Improving Channel Resilience for Task-Oriented Semantic Communications: A Unified Information Bottleneck Approach

Shuai Lyu, Yao Sun, *Senior Member, IEEE*, Linke Guo, *Senior Member, IEEE*, Xiaoyong Yuan, Fang Fang, *Senior Member, IEEE*, Lan Zhang, *Member, IEEE*, and Xianbin Wang, *Fellow, IEEE*

***Abstract*— Task-oriented semantic communications (TSC) enhance radio resource efficiency by transmitting task-relevant semantic information. However, current research often overlooks the inherent semantic distinctions among encoded features. Due to unavoidable channel variations from time and frequency-selective fading, semantically sensitive feature units could be more susceptible to erroneous inference if corrupted by dynamic channels. Therefore, this letter introduces a unified channel-resilient TSC framework via information bottleneck. This framework complements existing TSC approaches by controlling information flow to capture fine-grained feature-level semantic robustness. Experiments on a case study for real-time subchannel allocation validate the framework's effectiveness.**

***Index Terms*— Channel resilience, information bottleneck, task-oriented semantic communications, radio resource allocation.**

I. Introduction

FUTURE wireless networks are expected to support dramatically increased data traffic, driven primarily by the prevalence of sensing capabilities, distributed computing resources, and ongoing convergence with vertical applications, including transportation, online gaming and smart utilities. To support the unprecedented traffic growth with limited radio resources, task-oriented semantic communications (TSC) have garnered considerable interest [1], [2]. Unlike conventional bit-level communications, TSC leverages both distributed computing resources as well as the most recent artificial intelligence (AI) techniques to extract and transmit task-relevant semantic information, thereby reducing unnecessary traffic and enhancing radio resource utilization efficiency.

In TSC, the transceiver typically integrates semantic and channel coding functions to extract task-relevant semantic features and enable efficient transmissions [2]. Yang et al. introduced a novel compression method for AI tasks to improve transmission efficiency, demonstrated with a prototype for surface defect detection [3]. Instead of focusing solely on semantic coding, Shao et al. [4] and Sun et al. [5] proposed to jointly optimize semantic and channel coding based on information bottleneck to balance semantic distortion and transmission efficiency. Moreover, TSC problems have been investigated in multi-user and multi-modality scenarios [6], [7], [8].

Despite considerable efforts, existing TSC research primarily focuses on statistical channel conditions, assuming a certain channel condition when transmitting an input sample [2], [3], [4], [5], [6], [7]. However, the encoded features may be subject to various physical impairments during transmission. In orthogonal frequency division multiplexing (OFDM) systems, for example, signal transmission occurs across multiple subcarriers, each experiencing different frequency-selective fading over a wide bandwidth [9]. While recent work in [10] introduces adaptive subcarrier allocation for TSC, it primarily focuses on ensuring the delivery of semantic-critical feature units. As a result, if semantically sensitive feature units are assigned to subcarriers with poor performance, the resulting corruption of these features is more susceptible to erroneous task inference. While channel estimation techniques can evaluate instantaneous channel conditions in practice [9], the evaluation results cannot effectively guide resilient feature-level transmissions due to the neglected semantic distinctions among encoded feature units.

To bridge the gap between feature-level semantic distinctions and channel variations, this letter introduces an innovative TSC framework to improve channel resilience by evaluating and prioritizing encoded feature units of input data based on their robustness against channel variations. This framework is designed to be complementarily leveraged by existing TSC approaches to capture fine-grained feature-level semantic robustness, thereby adjusting transmission strategies for channel-resilient TSC. The primary contributions of this letter are summarized below.

