

Deep Learning-Based Estimation of the Unknown Road Profile and State Variables for the Vehicle Suspension System

GYUWON KIM¹, SOO YOUNG LEE¹, JONG-SEOK OH², AND SEUNGCHUL LEE^{1,3}

¹Department of Mechanical Engineering, Pohang University of Science and Technology (POSTECH), Pohang 37673, South Korea

²Division of Automotive and Mechanical Engineering, Kongju National University, Cheonan 31080, South Korea

³Graduate School of Artificial Intelligence, Pohang University of Science and Technology (POSTECH), Pohang 37673, South Korea

Corresponding author: Seungchul Lee (seunglee@postech.ac.kr)

This work was supported in part by the National Research Foundation of Korea (NRF) Grant funded by the Korean Government Ministry of Science and Information and Communications Technology (ICT) under Grant 2020R1A2C1009744, in part by the Institute for Information Communications Technology Promotion (IITP) Grant funded by the Korean Government [Ministry of Science, ICT, and Future Planning (MSIP)] [Artificial Intelligence Graduate School Program (POSTECH)] under Grant 2019-0-01906, and in part by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) Grant funded by the Korean Government [Ministry of Trade, Industry, and Energy (MOTIE)] under Grant 20206610100290.

ABSTRACT The vehicle suspension control unit serves as a critical component to the vehicle system, as it ensures the steering stability and sound ride quality of the vehicle. To effectively realize control strategies, it is essential to foreknowledge the road profile and the suspension system's internal state variables. While the mentioned variables are not practically measurable using commercial sensors, it is necessary to estimate the desired variables by utilizing observer systems. Conventional means have mainly employed model-based approaches, in which model uncertainties and high computational cost pose limitations for practical implementation. Herein, we propose a data-driven deep learning method as an alternative because no explicit physical modeling is required, and evaluation is computationally cheap. We first propose a novel encoder-decoder structured recurrent neural network model with a two-phase attention mechanism to estimate the unknown road profile and four state variables of the vehicle suspension system. Based on a simulated data set, we assess the proposed model's qualitative and quantitative results and demonstrate that our model can achieve highly accurate estimation results with fast computation time. Besides, we validate our black-box model's reliability by comparing its interpretation with the suspension system's actual physical characteristics. Furthermore, we compare the proposed model with existing baseline methods, and the results show that our proposed deep learning model significantly outperforms the baseline. Lastly, we experiment with our network's autoregressive capability and demonstrate the feasibility of estimating a *sequence* of future values, which has not been presented in previous works.

INDEX TERMS Deep learning, long short term memory (LSTM), observer systems, recurrent neural networks (RNNs), vehicle suspension control.

I. INTRODUCTION

The suspension system serves as a critical component to the vehicle system, as it ensures the steering stability and sound ride quality of the vehicle. Several means of control strategies have been developed to realize desirable suspension performance [1], [2], of which actuators (e.g., hydraulic actuators [3], magnetorheological dampers [4]) are utilized to adjust tunable system parameters according to different driving

conditions. For this purpose, it is necessary to foreknowledge the road profile and the suspension system's internal state variables. Means of direct measurement are available (e.g., road profile [5], [6], sprung mass displacement [7], unsprung mass displacement [8]), but limitations are apparent as they cannot be practically implemented. Therefore, it is necessary to *estimate* the desired variables using measurable sensor signals by utilizing *observer systems*.

Conventional means have mainly employed model-based approaches, among which the Kalman filter (KF) algorithm is broadly utilized [11]. The KF algorithm is based upon the

The associate editor coordinating the review of this manuscript and approving it for publication was Ah Hwee Tan.

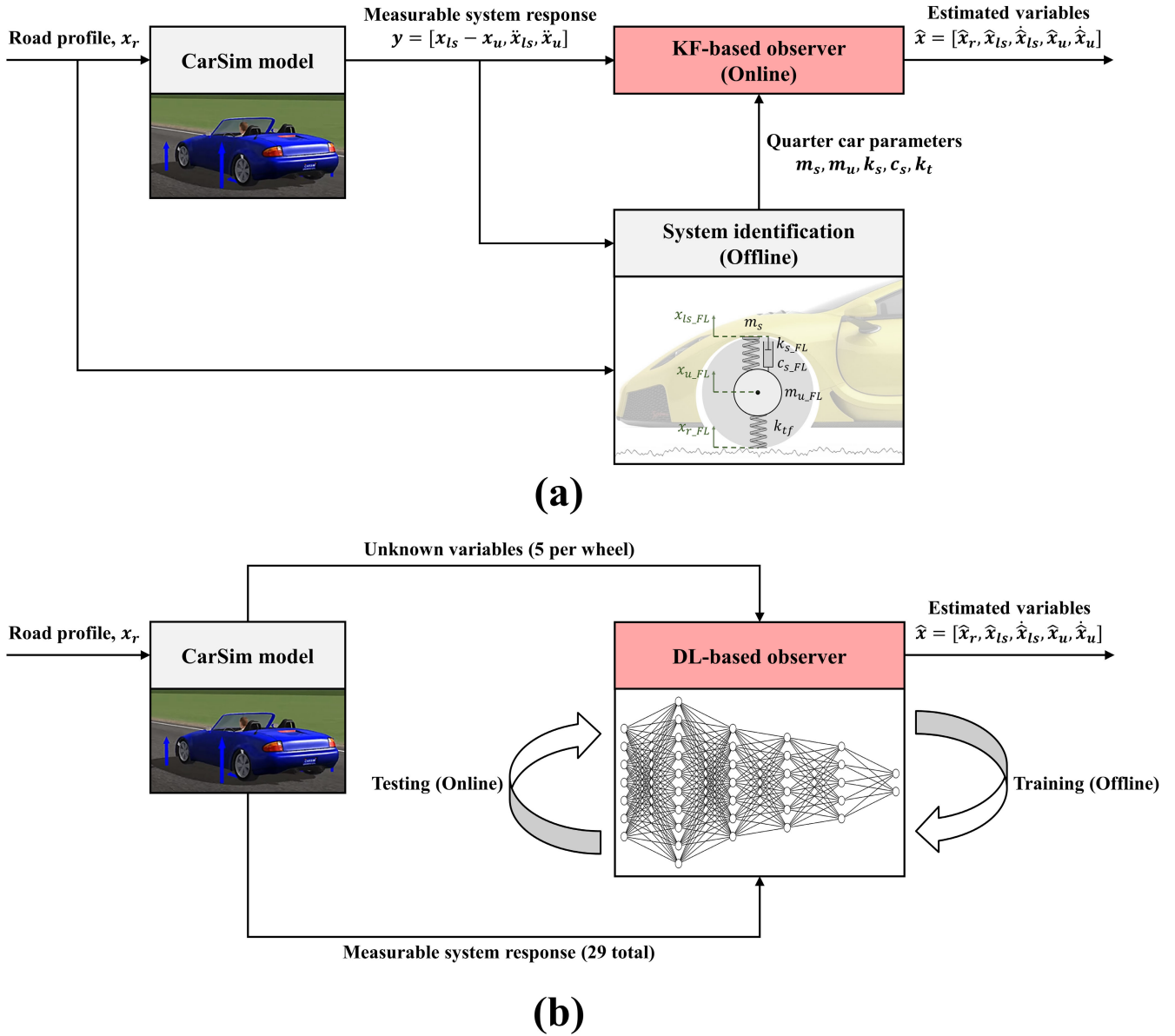


FIGURE 1. Strategies for establishing the observer systems used in this study: (a) KF-based (the EKF-UI [9]), (b) DL-based (our proposed model, the static NARX model [10]).

physical modeling of the vehicle suspension system, of which an offline procedure is required beforehand to identify the parameters of the modeled system. Subsequently, an online procedure is conducted to estimate the unknown variables by utilizing the modeled system and its related measurable sensor signals. A schematic of the above procedures is illustrated in Figure 1 (a). While studies have validated its usage with field test results [9], [12], several setbacks are inherent regarding its working principles. First, model uncertainty is evident as previous studies have mainly utilized the simplified linear quarter car suspension model [13], in which the nonlinearity in the suspension system is considered linear, and the pitch and roll motions of the vehicle body are neglected. Also, identified parameters vary under different driving conditions

(e.g., driving speed), thus hindering the estimation performance. Several adaptation laws have been proposed to overcome these issues [14], [15], but related algorithms are computationally burdensome, limiting its application for real-time implementation in embedded systems.

