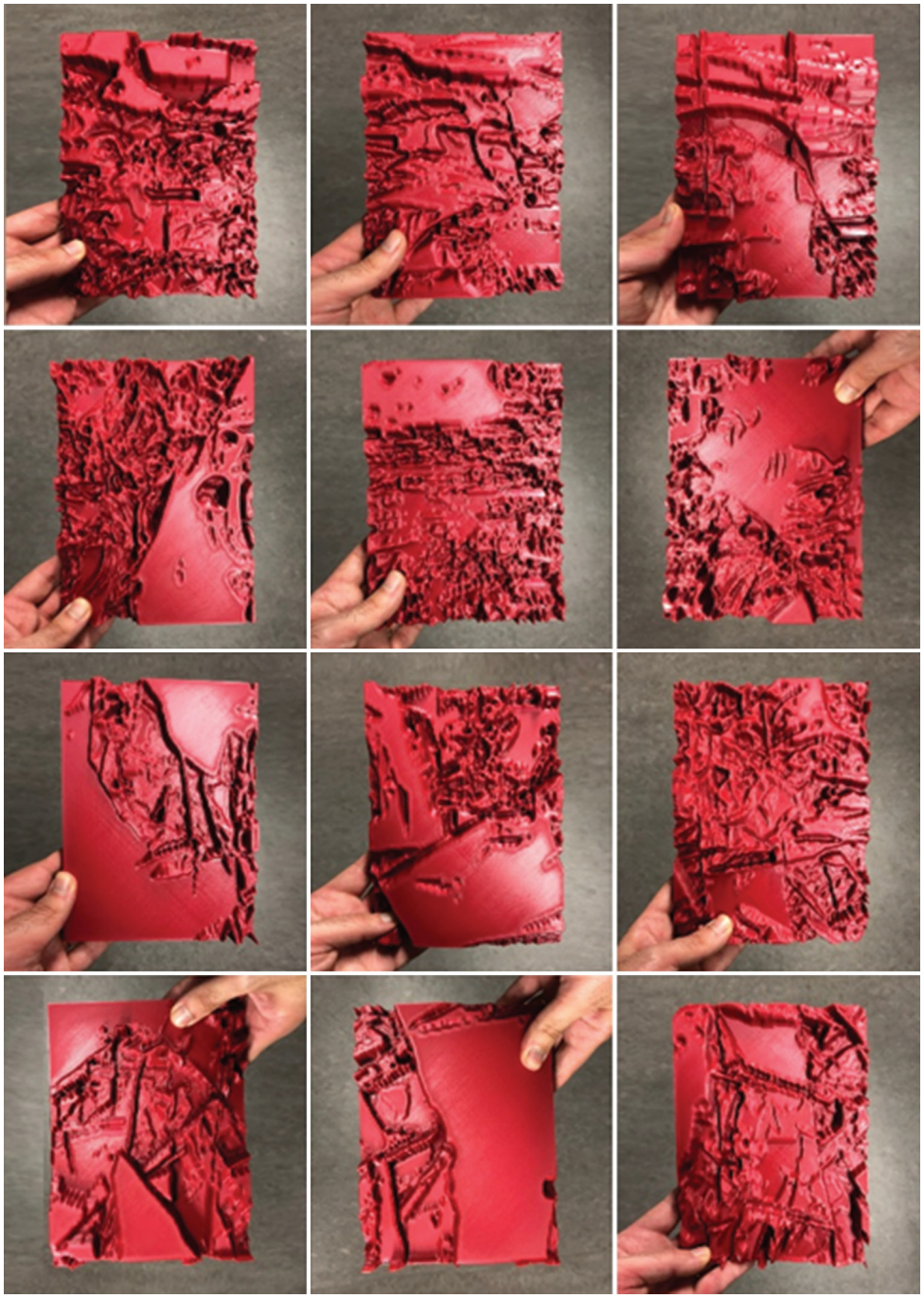


Figure 14. Set of 12 counter-mold parts produced through three-dimensional printing being poured over the assembly (Figure 15).[34]

was carried out in direct proportion to the volume of the artifacts, with quantities calculated in the Rhinoceros 3D modeling software (version 7). Using the amount of water as the reference parameter for scale-up, the volume of each artifact was set at approximately 2,000 mL, corresponding to a 20-fold increase in compositional scale. Table 4 shows the formulation.

