# Diffusion-Based Semantic Encoders for Secure and Efficient Edge Computing

Xue Qin, *Graduate Student Member, IEEE*, Lingshuang Liu, *Student Member, IEEE*, and Xuemin Shen, *Fellow, IEEE*

***Abstract*—Mobile Edge Computing (MEC) enables resource-constrained Internet of Things (IoT) devices to offload computation-intensive workloads to nearby edge servers, reducing latency and energy consumption. However, wireless offloading exposes transmitted data to eavesdropping, raising serious privacy concerns. Existing physical-layer security techniques either introduce additional overhead or fail to conceal semantic information. This paper proposes a diffusion-based semantic encoding framework for secure and efficient computation offloading in MEC systems. By applying a forward diffusion process, task inputs are transformed into approximately noise-like latent representations whose distribution is statistically close to a standard Gaussian distribution prior to transmission. Authorized edge servers equipped with learned reverse diffusion models can reliably recover task-relevant semantics, while intercepted representations reveal negligible information. We analyze the secrecy and robustness of diffusion-based encoding from an information-theoretic perspective and integrate the proposed encoder-decoder into a practical MEC offloading pipeline. Extensive experiments under realistic wireless conditions demonstrate that diffusion-based encoding significantly improves secrecy, reduces transmission overhead, and maintains high task performance. Compared with autoencoder and variational autoencoder baselines, the proposed approach offers substantially stronger resistance to reconstruction and inference attacks while preserving computational efficiency at the edge.**

***Index Terms*—Mobile edge computing, secure computing, task offloading, diffusion.**

## I. Introduction

THE rapid proliferation of the Internet of Things (IoT) has accelerated the demand for real-time data processing in various domains such as healthcare, transportation, and smart manufacturing. These applications frequently involve computationally intensive tasks, including image classification, sensor analytics, and autonomous control, that exceed the processing, memory, and battery capacities of typical IoT devices [1], [2]. Mobile Edge Computing (MEC) has emerged as a key enabler for next-generation intelligent IoT systems by relocating computation from distant cloud servers to edge nodes located near end devices [3], [4], [5]. This proximity-aware computing paradigm greatly mitigates the limitations of resource-constrained IoT hardware by allowing latency-sensitive and computation-intensive workloads to be offloaded to nearby edge servers. As a result, MEC significantly reduces end-to-end service delay, alleviates backbone traffic congestion, and improves energy efficiency by minimizing the computation burden on battery-powered devices. These benefits make MEC indispensable for emerging applications such as real-time perception, smart healthcare, and autonomous control, where responsiveness, reliability, and energy awareness are simultaneously required.

To support such stringent latency and energy requirements in dynamic wireless environments, recent MEC systems increasingly rely on machine learning techniques for computation offloading and resource management. Deep learning [6], [7] models have been used to approximate complex offloading and scheduling policies, while reinforcement learning [8], [9], [10], [11] enables autonomous decision making under time-varying channel conditions, workload fluctuations, and heterogeneous device capabilities. These learning-driven approaches significantly improve the system's adaptability and performance, but they primarily focus on optimizing latency, energy consumption, and throughput. The security implications of data offloaded over wireless channels, particularly the risk of information leakage, remain largely overlooked.

Secure computation offloading in MEC remains a major challenge. Task data must be transmitted over open and broadcast wireless channels, making them inherently vulnerable to intercept. A passive eavesdropper can capture raw features, sensor measurements, or even privacy-sensitive content such as medical images and personal identifiers, thus compromising user privacy and undermining the adoption of MEC-enabled services [12]. Existing physical-layer security (PLS) [13], [14] techniques, including friendly jamming, artificial-noise generation, and lightweight cryptography, provide partial mitigation but are constrained by practical limitations [15]. Friendly jamming consumes additional spectral and energy resources [16] and must be optimally coordinated to avoid harming legitimate receivers, while lightweight cryptography [17] still exposes feature-level structure and may remain too costly for highly resource-limited IoT devices [18]. Moreover, these mechanisms typically do not address the deeper issue that offloaded data, even if encrypted, may still reveal semantic information through statistical patterns exploitable by advanced inference attacks.

These limitations motivate the exploration of *generative semantic encoding* [19], [20] as a new paradigm for enhancing security in MEC systems. Recent advances in generative modeling enable high-dimensional data to be transformed into latent representations [21] that resemble noise-like signals,

Received 21 February 2026; revised 12 June 2026; accepted 22 July 2026. Date of publication 29 July 2026; date of current version 6 August 2026. The associate editor coordinating the review of this article and approving it for publication was C. Wu. *(Corresponding author: Xue Qin.)*

The authors are with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: x7qin@uwaterloo.ca; lingshuang.liu@uwaterloo.ca; sshen@uwaterloo.ca).

Digital Object Identifier 10.1109/TCCN.2026.3717940

making semantic content extremely difficult to recover without the appropriate decoder [22]. Diffusion models [23], [24], in particular, offer strong privacy-preserving characteristics: their forward process gradually perturbs inputs into Gaussian noise, and the resulting latent variables are statistically close to an isotropic noise distribution [25]. This property suggests that transmitting diffusion-generated latent vectors can effectively obscure sensitive information while preserving the features required for downstream computation. Autoencoder (AE) and variational autoencoder (VAE) encoders, although less expressive than diffusion models, provide meaningful baselines for analyzing the trade-off between privacy protection and task utility.

Recent studies have begun to explore diffusion models in the context of semantic communication and representation learning, where diffusion-based encoders are leveraged to improve robustness, generative fidelity, or task-oriented transmission efficiency [26], [27], [28]. In these works, diffusion processes are primarily used to enhance semantic reconstruction quality or to support end-to-end learning under noisy channels. However, existing diffusion-based semantic communication schemes largely focus on reliability and task performance, with limited attention to adversarial privacy, wiretapping resistance, or secure computation offloading in MEC environments. In particular, the potential of diffusion models to produce noise-like, statistically close to Gaussian-like latent representations for intrinsic privacy protection remains underexplored. Different from existing diffusion-based communication frameworks that primarily aim to improve semantic reconstruction quality or transmission reliability, this work leverages diffusion as a semantic obfuscation mechanism for secure MEC offloading. Specifically, diffusion is performed in the latent semantic space to suppress recoverable semantic information before transmission while preserving downstream task utility at authorized edge servers.

In this paper, we develop a diffusion-based semantic encoding mechanism as the foundation for secure and efficient computation offloading in MEC. Our main contributions are summarized as follows:

1) **Diffusion-based semantic encoder for secure edge offloading:** We design a lightweight diffusion-driven encoder that transforms task data into noise-like latent vectors whose statistical distribution closely approximates Gaussian noise, thereby offering strong privacy guarantees during wireless transmission. Unlike existing diffusion-based communication frameworks that primarily improve reconstruction fidelity or transmission reliability, the proposed approach employs diffusion as a semantic-level privacy protection mechanism for secure MEC offloading. Unlike conventional PLS techniques that primarily protect the wireless signal, the proposed approach directly protects task-relevant semantic information at the representation level.
2) **Semantic privacy analysis:** We analyze the privacy properties of diffusion-generated latent vectors using information-theoretic and adversarial reconstruction perspectives, quantifying the reduction in recoverable information under various wiretapping and inference attack models.
3) **Extensive performance evaluation and comparison:** We implement both AE and VAE encoders as baseline models to contextualize the security utility improvements achieved by diffusion-based semantic encoding. Through extensive simulations under realistic MEC configurations, we demonstrate that the proposed diffusion-based encoder enhances secrecy capacity, reduces transmission overhead, and maintains high task completion rates, outperforming conventional PLS techniques and generative baselines.

The remainder of this paper is organized as follows. Section II describes the system model and threat model, followed by problem formulation in III. Section IV presents the proposed diffusion-based semantic encoding framework for secure MEC offloading. Section V conducts the security analysis. Section VI provides experimental results and analysis. Finally, Section VII concludes this work.

## II. System Model

We consider a MEC system consisting of multiple edge servers in which a set of resource-constrained IoT or end devices offload computation-intensive tasks to nearby edge servers via wireless links, as shown in Fig. 1.

### A. Task Offloading Model

Each device generates a task

$$\Gamma_i = \{x_i, c_i, T_i^{\max}\},$$

where $x_i$ denotes the raw input (e.g., an image or sensor tensor), $c_i$ represents the required CPU cycles, and $T_i^{\max}$ is the deadline. Due to their limited computational capacity, end devices can offload tasks to one of $M$ edge servers, indexed by $j \in \mathcal{M}$, each with a maximum CPU frequency $f_j^{\max}$. To reduce transmission overhead and improve privacy, devices follow a *semantic offloading* paradigm in which only a latent representation of the task is transmitted. We consider task-oriented semantic offloading for supervised classification, where each input $x_i$ is associated with a discrete label $y_i \in \{1, \ldots, C\}$, and the edge server performs inference based on the decoded latent representation.

Before offloading, each IoT device encodes $x_i$ into a latent vector $z_i$ using a semantic encoder $\mathcal{E}(\cdot)$. The primary focus of this work is a diffusion-based encoder, while AE and VAE encoders are implemented solely as comparison baselines. The diffusion encoder generates a noise-like latent through a forward stochastic corruption process:

$$z_{i,t} = \sqrt{\bar{\alpha}_t}\, x_i + \sqrt{1-\bar{\alpha}_t}\, \epsilon, \qquad \epsilon \sim \mathcal{N}(0, I), \tag{1}$$

where $\bar{\alpha}_t$ is the cumulative noise schedule. After $T$ diffusion steps, $z_i = z_{i,T}$ becomes approximately noise-like and statistically close to Gaussian noise, thereby suppressing recoverable semantic information while preserving limited task-relevant features for reliable edge inference.

