Taming Signal-Dependent Counting Noise with Machine Learning for Molecular Communication

Yaqing Zhang[1], Min Luo[1], Chao Wang[1], Miaowen Wen[2], Fei Ji[2], and Yu Huang[1,2](✉)

[1] Research Center of Intelligent Communication Engineering, School of Electronics and Communication Engineering, Guangzhou University, Guangzhou, China
Yuhuang@gzhu.edu.cn

[2] Guangdong Provincial Key Laboratory of Short-Range Wireless Detection and Communication, School of Electronic and Information Engineering in South China University of Technology, Guangzhou, China

Abstract. Molecular communication (MC) is an emerging paradigm for exchanging information via chemical signals. It holds great promise for building nanoscale networks within biological organisms. However, the presence of counting noise originating from the molecular diffusion mechanism severely limits the signal detection performance. Currently, research on reducing such signal-dependent noise (SDN) in MC via diffusion (MCvD) remains scarcely explored. Besides, the lack of channel models in MCvD hinders the effects of model-based denoising approaches. Against this background, we resort to the data-driven machine learning (ML) methods that fit MCvD systems for noise mitigation. Additionally, the simulated numerical results show that such an ML-based scheme outperforms its classical model-based counterparts.

Keywords: Denoise · Diffusion · Machine learning · Molecular communication · Signal-dependent noise

1 Introduction

In recent years, molecular communication (MC) has garnered significant attention as a novel communication mode. Inspired by nature, the encoding, transmission, and decoding of information in MC are achieved through the utilization of information molecules, with excellent biocompatibility and efficient energy use, which has occupied an important position in nanonetwork engineering [1]. This new form of communication brings entirely new possibilities for nanoscale

This work was supported in part by the National Natural Science Foundation of China under Grant 62201161, in part by the Guangzhou Municipal Science and Technology Project under Grant 2023A04J1717, and in part by the Tertiary Education Scientific Research Project and Key Discipline Project of Guangzhou Education Bureau under Grants 202235019 and 202255467, respectively.

Y. Chen et al. (Eds.): BICT 2024, LNICST 592, pp. 76–86, 2025.
https://doi.org/10.1007/978-3-031-81599-7_8

network communication, providing revolutionary opportunities for advances in other fields such as medical care and environmental monitoring [2].

In MC, there are a variety of molecular transmission methods, including diffusion-based [3], flow-based, etc., among which MC via diffusion (MCvD) is one of the simplest methods to achieve MC, which simply depends on the Brownian motion of information molecules in the environment without further external energy. In the MCvD system, in addition to the common background noise, there exists counting noise, which is a type of signal-dependent noise (SDN) that arises due to random transmission of information molecules. Noise and inter-symbol interference (ISI) are inevitable during molecular transmission, exerting a significant impact on the accurate transmission and reception of information molecules, thereby substantially compromising system performance.

The initial research of SDN originated in the field of image processing [4], and then gradually extended to the field of visible light communication (VLC). In VLC, SDN has attracted a lot of attention, with research efforts focusing on modulation design [5,6], capacity boundary [7,8], and other challenges associated with SDN. Concurrently, researchers have also identified the presence of SDN in MC and conducted corresponding investigations. However, research in this area remains limited, particularly in terms of noise reduction methods for SDN. It is noteworthy that [9] inspired by stochastic resonance, a linear filter is designed to convert noise into useful signals and improve the signal-to-noise ratio, which effectively improves the performance of the system. In [10], a noise suppression filter, the I-filter, is proposed to reduce the influence of noise in MCvD by suppressing noise. Although these studies provide a certain foundation for SDN noise reduction methods, it is essential to explore novel noise reduction techniques further to achieve improved efficacy in reducing noise.

In the field of signal processing and communication, machine learning (ML) has great potential as a powerful tool. Notably, in MC, ML methods have attracted the attention and interest of researchers. Numerous researchers have extensively investigated the utilization of ML techniques for detecting received signals in MC systems [11–14]. Considering the non-stationary characteristics of SDN, traditional model-based noise reduction methods may have limitations in addressing this type of noise. In other words, denoising methods that are based on signal-independent noise models are not very effective in scenarios with SDN. In contrast, the data-driven approach of ML is more adaptable and flexible, allowing the model to adjust adaptively to specific data characteristics. By training models on large amounts of data, ML can gain a deeper understanding of the intricate relationship between signal and noise. Therefore, we contend that the application of ML techniques holds significant implications for addressing SDN challenges in MC. The main contribution of this paper lies in the application of ML methods based on the Long Short-Term Memory Recurrent Neural Network (LSTM-RNN) model to mitigate SDN in MCvD effectively.