- *Unified channel-resilience framework:* We develop a unified approach to analyze a well-trained TSC transceiver for channel resilience, providing a soft robustness mask for the encoded feature space without modifying the established TSC encoding and decoding functions. This mask will be utilized to prioritize robust feature units and adapt the transmission strategies against instantaneous channel variations in practice.
- *Robustness mask based on information bottleneck (IB):* We construct the robustness mask for encoded feature units by leveraging IB to regulate information flow with

Received 10 August 2024; revised 14 September 2024; accepted 16 September 2024. Date of publication 24 September 2024; date of current version 13 November 2024. The work was supported by the U.S. National Science Foundation (NSF) under Grants CNS-2418308, CCF-2427316, and CCF-2426318. The work of Linke Guo was supported by the U.S. NSF under Grants CNS-2008049 and CNS-2431440. The associate editor coordinating the review of this letter and approving it for publication was S. A. Tegos. *(Corresponding author: Lan Zhang.)*

Shuai Lyu, Linke Guo, Xiaoyong Yuan, and Lan Zhang are with the Department of Electrical and Computer Engineering, Clemson University, Clemson, SC 29634 USA (e-mail: slv@clemson.edu; linkeg@clemson.edu; xiaoyon@clemson.edu; lan7@clemson.edu).

Yao Sun is with the James Watt School of Engineering, University of Glasgow, G12 8QB Glasgow, U.K. (e-mail: Yao.Sun@glasgow.ac.uk).

Fang Fang is with the Department of Electrical and Computer Engineering and the Department of Computer Science, Western University, London, ON N6A 5B7, Canada (e-mail: fang.fang@uwo.ca).

Xianbin Wang is with the Department of Electrical and Computer Engineering, Western University, London, ON N6A 5B7, Canada (e-mail: xianbin.wang@uwo.ca).

Digital Object Identifier 10.1109/LCOMM.2024.3466995

explicitly added artificial noise. Based on the task inference sensitivity, this mask softly disentangles the encoded features into robust and non-robust from the semantic level.
- *Numerical evaluation.* We conduct experiments for real-time subchannel allocation problems as a case study. Evaluation results demonstrate the framework's effectiveness, especially under highly dynamic adverse channel conditions.

II. System Model

This letter considers a typical task-oriented semantic communications (TSC) system. As illustrated in Fig. 1, the transmitter comprises semantic and channel encoders. Given an input sample $x \in \mathcal{X}_T$ of task $\mathcal{T}$ with the inherent task-specific semantic information $y \in \mathcal{Y}_T$, *e.g.*, the target label, the transmitter first extracts the semantic information of x with the semantic encoder and then processes it via the channel encoder. The encoded features are given by $z = E_\varphi(x)$, which is also represented as $z = \{z_1, \ldots, z_m\}$, consisting m vectors. Note that we denote E_φ as the joint semantic and channel encoding function for ease of representation, which is consistent with many existing TSC works [4], [8], [11]. Encoded features z are then transmitted through the physical channel to the receiver, and the received signal can be given by $\hat{z} = Hz + n$, $\hat{z} = \{\hat{z}_1, \ldots, \hat{z}_m\}$, where H denotes the channel matrix and $n \sim \mathcal{N}(0, \gamma^2 I)$ denotes the additive white Gaussian noise (AWGN). The received signal is subsequently processed via a channel decoder and a semantic decoder. Similar to the transmitter, we denote D_θ as the joint channel and semantic decoding function and derive the reconstructed semantic information $\hat{y} = D_\theta(\hat{z})$.

The workflow of the above TSC system can be formulated as a probabilistic graphical model: $Y \leftrightarrow X \leftrightarrow Z \leftrightarrow \hat{Z} \leftrightarrow \hat{Y}$. In the following, we use upper-case letters, *e.g.*, X, and lower-case letters, *e.g.*, x, to represent random matrices and their realizations, respectively. Existing TSC research primarily focused on reducing the size of encoded features $z = \{z_1, \ldots, z_m\}$ while ensuring reconstruction performance of semantic information [2], [3], [4], [5], [6], [7]. To achieve this, encoding and decoding functions, *i.e.*, E_φ at the transmitter and D_θ at the receiver, are strategically optimized. We refer readers to recent TSC works [1], [2], [3], [4], [5], [6], [7] for more details.

