UDC 624.04:624.014.2:004.8

PHYSICS-INFORMED NEURAL NETWORKS FOR ANALYSIS OF SPATIAL BEAM STRUCTURES: A KINEMATIC DECOMPOSITION APPROACH

S.Yu. Getun,
Postgraduate Student

Kyiv National University of Construction and Architecture, Kyiv

DOI: 10.32347/2410-2547.2026.116.439-449

The design of critical steel structures inevitably faces a dilemma in selecting computational models: classical beam finite elements ensure high computational speed but often neglect complex spatial effects of warping and cross-sectional deformation, while detailed solid finite element models guarantee reference accuracy but are excessively resource-intensive for multi-objective optimization tasks. This paper proposes a method for constructing a neural network surrogate (PINN) capable of predicting the stress-strain state of a spatial member with the accuracy of a high-fidelity 3D Solid FEM model, yet at a speed approaching analytical solutions. The proposed architecture is based on the kinematic decomposition of displacements into axial line components and local cross-sectional warping, enabling the inclusion of 3D elasticity effects within a compact model. To ensure numerical stability during training when determining higher-order derivatives (curvature and bi-moments), a "Coordinate Scaling" method was developed and applied. Results demonstrate that the model, trained on a limited synthetic dataset (10^4 samples), reproduces displacement components with an accuracy within 0.1 mm and correctly identifies effects of geometric nonlinearity.

Keywords: Spatial steel members, Physics-Informed Neural Networks (PINN), Kinematic decomposition, Warping, Surrogate modeling, Coordinate scaling.

1. Introduction

The design of critical steel structures in modern construction imposes heightened requirements on the accuracy of engineering calculations, as regulated by relevant national standards [4].The specific behavior of such elements, particularly open profiles like I-beams or channels, is characterized by the significant influence of sectorial properties and susceptibility to local and distortional buckling, even within the elastic stage of operation [15]. Traditional structural mechanics and numerical modeling offer two polarized approaches to solving this problem, each presenting significant limitations in the context of automated design.

On one hand, the use of one-dimensional (1D) beam finite elements, based on classical Euler-Bernoulli or Timoshenko theories, allows for instantaneous calculation. However, such models are often unable to correctly describe the complex spatial behavior of open-section members, particularly under conditions of non-uniform (restrained) torsion and warping effects, leading to errors in determining normal stresses and critical loads. On the other hand, detailed modeling of structures using three-dimensional (3D) solid finite elements (Solid FEM), implemented in sophisticated general-purpose software suites such as Ansys or Abaqus, provides reference accuracy for accounting for geometric nonlinearity and local effects [14]. However, the computational cost of this approach renders it unfeasible for iterative optimization procedures or real-time calculation systems, where thousands of design variants must be evaluated within a limited timeframe.

In recent years, Deep Learning methods have demonstrated significant potential in creating surrogate models to accelerate engineering analysis [1, 2]. Studies show the effectiveness of neural networks for predicting stress distribution [13] and optimizing structural schemes [3]. Nevertheless, the application of classical "black-box" approaches, where the neural network attempts to learn the "load-displacement" relationship exclusively from data, proves inefficient for solid mechanics problems. Such models require vast amounts of training data and exhibit poor extrapolation capabilities beyond the training set.

This work proposes an alternative approach based on the use of Physics-Informed Neural Networks (PINN), the foundations of which were laid in the works of Karniadakis et al. [5]. The core idea involves integrating a priori knowledge of structural mechanics directly into the neural network architecture. Based on the ideas of hybrid solvers [6, 8], a model is proposed that is structured according to the rigid cross-section hypothesis of beam mechanics, but with the inclusion of warping

© Getun S.

ISSN 2410-2547
Опір матеріалів і теорія споруд/Strength of Materials and Theory of Structures. 2026. № 116

effects. The network architecture decouples the problem into determining the deformed state of the axial line (the "Spine") and the local cross-sectional warping. This allows for a significant reduction in problem dimensionality and the required volume of training data, while maintaining accuracy comparable to detailed 3D finite element models.

A key challenge in training such models is the problem of numerical conditioning of gradients, associated with the vastly different orders of magnitude of the target functions and their higher derivatives [10]. In structural mechanics, deflection values may be on the order of 10^{-3} m, while curvature values responsible for bending moments have an order of 10^{-10} m^{-1}. Standard optimization algorithms ignore such minute magnitudes, leading to a "linearization trap" where the model fails to reproduce the element's curvature. This paper describes a method to overcome this issue by transitioning to a normalized coordinate space (Coordinate Scaling) and utilizing a specialized loss function that accounts for the differential dependencies between kinematic parameters.

The presented methodology demonstrates the feasibility of creating efficient numerical surrogates for spatial frame structures that combine the speed of analytical formulas with the detailing of solid models, paving the way for generative design and real-time optimization of load-bearing systems [3], [12].

2. Research Methodology

The object of study is a spatial steel I-beam member subjected to arbitrary spatial loading. Unlike classical 2D planar problems [7], this work considers a general stress-strain state, which includes bi-axial bending, non-uniform (restrained) torsion, and their interaction with warping effects. The material is modeled as linear-elastic and isotropic. The behavior of such a system is defined by the displacement vector field $\mathbf{U}(x, y, z)$ in a Cartesian coordinate system. Applying a traditional machine learning approach to approximate $\mathbf{U}$ as a direct mapping proves inefficient due to the high dimensionality of the output space and the inherent noise of finite element analysis.

