International Journal of Artificial Intelligence and Communication Networks
Volume: 02, Issue: 01, January-March 2026, pp: 63-73
ISSN: 3071-3358

A Comparative Study of Generative AI Models in Semantic Communication Networks

Dr. Jameel Ahamed[1], Dr.Shaik Mohammed Rasool[2]
[1]Assistant Professor, College of Computing and Technology, University of Bisha, Saudi Arabia.
[2]Associate Professor, HOD, Department of Electronics and Communication Engineering, Lords Institute of Engineering and Technology, Hyderabad, India.
jameel.shaad@gmail.com, muhammadrasool45@gmail.com

Abstract
Communication networks are gradually shifting from traditional bit-oriented transmission toward semantic communication, where the emphasis is placed on delivering meaningful and context-aware information rather than transmitting complete data packets. This transition has encouraged the adoption of Generative Artificial Intelligence (Generative AI) models to improve semantic understanding, intelligent encoding, and adaptive information delivery. The present study provides a comparative examination of prominent generative models, including Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), Transformer architectures, Large Language Models (LLMs), and Diffusion Models within semantic communication networks. The comparison focuses on important performance aspects such as semantic accuracy, computational efficiency, latency, bandwidth optimization, contextual understanding, and adaptability in dynamic communication environments. In addition, the study discusses the applicability of these models in emerging technologies such as 5G, 6G, Internet of Things (IoT), edge intelligence, and autonomous communication systems. Particular attention is given to practical concerns, including privacy preservation, semantic ambiguity, computational cost, scalability, and security challenges associated with generative AI integration. Based on the comparative analysis, Transformer-based and large language models demonstrate stronger capabilities in semantic representation and intelligent communication, whereas hybrid generative frameworks show promising potential for improving efficiency and contextual accuracy in future communication ecosystems. The study concludes that Generative AI can significantly contribute to the development of intelligent, adaptive, and semantically aware communication infrastructures.

Keywords: *Generative Artificial Intelligence (Generative AI), Semantic Communication Networks, Generative Adversarial Networks (GANs), Variational Autoencoder (VAE), Large Language Models (LLMs), Transformers, Diffusion Models, Semantic Encoding, Intelligent Communication Systems, 5G Networks, 6G Networks, Internet of Things (IoT), Deep Learning, Context-Aware Communication, Quality of Service (QoS), Edge Intelligence.*

1. Introduction

The rapid expansion of digital technologies has transformed modern communication systems into highly interconnected and data-intensive environments. Traditional communication frameworks, which have been the foundation of telecommunications for several decades, are primarily designed to ensure the accurate transmission of bits across communication channels. Shannon's information theory established the mathematical basis for reliable data transmission by focusing on channel capacity, noise reduction, and error control mechanisms [1]. This model has successfully supported the development of successive wireless generations and remains central to contemporary communication networks. However, as communication applications become increasingly intelligent and context-dependent, the limitations of purely bit-oriented communication are becoming more apparent. Modern communication scenarios involve autonomous vehicles, smart healthcare systems, industrial automation, intelligent transportation, and massive Internet of Things (IoT) deployments. In such environments, transmitting every bit of information may not always be necessary. Instead, the primary objective is often to communicate the intended meaning of the information accurately and efficiently. This requirement has motivated researchers to explore semantic communication, a paradigm that shifts the communication focus from the transmission of symbols to the transmission of meaning [15].

Received: *10-01-2026*
Revised: *21-02-2026*
Accepted: *12-03-2026*
Published: *24-03-2026*
Citation: *"A Comparative Study of Generative AI Models in Semantic Communication Networks", ijaicn, vol. 2, no. 1, pp. 63–73, March. 2026,*

By emphasizing semantic relevance rather than exact bit-level reconstruction, semantic communication aims to improve communication efficiency while reducing bandwidth consumption and transmission overhead.

The emergence of semantic communication has attracted significant attention within both academia and industry, particularly in the context of next-generation wireless networks. Unlike conventional communication systems, semantic communication incorporates contextual understanding, knowledge representation, and intelligent information processing into the communication process. Semantic encoders extract meaningful information from source data, while semantic decoders reconstruct the intended meaning using contextual knowledge and machine learning techniques [16], [17]. Such capabilities are expected to play a critical role in future communication infrastructures where network resources must be utilized more efficiently and communication decisions must be made intelligently.

Artificial Intelligence (AI) has become a key enabling technology in the realization of semantic communication systems. Recent advances in deep learning have demonstrated exceptional capabilities in pattern recognition, feature extraction, language understanding, and knowledge representation. Deep neural networks have significantly improved the ability of machines to process complex data structures and capture semantic relationships within information [11], [18]. Consequently, AI-driven communication frameworks are increasingly being explored to address challenges related to network optimization, intelligent resource allocation, and context-aware information delivery [19], [24].

Among various AI technologies, Generative Artificial Intelligence (Generative AI) has emerged as a particularly promising approach for semantic communication. Generative AI models are designed not only to analyze data but also to generate meaningful representations that closely resemble real-world information. These models learn underlying data distributions and can reconstruct, predict, or generate content while preserving semantic characteristics. Such capabilities make them highly suitable for semantic encoding, semantic compression, contextual reasoning, and intelligent information reconstruction within communication networks [28].

