A Semantic Communication Model with Noise Reduction in a Cybernetic Avatar Teleoperation

Zie Eya Ekolle
Networks Research Institute
National Institute of Information and Communications Technology
Yokosuka, Japan

Homare Murakami
Networks Research Institute
National Institute of Information and Communications Technology
Yokosuka, Japan

Takeshi Matsumura
Networks Research Institute
National Institute of Information and Communications Technology
Yokosuka, Japan

***Abstract*—A Cybernetic Avatar (CA) teleoperation system is a robotic system designed to act as human surrogates in remote environments. This technology is an important innovation for achieving society 5.0 and industry 5.0. However, using conventional communication technologies, the communication between the teleoperator (TO) and CA will require more resources in terms of bandwidth usage and latency because the amount of data generated by the CA operation increases rapidly, especially in a multiple operators and avatars scenarios. This will lead to a communication bottleneck problem, especially in resource limited situations, where the bandwidth is not sufficient to efficiently transmit the large amount of data instantly. Many approaches have been proposed to address this challenge, but one promising approach is semantic communication, which consist of transmitting the meaning of data rather than the raw data. However, the presence of noise at the source of such communication drastically affects its reconstruction performance. In this paper, we proposed a semantic communication model with source noise reduction for a CA teleoperation system. We used this model to improve the performance of the denoising operation. Simulation results show an improvement on the reliability of the CA teleoperation in the presence of source noise.**

***Index Terms*—communication systems, semantic communication, noise reduction, denoising autoencoder, teleoperation system, cybernetic avatar.**

I. Introduction

In recent years, the need to advance the human society toward a state of freedom from the limitation of body, brain, space and time, has led to a shift in the focus toward Cybernetic Avatar (CA) development, because as explained in [1], CAs are considered as the future of the human society. CAs are robotic systems design to act as human surrogate in remote environments and can be used in many situations such as in disaster management, medical operations, and as assistance to the aging population.

Furthermore, with the advancement in 5G and beyond 5G (B5G) technologies, CA teleoperation systems will lead to an advancement in human civilization, where CA assistants will be present in almost all activities in human life to overcome our human limitations. More so, as explained in [2], the ultra-fast data rates, low latency, and massive network capacity for both real-time and near real-time communication using the 5G and B5G technologies will increase the availability and reliability of CA teleoperation services, creating a more seamless user experience.

However, as the number of CA and teleoperators (TO) increases on a given teleoperation scenario such as in [3], the amount of data generated to be transmitted between the CA and TO in real-time increases significantly, possibly leading to a communication bottleneck, where the bandwidth is not sufficient to efficiently transmit the large amount of data instantly. This is a crucial issue for CA teleoperation systems which depends on the reliability of real-time communication between the CA and TO. Many approaches have been proposed to address this issue but one promising approach is the use of semantic communication (SemCom) technology.

SemCom is a communication technique in which the meaning or intent of data rather than the raw data are transmitted. It drastically reduces the amount of data to be transmitted as described in [4], enabling the possibility of eliminating communication bottleneck in a resource limited and time-constrained situation, such as in a multiple CAs and TOs teleoperation. While the 5G and B5G technologies provide seamless connectivity, semantic communication technologies provide a resource economy to the CA and TO communication. A general architecture of semantic communication with source noise in a CA teleoperation is presented in Figure 1.

One of the major issue that can affect communication is noise at the source, which can corrupt the actual information to be communicated. Similar to conventional communication, noise introduce at the source of a SemCom model would affect the output at the receiving end of the the model, leading to a high reconstruction error. Therefore, the presence of corrupt information at the source would definitely lead to corrupt output if the noise is not eliminated or reduced. Reducing noise at the source is imperative in communication technology and it is generally done using denoising techniques.