Recently, data-driven methods have been introduced as an alternative approach to overcome the above problems. As illustrated in Figure 1 (b), data-driven methods differ in that a highly nonlinear mapping from the input to the output data is learned during the offline training process. Several merits are inherent in adopting data-driven methods. First, the model *learns* the system's dynamics directly from the data; thus, the need for explicit physical modeling can be avoided. Secondly, the forward path of evaluating a pre-trained deep

neural network consists solely of basic operations and non-linear activation functions. Hence, it is computationally much cheaper than the model-based Kalman filter algorithm.

Previous data-driven approaches have focused solely on estimating the road profile. Studies have been conducted to classify different road profiles using machine learning algorithms [16], [17]. Although validated with field test results, classification differs from the current research scope, as profile estimation poses a broader range of potential applications. Researches regarding profile estimation have mainly utilized nonlinear autoregressive exogenous (NARX) models using simple multi-layered perceptron (MLP) structures. Related studies are sub-categorized into static and dynamic NARX models, in which dynamic models are considered impractical as they require the actual values of the past road profile for the estimation of the current values [18], [19]. A study in [20] has utilized a fuzzy model along with a static NARX model to successfully estimate the road profile using simulated data, while a study in [10] validated the performance of a static NARX model with field test results.

Although the studies above have demonstrated the effectiveness of data-driven methods in road profile estimation, the following aspects require further investigation, which has motivated the current study.

- First, the estimation of the internal state variables is also desirable for control system applications. Since previous data-driven approaches have been limited to estimating the road profile, the feasibility of system state estimation should be investigated.
- Secondly, the utilization of a more sophisticated network structure should be considered. Although recurrent-structured networks are best suitable for dynamic data analysis, previous studies have only utilized simple MLP models. This poses clear limitations as it assumes that the data-driven inverse modeling of the suspension system can be conducted by static means.
- Lastly, input sensor selection should be conducted to determine the set of in-vehicle sensors best suited for the current application. While previous data-driven studies have mainly utilized acceleration signals, no valid basis has been provided for its sufficiency/redundancy. Input optimization should be considered as less required sensors imply a more cost-efficient and robust system.

This paper suggests a deep learning (DL)-based observer system with high accuracy and fast computation time to overcome the challenges described above. All data used in this study are acquired via simulation in the CarSim environment [21]. We first propose a novel encoder-decoder structured recurrent neural network (RNN) model with a two-phase attention mechanism to better characterize the suspension system's dynamic behavior. The model is trained to estimate the unknown road profile and four system state variables by utilizing measurable in-vehicle sensor signals. We search for a promising set of hyper-parameters via Bayesian optimization and then optimize the input sensors by observing the

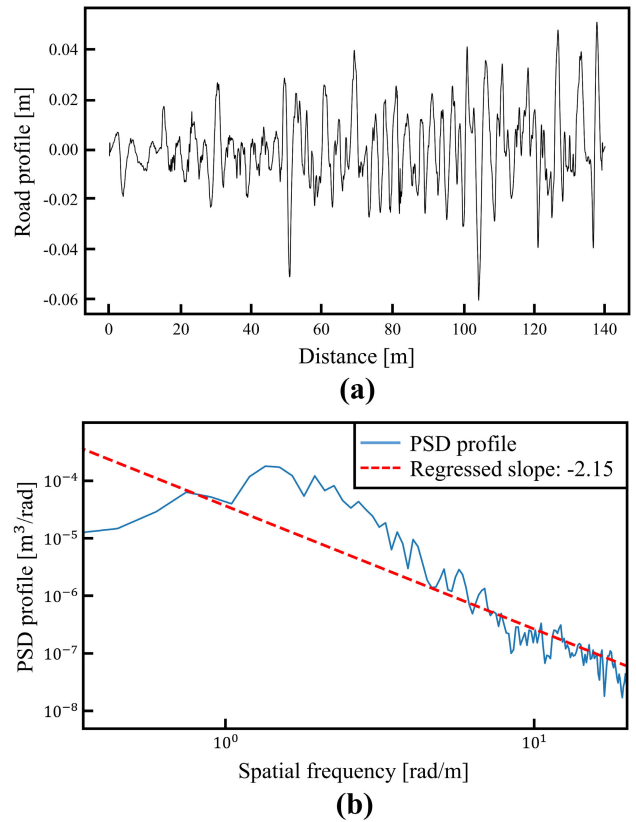


FIGURE 2. A random F class road profile designed as the road input: (a) road profile in the time domain, (b) its corresponding PSD profile.

attention values related to our proposed model's attention mechanism. Furthermore, we validate that the interpretation of the attention values is in-line with the suspension system's physical characteristics, thus providing a valid basis for model reliability. We evaluate qualitative and quantitative results with the extended Kalman filter with unknown inputs (EKF-UI) [9] and the static NARX model [10] to examine the estimation capability of our proposed method in comparison with the existing studies. Lastly, we demonstrate the feasibility of sequence estimation, which estimates a *sequence* of future values at each time step, and has not been proceeded by previous works.

II. MATERIALS AND METHODS

A. DATA ACQUISITION

All data used in this study are acquired via simulation in the CarSim environment [21], as it is reported to produce highly relevant results compared with actual field test experiments [9], [22], [23]. We first design a random F class road profile according to the ISO 8608 standard [24] for the input to the CarSim vehicle model. The produced road and its corresponding power spectral density (PSD) profile are presented in Figure 2 (a) and (b), respectively. In particular, the PSD profile's regressed slope is equal to -2.15 , which is close to the value -2 of the road roughness classified by

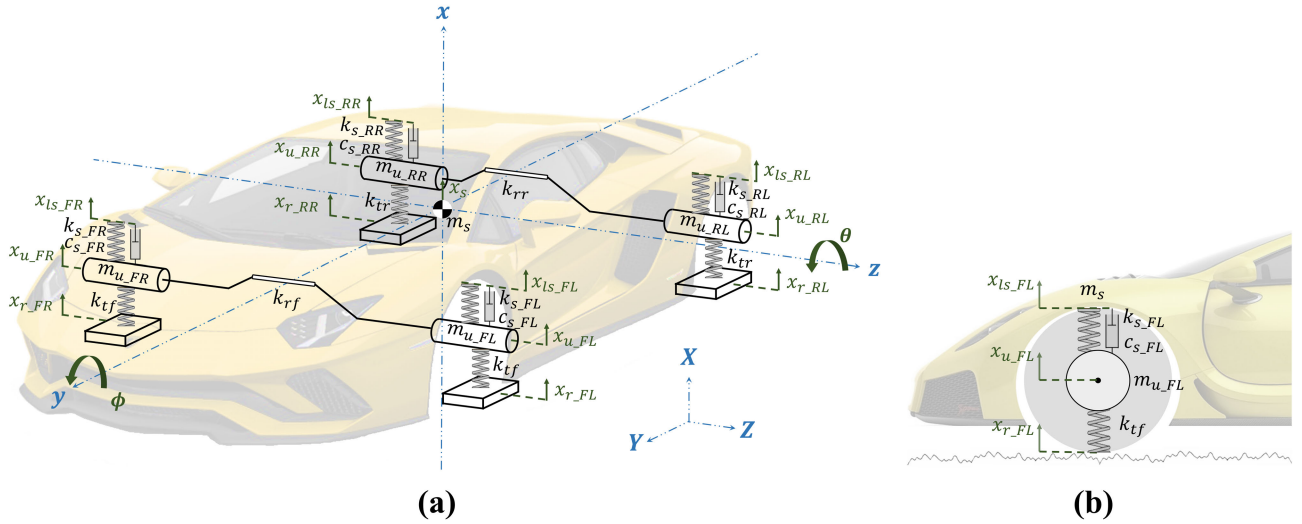


FIGURE 3. Two representative vehicle suspension models: (a) 7-DOF full car model, (b) 2-DOF quarter car model.

TABLE 1. Specifications of the input (measurable sensor signals) and output (unknown variables) data acquired via simulation in the CarSim environment.

Input data [Units]	Quantity
Local relative displacement [m]	4
Local sprung mass acceleration [m/s ²]	4
Unsprung mass acceleration [m/s ²]	4
Pitch & pitch rate [deg, deg/s]	2
Roll & roll rate [deg, deg/s]	2
Yaw & yaw rate [deg, deg/s]	2
Longitudinal acceleration [m/s ²]	5
Lateral acceleration [m/s ²]	5
Steering angle [deg]	1
Output data [Units]	Quantity
Road profile [m]	4
Local sprung mass displacement [m]	4
Local sprung mass velocity [m/s]	4
Unsprung mass displacement [m]	4
Unsprung mass velocity [m/s]	4

the international roughness index (IRI) [2]. A D-class sedan is used as the vehicle model, and virtual sensor signals are acquired with a sampling frequency of 1 kHz under 30 KPH driving speed.

