Feature-Importance-Aware Transmission Control in Semantic Communications

Seonghun Hong[a], Donghyun Lee[a], Dongwook Won[a], Wonjong Noh[b,*], Sungrae Cho[a,*]

[a] School of Computer Science and Engineering, Chung-Ang University, Seoul 06974, South Korea
[b] School of Software, Hallym University, Chuncheon 24252, South Korea

Abstract

Semantic communication shifts the focus from bit-level accuracy to task-relevant meaning. However, most methods assume equal importance across semantic units and rely on costly retraining, limiting scalability. This work proposes an importance-aware framework that accounts for unequal feature contributions. It introduces a new metric, importance-weighted semantic spectral efficiency (wSSE), to prioritize task-relevant features. It develops an empirically derived feature-accuracy matrix, inspired by saturation behavior. The framework enables importance-aware feature selection and subchannel allocation without online learning. This framework targets resource-constrained edge systems such as IoT cameras, UAV detection, and AR/VR. Experiments show up to 73.3% higher efficiency under constrained resources.

Keywords: Feature Importance, Importance-Weighted Semantic Spectral Efficiency, Semantic Communication

1. Introduction

The rapid development of data-driven services such as autonomous driving, smart healthcare [1, 2], and immersive AR/VR is pushing current communication systems to their limits [3, 4, 5]. Recent 6G visions emphasize AI-native control for intelligent, adaptive, and efficient wireless systems [6], prioritizing task-level performance and energy efficiency [7]. Conventional bit-level communication operates in a content-agnostic manner, transmitting all bits with full fidelity regardless of task relevance. This approach consumes significant bandwidth and energy to achieve task success [8].

Semantic communication (SC) has been proposed to overcome these limitations. Recent advances in deep learning have enabled the end-to-end extraction and encoding of semantic features [9, 10, 11, 12]. Instead of delivering every bit of data, SC extracts semantic information such as entities, relations, or other task-relevant features [13, 14]. It then transmits a meaningful form that the receiver can accurately interpret [15]. The goal is not to send exact words but to ensure mutual understanding of the intended message. In practice, SC uses a semantic encoder to detect and compress task-relevant meaning [16]. A semantic decoder then reconstructs or infers the conveyed intent, using shared knowledge bases and contextual information [17]. This demonstrates that it can maintain task accuracy while significantly reducing transmission overhead.

As SC evolves, some studies have introduced semantic metrics [18, 19]. These metrics evaluate task-level efficiency and user-centric utility rather than bit fidelity. However, they typically assume uniform importance across semantic features, overlooking that some features carry far greater weight in task performance. Especially, Cheng et al. emphasize that the importance of semantic features significantly impacts the task [20].

Inspired by [20], this work extends existing metrics by introducing importance-weighted semantic spectral efficiency (wSSE) to fill this gap. It explicitly accounts for unequal feature contributions. The value of redefining metrics is evident in other communication domains. New evaluation measures have reshaped how system performance is assessed [21]. Based on the proposed metric, an importance-aware end-to-end transmission control framework is also proposed. It enables efficient feature selection and subchannel assignment without expensive online learning. This is achieved by leverag-

*Corresponding author
Email addresses: shhong@uclab.re.kr (Seonghun Hong), dhlee@uclab.re.kr (Donghyun Lee), dwwon@uclab.re.kr (Dongwook Won), wonjong.noh@hallym.ac.kr (Wonjong Noh), srcho@cau.ac.kr (Sungrae Cho)

ing an empirically derived task accuracy matrix. It also achieves high accuracy while reducing the number of transmitted features and improving semantic efficiency under limited resources.

1.1. Related Work

Early research in SC introduced semantic-specific performance metrics. The semantic rate (S-R) and semantic spectral efficiency (S-SE) were proposed in [18, 19] to measure task-relevant meaning per unit time and bandwidth. Other semantic-aware metrics include semantic system throughput [22], semantic-aware quality of experience [23, 24] and compression-aware utility functions [25].

In the area of semantic-aware transmission control, a recent study [26] addressed transmission control by formulating a mixed-integer programming model that jointly allocates power and assigns subchannels, with semantic similarity serving as the optimization objective.

Zhang et al. [27] applied a deep reinforcement learning framework to adapt transmission decisions in real time, enabling robust semantic delivery under dynamically changing channel conditions. Li et al. [28] optimized the transmission pipeline at the training stage by allocating learning resources to enhance feature-level discrimination, indirectly improving the encoder's ability to support efficient transmission at inference time. In [29], a transformer-based encoder with a semantic-aware loss is proposed to produce representations resilient to transmission constraints. This design maintains higher task performance with limited resources. However, the studies mentioned above focused on accurately capturing meaning rather than distinguishing the degree of contribution to the task.

Han et al. [10] improves semantic-aware transmission via a patch-wise scheduler guided by an image-captioning attention map. For each input image, an attention heatmap is computed online and used to rank patches. Higher-scored regions are transmitted first so that caption quality is preserved with less bandwidth. However, it prioritizes image regions based on attention scores from an image captioning model, which limits generalization to other modalities or tasks.

