nature communications

Article

https://doi.org/10.1038/s41467-026-70361-y

Rejoining fragmented ancient bamboo slips with physics-driven deep learning

Received: 4 January 2026

Accepted: 24 February 2026

Published online: 06 March 2026

Check for updates

Jinchi Zhu[1,2,3,4,10], Zhou Zhao[5,10], Hailong Lei[3,4,10], Xiaoguang Wang[6,7], Jialiang Lu[3,4], Jing Li[3,4], Qianqian Tang[1], Jiachen Shen[1], Gui-Song Xia[1,2,8] ✉, Bo Du[1,2,8,9] ✉ & Yongchao Xu[1,2,6,8,9] ✉

Bamboo slips are a crucial medium for recording ancient civilizations in East Asia. However, many excavated bamboo slips have been fragmented into thousands of irregular pieces, making their rejoining a vital yet challenging step for understanding their content. Here we introduce WisePanda, a physics-driven deep learning framework designed to rejoin fragmented bamboo slips. Based on the physics of fracture and material deterioration, WisePanda automatically generates paired synthetic training data that captures the physical properties of bamboo fragmentations. This approach enables the training of a matching network without requiring manually paired samples, providing ranked suggestions to facilitate the rejoining process. Compared to the leading curve matching and modern generative methods, WisePanda yields a substantial and statistically significant increase in matching accuracy, accelerating archeologists' efficiency in rejoining fragmented bamboo slips. WisePanda provides a new paradigm for addressing data scarcity in ancient artifact restoration through physics-driven machine learning.

Bamboo slips, serving as a fundamental medium for documenting ancient East Asian civilizations, contain invaluable historical records spanning philosophy, law, and social life of the period[1–3]. Their durability has enabled these artifacts to survive millennia underground while retaining legible content, offering scholars unprecedented insights into historical societies. However, the excavation of these delicate artifacts presents a critical challenge[2,4] - many bamboo slips have been fragmented into thousands of pieces (Fig. 1), significantly complicating efforts to reconstruct and interpret their content. This fragmentation creates a fundamental obstacle in accessing the wealth of historical knowledge contained in these artifacts.

The rejoining of fragmented bamboo slips represents one of the most challenging problems in archeological preservation and cultural heritage studies[5]. The difficulty stems from multiple factors: First, the enormous number of potential fragment combinations makes manual matching extremely time-consuming, for instance, the Qin bamboo slips from the Shuihudi site comprise fragments numbering in the tens of thousands, with each piece potentially matching any of the others[6]. Second, environmental factors like moisture and pressure have caused extensive physical deterioration[7], distorting the original shapes and surfaces. Third, the sparse recorded characters often make it difficult to capture textual remnants at the fracture curves created by transverse breakage, rendering text-based matching approaches largely ineffective and necessitating reliance primarily on the morphological patterns of fracture curves for identification. The unique fiber structure of bamboo creates complex fracture patterns that traditional

[1]Institute of Artificial Intelligence, School of Computer Science, Wuhan University, Wuhan, China. [2]School of Artificial Intelligence, Wuhan University, Wuhan, China. [3]Center of Bamboo and Silk Manuscripts, Wuhan University, Wuhan, China. [4]Paleography and Chinese Civilization Inheritance and Development Program Collaborative Innovation Platform, Wuhan University, Wuhan, China. [5]School of Computer Science, Central China Normal University, Wuhan, China. [6]Intelligent Computing Laboratory for Cultural Heritage, Wuhan University, Wuhan, China. [7]School of Information Management, Wuhan University, Wuhan, China. [8]National Engineering Research Center for Multimedia Software, Wuhan University, Wuhan, China. [9]Hubei Key Laboratory of Multimedia and Network Communication Engineering, Wuhan University, Wuhan, China. [10]These authors contributed equally: Jinchi Zhu, Zhou Zhao, Hailong Lei. ✉e-mail: guisong.xia@whu.edu.cn; dubo@whu.edu.cn; yongchao.xu@whu.edu.cn

https://doi.org/10.1038/s41467-026-70361-y

Fig. 1 | The challenge of bamboo slip rejoining in archeological research. Some excavated bamboo slip fragments arranged in a spiral pattern (left) illustrate the overwhelming scale of the rejoining challenge faced by archeologists---with thousands of fragments at a single excavation site, each potentially matching with any other piece. A zoomed view of a representative fragment (upper right) reveals the characteristic irregular breakage pattern along the fracture curve. An example of a reconstructed complete slip (lower right) demonstrates how proper rejoining establishes morphological continuity, which is essential for the interpretation of ancient texts. The complex morphology of the fractures, coupled with deterioration from millennia underground, creates a matching problem of extraordinary complexity that has traditionally required intensive manual effort, with experts sometimes spending weeks to successfully match a single pair of fragments. The images of the ancient bamboo slips reused with permissions obtained from the Center of Bamboo and Silk Manuscripts, Wuhan University (the same applies to all subsequent figures of bamboo slips in this article).

curve-matching approaches struggle to analyze effectively. These challenges have made the restoration process highly labor-intensive, with experts sometimes requiring weeks to successfully match a single pair of fragments.

Here, we overcome these challenges through physics-driven deep learning. We propose WisePanda, a deep learning framework that leverages physical principles to overcome the data scarcity problem inherent in the fragmented bamboo slip rejoining task. Training deep learning models for fragment rejoining involves an inherent paradox: the manual fragment rejoining they seek to replace is itself the labor-intensive process traditionally needed to create their training data. We resolve this dilemma by resorting to the physics of fracture[8]. By modeling the physical properties of bamboo and the process that governs its degradation[7,9], we generate extensive synthetic training data that captures the essential characteristics of real paired fragment slips (Fig. 2). This physics-driven approach enables us to produce large-scale, realistic paired training data without requiring manual matching efforts, while ensuring the model learns meaningful patterns based on actual material properties rather than superficial features.

The generation of training data for bamboo slip rejoining requires a comprehensive understanding of both fracture formation and degradation processes[10]. Our physical model builds on two key observations: First, bamboo slips exhibit distinct “transverse” and “longitudinal” fracture modes due to their unique fiber structure[10], as illustrated in Supplementary Fig. 1, with transverse fractures accounting for approximately 70% of cases. Second, these fractures undergo complex corrosion processes over time, significantly altering their original patterns[2,7,11]. Building upon these insights, we developed a systematic framework to synthesize fragment pairs exhibiting transverse fractures, which constitute the predominant type of rejoining challenge.

The physics of fracture in bamboo slips is governed by their distinctive vertical fiber arrangement[12–14], as detailed in Supplementary Fig. 2. When stress extends horizontally across the slip, it creates characteristic fracture patterns as it propagates through consecutive fiber bundles[14,15]. We model this process through a probabilistic framework where the fracture angle of each fiber bundle is influenced by the stress field generated at the fracture endpoint of the preceding fiber bundle (Fig. 2a, b). This physical relationship can be expressed through a probability density function that determines likely fracture paths based on stress distribution and geometric configurations. The corrosion process is then simulated by analyzing geometric exposure - protruding areas and exposed surfaces are more susceptible to degradation through moisture absorption and microorganism activity[16]. To efficiently calculate the exposure area of each fiber bundle, we ingeniously employ ReLU functions that compare the height differences between adjacent fiber bundles, effectively capturing only the protruding portions that are more vulnerable to environmental deterioration (Fig. 2c). This computational approach allows us to precisely model the progressive and differential deterioration patterns observed in archeological samples. To ensure the generated data accurately reflects real bamboo slip characteristics, we optimize model parameters using a genetic algorithm[17]. The algorithm evaluates parameter sets by comparing the distribution of generated fragments