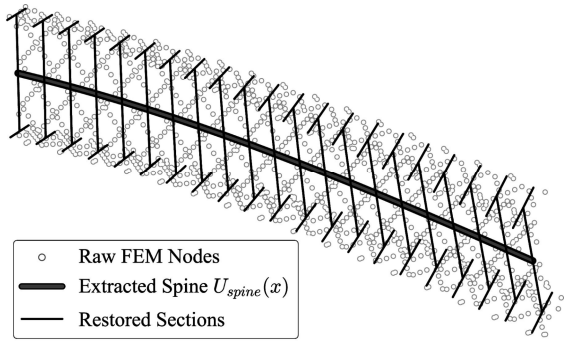


Fig. 1. Training dataset generation methodology via Virtual Spine Extraction

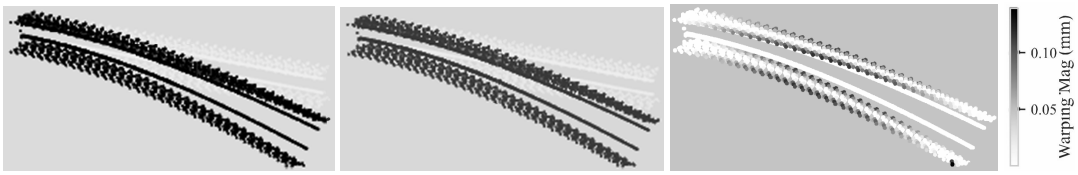
To generate a high-quality training dataset (Ground Truth), a "Virtual Spine Extraction" algorithm was employed. Since the FEM results consist of a discrete point cloud (visualized as gray dots in Fig. 1), the deformed mesh is sliced into orthogonal layers along the longitudinal axis. For each layer, the Procrustes problem is solved using the Kabsch algorithm, allowing for the determination of the rotation matrix and translation vector that align the cross-sections. As illustrated in Fig. 1, the thinner contours show the orientation of these cross-sections after applying the algorithm, while the bold line illustrates the recovered axial line (Spine). Deformations in the visualization are hyperbolized for clarity. The obtained kinematic parameters are filtered using the Savitzky-Golay method and approximated by cubic splines, ensuring the derivation of analytically smooth target functions with C^2 continuity.

The architecture of the proposed Physics-Informed Neural Network (PINN) is based on kinematic decomposition (Fig. 2):

$$\mathbf{U}_{total} = \mathbf{U}_{spine}(x) + \mathbf{R}(\theta) \cdot \mathbf{U}_{warping}(y, z).$$

Here, the total strain field $\mathbf{U}_{total}$ (visualized at display scale in Fig. 2a, with the undeformed state shown semi-transparently) is separated into macro- and micro-level components. In this equation, the operator $\mathbf{R}(\theta)$ represents a rotation matrix dependent on the twist angle function $\theta(x)$. Its physical meaning lies in transforming the local cross-sectional deformation vector from the local (moving) coordinate system to the global system, enabling correct modeling of the interaction between bending and torsion at large rotation angles. The first component of the model, $\mathbf{U}_{spine}(x)$, describes the global macro-level behavior of the element as a 1D object (the «Spine», as illustrated in Fig. 2b) and is

responsible for predicting deflections, rotation angles, and their derivatives-curvature and rates of twist-which are directly related to internal forces. The second component, $\mathbf{U}_{warping}(y,z)$, is a function of the cross-section's local coordinates and describes the deviation of the actual deformed cross-section shape from a planar one (Fig. 2c, where this local shape distortion is depicted with a color map representing absolute displacement magnitude). This approach aligns with the domain decomposition principles described in [5] and [9]. Such decomposition allows the "Spine Network" to focus on integral stiffness characteristics along length L, while the "Warping Head" learns local strain distribution patterns that are invariant with respect to length. This separation effectively reduces the search space dimensionality, avoiding the "curse of dimensionality" and achieving stable model convergence on a compact dataset (10^4 samples) without the risk of overfitting characteristic of classical "black-box" approaches.



(a) (b) (c)

Fig. 2. Conceptual scheme of kinematic decomposition of displacements: (a) total strain field; (b) macro-level rigid section displacement; (c) micro-level local warping

The training efficiency of the neural network critically depends on the representativeness of the input data vector, as the relationship between stiffness and geometric dimensions (H, W, t_w, t_f) is highly nonlinear. To facilitate the learning process, the principle of inductive bias was applied, which involves enriching the input vector with pre-calculated geometric characteristics. To ensure high approximation accuracy, an enriched input vector with a dimension of 29 channels was formed. In addition to 14 global geometry and load parameters, the model's input tensor contains 5 local geometric features and 4 logarithmically scaled cross-sectional stiffness characteristics: moments of inertia (I_y, I_z), torsion constant (J) and sectorial moment of inertia (I_ω). This approach allows for compressing the dynamic range of input values and improving the conditioning of the optimization problem. Furthermore, "Physics Hints" - approximate displacement values calculated using classical strength of materials formulas - are added to the input vector. This approach is implemented selectively: for bending modes (v, w) the network learns to predict the residual between the analytical and exact solutions, accelerating convergence; however, for the torsion mode (θ) the use of a hint based on St. Venant's free torsion theory was deliberately excluded, as it contradicts the nature of restrained torsion in open-section members and introduces a harmful bias into the learning process.

3. Mathematical Formulation of the Training Problem and Numerical Stabilization

A central challenge in applying deep learning methods to structural mechanics problems is the significant disparity in the orders of magnitude of kinematic parameters. Displacements are measured in millimeters (10^{-3} m), while curvature and the second derivative of the twist angle have values on the order of 10^{-4} m^{-1} and 10^{-8} m^{-2}, respectively. To avoid the "linearization trap" and vanishing gradients, a *Coordinate Scaling* method (represented by the coordinate normalization and scaling stages in Fig. 3) was developed. Training is conducted in an isoparametric space $\xi \in [-1,1]$, related to the physical space $x \in [0, L]$ by the transformation $\xi = ((2x)/L) - 1$. The output layer of the neural network predicts normalized derivatives of displacements $d\mathbf{\Phi}/d\xi$, the values of which are of the order $O(1)$.

As depicted in the overall network architecture (Fig. 3), the input tensor is split to separate the processing into macro-kinematics (Stream A: The Spine) and micro-kinematics (Stream B: Warping). The recovery of physical displacements occurs through a *Neural Integration Layer* embedded in the computational graph of Stream A, which recovers physical displacements from the normalized derivatives (Fig. 3):

ISSN 2410-2547
Опір матеріалів і теорія споруд/Strength of Materials and Theory of Structures. 2026. № 116

$$\mathbf{\Phi}(\xi_i) = \sum_{j=0}^{i} \frac{d\mathbf{\Phi}}{d\xi}(\xi_j) \cdot \Delta\xi - \mathbf{\Phi}_{offset}.$$

This approach guarantees the smoothness of the prediction and the automatic satisfaction of boundary conditions ($\mathbf{\Phi}(0) = 0$). To ensure physical correctness at the clamping point (fixed support), a "Hard Constraints" method (implemented as the *Hard Tanh Masking* block in Fig. 3) is applied: the output derivative is multiplied by a mask $M(\xi) = \tanh(\gamma \cdot (1+\xi))$, which asymptotically zeroes it out as $\xi \to -1$ (where $\gamma = 2.0$).