1.2. Contribution

Based on the above motivation, we propose an importance-aware framework. The contributions are summarized as follows:

- An importance-aware end-to-end transmission control framework is presented. It integrates per-feature importance into feature selection and subchannel assignment. It departs from uniform-importance or similarity-only policies. Combining selection and assignment allows for efficient, importance-aware control.
- wSSE is redefined by extending S-SE to weight features based on task relevance. wSSE provides an explicit efficiency-accuracy objective. It offers a task-grounded alternative to capacity-based or unweighted semantic metrics.
- To avoid online retraining to identify important features, the empirical accuracy surface is built as an offline table. The feature-importance vector is obtained through ablation and precomputed based on the SNR condition. This table is used to select transmission features. Then, it matches users to favorable subchannels. This two-step framework importance-aware control with negligible transmission overhead.
- Simulations show that the proposed framework achieves higher task accuracy and higher semantic efficiency than benchmarks at all power levels. These results underscore the practical value of prioritizing high-importance features.

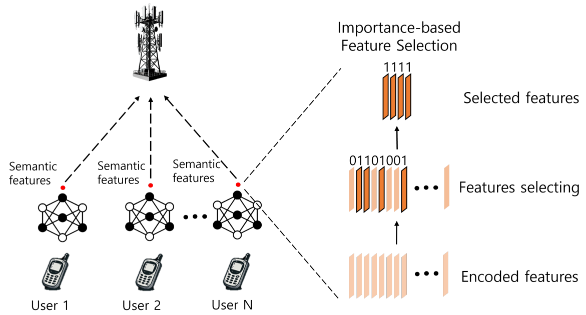

Figure 1. Structure of the importance-weighted semantic network. The source data is first encoded into semantic features, and each feature is assigned a weight according to its task relevance. High-importance features are selected and mapped to sub-channels, while low-importance features may be pruned. The receiver reconstructs the task output using the features.

2. System Model and Problem Formulation

This paper considers a cellular network comprising a base station (BS), with a set of N users denoted by $\mathcal{N} = \{1, 2, \ldots, i, \ldots, N\}$, as depicted in Fig. 1. A task-oriented DeepSC framework is adopted as the SC model, where the encoder and decoder are optimized to preserve the semantic information required for executing a specific task, such as image classification or object recognition.

Each user has a task-driven semantic transmitter that extracts high-level task-relevant features from the input source using deep neural networks (e.g., convolutional neural networks or transformer-based encoders [30, 31, 32, 33]). Instead of transmitting the entire original content, each user transmits only the task-critical semantic features to the BS. These representations are learned to be robust to subchannel distortion while maintaining task performance.

The task-oriented DeepSC transceiver is assumed to be pretrained off-line at the BS or on a cloud-based training server using task-specific loss functions. After training, the optimized semantic encoder is distributed to the users for local use.

The following section presents the semantic transmitter structure at the user side, the transmission model over the wireless subchannel, and the task-oriented semantic decoder deployed at the BS.

2.1. DeepSC Transmitter

In the proposed model, the ith user generates an image X_i, which is encoded into semantic features $f_{i,j}$ using a DeepSC-based encoder, as follows:

$$\mathbf{X}_i^{\text{enc}} = [f_{i,1}, f_{i,2}, f_{i,j}, \ldots, f_{i,F}], \tag{1}$$

where F denotes the number of features, depending on the encoder architecture. Assuming that the number of generated semantic features is fixed. Furthermore, the importance of each semantic feature varies depending on the task requested by the user. Each feature $f_{i,j}$ is associated with an importance weight $\omega_{i,j}$, formed as follows:

$$\boldsymbol{\omega}_i = [\omega_{i,1}, \omega_{i,2}, \ldots, \omega_{i,F}], \quad \text{where} \sum_{j=1}^{F} \omega_{i,j} = 1. \tag{2}$$

Let $\mathcal{M} = \{1, 2, \ldots, m, \ldots, M\}$ denote the set of available subchannels in the network, where M represents the number of subchannels. A feature selection vector $\boldsymbol{\beta} = \{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_N\}$ is defined to indicate whether the ith user transmits the jth semantic feature or not, where $\boldsymbol{\beta}_i = [\beta_{i,1}, \beta_{i,2}, \ldots, \beta_{i,F}]$, and $\beta_{i,j} = 1$ if feature $f_{i,j}$ is transmitted, and $\beta_{i,j} = 0$, otherwise. The selection rate k_i is given by the following:

$$k_i = \frac{\sum_{j=1}^{F} \beta_{i,j} |f_{i,j}|}{|X_i|}, \tag{3}$$

where $|f_{i,j}|$ and $|X_i|$ denote the (bit-level) size of feature $f_{i,j}$ and image X_i, respectively. Hence, each user transmits semantic representations derived from the original image X_i, which are semantically encoded to facilitate task execution.

The selection matrix $\mathbf{D}_i \triangleq \mathrm{diag}(\boldsymbol{\beta}_i) \in \{0, 1\}^{F \times F}$ formalizes the feature selection. With the encoder output $\mathbf{X}_i^{\text{enc}}$, the transmitted semantic vector becomes

$$\widetilde{\mathbf{X}}_i = \mathbf{X}_i^{\text{enc}} \mathbf{D}_i = [\, f_{i,1}\beta_{i,1}, \ \ldots, \ f_{i,F}\beta_{i,F} \,], \tag{4}$$

so that only features with $\beta_{i,j} = 1$ are sent, and the others are null.