https://doi.org/10.1038/s41467-026-70361-y

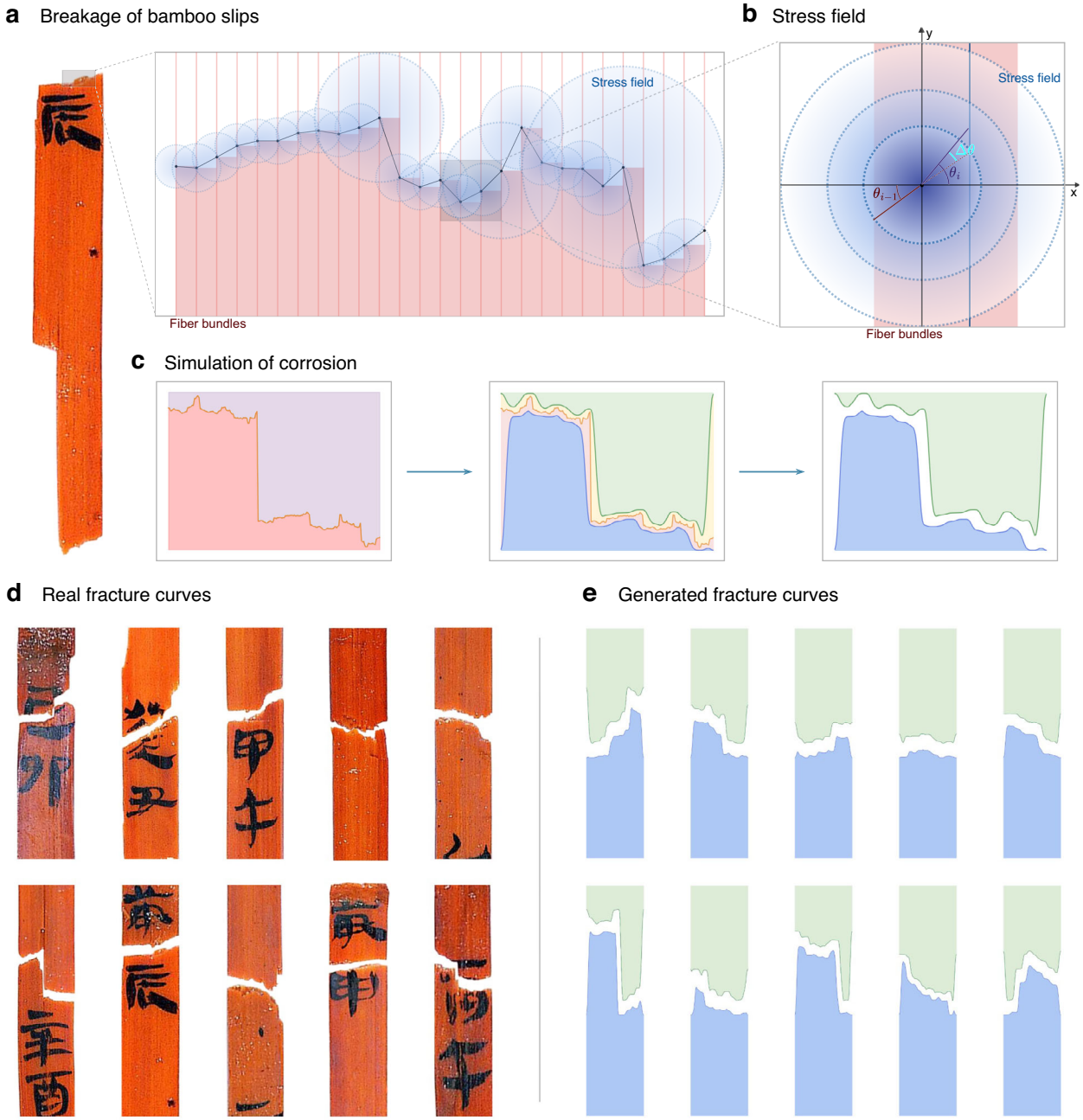

Fig. 2 | Physics-driven modeling of bamboo slip fracture and deterioration. **a** The breakage process of bamboo slips showing how fracture propagates across the bamboo's fiber structure, with the resulting irregular curve composed of black line segments and the corresponding stress field distribution (blue gradient). **b** Detailed stress field model illustrating the mathematical relationship between the fracture angles (θ_{i-1}, θ_i) and the stress propagation in the *x*-*y* coordinate system, where the blue dotted circles represent the stress field emanating from the fracture endpoint, showing how stress radiates outward and concentrates at fiber boundaries. **c** Time-sequential simulation of the corrosion process, demonstrating how the original fracture pattern (left) changes through environmental exposure (middle) to produce the final paired deteriorated curve morphologies (right), with protruding areas experiencing accelerated degradation. **d** Collection of real bamboo slip fracture curves extracted from archeological samples, exhibiting diverse breakage patterns that serve as a reference for model validation. **e** Synthetically generated paired fracture curves produced by our physics-driven model, displaying morphological characteristics highly similar to the real samples in (**d**). This approach enables the generation of extensive paired training data that captures both the physical properties of bamboo fragmentation and the effects of long-term degradation, effectively addressing the data scarcity challenge inherent in ancient artifact restoration.

against a reference set of 200 real fragment curves through dimensionality reduction and Silhouette analysis[18]. Our analysis shows high similarity between the synthetic and real data distributions. The resulting pipeline enables us to generate extensive paired training data (Fig. 2d, e) that captures both the physical properties of fracture formation and the effects of long-term degradation, effectively addressing the data scarcity challenge in rejoining fragmented bamboo slips.

WisePanda employs a TripletNet-based deep learning network[19] trained on physics-driven synthetic data to learn effective feature representations that distinguish matching and non-matching fragments, providing archeologists with ranked candidate suggestions to streamline the rejoining process (Fig. 3). WisePanda demonstrates substantial improvements in fragment matching across multiple datasets. On our primary test set of 118 expert-verified bamboo slip

https://doi.org/10.1038/s41467-026-70361-y

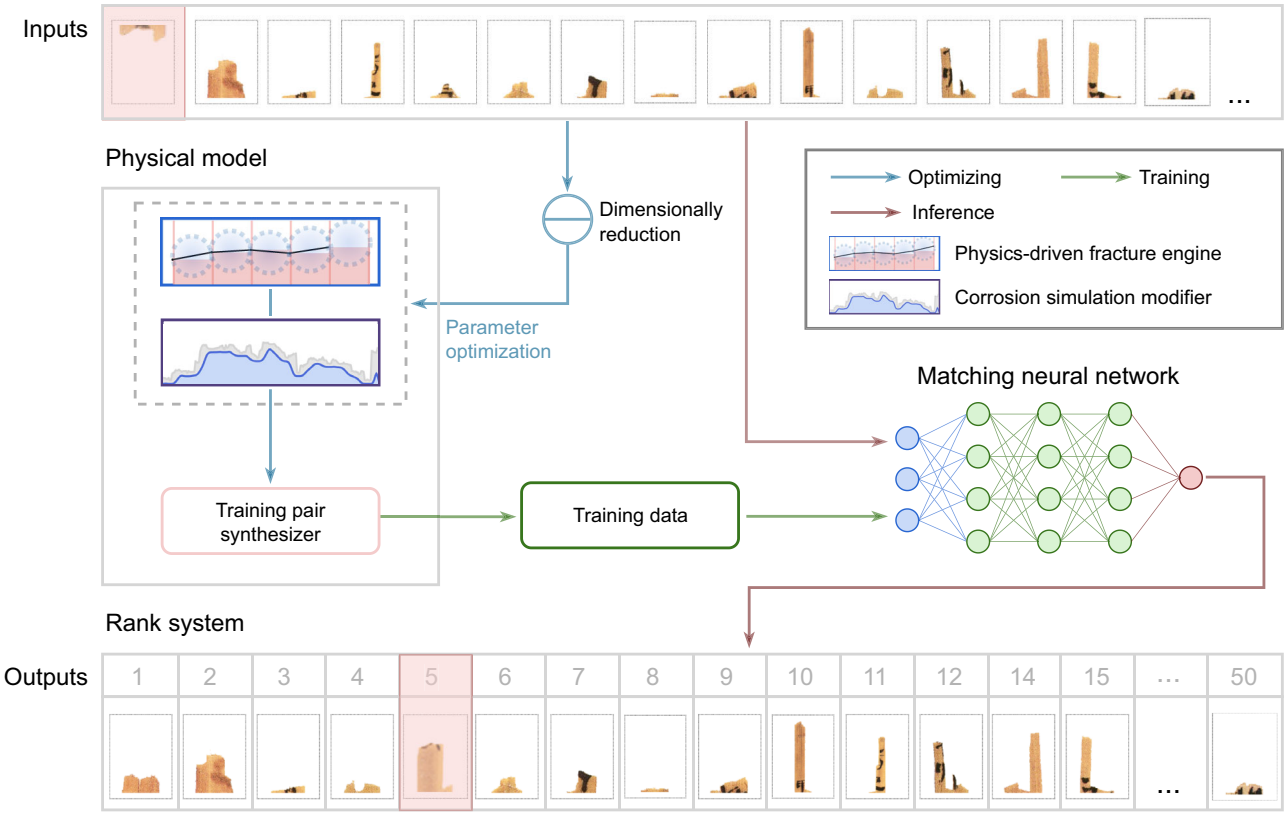

Fig. 3 | Framework of WisePanda integrating physics-driven deep learning for bamboo slip rejoining. The workflow diagram illustrates the complete pipeline of our system. The process begins with input bamboo slip fragments (top row), where the target fragment is highlighted in pink. In the physical model stage, we implement two key components: a physics-driven fracture engine that generates initial breakage patterns, and a corrosion simulation modifier that replicates long-term degradation effects. The dimensionality reduction module (center) enables comparison between real and synthetic fragments through statistical visualization techniques, facilitating parameter optimization to ensure realistic fracture simulation. The optimized physical model feeds into the training pair synthesizer, which generates matching fragment pairs for training the matching network. This matching network learns to effectively represent and compare fragment curves through a multi-layer architecture with specialized embedding capabilities. During inference (red arrows), the system compares the target fragment against all candidates, producing a ranked list of potential matches (bottom row). The ranking transforms thousands of possible matches into a manageable set, with the system highlighting the most probable candidates (position 5 shown in pink represents a correct match). This approach enables archeologists to focus their verification efforts on the most promising candidates, significantly accelerating the traditionally time-consuming rejoining process.

pairs (Bamboo236), WisePanda achieves 91.81% Top-50 accuracy, significantly outperforming leading curve matching methods (65.25 −85.03%) and modern generative approaches (70.48 −83.62%), with statistical significance confirmed across most metrics. When evaluated on challenging scenarios with increased interference fragments (Bamboo1350), WisePanda maintains 52.54% Top-50 accuracy compared to 12.71 −38.28% for competing methods, demonstrating robust performance under realistic archeological conditions. The framework also generalizes to wooden slip fragments, achieving 31.24% Top-50 accuracy despite material differences. Beyond numerical performance, WisePanda provides interpretability through its physics-driven foundation - unlike pure data-driven approaches that may overfit to superficial patterns, our model learns features grounded in actual fracture and degradation processes. We have developed a computer-assisted tool implementing WisePanda's capabilities that archeologists are now using to accelerate the traditional rejoining workflow. This work establishes a new paradigm for archeological fragment matching when paired training data is scarce but physical principles are well understood, with potential applications extending to other fragmented cultural artifacts.

Results

Dataset collection and preparation

The complete workflow from fragment collection through digitization to WisePanda's input is illustrated in Supplementary Fig. 4. For evaluating WisePanda's effectiveness, we collected several datasets of bamboo and wooden slip fragments. Our primary test set (Bamboo236) consists of 118 pairs of matched bamboo slip fragments, carefully verified and provided by archeological experts who successfully rejoined these fragments through meticulous manual work (achieved by three archeological experts within half a year). These bamboo slip artifacts are part of an ongoing archeological research project on ancient bamboo slips dating to the Western Han dynasty during Emperor Wen's reign (179–157 BCE), originated from Tomb 77 at Shuihudi, Yunmeng County, Hubei Province, excavated in 2006, with the cultural relics yet to be formally published. To simulate more challenging real-world scenarios, we created an extended dataset (Bamboo1350) by introducing 1114 additional interference fragments from the same set, expanding the candidate pool to 1,350 total fragments while maintaining the original matched pairs. To assess generalization capabilities across different materials, we also collected wooden slip datasets. The base wooden slip test set (Wood670) contains 335 expert-verified pairs, while its extended version (Wood3833) includes 3163 additional interference fragments. These diverse datasets enable comprehensive evaluation of fragment matching methods under varying conditions and material types.