Several generative models have gained prominence in recent years due to their remarkable performance across a wide range of applications. Generative Adversarial Networks (GANs) introduced an adversarial learning mechanism capable of generating realistic synthetic data and improving semantic reconstruction processes [3]. Variational Autoencoders (VAEs) provided efficient latent-space representations that support semantic compression and information recovery while reducing communication overhead [4]. Transformer architectures revolutionized sequence modeling by introducing self-attention mechanisms that effectively capture long-range contextual dependencies [5]. Building upon these advances, Large Language Models (LLMs) such as GPT and LLaMA have demonstrated exceptional capabilities in contextual understanding, reasoning, and knowledge-driven communication [6], [22], [23]. Additionally, Diffusion Models have emerged as powerful generative frameworks capable of producing high-quality content through iterative denoising processes, offering new opportunities for semantic reconstruction and multimedia communication applications [9], [10].

The relevance of generative AI extends beyond content generation and increasingly influences the design of future communication systems. Emerging technologies such as 5G, 6G, edge intelligence, and IoT networks require communication mechanisms that can adapt dynamically to varying network conditions while maintaining high levels of efficiency and reliability. Generative AI models provide the ability to understand context, compress semantic information, reconstruct missing data, and support intelligent decision-making processes, thereby enhancing the overall effectiveness of semantic communication systems [12], [13], [25].

Despite the growing interest in integrating generative AI into communication networks, significant differences exist among various generative models regarding their performance, computational requirements, semantic representation capabilities, and deployment feasibility. While Transformer-based architectures and Large Language Models exhibit superior contextual understanding, other models such as GANs, VAEs, and Diffusion Models offer distinct advantages in specific communication scenarios. Therefore, a systematic comparison of these models is essential to identify their strengths, limitations, and suitability for semantic communication applications [29], [30].

This article presents a comparative study of major generative AI models within the context of semantic communication networks. Specifically, the study examines Generative Adversarial Networks, Variational Autoencoders, Transformer architectures, Large Language Models, and Diffusion Models with respect to semantic accuracy, computational efficiency, latency, bandwidth optimization, contextual awareness, adaptability, and scalability. Furthermore, the article explores their applications in next-generation communication environments and discusses the challenges and future research opportunities associated with generative AI-enabled semantic communication. By providing a comprehensive comparative perspective, this study aims to contribute to the development of intelligent, adaptive, and semantically aware communication systems for future wireless networks.

2. Literature Review

Shannon established the foundation of modern communication theory through his pioneering work on information transmission. He introduced fundamental concepts such as entropy, channel capacity, and noise management, which enabled the development of reliable digital communication systems. Although Shannon's model focused primarily on the accurate transmission of bits, it laid the groundwork for subsequent research in communication efficiency and network optimization [1].

Weaver expanded Shannon's communication framework by emphasizing the broader dimensions of communication, including semantic interpretation and message effectiveness. His contribution highlighted the importance of understanding meaning in communication processes, thereby influencing later developments in semantic communication theory [2].

Goodfellow et al. introduced Generative Adversarial Networks (GANs), a novel deep learning framework consisting of a generator and a discriminator trained through adversarial learning. The proposed architecture demonstrated remarkable capabilities in generating realistic synthetic data and significantly influenced subsequent advancements in generative artificial intelligence. GANs have since been explored for semantic reconstruction and intelligent communication applications [3].

Kingma and Welling proposed the Variational Autoencoder (VAE), a probabilistic generative model capable of learning latent representations of complex datasets. Their work provided an efficient mechanism for data compression and reconstruction, making VAEs particularly suitable for semantic encoding and bandwidth-efficient communication systems [4].

Vaswani et al. revolutionized deep learning by introducing the Transformer architecture based entirely on self-attention mechanisms. Unlike recurrent neural networks, Transformers effectively captured long-range dependencies within data while enabling parallel processing. This architecture became the foundation for many advanced semantic communication and language understanding systems [5].

Brown et al. presented GPT-3, one of the earliest large-scale language models demonstrating powerful few-shot learning capabilities. The study showed that large transformer-based models could perform multiple language tasks with minimal task-specific training, thereby enhancing contextual understanding and semantic representation in intelligent communication environments [6].

Devlin et al. developed BERT, a bidirectional transformer model designed for deep language understanding. By learning contextual relationships from both directions of text sequences, BERT significantly improved performance across various natural language processing tasks and contributed to the advancement of semantic information extraction techniques [7].

Ho et al. introduced Denoising Diffusion Probabilistic Models (DDPMs), which generate high-quality data through a progressive denoising process. Their work demonstrated superior generation quality and robustness, establishing diffusion models as an emerging alternative to traditional generative architectures in semantic reconstruction and multimedia communication applications [10].

Lan et al. investigated the principles, architecture, and challenges of semantic communication systems. The authors emphasized that future communication networks should focus on transmitting semantic information rather than raw data, thereby improving communication efficiency while reducing unnecessary resource consumption [15].

Xie et al. proposed a deep learning-enabled semantic communication framework that utilized neural networks to extract and reconstruct semantic information. Experimental findings demonstrated that semantic communication systems could outperform conventional communication approaches under limited bandwidth conditions while maintaining meaningful information delivery [16].

Zhang et al. examined the role of semantic communication networks in future wireless systems. The study highlighted the integration of artificial intelligence, semantic encoding, and intelligent resource management as essential components for achieving efficient communication in next-generation network environments [17].