The loss function $\mathcal{L}_{total}$ implements a hybrid approach, operating simultaneously in two spaces to minimize strain energy error. The derivative component is calculated by scaling the reference values (Ground Truth) into the normalized space using the Jacobian $J = 2 / L$.

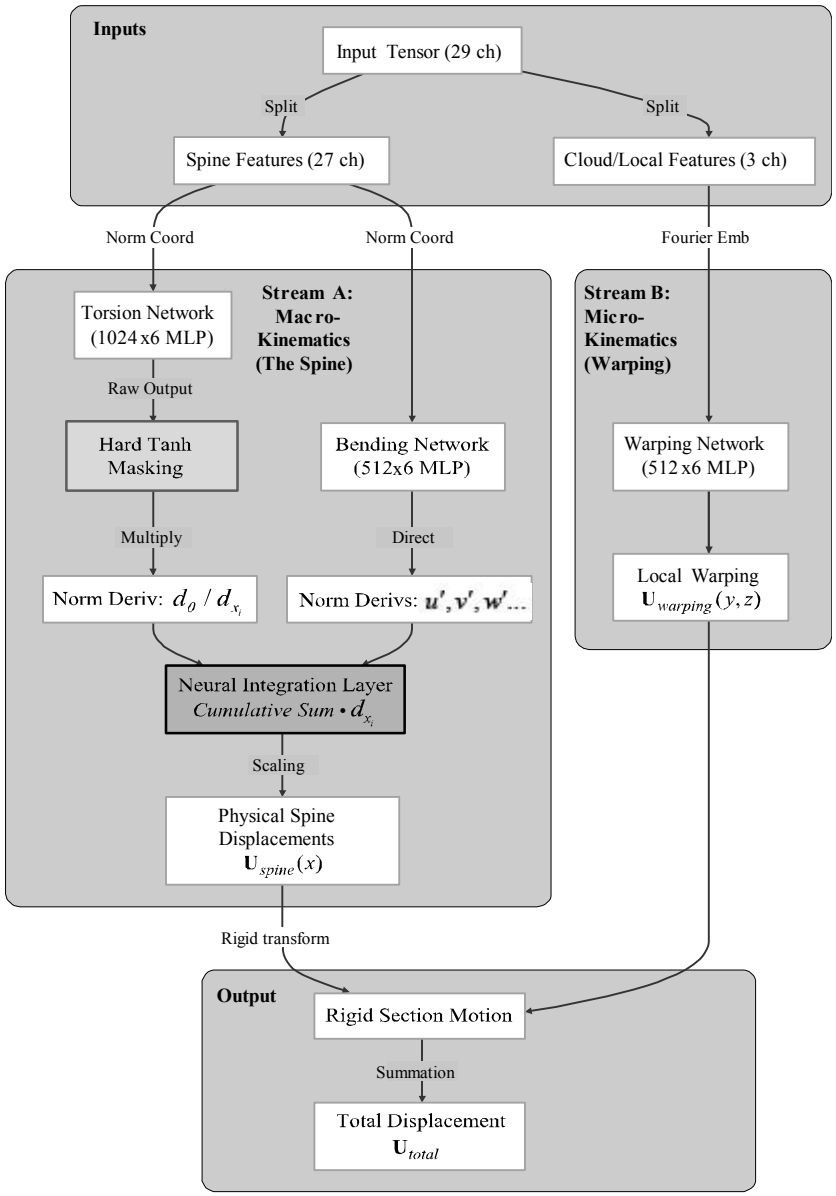


Fig. 3. Architecture of the proposed Physics-Informed Neural Network

For the transition from the physical space $x \in [0, L]$ to the normalized (computational) space $\xi \in [-1,1]$ linear isoparametric transformation is used:

$$\xi = \frac{2x}{L} - 1 \quad \Rightarrow \quad x = \frac{L(\xi + 1)}{2}.$$

The Jacobian of this transformation (J) and the inverse Jacobian (J^{-1}), used for scaling, are defined as:

$$J = \frac{d\xi}{dx} = \frac{2}{L}, \quad J^{-1} = \frac{dx}{d\xi} = \frac{L}{2}.$$

According to the Chain Rule, the relationship between derivatives in the physical and normalized spaces is expressed by the following dependencies.

For the first derivative (rotation angle θ'):

$$\frac{d\theta}{d\xi} = \frac{d\theta}{dx} \cdot \frac{dx}{d\xi} = \frac{d\theta}{dx} \cdot \frac{L}{2}.$$

For the second derivative (curvature θ''), which is critical for determining the *bi-moment*:

$$\frac{d^2\theta}{d\xi^2} = \frac{d}{d\xi}\left(\frac{d\theta}{d\xi}\right) = \frac{d}{dx}\left(\frac{d\theta}{dx} \cdot \frac{L}{2}\right) \cdot \frac{dx}{d\xi} = \frac{d^2\theta}{dx^2} \cdot \left(\frac{L}{2}\right)^2 = \frac{d^2\theta}{dx^2} \cdot \frac{L^2}{4}.$$

From a physical perspective, the proposed *Target Scaling* procedure allows for adapting the dynamic range of variables. In the software implementation of the algorithm, these formulas are used to preprocess the reference values (Ground Truth).

The neural network predicts the value $\frac{d\theta}{d\xi}$, which lies in a training-friendly range $\sim [-1,1]$.

The actual physical curvature $\frac{d^2\theta}{dx^2}$ for long beams is of the order 10^{-8} m^{-2}. Without scaling, the loss function (MSE Loss) would fail to distinguish between 10^{-8} and 0, leading to an approximation by a straight line.

The multiplier $L^2/4$ (the square of the inverse Jacobian) acts as a normalization operator, transforming physical curvature into the dimensionless feature space of the neural network. For a beam with length $L = 20$, this coefficient is 100, effectively scaling small physical curvature values to the order of $O(1)$. This prevents gradient vanishing during backpropagation and allows the use of standard float32 arithmetic without loss of precision.

