Best Kinematics for Shell Finite Elements using Convolutional Neural Networks

M. Petrolo and P. Iannotti

MUL^2 Lab, Department of Mechanical and Aerospace Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy

Revised Version of UMCM-2022-0942

Author for correspondence:
M. Petrolo, Associate Professor
MUL^2 Lab, Department of Mechanical and Aerospace Engineering,
Politecnico di Torino,
Corso Duca degli Abruzzi 24,
10129 Torino, Italy,
tel: +39 011 090 6845,
e-mail: marco.petrolo@polito.it

Abstract

The present work aims at exploring the capabilities of Convolutional Neural Networks (CNN) to identify the most influential generalized variables in 2D formulations for plates and shells. The outcome of CNN is the Best Theory Diagram (BTD), a graphical representation of the dependency between the model's accuracy and nodal degrees of freedom (DOF). The networks are trained with data derived from Finite Element (FE) computations and samples of reduced theories obtained as combinations of a given set of generalized displacement variables. Such samples are obtained through the Carrera Unified Formulation (CUF), a generalized approach to generating the governing equations for any structural model. Furthermore, the Node-Dependent Kinematics (NDK) included local refinements to lead to Best Theory Distributions of structural theories over an FE mesh, that is, identifying areas of a shell in which higher-order models are most necessary. The training data can refer to different analyses, e.g., static or free-vibration, whereas the network's input can include multiple structural parameters together with a sequence of expansion terms or theory distributions. The numerical results highlight a significant computational efficiency of CNN and its ability to identify the best models even for problem configurations not included in the training set.

Keywords: CUF, Shells, Finite Elements, Neural Networks, NDK

1 Introduction

The development of multilayered composites and their increasing adoption in many fields require new modeling strategies to adequately describe their complex mechanical behaviors. However, a model's accuracy can also significantly impact the computational costs, and an adequate compromise between accuracy and computational overhead must be found. Concerning plate and shell structures, many 2D theories were developed to provide a solution with accuracy comparable to 3D formulations but at lower costs.
The first and most known example is the Classical Lamination Theory (CLT), built as an extension of 2D theories for single-layer structures of isotropic materials to multilayered ones [1, 2]. CLT adopted the Kirchhoff hypotheses [3] and was later modified to include transverse shear strains, thus arriving at the definition of the First-order Shear Deformation Theory (FSDT) [4–6]. However, the FSDT assumes a linear variation of the in-plane displacement components along with the thickness and cannot satisfy the condition of null transverse shear stresses on the top and bottom faces of the structure, therefore, requiring the introduction of shear correction factors [7, 8]. Further developments on the FSDT [9–11] introduced higher-order terms to describe the in-plane displacement field, and, afterward, transverse displacement through higher-order expansions [12]. Other models followed in which the order of expansion adopted for each displacement component varied [13–21], including non-polynomial terms [22–25].
The abovementioned theories are called Equivalent Single Layer (ESL). To obtain a better description of the stress distributions, particularly the transverse ones, a selection process for the expansion terms is required [26]. The adoption of an infinite expansion would provide 3D solutions [27], but this approach is not practically feasible, and the use of the smallest number of variables to achieve adequate precision also contributes to reducing the computational costs. These difficulties in the description of the mechanical behavior of composite plates and shells led to new methodologies such as zig-zag models [28], and the layer-wise (LW) approach [29–32]. This paper focuses on ESL and the selection process for the expansion terms. This task can be performed in various ways. The two procedures historically adopted are the axiomatic and the asymptotic methods. In the first one, also referred to as the method of hypotheses, expansion terms are chosen according to mechanical behavior assumptions, as in the case of classical structural theories. The asymptotic method [33, 34], instead considers the 3D formulation as the starting point to derive a simplified model based on the influence of one or more characteristic parameters included in the governing equations, and whose associated terms are selectively suppressed, depending on their

influence on the accuracy of the solution.

A third approach has emerged in recent years to evaluate the effectiveness of higher-order theories; namely, the Axiomatic/Asymptotic Method (AAM) [35–39], and allowing the analysis of the influence of generalized displacement variables. The AAM can be used to obtain the Best Theory Diagram (BTD), where the best accuracy for a given number of nodal degrees of freedom can be read. AAM requires the numerical evaluation of arbitrary structural theories and uses the Carrera Unified Formulation [40], CUF, for this purpose. The introduction of CUF provided a hierarchical and generalized approach to obtaining the governing equations for structural models of any order and complexity. This methodology was also enriched by the development and successful implementation of the Node-Dependent Kinematics (NDK) [41, 42] in which each node of an FE model can have a different structural theory.

The combinations of available generalized variables are generally high, and the computational overhead to obtain a BTD can be high. A possible solution can be found in the adoption of Machine Learning (ML) techniques such as Neural Networks (NN) [43, 44]. The use and capabilities of these mathematical tools have been growing in the last years in many different fields, including structural mechanics [45–53]. In this work, Convolutional Neural Networks (CNN) [54–56] are used to evaluate structural models for plates and shells. This approach stems from the methodology and results described in [26, 57, 58], where NN was successfully employed to obtain BTD for different problem configurations and analyses. In this paper, CNN is employed for the first time to include more features in the training process and improve computational efficiency. The paper is structured as follows: Section 2 provides the description of the FE formulation methodology, Section 3 describes the evaluation procedure, the accuracy parameters employed, and the CNN adopted for this study, results are presented in Section 4 and the conclusions drawn in Section 5.

2 Carrera Unified Formulation and Finite Elements

In the present work, the FE results stemmed from the use of the CUF [59] applied to a bi-dimensional model. In this framework, considering the reference system employed in Fig. 1, the displacement field can be expressed as

$$\mathbf{u}(\alpha, \beta, z) = F_\tau(z)\mathbf{u}_\tau(\alpha, \beta) \qquad \tau = 1, \ldots, M \tag{1}$$

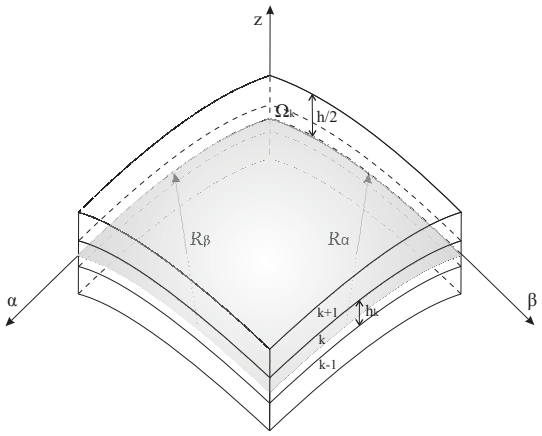


Figure 1: Reference geometry for shell models

The Einstein notation acts on τ and F_τ are the thickness expansion functions. $\mathbf{u}_\tau$ is the vector of the generalized unknown displacements and M is the number of expansion terms. For instance, a fourth-order model, referred to as N=4, is

$$\begin{aligned} u_\alpha &= u_{\alpha_1} + z\, u_{\alpha_2} + z^2\, u_{\alpha_3} + z^3\, u_{\alpha_4} + z^4\, u_{\alpha_5} \\ u_\beta &= u_{\beta_1} + z\, u_{\beta_2} + z^2\, u_{\beta_3} + z^3\, u_{\beta_4} + z^4\, u_{\beta_5} \\ u_z &= u_{z_1} + z\, u_{z_2} + z^2\, u_{z_3} + z^3\, u_{z_4} + z^4\, u_{z_5} \end{aligned} \tag{2}$$

and has fifteen nodal DOF. The order and type of expansion is a free parameter. The metric coefficients H_α^k, H_β^k and H_z^k of the k^{th} layer are

$$H_\alpha^k = A^k(1 + z_k/R_\alpha^k), \quad H_\beta^k = B^k(1 + z_k/R_\beta^k), \quad H_z^k = 1 \tag{3}$$