Li et al. explored the application of generative artificial intelligence within semantic communication frameworks. Their findings indicated that generative models can improve semantic representation, enhance information reconstruction, and reduce communication overhead by transmitting semantic features instead of complete datasets [28].

Wang et al. proposed a transformer-based semantic encoding mechanism for intelligent communication networks. The study demonstrated that self-attention-based architectures effectively capture contextual information, resulting in improved semantic accuracy and communication efficiency compared with conventional encoding techniques [29].

Rahman et al. conducted a comparative analysis of multiple generative AI models in semantic communication systems. Their results revealed that transformer-based models generally achieve superior semantic understanding and contextual awareness, whereas VAEs and GANs offer advantages in terms of computational efficiency and data reconstruction capabilities [30].

Singh et al. investigated the application of Large Language Models (LLMs) in semantic communication networks. The authors reported that LLMs provide enhanced contextual reasoning, semantic interpretation, and adaptive communication capabilities, making them promising candidates for intelligent communication infrastructures in future wireless environments [32].

3. Semantic Communication Networks and Generative AI Models

Traditional communication systems are primarily concerned with the accurate transmission of data from a source to a destination. In these systems, communication performance is typically evaluated based on bit error rate, channel capacity, throughput, and transmission reliability. While such metrics remain important, modern communication environments increasingly require systems capable of understanding the context and significance of transmitted information. This requirement has led to the emergence of semantic communication, which seeks to transmit the intended meaning of information rather than merely reproducing data symbols at the receiver end [15].

Semantic communication introduces a fundamental shift in communication objectives. Instead of transmitting complete datasets, semantic communication systems identify and transmit only the information that is relevant to a specific task or application. By focusing on meaning rather than data volume, semantic communication can reduce bandwidth requirements, improve communication efficiency, and support intelligent decision-making processes. Such capabilities are particularly important in next-generation communication environments where massive numbers of devices generate large quantities of data continuously [16], [17].

A typical semantic communication system consists of a semantic encoder, communication channel, semantic decoder, and a knowledge base. The semantic encoder extracts meaningful features from source information and converts them into compact semantic representations. These representations are transmitted through the communication channel and subsequently reconstructed by the semantic decoder using contextual knowledge and learned semantic relationships. The knowledge base plays a crucial role in enabling the receiver to interpret transmitted information accurately and recover the intended meaning even when some data is lost during transmission [15].

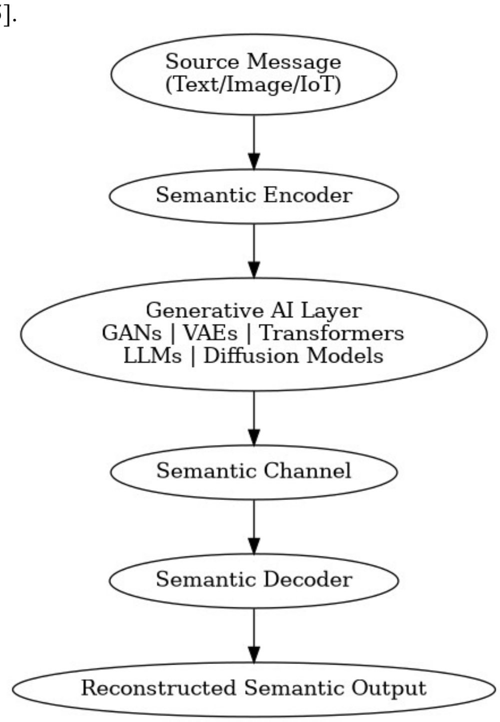

Figure 1 – Generative AI-Enabled Semantic Communication Framework

The growing adoption of artificial intelligence has significantly enhanced the capabilities of semantic communication systems. Deep learning models can identify complex patterns, capture contextual relationships, and generate meaningful representations from diverse data sources. As a result, AI-driven semantic communication frameworks have demonstrated superior performance in tasks involving language processing,

International Journal of Artificial Intelligence and Communication Networks
Volume: 02, Issue: 01, January-March 2026, pp: 63-73
ISSN: 3071-3358

image understanding, speech recognition, and intelligent information reconstruction [11], [19].

Generative Artificial Intelligence has emerged as one of the most influential branches of AI for semantic communication applications. Unlike discriminative models that primarily classify or predict information, generative models learn the underlying distribution of data and generate new outputs that preserve semantic characteristics. This capability enables efficient semantic encoding, information reconstruction, contextual understanding, and adaptive communication, making generative AI highly relevant for future communication systems [28].

Among the various generative architectures, Generative Adversarial Networks (GANs) have gained considerable attention for their ability to generate realistic synthetic data. GANs consist of two neural networks, namely a generator and a discriminator, which compete during the training process. The generator attempts to produce realistic outputs, while the discriminator distinguishes between real and generated samples. Through iterative optimization, GANs learn complex data distributions and support semantic reconstruction tasks in communication systems. Their ability to regenerate missing information makes them useful for bandwidth-constrained communication environments [3].

Variational Autoencoders (VAEs) represent another important class of generative models. VAEs learn compressed latent representations of data through probabilistic encoding and decoding mechanisms. By capturing essential semantic information within a low-dimensional latent space, VAEs facilitate efficient data compression and reconstruction. This characteristic makes them particularly suitable for semantic communication applications where communication resources are limited and transmission efficiency is a critical requirement [4].