The rest of the paper is organized as follows. Section 2 revisits the MCvD system model. In Sect. 3, we propose a denoising method based on machine learning using the LSTM-RNN model. In Sect. 4, the numerical results demonstrate that the LSTM-RNN model exhibits remarkable efficacy in mitigating SDN. Finally, conclusions are drawn in Sect. 5.

2 System Model

In this paper, we consider an MCvD system operating within a three-dimensional (3D) environment. As shown in Fig. 1, we consider a typical end-to-end model [15], which consists of a point source transmitter, a diffusion channel, and a passive spherical receiver with a radius of r, whose volume is $V_R = \frac{4}{3}\pi r^3$. The transmitter is located at the origin of the Cartesian coordinate system. The information molecules in the diffusion channel follow Brownian motion, where the spatiotemporal distribution of the average molecular concentration obeys Fick's second law of diffusion, which can be expressed as

$$\frac{1}{D}\frac{\partial c(\mathbf{d},t)}{\partial(t)} = \nabla^2 c(\mathbf{d},t), \tag{1}$$

where ∇^2 is the Laplace operator and D is the diffusion coefficient. Furthermore, $\mathbf{d}$ is the location of information molecules in the 3D Cartesian coordinates, which can be represented as $\mathbf{d} = [x, y, z]$. $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ in Cartesian coordinate, $c(\mathbf{d}, t)$ represents the average concentration of information molecules at coordinate $\mathbf{d}$ and time t. When the transmitter emits a pulse signal at time $t = 0$, the corresponding initial condition has the following form as

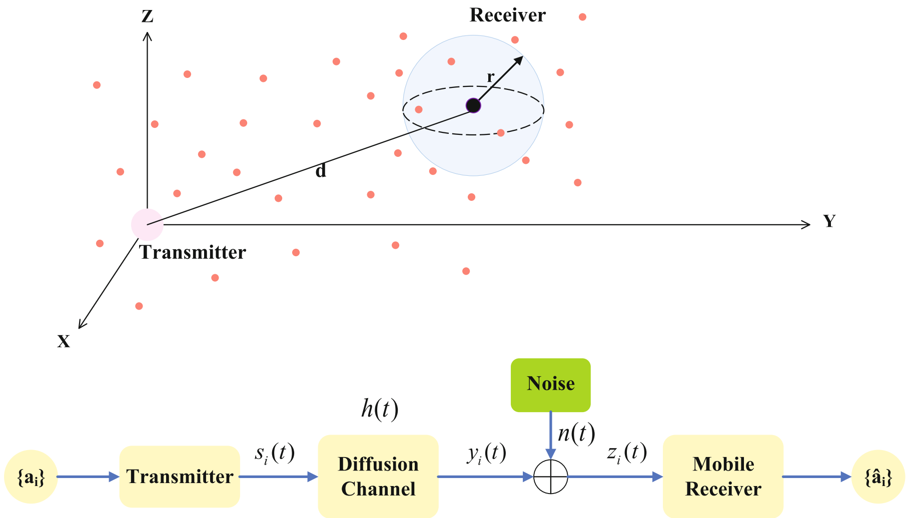

Fig. 1. MCvD system model.

$$c(\mathbf{0}, t \to 0) = Q\delta(\mathbf{d}), \tag{2}$$

where Q represents the number of information molecules in each pulse signal, and $\delta(\cdot)$ represents the Dirac delta function. Additionally, considering the unbounded environment, the boundary conditions can be expressed as

$$c\left(\|\mathbf{d}\| \to \infty, t\right) = 0 \tag{3}$$

Assuming that the received molecular concentration is uniform in the receiving space, molecular concentration can be solved (2) with (3) and (4)

$$c(\mathbf{d}, t) = \frac{Q}{(4\pi Dt)^{3/2}} \exp\left(-\frac{\|\mathbf{d}\|^2}{4Dt}\right). \tag{4}$$

