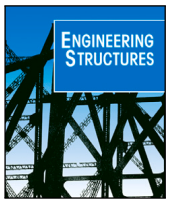
Engineering Structures 323 (2025) 119227

Contents lists available at ScienceDirect

Engineering Structures

journal homepage: www.elsevier.com/locate/engstruct



A real-time structural seismic response prediction framework based on transfer learning and unsupervised learning

Hongrak Pak *, Stephanie German Paal

Department of Civil and Environmental Engineering, Texas A&M University, College Station, 77843, TX, United States

ARTICLE INFO

Keywords:
Transfer learning
Long short-term memory (LSTM)
Structural seismic behavior
Unsupervised learning
SSR net

ABSTRACT

Conventional data-driven methods for predicting the seismic response of structures often require extensive data and computational resources. To address these challenges, a novel deep learning framework that can efficiently and accurately predict the structural seismic responses is proposed. The proposed framework overcomes the limitations of the conventional data-driven methods, by utilizing transfer learning based on the most relevant knowledge determined via the unsupervised learning technique. The framework leverages the seismic information history database to identify the most similar previous earthquake, and subsequently transfers the corresponding knowledge from the Structural Seismic Response network (SSR net) to predict structural responses caused by a new earthquake. This innovative method significantly reduces the need for extensive data collection and provides efficient predictions. Case studies demonstrate the framework's ability to predict seismic structural responses without extensive training or data collection. The framework can reliably capture the complex nonlinear dynamics of structures under seismic loads and offer significant potential for advancing seismic fragility analyses and reliability assessments. Future research will focus on expanding the framework's applicability to various structural types and further refining its prediction capabilities.

1. Introduction

Long-term records from the National Earthquake Information Center, since about 1900, reveal that there are on the order of 20,000 earthquakes all over the world in any given year, or approximately 55 per day. Moreover, sixteen significant earthquakes of magnitude seven or greater are statistically expected. In most catastrophic earthquakes, many buildings and infrastructure experience structural damage due to the strong ground motion. Entire cities have been completely destroyed, and there have been countless casualties as a result of catastrophic earthquakes, such as the Northridge earthquake in 1994, the Kobe earthquake in 1995, the Sichuan earthquake in 2008, and many others. To ensure the seismic stability of structures, of most significant interest is the deformation of the structural system or displacement of the mass induced by the ground shaking. [1]. As is well known, much research has focused on the development of advanced design methods [2–4], mitigation strategies [5–7], and retrofitting methods [8–10]. The research efforts during the last several decades have seen enormous resources dedicated to achieving more systematic and effective ways to ensure the safety of structures under seismic loads. The previous approaches can be mostly divided into experimental- and numerical- or analytical-based approaches. Experimental-based approaches have investigated characteristics of structural dynamic behavior, provided valuable evidence to clarify unknown relationships, and proposed state-of-the-art design philosophies. Because of the direct measurement of unidentified phenomena, experiments have been treated as the most reliable and decisive approach in structural engineering (second to real-world measurements). However, additional experiments are required to justify this conclusion for any modifications to existing structural designs or for new loading conditions. Additional experiments are not always feasible, considering the large scale of infrastructure and the expensive nature of such tests. Nonlinear dynamic analyses built upon various theoretical backgrounds are available for a wide range of structures with thousands of degrees of freedom. Extensive research based on numerical approaches has been conducted to simulate the time history of the nonlinear structural behavior and determine the seismic fragility curves for various types of infrastructure [11–13]. These nonlinear dynamic analysis methods have been extended to account for stochastic uncertainties of earthquakes and structural systems. Although such numerical methods have significantly improved the understanding of the seismic performance of buildings and infrastructure, there are still a few challenges. Assumptions made in numerical approaches may have an adverse effect on the results if they are in fact different from reality. Moreover, thousands of simulations varying the structural properties and loading conditions should be implemented

* Corresponding author.
E-mail address: hongrak822@tamu.edu (H. Pak).

https://doi.org/10.1016/j.engstruct.2024.119227
Received 22 May 2024; Received in revised form 6 October 2024; Accepted 23 October 2024
Available online 10 November 2024

to adopt stochastic uncertainties in seismic hazard analysis. A complex numerical model typically requires significant computational cost and time. In the worst case, the stability of nonlinear dynamic solutions may not be guaranteed due to the intrinsic convergence issue in the iteration procedure. An alternative way to address such drawbacks of the existing approaches mentioned earlier is machine learning (ML)- or deep learning (DL)-based approaches. By virtue of the ground motion recording systems, these approaches have provided noteworthy performance for a better understanding and accurate prediction of the dynamic behavior of structures in earthquake events. Earlier ML studies have used support vector machines or artificial neural networks (ANNs) to estimate structural system nonlinear behaviors under quasi-static or dynamic loading conditions [14–20]. However, using an ANN for estimating highly complex nonlinear dynamic behavior is often impractical because it assumes that the previous and current sequences are independent. Therefore, the temporal dependencies in time-series data cannot be considered, and the generalized nonlinear dynamic behavior may not be captured. The convolutional neural network (CNN) and recurrent neural network (RNN), which are more recent advancements in ML, may be more useful in this domain.

2. Literature review

Deep neural networks have demonstrated unprecedented performance in many engineering problems. In particular, CNNs and RNNs, variants of neural networks heavily used in image classification or natural language processing, have been gaining attention in the earthquake engineering realm as well. Several studies have utilized a CNN or RNN to estimate the structural response under dynamic loading conditions. Table 1 summarizes DL-based approaches to predict the building damage assessment or the time history of structural responses. There have been a few studies based on CNNs for predicting the dynamic behavior of a structure [21–23] and classifying seismic vibration versus ambient vibration [24]. These studies have shown that their proposed CNN-based models can deal with a large volume of signal data and automatically extract valuable spatial information from those signal data. As can be expected, many prior studies regarding seismic responses or damage assessments have focused on RNNs or long short-term memory (LSTM) models [25–28] since those model architectures are specifically designed to capture sequential dependencies between input and output variables. Ahmed et al. [29] used the overlapped data sequence to train their LSTM model. Some variations of RNN or LSTM architecture also have been reported in several studies [30–32]. Yu et al. [33] and Zhang et al. [34] embedded physics principles into DL-based approaches. Notably, Xu and Noh [35] adopted a physics-informed DL and domain adversarial network to diagnose building damages induced by earthquakes based on the data from different buildings.

Previous studies have introduced various types of DL-based architectures to improve seismic performance evaluation. Most of these previous studies developed models trained on the structural responses from numerical models, shake table tests, or field measurements. Although promising results have been reported with complicated models, additional studies should be further investigated to bridge the following research gaps.

First, the more complex a DL-based model architecture, the more training data is required to train a model and avoid overfitting. Generally, a large amount of data is necessary to train a complex DL-based model since it intrinsically has thousands of parameters to be trained. An obvious solution is to collect more data of interest; however, this is not always feasible considering the size of specimens in a shake table test or full-scale structures. Secondly, low-fidelity data generated by numerical models is not helpful for accurately predicting real-world situations, and high-fidelity data from more detailed numerical models is difficult to calibrate and computationally expensive. Moreover, the assumptions or approximations included in the more high-fidelity numerical models are likely to work as undesirable constraints when generating data. Thus, the generated data would be bound by those constraints, which may compromise data quality. Lastly, existing methods do not explain how to effectively deal with a new earthquake which is not represented in the utilized dataset. Based on long-term records, earthquakes have a wide variability of intensity, duration, and general appearance. Consequently, the existing DL-based approaches may not be able to provide reliable response prediction for a new earthquake outside of the established data.

To address the aforementioned issues in the existing DL-based approaches in structural engineering, this study proposes a novel framework based on LSTM networks that can rapidly and precisely predict the structural response under unknown seismic loading, by integrating transfer learning (TL) and unsupervised learning. TL aims to manipulate a model trained on one domain (source domain) to provide accurate predictions in another related domain (target domain). Thus, a robust prediction model can be achieved without excessive training data and computing resources, thanks to the knowledge transferred from a pre-trained model. While advanced TL methods such as domain adaptation or knowledge distillation techniques offer sophisticated capabilities [40,41], this study prioritized methods that would provide an efficient model with predictive accuracy as well as a straightforward algorithm for knowledge transfer. Such needs were driven by inevitable challenges in complex knowledge transfer methods, especially in real-time applications where rapid response is critical.

Although TL has shown great potential in the field of structural engineering [35,42–44], it is still most common in the CNN model architecture to share spatial features such as VGG-16 or ResNet50, and a few studies in other engineering applications have adopted TL to LSTM models [45–47]. This is because pre-trained temporal dependencies will no longer be valid on a new earthquake without an effort to find the best available relevant domain. When transferring knowledge from the source domain to the target domain, the performance of TL approaches depends on how similar the two domains are. Therefore, choosing an appropriate source domain closer to the target domain can maximize the performance of TL approaches. However, it can be challenging to determine the similarity between the domains, since there is no label information. Recognizing the challenge of determining this similarity without label information, this study incorporates an unsupervised learning algorithm to enhance the reliability of the proposed framework. In such cases, unsupervised learning, a class of ML algorithms that can learn and analyze patterns from unlabeled datasets, can intelligently provide an appropriate source domain to improve the performance of TL.

The objectives of this study are summarized as follows, with a focus on the novel contributions:

1. To propose a deep transfer learning framework: This study introduces a transfer learning framework specifically designed for predicting structural dynamic responses with minimal ground motion data. The proposed framework leverages the transferred knowledge to significantly reduce data requirements and addresses the common challenge of data scarcity in seismic engineering.
2. To handle ground motion correlations or unseen data: The proposed framework incorporates unsupervised learning approach to intelligently manage the correlation between recorded ground motions and new seismic events. This ensures that the most relevant historical data is utilized for making predictions, enhancing the model's accuracy and robustness even when dealing with novel seismic inputs.
3. To provide a practical pre-trained model: We aim to deliver a practical pre-trained model that can be readily used by engineers and researchers, similar to how VGG-16 or ResNet50 are utilized in image processing for spatial feature extraction.

Table 1
Summary of DL-based approaches for seismic response prediction or assessment.

Reference	Task	Network architecture	Response data source	Number of training ground motions[a]
[21]	Peak responses	CNN w/ TL	Numerical model	1199
[36]	Response prediction	NARX-RNN	Shaking table test	4/5
[25]	Response prediction	LSTM	Numerical model Field data	50/15/67
[22]	Response prediction	CNN	Numerical model Shaking table test	255/18
[33]	Response prediction	Physics-guided RNN	Numerical model	300/300
[34]	Response prediction	Physics-guided CNN	Numerical model Field data	10/50/15
[37]	Regional damage assessment	CNN	Numerical model	10,548
[26]	Response prediction	1D-CNN LSTM	Numerical model	7
[24]	Vibration classification	CNN w/ TL	Numerical model Field data	About 34,000
[30]	Response prediction	ConvLSTM	Field data	30
[35]	Building damage assessment	PhyMDAN	Numerical model Shaking table test	40 and more
[27]	Regional damage assessment	LSTM	Numerical model	42,192
[31]	Response prediction	Attention-based RNN	Numerical model Shaking table test	90/60
[28]	Response prediction	ANN, LSTM	Numerical model	1500
[29]	Building damage assessment	Stacked LSTM	Numerical model	256
[32]	Response prediction	Recursive LSTM	Numerical model	27/67
[23]	Response prediction	CNN	Numerical model	200
[38]	Response prediction	CNN	Numerical model	1499
[39]	Response prediction	Autoencoder w/ CNN	Numerical model Field data	200

[a] Multiple numbers indicate that more than one case was reported in the reference.