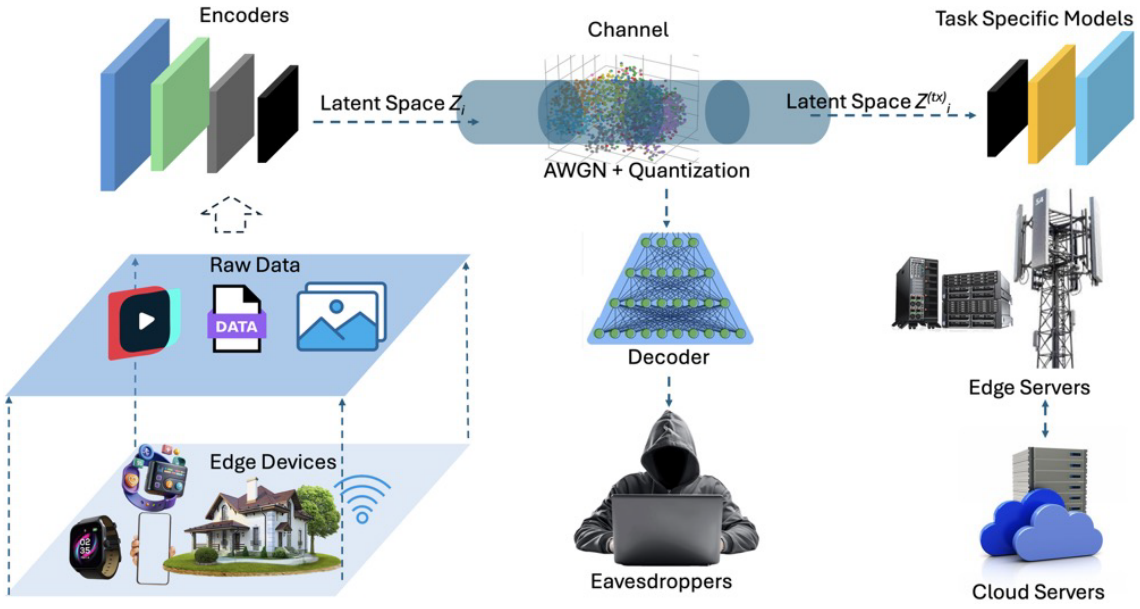



Fig. 1. System model.

For the baselines, the VAE encoder produces stochastic latents

$$z_i \sim \mathcal{N}\big(\mu(x_i), \sigma^2(x_i) I\big), \tag{2}$$

while the AE produces a deterministic embedding $z_i = f_\theta(x_i)$. Although both baselines provide compression, their latent representations typically preserve more semantic structure than diffusion-based latents.

To ensure consistent evaluation across all encoders, every latent vector, regardless of whether it originates from diffusion, VAE, or AE, is transmitted through an identical noisy and quantized communication channel. Specifically, the transmitted representation is

$$z_i^{(\mathrm{tx})} = Q(z_i + \eta_i), \qquad \eta_i \sim \mathcal{N}(0, \sigma_c^2 I), \tag{3}$$

where $Q(\cdot)$ denotes uniform quantization implemented with a straight-through estimator. This step models realistic bandwidth limitations and ensures that all encoder types experience the same channel perturbations.

During offloading, the edge server $j$ receives

$$y_{ij} = h_{ij}\sqrt{P_i}\, z_i^{(\mathrm{tx})} + n_{ij}, \tag{4}$$

where $P_i$ is the transmit power, $h_{ij}$ is the block-fading coefficient, and $n_{ij} \sim \mathcal{N}(0, N_0)$ is receiver noise. The achievable transmission rate is

$$r_{ij} = B \log_2\left(1 + \frac{P_i |h_{ij}|^2}{N_0}\right), \tag{5}$$

where $B$ is the allocated bandwidth. The transmission latency is therefore

$$T_{ij}^{\mathrm{tx}} = \frac{|z_i^{(\mathrm{tx})}|}{r_{ij}}, \tag{6}$$

where $|z_i^{(\mathrm{tx})}|$ denotes the encoded representation size. Diffusion-based encoders typically achieve high compressibility due to their latent structure, reducing transmission overhead.

Upon reception, the server decodes the noisy latent using the corresponding decoder $\mathcal{D}(\cdot)$ to reconstruct only task-relevant features, not the original input. The computation time is given by

$$T_{ij}^{\mathrm{comp}} = \frac{c_i}{f_j}, \tag{7}$$

with $0 < f_j \le f_j^{\max}$. $f_j$ is the CPU frequency allocated to the task at server $j$, and $f_j^{\max}$ is the maximum CPU frequency of edge server $j$.

The total task completion time is calculated as

$$T_i^{\mathrm{tot}} = T_{ij}^{\mathrm{tx}} + T_{ij}^{\mathrm{comp}}. \tag{8}$$

A task is successfully completed if

$$T_i^{\mathrm{tot}} \le T_i^{\max}. \tag{9}$$

We assume a passive eavesdropper observing the same wireless medium. Its received signal is

$$y_{ie} = h_{ie}\sqrt{P_i}\, z_i^{(\mathrm{tx})} + n_{ie}, \tag{10}$$

where $h_{ie}$ denotes the eavesdropper's channel gain, $z_i^{(\mathrm{tx})} = Q(z_i + \eta_i)$ is the transmitted latent space, and $n_{ie} = \mathcal{N}(0, N_0)$ is receiver noise. Let $z_i^{(\mathrm{eav})}$ denote the effective latent representation inferred by the eavesdropper from $y_{ie}$, obtained through any feasible signal processing or inference strategy available to the adversary. The adversary aims to reconstruct the original input $x_i$ or infer sensitive semantic attributes from $z_i^{(\mathrm{eav})}$.

Unlike conventional deterministic or lightly stochastic encoders, the proposed diffusion-based encoder applies a *finite-step forward diffusion process* that progressively injects noise into the latent representation. As the diffusion step index increases, the marginal distribution of the latent variable approaches an isotropic Gaussian. In practice, the diffusion process is terminated at a task-aware intermediate step $T$, yielding a latent representation that is *approximately noise-like* while still preserving sufficient information for recovery

by an authorized edge server equipped with the learned reverse diffusion model. In practice, $z_i^{(\mathrm{eav})}$ is approximated using a trained latent classifier or inversion network, providing an empirical lower bound on $I(x_i; z_i^{(\mathrm{eav})})$.

From the eavesdropper's perspective, which lacks access to the decoder $\mathcal{D}(\cdot)$ and the reverse diffusion dynamics, the inferred latent $z_i^{(\mathrm{eav})}$ appears statistically similar to a sample drawn from a high-entropy Gaussian distribution. Consequently, the semantic information leaked through passive interception is strongly suppressed, which can be characterized by

$$I(x_i; z_i^{(\mathrm{eav})}) \ll I(x_i; z_i^{(\mathrm{legit})}), \tag{11}$$

where $z_i^{(\mathrm{legit})}$ denotes the latent representation available to the authorized edge server prior to reverse diffusion decoding. This design achieves a principled trade-off between privacy and utility: sufficient corruption to prevent unauthorized inference, while maintaining reliable task recovery for legitimate receivers.

To quantify security, we evaluate three standard secrecy metrics: (i) *mutual information leakage* between the original input and the eavesdropper's observation,

$$I(x_i; y_{ie}) = \mathbb{E}_{p(x_i, y_{ie})}\left[\log \frac{p(y_{ie} \mid x_i)}{p(y_{ie})}\right], \tag{12}$$

which quantifies the amount of semantic information about $x_i$ recoverable from the intercepted signal $y_{ie}$. (ii) reconstruction error $\mathbb{E}\|x_i - \hat{x}_i^{(e)}\|^2$ at the eavesdropper, (iii) secrecy capacity, defined as the difference between legitimate and adversarial channel capacities. Diffusion-based semantic encoding improves all three metrics compared with AE, VAE, and conventional physical-layer security techniques, owing to its intrinsically noise-like latent representations and robustness to inversion.

### B. Threat Model

We consider a passive yet capable adversary whose goal is to compromise the privacy of offloaded data in the MEC system. The adversary does not manipulate transmissions, scheduling, or computation, but can overhear all uplink wireless communications from IoT devices. This aligns with widely studied eavesdropping scenarios in open wireless environments, where attackers may not disrupt service but can freely monitor the channel.

During offloading, the adversary intercepts the same noisy and quantized latent representation transmitted over the wireless channel. Specifically, the eavesdropper observes the received signal defined Eq. (10), corresponding to the transmission of $z_i^{(\mathrm{tx})} = Q(z_i + \eta_i)$ over a fading channel. The adversary's objective is to reconstruct the original input $x_i$ or infer sensitive semantic attributes from the intercepted representation. This includes reconstruction attacks based on generative modeling, feature-inference attacks, and classification attacks commonly considered in privacy-sensitive machine learning systems.

We assume a strong *white-box* adversary with full knowledge of the MEC architecture, encoder model classes, communication pipeline, and training procedures. However, the adversary does *not* have access to the decoder parameters $\mathcal{D}(\cdot)$ stored securely on authorized edge servers. Consequently, recovering $x_i$ from $z_i^{(\mathrm{tx})}$ requires solving an ill-posed inverse problem without access to the reverse diffusion dynamics or decoder network. For conventional AE and VAE baselines, the latent representations retain structured semantic features that may assist an adversary in partial reconstruction or attribute inference. In contrast, diffusion-based semantic encoding applies a finite-step forward corruption process that progressively increases the entropy of the latent representation, yielding latents that are *approximately noise-like* prior to transmission. After channel noise and quantization, the intercepted signal becomes even more weakly informative, further suppressing exploitable structure. Accordingly, the design objective of diffusion-based encoding is to significantly reduce the mutual information between the task input and the eavesdropper's observation. Following the latent-space privacy criterion in Eq. (11), and noting that both $y_{ij}$ and $y_{ie}$ are noisy transformations of the corresponding latent representations, the data-processing inequality implies that the same ordering holds at the signal level, i.e., $I(x_i; y_{ie}) \ll I(x_i; y_{ij})$, where $y_{ij}$ denotes the signal received by the legitimate edge server. Formal information-theoretic analysis and empirical validation of this property are provided in Sections V and VI, respectively. We do not assume the presence of cryptographic keys, cooperative jamming, or additional physical-layer security mechanisms; such enhancements are outside the scope of the core threat model. The adversary is restricted to passive observation and cannot alter or inject packets, consistent with standard assumptions in secure MEC and wireless privacy research.

## III. Problem Formulation

We now formalize the secure semantic offloading problem in the MEC system described earlier. Each IoT device offloads a computation task $\Gamma_i = \{x_i, c_i, T_i^{\max}\}$ by transmitting an encoded latent representation $z_i$ to an edge server. The system performance is characterized by three coupled aspects: end-to-end latency, transmission energy consumption, and semantic privacy against eavesdroppers. In this work, we do not perform joint resource allocation or scheduling; instead, we treat the choice of semantic encoder as the primary design variable and evaluate the induced trade-offs among these metrics under fixed MEC parameters.