Evaluation methodology and baseline comparison

WisePanda relied on manual pre-definition of fragment orientation, with archeologists easily pre-classifying fragments as upper or lower

Table 1 | Comparison of different fragment matching methods across different test scenarios

	Method	Top-1(%)	Top-5(%)	Top-10(%)	Top-20(%)	Top-50(%)	er[0.5] Top-100(%)
Bamboo236	DTW[20]	8.05 ± 0.42	22.88 ± 0.43	31.36 ± 0.43	48.73 ± 1.47	74.72 ± 2.00	94.63 ± 0.89
	FMM[21]	8.90 ± 0.85	18.64 ± 0.00	26.41 ± 0.65	37.71 ± 1.53	65.25 ± 1.70	94.35 ± 0.24
	SIS[22]	10.45 ± 0.24	21.47 ± 0.98	29.38 ± 2.34	44.07 ± 4.43	71.47 ± 2.45	94.07 ± 0.00
	Event-DTW[23]	7.63 ± 0.00	24.01 ± 0.65	31.35 ± 0.73	48.45 ± 1.07	75.42 ± 1.12	94.63 ± 0.25
	Drop-DTW[24]	8.34 ± 1.22	23.02 ± 0.65	31.78 ± 0.42	49.29 ± 0.25	75.56 ± 0.65	94.92 ± 0.00
	Modified COW[25]	8.05 ± 0.85	24.44 ± 0.25	38.42 ± 1.71	57.06 ± 1.22	85.03 ± 0.65	98.59 ± 0.65
	Two-stage DTW[26]	7.63 ± 0.00	23.45 ± 0.49	30.65 ± 0.24	49.44 ± 0.25	74.58 ± 0.00	95.34 ± 0.00
	GAN[27]	4.52 ± 1.07	17.94 ± 1.07	32.06 ± 3.66	54.52 ± 2.45	81.36 ± 0.85	96.61 ± 0.73
	seriesGAN[28]	4.38 ± 2.69	19.49 ± 4.04	32.49 ± 4.11	52.97 ± 3.88	83.62 ± 1.07	96.75 ± 0.65
	DDPM[29]	3.39 ± 0.42	10.03 ± 0.88	19.63 ± 2.55	34.75 ± 0.85	70.48 ± 5.52	96.05 ± 0.98
	Diffusion-TS[30]	3.39 ± 1.47	14.83 ± 1.27	27.26 ± 2.82	45.91 ± 2.82	80.37 ± 3.55	98.73 ± 0.85
	WisePanda	13.42 ± 1.07	37.15 ± 1.76	50.85 ± 1.53	66.81 ± 2.89	91.81 ± 2.01	100.00 ± 0.00
Bamboo1350	DTW[20]	3.95 ± 0.25	11.72 ± 0.24	14.12 ± 0.25	19.78 ± 0.65	29.80 ± 1.91	37.29 ± 1.12
	FMM[21]	5.79 ± 1.36	11.02 ± 0.00	12.85 ± 0.25	15.40 ± 0.49	23.45 ± 1.49	32.20 ± 1.53
	SIS[22]	7.34 ± 0.25	14.12 ± 1.29	17.94 ± 1.61	22.04 ± 1.85	32.35 ± 3.45	44.21 ± 4.53
	Event-DTW[23]	4.09 ± 0.49	12.85 ± 0.65	15.96 ± 1.07	22.46 ± 1.47	34.04 ± 0.88	42.94 ± 2.00
	Drop-DTW[24]	4.24 ± 0.43	11.16 ± 0.24	14.41 ± 0.43	19.92 ± 0.00	29.66 ± 0.00	37.29 ± 0.74
	Modified COW[25]	2.97 ± 0.85	9.18 ± 0.65	13.70 ± 1.48	19.35 ± 1.71	28.96 ± 1.71	41.53 ± 2.55
	Two-stage DTW[26]	4.10 ± 0.25	12.29 ± 0.00	14.41 ± 0.00	19.78 ± 0.65	29.38 ± 0.24	36.86 ± 0.43
	GAN[27]	0.56 ± 0.65	4.24 ± 0.85	7.77 ± 1.71	14.69 ± 3.60	26.27 ± 5.73	41.95 ± 8.35
	seriesGAN[28]	1.69 ± 1.47	6.78 ± 0.73	12.43 ± 1.91	21.33 ± 2.18	38.28 ± 3.53	49.15 ± 4.78
	DDPM[29]	0.42 ± 0.43	2.26 ± 0.49	3.53 ± 1.07	5.23 ± 1.07	12.71 ± 1.70	21.47 ± 2.59
	Diffusion-TS[30]	1.55 ± 1.30	4.52 ± 1.36	8.76 ± 2.09	15.54 ± 2.72	27.97 ± 1.12	43.08 ± 2.89
	WisePanda	5.37 ± 1.76	16.53 ± 1.70	25.42 ± 1.12	37.57 ± 0.88	52.54 ± 0.85	64.40 ± 2.24
Wood670	DTW[20]	2.60 ± 0.08	6.71 ± 0.37	11.45 ± 1.13	16.89 ± 1.28	23.61 ± 1.92	29.06 ± 1.18
	FMM[21]	2.19 ± 0.09	6.92 ± 0.22	9.35 ± 0.31	13.28 ± 0.00	21.64 ± 0.30	27.41 ± 0.17
	SIS[22]	2.34 ± 0.17	7.81 ± 0.17	10.75 ± 0.00	14.92 ± 0.51	23.08 ± 0.60	27.96 ± 0.00
	Event-DTW[23]	2.84 ± 0.25	7.06 ± 0.52	10.95 ± 0.08	15.87 ± 0.62	23.38 ± 0.23	28.81 ± 0.26
	Drop-DTW[24]	2.39 ± 0.30	6.87 ± 0.30	10.9 ± 0.15	15.42 ± 0.31	23.33 ± 0.17	28.76 ± 0.23
	Modified COW[25]	0.95 ± 0.22	2.89 ± 0.45	5.02 ± 0.62	7.86 ± 1.12	14.53 ± 1.45	22.59 ± 1.25
	Two-stage DTW[26]	2.74 ± 0.09	6.67 ± 0.31	10.70 ± 0.17	15.87 ± 0.23	23.13 ± 0.15	28.66 ± 0.15
	GAN[27]	0.60 ± 0.00	2.24 ± 0.26	4.28 ± 0.31	6.96 ± 0.34	14.48 ± 0.83	24.53 ± 0.90
	seriesGAN[28]	0.60 ± 0.15	2.94 ± 0.37	5.07 ± 0.91	8.41 ± 1.25	17.21 ± 0.75	26.97 ± 0.56
	DDPM[29]	0.40 ± 0.09	1.29 ± 0.23	2.59 ± 0.43	4.73 ± 0.70	10.45 ± 1.42	18.16 ± 0.99
	Diffusion-TS[30]	0.55 ± 0.23	2.14 ± 0.61	4.33 ± 0.65	7.86 ± 1.13	15.92 ± 1.04	25.17 ± 1.53
	WisePanda	1.89 ± 0.19	10.14 ± 0.21	17.25 ± 1.56	24.31 ± 2.70	31.24 ± 1.25	35.61 ± 1.65
	WisePanda[Wood3833]	1.11 ± 0.29	2.82 ± 0.19	6.67 ± 1.39	9.48 ± 1.78	16.72 ± 1.27	18.38 ± 1.08

Results show Top-k matching accuracy (percentage of correctly identified matches within the top k candidates) across bamboo and wooden slip datasets. The original bamboo test set (Bamboo236) contains 118 paired fragments, while the expanded dataset (Bamboo1350) expands the candidate pool with 1114 additional interference fragments. Similarly, for wooden slips, we compare performance on the original test set (Wood670) and an expanded dataset (Wood3833).

parts based on visual inspection of fracture patterns and text orientation. The evaluation process involved computing similarity scores between each fragment's fracture curve and all potential matching candidates from the pool.

We extracted features from both fracture curves of each fragment, ranked the computed similarity scores, and measured the probability of finding the correct match within the Top-k predictions.

We conducted comprehensive extensive experiments comparing our WisePanda with two categories of methods: (1) Curve matching methods developed from 1994 to 2025: including curve matching with Dynamic Time Warping (DTW)[20], Fast Matching Method (FMM)[21], Scale Invariant Signature (SIS)[22], Event-DTW[23], Drop-DTW[24], Modified COW[25], and Two-stage DTW[26]; (2) Modern generative approaches: including the pioneer GAN[27], seriesGAN[28] for sequential data, DDPM[29] representing the pioneer diffusion paradigm, and Diffusion-TS[30] representing conditional generation optimized for time series data.

For each compared method, we reported the mean and standard deviations of Top-k accuracy across three runs by varying each method's parameters (for curve matching methods) or random seeds (for deep-learning-based methods). We also reported P-values with permutation tests across all Top-k metrics comparing our WisePanda and the other methods. This comparative framework allows us to assess the relative performance of our physics-driven approach against established curve matching techniques and cutting-edge generative methods in the field of fragment matching.