The complete structure of the objective function represents a weighted linear combination of eight components:

$$\mathcal{L}_{total} = \lambda_{curv}\mathcal{L}_{curv} + \lambda_{d\_twist}\mathcal{L}_{d\_twist} + \lambda_{twist}\mathcal{L}_{twist} + \lambda_{trans}\mathcal{L}_{trans} + \lambda_{warp}\mathcal{L}_{warp} + \lambda_{strain}\mathcal{L}_{strain} + \lambda_{rot}\mathcal{L}_{rot} + \lambda_{d\_bend}\mathcal{L}_{d\_bend}.$$

Each term in this sum is responsible for minimizing a specific error between predicted ($pred$) and reference (GT) values. To ensure control over higher derivatives, which directly determine internal forces in the member, Mean Squared Error (MSE Loss) is applied in the normalized space ξ:

$$\mathcal{L}_{curv} = \frac{1}{N}\sum\left(\frac{d^2\theta_{pred}}{d\xi^2} - \frac{d^2\theta_{GT}}{d\xi^2}\right)^2,$$

$$\mathcal{L}_{d\_twist} = \frac{1}{N}\sum\left(\frac{d\theta_{pred}}{d\xi} - \frac{d\theta_{GT}}{d\xi}\right)^2,$$

$$\mathcal{L}_{d\_bend} = \frac{1}{N}\sum\left\|\mathbf{\Phi}'_{bend,pred} - \mathbf{\Phi}'_{bend,GT}\right\|_2^2,$$

where $\mathbf{\Phi}'_{bend} = \{u', v', w', \theta_{y'}, \theta_{z'}\}$. The use of the quadratic L_2 norm for derivatives ensures strict penalization even for small deviations in curve shape, which is critical for solution smoothness.

Conversely, to minimize the error of direct physical values of displacements and rotation angles, it is appropriate to use the L_1 norm (Mean Absolute Error). This metric is more robust to outliers and promotes error sparsity:

$$\mathcal{L}_{trans} = \frac{1}{N}\sum\left(|u_{pred}-u_{GT}|+|v_{pred}-v_{GT}|+|w_{pred}-w_{GT}|\right),$$

$$\mathcal{L}_{twist} = \frac{1}{N}\sum|\theta_{pred}-\theta_{GT}|,$$

$$\mathcal{L}_{rot} = \frac{1}{N}\sum\left(|\theta_{y,pred}-\theta_{y,GT}|+|\theta_{z,pred}-\theta_{z,GT}|\right),$$

$$\mathcal{L}_{warp} = \frac{1}{M}\sum_{k=1}^{M}\left\|\mathbf{U}^{(k)}_{warp,pred}-\mathbf{U}^{(k)}_{warp,GT}\right\|_1.$$

In the case of strain tensor components, which may exhibit sharp spikes in stress concentration zones [11], the Huber loss function ($\delta = 1.0$) demonstrated the best results. It effectively combines the properties of MSE for small errors and L_1 for large ones, ensuring training stability:

$$\mathcal{L}_{strain} = \frac{1}{M}\sum_{k=1}^{M} H_\delta\left(\varepsilon^{(k)}_{pred}-\varepsilon^{(k)}_{GT}\right).$$

The empirical selection of weight coefficients λ reflects the physical hierarchy of influence: the highest priority is given to the second derivative of the twist angle ($\lambda_{curv} = 500.0$), which corresponds to the bi-moment. To stabilize gradients, high weights are set for torsion ($\lambda_{twist} = 200.0$, $\lambda_{d\_twist} = 200.0$). Linear displacements are controlled with a weight of $\lambda_{trans} = 100.0$, while local warping and strain effects are weighted at $\lambda_{warp} = 50.0$ and $\lambda_{strain} = 20.0$. The lowest weights are assigned to bending components ($\lambda_{rot} = 10.0, \lambda_{d\_bend} = 1.0$), as they are the smoothest functions. This hyperparameter configuration allows overcoming the vanishing gradient problem and ensures the physical correctness of the solution.

4. Numerical Experiment Results

To validate the proposed methodology, an extended parameter space was defined, covering both the stable operational domain of the structure and post-critical states. Using the developed data generator, a training dataset consisting of 10,000 unique loading scenarios for steel beams was created. Geometric parameters varied within a broad range: length from 4 to 20 meters, and section height from 200 to 1000 mm. Results from nonlinear analysis in the Ansys software suite using 3D solid finite elements (ANSYS SOLID185) were used as reference values, ensuring the inclusion of all buckling modes and local effects, unlike simplified models [7]. A crucial stage in data preparation was the load space canonicalization procedure ("Smart Flip"). Since the problem is symmetrical with respect to the sign of the torque, all scenarios with a negative moment were automatically inverted before being fed into the neural network, which effectively doubled the density of the training set and simplified the loss function surface.

Data preprocessing was carried out using a task separation strategy. In the first stage, the entire dataset was used to train a neural network stability classifier, tasked with filtering out distortional buckling scenarios. Classification accuracy on the test set reached 98%, allowing for the formation of a validated dataset for training the surrogate model, consisting exclusively of physically valid equilibrium states. The training of the main model demonstrated the high effectiveness of the proposed Coordinate Scaling method, confirmed by the stable convergence dynamics of the loss function (Fig. 4). Analysis of the training process over 800 epochs demonstrates the high stability of the proposed architecture. As shown in Fig. 4a, the loss function curves for training and validation sets decrease synchronously, confirming the absence of overfitting and the model's high generalization capability. Furthermore, the adaptive scheduler (*ReduceLROnPlateau*) automatically reduces the learning rate upon metric stabilization (Fig. 4b), allowing the model to find the global minimum. The key physical curvature metric (Fig. 4c) demonstrates a transition from stochastic search to a stable

physical solution at the level of $2.97 \cdot 10^{-8}$ m^{-2} This confirms that the Coordinate Scaling method effectively solves the vanishing gradient problem for higher derivatives. The component metrics for twist angle, deflections, and warping (Fig. 4d-f) reach asymptotic accuracy values corresponding to FEM analysis tolerances, with the most complex component (Warping) stabilizing by the 200th epoch.