As shown in Fig. 1, R_α^k and R_β^k are the principal radii of the middle surface of the k^{th} layer, A^k and B^k the coefficients of the first fundamental form of Ω_k. In this paper, only constant radii of curvature were considered, $\mathrm{A}^k = \mathrm{B}^k = 1$. The geometrical strains can be written as

$$\begin{aligned} \boldsymbol{\epsilon}_p^k &= \left\{ \epsilon_{\alpha\alpha}^k, \epsilon_{\beta\beta}^k, \epsilon_{\alpha\beta}^k \right\}^T = (\boldsymbol{D}_p^k + \boldsymbol{A}_p^k)\boldsymbol{u}^k \\ \boldsymbol{\epsilon}_n^k &= \left\{ \epsilon_{\alpha z}^k, \epsilon_{\beta z}^k, \epsilon_{zz}^k \right\}^T = (\boldsymbol{D}_{n\Omega}^k + \boldsymbol{D}_{nz}^k - \boldsymbol{A}_n^k)\boldsymbol{u}^k \end{aligned} \tag{4}$$

where

$$\boldsymbol{D}_p^k = \begin{bmatrix} \frac{\partial_\alpha}{H_\alpha^k} & 0 & 0 \\ 0 & \frac{\partial_\beta}{H_\beta^k} & 0 \\ \frac{\partial_\beta}{H_\beta^k} & \frac{\partial_\alpha}{H_\alpha^k} & 0 \end{bmatrix} \quad \boldsymbol{D}_{n\Omega}^k = \begin{bmatrix} 0 & 0 & \frac{\partial_\alpha}{H_\alpha^k} \\ 0 & 0 & \frac{\partial_\beta}{H_\beta^k} \\ 0 & 0 & 0 \end{bmatrix} \quad \boldsymbol{D}_{nz}^k = \begin{bmatrix} \partial_z & 0 & 0 \\ 0 & \partial_z & 0 \\ 0 & 0 & \partial_z \end{bmatrix} \tag{5}$$

$$\boldsymbol{A}_p^k = \begin{bmatrix} 0 & 0 & \frac{1}{H_\alpha^k R_\alpha^k} \\ 0 & 0 & \frac{1}{H_\beta^k R_\beta^k} \\ 0 & 0 & 0 \end{bmatrix} \quad \boldsymbol{A}_n^k = \begin{bmatrix} \frac{1}{H_\alpha^k R_\alpha^k} & 0 & 0 \\ 0 & \frac{1}{H_\beta^k R_\beta^k} & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{6}$$

For the stress-strain relations, it follows that

$$\begin{aligned} \boldsymbol{\sigma}_p^k &= \left\{ \sigma_{\alpha\alpha}^k, \sigma_{\beta\beta}^k, \sigma_{\alpha\beta}^k \right\}^T = \boldsymbol{C}_{pp}^k \boldsymbol{\epsilon}_p^k + \boldsymbol{C}_{pn}^k \boldsymbol{\epsilon}_{pn}^k \\ \boldsymbol{\sigma}_n^k &= \left\{ \sigma_{\alpha z}^k, \sigma_{\beta z}^k, \sigma_{zz}^k \right\}^T = \boldsymbol{C}_{np}^k \boldsymbol{\epsilon}_{np}^k + \boldsymbol{C}_{nn}^k \boldsymbol{\epsilon}_n^k \end{aligned} \tag{7}$$

where

$$\begin{aligned} \boldsymbol{C}_{pp}^k &= \begin{bmatrix} C_{11}^k & C_{12}^k & C_{16}^k \\ C_{12}^k & C_{22}^k & C_{26}^k \\ C_{16}^k & C_{26}^k & C_{66}^k \end{bmatrix} \quad \boldsymbol{C}_{pn}^k = \begin{bmatrix} 0 & 0 & C_{13}^k \\ 0 & 0 & C_{23}^k \\ 0 & 0 & C_{36}^k \end{bmatrix} \\ \boldsymbol{C}_{np}^k &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ C_{13}^k & C_{23}^k & C_{36}^k \end{bmatrix} \quad \boldsymbol{C}_{nn}^k = \begin{bmatrix} C_{55}^k & C_{45}^k & 0 \\ C_{45}^k & C_{44}^k & 0 \\ 0 & 0 & C_{33}^k \end{bmatrix} \end{aligned} \tag{8}$$

The FE formulation uses a nine-node shell element based on the Mixed Interpolation of Tensorial Component (MITC) method [60]. The displacement vector becomes

$$\delta \boldsymbol{u}_s = N_j \delta \boldsymbol{u}_{sj}, \qquad \boldsymbol{u}_\tau = N_i \boldsymbol{u}_{\tau i} \qquad i, j = 1, \cdots, 9 \tag{9}$$

$\boldsymbol{u}_{\tau i}$ and $\delta \boldsymbol{u}_{sj}$ are the nodal displacement vector and its virtual variation, respectively. The strain expression becomes

$$\begin{aligned} \boldsymbol{\epsilon}_p &= F_\tau (\boldsymbol{D}_p + \boldsymbol{A}_p) N_i \boldsymbol{u}_{\tau i} \\ \boldsymbol{\epsilon}_n &= F_\tau (\boldsymbol{D}_{n\Omega} - \boldsymbol{A}_n) N_i \boldsymbol{u}_{\tau i} + F_{\tau,z} N_i \boldsymbol{u}_{\tau i} \end{aligned} \tag{10}$$

MITC contrasts the membrane and shear locking via a specific interpolation strategy for the strain components on the nine-node shell element, as follows:

$$\begin{aligned}
\boldsymbol{\epsilon}_p &= \begin{bmatrix} \epsilon_{\alpha\alpha} \\ \epsilon_{\beta\beta} \\ \epsilon_{\alpha\beta} \end{bmatrix} = \begin{bmatrix} N_{m1} & 0 & 0 \\ 0 & N_{m2} & 0 \\ 0 & 0 & N_{m3} \end{bmatrix} \begin{bmatrix} \epsilon_{\alpha\alpha_{m1}} \\ \epsilon_{\beta\beta_{m2}} \\ \epsilon_{\alpha\beta_{m3}} \end{bmatrix} \\
\boldsymbol{\epsilon}_n &= \begin{bmatrix} \epsilon_{\alpha z} \\ \epsilon_{\beta z} \\ \epsilon_{zz} \end{bmatrix} = \begin{bmatrix} N_{m1} & 0 & 0 \\ 0 & N_{m2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \epsilon_{\alpha z_{m1}} \\ \epsilon_{\beta z_{m2}} \\ \epsilon_{z z_{m3}} \end{bmatrix}
\end{aligned} \tag{11}$$

Strains $\epsilon_{\alpha\alpha_{m1}}$, $\epsilon_{\beta\beta_{m2}}$, $\epsilon_{\alpha\beta_{m3}}$, $\epsilon_{\alpha z_{m1}}$, and $\epsilon_{\beta z_{m2}}$ stem from 10 and

$$\begin{aligned}
N_{m1} &= [N_{A1}, N_{B1}, N_{C1}, N_{D1}, N_{E1}, N_{F1}] \\
N_{m2} &= [N_{A2}, N_{B2}, N_{C2}, N_{D2}, N_{E2}, N_{F2}] \\
N_{m3} &= [N_P, N_Q, N_R, N_S]
\end{aligned} \tag{12}$$

Subscripts m1, m2 and m3 indicate the point groups (A1,B1,C1,D1,E1,F1), (A2,B2,C2,D2,E2,F2), and (P,Q,R,S), respectively, see Fig. 2. Via the Principle of Virtual Displacements (PVD) for the static analysis, the equilibrium equation reads

$$\boldsymbol{k}^k_{\tau sij} \boldsymbol{u}^k_{\tau i} = \boldsymbol{p}^k_{sj} \tag{13}$$

