On the Performance of Channel Coding in Digital Semantic Communication

Seonjung Kim and Yo-Seb Jeon
Department of Electrical Engineering, POSTECH, Pohang, Gyeongbuk 37673, South Korea
Email: {seonjung.kim, yoseb.jeon}@postech.ac.kr

***Abstract*—Semantic communication aims to transmit task-relevant meaning, not just raw data. This paper evaluates the performance of repetition, polar, and low-density parity-check (LDPC) codes in a digital semantic communication system. The probabilities are obtained via end-to-end training with a fixed bit-flip probability. Experiments on image transmission over additive white Gaussian noise channels show that the trade-off between image reconstruction quality and energy efficiency is controlled by the bit-flip probability.**

I. Introduction

As a new approach, semantic communication transmits task-relevant information instead of raw data, leading to more efficient use of communication resources and enhanced robustness to noise and interference. [1]. The early research was focused on neural joint source-channel coding (JSCC). In JSCC, encoders and decoders are trained together in an end-to-end manner to minimize task-specific losses [2]. Although these analog JSCC methods proved effective, their reliance on continuous-valued outputs made them incompatible with modern digital systems.

To address this, digital semantic communication has gained attention for its compatibility and flexibility [3]–[5]. Various methods have enabled digital outputs through sampling-based approaches, Gumbel-softmax, or differentiable quantization, respectively. While these methods have shown promising results, most do not explicitly consider error characteristics of semantic representations, limiting their integration with advanced channel coding strategies.

Our study aims to investigate the performance of standard channel coding schemes in digital semantic communication systems, while a bit-flip probability is fixed during end-to-end training. We evaluate three coding strategies—repetition coding and low-density parity-check (LDPC) coding—under a fixed signal-to-noise ratio (SNR) in an additive white Gaussian noise (AWGN) channel. Our results reveal that there is a trade-off between image reconstruction quality and energy efficiency, which is controlled by the bit-flip probability.

II. System Model

In this section, we describe the digital semantic communication scenario considered in this work.

This work was supported 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 a grant of Korean ARPA-H Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare (No. RS-2025-25454431).

A. Training Stage

This study considers a digital semantic communication system designed for transmitting images. In this system, an input image, represented as $\boldsymbol{u} \in \mathbb{R}^U$, is first processed by a semantic encoder $f_{\boldsymbol{\theta}_{\text{enc}}}$ parameterized by $\boldsymbol{\theta}_{\text{enc}}$. This encoder extracts a key feature of the image as follows:

$$\boldsymbol{v} = f_{\boldsymbol{\theta}_{\text{enc}}}(\boldsymbol{u}) \in \mathbb{R}^M, \tag{1}$$

where $\boldsymbol{v}$ represents the semantic feature vector of length M. After encoding, each element of $\boldsymbol{v}$ is quantized using a uniform B-bit quantizer. The resulting quantized output is defined as

$$\boldsymbol{q}_i = \mathsf{Q}(v_i) \in \mathcal{Q}, \quad i \in [M], \tag{2}$$

where v_i denotes the i-th element of $\boldsymbol{v}$, $\mathcal{Q} = \{\tilde{\boldsymbol{q}}_1, \tilde{\boldsymbol{q}}_2, \dots, \tilde{\boldsymbol{q}}_{2^B}\}$ represents the quantizer codebook, and each $\tilde{\boldsymbol{q}}_i$ is a binary vector of length B. In this work, we refer to each binary entry of $\tilde{\boldsymbol{q}}_i$ as a *semantic* bit. The transmitter then constructs a *semantic* bit sequence as $\boldsymbol{b} = [\boldsymbol{q}_1^\top, \cdots, \boldsymbol{q}_M^\top]^\top \in \{0,1\}^K$, where $K = MB$ is the total number of semantic bits.