We assumed a fixed location of the receiver as $\mathbf{d} = [d, 0, 0]$, so the distance of the transceiver is a constant value, i.e., $\|\mathbf{d}\| = d$. Based on (4), the channel impulse response of the end-to-end MC system is

$$h(t) = \frac{c(\mathbf{d}, t)}{Q} = \frac{1}{(4\pi Dt)^{3/2}} \exp(-\frac{d^2}{4Dt}). \tag{5}$$

In terms of modulation, the on-off keying (OOK) approach is deployed. Let $\mathbf{a} = [a_1, \ldots, a_i, \ldots, a_I]$ represent a transmitted binary block of length I, where the value of the i-th entry a_i is either 0 or 1. In OOK, the transmitter emits Q information molecules in an impulsive manner for bit-1 transmission. Conversely, it keeps silent when transmitting bit-0. Therefore, the transmitted signal by the i-th symbol can be expressed as

$$s_i(t) = a_i Q\delta[t - (i-1)T], \tag{6}$$

where T represents the symbol interval. Due to the randomness in the transmission of information molecules, previously transmitted signal particles may remain in the channel, affecting the current received signal, known as ISI. In the absence of noise, the concentration of information molecules that reach the receiver after the i-th symbol propagates through the diffusion channel is

$$y_i(t) = s_i(t)h(t) = \sum_{i=1}^{\min\left\{\left\lfloor \frac{t}{T} \right\rfloor, I\right\}} a_i Q h[t - (i-1)T], \tag{7}$$

where $\lfloor \cdot \rfloor$ denotes the floor function. Noise interference exists in the MC system, including background noise and counting noise. In OOK modulation, the transmission of a bit-1 is typically affected by the signal-dependent counting noise rather than the conventional background noise that is signal-independent [15]. Since the noise reduction algorithm for background noise is relatively mature, we focus on the denoising of counting noise, which is rarely studied in the literature. Thus, the received signal of the i-th symbol in the diffusion-based MC can be formulated as follows

$$z_i(t) = y_i(t) + n(t). \tag{8}$$

The counting noise process, i.e., $n(t)$, is caused by the randomness in the molecular diffusion mechanism, which follows a normal distribution with a mean of zero [16]

$$n(t) \sim \mathcal{N}\big(0, \sigma^2(t)\big), \tag{9}$$

where the variance term is dependent on the signal strength and is expressed as

$$\sigma^2(t) = \frac{y(t)}{V_R}. \tag{10}$$

3 Machine Learning Based Denoising Method For SDN

MC systems are frequently subjected to intricate noise interference, which has a substantial impact on the quality and reliability of signals, thereby restricting system performance. Therefore, it is essential to propose a noise reduction method to eliminate noise interference. Traditional denoising techniques typically focus on noise that is unrelated to the signal. However, ML, as a powerful data-driven approach, can learn the characteristics of the noise model through training on data and accurately predict and remove noise. In contrast to traditional methods, ML not only effectively captures the characteristics of the SDN, improving system robustness and adaptability, but also demonstrates superior ability in managing complex noise interference within the MC system, thus offering new opportunities for optimizing system performance. Therefore, integrating ML into MC can effectively eliminate the SDN and subsequently improve the performance of the MC system. In ML, there is a wide variety of neural networks to choose from. Given that the signal in MC is time series data, we have chosen to use recurrent neural networks (RNNs) as the fundamental network structure. It is worth noting that in the field of MC, RNNs have been used as the basic network structure for sequence detection, and the effect is remarkable [14]. RNNs can capture temporal dependencies in time series data through recurrent connections. However, traditional RNNs suffer from the problem of vanishing gradients during training, which makes it challenging for the network to learn long-range temporal dependencies. Therefore, in this context, we choose an improved RNN architecture known as the LSTM network. This addresses the problem of gradient disappearance in traditional RNNS by introducing gating mechanisms. Compared to traditional RNNs, the LSTM-RNN model can effectively capture long-term dependencies in time series data through a gating mechanism. This enables it to model the characteristics of the SDN more accurately and improve the system's ability to suppress noise.