III. Channel-Resilient TSC Framework

A. Design Intuition

While TSC systems have been extensively studied [1], [2], [3], [4], [5], [6], [7], existing research mainly considers statistical channel conditions to optimize semantic and channel coding functions, assuming that all encoded features of an input sample are transmitted under the same channel condition. However, this assumption may not hold true in practice. For instance, in OFDM systems, multiple subcarriers are used to simultaneously transmit data across a wide band [9]. Hence, *encoded features may suffer distinct channel impairments* due to different frequency-selective fading between subcarriers. Although instantaneous channel state information (CSI) can be

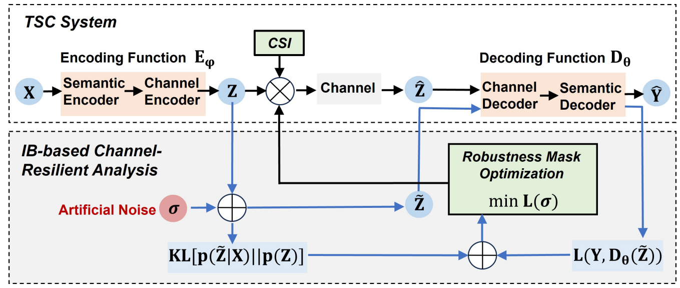


Fig. 1. Overview of channel-resilient TSC framework. Blue arrows indicate IB-based channel-resilient analysis; black arrows indicate transmission procedures in the TSC system.

monitored using estimation techniques, such as pilot symbols embedded in OFDM symbols, the adaptation of transmission strategies is designed to maximize the successful delivery of OFDM symbols. The focus remains on optimizing bit-level transmission performance, *i.e.*, duplicating all encoded features, rather than semantic inference. In other words, *feature units across the encoded signal are assumed equally robust against channel variations*. However, feature units may affect semantic inference differently, *e.g.*, the semantically sensitive or non-robust units are more susceptible to erroneous task inference if corrupted by poor channel conditions. To address these limitations, this letter introduces an innovative framework for channel-resilient TSC by evaluating and prioritizing encoded feature units based on their robustness against channel variations. The framework offers a *unified* solution to complement existing TSC approaches from two perspectives:
- The framework seamlessly integrates with existing TSC approaches by analyzing a well-trained TSC transceiver. Specifically, a soft robustness mask is created for the encoded feature space without modifying the established TSC encoding and decoding functions, *i.e.*, $E\varphi$ and D_θ.
- The robustness mask aims to align semantic-level feature units with instantaneous channel variations. The mask with feature-level semantic distinctions guides transceiver to efficiently adjust transmission strategies based on instantaneous CSI, aiming for task-specific semantic inference.

To achieve this goal, we leverage information bottleneck (IB) [12] to analyze the encoded feature space by explicitly adding artificial noise to synthesize channel variations. Based on the semantic inference sensitivity of artificial noise intervention, a soft feature robustness mask is generated to *indicate how encoded feature units corrupted by channel variations affect semantic inference with assigned information*. In the following, we first present the IB reformulation to control information flow for channel resilience purposes and then provide a tractable solution to obtain the robustness mask. Finally, a case study of leveraging the mask for subchannel allocation is introduced.

B. Information Bottleneck Reformulation

IB is an information theoretical design principle that aims to find the best tradeoff between accuracy and complexity [12]. Several recent works leveraged the IB principle in TSC to seek the tradeoff between transmission rate and semantic information distortion [4], [13]. Following the aforementioned probabilistic graphical model in Section II, $Y \leftrightarrow X \leftrightarrow Z \leftrightarrow \hat{Z} \leftrightarrow \hat{Y}$, the objective of optimizing the encoding and decoding functions, *i.e.*, φ and θ, based on the IB principle can be given by

$$\min L(\varphi,\theta) = -I(\hat{Z},Y) + \beta I(\hat{Z},X), \tag{1}$$

where I is the mutual information; β is the Lagrange multiplier that regulates the amount of information in feature space $\hat{Z}$. In (1), the first term minimizes semantic distortion by allowing the received encoded features $\hat{Z}$ to be predictive on semantic information Y, and the second term maximizes transmission rate by enforcing the compression from the input X to received encoded features $\hat{Z}$.