Under the system model in Section II, the end-to-end latency of task $\Gamma_i$ is given by the sum of transmission and computation delays, which depends on the size of the transmitted noisy, quantized latent representation $|z_i^{(\mathrm{tx})}|$, the achievable transmission rate $r_{ij}$, and the allocated CPU frequency $f_j$. Since $r_{ij}$ and $f_j$ are fixed in our evaluation, different encoders primarily influence latency through $|z_i^{(\mathrm{tx})}|$.

### A. Energy Consumption Model and Trade-Off Discussion

The device-side energy consumption is dominated by the transmission energy associated with sending the latent representation:

$$E_i^{\mathrm{tx}} = P_i T_{ij}^{\mathrm{tx}} = P_i \frac{|z_i^{(\mathrm{tx})}|}{r_{ij}}, \tag{13}$$

where $P_i$ is the transmission power. For fixed $P_i$ and $r_{ij}$, both transmission energy and latency scale linearly with the effective size of the transmitted latent representation $|z_i^{(\mathrm{tx})}|$. As a result, reducing the latent size simultaneously lowers transmission delay and energy consumption.

However, aggressive compression or noise injection may degrade the utility of the transmitted representation and adversely affect task accuracy. From a privacy perspective, stronger semantic obfuscation-achieved through stochastic encoding or increased corruption-reduces information leakage to eavesdroppers but may increase the difficulty of recovery at the edge server. Conventional AE and VAE encoders therefore face a trade-off between compactness, robustness, and privacy: reducing latency and energy via smaller latent representations often comes at the cost of weaker semantic privacy or degraded task performance.

Diffusion-based semantic encoding mitigates this trade-off by decoupling privacy from compactness. The forward diffusion process increases the entropy of the latent representation and suppresses semantic structure without requiring excessive dimensionality reduction. As a result, diffusion-based encoders can achieve low latency and energy consumption through compact latents while simultaneously providing strong semantic privacy and maintaining high task accuracy. This three-way trade-off among latency, energy efficiency, and privacy is quantitatively evaluated in Section VI.

### B. Semantic Privacy Model

To quantify semantic privacy leakage, we consider the mutual information between the original task data and the adversary's observation,

$$\mathcal{L}_i^{\mathrm{sec}} = I(x_i; y_{ie}), \tag{14}$$

which measures the amount of information about $x_i$ that can be inferred from the intercepted signal $y_{ie}$. By the data-processing inequality, this leakage is upper bounded by the mutual information between $x_i$ and the transmitted latent representation,

$$I(x_i; y_{ie}) \le I(x_i; z_i^{(\mathrm{tx})}), \tag{15}$$

since $y_{ie}$ is obtained from $z_i^{(\mathrm{tx})}$ through a noisy wireless channel.

In diffusion-based semantic encoding, the transmitted latent $z_i^{(\mathrm{tx})}$ is generated via a finite-step forward diffusion process that progressively injects Gaussian noise into the input. As the number of diffusion steps increases, the conditional distribution $p(z_i^{(\mathrm{tx})} \mid x_i)$ approaches an input-independent Gaussian distribution, thereby significantly reducing statistical dependence between $x_i$ and $z_i^{(\mathrm{tx})}$. As a result, the mutual information $I(x_i; z_i^{(\mathrm{tx})})$ becomes small, leading to

$$\mathcal{L}_i^{\mathrm{sec}} = I(x_i; y_{ie}) \approx 0, \tag{16}$$

where the approximation reflects negligible semantic leakage rather than an exact zero value.

In contrast, AE and VAE encoders map inputs to structured latent manifolds that retain class- or feature-dependent information, yielding larger values of $I(x_i; z_i^{(\mathrm{tx})})$ and consequently higher semantic leakage. Exact computation of $I(x_i; y_{ie})$ is intractable in high-dimensional settings with unknown joint distributions; therefore, following common practice in privacy-preserving learning and secure communication, we approximate semantic leakage through adversarial reconstruction and inference attacks, which provide empirical lower bounds on the recoverable information.

### C. Objective and Design Trade-Offs

Secure semantic offloading in MEC inherently involves balancing three coupled objectives: (i) low end-to-end latency, ensuring that the total task completion time $T_i^{\mathrm{tot}}$ satisfies the deadline constraint; (ii) energy efficiency, minimizing the transmission energy consumption $E_i^{\mathrm{tx}}$ at resource-constrained devices; and (iii) semantic privacy, suppressing information leakage to potential eavesdroppers as quantified by $\mathcal{L}_i^{\mathrm{sec}}$. These objectives are intrinsically interdependent. Reducing latency typically requires higher transmission rates or more compact representations; improving energy efficiency favors shorter encoded representations; and enhancing privacy often necessitates additional stochasticity or corruption in the latent space, which may affect task utility. Diffusion-based semantic encoding is designed to mitigate these tensions by producing compact, noise-like latent representations that simultaneously preserve task-relevant information for authorized edge servers and suppress semantic leakage to adversaries.

These considerations motivate the following conceptual multi-objective formulation:

$$\begin{aligned}
\min_{\mathcal{E},\mathcal{D}} \quad & \sum_{i\in\mathcal{N}} \Big(\alpha_1 T_i^{\mathrm{tot}}(\mathcal{E},\mathcal{D}) + \alpha_2 E_i^{\mathrm{tx}}(\mathcal{E}) + \alpha_3 \mathcal{L}_i^{\mathrm{sec}}(\mathcal{E})\Big) \\
\text{s.t.} \quad & z_i = \mathcal{E}(x_i), \quad z_i^{(\mathrm{tx})} = Q(z_i + \eta_i), \\
& T_i^{\mathrm{tot}}(\mathcal{E},\mathcal{D}) \le T_i^{\max}, \quad \forall i, \\
& 0 < f_j \le f_j^{\max}, \quad \forall j, \\
& \mathcal{L}_{\mathrm{task}}\Big(\mathcal{D}(z_i^{(\mathrm{tx})}),\, y_i\Big) \le \varepsilon, \quad \forall i, \\
& \mathcal{L}_i^{\mathrm{sec}}(\mathcal{E}) = I(x_i; y_{ie}), \quad \forall i,
\end{aligned} \tag{17}$$

where $\mathcal{L}_{\mathrm{task}}(\cdot,\cdot)$ denotes a task-dependent loss function (e.g., cross-entropy for classification), $\varepsilon$ specifies an acceptable task distortion level, and $\alpha_1, \alpha_2, \alpha_3 \ge 0$ weight the relative importance of latency, energy, and privacy.

We emphasize that Eq. (17) is not intended to be solved explicitly through joint resource allocation or online optimization. Instead, it serves as a conceptual formulation that clarifies the fundamental trade-offs among latency, energy efficiency, and semantic privacy in MEC offloading. In this work, we instantiate different encoder-decoder pairs, including AE, VAE, and diffusion-based semantic encoders, train them under identical MEC configurations, and empirically evaluate the resulting performance metrics. This empirical comparison provides an approximation of the Pareto trade-offs implied by Eq. (17) and highlights the advantages of diffusion-based semantic encoding for secure and efficient edge offloading.

## IV. Diffusion-Based Semantic Encoding

In this section, we present the proposed diffusion-based semantic encoding framework for secure computation

offloading. Diffusion models progressively convert structured data into noise through a forward stochastic process and then learn a reverse process that reconstructs task-relevant information from the noisy representation. This mechanism is well suited to MEC offloading because the transmitted latent vector is intentionally noise-like and privacy-preserving, yet remains decodable by authorized edge servers. We describe the forward diffusion process, reverse decoding, training objective, semantic privacy analysis, and the deployment pipeline for edge computing. Together, these components form a unified diffusion-based semantic encoding framework for secure and efficient MEC systems.

### A. Forward Diffusion Process

Given an input task sample $x_i$ (e.g., an image or sensor tensor), the forward diffusion process gradually corrupts $x_i$ into latent variables $\{z_{i,t}\}_{t=1}^{T}$ by adding Gaussian noise in a controlled manner. Let $\{\beta_t\}_{t=1}^{T}$ denote a predefined noise schedule with $\beta_t \in (0,1)$. Define

$$\alpha_t = 1 - \beta_t, \qquad \bar{\alpha}_t = \prod_{s=1}^{t} \alpha_s. \tag{18}$$

A single step of the forward process is given by

$$z_{i,t} = \sqrt{\alpha_t}\, z_{i,t-1} + \sqrt{1-\alpha_t}\, \epsilon_t, \qquad \epsilon_t \sim \mathcal{N}(0, I), \tag{19}$$

with $z_{i,0} = x_i$. Using standard reparameterization, the distribution at step $t$ admits a closed-form expression:

$$z_{i,t} = \sqrt{\bar{\alpha}_t}\, x_i + \sqrt{1-\bar{\alpha}_t}\, \epsilon, \qquad \epsilon \sim \mathcal{N}(0, I). \tag{20}$$

After sufficiently many steps $T$, the distribution of $z_{i,T}$ becomes statistically close to a standard Gaussian. We transmit this final noisy representation

$$z_i \triangleq z_{i,T}, \tag{21}$$

which inherently obfuscates semantic content and significantly reduces information leakage during wireless transmission.

### B. Reverse Decoding at the Edge Server

Authorized edge servers are equipped with a reverse diffusion decoder that reconstructs task-relevant features from the noise-like latent vector $z_i$. The reverse process is parameterized by a neural network $\epsilon_\theta(\cdot)$ trained to estimate the noise component at each diffusion step. The reverse update for step $t$ is

$$\hat{z}_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(\hat{z}_t - (1-\alpha_t)\, \epsilon_\theta(\hat{z}_t, t)\right) + \sigma_t\, \eta_t, \tag{22}$$

where $\eta_t \sim \mathcal{N}(0, I)$ is injected noise, and $\sigma_t$ is a variance term ensuring proper reverse sampling. Starting from $\hat{z}_T = z_i$, the decoder iteratively applies this update until $\hat{z}_0$ is produced. The output $\hat{z}_0$ is then passed through a lightweight projection network to recover task-relevant features suitable for inference:

$$\hat{x}_i = g_\phi(\hat{z}_0). \tag{23}$$

Importantly, only authorized servers possess $\epsilon_\theta$ and $g_\phi$; therefore, even if an adversary steals $z_i$, they cannot execute the reverse process.

