This is a preprint. The DOI for the final published version is available at https://doi.org/10.23919/ICCAS63016.2024.10773139.

Digital Twin Simulation of Teleoperation: Communication Delay Compensation Using LSTM Network

Hwanhee Kang[1], Hyunrok Cha[2], Myeonghwan Hwang[2], Seungha Yoon[2], and Eugene Kim*

[1]Robot Engineering, Korea National University of Science and Technology,
Seoul, 13391, South Korea (hwan@kitech.re.kr)
[2]Purpose-Based Mobility group, Seonam Division, Korea Institute of Industrial Technology, Gwangju 61012, South Korea

* Corresponding author with the Purpose-Based Mobility group, Seonam Division, Korea Institute of Industrial Technology, Gwangju 61012, South Korea(egkim@kitech.re.kr)

Abstract: Accidents related to autonomous driving are being reported one after another, and the commercialization and diffusion of autonomous driving technology may be further into the future than we think. From this perspective, vehicle teleoperation technology is being considered as a complement to fully autonomous driving technology and is being piloted in some countries such as the United States. Teleoperation is used when encountering an emergency during autonomous driving, delivering a shared vehicle, or loading a new car into a transport vehicle. Despite the efforts of many researchers, Wireless Teleoperation is not free from the effects of communication delay due to physical limitations. The communication delay makes it difficult to guarantee real-time, resulting in poor teleoperation performance. Therefore, in order to perform smooth teleoperation, it is necessary to reduce the communication delay sufficiently, even if it cannot be completely eliminated. This study used RNN-based Long Short-Term Memory(LSTM) to reduce the impact of inevitable delays in teleoperation. The proposed model predicted the current control command of the steering angle considering communication delay and resulted in a smaller error from the real value compared to the raw sample. Given these results, it is expected that this study will contribute to real-time teleoperation that is robust to the communication delay.

Keywords: Teleoperation, time-varying delay, delay compensation, autonomous driving, LSTM

1. INTRODUCTION

Autonomous driving technology may enter our daily lives later than expected [1], [2], [3]. According to the California Department of Transportation, hundreds of incidents are reported each year where autonomous driving being disabled, most of which could have resulted in serious accidents if the vehicle occupants had not taken over the driving [4]. One of the technologies being discussed from this perspective is Teleoperation, a takeover technology through remote control [5]. Many companies around the world are already trying to distribute commercialized technologies to ensure safety without humans in the vehicle, such as urban driving and loading dock UAV control using teleoperation technology [6].

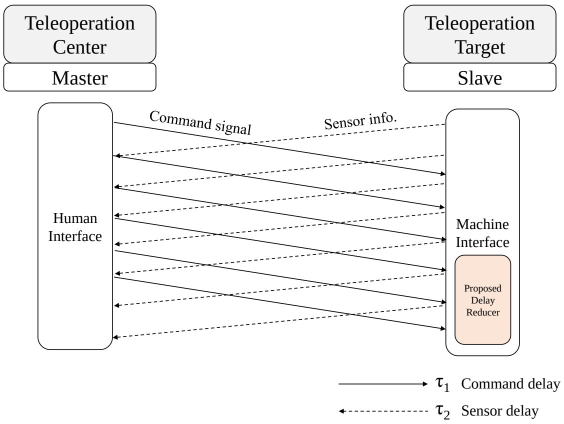

Fig. 1. Illustration of teleoperation system composed of a master-slave model. τ_1 is a command delay sent from the master, and τ_2 is a sensor delay received from the slave.

In the case of teleoperation, there are two types of communication delays: communication delay related to command signals, and communication delay related to sensor information [7]. Fig. 1 shows an overall structure of the teleoperation. First, in the case of command signals, there is a sense of control heterogeneity that occurs when the command is transmitted from a remote teleoperator to the target vehicle due to the physical constraints of wireless communication [8]. Especially when communication delays are large, control heterogeneity can increase the gap between the teleoperator's desired control and the outcome, thereby increasing the possibility of collisions between the vehicle and other objects. [9]. Second, the case of sensor information refers to image information for a remote teleoperator to recognize the current situation, but due to communication delay, there is a possibility that an incorrect command may be issued [10], [11]. Therefore, methods to reduce the impact of communication delay have been proposed, which are largely divided into predictive command signal and predictive display methods.