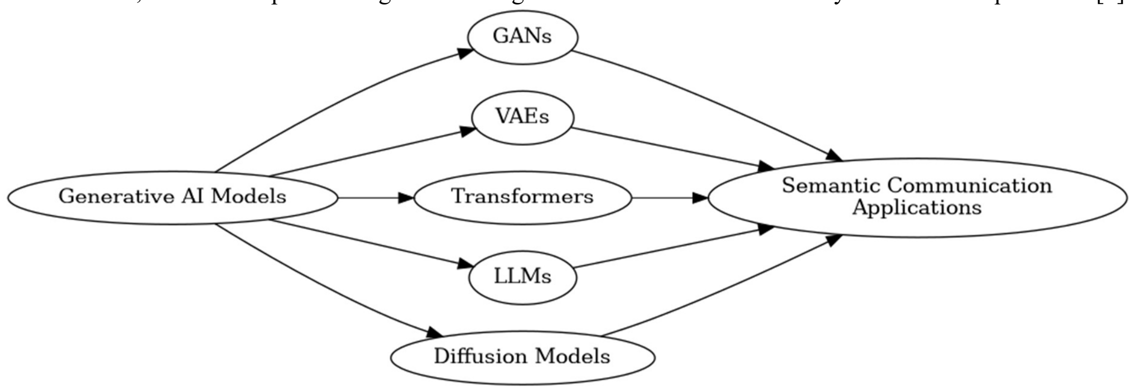

Figure 2 – Generative AI Models in Semantic Communication

Transformer architectures have transformed the field of machine learning by introducing self-attention mechanisms that capture long-range contextual dependencies. Unlike traditional sequential models, Transformers process information in parallel and can effectively model relationships among distant elements within a sequence. Their superior contextual understanding has led to widespread adoption in semantic encoding, language processing, and intelligent communication systems. Transformer-based semantic communication frameworks have demonstrated significant improvements in communication accuracy and contextual representation [5], [29].

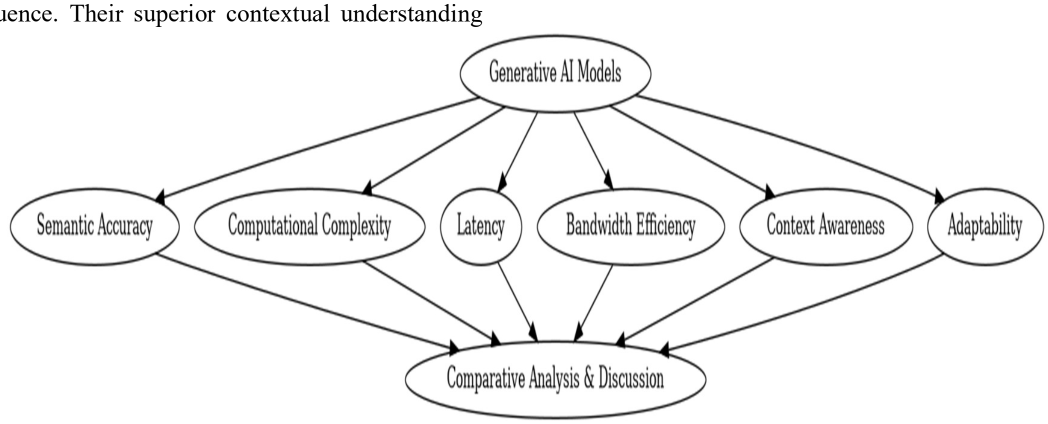

Figure 3 – Comparative Evaluation Framework

International Journal of Artificial Intelligence and Communication Networks
Volume: 02, Issue: 01, January-March 2026, pp: 63-73
ISSN: 3071-3358

Large Language Models represent an advanced extension of Transformer architectures. Trained on massive datasets, LLMs possess extensive knowledge representation capabilities and can perform sophisticated reasoning tasks. These models can interpret user intent, understand contextual information, and generate meaningful responses with remarkable accuracy. In semantic communication networks, LLMs can facilitate intelligent information exchange, semantic compression, and adaptive communication strategies, thereby enhancing communication effectiveness across diverse applications [6], [22], [23], [32].

Diffusion Models have recently emerged as powerful generative frameworks capable of producing high-quality content through iterative denoising processes. Unlike GANs, which rely on adversarial training, diffusion models gradually learn to reconstruct original data from noisy inputs. This approach often produces stable training outcomes and superior generation quality. In semantic communication systems, diffusion models can assist in semantic reconstruction, multimedia communication, and intelligent content recovery under challenging transmission conditions [9], [10], [33].

The integration of these generative models into semantic communication networks offers significant opportunities for improving communication efficiency, reducing bandwidth consumption, and enhancing contextual understanding. However, each model exhibits distinct strengths and limitations regarding computational complexity, latency, semantic accuracy, scalability, and deployment feasibility. Consequently, a systematic comparative analysis is necessary to determine the most appropriate generative architecture for different semantic communication scenarios. The following section presents a detailed comparison of GANs, VAEs, Transformers, Large Language Models, and Diffusion Models based on critical performance metrics relevant to future intelligent communication systems.

4. Comparative Analysis of Generative AI Models

The integration of Generative Artificial Intelligence into semantic communication networks has introduced new possibilities for efficient information transmission, semantic understanding, and intelligent communication. Different generative models exhibit distinct capabilities in semantic representation, computational efficiency, contextual awareness, and communication performance. Therefore, a systematic comparison is necessary to determine their suitability for next-generation communication systems.