Different denoising techniques have been developed for communication technology such as filtering based techniques (e.g., Median, Gaussian, Wiener, and Kalman filters), transform based techniques (e.g., Wavelet and Fourier transforms), and machine learning based techniques (e.g., CNNs and GANs). These denoising techniques reduce the source noise associated with the transmitted signal, enabling a denoised output at the receiving end of the communication. The major objective in all denoising techniques is to mitigate noise without distorting the original source signal. In this paper, we

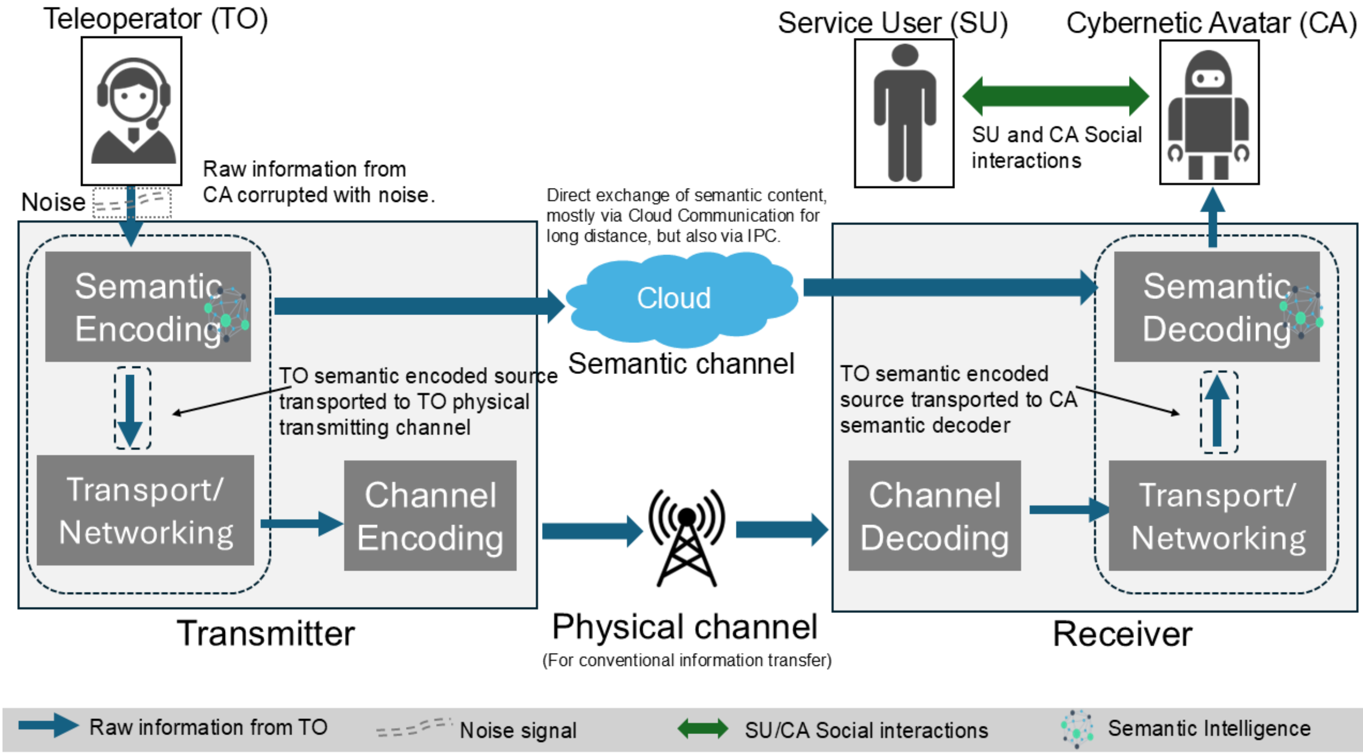

Fig. 1. General architecture of semantic communication with source noise in a CA teleoperation system

use a machine learning based denoising technique to reduce source noise in CA teleoperation.

