CIAMTIS
U.S. DOT Region 3 University Transportation Center

Utilizing daily traffic as a sensor network for infrastructure health monitoring

July, 2024

Prepared by:
S. Pakzad, M. Takac, Lehigh University

r3utc.psu.edu

LEHIGH UNIVERSITY

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

Technical Report Documentation Page

1. Report No. CIAM-COR-R50	**2. Government Accession No.**	**3. Recipient's Catalog No.**
4. Title and Subtitle Utilizing daily traffic as a sensor network for infrastructure health monitoring		**5. Report Date** July, 2024 **6. Performing Organization Code**
7. Author(s) Shamim N. Pakzad, Martin Takac		**8. Performing Organization Report No.**
9. Performing Organization Name and Address Lehigh University, 27 Memorial Drive West Bethlehem, PA 18015 USA		**10. Work Unit No. (TRAIS)** **11. Contract or Grant No.** [69A3551847103]
12. Sponsoring Agency Name and Address U.S. Department of Transportation Research and Innovative Technology Administration 3rd Fl, East Bldg E33-461 1200 New Jersey Ave, SE Washington, DC 20590		**13. Type of Report and Period Covered** Draft Final Report 2/15/2021 – 11/14/2022 **14. Sponsoring Agency Code**

15. Supplementary Notes

[Enter sponsor contact's Name, Email, Phone]

16. Abstract

Mobile sensing is a novel paradigm that offers numerous advantages over conventional stationary sensor networks for real time bridge monitoring. Mobile sensors have low setup costs, collect spatio-temporal information efficiently, and require no dedicated sensors to any particular structure. Most importantly, they can capture comprehensive spatial information using few sensors. The advantages of mobile sensing combined with the ubiquity of smartphones with internet of things (IoT) connectivity have motivated researchers to consider smartphones carried within vehicles as large-scale sensor networks that can contribute to the health assessment of structures. A practical implementation of mobile sensors has several challenges. Most notably, the signals collected within a vehicle's cabin is contaminated by the vehicle suspension dynamics and the road profile; therefore, the efficient extraction of bridge vibration from signals collected within the vehicle is of great importance. The majority of available approaches for addressing this are typically system specific and restricted by assumptions of linearity. This limits the scope of application since vehicles mostly act nonlinearly depending on their manufacturing specifications. In addition, the variety of vehicle systems and road conditions complicates the exploration for a unified method for this task. This project proposes deep learning frameworks with domain adaptability that enable vehicle signal decontamination in a more reliable and practical manner. This framework will transform vehicles into robust and high-quality vibration sensors for infrastructure monitoring. Furthermore, this will render smartphone-based vehicle sensing data a valuable source of information that will enable crowdsourcing and facilitate infrastructure condition assessment in real time at an unprecedented scale, rate and resolution.

17. Key Words Mobile sensing, fatigue life estimation, remaining useful life, transfer learning	**18. Distribution Statement** No restrictions. This document is available from the National Technical Information Service, Springfield, VA 22161

19. Security Classif. (of this report) Unclassified	**20. Security Classif. (of this page)** Unclassified	**21. No. of Pages** 25	**22. Price**

Form DOT F 1700.7 **(8-72) Reproduction of completed page authorized**

Table of Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

Indirect bridge health monitoring necessitates input estimation for vehicles because response that can be conveniently collected within the cabin is highly contaminated by the vehicle suspension. Additionally, road conditions and vehicle speed play a critical role in the quality of collected cabin data. Yang et al. [2020] demonstrated that by solving the inverse problem deconvolving the measured cabin accelerations to recover the tire-level input - the detrimental suspension effects could be mitigated. Eshkevari et al. [2020a] showed that under a linearity assumption for the vehicle suspension, the vehicle cabin signal can be deconvolved using a transfer function. Nayek and Narasimhan [2020] proposed a Gaussian process latent force model that successfully estimated the road input for a linear, damped, two degree of freedom (DOF) quarter car suspension model. However, all these models require complete knowledge about the system which is impractical for real-world applications. To address that, Eshkevari et al. [2020a] also proposed a data-driven method based on blind source separation (BSS) techniques. While avoiding the modeling complications, this method has alternate limitations stemming from the BSS assumptions pertaining to the system linearity and signal's frequency distribution.

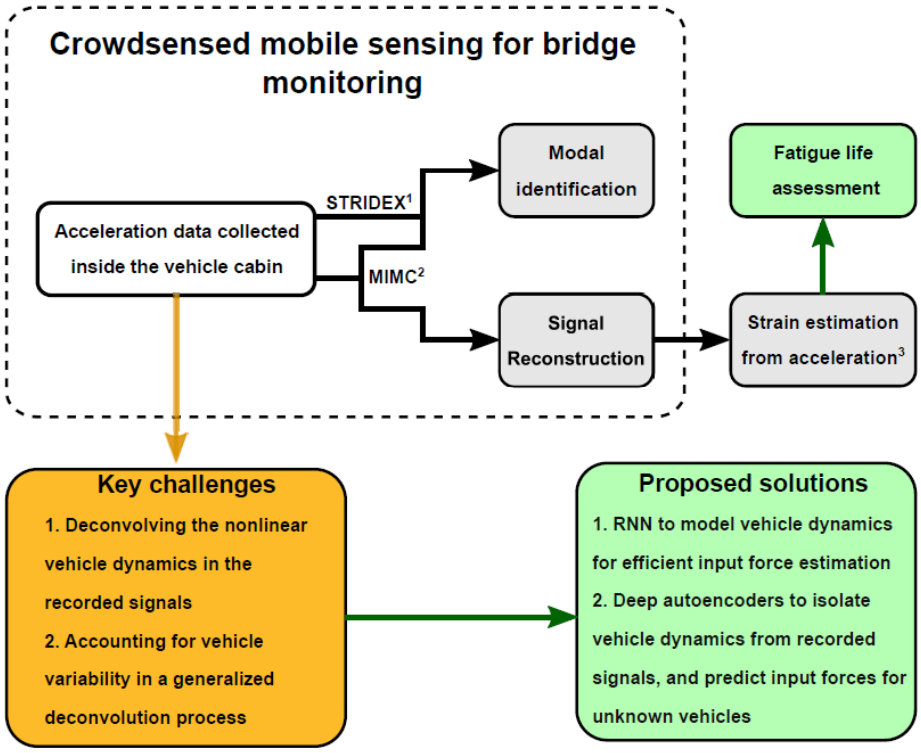

Figure 1. A general framework of the proposed methodology.

Our previous works demonstrate the efficacy of utilizing vehicles as mobile sensors for monitoring of bridges. In particular, we proposed multiple methods that can process a pool of mobile sensing data that are collected in controlled or uncontrolled settings and extract modal properties of a bridge [Eshkevari et al., 2020 b,c]. In this project we aim to generalize this paradigm such that the new framework will: (a) filter out the signal contamination caused by the vehicle dynamic system, (b) account for real-world nonlinear systems that constitute the majority of commercial vehicles by introducing a deep learning solution, and (c) extend the application of the learned neural network to various vehicles with minimal computational effort. The outcome of this project is a comprehensive framework for using vehicles as sensors for crowdsensed vibration-based bridge monitoring and modal identification. Finally, to demonstrate the application of the proposed framework, it will serve as a basis for bridge fatigue assessment. Figure 1 shows an overview of the project rationale.