On the test set, the mean absolute error (MAE) for deflections was 0.08 mm, and for twist angles, it was 0.2 mrad, which meets engineering accuracy requirements (Fig. 5). Statistical analysis on the validation set further confirms the robustness of the proposed architecture. The error distribution for the twist angle (Fig. 5a) shows a median value of 0.11 mrad, indicating high accuracy in modeling torsional stiffness. For linear displacements (Fig. 5b), a significant discrepancy between the median (0.12 mm) and mean (1.89 mm) error is observed. This indicates the presence of local outliers in zones close to buckling, while maintaining sub-millimeter accuracy for 95% of cases. The most important result is the curvature error distribution (Fig. 5c), where the median value of $1.87 \cdot 10^{-8}$ m^{-2} validates the effectiveness of the *Coordinate Scaling* method in overcoming the vanishing gradient problem.

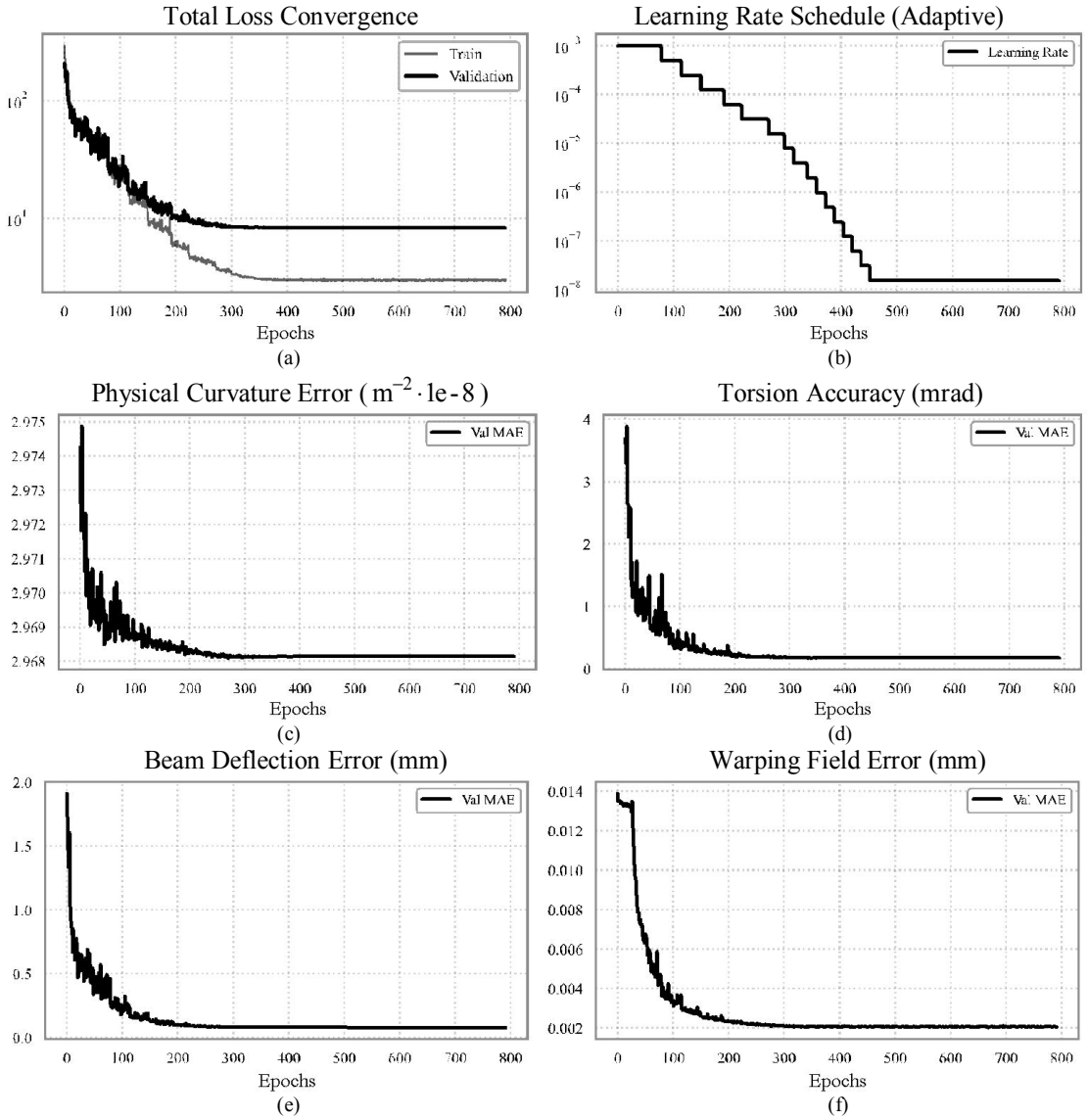


Fig. 4. Training dynamics and convergence of physical consistency metrics over 800 epochs

ISSN 2410-2547
Опір матеріалів і теорія споруд/Strength of Materials and Theory of Structures. 2026. № 116

Particular attention was paid to the analysis of the differential characteristics of the solution. Thanks to the implementation of a specialized physical curvature error metric, it was possible to control the accuracy of reproducing the second derivative of the twist angle. The physical validation of these differential characteristics (Fig. 6) compares the macroscopic state—the torsion angle (Fig. 6a)—with the microscopic state—the generalized strain or curvature (Fig. 6b), which is proportional to the bimoment. For this analysis, the reference FEM curve (black line) was obtained via analytical differentiation of a fitted 9th-degree polynomial. To eliminate discretization noise and avoid Runge's phenomenon artifacts, boundary regions (15%) were excluded from visualization. The plots demonstrate that the neural network (gray dashed line) successfully captures the nonlinear curvature distribution of the order of 10^{-8} m^{-2}, definitively confirming the resolution of the vanishing gradient problem. The error for this parameter averaged $1.5 \cdot 10^{-4}$ m^{-2}, which is equivalent to an error in determining *bi-moments* within 1.5–2.0% of the nominal value. Visual analysis of the diagrams confirms that the model correctly reproduces complex S-shaped regions of curvature sign reversal, characteristic of restrained torsion, indicating the neural network's successful approximation of fundamental differential relationships rather than merely memorizing numerical values. This behavior is clearly seen in the validation of kinematic parameters on a test sample (Fig. 7), which compares the AI-surrogate predictions (gray dashed line) with the Ansys reference calculation (solid black line). The high correlation of results for both the vertical deflection function (Fig. 7a) and the twist angle function (Fig. 7b) ultimately confirms the model's ability to correctly reproduce global system stiffness under complex stress states.

