Physical Communication 73 (2025) 102857

Contents lists available at ScienceDirect

Physical Communication

journal homepage: www.elsevier.com/locate/phycom

Full length article

A survey of domain generalization in AI-enabled semantic communication: Architecture, challenges and future opportunities

Muhammad Furqan Zia [a],*, Messaoud Ahmed Ouameur [a], Miloud Bagaa [a], Daniel Massicotte [a], Adlen Ksentini [b]

[a] Department of Electrical and Computer Engineering, University of Quebec at Trois-Rivieres (UQTR), QC G9A 5H7, Canada
[b] Department of Communication Systems, EURECOM, Sophia Antipolis, 06904, France

ARTICLE INFO

Keywords:
Domain generalization
Generalizable communication systems
Semantic communication
AI-enabled wireless systems

ABSTRACT

The growing integration of artificial intelligence (AI) into wireless communication systems is driving a shift toward semantic communication, an emerging paradigm that prioritizes the exchange of meaning over raw data. However, semantic communication systems face major challenges when deployed across diverse and unseen domains due to variations in language, context, and channel conditions. This survey provides a comprehensive overview of Domain Generalization (DG) as a key enabler for improving the robustness and adaptability of AI-enabled semantic communication. We explore the types of domain shifts and review the latest DG techniques applicable to semantic communication. Additionally, the paper discusses architectural considerations and real world applications across varied wireless scenarios. Unlike prior works, this survey brings together DG strategies specifically within the context of semantic communication, identifying open challenges and future research directions such as scalable adaptation, resource efficient deployment, and resilience in dynamic environments. It aims to serve as a timely resource for researchers and practitioners working to develop reliable, generalizable communication systems for next generation networks.

1. Introduction

The evolution of wireless networks from 1G to 6G represents an impressive journey through technological progress, driven by an ever growing demand for faster, more intelligent, and context aware communication systems. It commenced with first generation (1G) networks, which introduced mobile communication through basic voice calls with limited mobility. Subsequent generations, such as 2G and 3G, brought digital voice and data transmission capabilities, significantly expanding mobile communication possibilities. With the advent of 4G, mobile networks experienced a transformative leap, facilitating high speed data access, widespread internet usage, and app driven interactions. The emergence of 5G elevated connectivity further, delivering ultra fast data rates, exceptionally low latency, and supporting extensive IoT deployments [1]. Now, standing at the threshold of the 6G era, we anticipate groundbreaking advancements such as terahertz communication, real time immersive applications, and the deep integration of artificial intelligence (AI) to enable semantic communication which is a paradigm shift that focuses on transmitting the intended meaning rather than raw data [2,3]. This evolution signals a transition from merely connecting devices to enabling intelligent, adaptive, and semantically rich interactions. Fig. 1 illustrates this evolutionary trajectory from early mobile communication to the envisioned capabilities of 6G networks [4].

The continuous pursuit of increased bandwidth and reduced latency remains a cornerstone of current 5G networks and the forthcoming 6G, reflecting the need for more intelligent and efficient communication mechanisms. Modern digital services, such as high definition video streaming, cloud gaming, real time augmented reality (AR), and large scale Internet of Things (IoT) applications, impose unprecedented demands on network capacity and responsiveness [5]. While traditional communication systems rely on transmitting raw data, this approach becomes increasingly unsustainable as data volumes grow exponentially. The anticipated 6G architecture aims to address these challenges through technologies such as edge computing and network slicing, which optimize data routing and processing to minimize latency [6]. However, to achieve truly scalable and efficient communication, a fundamental shift is required in how information is transmitted. Semantic communication (SemCom) meets this need by focusing on the transmission of meaningful information rather than raw symbols, thereby

* Corresponding author.
E-mail addresses: Muhammad.Furqan.Zia@uqtr.ca (M.F. Zia), messaoud.ahmed.ouameur@uqtr.ca (M.A. Ouameur), miloud.bagaa@uqtr.ca (M. Bagaa), daniel.massicotte@uqtr.ca (D. Massicotte), adlen.ksentini@eurecom.fr (A. Ksentini).

https://doi.org/10.1016/j.phycom.2025.102857
Received 6 June 2025; Received in revised form 14 September 2025; Accepted 17 September 2025
Available online 24 September 2025

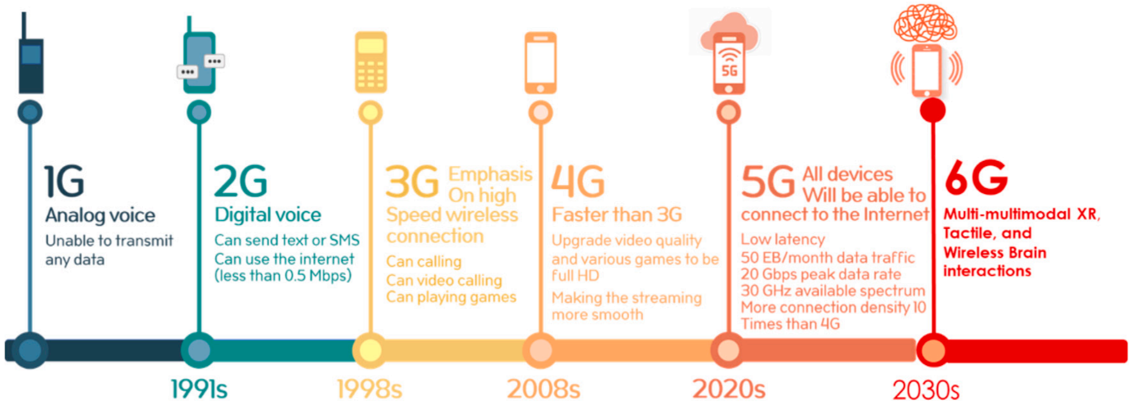

Fig. 1. Evolution of wireless networks.

reducing the amount of data that must be sent without sacrificing the integrity of the intended message. By enabling the receiver to infer the original meaning with fewer bits, SemCom significantly alleviates bandwidth pressure and shortens transmission times. In doing so, it not only enhances spectral and energy efficiency but also lays the foundation for a more responsive, context aware, and scalable communication framework critical for the demands of future 6G systems [7].

To fulfill the full potential of semantic communication within 6G networks, integrating artificial intelligence into communication designs becomes imperative [8]. While traditional wireless networks excel at transmitting vast volumes of data, they fall short in interpreting and contextualizing the transmitted content. This limitation becomes more pronounced with the emergence of 6G, where the need for intelligent, adaptive, and context aware communication systems intensifies [9]. AI-enabled semantic communication aligns naturally with this evolution, enabling systems to not only transmit but also understand and act on the semantic content of data. By embedding contextual awareness and cognitive reasoning into communication processes, AI-driven SemCom systems are positioned to meet the unique demands of 6G where applications extend beyond mere data exchange to include personalized, real time, and meaningful information interactions [10].

Within the landscape of 6G and semantic communication, domain generalization (DG) emerges as a critical challenge. Wireless communication networks in the 6G era will operate in dynamic and constantly evolving environments, requiring AI models to adapt effectively across diverse and unpredictable scenarios [11]. Domain generalization addresses the need for AI models to maintain their performance despite domain shifts caused by changes in network conditions, contexts, and operational scenarios. The adaptability provided by domain generalization significantly impacts the reliability and performance quality of semantic communication systems. Techniques that facilitate knowledge transfer across various domains are essential to ensure consistent AI performance in the heterogeneous environments envisioned for 6G [12].

Domain generalization thus represents a transformative capability for AI-enabled semantic communication systems. Just as experienced travelers adapt to new countries using insights from past experiences, AI models leveraging domain generalization can seamlessly adapt across diverse network conditions, from dense urban environments rich with data to remote areas with limited information availability. These capabilities foster resilient, versatile, and reliable communication systems capable of interpreting and conveying complex meanings clearly and effectively across varying contexts [13]. Ultimately, domain generalization promises to evolve communication networks into sophisticated ecosystems capable of maintaining robust, context aware interactions, significantly enhancing connectivity and interaction depth across global communication platforms [14].

The motivation behind this survey stems from the pressing need to address a fundamental challenge in AI-enabled semantic communication: ensuring consistent performance across diverse, dynamic, and previously unseen environments. As 6G networks evolve to support complex and heterogeneous communication scenarios, traditional AI models struggle to maintain generalization beyond their training domains. Domain Generalization (DG) offers a promising direction by equipping models with the ability to adapt to new contexts without explicit retraining or prior exposure. Despite its critical role, the application of DG in semantic communication remains underexplored, with limited efforts to consolidate current techniques, challenges, and use cases in a unified framework. This work aims to fill that gap by presenting a structured overview of DG strategies, analyzing their applicability in semantic communication systems, and highlighting future directions that can guide the design of more adaptable and robust 6G communication infrastructures.

1.1. Contributions and paper organization

To the best of our knowledge, this paper presents the first comprehensive survey that specifically investigates the role of Domain Generalization (DG) in semantic communication. The major contributions of this work are outlined as follows:

- **Focused Analysis of Domain Generalization in Semantic Communication:** This paper highlights the pivotal role of domain generalization in ensuring that AI-enabled semantic communication systems maintain robust performance across diverse and previously unseen domains. We explore how DG addresses the challenge of domain shifts and supports adaptability in real-world, heterogeneous communication environments.
- **Survey of State of the Art DG Techniques:** We provide a structured overview of recent advancements in DG, particularly within the context of semantic communication. Our survey categorizes and analyzes techniques such as data augmentation, meta-learning, adversarial training, and robust optimization in terms of their relevance and applicability to communication systems.
- **Identification of Application Scenarios and Challenges:** We present several real world use cases that benefit from DG-enhanced semantic communication and outline the practical challenges associated with deploying DG models in dynamic, resource constrained, or low data environments.
- **Roadmap for Future Research:** Based on the gaps identified in existing literature, we propose promising directions for future research. These include improving DG efficiency, minimizing training overhead, and enhancing generalization across multimodal and multilingual communication scenarios.