We consider all measurable system response of the 7-DOF full car suspension model adopted in [25] (Figure 3 (a)). The model consists of a sprung mass (m_s), four unsprung masses (m_u), four pairs of suspension coil springs and dampers (k_s and c_s), two anti-roll bars (k_r), and four tires modeled as spring elements (k_t). FL, FR, RL, and RR denotes forward-left, forward-right, rear-left, and rear-right, respectively. The sprung mass includes the vehicle chassis and body, while the unsprung mass includes the wheels, brakes, and calipers. Fixture locations of the suspension elements are designated as the local sprung mass locations, with its vertical displacements denoted as x_{ls} . Vertical displacements of the sprung mass, unsprung mass, and the road profile are denoted as x_s , x_u , and x_r , respectively. The sprung mass's angular motions include the pitch (θ) and the roll (ϕ).

Additional measurable sensor signals such as the yaw, yaw rate, longitudinal accelerations, lateral accelerations, and the steering angle are also considered. The pitch, roll, yaw, and its corresponding rates are measured from the sprung mass center of gravity, while the longitudinal and lateral accelerations are measured from the sprung mass center of gravity and each wheel center. In total, 29 measurable sensor signals are used as the input data to train the neural networks to estimate the unknown road profile and four system state variables per wheel (FL, FR, RL, RR), as illustrated in Figure 1 (b). All acquired sensor signals are split into training, validation, and test data to establish the observer models. Specifications of the input and output data are listed in Table 1.

It is worth mentioning that while the majority of the model-based approaches are limited to utilizing the simplified linear quarter car suspension model [13] (Figure 3 (b)), our proposed method can fully utilize all measurable in-vehicle sensor signals. Although the linear quarter car suspension model offers the advantage of low computational cost in KF algorithms, it incorporates model uncertainties as it only considers the vehicle's heave motion. This study validates that our proposed method outperforms the quarter car model-based EKF-UI algorithm [9] in estimation accuracy and computation time. Further details are provided in Section III-B.

B. PROPOSED MODEL

This study proposes a novel encoder-decoder structured deep recurrent neural network model with a two-phase attention mechanism to estimate the unknown road profile and four state variables of the vehicle suspension system. The architecture of the proposed model is illustrated in Figure 4. Our proposed model consists of three constituent parts, i.e., the encoder network, the two-phase attention layer, and the decoder network. Our goal is to train all sub-networks in an

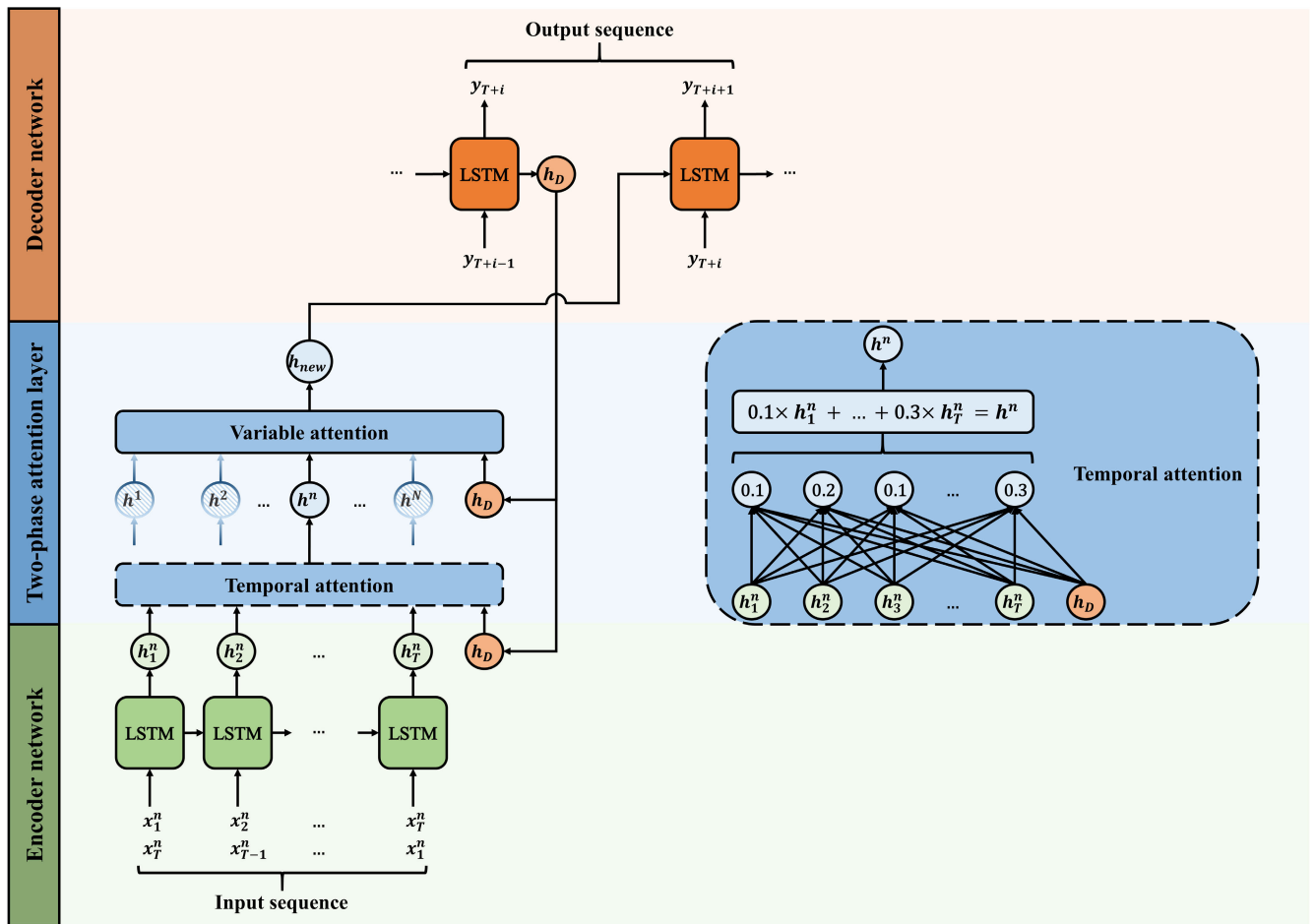


FIGURE 4. Our proposed network architecture consisted of the encoder network, the two-phase attention layer, and the decoder network. A detailed description of the temporal attention mechanism is illustrated.

end-to-end manner to establish a DL framework for precisely estimating the desired variables.

The encoder network is designed to extract high-level features from sequences of input signals. Unlike static NARX models, our network utilizes an RNN structure, which takes sequential data as the input to better analyze its dynamic characteristics. Among existing RNN modules, the long short term memory (LSTM) is adopted, as it has been proposed to overcome the gradient vanishing problem for long-term dependencies [26]. Besides, bidirectional encoding is utilized to enhance performance [27]. The entire encoder network consists of N parallel LSTM networks according to the number of input signals. x^n_t denotes the input sequence x at time t associated with the n^{th} input signal, and the same annotation applies to the hidden state h . T denotes the input sequence length.

Subsequently, the encoded hidden states are passed on to the two-phase attention layer. The key part of this paper is using the two-phase attention mechanism to enhance estimation accuracy and attain knowledge of the deep neural network's working mechanism. While conventional

encoder-decoder-structured RNNs suffer from the bottleneck issue [28], the attention mechanism has been proposed to alleviate the problem by exclusively focusing on relevant positions within the input sequence [29]. We propose to apply the attention mechanism in a two-phase manner to avoid the above problem and enhance the estimation accuracy. First, the attention mechanism is applied to the temporal-wise hidden states of each LSTM network. As illustrated, the encoded hidden states are fed forward in a fully-connected layer composed of weights, biases, and the hyperbolic tangent function ($\tanh$) as the nonlinear activation function. The softmax function is applied to the resulting values, and its output indicates the *temporal attention values*, as values associated with important positions within the sequence will be trained to be larger than the unimportant ones. The temporal attention values are linearly combined with the sequence of hidden states to obtain the summarized hidden state, h^n , for each input variable. Subsequently, variable-wise attention is applied in the same manner to the summarized hidden states to produce a single encoded hidden state, h_{new} , to pass on to the decoder network. Similarly, softmax values associated

with the variable-wise attention mechanism indicate the *variable attention values*, as values associated with the important input variables will be trained to be larger than that of the unimportant ones. It is worth mentioning that the interpretation of the deep neural network can be achieved by examining the attention values. We will utilize the attention values to optimize the relevant inputs and provide interpretations of our network's mechanism in Section III-A.