In a CA teleoperation, an effective denoising technique is necessary because the CA teleoperation requires a reliable end-to-end (E2E) communication as discussed in [3]. To achieve this, we enhance the CA teleoperation with a SemCom model, where meaning must be consistently preserved and interpreted across the entire communication path—from the TO's command interface to the CA's actuators and sensors, and even further to the service user interaction. This makes the problem far more complex, requiring not only noise reduction but also issues of timing, reliability, cross-layer semantics alignment, and multi-agent context sharing.

In this paper, we address these challenges by proposing a semantic communication model for noise reduction in CA teleoperation. The proposed model operates at the application layer of the TCP/IP protocol stack and can support both actuation and sensing signals. Our contributions are as follows:

- A denoising SemCom model at the application layer tailored for CA teleopeartion.
- Simulation and evaluation of the denoising effect of the SemCom model on CA teleoperation.

The rest of the paper is divided as follows; Section II discusses about related works in denoising techhniques and CA teleoperation system. Section III discusses about our proposed model. Section IV is about experiment and results. We conclude in Section V.

II. Related works

There are many researches on denoising techniques but very few on SemCom denoising application to teleoperation system.

Gul et al. [5], present a denoising autoencoder (DAE) to enhance sensing reliability by mitigating the effects of abnormal sensing reports and noise disturbances at a Fusion Center. The autoencoder produces cleaned energy data, which is fed into a machine learning classifier to estimate channel availability and accumulate global decisions. They further perform performance analysis which shows that a combination of DAE and Ensemble Classifier yields high accuracy, F1-score, and Matthew's Correlation Coefficient.

Sabera et al. [6], propose a GAN-based architecture for denoising ECG signals. They demonstrated its efficacy in removing both white noise and motion artefacts from ECG data. Results from their experiment indicate that GANs can outperform conventional denoising methods, leading to clearer and more reliable ECG signals for further clinical analysis.

Ramadhan et al. [7], proposed a convolutional autoencoder with sequential and channel attention (CAE-SCA) to address the issue of noise and desired signal frequencies overlapping in ECG signals. Experimental results give an average SNR value of 16.187 dB, RMSE of 0.059, and PRD value of 18.529 in the MIT-BIH database. While in the SHDB-AF dataset, the model obtained 15.308 dB of SNR, 0.049 of RMSE, and 19.220 of PRD. These results demonstrate their CAE-SCA outperforms the selected state-of-the-art methods across different metrics.

Mahaseni et al. [8], proposed a method for reconstructing EEG signals using a variant of the variational autoencoder (VAE) called beta-VAE. They evaluated their proposed model on DEAP dataset and show that it learns a compressed representation of the EEG signal in an unsupervised manner, and the reconstructed signal contains less artifact. Results shows that their approach has a potential for improving the analysis and understanding of EEG signals in clinical and research settings.

Cho et al. [9], propose GoonDAE, a novel denoising-based real-time driver assistance method that enables stable teleoperated off-road driving. They introduce a denoising autoencoder (DAE) based on a skip-connected long short-term memory (LSTM) to assist the unskilled driver control input through denoising. Experiments in the simulated off-road environment show that GoonDAE significantly improves the driving stability of unskilled drivers.

Samarathunga et al. [10], analyze a framework for semantic image transmission that independently and jointly models semantic and technical noise while employing Turbo codes for robust error detection and correction. They systematically examine their individual and combined impacts through seven experimental scenarios and benchmark the results against state-of-the-art image codecs. The results show that joint noise modeling not only improves semantic fidelity, but also enhances overall system robustness, and that Turbo coding proves effective in mitigating both noise types.

While these works focus on the use of denoising techniques to different applications, they are limited in their use of SemCom for performance improvement of CA teleoperation.

III. Modeling

Our proposed model uses a SemCom to enhanced CA teleoperation, where the input signal from the TO is transmitted to the CA as a semantic encoded signal. This signal can represents the actuation or sensing signal of the TO for the CA to execute. Upon receiving the signal, the CA semantically decode them and execute the actions of the TO. If noise is introduce at the source to corrupt the input signal of the TO, this corrupted signal will be transmitted to the CA, affecting the desire output of the TO on the CA.