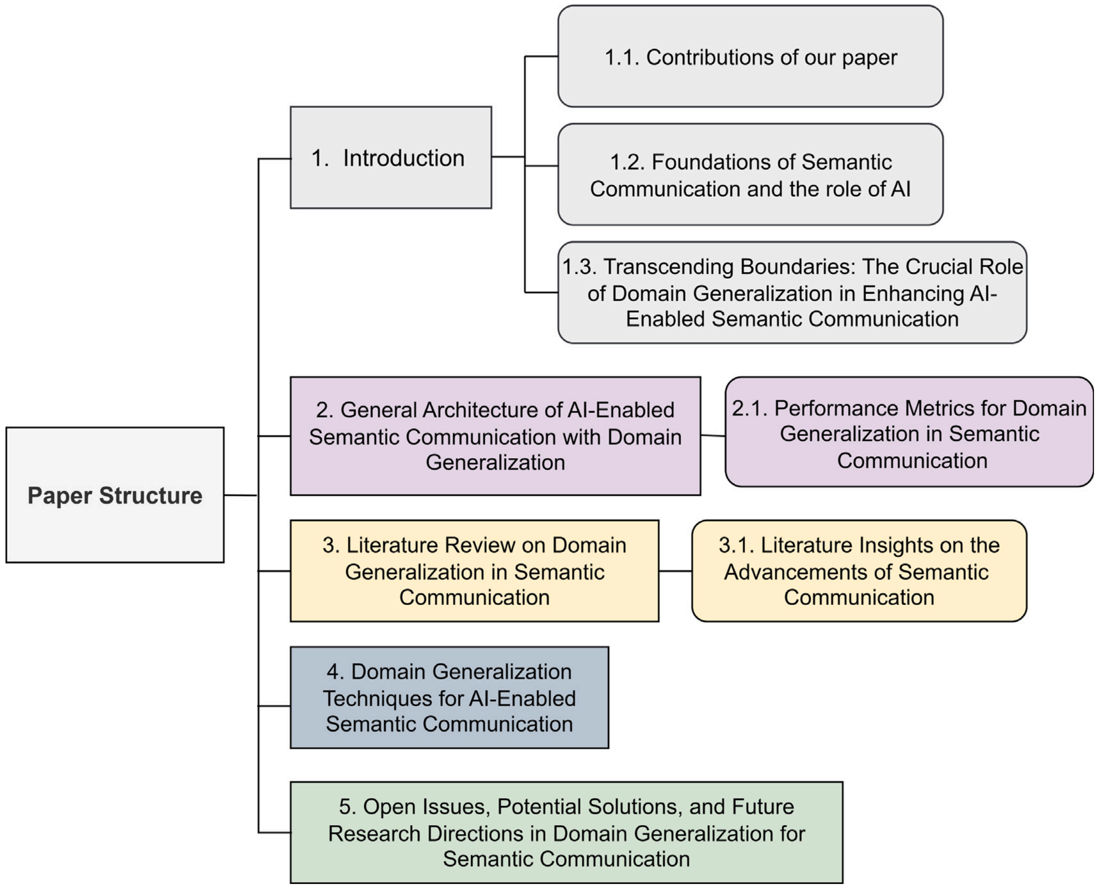

Fig. 2. Paper Structure.

For the reader's convenience, the main abbreviations used in this paper are listed in Table 1, and the paper's structure is illustrated in Fig. 2.

By the end of this survey, readers will gain a comprehensive understanding of how domain generalization contributes to building robust, adaptive, and scalable semantic communication systems ultimately supporting the development of resilient 6G wireless networks.

1.2. Foundations of semantic communication and the role of AI

Semantic communication represents a significant evolution in the realm of communication theory, addressing critical limitations inherent in traditional communication methodologies grounded in Shannon's classical information theory. Shannon's information theory primarily focuses on the technical aspects of data transmission, emphasizing accurate and efficient delivery of bits over communication channels. While revolutionary in ensuring reliable data transmission through advanced encoding and decoding techniques, such as CDMA, OFDM, and MIMO [15,16], Shannon's framework predominantly measures communication effectiveness in terms of bit level accuracy, leaving the meaningful interpretation and context of transmitted data largely unaddressed [17].

Weaver, recognizing these critical gaps in Shannon's theory, proposed a complementary communication model in 1949, which articulated the significance of semantic interpretation in effective communication [18]. Weaver introduced three distinct communication levels: technical, semantic, and efficacy, as depicted in Fig. 3 [19]. The technical level corresponds directly with Shannon's model, emphasizing the accurate transmission of signals and bits across communication channels. However, Weaver argued that successful communication requires not only accurate transmission but also precise understanding at the semantic level, where the meaning and context of transmitted messages are comprehensively interpreted. The efficacy level further encompasses the practical impact and usefulness of the received message on the receiver.

To mathematically illustrate the bandwidth efficiency achieved through semantic communication compared to Shannon's traditional approach, consider the following formulation. Let B_c denote the bandwidth used to transmit a full message of size N bits:

$$B_c \propto N. \tag{1}$$

Semantic communication, by contrast, only transmits essential semantic content. The required bandwidth B_s can be represented as:

$$B_s \propto \alpha N, \tag{2}$$

where $0 < \alpha < 1$ is the fraction of semantically essential information. The bandwidth saving S achieved is:

$$S = \frac{B_c - B_s}{B_c} = 1 - \alpha. \tag{3}$$

This reflects the proportionate bandwidth reduction achieved through semantic communication, with α indicating the efficiency of semantic content compression. Thus, the mathematical representation simplifies the concept and does not capture the complexity of semantic encoding and decoding processes. Still, it illustrates the fundamental principle of how focusing on transmitting meaning rather than raw data can lead to more efficient use of bandwidth. The Fig. 4 illustrates the conceptual difference between conventional communication and semantic communication, demonstrating how semantic communication enables intelligent and context aware interactions that result in lower bandwidth utilization. In conventional communication systems, a transmitted message such as "The weather today is very hot" would typically

Table 1
List of common acronyms.

Acronym	Description
AI	Artificial Intelligence
AR	Augmented Reality
AWGN	Additive White Gaussian Noise
BERT	Bidirectional Encoder Representations from Transformers
BLEU	Bilingual Evaluation Understudy Score
CIFAR-10	Canadian Institute For Advanced Research 10-class Dataset
COCO	Common Objects in Context
CSI	Channel State Information
DAEL	Domain Adaptive Ensemble Learning
DANN	Domain Adversarial Neural Network
DG	Domain Generalization
DIR	Decoding Information Resolution
DIV2K	DIVerse 2K Resolution Dataset
DRO	Distributionally Robust Optimization
D-SAM	Domain Specific Aggregation Module
eMBB	Enhanced Mobile Broadband
Europarl	European Parliament Proceedings Parallel Corpus
F1	F1-Score (harmonic mean of Precision and Recall)
FDSD	Fréchet Deep Speech Distance
FN	False Negative
FP	False Positive
GAN	Generative Adversarial Network
GLUE	General Language Understanding Evaluation
GRL	Gradient Reversal Layer
IoT	Internet of Things
JSCC	Joint Source Channel Coding
KB	Knowledge Base
LLM	Large Language Model
LPIPS	Learned Perceptual Image Patch Similarity
MAML	Model Agnostic Meta-Learning
MetaReg	Meta Regularization
MI	Mutual Information
Mini-ImageNet	Subset of ImageNet for Few-Shot Learning
MLDG	Model Agnostic Meta-Learning for Domain Generalization
MIMO	Multiple Input Multiple Output
mAP	Mean Average Precision
MS-SSIM	Multi-Scale Structural Similarity Index Measure
NLP	Natural Language Processing
NYU Depth V2	New York University Depth Dataset V2
OFDM	Orthogonal Frequency Division Multiplexing
OOD	Out-Of-Domain
PPL	Perplexity
PSNR	Peak Signal to Noise Ratio
ROUGE	Recall-Oriented Understudy for Gisting Evaluation
RSC	Representation Self Challenging
SemCom	Semantic Communication
SNR	Signal to Noise Ratio
SSL	Self-Supervised Learning
SSIM	Structural Similarity Index Measure
STM	Semantic Textual Similarity Measure
STS	Semantic Textual Similarity Benchmark
TN	True Negative
TP	True Positive
URLLC	Ultra Reliable Low Latency Communication
VCTK	Voice Cloning Toolkit Corpus
VLM	Vision Language Model
WER	Word Error Rate
WMT	Workshop on Machine Translation
XR	Extended Reality
XSum	Extreme Summarization Dataset

be delivered word for word using standard encoding and decoding processes. In contrast, semantic communication emphasizes the transfer of meaning rather than exact wording. For example, if the receiver already possesses contextual knowledge such as awareness that it is summer in a region known for high temperatures the message could be compressed to a shorter phrase like “It’s hot” or even represented by a symbolic value denoting “very hot” weather. The receiver then uses its own background knowledge, situational awareness, or prior information to fill in the missing details and recover the full meaning that the sender intended.

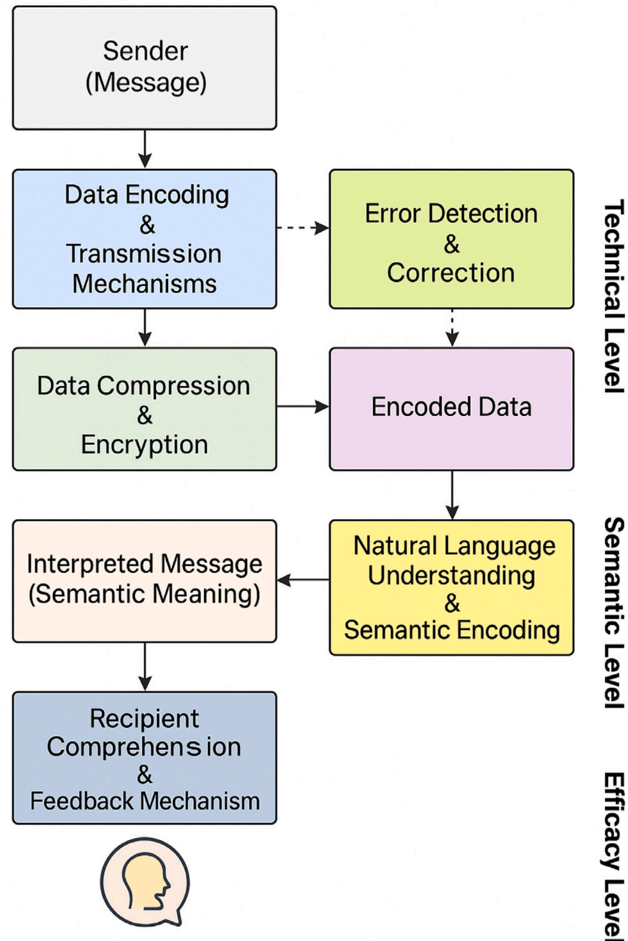

Fig. 3. The three major levels of semantic communications as introduced by Weaver.

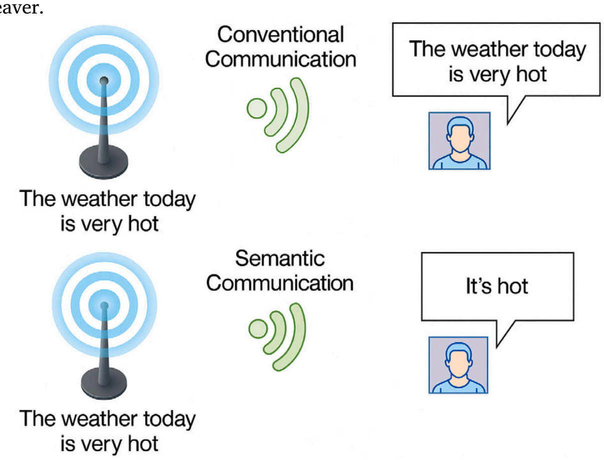

Fig. 4. Transmission via Normal/Conventional Communication vs Semantic Communication.