Performance analysis and statistical validation

The comprehensive comparison (Table 1) among different methods reveals insights about the nature of fragmented bamboo slip matching problem. On our original test set (Bamboo236), WisePanda demonstrated superior performance across all Top-k metrics, in particular achieving in average a Top-50 accuracy of 91.81%. The curve matching

methods and modern generative methods showed varying performances, ranging from 65.25% to 85.03% and 70.48% to 83.62%, respectively. To simulate real-world archeological scenarios where numerous unrelated fragments exist, we expanded our test dataset by introducing additional interference fragments. As shown in Supplementary Fig. 5, when the candidate pool was increased by a factor of approximately 10, WisePanda maintained a Top-50 accuracy of 52.54% in average compared to the original 91.81%, demonstrating robust performance even with a significantly increased search space. Indeed, on the Bamboo1350 dataset, we observed similar performance superiority of our WisePanda with 52.54% Top-50 accuracy over the other methods, with a Top-50 accuracy ranging from 12.71% to 38.28%. Statistical significance testing further confirmed WisePanda's superiority over all competing methods across most evaluation metrics (Supplementary Table 1, for detailed P-values). We also quantified the similarity between real fragmented curves and synthetic ones generated by different methods using the silhouette score defined by Equation (9). Our physics-driven data synthesis achieved -0.0082 silhouette score compared with much larger values for the other modern generative methods (GAN[27]: 0.4095, seriesGAN[28]: 0.4399, DDPM[29]: 0.4285, Diffusion-TS[30]: 0.3978), demonstrating WisePanda's superiority in generating more authentic fracture curves and thus explaining WisePanda's better performance over the other modern generative methods.

Interpretability of physics-driven approach

The fundamental challenge in generating training data for fragmented bamboo slip matching lies in the irreversible nature of the physical degradation process[7,31]. When archeologists uncover bamboo slip fragments, they observe only the final degraded curves[2] (Fig. 2c). The original fracture patterns are permanently altered by centuries of degradation and cannot be directly inferred from the preserved fragments[5]. This irreversibility poses a significant limitation for conventional data-driven approaches[32,33], which typically require paired samples of "before-degradation" and "after-degradation" states for effective training data that is inherently unavailable in archeological contexts. Even generative models like generative adversarial networks and diffusion models suffer from distribution bias when learning only from limited pairs of degraded real fragments, which is the usual case in the field of rejoining ancient fragmented bamboo slips. Whereas, incorporating the underlying physical principles that govern fracture formation and degradation may provide a surrogate solution.

Our physics-driven model addresses this challenge by modeling the forward process of fragmentation and degradation. We first simulate the initial fracture patterns based on the physics of bamboo fiber structures[9,10,34,35], then apply physical degradation processes to these patterns. This sequential modeling captures both the immediate fracture characteristics and their subsequent transformation through environmental exposure. The physical processes governing bamboo deterioration follow specific geometric principles, where protruding areas and exposed surfaces are more susceptible to degradation[7,31]. By incorporating these physical constraints, WisePanda generates realistic degradation patterns that closely mirror natural deterioration processes, effectively bridging the gap between the unknown original fracture patterns and the observed degraded edges.

These physical principles not only enhance the quality of synthetic training data but also provide interpretability to our WisePanda's behavior. The combination of authentic fracture patterns and physically-based degradation enables WisePanda to capture fundamental characteristics of bamboo slip fragmentation, generating abundant realistic training data and leading to more robust and generalizable performance. Unlike pure data-driven approaches that may overfit to superficial patterns in limited training samples, our physics-driven model learns meaningful features grounded in the actual physical processes of fracture and degradation, allowing it to generalize effectively to unseen fragments with diverse deterioration patterns.

Generalization to ancient wooden slip fragments

To examine WisePanda's generalization capability, we applied our approach to wooden slips - another important medium for ancient text preservation that shares similar characteristics with bamboo slips but exhibits distinct material properties. While wooden slips served similar documentary purposes, their different fiber arrangement and material structure present unique challenges for fragment matching[36]. We evaluated our method on a dataset of 335 pairs of wooden slip fragments, following the same experimental protocol as with bamboo slips.

As shown in Table 1, WisePanda demonstrated promising cross-material generalization with an average 31.24% Top-50 accuracy on wooden slips, outperforming the other methods. When tested with extended interference fragments in both materials, performance predictably decreased to 16.72% for wooden slips (Wood3833), which nevertheless remains valuable for assisting archeologists in rejoining ancient, fragmented wooden slips.

This performance difference can be attributed to two key factors. First, the fiber structure of wooden slips differs substantially from bamboo slips, leading to distinct fracture patterns that deviate from our physics-driven model's assumptions[37]. Second, the degradation process varies significantly due to diverse preservation conditions - factors such as burial duration, soil moisture levels, microbial activity, and regional soil composition all affect how materials deteriorate over time[38]. Even artifacts of the same material excavated from different sites or time periods may exhibit varying degradation patterns, resulting in different discriminative characteristics within different datasets and requiring different physical parameters for accurate modeling. These variations are particularly pronounced when comparing wooden and bamboo materials, as their distinct physical properties interact differently with environmental factors, requiring re-design and re-train the physics-driven model to get optimal performance for wooden slips.

Discussion

In this work, we present WisePanda, the first physics-driven deep learning framework designed for rejoining fragmented ancient bamboo slips. By incorporating principles from the physics of fracture into the generation of synthetic training data, our approach overcomes the fundamental challenge of paired data scarcity in AI for rejoining fragmented ancient bamboo slips, and significantly improves the efficiency of archeologists. The effectiveness of this approach is demonstrated through superior matching accuracy and the development of practical tools that are now assisting archeologists in their reconstruction work. Our framework's success in handling both fracture pattern generation and material degradation simulation highlights the advantage of combining physical principles with deep learning, particularly in scenarios where traditional data-driven approaches are constrained by limited training samples.

Beyond its current implementation, this research opens new horizons in bamboo slip reconstruction and broader cultural heritage preservation. For bamboo slips specifically, our framework can be extended to handle more complex scenarios by incorporating additional physical parameters that better account for regional variations and different preservation conditions. Furthermore, our approach establishes a new paradigm for fragment matching where paired training data is scarce. The incorporation of physics-driven mechanisms can be adapted to other archeological materials such as wooden slips, ceramics, metals, and pottery, by modeling their specific physics of fracture and degradation processes.

To situate our pairwise matching module within practical archeological workflows, WisePanda is readily incorporated into a hierarchical multi-fragment reassembly process, as illustrated in Supplementary Fig. 6. Since bamboo slips exhibit substantial variation in length while remaining narrow in width, fragments can be organized into two matching pools according to their horizontal completeness: longitudinally fractured pieces that lack full transverse width, and transversely fractured pieces that preserve it. This separation enables a structured assembly procedure in which longitudinal fragments are first matched and merged to restore horizontal integrity; newly formed horizontally complete composites are then transferred into the transverse pool, where further matching proceeds. After each assembly iteration, archeologists conduct a brief manual assessment to decide whether a composite should continue in the matching cycle or be finalized. In the present work, WisePanda primarily targets transversely fractured slips, where edge morphology offers a stable and consistently observable cue that can be modeled through our physics-driven framework. By contrast, longitudinal rejoining typically relies on complementary signals such as character continuity and textual layout, which require a distinct computational strategy. For this reason, the longitudinal stage in the workflow is designed as an independent module: the current manuscript focuses on the physics-based transverse matching component, while explicitly reserving an interface for plugging in specialized longitudinal-matching methods. We have already conducted preliminary experiments with a character-continuity-based module for longitudinal fractures, made available alongside the code as an extended component, but its technical details are kept outside the scope of this paper to preserve the conceptual coherence of the proposed physics-driven paradigm.

While WisePanda achieves promising results, certain challenging cases remain for future exploration. Fragments with transversely straight and relatively flat fracture edges lack distinctive geometric features, making it difficult to distinguish among many similar candidates (Supplementary Fig. 10). For such geometrically ambiguous cases, incorporating additional cues such as character continuity, texture patterns, and ink stroke characteristics represents a promising direction for enhancement, particularly when coupled with more explicit human-AI collaborative workflows. From a methodological perspective, integrating expert knowledge more deeply into the learning process presents another opportunity - future systems might implement reinforcement learning paradigms where archeologist feedback continuously refines the matching models. Additionally, exploring self-supervised or few-shot learning approaches could further reduce the dependence on synthetic data, potentially capturing even more nuanced aspects of artifact deterioration that are challenging to model explicitly.

In conclusion, WisePanda represents not merely a physics-informed methodological framework in fragmented bamboo slip rejoining but rather a transformative approach to preserving and studying cultural heritage. By combining physical principles with artificial intelligence, our method provides a robust solution when paired training data is limited, but physical principles are well understood. This synergy of physics and deep learning opens new possibilities for archeological fragment matching, potentially revolutionizing how we recover and interpret artifacts from our past.

Methods

Previous work

In recent years, various approaches have been proposed for restoring and analyzing fragmented ancient artifacts[5,39]. Traditional methods have primarily focused on geometric feature analysis, particularly curve matching methods. Modern approaches mainly leverage artificial intelligence techniques[40,41]. The challenge of reassembling fragmented ancient texts[42] has been addressed across diverse cultural materials and time periods. Early computational attempts include Tyndall's work on Hittite cuneiform tablets using naive Bayes and maximum entropy classifiers[43], followed by Collins et al.'s matching algorithm for 3D scans of cuneiform tablets[44]. Notable recent work includes the restoration of Pompeii's frescoes using a robotic system for automated fragment matching[45], Ithaca, a deep neural network system that successfully restored and attributed ancient Greek texts[46], Aeneas, recent advance in contextualizing ancient texts using generative neural networks[47], and the reunion helper for Dunhuang manuscript fragments[48,49]. The adoption of deep learning architectures has marked a significant shift in the field. Pirrone et al. employed a Siamese network to match papyrus fragments from the APIS UM Papyrology Database, achieving high accuracy on synthetic datasets spanning multiple ancient scripts[50]. Abitbol et al. developed a more sophisticated system for Dead Sea Scrolls reassembly, combining CNN-based fiber pattern recognition with geometric alignment and random forest classifiers[51]. For oracle bone fragments, researchers have explored using GAN for augmenting training pairs and self-supervised learning for restoration tasks[52]. These works demonstrate the diverse applications of computational methods across different cultural heritage materials.