Following this scaling procedure, the three artifacts were produced by casting the pastes into the molds, allowing them to cure for 12 h before demolding.

3. Results and discussion

Figure 22 presents the results of the materialization of narratives from participants A, B, and C into design artifacts.

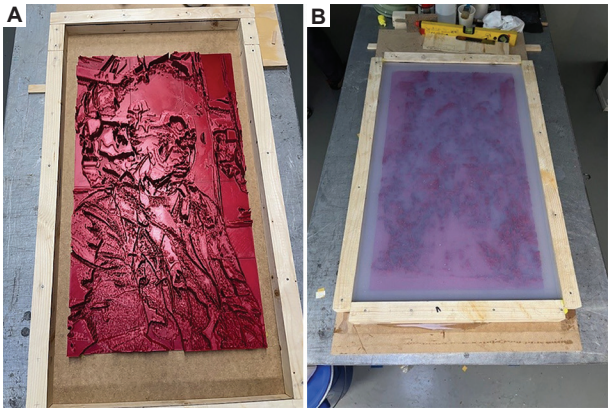


Figure 15. Silicon mold production. (A) The positioning of artifact components on a wooden base. (B) The industrial silicon pouring process

Figure 16. Prepared silicon molds ready for use

Figure 17. Experiment using composition 1

The artifacts developed in the presented project aim to evoke viewers' own modes of perception of the artworks co-created with the participants through the semiotic potential of the intersection of material, form, and image in a de-representative esthetics. In doing so,

Figure 18. Experiment using composition 2

Figure 19. Experiment using composition 3

Figure 20. Experiment using shredded plastic bags in composition 3

Figure 21. Experiment using hair in composition 3

they avoid symbolic signification or the fixed meanings of signs imposed by established representational systems. The artifacts thus propose, at least from a philosophical standpoint, a poetic provocation to symbolic signification in the loss of fixed meaning within the subject–object domain. This is achieved by evoking a form of affect that is unconditioned by the established sets of symbols familiar to our common form of world perception.[36,37] Within this process, combined with biomaterial exploration and participatory co-creative methodologies, AI is not intended to be employed as an instrumental tool, but rather as a mode of revealing something, in line with a

Figure 22. Biomaterial design artifacts. Outcomes of (A) participant A, (B) participant B, and (C) participant C

Table 3. Summary of biomaterial experiences until arriving at the multivalent formula

Substances	Composition		
	1	2	3
Base substance			
Wood ashes, shredded plastic bags, or hair			
Base solvent			
Water (mL)	300	150	100
Thickener agents			
Sodium alginate (g)	45	-	-
Gelatin (g)	15	60	30
Cornstarch (g)	-	30	15
Sugar (g)	-	8	-
Plasticizer			
Glycerin (mL)	6	-	12
Antifungal agent			
Sodium bicarbonate (g)	-	4	4

Table 4. Scaled composition of composition 3

Substance	Amount
Water	2,000 mL
Base substance	2,400 g
Glycerin	240 mL
Sodium bicarbonate	80 g
Gelatin	600 g
Cornstarch	300 g

Heideggerian sense of technology. Accordingly, the act of poetic creation at the intersection of critical thinking, social concerns, and technology becomes a space for exploring perception, stretching beyond the limits of representation. In such an endeavor, apprehension may potentially move beyond signification toward another mode of meaning-making grounded in experience, wherein the intensity of what is felt by the participant provokes a direct affect in the viewer.