Instead, this letter intends to improve the channel resilience of a well-trained TSC transceiver, whose encoding and decoding functions (φ,θ) are frozen.[1] To estimate how encoded features corrupted by channel variations affect TSC, we reuse the objective in (1) by introducing the artificial noise σ to simulate channel impairments, which control the information flow and estimate the robustness of the encoded feature space. Hence, the estimation for channel-resilient TSC can be formulated as

$$\begin{aligned} \min \quad & L(\sigma) = -I(\tilde{Z},Y|\varphi,\theta) + \beta I(\tilde{Z},X|\varphi,\theta),\\ s.t. \quad & \tilde{Z} = Z + \sigma\cdot\epsilon, \end{aligned} \tag{2}$$

where $\tilde{Z}$ stands for the encoded features corrupted by channel impairments, such as fading and noise. To model the effect of channel impairments, we inject artificial noise $\sigma\cdot\epsilon$ to the encoded features $Z = E_\varphi(X)$, where the operator $\cdot$ denotes the Hadamard product and ϵ represents the Gaussian noise sampled from $\mathcal{N}(0,I)$. Thus, we obtain $\tilde{Z} \sim \mathcal{N}(E_\varphi(X),\sigma^2)$. Here, the noise variation measures the correlation between the encoded features and the inferred semantic information based on the fact that robustness refers to a high correlation on semantic inference under impairments and non-robustness is the opposite [14]. As (φ,θ) are frozen throughout the optimization for channel resilience analysis, we exclude φ and θ in the following notation for simplicity.

C. Upper Bound of the IB-Based Robustness Estimation

To calculate the achievable artificial noise, we resolve the difficulty of mutual information computation in (2) by deriving the upper bound of the objective function. We start with the first term $I(\tilde{Z},Y)$. Writing it out in full, this becomes

$$\begin{aligned} I(\tilde{Z},Y) &= \int p(y,\tilde{z})\log\frac{p(y,\tilde{z})}{p(y)p(\tilde{z})}dyd\tilde{z}\\ &= \int p(y,\tilde{z})\log p(y|\tilde{z})dyd\tilde{z} + H(Y), \end{aligned} \tag{3}$$

where $p(y)$ and $p(\tilde{z})$ are the probability of the semantic information and encoded features with artificial noise, respectively. $H(Y)$ is the entropy of the semantic information that is independent of the optimization and thus ignored. Recalling the probabilistic graphical model $Y \leftrightarrow X \leftrightarrow Z$, we rewrite $p(y,\tilde{z})$ based on the underlying characteristics of the Markov chain,

$$p(y,\tilde{z}) = \int p(x)p(z|x)p(\tilde{z}|z)p(y|\tilde{z})dxdz. \tag{4}$$

Then, we derive the upper bound of $-I(\tilde{Z},Y)$ as

$$\begin{aligned} -I(\tilde{Z},Y) &\le -\int p(y,\tilde{z})\log p(y|\tilde{z})dyd\tilde{z}\\ &= \mathop{\mathbb{E}}_{X\sim p(X),Z\sim p_\varphi(Z|X),\tilde{Z}\sim p(\tilde{Z}|Z)}[\mathcal{L}(Y,D_\theta(\tilde{Z}))], \end{aligned} \tag{5}$$

where $\mathcal{L}$ is the cross-entropy loss. Since we evaluate the robustness of encoded features from a well-trained TSC transceiver, we have $p_\varphi(Z|X)$ and $D_\theta(\tilde{Z})$. Besides, the corrupted features with artificial noise are given by $\tilde{Z} \sim \mathcal{N}(E_\varphi(X),\sigma^2)$.