Though Weaver’s concept of semantic level communication was fascinating, unfortunately, it could not become the need of the hour and due to technical constraints further research in this area was silenced. On the other hand, Shannon’s information theory, which primarily focuses on the technical aspect of information transmission, which is about how to encode and transmit information efficiently and reliably over a noisy channel received more focus on developing different

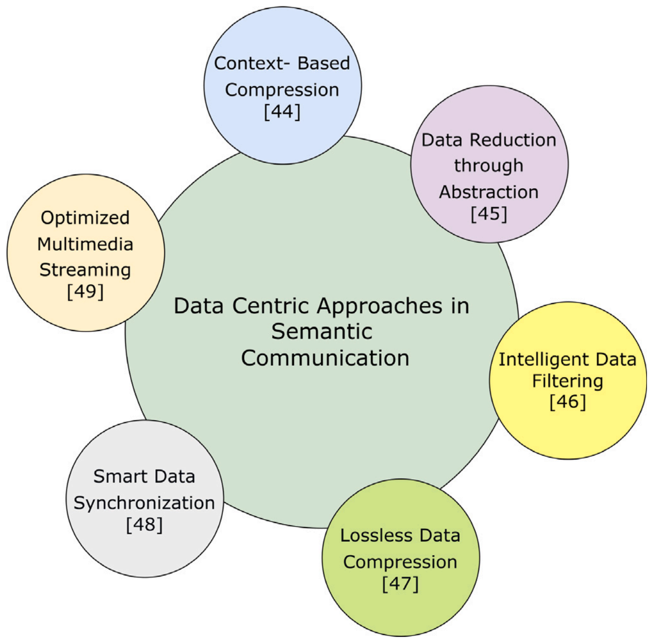

Fig. 5. Data Centric Semantic Communication approaches.

communication techniques [20]. The good news is with the recent advancement in AI, prerequisite techniques like semantic communication have started gaining popularity to add the ingredient of intelligence in communication systems. In the past few years, the topic of semantic communication has gained the attention of many researchers worldwide. As a result, numerous system models and transceiver and receiver designs have recently been presented [21–25] which proves that semantic communication is a promising enabler for the technologies that will be using 6G and beyond applications.

Semantic communication can help save bandwidth by optimizing the amount of information transmitted and reducing unnecessary data transmission. In Fig. 5 we display the data centric semantic communication approaches and now we further describe how it works:

1. Context Based Compression: Semantic communication enables the exchange of meaningful messages that convey the intended information more efficiently. Instead of transmitting raw or redundant data, the communication can be focused on conveying the essential semantic meaning. For example, in a chat conversation, using abbreviations or emojis to represent specific ideas can convey the same message with fewer characters, saving bandwidth [26].

2. Data Reduction through Abstraction: By conveying the underlying meaning of information, semantic communication can use abstract representations rather than detailed, data heavy descriptions. For instance, sending a compressed summary or the key points of a document instead of the entire document can save bandwidth while conveying the essential information [27].

3. Intelligent Data Filtering: In systems with AI or machine learning capabilities, semantic communication allows for more intelligent data filtering. Instead of sending all data to a central processing unit, only the relevant and semantically important data can be selected and transmitted, reducing unnecessary data traffic and saving bandwidth [28].

4. Lossless Data Compression: Semantic communication can enable lossless data compression techniques that maintain the exact meaning of the information while reducing its size. By utilizing appropriate algorithms, the communication can achieve compression without losing any vital details, thus saving bandwidth during data transmission [29].

5. Smart Data Synchronization: In scenarios where multiple devices or systems need to synchronize data, semantic communication can ensure that only the changes or updates to the data are transmitted, rather than sending the entire datasets repeatedly. This selective synchronization helps save bandwidth by minimizing the amount of data transferred [30].

6. Optimized Multimedia Streaming: In media streaming applications, semantic communication can facilitate adaptive bitrate streaming, where the quality of the media (e.g., video or audio) is adjusted based on the user's device capabilities and network conditions. This optimization ensures that the most relevant and suitable media content is transmitted, saving bandwidth and providing a smoother user experience [31].

These advancements collectively illustrate how semantic communication, rooted in Weaver's foundational ideas and empowered by AI, transcends the traditional limitations imposed by Shannon's model. It enables not only efficient transmission but also meaningful understanding across diverse and evolving communication scenarios. However, the practical realization of semantic communication systems in dynamic 6G environments introduces a new challenge: ensuring consistent semantic fidelity across unseen or varying domains. This calls for robust learning frameworks specifically, domain generalization techniques that empower semantic models to adapt and perform reliably under domain shifts without requiring retraining. The next sub-section explores this critical frontier, highlighting how domain generalization forms a cornerstone in achieving scalable, resilient, and intelligent semantic communication.

1.3. Transcending boundaries: The crucial role of domain generalization in enhancing AI-enabled semantic communication

Domain generalization (DG) stands as a foundational pillar at the intersection of artificial intelligence and semantic communication, particularly within the heterogeneous and dynamic environments of 6G and beyond. Unlike conventional AI models that rely heavily on domain specific training data, DG equips models with the ability to transfer learned representations to novel, unseen domains without retraining [32]. This adaptability is crucial for preserving semantic fidelity across fluctuating linguistic, cultural, contextual, and network specific conditions that characterize modern communication systems [33].

The role of DG is not merely technical but systemic. As AI transforms communication from raw data exchange into semantically rich and context aware interaction, semantic models are increasingly exposed to diverse operating conditions. Without DG, these models risk brittleness and inconsistency when deployed outside their original training environments. By contrast, DG supports resilience by ensuring that semantic encoders and decoders maintain reliability across variations in language, modality, and user preferences, while retaining the bandwidth and latency efficiencies inherent in semantic compression [34]. This capacity is particularly critical for cross lingual dialogues, multimodal interactions, and adaptive communication in highly dynamic 6G networks.

Beyond enabling robust message interpretation, DG strengthens the scalability and inclusivity of semantic communication. Systems designed with DG are better prepared to handle global deployment scenarios where linguistic diversity, cultural nuance, and shifting social contexts come into play. This makes them suitable not only for multilingual collaboration and immersive applications but also for socially attuned use cases such as assistive technologies and emergency communication, where trustworthiness and accessibility are paramount [35]. In this way, DG extends the benefits of semantic communication to broader populations and ensures equitable access to intelligent systems.

1.3.1. Domain shifts in AI-enabled semantic communication

Domain shifts in AI-enabled semantic communication represent a pivotal challenge as networks transition from controlled settings to real world deployment across diverse and non-stationary environments. These shifts refer to changes in data distribution, communication context, or operating conditions that can significantly affect semantic fidelity. Because semantic communication systems depend on AI models

Table 2
Domain shifts in semantic communication: examples and challenges.

Domain shifts	Example	Domain generalization challenge
Multimodal Communication	AI interpreting a video with text captions.	Generalizing understanding and integration skills across various sensory data types.
Linguistic Variability	AI adapting to regional dialects or slang.	Enabling models to work effectively across diverse linguistic expressions.
Cultural and Contextual Shifts	AI adjusting communication for cultural norms.	Aligning communication strategies with varied cultural norms and cues.
Technological Evolution	AI using new protocols in smart homes.	Adapting to evolving technologies without retraining.
Network Dynamics	AI functioning over different internet speeds.	Maintaining communication quality across variable connectivity.
Physical Layer Variability	AI adapting to Rayleigh fading or hardware impairments.	Ensuring semantic fidelity across unseen channel and hardware conditions.

to understand, compress, and reconstruct meaning, robustness against domain variability is essential. DG addresses this by enabling AI models to generalize across unseen domains without retraining [33].

To better understand these challenges, Table 2 categorizes common sources of domain shift and illustrates their impact on semantic communication. It highlights how DG mitigates their effects to maintain semantic consistency across heterogeneous contexts.

a. Multimodal Communication: The rise of multimodal communication combining text, audio, images, and contextual cues requires AI models to infer meaning across heterogeneous inputs. Shifts in modality (e.g., moving from text to video, or from clean to noisy audio) can disrupt semantic interpretation. DG provides cross modal resilience, enabling models to transfer knowledge across modalities while maintaining coherence [36].

b. Linguistic Variability: Semantic communication must support diverse languages, dialects, and registers. Regional expressions, evolving slang, and technical jargon introduce variability that can degrade understanding. DG mitigates this by abstracting high level semantics, allowing models to generalize across linguistic distributions and ensure consistent interpretation in multilingual and informal contexts [37].

c. Cultural and Contextual Shifts: Cultural norms and contextual expectations shape how meaning is conveyed and received. Domain shifts occur when AI models encounter unfamiliar values or implicit conventions. DG addresses this by leveraging transferable semantic representations, fostering culturally adaptive and globally relevant communication [38].

d. Technological Evolution: Evolving devices, codecs, and communication standards create shifting technological environments. Without DG, semantic systems risk obsolescence when infrastructure changes. DG enables continuity by allowing AI models to generalize knowledge across new platforms, ensuring long term reliability and future readiness [39].

e. Network Dynamics: Bandwidth fluctuations, latency, jitter, and signal variability are inherent to wireless networks. Semantic fidelity must be maintained despite these constraints. DG enhances resilience by enabling models to adjust semantic compression and reconstruction strategies dynamically, preserving communication quality in volatile conditions [40].

f. Physical Layer Variability: Beyond higher layer contextual and linguistic factors, many critical domain shifts originate at the physical layer. Diverse channel conditions (e.g., AWGN, Rayleigh, Rician, and Nakagami fading), fluctuations in signal to noise ratio (SNR), and hardware impairments such as IQ imbalance, phase noise, and nonlinearities in RF front ends can all introduce distributional shifts that directly affect semantic fidelity. From a DG perspective, these impairments can be treated as distinct domains, where models trained under one set of channel or hardware conditions may fail under others. DG techniques such as meta-learning, adversarial domain alignment, and distributionally robust optimization provide mechanisms to enhance resilience by simulating or aligning across such conditions during training [41]. Viewing physical layer variability as domain shifts not only strengthens semantic robustness but also broadens the relevance of DG for physical communication research, highlighting the importance of cross layer designs that integrate PHY level phenomena into semantic communication frameworks.

Within 6G networks, where semantic communication will span highly dynamic and heterogeneous contexts, DG is not optional, it is foundational. By preparing systems to handle domain shifts systematically, DG ensures robustness, scalability, and trustworthiness across diverse operational landscapes.

1.3.2. Strategic impacts of domain generalization on semantic communication

DG holds transformative potential for advancing AI-enabled semantic communication. As shown in Table 2, domain shifts range from linguistic and cultural diversity to network and technological variability. DG enhances the robustness, scalability, and inclusivity of these systems, ensuring performance across unseen conditions.

i. Robustness to Environmental Variability: DG equips semantic models to remain reliable across cultural, linguistic, and situational variability, enabling accurate meaning exchange in real world applications such as international negotiation, virtual assistants, and context aware services [33,42,43].

ii. Efficiency and Scalability: Conventional retraining across every new domain is impractical. DG supports unified and reusable models, reducing computational cost and accelerating deployment in emerging use cases, thereby promoting sustainable scaling of semantic communication systems [44–46].

iii. Enhanced Generalization Capabilities: Semantic communication seeks to transmit meaning across diverse contexts. DG guides models to focus on invariant semantic patterns that persist despite shifts, supporting consistent interpretation across languages, modalities, and cultures [38,47].

iv. Broader Applicability and Accessibility: DG extends semantic communication to underserved contexts, including cross lingual collaboration, accessibility for differently abled users, and deployment in resource constrained environments [48,49]. This inclusivity ensures equitable participation in future communication ecosystems.

Therefore, DG transcends algorithmic performance, representing a paradigm shift toward communication systems that are adaptive, inclusive, and globally deployable. As 6G and beyond evolve, DG provides the foundation for resilient, trustworthy, and human centric semantic communication that can operate effectively across domains and contexts.