4. Conclusion

This study explored the intersection of generative AI, biomaterial experimentation, and participatory and co-creative methodologies in a practical design project, providing a tangible way to reimagine lived experiences shaped by identity-based exclusionary processes. Instead of focusing on the AI-driven discovery and design of new technological materials, this study highlighted design as a mediating process among human experience, critical thinking, biomaterial exploration, craft practice, and digital fabrication processes. Through the integration of co-creation, fabrication, and materialization, the study used AI as a critical collaborator/mediator between subjective narratives and tangible speculative artifacts.

During the co-creation stage, participants expressed their critical perspectives on personal experiences as visually reimagined scenarios. This was achieved through a process that included data collection through open-ended surveys, visualization with text-to-image generative AI, and the conversion of these into digital objects using algorithmic 3D modeling. This process granted participants speculative agency in critically shaping alternative imaginaries of their lived experiences, showing how generative AI can potentially enable new forms of social engagement when aligned with participatory and co-creative design methodologies. The digital fabrication

stage converted digital objects into physical counter-molds using 3D printing. Subsequently, silicon casting from these counter-molds was employed to facilitate an easier removal of the final molds, reinforcing the reciprocal relationship between digital and artisanal creation in contemporary craft and design. Finally, the materialization stage established a direct link between biomaterials and participants' critical narratives by incorporating symbolically charged substances (hair, wood ash, and plastic bags) into exploratory biomaterial compositions. Grounded in principles of biodesign and circular economy, these compositions demonstrated the expressive and affective potential of biomaterials as carriers of identity, memory, and speculative critique.

The resulting artifacts encourage the viewer to engage with meaning by inciting a singular interpretation rather than relying on fixed symbolic representation, using AI not as a mere tool but as a mode of revealing, which aligns with Heidegger's critique of technology. By integrating generative AI into a co-creative and participatory design process exploring biomaterials, the project demonstrates how technological mediation can be used, both critically and artistically, to support socially engaged, embodied, and ethically conscious practices of creation.

This study thus contributes to the ongoing interdisciplinary discussion on digital fabrication, craft making, AI, and design practices, highlighting the importance of participation, material agency, and subjectivity at a time when technological progress continues to be prioritized over human experience and social justice, rather than ethically developed alongside them. This study emphasizes that, despite the era of prominent technological advancements, there remains significant value in reconnecting design technologies with lived narratives, meaning-making, critical thinking, and speculative imagination. Future work may involve enhancing the methodological approach by further increasing participants' engagement and extending this approach to wider collective settings, integrating additional forms of biological or waste materials, and further investigating how generative AI can enhance co-creative authorship and biomaterial exploration within ecologically oriented and socially engaged design contexts.

Acknowledgments

The authors express their gratitude to the Associate Laboratory of Energy, Transport, and Aerospace (LAETA) and the Institute of Science and Innovation in Mechanical and Industrial Engineering (INEGI), whose resources were essential for the development of this work.

Funding

None.

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: Guilherme Giantini
Data curation: Guilherme Giantini
Investigation: All authors
Methodology: All authors
Project administration: All authors
Resources: Lígia Lopes, Jorge Lino Alves
Supervision: Lígia Lopes, Jorge Lino Alves
Validation: Lígia Lopes
Visualization: Guilherme Giantini
Writing–original draft: All authors
Writing–review & editing: All authors

Ethics approval and consent to participate

No specific ethical approval was required for this study, as the master's committee responsible for reviewing the original research determined that the survey's statement would sufficiently cover participants' consent. Permission was obtained from participants in written format. By participating in the anonymized survey, participants agreed to provide and grant permission for the use of their shared data. These data are limited to subjective information (i.e., subjective perceptions, opinions, and claims).

Consent for publication

By participating in the anonymized survey, the participants granted their permission for the use of their shared subjective data, which includes publication for scientific and artistic purposes.

Availability of data

The original data presented in this study are openly accessible in *Repositório Aberto da Universidade do Porto* at https://hdl.handle.net/10216/164287.

Further disclosure

This is the first English-language version of this work. The research underpinning this paper was originally presented in a Portuguese-language master's level thesis (December 2024), which is available at https://hdl.handle.net/10216/164287.

