Data-Driven Hybrid Neural Network Under Model-Driven Supervised Learning for Structural Dynamic Impact Localization

Yingxin Luan, Teng Li[(✉)], Ran Song, and Wei Zhang

School of Control Science and Engineering, Shandong University, Jinan 250014, China
li.teng@sdu.edu.cn

Abstract. AI-oriented schemes, in particular deep learning schemes, provide superior capabilities of representative learning that leads to innovative estimation paradigm for structural health monitoring (SHM) applications. This paper introduces a model-driven and deep learning-enabled framework for localizing dynamic impact loads on structures. In this paper, finite element modeling (FEM) is conducted to generate enough labeled data for supervised learning. Meanwhile, a hybrid deep neural network (DNN) is established by integrating attentive and recurrent neural networks to exploit the latent features over both sensor-wise and temporal scales. The proposed DNN model is implemented to reveal the multivariate and temporal hidden correlations among complex time-series measurements and to estimate impact localization on structures. The experimental results from both numerical and physical tests demonstrate the superior performance of the proposed methodology.

Keywords: Deep learning · Attention mechanism · Structural health monitoring · Impact localization

1 Introduction

With the development of structure and infrastructure industries, more efforts have been made to detect the impact or damage to critical structural conditions [3,18]. In particular, for some severe weather conditions or natural disasters, the robustness of complex structures directly determines the safety of our human community and urban society.

Structural health monitoring (SHM) techniques refer to continuous condition monitoring of structures or structural components by measuring impact or damage-sensible data from instrumented sensors [14]. In traditional SHM applications, vibration-based approaches were able to complete damage detection of rotating machinery [11] and frame structures [4]. With superior performance

of representative and supervised learning [17], the implementation of AI-based techniques empowers the traditional SHM field. Particularly, SHM methods combined with finite element (FE) model and machine/deep learning schemes have achieved great advantages in recent studies [5,13]. Deep neural network (DNN) prediction models trained through a large amount of model-based structural data has proven their performance of efficient operation and high accuracy [6,9].

Localization of dynamic impact is great significant to detect potential structural malfunction or damage in SHM application scenarios. Many recent work has focused on this research topic. For example, [12] investigated similarity metrics in the time domain to determine the specific locations of impact loads based on the differences in structural responses of two impact load groups. This type of time domain process generally leads to heavy computational cost for positioning estimation. Liu *et al.* [10] studied and compared a series of machine learning methods for low-velocity impact localization. Established upon DNN schemes, Zhou et al. [19] proposed a recurrent neural network to identify the time-series responses of structural impact loads, without providing the actual position of a dynamic impact load. Zargar and Yuan [16] proposed a unified CNN and RNN network architecture for impact diagnosis, which deploying a high-speed camera to collect plenty of wavefield images.

Due to the lack of on-site collections at the time instant of dynamic impacts using limited sensors, it is extremely challenging and costly to directly capture real structural data samples from physical implementation. Different from the existing vibration-based approaches, a combination of model-driven and data-driven learning methodology is proposed in this paper. Specifically, enough data samples are generated with impact load positions and corresponding structural responses through numerical simulation using structural finite element modeling (FEM). In addition, to achieve better representation of multivariate and temporal sensory measurements, an attention-based hybrid DNN model is developed to regress the structural data and provide estimation of impact locations. Summarily, the contribution of the proposed methodology are in three folds:

(1) A novel systematic workflow for structural dynamic impact localization is proposed, which integrates both model-driven FEM and data-driven deep learning approaches.
(2) An attention-based hybrid DNN model is proposed to exploit complex hidden features among temporal and multivariate time series.
(3) Both numerical and real-world experimental testing are designed and conducted to validate the estimation performance of the proposed methodology.

In the rest of the paper, Sect. 2 formulates the research problem. Section 3 presents the dataset generation using model-driven approach while proposes the data-driven DNN model for impact load localization. Experiments are given in Sect. 4 and the last section concludes the paper.