Generative Adversarial Networks (GANs) are widely recognized for their ability to generate realistic synthetic data through adversarial learning between a generator and a discriminator. In semantic communication environments, GANs can effectively reconstruct missing information and improve data recovery under constrained communication conditions. Their ability to learn complex data distributions enables efficient semantic reconstruction. However, GANs often suffer from training instability, mode collapse issues, and significant computational requirements, which may limit their deployment in real-time communication scenarios [3], [30].

Variational Autoencoders (VAEs) provide an efficient framework for semantic compression and latent representation learning. By encoding information into a lower-dimensional latent space, VAEs reduce communication overhead while preserving essential semantic content. Their probabilistic architecture facilitates robust information reconstruction even in noisy communication environments. Compared with GANs, VAEs generally offer more stable training and lower computational complexity. Nevertheless, the generated outputs may lack the fidelity and semantic richness achieved by more advanced generative models [4], [30].

Transformer architectures have emerged as one of the most influential technologies in semantic communication. Their self-attention mechanism enables effective modeling of long-range dependencies and contextual relationships within transmitted information. Transformers can capture semantic meaning more accurately than traditional neural network architectures, making them highly effective for semantic encoding and decoding tasks. Furthermore, their parallel processing capability supports faster training and inference compared with recurrent architectures. These advantages have positioned Transformers as a foundational technology for intelligent communication systems [5], [29].

Large Language Models (LLMs) extend Transformer architectures by leveraging massive-scale pretraining on diverse datasets. These models possess advanced reasoning, contextual understanding, and knowledge representation capabilities. Within semantic communication networks, LLMs can interpret user intent, compress semantic information, and reconstruct messages with remarkable contextual accuracy. Their ability to perform semantic inference enables efficient communication even when partial information is received. However, LLMs require substantial computational resources, memory capacity, and energy consumption, creating challenges for deployment in resource-constrained environments such as edge devices and IoT networks [6], [22], [23], [32].

Diffusion Models represent a recent advancement in generative AI and have demonstrated impressive

International Journal of Artificial Intelligence and Communication Networks
Volume: 02, Issue: 01, January-March 2026, pp: 63-73
ISSN: 3071-3358

performance in content generation and reconstruction tasks. Through iterative denoising processes, diffusion models generate high-quality outputs while maintaining semantic consistency. Their training stability often exceeds that of GANs, and they can effectively recover semantic information from degraded inputs. Despite these advantages, diffusion models typically require longer inference times due to their iterative generation mechanism, which may affect latency-sensitive communication applications [9], [10], [33].

The effectiveness of a generative model in semantic communication can be evaluated using several performance indicators. Semantic accuracy measures the model's ability to preserve and reconstruct intended meaning. Computational complexity reflects the processing resources required during training and deployment. Latency represents the time needed for semantic encoding, transmission, and reconstruction. Bandwidth efficiency indicates the model's capability to reduce transmitted information while maintaining semantic integrity. Contextual understanding evaluates how effectively the model captures relationships and meaning within information, whereas adaptability assesses performance across varying communication environments [16], [17], [28].

Among the examined models, Transformer architectures and Large Language Models demonstrate the strongest semantic understanding capabilities because of their attention-based contextual learning mechanisms. They consistently outperform other architectures in tasks requiring semantic reasoning, language understanding, and context-aware communication. VAEs perform particularly well in semantic compression applications, while GANs offer strong data reconstruction capabilities. Diffusion Models provide high-quality semantic reconstruction but may experience higher inference latency compared with Transformer-based approaches [29], [30], [32], [33].

The comparative findings indicate that no single generative model is universally optimal for all communication scenarios. Instead, model selection depends on application requirements, available computational resources, latency constraints, and desired communication performance. Hybrid architectures that combine the semantic understanding capabilities of Transformers and LLMs with the efficiency advantages of VAEs or the reconstruction strengths of GANs may provide a promising direction for future semantic communication systems [34], [35].

Table 1. Comparative Analysis of Generative AI Models in Semantic Communication Networks

Parameter	**GANs**	**VAEs**	**Transformers**	**LLMs**	**Diffusion Models**
Semantic Accuracy	Moderate	Moderate	High	Very High	High
Computational Complexity	High	Moderate	High	Very High	High
Training Stability	Low	High	High	High	Very High
Latency	Moderate	Low	Low	Moderate	High
Bandwidth Efficiency	Moderate	High	High	Very High	Moderate
Contextual Understanding	Low	Moderate	High	Very High	High
Adaptability	Moderate	Moderate	High	Very High	High
Scalability	Moderate	High	High	High	Moderate
Resource Requirement	High	Moderate	High	Very High	High
Suitability for Semantic Communication	Good	Good	Excellent	Excellent	Very Good

The comparative assessment demonstrates that Transformer-based architectures and Large Language Models currently provide the most effective solutions for semantic communication networks because of their superior contextual awareness, semantic representation capabilities, and adaptability. Nevertheless, GANs, VAEs, and Diffusion Models continue to offer valuable advantages in specialized communication tasks, suggesting that hybrid generative frameworks may represent the future direction of intelligent semantic communication systems [34], [35].