During the training stage, we model the channel by assuming that each semantic bit in the sequence $\boldsymbol{b}$ is independently flipped with a fixed bit-flip probability, denoted by μ. This process results in a received bit vector $\hat{\boldsymbol{b}}$. This fixed bit-flip probability, μ, serves as a hyperparameter during the end-to-end training process, effectively capturing the trade-off between image reconstruction quality and energy efficiency. For instance, training with a low μ encourages the neural network to learn a more detailed semantic representation, assuming a highly reliable channel. This results in higher task performance, i.e., better image reconstruction quality. However, achieving such a low-bit probability in a practical communication system requires significant resources, such as higher transmission power or larger channel codes, leading to lower energy efficiency.

From the received bit sequence $\hat{\boldsymbol{b}}$, the receiver constructs the semantic feature vector $\hat{\boldsymbol{v}}$. Specifically, $\hat{\boldsymbol{b}}$ is first partitioned into M binary vectors, $\hat{\boldsymbol{q}}_i$, each of length B. Then, each binary vector is dequantized to recover the corresponding semantic feature element:

$$\hat{v}_i = Q^{-1}(\hat{\boldsymbol{q}}_i) \in \mathbb{R}, \quad i \in [M]. \tag{3}$$

The resulting feature vector is then fed into the semantic decoder to reconstruct the original input data:

$$\hat{\boldsymbol{u}} = f_{\boldsymbol{\theta}_{\text{dec}}}(\hat{\boldsymbol{v}}) \in \mathbb{R}^U, \tag{4}$$

where $f_{\boldsymbol{\theta}_{\text{dec}}}$ denotes the semantic decoder.

B. Inference Stage

In the inference stage, the transmitter first transforms an input image $\boldsymbol{u}$ into a semantic bit sequence $\boldsymbol{b} \in \{0,1\}^K$ using the trained semantic encoder $f_{\boldsymbol{\theta}_{\text{enc}}}$ and the quantizer Q. To improve the reliability of the transmission, a channel coding process is utilized, which adds redundant bits to $\boldsymbol{b}$ to correct errors. The bit sequence that results from channel encoding is denoted as

$$\tilde{\boldsymbol{b}} = g_{\text{enc}}(\boldsymbol{b}) \in \{0,1\}^{N_{\text{tot}}}, \tag{5}$$

where g_{enc} denotes the channel encoder.

The wireless link between the transmitter and receiver is modeled as an additive white Gaussian noise (AWGN) channel. In AWGN channel, the received signal at time slot t is represented as

$$y_t = \sqrt{p_t} x_t + v_t, \quad t \in [T], \tag{6}$$

where $v_t \sim \mathcal{CN}(0, \sigma^2)$ denotes Gaussian noise with zero mean and variance σ^2, and p_t is the transmission power allocated to the t-th symbol.

After receiving the signal, the receiver attempts to recover a semantic bit sequence $\hat{\boldsymbol{b}} = [\hat{\boldsymbol{q}}_1^\top, \cdots, \hat{\boldsymbol{q}}_M^\top]^\top \in \{0,1\}^K$ using channel decoding, as $\hat{\boldsymbol{b}} = g_{\text{dec}}(\boldsymbol{y})$, where g_{dec} is the channel decoder, and $\boldsymbol{y} = [y_1, \cdots, y_T]^\mathsf{T}$ is a received signal vector.

Finally, the reconstructed semantic bit sequence $\hat{\boldsymbol{b}}$ is dequantized to recover the semantic feature vector $\hat{\boldsymbol{v}}$. This vector is then fed into the trained semantic decoder $f_{\boldsymbol{\theta}_{\text{dec}}}$ to reconstruct the original data.

III. Simulation Results

In this section, we evaluate the performance of standard channel coding schemes for an image transmission task with MNIST dataset. Convolutional neural network (CNN)-based autoencoder architectures is adopted, where the network design is guided by [2]. The quantization bit for feature vector is 8 bits. The symbol is transmitted with power $P_{\text{trans}} = 0$ dBW, and the SNR is fixed at 0 dB under an AWGN channel. In the simulations, we compare the following channel coding frameworks:

- **Repetition:** All bits are encoded using the repetition code with a fixed repetition number.The repetition number is determined as the smallest integer R which satisfies

$$\sum_{j=\lceil R/2 \rceil}^{R} \binom{R}{j} \epsilon^j (1-\epsilon)^{R-j} \le \mu, \tag{7}$$

  where $\epsilon = Q(\sqrt{2\mathsf{SNR}})$.
- **LDPC:** All bits are encoded using the LDPC code with a fixed rate. The rate is determined as the smallest rate whose resulting bit error rate is lower than μ among the candidates $\{3/4,\ 2/3,\ 1/2,\ 1/3\}$.
- **Genie:** This ideal baseline assumes perfect transmission with no bit-flips, i.e., all bits are received without error.