The manual restoration of bamboo and wooden slips has a long history in China[53–55]. Wang Guowei's pioneering work *Liusha Zhuijian*[56] systematically documented traditional techniques for matching and rejoining ancient bamboo and wooden manuscripts, establishing foundational methodologies that are still valuable today. However, manual restoration remains an extremely time-consuming and labor-intensive task, often requiring weeks or months to successfully match a single pair of fragments. To develop computational tools that can assist in this process, a thorough understanding of bamboo's material structure, fracture mechanisms, and degradation patterns is essential. Bamboo's hierarchical fiber structure[57] produces distinctive transverse fracture patterns influenced by fiber orientation[15] and centuries of environmental degradation. Principles from fracture mechanics[8,58] and computational modeling[59] provide theoretical foundations for simulating these physical processes in data generation and fragment matching.

With recent technological advancements, computational approaches and artificial intelligence have emerged as powerful tools to assist archeologists in this complex process, significantly accelerating the restoration workflow while maintaining the essential role of expert knowledge. While deep learning approaches have shown promise in cultural relic restoration, the application to fragmented bamboo slips faces significant challenges due to the lack of sufficient paired training data and the complex physical degradation patterns unique to bamboo materials. Recent advances in physics-informed learning have demonstrated the potential of integrating domain knowledge with data-driven approaches. Physics-informed neural networks successfully encode physical laws into deep learning architectures[60,61], enabling more accurate predictions with limited training data. In the cultural heritage domain, computational systems for fragment matching and reassembly[62] have demonstrated the potential of digital approaches in archeological restoration. These developments suggest that incorporating physical principles into data-driven methods can enhance both performance and interpretability. Our work addresses these limitations by introducing physics-driven data generation. Unlike previous approaches, we integrate fracture mechanics and material degradation principles into the restoration process.

This physics-driven methodological framework allows us to generate physically plausible paired training data that reflects the actual physical processes of how bamboo slips break and degrade over time.

Moreover, our physics-driven approach not only provides interpretable results that align with archeological domain knowledge but also demonstrates strong extensibility to various archeological fragment restoration tasks.

Our contribution lies in establishing the physics-driven methodological paradigm that transforms the data requirement from relying on authentic paired datasets to leveraging physical principles to synthesize paired datasets for training-based methods, distinguishing our approach from previous geometric curve matching and purely data-driven generative methods. The framework we propose may open a new paradigm for applying deep learning in archeological restoration[63], particularly valuable when paired training data is scarce but physical principles are well understood. This methodology can potentially be adapted to restore other types of archeological fragments by modeling their specific physical properties and degradation processes[64,65], offering a generalizable solution for the broader field of cultural heritage preservation.

Physical model for data generation

Bamboo slip fragments exhibit two distinct fracture modes: transverse and longitudinal, with transverse fractures accounting for approximately 70% of cases[15]. Due to bamboo's unique fiber structure, the transverse fractures occur when stress cuts across the bamboo slip's cross-section, creating uneven, wavy edge curves. The longitudinal fractures happen along the fiber direction, resulting in straight edges, as illustrated in Supplementary Fig. 1. In this work, we focus on transverse fractures due to their prevalence and complexity. Although less common, some fragments exhibit hybrid transverse-longitudinal characteristics. In such cases, maintaining at least one region of horizontal integrity ensures consistent transverse scaling during curve extraction, whereas the absence of such a reference can lead to horizontal stretching and scale distortion, as illustrated in Supplementary Fig. 7 To systematically model the bamboo slip structure, we establish a coordinate system where the x-axis represents the horizontal direction along the bamboo slip's width and the y-axis represents the vertical direction corresponding to fiber bundle heights. We abstract the bamboo slip as a series of N_f vertically aligned fiber bundles, as shown in Supplementary Fig. 2(a). Each fiber bundle is represented as a discrete structural unit with a uniform width of δ_x, forming a continuous arrangement along the horizontal axis. The initial height for the front of ith fiber bundle before environmental degradation is denoted as y_i^{init}. The final height after corrosion processes is represented as y_i^{final}. This discretization allows us to precisely model the propagation of fracture across the material's anisotropic structure. The vertical positions of these fiber bundles determine the morphology of the fracture curve, with the height of each bundle representing its position on the fracture curve. When modeling the fracture process, we observe that bamboo slips were typically buried in a loosely rolled form, causing them to experience predominantly Mode III (out-of-plane shear) stress during burial compression and subsequent deterioration. This specific stress mode occurs when opposing forces act parallel to each other but in different planes, creating a tearing effect particularly relevant to rolled manuscripts. We model the bamboo slip structure as consecutive fiber bundles[12–14,66], where stress propagates from one bundle to another one (Fig. 2a). At the terminal point of each fractured fiber, a stress component K_{III} continues to propagate[67], following the formula:

$$K_{III} = \sqrt{2\pi}\sigma\sqrt{\alpha}, \tag{1}$$

where σ represents the far-field shear force and α denotes the cumulative stress extension path length. The stress field at the fracture propagation front $\sigma_{yz}(r, \theta)$ can be described by:

$$\sigma_{yz}(r,\theta) = \frac{K_{III}}{\sqrt{2\pi r}}\cos\left(\frac{\theta}{2}\right), \tag{2}$$

where r is the radial distance from the fracture propagation front and θ is the angular position relative to the direction of stress propagation, as shown in Supplementary Fig. 2(b). This equation describes how the stress force decreases with distance from the fracture propagation front while also varying with direction.

For the xy plane of the writing surface, we develop a probabilistic model to describe how fracture propagates across consecutive fiber bundles. As illustrated in Supplementary Fig. 2(b), when a fracture reaches point P_2 at the end of fiber bundle f_{i-1}, it generates a stress field that determines the probable direction of continued fracture propagation into the next fiber bundle f_i. The blue dotted circles emanating from point P_2 represent this stress field distribution, with point P_1 marking the beginning of the fracture on bundle f_{i-1} and point P_3 indicating the potential endpoint of the fracture on bundle f_i. The stress field on the xy plane is described by:

$$\sigma_{xy}(r,\theta) = V_{xy}\cdot\frac{K_{III}}{\sqrt{2\pi r}}\cos\left(\frac{\theta}{2}\right), \tag{3}$$

where V_{xy} is a dimensionless constant representing the projection of out-of-plane (Mode III) stress onto the writing surface plane, r is the radial distance from the fracture propagation front at point P_2, and θ is the angular position in the stress field. This formulation accounts for bamboo's anisotropic fiber structure in translating three-dimensional stress into planar fracture propagation. The angles θ_{i-1} and θ_i, measured relative to the horizontal axis, define the fracture directions in consecutive fiber bundles, with $\theta_i = \theta_{i-1} + \Delta\theta$. The stress field determines the probability distribution of angle change $\Delta\theta$, which governs how the fracture direction transitions between fiber bundles. This probability density function can be expressed as:

$$p(\Delta\theta|\theta_{i-1}) = \frac{V_{xy}\cdot K_{III}\cdot\cos(\frac{\Delta\theta}{2})}{\sqrt{2\pi}(\delta_x\cos(\theta_{i-1}+\Delta\theta))}. \tag{4}$$

By substituting $K_{III} = \sqrt{2\pi}\sigma\sqrt{\alpha}$, we simplify the above equation to:

$$p(\Delta\theta|\theta_{i-1}) = \frac{V_{xy}\cdot\sigma\sqrt{\alpha}\cdot\cos\left(\frac{\Delta\theta}{2}\right)}{\delta_x\cos(\theta_{i-1}+\Delta\theta)}. \tag{5}$$

While V_{xy} and σ are unmeasurable constants that scale the probability distribution, they do not affect the relative probabilities of different fracture paths. Once the fracture angle is determined, the initial vertical height relationship between adjacent fiber bundles follows:

$$y_{i+1}^{init} - y_i^{init} = \tan(\theta_i)\cdot\delta_x. \tag{6}$$

This mathematical modeling enables us to construct a Markov chain-like model for fracture propagation, where each state represents a fiber bundle's initial front height y_i^{init}, and transitions between states are governed by the probability density function $p(\Delta\theta|\theta_{i-1})$. This results in a physically plausible model that captures how Mode III stress-induced fracture propagates through bamboo's anisotropic structure, allowing us to generate diverse yet authentic synthetic fracture patterns that serve as the foundation for subsequent corrosion simulation to obtain the final degraded fracture curves.

After the initial fracture curves are generated, the bamboo slip fragments undergo long-term environmental degradation that significantly alters their original morphology. The degradation process of bamboo slip fragments is highly dependent on their geometric features, where sharp edges and protrusions are more susceptible to corrosion compared to smooth surfaces[16,68], even with identical perimeter lengths. This geometric dependency is primarily due to stress concentration at sharp points and increased surface area exposure to environmental factors[68]. We observe that in archeological samples, protruding areas of bamboo slip edges show significantly more deterioration than recessed areas, consistent with their greater environmental exposure. To mathematically capture this geometric dependency in our corrosion model, we employ the ReLU function to

calculate the structural exposure area for each fiber bundle. We define the structural exposure area S_i as:

$$S_i = ReLU(y_i^{init} - y_{i-1}^{init}) + ReLU(y_i^{init} - y_{i+1}^{init}). \quad (7)$$

The erosion process is then modeled by reducing the fiber height based on the structural exposure area and a corrosion rate coefficient c_e. The final degraded height for the front of ith fiber bundle y_i^{final} is given by:

$$y_i^{final} = y_i^{init} - (S_i \times c_e). \quad (8)$$

This approach ensures that regions with higher geometric exposure experience accelerated degradation, accurately simulating how moisture and microorganisms propagate through the bamboo structure[16,68,69]. To capture the progressive nature of long-term degradation, we implement an iterative corrosion process where the degradation is applied over multiple time steps, with each iteration representing a period of environmental exposure. The synchronized update mechanism ensures that all fiber heights are calculated based on the current state before updating to the next iteration, preventing cascade effects where early updates could disproportionately influence later calculations within the same corrosion cycle. This multi-step simulation accurately reflects the physical degradation process observed in actual bamboo slips, where environmental factors affect the entire surface concurrently rather than sequentially, ultimately generating realistic synthetic fragment pairs that capture both the initial fracture characteristics and their subsequent transformation through centuries of environmental exposure.