Next, we focus on the second mutual information, $I(\tilde{Z},X)$, in (2) and have

$$\begin{aligned} I(\tilde{Z},X) &= \int p(\tilde{z},x)\log\frac{p(\tilde{z},x)}{p(\tilde{z})p(x)}d\tilde{z}dx\\ &= \int p(\tilde{z},x)\log\frac{p(\tilde{z}|x)}{p(z)}dzd\tilde{z}dx\\ &\quad + \int p(\tilde{z},x)\log\frac{p(z)}{p(\tilde{z})}dzd\tilde{z}dx\\ &= KL[p(\tilde{Z}|X)||p(Z)] - KL[p(\tilde{Z})||p(Z)], \end{aligned} \tag{6}$$

where KL represents Kullback–Leibler divergence that measures the difference between two probability distributions. Since $KL[p(\tilde{Z})||p(Z)] \ge 0$, we have

$$\begin{aligned} I(\tilde{Z},X) &\le KL[p(\tilde{Z}|X)||p(Z)]\\ &= \frac{1}{2}\sum_{k=1}^{m}[\frac{\sigma_k^2}{\delta_k^2} + \log\frac{\delta_k^2}{\sigma_k^2} - 1], \end{aligned} \tag{7}$$

where k denotes the index of the artificial noise variation added to the kth encoded feature unit, *i.e.*, $\sigma = [\sigma_1,\cdots,\sigma_m]$, and $\delta = [\delta_1,\cdots,\delta_m]$ represents the inherent variation of the encoded feature over the task, reflecting its natural fluctuations without channel impairments. Built on the above derivation, we obtain the non-negative upper bound of the objective function in (2)

$$\begin{aligned} L(\sigma) &\le L'(\sigma)\\ &= \mathop{\mathbb{E}}_{X\sim p(X),Z\sim p_\varphi(Z|X),\tilde{Z}\sim\mathcal{N}(E_\varphi(X),\sigma^2)}[\mathcal{L}(Y,D_\theta(\tilde{Z}))]\\ &\quad + \beta\left(\frac{1}{2}\sum_{k=1}^{m}[\frac{\sigma_k^2}{\delta_k^2} + \log\frac{\delta_k^2}{\sigma_k^2} - 1]\right). \end{aligned} \tag{8}$$

Therefore, by propagating $\tilde{Z}$ through the decoding function, the artificial noise can be optimized by $\sigma = \sigma - \frac{\partial L'(\sigma)}{\partial\sigma}$.

D. Robustness Mask and Case Study

After optimizing artificial noise σ, we analyze the artificially corrupted encoded features $\tilde{Z}$ and assess the robustness of each feature unit based on its sensitivity to semantic inference. Define the encoded feature variation of task $\mathcal{T}$ by $R = \max(\delta^2)$, which represents the maximal input variations mapping to the encoded feature space Z. Intuitively, the injected artificial noise should be restricted below R to ensure

[1]Note unlike adversarial training [14] that updates parameters for model robustness, we analyze the fixed model and estimate the feature-level robustness.

inference reliability. However, recent research indicates that the correlation between different units in the feature space and inference performance are different [15]. Thus, feature unit $z_k \in z = \{z_1, \ldots, z_m\}$ with a high correlation to inference results should have $\sigma_k^2 > R$, *i.e.*, robust against channel impairment, and z_k with $\sigma_k^2 < R$ is non-robust since small channel impairment behaves as a strict restriction to retain semantic inference performance. Therefore, we explicitly disentangle encoded features into robust and non-robust against channel variations for TSC.

Remark: By optimizing $\sigma_k \in \sigma$ for each unit of the encoded features, the IB-based problem formulation in (2) controls the information flow to the decoding function and evaluates feature unit robustness. We approximate the distribution of task $\mathcal{T}$ with empirical risk minimization over the entire dataset $\mathcal{X}_T$. Hence, the soft robustness mask of feature unit $z_k \in z = \{z_1, \ldots, z_m\}$ is given by

$$r_k = \frac{\sum_{x_i \in \mathcal{X}_T} \sigma_k^i}{\sum_{x_j \in \mathcal{X}_T} \sum_{l=1}^{m} \sigma_l^j}, \quad \sum_{k=1}^{m} r_k = 1. \tag{9}$$

This mask can be leveraged by a well-trained TSC transceiver (*i.e.*, given E_φ and D_θ) for task $\mathcal{T}$ to accommodate instantaneous channel variations. Based on the robustness score, priority is provided for transmitting the encoded feature units to achieve reliable semantic inference.