In Fig. 1, we compare the PSNR and corresponding total blocklength for various channel coding frameworks, using the MNIST dataset with $\mu = 10^{-4}$ and $\mu = 10^{-2}$. Both the repetition and LDPC coding schemes achieve a PSNR that approaches the genie baseline, which serves as the performance upper bound. The results also demonstrate the trade-off controlled by μ. Targeting a smaller bit-flip probability allows the system to achieve a significantly higher PSNR, but this requires a longer total block length to meet the stricter reliability constraint. Furthermore, a comparison between the coding schemes shows that LDPC consistently uses much less blocklength than repetition coding while maintaining the similar PSNR performance.

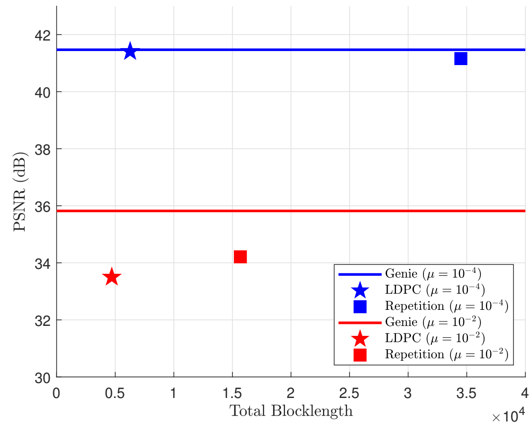

Fig. 1. Comparison of PSNR versus total blocklength for various channel coding frameworks using the MNIST dataset ($K = 3136$).

IV. Conclusion

In this paper, we investigated the performance of standard channel coding schemes in a digital semantic communication system trained with a fixed target bit-flip probability. We demonstrated that this training methodology effectively captures the trade-off between image reconstruction quality and energy efficiency, controlled by the hyperparameter, μ. Our key finding is that under fixed channel conditions, both repetition and LDPC codes can approach the ideal performance upper bound set by a genie-aided system. Furthermore, our results highlight that advanced channel codes like LDPC offer significantly better efficiency, achieving a similar reconstruction quality to repetition coding but with a much shorter total blocklength.

References

[1] X. Luo, H.-H. Chen, and Q. Guo, "Semantic communications: Overview, open issues, and future research directions," *IEEE Wireless Commun.*, vol. 29, no. 1, pp. 210–219, Feb. 2022.

[2] E. Bourtsoulatze, D. B. Kurka, and D. Gündüz, "Deep joint source-channel coding for wireless image transmission," *IEEE Trans. Cogn. Commun. Netw.*, vol. 5, no. 3, pp. 567–579, Sep. 2019.

[3] K. Choi, K. Tatwawadi, A. Grover, T. Weissman, and S. Ermon, "Neural joint source-channel coding," in *Proc. Int. Conf. Mach. Learn. (ICML)*, Long Beach, CA, USA, Jun. 2019, pp. 1182–1192.

[4] S. Yao, K. Niu, S. Wang, and J. Dai, "Semantic coding for text transmission: An iterative design," *IEEE Trans. Cogn. Commun. Netw.*, vol. 8, no. 4, pp. 1594–1603, Dec. 2022.

[5] T.-Y. Tung, D. B. Kurka, M. Jankowski, and D. Gündüz, "DeepJSCC-Q: Constellation constrained deep joint source-channel coding," *IEEE J. Sel. Areas in Inf. Theory*, vol. 3, no. 4, pp. 720–731, Dec. 2022.