To fully exploit the temporal characteristics of LSTM-RNN models, we opt to employ mutual information as a metric for assessing correlation. Mutual information $I(X;Y)$ quantifies the relationship between two random variables X and Y, where X and Y represent noisy signal and noiseless signal, respectively.

$$I(X;Y) = \sum_{x}\sum_{y} P(X=x, Y=y) \log\left(\frac{P(X=x, Y=y)}{P(X=x)P(Y=y)}\right). \tag{11}$$

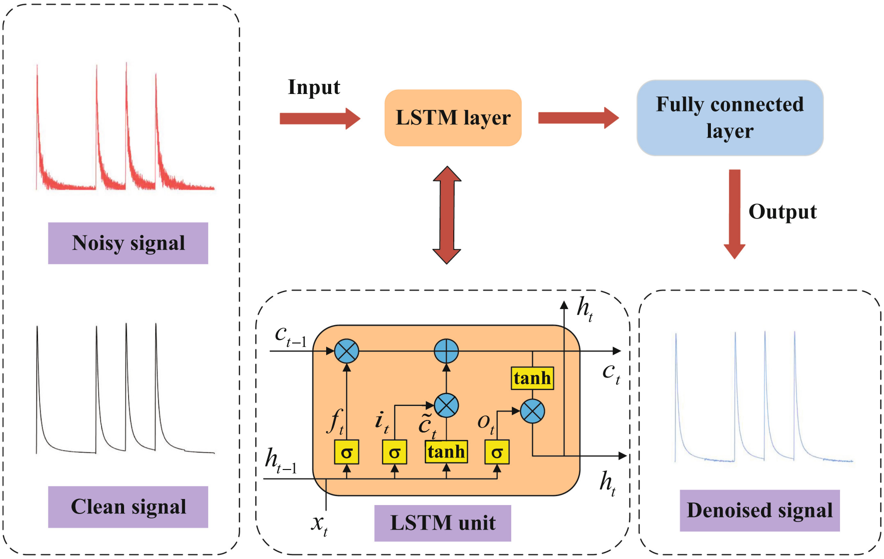

Fig. 2. LSTM-RNN model architecture for denoising.

When the value of $I(X;Y)$ is zero, it indicates a lack of correlation between the noisy signal and the non-noisy signal. Conversely, as the value of $I(X;Y)$ increases, a stronger correlation is observed between these signals. This suggests a more significant information correlation exists between them, with one variable being more predictive of the other's value. Based on this observation, we select an appropriate memory length and train the LSTM-RNN model to capture the temporal correlation between noisy and non-noisy signals effectively.

In Fig. 2, the proposed LSTM-RNN model architecture for denoising consists of an input layer, an LSTM layer, a fully connected layer, and an output layer. The LSTM unit [17] consists of the input gate, the forget gate, and the output gate. These gates control the updating of the LSTM cell state. Specifically, the input gate controls whether new information should be added to the cell state, the amnesia gate controls whether information from previous cell states should be discarded, and the output gate controls the extent to which information in the cell state is output in the current hidden state of the time step. The LSTM unit employs two activation functions, namely the Sigmoid function $\sigma(\cdot)$ and the hyperbolic tangent function $\tanh(\cdot)$. $\sigma(\cdot)$ converts the input to a range of (0, 1) and is commonly employed to control the on-off state of the door. The function $\tanh(\cdot)$ converts the input into the range $(-1, 1)$ and is primarily used to calculate new candidate state values. In (12)–(17), x_t represents the input of the current time step, $\mathbf{W}$ and $\mathbf{U}$ represent input weight matrices and recurrent weight matrices, and b represents bias. The activations of the gates, f_t, i_t, and o_t are calculated for a timestep t using the activation of the cell at $t-1$ (c_{t-1}) according to Eq. (12), (13), and (14).

$$f_t = \sigma(\mathbf{W}_f x_t + \mathbf{U}_f h_{t-1} + b_f), \tag{12}$$

$$i_t = \sigma(\mathbf{W}_i x_t + \mathbf{U}_i h_{t-1} + b_i), \tag{13}$$

$$o_t = \sigma(\mathbf{W}_o x_t + \mathbf{U}_o h_{t-1} + b_o). \tag{14}$$

Formula (15) describes the process of calculating the candidate cell state $\tilde{c}_t$ based on the current input and the previous hidden state.

$$\tilde{c}_t = \tanh(\mathbf{W}_c x_t + \mathbf{U}_c h_{t-1} + b_c). \tag{15}$$

Formulas (16) and (17) describe the updating process of the cell state c_t and the hidden state h_t, obtained by multiplying the value of the forgetting gate by the previous cell state, and the candidate cell state by the value of the input gate, and adding them. The cell state is then processed by the tanh function and multiplied with the output gate to get the hidden state.

$$c_t = f_t \otimes c_{t-1} + i_t \otimes \tilde{c}_t, \tag{16}$$

$$h_t = o_t \otimes \tanh(c_t), \tag{17}$$

where $\otimes$ represents the element-by-element product.