2. Architecture and performance metrics for domain generalization in AI-enabled semantic communication

AI-enabled semantic communication marks a paradigm shift from conventional data transmission by focusing on conveying intended meaning rather than transmitting raw symbols. To support real world applications, such systems must be resilient to domain shifts variations in input modalities, user contexts, linguistic styles, or environmental conditions. Therefore, a general architecture that incorporates domain generalization (DG) mechanisms is essential for achieving robust

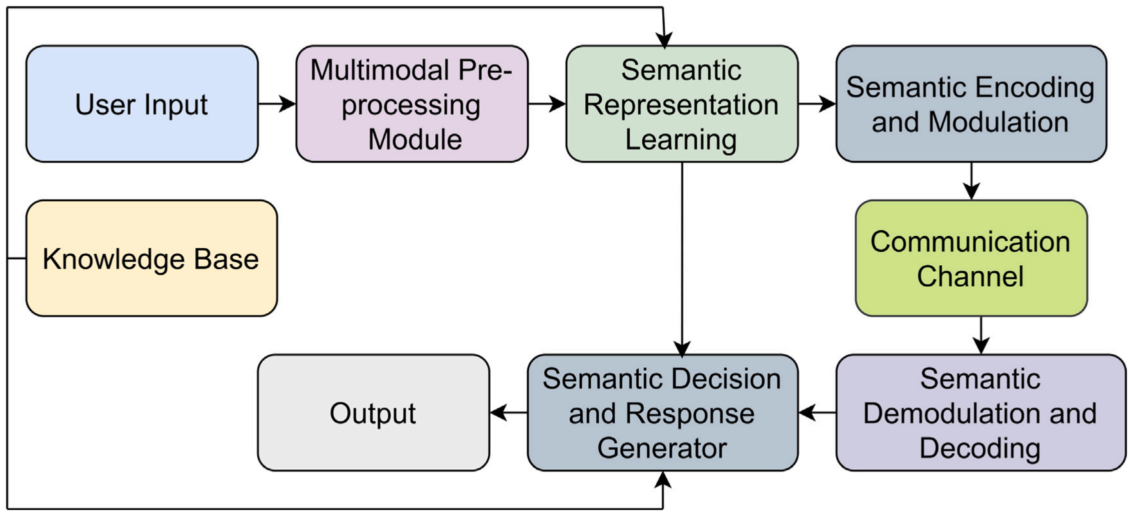

Fig. 6. General architecture of AI-enabled semantic communication highlighting key modules that facilitate domain generalization through multimodal understanding, adaptive semantic encoding, and context aware decision generation.

Table 3
Functional roles and domain generalization contributions of key modules in AI-enabled semantic communication architecture.

Architecture module	Function in semantic communication	Contribution to domain generalization (DG)
User Input (Text/Speech/Image)	Accepts raw input from users or sensors (e.g., text queries, voice, camera feeds). Forms the multimodal basis for semantic interpretation.	Not directly part of DG, but motivates the need for generalization due to inherent input diversity across domains, languages, or modalities.
Multimodal Pre-processing Module	Transforms raw input into structured embeddings using modality specific techniques (e.g., tokenizers, CNNs, spectrograms).	**Supports DG**: Normalizes input structure across domains, helping the downstream representation learner cope with input heterogeneity.
Semantic Representation Learning	Learns domain invariant semantic embeddings $\mathbf{z} \in \mathcal{Z}$ using neural encoders (e.g., transformers, contrastive models).	**Core DG Enabler**: Employs strategies like adversarial training, meta-learning, or DRO to produce embeddings that generalize across domains.
Knowledge Base (KB)	Provides domain independent and contextual world knowledge for grounding semantic decisions and disambiguation.	**DG Supportive**: Anchors semantic reasoning in shared knowledge; enables generalization by reducing dependence on domain specific features.
Semantic Encoding and Modulation	Compresses and prepares semantic vectors for transmission using joint source channel encoding (e.g., variational encoders, autoencoders).	DG indirect: Ensures robustness to channel level variation; less focused on semantic domain adaptation but critical for transmission integrity.
Communication Channel (AWGN/Fading)	Simulates environmental noise, fading, or interference; challenges the system to maintain semantic fidelity during real world transmission.	Not a DG module per se, but **exposes the system to domain shifts at the channel level**, indirectly enhancing robustness of downstream decoding.
Semantic Demodulation and Decoding	Recovers semantic embeddings $\hat{\mathbf{z}}$ from noisy signals using domain robust decoders (e.g., GRL-based or self-supervised models).	**DG-Linked**: Critical for reconstructing semantics across different channel conditions or encoder variants; can leverage self-supervised pretraining.
Semantic Decision and Response Generator	Uses semantic vectors and contextual inputs to generate task specific responses (e.g., translations, actions, summaries).	**DG-Linked**: Supports generalization to unseen tasks, styles, or domains using meta-learning or adaptive response models.
AI-Enhanced Output	Generates the final system output in a format tailored to users or applications (e.g., speech, structured message, visualization).	Indirect: Reflects the success of upstream DG modules; ensures semantic consistency across usage contexts.
Receiver Endpoint	The endpoint that consumes the output (human user or automated system). Closes the communication loop.	Not applicable to DG directly; receives benefits from DG-driven processing upstream.

and adaptable semantic communication in dynamic wireless environments [21]. The Fig. 6 illustrates a conceptual architecture that outlines the key building blocks of AI-enabled semantic communication systems enhanced with domain generalization. This framework is not a concrete implementation but rather a generic blueprint intended to guide future research and system design [50–52]. Table 3 provides a detailed mapping of the functional roles of each architectural component in the AI-enabled semantic communication system shown in Fig. 6. It highlights how each module contributes to domain generalization within the system.

Beginning with the User Input and Multimodal Preprocessing Module, the architecture ensures that data from heterogeneous sources can be consistently transformed into representations suitable for semantic analysis. The Semantic Representation Learning block plays a pivotal role in enabling DG by extracting domain invariant features that remain robust across various contexts, languages, and modalities. This capability is further reinforced by the Knowledge Base, which provides contextual grounding and supports generalization by encoding shared knowledge applicable across domains. The Semantic Encoding and Modulation and its counterpart, Semantic Demodulation and Decoding, maintain the integrity of semantic content under different channel conditions, thus contributing to generalization in noisy or degraded transmission environments. Meanwhile, the Semantic Decision and Response Generator employs adaptive reasoning mechanisms to uphold semantic consistency and task performance even in unfamiliar communication scenarios. Finally, the AI-Enhanced Output and Receiver Endpoint ensure that the system's response remains coherent, informative, and contextually relevant across diverse user profiles and interaction settings [33,53,54].

The Table 3 complements the system diagram and clarifies how each module contributes to building domain generalizable semantic communication systems for next generation wireless networks. The following section elaborates on the operation of each module, supported by mathematical formulations that explain their roles in enabling domain generalization in AI-enabled semantic communication.

- **Multimodal User Input:** Accepts user generated input data in various formats, including text, audio, images, and sensor readings. These inputs serve as the foundation for semantic inference and often vary in structure and context across different domains.
- **Multimodal Pre-processing Module:** Transforms raw input data into structured feature vectors using modality specific deep learning models. For example, BERT may be used for text embeddings, ResNet or Vision Transformers for images, and spectrogram based CNNs for speech. This step ensures uniform representations suitable for semantic interpretation [55].
- **Semantic Representation Learning:** This is the core component of the AI-enabled semantic communication architecture, responsible for generating robust and domain invariant semantic embeddings from heterogeneous input features. Given input data $\mathbf{x} \in \mathcal{X}$, which may consist of text, speech, image, or multimodal signals, the encoder function $f_\theta : \mathcal{X} \to \mathcal{Z}$ learns to produce a semantic embedding $\mathbf{z} = f_\theta(\mathbf{x}) \in \mathcal{Z}$ that captures the underlying meaning while abstracting away domain specific information. This abstraction is crucial for ensuring semantic coherence when the system operates across different domains characterized by variations in modality, context, or communication conditions [56, 57].

  The training objective for this module incorporates a dual component loss function designed to enhance both semantic accuracy and generalization. Formally, the objective is defined as:

$$\mathcal{L}_{\text{DG}}(\theta) = \sum_{i=1}^{N} \mathbb{E}_{(\mathbf{x},y)\sim D_i} \left[\ell(f_\theta(\mathbf{x}), y)\right] + \lambda \mathcal{R}(\theta) \tag{4}$$

  where D_i represents the data distribution from the ith domain (e.g., different speakers, languages, or environments), $\ell(\cdot)$ denotes a task specific loss function such as cross entropy or BLEU score [58]. Here, ℓ is computed after mapping $f(x)$ to a task specific prediction (e.g., via a classifier or decoder), and then compared with the label y, and $\mathcal{R}(\theta)$ is a domain generalization regularizer. This regularizer can take various forms depending on the adopted technique: adversarial losses (e.g., DANN) aim to confuse a domain discriminator, thereby enforcing domain invariance; statistical approaches (e.g., Maximum Mean Discrepancy or CORAL) align feature distributions across domains; and meta-learning strategies (e.g., MLDG, MetaReg) simulate domain shifts during training to enhance adaptability. The scalar λ is a hyperparameter that controls the trade off between semantic fidelity and domain robustness [59].

  This formulation ensures that the encoder not only learns to predict semantically meaningful outputs across all training domains but is also explicitly regularized to suppress domain specific variance. In practice, this allows the system to generalize to previously unseen domains with minimal or no fine tuning, which is critical in dynamic and diverse real world communication environments. By learning domain agnostic embeddings, the Semantic Representation Learning module thus serves as the foundation for robust downstream modules such as semantic encoding, reasoning, and response generation in AI-enabled semantic communication systems.
- **Knowledge Base:** Acts as a static or dynamic source of contextual and domain specific semantic information. It supports both training and inference by supplying background knowledge necessary for accurate disambiguation and semantic interpretation.
- **Semantic Encoding and Modulation:** Transforms the semantic vector $\mathbf{z}$ into a transmittable signal $\mathbf{s} \in \mathcal{S}$ using a joint source channel encoder g_ϕ:

$$\mathbf{s} = g_\phi(\mathbf{z}). \tag{5}$$

  It is worth noting that semantic communication systems can be realized in two main ways i.e. through separated source channel coding (SSCC), where semantic representation and channel coding are designed and optimized independently, or via joint source channel coding (JSCC), which integrates semantic extraction and communication into a unified end to end framework. SSCC supports modularity and allows domain generalization techniques to be applied separately to semantic or channel components, whereas JSCC especially learning based variants has demonstrated superior robustness and scalability under challenging and dynamic channel conditions [60–62]. These architectural differences could lead to significantly distinct interactions with DG strategies as the DG may act on individual modules in SSCC, yet must jointly account for both semantic and channel shifts in JSCC designs.
- **Communication Channel:** Models environmental noise during transmission. The transmitted signal $\mathbf{s}$ is corrupted by Gaussian noise $\mathbf{n} \sim \mathcal{N}(0, \sigma^2\mathbf{I})$:

$$\mathbf{r} = \mathbf{s} + \mathbf{n}. \tag{6}$$

- **Semantic Demodulation and Decoding:** Recovers the original semantic embedding $\hat{\mathbf{z}}$ from the noisy received signal $\mathbf{r}$ using the decoder h_ψ:

$$\hat{\mathbf{z}} = h_\psi(\mathbf{r}). \tag{7}$$

  The quality of reconstruction is evaluated using:

$$\mathcal{L}_{\text{rec}} = \mathbb{E}_{\mathbf{z}} \left[\|\hat{\mathbf{z}} - \mathbf{z}\|_2^2\right]. \tag{8}$$

- **Semantic Decision and Response Generator:** This module plays a pivotal role in interpreting the domain invariant semantic representation $\hat{\mathbf{z}}$ and generating the final system output $\hat{y}$ in an adaptive, context aware manner. This module integrates decoded semantic features from prior stages with auxiliary context retrieved from the *Knowledge Base* K, enabling the model to reason semantically and make informed decisions across diverse domains and scenarios [63]. The semantic decision function is defined as:

$$\hat{y} = d_\omega(\hat{\mathbf{z}}, K), \tag{9}$$

  where $d_\omega(\cdot)$ denotes the decision function or classifier parameterized by ω, responsible for interpreting semantic inputs and generating outputs relevant to the given task (e.g., classification, translation, summarization, or response generation). The incorporation of K allows the model to contextualize $\hat{\mathbf{z}}$, enriching decision making with domain specific facts, rules, or background knowledge essential for semantic consistency [64].