These objectives are achieved through the proposed framework consisting of four parts, which will be detailed in the subsequent sections. This innovative approach not only advances the current state-of-the-art in seismic response prediction but also offers practical solutions for the challenges faced in real-world applications.

3. A novel real-time structural seismic response prediction framework

In a general sense, the seismic response of a structure is affected by many factors, for example, the spatial information for the earthquake epicenter, earthquake intensity, structural characteristics, etc. Thus, most existing approaches for predicting the nonlinear structural response induced by an earthquake require a large dataset to maintain prediction performance even in an unseen earthquake. The dataset typically includes a variety of ground motions, structural responses, and material properties. However, the generalization capability of a trained model, which is its ability to appropriately handle new data, is not always ensured. In such situations, a trained model that consumed a significant amount of computational cost and time for training would be useless.

This study proposes a novel deep learning framework to rapidly and accurately predict the seismic response of a structure in only a few seconds (from data processing to prediction). The proposed framework needs two different data sequences: the time history of the ground motion recorded during the earthquakes and the corresponding structural displacement record. By adopting transfer learning and unsupervised learning, only two sequences of the ground motion and corresponding structural displacements are required to train and test the deep learning model. The proposed framework consists of four parts: the seismic information history database, the Structural Seismic Response Network (SSR net), the unsupervised nearest neighbor algorithm, and the knowledge transfer strategy. Fig. 1 shows the schematic procedure of the four parts in the proposed framework, and a detailed explanation of each part is thoroughly explained in the following sections.

3.1. Part 1: Seismic information history database

The aim of the first part of this framework is to build a seismic information database for a specific structure, consisting of several important values extracted from a number of ground motions that have been previously recorded. First, as can be seen in Fig. 1, important values are extracted from the previously recorded n different ground motions, where n is the number of ground motions that have been recorded at a seismograph station or the first floor of a structure. The extracted values from a previously recorded ground motion are representative of each earthquake (e.g., focal depth, epicentral distance, peak ground acceleration (PGA), peak velocity, peak displacement, peak structure acceleration, spectral acceleration (S_a)). The database established with the extracted values in this part will be used as the input to identify the most relevant previous earthquake for a new earthquake in Part 3.

3.2. Part 2: Structural Seismic Response network (SSR net)

In this part of the proposed framework, a group of LSTM networks, referred to as the Structural Seismic Response Network (SSR net), is established on the previously recorded ground motions and displacements. Instead of training the entire n different ground motions on a single LSTM network, the SSR net is composed of n different LSTM networks, each of which is trained on a ground motion during a specific earthquake and the corresponding displacements. Each

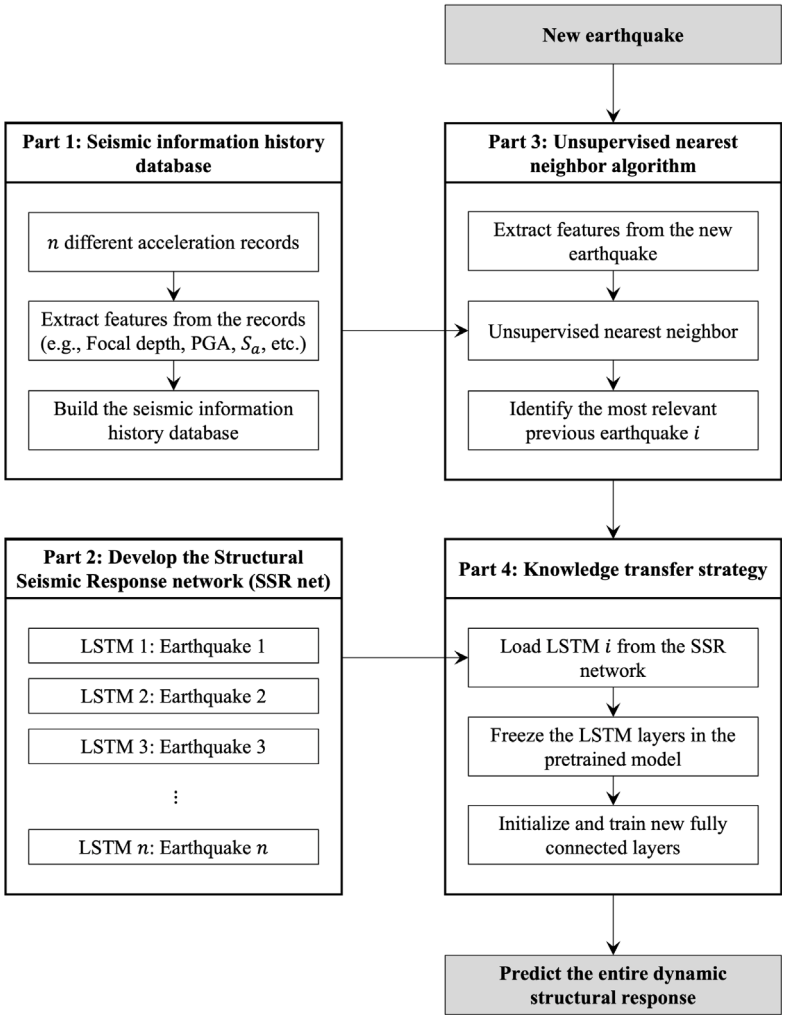


Fig. 1. The overview of the proposed framework for leveraging knowledge from previous earthquakes.

LSTM network is expected to understand how to predict the time history of the structural displacement based on a single earthquake ground motion. Therefore, this approach results in much simpler and less resource-intensive LSTM networks than the models developed in previous studies. This is because an individual LSTM network in the SSR net does not need to capture the generalized relationship for all available earthquake records. As mentioned earlier, an extensive dataset and complex model architecture is required to ensure prediction performance in traditional deep learning models. However, the proposed framework can make more accurate and efficient predictions with much less data and parameters that need to be trained. Even though an individual LSTM network in the SSR net is only trained on a specific ground motion, the ability of this framework to accurately predict the structural responses caused by arbitrary earthquakes will be explained in Parts 3 and 4. The n different LSTM networks established in this part provide a stable and reliable foundation for leveraging common knowledge to predict the structural seismic response caused by an unseen earthquake. More information on training an individual network will be provided in later sections.

3.2.1. Data preprocessing

The dataset fed into the ith LSTM network in the SSR net contains the time histories of the ground motion during an ith previous earthquake, $\mathbf{x}_i = [x_1, x_2, \dots, x_t]^\mathsf{T} \in \mathbb{R}^{t\times 1}$, and the corresponding displacement vector, $\mathbf{y}_i = [y_1, y_2, \dots, y_t]^\mathsf{T} \in \mathbb{R}^{t\times 1}$, where t is the number of time steps. Note that only one ground motion and the corresponding structural displacement record is needed to train and test an individual LSTM network in the SSR net. Generally, the initial format of the data acquired from sensors may not be adequate for training and testing an LSTM network, so a few steps are necessary before training or testing the model. First, to minimize adverse effects caused by scale and to easily learn the problem task, it is common to scale the dataset before training or testing a model. The dataset used in this study is transformed by RobustScaler in the Python library 'scikit-learn'. Interestingly, it uses the median and the interquartile range instead of the mean and standard deviation. In this way, it allows an ML model to be more powerful against outliers than typical scaling methods. Subsequently, the time series data was segmented into a number of fixed length vectors. The original dimension of the input and output variables, $\mathbb{R}^{t\times 1}$, should be converted to the proper format, $\mathbb{R}^{t'\times w}$, where w is the length of the input sequence in one sample. The number of samples fed into an LSTM network, t', could vary depending on w and the stride size, s. Thus, the kth prediction, y_k, can be made based on the input vector of $[x_{k-w}, \dots, x_{k-2}, x_{k-1}]$. Finally, the format of the input data should be reshaped to a 3-dimensional array for the LSTM layers used in the proposed framework. The data preprocessing in the proposed framework enables a network to take in input sequences with a uniform length and to be trained efficiently.

3.2.2. Individual LSTM networks in SSR net

An LSTM network proposed by Hochreiter and Schmidhuber [48] is a specialized type of RNN that handles long-term temporal dependencies between the input and output variables particularly well. In

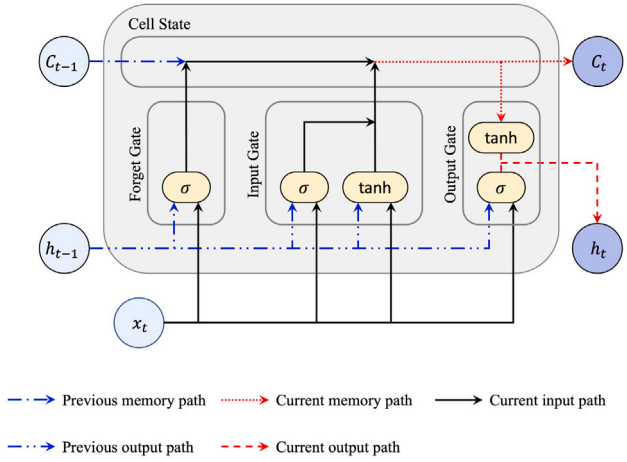


Fig. 2. Network architecture and data flow inside of an LSTM unit.

a standard RNN architecture, it is assumed that the current output variable depends on the prior information based on the sequence of the input variables. Thus, unlike the ANN architecture, the recurrent feedback loops allow RNNs to maintain prior temporal information and analyze the sequential data. However, a standard RNN often encounters a problem if the first derivative of the loss function decays exponentially, so-called the vanishing gradient problem. In this situation, the gradient of the loss function gets smaller and smaller as it passes through the deeper layers; consequently, the algorithm cannot update the weight parameters, and the trained model is essentially useless. By introducing the input and forget gates, however, an LSTM layer can alleviate the vanishing gradient problem and show better performance on long-term temporal dependencies. A single LSTM unit has four gates consisting of a forget gate, f_t, an input gate, i_t, an output gate, o_t, and an internal memory cell, C_t. They are designed to control how long information will be stored and how much information will be passed to the next time step. The network architecture of an LSTM unit and internal data flow are depicted in Fig. 2. Based on the information from the current input, x_t, and the previous hidden state, h_{t-1}, f_t decides which information in C_{t-1} should be maintained or discarded, as shown in Eq. (1). Then, i_t calculates the appropriate amount of the added information, and then, C_{t-1} is updated to C_t, as can be seen in Eqs. (2) through (4). Lastly, the current hidden state, h_t, is determined based on Eqs. (5) and (6). More details of the fundamental LSTM architecture can be found in the literature [48].

$$f_t = \sigma\left(w_f \cdot \left[h_{t-1}, x_t\right] + b_f\right) \tag{1}$$

$$i_t = \sigma\left(w_i \cdot \left[h_{t-1}, x_t\right] + b_i\right) \tag{2}$$

$$\tilde{C}_t = \tanh\left(w_c \cdot \left[h_{t-1}, x_t\right] + b_c\right) \tag{3}$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{4}$$

$$o_t = \sigma\left(w_o \cdot \left[h_{t-1}, x_t\right] + b_o\right) \tag{5}$$

$$h_t = o_t \odot \tanh\left(C_t\right) \tag{6}$$

where w_f, w_i, w_c, and w_o are the weight parameters, b_f, b_i, b_c, and b_o are the bias, x_t is the current input, $\sigma(\cdot)$ denotes the sigmoid function, and $\odot$ means the element-wise multiplication between two vectors.