4 Numerical Results

In this section, we aim to evaluate the noise reduction performance of the proposed LSTM-RNN network model in the MCvD system. Specifically, simulation experiments are conducted on simulated signal data to verify the effect of the network model in reducing noise.

4.1 Experimental Setup

The system parameters are set as follows: the diffusion coefficient is $D = 2.2 \times 10^{-9}\,\mathrm{m}^2/\mathrm{s}$, and the number of information molecules released by the transmitter is $Q = 2 \times 10^9$. The distance between the transceivers is $d = 20\,\mu\mathrm{m}$. The simulation experiment is set up as follows: The dataset for the simulation experiment consists of a continuous transmission of a sequence comprising 1000 randomly generated symbols, either "0" or "1". The dataset is divided into a ratio of $7:2:1$, with 70% of the symbols allocated to the training set, 20% to the validation set, and 10% to the test set. This scaling is a common practice in ML to make the most of the data, validate the model's performance, and evaluate its ability to generalize. The training set is used to train the neural network model, while the validation set is employed to tune the model's hyperparameters and monitor the training process. Finally, the test set is used to evaluate the model's performance in reducing noise. The major model hyperparameters are summarized in Table 1.

Table 1. LSTM network hyperparameters

Hyperparameter	Value
Number of layers	5
Hidden state size	4
Epoch	60
Batch size	32
Initial learn rate	0.01
Learn rate drop factor	0.9
Learn rate drop period	10

4.2 Simulation Results

We test the noise reduction performance of this network model at various symbol intervals ($T = 0.4 \sim 2.0$ s) and use the symbol interval $T = 1$ s as an example. In this paper, we use the root mean square error (RMSE) as an indicator to evaluate noise reduction performance, which is defined as follows

$$RMSE = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(\hat{r}(j) - r(j)\right)^2}, \tag{18}$$

where N is the signal length, $r(j)$ represents the normalized denoised signal vector, and $\hat{r}(j)$ represents the normalized noiseless signal vector.

Figure 3 illustrates the noise reduction effect of certain symbols in the test set, which clearly shows that the denoised signal is almost perfectly restored to the noiseless signal, highlighting the excellent performance of the model in suppressing SDN. In addition, the RMSE value for the test set is measured at 0.005, further confirming the excellent performance of the model in suppressing the SDN. In general, the model demonstrates excellent performance in reducing noise and has a remarkable effect on suppressing the SDN.

The FIR low-pass filter and wavelet denoising are two classical methods widely used in signal processing. FIR low-pass filters are well-known for their simple structure and effective filtering capabilities. It can filter out high-frequency noise while retaining low-frequency information, making it a classical method commonly used for signal denoising. On the other hand, wavelet denoising can offer improved local feature representation in the time-frequency domain. Its multi-scale analysis characteristics enable it to capture instantaneous features and frequency information in the signal. Both methods have distinct advantages in addressing noise. Therefore, we compare these two model-based methods with the data-driven LSTM-RNN models, focusing on their noise reduction performance. To ensure optimal performance, we utilize a brute-force search method to determine the optimal parameters for the FIR low-pass filter and wavelet transform. Specifically, the FIR low-pass filter is implemented using the fir1 function, the filtering order of 4, and the cutoff frequency of 0.001. For the