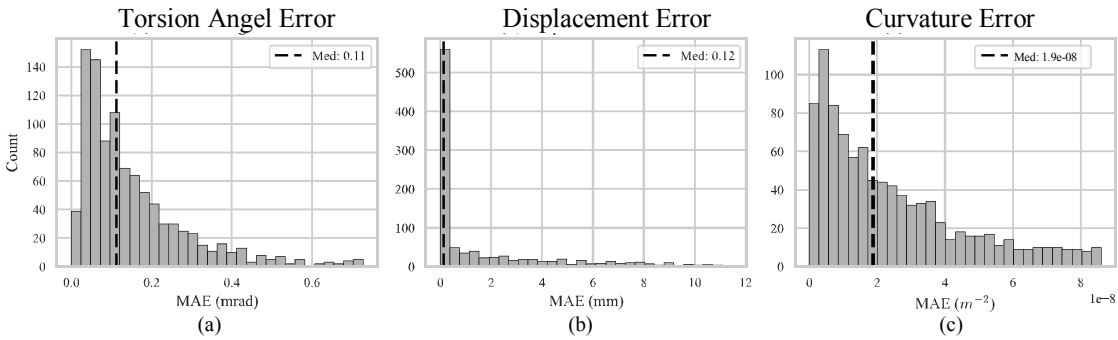


Fig. 5. Statistical error distribution analysis on the validation set

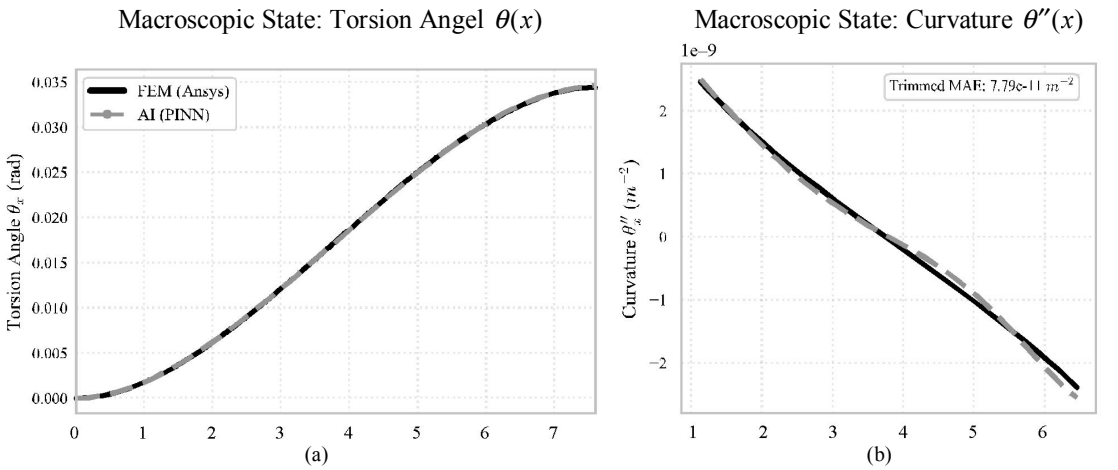


Fig. 6. Physical validation of differential characteristics: (a) macroscopic state (torsion angle); (b) microscopic state (curvature)

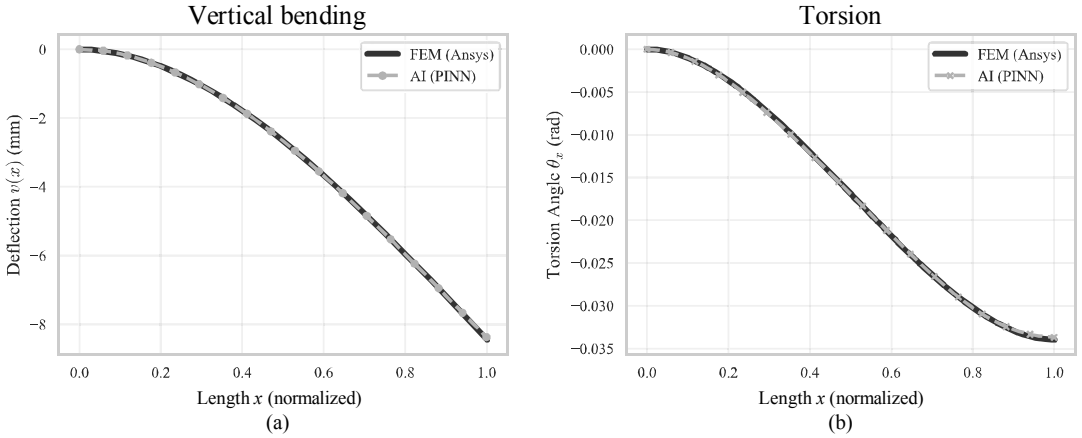


Fig. 7. Validation of kinematic parameters on a test sample: (a) vertical deflection; (b) twist angle

5. Discussion and Limitations

The presented results allow for formulating a series of important observations regarding the architecture of neural network surrogates in solid mechanics. First, the hypothesis regarding the existence of a "linearization trap" when training on multi-scale data was confirmed, aligning with the conclusions of [10] and [16]. It was experimentally proven that without the application of Coordinate Scaling and a specialized loss function with second-derivative control, gradient optimization methods fail to tune the network weights to reproduce small curvature values, reducing the solution to a linear interpolation of displacements.

Second, a significant result was obtained regarding the use of a priori knowledge in the form of "physics hints." The study established that adding the analytical solution for St. Venant's free torsion as an input feature worsened model convergence in restrained torsion problems. This is explained by the fact that the simplified analytical formula creates a strong inductive bias that contradicts the complex mechanical behavior of warping in a clamped member. Discarding this hint in favor of learning from "scratch" reduced prediction error. Conversely, the use of architectural constraints in the form of hyperbolic tangent output masking proved critically important, ensuring precise satisfaction of boundary conditions in the fixed end regardless of the network's weight values.