  The optimization objective for this module focuses on minimizing the expected semantic prediction loss between the generated output $\hat{y}$ and the target label y:

$$\mathcal{L}_{\text{decision}} = \mathbb{E}_{(\hat{\mathbf{z}},y)} \left[\ell(d_\omega(\hat{\mathbf{z}}, K), y)\right], \tag{10}$$

where $\ell(\cdot)$ is a task specific loss function, such as cross entropy for classification or mean squared error for regression based outputs. This objective ensures that the model not only learns accurate mappings from semantic embeddings to outputs but also generalizes well across different task domains, even when presented with previously unseen input patterns [64].
From a domain generalization (DG) standpoint, this module is critical for ensuring output level consistency and robustness. Since it operates on representations that may originate from a variety of domains (e.g., different users, languages, or modalities), the decision function must be resilient to contextual variation. This can be further supported through techniques such as domain mixed batch training, attention based integration of K, and architectural strategies that promote interpretability and transferability of semantic reasoning. By learning how to generate coherent, context sensitive responses across domains, the Semantic Decision and Response Generator becomes central to fulfilling the mission of adaptable AI-enabled semantic communication systems [65].
- **Output:** The final output $\hat{y}$ is rendered in natural language, visual, or symbolic form and delivered to the user or system endpoint.

The entire architecture is trained end to end using a composite objective:

$$\min_{\theta,\phi,\psi,\omega} \quad \mathcal{L}_{\text{DG}} + \alpha\mathcal{L}_{\text{rec}} + \beta\mathcal{L}_{\text{decision}}, \tag{11}$$

where α, β are hyperparameters balancing the trade offs between generalization, semantic preservation, and task accuracy.

This general framework provides a technically sound foundation for designing domain agnostic semantic communication systems. It offers guidance for integrating domain generalization objectives across both the semantic and transmission layers, particularly relevant for robust AI-enabled systems under the stringent requirements of future 6G and beyond wireless networks.

2.1. Performance metrics for domain generalization in semantic communication

In AI-enabled semantic communication systems with domain generalization, rigorous performance evaluation is essential for benchmarking models, guiding design decisions, and validating transmission fidelity, semantic preservation, and robustness under distribution shifts. Table 4 provides an overview of the principal metrics employed in this context, covering semantic fidelity (e.g., BLEU, ROUGE, semantic similarity), generative quality (e.g., perplexity), information retrieval effectiveness (e.g., precision, recall, F1-score, mAP), and communication robustness (e.g., accuracy, PSNR). These metrics collectively establish a multidimensional foundation for evaluating semantic communication models. In the subsequent discussion, each metric is examined in detail, with emphasis on its applicability, limitations, and extensions under domain generalization.

Let $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N}$ denote a dataset of N samples, where x_i represents the input and y_i the corresponding ground-truth label or target output. The model prediction is expressed as $\hat{y}_i = f(x_i;\theta)$, with θ denoting the learnable parameters of the system. Within this framework, several performance metrics are employed to assess semantic communication models under both in-domain and out-of-domain (OOD) conditions, thereby highlighting their ability to generalize across distributional shifts.

- **Precision** ($\mathcal{P}$) is a key metric that quantifies the relevance of retrieved or predicted information. In semantic communication with domain generalization, precision evaluates how accurately the system preserves semantic fidelity when exposed to unseen domains. A DG-robust model should maintain high precision not only in the source domain but also under out-of-domain (OOD) conditions, minimizing false positives even when channel characteristics, languages, or modalities differ from training. Formally, precision is defined as:

$$\mathcal{P} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{12}$$

where TP denotes the number of true positives (correct positive predictions) and FP represents the number of false positives (incorrect positive predictions). In the DG setting, reporting precision across both ID and OOD domains enables quantification of semantic reliability under distributional shifts [66].

- **Recall** ($\mathcal{R}$) complements precision by quantifying the system's ability to capture all relevant semantic information. In DG-enabled semantic communication, recall indicates how effectively a model retrieves or preserves critical semantic content when tested on domains that differ from training (e.g., unseen channels, dialects, or modalities). High recall ensures comprehensive semantic coverage and robustness against information loss under OOD conditions. Formally, recall is defined as:

$$\mathcal{R} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{13}$$

where TP denotes the number of true positives (correct positive predictions) and FN represents the number of false negatives (missed relevant instances). In practice, recall under DG evaluates a system's resilience against under-representation of domain-specific features and its ability to maintain semantic completeness [67].

- **F1-Score** (F_1) is a composite metric that harmonizes precision and recall into a single value. In the context of domain generalization for semantic communication, the F1-score is particularly valuable because it balances the trade-off between accuracy (high precision) and completeness (high recall) under both in-domain and out-of-domain (OOD) conditions. A DG-robust system should maintain a stable F1-score across distributional shifts, reflecting its ability to deliver semantically accurate outputs while avoiding information loss.
Mathematically, the F1-score is defined as the harmonic mean of precision and recall:

$$F_1 = 2 \cdot \frac{\mathcal{P} \cdot \mathcal{R}}{\mathcal{P} + \mathcal{R}}, \tag{14}$$

where $\mathcal{P}$ and $\mathcal{R}$ denote precision and recall, respectively. In DG evaluation, reporting the F1-score across both ID and OOD domains highlights the model's resilience in balancing semantic reliability with comprehensive coverage, even under noisy or imbalanced conditions [68].

- **Accuracy** ($\mathcal{A}$) is a fundamental metric that measures the proportion of correctly classified instances among all evaluated samples. In semantic communication, accuracy is often applied to tasks such as sentiment analysis, topic classification, or intent recognition, where categorical predictions are required. Within the domain generalization setting, accuracy provides insight into how consistently a model maintains correct predictions when evaluated across both in-domain and out-of-domain (OOD) distributions. A DG-robust model should sustain high accuracy even under varying channel conditions, linguistic diversity, or unseen modalities, ensuring reliable semantic preservation.
Formally, accuracy is defined as:

$$\mathcal{A} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}, \tag{15}$$

where TP and TN represent the number of true positives and true negatives, while FP and FN denote false positives and false negatives, respectively. In DG evaluation, accuracy must be reported not only as an aggregate value but also in terms of performance gaps between ID and OOD domains, reflecting the system's robustness and generalization capacity [69].

Table 4
Summary of performance metrics for domain generalization in AI-enabled semantic communication with example datasets.

Metric	Best-suited modality	Common datasets/Benchmarks	Mathematical expression	DG-aware description
Precision	Text, Multimodal	CIFAR-10, SVHN, Europarl	$\mathcal{P} = \frac{TP}{TP+FP}$	Assesses prediction relevance across domains.
Recall	Text, Multimodal	CIFAR-10, SVHN, Europarl	$\mathcal{R} = \frac{TP}{TP+FN}$	Evaluates completeness of retrieval under OOD.
F1-Score	Text, Multimodal	CIFAR-10, SVHN, Multimodal fusion sets	$F_1 = 2 \cdot \frac{\mathcal{P} \cdot \mathcal{R}}{\mathcal{P}+\mathcal{R}}$	Balances precision and recall for DG robustness.
Accuracy	Text, Image, Multimodal	CIFAR-10, Mini-ImageNet, NYU Depth V2	$\mathcal{A} = \frac{TP+TN}{TP+FP+TN+FN}$	Classification reliability across ID/OOD domains.
Mean Average Precision (mAP)	Text, Multimodal (retrieval)	COCO-Stuff, ImageNet, Multimodal retrieval datasets	$\mathrm{mAP} = \frac{1}{Q}\sum_{q=1}^{Q}\int_0^1 \mathcal{P}_q(r)\,dr$	Assesses ranking performance and retrieval stability.
BLEU Score	Text	Europarl, WMT, OpenSubtitles	$\mathrm{BLEU} = \mathrm{BP} \cdot \exp\left(\sum_{n=1}^{N} w_n \log p_n\right)$	Evaluates semantic fidelity via n-gram overlap.
ROUGE Score	Text	CNN/DailyMail, XSum, Europarl	$\mathrm{ROUGE}_n = \frac{\sum_{\mathrm{ref}} \mathrm{Count}_{\mathrm{match}}(n\text{-gram})}{\sum_{\mathrm{ref}} \mathrm{Count}(n\text{-gram})}$	Assesses summarization/dialogue fidelity under DG.
Perplexity	Text	Penn Treebank, Europarl, Wikitext	$\mathcal{PPL} = \exp\left(-\frac{1}{N}\sum_{i=1}^{N} \log P(y_i \mid y_{<i})\right)$	Measures fluency and stability under OOD inputs.
Word Embedding Evaluation	Text, Multimodal	WordSim353, GLUE, COCO captions	$\cos(\theta) = \frac{\mathbf{w}_1 \cdot \mathbf{w}_2}{\|\mathbf{w}_1\|\,\|\mathbf{w}_2\|}$	Tests whether embeddings preserve semantic relations across domains.
Semantic Similarity	Text, Multimodal	STS Benchmark, Europarl, COCO captions	$\mathrm{sim}(\mathbf{u}, \mathbf{v}) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}$	Quantifies meaning preservation across ID/OOD.
PSNR	Image, Audio	COCO-Stuff, DIV2K, ImageNet, VCTK (audio)	$\mathrm{PSNR} = 10 \cdot \log_{10}\left(\frac{\mathrm{MAX}^2}{\mathrm{MSE}}\right)$	Evaluates reconstruction quality under unseen channel noise.

- **Mean Average Precision (mAP)** is a widely used metric in information retrieval that evaluates performance across multiple recall levels and queries. It provides a comprehensive measure of how consistently a system maintains precision as recall increases. In semantic communication, mAP is particularly relevant for tasks such as semantic retrieval, multimodal search, or knowledge-based message reconstruction, where relevance must be preserved across diverse inputs.
  In the context of domain generalization, mAP becomes an important tool for assessing robustness across unseen domains. A DG-robust system should achieve stable mAP values not only on the source domain but also under out-of-domain (OOD) conditions, ensuring that semantic relevance is consistently maintained despite distributional shifts in channel conditions, modalities, or linguistic domains.
  Formally, mAP is defined as:
$$\mathrm{mAP} = \frac{1}{Q}\sum_{q=1}^{Q}\int_0^1 \mathcal{P}_q(r)\,dr, \tag{16}$$
  where Q denotes the total number of queries and $\mathcal{P}_q(r)$ represents the precision as a function of recall r for query q. Reporting mAP across both ID and OOD domains provides a robust indicator of generalization capacity, reflecting the system's ability to maintain relevance under distributional variations [70].