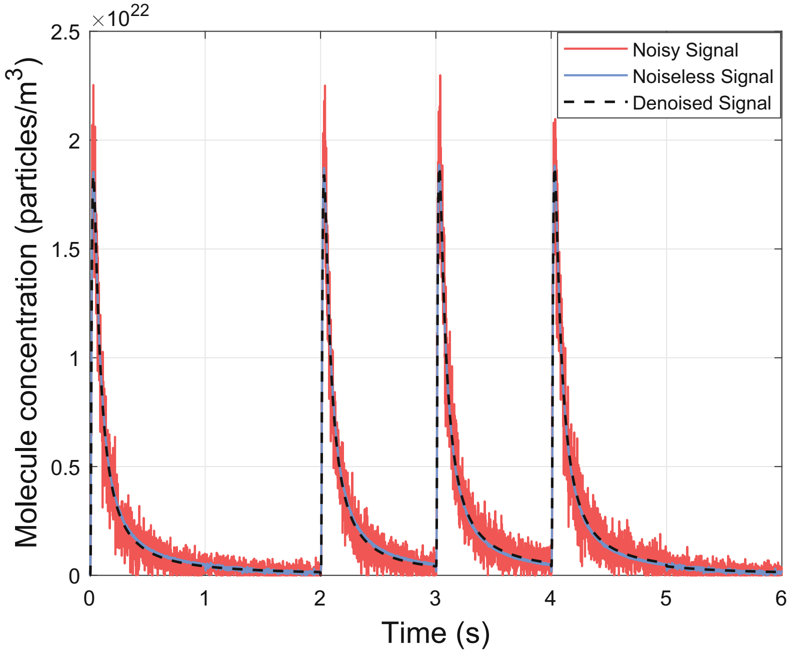

Fig. 3. The denoising effect of LSTM-RNN model in MC.

wavelet denoising, we choose rbio1.3 as the wavelet basis function. In the experiments, we compare the performance of three methods under various symbol intervals ($T = 0.4 \sim 2.0$ s) and evaluate quantitatively by RMSE. By observing the comparison results in Fig. 4, we note that the RMSE values of all three methods decrease as the symbol interval increases. Furthermore, the noise reduction method employing the LSTM-RNN model demonstrates a smaller RMSE value at various symbol intervals compared to the traditional FIR low-pass filter and wavelet denoising, which indicates that the LSTM-RNN model performs better in mitigating SDN.

The superior performance of the LSTM-RNN model can be attributed to its robust ML capabilities and adaptive network architecture. The proposed model has powerful ML capabilities to learn complex patterns and Temporal characteristics in data. In contrast, traditional FIR low-pass filters and wavelet denoising methods heavily rely on mathematical modeling and signal processing, making it challenging to adapt to the complex characteristics of the SDN flexibly. More importantly, the architecture of the LSTM-RNN model is well-suited for processing time series data, effectively capturing and learning long-term dependencies within the data. Furthermore, the flexible network structure enables better adaptation to the non-stationary characteristics of SDN. In contrast, traditional methods may be limited by fixed filter characteristics or wavelet basis functions, which makes them less adaptable to changing noise characteristics. Therefore,

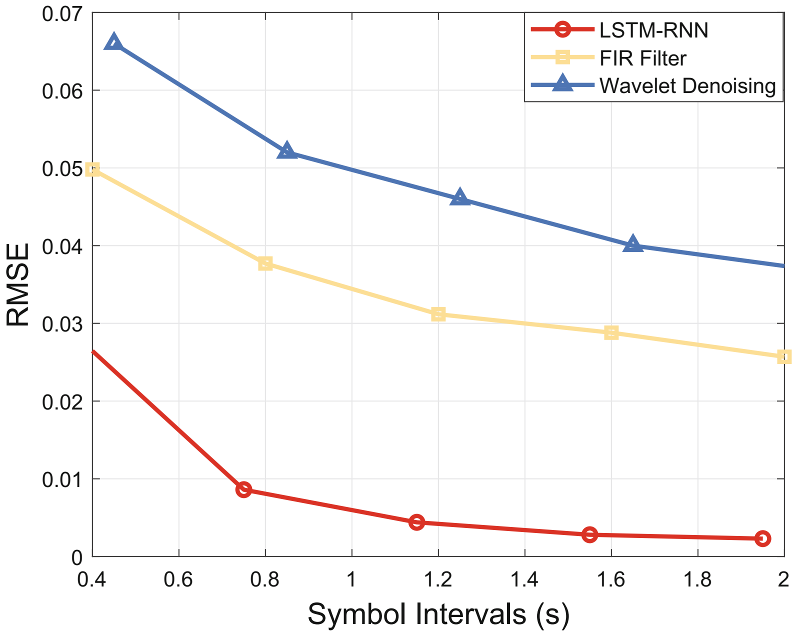

Fig. 4. Comparison of the noise reduction effect of three methods at different intervals.

our proposed LSTM-RNN model demonstrates superior performance in reducing noise while mitigating the SDN.