BACKGROUND

Crowdsensed vibrations are typically recorded from inside the cabin of a vehicle. These responses are contaminated by vehicle and pavement induced dynamics. For all practical purposes, the vehicle dynamics of all modern commercial vehicles are nonlinear in nature (Maher and Young [2011]). Hence, effective deconvolution of recorded responses from vehicle dynamics is imperative prior to its use for bridge monitoring applications. This gives rise to an input force identification problem, wherein, the goal is to efficiently estimate the input force that a vehicle is subjected to from responses recorded from within the vehicle cabin. Inverse dynamic problems are of great importance for estimating demands that dynamical systems are subjected to. The inherent complexity of such problems promoted the development of black-box solutions. Deep learning has shown great potential as universal function approximators for learning complex relationships and patterns from raw data. Sarego et al. [2018] utilized a multi-layer perceptron (MLP) with the real and complex values of the response's FFT as an input for force reconstruction of impact loads for composite materials. Zhou et al. [2019] developed a long short-term memory (LSTM) network to recover an impulse with only the acceleration response as input. Due to the nonlinearity and complexity of realistic dynamic systems, it is preferred to design an approach that accomplishes the input estimation with no baseline model or restrictive assumptions.

OBJECTIVES

The outcome of this project will transform the concept of mobile sensing for infrastructure monitoring from its current nascent form to a practical real-time tool for large scale infrastructure monitoring at the city scale. The proposed algorithms will facilitate the use of crowdsensed vehicle response data for continuous monitoring. This will enable bridge owners circumvent tedious and costly fixed sensor deployment on civil infrastructure and harness the power of big-data for structural condition assessment. Subsequently, this framework can be used for vulnerability assessment, life cycle analysis and in formulation of resource allocation policies, paving the way for smart resilient infrastructure systems.
The deliverables of the project include:

1. A recurrent neural network framework for efficient input force identification, crucial for deconvolving nonlinear dynamic response of vehicle suspension systems from response collected from inside a vehicle.
2. An autoencoder architecture that generalizes the deconvolution process. The proposed architecture will be trained on a set of known vehicles and subsequently be used for vehicles whose system dynamics are unknown.
3. A variational setting for tuning deconvolution results based on variations in mechanical properties (e.g., vehicle weight) and data collection positioning (e.g., sensor angles and mounting location).

DATA AND DATA STRUCTURES

The project evaluates the performance of the proposed RNN framework through real-world and numerical case studies. For objective 1, an experiment was designed and conducted in order to estimate the input of a real-world vehicle using its cabin acceleration data. The details are provided in the subsequent sections of this report. For objective 2, a numerical case study, using response from five different vehicles modeled as nonlinear two degree of freedom (DOF) quarter cars has been used.

For objectives 3, the project shows an extensive real-world study on bridge monitoring with crowdsourced smartphone-vehicle trips in which absolute value mode shapes are estimated. Data collected from the Gene Hartzell Bridge over Lehigh River have been exploited to validate the results related to the objective 3.

CHAPTER 2

Methodology

PROPOSED RNN FRAMEWORK

Recurrent Neural Network: Data-driven methods for learning nonlinear dynamic systems

Overview

In this context, we developes a recurrent neural network (RNN) framework that learns the nonlinear transformation between the input and output of such dynamic systems, with particular emphasis on vehicle systems (Figure 2). The major contributions of this section include: (a) implementing the network on realistic vehicles subjected to various road conditions and achieving promising input estimations, (b) proposing a confidence measure for the input regression task, and (c) designing a practical and inexpensive sensor setup for data collection. In this phase, we conducted real-world experiments to show the adaptability of the method, ease of implementation, and reliability. This case study is directed toward indirect bridge health monitoring with the objective of estimating vehicle contact point input given its cabin vibrations.

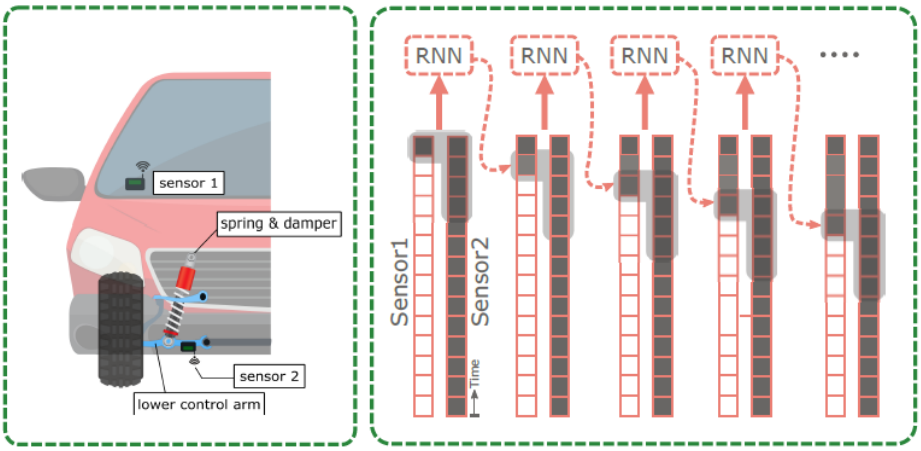

Figure 2. RNN network architecture and inference.

(Eshkevari et al. 2022. Input estimation of nonlinear systems using probabilistic neural network)

Architechture and training

In linear time invariant dynamic systems, a finite impulse response (FIR) filter can replace the exact transfer function of the system. This simplicity however is not practical for nonlinear systems owing to input amplitude dependency of the transfer functions. Yet, the state transition equation of a nonlinear time invariant dynamic system can be represented as in Eq. (1).

$$x_{k+1} = f(x_k, u_k) + v_k \tag{1}$$

where x_k is the full state vector of the system at time step k, $f(.)$ is a characteristic function of the system, u_k is the applied load at time step k, and v_k is the process noise. This state equation usually pairs with an observation equation using which the measurement vectors and full state vectors are related. In the direct dynamic problems, given u_k and initial state x_0, the full state space can be estimated in a data driven fashion with no need for available estimates of $f(.)$. Recurrent neural architectures have been widely proposed for signal regression tasks due to their capability of learning temporal dependencies and dynamic equations. In this project, the main objective is to tackle the inverse problem: given x_k for $k \in 1:T$ and a prior estimate for u_T , it is desired to estimate u_k for $k \in 1:T-1$. This problem is equivalent to the response deconvolution of a dynamic system (linear or nonlinear) without using the prior knowledge about the system. Note that in case of linear systems, the methods proposed in (Eshkevari et al. 2020) or (Yang et al. 2018) are valid candidates, while the proposed neural network in this project approximates a nonlinear deconvolution filter to estimate the system input. Fig. 2 presents a schematic overview of the inference using the proposed framework. In this figure the neural network is represented as an RNN block with an inverted L-shape input; at each time step the RNN block processes the input and output values inside the L-shape binder to predict the one-step backward estimation of the input. This process is repeated until the maximum possible length of the input signal is estimated. In this framework, the input signals are associated to the CP (contact point) or ground motion of the vehicle or building systems, respectively. Besides, cabin signal and story accelerations are systems' outputs. Note that the figure depicts a single input single output (SISO) case in which the number of response channels equals to one. However, in MDOF systems, the network dimensions adapt accordingly with no substantial change in the proposed structure or the pipeline. As it is shown, the network unravels the input signal in a reverse order. This is not a necessity, however, the RNN requires one initial input for initiating the recurrent process and it is more practical to estimate the force when the signal collection ends (i.e., when system returns to at-rest condition). On the other hand, the forward-unraveling option is more desired for real-time input estimation. In this study, the reverse-unraveling approach is adopted due to its slightly better performance. The neural architecture of the RNN block incorporates fully-connected layers for transitioning between two consecutive time steps. The data flows through two stages in the network: shared layers and channel-specific layers. Channels are defined as input signals collected from different axes (i.e., x, y, z) or at different degrees of freedom (i.e., story one). By using the shared layers, the network is constrained to learn information that are equally useful for input estimations in different channels. In addition, the number of learning variables is reduced significantly. In contrast, our preliminary experiments showed that by merely utilizing shared layers, the performance is noticeably lower suggesting the use of a few channel specific layers at the terminal end of the network flow.