To ensure our physical model accurately reflects real bamboo slip characteristics, we develop a parameter optimization strategy driven by comparison between real and synthetic data distributions. We first select a set of 200 real fragment curves (set A) extracted from archeological bamboo slips, which represent the final state after both fracture formation and corrosion processes, with curves characterized by their y^{final} coordinates along the fragment curve. Concurrently, we generate a set of 200 simulated fragment curves (set B) using randomly initialized parameters in our physical model, where the synthetic fragments undergo the complete process from initial fracture generation (y^{init} heights) through corrosion simulation to obtain final degraded curve profiles (y^{final} heights). Both sets are subjected to t-SNE (t-Distributed Stochastic Neighbor Embedding)[18] dimensionality reduction, projecting the high-dimensional curve data into a 2D space to facilitate comparison of their distributions. The silhouette score s for these reduced data sets is calculated as:

$$s(j) = \frac{b(j) - a(j)}{\max\{a(j), b(j)\}}, \quad (9)$$

where $a(j)$ represents the mean distance between point j and all other points within its own set (either set of real fragments A or set of synthetic fragments B), and $b(j)$ is the mean distance between point j and all points in the other set. The silhouette score lies in the range of [-1, 1], and measures the distributional distinction between the two datasets. A smaller silhouette score indicates better similarity between the simulated and authentic datasets. In particular, a negative silhouette score signifies that synthetic fragments exhibit characteristics closer to the authentic specimens than to other synthetic samples within their own distribution, demonstrating good replication of real fracture patterns.

To find the optimal set of parameters for our physical model, we employ a genetic algorithm[17] with the fitness function F as:

$$F = \frac{1}{|S|} + \lambda \cdot \text{diversity}, \quad (10)$$

where λ is a weight coefficient that balances optimization precision (minimizing silhouette score) with exploration capability (maintaining genetic diversity). The genetic algorithm manages a population of parameter configurations $\mathcal{P} = \{p_1, \ldots, p_n\}$, where each parameter set p_i contains specific values for both fracture physics and corrosion simulation components of our model. During each optimization iteration, new candidate parameter sets are generated through controlled recombination:

$$p_{new} = \beta \cdot p_1 + (1 - \beta) \cdot p_2 + \epsilon, \quad (11)$$

where p_1 and p_2 are parent parameter sets selected based on their fitness scores, $\beta \in [0, 1]$ is a randomly generated crossover weight that varies with each recombination operation, and ϵ represents a small random mutation noise term added to maintain genetic diversity. This randomized crossover mechanism allows the algorithm to explore different combinations of parameter values while gradually converging towards configurations that produce simulated fracture patterns closely matching the real archeological data[17].

Data preparation and augmentation

Based on our analysis of Mode III (out-of-plane shear) stress-induced fracture mechanisms in bamboo slips[8], which typically occur when rolled bamboo manuscripts experience perpendicular forces during burial and subsequent deterioration, we implement a probabilistic framework to generate initial fracture patterns. The generation process leverages Eq. (5)–(7) to construct a Markov chain-like structure, where each state represents the initial height y_i^{init} of ith fiber bundle (Supplementary Algorithm 1). This algorithm generates fracture curves that closely mimic the physical behavior of bamboo fiber bundles under stress. The probabilistic nature of the generation process, governed by the Probability Density Function (PDF) derived from stress field analysis in Eq. (5), ensures that each generated curve exhibits realistic variations while maintaining the fundamental characteristics of bamboo slip fracture. These generated curves serve as the foundation for subsequent corrosion simulation.

Building upon Eqs. (6)–(8), we developed a multi-scale corrosion model that quantifies degradation based on the geometric exposure of individual fibers and their spatial relationships with adjacent fibers. The ReLU function effectively captures only the protruding portions that are more vulnerable to environmental degradation (Supplementary Algorithm 2). Supplementary Algorithm 2. implements a synchronized update mechanism where all fiber heights' expected changes are calculated first, and then all heights are updated simultaneously at each time step. This approach prevents cascade effects where early fiber updates could disproportionately influence later calculations within the same corrosion cycle. The number of iterations N_c and corrosion rate c_e are determined through parameter optimization to match observed degradation patterns in archeological samples. This two-phase process - height calculation followed by simultaneous update - ensures that the corrosion simulation accurately reflects the physical degradation process observed in actual fragmented bamboo slips, where environmental factors affect the entire surface concurrently rather than sequentially. The algorithm transforms the initial fracture curves into realistic degraded fragments, completing the synthetic data generation pipeline.

To ensure our physical model accurately reflects real bamboo slip characteristics, we developed a parameter optimization strategy driven by comparison between real and synthetic data distributions. The parameters in our physical model have physical interpretations that can be loosely connected to real-world properties[70,71], such as bamboo fiber structure, material characteristics, and environmental degradation factors. While these parameters are not direct measurements of physical properties, they collectively contribute to modeling the complex fracture and corrosion processes observed in archeological

bamboo slip fragments[72,73]. We collect curves from real fragmented bamboo slips, which represent the final state after both fracture and corrosion processes, characterized by their y^{final} coordinates along the fragment boundary. These curves undergo dimensionality reduction using t-SNE to create a reference distribution. Concurrently, we generate synthetic data using randomly initialized parameters and apply the same dimensionality reduction process. Our synthetic fragments undergo the complete physical simulation pipeline: starting with initial fracture generation to obtain y^{init} coordinates, followed by corrosion simulation to produce final degraded profiles with y^{final} coordinates that can be directly compared with real archeological data. The optimization objective is to minimize the distributional difference between real and synthetic data, quantified by the silhouette coefficient in Eq. (9). Our genetic algorithm iteratively refines the parameter set by maintaining a population of parameter configurations $\mathcal{P} = \{p_1, \ldots, p_n\}$, where each p_i contains parameters for both fracture and corrosion models[17]. The visualization in Supplementary Fig. 8 confirms that optimizing these physical parameters directly improves the model's ability to identify correct fragment matches. The fitness function follows Eq. (10), which combines the silhouette score with a diversity term for balancing between distribution matching and population diversity. New parameter sets are generated through Eq. (11), where p_1 and p_2 are selected based on fitness scores, β is a random crossover weight, and ϵ represents mutation noise. The effectiveness of our parameter optimization approach is visually demonstrated in Supplementary Fig. 8, where the distribution of synthetic data (blue points) progressively aligns with the real data distribution (green points) through iterations. The corresponding heatmap of matching quality shows stronger correlation patterns as the silhouette coefficient improves, validating our parameter optimization strategy[74]. This optimization ensures that our synthetic fragments accurately capture both the initial fracture characteristics and the subsequent degradation patterns, enabling the deep learning model to learn features grounded in actual physical processes.

WisePanda architecture

To avoid interference from low-level visual features such as color variations and surface texture, our approach employs a two-stage preprocessing: digitization with uniform backgrounds and geometry-only curve extraction that eliminates color and texture information while preserving fracture boundary morphology. Each input to WisePanda consists of bamboo slip fragment curve data encoded as 64-dimensional vectors. We extract these vectors through a straightforward sampling process: after detecting and isolating the fracture curve from each fragment image, we normalize the curve length and sample 64 equidistant points along the curve from left to right. The resulting sequence of 64 heights captures the distinctive geometric profile of each fracture curve while providing a standardized input format for the following neural network. This fixed-dimensional representation ensures consistent processing across fragments of different sizes and complexities while preserving the key morphological features that distinguish potential matching pairs.

The core of WisePanda's architecture is a TripletNet-based deep learning network designed for similarity learning between fragment pairs. WisePanda's network employs a two-stage architecture (Supplementary Fig. 9): a Feature Encoder that transforms input curves into geometric feature embeddings, followed by a Distance Metric Network that computes learnable pairwise distances through additional metric layers. The network processes triplets during training: an anchor fragment c_x, a positive match c_x^+, and a negative non-match c_x^-. Each branch shares identical Feature Encoder weights to ensure consistent feature extraction. The Distance Metric Network receives any feature pair and outputs *dissimilarity score* as distance via sigmoid activation, where a value approaching 0 indicates a likely match and a value approaching 1 indicates a non-match. During inference, the network computes similarity scores as 1 − *dissimilarity score* for each candidate pair, enabling ranked matching recommendations. This learnable distance mechanism captures bamboo-specific fracture patterns more effectively than fixed distance metrics.

To facilitate archeological fragment rejoining, WisePanda implements a ranking-based matching system that provides archeologists with a manageable set of potential matches[75,76]. For each input fragment, the network computes similarity scores with all candidate fragments in the dataset. These scores are then sorted to generate a ranked list of the Top-k most probable matches, typically $k = 50$, significantly reducing the search space from thousands of possibilities to a practical subset for further manual verification. The ranking process is guided by the predicted similarity metric from our TripletNet network architecture. When a new fragment is presented, the system: (1) extracts the fracture curve and generates the 64-dimensional feature vector; (2) computes pair-wise similarity scores with existing fragments through the trained network; (3) ranks all candidates based on their similarity scores; and (4) returns the Top-k candidates with the highest matching probabilities. This approach transforms WisePanda from a theoretical matching model into a practical archeological tool that effectively guides the rejoining process while involving human expertise in the final verification steps.