2.2. Transmission Model

In this paper, considering an orthogonal frequency division multiplexing (OFDM) based system where each user is assigned to at most one orthogonal subchannel. The wireless subchannel between user i and the BS includes large-scale fading g_i and small-scale Rayleigh fading $h_{i,m}$. Thus, the SNR is given by

$$\gamma_{i,m} = \frac{P_i g_i |h_{i,m}|^2}{W N_0}, \tag{5}$$

where P_i, W, and N_0 denote the transmit power, bandwidth, and noise power spectral density, respectively. Furthermore, the subchannel allocation is managed using a binary variable $\boldsymbol{\alpha} = \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_N\}$, where $\boldsymbol{\alpha}_i = [\alpha_{i,1}, \alpha_{i,2}, \ldots, \alpha_{i,M}]^T$. Further, $\alpha_{i,m} \in \{0, 1\}$, that is, $\alpha_{i,m} = 1$ indicates that user i is assigned with subchannel m, and $\alpha_{i,m} = 0$, otherwise. Each user and subchannel is limited to a one-to-one pairing:

$$\sum_{i=1}^{N} \alpha_{i,m} \leq 1, \ \forall m \quad \text{and} \quad \sum_{m=1}^{M} \alpha_{i,m} \leq 1, \ \forall i. \tag{6}$$

2.3. Task-Oriented DeepSC Receiver

At the BS, the signal from the ith user assigned to subchannel m can be expressed as follows:

$$\mathbf{Y}_i = \sqrt{g_i} h_{i,m} (\widetilde{\mathbf{X}}_{i,m}) + \mathbf{z}, \tag{7}$$

where $\mathbf{z}$ denotes the additive white Gaussian noise , and each element of $\mathbf{z}$ follows $\mathcal{CN}(0, N_0)$. The received signal is decoded using the subchannel decoder and is further processed using the semantic decoder to reconstruct the original semantic content X_i.

2.4. Task Performance Metric

To evaluate the semantic transmission performance, a task accuracy metric ξ is defined as the average outcome of the binary task performance function $f_{\text{succ}}(\cdot)$ applied to each received signal $\mathbf{Y}$ under (q, γ) conditions. Here,

q denotes the importance level, defined as the cumulative sum of the importance weights of the selected features (i.e., $q = \boldsymbol{\beta}\boldsymbol{\omega}^\top$)

$$\xi(q,\gamma) = \mathbf{E}_{\mathbf{Y}\sim prob(q,\gamma)}[f_{\text{succ}}(\mathbf{Y})] \approx \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} f_{\text{succ}}(\mathbf{Y_t}), \tag{8}$$

where T denotes the number of samples , and $f_{\text{succ}}(\cdot)$ outputs 1 if the task is successful and outputs 0 otherwise.

The task accuracy ξ is in the range $[0,1]$, where $\xi = 1$ indicates perfect alignment in task-relevant features, and $\xi = 0$ indicates a complete mismatch with respect to the task objective.

2.5. *Importance-Weighted Semantic Information*

S-SE is extended by introducing wSSE, which evaluates each feature's task contribution and prioritizes high-utility features under constraints. The wSSE $\Phi_{i,m}$ is defined using the task-accuracy function $\xi(q_{i,m},\gamma_{i,m})$ as follows:

$$\Phi_{i,m} = \frac{I}{Fk_i}\xi_{i,m}(\boldsymbol{\beta_i}\boldsymbol{\omega_i^T},\gamma_{i,m}),. \tag{9}$$

where $I = \sum_{j=1}^{F} -P(f_{i,j})\log P(f_{i,j})$ denotes the expected amount of semantic information contributed by the features, and $P(f_{i,j})$ represents the probability of occurrence of feature $f_{i,j}$ in the dataset or task context.

2.6. *Problem Formulation*

The objective is to maximize the system-wide wSSE, measuring task performance per unit bandwidth.

$$\textbf{(P1)} \quad \max_{\alpha,\beta} \sum_{i=1}^{N}\sum_{m=1}^{M} \alpha_{i,m}\Phi_{i,m}(\boldsymbol{\beta_i}\boldsymbol{\omega_i^T},\gamma_{i,m}), \tag{10a}$$

$$\text{s.t.} \quad C_1: \quad \alpha_{i,m}\in\{0,1\}, \quad \forall i\in\mathcal{N},\ \forall m\in\mathcal{M}, \tag{10b}$$

$$C_2: \quad \sum_{i=1}^{N}\alpha_{i,m}\le 1, \quad \forall m\in\mathcal{M}, \tag{10c}$$

$$C_3: \quad \sum_{m=1}^{M}\alpha_{i,m}\le 1, \quad \forall i\in\mathcal{N}, \tag{10d}$$

$$C_4: \quad \beta_{i,j}\in\{0,1\}, \quad \forall i\in\mathcal{N}, \forall j\in\{1,\dots,F\}, \tag{10e}$$

$$C_5: \quad \xi_{i,m}\ge\xi_{\text{th}}, \quad \forall i\in\mathcal{N},\ \forall m\in\mathcal{M}. \tag{10f}$$

In **(P1)**, (10a) denotes the objective function that finds optimal subchannel allocation and feature selection.