We conduct a case study to leverage the robustness mask r for subchannel allocation. Feature units with a small robustness score are considered non-robust against channel impairments, which should be assigned to high-quality subchannels that are measured by channel estimation techniques [9]. Given a set of available subchannels with distinct channel conditions, we apply a greedy-based method to assign subchannels to the m encoded feature units following their pre-sorted robustness mask $r = [r_1, \ldots, r_m]$. Note that the pre-known r will be reused for future subchannel allocations, and the computational complexity of each subchannel allocation can be given as $O(m)$. The pseudocode of the robustness mask-based subchannel allocation is given in Algorithm 1.

IV. Numerical Evaluation

A. Evaluation Setup

We consider a TSC task using the CIFAR-10 [16] dataset for image classification. We adopt the VGG16 [17] architecture as the encoding function E_φ and three fully connected layers as the decoding function D_θ. We consider AWGN and fading channel conditions to train two types of TSC transceivers, respectively. The transceivers are frozen for channel resilience analysis. For testing, we evaluate performance under frequency-selective channel conditions, specifically considering a typical OFDM symbol structure with 272 subcarriers, including 16 pilots evenly distributed across 256 data subcarriers [10]. The number of encoded feature units is 512. The hyperparameter β of IB is empirically set as 0.3.

B. Evaluation Results

We first evaluate the effectiveness of the robustness mask for encoded feature units using the AWGN-trained transceiver as an example. According to the robustness scores, we rerank encoded feature units. The units in the first half, which have higher scores, are expected to be more resilient against channel variations than those in the second half, which have lower scores. Fig. 2 visualizes the feature-level inference performance between ideal and noisy channel conditions using 2D t-SNE. In ideal channel environments, *i.e.*, no noise, both robust (first-half) feature units and non-robust (second-half) feature units achieve comparable inference performance. However, under noisy conditions, *i.e.*, SNR $= 0$, the performance

---

Algorithm 1 Channel-Resilient TSC: Subchannel Allocation

Input: (1) a well-trained TSC transceiver with encoding function E_φ and decoding function D_θ; (2) input data and its corresponding semantic information $(x, y) \in (\mathcal{X}_T, \mathcal{Y}_T)$; (3) instantaneous $CSI = \{CSI_1, \ldots, CSI_s\}$ for s subchannels;

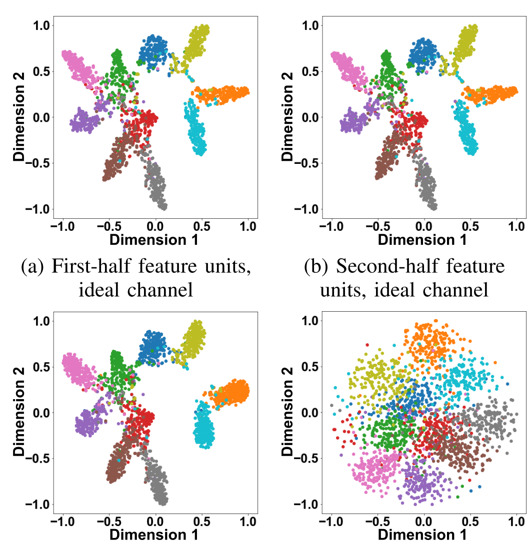
Output: (1) robustness mask (2) subchannel assignment

// *Robustness mask generation*

1) **Encode**: $z = E_\varphi(x) = \{z_1, \ldots, z_m\}$, $x \in \mathcal{X}_T$.
2) **Initialize artificial noise**: $\sigma = 0$ with a neutral noise.
3) **For each iteration**:
   a) **Adjust noise**: $\sigma = \log(1+\exp(\sigma))$, ensure positive through SoftPlus.
   b) **Inject noise**: $\tilde{z} = z + \sigma \cdot \epsilon$ based on (2).
   c) **Decode**: $\hat{y} = D_\theta(\tilde{z})$.
   d) **Update**: $\sigma = \sigma - \frac{\partial L(\sigma)}{\partial \sigma}$ based on (8).
4) **End For**
5) **Robustness mask**: $r = \{r_1, \ldots, r_m\}$ based on (9).