### C. Training Objective

The diffusion encoder and decoder are trained jointly using a denoising score-matching objective. For a randomly sampled diffusion timestep $t$, the training target is to predict the noise $\epsilon$ used in the forward process:

$$\mathcal{L}_{\text{diff}}(\theta) = \mathbb{E}_{x_i,\, t,\, \epsilon}\left[\left\|\epsilon - \epsilon_\theta\left(\sqrt{\bar{\alpha}_t}\, x_i + \sqrt{1-\bar{\alpha}_t}\, \epsilon,\ t\right)\right\|_2^2\right]. \tag{24}$$

Minimizing $\mathcal{L}_{\text{diff}}$ ensures that the decoder can accurately model the reverse denoising dynamics, enabling high-quality semantic reconstruction at the edge. Since the transmitted latent $z_i = z_{i,T}$ is statistically close to a Gaussian distribution, the diffusion model naturally enforces semantic obfuscation without the need for explicit adversarial training.

**Algorithm 1** Training of Diffusion-Based Semantic Encoder-Decoder

**Input:** Training dataset $\mathcal{D} = \{(x_i, y_i)\}$, diffusion steps $T$, noise schedule $\{\beta_t\}_{t=1}^{T}$, denoising network $\epsilon_\theta(\cdot)$, task head $g_\phi(\cdot)$
**Output:** Trained parameters $\theta$ and $\phi$
Initialize $\theta$ and $\phi$
Compute $\alpha_t = 1 - \beta_t$ and $\bar{\alpha}_t = \prod_{s=1}^{t} \alpha_s$
**while** *not converged* **do**
  Sample mini-batch $(x_i, y_i) \sim \mathcal{D}$
  Sample diffusion step $t \sim \text{Uniform}\{1, \ldots, T\}$
  Sample noise $\epsilon \sim \mathcal{N}(0, I)$
  Generate noisy latent representation $z_{i,t} = \sqrt{\bar{\alpha}_t} x_i + \sqrt{1-\bar{\alpha}_t}\epsilon$
  Predict noise $\hat{\epsilon} = \epsilon_\theta(z_{i,t}, t)$
  Compute diffusion loss $\mathcal{L}_{\text{diff}} = \|\epsilon - \hat{\epsilon}\|_2^2$
  Perform reverse diffusion to obtain $\hat{z}_{i,0}$
  Predict task output $\hat{y}_i = g_\phi(\hat{z}_{i,0})$
  Compute task loss $\mathcal{L}_{\text{task}} = \ell(\hat{y}_i, y_i)$
  Compute total loss $\mathcal{L} = \mathcal{L}_{\text{diff}} + \lambda \mathcal{L}_{\text{task}}$
  Update $\theta$ and $\phi$ using gradient descent
**return** $\theta, \phi$

The training procedure of the proposed diffusion-based semantic encoder-decoder and the overall semantic offloading framework are summarized in Algorithms 1 and 2, respectively.

### D. Semantic Obfuscation and Privacy Preservation

The structure of the diffusion process provides inherent privacy benefits. Because $z_{i,T}$ converges toward a sample from $\mathcal{N}(0, I)$ regardless of the input $x_i$, the transmitted vector leaks minimal semantic information. Formally, as $T$ increases,

$$z_{i,T} \xrightarrow{d} \mathcal{N}(0, I), \tag{25}$$

and therefore the adversary's mutual information satisfies

$$I(x_i; z_{i,T}) \to 0. \tag{26}$$

This is in contrast to AE and VAE baselines, where latent vectors retain structured information that can be exploited by reconstruction or inference attacks. The diffusion encoder thus

**Algorithm 2** Diffusion-Based Semantic Offloading in MEC

**Input:** Task $\Gamma_i = \{x_i, c_i, T_i^{\max}\}$, trained $\epsilon_\theta$, trained $g_\phi$, diffusion depth $T$
**Output:** Inference result $\hat{y}_i$
IoT device receives task input $x_i$
**for** $t = 1, \ldots, T$ **do**
  Sample $\epsilon_t \sim \mathcal{N}(0, I)$
  Apply forward diffusion:
  $z_{i,t} = \sqrt{\alpha_t} z_{i,t-1} + \sqrt{1-\alpha_t}\epsilon_t$
Set transmitted latent $z_i = z_{i,T}$
Transmit through wireless channel: $z_i^{(\mathrm{tx})} = Q(z_i + \eta_i)$
Edge server receives $z_i^{(\mathrm{tx})}$
Initialize $\hat{z}_T = z_i^{(\mathrm{tx})}$
**for** $t = T, T-1, \ldots, 1$ **do**
  Estimate noise $\hat{\epsilon}_t = \epsilon_\theta(\hat{z}_t, t)$
  Perform reverse diffusion:
  $\hat{z}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\hat{z}_t - (1-\alpha_t)\,\hat{\epsilon}_t\right) + \sigma_t \eta_t$
Recover task representation $\hat{z}_0$
Perform downstream inference $\hat{y}_i = g_\phi(\hat{z}_0)$
**return** $\hat{y}_i$

provides a principled mechanism for generating noise-like yet decodable representations suitable for secure offloading.

The overall MEC offloading pipeline operates as follows. Each IoT device applies the forward diffusion process to transform its task input into an approximately noise-like latent representation, which is then transmitted over the wireless channel to the selected edge server. Upon reception, the edge server applies the learned reverse diffusion model to recover task-relevant features and executes the task using its available computational resources. This design ensures that the transmitted signal remains compact and privacy-preserving while enabling accurate task execution at authorized edge servers. The proposed framework employs latent-space diffusion rather than computationally intensive pixel-space diffusion. Moreover, only the forward diffusion process is executed on the IoT device, while reverse diffusion decoding is performed at the edge server. Therefore, the computational burden on IoT devices remains manageable for practical MEC-IoT deployments.

## V. Security Analysis

In this section, we analyze the security guarantees of the proposed diffusion-based semantic encoding framework. We examine privacy leakage from an information-theoretic perspective, characterize the indistinguishability of transmitted latent vectors, and evaluate robustness against reconstruction and inference attacks commonly encountered in MEC offloading. Throughout, we compare diffusion-based encoding with AE and VAE baselines to highlight the fundamental security advantages induced by the diffusion corruption process.

### A. Information-Theoretic Leakage

The core security objective is to minimize the mutual information between the raw input $x_i$ and the eavesdropper's observation $y_{ie}$. As defined in Eq. (14), the semantic leakage associated with task $\Gamma_i$ defined Eq. (14) and $y_{ie}$ follows the intercepted signal model in Eq. (10) and depends on the transmitted latent representation $z_i$.

For diffusion-based semantic encoding, the transmitted latent corresponds to the terminal forward-diffusion variable $z_{i,T}$. As discussed in Section II, the cumulative diffusion process progressively injects Gaussian noise, driving the latent distribution toward an input-independent Gaussian. Consequently, the statistical dependence between $x_i$ and $z_{i,T}$ becomes weak, yielding $I(x_i; z_{i,T}) \approx 0$ for sufficiently large diffusion depths.

Since the eavesdropper observes only a noisy linear transformation of $z_{i,T}$, the data-processing inequality implies

$$I(x_i; y_{ie}) \leq I(x_i; z_{i,T}) \approx 0, \tag{27}$$

establishing an information-theoretic upper bound on semantic leakage. This bound holds independently of channel conditions or the adversary's computational capability. In contrast, AE- and VAE-based encoders produce structured latent manifolds that retain class- or feature-dependent information, resulting in strictly larger leakage and weaker privacy guarantees.

### B. Semantic Indistinguishability

We evaluate whether $z_i$ is statistically distinguishable from pure Gaussian noise. Let $p_z$ denote the distribution of encoded latent vectors and $p_{\mathrm{noise}} = \mathcal{N}(0, I)$ denote the target distribution. A common measure of distributional distinguishability is the Kullback-Leibler (KL) divergence between the latent distribution and a reference noise distribution,

$$D_{\mathrm{KL}}(p_z \,\|\, p_{\mathrm{noise}}) = \mathbb{E}_{p_z(z)}\left[\log \frac{p_z(z)}{p_{\mathrm{noise}}(z)}\right], \tag{28}$$

which quantifies how easily an adversary can statistically distinguish the transmitted latent representation from pure noise. Diffusion-based encoding explicitly drives $p_z$ toward $p_{\mathrm{noise}}$ through the cumulative forward corruption process, ensuring

$$D_{\mathrm{KL}}(p_z \,\|\, p_{\mathrm{noise}}) \to 0. \tag{29}$$

This implies that an eavesdropper cannot reliably distinguish transmitted messages from channel noise, leading to a form of semantic deniability. Conversely, AE and VAE latents cluster around a low-dimensional manifold with identifiable structure, making them easily separable from Gaussian noise via likelihood tests or simple classifiers.

### C. Robustness to Reconstruction Attacks

We analyze robustness against reconstruction attacks, in which an adversary attempts to estimate the original input $x_i$ from the intercepted signal $y_{ie}$ using neural inversion, denoising, or generative priors. Let $\hat{x}_i^{(e)}$ denote the adversary's estimate of $x_i$, and define the expected reconstruction error as

$$\mathcal{R}_i = \mathbb{E}\left[\|x_i - \hat{x}_i^{(e)}\|_2^2\right]. \tag{30}$$

The attacker's inference problem can be expressed as

$$\hat{x}_i^{(e)} = \arg\min_x \|y_{ie} - h_{ie}\sqrt{P_i}\,\mathcal{E}(x)\|^2 + \lambda\,\Omega(x), \tag{31}$$

where $\Omega(\cdot)$ represents any structural or generative prior available to the adversary.