In general, learning-based regression models are designed to directly estimate regressed outputs with no confidence quantification (e.g., estimate scalars or a set of values). In contrast, the estimated outputs of a classifier are class probabilities which are also useful for uncertainty analyses. For example, in the case of risk-averse problems, by setting higher bars for the classification probabilities, it is possible to enhance the accuracy of the classification predictions at the expense of lowering the recall. Such estimation confidence analyses are not possible when the neural network outputs deterministic values (i.e., regression models). To address that, we introduce a probabilistic learning-based regression model that estimates a normal distribution for the regressed values, instead of estimating actual values.

During the training process, the optimization objective is designed to push the mean values of the normal distributions to the actual regression values and shrink the variance for reaching higher confidence.

For inference, the means of distributions are considered as actual regressed predictions while standard deviations indicate the prediction confidence (e.g., low standard deviation means a narrow normal distribution which translates to high confidence). The last layer of each channel-specific network predicts μ and σ constructing the normal distribution for that particular channel. To certify that $\sigma > 0$, the value of the associated output node is passed through an exponential function. The proposed network is termed as a probabilistic regression model (consistent with (Nix et al. 1994 and Lakshminarayanan et al. 2016)) since its final product is not a deterministic regression value but a Gaussian probability distribution. Therefore, the network not only estimates the regression value, but also quantifies its probability (which is inversely proportional to uncertainty and aleatoric noise level as shown in (Lakshminarayanan et al. 2016). For inference, the output inevitably collapses to the mean values of the distributions to enable the recurrent feedback. The proposed methodology consists of two steps: training and inference. In the training phase, multiple input and output signals from the system of interest are required so that the dynamical system can be learned by the RNN model. This is potentially made possible by temporarily sensing the system's input or using finite element surrogate models for simulation. In the inference stage, however, the only input value that should be available is the input at the terminal state (the systems' input at the last discrete value of the signal). This input value in many applications can be simply set to zero considering an at-rest condition at the end of the sensing period (e.g., buildings after an earthquake will return to zero acceleration). Given this trivial input state, the system can unravel the previous inputs by processing the outputs that are fully available. To train the proposed probabilistic regression model, conventional error minimizing loss functions are not applicable since these functions incorporate deterministic values rather than distributions. Instead, the loss function has to directly incorporate the negative log likelihood of the observations given the model parameters. In this context, the RNN block is parameterized by θ and the goal is to maximize the probability of correctly estimating targets y_i given system inputs x_i under the trained parameters. With this definition, the loss function $L(\theta)$ is defined as the following:

$$L(\theta) = -\log(p(y_i\,|x_i, \theta)) + L_{proj}\,(\theta, n_{proj}) \tag{2}$$

Where $p(y_i|x_i, \theta)$ is the probability of drawing system input y_i given system output x_i and model parameters θ. The second term of the loss function is the projection loss with a projection length of n_{proj} (defined in Eshkevari et al. 2021) and based on the parameterized model $f(\theta)$. As the likelihood term becomes smaller, we ensure that the network's output distributions are more likely to predict values that are close to the actual outputs. The second term of the loss function also attempts to enhance the regression accuracy for longer projections in a conventional mean squared error (MSE) minimization manner. In this term, a strictly increasing geometric factor is element-wise multiplied to the outputs in the trajectory to put more weight on the accuracy of more distant estimations. To train the network using this loss function, Newton trust-region approach is adopted due to its outstanding performance compared to linear methods. The training process consists of two phases: 50 epochs with projection length of five and 20 epochs with projection length of 50. Depending on the case study and its characteristics, the architecture of the network is modified in order to adapt with the proper number of output channels. This variation results in different number of trainable parameters. Regardless, the dimension of the RNN variable space is between 5,000 and 15,000 in all cases tested in this study. To avoid overfitting, the performance of the models is evaluated using the cross validation technique. The inclusion of the second term in the loss function is effective in preventing overfitting for its emphasis on long-run projected signals rather than one-step estimations. In addition, the compact design of the network helps to avoid overfitting by reducing the model variance.

Validation: Real world vehicles

An experiment was designed and conducted in order to estimate the input of a real-world vehicle using its cabin acceleration data. In this experiment, the data is collected in two locations: inside the vehicle cabin and in proximity to the CP. Note that the actual vehicle's CP is practically inaccessible for a sensor device. Therefore, the lower control arm was selected as a feasible location, and a manually assembled sensor bundle was attached to that location. The sensors were wirelessly communicating with a computer which was held by the operator in the passenger's front seat. The cabin sensor was attached to the dashboard of the vehicle. The sensor layout is presented schematically in Fig. 3(a). As it is shown in this figure, sensor 2 is mounted on the lower control arm which is found to be a suitable location for the vehicle input data collection and is not affected by the suspension springs. The arm is a solid beam attached to the rim and is located right before the spring and the shock absorber on the load path from the tire to the vehicle cabin. The sensor bundle used for vehicle data collection is shown in Fig. 3(b) (similar configuration is used in both locations). The bundle consists of three components: (1) a Raspberry Pi zero board, (2) an ADXL345 accelerometer, and (3) a power source. The Raspberry Pi is selected for its data processing and storage functionality as well as its low cost, easy programming, and wireless connectivity. ADXL345 is a three-axis accelerometer which is compatible with Raspberry Pi and collects data with high rate. The acceleration range and sampling frequency can be tuned based on the application and required accuracy. To select these parameters, a lab-scale experiment was conducted on a single degree of freedom system and the accuracy of the neural network predictions was compared for data collected from different sensor settings. Based on this preliminary study, the sampling frequency of 500 Hz and acceleration range of ±16.0 g were set for the final experimental trial. Note that the adjusted frequency is an upper bound for the sensor and in practice, the sensor collects data with nonuniform time intervals and lower rates. This is found affected by the throughput rate of the Raspberry Pi and its wireless communication.

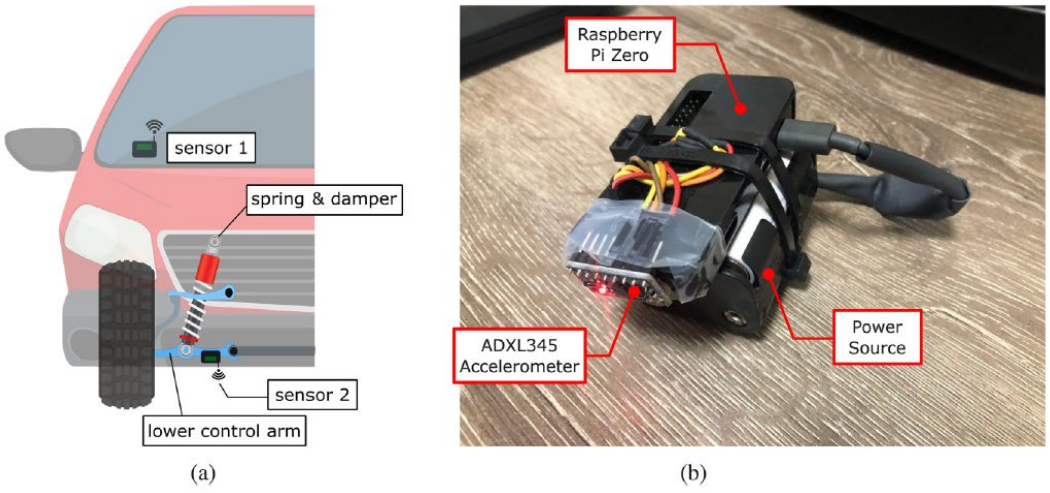

Figure 3. (a) Schemati view of the car and sensor layout, (b) sensor setup used in the experiment: the main board is a Raspberry Pi zero and the sensing device is an ADXL345 accelerometer.