The 3×3 matrix $\boldsymbol{k}^k_{\tau sij}$ is the fundamental mechanical nucleus whose expression is independent of the order of the expansion. $\boldsymbol{p}^k_{sj}$ is the load vector. In a similar way, for the dynamic case,

$$\boldsymbol{m}^k_{\tau sij} \ddot{\boldsymbol{u}}^k_{\tau i} + \boldsymbol{k}^k_{\tau sij} \boldsymbol{u}^k_{\tau i} = 0 \tag{14}$$

where $\boldsymbol{m}^k_{\tau sij}$ is the fundamental nucleus of of the mass matrix. The assembly over all nodes and elements and the introduction of the harmonic solution lead to the eigenvalue problem,

$$(-\omega_n^2 \boldsymbol{M} + \boldsymbol{K}) \boldsymbol{U}_n = 0 \tag{15}$$

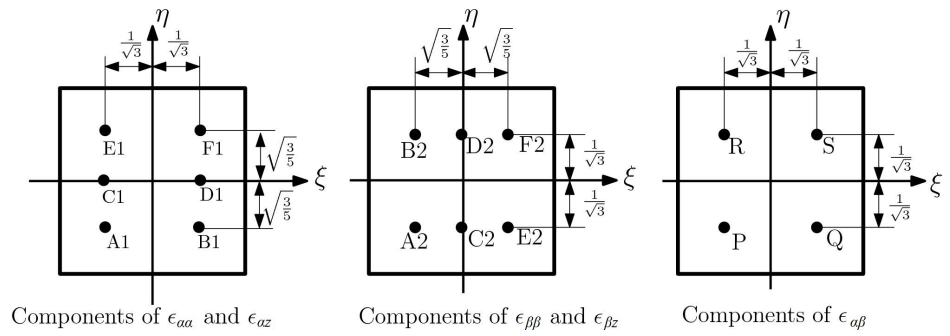


Figure 2: MITC9 tying points

In the CUF, through the NDK approach [42, 61], each node can have a different shell theory, and, in the present work, the same one is adopted for every node on a single element. A more detailed description of the procedures deployed by the CUF can be found in [62].

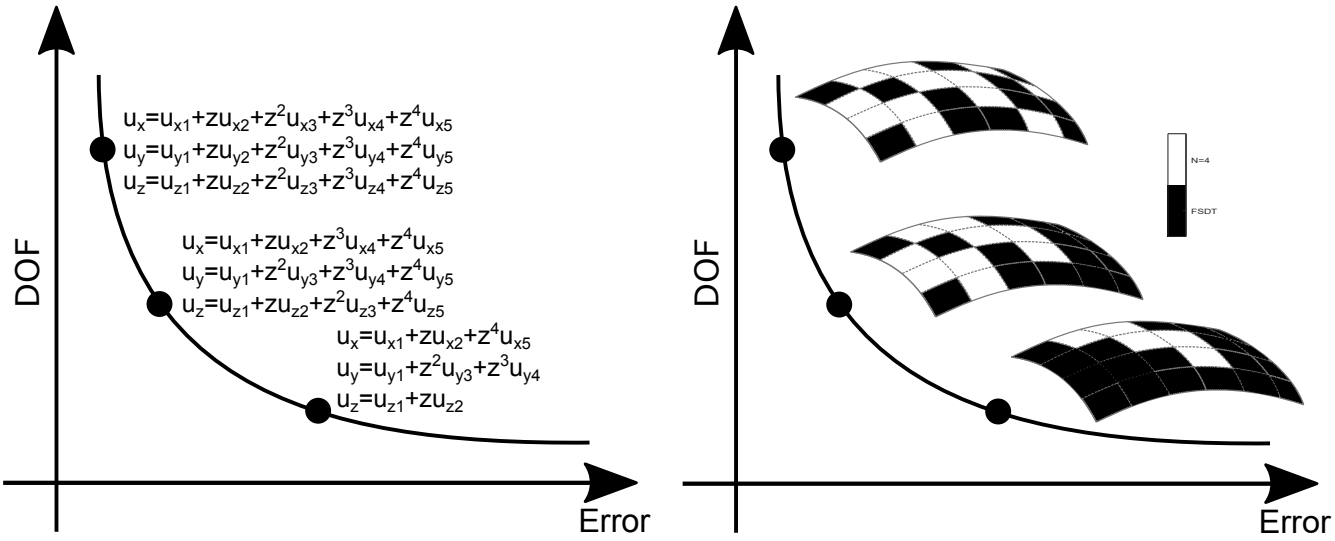
3 Best Theory Diagrams and Convolutional Neural Networks

The starting point of the evaluation procedure known as AAM is the axiomatic choice of the maximum order of the expansion. From there, all the possible combinations of active terms up to said order are compared to a reference model. Depending on the complexity of the theory, i. e. the number of expansion terms, different levels of accuracy can be achieved. One of the outcomes is the Best Theory Diagram (BTD). For each number of active terms, also referred to as nodal degrees of freedom (DOF), the BTD reports best accuracy obtainable, as in Fig. 3a. As the reference solution, a full fourth-order expansion (N=4) was considered, similarly to what was presented in previous works [57, 58]. An N=4 model has fifteen generalized displacement variables; therefore, 2^{15} reduced models stem by combining all terms. In this work, 2^{12} combinations were considered as the three zeroth-order terms were kept active due to their high importance for the accuracy of the solution. To evaluate the accuracy of each model, two different control parameters were used. For the static case,

$$Error = 100 \times \frac{|u_z - u_z^{N=4}|}{|u_z^{N=4}|} \tag{16}$$

u_z is the maximum displacement, and $u_z^{N=4}$ is the reference value. The average relative errors for each of the ten computed natural frequencies for free-vibration analyses.

$$Error = \sum_{i=1}^{10} \frac{f_i / f_i^{N=4}}{10} \tag{17}$$



(a) Best Theory Diagram

(b) Best Theory Distribution

Figure 3: Best Theory Diagrams.

These parameters were also used to build Best Theory Distributions via NDK. As detailed in the result section, a 4X4 mesh was used in this case. Two structural models, FSDT and N=4, were considered, leading to 2^{16} combinations of mesh distributions. In this case, BTD can be seen as the Best Theory Distribution, i.e., the curve composed of all meshes with a given number of N=4 and FSDT elements providing the minimum error, see Fig. 3b. In the case of sixteen elements, the BTD will have sixteen models, the first one with all N=4 and the last one with all FSDT.

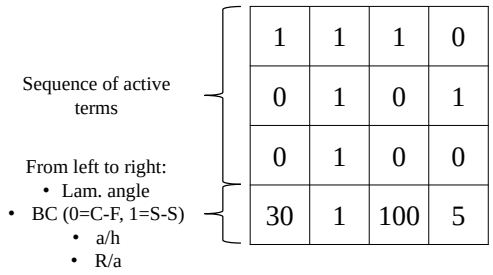
Results from FEM served as training sets for the CNN employed. Developed from the basic multi-layer perceptron (MLP), CNN can handle tensorial inputs of multiple orders and detect interactions among input features. In the case of a convolutional layer, the equivalent of a single neuron of a common NN is represented by a convolutional kernel, often referred to as a filter, which is a tensor of the same order as the input containing the model weights. The convolutional operation can be interpreted as the sum of the products between corresponding elements of inputs and filters, with the latter sliding over the input matrix with a specific step, or stride, in each direction, starting from the upper-left corner. Depending

on the choice of the stride, the dimensions of the output may vary, passing through a convolutional layer. One way to preserve the original size of the input is by adding rows and columns, typically of zeros, around the perimeter of each of the input tensor components. This technique, called padding, is used when loss of information must be avoided, whereas increasing strides can be compared to a down-sampling procedure. The training of the network is performed through the minimization of a loss function; in the case of this paper, the mean squared error between predictions and expected output. The outcome of this procedure is represented by the network parameters, weights, and biases, which are updated through backpropagation. The optimization algorithm used is Adam [63], a stochastic gradient descent method vastly employed in the ML field, given its computational efficiency and robustness. For further details on the mathematical background for CNNs, the reader may refer to [64].