For AE and VAE encoders, the mapping from $x$ to $z$ is continuous and retains structured semantic information, making partial inversion feasible under favorable conditions. In contrast, diffusion-based encoding progressively transforms the transmitted latent $z_{i,T}$ into an approximately noise-like representation that is statistically close to a Gaussian distribution, substantially weakening the statistical dependence between $x_i$ and $z_{i,T}$ prior to transmission. However, the latent representation is not completely structure-free, since preserving limited task-relevant information is necessary to maintain reliable downstream task performance at authorized edge servers. Consequently, the attacker is faced with a severely ill-posed inverse problem in which many candidate inputs are equally consistent with the observed signal.

From this perspective, diffusion-based semantic encoding is expected to induce larger reconstruction uncertainty than AE- and VAE-based encoders, leading to the qualitative ordering

$$\mathbb{E}[\mathcal{R}_i^{\text{diffusion}}] \gtrsim \mathbb{E}[\mathcal{R}_i^{\text{VAE}}] \geq \mathbb{E}[\mathcal{R}_i^{\text{AE}}]. \tag{32}$$

We emphasize that this ordering is not claimed as a universal theoretical guarantee, but as a theory-guided expectation under comparable architectures and attacker models, which is empirically validated in Section VI.

### D. Resistance to Feature-Inference Attacks

Beyond reconstructing raw data, adversaries may attempt to infer task-relevant features such as class labels, object categories, or sensitive attributes. Let $\hat{y}_i^{(e)}$ denote the adversary's predicted task label. Feature-inference risk is measured by the adversarial classification accuracy

$$\mathcal{A}_i^{\text{adv}} = \Pr\left(\hat{y}_i^{(e)} = y_i\right). \tag{33}$$

Because diffusion-based encoding obscures structure in the input, renders $z_{i,T}$ isotropic, and eliminates low-frequency spectral components essential for classification, adversarial accuracy is reduced compared with structured AE/VAE latents:

$$\mathcal{A}_i^{\text{adv}} \approx \frac{1}{K}, \tag{34}$$

for a $K$-class task. For AE and VAE encodings, significant class-discriminative information remains in $z_i$, enabling nontrivial adversarial classification.

### E. Comparison With Physical-Layer Security

Conventional PLS techniques such as friendly jamming improve legitimate link quality relative to the eavesdropper but do not eliminate semantic leakage, since the transmitted waveform still contains structured transforms of $x_i$. Diffusion-based encoding is complementary to PLS yet fundamentally stronger in semantic privacy: the transmitted signal is noise-like even before modulation, making it robust to perfect-channel adversaries and independent of SNR differences. This property enables secure offloading even in adverse or high-risk environments where PLS assumptions fail.

In summary, the diffusion-based semantic encoder provides strong information-theoretic privacy, noise indistinguishability, and robustness to both reconstruction and inference attacks. These guarantees arise intrinsically from the diffusion corruption process and do not rely on cryptographic keys, adversarial training, or cooperative jamming. Compared with AE, VAE, and conventional PLS schemes, diffusion-based encoding offers substantially stronger protection while maintaining full utility at authorized MEC servers.

## VI. Results and Analysis

We now present a comprehensive empirical evaluation of the proposed diffusion-based semantic encoder and its baselines (AE and VAE). All models are implemented in Python and PyTorch and are trained under identical settings, with the same dataset, optimizer, learning rate schedule, and channel parameters (additive white Gaussian noise (AWGN) + quantization). All experiments are conducted on the CIFAR-10 image classification dataset. Unless otherwise stated, the latent dimension is fixed at $d = 64$. All models are trained at an SNR of 20 dB and evaluated under SNRs of $\{0, 5, 10, 20\}$ dB. The default quantization resolution is 8 bits. The semantic encoder compresses the input data into a latent representation that is transmitted to the edge server, where a lightweight classifier operates on the decoded latent. Task accuracy is defined as the fraction of correctly classified samples over the test set and therefore takes values in the range $[0, 1]$. Training and validation accuracy curves are reported to assess convergence behavior and generalization performance.

Our analysis focuses on five key aspects: (i) task performance, (ii) training dynamics, (iii) latent-space geometry, (iv) robustness to channel noise and adversarial reconstruction, and (v) semantic privacy leakage. These collectively approximate the performance-privacy-efficiency trade-offs described in Section III.

To ensure fair comparison, AE, VAE, and diffusion-based semantic encoders are implemented using the same encoder backbone and latent dimensionality. Specifically, the latent dimension is fixed to $d = 64$ for all models. Both AE and VAE employ the same convolutional encoder consisting of four convolutional blocks with channel dimensions $3 \to 32 \to 64 \to 128 \to 256$, where each block contains a convolution layer, batch normalization, and ReLU activation, followed by adaptive average pooling and latent projection layers that map the extracted features to a 64-dimensional latent representation. The AE and VAE encoders employ the same latent classifier head consisting of LayerNorm, a fully connected layer with 256 hidden units, ReLU activation, dropout with rate 0.1, and a final classification layer. The VAE additionally introduces stochastic latent sampling through the reparameterization trick.

The diffusion-based semantic encoder adopts a standard DDPM-style diffusion process with a linear noise schedule consisting of 1000 diffusion steps, where the noise variance increases linearly from $10^{-4}$ to $2\times10^{-2}$. The task-conditioned

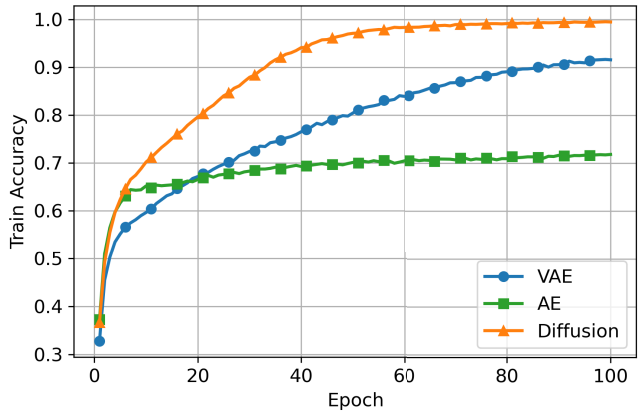



Fig. 2. Training accuracy of AE, VAE, and diffusion-based semantic encoders.

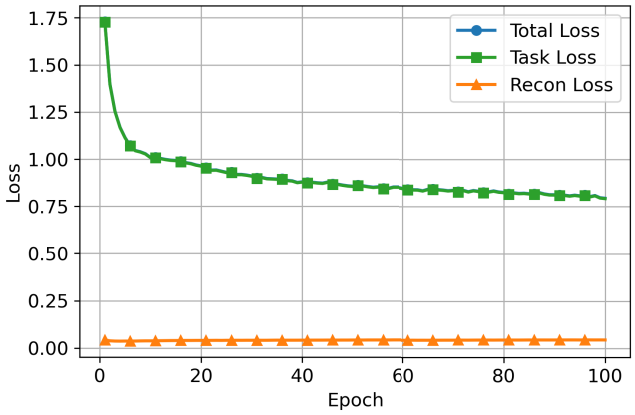



Fig. 4. Training loss curves for the AE-based semantic encoder.

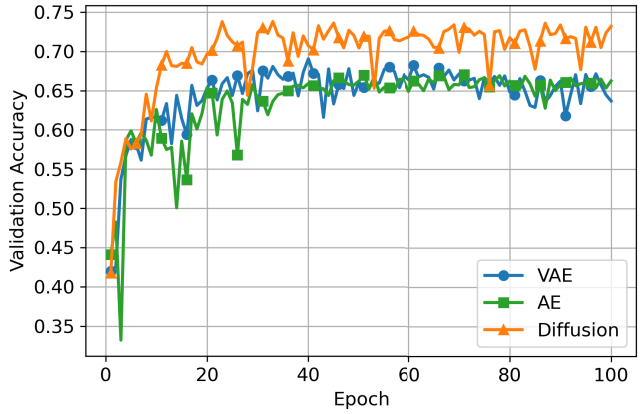



Fig. 3. Validation accuracy comparison across semantic encoders.

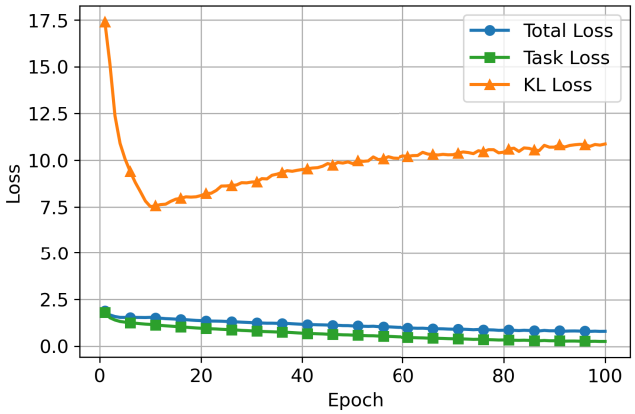



Fig. 5. Training loss curves for the VAE-based semantic encoder.

denoising network employs a hidden dimension of 512, time-embedding dimension of 64, and task-conditioning embedding dimension of 128.

To evaluate semantic leakage, a separate attacker network is trained for each encoder type. The attacker is implemented as a three-layer multilayer perceptron (MLP), consisting of two hidden layers with 128 units and ReLU activations, followed by a classification output layer that predicts the ground-truth class label from the latent representation. Following standard practice, the attacker is trained after the semantic encoder is frozen using the Adam optimizer with learning rate $10^{-3}$, batch size 128, and 3 training epochs. To estimate an upper bound on semantic leakage, the attacker is trained on clean latent representations and evaluated on transmitted latents after AWGN corruption and quantization. All hyperparameters are kept identical unless otherwise specified.

### A. Task Accuracy Across Semantic Encoders

Figures 2 and 3 compare the training and validation accuracy of the three semantic encoders under identical MEC settings. The diffusion-based encoder consistently achieves the highest accuracy throughout training and maintains its advantage on the validation set. This is notable because the diffusion latent is intentionally noise-like; yet, its structured forward corruption process yields representations that remain both expressive and robust to channel perturbations.

The AE baseline exhibits the lowest performance: its deterministic latent representation is highly sensitive to AWGN and quantization, leading to early saturation and limited generalization. The VAE baseline performs better due to stochastic sampling, which regularizes the latent space and improves robustness; however, its accuracy still lags behind diffusion. Overall, diffusion-based encoding provides the most stable and robust task performance across the entire training trajectory.