(Eshkevari et al. 2022. Input estimation of nonlinear systems using probabilistic neural network)

For the road test, a KIA Forte 2020 was equipped with the sensor sets. According to the vehicle's official specifications, the vehicle suspension is equipped with nonlinear suspension systems in front and rear positions. In particular, the suspension system consists of MacPherson strut and twin tube shock absorbers that both exhibit nonlinear behaviors. The instrumented vehicle was driven over roads with

different roughness conditions, including recently paved, poor condition, and gravel roads near Lehigh University campus. In total, 23 scans of 50,000 samples were collected. The vehicle speed was mostly kept within 10–12.5 mph, however, in rare situations of traffic congestion in the testing area, the speed varied. The collected data was then preprocessed for training which included the following steps: (1) signal resampling in order to even the time intervals between samples, (2) signal filtering using a band-limited filter, and (3) downsampling to 100 Hz. Filtering and downsampling steps reduce high frequency noise as well as measurement drifts in the collected signals.

After preprocessing, signals were normalized linearly using the previously explained approach. This approach for normalization is found to yield better performances compared to other conventional methods (e.g., based on maximum absolute value). The training process of the real-world vehicle experiment is the same as the previous case studies. From 23 scans, 10, 1, and 12 samples are randomly picked for training, evaluation, and testing, respectively. Note that the majority of data are kept for testing for better performance assessment. As an important note, in this experiment the vehicle's suspension system is assumed as a quarter-car and the objective is to predict three-axis vibrations at sensor 2 given collected responses at sensor 1 (see Fig. 3(a)). However in reality, the vehicle receives input from all four tires, meaning that a more realistic scenario would be to train the network based on inputs from four tires and outputs from a complete set of vehicle cabin vibrations (including translational and rotational accelerations). In this study, the simplified version of the problem is investigated at the expense of a lower input prediction accuracy. The attached sensors were not perfectly oriented with respect to the global axes and as a result, the local axes of the sensor do not match the global orientations.

TRANSFER LEARNING FOR INPUT ESTIMATION OF VEHICLES SYSTEMS

Overview

The recurrent neural network implementation described in the previous section requires independent training for distinct vehicles. In this project, we further generalize the framework using Transfer learning (TL), an emerging and extremely popular domain in machine learning (ML), that deals with transferring knowledge learned for a specific task to other models with similar but not exactly the same task. In detail, we construct a deep autoencoder that will allow us to adapt a previously learned network for a specific vehicle to predict the response of an unseen vehicle when it traverses over a pavement. This will ensure robustness of the crowdsensed monitoring framework to vehicle variability and transform it into a large-scale and practical infrastructure monitoring paradigm.

Methodology

Figure 4 presents the schematic architecture for this novel framework for tire-level input estimation for a variety of suspension systems through domain adaptation. The network will separate the road and vehicle information from the recorded cabin vibration signals to create two physically meaningful latent features: road (RLF) and vehicle (VLF). RLF will then be decoded for accurate tire-level input estimation for unknown vehicle suspension systems by cross-inferring the two autoencoders. Furthermore, VLF will allow for vehicle type classification based on its dynamic characteristics. This network will ensure robustness of a crowdsensed bridge health monitoring framework to variations in vehicle type and road profiles.

In detail, autoencoders are used to extract features such that the road (system input) latent features are void of information about the vehicle type (system). To extract these features, the network has two autoencoders for the road signal and cabin signal with the road latent feature being shared between the two networks. An adversarial training setup forces the road specific latent feature to be invariant to the vehicle

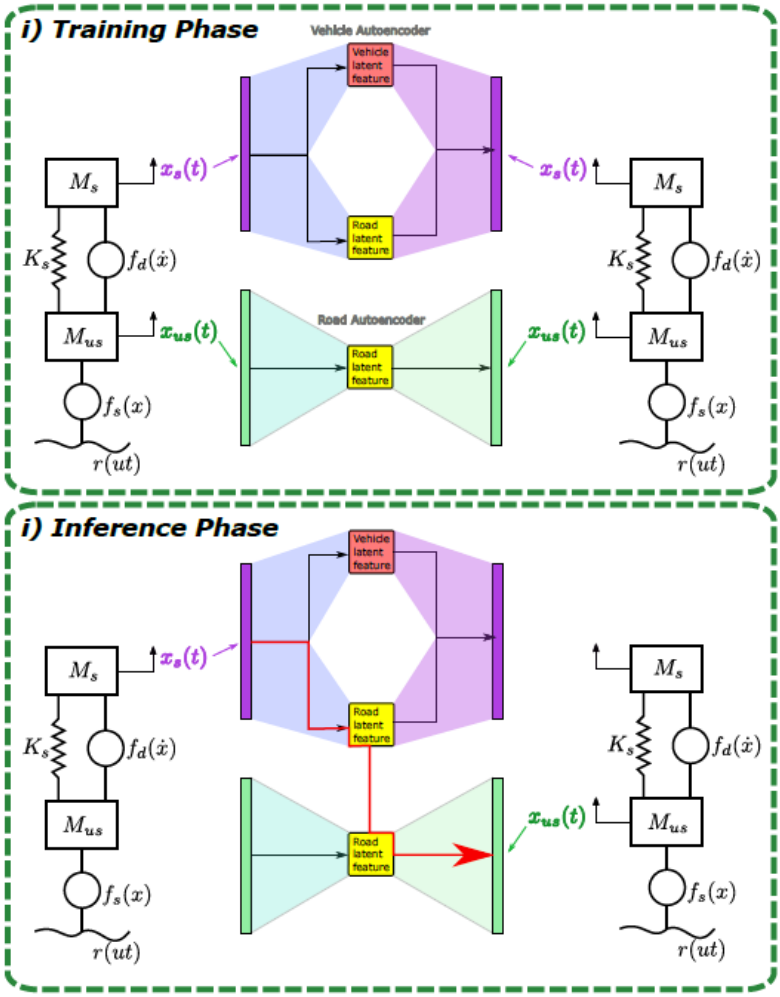

Figure 4. Architecture of the proposed transfer learning (TL) framework.

(Cronin et al. 2021. Transfer learning for input estimation of vehicle systems)

type, meaning the only information left is that of the road profile. The network consists of two main sections: road autoencoder and vehicle autoencoder. The road encoder (ECR) transforms the road signal to an arbitrary latent space creating the road latent feature (RLF). Using RLF, the road decoder (DCR) attempts to reconstruct the original road signal.

The reconstruction loss for the road is the mean squared error (MSE) between the input and reconstruction (L1). The vehicle autoencoder also contributes a corresponding reconstruction loss, L2. RLF has an additional constraint from vehicle to the road encoder (R-ECV), which tries to map the cabin signal to the same latent space. This is achieved by minimizing the difference between RLF created by ECR and R-ECV (L3). Subsequently, we use R-ECV and DCR for estimating the vehicle input (i.e., network cross-inference).

The primary purpose of this network is for the tire-level estimation, however, additional information can also be recovered about the type of vehicle and the expected accuracy of the network, which is directly related to the dynamics of the vehicle. An additional cabin signal encoder (V -ECV) is used to extract vehicle information (V LF) from the cabin signal which is used as an input for the vehicle classifier (CL) and DCR. CL, with classification loss L5, attempts to classify vehicle type. The cabin signal can be reconstructed by combining the road and vehicle information through concatenating RLF and VLF as an input to the vehicle decoder (DCv). However, there is no guarantee without additional constraints that information from the car is not embedded in RLF as well from this optimization process. An adversarial training process is used to apply this constraint to the latent feature. The networks in Figure 4 are trained in an adversarial process to make the vehicle type indiscernible to the network in RLF. Using an adversarial training process, all other parameters are trained with the loss shown in Equation 3 while the weights of CL_{adv} are not trainable. As shown in this equation, the classification loss of CL_{adv}, L4 appears with a negative sign, implying the intention for maximization. Every K (user defined parameter) epochs, all parameters of the networks except for CL_{adv} weights are frozen and CL_{adv} parameters are optimized for L4 minimization. By iterating this process, the network makes the latent feature invariant to the car type and the adversarial classifier is unable to predict vehicle classes with accuracy greater than 1/n with n being the number of cars in the dataset.

$$L_{total} = L_1 + L_2 + L_3 + L_4 + L_5 \quad (3)$$

We demonstrate the efficacy of the proposed architecture through a numerical case study, wherein we train our autoencoders using response from five different vehicles modeled as nonlinear two degree of freedom (DOF) quarter cars. The nonlinearity stems from a bilinear damping of the suspension and a second order polynomial stiffness of the tire . Figure 5 shows the quarter car model along with its parameters.

