

Article

Not peer-reviewed version

Fuzzy-Based Convolutional Neural Network Model for Structural Response Prediction Under Seismic Excitation

[Mohammad Sadegh Barkhordari](#)^{*} and Mohammad Mahdi Barkhordari

Posted Date: 4 November 2024

doi: 10.20944/preprints202410.0956.v2

Keywords: Seismic excitation; Fuzzy logic-based network; Convolutional neural network; Seismic response prediction



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Fuzzy-Enhanced Convolutional Neural Network for Predicting Structural Responses to Seismic Excitations

Mohammd Sadegh Barkhordari ^{1,*} and Mohammad Mahdi ²

¹ AmirKabir university, Iran

² School of Medicine, Kerman University of Medical Sciences, Kerman 7616914115, Iran

* Correspondence: m.s.barkhordari@aut.ac.ir

Abstract: This study tackles the challenge of accurately predicting structural behavior under seismic excitation—a critical need in designing and assessing resilient buildings and infrastructure. Traditional approaches, including numerical modeling and differential equation solvers, often encounter substantial computational challenges, particularly when addressing nonlinear hysteretic behavior or large-scale structures. To address these issues, we present a novel Fuzzy-Based Convolutional Neural Network (FuzzyCNN) model, which combines fuzzy logic with convolutional neural networks to effectively capture and manage the uncertainties and complexities of soil-structure interaction under seismic loading. The model's efficacy is validated through both numerical simulations and experimental data from a mid-rise concrete building subjected to seismic events. Comparative analysis reveals that the FuzzyCNN outperforms a conventional Physics-Informed CNN (PhyCNN) model, offering superior accuracy and robustness in predicting seismic responses. Key findings indicate that the FuzzyCNN not only enhances predictive accuracy but also better handles uncertainties than the PhyCNN. This advancement has significant implications for engineering design, seismic risk assessment, and the development of resilient infrastructure, promising improved efficiency and accuracy in dynamic response prediction.

Keywords: fuzzy logic-based network; convolutional neural network; seismic response prediction; seismic excitation

1. Introduction

Dynamic response prediction of structural systems plays a pivotal role in both the design and evaluation of individual buildings and the reliability assessment of infrastructure and large urban areas [1]. Conventionally, this process entails the creation of numerical models for dynamic systems, with response prediction facilitated by numerical differential equation solvers like the Newmark- β method [2]. The finite element method (FEM) stands out as a prominent technique in this domain, extensively employed across various engineering disciplines such as civil, mechanical, and aeronautical engineering [3]. Despite significant advancements in computational capabilities, the escalating complexity of numerical models poses formidable challenges, especially in dealing with large-scale engineering problems characterized by nonlinear hysteretic behaviors under dynamic loads [4]. The computational burden becomes particularly onerous when addressing optimization tasks and handling stochastic uncertainties associated with external loads. Techniques like Monte Carlo simulations and incremental dynamic analysis (IDA) are indispensable in accounting for these uncertainties, but they significantly augment the computational cost [5].

When assessing structures under seismic excitation, it is crucial to account for the variability in frequency content across different recorded earthquakes [6]. Each earthquake has unique characteristics, and relying on a single record may not capture the full range of possible seismic responses. Therefore, in order to perform a thorough evaluation, it is advisable to use a large set of

records, for instance 80 different earthquake recordings [2]. Additionally, to understand the structure's performance under increasing seismic demand, the intensity of each earthquake record should be scaled multiple times, for example, ten intensity levels per record. This process helps capture how a structure behaves under a wide spectrum of seismic intensities. For instance, consider an earthquake record that lasts 50 seconds and is sampled at a time step of 0.005 seconds. This results in 10,000 data points for that single record. When performing an (IDA) using 80 earthquake records and scaling each record 10 times, the total number of data points grows significantly. Specifically, the analysis will involve $80 \times 10 \times 10,000 = 8,000,000$ data points[7].

In addition to the large number of data points generated from the earthquake records, the structural models themselves involve complex equations that include large matrices representing stiffness, mass, damping, and other dynamic properties [8]. These matrices are essential for accurately capturing the behavior of structures under seismic loads. For example, the stiffness matrix defines how resistant a structure is to deformation, while the mass matrix accounts for how inertia affects the structure's response to motion. The damping matrix, meanwhile, represents energy dissipation mechanisms within the structure. This challenge persists despite the availability of high-performance computing clusters or facilities. Consequently, the quest for efficient computational methodologies to tackle these complexities remains a pressing concern in the realm of structural dynamic analysis and design [9].

In response to the computational challenges posed by complex engineering simulations, researchers have delved into the realm of metamodeling as a means to alleviate the computational burden [10]. Metamodels serve as reduced-fidelity surrogate models of high-fidelity simulations, aiming to capture the input-output relationships of systems while significantly reducing computational costs [10]. Traditionally, techniques such as regression and response surface methodology (RSM) have been prevalent in metamodeling, relying on polynomial least-square fitting [11]. However, the simplicity of these methods and their dependence on second-order polynomials often lead to insufficient accuracy, particularly in capturing highly nonlinear system behaviors [11]. To address these limitations, alternative metamodeling approaches have emerged, including Kriging, support vector machine, and polynomial chaos expansions, among others [12–14]. These methods offer promising avenues for uncertainty quantification and have found applications in various engineering domains.

However, the calibration of high-fidelity models, especially those with numerous parameters, demands significant computational resources [15]. To mitigate this challenge, model order reduction techniques have been developed [15]. These techniques enable the creation of reduced-fidelity metamodels that approximate the behavior of complex systems while substantially reducing computational costs. Despite these advancements, current methodologies often struggle to address highly nonlinear structures under non-stationary conditions [15]. This limitation underscores the ongoing need for innovative approaches to metamodeling, particularly in the context of complex engineering systems subjected to dynamic loads and nonlinear behaviors. Efforts to bridge this gap are vital for advancing the efficiency and accuracy of computational simulations in engineering design and analysis. To extend the applicability and accuracy of dynamic response predictions, ongoing research endeavors are exploring innovative methodologies. These include leveraging machine learning (ML) techniques to refine model updating processes and enhance the predictive capabilities of numerical simulations.

Zhanga et al. [15] have developed an approach to metamodeling nonlinear structural systems with limited data by integrating partial physics knowledge into deep long short-term memory (LSTM) networks. Their framework ensures precise model training by incorporating physics constraints into the loss function, effectively capturing system nonlinearity. Tailored for dynamic structures, the framework considers equations of motion, state dependence, and hysteretic constitutive relationships. The embedded physics mitigate overfitting, reduce dependency on extensive training datasets, and enhance model robustness, surpassing that of traditional data-driven neural networks. Eshkevari et al. [16] present a physics-informed recurrent neural network (RNN) model designed to predict the dynamics of multi-degree-of-freedom systems subjected to diverse

ground motions. The model predicts displacement, velocity, acceleration, and internal forces, outperforming existing models with improved predictive accuracy and fewer parameters. Inspired by differential equation solvers, the recurrent block architecture aims for more generalized solutions. Innovative training techniques such as hard sampling, trajectory loss function, and trust-region optimization expedite learning with limited datasets. Zhang et al. [17] present a novel approach using a physics-guided convolutional neural network (PhyCNN) for data-driven modeling of structural seismic responses. Their method integrates physics constraints into deep learning models, allowing for accurate prediction of structural responses with limited seismic datasets. Numerical simulations and experiments confirm PhyCNN's acceptable performance in predicting seismic responses.

Xu et al. [18] introduce an innovative physics-informed ML approach leveraging Convolutional Neural Networks (CNNs) to accurately reconstruct both dynamic and static displacements of plate-beam composite structures. Integrating past acceleration-related physics knowledge into the deep learning framework is the core concept. To improve physics-based training efficiency, a dual-branch network architecture is developed to extract extensive details from inputs. Multi-task weights are adjusted using carefully selected scales and punishment factors. A comparison of their suggested methodology with computer simulations and empirical validations in the statics and structural dynamics domains demonstrates its superior performance. Incorporating acceleration data improves the trained model's robustness for more accurate displacement predictions while also drastically lowering measurement costs. Hu et al. [19] presented a new neural network, merging physics-informed neural networks with model-based transfer learning, for better seismic response predictions with limited data. Integrating physics knowledge, represented by a new solver, enhances the network's ability to capture structural nonlinearities. Model-based transfer learning improves generalization by transferring features from a source to target buildings. Their model outperforms data-driven networks in predicting seismic responses across target buildings from numerical models to physical shaking table tests.