5. Applications in Next-Generation Communication Systems

The convergence of Generative Artificial Intelligence and semantic communication has created significant opportunities for the development of intelligent communication infrastructures. As communication networks continue to evolve toward higher levels of automation, connectivity, and intelligence, conventional data transmission mechanisms are increasingly being supplemented by semantic-aware communication approaches. Generative AI models play an important role in this transition by enabling efficient semantic extraction, contextual understanding, adaptive decision-making, and intelligent information reconstruction. Consequently, semantic communication supported by generative AI is expected to become a fundamental component of future communication ecosystems [12], [13], [24].

One of the most significant application areas is fifth-generation (5G) wireless communication networks. Although 5G technologies provide enhanced data rates, low latency, and massive device connectivity, increasing network traffic continues to create challenges related to spectrum utilization and resource management. Generative AI-assisted semantic communication can reduce the amount of transmitted data by focusing on meaningful information rather than complete datasets. This capability improves bandwidth utilization while maintaining communication effectiveness, thereby supporting high-density communication environments and advanced mobile applications [20], [24].

The importance of semantic communication becomes even more pronounced in sixth-generation (6G) networks. Unlike previous communication generations, 6G is envisioned as an intelligence-native communication ecosystem where artificial intelligence is deeply integrated into network operations. Future 6G systems are expected to support holographic communications, immersive virtual environments, digital twins, and autonomous machine interactions. Such applications require communication systems capable of understanding context, predicting user requirements, and dynamically adapting communication strategies. Generative AI models, particularly Transformer-based architectures and Large Language Models, provide advanced semantic reasoning capabilities that align closely with the objectives of intelligent 6G communication networks [12], [13], [14], [25].

The Internet of Things (IoT) represents another critical application domain for semantic communication. Modern IoT environments consist of billions of interconnected sensors and devices that continuously generate large volumes of data. Transmitting all collected information to centralized servers often results in network congestion, increased energy consumption, and communication delays. Semantic communication enables IoT devices to transmit only relevant and meaningful information, thereby reducing communication overhead. Generative AI models further enhance this process through intelligent semantic compression, context-aware filtering, and adaptive information reconstruction, resulting in more efficient IoT communication frameworks [19], [31].

Edge intelligence has emerged as a key technology for supporting real-time applications that require low latency and localized processing. Traditional cloud-based communication architectures may not always satisfy the stringent latency requirements of applications such as industrial automation, healthcare monitoring, and autonomous systems. By integrating generative AI models into edge devices, semantic information can be processed closer to the data source, reducing transmission delays and improving response times. Lightweight Transformer models and compressed generative architectures are increasingly being investigated to enable semantic communication at the network edge while minimizing computational requirements [26], [27].

Autonomous vehicles and intelligent transportation systems also benefit significantly from semantic communication technologies. Autonomous vehicles continuously exchange information regarding road conditions, traffic patterns, pedestrian movements, and navigation decisions. Conventional communication methods often generate substantial communication overhead due to the transmission of raw sensor data. Semantic communication allows vehicles to exchange only essential contextual information, thereby reducing bandwidth requirements and supporting faster decision-making. Generative AI models can further enhance communication reliability by reconstructing incomplete information and predicting relevant environmental conditions [13], [24].

Smart city infrastructures represent another promising application area. Modern smart cities rely on interconnected systems for traffic management, energy distribution, environmental monitoring, public safety, and urban planning. These systems generate enormous amounts of heterogeneous data that must be communicated efficiently among various stakeholders. Semantic communication frameworks can improve interoperability and communication efficiency by prioritizing meaningful information exchange. Generative AI techniques contribute to intelligent data interpretation, semantic integration, and adaptive communication across diverse smart city services [19], [27].

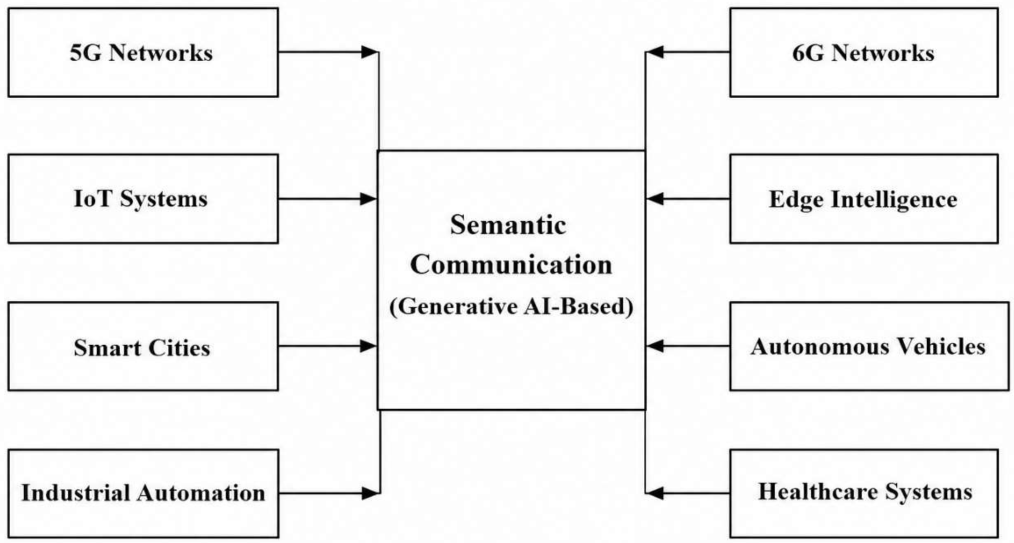

Figure 4. Applications of Generative AI-Based Semantic Communication in Next-Generation Networks

Industrial automation and Industry 4.0 environments increasingly depend on machine-to-machine communication, robotics, and intelligent control systems. Manufacturing facilities require communication systems capable of supporting real-time monitoring, predictive maintenance, and autonomous decision-making. Semantic communication enables industrial devices to exchange task-relevant information while minimizing unnecessary data transmission. Generative AI models support anomaly detection, semantic reconstruction, and adaptive process optimization, thereby improving operational efficiency and communication reliability within industrial networks [24], [31].