Training details

WisePanda was trained using two NVIDIA RTX 4090 GPUs with a batch size of 100. The network was trained using the Adam optimizer with an initial learning rate of 1e−3 and a one-cycle learning rate scheduler[77]. We employed an extended triplet loss that optimizes distance metrics directly. Specifically, we implemented a combined loss function that extends the concept of triplet loss to a more direct optimization objective:

$$\mathcal{L} = E[(d_{pos} - 0)^2 + (d_{neg} - 1)^2] \tag{12}$$

where d_{pos} represents the distance between matching pairs (which we want to push toward 0) and d_{neg} represents the distance between non-matching pairs (which we want to push toward 1). Our extended triplet loss incorporates two key innovations: (1) a learnable Distance Metric Network that computes task-specific distances through metric layers, rather than using fixed distance measures like Euclidean distance, and (2) a squared error formulation that provides smoother gradients than standard max-margin triplet loss. Ablation studies (Supplementary Table 2) validate these design choices, showing our extended triplet loss substantially outperforms classic triplet loss (Top-50: 91.81% vs. 49.44%), contrastive loss (41.67%), and BCE loss (41.39%).

To address the inherent class imbalance problem in fragment matching, we employed a controlled sampling strategy: for each training iteration, we extracted one matched pair from our physics-driven synthetic data as anchor and positive samples, then selected a single fragment from the remaining pool as the negative sample. This controlled sampling effectively mitigates class imbalance without requiring complex reweighting schemes. We utilized PReLU activation functions for non-linear transformations, which adapt to the data by learning the optimal negative slope parameter rather than using a fixed value as in standard ReLU.

The input fragmented curves were normalized by rescaling both x and y^{final} coordinates to fall within the [0, 1] range before being sampled and fed into the network. The training process ran for 150 epochs to ensure convergence, with model parameters being updated approximately 100,000 times. To ensure reproducibility, model training was conducted with fixed random seeds. For inference, we employ an efficient ranking strategy where each fragment is compared with the entire database using the trained similarity metric. The matches are then sorted by matching probabilities to generate the Top-k recommendations for archeological verification.

Dataset details

The bamboo slips in this study originate from Tomb 77 at Shuihudi, Yunmeng County, Hubei Province, excavated in 2006. The tomb owner, Yue Ren, served as a local official in Anlu County. The slips document legal codes, administrative records, work logs, mathematical texts, and almanacs dating to the Western Han dynasty during Emperor Wen's reign (179–157 BCE). These bamboo slips fractured due to external forces, with the waterlogged soil environment causing mild corrosion through microbial activity and moisture-induced degradation of bamboo fibers. Following the accidental discovery in 2006, archeologists conducted scientific excavation and artifact recovery under strict cultural heritage protection regulations. Importantly, both Bamboo236 and Bamboo1350 consist of high-resolution digitized images of real excavated fragments rather than idealized simulations; these are the same digitized materials used by experts in current manual rejoining practice, since physical manipulation of original artifacts is restricted for conservation reasons.

Baseline methods

To evaluate WisePanda's effectiveness, we compare it against multiple baseline approaches spanning traditional curve matching methods and modern generative models. To ensure statistical reliability, all methods were evaluated with three independent runs, with results reported as mean ± standard deviation. For curve matching methods, three runs were conducted with different parameter configurations. For generative model-based methods and WisePanda, training runs were performed with different random seeds to control the matching network training process.

We implemented seven curve matching methods from 1994 to 2025. The traditional curve matching approach employs dynamic time warping (DTW)[20] to measure the geometric similarity between fragmented curves. This method directly compares the normalized (x, y^{final}) coordinate sequences of two fragments, characterizing their alignment cost as a measure of potential matching. SIS[22] transforms fragmented curves into scale-invariant representations by computing local curvature features along the fragmented curve, enabling comparison through simple distance metrics. FMM[21] accelerates the matching process by decomposing fragmented curves into key geometric features and employing a hierarchical matching strategy. Event-DTW[23] extends DTW by identifying and aligning event patterns within time series data. Drop-DTW[24] improves robustness through selective feature dropping during alignment. Modified Correlation Optimized Warping (Modified COW)[25] optimizes segment-based warping with correlation constraints. Two-stage DTW[26] employs hierarchical matching by first coarse alignment followed by fine-grained refinement.

We implemented four representative generative approaches: The original GAN[27] generates synthetic samples through adversarial training between generator and discriminator networks. SeriesGAN[28] adapts the GAN framework specifically for sequential data generation. Denoising Diffusion Probabilistic Model (DDPM)[29] generates samples through iterative denoising processes. Diffusion-TS[30] optimizes the conditional diffusion paradigm for time series and curve data generation.

Evaluation metrics

To assess WisePanda's performance in bamboo slip rejoining, we employ the Top-k accuracy metric, which measures the percentage of correctly identified matches within the k highest-ranked suggestions. For each query fragment, a match is considered successful if the true matching fragment appears among the Top-k candidates proposed by the model. We report results for $k \in 1, 5, 10, 20, 50, 100$, providing a comprehensive view of the model's ranking capability at different thresholds of consideration. This metric directly reflects the system's practical utility in archeological workflows, where experts typically examine a limited number of the most probable matches.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data supporting the findings of this study are available within the paper, its Supplementary Information, Code Ocean, and from the author upon request.

Code availability

The code for the WisePanda framework, including the physics engine, deep learning training, and comparative evaluation pipelines for bamboo slip fragment rejoining, is available on Code Ocean: https://codeocean.com/capsule/0259823/tree.

References

1. Allan, S. *Buried ideas: Legends of abdication and ideal government in early Chinese bamboo-slip manuscripts*. State University of New York Press, (2015).
2. Liu, G. *Approaching the Tsinghua Bamboo Slips*. Taylor & Francis, (2024).
3. Liu, X. From bamboo slips to received versions: Common features in the transformation of the" Laozi". *Harv. J. Asiat. Stud.* **63**, 337–382 (2003).
4. Johnson, M. *Archaeological theory: an introduction*. John Wiley & Sons, (2019).
5. Zhou, J. et al. Identifying designs from incomplete, fragmented cultural heritage objects by curve-pattern matching. *J. Electron. Imaging* **26**, 011022 (2017).
6. Shuihudi Qin Tomb Bamboo Slips Compilation Group, editor. *Shuihudi Qin Mu Zhujian (Shui hu di qin mu zhu jian)*. Wenwu Chubanshe (Cultural Relics Publishing House), Beijing, (1990). Title translates to: Bamboo Slips from the Qin Tomb at Shuihudi. [in Chinese].
7. Cha, M.Y., Lee, K.H. & Kim, Y.S. Micromorphological and chemical aspects of archaeological bamboos under long-term waterlogged condition. *Int. Biodeterior. Biodegrad.* **86**, 115–121 (2014).
8. Anderson, T. L. & Anderson, T. L. *Fracture mechanics: fundamentals and applications*. CRC Press, (2005).
9. Tan, T. et al. Mechanical properties of functionally graded hierarchical bamboo structures. *Acta Biomater.* **7**, 3796–3803 (2011).
10. Behera, D. et al. A review on bamboo fiber reinforced composites and their potential applications. *Emerg. Mater.* **8**, 157–172 (2025).
11. Chen, Z., Ma, R., Du, Y. & Wang, X. State-of-the-art review on research and application of original bamboo-based composite components in structural engineering. In *Structures*, volume 35, pages 1010–1029, (2022).
12. Dixon, P. G. & Gibson, L. J. The structure and mechanics of moso bamboo material. *J. R. Soc. Interface* **11**, 20140321 (2014).
13. Ramful, R. & Sakuma, A. Investigation of the effect of inhomogeneous material on the fracture mechanisms of bamboo by finite element method. *Materials* **13**, 5039 (2020).
14. Chen, M. et al. Fracture modes of bamboo fiber bundles in three-point bending. *Cellulose* **26**, 8101–8108 (2019).
15. Ramful, R. Investigation of the transverse fracture mechanisms of bamboo by the finite element method. *J. Mater. Sci.* **57**, 6233–6248 (2022).
16. Blanchette, R. A., Haight, J. E., Koestler, R. J., Hatchfield, P. B. & Arnold, D. Assessment of deterioration in archaeological wood from ancient Egypt. *J. Am. Inst. Conserv.* **33**, 55–70 (1994).
17. Goldberg, D. E. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, Reading, MA, (1989).
18. Van der Maaten, L. & Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* 9, (2008).
19. Hoffer, E. & Ailon, N. Deep metric learning using triplet network. In *Proc. Int. Worksh. on Simil.-Based Pattern Recognit.* 84–92, (2015).