2   Problem Formulation

The research goal of this paper is to study a modeling process to generate numerical simulations of structural impact responses while establishing a DNN model to regress the responses and impact position and estimate the impact location.

The present paper focuses on the localization of dynamic impact on structures using a systematic workflow with model-driven approach and deep learning approach. First, due to the lack of on-site measurements from real structures. The training data for the proposed DNN via supervised learning is generated using finite element modeling. Afterwards, the data is utilized to train the established DNN model. Finally, the trained model is deployed to estimate impact locations.

The research motivation of the studied problem can be formulated as:

$$(x, y) = \mathcal{F}(\mathbf{A}_{1:T}^1, \mathbf{A}_{1:T}^2, \ldots, \mathbf{A}_{1:T}^M) \tag{1}$$

where $\mathbf{A}_{1:T}^m, m = 1, 2, ..., M$ represents the measurements of the mth instrumented sensors over T time period. (x, y) indicates the coordinates of the location of an impact load on a target structure. $\mathcal{F}$ denotes a DNN model that can accurately regress between the measurements $\mathbf{A}_{1:T}^{1:M}$ and the impact location (x, y), which is the core objective of this work.

To achieve the aforementioned motivation, the proposed methodology can be summarized as three main procedures.

(1) **Model-driven data generation:** FEM is operated to generate volumes of data samples for model-driven supervised learning. For each impact load, generating structural response $\mathbf{A}_{1:T}^{1:M}$ and its label location (x, y).
(2) **Data-driven model training:** A data-driven DNN model $\mathcal{F}$ is developed and trained using the FEM data to regress among impact locations and the corresponding multivariate time series.
(3) **Data-driven model testing:** The trained DNN model $\mathcal{F}$ is deployed for localization of dynamic impacts on structures using measurements from instrumented sensors on structures.

3   Methodology

This section introduces the main methodology of the proposed model-driven as well as deep-learning driven framework in detail, including FEM data generation and DNN model development.

3.1   FEM Data Generation

To generate sufficient training data under supervised learning, an accurate FE model of an objective structure is developed first. Depending on the scale and complexity of the target structure, different levels of modeling details should be

considered, ranging from component level continuum solid elements to space rod system modeling.

With the developed model, a dynamic time history analysis is performed by applying a impact load case to any location of interest on the model. The impact load is often characterized as a triangular pulse-like loading pattern with an amplitude and a short duration in several milliseconds. The analysis will generate structural dynamic responses (e.g., displacement, velocity, acceleration, etc.) at any node and element in time domain.

The data that is generated by FEM can be formulated as:

$$m\ddot{x} + c\dot{x} + kx = p(t) \tag{2}$$

where m, c and k denote the mass, damping, and stiffness of the system, respectively. The impact load $p(t)$ can be expressed as:

$$p(t) = \begin{cases} 2p_o \frac{t}{t_d} & \text{for } t \leq \frac{1}{2}t_d \\ 2p_o \frac{t_d - t}{t_d} & \text{for } t \geq \frac{1}{2}t_d \end{cases} \tag{3}$$

where p_o is the amplitude of the triangular pulse and t_d is the pulse duration.

In this study, a large amount of loading cases with different amplitude and duration are randomly applied to the FE model. In each loading case, the location of the impact load is recorded as input data, while the acceleration responses at locations with instrumented sensors are monitored as output data.

3.2   Neural Network Design

The developed model considers both the correlationship among multi-sensor measurements and the temporal interrelationship in sequential data. The model conducts the regression process between the multivariate time-series inputs and the impact locations.