References

1. Ng WL, Goh GL, Goh GD, Ten JSJ, Yeong WY. Progress and opportunities for machine learning in materials

and processes of additive manufacturing. *Adv Mater*. 2024;36(34):2310006.

doi: 10.1002/adma.202310006

2. Friedrich J, Brückner A, Mayan J, Schumann S, Kirschenbaum A, Zinke-Wehlmann C. Human-centered AI development in practice-insights from a multidisciplinary approach. *Z Arb Wiss*. 2024;78(3):359-376.

doi: 10.1007/s41449-024-00434-5

3. Guo K, Yang Z, Yu CH, Buehler MJ. Artificial intelligence and machine learning in design of mechanical materials. *Mater Horiz*. 2021;8(4):1153-1172.

doi: 10.1039/D0MH01451F

4. Bai X, Zhang X. Artificial intelligence-powered materials science. *NanoMicro Lett*. 2025;17(1):135.

doi: 10.1007/s40820-024-01634-8

5. Hang R, Yao X, Bai L, Hang R. Evolving biomaterials design from trial and error to intelligent innovation. *Acta Biomater*. 2025;197:29-47.

doi: 10.1016/j.actbio.2025.03.013

6. Bryan-Kinns N, Liu Y, Zhang D, Tan Z, Tan H. The role of digital technologies in contemporary craft practice from UK-China insights. *J Contemp Chin Art*. 2024;11(1):75-93.

doi: 10.1386/jcca_00093_1

7. Sanoff H. Participatory design. *J Des Plan Aesthet Res*. 2022;1(2):1-12.

doi: 10.55755/DepArch.2022.8

8. Gautam A. Reconfiguring participatory design to resist AI realism. *Assoc Comp Mach*. 2024;2:31-36.

doi: 10.1145/3661455.3669867

9. Buš P, Dong Z. DeepCraft: Imitation learning method in a cointelligent design to production process to deliver architectural scenarios. *Archit Intell*. 2024;3(1):12.

doi: 10.1007/s44223-024-00055-2

10. Buš P. DeepCraft: Co-intelligent architecture and human and ai-driven craftsmanship in design-to-production pipelines. In: Yan C, Chai H, Sun T, Yuan PF, editors. *Phygital Intelligence*. Singapore: Springer Nature; 2024. p. 368-378.

doi: 10.1007/978-981-99-8405-3_31

11. Song MJ. Craftspeople's new identity: The impact of digital fabrication technologies on craft practices. *Int J Technol Des Educ*. 2022;32(4):2365-2383.

doi: 10.1007/s10798-021-09687-1

12. Vartiainen H, Tedre M. Using artificial intelligence in craft education: Crafting with text-to-image generative models. *Digit Creat*. 2023;34(1):1-21.

doi: 10.1080/14626268.2023.2174557

13. Messer U. Co-creating art with generative artificial intelligence: Implications for artworks and artists. *Comput Hum Behav Artif Hum*. 2024;2(1):100056.

doi: 10.1016/j.chbah.2024.100056

14. Gokcekuyu Y, Ekinci F, Guzel MS, Acici K, Aydin S, Asuroglu T. Artificial intelligence in biomaterials: A comprehensive review. *Appl Sci*. 2024;14(15):6590.

doi: 10.3390/app14156590

15. Chandrasekaran S, Stojcevski A, Littlefair G, Joordens M. Project-oriented design-based learning: Aligning students' views with industry needs. *Int J Eng Educ*. 2013;29(5):1109-1118.