As can be seen in Fig. 1, the SSR net consists of n different LSTM networks trained on n different historically recorded earthquakes. Each model is randomly initialized and trained on a single earthquake event, rather than all previous earthquakes. In every LSTM network in the SSR net, the recorded ground motion is used as the input variable and the corresponding structural displacement record is defined as the response variable. The training and test sets are mutually exclusive from each other to avoid data leakage problems and to precisely evaluate the performance of each model. The first 50% of the recorded earthquake ground motion is used for training, and the test set is set to be the remaining 50% of the data. During the training process, the Adam optimizer is adopted to find the optimal weight parameters and biases such that the loss function is minimized. The mean squared error (MSE) was chosen as the loss function for model training. This is due to its effectiveness in measuring the average squared difference between the predicted and actual values, which can penalize larger errors to a greater extent than smaller ones. Since the Adam optimizer is a combination of the gradient descent with momentum and root mean square propagation algorithms, it has been widely used to train deep neural network models. The performance of each individual LSTM network in the SSR net is monitored to avoid the overfitting problem during the training procedure and evaluated on the test set after training the model. Three metrics, MSE, root mean square error (RMSE), and the coefficient of determination (R^2) are employed to estimate the model performance and errors. The following equations show their mathematical definitions:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{7}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{8}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{9}$$

where n is the number of samples, y_i is the true response value, $\hat{y}_i$ is the predicted response value, and $\bar{y}$ denotes the mean of the actual response values.

3.2.3. Model architecture

Each individual LSTM model consists of an input layer, two LSTM layers, one fully connected layer, and an output layer, which were selected after conducting a sufficient preliminary test. Since the performance of a model is highly dependent on the hyperparameters, the best combination of hyperparameters should be investigated prior to training and testing a model. To effectively find the optimal hyperparameters to maximize the performance of each trained model, the Bayesian optimizer [49] has been implemented. According to the hyperparameter tuning results, the number of neurons in each layer and the length of the input sequence, w, can be determined. The model architecture determined by the extensive hyperparameter tuning process is consistently maintained for the n LSTM networks in the SSR net. The model architecture may appear oversimplified to provide accurate predictions on an unseen earthquake. Each individual LSTM network, however, is only expected to learn about a single earthquake event. The generalized prediction ability to unknown earthquakes will be achieved by transferring the acquired knowledge across the individual earthquakes. Thus, such a simple model architecture with a few layers is enough, and this will be supported by the results presented in Section 4.

3.3. Part 3: Unsupervised nearest neighbor algorithm

The seismic information history database developed in Part 1 and the SSR net constructed in Part 2 are based on the n previously recorded earthquake events. The generalized prediction performance, however, is not yet guaranteed because the individual LSTM network in the SSR net is only trained on a single earthquake ground motion and the corresponding structural response. To ensure the generalization capabilities of the proposed framework and to address the potential gap between the training data and unseen data, the aim of Part 3 is to

identify the most relevant previous earthquake when a new earthquake occurs. This process bridges the gap between known and unknown data by transferring knowledge closely aligned with the unseen data. By employing the unsupervised nearest neighbor (UNN) algorithm, the framework intelligently identifies the most relevant past earthquake, enhancing the robustness and adaptability of the predictive model. Once the most relevant previous earthquake, i, is chosen through Part 3, knowledge acquired from the earthquake i will be transferred to predict the structural displacement caused by the new earthquake. If there are more overlapped characteristics between the previous earthquake i and the new earthquake, the effectiveness of the knowledge transfer strategy proposed in this study will be maximized. The performance of the proposed procedure will decrease when the previous earthquake i and the new earthquake are less relevant to one other. Thus, the previous earthquake, i, should be appropriately chosen from all n previously recorded earthquakes.

To identify the most relevant earthquake, the seismic information history database established in Part 1 is employed here. First, similar to the procedures introduced in Part 1, several important values are extracted from the new earthquake that are representative of the new earthquake (e.g., focal depth, epicentral distance, peak ground acceleration (PGA), peak velocity, peak displacement, peak structure acceleration, spectral acceleration (S_a)). Subsequently, an ML model trained on the seismic information history database is employed to decide the most relevant previous earthquake to the new one. In other words, the seismic information history database established in Part 1 is used as the training set, and a single instance consisting of several important values extracted from the new earthquake is used as the test set. Because there is no response variable quantifying how similar or dissimilar earthquakes are, the UNN algorithm is included in the proposed framework. Based on the information in the seismic information history database, the UNN algorithm identifies the most relevant previous earthquake in the high-dimensional Euclidean space. The UNN algorithm can be mathematically represented as:

$$unn(\mathbf{x}) = \underset{i=1,\ldots,n}{\operatorname{argmin}} \|\mathbf{x} - \mathbf{x}_i\|_p \tag{10}$$

where $\mathbf{x}_i$ is the input vector, $\mathbf{x}$ is the test instance, and $\|\cdot\|_p$ is the p-norm of a vector.

Based on Eq. (10), the Euclidean distance ($p = 2$) is computed between the new earthquake and every other earthquake in the seismic information database. The pseudocode provided in Algorithm 1 outlines the steps for intelligently identifying the most relevant previous earthquake. The algorithm appropriately selects the most relevant previous earthquake i based on the similarity between the ground motions, and the LSTM network trained on the earthquake i will be used as a source LSTM network in Part 4.

Algorithm 1 Part 3: Unsupervised nearest neighbor algorithm

```
Input: Seismic information history database, New earthquake, p
Output: Most relevant earthquake index i
Extract features (e.g., focal depth, epicentral distance, PGA, etc.)
Initialize minimum distance as infinity
for each earthquake e_i in Seismic information history database do
    Extract features from e_i
    Calculate the p-norm between features of new earthquake and e_i
    if calculated norm < minimum distance then
        Update minimum distance
        Set i = e_i
    end if
end for
return the index i of the most relevant earthquake
```

3.4. Part 4: Knowledge transfer strategy

The generalized prediction ability is the most critical component of a trained LSTM network in earthquake engineering because it should maintain good performance when a new earthquake occurs. One problem is that the new earthquake will produce new ground motions and corresponding structural displacements that have never been experienced before. So far, the individual LSTM network in the SSR net ensures the ability to predict the structural displacement induced by a single ground motion. However, due to the intrinsic variability of tectonic activity and the lack of training data, it is still challenging for a prediction model to learn everything associated with the earthquake ground motion and the corresponding structural displacements. Thus, an individual LSTM network in the SSR net may not work well when a new earthquake occurs. Such a drawback can be resolved by adopting the knowledge transfer strategy developed in this work.

The purpose of Part 4 is to maintain or increase the performance of the proposed framework for an unseen earthquake event. Although no individual LSTM model in the SSR net will be trained on the unseen earthquake, the proposed framework is expected to accurately predict the structural displacement record by transferring knowledge from the most relevant previous earthquake identified in Part 3. TL wisely utilizes knowledge gained from one domain or task to improve prediction on another related domain or task, while traditional ML is isolated to prediction on a specific domain or task. This approach ensures that the model can generalize well to new earthquakes by leveraging similarities between previous and new events, even when they are not identical. Depending on the characteristics of the domains and tasks, TL can be broadly categorized into four techniques: instance-based transfer, feature-based transfer, relation-based transfer, and model-based transfer [50]. In this part of the framework, the model parameter-based transfer strategy has been utilized when transferring knowledge gained from the most relevant earthquake. Such a procedure assumes that a new earthquake and the most relevant earthquake identified in Part 3 have a great deal in common, or at least to some extent. Therefore, the underlying relationship between the most relevant previous earthquake and the corresponding displacement record can be a reliable foundation for predicting the structural displacement when a new earthquake occurs. The knowledge information in the ith LSTM network, which is trained on the most relevant earthquake event, i, is utilized to predict the structural displacements induced by a new earthquake. It should be denoted that only one previous earthquake has been used to train the ith LSTM network. Thanks to the knowledge transfer strategy in the proposed framework, an LSTM model for a new earthquake can easily build a generalized prediction from the LSTM network trained on the most relevant previous earthquake. The transferred LSTM layers are frozen during the remaining procedures of the proposed framework since the long-term dependencies of two earthquakes are assumed to be sufficiently correlated with one another. On the other hand, a fully connected layer is randomly initialized and placed after the transferred LSTM layers. This strategy enables the transferred knowledge to be calibrated to the new data and enhance the predictive accuracy. To train the randomly initialized fully connected layer, the proposed framework uses a very short initial section of the acceleration time history caused by a new earthquake. It should be noted that achieving robust performance from a separate LSTM model trained solely on the short initial section of acceleration time history is impossible. By virtue of those transferred layers, a substantial portion of the parameters in the network does not need to be trained from scratch. Therefore, the proposed framework can significantly reduce the model complexity and the number of training samples needed to obtain an accurate predictive model.

4. Case studies

The proposed framework's robustness was verified through three case studies using real-world earthquake records. This study has utilized the real-world data provided by the Center for Engineering Strong Motion Data (CESMD) [51]. The CESMD is a cooperative center established by the US Geological Survey (USGS) and the California Geological Survey (CGS) to integrate strong earthquake motions from several data sources. The CESMD serves as a comprehensive repository for strong-motion data, including ground acceleration records, structural response data, and related seismic information. It offers detailed records from a wide range of seismic monitoring networks, covering various structures and infrastructures, such as buildings, bridges, and free-field sites. For real-world implementations, an existing six-story building, a highway bridge, and a high-rise building were chosen to evaluate the performance of the proposed framework. The acceleration time history recorded at the ground level or on the first floor is fed into an LSTM network, and the trained LSTM network is expected to precisely estimate the dynamic displacements caused by an unseen ground motion. It should be noted that implementing the entire procedure from the beginning to the end takes only a few seconds. This is because the proposed framework is trained on only the one earthquake event that is identified as the most relevant earthquake to the new event, and the introduction of the knowledge transfer strategy significantly reduces the model complexity to predict the seismic displacement caused by an unseen earthquake. For each case study, the hyperparameter tuning process needs to be carried out, and the resulting hyperparameters are used as key parameters in the architecture of the LSTM network. The dataset, code, and in-depth results associated with this study can be accessed through the NHERI DesignSafe-CI portal [52].