20. Berndt, D. J. & Clifford, J. Using dynamic time warping to find patterns in time series. In *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, 359–370, (1994).
21. Frenkel, M. & Basri, R. Curve matching using the fast marching method. In *Proc. Int. Worksh. on Energy Minimiz. Methods in Comput. Vis. Pattern Recognit.* 35–51, (2003).
22. Calabi, E., Olver, P. J., Shakiban, C., Tannenbaum, A. & Haker, S. Differential and numerically invariant signature curves applied to object recognition. *Int. J. Comput. Vis.* **26**, 107–135 (1998).
23. Jiang, Y. et al. Eventdtw: An improved dynamic time warping algorithm for aligning biomedical signals of nonuniform sampling frequencies. *Sensors* **20**, 2700 (2020).
24. Dvornik, M., Hadji, I., Derpanis, K. G., Garg, A. & Jepson, A. Drop-dtw: Aligning common signal between sequences while dropping outliers. *Proc. Adv. Neural Inf. Process. Syst.* **34**, 13782–13793 (2021).
25. Khosravi, M., Soleimanmeigouni, I., Ahmadi, A., Nissen, A. & Xiao, X. Modification of correlation optimized warping method for position alignment of condition measurements of linear assets. *Measurement* **201**, 111707 (2022).
26. Sun, B., Li, H. & Wang, J. Two-stage dynamic time warping for intelligent fault detection of rotating machinery under variable speed. *Meas. Sci. Technol.* **36**, 056124 (2025).
27. Goodfellow, I. J. et al. Generative adversarial nets. *Proc. Adv. Neural Inf. Process. Syst.* **27**, (2014).
28. Eskandari N., Mohammad R., Hamdi, S.M., & Boubrahimi, S.F. SeriesGAN: Time series generation via adversarial and autoregressive learning. In *IEEE International Conference on Big Data*, 860–869, (2024).
29. Ho, J., Jain, A. & Abbeel, P. Denoising diffusion probabilistic models. *Proc. Adv. Neural Inf. Process. Syst.* **33**, 6840–6851 (2020).
30. Yuan, X. and Qiao, Y. Diffusion-TS: Interpretable diffusion for general time series generation. In *Int. Conf. Learn. Represent* (2024).
31. Li, M. -Y. et al. Investigation into the deterioration process of archaeological bamboo strips of China from four different periods by chemical and anatomical analysis. *Polym. Degrad. Stab.* **109**, 71–78 (2014).
32. Eslami, D., Di Angelo, L., Di Stefano, P. & Pane, C. Review of computer-based methods for archaeological ceramic sherds reconstruction. *Virtual Archaeol. Rev.* **11**, 34–49 (2020).
33. Hou, M. et al. Novel method for virtual restoration of cultural relics with complex geometric structure based on multiscale spatial geometry. *ISPRS Int. J. Geo-Inf.* **7**, 353 (2018).
34. Amada, S., Ichikawa, Y., Munekata, T., Nagase, Y. & Shimizu, H. Fiber texture and mechanical graded structure of bamboo. *Compos. B: Eng.* **28**, 13–20 (1997).
35. Low, I. M., Che, Z. Y. & Latella, B. A. Mapping the structure, composition and mechanical properties of bamboo. *J. Mater. Res.* **21**, 1969–1976 (2006).
36. Staack, T. Reconstruction of early Chinese bamboo and wood manuscripts: A review (1900–2010). *CSMC Occas. Pap.* **5**, (2016).
37. Haiyan, T. Damage of bamboo and wooden materials based on linear elastic fracture mechanics in garden design. *Fract. Struct. Integr.* **10**, 472–480 (2016).
38. Singh, A. P., Kim, J. ongS. ik, Möller, R., Chavan, R. R. & Kim, Y. oonS. oo The pivotal role of microscopy in unravelling the nature of microbial deterioration of waterlogged wood: A review. *Forests* **15**, 889 (2024).
39. Zhang, Y., Li, K., Chen, X., Zhang, S. & Geng, G. A multi feature fusion method for reassembly of 3d cultural heritage artifacts. *J. Cult. Herit.* **33**, 191–200 (2018).
40. Das, B.R., Maringanti, H.B., & Dash, N.S. Role of artificial intelligence in preservation of culture and heritage. In *Digital. Cult. Technol.*, pages 92–97. Routledge, (2022).
41. Tsatsanashvili, A. Artificial intelligence in the protection of intangible cultural heritage. *Eur. J. Turk. Stud.* **12**, 163–178 (2024).
42. Sommerschield, T. et al. Machine learning for ancient languages: A survey. *Comput. Linguist.* **49**, 703–747 (2023).
43. Tyndall, S. Toward automatically assembling Hittite-language cuneiform tablet fragments into larger texts. In *Proc. of ACL*, pages 243–247, (2012).
44. Collins, T. et al. Computer-assisted reconstruction of virtual fragmented cuneiform tablets. In *Proc. VSMM*, pages 70–77, (2014).
45. Tsesmelis, T. et al. Re-assembling the past: The repair dataset and benchmark for real world 2d and 3d puzzle solving. *Proc. Adv. Neural Inf. Process. Syst.* **37**, 30076–30105 (2024).
46. Assael, Y. et al. Restoring and attributing ancient texts using deep neural networks. *Nature* **603**, 280–283 (2022).
47. Assael, Y. et al. Contextualizing ancient texts with generative neural networks. *Nature*, 1–7, (2025).
48. Zheng, Y., Li, X. & Weng, Y. Reunion helper: an edge matcher for sibling fragment identification of the Dunhuang manuscript. *Herit. Sci.* **12**, 52 (2024).
49. Zhang, Y. et al. Llmco4mr: Llms-aided neural combinatorial optimization for ancient manuscript restoration from fragments with case studies on Dunhuang. In *Proc. Eur. Conf. on Comput. Vis.*, 253–269, (2024).
50. Pirrone, A., Aimar, Marie Beurton, and Journet, N. Papy-s-net: A siamese network to match papyrus fragments. In *Proc. of HIP*, 78–83, (2019).
51. Abitbol, R., Shimshoni, I. & Ben-Dov, J. Machine learning based assembly of fragments of ancient papyrus. *J. Comput. Cult. Herit.* **14**, 1–21 (2021).
52. Zhang, C. et al. Data-driven oracle bone rejoining: A dataset and practical self-supervised learning scheme. In *Proc. ACM SIGKDD Int. Conf. Knowledge Discovery Data Mining*, pages 4482–4492, (2022).
53. Staack, T. Bindings of ancient Chinese bamboo and wood scrolls. *Stud. Manusc. Cult.*, page 39, (2023).
54. Chang, Chun-shu. Written on bamboo and silk: The beginning of Chinese books and inscriptions, (1964).
55. Needham, J. *Science and Civilisation in China: Volume 5, Chemistry and Chemical Technology, Part 5, Spagyrical Discovery and Invention: Physiological Alchemy*, volume 5. Cambridge University Press, (1983).
56. Luo, Z. & Wang, G.*Liusha Zhuijian (Liu sha zhui jian)*. Shangyu Luo shi yin xing, (1914).
57. Yu, Y., Tian, G., Wang, H., Fei, B. & Wang, G. Mechanical characterization of single bamboo fibers with nanoindentation and microtensile technique. *Holzforschung*, **65**, (2011).
58. Knott, J. F. *Fundamentals of fracture mechanics*. Gruppo Italiano Frattura (1973).
59. Silling, S. A. Reformulation of elasticity theory for discontinuities and long-range forces. *J. Mech. Phys. Solids* **48**, 175–209 (2000).
60. Raissi, M., Perdikaris, P. & Karniadakis, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **378**, 686–707 (2019).
61. Raissi, M., Yazdani, A. & Karniadakis, G. E. Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations. *Science* **367**, 1026–1030 (2020).
62. Brown, B. J. et al. A system for high-volume acquisition and matching of fresco fragments: Reassembling theran wall paintings. *ACM Trans. Graph.* **27**, 1–9 (2008).
63. Ghaith, K. & Hutson, J. A qualitative study on the integration of artificial intelligence in cultural heritage conservation. *Metaverse*, **5**, (2024).
64. Aldakheel, F., Elsayed, E. S., Heider, Y. & Weeger, O. Physics-based machine learning for computational fracture mechanics. *Mach. Learn. Comput. Sci. Eng.* **1**, 18 (2025).
65. LeSar, R. *Introduction to computational materials science: fundamentals to applications*. Cambridge University Press, (2013).

66. Huang, Y. & Fei, B. Comparison of the mechanical characteristics of fibers and cell walls from moso bamboo and wood. *BioResources* **12**, 8230–8239 (2017).
67. Gdoutos, E. E. *Fracture mechanics: an introduction*, volume 263. Springer Nature (2020).
68. Forest Products Laboratory (US). *Wood handbook: wood as an engineering material*. Number 72. The Laboratory (1987).
69. Liu, L. et al. Cracks repairing and resistance to water penetration properties of microbial self-healing cement. *Eng. Life. Sci.* **25**, e70010 (2025).
70. Hussain, A., Sakhaei, A.H., & Shafiee, M. Machine learning-based constitutive modelling for material non-linearity: A review. *Mech. Adv. Mater. Struct*. pages 1–19, (2024).
71. Kutz, J. N. Machine learning for parameter estimation. *Proc. Natl. Acad. Sci. USA* **120**, e2300990120 (2023).
72. Brynjarsdottir, J. & O'Hagan, A. Learning about physical parameters: The importance of model discrepancy. *Inverse Probl.* **30**, 114007 (2014).
73. Matthiesen, H. Detecting and quantifying ongoing decay of organic archaeological remains: a discussion of different approaches. *Quat. Int.* **368**, 43–50 (2015).
74. Rousseeuw, P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **20**, 53–65 (1987).
75. Funkhouser, T. et al. Learning how to match fresco fragments. *ACM J. Comput. Cult. Herit.* **4**, 1–13 (2011).
76. Di Angelo, L., Di Stefano, P. & Guardiani, E. A review of computer-based methods for classification and reconstruction of 3D high-density scanned archaeological pottery. *J. Cult. Herit.* **56**, 10–24 (2022).
77. Smith, L. N. & Topin, N. Super-convergence: Very fast training of neural networks using large learning rates. In *AI ML Multi-Domain Oper. Appl*. 11006, 369–386, (2019).

Acknowledgements

This work is supported in part by NSFC under Grant 62225113 (B.D.), Grant 62325111 (G.X.), Grant 62222112 (Y.X.), and Grant 62176186 (Y.X.); and in part by the National Social Science Fund of China under grant 16ZDA115 (H.L., J.L.L., and J.L.) and 21&ZD334 (X.W.), the Wuhan University Education & Development Foundation Lei Jun Education Fund (Y.X.), the WHU-Kingsoft Joint Lab (B.D.), the Innovative Research Group Project of Hubei Province under Grant 2024AFA017 (B.D.), and the New Cornerstone Science Foundation through the XPLORER PRIZE (B.D.).

Author contributions

Y.X. conceived the study. J.Z. and Z.Z performed experiments, analyzed data, wrote the code, and manuscript. H.L., X.W., J.L.L., and J.L. provided archeological data and domain knowledge. J.Z., Q.T., and J.S. contributed to the literature review and software implementation. G.X. participated in methodology discussions, data analysis, and manuscript organization. B.D. participated in methodology discussions, data analysis, and manuscript editing. Y.X. participated in methodology discussions, experimental designs, data analysis, manuscript organization, and revision.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-026-70361-y.

Correspondence and requests for materials should be addressed to Gui-Song Xia, Bo Du or Yongchao Xu.

Peer review information : *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.