Constraints (10b), (10c), and (10d) are subchannel assignment constraints.

Constraint (10e) represents binary feature selection, and (10f) guarantees that the success rate of the task, $\xi_{i,m}$, for reliable task execution meets a minimum threshold ξ_{th}.

However, due to the non-convexity of **(P1)**, it is challenging to solve using conventional convex optimization methods.

3. Proposed Approach

A semantic-importance-aware framework is designed to integrate feature selection with subchannel assignment for the formulated optimization problem.

3.1. *Feature-Importance Task-Accuracy Table*

A task-accuracy table is first constructed that maps combinations of semantic-importance levels and SNR values to task accuracy, enabling semantic-aware optimization.

3.1.1. *Feature-Importance Extraction*

A feature-importance vector $\boldsymbol{\omega_i}$ is extracted using a pretrained semantic transceiver. An ablation-based estimation is employed, where the importance $\omega_{i,j}$ of feature $f_{i,j}$ is measured as the normalized decline in task performance when that feature is masked. This method, widely employed in interpretable deep learning [34], eludes gradient backpropagation and yields stable, task-specific importance values. Concretely, a one-at-a-time ablation is performed with the transceiver frozen at a nominal SNR to decouple channel randomness. For each feature $f_{i,j}$, let ξ_{full} denote the batch-averaged task accuracy when all features are used, and ξ_{-j} the accuracy when $f_{i,j}$ is masked (set to zero). The feature's contribution is defined as the non-negative drop in accuracy caused by masking. Contributions across all features are then normalized so that the resulting importance scores are comparable and sum to one, yielding a probability-like importance vector. To improve stability and reproducibility, the ablation is repeated on randomly drawn mini-batches, and the final importance scores are obtained by averaging over these runs [35].

3.1.2. *Quantization of Total Importance*

The total importance score q is discretized into $|Q|$ representative levels, $Q = \{\tilde{q}_1,\tilde{q}_2,\tilde{q}_l,\dots,\tilde{q}_{|Q|}\}$.

For each $\tilde{q}\in Q$, the corresponding selected feature subset $\mathcal{L}_{i,\tilde{q}_l}$ is determined, whose cumulative feature importance is closest to $\tilde{q}$.

If multiple subsets satisfy this condition, the subset with the smallest number of features is selected to minimize transmission costs and enhance efficiency, while still meeting the intended importance level.

The set $\mathcal{L}_i = \{\mathcal{L}_{i,\tilde{q}_l} \mid \tilde{q}_l \in Q\}$ represents the *feature combination table*, mapping each quantized importance level to its selected feature set.

3.1.3. Construction of Task-Accuracy Table

This approach simulates transmission over varying SNR conditions $\gamma \in \mathcal{G}$ using a deep learning-based semantic transceiver (e.g., DeepSC [23]). For each $(\tilde{q}, \gamma)$ pair, the simulation is performed using the corresponding feature subset $\mathcal{L}_{i,\tilde{q}}$, and the resulting value of the semantic task-accuracy function $\xi(\tilde{q}, \gamma)$ is recorded in a *task-accuracy table* $\mathcal{T}_i(q, \gamma)$. This method ensures that the table values are empirically grounded rather than based on synthetic models. Avoiding closed-form evaluation or real-time decoding is suitable for delay-sensitive applications. Empirically, the task-accuracy function $\xi(q, \gamma)$ exhibits saturation-like characteristics: accuracy increases sharply in the low-to-mid ranges of importance sum q or SNR γ, and gradually flattens at higher levels. This S-shaped behavior, consistent with observations in semantic similarity studies [19, 36]. It motivated the choice of quantization levels to capture both the steep growth and plateau regions. The constructed table reflects realistic diminishing returns in semantic task performance. Algorithm 1 summarizes this process.

3.2. Proposed Optimization Framework

Based on the constructed feature-importance-based task-accuracy table, this section describes the proposed algorithm that maximizes the total wSSE of the system under the given constraints. Algorithm 2 summarizes the complete optimization pipeline.

3.2.1. Optimal Semantic Feature Selection (Phase-1)

A wSSE-feature matrix $\mathcal{F}$ is constructed for all users and subchannels, where $\mathcal{F}$ of $(\mathbb{R}, \text{feature set})^{N \times M}$. For each subchannel m of user i, the SNR $\gamma_{i,m}$ is identified, and the candidates quantized importance levels $\tilde{q}_l \in Q$ are retrieved from the precomputed table $\mathcal{L}_i$ to satisfy the accuracy constraint in (10f), that is, $\mathcal{T}(\tilde{q}_l, \gamma_{i,m}) \geq \xi_{\text{th}}$.