There are persistent challenges in training a dependable deep learning model for estimating the response of nonlinear structures subjected to seismic excitation. Notably, the process of solving partial differential equations (PDEs) or governing equations within Physics-Informed Neural Networks (PINNs) often entails iterative numerical methods, which incur significant computational costs, particularly for large-scale structural configurations. Consequently, this computational overhead can impede real-time applications and the execution of extensive simulations crucial for comprehensive seismic risk evaluation. Moreover, the application of physics-informed neural networks to intricate and inherently chaotic systems, such as those governed by soil-structure interaction or characterized by rocking behavior [20], remains arduous due to the inherent limitations in handling the heightened complexity of the data, surpassing the capabilities of current physics-guided models [19]. In addition, while previous studies introduced techniques that excel in generating intricate reconstructions of the time-history response based on their training datasets, they primarily cater to relatively straightforward structures. These structures typically include single-degree-of-freedom systems, shear-type multi-degree-of-freedom structures with hysteresis modeled numerically, or low-rise buildings with fewer than three stories. However, their applicability may be limited when it comes to capturing the intricate behaviors arising from material and geometric nonlinearities across a spectrum of building heights.

In this study, a novel approach is proposed by integrating fuzzy logic principles into a CNN framework, termed FuzzyCNN, to effectively model and predict the dynamic behavior of soil-structure systems under seismic excitations. The FuzzyCNN model is designed to handle the uncertainties and complexities inherent in such systems, leveraging the interpretability of fuzzy logic combined with the robust feature extraction capabilities of CNNs. The primary objectives of this research are to develop a FuzzyCNN model capable of accurately predicting the responses of structures subjected to earthquake loads and to compare its performance against traditional physics-informed CNN (PhyCNN) models. Specifically, we aim to address the following research questions: How does the integration of fuzzy logic into a CNN affect the model's accuracy and robustness? Can the FuzzyCNN model outperform existing PhyCNN models in terms of prediction accuracy and

handling of uncertainties? The remainder of this paper is structured as follows: Section 2 details the methodology, including the development of the FuzzyCNN model and the numerical techniques employed for its validation. Section 3 presents the results and discussion, providing a comparative analysis of the FuzzyCNN and PhyCNN models based on various performance metrics. Finally, Section 4 concludes the paper, highlighting the key findings and potential future research directions.

2. Problem Definition

Dynamic equilibrium equations considering soil-structure interaction and seismic excitation typically involve accounting for the behavior of both the structure and the underlying soil medium. These equations describe the balance of forces acting on the system while considering the dynamic response to seismic loading. They can be expressed in matrix form as [21]:

$$[M]\{\ddot{r}(t)\} + [C]\{\dot{r}(t)\} + [K_{fs}]\delta + [K_s]u = \{R(t)\} \quad (1)$$

where $[M]$ and $[C]$ represent the mass and damping matrices of the system, respectively. K_s is the stiffness matrix of the structure, K_{fs} is the interaction stiffness matrix between the structure and the soil, u denotes the displacement vector of the structure, and δ denotes the relative displacement vector between the structure and the soil. \ddot{r} and \dot{r} denote the acceleration and velocity vectors, while $R(t)$ signifies the vector of reactions from the time-dependent springs and dampers that have replaced the soil or the external force vector due to seismic excitation. It's possible to merge $[K_{fs}]$ and $[K_s]$ into a single term, representing the total stiffness effect on the system. This combined stiffness term accounts for both the intrinsic stiffness of the structure and the stiffness due to interaction with the soil. It simplifies the dynamic equilibrium equations for a system subjected to seismic excitation, depicted in Figure 1 [22]:

$$[M]\{\ddot{r}(t)\} + [C]\{\dot{r}(t)\} + [K]\{r(t)\} = \{R(t)\} \quad (2)$$

where $[K]$ represents the total stiffness matrix, combining both the structural stiffness and the interaction stiffness.

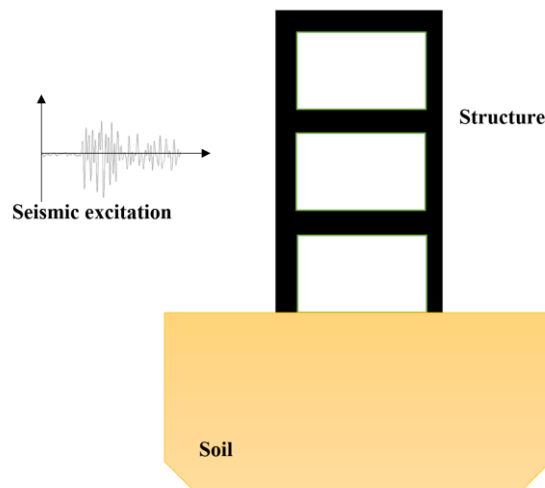


Figure 1. Diagram illustrating the dynamic interaction between soil and structure.

The dimensions of these matrices and vectors correspond to the total degrees of freedom in the system. Notably, the coefficients of the springs and dampers are analogous to the support's flexibility concerning vibration frequency. Hence, a transformation from the time domain to the frequency domain facilitates an easier solution of Equation (1). This transformation can be achieved through the use of the Fourier transform (F_r), as depicted in Equation (3) [23]:

$$\begin{aligned} F_r\{r(t)\} &= \{u(\omega)\} \\ F_r\{R(t)\} &= \{P(\omega)\} \end{aligned} \quad (3)$$

Moreover, leveraging the properties of the Fourier transform, Equation (4) emerges [24]:

$$\begin{aligned} (-\omega^2[M] + i\omega[C] + [K])\{u(\omega)\} &= \{P(\omega)\} \\ [S(\omega)] &= (-\omega^2[M] + i\omega[C] + [K]) \quad (4) \\ [S(\omega)]\{u(\omega)\} &= \{P(\omega)\}, i = \sqrt{-1} \end{aligned}$$

where $[S(\omega)]$ represents the dynamic stiffness matrix of the structure. In addition, Rayleigh damping is commonly assumed, where damping is proportional to both velocity and stiffness, formulated as [25]:

$$[C] = (\alpha[M] + \beta[K]) \quad (5)$$

Here, α and β are constant coefficients derived from the system's dynamic characteristics. In the domain of structural-soil interaction, a mixed stiffness matrix is employed in the frequency domain, comprising both real and imaginary components [25,26]. This mixed stiffness matrix signifies a phase discrepancy between the applied load and the resulting reaction, indicating a delay in the reaction relative to the applied load due to system damping. Upon comparing equations (1) and (3), it becomes apparent that solving the dynamic equilibrium equation of the system in the frequency domain, i.e., is notably more straightforward using fundamental principles of algebraic equations [25,26]. However, a significant challenge lies within these equations: besides the unknown displacements $\{u(\omega)\}$, the Fourier transform of the reaction vector $R(t)$ is also unknown, as $R(t)$ represents the forces acting on the flexible support, inherently ambiguous. Consequently, the system presents one equation with two unknowns, resulting in an indeterminate problem [25,26].

3. Methodology

3.1. Fuzzy-Based Convolutional Neural Network (FuzzyCCN)

This research examines soil-structure systems undergoing forced vibrations. The system is characterized by: (1) Input: a set of excitations represented by the vector ($a(t)$); and (2) Output: a set of responses represented by the vector ($d(t)$). Here, $a(t)$ is ground acceleration and $d(t)$ is the relative displacements. Assuming that we have no knowledge of the actual physical structure or mechanics of the system (treating it as a "black box"), a FuzzyCNN architecture is developed to create a substitute model that mimics the behavior of the original system. This approach suggests using a data-driven method to model a complex system whose internal workings are unknown. The FuzzyCNN will attempt to learn the relationship between inputs and outputs, effectively creating a surrogate model that can predict the system's behavior without understanding its underlying physical principles.