We generate the road profile as a stochastic field from the power spectral density $G(\lambda) = G_{\lambda 0}(\lambda/\lambda_0)^{\nu}$, with $\nu = -2$,$G_{\lambda 0} = \gamma * 6 \times 10^{-3}, \gamma \sim U(0.1)$, and λ being the spatial frequency. Table 1 lists the model parameters for each vehicle class used for training. The other parameters of the model are assumed the be constant for all the vehicles with the following values: $\beta_1 = 4.23,\ \beta_{12} = 15.49,\ v_+ = 0.0045\frac{m}{s}, v_- = 0.005\frac{m}{s}, \alpha = 0.1\ m^{-1}, and\ u = \frac{5m}{s}$.

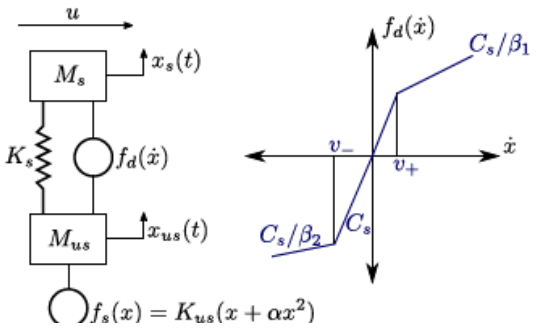

Figure 5. Non linear quarter car model.

(Cronin et al. 2021. Transfer learning for input estimation of vehicle systems)

Table 1. Vehicle model parameters for each class

Class	Ms (kg)	Mus (kg)	Cs (N/m/s)	Ks (N/m)	Kus (N/m)
1	305.6	150.6	1.61 x 10^3	1.77 x 10^4	1.48 x 10^5
2	429.8	13.39	2.46 x 10^3	3.48 x 10^4	2.68 x 10^5
3	487.2	129.7	5.79 x 10^3	2.32 x 10^4	3.11 x 10^5
4	2141	74.08	4.34 x 10^3	1.22 x 10^4	4.93 x 10^5
5	372.9	206.5	2.76 x 10^3	1.76 x 10^4	1.69 x 10^5

We generate 1, 000 samples from the model including road input acceleration ($\ddot{r}(ut)$), unsprung mass ($\ddot{x}_{us}(t)$), and sprung mass ($\ddot{x}_s(t)$),, acceleration signals. We split the data 7 : 3 into training and evaluation datasets and subsequently train the road and vehicle autoencoders.

APPLICATIONS IN BRIDGE HEALTH MONITORING

Overview

We apply the developed methodologies for different phases of a crowdsensing-based bridge health monitoring framework (the framework was previously designed and published by our research team (Matarazzo and Pakzad, 2018, Eshkevari et al., 2020b). We demonstrate the efficacy of the proposed frameworks through experiments. This project exploits mobile data collection from lab-scale and real-world bridges using smartphones as well a novel low-cost sensing bundle.

Methodology

We performed the most extensive real-world study to date on bridge monitoring with crowdsourced smartphone-vehicle trips in which we estimate absolute value mode shapes (AMS) and simulate damage detection capabilities. The methodology is based on the synchrosqueezed wavelet transform converting each acceleration time series to the time-frequency domain; in this way the instantaneous magnitudes are calculated. The bridge is divided into overlapping segments; the width (Δ) and spatial stride (S) are two parameters of the method. Averaging the magnitudes in each segment for each trip at the modal frequencies yields a distribution of magnitudes at each location, the mean of which is the AMS estimate. Including a small bandwidth (ε) around the modal frequency in the averaging process leads to robust results on noisy datasets. The method analyzes over 500 trips across four bridges with main spans ranging from 30 to 1300 meters in length, representing about one-quarter of US bridges. We demonstrate a bridge health monitoring platform compatible with ride-sourcing data streams that check conditions daily. The result is the potential to commodify data-driven structural assessments globally. Included among the case studies is the Gene Hartzell Bridge. The SVT data called GHMB is a controlled dataset of 332 trips collected by the research team's vehicles and the Good Vibrations App. For data collection, the team drove at a set of prescribed speeds. The mean speed across all trips was 82.25 km/h. For analysis, the dataset was subdivided to consider the slowest 102 trips, with the intention of dividing large datasets into potentially more informative subsets.

CHAPTER 3

Findings

RESULTS FOR RNN FRAMEWORK

To evaluate the performance of the network for input estimation, the reconstructed input signals for one of the testing samples are presented in Figs. 6 and 7.

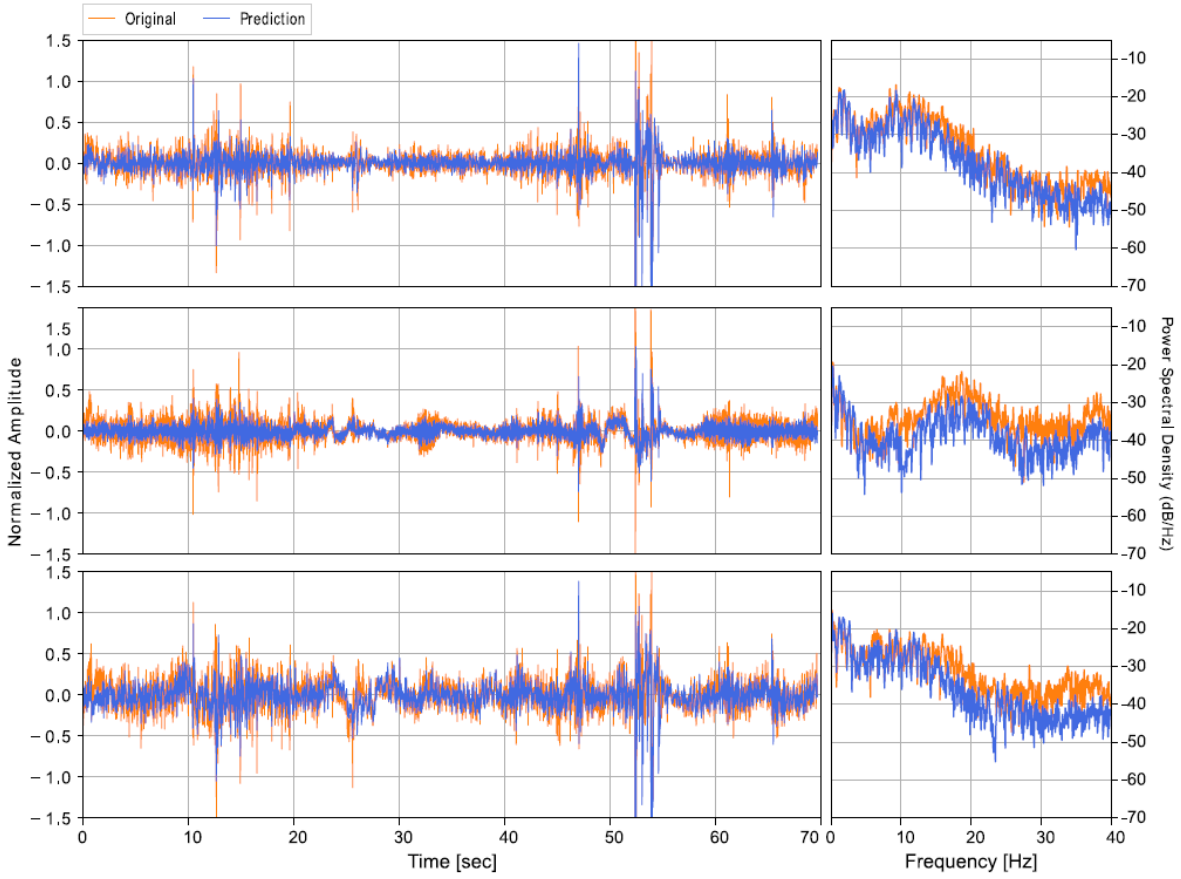

Figure 6. Vehicle input signal predictions in three axes. Long time projection is presented (70 s).