In the case of the predictive command signal, it is a method of supplementing the command signal under the assumption that the communication delay can be known

---

*This work was supported by “Development of Core Technologies for a Working Partner Robot in the Manufacturing Field” (KITECH EO-24007)

This is a preprint. The DOI for the final published version is available at https://doi.org/10.23919/ICCAS63016.2024.10773139.

accurately or with a small error [8]. A representative example is the study by Zheng et.al. In particular, research results were reported to increase the teleoperation performance as assessed by track completion time, track keeping error, and steering control effort using a predictor-based framework targeting high-speed UGV [9], [12]. In the case of predictive display, assuming that the communication delay is known, the video information shown on the teleoperator is predicted by artificially zooming or shifting previous frames along the movement path [10]. A representative example is the study by Prakash et.al. In particular, research results reported that predictive image information based on the expected movement path of the controlled vehicle obtained by FOV conversion of the camera using a depth camera improves the teleoperation performance [13], [11]. In this way, it is expected that many derivative technologies will be created thanks to the dedicated efforts of teleoperation.

The motivation for this study begins with the question of whether the predictive command signal method can be replaced by a data-driven delay compensation approach. The conventional predictive command signal method essentially systematically minimizes the influence of the communication delay on the teleoperation command signal [14], [8], [7]. However, there are many cases where it is not possible to improve performance according to the constantly changing communication environment systematically. Therefore, the purpose of this study is to improve the performance of predictive command signals through artificial intelligence learning. As an artificial intelligence technique that can be used on time-series sample, an LSTM(Long Short-Term Memory) of the RNN(Recurrent Neural Network) series is well known for its excellent performance.

1.1 Contribution

Our study contributes to a communication delay compensation approach using the data-driven command signal prediction method, especially for the time-varying delay following the Gaussian distribution. The result shows that the proposed method is effective not only when compensating for communication delays following a single Gaussian distribution, but also for communication delays mixed with two Gaussian distributions.

2. METHODOLOGY

In this section, the basic LSTM model and a proposed method are introduced.

2.1 Long Short Term Memory (LSTM) models

RNNs are used in time series prediction models because they have information about previous inputs [15]. However, there was a limitation that prediction performance decreased as the input length increased. To overcome the limitation of RNN, LSTM was developed which has three gates in one cell, of which the forgetting gate determines how much past information is to be reflected. The forget gate is the section that decides whether to keep or remove the past information [16]:

$$f_t = \sigma(W_{f_h}[h_{t-1}], W_{f_x}[x_t], b_f), \tag{1}$$

where the decision is based on the value of h_{t-1} and $x = (x_1, x_2, ..., x_t)$, which is the hidden state of the previous cell and the current input value respectively. The W_{f_h} and W_{f_x} are the weight matrices for each h and x respectively. Finally, the σ represents the non-linear function such as sigmoid, ReLU(Rectified Linear Unit), or hyperbolic tangent, and b_f denotes a constant bias for the f. The output f_t that passes through the sigmoid function is calculated as a value in the range of 0 to 1. If it is closer to 0, more of the existing information is forgotten, and if it is closer to 1, more is kept.

Next is an input gate which is well known as:

$$i_t = \sigma(W_{i_h}[h_{t-1}], W_{i_x}[x_t], b_i), \tag{2}$$

$$\tilde{c}_t = \tanh(W_{c_h}[h_{t-1}], W_{c_x}[x_t], b_c), \tag{3}$$

where the i_t is a value indicating whether new information should be updated, and $\tilde{c}_t$ is a vector indicating the information to be updated. Those two i_t and $\tilde{c}_t$ should be updated by using:

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t, \tag{4}$$

where c_{t-1} is the information vector of the previous time, which is maintained only as much as the output of the forget gate, and the new information vector is reflected as much as i_t to calculate c_t.

As a result, an output gate that uses two non-linear layers can be written as:

$$o_t = \sigma(W_{o_h}[h_{t-1}], W_{o_x}[x_t], b_o) \tag{5}$$

$$h_t = o_t * \tanh(c_t), \tag{6}$$

where o_t is the output value, and h_t is the hidden value ranging from -1 to 1.