4.1. Case study 1: Six-story building

An existing six-story building in San Bernadino, California, was used to investigate the performance of the proposed framework. This reinforced concrete (RC) structure has a typical rectangular plan shape, and the RC wall is the primary lateral force-resisting system. To monitor the structural seismic response, nine accelerometers are located on the first, third, and top floors in different directions. A more detailed description of the structural configuration, sensor locations, and their orientations is shown in Fig. 3. This existing RC building has been exposed to several strong ground motions from 1992 to 2021, and the measured accelerations with the corresponding structural displacement records are available through the CESMD website. In this case, the 'Devore' earthquake was assumed as the new earthquake (the target earthquake), and the proposed framework is implemented by following the procedures introduced in Section 3. The first step of the proposed framework is to establish the seismic information history database for this six-story RC building. It has been exposed to 42 strong ground motions since 1992. Several important values, which are representative of an earthquake, are extracted from each ground motion. The statistics of the seismic information history database for the six-story RC building are listed in Table 2.

The second step of the proposed framework is to establish the SSR net introduced in Section 3.2. Each LSTM network in the SSR net is specialized for an individual earthquake that has been experienced before. Prior to training and testing multiple LSTM models, the hyperparameter tuning process was implemented. According to the results from the hyperparameter tuning procedure, the most stable and highest performance can be found when the length of the input sequence, w is 10, the number of units in the first LSTM layer is 200, the number of units in the second LSTM layer is 100, and the number of units in a fully connected layer is 30. More details of the hyperparameter tuning results can be accessed through the NSF NHERI DesignSafe-CI portal. With those chosen hyperparameters, the SSR net consisting of multiple LSTM networks is established. The acceleration time history at the first floor, which is assumed to be the ground acceleration, was used as the input for each LSTM network. Nine LSTM networks are developed for each earthquake since nine sensors are located throughout the building. Thus, each network is expected to understand the dynamic relationships and accurately predict the structural displacements measured from an individual sensor. The first 50% of the time histories of acceleration and displacements were used as the training set, and the remaining 50% was used as the test set. In Part 3, the most relevant earthquake to a new earthquake was selected from the seismic information history database established in the first step. In this case, the UNN algorithm introduced in Section 3.3 has confirmed that the 'Chinohills' earthquake (the source earthquake) is the most relevant and similar earthquake to the Devore earthquake. Thus, in Part 4, knowledge acquired from the Chinohills earthquake is transferred into a new LSTM network to predict the seismic response caused by the Devore earthquake. By virtue of the transferred knowledge, the number of parameters that should be trained can be remarkably reduced from 488,641 to 6,061. Furthermore, notably, only 5% of the acceleration time history was used to fine-tune the LSTM network with transferred knowledge, which would be impossible if a conventional LSTM network was used. The remaining 95% of the time history sequence was used to evaluate the performance of the proposed framework.

Table 2
Statistics of the seismic information history database for the six stories RC building.

Parameter	Unit	Average	Standard deviation	1st quartile	3rd quartile
Magnitude		4.714	1.122	4.0	5.4
Focal depth	km	10.81	4.84	7.175	14.75
Epicentral distance	km	62.871	67.509	12.425	91.975
Ground PGA	g	0.025	0.026	0.008	0.03
Acceleration peak	g	0.02	0.019	0.008	0.027
Velocity peak	mm/s	13.874	18.855	3.25	13.2
Displacement peak	mm	3.895	12.47	0.013	2.0
S_a at 0.3 s	g	0.041	0.044	0.014	0.058
S_a at 1 s	g	0.011	0.016	0.001	0.013
S_a at 3 s	g	0.002	0.003	0.0	0.002
Structure PGA	g	0.073	0.077	0.021	0.097

Table 3
Prediction results for the source earthquake (Chinohills).

Input sensor	Output sensor	Entire history			Peak values		
		RMSE [mm]	R^2	MRE [%]	RMSE [mm]	Pred/True	SMAPE [%]
1	1	0.036	0.995	7.890	0.015	0.977	2.623
	4	0.037	0.995	8.164	0.019	1.010	5.349
	6	0.031	0.996	6.174	0.036	1.064	6.141
	7	0.031	0.996	6.572	0.023	1.033	4.176
	9	0.028	0.997	8.394	0.028	0.975	6.758
2	2	0.022	0.994	8.309	0.013	0.971	2.990
3	3	0.034	0.996	5.974	0.015	0.987	1.490
	5	0.033	0.996	6.149	0.009	0.994	1.708
	8	0.022	0.999	3.511	0.026	0.972	3.449

Table 4
Prediction results for the target earthquake (Devore).

Input sensor	Output sensor	Entire history			Peak values		
		RMSE [mm]	R^2	MRE [%]	RMSE [mm]	Pred/True	SMAPE [%]
1	1	0.033	0.990	8.664	0.022	0.989	1.688
	4	0.035	0.990	8.222	0.017	0.996	0.879
	6	0.024	0.996	5.245	0.018	0.996	0.579
	7	0.050	0.987	7.587	0.042	0.979	2.160
	9	0.052	0.987	7.792	0.067	0.982	3.855
2	2	0.011	0.991	13.234	0.008	0.971	3.559
3	3	0.028	0.994	6.774	0.011	1.021	2.218
	5	0.035	0.993	6.690	0.029	0.995	0.802
	8	0.096	0.991	6.096	0.140	0.984	2.270

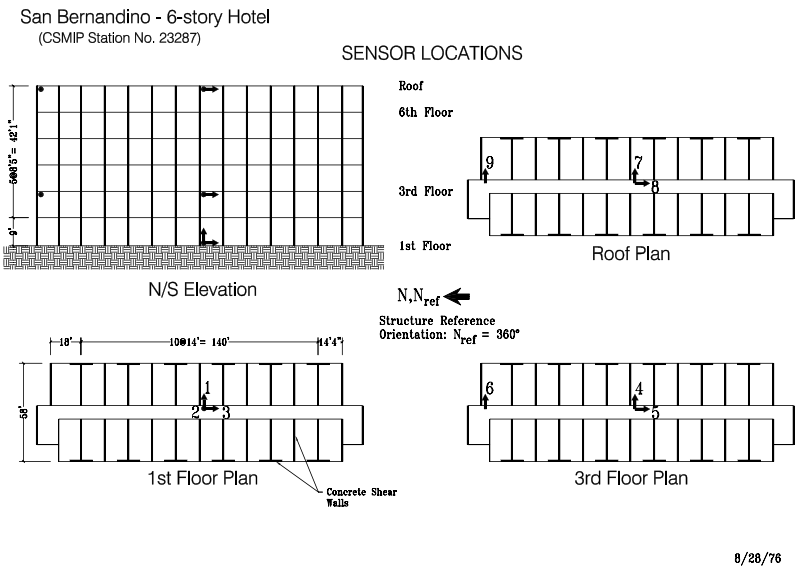
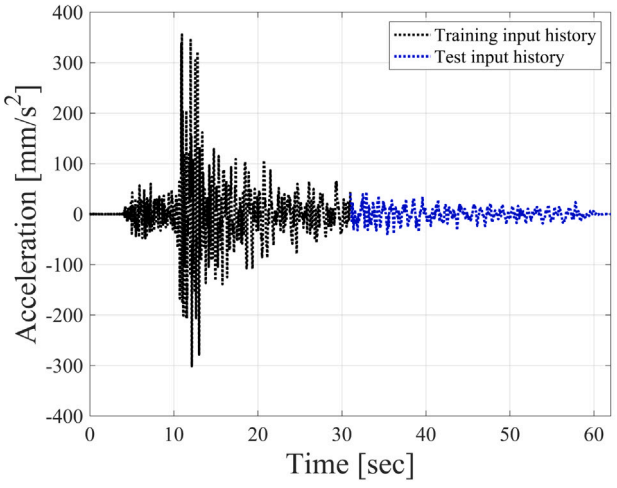
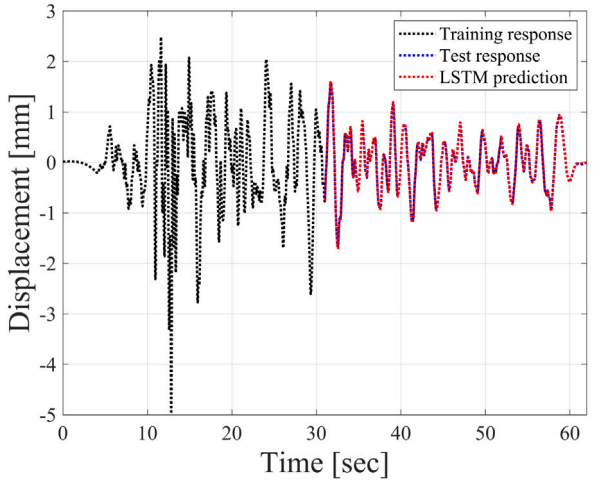


Fig. 3. Structural configuration and sensor locations in the six stories building [51].

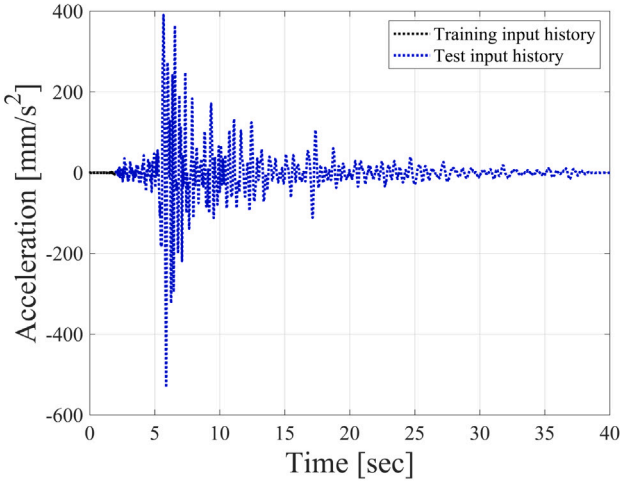


(a) Time history of acceleration for training and testing.

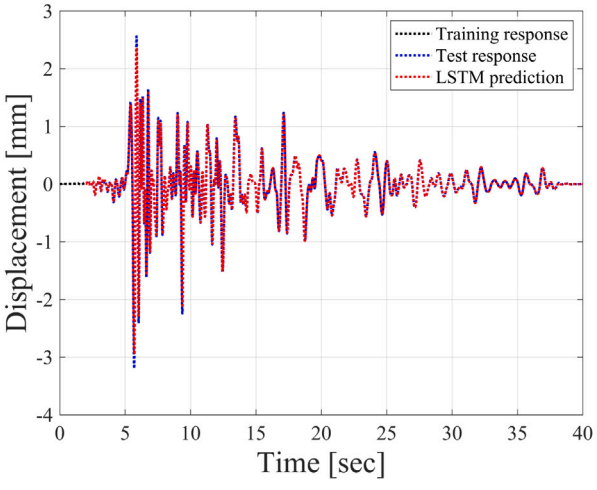


(b) Time history of displacement for training and testing.