### B. Training Dynamics and Loss Curves

Figures 4-6 show the evolution of total loss, task loss, and model-specific auxiliary losses across training epochs for the AE, VAE, and diffusion-based encoders. The diffusion model demonstrates the smoothest and most stable convergence: both its task loss and diffusion-denoising loss decrease steadily, indicating effective learning of task-relevant features while maintaining a noise-like latent space.

In contrast, the AE baseline converges quickly but plateaus at a higher loss. Because AE latent vectors are deterministic and highly sensitive to channel perturbations, the model struggles to refine task-relevant information once quantization and

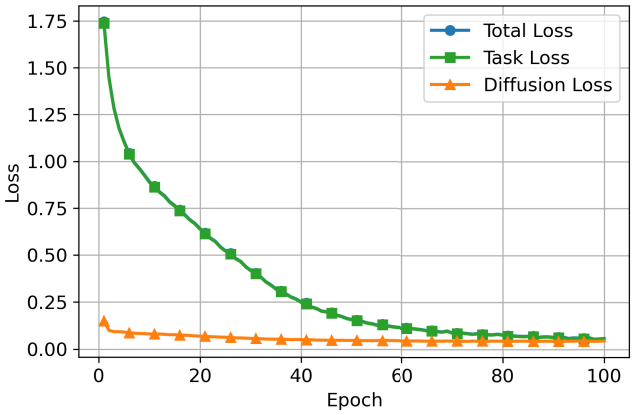



Fig. 6. Training loss curves for the diffusion-based semantic encoder.

AWGN are applied. This early saturation reflects its limited expressive capacity in noisy offloading scenarios.

The VAE exhibits intermediate behavior. The stochastic sampling mechanism helps regularize the latent space and initially accelerates optimization, but the KL term grows throughout training, causing the total loss to follow a characteristic U-shape. This behavior is typical of VAEs and indicates a trade-off between latent regularity and reconstruction quality.

Overall, diffusion-based training yields the most stable and consistent loss profile, suggesting inherent robustness to channel noise and a better alignment between latent structure and downstream task objectives. The gradual corruption process in diffusion acts as an implicit regularizer, reducing overfitting and enabling superior task performance.

### C. Latent-Space Geometry: t-SNE and UMAP

To qualitatively analyze the geometric structure and semantic separability of the learned latent representations, we employ two widely used nonlinear dimension-reduction techniques: t-distributed Stochastic Neighbor Embedding (t-SNE) [29] and Uniform Manifold Approximation and Projection (UMAP) [30]. These methods project high-dimensional latent vectors into low-dimensional spaces while preserving local neighborhood structure (t-SNE) and global manifold relationships (UMAP), respectively, enabling visual assessment of semantic clustering and potential information leakage. To characterize how each semantic encoder organizes information in the latent space, we visualize the encoded representations $z_i$ from the CIFAR-10 test set using t-SNE and UMAP in both 2D and 3D. All models are evaluated on the same unseen data, and all latent vectors are extracted from the final trained AE-MEC, VAE-MEC, and Diffusion-MEC encoders. For each encoder, the latent matrix is standardized and projected using t-SNE (with three or two components) or UMAP (with $n_{\text{neighbors}} = 15$, *min_dist* = *0.1*, Euclidean metric), providing comparable low-dimensional views of their geometric and topological structure. Figures 7 and 8 present the resulting 3D (top row) and 2D (bottom row) embeddings.

Figures 7 and 8 show the t-SNE and UMAP projections, respectively, with the top row presenting 3D embeddings and the bottom row their 2D counterparts. These plots highlight the extent to which each encoder preserves or destroys semantic structure in the latent space. Across both visualization methods, the AE-MEC latent space exhibits clear, compact, and well-separated clusters corresponding to semantic classes. The structure is highly discriminative but also highly inversion-prone, as the latent vectors remain strongly correlated with class identity. The VAE-MEC latent space is more diffuse: stochastic sampling spreads the representations and partially overlaps different classes, offering moderate privacy while retaining some semantic organization.

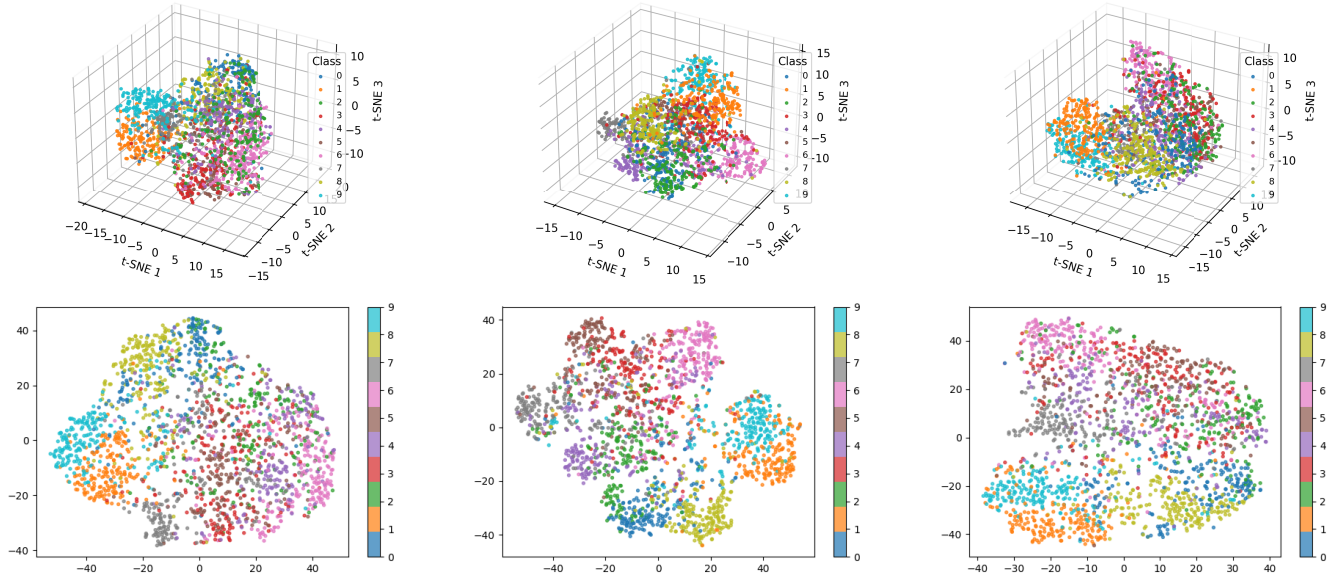
In contrast, the Diffusion-MEC latent space appears fundamentally different. Both 3D and 2D embeddings collapse into a mixed, amorphous distribution with no discernible class-dependent geometry. Classes are uniformly interwoven, and neither local nor global manifold structures are preserved. This isotropic, noise-like latent organization is consistent with the behavior expected from a forward diffusion process and aligns with the information-theoretic goal of minimizing $I(x_i; z_i)$. The result is a latent representation that is highly resistant to reconstruction or feature inference, providing strong semantic obfuscation while remaining compatible with downstream decoding on authorized edge servers.

These latent-space observations directly support the privacy-leakage analysis presented later and demonstrate that diffusion-based semantic encoders inherently destroy class-level structure that AE and VAE models still retain.

### D. Privacy Leakage Under Adversarial Attacks

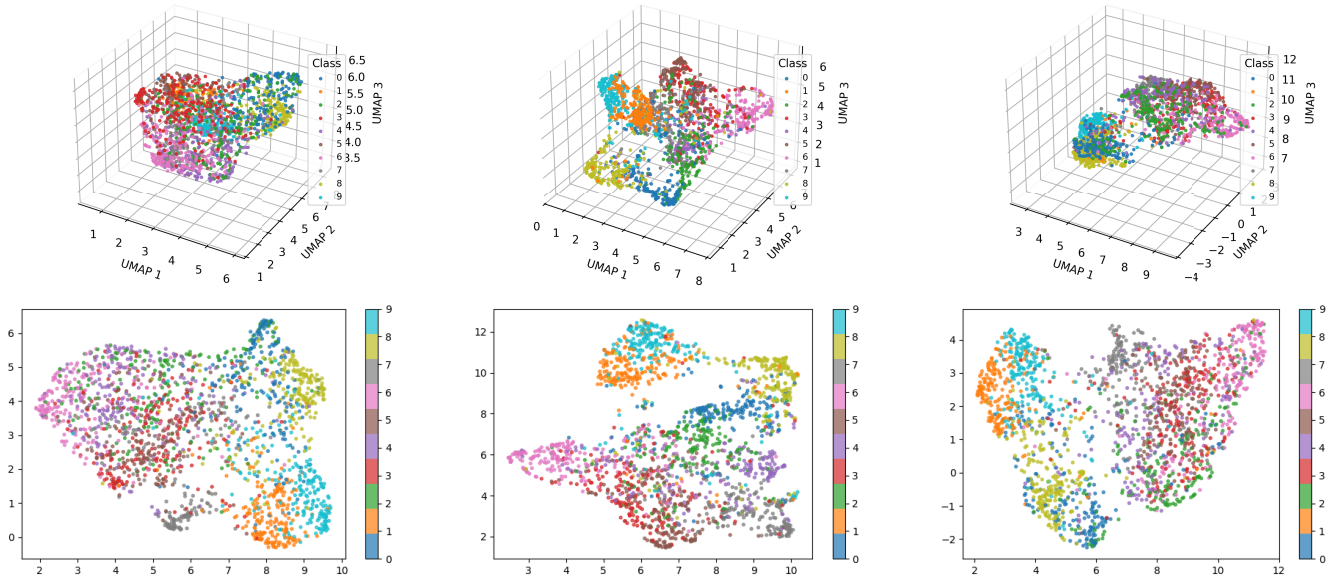
We quantify semantic leakage by training a separate latent-space attacker for each encoder type. The attacker is implemented as a three-layer MLP with hidden dimension 128 and ReLU activations. Following standard practice, the attacker is trained after the semantic encoder is frozen using the Adam optimizer with learning rate $10^{-3}$, batch size 128, and 3 training epochs. To estimate an upper bound on semantic leakage, the attacker is primarily trained on clean latent representations and evaluated on transmitted latents after AWGN corruption and quantization. The attacker attempts to infer ground-truth class labels from latent vectors. Accordingly, attacker accuracy directly reflects privacy leakage: higher values indicate more residual semantic structure exposed to an adversary. Figure 9 (bottom row) reports attacker accuracy across quantization levels (4−, 6−, and 8-bit) and SNR values. Across all settings, AE-MEC exhibits the lowest attacker accuracy, reflecting its limited latent expressiveness and the substantial distortion induced by channel noise and quantization. However, this privacy benefit comes at the cost of significantly degraded task accuracy, especially at low SNR, indicating that AE-MEC achieves privacy primarily by collapsing useful semantic information.

