Review Report: Application of Generative AI in Semantic Communication

Lingyu Wang

2024.06.16

1 Introduction

Semantic communication (SemCom) is an emerging paradigm that can significantly save spectrum resources and improve information interaction efficiency. However, existing SemCom systems have significant limitations in context reasoning and background knowledge provisioning. Generative Artificial Intelligence (GAI) technology has shown great potential in automatically creating valuable, diverse, and personalized content. This paper proposes a new GAI-assisted Semantic Communication Network (GAI-SCN) framework to achieve more efficient and reliable semantic reasoning and resource utilization.

2 Background

Existing SemCom systems are limited by the lack of context reasoning and background knowledge provisioning. GAI technology has shown great potential in generating multimodal content and understanding context, which can significantly improve training efficiency, context reasoning capabilities, and spectrum utilization in semantic communication.

3 Challenges

1. **Data Demand**: SemCom requires vast amounts of data for background knowledge construction and model pre-training.

2. **Semantic Accuracy**: Current systems lack sufficient accuracy in semantic calibration and recovery when dealing with complex context fragments.

3. **Computational and Storage Resources**: Large GAI models, such as ChatGPT-3, require significant computational and storage resources.

4. **Content Uncertainty and Latency**: GAI-generated content may introduce uncertainty and increase data processing and transmission latency.

4 Solution

This paper proposes a GAI-assisted Semantic Communication Network (GAI-SCN) framework, which collaborates across cloud, edge, and mobile layers. It uses global and local GAI models to provide multimodal semantic content, joint source-channel coding (JSCC), and AI-generated content (AIGC) acquisition to maximize the efficiency and reliability of semantic reasoning and resource utilization.

- Deploy large GAI models in the cloud for computationally intensive tasks.
- Deploy semantic encoders at the edge layer to improve system resilience through JSCC.
- Embed lightweight GAI models in terminal devices for local processing and personalization.

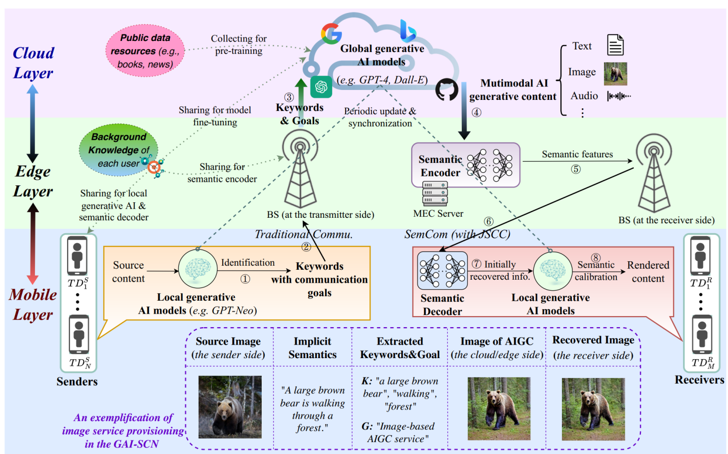

Figure 1: Illustration of the proposed GAI-SCN framework in a collaborative cloud-edge-mobile design, where an exemplification of image service provisioning is presented.

5 Experimental Setup

1. **Model Selection**:
   - **Local GAI Model**: ViT combined with GPT-2 for image-to-text conversion, keyword extraction, and goal identification.
   - **Global GAI Model**: Stable Diffusion 2.1 for generating AI-generated images from prompts.

2. **Semantic Communication Setup**:
    - Use deep convolutional network (Observation-Centric Sort) and Transformer-based semantic decoder for segmentation and recovery.
    - Train models with additive white Gaussian noise channel at 0 dB SNR, testing with 327 images.
3. **Optimizer**:
    - Adam optimizer with an initial learning rate of 5×10.
4. **Benchmarks**:
    - **GAI-Assisted Traditional Communication**: Encodes AIGC into bits using predefined rules.
    - **Typical SemCom**: Transmits original images solely through semantic coding models.

6 Results Analysis

Experimental results show that the complexity of images significantly impacts the recovery performance when using the GAI-SCN framework. Specifically, as the number of observable objects in the original images increases:

- **Semantic similarity** decreases.
- **Object quantity discrepancy** increases.
- **Recovery ratio of original objects** declines.

This indicates that the framework faces greater challenges in maintaining recovery accuracy and semantic consistency with more complex images.

7 Open Research Issues and Outlooks

1. **Limited Device Resources**: Implementing sophisticated AI models on devices with limited storage, memory, and computational power.
    - **Potential Solution**: Use model compression and acceleration techniques like knowledge distillation, parameter pruning, and quantization to reduce model complexity and size.
2. **Randomness in GAI Content**: Variability in AI-generated content and semantic decoder outputs, leading to inconsistencies.
    - **Potential Solution**: Investigate granularity tuning for keyword extraction and semantic calibration to mitigate randomness.

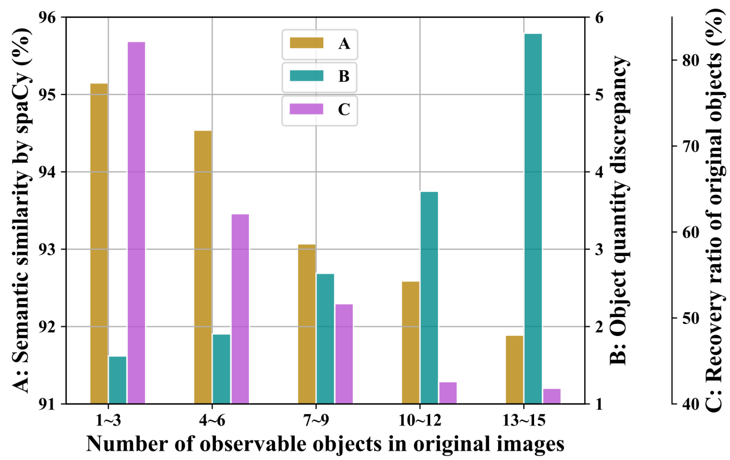

Figure 2: Comparisons between original and recovered images by the proposed GAI-SCN framework in terms of three metrics: A) Semantic similarity by spaCy; B) Object quantity discrepancy; C) Recovery ratio of original objects.

3. **Inactive Sharing of Knowledge and Preferences**: Encouraging users to share personal preferences and background knowledge necessary for customized AI and SemCom services.

    - **Potential Solution**: Develop incentive mechanisms, such as rewards or benefits, to motivate users to contribute their data for system improvements.

8 Conclusion

Integrating generative AI with semantic communication offers a promising approach to enhancing communication efficiency and resource utilization. The GAI-SCN framework, through the use of global and local GAI models, achieves multimodal content provisioning, joint source-channel coding (JSCC), and AI-generated content (AIGC) acquisition, providing an innovative solution for future semantic communication networks.