Inspired by the attention mechanism introduced in [1,15], considering the input of multivariate time series $\mathbf{A}_{1:T}^{1:M}$, a self-attention network is embedded in the proposed neural network to capture the sensor-wise dependencies within the multivariate time-series data. The attention layers can highlight the core features within the large volume of time-series data. The attentively extracted features is formulated as:

$$\begin{aligned} \mathbf{H} &= \text{softmax}(\frac{\mathbf{Q}\mathbf{K}^\top}{||\mathbf{Q}||||\mathbf{K}||})\mathbf{V} \\ &= \text{softmax}(\frac{(\mathbf{A}_{1:T}^{1:M}\mathbf{W}_Q)(\mathbf{A}_{1:T}^{1:M}\mathbf{W}_K)^\top}{||\mathbf{A}_{1:T}^{1:M}\mathbf{W}_Q||||\mathbf{A}_{1:T}^{1:M}\mathbf{W}_K||})(\mathbf{A}_{1:T}^{1:M}\mathbf{W}_V). \end{aligned} \tag{4}$$

where $\mathbf{Q}$, $\mathbf{K}$ and $\mathbf{V}$ denote the query, key, and value matrices, respectively, which are projected by the input $\mathbf{A}_{1:T}^{1:M}$ with learnable weight matrices $\mathbf{W}_Q$, $\mathbf{W}_K$, and $\mathbf{W}_V$. The dot product between the query matrix $\mathbf{Q}$ and the key matrix $\mathbf{K}$ leads to the attention matrix $\boldsymbol{\alpha}$ after normalizing by a softmax function. The resulting attention scores in the attention matrix are designated to the value matrix $\mathbf{V}$.

The final weighted output is the attentively extracted hidden feature $\mathbf{H}$, which integrates the extracted features within measurements of $\mathbf{M}$ sensors. The dot product $||\mathbf{Q}||||\mathbf{K}||$ is used to adjust a large dot-product value that may cause small gradients when operating backpropagation in neural network training.

To further extract the sequential characteristics, recurrent layers are implemented onto the attentively exploited features. This paper utilizes stacks of gated recurrent unit (GRU) [2,8] as recurrent layers in the network, updated as:

$$
\begin{aligned}
\mathbf{z}_n^{(r)} &= \text{sigmoid}(\text{Concat}[\mathbf{h}_{n-1}^{(r)}, \mathbf{Z}_n^{(r)}] \cdot \mathbf{W}_z^{(r)}) \\
\mathbf{r}_n^{(r)} &= \text{sigmoid}(\text{Concat}[\mathbf{h}_{n-1}^{(r)}, \mathbf{Z}_n^{(r)}] \cdot \mathbf{W}_z^{(r)}) \\
\tilde{\mathbf{h}}_n^{(r)} &= \tanh(\text{Concat}[\mathbf{r}_n^{(r)} \odot \mathbf{h}_{n-1}^{(r)}, \mathbf{Z}_n^{(r)}] \cdot \mathbf{W}_h^{(r)}) \\
\mathbf{h}_n^{(r)} &= (1 - \mathbf{z}_n^{(r)}) \odot \mathbf{h}_{n-1}^{(r)} + \mathbf{z}_n^{(r)} \odot \tilde{\mathbf{h}}_n^{(r)}
\end{aligned}
\tag{5}
$$

where $r = 1, 2, ..., R$ denotes the rth recurrent stack in the GRU stacks. $n = 1, 2, ..., N$ denotes the nth GRU module in a GRU stack. $\mathbf{Z}_n^{(r)}$ is denoted as:

$$
\mathbf{Z}_n^{(r)} = \begin{cases} \mathbf{H}_n \subset \mathbf{H} = \text{Concat}[\mathbf{H}_1, \mathbf{H}_2, ..., \mathbf{H}_N] & r = 1 \\ \mathbf{h}_n^{(r-1)} & r = 2, 3, ..., R \end{cases}
\tag{6}
$$

Finally, the estimated output is generated via fully-connected (FC) layers as:

$$
(\hat{x}, \hat{y}) = \text{FC}(\mathbf{h}_N^{(R)})
\tag{7}
$$

The overall network architecture is integrated by the attention subnetwork and the recurrent subnetwork. Figure 1 shows the overall architecture of the proposed DNN model. The training of the proposed architecture is conducted using the data and labels that are generated by FEM in Sect. 3.1.