Fig. 4. Comparison of the measured and predicted time history at Sensor 7 during the source earthquake (Chinohills).



(a) Time history of acceleration for training and testing.



(b) Time history of displacement for training and testing.

Fig. 5. Comparison of the measured and predicted time history at Sensor 7 during the target earthquake (Devore).

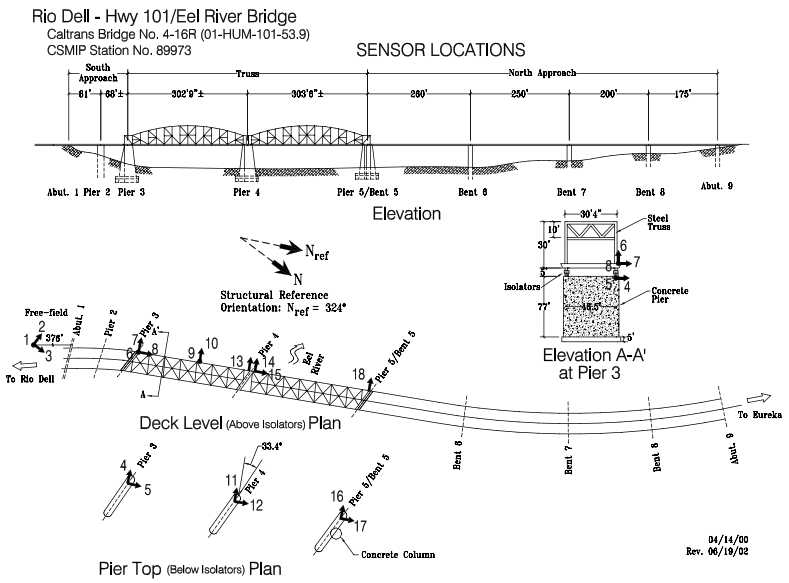


Fig. 6. Structural configuration and sensor locations on the highway bridge [51].

The RMSE and R^2 values were calculated based on the measured and predicted time histories of displacement at each sensor. To more comprehensively evaluate the performance of the proposed framework, peak values in each vibration cycle were extracted from the entire measured time history of displacement and compared with those values precited by the proposed framework. When comparing the measured and predicted responses and peak values, the median relative error (MRE) and symmetric mean absolute percent error (SMAPE) are additionally considered. The performance of the LSTM network designed for the Chinohills earthquake is summarized in Table 3. Since the LSTM network designed for the Chinohills earthquake showed accurate predictive performance, the knowledge from this network will be a valuable resource to predict the displacements caused by the Devore earthquake. This hypothesis is well demonstrated based on the results listed in Table 4, which show the performance of the LSTM network designed for the Devore earthquake. According to the comparison of the measured and predicted displacement time histories at Sensor 7, which is depicted in Figs. 4 and 5, the proposed methodology has shown outstanding performance. It should be noted that Sensor 7 has the lowest R^2 value. Interestingly, such a remarkable performance can be achieved with only 5% of the acceleration history along with the transferred knowledge from the LSTM network trained on the most relevant earthquake. Because of the transferred knowledge, the model complexity can be drastically reduced, and the time history of the displacement can be predicted in just a few seconds.

4.2. Case study 2: Highway bridge

To extend the proposed approach to another type of structure, a highway bridge has been utilized as the second case study. This bridge has two prestressed concrete girders, two steel truss spans, and four concrete box girders. The total length of the bridge is approximately 500 m, and the bridge spans range from 20 m to 93 m. There have been 15 sensors installed on the bridge and 3 sensors at a free-field site since 2001. More details regarding the structural configuration, sensor locations, and their orientations are depicted in Fig. 6. Table 5 shows the statistics of the seismic information history database for the highway bridge. The database consists of 33 strong ground motions measured from 2005 to 2021. In this case, the earthquake named 'Petrolia' was assumed to be the new earthquake for the sake of evaluating this framework, and the proposed methodology is expected to provide accurate displacement predictions during the new earthquake with significantly less complex model.

Table 5
Statistics of the seismic information history database for the highway bridge.

Parameter	Unit	Average	Standard deviation	1st quartile	3rd quartile
Magnitude		5.033	0.887	4.3	5.6
Focal depth	km	17.515	7.899	10.0	24.3
Epicentral distance	km	52.436	40.403	30.7	73.6
Ground PGA	g	0.040	0.066	0.013	0.034
Acceleration peak	g	0.039	0.066	0.013	0.034
Velocity peak	mm/s	22.933	44.764	5.8	20.2
Displacement peak	mm	2.333	4.929	0.021	2.0
S_a at 0.3 s	g	0.123	0.225	0.029	0.095
S_a at 1 s	g	0.014	0.027	0.003	0.011
S_a at 3 s	g	0.001	0.003	0.0	0.001
Structure PGA	g	0.105	0.137	0.042	0.107

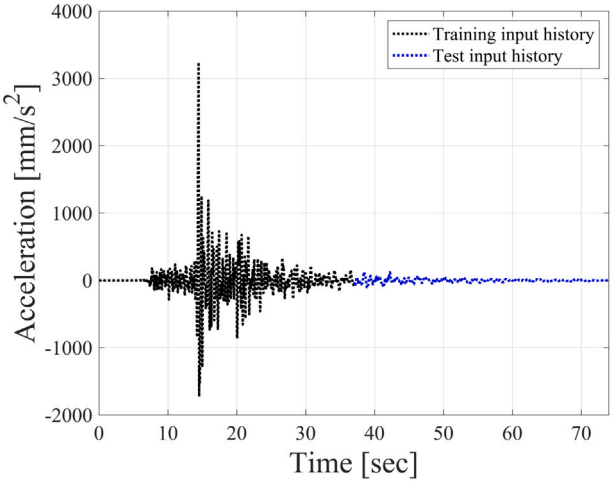
Based on the hyperparameter tuning results, the optimal combination of hyperparameters has been found in a given search space. The length of the input sequence is 10, the number of units in the first LSTM layer is 200, the number of units in the second LSTM layer is 200, and the number of units in a fully connected layer is 10. Using the unsupervised learning process outlined in Part 3, the earthquake named 'Ferndale' was selected as the most relevant earthquake to the new earthquake. Thus, knowledge acquired from the Ferndale earthquake is transferred to the LSTM network designed for the Petrolia earthquake. Similar to the first case, the number of parameters that should be trained can be drastically reduced from 488,441 to 4,041. In Case Study 2, the acceleration time histories measured at Sensors 1, 2, and 5, which were used as ground accelerations in their respective directions, were used as inputs for each LSTM network. Each network was trained to understand the dynamic relationships between the input ground motion and the structural response. Since there is no data measured from Sensor 3, the data measured from Sensor 5 has been utilized to predict the seismic response and then is compared with the measured data in the corresponding directions. The predicted displacements based on the proposed methodology have been compared with the measured data in corresponding directions, and the results are summarized in Tables 6 and 7. Table 6 shows the performance of the LSTM model designed for the Ferndale earthquake, and Table 7 shows the performance on the Petrolia earthquake. The proposed framework has also demonstrated outstanding performance on the bridge structure, as evident from the results presented in Tables 6, 7, Figs. 7, and 8. Sensor 11 exhibits the biggest difference between the measured and predicted values; however, the R^2 value on this sensor is still higher than 0.98.

Table 6
Prediction results for the source earthquake (Ferndale).

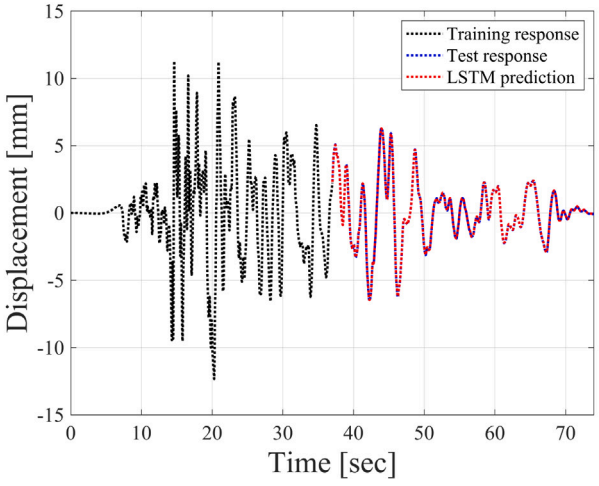
Input sensor	Output sensor	Entire history RMSE [mm]	R^2	MRE [%]	Peak values RMSE [mm]	Pred/True	SMAPE [%]
1	1	0.044	0.999	2.837	0.050	1.029	2.872
	6	0.033	0.999	2.816	0.034	0.999	4.461
	9	0.048	0.999	2.666	0.052	0.991	5.435
2	2	0.178	0.996	5.453	0.228	0.950	5.741
	4	0.037	0.999	1.737	0.026	1.001	1.669
	7	0.046	0.999	2.908	0.048	0.997	5.543
	10	0.061	0.999	1.782	0.062	0.998	2.693
	11	0.026	0.999	1.532	0.025	0.993	2.178
	13	0.074	0.999	4.719	0.078	0.982	5.251
	14	0.047	0.999	1.688	0.046	1.008	1.661
	16	0.038	0.999	1.394	0.042	0.989	1.205
	18	0.114	0.998	4.618	0.134	1.010	7.028
5	5	0.077	0.999	4.009	0.101	0.983	3.921
	8	0.110	0.998	5.455	0.135	1.037	6.473
	12	0.183	0.995	4.873	0.190	1.042	5.175
	15	0.097	0.999	4.881	0.113	1.017	7.793
	17	0.089	0.999	4.578	0.101	0.976	3.762

Table 7
Prediction results for the target earthquake (Petrolia).

Input sensor	Output sensor	Entire history RMSE [mm]	R^2	MRE [%]	Peak values RMSE [mm]	Pred/True	SMAPE [%]
1	1	0.057	0.993	4.388	0.020	1.000	2.142
	6	0.075	0.987	6.804	0.022	1.016	2.655
	9	0.204	0.995	4.667	0.186	1.035	4.384
2	2	0.140	0.991	3.288	0.138	1.001	1.526
	4	0.128	0.991	3.684	0.055	1.006	2.452
	7	0.154	0.990	3.672	0.150	1.023	6.368
	10	0.142	0.997	4.056	0.092	1.007	3.630
	11	0.169	0.983	5.446	0.131	0.986	4.231
	13	0.130	0.994	5.108	0.083	0.977	4.637
	14	0.140	0.996	4.144	0.100	1.006	2.964
	16	0.127	0.992	4.058	0.093	1.003	3.901
	18	0.206	0.988	6.445	0.114	0.975	4.417
5	5	0.199	0.992	3.177	0.251	0.999	1.448
	8	0.193	0.994	4.645	0.151	1.010	1.931
	12	0.259	0.988	4.456	0.238	1.016	2.403
	15	0.167	0.997	3.579	0.130	0.999	1.526
	17	0.213	0.993	3.192	0.250	1.009	3.516



(a) Time history of acceleration for training and testing.



