# Digital Semantic Communication with Neural Image Compression

Long V. Nguyen*, Tuan T. Nguyen†, Octavia A. Dobre* and Trung Q. Duong*

* Memorial University, Canada, e-mail: {lnguyen, odobre, tduong}@mun.ca
† University of Greenwich, UK, e-mail: tuan.nguyen@greenwich.ac.uk

***Abstract*—Although analog semantic communication systems have attracted significant attention recently, there has been relatively less focus on digital semantic communication systems. In this work, we present a neural image compression-enabled semantic communication for digital image transmission to enhance the efficiency of image transmission, named NCSC. By employing an accurate and adaptable entropy model, NCSC obtains the efficiently compressed bitstreams, which are compatible with digital communication systems. Incorporating with the well-established digital components, our system trained on the MS-SSIM metric can achieve a significant bandwidth compression ratio of 0.002 at low SNR, reducing remarkably transmission overhead. Extensive simulations show that our approach outperforms traditional digital communication systems in terms of perceptual quality and bandwidth efficiency under challenging channel conditions.**

***Index Terms*—Semantic communication, digital communication, learned image compression.**

## I. Introduction

The field of communication has undergone a paradigm shift with the recent emergence of semantic communication. Unlike traditional methods that focus solely on the reliable transmission of bits, semantic communication [1] aims to convey the meaning and intent of information, promising enhanced efficiency and robustness in varying environments. While early semantic communication research explored analog semantic transmission schemes, practical deployment necessitates a transition towards digital semantic communication systems to leverage existing infrastructure and benefit from well-established digital signal processing techniques. This transition inherently offers better scalability and flexibility, making them suitable for a wide range of applications, including the rapidly growing domains of IoT, 5G, and beyond.

Neural data compression [2], powered by deep learning architectures, has emerged as a groundbreaking approach in image compression, demonstrating remarkable capabilities in achieving unprecedented compression ratios while maintaining high perceptual quality. Unlike traditional compression methods such as JPEG, JPEG2000, which rely on hand-crafted features and predetermined transformation techniques, neural compression leverages learned representations to capture complex patterns and semantic features within the data. Neural compression techniques can achieve higher compression than conventional methods while maintaining comparable or better image quality. This capability is not only crucial for efficient image storage, but also has profound implications for communication systems, enabling the transmission of richer and more complex data at reduced bandwidth requirements.

This paper introduces a novel digital semantic communication system that integrates the advantages of neural image compression. By employing a neural compression-based entropy coder, our proposed system generates a highly compressed bitstream of the original image for transmission. These compressed representations can be efficiently transmitted and combined with widely adopted digital components in practice. At the receiver, the image is reconstructed from the transmitted bitstreams by the learned decoder. This allows us to achieve highly efficient data compression while maintaining the semantic content of the transmitted information, leading to more efficient and robust communication systems. We demonstrate that this approach can minimize the required transmission overhead significantly while attaining the visually pleasant reconstructed images, providing a pathway towards more efficient and intelligent digital communication systems.

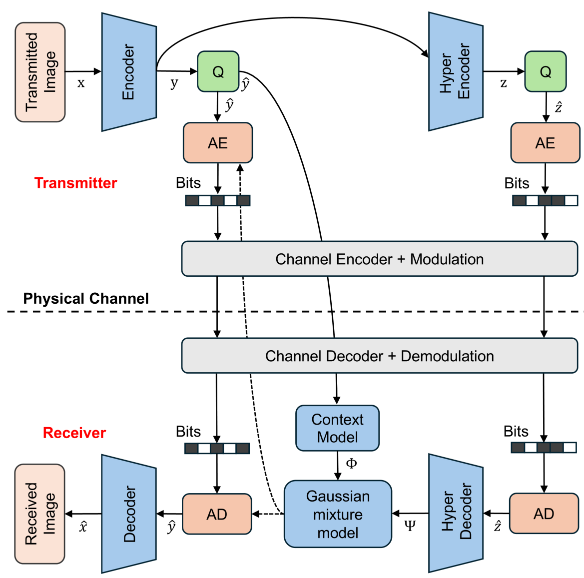


Fig. 1: Architecture of neural compression-based digital semantic communication (NCNS)

## II. System Model

At the transmitter, the encoder maps an input RGB image $x$ into a low-dimensional latent representation, followed by a quantization to generate $\hat{y}$. An arithmetic coder is utilized to encode the quantized codes into a bitstream. A similar process is applied to the Hyper Encoder, a subnetwork for learning a probabilistic model over quantized latents used for entropy coding. These compressed bitstreams are then transmitted through the wireless channel, which is compatible with practical digital communication systems. We define the ratio of the number of compressed bits at the channel input to that of the original source bits as the bandwidth compression ratio (BCR), which quantifies the transmission overhead.