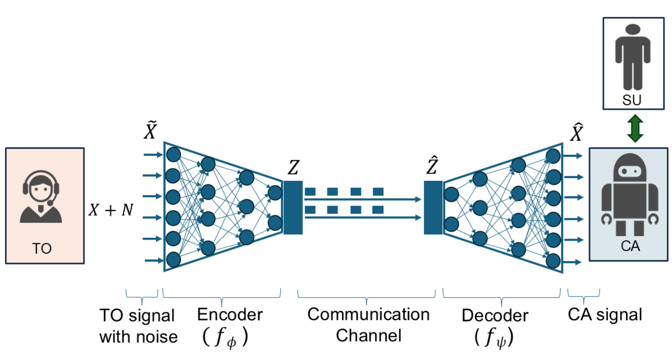

Fig. 2. Proposed architecture for a SemCom model with source noise

Figure 2 illustrate the SemCom model in a CA teleoperation system with source noise, N, and TO signal, X. The observable space, $\tilde{X}$, of the SemCom model representing the corrupt signal from TO. These are compressed to the latent space, Z, and transmitted via a communication channel to the CA. The CA then decode the transmitted semantic information to execute the actions of the TO. Assuming that the communication channel is perfectly noiseless with no information loss, then the transmitted semantic code Z at the transmitter will be equal to $\widehat{Z}$ at the receiver.

The objective of the SemCom model is to make sure that the input signal X and output signal $\widehat{X}$ are identical in the presence of noise N. To model the semantic encoder and decoder for this case study, we use a denoising autoencoder in which an encoder, f_ϕ, encode and compress $\tilde{X}$ to Z, which is then transmitted to the CA and receive at decoder, f_ψ, as $\widehat{Z}$. The decoder reconstruct X by generating $\widehat{X}$ from $\widehat{Z}$.

Mathematically, the encoder is represented as,

$$f_\phi : \tilde{X} \rightarrow Z \tag{1}$$

$$Z = f_\phi(\tilde{X}) \tag{2}$$

where $\tilde{X} = X + N$, $N \sim \mathcal{N}(0, \sigma^2 I)$, f_ϕ is the encoding function with parameter ϕ, $X \in \mathbb{R}^n$ represent the signals of the TO, $Z \in \mathbb{R}^m$ represented the encoded semantic data, and $n > m$ with $n, m \in \mathbb{N}$.

Similarly, the decoder is represented as,

$$f_\psi : \widehat{Z} \rightarrow \widehat{X} \tag{3}$$

$$\widehat{X} = f_\psi(\widehat{Z}) \tag{4}$$

where f_ψ is the decoding function with parameter ψ, $\widehat{Z} \in \mathbb{R}^m$ represented the transmitted semantic code, $\widehat{X} \in \mathbb{R}^n$ represent the reconstructed output of the decoder, and $n > m$ with $n, m \in \mathbb{N}$.

In this paper, we consider a perfectly noiseless channel, hence $Z = \widehat{Z}$. However, in a noisy channel where most likely $Z \neq \widehat{Z}$, the amount of channel noise would further affect the reconstruction performance of the model as discussed in [4].

The introduction of noise at the source will definitely also affect the reconstruction performance and decoded output of the SemCom model. The objective of this proposed SemCom model is to reduce the reconstruction error due to source noise during teleoperation training. This is because while the optimization of the SemCom model will improve its reconstruction performance, any noise introduce at the source can affect this reconstruction performance.

The optimization strategy consist of searching the optimal parameters of the encoder and decoder that minimize the error in the SemCom such that $\widehat{X} = X$. In addition to source noise, any difference between $\widehat{X}$ and X is considered to be caused by a reconstruction error in the model. This error is generally measured using a loss function defined over $\widehat{X}$ and X to optimize the parameters of the encoder and decoder. Reducing this error is essential in improving the reconstruction accuracy.