16. Gomes Â, Rangel B, Carneiro V, Lino J. "Learning by Doing" integrated project design in a master program on product and industrial design. In: Nascimento MM, Alves GR, Morais EVA, editor. *Contributions to Higher Engineering Education*. Singapore: Springer; 2018. p. 105-134.

doi: 10.1007/978-981-10-8917-6_5

17. Schoch K, Bickel M, Liedtke C, Hemmert F. Circular economy from scratch: A novel project-based learning method to increase motivation in metal recycling among industrial design students. *MethodsX*. 2025;14:103137.

doi: 10.1016/j.mex.2024.103137

18. Den Hollander MC, Bakker CA, Hultink EJ. Product design in a circular economy: Development of a typology of key concepts and terms. *J Ind Ecol*. 2017;21(3):517-525.

doi: 10.1111/jiec.12610

19. Alves JL, Alexandre R, Lopes L. Competencias convergentes de los estudiantes: Experiencia educativa con futuros diseñadores industriales e ingenieros mecánicos [Students' converging skills: Educational experience with future industrial designers and mechanical engineers]. *Cuad Cent Estud En Diseño Comun Ens*. 2022;(166):231-238.

doi: 10.18682/cdc.vi166.7055

20. Alves JL, Facca CA, Fernandes AP, Rangel B, Barbosa AM. Interdisciplinariedad entre diseño e ingeniería: Nuevas competencias en la docencia de proyectos para la innovación circular [Interdisciplinarity between design and engineering: New skills in teaching projects for circular innovation]. *Cuad Cent Estud En Diseño Comun Ens*. 2020;(114):179-203.

doi: 10.18682/cdc.vi114.4121

21. Guo X, Xiao Y, Wang J, Ji T. Rethinking designer agency: A case study of co-creation between designers and AI. *Int Assoc Soc Des Res*. 2023;170:1-18.

doi: 10.21606/iasdr.2023.478

22. Simonsen J, Robertson T. *Routledge International Handbook of Participatory Design*. 1st ed. Routledge; 2013. Available from: https://www.routledge.com/Routledge-International-Handbook-of-Participatory-Design/Simonsen-Robertson/p/book/9780415720212 [Last accessed on 2025 Aug 22].

23. Bossen C, Dindler C, Iversen OS. Evaluation in participatory design: A literature survey. *Assoc Comp Mach*. 2016;1:151-160.

doi: 10.1145/2940299.2940303

24. Smith RC, Bossen C, Kanstrup AM. Participatory design in an era of participation. *Int J CoCreat Des Arts*. 2017;13(2):65-69.

doi: 10.1080/15710882.2017.1310466

25. Sanders EBN, Stappers PJ. Co-creation and the new landscapes of design. *Int J CoCreat Des Arts*. 2008;4(1):5-18.

doi: 10.1080/15710880701875068

26. Roberts DL, Darler W. Consumer co-creation: An opportunity to humanise the new product development process. *Int J Mark Res*. 2017;59(1):13-33.

doi: 10.2501/IJMR-2017-003

27. Ramaswamy V, Ozcan K. What is co-creation? An interactional creation framework and its implications for value creation. *J Bus Res*. 2018;84:196-205.

doi: 10.1016/j.jbusres.2017.11.027

28. Hyysalo S, Dorta T. Design participation and codesign: New forms, application domains, and representational systems. *Designing*. 2025;1(1):24-31.

doi: 10.1177/30497671251392167

29. Oppenlaender J. *The Creativity of Text-to-Image Generation*. New York: Association for Computing Machinery; 2022. p. 192-202.

doi: 10.1145/3569219.3569352

30. Heidegger M. *The Question Concerning Technology and Other Essays*. New York: Garland Publishing; 1977. Available from: https://philpapers.org/rec/HEITQC [Last accessed on 2025 Dec 08].

31. Thomson, ID. Heidegger on technology's danger and promise in the age of AI. In: *Elements in the Philosophy of Martin Heidegger*. United Kingdom: Cambridge University Press; 2025.

doi: 10.1017/9781009629423

32. Afonso JA, Alves JL, Caldas G, Gouveia BP, Santana L, Belinha J. Influence of 3D printing process parameters on the mechanical properties and mass of PLA parts and predictive models. *Rapid Prototyp J*. 2021;27(3):487-495.

doi: 10.1108/RPJ-03-2020-0043

33. Alves JL, São Simão M, Neto RJL, Duarte TMGP. *PROTOCLICK - Prototipagem Rápida*. Protoclick; 2001. Available from: https://www.booki.pt/loja/prod/protoclick-prototipagem-rapida/9729537615 [Last accessed on 2025 Nov 17].