(Eshkevari et al. 2022. Input estimation of nonlinear systems using probabilistic neural network)

Figures generally confirm the efficacy of the input estimation in all three axes. The original input signal is highly nonstationary which is caused by irregular road conditions (such as road bumps or potholes) that complicates the process of learning. Yet, the trained network successfully estimated the overall patterns. From Fig. 7, the network tends to estimate input samples with low confidence (implied by the wider CI) and the number of missed points has increased compared to the previous case studies. This is also evident from the residuals shown with the shaded area in Fig. 7. Note that despite higher residuals, the number of missed points is not unreasonably high. This supports the fact that the network is aware of the high uncertainty in this domain. In fact, the network may not be able to exactly estimate the input signal (suggested by high residual), but the expected range is learned with high accuracy (limited number of missed points). To quantify the accuracy of the input estimations, random signal slices from the test dataset were estimated and metrics were calculated. In this analysis, the lower bound for the slice length is set to 200. For each estimation, the accuracy is measured in terms of the correlation coefficient, the MSE of the signals in time domain, and the MAE of the signals in frequency domain and presented in Table 2. From this table, the mean and median of correlation coefficients vary between channels with maximum in channel 2. These numbers are lower for channel 0 and channel 1, respectively.

Table 2. Statistical metric drived by processing the testing data

Channel	**Correlation Coefficient**		**MSE in time**		**MAE in frequency**	
	Mean	Median	Mean	Median	Mean	Median
0	0.6690	0.6863	0.0133	0.0078	2.5770	2.3783
1	0.5570	0.5554	0.0161	0.0111	3.2046	3.1082
2	0.7424	0.7473	0.0157	0.0115	2.9958	2.7669

From Table 2, it is hard to distinguish the channel with the highest accuracy in the time and frequency predictions since the range and mean values are quite similar. This implies that the trained network should have similar confidence for predicting regression values in these three channels. To investigate that, the histograms of standard deviations are presented in Fig. 8. In this figure the standard deviations range very similarly over 0.05 to 0.3 in different channels. One may notice that channel 0 is marginally inclined to the left (implying higher confidence in this channel), however, the difference is not strong. In other words, given the histograms of Fig. 8, it is expected that the performance of the trained network in all three axes is nearly analogous. More importantly, this finding is not a result of post-analysis using an available or labeled dataset, but it is readily available through the inference of the probabilistic network.

Following the same procedures as previous sections, Fig. 7 presents the predicted and original signals and also highlights the 95% confidence interval of the predictions. The strong majority of the predicted samples fall inside the confidence band, confirming the strength of the trained network. The outliers marked with circles indicate samples that the model cannot predict successfully, even considering the confidence interval. The occurrence frequency of such points can also be a good measure of the network performance. Finally, the interpretability of the trained network is analyzed. Fig. 9 presents the importance factors of the network's input features. As expected, due to the highly nonlinear nature of the vehicle suspension system, nonstationary input data, and the contribution of the vehicle inputs from other tires, the importance patterns are not as explainable as the previous cases. Yet, in all three channels the importance of the most distant time steps decreases. This is consistent with our general expectation regarding causality of dynamic systems. In addition, the cross-channel contributions are quite common in this experimental case (i.e., the contributions of channels perpendicular to a certain output channel are high). For instance, all three channels of the vehicle's input are strongly contributing in predicting outputs at channel 1 (Fig. 9(b)) which is unlike to the findings of the previous test cases.

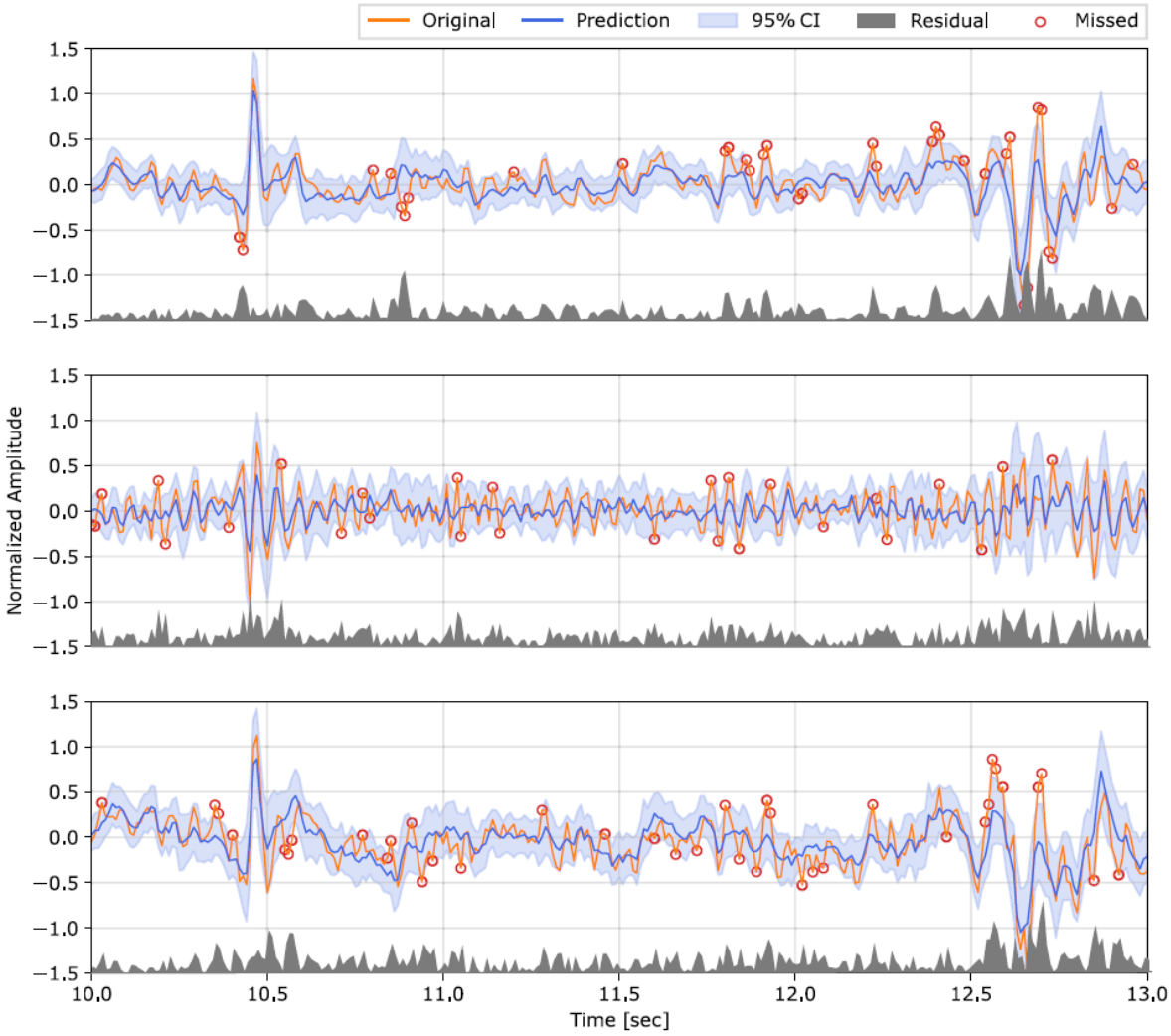

Figure 7. Vehicle input signal predictions in three axes. Short time projection is presented (3s).