2.2 Prediction based on delay compensation

The main concern in teleoperation is a large communication delay because it degrades real-time performance. The influence of the communication delay gets stronger, especially in cases when the delay is time-variant. The command signal samples that are treated as inputs to the LSTM network should be in the form of a sequence containing the values corresponding to multiple points in time. In this study, considering the size of the input or output may not be necessarily the same, a seq-2-seq architecture was introduced [17].

As Fig. 2 shows, the proposed model's encoder consists of LSTM trained the pattern or connectivity among the input sequence. Then, the decoder predicted the current steering angle by forecasting the unknown delay. What should be noticed is that the input vector consists of two vectors such as a steering angle vector and a communication delay vector. It is important to know the communication delay time in advance by comparing the timestamp when the signal was transmitted with the timestamp

This is a preprint. The DOI for the final published version is available at https://doi.org/10.23919/ICCAS63016.2024.10773139.

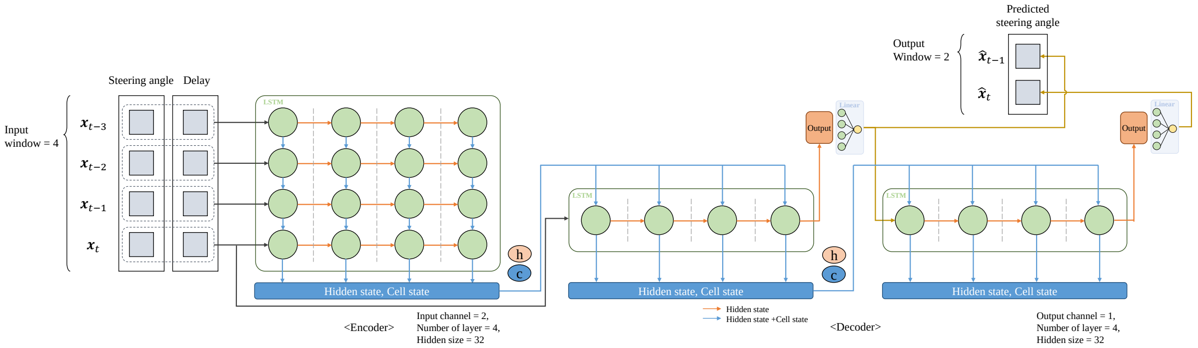

Fig. 2. Architecture of a proposed model which is composed of seq to seq model compensating the communication delay for the command signals. In this study two input vectors were used; steering angle, and communication delay (input window length: 4, output window length: 2)

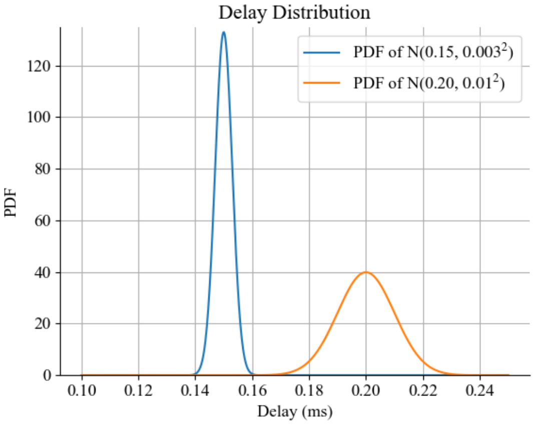

Fig. 3. Actual communication delay distribution consists of the mixture of the passive communication delay(blue line) always existing and the temporal communication delay(orange line), which is the same as the outlier[18]. In real teleoperation, the passive communication delay accounts for 90% and the temporal communication delay occupies the remainder.

when it was received. However, since the internal clocks of the master and slave machines are inherently independent of each other, those two clocks must be precisely synchronized in advance, such as using an NTP server. Therefore, in this study, we assume that the communication delay can be precisely measured.

3. EXPERIMENT

In this section, general experimental conditions are described.

3.1 Digital twin delay model

In this study, a realistic communication delay distribution reproducing that of real teleoperation conditions was generated by our previous study [18]. Fig. 3 shows two probability density functions following different Gaussian distributions which constitute the actual communication delay distribution. Adjusting the mixing ratio of these two probability density functions, several transformed distributions were used in prediction to achieve application in various situations:

$$\tau_1 \sim (1-\rho)\mathcal{N}(\mu_1, \sigma_1^2) + \rho\mathcal{N}(\mu_2, \sigma_2^2), \qquad (7)$$

where τ_1 is a command delay in the communication delay, μ_i and σ_i denote a mean and a standard deviation of the delay distribution that follows Gaussian. Especially ρ indicates a contamination ratio that scales the proportion of the mixed Gaussian distribution which stays between 0 and 1. In this study, we assume the mixture of the Gaussian delay distribution model which has 2-dimension which is: $\mu_1 = 0.15, \mu_2 = 0.20, \sigma_1 = 0.003, \sigma_2 = 0.01$ and ρ depends on conditions.