// *Robustness mask implementation: assign subchannels*

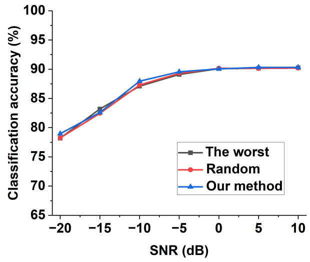
1) **For** $l \in \{1, \ldots, m\}$ **do**:
   a) Assign z with the smallest r_i to subchannel with the best CSI_j
   b) Remove jth subchannel
2) **End for**



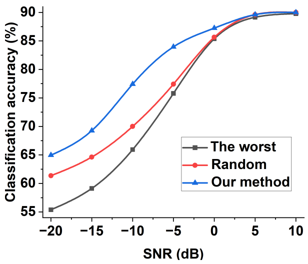
(a) First-half feature units, ideal channel (b) Second-half feature units, ideal channel (c) First-half feature units, noisy channel (d) Second-half feature units, noisy channel

Fig. 2. Comparison of feature-level channel resilience between the ideal (noise = 0) and noisy (SNR = 0) channel conditions.

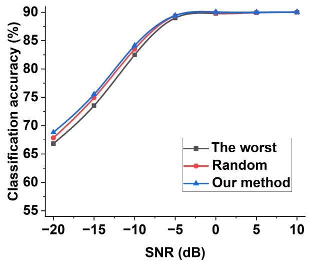
(a) High Variance, AWGN-Trained Transceiver



(b) Low Variance, AWGN-Trained Transceiver



(c) High Variance, Fading-Trained Transceiver



(d) Low Variance, Fading-Trained Transceiver

Fig. 3. Comparison of inference performance between dynamic (high variation, 15) and stable (low variation, 2) subchannel environments with transceivers trained under AWGN and fading conditions on CIFAR-10 dataset.

disparity between the robust and non-robust feature units becomes apparent. Despite channel impairments, the first half of the units with higher robustness scores maintain performance levels similar to those in ideal conditions ((a) vs. (c)), whereas the performance of the second half deteriorates significantly ((b) vs. (d)). This contrast demonstrates the effectiveness of the robustness mask in evaluating feature-level channel resilience.

We then evaluate the effectiveness of using the robustness mask for subchannel allocation with AWGN-trained transceiver and fading-trained transceiver by low and high variations. The low variation represents stable subchannel environments, while the high variation denotes highly dynamic discrepancies between subcarriers, regardless of their average performance. Fig. 3 illustrates the inference accuracy with on-average subchannel performance (SNR) between different channel variances with different transceivers. Specifically, Fig. 3(a) and Fig. 3(c) present inference performance under highly dynamic subchannel conditions. We compare our method against two baselines: random and worst-case allocations, the latter assigning the lowest-quality subchannels (low SNR) to the least robust feature units (low robustness score). We observe that our method consistently outperforms the baselines across all SNR levels and with different transceivers, which demonstrates its effectiveness. As the SNR increases, the performance gap between our method and the baselines narrows on both datasets. This suggests that satisfactory average subchannel performance can compensate for subchannel variations. *The strength of our method is particularly apparent in highly dynamic and challenging channel conditions.* Furthermore, by comparing Fig. 3(b) and Fig. 3(d), we observe a marginal difference between the three methods in stable channel environments. This consistently validates our finding above.