- **BLEU Score** is a standard metric in natural language processing, originally developed for machine translation, and widely applied in semantic communication tasks that involve text generation, such as conversational agents, summarization, or semantic message reconstruction. BLEU evaluates the quality of generated text by measuring the n-gram overlap between model outputs and reference sequences. Higher BLEU scores indicate stronger alignment with reference semantics, ensuring that generated responses are both linguistically and semantically coherent [71].
  In the context of domain generalization, BLEU plays an important role in quantifying semantic preservation under unseen linguistic or channel conditions. A DG-robust semantic communication system should sustain consistent BLEU scores across both in-domain and out-of-domain (OOD) test sets, reflecting its ability to maintain semantic fidelity even when facing shifts in vocabulary distributions, dialects, or communication environments.
  Formally, BLEU is defined as:
$$\mathrm{BLEU} = \mathrm{BP} \cdot \exp\left(\sum_{n=1}^{N} w_n \log p_n\right), \tag{17}$$
  where p_n represents the modified precision of n-grams, w_n is the weight assigned to n-grams of size n, and BP is the brevity penalty applied to discourage excessively short outputs. Reporting BLEU scores across ID and OOD domains allows researchers to evaluate not only generation quality but also robustness to domain shifts in semantic communication systems.

- **ROUGE Score** is a widely adopted metric for evaluating the quality of generated text summaries and responses, making it highly relevant for semantic communication systems such as chatbots, dialogue agents, and content summarization tools. ROUGE measures the degree of overlap between generated and reference summaries in terms of n-grams, longest common subsequences, or skip-grams. Higher ROUGE values indicate that the generated output is more accurate, contextually relevant, and semantically aligned with the reference content [68].
  Within the domain generalization setting, ROUGE provides a means to assess semantic preservation when models are applied to unseen domains. For example, in cross-lingual summarization or communication over noisy/unseen channels, a DG-robust system should sustain consistent ROUGE scores, demonstrating its ability to generate contextually faithful and semantically coherent outputs despite distributional shifts.

Formally, ROUGE is expressed as:

$$\mathrm{ROUGE}_n = \frac{\sum_{\mathrm{ref}} \mathrm{Count}_{\mathrm{match}}(\text{n-gram})}{\sum_{\mathrm{ref}} \mathrm{Count}(\text{n-gram})}, \tag{18}$$

where $\mathrm{Count}_{\mathrm{match}}$ denotes the number of overlapping n-grams between the generated output and the reference. Reporting ROUGE across both in-domain (ID) and out-of-domain (OOD) scenarios enables researchers to capture the system's robustness in generating semantically faithful summaries under varying conditions.

- **Perplexity** ($\mathcal{PPL}$) is a widely used metric for evaluating the quality and fluency of language models. It quantifies how well a model predicts the next token in a sequence, with lower perplexity values indicating better predictive capability and greater linguistic coherence. In semantic communication, perplexity is particularly important for tasks involving text generation or completion, such as chatbot responses, semantic message reconstruction, or auto-completion systems [72].
Within the domain generalization context, perplexity serves as an indicator of a model's robustness to unseen distributions. A DG-robust semantic communication system should maintain low perplexity not only on in-domain (ID) test sets but also on out-of-domain (OOD) scenarios, such as unseen languages, dialects, or noisy channel conditions. Stable perplexity across distributional shifts suggests that the system generates fluent and context-aware outputs despite exposure to novel domains.
Formally, perplexity is defined as:

$$\mathcal{PPL} = \exp\left(-\frac{1}{N}\sum_{i=1}^{N}\log P(y_i|y_{<i})\right), \tag{19}$$

where y_i denotes the ground-truth token at position i, $y_{<i}$ represents the preceding sequence of tokens, and $P(y_i|y_{<i})$ is the conditional probability assigned by the model. In DG evaluation, reporting perplexity across ID and OOD domains provides insights into the model's generative stability and semantic fluency under distributional shifts.

- **Word Embedding Evaluation** plays a crucial role in semantic communication, as embeddings such as Word2Vec, GloVe, or transformer-based contextual representations capture semantic relationships between words and phrases. Evaluation tasks typically involve computing cosine similarity between vectors, completing analogy tests (e.g., "king – man + woman = queen"), or assessing the clustering of word vectors in semantic space. High-quality embeddings enhance the semantic communication system's ability to represent meaning accurately, thereby improving encoding, transmission, and decoding processes [73].
In the domain generalization setting, embedding evaluation is extended to assess robustness under distributional shifts. A DG-robust embedding space should maintain consistent semantic relationships across both in-domain (ID) and out-of-domain (OOD) conditions, such as unseen languages, new channel characteristics, or domain-specific terminology. Stability of embeddings across these variations indicates strong generalization in semantic representation learning, which is fundamental for reliable semantic communication.
A common metric for evaluating embedding similarity is cosine similarity, expressed as:

$$\cos(\theta) = \frac{\mathbf{w}_1 \cdot \mathbf{w}_2}{\|\mathbf{w}_1\|\|\mathbf{w}_2\|}, \tag{20}$$

where $\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{R}^d$ represent word embeddings in a d-dimensional space. In DG evaluation, cosine similarity can be measured across ID and OOD embedding spaces to determine the extent to which semantic consistency is preserved under domain shifts.

- **Semantic Similarity** metrics, such as cosine similarity, Jaccard index, or Euclidean distance, are employed to quantify the degree of relatedness between text documents, sentences, or embedding representations. In semantic communication systems, these measures are critical for evaluating how closely the reconstructed output preserves the meaning of the original input. High semantic similarity ensures that transmitted messages align with user intent, thereby enhancing reliability and overall system performance [74].
Semantic similarity is often regarded as a practical measure of *semantic fidelity*, since both concepts aim to assess the preservation of intended meaning during transmission. While semantic fidelity is a broader notion encompassing the overall accuracy of semantic preservation (including task-level correctness), semantic similarity metrics provide a quantifiable approach to approximating this fidelity.
In the context of domain generalization, semantic similarity provides an essential means of assessing robustness to distributional shifts. A DG-robust model should maintain high similarity scores not only in-domain but also when exposed to out-of-domain (OOD) conditions such as new communication environments, languages, or modalities. This demonstrates the system's capacity to preserve semantic content even when the test distribution deviates from the training domain.
A common approach is to compute cosine similarity between the vector representations of input and reconstructed messages:

$$\mathrm{sim}(\mathbf{u}, \mathbf{v}) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}, \tag{21}$$

where $\mathbf{u}, \mathbf{v}$ denote the embedding representations of the input and the reconstructed messages, respectively. In DG evaluation, this metric is particularly useful for quantifying semantic fidelity across ID and OOD domains, providing a direct measure of meaning preservation under domain shifts.

- **Peak Signal-to-Noise Ratio (PSNR)** is a classical metric used to evaluate the fidelity of reconstructed signals, particularly in image or audio transmission tasks. PSNR measures the ratio between the maximum possible signal power and the power of the reconstruction error, thereby quantifying distortion at the signal level. Higher PSNR values indicate closer alignment between the transmitted and reconstructed signals, reflecting lower levels of degradation and better reconstruction quality. In semantic communication, PSNR complements semantic-oriented metrics by ensuring that low-level distortions do not undermine the interpretability or utility of the transmitted content.
Within the domain generalization setting, PSNR serves as an important robustness measure. A DG-robust system should maintain stable PSNR values not only in-domain but also when operating under out-of-domain (OOD) conditions, such as unseen channel environments (e.g., Rayleigh or Rician fading) or novel noise distributions. Consistency in PSNR across these scenarios demonstrates resilience in maintaining reconstruction quality despite distributional shifts.
Formally, PSNR is expressed as:

$$\mathrm{PSNR} = 10 \cdot \log_{10}\left(\frac{\mathrm{MAX}^2}{\mathrm{MSE}}\right), \tag{22}$$

where MAX denotes the maximum possible pixel (or signal) value and MSE is the mean squared error between the original and reconstructed signals. In DG evaluation, PSNR is often reported alongside semantic metrics such as BLEU or semantic similarity, providing a holistic view of both signal-level fidelity and semantic preservation under domain shifts.

In summary, the set of DG-aware performance metrics offers a multidimensional perspective for the evaluation of AI-enabled semantic

Table 5
Comparison of semantic communication surveys: Coverage of domain generalization aspects.

Ref.	DG techniques	DG taxonomy	Applications/Use cases	SemCom framework for DG	DG challenges	Paper scope
[75]	✗	✗	✓	✗	✗	Theory, Metrics, Challenges
[76]	✗	✗	✓	✗	✗	Task driven SemCom, AI for IoT
[77]	✗	✗	✗	✓	✗	Generative AI, Architecture
[20]	✗	✗	✓	✗	✗	6G Theory, Design Vision
[78]	✗	✗	✗	✗	✗	Historical Review, Principles
[79]	✗	✗	✓	✗	✗	Security, Network Stack, Protocols
[80]	✗	✗	✓	✓	✗	SemCom Architecture
[34]	✓	✗	✓	✗	✗	AI Techniques, Efficiency
[81]	✗	✗	✗	✗	✗	Metrics, Evaluation, Theory
[82]	✗	✗	✓	✗	✗	Bandwidth Efficiency, Applications
Our Survey	✓	✓	✓	✓	✓	Cross domain DG in SemCom

Note: This table shows that while previous surveys contribute to selective areas of semantic communication, they largely omit comprehensive coverage domain generalization in SemCom. Our survey addresses all key dimensions in a unified manner.

communication systems, capturing accuracy, completeness, semantic fidelity, generative quality, and resilience to distributional shifts. The selection of appropriate metrics should be carefully aligned with the specific objectives and modalities of the target application, whether text, image, or multimodal. Furthermore, the integration of these quantitative measures with explainability techniques enhances interpretability, transparency, and trust, thereby facilitating the development of semantic communication models that are not only high performing but also reliable and generalizable across heterogeneous environments.

3. Literature review on domain generalization in semantic communication

In this section, we review the evolving literature on AI-enabled semantic communication with a focused lens on *domain generalization*. Semantic communication is an emerging paradigm designed to enhance communication efficiency by enabling systems to understand and transmit meaning rather than raw data. As these systems are increasingly deployed in real world environments characterized by domain shifts such as variations in user context, language, or network conditions the importance of domain generalization becomes paramount [83].

Domain generalization addresses a core limitation of traditional AI models in semantic communication: the inability to maintain performance across unseen or dynamic domains without retraining. Despite the increasing interest in semantic communication, our investigation reveals a substantial gap there is currently no comprehensive survey that deeply explores the role of domain generalization in making semantic communication systems robust, context aware, and scalable. Our work aims to fill this gap by systematically analyzing the current state of domain generalization techniques as applied to semantic communication, providing a forward looking perspective on how these methods can shape the adaptability and reliability of future 6G communication systems.

While early works have laid foundational concepts for semantic communication, few have addressed the intersection of domain generalization and semantic understanding in communication systems. Moreover, existing surveys often treat domain generalization only as a peripheral concern or focus exclusively on general purpose vision and NLP tasks. As shown in Table 5, previous reviews have largely overlooked the unique challenges and opportunities that domain shifts present in AI-enabled semantic transmission. This oversight highlights the need for a focused, in depth survey that examines domain generalization as a central pillar in the development of scalable and intelligent semantic communication models.