The proposed framework consists of a series of 1D-convolution layers with fuzzy membership functions and ReLU activations. The 1D convolutional layer can be expressed as:

$$y(t) = \sum_{k=0}^{K-1} a(t-k) \cdot w(k) + b \quad (6)$$

where $a(t)$ is the input, w are the filter weights, b is the bias, and K is the filter size. This operation is repeated across the input sequence. The Gaussian fuzzy membership function applied to the output of each convolutional layer is given by:

$$G(y) = \exp(-0.5 \cdot (\frac{y - \mu_x}{\sigma_x + \varepsilon})^2) \quad (7)$$

where μ_x is the mean of y , σ_x is the standard deviation of y , and ε is a small constant to avoid division by zero. The fuzzy membership function enhances the interpretability and robustness of the model by incorporating fuzzy logic principles [27,28]. After the convolution and fuzzy membership operations, the ReLU (Rectified Linear Unit) activation function is applied:

$$\text{ReLU}(x) = \max(0, G(y)) \quad (8)$$

Combining these steps, each layer in the fuzzy-CNN model can be represented as:

$$x^{(l+1)} = \text{ReLU}(G(w^l * a^l + b^l)) \quad (9)$$

where $*$ denotes the convolution operation, and l denotes the layer index.

The transformation matrix T_t is used to compute the first and second derivatives of the outputs of FuzzyCNN ($d(t)$). The finite difference method is a numerical technique for estimating derivatives of functions [29]. Given discrete data points $d_i = d(t_i)$ at time t_i , the derivatives of d can be approximated using finite differences [29]. Here, both the first-order and second-order derivatives are approximated. The first and second derivatives of d are computed as [29]:

$$\begin{aligned} \dot{d} &= T_t d \\ \ddot{d} &= T_t \dot{d} = T_t (T_t d) = T_t^2 d \end{aligned} \quad (10)$$

The form of the T_t matrix that approximates the derivative of the time series data $d(t)$ is [29]:

$$T_t = \frac{1}{\Delta t} \begin{bmatrix} -\frac{3}{2} & 2 & -\frac{1}{2} & 0 & \dots & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} & 0 & \dots & 0 \\ 0 & -\frac{1}{2} & 0 & \frac{1}{2} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & -\frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \dots & 0 & \frac{1}{2} & -2 & \frac{3}{2} \end{bmatrix} \quad (11)$$

Where Δt is the time step size. Each row of T_t represents the coefficients used to approximate the derivative at a particular point in the time series. The first row uses forward differences, the last row uses backward differences, and the interior rows use central differences [29].

The loss function combines the mean squared error between the predicted and true \ddot{d} , and a regularization term that penalizes the first values of the predicted d :

$$\text{loss} = \frac{1}{N} \sum_{i=1}^N (\ddot{d}_i^{\text{true}} - \ddot{d}_i^{\text{pred}})^2 + \sum_{i=1}^{N/100} (d_i^{\text{pred}})^2 \quad (12)$$

where N is the number of data points. The regularization term ($\sum_{i=1}^{N/100} (d_i^{\text{pred}})^2$) based on the first

$N/100$ values of d^{pred} is designed to constrain the model predictions, ensuring they do not deviate significantly from expected values or exhibit undesired behavior [30,31]. Regularization is a technique used in ML to prevent overfitting by adding a penalty to the loss function [31]. This term can also guide the model to learn specific properties of the data, especially when dealing with physical systems where certain behaviors or values are expected [31].

The Adam optimizer updates the model parameters based on the gradients of the loss function with respect to the parameters. The update rule for parameter θ is [32]:

$$\begin{aligned}
m_t &= \beta_1 m_{t-1} + (1 - \beta_1) g_t, \quad v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \\
\hat{m}_t &= \frac{m_t}{1 - \beta_1^t}, \quad \hat{v}_t = \frac{v_t}{1 - \beta_2^t} \\
\theta_{t+1} &= \theta_t - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \varepsilon}
\end{aligned} \tag{13}$$

where m_t and v_t are moving averages of the gradient and its square, respectively, g_t is the gradient at time step t , α is the learning rate (in this case, a value of 0.001 is used), β_1 is the exponential decay rate for the first moment estimates, β_2 is the exponential decay rate for the second moment estimates, and ε is a small constant (e.g., 10^{-8}) to avoid division by zero. Typically, default exponential decay rate values are $\beta_1 = 0.9$ and $\beta_2 = 0.999$ [32].

Figure 2 illustrates the architecture of the proposed FuzzyCNN model, along with the associated training and testing framework. This configuration is specifically designed to effectively extract temporal features from univariate time series data. Optimal parameters are determined using the GridSearch algorithm. The proposed model comprises four series of 1D convolutional layers, each incorporating fuzzy membership functions and utilizing ReLU (Rectified Linear Unit) activations. The final layer of the framework is Dense layers with 50 neurons, which serves to consolidate the extracted features and produce the desired output. Table 1 provides a comprehensive overview of the FuzzyCNN layers, detailing their specifications and parameters.

Table 1. The characteristics of the FuzzyCNN layers.

Parameter	Value
Activation functions	ReLU
Number of filters	64
Size of the convolutional filter	50
Stride of the convolution	1
Padding	Same
Use bias	True
Last layer	Dense with 50 neurons

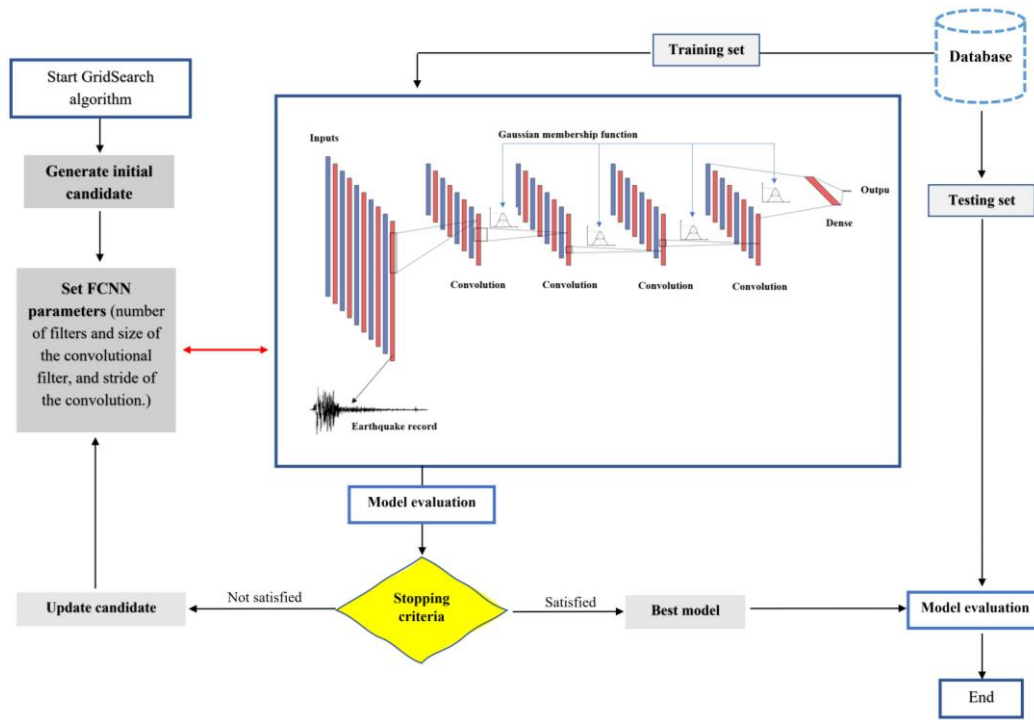


Figure 2. Architecture of the proposed FuzzyCNN model and training/testing framework.