The Euclidean distance is utilized as the loss function to quantify the error distance between the estimated location and the ground truth location, i.e.:

$$
loss = \sqrt{(x - \hat{x})^2 + (y - \hat{y}^2)}
\tag{8}
$$

where $\hat{x}$ and x represent the predicted and ground truth horizontal coordinate of the impact location, respectively. $\hat{y}$ and y represent the predicted and ground truth vertical coordinate of the impact location, respectively. Adam optimizer [7] is utilized to train the proposed DNN model with the Euclidean distance as the training loss.
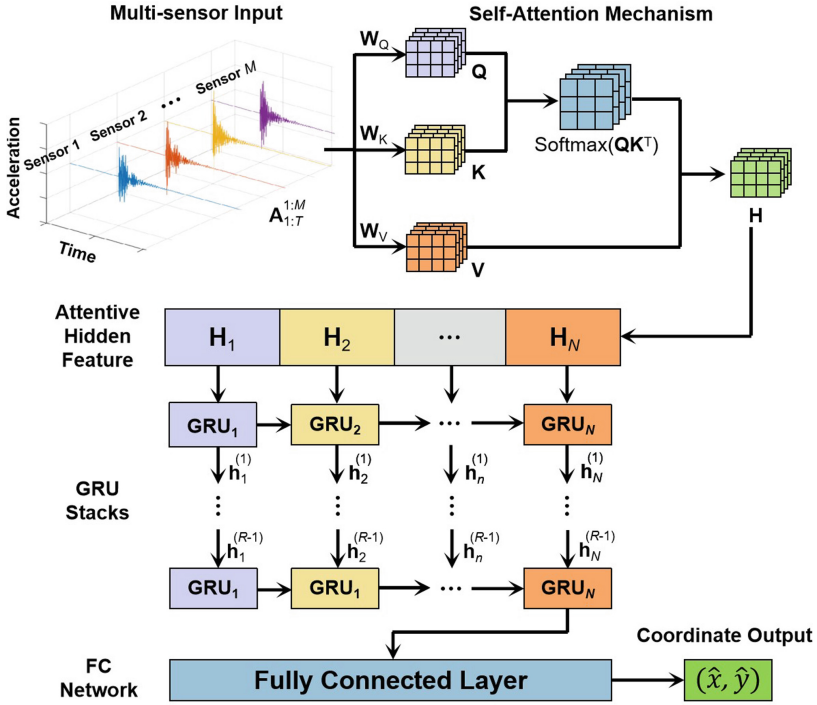
Fig. 1. Overall architecture of the proposed deep neural network.

4   Experiments

In this section, the performance of the proposed neural network is conducted and validated on real-world experiments. The structural setup of the physical experiments, the data generation setup of the dataset, the neural network setup, and the experimental performance of the proposed model are presented in detail.

4.1   Experimental Settings

Structural Setup. The experiment was conducted on an square aluminium plate structure with clamped boundary condition on two sides, as seen in Fig. 2. The plate had a dimension of 270 mm by 300 mm with a thickness of 6 mm. The density and the Young's modulus of the aluminum material were 69 GPa and 2.7 g/cm^3 with a Poisson ratio of 0.3.
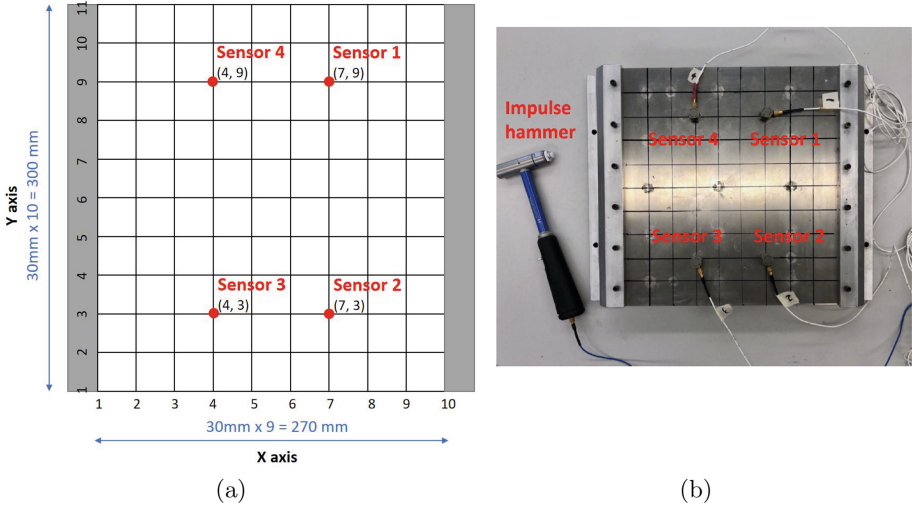
Fig. 2. Structural setup. (a) Plate specification and sensor placement. (b) Physical structure.