By offering a detailed synthesis of this emerging intersection, our survey provides a roadmap for building semantic communication frameworks that can operate reliably across diverse, unseen, and evolving communication environments. In doing so, we contribute to both theoretical development and practical system design, guiding future research toward robust, adaptable, and real world ready semantic communication infrastructures.

3.1. Literature insights on the advancements of semantic communication

The literature on semantic communication has increasingly recognized the importance of robustness and adaptability, especially in dynamic wireless environments. In this context, domain generalization techniques have emerged as key enablers of generalizable semantic inference, helping models maintain stable performance under domain shifts. While research in this area remains nascent, a growing body of work has explored how semantic transceivers can be made more resilient to variability in linguistic, contextual, and network specific domains.

This subsection presents a curated set of recent contributions in semantic communication and highlights how these works incorporate or could benefit from domain generalization principles. Table 6 summarizes the methodologies, input types, optimization algorithms, performance metrics and limitations in selected studies, along with their relevance to DG techniques where applicable. These insights are further discussed to emphasize how domain generalization elevates the scalability, reliability, and applicability of AI-enabled semantic systems across real world communication scenarios.

The study presented in [23] introduces an innovative end to end image transmission system that leverages semantic communication and integrates AI technologies with 6G communication networks. This system, consisting of a transmitter (encoder) and a receiver (decoder), communicates over a physical channel. The transmitter performs a semantic segmentation task on the image to extract its semantic meaning, transmitting the segmentation map instead of the original image. At the receiving end, a pre-trained GAN network reconstructs a realistic image from the received segmentation map. Using the COCO Stuff dataset for training, both the transmitter and receiver share a common knowledge base. The research also examines the impact of physical channel distortions and quantization noise on the transmission of multimedia content via semantic communication. Experimental results highlight the system's efficiency, demonstrating a compression ratio of approximately 20 compared to traditional image transmission methods.

A DeepSC model is presented in [84], which is an intelligent end to end communication system, operates on two key levels: semantic and transmission. The semantic level focuses on processing information for encoding and decoding, while the transmission level ensures accurate exchange over the medium. DeepSC aims to maximize system capacity, emphasizing sentence meaning recovery over traditional bit or symbol error concerns. Utilizing transfer learning, it adapts to diverse communication environments, demonstrating superior performance in low signal to noise ratio scenarios and increased resilience to channel variation. The comparisons with traditional systems highlight DeepSC's robustness, particularly in low SNR regimes, across varied channel conditions. Introducing the novel metric "sentence similarity", the system precisely measures semantic communication performance, justifying DeepSC's efficacy in meaning recovery and semantic error reduction. An analysis of loss value, mutual information, BLEU score,

Table 6
State of the art AI models and generalization methods in semantic communication with DG-specific limitations.

Study	Input feature	Model	Performance metric	DG methodology	DG-specific limitations
[23]	Image	Transformer	PSNR	GAN	Training instability; adversarial vulnerability
[84]	Text	Transformer, DNN	BLEU, MI	Transfer Learning	Needs large pretraining; weak zero/few-shot OOD
[85]	Multimodal (Text, Image)	CNN (ResNet)	MAP	GAN	Modality gap persists; limited interpretability
[76]	Image	ResNet	Accuracy	Self-supervised Learning	Data hungry; sensitive to imbalance
[86]	Multimodal Images	ResNet	Accuracy	Self-supervised Learning	Scalability/compute overhead; adversarial robustness unverified
[87]	Images	CNN	PSNR, SSIM	Meta-Learning	High compute; sensitive to task design
[88]	Multimodal Images	AE, VAE	PSNR, Accuracy	Federated + MAML	Comm. overhead; privacy leakage risk
[89]	Images	Swin Transformer	PSNR	Transfer Learning	Drops in dynamic domains; depends on pretraining
[90]	Text	Transformer	BLEU, Similarity	GAN	Semantic drift under attack; limited OOD robustness
[43]	Text	DNN	WER, FDSD, Accuracy	Unified Feature Learning	Overfitting risk; lacks adversarial defense
[91]	Image	Swin Transformer + Diffusion	PSNR, LPIPS	Generative Refinement	High compute; data bias sensitivity
[92]	Image	VLM + Transformer	BLEU, SSQ	Continual Learning	Catastrophic forgetting; memory overhead
[93]	Image	Diffusion (DDPM)	PSNR, MS-SSIM	Noise Modeling + Denoising	Resource intensive; adversarial robustness unproven
[94]	Multimodal (Text, Image)	ResNet-50, CLIP, Diffusion	PSNR, MS-SSIM	Multimodal Fusion + Selection	Fusion overhead; cross-modal inconsistencies
[95]	Image	ResNet + Diffusion (DDPM/DDIM)	PSNR, SSIM, Semantic Fidelity	Semantic Priors	High complexity; priors may leak info

and SNR underscores DeepSC's performance, revealing the impact of training parameters on mutual information and BLEU score. While the sources lack explicit overall results, they provide valuable insights into DeepSC's enhanced performance and the metrics used for evaluation.

The paper [85], introduces a comprehensive system model for cross modal retrieval and semantic representation learning. It incorporates generative parts for image and text representation, classification parts for multi label classification, and an inter modal discriminative network to bridge the modality gap. Adversarial learning is employed to iteratively train generative and discriminative models, aiming to generate common semantic representations and distinguish differences between modalities. The joint optimization of the generative adversarial network and classification network enhances cross modal retrieval. The proposed model demonstrates superior performance on widely used datasets, achieving indistinguishable and close representations for data with similar semantics through adversarial learning. The inter modal discriminative network effectively contributes to overcoming the modality gap by distinguishing between image and text modalities.

Tilahun et al. propose a self supervised learning based framework for task oriented semantic communication under limited label availability, aimed at enhancing generalization across varying channel conditions and data distributions [76]. The system, termed SLSCom, integrates deep learning and the Information Bottleneck (IB) principle to extract semantically meaningful representations using both labeled and unlabeled data. A two stage training strategy is employed: self supervised pre-training with unlabeled samples via pretext tasks (e.g., contrastive InfoNCE loss and reconstruction), followed by supervised fine tuning on a small set of labeled examples. Experiments on CIFAR-10 and SVHN datasets demonstrate that SLSCom maintains high classification accuracy even with minimal supervision and under domain shifts such as SNR variations. By leveraging self supervision, the framework significantly improves semantic representation robustness and effectively addresses the domain generalization challenge in low label semantic communication systems.

Zhao et al. introduce a self-supervised, multi modal semantic communication framework that addresses domain generalization through task agnostic pre-training and label efficient learning [86]. The proposed system operates across multiple modalities (e.g., RGB and depth data) using separate encoders at edge devices and a central decoder for downstream classification tasks. A two stage training scheme is adopted: in the first stage, a novel self supervised learning method is used to capture both shared and unique semantic features across modalities via intra modal and cross modal contrastive objectives. In the second stage, supervised fine tuning adapts the system to specific tasks with minimal labeled data. This pre-training strategy significantly reduces training related communication overhead while maintaining robustness across different channel conditions and label scarcity. Experimental results on the NYU Depth V2 dataset demonstrate that the framework outperforms fully supervised and other self supervised baselines in both efficiency and accuracy. By enabling task agnostic semantic representation learning, this approach effectively enhances domain generalization capabilities in multi modal semantic communication.

Chen et al. propose a scalable semantic communication system tailored for industrial scenarios, where the dynamic addition of new transmission tasks imposes challenges for adaptability and training efficiency [87]. To address this, the authors introduce a joint source channel coding (JSCC) semantic communication framework integrated with meta-learning and a novel metric called Decoding Information Resolution (DIR) to assess scalability. The system employs a Convolutional Neural Network (CNN) for semantic encoding and decoding and is trained using a model agnostic meta-learning algorithm that operates in three stages: offline training, quasi-online adaptation, and online deployment. The framework is validated using the mini ImageNet dataset and tested under various SNR conditions. Results demonstrate that the proposed method significantly improves scalability and performance, particularly in low SNR environments and during the addition of new tasks. This work highlights the effectiveness of meta-learning in enhancing domain generalization and scalability in semantic communication systems.

Bian et al. present a federated semantic communication framework tailored for the Metaverse, aiming to balance communication efficiency, privacy, and adaptability in large scale virtual environments [88]. The framework integrates deep learning based semantic communication with federated learning to enable collaborative model training across distributed nodes without sharing raw data. Core components include semantic encoders/decoders (e.g., autoencoders and variational autoencoders) and semantic digital twins (SCDTs) for efficient semantic representation and reconstruction. The model is evaluated on diverse datasets (MNIST, KMNIST, CIFAR-10), showing improved performance in both PSNR and classification accuracy under noisy and compressed settings. To enhance robustness across non-IID data distributions, the system incorporates Model Agnostic Meta-Learning (MAML) into the federated training loop. This approach significantly boosts generalization across domains and tasks, making it a strong candidate for enabling scalable and adaptive semantic communication systems in the Metaverse.

Nguyen et al. address the challenge of semantic communication in a multi user setting with users having diverse computing capacities [89]. They propose a Swin Transformer based semantic communication system for image transmission, where a single encoder at the base station communicates with multiple decoders of varying complexity. To improve training efficiency and generalization for low computing users, they introduce a two stage training strategy: (1) pairing the encoder with a high capacity decoder for initial training and (2) reusing the encoder while training the low capacity decoder. Two techniques enhance generalization: transfer learning, where parameters are transferred between compatible layers of decoders, and knowledge distillation, where the high capacity decoder acts as a teacher guiding the low capacity decoder. Experiments on the DIV2K dataset demonstrate improvements in PSNR under varied SNR conditions, showing the system's effectiveness in adapting to resource constrained users. This approach highlights the role of cross user transfer and teacher student learning in promoting domain generalization in semantic communication.

Mao et al. propose a novel GAN based semantic communication framework (Ti-GSC) designed for text transmission without relying on Channel State Information (CSI), addressing a major limitation in existing SC systems over fading channels [90]. The proposed framework integrates two core modules: (1) an autoencoder based encoder decoder module (AEDM) built upon Transformers for semantic extraction and reconstruction, and (2) a GAN-based signal distortion suppression module (GSDSM), which learns to remove syntactic and semantic noise without prior CSI. The system is trained using a joint optimization strategy (JOT) involving a composite loss function that includes cross entropy, adversarial, syntactic, and semantic distortion terms. The GSDSM is implemented using a U-Net generator and a convolutional discriminator, enabling semantic alignment of distorted signals. Evaluated on the Europarl text dataset across AWGN, Rician, and Rayleigh channels, the Ti-GSC significantly improves BLEU and sentence similarity scores compared to baselines, even in the absence of CSI. The paper also performs generalization experiments on Nakagami-m fading channels, confirming the model's robustness in unseen channel conditions thereby demonstrating strong domain generalization capabilities.