3.2. Numerical Validation

The efficiency of the suggested FuzzyCNN model is initially evaluated through a numerical case study adapted from the research conducted by Zhang and colleagues [17]. The case study focuses on a one-degree-of-freedom (SDOF) structural system, which is subjected to earthquake records. The dynamic behavior of this nonlinear system is mathematically described by the following equation of motion [17]:

$$m\ddot{u} + c\dot{u} + k_1u + k_2u^3 = -m\ddot{u}_g \quad (14)$$

where $m = 1 \text{ kg}$; $c = 1 \text{ Ns} / \text{m}$; $k_1 = 20 \text{ N} / \text{m}$; $k_2 = 200 \text{ N} / \text{m}$

where m is the mass; u , \dot{u} , and \ddot{u} are the relative displacement, velocity, and acceleration; and \ddot{u}_g is the ground motion acceleration. Eighty-five samples, each of which represents a separate seismic sequence, are created artificially through numerical simulations of a nonlinear system. Selected earthquake recordings that represent a 10% probability of exceedance during a 50-year span are taken from the PEER database [33,34]. With a sampling frequency of 20 Hz, each simulation lasts 50 seconds, yielding 1,001 data points for each record. Fifteen percent of these datasets are used as a test and validation set to gauge prediction accuracy, while seventy percent are randomly selected for training. The FuzzyCNN model is also tested against Zhang et al.'s PhyCNN model [17], which incorporates physics-informed loss functions to improve model accuracy.

3.3. PhyCNN Model

Zhang et al. presented the PhyCNN framework designed for seismic response modeling, which integrates deep learning with fundamental physical principles, specifically leveraging dynamic equations to enhance prediction accuracy in scenarios with limited datasets. PhyCNN bridges data-driven methods and first-principles modeling by embedding physical constraints—namely, the equations of motion—directly into the learning process.

The structural dynamic behavior under seismic excitation is governed by Newton's second law, expressed as:

$$M\ddot{x}(t) + h(t) = -M\Gamma\ddot{x}_g(t) \quad (15)$$

where M represents the mass matrix, $\ddot{x}(t)$ is the relative acceleration vector, $h(t)$ denotes the generalized restoring force, Γ is the influence vector representing how ground motion is distributed across the structure, and $\ddot{x}_g(t)$ is the external ground acceleration. This equation forms the backbone of the PhyCNN model, guiding its predictions of structural displacements, velocities, and restoring forces. By embedding this dynamic equation into the architecture, PhyCNN ensures that predictions adhere to the underlying physics of the system, rather than relying purely on empirical data. The PhyCNN employs a composite loss function that combines both data-driven and physics-informed components. The data loss, $J_D(\theta)$, quantifies the discrepancy between the predicted and observed structural states (displacement, velocity, and restoring force) as:

$$J_D(\theta) = \frac{1}{N} \sum_{i=1}^N (\|x_c - x_o\|^2 + \|\dot{x}_c - \dot{x}_o\|^2 + \|RF_p - RF_o\|^2) \quad (16)$$

In addition, the physics loss, $J_{phy}(\theta)$, imposes constraints that ensure the predicted states are consistent with the system's dynamic behavior:

$$J_{phy}(\theta) = \frac{1}{N} \sum_{i=1}^N (\|\ddot{x} + \Gamma \ddot{x}_g\|^2 + \underbrace{\frac{1}{M} \|\dot{x} + x\|^2}_{h(t)}) \quad (17)$$

where x_c , \dot{x}_c , and RF_p are the computed displacement, velocity, and restoring force, respectively, and x_o , \dot{x}_o , and RF_o represent the measured values. The overall loss function used for optimization is a weighted sum of these components:

$$J(\theta) = J_{phy}(\theta) + J_D(\theta) \quad (18)$$

This composite loss function integrates both the empirical data and the governing physical laws into the training process. By doing so, the PhyCNN framework try to enforces physically plausible predictions that respect the system's dynamic behavior while simultaneously minimizing prediction error with respect to the available data.

3.4. Experimental Validation

Field sensing data is used to further evaluate the performance of the FuzzyCNN model. A six-story mid-rise concrete hotel building located in San Bernardino, California, constructed in 1970, is chosen [35]. Accelerometers installed on the third floor in both longitudinal and transverse directions are employed, with their specific placement detailed in Figure 3 [35]. Twenty-three seismic events recorded by sensors from 1987 to 2018 are used in this study [17,35]. The acceleration data are processed using a specific type of filter to eliminate low-frequency noise and artifacts, ensuring that only the relevant higher-frequency seismic data are retained for analysis [17,35]. Here, a Butterworth high-pass filter is used [35]. A Butterworth filter is a type of signal-processing filter designed to have a flat frequency response in the passband [17,35]. Two reactive components (2-pole filter) determine its frequency characteristics. A high-pass filter allows high-frequency signals to pass through while attenuating (reducing) low-frequency signals. The cutoff frequency is the threshold frequency where the filter begins to attenuate frequencies below this value. In this case, any frequency components of the acceleration data below 0.1 Hz will be significantly reduced [17,35]. Low-frequency behavior in this context refers to unwanted low-frequency components in the measured acceleration data, such as noise or drift, which can obscure the meaningful high-frequency data related to seismic events. Moreover, given that this example incorporates sensor data from an actual building, the response of the structure inherently includes the effects of the soil-structure interaction and the influence of the surrounding soil on the foundation. This integration ensures that the dynamic response captured by the sensors reflects realistic conditions, accounting for the complex interplay between the building and its subsoil environment. Detailed analysis of these effects is elaborated in section 2 of the article.

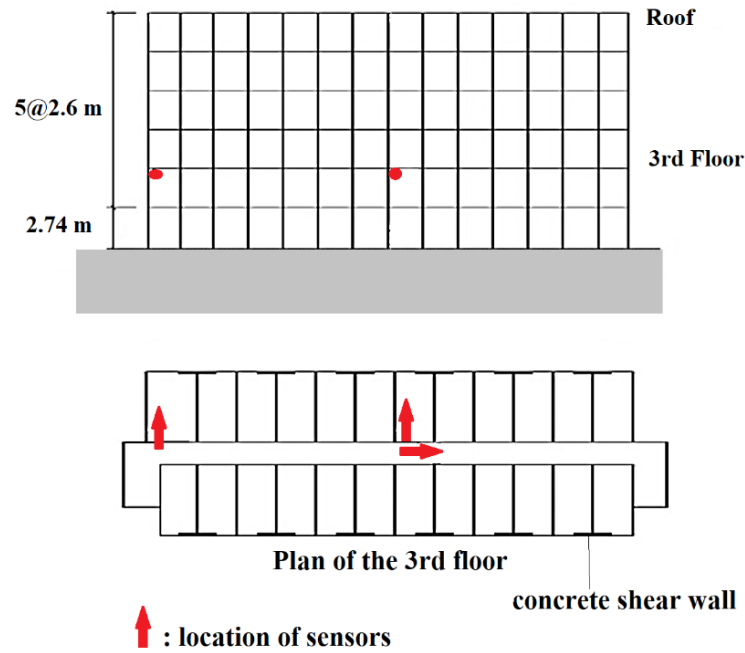


Figure 3. Sensor locations in the 6-story building.

4. Results and Discussion

Figures 4 and 5 show the architecture of the FuzzyCNN and PhyCNN models. The number of parameters in a neural network model is a critical factor that can affect the model's complexity, training time, and risk of overfitting. The proposed model (FuzzyCNN) has 621,157 parameters, while PhyCNN has 828,571 parameters. The PhyCNN model has approximately 207,414 more parameters than the FuzzyCNN model. This larger parameter count suggests that PhyCNN incurs increased computational costs and carries a potentially higher risk of overfitting, particularly when dealing with limited data [36].

Figure 6 illustrates the predicted behavior of the SDOF structural system using the FuzzyCNN model. The plot represents the displacement of the structural response. The figure demonstrates the model's ability to capture the nuanced responses of an SDOF system under various conditions. Figure 7 presents the prediction results of the PhyCNN model by Zhuang and colleagues [17] for the same SDOF structural system as in Figure 6. It provides a comparative basis for evaluating the performance of the FuzzyCNN model. The PhyCNN model is another neural network-based approach that focuses on leveraging physical principles in the prediction process. The comparison between FuzzyCNN and PhyCNN predictions is essential to understand the advantages and limitations of each model. While the PhyCNN might excel in scenarios where physical laws are well-defined and can be directly incorporated, it may struggle with uncertainties or data-driven nuances that the FuzzyCNN can address through its fuzzy logic components.