Considering the edge of the structural plate, the effective experimental area was divided by 9×10 small squares, learning to a square grid pattern. Four Brüel & Kjær type 4395 accelerometers were distributed over the plate at coordinates: Sensor 1 (7, 9), Sensor 2 (7, 3), Sensor 3 (4, 3), Sensor 4 (4, 9). The impact load was applied by using a Dytran impulse hammer. The data was transmitted using CoCo-80 acquisition system via Ethernet. All measurements were performed in a 5.76 kHz frequency range.

Data Generation Setup. To generate training dataset, the FEM of the test specimen was developed in commercial software SAP2000. The plate was modeled and meshed with thin shell element. Smaller grid tessellation can lead to higher resolution in terms of identification of impact location. In the experiment, a total of 86 grid nodes were designated in FEM, as seen in Fig. 3. Except the sensor locations and clamped edges, an impact load were randomly applied among the grid nodes with different amplitude p_o and duration t_d.

According to the actual test data, a range of 1 ms to 3 ms was assumed for t_d and the p_o was ranging from 20 N to 150 N. In total, 1,000 loading scenarios were simulated by performing dynamic time history analyses on the developed model, where the acceleration time history responses and impact locations were recorded. For each analysis, a complete loading duration was 0.5 s with a sampling rate 12800 Hz. As an illustration, Fig. 3(c) shows the acceleration time histories at four instrumented locations subjected to a loading example.

(a)   (b)

(c)
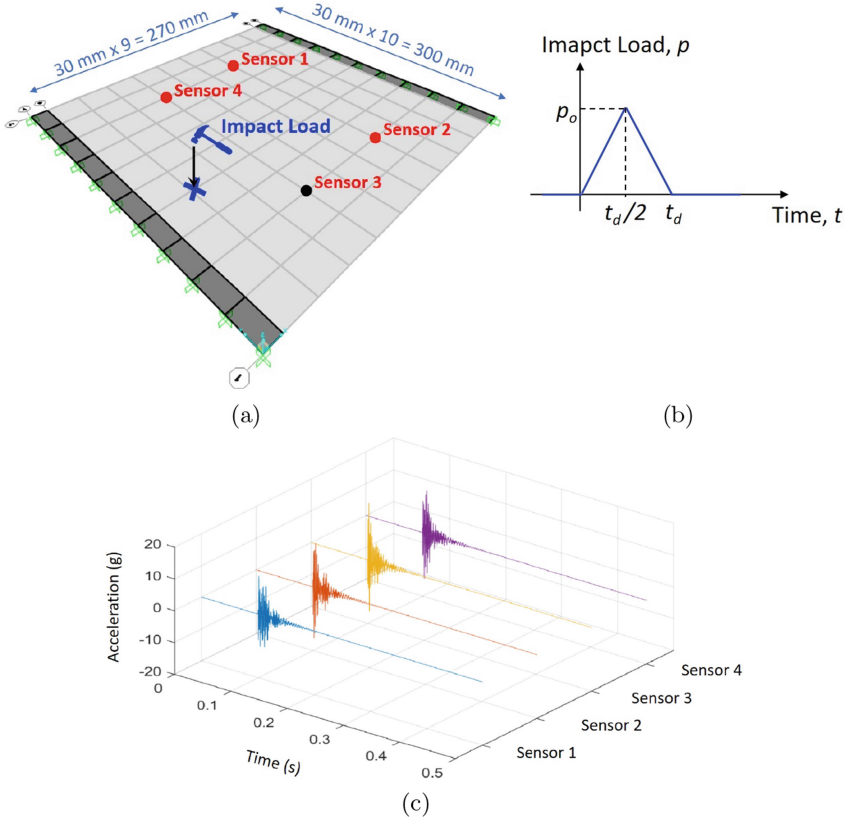
Fig. 3. FEM setup. (a) Finite element model for data generation, (b) Execution example of an impact load input, (c) Corresponding response data of acceleration time series.