For a given $\gamma_{i,m}$, among the feature sets corresponding to these candidates $\tilde{q}_l$, this approach selects the one that yields the maximum wSSE, preferring subsets with fewer features when multiple sets achieve the same wSSE value. Then, the selected feature set and its wSSE value are stored in $\mathcal{F}_{i,m}$. For example, if $\tilde{\boldsymbol{q}}$ is quantized as $[0, \ldots, 0.30, 0.32, 0.34]$, $q = 0.31$, $\mathcal{L}_{i,0.32}$ contains three features, and $\mathcal{L}_{i,0.34}$ contains two features, the latter may be selected if it has a higher wSSE value.

Algorithm 1 Feature-Importance Task-Accuracy Table Construction

---

Input: Pretrained semantic transceiver, importance vector $\boldsymbol{\omega}_i = [\omega_{i,1}, \omega_{i,2}, \omega_{i,j}, \ldots, \omega_{i,F}]$, SNR set $\mathcal{G}$, quantization levels $Q = \{\tilde{q}_1, \tilde{q}_2, \tilde{q}_l, \ldots, \tilde{q}_{|Q|}\}$
Output: Task-accuracy table $\mathcal{T}_i(q, \gamma)$ for all $q \in Q$ and $\gamma \in \mathcal{G}$
for each quantized importance sum $\tilde{q} \in Q$ **do**
    Determine the feature subset $\mathcal{L}_{i,\tilde{q}}$ whose total importance is closest to $\tilde{q}$ and has the minimum number of features among such subsets
end for
for each SNR value $\gamma \in \mathcal{G}$ **do**
    for each quantized sum of importance $\tilde{q} \in Q$ **do**
        Simulate the transmission using selected semantic features $\mathcal{L}_{i,\tilde{q}}$ at SNR γ
        Compute the task-accuracy function $\xi(q, \gamma)$ from the transceiver output
    end for
end for
Return: Feature-importance-based task-accuracy table $\mathcal{T}_i(q, \gamma)$

Then, $\mathcal{F}$ is again reduced to $\overline{\mathcal{F}} \in \mathbb{R}^{N \times M}$ so that each element $\overline{\mathcal{F}}_{i,m}$ maintains only the best wSSE value. Therefore, for all user-channel pairs, the optimal semantic feature selection is determined.

3.2.2. Optimal Subchannel Assignment (Phase-2)

The optimal user-channel matching is determined based on the wSSE matrix $\overline{\mathcal{F}}$ obtained in Phase-1. This problem can be considered a standard maximum weight assignment problem; therefore, the Hungarian algorithm can be employed to obtain the optimal assignment. The optimal subchannel assignment is saved in a matrix Γ (for $N \neq M$, zero-weight dummy nodes are added to convert the matrix to the square form). Each user-subchannel pair with $\Gamma_{i,m} = 1$ is scheduled, and its corresponding optimal feature set stored in $\mathcal{F}$ is transmitted.

Remark 1. By nature of OFDM links, a high peak-to-average power ratio forces power-amplifier backoff, limiting usable transmit power. This constraint makes feature prioritization particularly useful in budget-constrained situations [37]. In practical, low-power industrial Internet of Things anomaly detection [38, 39, 40], transmitting only the most important features can

Algorithm 2 Importance-based Subchannel Allocation

Require: Precomputed feature-importance-based task-accuracy table $\mathcal{T}(q, \gamma)$ by Algorithm 1; SNR values $\gamma_{i,m}$ for all $i \in \mathcal{N}, m \in \mathcal{M}$; importance vector $\boldsymbol{\omega}_i$ for each user i; threshold ξ_{th}

Ensure: Optimal subchannel assignment $\boldsymbol{\alpha}$ and feature selection mask $\boldsymbol{\beta}_i$ for each user

Initialize the wSSE matrix $\overline{\mathcal{F}} \in \mathbb{R}^{N \times M}$
for each user $i \in \mathcal{N}$ **do**
  for each subchannel $m \in \mathcal{M}$ **do**
    Retrieve $\mathcal{T}_i(\tilde{q}_l, \gamma_{i,m})$ for all $q_l \in \mathcal{Q}$ from the precomputed table $\mathcal{L}_i$
    Identify feasible q_l values satisfying $\mathcal{T}_i(q_l, \gamma_{i,m}) \geq \xi_{\text{th}}$
    Select $\tilde{q}_l$ that maximizes $\Phi_{i,m} = \frac{\mathcal{T}_i(\tilde{q}_l, \gamma_{i,m})}{\sum_{j=1}^{F} \beta_{i,j}}$
    Set feature mask $\boldsymbol{\beta}_i \leftarrow \mathcal{L}_{i,\tilde{q}_l}$
    Compute set $\overline{\mathcal{F}}_{i,m} \leftarrow \Phi_{i,m}^{max} = \frac{\mathcal{T}_i(\tilde{q}_l, \gamma_{i,m})}{\sum_{j=1}^{F} \beta_{i,j}}$
  end for
end for
Apply the Hungarian algorithm on matrix $\overline{\mathcal{F}}$ to solve the optimal subchannel assignment
Return subchannel assignment matrix Γ and selected feature masks $\boldsymbol{\beta}_i$ for each user

accurately convey the meaning of the detection data [41]. This reduces latency and energy consumption simultaneously. Furthermore, in vehicle-to-everything cooperative perception, prioritizing hazard-relevant features sustains situational awareness under congestion, adapts quickly to SNR/load changes via lookup, and avoids online retraining [42].