To facilitate a more robust comparison of the results between two models, the correlation coefficient has been employed. The correlation coefficient is a statistical measure used to assess the strength and direction of a linear relationship between two variables.

$$R^2 = 1 - \frac{\sum_{i=1}^n (a_i - p_i)^2}{\sum_{i=1}^n (a_i - \bar{a})^2} \quad (15)$$

where n represents the number of data points, a_i is the recorded values, p_i is the computed values, and \bar{a} is the mean of the recorded values. In this context, it measures the correlation between predicted and actual values for the structural system's response.

Layer (type)	Output Shape	Param #
conv1d_54 (Conv1D)	(None, None, 64)	3,264
lambda_16 (Lambda)	(None, None, 64)	0
activation_69 (Activation)	(None, None, 64)	0
conv1d_55 (Conv1D)	(None, None, 64)	204,864
lambda_17 (Lambda)	(None, None, 64)	0
activation_70 (Activation)	(None, None, 64)	0
conv1d_56 (Conv1D)	(None, None, 64)	204,864
lambda_18 (Lambda)	(None, None, 64)	0
activation_71 (Activation)	(None, None, 64)	0
conv1d_57 (Conv1D)	(None, None, 64)	204,864
lambda_19 (Lambda)	(None, None, 64)	0
activation_72 (Activation)	(None, None, 64)	0
dense_26 (Dense)	(None, None, 50)	3,250
activation_73 (Activation)	(None, None, 50)	0
dense_27 (Dense)	(None, None, 1)	51

Total params: 621,157 (2.37 MB)
 Trainable params: 621,157 (2.37 MB)
 Non-trainable params: 0 (0.00 B)

Figure 4. Overview of the FuzzyCNN's architecture.

Layer (type)	Output Shape	Param #
conv1d_49 (Conv1D)	(None, None, 64)	3,264
activation_62 (Activation)	(None, None, 64)	0
conv1d_50 (Conv1D)	(None, None, 64)	204,864
activation_63 (Activation)	(None, None, 64)	0
conv1d_51 (Conv1D)	(None, None, 64)	204,864
activation_64 (Activation)	(None, None, 64)	0
conv1d_52 (Conv1D)	(None, None, 64)	204,864
activation_65 (Activation)	(None, None, 64)	0
conv1d_53 (Conv1D)	(None, None, 64)	204,864
activation_66 (Activation)	(None, None, 64)	0
dense_23 (Dense)	(None, None, 50)	3,250
activation_67 (Activation)	(None, None, 50)	0
dense_24 (Dense)	(None, None, 50)	2,550
activation_68 (Activation)	(None, None, 50)	0
dense_25 (Dense)	(None, None, 1)	51

Total params: 828,571 (3.16 MB)
 Trainable params: 828,571 (3.16 MB)
 Non-trainable params: 0 (0.00 B)

Figure 5. Overview of the PhyCNN's architecture (Zhuang and colleagues [17]).

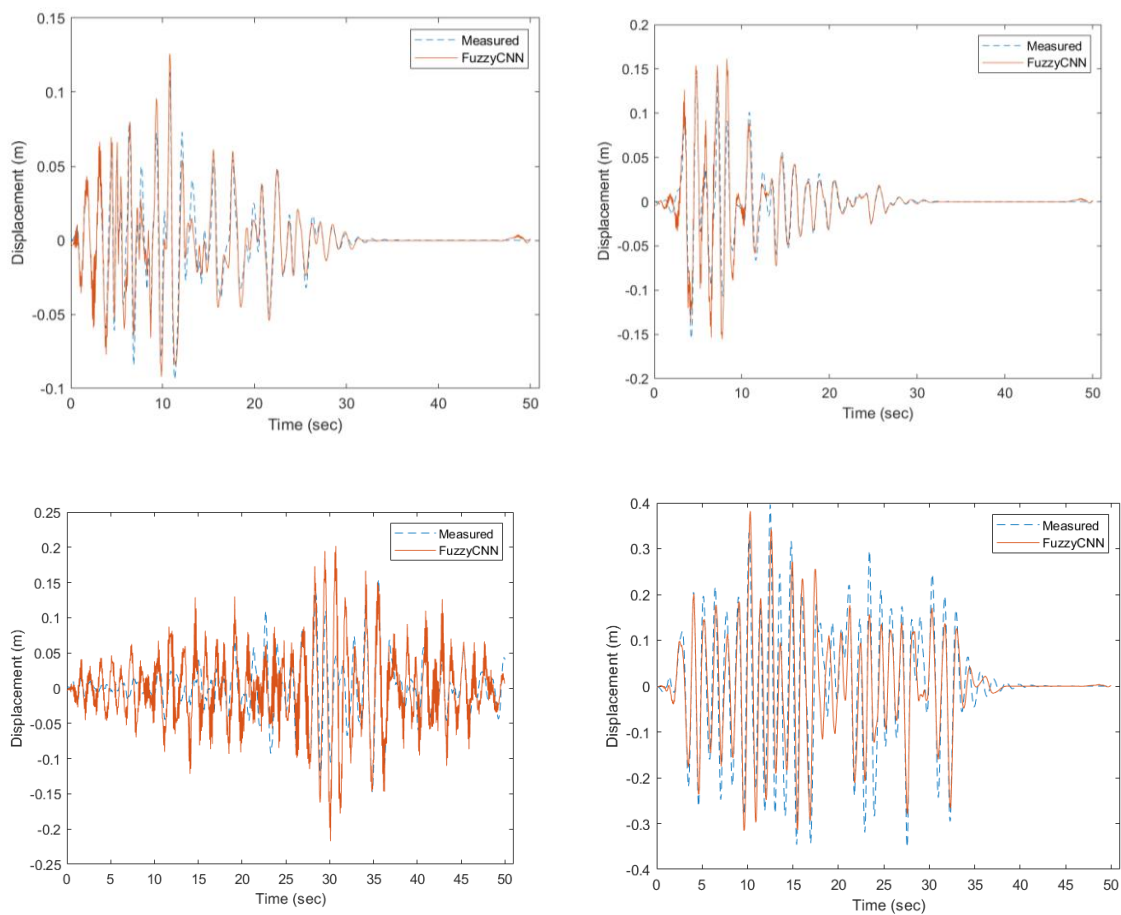
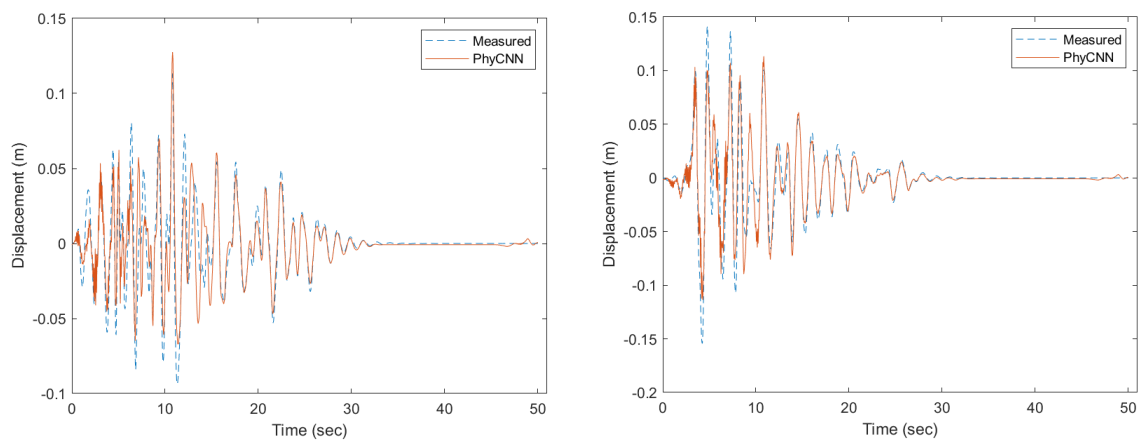


Figure 6. Example of the proposed FuzzyCNN model's prediction - SDOF structural system