(b) Time history of displacement for training and testing.

Fig. 7. Comparison of the measured and predicted time history at Sensor 11 during the source earthquake (Ferndale).

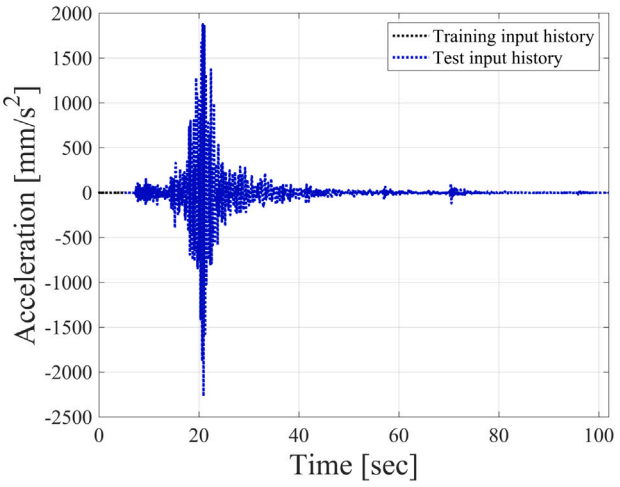
4.3. Case study 3: High-rise building

The third case study investigates a 12-story high-rise building located in Ventura, California, to provide sufficient evidence of the proposed framework's generalizability and to demonstrate its effectiveness for a high-rise building. This rectangular-plan structure is supported by RC columns and shear walls, with a foundation of 60-foot RC piles. Fifteen sensors installed across four levels provide detailed response data for various seismic events. This study focuses on predicting the structural displacements during earthquakes, leveraging the proposed framework to demonstrate its effectiveness on a high-rise building. More details regarding the structural configuration, sensor locations, and their orientations are depicted in Fig. 9. Table 8 shows the statistics of the seismic information history database for this building. The database consists of 24 ground motions measured from 1978 to 2024. In this case, the earthquake named 'Ojai' on November 03, 2023 was assumed to be the new earthquake for the sake of evaluating this framework, and the proposed methodology is expected to provide accurate displacement predictions during the new earthquake.

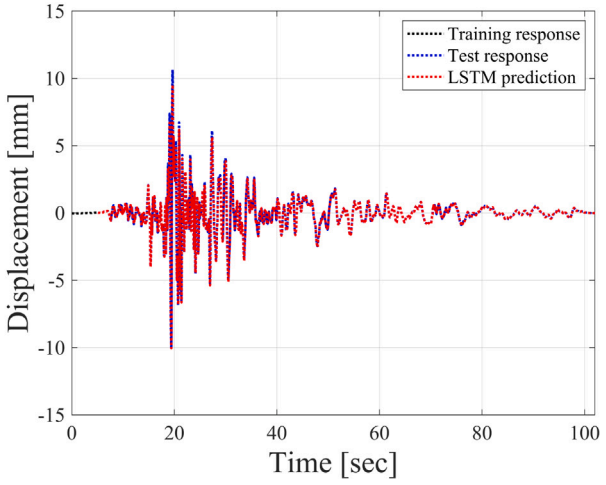
Following the proposed framework introduced in Section 3, the structural seismic information history database was established, and the SSR net was composed of multiple LSTM models for this structure. The optimal combination of hyperparameters was determined through a rigorous Bayesian optimization tuning process. Specifically, the input sequence was set to 20, and both the first and second LSTM layers have 150 units, followed by a fully connected layer with 20 units. Based on the seismic information history database, the UNN algorithm identified the 'Ojai' earthquake on August 20, 2023, as the most relevant previous earthquake. By leveraging the acquired knowledge from the most relevant previous earthquake, the proposed framework can reduce the model complexity and maintain accurate predictive performance. The results are summarized in Tables 9 and 10. These tables show

Table 8
Statistics of the seismic information history database for the 12-story building.

Parameter	Unit	Average	Standard deviation	1st quartile	3rd quartile
Magnitude		4.613	1.158	3.775	5.125
Focal depth	km	12.18	5.09	9.425	15.975
Epicentral distance	km	67.025	98.084	19.5	64.55
Ground PGA	g	0.026	0.032	0.009	0.03
Acceleration peak	g	0.026	0.032	0.009	0.031
Velocity peak	mm/s	1.978	2.498	0.3	1.96
Displacement peak	mm	0.842	2.125	0.0	0.23
S_a at 0.3 s	g	0.046	0.067	0.011	0.052
S_a at 1 s	g	0.013	0.017	0.002	0.014
S_a at 3 s	g	0.003	0.006	0.0	0.002
Structure PGA	g	0.079	0.097	0.025	0.078



(a) Time history of acceleration for training and testing.



(b) Time history of displacement for training and testing.

Fig. 8. Comparison of the measured and predicted time history at Sensor 11 during the target earthquake (Petrolia).

the predictive performance of the entire displacement time history and peak values. The performance is measured using RMSE, R^2, MRE, and SMAPE. Table 9 shows the performance of the LSTM model trained on the source earthquake. The results presented in Table 9 indicate the LSTM model designed for the source earthquake can be served as a solid foundation for transferring knowledge to predict the structural response during the target earthquake. According to Table 10, the proposed framework maintains good performance despite the differences between two events. Notably, the model used for the target earthquake was much less complex since the transferred knowledge from the source earthquake drastically reduced the number of parameters that needed to be trained. Figs. 10 and 11 demonstrate that the proposed framework successfully leverages the transferred knowledge from the previous event.

5. Discussion

5.1. Impact of data preprocessing methods

To rigorously assess the impact of data preprocessing techniques, this study investigates their effects on model performance by measuring the R^2 and RMSE values from 10 repeated trials. These metrics were

Table 9
Prediction results for the source earthquake (Ojai on Aug 20, 2023).

Input sensor	Output sensor	Entire history RMSE [mm]	R^2	MRE [%]	Peak values RMSE [mm]	Pred/True	SMAPE [%]
13	3	0.010	0.994	9.146	0.013	0.945	5.878
	4	0.011	0.994	15.761	0.010	0.999	7.588
	5	0.006	0.997	6.979	0.008	0.973	4.116
	7	0.003	0.999	4.254	0.004	1.020	1.978
	8	0.003	0.998	6.339	0.005	1.024	3.577
	10	0.003	0.997	5.866	0.006	1.060	5.747
	11	0.003	0.997	10.127	0.004	1.035	3.663
	12	0.002	0.998	7.622	0.003	0.965	3.598
	13	0.003	0.993	15.111	0.005	0.946	5.603
14	1	0.001	0.999	3.608	0.002	0.980	2.043
	14	0.001	0.996	10.208	0.002	0.971	3.535
15	2	0.011	0.996	6.874	0.015	0.938	6.447
	6	0.007	0.994	8.137	0.010	0.926	7.650
	9	0.002	0.998	3.876	0.002	1.028	2.775
	15	0.001	0.997	8.028	0.003	0.991	3.283

Table 10
Prediction results for the target earthquake (Ojai on Nov 03, 2023).

Input sensor	Output sensor	Entire history RMSE [mm]	R^2	MRE [%]	Peak values RMSE [mm]	Pred/True	SMAPE [%]
13	3	0.059	0.978	9.960	0.073	0.987	4.044
	4	0.052	0.988	8.164	0.069	0.965	3.653
	5	0.039	0.984	10.889	0.045	0.971	3.117
	7	0.031	0.983	12.067	0.024	1.003	1.326
	8	0.030	0.981	9.513	0.049	0.962	4.070
	10	0.019	0.979	8.700	0.037	0.956	5.672
	11	0.021	0.969	14.179	0.035	0.957	4.561
	12	0.008	0.989	7.261	0.016	0.994	1.360
	13	0.013	0.972	5.420	0.022	1.022	3.862
14	1	0.003	0.996	4.172	0.008	0.976	3.790
	14	0.004	0.988	9.054	0.005	1.006	3.556
15	2	0.047	0.982	10.472	0.054	0.951	5.234
	6	0.026	0.979	12.585	0.028	0.983	4.535
	9	0.014	0.982	7.265	0.019	0.981	4.079
	15	0.011	0.972	6.496	0.026	1.016	3.539

evaluated for Sensor 11 in Case Study 2, and three data preprocessing methods were considered: Robust scaling, Standard scaling, and Minmax scaling. The Robust scaling transforms the original data based on the interquartile range, and the Standard scaling transforms the data to have a mean of 0 and a standard deviation of 1. The Minmax scaling uses the maximum and minimum values of the original data.

Fig. 12 shows the box plots of R^2 and RMSE values from different data preprocessing methods in 10 repeated trials. According to Fig. 12, the Robust scaling shows the highest R^2 and the lowest RMSE values, while the Standard scaling also demonstrates stable performance. In contrast, the Minmax scaling has a weaker performance, as can be seen from the box plots of R^2 and RMSE values. In addition, two-sample t-tests were conducted to statistically further validate these observations. The results are summarized in Table 11. According to t-statistics and p-values from the comparisons, the difference between the Robust scaling and Standard scaling methods is not statistically significant. However, the Robust vs. Minmax and the Standard vs. Minmax show significant differences between the methods. Thus, the Robust and Standard scaling methods provide better performance than the Minmax scaling method, and no significant difference was found between the Robust and Standard methods.