4. Experiments

Feature importance was calculated using an ablation, one feature at a time, with a fixed SNR in Python (Spyder). Pytorch was used for inference. CIFAR-100 was used as the source dataset. The importance of each feature was normalized. Feature selection and channel assignment simulations were run multiple times using a fixed seed in Python (Colab), and the results are measured as a trial average. To improve stability, this procedure was repeated for multiple randomly sampled mini-batches using a fixed seed, and all values were averaged. To assess the effectiveness of the proposed importance-aware transmission strategies, the proposed scheme is compared with the following baselines. Simulation parameters are summarized in Table 2.

Table 1
Simulation Parameters

Parameter	Value
Number of users, N	5
Number of subchannels, M	5
Number of encoded semantic features, F	64
Channel bandwidth, W	180 KHz
Noise power spectral density, N_0	-174 dBm/Hz
Pathloss model	$128.1 + 37.6 \log_{10}(d\ [\text{km}])$ dB
Shadow effect factor	6 dB
Transmit power, P	10 dBm
Task-accuracy threshold, ξ_{th}	0.8

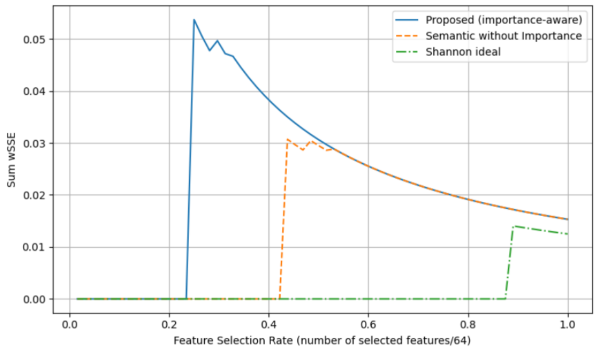

Figure 2. Sum of wSSE versus the selection rate.

- Semantic-Without-Importance (SWI): It assumes that all semantic features have identical importance. For each subchannel condition, the system dynamically determines the number of features to transmit to maximize the wSSE metric under a uniform-importance constraint.
- Shannon-Ideal (SI): Following the approach in [43], it employs the widely used lossy compression (e.g., JPEG), and classification accuracy was measured across varying compression ratios and SNRs. The resulting accuracy values were converted into the wSSE form using the same computation as for the semantic scheme.

Fig. 2 plots the sum wSSE against the feature-selection rate. SWI, which ignores feature importance, requires a selection rate of 0.44 to meet the threshold and produces lower peaks. This is because it fails to distinguish between features that actually contribute to the task. SI requires significantly more features to reach the task success threshold. Therefore, wSSE remains close to 0 until the selection rate increases. Meanwhile, the proposed method selects features with high importance first. Therefore, a small subset of features quickly reaches the importance required for task success. This exceeds the accuracy threshold at a selection rate of 0.25. After reaching the peak, most of the additional fea-

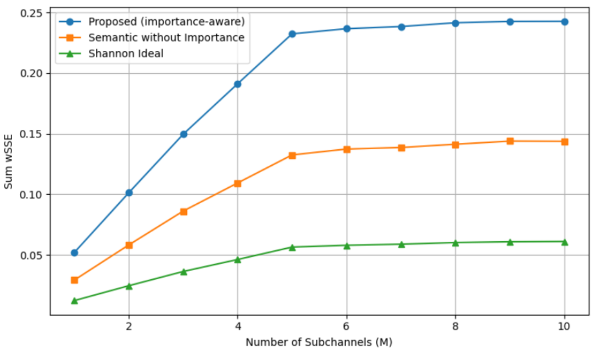

Figure 3. Sum of wSSE versus the number of subchannels.

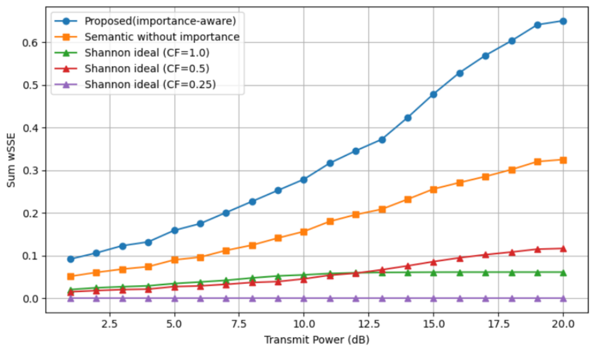

Figure 4. Sum of wSSE versus the transmission power.

tures contribute less to the task, so wSSE decreases as the selection set grows. That is, the proposed approach meets the accuracy threshold, much earlier than SWI, and achieves a peak wSSE 73.3% higher.

Fig. 3 shows the sum of wSSE versus the number of subchannels M. All methods improve with an increasing M value due to the greater flexibility in user-channel matching, but the proposed approach consistently achieves the highest performance by integrating importance-aware feature selection. SWI saturates at relatively low wSSE because irrelevant features waste resources. While SI remains lower overall since it ignores semantic relevance.

Fig. 4 shows wSSE versus transmission power. The proposed approach yields the highest values, with sharp gains at low-to-mid power by prioritizing important features. SWI grows more slowly since power is wasted on less-relevant features, while SI under different compression factors performs worst, requiring more power to satisfy accuracy.