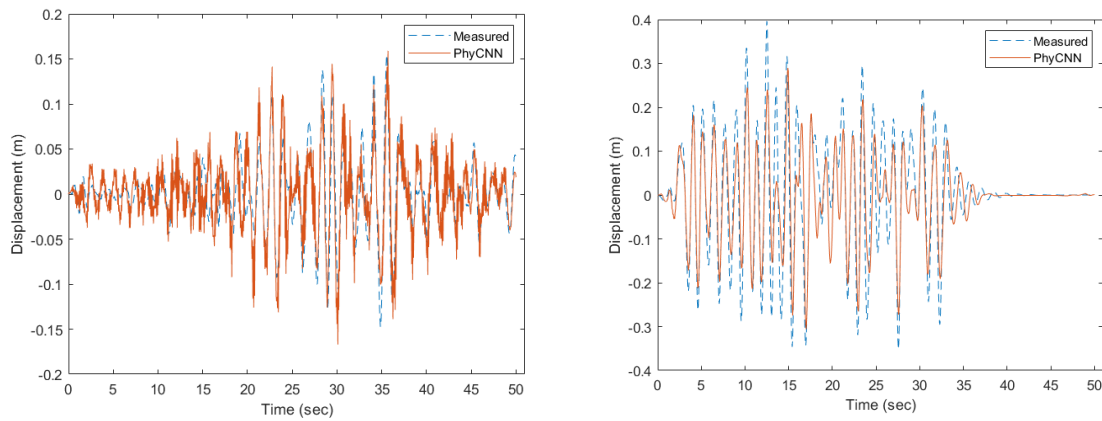


Figure 7. Predictions of the Zhuang and colleagues' model- SDOF structural system

The histograms (Figure 8) provide a visual representation of the distribution of correlation coefficients obtained from various predictions made by the models. The histogram for the FuzzyCNN model (Figure 8a) shows how well this model's predictions correlate with actual data. A peak towards higher correlation values indicates better model accuracy. The PhyCNN model's histogram (Figure 8b) serves a similar purpose. Figure 8c allows for a direct comparison between the two models, highlighting any differences in performance. For instance, a broader distribution in one histogram might suggest that a model performs inconsistently across different scenarios (like PhyCNN), while a tighter distribution with higher correlation values would indicate consistent and accurate predictions (like FuzzyCNN). It appears that nearly 5% of the predictions made by the PhyCNN model have an R^2 value lower than 0.4. An R^2 value of 0.4 or lower signifies that the model explains less than 40% of the variance in the dependent variable based on the independent variables. In practical terms, this means that the model's predictions are not closely aligned with the actual outcomes for those specific cases. Such a low R^2 indicates a weak predictive power, suggesting that the model's performance is inconsistent across the dataset. The PhyCNN model's dependence on directly integrating physical concepts into the neural network can be advantageous and disadvantageous at the same time [37]. This method guarantees that forecasts follow established physical rules, but if the physics is not precise or complete for all possible circumstances, it might not fully represent the complexity of the data. The PhyCNN model's overall dependability and credibility may be impacted if a subgroup of predictions has such low R^2 values. Even a tiny fraction of inaccurate forecasts can have a big impact on crucial applications like safety evaluations or structural engineering, sometimes resulting in dangerous choices or wrong conclusions [38].

Based on Figure 8, it appears that for the PhyCNN model, almost 17 percent of predictions have an R^2 higher than 0.9, while for the FuzzyCNN model, almost 24 percent of predictions have an R^2 higher than 0.9. The higher percentage of predictions with R^2 values above 0.9 in the FuzzyCNN model suggests it has a more flexible architecture, capable of adapting to a wider variety of data patterns. This flexibility may be attributed to the integration of fuzzy logic, which helps manage uncertainties and non-linearities in the data. One reason for the poorer performance of PhyCNN could be its sensitivity to noise present in earthquake records [38]. Noise in earthquake recordings can significantly impact the performance and reliability of PhyCNN models. These models are designed to leverage physical principles, making them sensitive to data quality. Earthquake data, however, is often contaminated with various types of noise, including ambient environmental noise, instrument noise, and interference from human activities or other seismic events [38]. This noise can obscure critical features within the seismic waveforms, such as amplitude, frequency content, and phase, which are essential for accurate modeling and interpretation [38]. Consequently, when the PhyCNN model encounters noisy data, it may struggle to extract meaningful patterns, leading to reduced accuracy in its predictions [38].

Additionally, the PhyCNN model contains approximately 207,414 more parameters than the FuzzyCNN model. This significantly larger parameter count suggests that PhyCNN may have a

higher capacity for learning complex patterns, but it also carries a potentially increased risk of overfitting, particularly when dealing with limited datasets. The trade-off between model complexity and generalization ability should be carefully considered in the context of the available training data and the specific task at hand [36]. In other words, the PhyCNN may unintentionally learn to identify noise patterns as meaningful features if the training set contains noisy recordings. This overfitting weakens the model's robustness in real-world applications, where noise characteristics can differ greatly, as well as its capacity to generalize to new, unseen data [39].

Figure 9 shows the performance of FuzzyCNN for the prediction of the displacement of the six-story mid-rise hotel building. Figure 9a displays examples of the estimated displacement time histories on the 3rd floor. The predictions made by the FuzzyCNN show good agreement with the reference data in terms of residual displacement suggestive of plastic deformation, phase alignment, and amplitude. These characteristics of the model's capturing accuracy indicate that it is a good emulation of the underlying physical phenomena and structural reactions. Moreover, as Figure 9b illustrates, the correlation coefficients primarily surpass 0.85, highlighting the strong prediction model precision. This high degree of accuracy shows that the FuzzyCNN is a reliable tool for sophisticated structural analysis and simulations because it can accurately predict the complicated, nonlinear behaviors that are inherent in materials undergoing plastic deformation. The model's resilience and dependability in real-world applications are highlighted by the consistency in both magnitude and phase as well as the rarely observed differences in residual drifts.

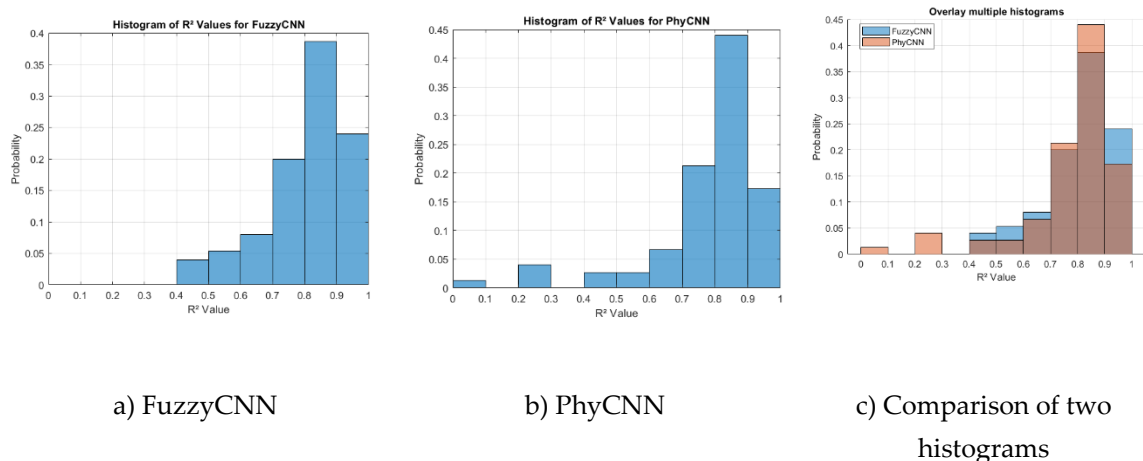
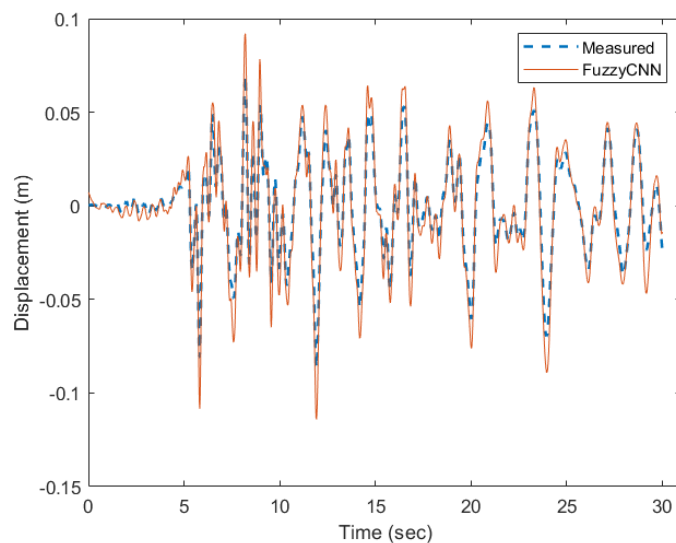
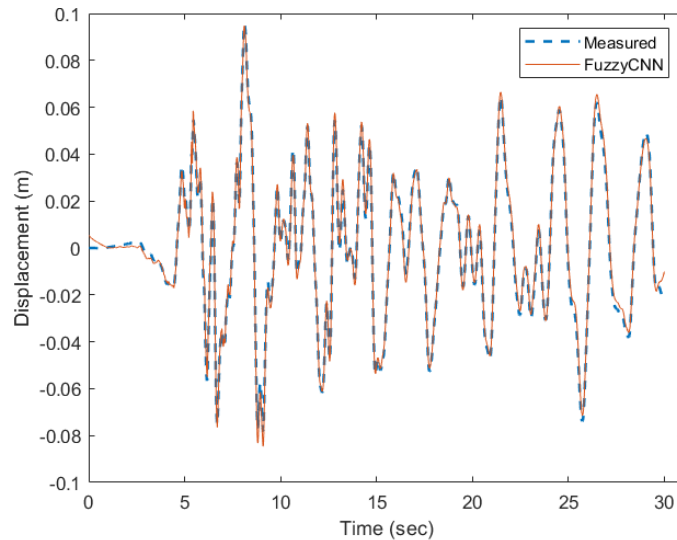