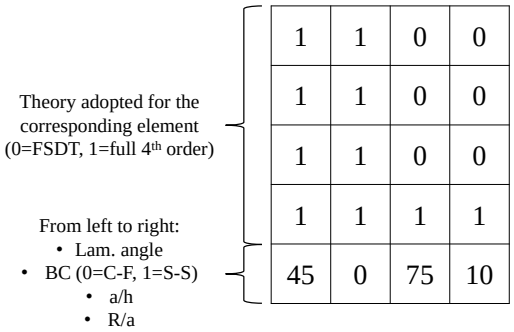
The CNN architecture used in this paper is summarized in Table 1 and chosen after multiple tests on various configurations to minimize computational costs and maximize accuracy. The CNN adopted combines convolutional and dense layers, in both cases of dimension 128, and a Rectified Linear Unit (ReLU) was used as the activation function. In the output layer, a Sigmoid function was employed. Using CNN allowed multi-dimensional inputs and involved different structural features in the training phase. The input was built by encoding the sequence of active terms representing a specific structural model into a sequence of 0 and 1 and corresponding to deactivated and active terms of the expansion, respectively. Additional information such as thickness ratio, stacking sequence, and boundary conditions was conveyed by appending them to the same array in an adequately coded format. The final input form required re-shaping the assembled array in a matrix format. In the case of mesh distributions over a 4X4 mesh, the first part of the inputs consisted of 16 terms, in which 0 and 1 corresponded to FSDT and N=4, respectively. Two examples of inputs, one for each case, are provided in Fig. 4. The training of the network used some 10% of all possible models evaluated for each structural case, randomly selected.

Table 1: Parameters and architecture of adopted CNN

Layer	Filters (Size) /Nodes	Activation Function
Convolutional	128 (3x3)	ReLU
Convolutional	128 (3x3)	ReLU
Convolutional	128 (3x3)	ReLU
Dense	128	ReLU
Dense	128	ReLU
Ouput	1	Sigmoid



(a) Single theory - Input for a simply-supported shell with $\theta =$ 30, a/h=100, R/a=5.



(b) Mesh distribution - Input for a clamped-free shell with $\theta =$ 45, a/h=75, R/a=10.

Figure 4: Input format

4 Numerical Results

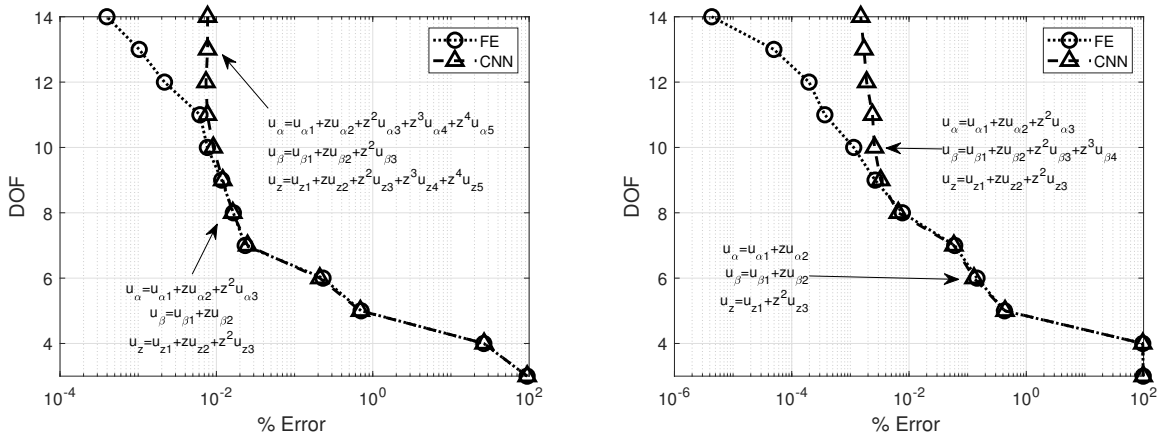
The numerical results presented in this section were all obtained considering the same base geometry of a square shell, with a=b=1 m and the same curvatures. Different thickness ratios, a/h, stacking sequences and curvature radii, R/a, were used. Two boundary condition sets were considered: two opposite edges clamped and two free (C-F), and all simply-supported (S-S). The following material properties were adopted: E_1/E_2=25, G_{12}/E_2=G_{13}/E_2=0.5, G_{23}/E_2=0.2, ν=0.25. Static and free-vibration analyses were carried out using FEM. The former considered a bi-sinusoidal load applied on the upper surface of the shell, and the output consisted of the maximum vertical displacement u_z. The latter gave the first ten natural frequencies. A 2D model was used throughout this work, and, to lower the global computational cost, a quarter of the shell with two symmetry boundary conditions was considered; thus, only symmetric modes were computed. A 4X4 nine-node (Q9) mesh was employed.

4.1 Verification of the CNN, static analysis

The first set of results aims at verifying the accuracy of the CNN concerning the static analysis. Numerical cases were retrieved from [57] with R/a=5, and in which a different neural network was used, namely, a Fully Connected Neural Network (FCNN), multilayer feed-forward, with Levenberg–Marquardt training functions.

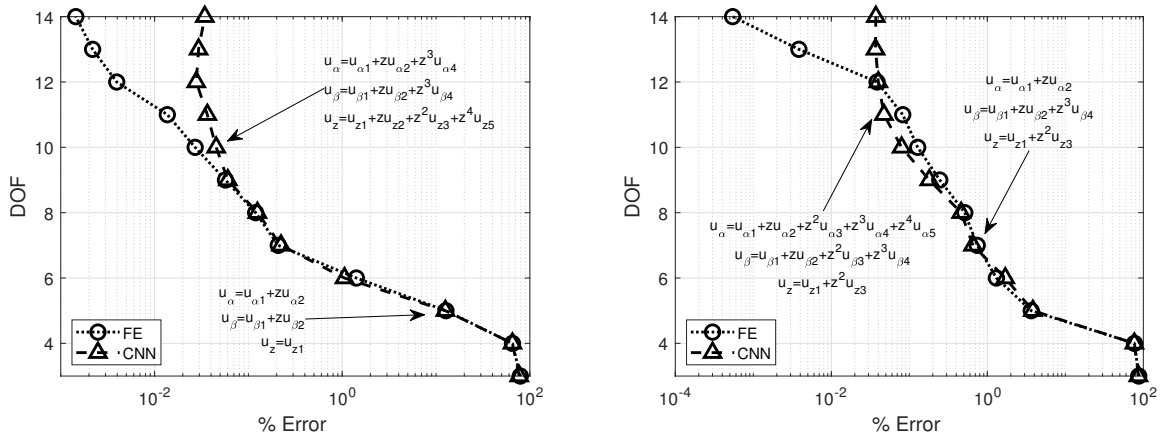
Figures 5 and 6 show BTD as obtained with FE and CNN. For the sake of clarity, in all the following sections, BTD do not report the error associated with the reference model, i.e., the 15 DOF model with null error, and the corresponding value from CNN. The FE results were obtained by considering all combinations of structural theories and selecting those with the minimum error per given nodal DOF. CNN, instead, used 10% of FE results for training and, then, was employed to evaluate the error of all combinations. Some of the best models are indicated. Tables 2 and 3 compare the error provided by FE, CNN and FCNN for best models. Each row refers to different nodal DOF. For instance, the second row refers to the best model with fourteen DOF per node. The results show that

- In all cases, there is a perfect match between FE and CNN. Moreover, the CNN is, at least, as accurate as FCNN.
- Five to seven expansion terms are required to achieve an error lower than 1%. As well-known, first-order terms are essential to have a good accuracy with in-plane cubic and transverse parabolic terms following.



(a) C-F (b) S-S

Figure 5: BTD for shells with 0/90/0, a/h=100 and R/a=5, static analysis.



(a) 0/90/0 (b) 90/0

Figure 6: BTD for S-S shells, a/h=10 and R/a=5, static analysis.

Table 2: Comparisons of FE, CNN and FCNN for S-S shell, 0/90/0, a/h=25, R/a=5, static analysis.

DOF	% Err. (FE)	% Err. (CNN)	% Err. [57]
15	0	0.2785	0.5690
14	0.0003	0.2512	0.6008
13	0.0022	0.2558	0.6424
12	0.0250	0.3195	0.7215
11	0.2787	0.3705	0.8150
10	0.4637	0.5079	0.9235
9	0.7149	0.7318	0.9752
8	1.0095	1.1749	1.0393
7	1.3902	1.3060	1.4889
6	1.8457	1.7049	2.0527
5	2.8213	2.8911	2.6410
4	89.8768	90.26311	85.1964

Table 3: Comparisons of FE, CNN and FCNN for S-S shell, 0/90/0, a/h=75, R/a=5, static analysis.