Neural Network Setup. The data set is generated via FEM. The experimental data composes of training set, validation set and test set. Specifically, the dimensions of training set, validation set, test set are provided in Table 1. To train the proposed model, the batch size is set to 48 with the learning rate as 0.001. The dropout rate is set to 0.05. In addition, the hyperparameter settings are given in Table 2.

Table 1. Dataset setting.

Dataset	Source	Sample no.	Dimension
Training set	FEM	800	(800, 4, 6400)
Validation set	FEM	100	(100, 4, 6400)
Test set	FEM	100	(100, 4, 6400)

Table 2. Hyperparameter settings.

Network	Symbol	Setting	Description
Attention network	T	6400	Size of input time window
	d_q, d_k	[32, 64, 96, 128]	Dimension of the query and key state
	d_v	[20, 40, 60, 80]	Dimension of the value state
Recurrent network	N	[5, 10, 15, 20, 25]	Number of GRU blocks of each stack
	R	[1, 2, 3, 4, 5]	Number of GRU stacks
	d_r	[64, 96, 128, 256]	Dimension of the GRU hidden state
	L	[1, 2, 3, 4, 5]	Number of FC layers
	d_l	[8, 16, 32, 64]	Dimension of FC layer

4.2   Experimental Results

To evaluate the performance of the proposed model, four metrics are used to evaluate the localization accuracy. Specifically, mean absolute error of X axis (MAE of X), mean absolute error of Y axis (MAE of Y), Euclidean distance, and accuracy rate are the selected metrics in the experiments. These evaluation metrics are defined as:

$$\text{MAE of X: } e_x = \frac{1}{Total}\Sigma_{\text{test}}|x - \hat{x}| \tag{9}$$

$$\text{MAE of Y: } e_y = \frac{1}{Total}\Sigma_{\text{test}}|y - \hat{y}| \tag{10}$$

$$\text{Euclidean distance: } dis = \frac{1}{Total}\Sigma_{\text{test}}\sqrt{(x - \hat{x})^2 + (y - \hat{y}^2)} \tag{11}$$

$$\text{Accuracy rate: } acc = \frac{True}{Total} \times 100\% \tag{12}$$

where $\hat{x}$ and $\hat{y}$ represent the predicted horizontal and vertical coordinates of the impact location, respectively. x and y represent the ground-truth horizontal and vertical coordinates of the impact, respectively. $True$ and $Total$ indicate the number of the correct estimation (the estimated locations exactly match the label locations) and the total tests, respectively.

To validate the estimation performance of the proposed attention-based GRU stacks model (AttnGRUS), the experiments also implement the-state-of-the-art models for a comparative study. Specifically, covolutional neural network (CNN), long short-term memory stacks (LSTMS), GRU stacks (GRUS), CNN-LSTMS, and CNN-GRUS are chosen as the benchmark models for the comparative study. The experiments are conducted on the Pytorch framework with a 2.6 GHz Intel(R) Xeon(R) E5-2670 CPU and three GeForce GTX 1080 Ti GPUs.