3.2 Performance estimation

Table 1. Parametric description consisting (7) where ρ is the contamination ratio.

	Con.#1	Con.#2	Con.#3	Con.#4
ρ	0	0.1	0.3	0.5

The prediction performance of the proposed network was compared with four communication delay models having different contamination ratios. As shown in Table 1, the contamination ratio differed from 0, 0.1, 0.3, and 0.5, respectively. The generated delay was added to the timestamp of the learning dataset [1] to create a delayed input time.

3.3 Sample preprocessing

Figure 4 demonstrates how steering angle samples are preprocessed and synchronized with time using the linear interpolation method. It should be noted that without the time synchronization, the original and the delayed steering angle samples are aligned at different points in time. After preprocessing, the actual steering angle corresponding to the delayed time is determined and used as the label for the delayed steering angle when entered into

---

[1] Comma ai steering dataset: `https://huggingface.co/datasets/commaai/commaSteeringControl`

the LSTM network. The total number of samples used for learning was 3500, and each sample contains 600 steering angle values collected 0.1 second intervals for 1 minute. As Fig. 5 shows, the input window was set to 4 and the output window to 2, which means that the LSTM network took four time points as the input at once and derived prediction for two time points. The stride was set to 1 so that the dataset was slid by 1 to form the training dataset. Finally, actual steering angles for two time points corresponding to the prediction of LSTM entered the LSTM network as the label.

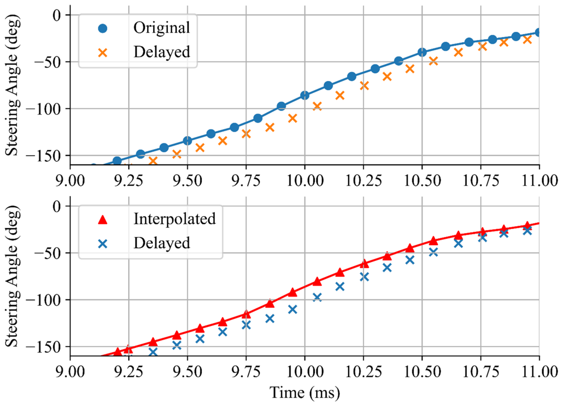

Fig. 4. Steering angle sample interpolation considering time synchronization. Since the delayed steering angle is already a past command signal when received, interpolation is used to obtain the actual steering angle to be achieved at that time.

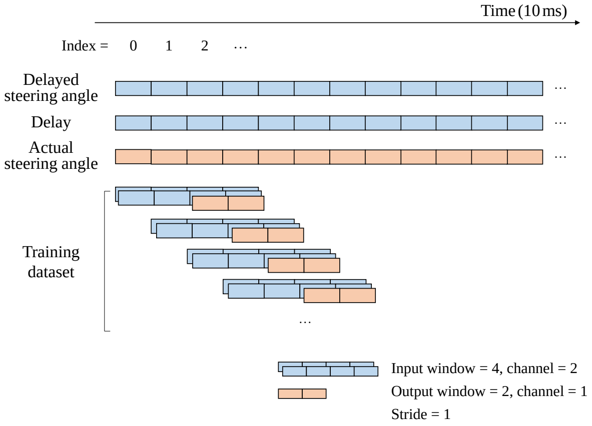

Fig. 5. Schematic diagram showing the training dataset generation using a slide window method. Actual steering angle sets consisted of interpolated steering angles in Fig. 4. The input window and channel are 4 and 2, while the output window and channel are 2 and 1, respectively.

4. RESULT & DISCUSSION

This section shows the results of the proposed method applied to the samples having four different delay models. As shown in table2 2, the RMSE values of predicted samples were reduced by more than 50% compared to the RMSE of raw samples through the communication delay compensation in condition#1∼#3 despite both the RMSE of raw samples and predicted samples increased as the contamination ratio increased. On the other hand, such delay compensation performance rapidly decreased in condition#4 where theρ value is 0.5, indicating that large communication delay outliers make prediction difficult.