DOF	% Err. (FE)	% Err. (CNN)	% Err. [57]
15	0	0.1320	0.0176
14	0.0055	0.0116	0.0215
13	0.0163	0.0139	0.0313
12	0.0235	0.0230	0.0447
11	0.0343	0.0335	0.0597
10	0.0556	0.0514	0.0836
9	0.0689	0.0746	0.1058
8	0.0931	0.0895	0.1321
7	0.1433	0.1433	0.1758
6	0.1971	0.1991	0.2394
5	0.5232	0.5321	0.6154
4	96.8974	96.3548	96.9259

4.2 Verification of the CNN, free-vibration analysis

This section concerns the verification of the CNN for free-vibration analyses using cases from [26, 58]. Figures 7 and 8 show BTD from FE and CNN. Concerning the use of NDK to build best theory

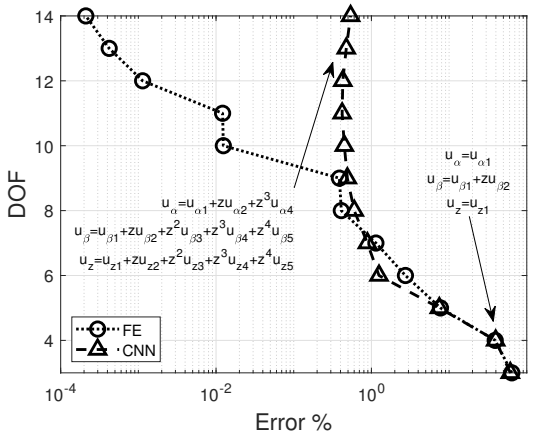
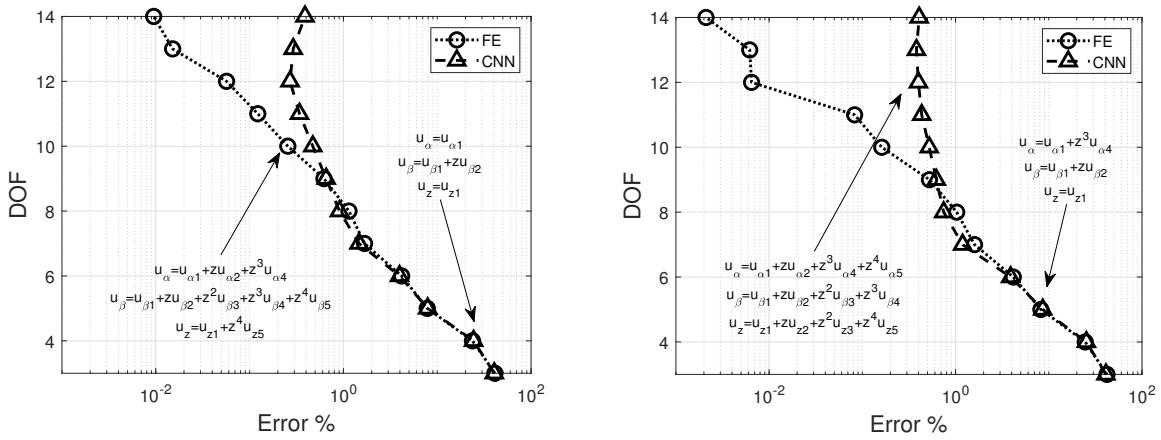


Figure 7: BTD for C-F shells, 0/90/0, a/h=20, R/a=10, free-vibration analysis.



(a) R/a=2 (b) R/a=5

Figure 8: BTD for S-S shells, a/h=10, 0/90/0, R/a=2 and 5, free-vibration analysis.

distributions, Figs. 9 and 10 show the BTD and the best distributions over the mesh of FSDT and N=4 structural theories. The vertical axis of Fig. 9 reports the number of elements with N=4, whereas the horizontal axis shows the error. Three distributions are presented in Fig. 10. The results suggest that

- As in the static case, there is a perfect match between FE and CNN results. The CNN can identify the BTD for structural theories and distributions.

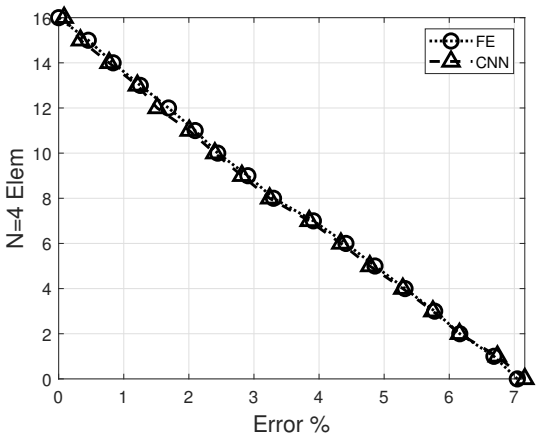
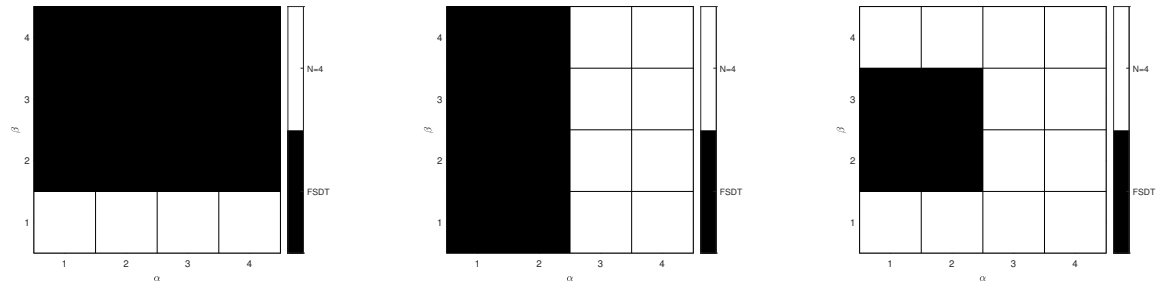
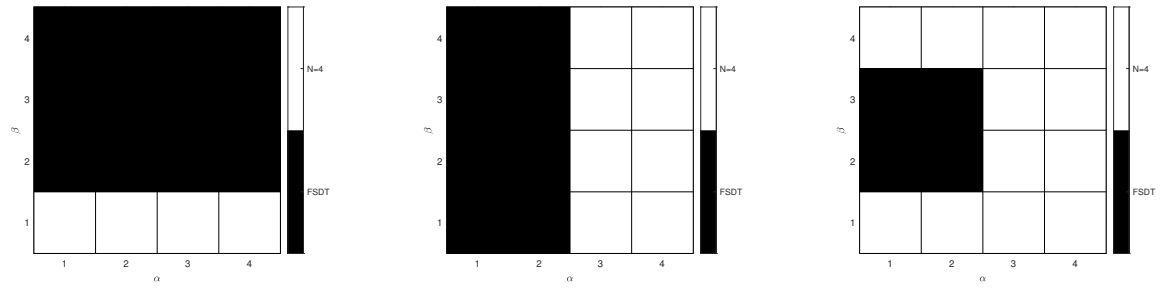


Figure 9: BTD for C-F shell, 90/0, a/h=10, R/a=5, NDK, free-vibration analysis.



(a) 12 FSDT, E=5.287% (CNN) (b) 8 FSDT, E=3.241% (CNN) (c) 4 FSDT, E=1.513% (CNN)



(d) 12 FSDT, E=5.321% (FE) (e) 8 FSDT, E=3.301% (FE) (f) 4 FSDT, E=1.687% (FE)

Figure 10: Comparison between FE and CNN mesh distributions and errors for the C-F shell, 90/0, a/h=10, R/a=5, free-vibration analysis.

- To have errors lower than 1%, some nine terms are required for the considered cases. First and third-order terms are the most influential.
- Concerning the distributions of theories over the FE mesh, the position of constraints play a fundamental role as pointed out in [58].