Different loss functions can be used to model the reconstruction error such as the mean square error, binary cross entropy, cosine loss, and so on. As shown in Figure 3, binary

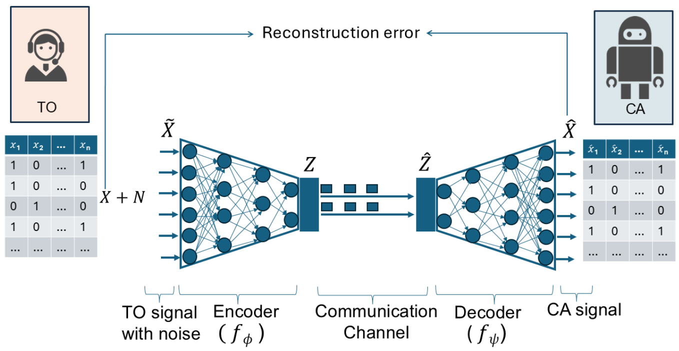

Fig. 3. SemCom with binary signals at the TO and CA

data with additive white Gaussian noise (AWGN) were used to represent the corrupted TO signals. Further more, to account for the occurrence of sparsity in the data, we use cosine loss. The cosine loss measures the dissimilarity between vectors and it is expressed with respect to our case study as follows,

$$\mathcal{L}(X, \widehat{X}) = \quad 1 - \frac{X \cdot f_\psi(f_\phi(X))}{|X||f_\psi(f_\phi(X))|} \tag{5}$$

where $\mathcal{L}(X, \widehat{X}) \in [0, 1]$ is the cosine loss function, X is the source information, $f_\psi(f_\phi(X)) = \widehat{X}$ is the reconstructed information, ϕ represents the parameters of the encoder, and ϕ represents the parameters of the decoder.

Optimizing the model to reduce error entails minimizing $\mathcal{L}(X, \widehat{X})$. We achieve this by searching the optimal parameters of the encoder and decoder that minimize $\mathcal{L}(X, \widehat{X})$, thereby increasing the similarity between X and $\widehat{X}$. The optimization process of $\mathcal{L}(X, \widehat{X})$ is expressed as follows,

$$\min_{\phi,\psi} \mathcal{L}(X, \widehat{X}) = \min_{\phi,\psi} \quad 1 - \frac{X \cdot f_\psi(f_\phi(X))}{|X||f_\psi(f_\phi(X))|} \tag{6}$$

$$\text{s.t.} \quad \widehat{X} = X \tag{7}$$

It is worth noting that the loss function does not explicitly take into account the corrupted TO signal $\tilde{X}$ but instead the uncorrupted signal X. This implies that the model is trained to reduce the source noise by adapting its effects on the parameters of the decoder through optimization of the reconstruction error.

An important aspect of this optimization is that unlike in most machine learning models where over-fitting of the model to the data is avoided, in the case of CA teleoperation where all the commands and their respective binary codes are known and well defined, over-fitting becomes an advantage to the model. In this case, the optimal parameters will be those at $\mathcal{L}(X, \widehat{X}) = 0$. However, if some unknown instances of the commands exist and can be initiated by the TO, such as new scenario, then generalization of the model over the input space is essential, as it is done in most autoencoder optimization and machine learning processes.

Evaluating the reliability in performance of the SemCom model is an important aspect in our research. Since the noise is reduced by adapting the decoder parameters to reduce the reconstruction error, we focus more on evaluating the noise reduction performance of the model, specifically the reconstruction performance of the decoder. The encoding and communication performances of the SemCom model where not considered in this paper. Hence, the compression performance of the model such as compression size, ratio, gain, and factor were not considered.

The denoising performance metrics of the model are group into two main categories; (1) evaluation metrics useful for assessing the model's ability to correctly classify each data point as "original signal" (positive) or "noise" (negative) and, (2) evaluation metrics that focus on the quality of the denoised signal compared to the original, clean data.