The decoder network then transforms the encoded information into the values of interest. Note that a sequence of future values can be estimated at each time step by utilizing the decoder network in an autoregressive manner [30] as illustrated. This concept of *sequence estimation* has not been proceeded by previous works. We demonstrate the feasibility of sequence estimation in Section III-C, as it poses potentials for a broad range of applications in predictive control strategies.

C. MODEL TRAINING AND EVALUATION

Several methods regarding the training strategies are employed to train our network and optimize its trainable parameters. As for the loss function L , the mean squared error (MSE) is used and is defined as follows:

$$L = \frac{\sum (y - \hat{y})^2}{N} \quad (1)$$

where y and $\hat{y}$ denotes the actual and the estimated value, respectively, and N denotes the number of training data. Trainable parameters are optimized using the Adam optimizer [31] to minimize the loss function. Besides, the step decay scheduler [32] is adopted for faster convergence. We apply Xavier initialization [33] to initialize the network weights, as most nonlinear activation functions utilize the hyperbolic tangent function. A dropout method regarding recurrent layers [34] is employed to prevent overfitting. All model training is conducted on an NVIDIA TITAN Xp GPU.

The mean absolute percentage error (MAPE) for the PSD profile is considered for the evaluation metric and is defined as follows [9]:

$$e_{rms} = \frac{1}{N} \sum \frac{|Y - \hat{Y}|}{Y_{rms}} \times 100 \quad (\%) \quad (2)$$

where Y and $\hat{Y}$ denotes the actual and the estimated PSD profile, respectively, Y_{rms} denotes the root mean squared (RMS) value of the actual PSD profile, and N denotes the number of evaluation data. All model evaluation is conducted on an Intel®Core™ i5-8400 x64 CPU 16.0GB RAM CPU.

D. MODEL OPTIMIZATION

The Bayesian optimization strategy [35] is first conducted to optimize the hyper-parameters regarding the structure and the training configuration of the proposed model. Unlike existing systematic approaches (i.e., grid search and random search [36]), Bayesian optimization considers prior knowledge of the training process and reduces the computational

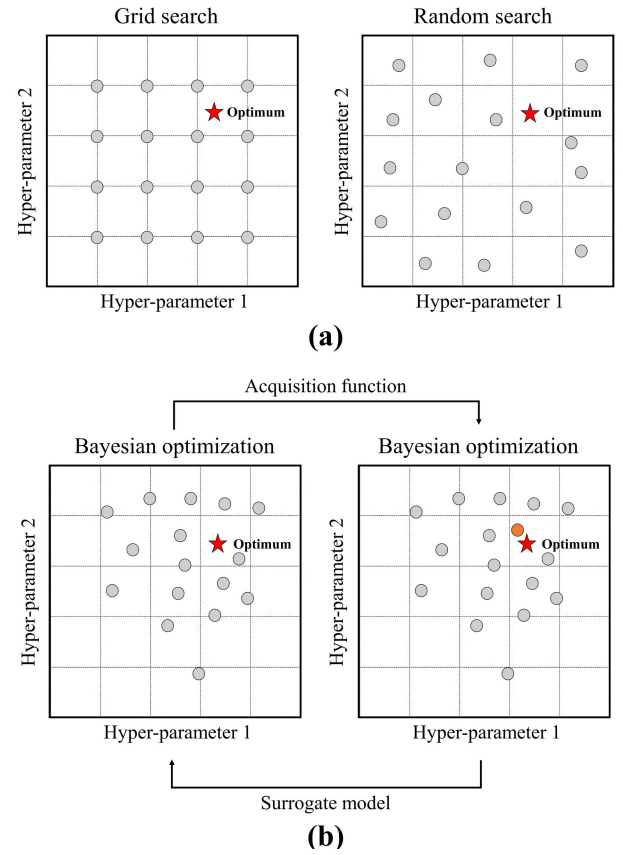


FIGURE 5. A schematic description of different hyper-parameter optimization strategies: (a) grid search and random search, (b) the Bayesian optimization strategy.

cost. A schematic of the Bayesian optimization procedure compared with grid search and random search is illustrated in Figure 5. Bayesian optimization is conducted based on the simultaneous interaction of the surrogate model and the acquisition function. The surrogate model probabilistically models the unknown objective function $f(\cdot)$ that maps the set of hyper-parameters λ_i to the evaluation result, by considering the prior evaluated samples $(\lambda_1, f(\lambda_1)), \dots, (\lambda_N, f(\lambda_N))$. The acquisition function then suggests a new set of hyper-parameters λ_{N+1} based on the exploration-exploitation strategies regarding the established surrogate model. This procedure can be expressed as follows:

$$\lambda^* = \operatorname{argmin}_{\lambda \in \Lambda} f(\lambda) \quad (3)$$

where Λ denotes the entire hyper-parameter search space, and λ^* denotes the optimal set of hyper-parameters derived from the optimization process. The tree Parzen estimator (TPE) is utilized as the surrogate model [37], and the expected improvement (EI) algorithm is utilized as the acquisition function [38].

Subsequently, input optimization is conducted by utilizing the attention values, as mentioned in Section II-D. Temporal-variable attention values are utilized to interpret which input variables, and more specifically, which temporal indices

within each variable contributed the most during the network training. The temporal-variable attention values are obtained by multiplying each variable attention value to its corresponding temporal attention values across different time indices. Relevant input sensors are selected to retrain the network by observing the obtained attention values, and we further validate its interpretation by referencing the suspension system's actual physical characteristics. Detailed discussions are presented in Section III-A.

III. RESULTS

A. MODEL OPTIMIZATION

Based on the searching spaces of the hyper-parameters defined in Table 2, the Bayesian optimization is conducted to explore the most promising set of hyper-parameters for our proposed model. The FL estimation model is used for the hyper-parameter optimization process, and its resulting hyper-parameters are applied to all wheel models during the training process. We use quantized uniform distribution as the searching space of each hyper-parameter, except for the learning rate, the epoch scheduler rate, and the dropout rate, in which log-uniform distribution is used. Besides, the combined validation MSE of the five estimation variables is used for the optimization metric. As previously discussed, Bayesian optimization utilizes prior knowledge of the optimization process and thereby effectively searches for new sets of hyper-parameters. This is observed in Figure 6, as the validation loss gradually decreases over the optimization iteration, where early stages in the optimization process show high validation loss with high variance and the latter show low validation loss that seems relatively convergent. The optimal hyper-parameter set with the lowest validation loss is found in the 80th optimization iteration, and its corresponding values are listed in Table 2.

TABLE 2. The optimal hyper-parameter set regarding the model structure, training, and regularization, obtained via 100 iterations of Bayesian optimization.

Hyper-parameter	Optimization type	Search space	Optimal parameter
Input sequence length	Model structure	[5, 50]	19
Hidden dimension	Model structure	[32, 128]	64
Batch size	Model training	[32, 128]	33
Learning rate	Model training	[1e-4, 1e-2]	0.004
Epoch scheduler step	Model training	[5, 20]	9
Epoch scheduler rate	Model training	[0.6, 0.9]	0.63
Dropout rate	Regularization	[0.1, 0.7]	0.38

An important hyper-parameter worth mentioning is the input sequence length. Having long input sequences incorporates the bottle-neck problem to the encoder-decoder model as it burdens the encoder to compress all necessary information into a fixed-dimensional tensor [28]. The use of the attention mechanism can relieve this issue, but it is not a complete solution. Therefore, a trade-off is requisite to find an optimal input sequence length that is long enough to contain sufficient information, but short enough to be effectively encoded. The input sequence length of 19 is found to be the optimal value via Bayesian optimization.

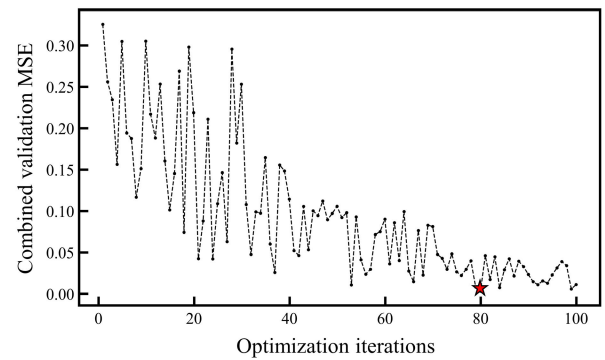


FIGURE 6. Progress plot illustrating the hyper-parameter optimization process over 100 iterations of Bayesian optimization. The red star denotes the lowest validation loss obtained during the optimization process.