VAE-MEC displays moderate attacker accuracy, revealing more latent structure than AE-MEC while providing better task performance. In contrast, Diffusion-MEC yields moderately higher attacker accuracy on the transmitted latents because the diffusion latent representations remain partially task-separable in order to preserve downstream inference utility. Although the latent space becomes approximately noise-like, it still retains




(a) Autoencoder (AE) (b) Variational Autoencoder (VAE) (c) Diffusion-based encoder

Fig. 7. 3D (top row) and 2D (bottom row) t-SNE visualizations of the latent representations for (a) AE, (b) VAE, and (c) diffusion-based encoder.




(a) Autoencoder (AE) (b) Variational Autoencoder (VAE) (c) Diffusion-based encoder

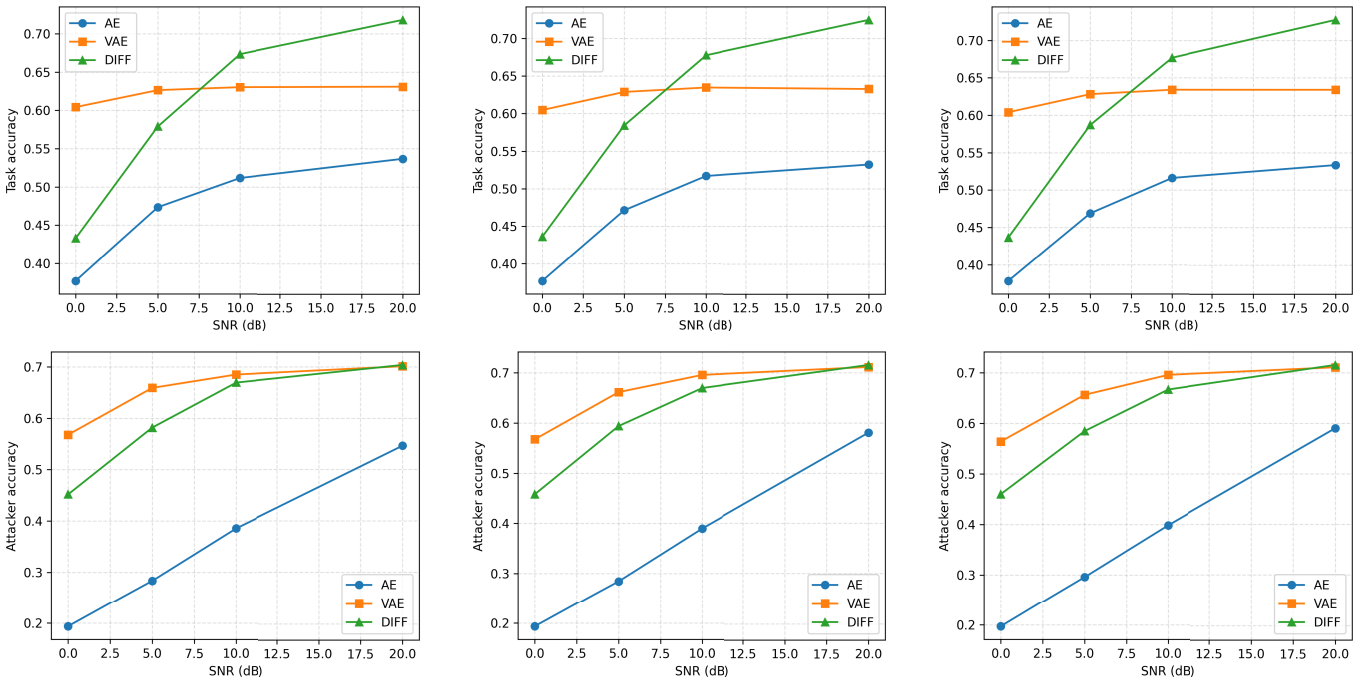
Fig. 8. 3D (top row) and 2D (bottom row) UMAP visualizations of the latent representations for (a) AE, (b) VAE, and (c) diffusion-based encoder.

limited task-relevant semantic structure required for reliable edge inference. Nevertheless, diffusion consistently delivers the best task accuracy across all SNR and quantization levels, demonstrating a far more favorable utility-privacy trade-off than AE-MEC and VAE-MEC.

These findings align with the latent-geometry visualizations reported earlier: AE latents collapse into tight, noisy clusters, VAE latents maintain moderate class structure, and diffusion latents remain the most stable under perturbations. Together, the results confirm that while diffusion does not minimize attacker accuracy, it provides the strongest overall robustness under noisy MEC conditions, preserving semantic utility while tolerating channel distortions far better than AE-MEC and VAE-MEC. Overall, these results indicate that diffusion-based semantic encoding does not completely eliminate latent structure; instead, it achieves a favorable privacy-utility trade-off by significantly suppressing recoverable semantic information while maintaining robust downstream inference performance.

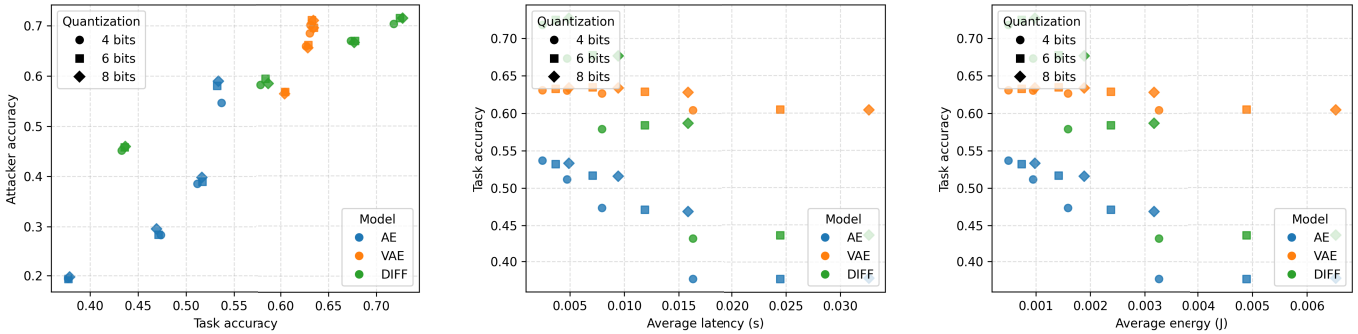
### *E. Communication Efficiency in Terms of Latency and Energy*

Communication efficiency is determined by both the quality of the transmitted latent and the number of bits required to




(a) 4-bit quantization (b) 6-bit quantization (c) 8-bit quantization

Fig. 9. Performance and attacker reconstruction across different quantization levels. Top row: task accuracy vs. SNR. Bottom row: attacker reconstruction accuracy vs. SNR.




(a) Accuracy vs. attacker performance (b) Accuracy vs. latency (c) Accuracy vs. energy consumption

Fig. 10. Trade-off across security, latency, and energy: (a) attacker performance, (b) latency, and (c) energy consumption.

represent it. For latent dimension $d$ and quantization width $b$, the transmitted payload is $L_i = db$ bits, with latency

$$T_i^{\text{tx}} = \frac{L_i}{B \log_2(1+\gamma)},$$

using a 1 MHz bandwidth and SNR $\gamma$. Transmission energy follows $E_i^{\text{tx}} = P_{\text{tx}} T_i^{\text{tx}}$ with $P_{\text{tx}} = 0.2$ W. Figure 10 illustrates the trade-offs between task accuracy, attacker accuracy, latency, and energy across all models.

Diffusion-MEC consistently achieves the most favorable accuracy-latency-energy trade-off. Even at 4-bit quantization, Diffusion-MEC maintains competitive accuracy while substantially reducing communication cost, whereas AE-MEC and VAE-MEC require higher bit-widths to preserve reliability. AE-MEC is especially sensitive to aggressive quantization and low SNR, leading to both degraded task performance and higher relative energy consumption for equivalent accuracy targets. VAE-MEC performs more robustly but still falls short of diffusion in all three dimensions.

### F. Latent Compression and Dimensionality

Latent dimensionality plays a critical role in communication-efficient semantic offloading, as smaller representations directly reduce transmission latency and energy consumption. In principle, aggressive dimensionality reduction may also remove task-relevant information and amplify privacy–utility trade-offs, making robustness under compression a key design requirement for MEC systems.

In the current implementation, all encoders are trained with a fixed latent dimension. A complete dimensionality sweep (e.g., $d \in \{16, 32, 64, 128\}$) would require retraining each

encoder–decoder pair and re-evaluating the resulting accuracy, privacy leakage, and communication costs, which we leave as future work. Nevertheless, the empirical results presented in this section provide consistent qualitative evidence regarding compression robustness.

Specifically, the diffusion-based encoder maintains strong task accuracy and low privacy leakage under aggressive quantization and noisy channel conditions (Figures 9-10), indicating that its noise-regularized latent manifold is resilient to information loss. In contrast, AE-MEC and VAE-MEC rely on more structured, class-aligned latent representations that are inherently more sensitive to compression and channel distortion. This behavior is further supported by the latent-space visualizations in Figures 7 and 8, where diffusion latents exhibit substantially weakened semantic clustering.

Together, these observations suggest that diffusion-based semantic encoding is well suited for bandwidth-limited MEC settings, where compact representations are essential. We will extend this analysis with explicit latent-dimension sweeps and corresponding plots in future work.