5. Conclusion

This paper proposes an importance-aware semantic transmission framework that jointly optimizes feature selection and subchannel assignment under constraints. A new metric, wSSE, captures the task relevance of transmitted features, supported by an empirical accuracy table enabling real-time decisions. Simulations result demonstrate that the proposed method achieves higher task accuracy and wSSE than both uniform-importance and Shannon-based baselines, particularly under stringent transmission budgets.

Acknowledgments

This work was supported in part by the IITP (Institute of Information & Communications Technology Planning & Evaluation) - ITRC (Information Technology Research Center) (IITP-2025-RS-2024-00436887, 50%) grant funded by the Korea government (Ministry of Science and ICT), in part by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2024-00453301), and in part by the Chung-Ang University Research Scholarship Grants in 2024.

References

[1] S. Rani, S. Kumar, A. Kataria, H. Min, Smarthealth: An intelligent framework to secure iomt service applications using machine learning, ICT Express 10 (2) (2024) 425–430.

[2] S.-W. Lee, J.-H. Kwon, D. Kim, E.-J. Kim, Research category classification of scientific articles on human health risks of electromagnetic fields using pre-trained bert, ICT Express 10 (2) (2024) 336–341.

[3] G. C. Amaizu, J. N. Njoku, J.-M. Lee, D.-S. Kim, Metaverse in advanced manufacturing: Background, applications, limitations, open issues & future directions, ICT Express 10 (2) (2024) 233–255.

[4] B. Saoud, I. Shayea, A. E. Yahya, Z. A. Shamsan, A. Alhammadi, M. A. Alawad, Y. Alkhrijah, Artificial intelligence, internet of things and 6g methodologies in the context of vehicular ad-hoc networks (vanets): Survey, ICT Express 10 (4) (2024) 959–980.

[5] M. Ergen, B. Saoud, I. Shayea, A. A. El-Saleh, O. Ergen, F. Inan, M. F. Tuysuz, Edge computing in future wireless networks: A comprehensive evaluation and vision for 6g and beyond, ICT Express 10 (5) (2024) 1151–1173.

[6] A. Islam, K. Chang, Navigating the future of wireless networks: A multidimensional survey on semantic communications, ICT Express 10 (4) (2024) 747–773.

[7] K. K. Vaigandla, A systematic survey on artificial intelligence in 6g wireless networks: Security, opportunities, applications, advantages, future research directions and challenges, Babylonian Journal of Artificial Intelligence 2025 (2025) 99–106.

[8] J. Kim, G. Caire, A. F. Molisch, Quality-aware streaming and scheduling for device-to-device video delivery, IEEE/ACM Transactions on Networking 24 (4) (2016) 2319–2331.

[9] H. Zhang, S. Shao, M. Tao, X. Bi, K. B. Letaief, Deep learning-enabled semantic communication systems with task-unaware transmitter and dynamic data, IEEE Journal on Selected Areas in Communications 41 (1) (2022) 170–185.

[10] X. Han, B. Feng, Y. Shi, Y. Wu, W. Zhang, Semantic-aware resource allocation for wireless image transmission, in: 2024 IEEE/CIC International Conference on Communications in China (ICCC), 2024, pp. 2071–2076.

[11] Z. Weng, Z. Qin, Semantic communication systems for speech transmission, IEEE Journal on Selected Areas in Communications 39 (8) (2021) 2434–2444.

[12] T.-Y. Tung, D. Gündüz, Deepwive: Deep-learning-aided wireless video transmission, IEEE Journal on Selected Areas in Communications 40 (9) (2022) 2570–2583.

[13] M. Awais, J. Choi, J. Park, Y. H. Kim, Intelligent data-aided semantic sensing with variational deep embedding, ICT Express 10 (4) (2024) 824–830.

[14] D. Won, G. Woraphonbenjakul, A. B. Wondmagegn, A. T. Tran, D. Lee, D. S. Lakew, S. Cho, Resource management, security, and privacy issues in semantic communications: A survey, IEEE Communications Surveys & Tutorials (2024).

[15] I. M. S. Putra, D. Siahaan, A. Saikhu, Recognizing textual entailment: A review of resources, approaches, applications, and challenges, ICT Express 10 (1) (2024) 132–155.

[16] R. Gipiškis, C.-W. Tsai, O. Kurasova, Explainable ai (xai) in image segmentation in medicine, industry, and beyond: A survey, ICT Express 10 (6) (2024) 1331–1354.

[17] G. Shi, Y. Xiao, Y. Li, X. Xie, From semantic communication to semantic-aware networking: Model, architecture, and open problems, IEEE Communications Magazine 59 (8) (2021) 44–50.

[18] Q. Lan, D. Wen, Z. Zhang, Q. Zeng, X. Chen, P. Popovski, K. Huang, What is semantic communication? a view on conveying meaning in the era of machine intelligence, Journal of Communications and Information Networks 6 (4) (2021) 336–371.

[19] L. Yan, Z. Qin, R. Zhang, Y. Li, G. Y. Li, Resource allocation for text semantic communications, IEEE Wireless Communications Letters 11 (7) (2022) 1394–1398.