5 Conclusion

In this paper, we proposed an innovative solution for the SDN in MCvD, specifically the LSTM-RNN model-based ML method for denoising. The proposed method's performance is evaluated through simulation experiments, comparing the RMSE of an ML-based noise reduction network with that of traditional noise reduction methods across various symbol intervals. These results confirm the excellent potential of the LSTM-RNN model in mitigating the SDN. This approach provides an effective solution for enhancing the performance and reliability of the MC system. Future research could explore and optimize other network structures within ML to tackle more intricate and realistic noise scenarios.

References

1. Lu, Y., Ni, R., Zhu, Q.: Wireless communication in nanonetworks: current status, prospect and challenges. IEEE Trans. Mol. Biol. Multi-Scale Commun. **6**(2), 71–80 (2020)

2. Kong, L., et al.: A survey for possible technologies of micro/nanomachines used for molecular communication within 6G application scenarios. IEEE Internet Things J. **10**(13), 11240–11263 (2023)
3. Farsad, N., Yilmaz, H.B., Eckford, A., Chae, C.B., Guo, W.: A comprehensive survey of recent advancements in molecular communication. IEEE Commun. Surveys Tuts. **18**(3), 1887–1919 (2016)
4. Kuan, D.T., Sawchuk, A.A., Strand, T.C., Chavel, P.: Adaptive noise smoothing filter for images with signal-dependent noise. IEEE Trans. Pattern Anal. Mach. Intell. **PAMI-7**(2), 165–177 (1985)
5. Gao, Q., Qaraqe, K., Serpedin, E.: Improving the modulation designs for visible light communications with signal-dependent noise. IEEE Commun. Mag. **58**(5), 26–32 (2020)
6. Wei, J., Wang, Y., Gong, C., Huang, N.: Noise analysis and modulation optimization for nonlinear visible light communication system with signal-dependent noise. IEEE Photon. J. **15**(6), 1–11 (2023)
7. Wang, J.Y., Fu, X.T., Lu, R.R., Wang, J.B., Lin, M., Cheng, J.: Tight capacity bounds for indoor visible light communications with signal-dependent noise. IEEE Trans. Wireless Commun. **20**(3), 1700–1713 (2021)
8. Wang, J.Y., et al.: Secrecy-capacity bounds for visible light communications with signal-dependent noise. IEEE Trans. Wireless Commun. **22**(11), 7227–7242 (2023)
9. Zheng, R., Lin, L., Yan, H.: A noise suppression filter for molecular communication via diffusion. IEEE Wireless Commun. Lett. **10**(3), 589–593 (2021)
10. Li, B., et al.: CSI-independent non-linear signal detection in molecular communications. IEEE Trans. Signal Process. **68**, 97–112 (2020)
11. Huang, Y., Ji, F., Wei, Z., Wen, M., Guo, W.: Signal detection for molecular communication: model-based vs. data-driven methods. IEEE Commun. Mag. **59**(5), 47–53 (2021)
12. Boulogeorgos, A.A., Trevlakis, S.E., Tegos, S.A., Papanikolaou, V.K., Karagiannidis, G.K.: Machine learning in nano-scale biomedical engineering. IEEE Trans. Mol. Biol. Multi-Scale Commun. **7**(1), 10–39 (2021)
13. Torres Gómez, J., Hofmann, P., Fitzek, F.H.P., Dressler, F.: Explainability of neural networks for symbol detection in molecular communication channels. IEEE Trans. Mol. Biol. Multi-Scale Commun. **9**(3), 323–328 (2023)
14. Farsad, N., Goldsmith, A.: Neural network detection of data sequences in communication systems. IEEE Trans. Signal Process. **66**(21), 5663–5678 (2018)
15. Jamali, V., Ahmadzadeh, A., Wicke, W., Noel, A., Schober, R.: Channel modeling for diffusive molecular communication-a tutorial review. Proc. IEEE **107**(7), 1256–1301 (2019)
16. Farahnak-Ghazani, M., Mirmohseni, M., Nasiri-Kenari, M.: Interference alignment using reaction in molecular interference channels. IEEE Trans. Nanobiosci. **22**(2), 259–267 (2023)
17. Kong, W., Dong, Z.Y., Jia, Y., Hill, D.J., Xu, Y., Zhang, Y.: Short-term residential load forecasting based on LSTM recurrent neural network. IEEE Trans. Smart Grid **10**(1), 841–851 (2019)