The current implementation of the model has certain limitations due to the problem formulation. The study was restricted to considering the material as an ideal elastic medium, ignoring the development of plastic deformations and the formation of plastic hinges. To ensure maximum accuracy of the verification data, modeling using three-dimensional solid elements was employed. This approach allowed for capturing the full stress tensor across the cross-section thickness without the kinematic simplifications inherent in shell theory, accepting the increase in computational costs as a justified compromise for a volumetric data representation. Additionally, the cross-sectional topology within this experiment was fixed to an I-beam profile. It is worth noting that the proposed architecture is universal for the class of open-section members: adapting the model for other open section types (channels, angles) does not require fundamental changes to the macro-level (*Spine Network*) and reduces to retraining only the lightweight "*Warping Head*" on the corresponding data. This distinguishes our approach from models limited to specific structure types, as in [12]. Furthermore, the model guarantees accuracy only within the stable equilibrium domain, relying on an external classifier to identify bifurcation states.

6. Conclusions

This work proposes and software-implements a Physics-Informed Neural Network architecture for modeling the stress-strain state of spatial beam structures. The scientific novelty of the research lies in combining the *Kinematic Decomposition* method, which separates the problem into predicting the deformed axial line and the field of local cross-sectional deformations, with a neural integration technique, where the numerical integration layer is embedded directly into the network architecture. This approach ensured the mathematical smoothness of the solution and guaranteed satisfaction of

boundary conditions. A key technical achievement was the development of the *Coordinate Scaling* method for derivatives, which solved the vanishing gradient problem and enabled the neural network to correctly manipulate curvature values of the order 10^{-8}m^{-2}. Numerical experiment results confirm that the created AI-surrogate is capable of replacing resource-intensive calculations of building structures based on solid finite elements, providing a speed increase of several orders of magnitude while maintaining engineering accuracy in determining internal forces within the structures. This paves the way for using such models in generative design tasks and real-time optimization of building load-bearing systems [3, 15].

REFERENCES

1. D. Samadian, I. B. Muhit, and N. Dawood, “Application of Data-Driven Surrogate Models in Structural Engineering: A Literature Review,”*Archives of Computational Methods in Engineering*, vol. 31, no. 2, pp. 735–784, 2024. DOI: 10.1007/s11831-024-10152-0.
2. H.-T. Thai, “Machine learning for structural engineering: A state-of-the-art review,”*Structures*, vol. 38, pp. 448–491, 2022. DOI: 10.1016/j.istruc.2022.02.003.
3. S. Getun, H. Ivanchenko, S. Botvinovska, G. Getun, and A. Solomin, “Optimization of structural computational models using neural networks: a systematic review of current approaches and prospects,”*Applied Geometry and Engineering Graphics*, vol. 108, pp. 72–96, 2025. DOI: 10.32347/0131-579X.2025.108.72-96.
4. *Steel structures. Design standards,* DBN V.2.6-198:2014, Ministry for Communities and Territories Development of Ukraine, Kyiv, 2022.
5. A. D. Jagtap and G. E. Karniadakis, “Extended physics-informed neural networks (XPINNs): A generalized space-time domain decomposition based deep learning framework for nonlinear partial differential equations,”*Communications in Computational Physics*, vol. 28, no. 5, pp. 2002–2041, 2020. DOI: 10.4208/cicp.OA-2020-0164.
6. D. W. Abueidda, S. Koric, N. A. Sobh, and H. Sehitoglu, “Accelerating Multiscale Modeling with Hybrid Solvers: Coupling FEM and Neural Operators with Domain Decomposition,”*arXiv preprint arXiv:2504.11383*, 2024. DOI: 10.48550/arXiv.2504.11383.
7. P. A. F. Enabe and R. Provasi, “A hybrid virtual element method and deep learning approach for solving one-dimensional Euler-Bernoulli beams,”*arXiv preprint arXiv:2501.06925*, 2025. DOI: 10.48550/arXiv.2501.06925.
8. H. Li, Y. Miao, Z. S. Khodaei, and M. H. Aliabadi, “Finite-PINN: A Physics-Informed Neural Network Architecture for Solving Solid Mechanics Problems with General Geometries,”*arXiv preprint arXiv:2412.09453*, 2024. DOI: 10.48550/arXiv.2412.09453.
9. M. Yan, S. Deng, and H. You, “Initialization-enhanced physics-informed neural network with domain decomposition (IDPINN),”*Journal of Computational Physics*, vol. 530, 113914, 2025. DOI: 10.1016/j.jcp.2025.113914.
10. K. Xie, Y. Huo, Z. Li, and Z. Wu, “ODE-DSN: A surrogate model for dynamic stiffness in microscopic RVE problems under nonuniform time-step strain inputs,”*Journal of Computational Design and Engineering*, vol. 12, no. 2, pp. 49–60, 2025. DOI: 10.1093/jcde/qwaf012.
11. Z. Han, J. Ou, and K. Koyamada, “A Physics-informed neural network-based Surrogate Model for Analyzing Elasticity Problems in Plates with Holes,”*Journal of Advanced Simulation in Science and Engineering*, vol. 11, no. 1, pp. 21–31, 2024. DOI: 10.15748/jasse.11.21.
12. H. Wu, Y.-C. Wu, P. Zhi, X. Wu, and T. Zhu, “Structural optimization of single-layer domes using surrogate-based physics-informed neural networks,”*Heliyon*, vol. 9, no. 10, e20867, 2023. DOI: 10.1016/j.heliyon.2023.e20867.
13. L. Liang, M. Liu, C. Martin, and W. Sun, “A deep learning approach to estimate stress distribution: a fast and accurate surrogate of finite-element analysis,”*Journal of The Royal Society Interface*, vol. 15, no. 138, 20170844, 2018. DOI: 10.1098/rsif.2017.0844.
14. V. A. Bazhenov, O. S. Sakharov, O. I. Hulyar, S. O. Pyskunov, and Y. V. Maksymyuk, “Features of using the moment finite element scheme (MFES) in nonlinear calculations of shells and plates,”*Strength of Materials and Theory of Structures*, no. 92, pp. 3–16, 2014.
15. V. Yurchenko and I. D. Peleshko, “Searching for a compromise solution in cross-section size optimization problems of cold-formed steel structural members,”*Strength of Materials and Theory of Structures*, no. 109, pp. 72–92, 2022. DOI: 10.32347/2410-2547.2022.109.72-92.
16. H. Ivanchenko, G. Getun, I. Sklyarov, A. Solomin, and S. Getun, “Application of the low-rank adaptation method on the example of fine-tuning a latent diffusion model,”*Strength of Materials and Theory of Structures*, no. 114, pp. 299–310, 2025. DOI: 10.32347/2410-2547.2025.114.299-310.
17. H. Ivanchenko, G. Getun, I. Bezklubenko, A. Solomin, and S. Getun, “Mathematical model of the stress-strain state of multilayered structures with different elastic properties,”*Strength of Materials and Theory of Structures*, no. 113, pp. 131–138, 2024. DOI: 10.32347/2410-2547.2024.113.131-138.