[20] R. Cheng, G. Balasubramaniam, Y. He, Y.-H. H. Tsai, H. Zhao, Greedy modality selection via approximate submodular maximization, arXiv preprint arXiv:2210.12562 (2022).

[21] S. H. O. Raghad Tariq Abdulrazzak Alhassani, Evaluation fiber optic communication network performance through manchester code, Babylonian Journal of Networking 2024 (2023) 11–14.

[22] L. Xia, Y. Sun, X. Li, G. Feng, M. A. Imran, Wireless resource management in intelligent semantic communication networks, in: IEEE INFOCOM 2022-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), IEEE, 2022, pp. 1–6.

[23] L. Yan, Z. Qin, R. Zhang, Y. Li, G. Y. Li, Qoe-aware resource allocation for semantic communication networks, in: GLOBECOM 2022-2022 IEEE Global Communications Conference, IEEE, 2022, pp. 3272–3277.

[24] C. Liu, C. Guo, Y. Yang, N. Jiang, Adaptable semantic compression and resource allocation for task-oriented communications, IEEE Transactions on Cognitive Communications and Networking 10 (3) (2023) 769–782.

[25] W. Li, H. Liang, C. Dong, X. Xu, P. Zhang, K. Liu, Non-orthogonal multiple access enhanced multi-user semantic communication, IEEE Transactions on Cognitive Communications and Networking (2023). arXiv:2303.06597.

[26] L. Wang, W. Wu, F. Zhou, Z. Yang, Z. Qin, Q. Wu, Adaptive resource allocation for semantic communication networks, IEEE Transactions on Communications 72 (11) (2024) 6900–6916.

[27] H. Zhang, H. Wang, Y. Li, K. Long, A. Nallanathan, Drl-driven dynamic resource allocation for task-oriented semantic communication, IEEE Transactions on Communications 71 (7) (2023) 3992–4004.

[28] Y. Li, X. Zhou, J. Zhao, Resource allocation for the training of image semantic communication networks, IEEE Transactions on Wireless Communications 24 (4) (2025) 2968–2984.

[29] H. Zhang, H. Wang, Y. Li, K. Long, V. C. M. Leung, Toward intelligent resource allocation on task-oriented semantic communication, IEEE Wireless Communications 30 (3) (2023) 70–77.

[30] L. L. Janneh, Y. Zhang, M. Hydara, Z. Cui, Deep learning-based hybrid feature selection for the semantic segmentation of crops and weeds, ICT Express 10 (1) (2024) 118–124.

[31] S. Khan, S. Khan, A. Sulaiman, M. S. A. Reshan, H. Alshahrani, A. Shaikh, Deep neural network and trust management approach to secure smart transportation data in sustainable smart cities, ICT Express 10 (5) (2024) 1059–1065.

[32] M. Umair, T. Sultana, Y.-K. Lee, Pre-trained language models for keyphrase prediction: A review, ICT Express 10 (4) (2024) 871–890.

[33] J. Park, J. Ko, Fedhm: Practical federated learning for heterogeneous model deployments, ICT Express 10 (2) (2024) 387–392.

[34] W. Samek, T. Wiegand, K.-R. Müller, Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models, ITU Journal: ICT Discoveries 1 (1) (2017) 1–10.

[35] J. Shin, J. Won, H.-S. Lee, J.-W. Lee, A review on label cleaning techniques for learning with noisy labels, ICT Express 10 (6) (2024) 1315–1330.

[36] X. Mu, Y. Liu, Semi-noma enabled coexisting semantic and bit communications, in: 2022 International Symposium on Wireless Communication Systems (ISWCS), 2022, pp. 1–6.

[37] K. K. Vaigandla, A comprehensive review on ofdm, 5g and various papr minimization techniques based on machine learning, Babylonian Journal of Networking 2025 (2025) 43–58.

[38] R. F. Pamungkas, I. B. K. Y. Utama, K. Hindriyandhito, Y. M. Jang, A hybrid approach of convlstmbnn-dt and gpt-4 for real-time anomaly detection decision support in edge–cloud environments, ICT Express 10 (5) (2024) 1026–1033.

[39] D. Lim, J. Kim, H. Kim, Efficient robot tracking system using single-image-based object detection and position estimation, ICT Express 10 (1) (2024) 125–131.

[40] R. Khan, M. Sohail, I. Usman, M. Sandhu, M. Raza, M. A. Yaqub, A. Liotta, Comparative study of deep learning techniques for deepfake video detection, ICT Express 10 (6) (2024) 1226–1239.

[41] A. S. M. Thabit, C. A. Kerrache, C. T. Calafate, A survey on monitoring and management techniques for road traffic congestion in vehicular networks, ICT Express 10 (6) (2024) 1186–1198.

[42] J. Yin, S. Zhao, S.-H. Hwang, Dynamic separate resource pool algorithm for vehicle-to-infrastructure communication, ICT Express 10 (6) (2024) 1259–1265.

[43] G. Nayak, G. Friedland, Deep layers beware: Unraveling the surprising benefits of jpeg compression for image classification preprocessing, 2023 IEEE International Symposium on Multimedia (ISM) (2023) 182–185.