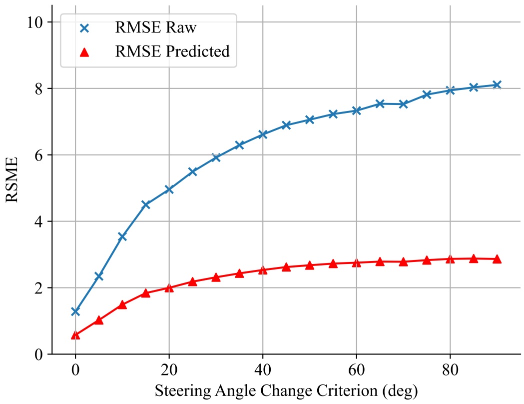

Fig. 6. RMSE of predicted samples applying the proposed method when the delay was following the condition#1 (ρ : 0).

Table 2. Result of communication delay compensation when the delay was following the condition#1∼4 (ρ : $0, 0.1, 0.3, 0.5$)(7) where ρ is the contamination ratio.

	Con.#1	Con.#2	Con.#3	Con.#4
RMSE Raw	1.28	1.41	1.56	1.63
RMSE Predicted	0.58	0.69	0.78	1.06
Reduction	54.69%	51.06%	50%	34.97%

In order to show the impact of the performance, the result was aligned according to the steering angle change. Fig. 6 shows the RMSE of predicted samples and raw samples in condition#1 where the ρ was 0. As the result shows, the RMSE of predicted samples was able to compensate for the communication delay compared to that of raw, delayed steering angle samples. Moreover, it can be known that the proposed method is more robust when the samples have large steering angle changes.

Fig. 7 presents all results when applied with the proposed model when the ρ was from 0 to 0.5. As a result, the proposed method significantly reduced the communication delay, and the larger the steering angle change, the greater the difference between the RMSE of predicted samples and raw samples.

This is a preprint. The DOI for the final published version is available at https://doi.org/10.23919/ICCAS63016.2024.10773139.

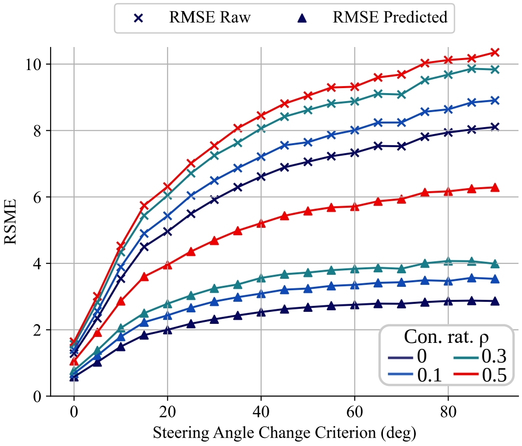

Fig. 7. RMSE of predicted samples applying the proposed method when the delay was following the condition#1∼4 (ρ : $0, 0.1, 0.3, 0.5$).

5. CONCLUSION

From the viewpoint of teleoperation, the communication delay has to be compensated in the sense of reducing the heterogeneity between the sent command signal and the actual actuation of the remote target. To the best of the author's knowledge, there were few approaches were reported constructed based on data-driven command signal compensator. In this study, an approach toward replacing the model and a non-model-based systematic approach was simulated by the use of the Long Short Term Memory model that has command signal vectors and instant communication delay vectors as input. As for the comparison, based on an actual communication delay, a digital twin delay was generated with four different outlier contamination ratios and was applied to a steering command signal input time. The result shows that regardless of the magnitude of the communication delay outlier, the model showed that it can not only compensate for the communication delay but also resist the communication delay outliers. Our future study includes expanding the proposed model to diverse application scenarios such as surgical robots, and the aerospace field, and evaluating of the performance with further extreme conditions to secure the stability of the teleoperation control.

REFERENCES

[1] A. Bhardwaj, A. H. Ghasemi, Y. Zheng, H. Febbo, P. Jayakumar, T. Ersal, J. L. Stein, and R. B. Gillespie, "Who's the boss? arbitrating control authority between a human driver and automation system," *Transportation research part F: traffic psychology and behaviour*, vol. 68, pp. 144–160, 2020.