The experiments are first carried out by numerical testing. Table 3 gives the experimental results on the numerical test set. For the proposed model, the network hyperparameters of AttentionGRUS are: $d_q, d_k = 64$, $d_v = 20$, $N = 10$, $R = 2$, $d_r = 256$, $L = 2$, $d_l = 16$. For the compared models, the network hyperparameters of the compared models are searched and set over grid search to achieve their best accuracy. As can be seen from the table, the proposed model AttentionGRUS achieves the best performance (highlighted gray cells)

Table 3. Numerical experimental results of the proposed and compared models.

| Model | Localization Performance | | | |
	e_x (mm)	e_y (mm)	dis (mm)	acc (%)
CNN	3.75	0.15	3.81	87
LSTMS	8.55	0.75	8.94	73
GRUS	7.2	0.75	7.86	82
CNN-LSTMS	2.1	1.2	3.12	90
CNN-GRUS	2.25	0.9	3	90
AttentionGRUS	0.15	0.3	0.45	98
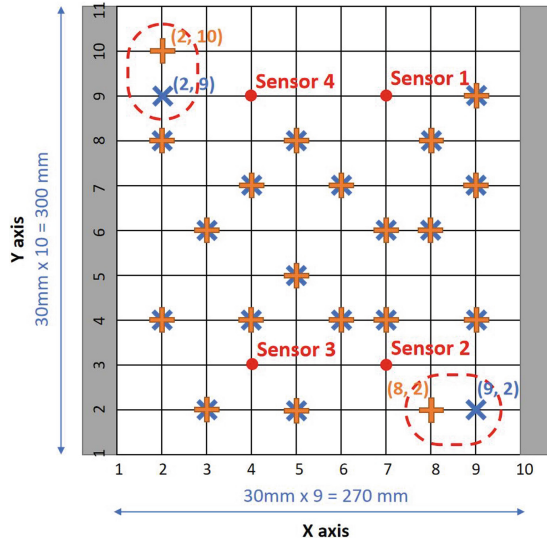
Fig. 4. Estimated locations (orange '+') and ground-truth locations (blue '×') of physical testing on structural plate. (Color figure online)

on the numerical test set in comparison with the compared models, with the accuracy rate of 98% and with the lowest estimation error regarding both X axis, Y axis, and the overall Euclidean distance. AttentionGRUS outperforms the compared benchmark models on the test set.

Figure 4 displays the experimental results of the physical testing on the structural plate using AttnGRUS. With a total of 20 randomly selected impact loads, the figure shows the estimations using AttnGRUS and the ground-truth impact locations. There are two samples not exactly matched (location (2, 9) and location (9, 2)) among the total of 20 tests. In this physical testing, the accuracy rate is 90% with e_x of 1.5 mm, e_y of 1.5 mm, and dis of 3 mm. It can be seen that the numerical localization performance of AttnGRUS in Table 3 is relatively better than the performance in the real-world testing. This difference is reasonable

due to the potential noise or uncertainty between the numerical and real-world experiments. Given both the numerical and the physical tests, the experimental results validate the effectiveness and the accuracy of the proposed method.

5   Conclusion

This paper proposes a supervised learning frame that integrates both model-driven method for dataset generation and a data-driven method for localization of structural impact loads. Aiming at accurate localization of transient impact, a novel attention-based GRU stacks model is proposed. This deep neural network model can effectively handle a large scale of multivariate time-series signals and estimate the location of the excitation source accurately. In the future work, transfer learning will be implemented to finely tune the proposed model for real-world deployment. In addition, the optimization of sensor placement can be determined by further investigating the attentive importance of each structural place in a regression process.