## VII. Conclusion

We have proposed a diffusion-based semantic encoding framework for secure and efficient computation offloading in MEC systems. The framework has transformed raw inputs into noise-like representations that are statistically close to Gaussian noise, providing intrinsic protection against eavesdropping while preserving task-relevant semantics for authorized edge servers. A unified MEC pipeline incorporating forward diffusion, channel perturbation, and reverse diffusion decoding has been developed and evaluated under realistic wireless impairments and quantization constraints. Extensive experimental results have shown that diffusion-based encoding achieves superior robustness to SNR degradation, aggressive quantization, and latent compression, while maintaining strong task performance and favorable latency-energy trade-offs. Latent-space analyses using t-SNE and UMAP have further confirmed that diffusion latents exhibit minimal geometric structure, leading to substantially lower adversarial inference accuracy compared with AE-MEC and VAE-MEC. The results further demonstrate that diffusion-based semantic encoding achieves a favorable privacy-utility trade-off by significantly suppressing recoverable semantic information while preserving reliable downstream inference performance. These findings have demonstrated that diffusion-driven semantic encoding is a promising mechanism for secure MEC systems, offering an effective balance among accuracy, communication efficiency, and semantic privacy in resource-constrained wireless environments. Nevertheless, the proposed framework has several limitations. Performance may degrade under extremely low SNR conditions due to severe channel corruption, and stronger adversaries equipped with surrogate reverse diffusion decoders may recover more semantic information than the lightweight attacker considered in this work. For future work, we will investigate robust defenses against such advanced adversaries, as well as multimodal tasks, adaptive diffusion scheduling under dynamic channel conditions, and lightweight reverse diffusion deployment on edge hardware.

## References


[1] M. Li, J. Gao, L. Zhao, and X. Shen, "Deep reinforcement learning for collaborative edge computing in vehicular networks," *IEEE Trans. Cognit. Commun. Netw.*, vol. 6, no. 4, pp. 1122–1135, Dec. 2020.

[2] L. Ale, N. Zhang, X. Fang, X. Chen, S. Wu, and L. Li, "Delay-aware and energy-efficient computation offloading in mobile-edge computing using deep reinforcement learning," *IEEE Trans. Cognit. Commun. Netw.*, vol. 7, no. 3, pp. 881–892, Sep. 2021.

[3] X. Lu et al., "Blockchain-enabled secure offloading for VEC: A multi-agent reinforcement learning approach," *IEEE Trans. Dependable Secure Comput.*, vol. 22, no. 3, pp. 2978–2995, May 2024.

[4] J. Liu, J. Ren, Y. Zhang, X. Peng, Y. Zhang, and Y. Yang, "Efficient dependent task offloading for multiple applications in MEC-cloud system," *IEEE Trans. Mobile Comput.*, vol. 22, no. 4, pp. 2147–2162, Apr. 2023.

[5] L. Ale, N. Zhang, S. A. King, and D. Chen, "Empowering generative AI through mobile edge computing," *Nature Rev. Electr. Eng.*, vol. 1, no. 7, pp. 478–486, Jun. 2024, doi: 10.1038/s44287-024-00053-6.

[6] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[7] D. Han et al., "Two-timescale learning-based task offloading for remote IoT in integrated satellite–terrestrial networks," *IEEE Internet Things J.*, vol. 10, no. 12, pp. 10131–10145, Jun. 2023.

[8] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: A Bradford Book, 2018.

[9] D. Silver et al., "Mastering the game of go without human knowledge," *Nature*, vol. 550, no. 7676, pp. 354–359, Oct. 2017, doi: 10.1038/nature24270.

[10] T. P. Lillicrap et al., "Continuous control with deep reinforcement learning," 2015, *arXiv:1509.02971*.

[11] L. Ale, S. A. King, N. Zhang, A. R. Sattar, and J. Skandaraniyam, "D3PG: Dirichlet DDPG for task partitioning and offloading with constrained hybrid action space in mobile-edge computing," *IEEE Internet Things J.*, vol. 9, no. 19, pp. 19260–19272, Oct. 2022.

[12] J. Si, Z. Cheng, Z. Li, J. Cheng, H.-M. Wang, and N. Al-Dhahir, "Cooperative jamming for secure transmission with both active and passive eavesdroppers," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5764–5777, Sep. 2020.

[13] X. Qin and X. S. Shen, "Physical layer security-assisted partial computation offloading in mobile edge computing," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Aug. 2024, pp. 874–878.

[14] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[15] A. Castiglione, M. Nappi, F. Narducci, and C. Pero, "Fostering secure cross-layer collaborative communications by means of covert channels in MEC environments," *Comput. Commun.*, vol. 169, pp. 211–219, 2021.

[16] S. Guo, J. Liu, Y. Yang, B. Xiao, and Z. Li, "Energy-efficient dynamic computation offloading and cooperative task scheduling in mobile cloud computing," *IEEE Trans. Mobile Comput.*, vol. 18, no. 2, pp. 319–333, Feb. 2019.

[17] V. Bhagat, S. Kumar, S. K. Gupta, and M. K. Chaube, "Lightweight cryptographic algorithms based on different model architectures: A systematic review and futuristic applications," *Concurrency Comput., Pract. Exper.*, vol. 35, no. 1, p. 7425, Jan. 2023. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/cpe.7425

[18] Y. Wu, G. Ji, T. Wang, L. Qian, B. Lin, and X. Shen, "Non-orthogonal multiple access assisted secure computation offloading via cooperative jamming," *IEEE Trans. Veh. Technol.*, vol. 71, no. 7, pp. 7751–7768, Jul. 2022.

[19] D. Huang, F. Gao, X. Tao, Q. Du, and J. Lu, "Toward semantic communications: Deep learning-based image semantic coding," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 55–71, Jan. 2023.

[20] Z. Bao et al., "SDAC—Semantic digital analog converter for semantic communications," *IEEE Trans. Commun.*, vol. 73, no. 11, pp. 11061–11077, Nov. 2025.

[21] H. Xie, Z. Qin, X. Tao, and K. B. Letaief, "Task-oriented multi-user semantic communications," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 9, pp. 2584–2597, Sep. 2022.

[22] D. Gündüz et al., "Beyond transmitting bits: Context, semantics, and task-oriented communications," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 5–41, Jan. 2023.
[23] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in *Proc. 34th Int. Conf. Neural Inf. Process. Syst.*, vol. 33, 2020, pp. 6840–6851.
[24] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, "Score-based generative modeling through stochastic differential equations," 2020, *arXiv:2011.13456*. [Online]. Available: https://openreview.net/forum?id=PxTIG12RRHS
[25] N. Carlini et al., "Extracting training data from diffusion models," in *Proc. 32nd USENIX Conf. Secur. Symp.*, 2023, pp. 5253–5270.
[26] E. Grassucci, G. Pignata, G. Cicchetti, and D. Comminiello, "Lightweight diffusion models for resource-constrained semantic communication," *IEEE Wireless Commun. Lett.*, vol. 14, no. 9, pp. 2743–2747, Sep. 2025.
[27] Y. Wang, S. Zhang, A. Zhang, S. Dang, H. Zhang, and S. Guo, "Sc-diffusion: Parameter generation for task-oriented semantic communication systems via conditional diffusion model," *IEEE Trans. Mach. Learn. Commun. Netw.*, vol. 3, pp. 1108–1120, 2025.
[28] B. Xu et al., "Semantic prior aided channel-adaptive equalizing and denoising semantic communication system with latent diffusion model," *IEEE Trans. Wireless Commun.*, vol. 24, no. 6, pp. 4614–4630, Jun. 2025.
[29] L. van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, no. 11, 2008.
[30] L. McInnes, J. Healy, N. Saul, and L. Großberger, "UMAP: Uniform manifold approximation and projection," *J. Open Source Softw.*, vol. 3, no. 29, p. 861, Sep. 2018, doi: 10.21105/joss.00861.

**Xue Qin** (Graduate Student Member, IEEE) received the B.Eng. degree from the North University of China (NUC) in 2008 and the M.A.Sc. degree from the University of Windsor, Windsor, ON, Canada, in 2022. She is currently pursuing the Ph.D. degree in electrical and computer engineering with the University of Waterloo, Waterloo, ON, Canada. Her research interests include mobile edge computing, AI, and network security.

**Lingshuang Liu** (Student Member, IEEE) received the B.Eng. and M.S. degrees from the School of Computer Science and Engineering, University of Electronic Science and Technology of China, in 2014 and 2017, respectively. She is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Waterloo, Canada. Her research interests include security and privacy in communication networks and artificial intelligence.

**Xuemin (Sherman) Shen** (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Rutgers University, New Brunswick, NJ, USA, in 1990.

He is currently an University Professor with the Department of Electrical and Computer Engineering, University of Waterloo, Canada. His research focuses on network resource management, wireless network security, AI for network—network for AI, 5G and beyond, and vehicular networks. He is also a registered Professional Engineer of Ontario, Canada, an Engineering Institute of Canada Fellow, a Canadian Academy of Engineering Fellow, a Royal Society of Canada Fellow, a Chinese Academy of Engineering Foreign Member, and an International Fellow of the Engineering Academy of Japan. He received the A. G. L. McNaughton Award in 2026 from IEEE Canada, "West Lake Friendship Award" from Zhejiang Province in 2023, President's Excellence in Research from the University of Waterloo in 2022, Canadian Award for Telecommunications Research from Canadian Society of Information Theory (CSIT) in 2021, the R. A. Fessenden Award in 2019 from IEEE Canada, Award of Merit in 2019 from the Federation of Chinese Canadian Professionals (Ontario), James Evans Avant Garde Award in 2018 from the IEEE Vehicular Technology Society, Joseph LoCicero Award in 2015 and Education Award in 2017 from the IEEE Communications Society (ComSoc), and Technical Recognition Award from Wireless Communications Technical Committee (2019) and AHSN Technical Committee (2013). He has also received the Excellent Graduate Supervision Award in 2006 from the University of Waterloo and the Premier's Research Excellence Award (PREA) in 2003 from the Province of Ontario, Canada. He serves/served as the General Chair for the 6G Global Conference'23 and ACM Mobihoc'15, the Technical Program Committee Chair/Co-Chair for IEEE Globecom'07, 16, and 24, IEEE Infocom'14, IEEE VTC'10 Fall, and the Chair for the IEEE ComSoc Technical Committee on Wireless Communications. He is the Past President of the IEEE ComSoc, the Vice President for Technical & Educational Activities, the Vice President for Publications, a Member-at-Large on the Board of Governors, Chair of the Distinguished Lecturer Selection Committee, and a member of IEEE Fellow Selection Committee of the ComSoc. He served as the Editor-in-Chief for the IEEE INTERNET OF THINGS JOURNAL, *IEEE Network*, PPNA, and *IET Communications*.