Healthcare communication systems also present valuable opportunities for the application of semantic communication technologies. Telemedicine platforms, remote patient monitoring systems, and intelligent diagnostic applications require efficient and reliable information exchange between healthcare providers and patients. Semantic communication can prioritize clinically relevant information while reducing communication overhead. Generative AI models assist in interpreting medical data, summarizing patient information, and reconstructing incomplete healthcare records, thereby enhancing the quality and efficiency of healthcare communication services [19], [24].

Overall, the integration of Generative AI into semantic communication networks provides substantial benefits across diverse application domains. The ability to understand context, compress semantic information, reconstruct missing data, and support intelligent communication decisions makes generative models particularly valuable for future communication ecosystems. As communication technologies continue to evolve toward AI-native architectures, semantic communication is expected to become a critical enabler of efficient, adaptive, and intelligent network operations [28], [34], [35].

6. Challenges and Future Research Directions

The integration of Generative Artificial Intelligence into semantic communication networks offers substantial benefits; however, several challenges must be addressed before large-scale deployment becomes feasible. One of the primary challenges is **semantic ambiguity**, where the intended meaning of transmitted information may vary depending on context, user knowledge, and application requirements. Ensuring consistent semantic interpretation remains a critical concern in intelligent communication systems [15], [16].

Another significant issue is **privacy and security**. Semantic communication often involves the processing of sensitive information, making systems vulnerable to privacy leakage, adversarial attacks, and unauthorized access. Developing secure and privacy-preserving communication frameworks is therefore essential for practical implementation [21], [24].

The **computational complexity** of advanced generative models, particularly Large Language Models and Diffusion Models, presents additional challenges. These models require substantial processing power, memory, and energy resources, limiting their deployment in resource-constrained environments such as IoT devices and edge networks [22], [23], [33].

Furthermore, **scalability and standardization** remain open research issues. Future communication networks will involve billions of connected devices, requiring semantic communication systems that can operate efficiently at large scales while maintaining interoperability across diverse platforms and applications [12], [25].

Future research should focus on developing **hybrid generative architectures** that combine the strengths of multiple AI models to improve communication efficiency and semantic accuracy. Additional attention should be directed toward **federated semantic learning**, **lightweight Large Language Models**, **privacy-preserving communication frameworks**, and **AI-native 6G networks**. These advancements are expected to enhance the reliability, scalability, and practical adoption of semantic communication technologies in future intelligent communication ecosystems [31], [34], [35].

7. Conclusion

The evolution of communication systems has gradually shifted the focus from transmitting large volumes of data to delivering meaningful and relevant information. In this context, semantic communication has emerged as a promising approach for improving communication efficiency by emphasizing the understanding and transmission of meaning rather than the exact reproduction of data. The integration of Generative Artificial Intelligence into semantic communication networks has further strengthened this paradigm by enabling intelligent information processing, semantic compression, contextual understanding, and adaptive communication.

This study examined the role of major generative AI models, including Generative Adversarial Networks, Variational Autoencoders, Transformer architectures, Large Language Models, and Diffusion Models, in semantic communication environments. The comparative analysis demonstrated that each model offers unique advantages depending on the communication requirements. While some models are more suitable for semantic compression and information reconstruction, others excel in contextual understanding and intelligent decision-making. Among the examined approaches, Transformer-based architectures and Large Language Models show considerable potential for supporting advanced semantic communication tasks due to their ability to capture contextual relationships and process complex information effectively.

The study also highlighted the growing relevance of semantic communication in emerging technologies such as next-generation wireless networks, Internet of Things ecosystems, edge computing environments, autonomous systems, smart cities, and industrial automation. As these technologies continue to generate vast amounts of data, the need for intelligent communication mechanisms that can efficiently exchange meaningful information becomes increasingly important.

Despite the promising opportunities, several challenges remain, including semantic ambiguity, privacy protection, security concerns, computational requirements, and scalability issues. Addressing these challenges will require continued research and collaboration among researchers, industry stakeholders, and communication technology developers. Future developments in hybrid generative models, lightweight AI architectures, and intelligent network design are expected to further enhance the effectiveness of semantic communication systems.

Overall, Generative Artificial Intelligence has the potential to significantly influence the future of communication networks by enabling more intelligent, adaptive, and context-aware information exchange. As research in this area continues to mature, semantic communication is likely to become a key component of future digital communication infrastructures, supporting more efficient and meaningful interactions between humans, machines, and connected devices.

References

[1]. Shannon, Claude E. (1948). “A Mathematical Theory of Communication.” *Bell System Technical Journal*, 27(3), 379–423.

[2]. Weaver, Warren (1949). *The Mathematical Theory of Communication*. University of Illinois Press.

[3]. Goodfellow, Ian et al. (2014). “Generative Adversarial Nets.” *Advances in Neural Information Processing Systems (NeurIPS)*, 2672–2680.