4.3 Multi-channel CNN with structural features as inputs

This section presents assessments concerning the use of CNN to include structural parameters in the training process. The network was trained considering the thickness, boundary conditions, the lamination angle. For each combination of thickness, lamination angle, and boundary condition, the network was trained using data from four different configurations referring to shell models with various thickness ratios and lamination angles. For the first case, the simply-supported one with a lamination angle of 30, the reference model used are listed in Table 4. The results relative to the second case, the clamped-free with lamination angle of 45, were obtained using data relative to the structural configurations of Table 5. Finally, for the simply-supported with lamination angle of 60, the employed data were relative to those described in Table 6.

Table 4: Case 1 - Features of training data.

Configuration	BC	a/h	R/a	Lam. angle (°)
1	Clamped-Free	50	5	15
2	Simply-supported	50	5	45
3	Clamped-free	100	5	15
4	Simply-supported	100	5	45

Table 5: Case 2 - Features of training data.

Configuration	BC	a/h	R/a	Lam. angle (°)
1	Clamped-Free	50	5	30
2	Simply-supported	50	5	60
3	Clamped-free	100	5	30
4	Simply-supported	100	5	60

Table 6: Case 3 - Features of training data.

Configuration	BC	a/h	R/a	Lam. angle (°)
1	Clamped-Free	50	5	45
2	Simply-supported	50	5	75
3	Clamped-free	100	5	45
4	Simply-supported	100	5	75

For each input, only the 10% of the reference data was used during the training of the network. The BTDs obtained are represented in Figures 11, 12, and 13, with their respective errors detailed in Tables 7, 9, and 11. The best theories are also presented, being described in Tables 8, 10, and 12.

The results suggest that

- The CNN was able to accurately reproduce the BTD, demonstrating that a more generalized use

Table 7: Multi-channel CNN, accuracy of BTD for S-S shell, 0/30/0, a/h=75, R/a=5, free-vibrations.

DOF	% Err. (FE)	% Err. (CNN)
15	0	0.1301
14	0.0004	0.1099
13	0.0011	0.1035
12	0.0025	0.0986
11	0.0048	0.0932
10	0.0182	0.1070
9	0.0333	0.1156
8	0.6419	0.3579
7	1.1381	1.1874
6	2.273	2.3549
5	4.8530	4.5359
4	61.7168	61.4835
3	90.7338	90.8009

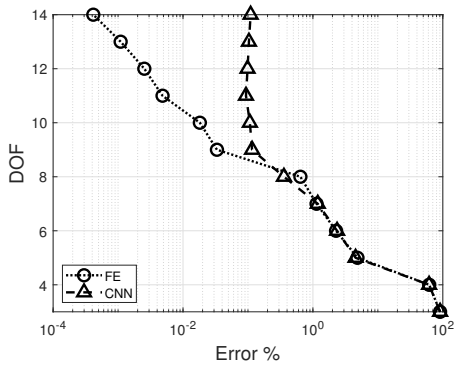


Figure 11: Multi-channel CNN, BTD for S-S shell, 0/30/0, a/h=75, R/a=5, free-vibrations.

Table 8: Multi-channel CNN, best theories for S-S shell, 0/30/0, a/h=75, R/a=5, free-vibrations.

DOF	$u_{\alpha 1}$	$u_{\beta 1}$	u_{z1}	$u_{\alpha 2}$	$u_{\beta 2}$	u_{z2}	$u_{\alpha 3}$	$u_{\beta 3}$	u_{z3}	$u_{\alpha 4}$	$u_{\beta 4}$	u_{z4}	$u_{\alpha 5}$	$u_{\beta 5}$	u_{z5}
15	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
14	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	△
13	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	△	▲	▲	△
12	▲	▲	▲	▲	▲	▲	△	△	▲	▲	▲	▲	▲	△	▲
11	▲	▲	▲	▲	▲	▲	△	△	▲	▲	▲	▲	▲	△	△
10	▲	▲	▲	▲	▲	▲	△	△	▲	▲	▲	▲	△	△	△
9	▲	▲	▲	▲	▲	▲	△	△	▲	▲	▲	△	△	△	△
8	▲	▲	▲	▲	▲	△	△	△	▲	▲	▲	△	△	△	△
7	▲	▲	▲	▲	▲	△	△	△	▲	▲	△	△	△	△	△
6	▲	▲	▲	▲	▲	△	△	△	▲	△	△	△	△	△	△
5	▲	▲	▲	▲	▲	△	△	△	△	△	△	△	△	△	△
4	▲	▲	▲	▲	△	△	△	△	△	△	△	△	△	△	△
3	▲	▲	▲	△	△	△	△	△	△	△	△	△	△	△	△

of this technique for the definition of the structural best theories is possible. The use of 10% of all combinations of FE analyses is enough for the CNN training.

Table 9: Multi-channel CNN, accuracy of BTD for C-F shell, 0/45/0, a/h=75, R/a=5, free-vibrations.

DOF	% Err. (FE)	% Err. (CNN)
15	0	0.1555
14	0.0006	0.1546
13	0.0015	0.1322
12	0.0034	0.1422
11	0.0061	0.1451
10	0.0332	0.2071
9	0.0659	0.2606
8	0.4919	0.8625
7	1.0434	1.6808
6	3.1792	2.4268
5	5.4155	4.0571
4	68.4405	71.7408
3	89.7576	90.5725

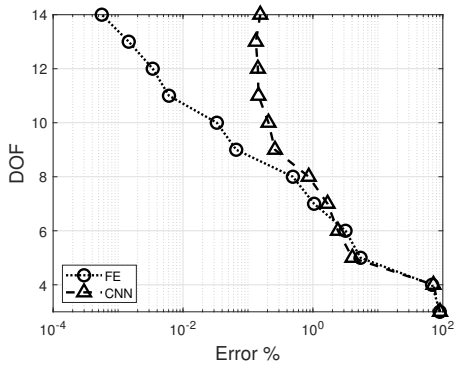


Figure 12: Multi-channel CNN, BTD for C-F shell, 0/45/0, a/h=75, R/a=5, , free-vibrations.

Table 10: Multi-channel CNN, best theories for C-F shell, 0/45/0, a/h=75, R/a=5, , free-vibrations.

DOF	$u_{\alpha 1}$	$u_{\beta 1}$	u_{z1}	$u_{\alpha 2}$	$u_{\beta 2}$	u_{z2}	$u_{\alpha 3}$	$u_{\beta 3}$	u_{z3}	$u_{\alpha 4}$	$u_{\beta 4}$	u_{z4}	$u_{\alpha 5}$	$u_{\beta 5}$	u_{z5}
15	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
14	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	△	▲	▲
13	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	△	▲	△
12	▲	▲	▲	▲	▲	▲	△	▲	▲	▲	▲	▲	△	▲	△
11	▲	▲	▲	▲	▲	▲	△	▲	▲	▲	▲	▲	△	△	△
10	▲	▲	▲	▲	▲	▲	△	△	▲	▲	▲	△	△	▲	△
9	▲	▲	▲	▲	▲	▲	△	△	▲	▲	▲	△	△	△	△
8	▲	▲	▲	▲	▲	△	△	△	▲	▲	▲	△	△	△	△
7	▲	▲	▲	▲	▲	△	△	△	▲	▲	△	△	△	△	△
6	▲	▲	▲	▲	▲	△	△	△	▲	△	△	△	△	△	△
5	▲	▲	▲	▲	▲	△	△	△	△	△	△	△	△	△	△
4	▲	▲	▲	▲	△	△	△	△	△	△	△	△	△	△	△
3	▲	▲	▲	△	△	△	△	△	△	△	△	△	△	△	△

- Lamination, boundary conditions and thickness can be used in the training process to build networks able to receive, as inputs, various values of these features and, thus, avoiding new FE analysis.

Table 11: Multi-channel CNN, accuracy of BTD for S-S shell, 0/60/0, a/h=75, R/a=5, , free-vibrations.