5.2. Impact of distance metrics

A further sensitivity analysis was performed on the methods for calculating norms or distances in the unsupervised learning algorithm. Four different distance metrics were examined: Euclidean, Manhattan,

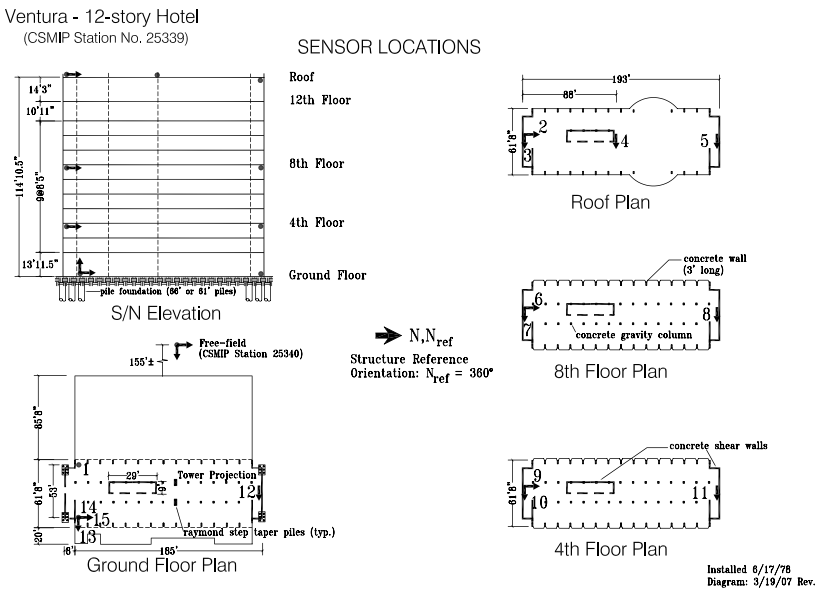
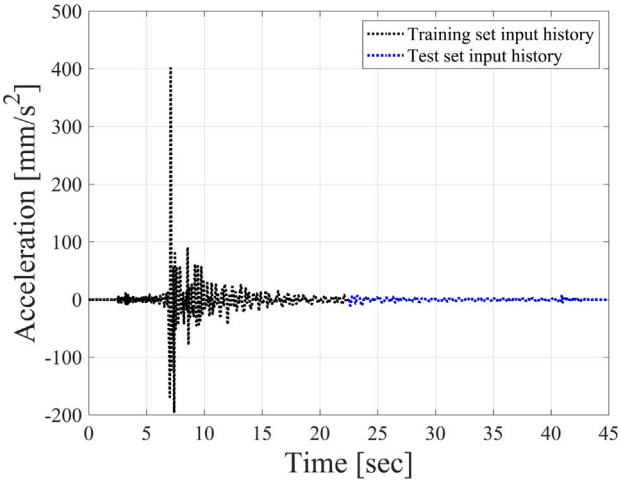
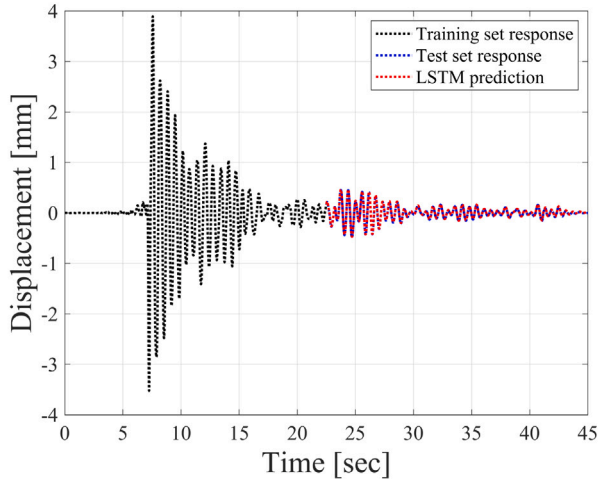


Fig. 9. Structural configuration and sensor locations in the 12-story city hall [51].

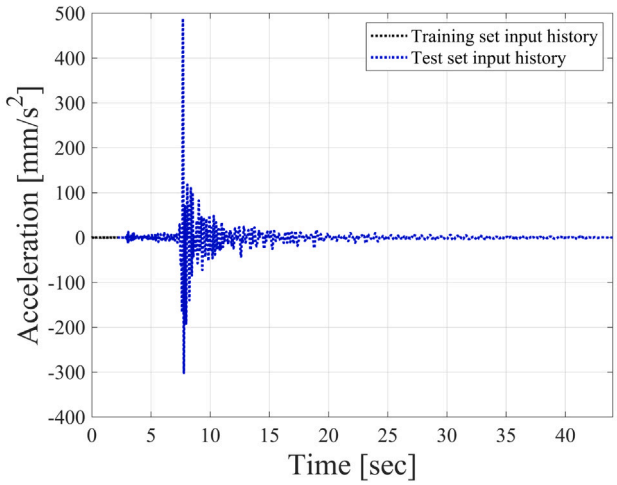


(a) Time history of acceleration for training and testing.

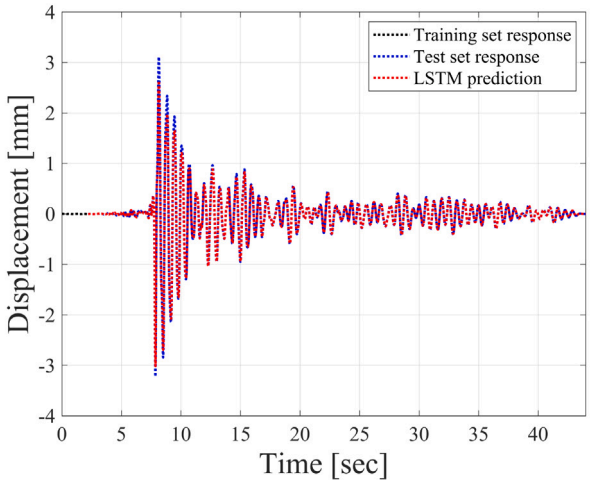


(b) Time history of displacement for training and testing.

Fig. 10. Comparison of the measured and predicted time history at Sensor 4 during the source earthquake.

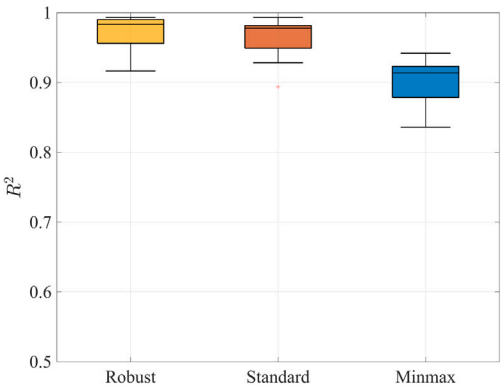


(a) Time history of acceleration for training and testing.

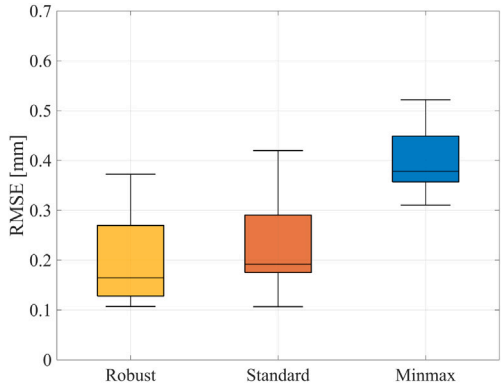


(b) Time history of displacement for training and testing.

Fig. 11. Comparison of the measured and predicted time history at Sensor 4 during the target earthquake.



(a) The coefficient of determination (R^2).



(b) The root mean square error (RMSE).

Fig. 12. Performance comparison of three different data preprocessing method.

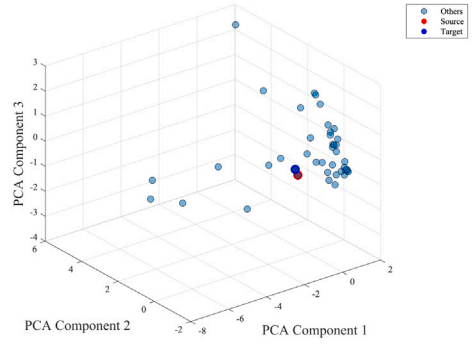
Table 11
Summary of two-sample t-test for data preprocessing methods.

Comparison	R^2 p-value	RMSE p-value
Robust vs. Standard	0.44937	0.42098
Robust vs. Minmax	0.00006	0.00002
Standard vs. Minmax	0.00070	0.00037

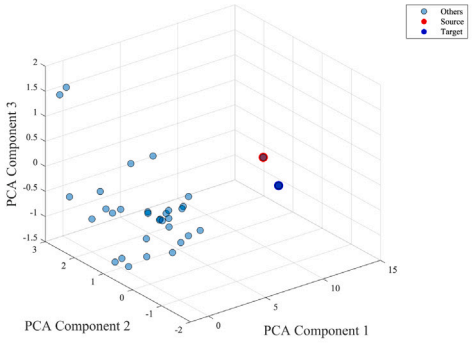
Minkowski, and Chebyshev distances. Euclidean distance, the most commonly used metric, calculates the straight line distance between two samples. Manhattan distance, also known as the L^1-norm or city block distance, sums the absolute differences between the coordinates of two samples. Minkowski distance, a generalization of both Euclidean and Manhattan distances, is calculated as the pth root of the sum of the absolute differences. Chebyshev distance calculates the maximum absolute difference between the coordinates of two samples. The results of the sensitivity analysis on distance metrics indicated that, despite the differences in these distance metrics, the UNN algorithm consistently selected the same previous earthquake as the most relevant for predicting the structural response to a new seismic event. This consistency suggests that the choice of distance metric, while mathematically substantial, may have a limited impact on the proposed framework, particularly in the case studies presented in this study.

5.3. Impact of the UNN algorithm

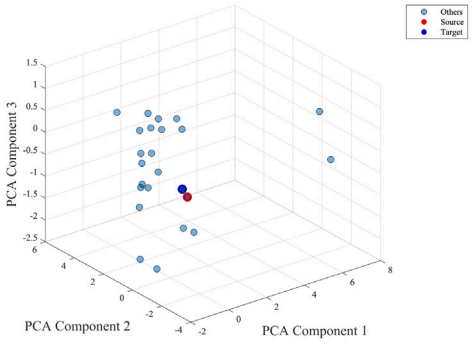
To evaluate the similarity between the source and target earthquakes used in the case studies, Principal Component Analysis (PCA) was employed. This will show the effectiveness of the UNN algorithm in the proposed framework. PCA is a dimensionality reduction technique



(a) 3D PCA visualization of earthquakes for Case Study 1.



(b) 3D PCA visualization of earthquakes for Case Study 2.



(c) 3D PCA visualization of earthquakes for Case Study 3.

Fig. 13. 3D PCA visualization of earthquakes. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

that can transform a high-dimensional dataset into a lower-dimensional space while maintaining the original information as much as possible. Thus, the results of PCA provide a visual insight on the similarity between the source and target earthquakes. Fig. 13 displays the results of PCA in three-dimensional space to provide a clear visualization of the relative positions of the source and target earthquakes compared to other events. The source earthquake is marked in red, the target earthquake in blue, and all other earthquakes in light blue. The proximity of the red and blue dots indicates the similarity between the source and target earthquakes. In all three figures, the source and target earthquakes are closely clustered, and this shows that the proposed framework can reliably identify the relevant knowledge.