Subsequently, we conduct input sensor optimization by observing the attention values of our hyper-parameter optimized models. From its observations, insights into our DL model's working mechanism can be achieved and are presented as follows. A heatmap of the FL estimation model's temporal-variable attention values is presented in Figure 7 (a). Regarding its temporal characteristics, it is observed that our estimation model focuses mostly on the current input values. This is similar to the working principles of the KF algorithm, which utilizes the *a priori* values to estimate the *a posteriori* values. As for its variable-wise importance, the pitch signal is shown to contribute the most for the network training, followed by the FL relative displacement and unsprung mass acceleration. The RL relative displacement and unsprung mass acceleration appear to have notable contributions also. Other signals seem of less importance for the training of the FL estimation model. The above observations are reasonable as signals from the left side of the vehicle are shown to contribute the most to establishing the FL estimation model, while the pitch signal is of general importance.

Likewise, variable attention values of all wheel models have been combined, and the following observations are made regarding Figure 7 (b). First, the pitch and the pitch rate are considered important, implying that utilizing the full car suspension model better represents the actual suspension system's underlying physics than the linear quarter car suspension model. In contrast to the pitch and the pitch rate, the roll and the roll rate seem of little importance. While longitudinal dimensions of the vehicle are longer than the lateral, pitch motion is usually more dominant than the roll motion, except for extreme driving conditions such as off-road driving and sharp cornering. Simulation procedures do not include any of the above, as the vehicle simulation is conducted on a straight road without a change of directions. Besides, the effect of the anti-roll bars also contributes to the dampening of the roll motion.

Secondly, it is observed that the sprung mass acceleration does not seem to contribute much to the training of our model, as its combined variable attention values indicate no significant importance. Given the local relative displacement and

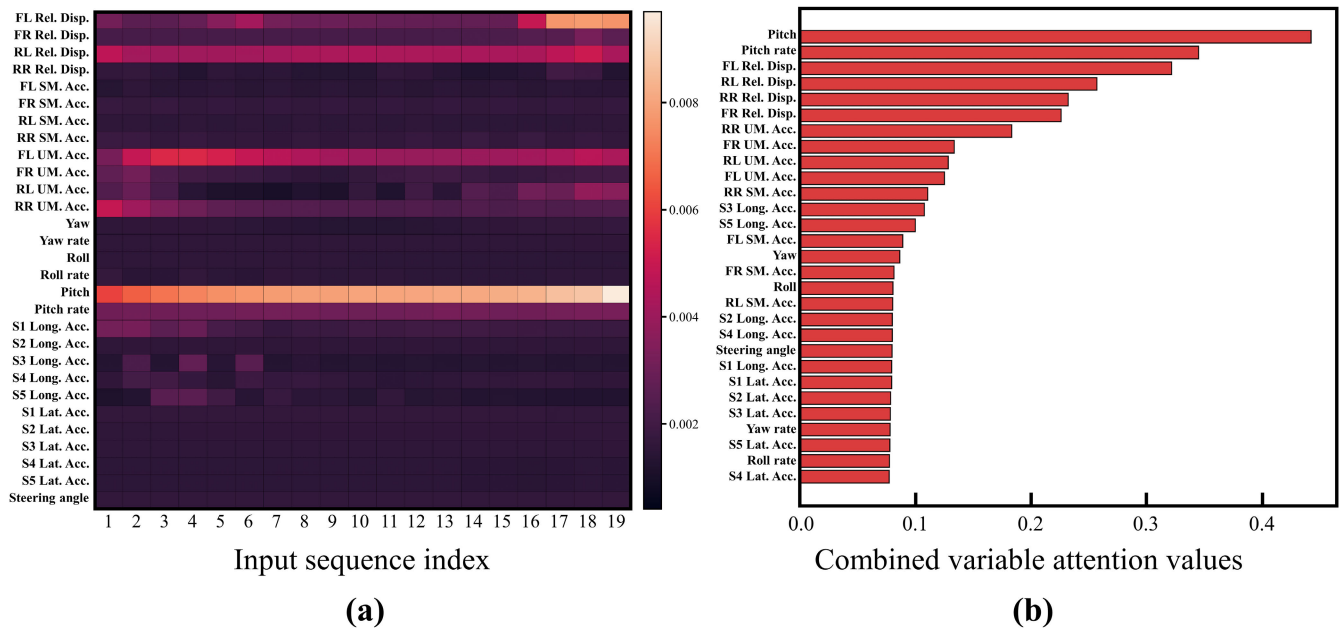


FIGURE 7. Our proposed model's attention values: (a) heatmap of the FL estimation model's temporal-variable attention values, (b) combined variable attention values of all wheel models.

the unsprung mass acceleration, the sprung mass acceleration can be regarded as redundant if the DL model can learn the relationship between the three variables. The unsprung mass acceleration is preferred over the sprung mass acceleration because wheels are closer to the road than the vehicle body. Therefore, the unsprung mass acceleration has higher correlations with the road profile than the sprung mass acceleration.

Similar observations are made regarding the longitudinal and lateral acceleration signals, as combined variable attention values show little importance. Although the above variables are considered to analyze the 6-DOF motion of the vehicle (including 3-DOF translational and 3-DOF rotational), it is known that its effects on the vertical motion of the sprung mass are negligible [39]. Other variables unrelated to the full car suspension model (i.e., yaw, yaw rate, steering angle) are also considered unimportant, which implies that our deep learning model provides physically relevant interpretation results.

We conduct input sensor optimization based on the above observations. The pitch, pitch rate, local relative displacement, and the unsprung mass acceleration signals (the top ten variables with the highest combined attention values) are chosen as the input sensors. The FL estimation model is retrained using the selected sensor signals, and its performance is evaluated in the following section.

B. MODEL EVALUATION

In this section, we proceed to evaluate our optimized network and compare its performance with baseline models in terms of estimation accuracy and computation time. Two comparative models are adopted in our study: the model-based EKF-UI

[9] and the MLP-based static NARX model [10]. Two hyperparameters are inherent for the static NARX model (the batch size and the learning rate), and its values are selected to be 128 and 0.001 via the grid search strategy. Our hyperparameter and input optimized model will be denoted as DL optimized for the remainder of the paper. We discuss the qualitative difference in terms of the established models' performance by comparing the actual and the estimated profiles. Besides, the estimation results are quantitatively assessed based on the evaluation metric in Section II-C.

Figure 8 and Table 3 presents the qualitative and quantitative results of the estimated profiles of different observer models. Figure 8 clearly illustrates that our DL optimized best follows the actual profile (denoted as CarSim) among the comparative models. While it is known that the EKF-UI's performance deteriorates under rough road conditions [9], it is observed that the EKF-UI produces severely noisy profiles as we utilize an F-class road input. Besides, the static NARX model shows the poorest performance, indicating that its shallow network structure is insufficient to capture the high nonlinearity of the dynamic suspension system.

It is also worth noting that our proposed model's performance improves after optimizing the inputs, as presented in Table 3. This is against intuition since it can be argued that more input data to reference from will result in better estimation performance. While 29 input variables are used in our initial training, the vast majority are considered unimportant, and the irrelevant variables are assigned with small variable attention values (Figure 7 (b)). It can be inferred that this hinders the training by causing *dispersion* of attention, as the network's ability to focus on relevant variables is dispersed to unimportant variables that outnumber. From this observation,