Regarding the performance metrics useful for assessing the model's ability to correctly classify each data point, we considered the denoising accuracy, precision, recall, and F-score as define below.

$$\text{Denosing accuracy} \triangleq \frac{\text{TP} + \text{FP}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \tag{8}$$

$$\text{Denoising precision} \triangleq \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{9}$$

$$\text{Denoising recall} \triangleq \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{10}$$

$$\text{Denoising F-score} \triangleq \frac{2(\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{11}$$

where TP is True positive, FP is False positive (Type I Error), TN is True negative, and FN is False negative (Type II Error).

Regarding the performance metrics useful for assessing the quality of the denoised signal compared to the clean data, we mainly considered the cosine loss. For performance comparative analysis, we consider other metrics Signal-to-Noise-Ratio (SNR), SNR improvement (SNRI), Mean square error (MSE), and Binary cross entropy (BCE) as define below.

$$\text{SNR} \triangleq \frac{1}{\sigma^2} \quad \text{(normalized linear signal power)} \tag{12}$$

$$\text{SNRI} \triangleq \frac{\text{SNR}_{\text{denoised}}}{\text{SNR}_{\text{noisy}}} \tag{13}$$

$$\text{MSE} \triangleq \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{14}$$

$$\text{BCE} \triangleq -\frac{1}{n} \sum_{i=1}^{n} (y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)) \tag{15}$$

where σ^2 is noise variance, y is the clean signal, $\hat{y}$ is the denoised signal, and n is the number of samples.

More so, apart from model performance metrics, system performance metrics of the denoising operation such as the denoising time, size (space), and rate are considered and defined below. The denoising time and space captures the total system time and memory used for the denoising operation. For simplicity, only the memory associated to the datasize of the signals was considered in this paper.

$$\text{Denosing time (T)} \triangleq \sum_{i=1}^{i=n} t'_i - t_i \tag{16}$$

$$\text{Denosing size (S)} \triangleq \sum_{i=1}^{i=n} s_i \tag{17}$$

$$\text{Denoising rate} \triangleq \frac{S}{T} \tag{18}$$

where n is number of samples, t is time before denoising a sample, t' is time after denoising a sample, T is total denoising time, s is data size per sample, S is total datasize.

Depending on the performance expectations of the experiment and characteristics of the denoising model, other performance metrics such as Structural Similarity Index (SSIM), Correlation Coefficient, Peak Signal-to-Noise Ratio (PSNR), Root Mean Square Error (RMSE), and Percent Root Mean Square Difference (PRD) can be used.

IV. Experiment and Result

The proposed denoising SemCom model was trained and tested on a simulated CA teleoperation system with a Gaussian source-noise. Backpropagation training technique with stochastic gradient descent (SGD) optimization were used to minimize the cosine loss in the model, thereby reducing the reconstruction error.

A. Experiment description

The experiment consists of a simulated TO and CA, logically connected via a simulated application layer communication link, using a publisher-subscriber communication model, mainly MQTT protocol. MQTT was used due to its adaptation to resource-constrained devices and networks, such as those associated with the TO and CA sensing, communication, and control operations.

An autoencoder architecture was used to model the proposed SemCom, and the simulation parameters of the model are presented in Table I. The data was generated from the simulated TO teleoperation, then an AWGN was introduce to corrupt the data before being encoded by the SemCom model and transmitted to the CA.

The results of the simulation together with performance evaluation are presented in the next section. Comparitive performance analysis were done with conventional binary denoising techniques. The conventional binary denoising techniques used include Median filter (MF), Least Mean Squares (LMS) fitler, Denoising Generative Adverserial Network (DnGAN), and Binary Discrete Fourier Transform (BDFT).

B. Results and Explanation

After training the proposed denoising SemCom model with data generated from the simulated CA teleoperation system, we evaluate the reconstruction error using cosine loss for different test scenarios with specific noise variances as shown in Figure 4. From the experiment, a higher noise value produces a higher reconstruction error, which indicates a low performance value of the denoising operation and model generalization.