Стаття надійшла 21.01.2026

Гетун С.Ю.

ФІЗИЧНО-ІНФОРМОВАНІ НЕЙРОННІ МЕРЕЖІ ДЛЯ РОЗРАХУНКУ ПРОСТОРОВИХ СТРИЖНЕВИХ КОНСТРУКЦІЙ:ПІДХІД КІНЕМАТИЧНОЇ ДЕКОМПОЗИЦІЇ

Проєктування відповідальних сталевих конструкцій неминуче стикається з дилемою вибору розрахункової схеми: класичні стрижневі скінченні елементи забезпечують високу швидкість обчислень, але часто нехтують складними просторовими ефектами депланації та деформації перерізу, тоді як детальні об'ємні скінченно-елементні моделі гарантують еталонну точність, проте є надмірно ресурсомісткими для задач багатокритеріальної оптимізації. У цій статті запропоновано метод побудови нейромережевого сурогата (PINN), здатного передбачати напружено-деформований стан просторового стрижня з точністю високополігональної 3D FEM-моделі, але зі швидкістю, що наближається до аналітичних рішень. Запропонована архітектура базується на кінематичній декомпозиції переміщень на компоненти вісьової лінії та локальної депланації перерізу, що дозволяє врахувати ефекти 3D-пружності в рамках компактної моделі. Для забезпечення чисельної стабільності навчання при визначенні вищих похідних (кривизни та бімоментів) розроблено та застосовано метод координатного масштабування (Coordinate Scaling). Результати демонструють, що модель, навчена на обмеженому синтетичному наборі даних (10^4 зразків), відтворює компоненти переміщень з точністю, що не перевищує 0.1 мм та коректно ідентифікує ефекти геометричної нелінійності.

Ключові слова: просторові сталеві елементи, фізично-інформовані нейронні мережі (PINN), кінематична декомпозиція, деформація, сурогатне моделювання, масштабування координат.

Getun S.Yu.

PHYSICS-INFORMED NEURAL NETWORKS FOR ANALYSIS OF SPATIAL BEAM STRUCTURES: A KINEMATIC DECOMPOSITION APPROACH

The design of critical steel structures faces a persistent dilemma between computational speed and physical fidelity. While classical 1D beam elements are computationally efficient, they often fail to capture complex spatial effects like non-uniform torsion and cross-sectional distortion, whereas detailed 3D solid finite element models (FEM) offer reference accuracy but come at a prohibitive computational cost, making them unsuitable for real-time generative design or multi-objective optimization. This study proposes a novel Physics-Informed Neural Network (PINN) architecture designed to function as a real-time "AI-Surrogate" capable of predicting the stress-strain state of spatial members with the accuracy of a high-fidelity 3D FEM model but at analytical speeds. The proposed approach utilizes a Kinematic Decomposition strategy, separating the displacement field into a macroscopic "spine" behavior and a field of local cross-sectional deformations. This effectively reduces the dimensionality of the problem and allows for training on a compact dataset of 10,000 samples. To address the "linearization trap" and the vanishing gradient problem associated with predicting higher-order derivatives (curvature and bi-moments), we introduce a Coordinate Scaling technique. This method normalizes the derivative space, ensuring numerical stability and physical consistency of the solution. Validated against nonlinear 3D solid FEM simulations (Ansys), the model demonstrates high precision, achieving a Mean Absolute Error (MAE) of 0.08 mm for deflections and 0.2 mrad for torsion angles. Furthermore, the specialized physics-informed loss function successfully minimizes the curvature error to 1.5×10^{-4} m^{-2}, ensuring the accurate recovery of internal forces. The results confirm that the proposed PINN architecture effectively bridges the gap between the speed of beam theories and the accuracy of volumetric models. The introduced Coordinate Scaling method proves critical for learning differential relationships in mechanics, paving the way for the next generation of real-time structural analysis tools.

Keywords: Spatial steel members, Physics-Informed Neural Networks (PINN), Kinematic decomposition, Warping, Surrogate modeling, Coordinate scaling.

УДК624.04:624.014.2:004.8
Гетун С.Ю. **Фізично-інформовані нейронні мережі для розрахунку просторових стрижневих конструкцій: Підхід кінематичної декомпозиції //** Опір матеріалів та теорія споруд: наук.-тех. збірн. – К.: КНУБА. 2026. – Вип. 116. – С. 439-449.
Запропоновано метод побудови фізично-інформованого нейромережевого сурогата (PINN), здатного передбачати напружено-деформований стан просторового стрижня з точністю високополігональної 3D FEM-моделі, але зі швидкістю, що наближається до аналітичних рішень.
Таб. 0. Рис. 7. Бібліогр. 17 назв.

UDC 624.04:519.85:004.85
Getun S.Yu. **Physics-Informed Neural Networks for Analysis of Spatial Beam Structures: A Kinematic Decomposition Approach** // Strength of Materials and Theory of Structures: Scientific-and-technical collected articles. – K.: KNUBA. 2026. – Issue 116. – P. 439-449.
A method for constructing a neural network surrogate (PINN) capable of predicting the stress-strain state of a spatial member with the accuracy of a high-fidelity 3D Solid FEM model, yet at a speed approaching analytical solutions, is proposed.
Tabl. 2. Fig. 4. Refs. 17.

Автор: аспірант кафедри будівельної механіки Гетун Сергій Юрійович
Адреса робоча: 03680 Україна, м. Київ, проспект Повітряних Сил, 31, Київський національний університет будівництва і архітектури, Гетуну Сергію Юрійовичу
E-mail: hetun_sy-2024@knuba.edu.ua
ORCID ID: https://orcid.org/0009-0001-2269-7035