References

1. Bahdanau, D., Cho, K., Bengio, Y.: Neural machine translation by jointly learning to align and translate. In: International Conference on Learning Representations (2015)
2. Cho, K., et al.: Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078. https://doi.org/10.48550/arXiv.1406.1078 (2014)
3. Das, S., Saha, P., Patro, S.K.: Vibration-based damage detection techniques used for health monitoring of structures: a review. J. Civ. Struct. Health Monit. **6**(3), 477–507 (2016). https://doi.org/10.1007/s13349-016-0168-5
4. Fatahi, L., Moradi, S.: Multiple crack identification in frame structures using a hybrid Bayesian model class selection and swarm-based optimization methods. Struct. Health Monit. **17**(1), 39–58 (2018). https://doi.org/10.1177/1475921716683360
5. Flah, M., Nunez, I., Ben Chaabene, W., Nehdi, M.L.: Machine learning algorithms in civil structural health monitoring: a systematic review. Arch. Comput. Methods Eng. **28**(4), 2621–2643 (2020). https://doi.org/10.1007/s11831-020-09471-9
6. Kim, T., Kwon, O.S., Song, J.: Response prediction of nonlinear hysteretic systems by deep neural networks. Neural Netw. **111**, 1–10 (2019). https://doi.org/10.1016/j.neunet.2018.12.005
7. Kingma, D., Ba, J.: Adam: a method for stochastic optimization. In: International Conference on Learning Representations, December 2014. https://doi.org/10.48550/arXiv.1412.6980
8. Li, T., Pan, Y., Tong, K., Ventura, C.E., de Silva, C.W.: A multi-scale attention neural network for sensor location selection and nonlinear structural seismic response prediction. Comput. Struct. **248**, 106507 (2021). https://doi.org/10.1016/j.compstruc.2021.106507

9. Li, T., Pan, Y., Tong, K., Ventura, C.E., de Silva, C.W.: Attention-based sequence-to-sequence learning for online structural response forecasting under seismic excitation. IEEE Trans. Syst. Man Cybern. Syst. **52**(4), 2184–2200 (2022). https://doi.org/10.1109/TSMC.2020.3048696

10. Liu, Q., Wang, F., Li, J., Xiao, W.: A hybrid support vector regression with multi-domain features for low-velocity impact localization on composite plate structure. Mech. Syst. Sig. Process. **154**, 107547 (2021). https://doi.org/10.1016/j.ymssp.2020.107547

11. Obuchowski, J., Wyłomańska, A., Zimroz, R.: Selection of informative frequency band in local damage detection in rotating machinery. Mech. Syst. Sig. Process. **48**(1), 138–152 (2014). https://doi.org/10.1016/j.ymssp.2014.03.011

12. Qiu, B., Zhang, M., Xie, Y., Qu, X., Li, X.: Localisation of unknown impact loads on a steel plate using a pattern recognition method combined with the similarity metric via structural stress responses in the time domain. Mech. Syst. Sig. Process. **128**, 429–445 (2019). https://doi.org/10.1016/j.ymssp.2019.04.015

13. Seventekidis, P., Giagopoulos, D., Arailopoulos, A., Markogiannaki, O.: Structural health monitoring using deep learning with optimal finite element model generated data. Mech. Syst. Sig. Process. **145**, 106972 (2020). https://doi.org/10.1016/j.ymssp.2020.106972

14. Sony, S., Laventure, S., Sadhu, A.: A literature review of next-generation smart sensing technology in structural health monitoring. Struct. Control. Health Monit. **26**(3), e2321 (2019). https://doi.org/10.1002/stc.2321

15. Vaswani, A., et al.: Attention is all you need. In: Advances in Neural Information Processing Systems, vol. 30 (2017)

16. Zargar, S.A., Yuan, F.G.: Impact diagnosis in stiffened structural panels using a deep learning approach. Struct. Health Monit. **20**(2), 681–691 (2021). https://doi.org/10.1177/1475921720925044

17. Zhang, W., Zhang, Y., Song, R., Liu, Y., Zhang, W.: 3D layout estimation via weakly supervised learning of plane parameters from 2D segmentation. IEEE Trans. Image Process. **31**, 868–879 (2021). https://doi.org/10.1109/TIP.2021.3131025

18. Zhao, B., Cheng, C., Peng, Z., Dong, X., Meng, G.: Detecting the early damages in structures with nonlinear output frequency response functions and the CNN-LSTM model. IEEE Trans. Instrum. Meas. **69**(12), 9557–9567 (2020). https://doi.org/10.1109/TIM.2020.3005113

19. Zhou, J., Dong, L., Guan, W., Yan, J.: Impact load identification of nonlinear structures using deep recurrent neural network. Mech. Syst. Sig. Process. **133**, 106292 (2019). https://doi.org/10.1016/j.ymssp.2019.106292