DOF	% Err. (FE)	% Err. (CNN)
15	0	0.0399
14	0.0007	0.0350
13	0.0018	0.0243
12	0.0041	0.0196
11	0.0067	0.0258
10	0.0431	0.0317
9	0.0952	0.0383
8	0.3641	0.1637
7	0.9189	0.5715
6	2.5535	2.4340
5	5.9604	5.4471
4	73.0429	75.9666
3	89.3867	92.3111

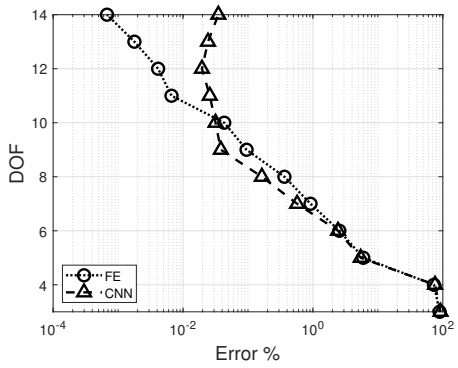


Figure 13: Multi-channel CNN, BTD for S-S shell, 0/60/0, a/h=75, R/a=5, , free-vibrations.

Table 12: Multi-channel CNN, best theories for S-S shell, 0/60/0, a/h=75, R/a=5, free-vibrations.

DOF	$u_{\alpha 1}$	$u_{\beta 1}$	u_{z1}	$u_{\alpha 2}$	$u_{\beta 2}$	u_{z2}	$u_{\alpha 3}$	$u_{\beta 3}$	u_{z3}	$u_{\alpha 4}$	$u_{\beta 4}$	u_{z4}	$u_{\alpha 5}$	$u_{\beta 5}$	u_{z5}
15	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
14	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	△	▲	▲
13	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	△	△	▲
12	▲	▲	▲	▲	▲	▲	△	▲	▲	▲	▲	▲	△	△	▲
11	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	△	△	△	△
10	▲	▲	▲	▲	▲	▲	△	△	▲	▲	▲	▲	△	△	△
9	▲	▲	▲	▲	▲	▲	△	△	▲	▲	▲	△	△	△	△
8	▲	▲	▲	▲	▲	▲	△	△	▲	▲	△	△	△	△	△
7	▲	▲	▲	▲	▲	△	△	△	▲	▲	△	△	△	△	△
6	▲	▲	▲	▲	▲	△	△	△	△	▲	△	△	△	△	△
5	▲	▲	▲	▲	▲	△	△	△	△	△	△	△	△	△	△
4	▲	▲	▲	▲	△	△	△	△	△	△	△	△	△	△	△
3	▲	▲	▲	△	△	△	△	△	△	△	△	△	△	△	△

- Overall, the results confirm that the proper refinement of shell models should prioritize the inclusion of in-plane third-order terms and transverse parabolic ones.

5 Conclusions

This paper has presented a new application of Convolutional Neural Networks (CNN) as surrogate models to evaluate the accuracy of structural theories and determine the best models. The focus was on shell structures made of composite materials. Through the Carrera Unified Formulation (CUF), the equations for static and free-vibration analyses were obtained, and the maximum order of the theories considered was four. The Node-Dependent Kinematics (NDK) approach was used to assign different structural theories to each node. The network's input consists of a combination of active expansion terms or theories distribution and configuration parameters such as thickness ratio, stacking sequence, and boundary conditions. The outputs were maximum displacements or natural frequencies. The results presented in this paper show that

- CNN performed better than fully-connected networks employed in previous works. CNN can handle more input features and has superior computational efficiency.
- CNN provided results for cases with features laying outside the ranges used in the training phase.
- The efficiency and versatility of CNN open to more generalized use of this technique. With further developments, the selection process of the best theory could completely avoid the necessity of FE computations for entire groups of structural configurations.

References

[1] E. Reissner and Y. Stavsky. Bending and stretching of certain types of heterogeneous aeolotropic elastic plates. *Journal of Applied Mechanics*, 28(3):402–408, 09 1961.

[2] J. M. Whitney and A. W. Leissa. Analysis of heterogeneous anisotropic plates. *Journal of Applied Mechanics*, 36(2):261–266, 06 1969.

[3] G. Kirchhoff. Über das gleichgewicht und die bewegung einer elastischen scheibe. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 1850(40):51–88, 1850.

[4] P.C. Yang, C. H. Norris, and Y. Stavsky. Elastic wave propagation in heterogeneous plates. *International Journal of Solids and Structures*, 2(4):665–684, 1966.

[5] E. Reissner. The effect of transverse shear deformation on the bending of elastic plates. *Journal of Applied Mechanics*, 12:69–77, 1945.

[6] R. D. Mindlin. Influence of rotatory inertia and shear on flexural motions of isotropic, elastic plates. *Journal of Applied Mechanics*, 18(1):31–38, April 2021.

[7] I. Babuska, J.M. d'Harcourt, and C. Schwab. Optimal shear correction factors in hierarchical plate modelling. Technical report, Maryland Univ., College Park. Inst. for Physical Science and Technology, 1991.

[8] S. Vlachoutsis. Shear correction factors for plates and shells. *International Journal for Numerical Methods in Engineering*, 33(7):1537–1552, 1992.

[9] B.F. Vlasov et al. On the equations of bending of plates. *Dokla Ak Nauk Azerbeijanskoi-SSR*, 3:955–979, 1957.

[10] J. N. Reddy. A simple higher-order theory for laminated composite plates. *Journal of Applied Mechanics*, 51(4):745–752, December 1984.

[11] J.N. Reddy and N.D. Phan. Stability and vibration of isotropic, orthotropic and laminated plates according to a higher-order shear deformation theory. *Journal of Sound and Vibration*, 98(2):157–170, 1985.

[12] F.B Hildebrand, E. Reissner, and G. B. Thomas. Notes on the foundations of the theory of small displacements of orthotropic shells. Technical report, Massachusetts Institute of Technology, 1949.

[13] R. K. Khare, V. Rode, A. K. Garg, and S. P. John. Higher-order closed-form solutions for thick laminated sandwich shells. *Journal of Sandwich Structures & Materials*, 7(4):335–358, 2005.

[14] A. K. Garg, R. K. Khare, and T. Kant. Higher-order closed-form solutions for free vibration of laminated composite and sandwich shells. *Journal of Sandwich Structures & Materials*, 8(3):205–235, 2006.

[15] H. Biglari and A. A. Jafari. High-order free vibrations of doubly-curved sandwich panels with flexible core based on a refined three-layered theory. *Composite Structures*, 92(11):2685–2694, 2010.

[16] R. K. Khare, A. K. Garg, and T. Kant. Free vibration of sandwich laminates with two higher-order shear deformable facet shell element models. *Journal of Sandwich Structures & Materials*, 7(3):221–244, 2005.

[17] C. P. Wu and C. C. Liu. Stress and displacement of thick doubly curved laminated shells. *Journal of Engineering Mechanics - ASCE*, 120(7):1403–1428, July 1994.

[18] P. H. Shah and R. C. Batra. Stretching and bending deformations due to normal and shear tractions of doubly curved shells using third-order shear and normal deformable theory. *Mechanics of Advanced Materials and Structures*, 25(15-16):1276–1296, 2018.

[19] S.M.R. Khalili, S. Tafazoli, and K. Malekzadeh Fard. Free vibrations of laminated composite shells with uniformly distributed attached mass using higher order shell theory including stiffness effect. *Journal of Sound and Vibration*, 330(26):6355–6371, 2011.

[20] P. Desai and T. Kant. On numerical analysis of axisymmetric thick circular cylindrical shells based on higher order shell theories by segmentation method. *Journal of Sandwich Structures & Materials*, 17(2):130–169, 2015.

[21] W. Zhen and C. Wanji. A global-local higher order theory for multilayered shells and the analysis of laminated cylindrical shell panels. *Composite Structures*, 84:350–361, August 2008.

[22] C. Ossadzow and M. Touratier. An improved shear-membrane theory for multilayered shells. *Composite Structures*, 52(1):85–95, 2001.