[2] D. Omeiza, H. Webb, M. Jirotka, and L. Kunze, "Explanations in autonomous driving: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 10142–10162, 2021.

[3] M. Sadaf, Z. Iqbal, A. R. Javed, I. Saba, M. Krichen, S. Majeed, and A. Raza, "Connected and automated vehicles: Infrastructure, applications, security, critical challenges, and future aspects," *Technologies*, vol. 11, no. 5, p. 117, 2023.

[4] State of california department of motor vehicle, "Autonomous vehicle disengagement reports 2019-2020," 2020.

[5] D. Majstorović, S. Hoffmann, F. Pfab, A. Schimpe, M.-M. Wolf, and F. Diermeyer, "Survey on teleoperation concepts for automated vehicles," in *2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 1290–1296, IEEE, 2022.

[6] E. Kim, T. Lim, M. Hwang, and H. Cha, "State estimation predictor with stochastic time-delay based framework for teleoperated vehicle," in *2023 IEEE 6th International Conference on Knowledge Innovation and Invention (ICKII)*, pp. 119–122, IEEE, 2023.

[7] E. Kim, M. Hwang, T. Lim, C. Jeong, S. Yoon, and H. Cha, "Communication delay outlier detection and compensation for teleoperation using stochastic state estimation," *Sensors*, vol. 24, no. 4, p. 1241, 2024.

[8] Y. Zheng, M. J. Brudnak, P. Jayakumar, J. L. Stein, and T. Ersal, "A predictor-based framework for delay compensation in networked closed-loop systems," *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 5, pp. 2482–2493, 2018.

[9] Y. Zheng, M. J. Brudnak, P. Jayakumar, J. L. Stein, and T. Ersal, "A delay compensation framework for predicting heading in teleoperated ground vehicles," *IEEE/ASME Transactions on Mechatronics*, vol. 24, no. 5, pp. 2365–2376, 2019.

[10] M. J. Brudnak, "Predictive displays for high latency teleoperation," in *Proc. NDIA Ground Veh. Syst. Eng. Technol. Symp.*, pp. 1–16, 2016.

[11] J. Prakash, M. Vignati, D. Vignarca, E. Sabbioni, and F. Cheli, "Predictive display with perspective projection of surroundings in vehicle teleoperation to account time-delays," *IEEE Transactions on Intelligent Transportation Systems*, 2023.

[12] Y. Zheng, M. J. Brudnak, P. Jayakumar, J. L. Stein, and T. Ersal, "Evaluation of a predictor-based framework in high-speed teleoperated military ugvs," *IEEE Transactions on Human-Machine Systems*, vol. 50, no. 6, pp. 561–572, 2020.

[13] J. Prakash, M. Vignati, E. Sabbioni, and F. Cheli, "Vehicle teleoperation: Human in the loop performance comparison of smith predictor with novel successive reference-pose tracking approach," *Sensors*, vol. 22, no. 23, p. 9119, 2022.

[14] S. Lu, M. Y. Zhang, T. Ersal, and X. J. Yang, "Workload management in teleoperation of unmanned ground vehicles: Effects of a delay compensation aid on human operators' workload and teleoperation performance," *International Journal of Human–Computer Interaction*, vol. 35, no. 19, pp. 1820–

This is a preprint. The DOI for the final published version is available at https://doi.org/10.23919/ICCAS63016.2024.10773139.

1830, 2019.
[15] A. Sherstinsky, “Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network,” *Physica D: Nonlinear Phenomena*, vol. 404, p. 132306, 2020.
[16] S. Siami-Namini, N. Tavakoli, and A. S. Namin, “The performance of lstm and bilstm in forecasting time series,” in *2019 IEEE International conference on big data (Big Data)*, pp. 3285–3292, IEEE, 2019.
[17] I. Sutskever, O. Vinyals, and Q. V. Le, “Sequence to sequence learning with neural networks,” *Advances in neural information processing systems*, vol. 27, 2014.
[18] E. Kim, H. Cha, T. Talluri, C. Jeong, H. Kim, S. Yoon, and M. Hwang, “Delay compensation and outlier removal in command signal for teleoperated vehicle using asynchronous filter,” *IEEE Access*, 2024.