(Eshkevari et al. 2022. Input estimation of nonlinear systems using probabilistic neural network)

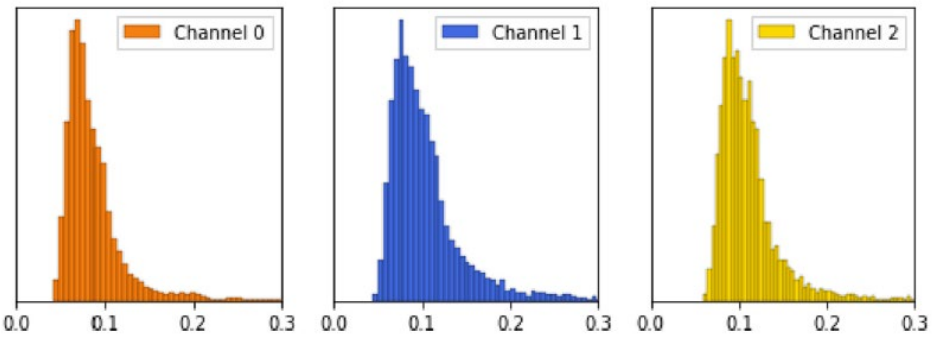

Figure 8. The histogram of standard deviations for three axes.

(Eshkevari et al. 2022. Input estimation of nonlinear systems using probabilistic neural network)

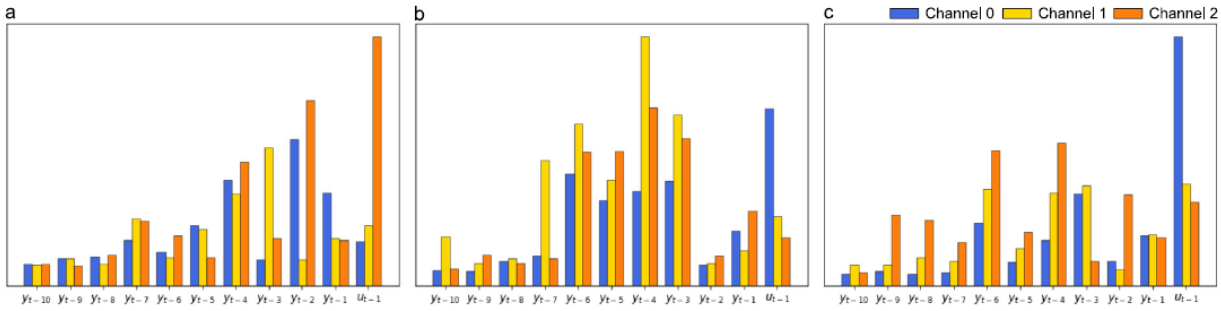

Figure 9. Importance map of the network features for different input channels: (a) Channel 0, (b) Channel 1, and (c) Channel 2. Legend refer to the channels of the vehicle output (cabin data).

(Eshkevari et al. 2022. Input estimation of nonlinear systems using probabilistic neural network)

As explained, the alignment of sensors at the two locations (Fig. 3(a)) was dictated by the best available attachment surface; as a result, the sensors were mounted with arbitrary orientations with respect to each other. Therefore, projections of multiple channels are expected to influence the prediction of any input channel in the vehicle's cabin.

RESULTS FOR TRANSFER LEARNING

One of the goals of this network is to separate information about the road and vehicle into RLF and V LF, respectively. Figure 10 is a two dimensional visualization of V LF using t-SNE. Not only do the five cars used in training produce distinctive clusters, but also cars with perturbed (1%, 5%, 10%, 15%, 20%, 30% and 40% variations) mechanical properties have similar latent representations to those most similar used in training. This confirms that the vehicle latent feature space is a continuous domain with similar cars closely spaced. In addition, having split and distinctive clusters enables CL to achieve high classification accuracy (95%, 93%, 91%, 86%, 87%, 76%, and 68%, respectively.).

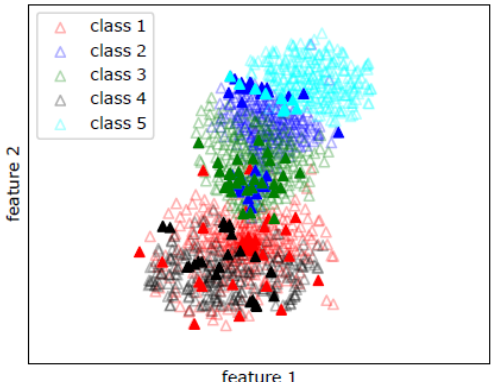

Figure10. Two-dimensional visualization using t-SNE of VLF. The open and filled markers represent the five vehicles used in training and vehicles with 40% perturbed mechanical properties, respectively.

(Cronin et al. 2021. Transfer learning for input estimation of vehicle systems)

This is important since not all cars achieve the same tire-level estimation accuracy, and a robust classifier enables to cherry pick realworld samples that belong to the vehicle classes with high tire-level estimation accuracy (Figure 11). The top portion of Figure 11 depicts the tire-level estimation accuracy. The histograms were generated using the network to estimate 1,580 samples of testing data with the correlation coefficient as a metric for accuracy. The plot shows coefficients ranging from 0.20 up to 0.98 for different classes. The best results are found for class 2 with median of 0.92. Note that the tire-level estimations are made by a crossinference of two autoencoders. In other words, the networks are not directly trained for estimating inputs from outputs and still the accuracy results are promising. Interestingly, the correlation coefficients estimates of each class of vehicles creates unimodal distributions implying that the properties of the sensing vehicle determines the accuracy of the tire-level input estimation. The bottom portion of Figure 11 shows the nonlinear transfer functions for the five vehicles used in training. Vehicles have a variety of suspension systems that affect the input signals differently. In other words, the output signal is a convolved version of the input signal with respect to these nonlinear transfer functions. Therefore, the vehicle classes that have low amplitude in a wide range of frequencies (e.g., class 1, 3, and 4), and have sharp low-damping spikes in the frequency domain (e.g., class 4 and 5) filter out the significant fraction of the frequency contents in the input. As a result, the filtered signal is not informative and input retrieval task becomes nearly impossible. This can explain the lower accuracy of classes 1, 4, and 5 with respect to classes 2 and 3. We believe the reduced accuracy for input recovery in some classes is inevitable and cannot be improved due to physical limitations. However, using CL (Figure 4) we can always select samples from reliable and high accuracy classes.

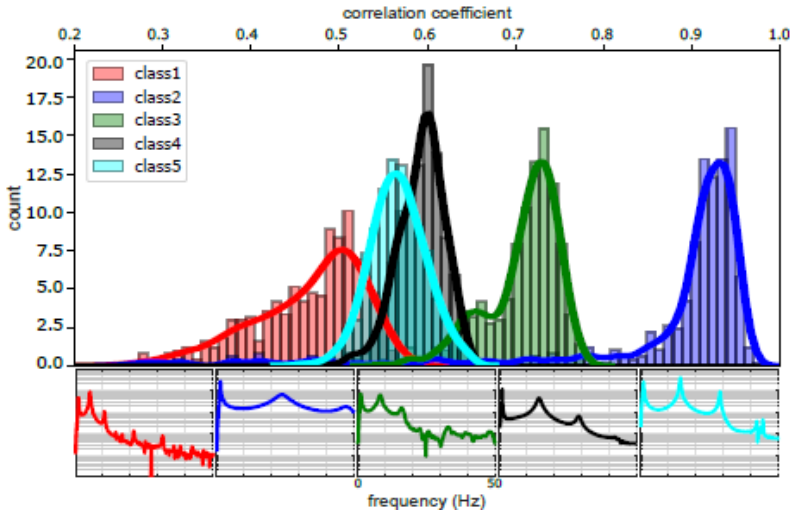

Figure 11. Top: histogram of correlation coefficients on the estimation results of 1,580 samples from the five vehicles used in training. Bottom: nonlinear transfer functions of five vehicle classes derived by averaging the input-output relationship for a range of input impulse amplitude.