[23] A.J.M. Ferreira, E. Carrera, M. Cinefra, C.M.C. Roque, and O. Polit. Analysis of laminated shells by a sinusoidal shear deformation theory and radial basis functions collocation, accounting for through-the-thickness deformations. *Composites Part B: Engineering*, 42(5):1276–1284, 2011.

[24] J. L. Mantari, A. S. Oktem, and C. Guedes Soares. Bending and free vibration analysis of isotropic and multilayered plates and shells by using a new accurate higher-order shear deformation theory. *Composites Part B: Engineering*, 43(8):3348–3360, 2012.

[25] A. S. Sayyad and Y. M. Ghugal. Static and free vibration analysis of laminated composite and sandwich spherical shells using a generalized higher-order shell theory. *Composite Structures*, 219:129–146, 2019.

[26] M. Petrolo and E. Carrera. Methods and guidelines for the choice of shell theories. *Acta Mechanica*, 231:395–434, 2020.

[27] K. Washizu. *Variational methods in elasticity and plasticity.* Pergamon, Oxford, 1968.

[28] E. Carrera. Historical review of Zig-Zag theories for multilayered plates and shells . *Applied Mechanics Reviews*, 56(3):287–308, May 2003.

[29] J. N. Reddy. *Mechanics of laminated composite plates and shells: theory and analysis.* CRC press, 2003.

[30] D. H. Robbins, Jr. and J. N. Reddy. Modelling of thick composites using a layerwise laminate theory. *International Journal for Numerical Methods in Engineering*, 36(4):655–677, 1993.

[31] A. Dasgupta and K. H. Huang. A layer-wise analysis for free vibrations of thick composite spherical panels. *Journal of Composite Materials*, 31(7):658–671, 1997.

[32] M. Yaqoob Yasin and S. Kapuria. An efficient layerwise finite element for shallow composite and sandwich shells. *Composite Structures*, 98:202–214, 2013.

[33] P. Cicala. *Systematic approximation approach to linear shell theory.* Libreria Editrice Universitaria Levrotto e Bella, 1965.

[34] E. Antona and G. Frulla. Cicala's asymptotic approach to the linear shell theory. *Composite Structures*, 52:13–26, April 2001.

[35] E. Carrera and M. Petrolo. Guidelines and recommendations to construct theories for metallic and composite plates. *AIAA Journal*, 48(12):2852–2866, 2010.

[36] E. Carrera and M. Petrolo. On the effectiveness of higher-order terms in refined beam theories. *Journal of Applied Mechanics*, 78(2), 2011.

[37] S. Candiotti, J.L. Mantari, J. Yarasca, M. Petrolo, and E. Carrera. An axiomatic/asymptotic evaluation of best theories for isotropic metallic and functionally graded plates employing non-polynomic functions. *Aerospace Science and Technology*, 68:179–192, 2017.

[38] M. Petrolo, M. Cinefra, A. Lamberti, and E. Carrera. Evaluation of mixed theories for laminated plates through the axiomatic/asymptotic method. *Composites Part B: Engineering*, 76:260–272, 2015.

[39] D. S. Mashat, E. Carrera, A. M. Zenkour, and S. A. Al Khateeb. Axiomatic/asymptotic evaluation of multilayered plate theories by using single and multi-points error criteria. *Composite Structures*, 106:393–406, 2013.

[40] E. Carrera. Theories and finite elements for multilayered plates and shells: A unified compact formulation with numerical assessment and benchmarking. *Archives of Computational Methods in Engineering*, 10:215–296, September 2003.

[41] E. Carrera and E. Zappino. One-dimensional finite element formulation with node-dependent kinematics. *Computers & Structures*, 192:114–125, 2017.

[42] E. Carrera, A. Pagani, and S. Valvano. Shell elements with through-the-thickness variable kinematics for the analysis of laminated composite and sandwich structures. *Composites Part B: Engineering*, 111:294–314, 2017.

[43] B. Cheng and D. M. Titterington. Neural networks: A review from a statistical perspective. *Statistical Science*, 9(1):2 – 30, 1994.

[44] M. Zakaria, M. Al-Shebany, and S. Sarhan. Artificial neural network: a brief overview. *International Journal of Engineering Research and Applications*, 4(2):7–12, 2014.

[45] H. T. Thai. Machine learning for structural engineering: A state-of-the-art review. *Structures*, 38:448–491, 2022.

[46] G. Balokas, S. Czichon, and R. Rolfes. Neural network assisted multiscale analysis for the elastic properties prediction of 3d braided composites under uncertainty. *Composite Structures*, 183:550–62, 2018.

[47] P. A. Gustafson, E. J. Pineda, T. M. Ricks, B. A. Bednarcyk, B. L. Hearley, and J. Stuckner. A convolutional neural network for enhancement of multi-scale localization in granular metallic representative unit cells. In *AIAA SCITECH 2022 Forum*, page 0078, 2022.

[48] MG. Vineela, A. Dave, and P.K. Chaganti. Artificial neural network based prediction of tensile strength of hybrid composites. *Materials Today: Proceedings*, 5(9, Part 3):19908–15, 2018.

[49] A. Pagani, M. Enea, and E. Carrera. Component-wise damage detection by neural networks and refined fes training. *Journal of Sound and Vibration*, 509:116255, 2021.

[50] J. Gajewski, P. Golewski, and T. Sadowski. Geometry optimization of a thin-walled element for an air structure using hybrid system integrating artificial neural network and finite element method. *Composite Structures*, 159:589–599, 2017.

[51] F. Tao, X. Liu, H. Du, S. Tian, and W. Yu. Discover failure criteria of composites from experimental data by sparse regression. *Composites Part B: Engineering*, 239:109947, 2022.

[52] C.A. Yan, R. Vescovini, and L. Dozio. A framework based on physics-informed neural networks and extreme learning for the analysis of composite structures. *Computers and Structures*, 265:106761, 2022.

[53] K. D. Humfeld, D. Gu, G. A. Butler, K. Nelson, and N. Zobeiry. A machine learning framework for real-time inverse modeling and multi-objective process optimization of composites for active manufacturing control. *Composites Part B: Engineering*, 223:109150, 2021.

[54] K. O'Shea and R. Nash. An introduction to convolutional neural networks. *ArXiv*, abs/1511.08458, 2015.

[55] S. Albawi, T. A. Mohammed, and S. Al-Zawi. Understanding of a convolutional neural network. In *2017 International Conference on Engineering and Technology (ICET)*, pages 1–6, 2017.

[56] Jiuxiang Gu, Zhenhua Wang, Jason Kuen, Lianyang Ma, Amir Shahroudy, Bing Shuai, Ting Liu, Xingxing Wang, Gang Wang, Jianfei Cai, and Tsuhan Chen. Recent advances in convolutional neural networks. *Pattern Recognition*, 77:354–377, 2018.

[57] M. Petrolo and E. Carrera. On the use of neural networks to evaluate performances of shell models for composites. *Advanced Modeling and Simulation in Engineering Sciences*, 7(31), 2020.

[58] M. Petrolo and E. Carrera. Selection of element-wise shell kinematics using neural networks. *Computers and Structures*, 244(106425), 2021.

[59] Carrera E. Theories and finite elements for multilayered plates and shells: a unified compact formulation with numerical assessment and benchmarking. *Archives of Computational Methods in Engineering*, 10(3):216–296, 2003.

[60] K. J. Bathe and E. N. Dvorkin. A formulation of general shell elements—the use of mixed interpolation of tensorial components. *International Journal for Numerical Methods in Engineering*, 22(3):697–722, 1986.

[61] G. Li, E. Carrera, M. Cinefra, A.G. de Miguel, A. Pagani, and E. Zappino. An adaptable refinement approach for shell finite element models based on node-dependent kinematics. *Composite Structures*, 210:1–19, 2019.

[62] E. Carrera, M. Cinefra, M. Petrolo, and E. Zappino. *Finite Element Analysis of Structures through Unified Formulation*. John Wiley & Sons, Chichester, 2014.

[63] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980, 2015.

[64] J. Wu. Introduction to convolutional neural networks. *National Key Lab for Novel Software Technology. Nanjing University. China*, 5(23):495, 2017.