[4]. Kingma, Diederik P. & Welling, M. (2014). “Auto-Encoding Variational Bayes.” *International Conference on Learning Representations (ICLR)*.

[5]. Vaswani, Ashish et al. (2017). “Attention Is All You Need.” *NeurIPS*, 5998–6008.

[6]. Brown, Tom B. et al. (2020). “Language Models are Few-Shot Learners.” *NeurIPS*, 33, 1877–1901.

[7]. Devlin, Jacob et al. (2019). “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.” *NAACL-HLT*, 4171–4186.

[8]. Radford, Alec et al. (2019). "Language Models are Unsupervised Multitask Learners." *OpenAI Technical Report*.
[9]. Rombach, Robin et al. (2022). "High-Resolution Image Synthesis with Latent Diffusion Models." *CVPR*, 10684–10695.
[10]. Ho, Jonathan et al. (2020). "Denoising Diffusion Probabilistic Models." *NeurIPS*, 33, 6840–6851.
[11]. LeCun, Yann, Bengio, Y., & Hinton, G. (2015). "Deep Learning." *Nature*, 521(7553), 436–444.
[12]. Zhang, Z. et al. (2019). "6G Wireless Networks: Vision, Requirements, Architecture, and Key Technologies." *IEEE Vehicular Technology Magazine*, 14(3), 28–41.
[13]. Saad, Walid, Bennis, M., & Chen, M. (2020). "A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems." *IEEE Network*, 34(3), 134–142.
[14]. Strinati, Emilio C. et al. (2019). "6G: The Next Frontier—From Holographic Messaging to Artificial Intelligence." *IEEE Vehicular Technology Magazine*, 14(3), 42–50.
[15]. Lan, Tian et al. (2021). "Semantic Communication: Principles and Challenges." *IEEE Network*, 35(2), 210–217.
[16]. Xie, Hui et al. (2021). "Deep Learning Enabled Semantic Communication Systems." *IEEE Transactions on Signal Processing*, 69, 2663–2675.
[17]. Zhang, Peng et al. (2022). "Semantic Communication Networks for Future Wireless Systems." *IEEE Communications Magazine*, 60(1), 30–36.
[18]. Qin, Zhijin et al. (2019). "Deep Learning in Physical Layer Communications." *IEEE Wireless Communications*, 26(2), 93–99.
[19]. Liu, Y. et al. (2020). "Artificial Intelligence for Next-Generation Wireless Networks: A Survey." *IEEE Communications Surveys & Tutorials*, 22(4), 2237–2275.
[20]. Bennis, Mehdi et al. (2018). "Ultra-Reliable and Low-Latency Wireless Communication." *Proceedings of the IEEE*, 106(10), 1834–1853.
[21]. Bommasani, Rishi et al. (2021). "On the Opportunities and Risks of Foundation Models." *Stanford CRFM Report*.
[22]. Touvron, Hugo et al. (2023). "LLaMA: Open and Efficient Foundation Language Models." *arXiv Preprint*.
[23]. OpenAI (2023). "GPT-4 Technical Report." *OpenAI Research*.
[24]. Chen, Min et al. (2021). "Artificial Intelligence in 6G Networks: Intelligence, Security and Semantic Communication." *IEEE Network*, 35(2), 34–42.
[25]. You, Xiaohu et al. (2021). "Towards 6G Wireless Communication Networks: Vision, Enabling Technologies, and New Paradigm Shifts." *Science China Information Sciences*, 64(1), 1–74.
[26]. Park, J. et al. (2021). "Communication-Efficient Distributed Learning over Wireless Networks." *IEEE Communications Magazine*, 59(4), 16–22.
[27]. Tang, Fei et al. (2021). "AI-Empowered Intelligent Communication for Future Networks." *IEEE Wireless Communications*, 28(6), 54–61.
[28]. Li, X. et al. (2022). "Generative AI-Assisted Semantic Communication Models." *IEEE Access*, 10, 88901–88919.
[29]. Wang, T. et al. (2022). "Transformer-Based Semantic Encoding for Intelligent Communication Networks." *Computer Communications*, 188, 64–78.
[30]. Rahman, M. A. et al. (2023). "Comparative Analysis of Generative Models in Semantic Communication Systems." *Wireless Networks*, 29(7), 2551–2570.
[31]. Kumar, R. et al. (2023). "AI-Driven Semantic Compression in Next-Generation Networks." *Future Generation Computer Systems*, 143, 210–224.
[32]. Singh, A. et al. (2024). "Large Language Models for Semantic Communication in Intelligent Networks." *Journal of Network and Computer Applications*, 225, 103715.
[33]. Patel, S. et al. (2024). "Diffusion Models for Adaptive Semantic Communication Networks." *Sensors*, 24(2), 870.
[34]. Verma, P. et al. (2024). "Generative AI-Based Communication Optimization in Smart Networks." *IEEE Access*, 12, 44112–44130.
[35]. Vishwanath Nikhil , Dr. Mohammed Abdul Bari ," Real Time Powerlifting Form Assessment using Yolov5 and Mediapipe", International Journal of Information Technology and Computer Engineering, SSN 2347–3657, Volume 13, Special Issue 2s, 2025, Pages 574-584
[36]. Zhou, X. et al. (2025). "Hybrid Generative Models for Semantic Wireless Communication." *Computer Networks*, 245, 110512.