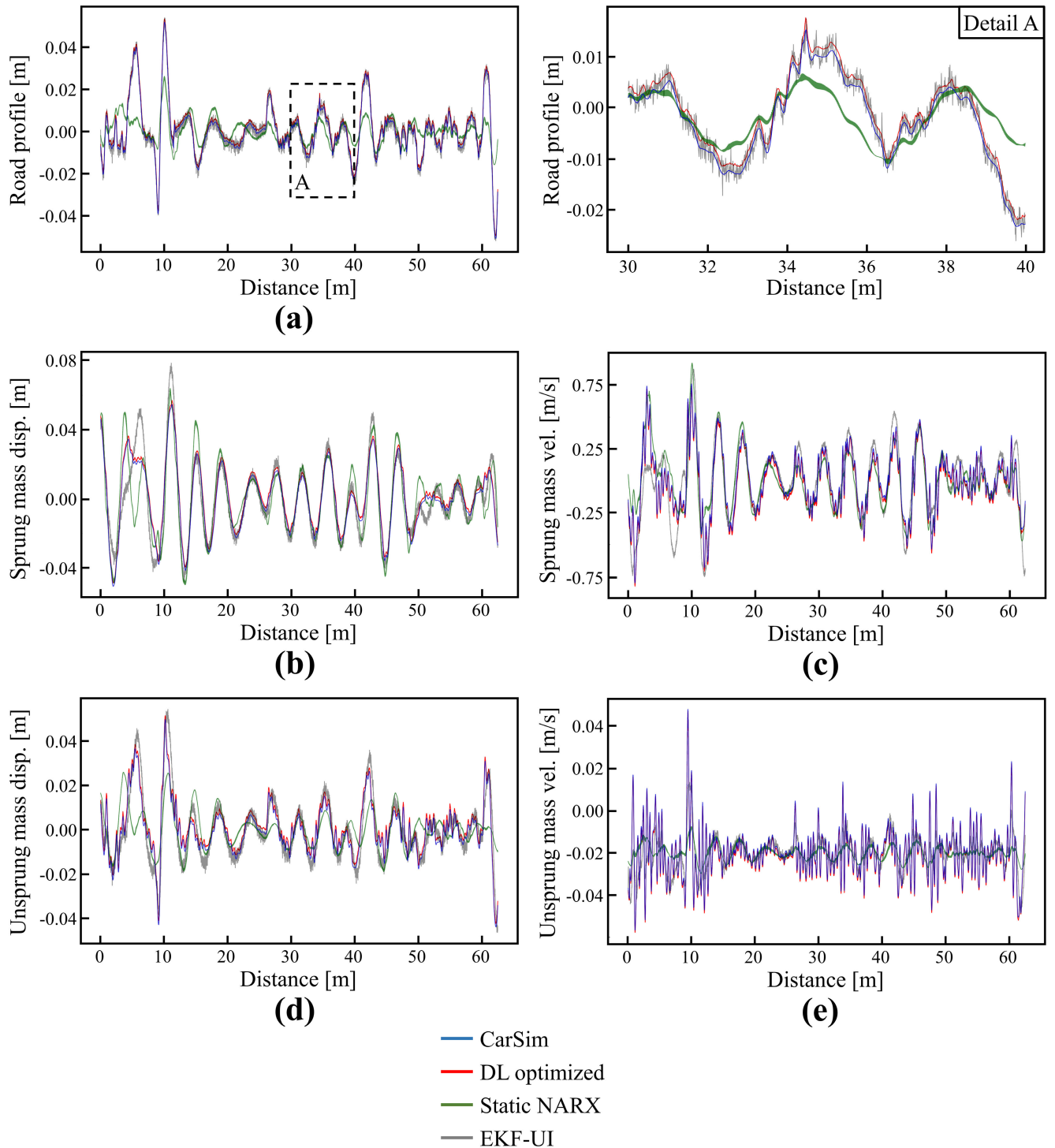
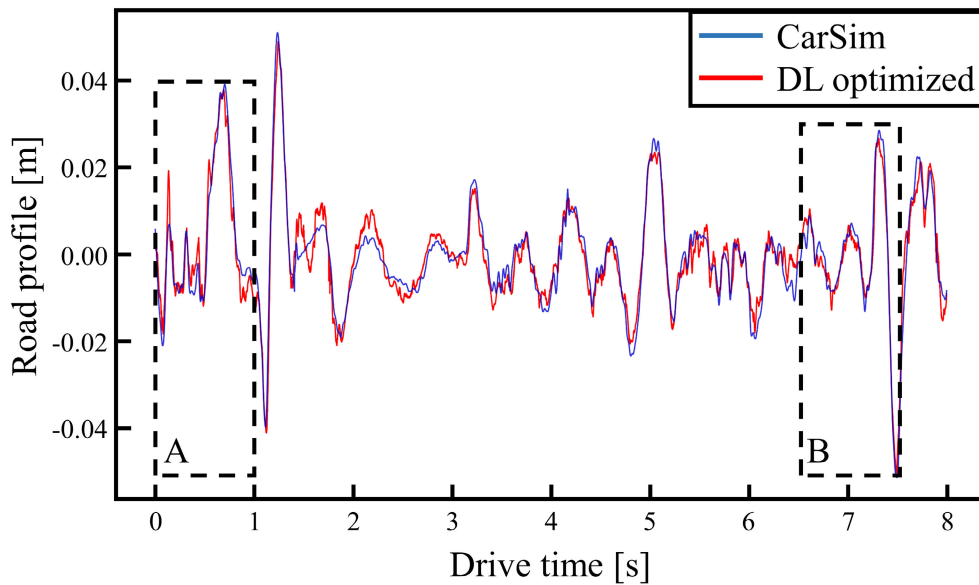


FIGURE 8. Qualitative results of the estimation performance of different observer models: (a) the road profile, with section A in zoomed details, (b) the sprung mass displacement, (c) the sprung mass velocity, (d) the unsprung mass displacement, (e) the unsprung mass velocity.

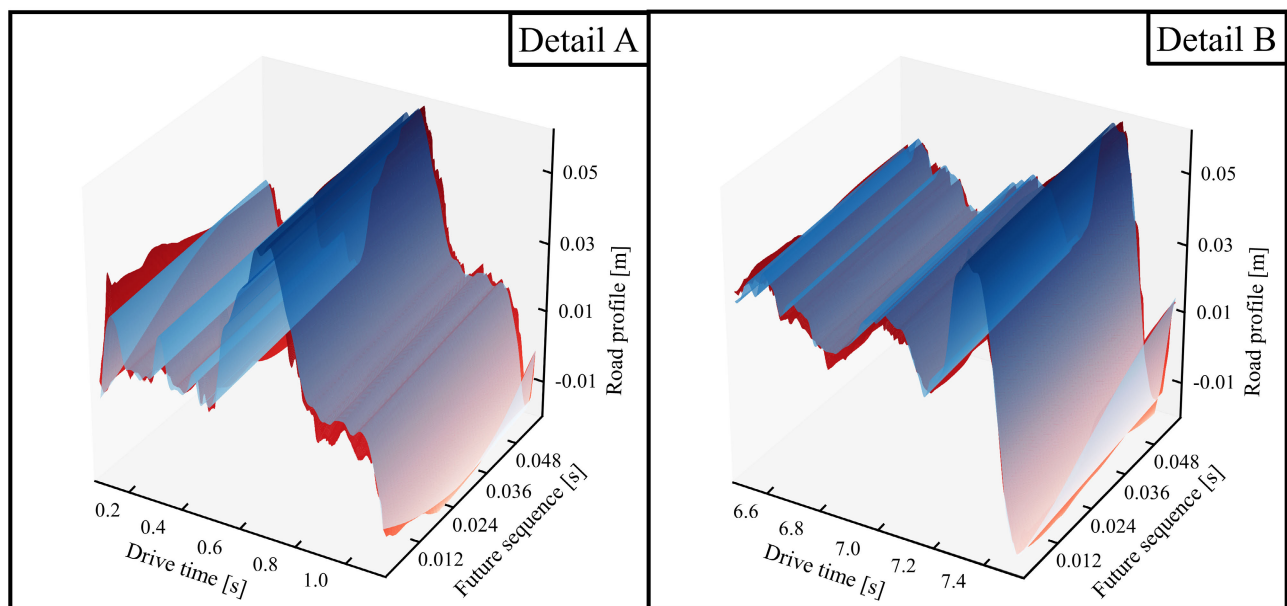
we emphasize that a careful selection of relevant input sensors is necessary when taking a data-driven approach to modeling dynamic systems.

The computation time is also compared and evaluated on the test data set consisted of 7996 sampling time steps.

+ The static NARX model shows the fastest computation time since it is a shallow network model with very few parameters compared to our DL optimized. Although it offers the fastest computation time, it shows the worst estimation performance. Our DL optimized shows the second-best performance in



(a)



(b)

FIGURE 9. Qualitative results of the sequence estimation performance of the road profile, estimating a future sequence of 50 sampling time steps (0.05 sec): (a) initial estimation at $(t + 0 \times dt)$, (b) sequence of estimated profiles at each time step $(t + n \times dt)$, corresponding to the zoomed-in section of A and B.

computation speed, approximately eight times faster than the EKF-UI, while posing the best estimation performance for all variables.