The reverse operations are deployed at the receiver to reconstruct the original image. The Gaussian mixture model (GMM) is shared between the transmitter and receiver for entropy coding, conditioned on the previously learned hyperprior. GMM is chosen for its ability to model complex distributions in the latent space, ensuring robust entropy coding.

The neural compression-based digital semantic communication framework is illustrated in Fig. 1, named NCSC. The system design of NCSC is inspired by [3], in which we eliminate the attention modules to decrease the system's complexity without significantly compromising the performance. The integration of neural compression allows the system to reduce the redundancy of the source images significantly by adopting the accurate and flexible entropy model. Our model is optimized using two quality metrics: mean square error (MSE) and MS-SSIM. This results in two achieved frameworks, called MSE-NCSC and MSS-NCSC, respectively. We evaluate our approaches on the widely used Kodak dataset with 24 uncompressed 768 × 512 images. For comparison, the reconstruction performance is measured using various metrics and compared with other traditional methods, as shown in Table I.

## III. Simulation Results and Discussions

TABLE I: Performance Comparison of Different Methods

| Method | TO[1] (symbols) ↓ | BPP ↓ | PSNR ↑ | MS-SSIM ↑ | FID ↓ |
|---|---|---|---|---|---|
| Raw Image | 7077888 | 24 | - | - | - |
| JPEG | 51064 | 0.173 | 21.41 | 0.712 | 327.3 |
| WEBP | 36773 | 0.125 | 26.53 | 0.877 | 126.4 |
| BPG | 31026 | 0.117 | 27.83 | 0.900 | 108.6 |
| MSE-NCSC | 17605 | 0.117 | 28.10 | 0.923 | 115.4 |
| MSS-NCSC | 14482 | **0.100** | 25.29 | 0.942 | 97.2 |

[1]Transmission Overhead regarding the number of transmitted symbols per image.

We simulate our system under the AWGN channel. The 16-QAM is adopted for digital modulation and the (1536, 4608) LDPC codes are selected with a code rate of 1/3. As illustrated in Table I, our proposed method, MSS-NCSC, achieved a notably lower bits-per-pixel rate, minimizing the transmission overhead to only 14482 symbols per image and yielding a BCR of 0.002. Additionally, the method indicated superior performance across various perceptual quality metrics, including MS-SSIM and FID, particularly at a low SNR of 3 dB.

The Fig. 2 shows the image kodim07 reconstructed by different methods. NCSC trained on specified metrics significantly outperforms traditional methods in terms of visual perception and bandwidth compression ratio at 3 dB SNR. This underscores the resource efficiency and semantic preservation of our frameworks.

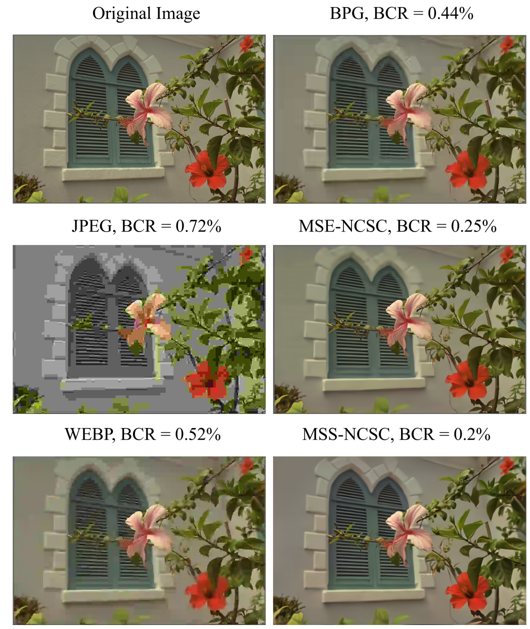


Fig. 2: Visualization of reconstructed images Kodim07 from Kodak dataset for the SNR = 3 dB

However, our system fails to recover the transmitted image under severely challenging channel conditions (below 3 dB). This is because the bitstreams are very susceptible to bit-flips caused by noise within the physical channel. Although this can be handled by the channel encoder to some extent, it may lead to higher computational complexity and latency, along with an increase in the number of bits transmitted. This problem will be further investigated in our future work.

## IV. Conclusion

By integrating an accurate entropy model based on neural compression with digital communication systems, we proposed a novel NCSC framework that achieves significant bandwidth compression while maintaining the visually pleasant quality of the reconstructed image at low SNR. The extensive results demonstrate the proposed approach's superiority over traditional methods in various factors, marking a step toward efficient, intelligent communication systems.

## 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, 2022.

[2] Y. Yang, S. Mandt, L. Theis *et al.*, "An introduction to neural data compression," *Foundations and Trends® in Computer Graphics and Vision*, vol. 15, no. 2, pp. 113–200, 2023.

[3] Z. Cheng, H. Sun, M. Takeuchi, and J. Katto, "Learned image compression with discretized gaussian mixture likelihoods and attention modules," in *Proc. IEEE/CVF Conf. Comput. Vision Pattern Recogn. (CVPR)*, 2020, pp. 7936–7945.