TABLE I
Simulation Parameters

Encoder/Decoder	
Model type	DNN
Observation dimension	15
Latent dimension	5
Number of hidden layers	3
Symbol rate (symbol/sec)	5 Bd
Optimization logic	SGD
Loss function	Cosine loss
Learning rate	0.001
Learning epoch	50
Learning approach	Back-prop
Source noise model	AWGN
Source noise intensity (variance)	[0,0.1]
Communication channel	
Model type	Noiseless
Protocol model	Pub-Sub
Protocol standard	MQTT
Capacity per channel (bit/symbol)	1
Packet loss probability per channel	0
Number of parallel channels	5
Dataset	
Total observation instances	1400
Training instances	500
Test instances per test scenario	300
Number of test scenario	3
Observation features	15
Total transmission instances	1400
Packets per transmission	5
Semantic features per transmission	5
Encoded label per transmission	1

While the model generalization performance can be improved using techniques such as cross-validation and regularization during training, improving the denoising performance requires further minimization of the reconstruction error, which can lead to overfitting. Hence, a trade-off arises between improving denoising performance and reducing overfitting. We simply prioritize the former in this paper.

Furthermore, assuming a normalized signal power, we compare the quality of the denoised signal from our model to that of some conventional binary denoisers techniques using the denoised SNR (DnSNR), SNRI, MSE, and BCE metrics for a noisy input with a variance of 0.05 and corresponding denoised output as shown in Figure 5. In the figure, our SemCom model has a denoised SNR of 110 better than the other models.

Lastly, on Table II, we compare the performance of our denoising model with conventional binary denoising techniques using all the metrics presented in section III. As shown on Table II, our model has a DnSNR of 110, SNRI of 5.5, and denoising accuracy of 94.5%, better than the other models.

V. Conclusion

In this paper, we proposed and simulated a SemCom model with source noise reduction in a CA teleoperation. We perform experiments and evaluate its performance with respect to conventional binary denoising techniques. The results show that the introduction of noise highly impact the performance of CA teleoperation. However, our proposed SemCom model shows a DnSNR of 110, SNRI of 5.5, denoising accuracy of 94.5%, and F-score 91.6% better than the other models.

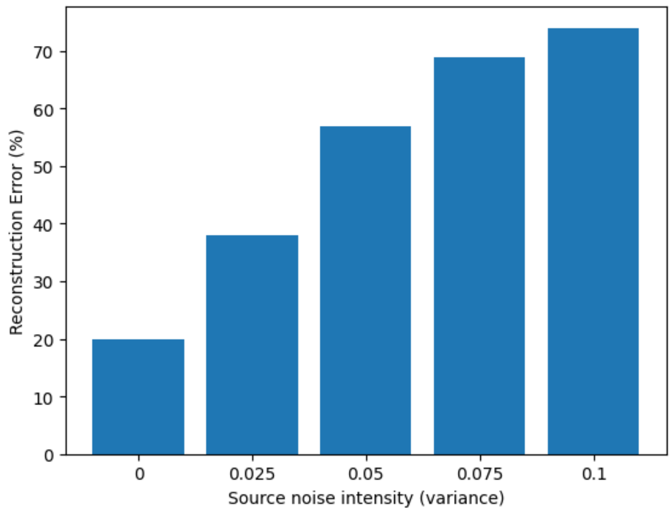

Fig. 4. Reconstruction errors of our model for different noise variance

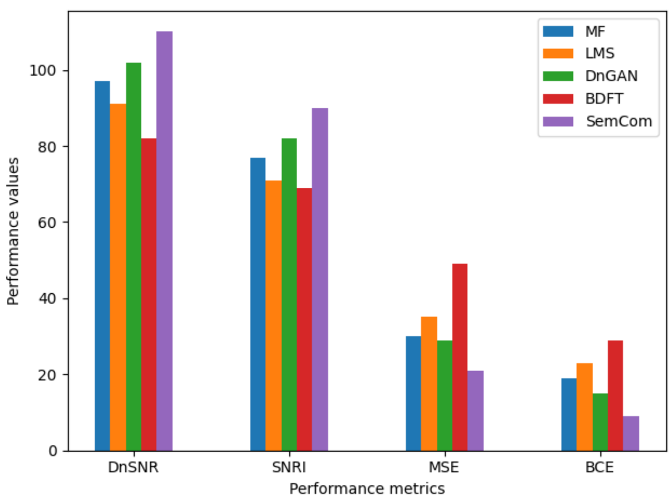

Fig. 5. Performance comparison of different denoising techniques based on denoised SNR, SNRI, MSE, and BCE