We further experiment with additional data simulated under 50 KPH driving conditions to demonstrate our proposed method's feasibility under different driving conditions. The same set of optimal hyper-parameters and input

sensors are utilized to retrain our proposed model. Evaluation results are compared with the EKF-UI and the static NARX model, as presented in Table 4. Our DL optimized shows slightly improved performance metrics compared to that of the 30 KPH data, except for the sprung mass velocity, which remained the same. On the contrary, an overall deterioration is observed for the EKF-UI and the static NARX model. While

TABLE 3. Quantitative results of the estimation performance of different observer models. All metrics are given in MAPE (%) for the PSD profile (except for the computation time).

	DL 29 inputs	DL optimized	Static NARX [10]	EKF-UI [9]
Road profile [m]	0.10	0.06	4.66	0.30
Local sprung mass displacement [m]	0.03	0.02	2.87	1.53
Local sprung mass velocity [m/s]	0.09	0.06	1.40	2.09
Unsprung mass displacement [m]	0.10	0.08	5.06	5.93
Unsprung mass velocity [m/s]	0.22	0.18	12.90	13.77
Computation time [s]	2.44	2.01	1.03	15.98

TABLE 4. Quantitative results of the estimation performance of different observer models under 50 KPH driving conditions. All metrics are given in MAPE (%) for the PSD profile (except for the computation time).

	DL optimized	Static NARX [10]	EKF-UI [9]
Road profile [m]	0.04	7.56	4.93
Local sprung mass displacement [m]	0.01	3.51	1.30
Local sprung mass velocity [m/s]	0.06	7.57	6.54
Unsprung mass displacement [m]	0.05	9.56	8.38
Unsprung mass velocity [m/s]	0.06	14.71	13.19
Computation time [s]	2.19	1.02	16.03

TABLE 5. Quantitative results of the sequence estimation performance of all variables, estimating a future sequence of 50 sampling time steps (0.05 sec). All metrics are given in averaged MAPE (%) for the PSD profile (except for the computation time).

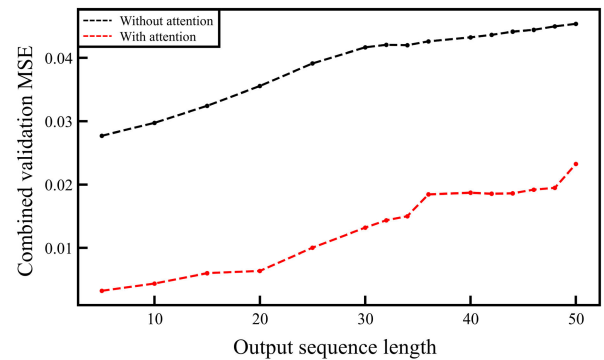
	DL optimized
Road profile [m]	0.62
Local sprung mass displacement [m]	0.27
Local sprung mass velocity [m/s]	0.38
Unsprung mass displacement [m]	0.92
Unsprung mass velocity [m/s]	4.31
Computation time [s]	10.1

the EKF-UI utilizes the linear quarter car suspension model, its accuracy deteriorates at high-speed driving conditions due to the pitch and roll motion effect. Also, it is known that identified quarter car parameters, especially the sprung mass and the tire stiffness, vary under different driving conditions and deteriorates the estimation performance [9]. As for the static NARX model, results indicate that its performance suffers from the lack of network capacity to capture the high nonlinearities incorporated at higher driving speeds. On the contrary, we validate that our DL optimized can accurately conduct the desired action even at high-speed driving conditions with relatively fast computation time.

C. SEQUENCE ESTIMATION

As mentioned in Section II-B, we proceed to further extend our decoder model's architecture to an autoregressive structure to test the feasibility of sequence estimation for all variables. Sequence estimation implies estimating a *sequence* of future values at each time step. It is worth mentioning that sequence estimation has not been conducted in previous studies, as existing model-based and MLP-based approaches are not capable of its demonstration.

Figure 9 presents the sequence estimation results of the road profile, estimating a future sequence of 50 sampling time steps (0.05 sec). Figure 9 (a) presents the initial road profile estimated at $(t + 0 \times dt)$, with dt denoting the sampling time step. Figure 9 (b) illustrates the sequence of estimated profiles at each time step corresponding to the zoomed-in section of A and B. Table 5 presents the averaged performance index

**FIGURE 10.** Sequence estimation performance of the models with and without the two-phase attention mechanism for various output sequence lengths.

(MAPE for the PSD profile) for all estimated profiles at $(t + n \times dt)$, with n ranging from zero to 50. Based on the qualitative and quantitative results, it is observed that our model is capable of sequence autoregression for all variables, as all quantitative metrics except the road profile show better performance than the one-point estimation result of the EKF-UI. Regarding computation time, our model still tops the EKF-UI, even when having the disadvantage of calculating an additional sequence of 49 data points per estimation. Subsequently, we compare the models' estimation results with and without the two-phase attention mechanism to investigate its impact on the sequence estimation performance. While estimation results deteriorate at longer output sequence lengths [28], we validate that utilizing the two-phase attention mechanism can enhance performance even when estimating long output sequences. The baseline model is an encoder-decoder structured LSTM network without attention mechanism, as presented in [30]. Its hyper-parameters include the batch size, the learning rate, and the hidden dimension size. Bayesian optimization is conducted as previously presented, and the hyper-parameter values of 38, 0.0001, and 46 are chosen as the promising values. Figure 10 presents the combined validation MSE values of the two models for various output sequence lengths. It can be observed that our proposed model with the two-phase attention mechanism significantly outperforms the baseline network, as the combined MSE value of output sequence length of 50 is lower than that of the one-point estimation MSE of the baseline network.

IV. CONCLUSION

This study proposed a deep learning approach to accurately estimate the unknown road profile and four state variables of the vehicle suspension system with fast computation time. The main contributions of the work are summarized as follows.

- First, we proposed a novel encoder-decoder structured RNN model with a two-phase attention mechanism and optimized its hyper-parameters and input sensors. In particular, we analyzed that a careful selection of relevant input sensors is necessary when utilizing

data-driven methods to model dynamic systems. Results have shown that our optimized model trained with ten inputs outperforms the pre-optimized model trained with all 29 inputs, indicating that more input sensors do not necessarily result in better performance.

- Secondly, we verified that our proposed DL model could improve both qualitative and quantitative results. While we compared the estimation performance with two baseline models (i.e., the EKF-UI [9] and the static NARX [10] model), qualitative results have shown that our model best estimates the true profile while others lacked accuracy and suffered from noise. Our model also exhibited the best quantitative result, achieving an averaged error metric of 0.008 (%) while posing the second-fastest computation time. Furthermore, we validated our model's capability of operating under various driving conditions by experimenting with the 50 KPH data set, while both baseline models failed to produce desirable results.
- Lastly, we extended our network structure to an autoregressive model and demonstrated the feasibility of sequence estimation. While we compared the models' estimation results with and without the two-phase attention mechanism, results have indicated that our proposed model with the two-phase attention mechanism significantly outperforms the baseline. Our model achieved an averaged error metric of 1.3 (%) when estimating a future sequence of 50 sampling time steps (0.05 sec) and conducted the action with reasonably fast computation time.

For the future work of the current research, we propose to train and evaluate the deep neural network with actual in-vehicle sensor data under free-ride driving conditions, including various driving speed intervals and road conditions. As feasibility has been proven with simulation data, we aim to utilize the trained network for actual active suspension control strategies.