V. Conclusion

This letter introduced an innovative framework for channel-resilient task-oriented semantic communications (TSC), which analyzes the encoded feature space of a well-trained TSC transceiver. A robustness mask for encoded feature units is created based on the information bottleneck, which was implemented for subchannel allocation as a case study. The effectiveness of this framework was validated, especially under highly dynamic adverse channel conditions. To the best of our knowledge, this letter represents the first effort to enhance channel resilience for TSC using the IB principle. We aim to pave the way for exploring other channel resilience techniques for TSC from information-theoretic and signal-processing perspectives.

References

[1] T. M. Getu, G. Kaddoum, and M. Bennis, "A survey on goal-oriented semantic communication: Techniques, challenges, and future directions," *IEEE Access*, vol. 12, pp. 51223–51274, 2024.

[2] 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.

[3] Y. Yang et al., "Semantic communications with artificial intelligence tasks: Reducing bandwidth requirements and improving artificial intelligence task performance," *IEEE Ind. Electron. Mag.*, vol. 17, no. 3, pp. 4–13, Mar. 2022.

[4] J. Shao, Y. Mao, and J. Zhang, "Learning task-oriented communication for edge inference: An information bottleneck approach," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 1, pp. 197–211, Jan. 2022.

[5] L. Sun, Y. Yang, M. Chen, C. Guo, W. Saad, and H. V. Poor, "Adaptive information bottleneck guided joint source and channel coding for image transmission," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 8, pp. 2628–2644, Aug. 2023.

[6] H. Xie, Z. Qin, and G. Y. Li, "Task-oriented multi-user semantic communications for VQA," *IEEE Wireless Commun. Lett.*, vol. 11, no. 3, pp. 553–557, Mar. 2021.

[7] C. Liu, C. Guo, Y. Yang, and N. Jiang, "Adaptable semantic compression and resource allocation for task-oriented communications," *IEEE Trans. Cognit. Commun. Netw.*, vol. 10, no. 3, pp. 769–782, Jun. 2024, doi: 10.1109/TCCN.2023.3346481.

[8] F. Zhao, G. Bagwe, E. Mohammed, L. Feng, L. Zhang, and Y. Sun, "Joint computing resource and bandwidth allocation for semantic communication networks," in *Proc. IEEE 98th Veh. Technol. Conf. (VTC-Fall)*, Oct. 2023, pp. 1–5.

[9] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," *IEEE Trans. Broadcast.*, vol. 48, no. 3, pp. 223–229, Sep. 2002.

[10] C. Liu, C. Guo, Y. Yang, W. Ni, and T. Q. S. Quek, "OFDM-based digital semantic communication with importance awareness," *IEEE Trans. Commun.*, early access, May 7, 2024, doi: 10.1109/TCOMM.2024.3397862.

[11] S. Ma et al., "Task-oriented explainable semantic communications," *IEEE Trans. Wireless Commun.*, vol. 22, no. 12, pp. 9248–9262, Dec. 2023.

[12] N. Tishby, F. C. Pereira, and W. Bialek, "The information bottleneck method," 2000, *physics/0004057*.

[13] S. Xie, S. Ma, M. Ding, Y. Shi, M. Tang, and Y. Wu, "Robust information bottleneck for task-oriented communication with digital modulation," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 8, pp. 2577–2591, Aug. 2023.

[14] A. Ilyas, S. Santurkar, D. Tsipras, L. Engstrom, B. Tran, and A. Madry, "Adversarial examples are not bugs, they are features," in *Proc. Adv. neural Inf. Process. Syst.*, vol. 32, 2019, pp. 1–12.

[15] D. Tsipras, S. Santurkar, L. Engstrom, A. Turner, and A. Madry, "Robustness may be at odds with accuracy," 2018, *arXiv:1805.12152*.

[16] A. Krizhevsky and G. Hinton, "Learning multiple layers of features from tiny images," Univ. Toronto, Toronto, ON, Canada, Tech. Rep., 2009. [Online]. Available: https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf

[17] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.