(Cronin et al. 2021. Transfer learning for input estimation of vehicle systems)

RESULTS FOR BRIDGE HEALTH MONITORING-MODAL IDENTIFICATION

Four case studies demonstrated a wide range of real-world examples in which important spatial vibration characteristics of bridges, i.e., absolute value mode shapes (AMSs), can be extracted successfully from crowdsourced smartphone-vehicle trip (SVT) data. Bridges from three general classes were considered: short-span, medium-to-long-span, and long-span, and three types of crowdsourced SVT datasets were used: controlled, partially controlled, and uncontrolled. The successful applications on the Golden Gate Bridge (GGB) show capabilities for identifying AMSs of long-span bridges in both vertical and horizontal directions (see Figure 12a-e). The studies on the Cadore Bridge (CAD) and Gene Hartzell Memorial Bridge (GHMB) confirmed the method's applicability to a larger group of existing bridges with maximum spans between 50−500 m, i.e., medium- to long-span bridges. According to the national bridge inventory (NBI) curated by the Federal Highways Administration (FHWA), about 12,000 bridges (2%) in the US belong to this class. Similarly, the successful study on the short-span Ciampino Bridge (CMP) represents the applicability of the method to US bridges having maximum spans between 15 m and 50 m, which includes nearly 163,000 bridges (26%). Even though our cases were diverse, by span length, they only represent about a quarter of all bridges, and additional work needs to be done to discover the limitations of the applicability of the method

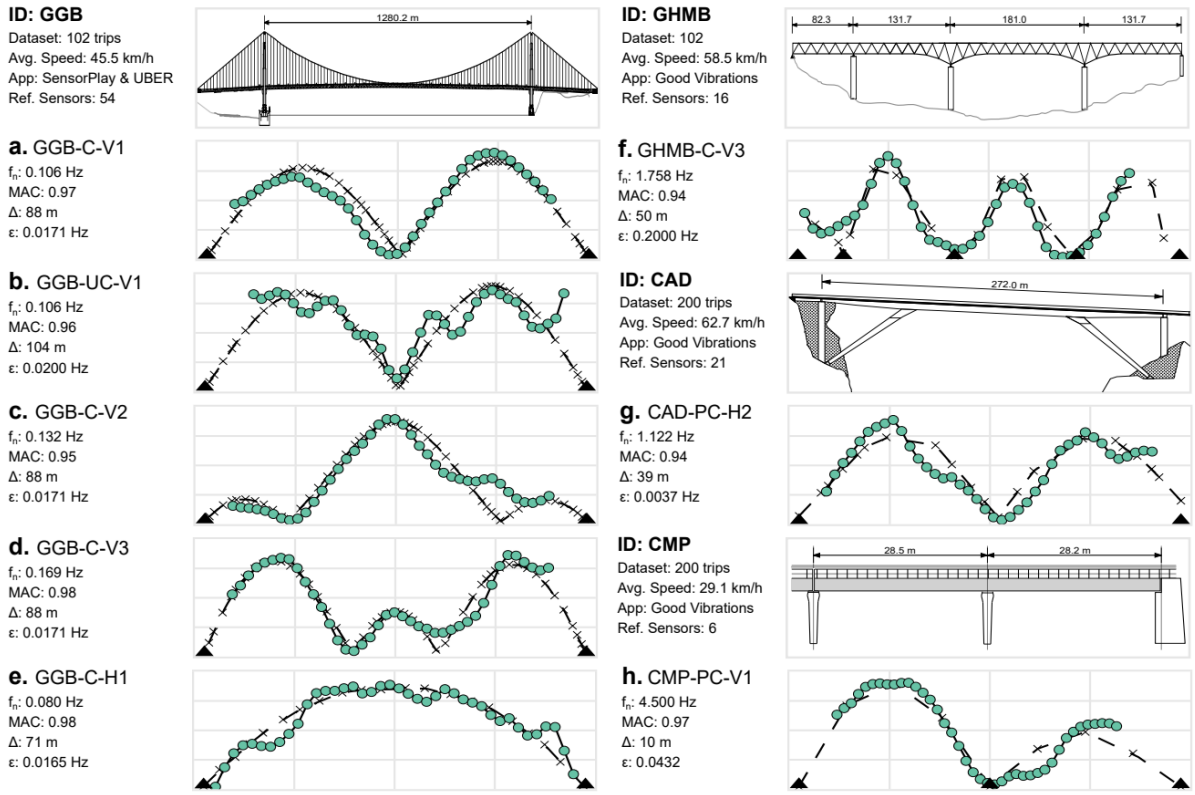

Figure 12. AMS estimates for all case studies: (a-e) Golden Gate Bridge, (f) Gene Hartzell Memorial Bridge, (g) Cadore Bridge, and (h) Ciampino Bridge (add citation).

(*Cronin et al. 2024 Bridging the Gap: commodifying infrastructure spatial dynamics with crowdsourced smartphone data*)

CHAPTER 4

Recommendations

FUTURE DIRECTIONS OF RESEARCH

In this project we demonstrate the efficacy of a RNN-based framework with domain adaptability that enable signal decontamination in a more reliable and practical manner. In addition to that we proposed and validated methodologies to enable crowdsourcing and facilitate infrastructure condition assessment in real time at an unprecedented scale, rate and resolution.

In future, the research team plans to:

- Develop methods to achieve complete modal identification – The current method effectively ascertains absolute mode shapes instead of complete mode shapes. In the future, the team plans to develop a framework that can perform complete bridge modal identification using field data.
- Extend the developed networks to facilitate the development of digital twins for bridge assessment – the proposed artificial neural network-based approaches are extremely effective at input force identification. However, their interpretability and generalizability is still limited by the networks employed. In the future, the team plans to explore alternative deep neural network architectures that can improve on both interpretability and generalizability, which in turn will facilitate enhanced decision-making and uncertainty quantification.
- Utilize Transfer Learning techniques to transfer knowledge gained from a specific bridge to similar ones – the existing neural network approaches that have been devised are case specific and hence, cannot be used for multiple bridges in the absence of extensive data, thus affecting scalability of proposed approaches. The team plans to explore the potential for the use of transfer learning that will ensure the potential for reuse of trained machine learning-based bridge meta-models for a set of "similar" bridges. This will consequently facilitate the development of digital twins for bridge inventories.

References

Eshkevari, S.S., Matarazzo, T. J., and Pakzad, S.N. (2020a). Bridge modal identification using acceleration measurements within moving vehicles. *Mechanical Systems and Signal Processing*, 141, 106733.

Eshkevari, S.S., Cronin, L., Pakzad, S.N., and Matarazzo, T. J.. (2020b). Wavelet platform for crowdsensed modal identification of bridges. arXiv preprint arXiv: 2007.09249.

Yang, Y.B., Zhang, B., Qian, Y., Wu, and Yuntian. (2018) Contact-point response for modal identification of bridges by a moving test vehicle. *Int. J. Struct. Stab. Dyn*., 18, 1850073.

Nix, D.A., and Weigend, A.S. (1994). Estimating the mean and variance of the target probability distribution. *Proceedings of 1994 Iee International Conference of Neural Networks (ICNN'94),* 1, 55-60.

Lakshminarayanan, B., Pritzel, A., and Blundell, C. (2016). Simple and scalable predictive uncertainty estimation using deep ensembles. *arXiv.*

Eshkevari, S.S., Takac, M., T. J., Pakzad, S.N., and Jahani, M. (2021). DynNet: Physics-based neural architecture design for nonlinear structural response modeling and prediction. *Eng. Struct*, 229, 111582.

Matarazzo, T.J., and Pakzad, S.N. (2018). Scalable structural modal identification using dynamic sensor network data stridex. *Computer-Aided Civil and Infrastructure Engineering*, 22, 4-20.

Eshkevari, S.S., Pakzad, S.N., Takac, M.,, and Matarazzo, T. (2020c). Modal Identification of bridges using mobile sensors with sparse vibration data. *Journal of Engineering Mechanics*, 146, 04020011.