Figure 8. Histograms of the correlation coefficient.





a) Displacement time histories

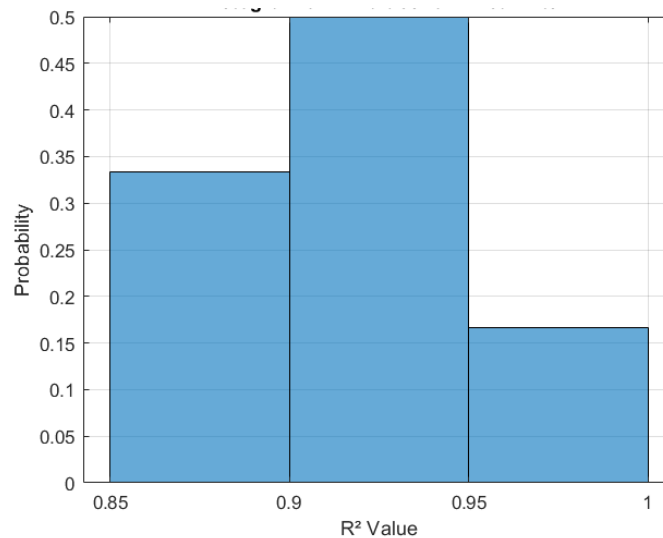
b) Histogram of R²

Figure 9. The FuzzyCNN model prediction of the displacement of the six-story mid-rise hotel building.

5. Conclusions

This study addresses the critical challenge of accurately predicting the dynamic response of nonlinear structural systems to seismic events. Traditional methods often struggle with computational intensity and the ability to generalize from limited datasets, necessitating the development of more efficient and robust approaches. The research introduced the Fuzzy Convolutional Neural Network (FuzzyCNN) model, which integrates fuzzy logic with convolutional neural networks to enhance the prediction accuracy of structural responses under seismic loading. Compared to the PhyCNN model, FuzzyCNN demonstrated superior performance, achieving higher correlation coefficients and greater robustness to noise in seismic data. Specifically, 24% of FuzzyCNN predictions had coefficient of determination (R^2) values above 0.9, indicating a strong predictive capability, while only 17% of PhyCNN predictions reached the same threshold.

Additionally, about 5% of the predictions from the PhyCNN model had R^2 values lower than 0.4, highlighting its comparative underperformance in certain scenarios. The FuzzyCNN model's integration of fuzzy logic enables it to manage uncertainties and nonlinearities more effectively than traditional physics-informed neural networks. This adaptability allows FuzzyCNN to provide more consistent and accurate predictions across diverse seismic events, highlighting its potential as a valuable tool in structural engineering. The study's findings suggest that incorporating fuzzy logic principles into neural network architectures can significantly enhance their performance in modeling complex dynamic systems. Also, the field measurements of the six-story concrete structure in San Bernardino used in this research indicate that the predictions align closely with historical sensing data for earthquakes of varying magnitudes and frequency. This suggests a strong correlation between the modeled predictions and actual seismic activity observed in the area, demonstrating the effectiveness of the predictive models employed in the study.

In conclusion, the FuzzyCNN model represents a significant advancement in predicting structural responses under seismic loading. By effectively combining fuzzy logic with convolutional neural networks, it offers a robust and adaptable solution for structural engineers. Ongoing refinements aim to ensure its practical application in mitigating seismic risks, highlighting the potential of advanced machine learning techniques integrated with traditional engineering principles for the future of structural dynamics.

Author Contributions: Conceptualization, M.S.B. and S.K.; methodology, M.S.B. and S.K.; software, M.S.B. and S.K.; validation, S.K.; formal analysis, S.K.; investigation, M.S.B. and S.K.; resources, M.S.B. and S.K.; data curation, S.K.; writing—original draft preparation, S.K.; writing—review and editing, M.S.B. and S.K.; visualization, S.K.; supervision, M.S.B. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability Statement: The data are available upon request.

Funding: This research received no funding.

References

- 1 Massone LM, Bedecarratz E, Rojas F, Lafontaine M. Nonlinear modeling of a damaged reinforced concrete building and design improvement behavior. *Journal of Building Engineering*. 2021;41:102766. <https://doi.org/10.1016/j.jobe.2021.102766>
- 2 Barkhordari M, Tehranizadeh M. Ranking passive seismic control systems by their effectiveness in reducing responses of high-rise buildings with concrete shear walls using multiple-Criteria decision making. *International Journal of Engineering*. 2020;33(8):1479-90. <https://doi.org/10.5829/ije.2020.33.08b.06>
- 3 Doan QH, Keshtegar B, Kim S-E, Thai D-K. Generative adversarial networks for overlapped and imbalanced problems in impact damage classification. *Information Sciences*. 2024;675:120752. <https://doi.org/10.1016/j.ins.2024.120752>
- 4 Tahaii SM, Hamidi H, Vaseghi Amiri J. Inelastic Seismic Demand of Steel-Plate Shear Wall Structures: Emphasis on the PTD Effect. *International Journal of Civil Engineering*. 2022;20(10):1145-63. <https://doi.org/10.1007/s40999-022-00716-1>
- 5 Mai SH, Dang H-K, Nguyen VT, Thai D-K. Stochastic nonlinear inelastic analysis for steel frame structure using Monte Carlo sampling. *Ain Shams Engineering Journal*. 2023;14(11):102527. <https://doi.org/10.1016/j.asej.2023.102527>
- 6 Thai D-K, Le D-N, Doan QH, Pham T-H, Nguyen D-N. A hybrid model for classifying the impact damage modes of fiber reinforced concrete panels based on XGBoost and Horse Herd Optimization algorithm. *Structures*. 2024;60:105872. <https://doi.org/10.1016/j.istruc.2024.105872>
- 7 Aminian FM, Khojastehfar E, Ghanbari H. Effects of near-fault strong ground motions on probabilistic structural seismic-induced damages. *Civil Engineering Journal*. 2019;5(4):796-809. <https://doi.org/10.28991/cej-2019-03091289>
- 8 Dang H-K, Thai D-K, Kim S-E. Stochastic analysis of steel frames considering the material, geometrical and loading uncertainties. *Advances in Engineering Software*. 2023;179:103434. <https://doi.org/10.1016/j.advengsoft.2023.103434>