6. Conclusions

In developing an accurate prediction model for real-time use in large civil infrastructures, it is crucial to consider the speed of processing and computational time. By adopting a knowledge transfer strategy into the proposed framework, this study can minimize the number of parameters in the network architecture and greatly increase the model efficiency. Furthermore, the UNN algorithm has been developed to wisely discover the most helpful knowledge to avoid performance deterioration due to the simple model architecture. Because of the knowledge transferred from the most relevant earthquake event, the proposed framework can maximize an LSTM network's ability to generalize the seismic response even for new, previously unseen earthquakes. In accordance with the results, the following conclusions are drawn:

- The proposed framework provides reliable and efficient predictions for the displacement time history caused by unknown earthquakes. It takes only a few seconds to conduct the entire procedure, including training.
- Since the proposed framework leverages knowledge acquired from the most relevant previous event, it is particularly useful in scenarios where labeled data for new earthquakes is limited or unavailable, a challenge not extensively addressed by the existing literature.
- The proposed methodology is flexible to integrate as many available earthquake records as possible to wisely select the most helpful knowledge for a new earthquake. Therefore, it will provide enormous practicality without rigorous training of a model from scratch in each instance. As more earthquake records are included in the seismic information history database, the performance of the proposed framework will be enhanced.
- This study has significant potential in seismic fragility or reliability assessments for any type of structure. Without performing nonlinear time history analyses, engineers will be able to effectively estimate the dynamic response of a structure caused by ground motions. It is also suited for scenarios where rapid seismic response predictions are critical, such as in emergency response and for continuous monitoring of critical infrastructure.

Future research should involve conducting in-depth sensitivity analyses, validating the framework under more severe earthquake conditions, and exploring the use of low-quality data to enhance the framework's robustness and applicability. Additionally, integrating structural characteristics into the framework, incorporating zero-shot learning approaches or using cutting-edge models could further expand its generalizability across a wider range of structures.

CRediT authorship contribution statement

Hongrak Pak: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Stephanie German Paal:** Writing – review & editing, Supervision, Funding acquisition.

Declaration of competing 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.

Acknowledgments

This material is based in part upon work supported by the National Science Foundation under the grant CMMI #1944301. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Data availability

Data will be made available on request.

References

[1] Chopra A. Dynamics of structures: Theory and applications to earthquake engineering. Pearson; 2017.
[2] Whittaker A, Hart G, Rojahn C. Seismic response modification factors. J Struct Eng 1999;125(4):438–44.
[3] Sezen H, Moehle JP. Shear strength model for lightly reinforced concrete columns. J Struct Eng 2004;130(11):1692–703.
[4] Takagi J, Wada A. Recent earthquakes and the need for a new philosophy for earthquake-resistant design. Soil Dyn Earthq Eng 2019;119:499–507.
[5] Dyke S, Spencer Jr B, Sain M, Carlson J. An experimental study of MR dampers for seismic protection. Smart Mater Struct 1998;7(5):693.
[6] Buckle I, Nagarajaiah S, Ferrell K. Stability of elastomeric isolation bearings: Experimental study. J Struct Eng 2002;128(1):3–11.
[7] Yang T, Konstantinidis D, Kelly JM. The influence of isolator hysteresis on equipment performance in seismic isolated buildings. Earthq Spectra 2010;26(1):275–93.
[8] Thermou G, Elnashai A. Seismic retrofit schemes for RC structures and local-global consequences. Prog Struct Eng Mater 2006;8(1):1–15.
[9] Hueste MBD, Bai J-W. Seismic retrofit of a reinforced concrete flat-slab structure: Part I—seismic performance evaluation. Eng Struct 2007;29(6):1165–77.
[10] Williams RJ, Gardoni P, Bracci JM. Decision analysis for seismic retrofit of structures. Struct Saf 2009;31(2):188–96.
[11] Nagarajaiah S, Xiaohong S. Response of base-isolated USC hospital building in Northridge earthquake. J Struct Eng 2000;126(10):1177–86.
[12] Simon J, Bracci JM, Gardoni P. Seismic response and fragility of deteriorated reinforced concrete bridges. J Struct Eng 2010;136(10):1273–81.
[13] Lim H-K, Kang JW, Pak H, Chi H-S, Lee Y-G, Kim J. Seismic response of a three-dimensional asymmetric multi-storey reinforced concrete structure. Appl Sci 2018;8(4):479.
[14] Chen S, Billings SA. Neural networks for nonlinear dynamic system modelling and identification. Int J Control 1992;56(2):319–46.
[15] Zhou C, Chase JG, Rodgers GW. Degradation evaluation of lateral story stiffness using HLA-based deep learning networks. Adv Eng Inform 2019;39:259–68.
[16] Yinfeng D, Yingmin L, Ming L, Mingkui X. Nonlinear structural response prediction based on support vector machines. J Sound Vib 2008;311(3–5):886–97.
[17] Luo H, Paal SG. Artificial intelligence-enhanced seismic response prediction of reinforced concrete frames. Adv Eng Inform 2022;52:101568.
[18] Hariri-Ardebili MA, Chen S, Mahdavi G. Machine learning-aided PSDM for dams with stochastic ground motions. Adv Eng Inform 2022;52:101615.
[19] Kazemi F, Asgarkhani N, Jankowski R. Machine learning-based seismic fragility and seismic vulnerability assessment of reinforced concrete structures. Soil Dyn Earthq Eng 2023;166:107761.
[20] Asgarkhani N, Kazemi F, Jankowski R. Machine learning-based prediction of residual drift and seismic risk assessment of steel moment-resisting frames considering soil-structure interaction. Comput Struct 2023;289:107181.
[21] Kim T, Kwon O-S, Song J. Response prediction of nonlinear hysteretic systems by deep neural networks. Neural Netw 2019;111:1–10.
[22] Oh BK, Park Y, Park HS. Seismic response prediction method for building structures using convolutional neural network. Struct Control Health Monit 2020;27(5):e2519.
[23] Wen W, Zhang C, Zhai C. Rapid seismic response prediction of RC frames based on deep learning and limited building information. Eng Struct 2022;267:114638.
[24] Liao W, Chen X, Lu X, Huang Y, Tian Y. Deep transfer learning and time-frequency characteristics-based identification method for structural seismic response. Front Built Environ 2021;10.
[25] Zhang R, Chen Z, Chen S, Zheng J, Büyüköztürk O, Sun H. Deep long short-term memory networks for nonlinear structural seismic response prediction. Comput Struct 2019;220:55–68.
[26] Huang P, Chen Z. Deep learning for nonlinear seismic responses prediction of subway station. Eng Struct 2021;244:112735.
[27] Xu Y, Lu X, Cetiner B, Taciroglu E. Real-time regional seismic damage assessment framework based on long short-term memory neural network. Comput-Aided Civ Infrastruct Eng 2021;36(4):504–21.
[28] Li C, Li H, Chen X. Fast seismic response estimation of tall pier bridges based on deep learning techniques. Eng Struct 2022;266:114566.
[29] Ahmed B, Mangalathu S, Jeon J-S. Seismic damage state predictions of reinforced concrete structures using stacked long short-term memory neural networks. J Build Eng 2022;46:103737.
[30] Torky AA, Ohno S. Deep learning techniques for predicting nonlinear multi-component seismic responses of structural buildings. Comput Struct 2021;252:106570.

[31] Li T, Pan Y, Tong K, Ventura CE, de Silva CW. Attention-based sequence-to-sequence learning for online structural response forecasting under seismic excitation. IEEE Trans Syst Man Cybern 2021;52(4):2184–200.
[32] Xu Z, Chen J, Shen J, Xiang M. Recursive long short-term memory network for predicting nonlinear structural seismic response. Eng Struct 2022;250:113406.
[33] Yu Y, Yao H, Liu Y. Structural dynamics simulation using a novel physics-guided machine learning method. Eng Appl Artif Intell 2020;96:103947.
[34] Zhang R, Liu Y, Sun H. Physics-guided convolutional neural network (PhyCNN) for data-driven seismic response modeling. Eng Struct 2020;215:110704.
[35] Xu S, Noh HY. PhyMDAN: Physics-informed knowledge transfer between buildings for seismic damage diagnosis through adversarial learning. Mech Syst Signal Process 2021;151:107374.
[36] Perez-Ramirez CA, Amezquita-Sanchez JP, Valtierra-Rodriguez M, Adeli H, Dominguez-Gonzalez A, Romero-Troncoso RJ. Recurrent neural network model with Bayesian training and mutual information for response prediction of large buildings. Eng Struct 2019;178:603–15.
[37] Lu X, Xu Y, Tian Y, Cetiner B, Taciroglu E. A deep learning approach to rapid regional post-event seismic damage assessment using time-frequency distributions of ground motions. Earthq Eng Struct Dyn 2021;50(6):1612–27.
[38] Kim T, Kwon O-S, Song J. Deep learning based seismic response prediction of hysteretic systems having degradation and pinching. Earthq Eng Struct Dyn 2022.
[39] Zhang C, Wen W, Zhai C, Jia J, Zhou B. Structural nonlinear seismic time-history response prediction of urban-scale reinforced concrete frames based on deep learning. Eng Struct 2024;317:118702.
[40] Liu N, Li Z, Liu R, Zhang H, Gao J, Wei T, Si J, Wu H. ASHFormer: axial and sliding window based attention with high-resolution transformer for automatic stratigraphic correlation. IEEE Trans Geosci Remote Sens 2023.
[41] Liu N, Zhang Y, Yang Y, Liu R, Gao J, Zhang N. Application of sparse s transform network with knowledge distillation in seismic attenuation delineation. Pet Sci 2024.
[42] Luo H, Paal SG. Reducing the effect of sample bias for small data sets with double-weighted support vector transfer regression. Comput-Aided Civ Infrastruct Eng 2021;36(3):248–63.
[43] Pak H, Paal SG. Evaluation of transfer learning models for predicting the lateral strength of reinforced concrete columns. Eng Struct 2022;266:114579.
[44] Pak H, Leach S, Yoon SH, Paal SG. A knowledge transfer enhanced ensemble approach to predict the shear capacity of reinforced concrete deep beams without stirrups. Comput-Aided Civ Infrastruct Eng 2023;38(11):1520–35.
[45] Fong IH, Li T, Fong S, Wong RK, Tallon-Ballesteros AJ. Predicting concentration levels of air pollutants by transfer learning and recurrent neural network. Knowl-Based Syst 2020;192:105622.
[46] Tan Y, Zhao G. Transfer learning with long short-term memory network for state-of-health prediction of lithium-ion batteries. IEEE Trans Ind Electron 2019;67(10):8723–31.
[47] Tariq S, Lee S, Woo SS. Cantransfer: Transfer learning based intrusion detection on a controller area network using convolutional LSTM network. In: Proceedings of the 35th annual ACM symposium on applied computing. 2020, p. 1048–55.
[48] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997;9(8):1735–80.
[49] Bergstra J, Yamins D, Cox D. Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In: International conference on machine learning. PMLR; 2013, p. 115–23.
[50] Pan SJ, Yang Q. A survey on transfer learning. IEEE Trans Knowl Data Eng 2009;22(10):1345–59.
[51] Haddadi H, Shakal A, Stephens C, Savage W, Huang M, Leith W, Parrish J, Borcherdt R. Center for engineering strong-motion data (CESMD). In: Proceedings of the 14th world conference on earthquake engineering, Beijing, October. 2008, p. 12–7.
[52] Pak H, Paal S. Code and data for a real-time structural seismic response prediction framework based on transfer learning and unsupervised learning. Designsafe-CI 2023. http://dx.doi.org/10.17603/ds2-fx45-dd16.