34. Liravi F, Toyserkani E. Additive manufacturing of silicone structures: A review and prospective. *Addit Manuf*. 2018;24:232-242.

doi: 10.1016/j.addma.2018.10.002

35. Giantini G, Lopes L, Alves JL. Biotextiles de residuos de la industria alimenticia: Una experiencia de Biodiseño para tote bags sostenibles [Biotextiles from the waste of the food industry: A biodesign experience for sustainable tote bags]. *Base Diseño E Innov*. 2022;7(7):71-88.

doi: 10.52611/bdi.num7.2022.792

36. Heidegger M. *Off the Beaten Track*; 2002. Available from: https://www.cambridge.org/universitypress/subjects/philosophy/philosophy-texts/heidegger-beaten-track [Last accessed on 2025 Nov 17].

37. Merleau-Ponty M. *Phénoménologie de La Perception [Phenomenology of Perception]*. 3rd ed., Vol. 10. Gallimard; 1945. Available from: https://www.gallimard.fr/catalogue/phenomenologie-de-la-perception/9782070293377 [Last accessed on 2025 Nov 24].

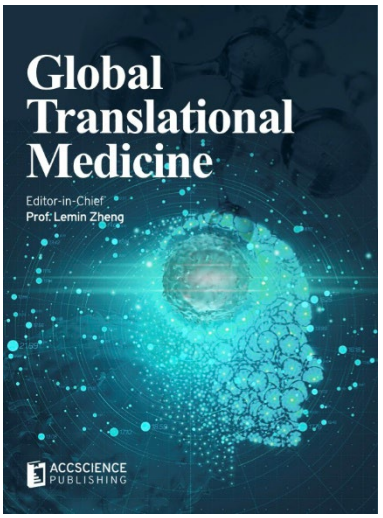
Our Journals

Tumor Discovery is a peer-reviewed and open-access journal that aims to present new cancer research with strong emphasis on fundamental and translational studies. *Tumor Discovery* covers topics, including but not limited to the following:

- Etiology and pathogenesis of cancer
- Mechanisms and molecular pathways underlying cancer initiation and progression
- Tumor metastasis
- Tumor evolution and heterogeneity
- Tumor microenvironment and tumor-host interactions
- Cancer genetics and genomics
- Cancer characterization using omics approaches
- Discovery and validation of cancer biomarker
- Discovery of new therapeutic targets
- New approaches of diagnostic and treatment modalities
- Statistical methods in cancer research

Global Translational Medicine is a quarterly journal that focuses on medicine, biological sciences, and biomaterials engineering. The goal of *Global Translational Medicine* is to provide a platform to researchers for showcasing their latest research works in translational medicine so as to advance the field towards the betterment of human health. Despite the advancement of omics and new technologies, the process of transforming these technologies and scientific research results into effective therapies and putting them into clinical use still has a long way to go. *Global Translational Medicine* provides a platform to fill the gaps in preclinical and inter-disciplinary research, to promote clinical translation of scientific research results, and to contribute to the conception of new and improved preventive measures as well as diagnostic and therapeutic techniques of diseases.

Global Translational Medicine covers the following themes: cardiovascular disease, metabolism/diabetes/obesity, neuroscience/neurology, cancer, biomaterials and their applications in medicine, proteomics/metabolomics, pharmacogenomics, biomarkers, bioinformatics and data mining, animal and clinical research, and medical methods arising from interdisciplinary crossover.



Start a new journal

Write to us via email if you are interested to start a new journal with AccScience Publishing. Please attach your CV, professional profile page and a brief pitch proposal in your email. We shall inform you of our decision whether we are interested to collaborate in starting a new journal.

Contact: info@accscience.com

https://accscience.com/journal/IJAMD

Contact

www.accscience.com

9 Raffles Place, Republic Plaza 1 #06-00 Singapore 048619

E-mail: editorial@accscience.com

Phone: +65 8182 1586