TABLE II
PERFORMANCE COMPARISON OF DIFFERENT MODELS

Performance metrics	**MF**	**LMS**	**DnGAN**	**BDFT**	**SemCom**
DnSNR	97	91	102	89	110
SNRI	4.85	4.55	5.1	4.45	5.5
MSE	30	35	29	49	21
BCE	19	23	15	29	9
Denoising accuracy	84%	82%	87.9%	81%	94.5%
Denoising precision	86.1%	85%	89%	80.4%	92%
Denoising recall	87.8%	86.1%	88%	82%	91.1%
Denoising F-score	86.9%	85.5%	88.5%	81.2%	91.6%
Denoising time	2 sec	3 sec	10 sec	6 sec	4 sec
Denoising size	5kb	5kb	5kb	5kb	5kb
Denoising rate	2.5kbps	1.7kbps	500bps	833bps	1.25kbps

ACKNOWLEDGMENT

A part of this research has been supported by JST [Moonshot R&D][Grant Number JPMJMS2216].

REFERENCES

[1] N. Hagita, et al., "Cybernetic avatars: Teleoperation technologies from in-body monitoring to social interaction," Sci. Robot., vol. 9, no. 96, eadg1842, Nov. 2024.

[2] H. Murakami, et al., "Towards Stable Teleoperation of Cybernetic Avatars: Design and Evaluation of a Hybrid Traffic Management Platform," IEEE GLOBECOM 2025.

[3] T. Matsumura, H. Murakami, A. Wakayama and K. Ibuka, "Toward Reliable End-to-End Communication for Global Cybernetic Avatar Teleoperation," 2025 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN), London, United Kingdom, pp. 1-6, 2025.

[4] Z. E. Ekolle, H. Murakami and T. Matsumura, "A Goal-Oriented Semantic Communication for A Reliable Cybernetic Avatar Teleoperation System," 2025 IEEE Future Networks World Forum (FNWF), Bangalore, India, pp. 1-6, 2025.

[5] Gul N., Kim S.M., Akbar S., Elahi A., Kim J., "Enhanced Sensing Performance through the Integration of Denoising Autoencoder and Ensembling Techniques", Scientific Reports, vol. 16, 535, 2026.

[6] G. Sabera, S. H. Naga Ketan and M. Mohebbanaaz, "Generative Adversarial Networks for Noise Reduction in Electrocardiogram Signals," 2025 IEEE 14th International Conference on Communication Systems and Network Technologies (CSNT), Bhopal, India, pp. 552-556, 2025.

[7] A. W. Ramadhan et al., "Convolutional Autoencoder With Sequential and Channel Attention for Robust ECG Signal Denoising With Edge Device Implementation," IEEE Access, vol. 13, pp. 54407-54422, 2025.

[8] B. Mahaseni and N. M. Khan, "EEG Signal Denoising Using Beta-Variational Autoencoder," 2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, pp. 1-4, 2024.

[9] Y. Cho, H. Yun, J. Lee, A. Ha and J. Yun, "GoonDAE: Denoising-Based Driver Assistance for Off-Road Teleoperation," IEEE Robotics and Automation Letters, vol. 8, no. 4, pp. 2405-2412, 2023.

[10] P. Samarathunga et al., "Semantic and Technical Noise Modeling in Semantic Image Communication Over Wireless Channels," IEEE Access, vol. 13, pp. 150872-150900, 2025.