- 9 Raoufy AA, Kheyroddin A, Naderpour H. Seismic Vulnerability Assessment of Reinforced Concrete Hospital Buildings Using Rapid Visual Screening Method According to FEMA P-154 criteria and Iranian Code# 364. *Civil Infrastructure Researches*. 2023;9(2):77-93. <https://doi.org/10.22091/cer.2023.9178.1462>
- 10 Bond RB, Ren P, Sun H, Hajjar JF, editors. *Physics-Guided Machine Learning for Structural Metamodeling and Fragility Analysis*. International Conference on the Behaviour of Steel Structures in Seismic Areas; 2024: Springer.
- 11 Alizadeh R, Allen JK, Mistree F. Managing computational complexity using surrogate models: a critical review. *Research in Engineering Design*. 2020;31(3):275-98. <https://doi.org/10.1007/s00163-020-00336-7>
- 12 Kleijnen JP. Kriging metamodeling in simulation: A review. *European journal of operational research*. 2009;192(3):707-16. <https://doi.org/10.1016/j.ejor.2007.10.013>
- 13 Spiridonakos MD, Chatzi EN. Metamodeling of dynamic nonlinear structural systems through polynomial chaos NARX models. *Computers & Structures*. 2015;157:99-113. <https://doi.org/10.1016/j.compstruc.2015.05.002>
- 14 Clarke SM, Griebisch JH, Simpson TW. Analysis of support vector regression for approximation of complex engineering analyses. 2005. <https://doi.org/10.1115/1.1897403>
- 15 Zhang R, Liu Y, Sun H. Physics-informed multi-LSTM networks for metamodeling of nonlinear structures. *Computer Methods in Applied Mechanics and Engineering*. 2020;369:113226. <https://doi.org/10.1016/j.cma.2020.113226>
- 16 Eshkevari SS, Takáč M, Pakzad SN, Jahani M. DynNet: Physics-based neural architecture design for nonlinear structural response modeling and prediction. *Engineering Structures*. 2021;229:111582. <https://doi.org/10.1016/j.engstruct.2020.111582>
- 17 Zhang R, Liu Y, Sun H. Physics-guided convolutional neural network (PhyCNN) for data-driven seismic response modeling. *Engineering Structures*. 2020;215:110704. <https://doi.org/10.1016/j.engstruct.2020.110704>
- 18 Xu K, Wang Q, Yang X, Ding D, Zhao Z, Hu Z, et al. Novel physics-informed neural network approach for dynamic and static displacement reconstruction via strain and acceleration. *Measurement*. 2024;114588. <https://doi.org/10.1016/j.measurement.2024.114588>
- 19 Hu Y, Guo W, Shi C. Physics knowledge-based transfer learning between buildings for seismic response prediction. *Soil Dynamics and Earthquake Engineering*. 2024;177:108420. <https://doi.org/10.1016/j.soildyn.2023.108420>
- 20 Shen S, Málaga-Chuquitaype C. Physics-informed artificial intelligence models for the seismic response prediction of rocking structures. *Data-Centric Engineering*. 2024;5:e1. <https://doi.org/10.1017/dce.2023.26>
- 21 Behnamfar F, Hadian A, Farazmand M. Ductility demand estimation of high-rise steel special moment frames due to structure-pile-soil interaction. *Geomechanics and Geoengineering*. 2024;1-13. <https://doi.org/10.1080/17486025.2024.2357186>
- 22 Farazmand M, Behnamfar F, Aziminejad A. Effects of the vertical and horizontal components of near-field ground motions on the seismic behavior of buildings considering soil-structure interaction. Available at SSRN 4602706. 2023. <https://ssrn.com/abstract=4602706>
- 23 Beyki Milajerdi M, Behnamfar F. Soil-structure interaction analysis using neural networks optimised by genetic algorithm. *Geomechanics and Geoengineering*. 2022;17(5):1369-87. <https://doi.org/10.1080/17486025.2021.1940313>
- 24 Akhoondi MR, Behnamfar F. Seismic fragility curves of steel structures including soil-structure interaction and variation of soil parameters. *Soil Dynamics and Earthquake Engineering*. 2021;143:106609. <https://doi.org/10.1016/j.soildyn.2021.106609>
- 25 madani b, Behnamfar F. Parametric Study of Structure-Soil-Structure Interaction in Time and Frequency Domains. *Amirkabir Journal of Civil Engineering*. 2021;52(11):2761-78. <https://doi.org/10.22060/ceej.2019.16156.6143>
- 26 Kermani M, Saadatpour MM, Behnamfar F, Ghandil M. Effects of seismic pounding between adjacent structures considering structure-soil-structure interaction. *Scientia Iranica*. 2020;27(5):2230-46. <https://doi.org/10.24200/sci.2019.5405.1255>
- 27 Chen M-S, Wang S-W. Fuzzy clustering analysis for optimizing fuzzy membership functions. *Fuzzy Sets and Systems*. 1999;103(2):239-54. [https://doi.org/10.1016/S0165-0114\(98\)00224-3](https://doi.org/10.1016/S0165-0114(98)00224-3)
- 28 Kumar M, Stoll R, Stoll N. A robust design criterion for interpretable fuzzy models with uncertain data. *IEEE Transactions on Fuzzy Systems*. 2006;14(2):314-28. <https://doi.org/10.1109/TFUZZ.2005.861614>
- 29 Thomas JW. *Numerical Partial Differential Equations: Finite Difference Methods*. Midtown Manhattan, New York City: Springer New York, NY; 2013.
- 30 Barkhordari MS, Armaghani DJ, Asteris PG. Structural damage identification using ensemble deep convolutional neural network models. *Comput Model Eng Sci*. 2022;134(2). <https://doi.org/10.32604/cmesci.2022.020840>
- 31 Tian Y, Zhang Y. A comprehensive survey on regularization strategies in machine learning. *Information Fusion*. 2022;80:146-66. <https://doi.org/10.1016/j.inffus.2021.11.005>

- 32 Yi D, Ahn J, Ji S. An effective optimization method for machine learning based on ADAM. *Applied Sciences*. 2020;10(3):1073. <https://doi.org/10.3390/app10031073>
- 33 Goulet CA, Kishida T, Ancheta TD, Cramer CH, Darragh RB, Silva WJ, et al. PEER NGA-east database. *Earthquake Spectra*. 2021;37(1_suppl):1331-53. <https://doi.org/10.1177/87552930211015695>
- 34 Ancheta TD, Darragh, R. B., Stewart, J. P., Seyhan, E., Silva, W. J., Chiou, B. S. J., Wooddell, K. E., Graves, R. W., Kottke, A. R., Boore, D. M., Kishida, T., & Donahue, J. L. PEER NGA-West2 Database, PEER Report 2013-03. University of California, Berkeley, CA; 2013.
- 35 Haddadi H, Shakal A, Huang M, Parrish J, Stephens C, Savage W, et al., editors. Report on progress at the center for engineering strong motion data (CESMD). The 15th world conference on earthquake engineering Lisbon, Portugal; 2012.
- 36 Han Z, Gao C, Liu J, Zhang SQ. Parameter-efficient fine-tuning for large models: A comprehensive survey. arXiv preprint arXiv:240314608. 2024.
- 37 Naderpour H, SoltaniMatin A, Kheyroddin A, Fakharian P, Ezami N. Optimizing Seismic Performance of Tuned Mass Dampers at Various Levels in Reinforced Concrete Buildings. *Buildings* [Internet]. 2024; 14(8).
- 38 Guo J, Enokida R, Li D, Ikago K. Combination of physics-based and data-driven modeling for nonlinear structural seismic response prediction through deep residual learning. *Earthquake Engineering & Structural Dynamics*. 2023;52(8):2429-51. <https://doi.org/10.1002/eqe.3863>
- 39 Ghanizadeh AR, Aziminejad A, Asteris PG, Armaghani DJ. Soft Computing to Predict Earthquake-Induced Soil Liquefaction via CPT Results. *Infrastructures* [Internet]. 2023; 8(8).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.