REFERENCES

- [1] H. E. Tseng and D. Hrovat, "State of the art survey: Active and semi-active suspension control," *Vehicle Syst. Dyn.*, vol. 53, no. 7, pp. 1034–1062, Jul. 2015.
- [2] D. Hrovat, "Survey of advanced suspension developments and related optimal control applications," *Automatica*, vol. 33, no. 10, pp. 1781–1817, 1997.
- [3] S. Salleh, M. F. Rahmat, S. M. Othman, and K. A. Danapalasingam, "Review on modeling and controller design of hydraulic actuator systems," *Int. J. Smart Sens. Intell. Syst.*, vol. 8, no. 1, pp. 338–367, 2015.
- [4] X. Zhu, X. Jing, and L. Cheng, "Magnetorheological fluid dampers: A review on structure design and analysis," *J. Intell. Mater. Syst. Struct.*, vol. 23, no. 8, pp. 839–873, May 2012.
- [5] J.-Y. Han, A. Chen, and Y.-T. Lin, "Image-based approach for road profile analyses," *J. Surveying Eng.*, vol. 142, no. 1, Feb. 2016, Art. no. 06015003.
- [6] P. Kumar, P. Lewis, C. P. McElhinney, and A. A. Rahman, "An algorithm for automated estimation of road roughness from mobile laser scanning data," *Photogramm. Rec.*, vol. 30, no. 149, pp. 30–45, Mar. 2015.
- [7] J. Hwang, H. Yun, S.-K. Park, D. Lee, and S. Hong, "Optimal methods of RTK-GPS/accelerometer integration to monitor the displacement of structures," *Sensors*, vol. 12, no. 1, pp. 1014–1034, Jan. 2012.
- [8] S. J. Kim, K.-S. Kim, and Y.-S. Yoon, "Development of a tire model based on an analysis of tire strain obtained by an intelligent tire system," *Int. J. Automot. Technol.*, vol. 16, no. 5, pp. 865–875, Oct. 2015.
- [9] G.-W. Kim, S.-W. Kang, J.-S. Kim, and J.-S. Oh, "Simultaneous estimation of state and unknown road roughness input for vehicle suspension control system based on discrete Kalman filter," *Proc. Inst. Mech. Eng. D, J. Automobile Eng.*, vol. 234, no. 6, pp. 1610–1622, May 2020.
- [10] H. M. Ngwangwa, P. S. Heyns, H. G. A. Breytenbach, and P. S. Els, "Reconstruction of road defects and road roughness classification using artificial neural networks simulation and vehicle dynamic responses: Application to experimental data," *J. Terramech.*, vol. 53, pp. 1–18, Jun. 2014.
- [11] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng.*, vol. 82, no. 1, pp. 35–45, Mar. 1960.
- [12] S.-W. Kang, J.-S. Kim, and G.-W. Kim, "Road roughness estimation based on discrete Kalman filter with unknown input," *Vehicle Syst. Dyn.*, vol. 57, no. 10, pp. 1530–1544, 2019.
- [13] T. Nguyen, B. Lechner, and Y. D. Wong, "Response-based methods to measure road surface irregularity: A state-of-the-art review," *Eur. Transp. Res. Rev.*, vol. 11, no. 1, p. 43, Dec. 2019.
- [14] Z. Wang, M. Dong, Y. Qin, Y. Du, F. Zhao, and L. Gu, "Suspension system state estimation using adaptive Kalman filtering based on road classification," *Vehicle Syst. Dyn.*, vol. 55, no. 3, pp. 371–398, Mar. 2017.
- [15] Y. Qin, R. Langari, Z. Wang, C. Xiang, and M. Dong, "Road profile estimation for semi-active suspension using an adaptive Kalman filter and an adaptive super-twisting observer," in *Proc. Amer. Control Conf. (ACC)*, May 2017, pp. 973–978.
- [16] Y. Qin, C. Wei, X. Tang, N. Zhang, M. Dong, and C. Hu, "A novel nonlinear road profile classification approach for controllable suspension system: Simulation and experimental validation," *Mech. Syst. Signal Process.*, vol. 125, pp. 79–98, Jun. 2019.
- [17] Y. Qin, Z. Wang, C. Xiang, E. Hashemi, A. Khajepour, and Y. Huang, "Speed independent road classification strategy based on vehicle response: Theory and experimental validation," *Mech. Syst. Signal Process.*, vol. 117, pp. 653–666, Feb. 2019.
- [18] H. M. Ngwangwa, P. S. Heyns, F. J. J. Labuschagne, and G. K. Kululanga, "Reconstruction of road defects and road roughness classification using vehicle responses with artificial neural networks simulation," *J. Terramech.*, vol. 47, no. 2, pp. 97–111, Apr. 2010.
- [19] A. Solhmirzaei, S. Azadi, and R. Kazemi, "Road profile estimation using wavelet neural network and 7-DOF vehicle dynamic systems," *J. Mech. Sci. Technol.*, vol. 26, no. 10, pp. 3029–3036, Oct. 2012.
- [20] Y. Qin, R. Langari, and L. Gu, "The use of vehicle dynamic response to estimate road profile input in time domain," in *Proc. Dyn. Syst. Control Conf.*, vol. 46193, New York, NY, USA: American Society of Mechanical Engineers, 2014, Art. no. V002T27A002.
- [21] *CarSim Reference Manual*, Mech. Simul. Corp., Ann Arbor, MI, USA, 2006.
- [22] Y. Chen and J. Wang, "Adaptive vehicle speed control with input injections for longitudinal motion independent road frictional condition estimation," *IEEE Trans. Veh. Technol.*, vol. 60, no. 3, pp. 839–848, Mar. 2011.
- [23] D. Li, S. Du, and F. Yu, "Integrated vehicle chassis control based on direct yaw moment, active steering and active stabiliser," *Vehicle Syst. Dyn.*, vol. 46, no. 1, pp. 341–351, Sep. 2008.
- [24] G. Rill, *Road Vehicle Dynamics: Fundamentals and Modeling*. Boca Raton, FL, USA: CRC Press, 2011.
- [25] D.-S. Yoon, G.-W. Kim, and S.-B. Choi, "Response time of magnetorheological dampers to current inputs in a semi-active suspension system: Modeling, control and sensitivity analysis," *Mech. Syst. Signal Process.*, vol. 146, Jan. 2021, Art. no. 106999.
- [26] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [27] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2673–2681, Nov. 1997.
- [28] K. Cho, B. van Merriënboer, D. Bahdanau, and Y. Bengio, "On the properties of neural machine translation: Encoder-decoder approaches," 2014, *arXiv:1409.1259*. [Online]. Available: <http://arxiv.org/abs/1409.1259>
- [29] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," 2014, *arXiv:1409.0473*. [Online]. Available: <http://arxiv.org/abs/1409.0473>
- [30] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 3104–3112.

- [31] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: <http://arxiv.org/abs/1412.6980>
- [32] R. Ge, S. M. Kakade, R. Kidambi, and P. Netrapalli, "The step decay schedule: A near optimal, geometrically decaying learning rate procedure for least squares," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, pp. 14977–14988.
- [33] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proc. 13th Int. Conf. Artif. Intell. Statist.*, 2010, pp. 249–256.
- [34] Y. Gal and Z. Ghahramani, "A theoretically grounded application of dropout in recurrent neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2016, pp. 1019–1027.
- [35] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas, "Taking the human out of the loop: A review of Bayesian optimization," *Proc. IEEE*, vol. 104, no. 1, pp. 148–175, Jan. 2016.
- [36] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *J. Mach. Learn. Res.*, vol. 13, no. 1, pp. 281–305, Feb. 2012.
- [37] J. S. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl, "Algorithms for hyper-parameter optimization," in *Proc. Adv. Neural Inf. Process. Syst.*, 2011, pp. 2546–2554.
- [38] J. Wilson, F. Hutter, and M. Deisenroth, "Maximizing acquisition functions for Bayesian optimization," in *Proc. Adv. Neural Inf. Process. Syst.*, 2018, pp. 9884–9895.
- [39] R. N. Jazar, *Vehicle Dynamics: Theory and Application*. Cham, Switzerland: Springer, 2017.



GYUWON KIM received the B.S. degree from Inha University, Incheon, South Korea, in 2020. He is currently pursuing the M.S. degree with the Industrial Artificial Intelligence Laboratory, POSTECH, South Korea. His research interests include artificial intelligence integrated sensor systems for mechanical, electrical, and biomedical applications.



SOO YOUNG LEE received the B.S. degree from Chung-Ang University, Seoul, South Korea, in 2019. He is currently pursuing the M.S./Ph.D. degree with the Industrial Artificial Intelligence Laboratory, POSTECH, South Korea. His research interests include industrial artificial intelligence with mechanical systems and applications of AI-based smart manufacturing.



JONG-SEOK OH received the B.S., M.S., and Ph.D. degrees in mechanical engineering from Inha University, in 2009, 2011, and 2015, respectively. He is currently an Assistant Professor with Kongju National University. His research interests include robust controller design and control of various systems using smart actuators, such as magnetorheological (MR), electrorheological (ER) fluids, and piezo actuators.



SEUNGCHUL LEE received the M.S. and Ph.D. degrees from the University of Michigan, Ann Arbor, MI, USA, in 2008 and 2010, respectively. He was an Assistant Professor with the Ulsan National Institute of Science and Technology, South Korea, and a Postdoctoral Research Fellow with the University of Michigan. He has been an Assistant Professor with the Department of Mechanical Engineering, POSTECH, since 2018. His research interests include industrial artificial intelligence with mechanical systems, deep learning for machine health-care, IoT-based smart manufacturing, and manufacturing data visualization. He extends his research work to developing self-sustainable systems via an intelligent, informatics, and the industrial IoT system design at POSTECH.

• • •