Qin et al. propose a generalized semantic communication (GSC) framework that extends semantic processing to both the source and the wireless channel, aiming to optimize efficiency across diverse multimodal data types and channel environments [43]. The system supports semantic transmission of text, speech, image, and video, employing deep learning (DL) models for joint semantic channel coding. In addition to conventional source semantics, the framework introduces environment semantics, where semantic features (e.g., object layout and categories) extracted from images assist in channel estimation and precoding tasks. This reduces reliance on pilot signals and explicit channel state information (CSI). A case study on speech transmission (DeepSC-ST) demonstrates enhanced performance under Rayleigh channels using semantic representations rather than raw signals. For the channel side, an environment semantics aided communication (ESAC) module predicts beamforming vectors directly from semantic scene features. The system emphasizes generalization across source tasks and channel conditions, enabled by unified DL-based abstraction and feature selection strategies, making it a compelling approach for domain adaptive semantic communication.

Zhang et al. propose a next generation semantic communication system called GSC (Generative AI Semantic Communication), which integrates a Swin Transformer based semantic encoder with a Diffusion Model (DM) powered refinement module to enhance reconstruction fidelity, particularly under challenging channel conditions [91]. The system introduces a novel semantic successive refinement strategy, wherein an initial coarse semantic reconstruction is first generated by the base decoder. This is then passed to a pre-trained Diffusion Model, which acts as a powerful generative prior to iteratively refine and reconstruct the output with high semantic consistency and perceptual quality. This generative refinement enables the system to correct semantic distortions that may arise due to noise or compression during transmission. The authors further extend their design to a multi user asynchronous setting (MU-GSC), supporting scalable and parallel transmissions. Experiments conducted under AWGN and Rayleigh fading channels show that GSC significantly outperforms DeepJSCC and conventional LDPC based baselines in terms of PSNR and LPIPS, even at low SNRs. While the framework does not use traditional DG algorithms, the diffusion based generative enhancement enables strong generalization across varying channel conditions, making it an effective domain adaptive solution for future semantic communication networks.

Jiang et al. propose a novel Visual Language Model based Cross modal Semantic Communication (VLM-CSC) system to overcome key limitations of traditional image semantic communication (ISC) systems, such as low information density, catastrophic forgetting, and SNR variability [92]. Their framework introduces three core components: (1) a Cross modal Knowledge Base (CKB) using pretrained BLIP and Stable Diffusion (SD) models for transforming images into high density text at the transmitter and reconstructing them at the receiver, (2) a Memory Assisted Encoder/Decoder (MED) combining short term and long term memory to enable continual learning, and (3) a Noise Attention Module (NAM) that adaptively allocates importance to semantic and channel coding based on SNR feedback. The system uses transformer based encoders/decoders for text processing and integrates pretrained vision language models to boost semantic alignment. Evaluated across diverse datasets (e.g., CIFAR, CATSvsDOGS) and SNR levels, VLM-CSC demonstrates high BLEU scores and Semantic Service Quality (SSQ), while maintaining robustness and adaptability. The MED mechanism particularly enhances domain generalization by preserving past knowledge while learning new domain shifts, making VLM-CSC a powerful DG capable architecture for multimodal communication.

The study in [93] proposes a Joint Source Channel Noise Adding with Adaptive Denoising (JSCNA-AD) framework, developed on the foundation of the Denoising Diffusion Probabilistic Model (DDPM) for semantic communication. The proposed method incorporates channel noise directly into the forward diffusion process and applies attention guided adaptive denoising at the receiver, thereby enabling high fidelity reconstruction of semantically important regions while minimizing redundant computations. Experimental evaluations on ImageNet-256 and STL-10 datasets over AWGN and Rayleigh channels demonstrate performance improvements of up to 13.3 dB in PSNR, alongside superior MS-SSIM, reduced FLOPs, and lower inference time compared

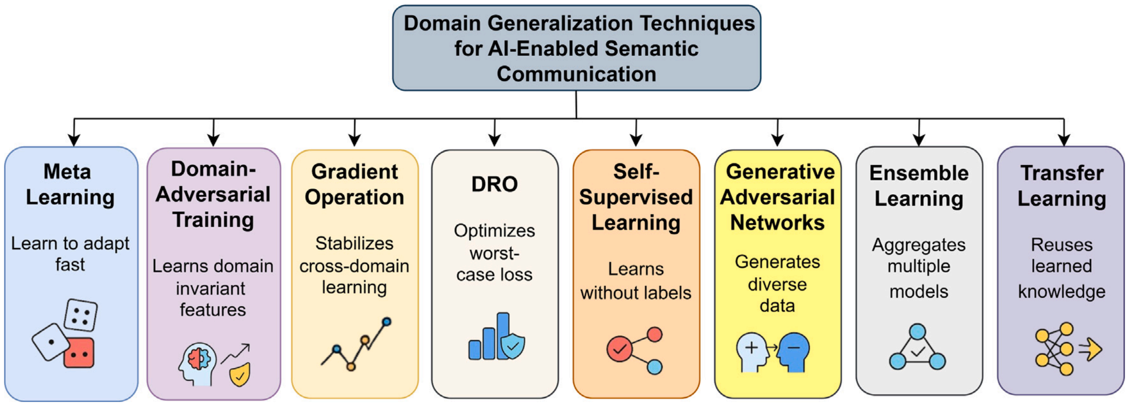

Fig. 7. State of the art domain generalization techniques for AI-Enabled Semantic Communication.

with JPEG2000+LDPC, DeepJSCC, and WITT baselines. These findings suggest that integrating channel aware noise modeling within DDPM based diffusion processes, coupled with adaptive denoising, can substantially enhance reconstruction fidelity, robustness under diverse channel conditions, and computational efficiency in semantic communication.

The study in [95] presents an end to end semantic communication system designed for efficient image transmission under bandwidth limited conditions. At the transmitter, a deep learning classifier (ResNet-18) predicts image category labels, and dictionary learning with sparse coding is employed to extract compact semantic features. At the receiver, a diffusion model based conditional generator synthesizes candidate images from the received category, while dictionary learning is again used to match and select the image most consistent with the transmitted features. To assess reconstruction quality, the authors introduce a Semantic Fidelity Index (SFI) that jointly incorporates mutual information and CNN feature similarity, providing a more reliable evaluation than conventional metrics such as PSNR and SSIM. Experimental results demonstrate that the proposed system achieves superior semantic fidelity and visual quality compared to JPEG2000 and generative baselines (GAN, VAE, StyleGAN2), while operating under reduced bandwidth constraints. Moreover, the diffusion model yields better convergence behavior and higher reconstruction quality than alternative generative approaches at lower complexity, indicating that integrating dictionary learning with diffusion based generative reconstruction and a semantic fidelity metric enables more efficient, semantically faithful, and robust image transmission than traditional codecs or alternative generative models.

The authors in [94] presented a Multimodal Semantic Communication (MMSemCom) framework that jointly exploits image and text features for supervised image generation. At the transmitter, semantic features are extracted using ResNet-50 for image representations and CLIP for text prompts, which are subsequently fused into a unified MultiSem representation via cross attention. At the receiver, a diffusion model (DM) generates candidate images, followed by a supervised selection mechanism either sequential or simultaneous that aligns the generated outputs with the transmitted multimodal semantics. Extensive experiments on CIFAR-100, STL-10, and ImageNet-256 under AWGN and Rayleigh channels demonstrate that MMSemCom consistently outperforms traditional codecs (BPG, WebP) and deep learning baselines (DeepJSCC, WITT, LaMoSC). The sequential variant achieves higher robustness in low SNR conditions, reaching up to 35.6 dB PSNR, whereas the simultaneous variant offers improved efficiency by achieving strong fidelity with lower overhead. Furthermore, the framework exhibits strong generalization to out of distribution datasets such as FGVC-Aircraft and Stanford Cars, underscoring its resilience to domain shifts. Collectively, these results indicate that combining multimodal semantic features with diffusion based generation and supervised selection not only improves reconstruction fidelity and robustness to channel noise but also enhances generalization to previously unseen domains.

In summary, Section 3 and Table 6 together show that domain generalization in semantic communication is advanced by complementary building blocks that map onto the SemCom pipeline: strong representation learners (transfer and self supervised) establish versatile feature bases across modalities and channel conditions; adversarial and diffusion based components repair residual distribution gaps and enhance semantic fidelity; meta learning enables rapid, label efficient adaptation when environments change; federated and continual variants preserve performance across clients and over time without centralizing data; and unified feature learning with semantic priors imposes structure that preserves meaning under perturbations. In practice, the most effective systems compose two or more families, for example self or transfer pretraining for general features, a lightweight DG head (meta learning or domain adversarial), and selective generative refinement, balancing robustness with deployability. Table 6 makes these trade offs explicit by aligning each study's modality, model, metrics, DG methodology, and limitations from the literature offering a clear blueprint for designing resilient SemCom systems under real world constraints.

The recent studies reviewed in this section highlight substantial progress in strengthening the robustness and adaptability of semantic communication through diverse domain generalization strategies, including self-supervised learning, meta-learning, transfer learning, and generative refinement techniques. Collectively, these approaches establish a solid foundation for advancing the field while pointing toward promising future directions. They underscore the importance of sustained innovation and interdisciplinary collaboration to realize scalable, efficient, and reliable semantic communication systems capable of meeting the evolving demands of next generation networks.

4. Domain generalization techniques for AI-enabled semantic communication

This section presents an analysis of current techniques that enhance the domain generalization capabilities of AI-enabled semantic communication systems. Our contribution addresses a key gap in the literature by focusing on how modern models can generalize effectively across varied and unseen domains such as changes in user context, modality, network conditions, and linguistic styles without requiring retraining.

We examine a range of diverse methods aimed at improving generalization performance. Fig. 7 highlights key domain generalization techniques, which we discuss below.

1. Meta-Learning:

Meta-learning, also known as "learning to learn", is a powerful approach aimed at enhancing domain generalization by enabling models to quickly adapt to new tasks with limited data. This is particularly beneficial in semantic communication, where models must efficiently understand and convey meaning across varying domains such as different languages, modalities, communication styles, and user contexts [96]. Meta-learning frameworks such as Model Agnostic Meta-Learning (MAML) [97], First Order MAML [98], Model Agnostic Meta-Learning for Domain Generalization (MLDG) [99], and MetaReg [100] have shown effectiveness in enabling fast adaptation and improved generalization capabilities.

These methods typically follow a two level learning scheme: a base learner is trained on individual tasks using conventional optimization techniques, while a meta learner optimizes the base learner's ability to perform well across a distribution of tasks or domains. In the context of semantic communication, this setup allows the model to acquire a strong initialization that can be fine tuned on unseen semantic tasks using just a few gradient steps. Tasks can include translation, summarization, or classification in different linguistic or domain settings, where the semantic structure of the input varies.

This generalization capability becomes especially critical when adapting semantic models to emerging environments, such as low resource languages or newly introduced modalities. For instance, in cross domain sentiment analysis on social media platforms, each platform (e.g., Twitter, Facebook, Instagram) features distinct writing styles, vocabularies, and context. Meta-learning algorithms like MLDG and MetaReg allow models to learn domain invariant features across these tasks and adapt rapidly without retraining from scratch. The models not only learn the core patterns of language understanding but also effectively handle style and dialect variation.

To evaluate the performance of meta-learning in semantic communication, metrics such as accuracy, BLEU scores for translation, ROUGE for summarization, and adaptation speed are commonly used [101]. These metrics assess both the quality of semantic preservation and the efficiency of adaptation across tasks. For example, a multilingual customer support system equipped with a meta learned semantic encoder could generalize across different customer service scenarios, quickly adapting to new query patterns or domain specific language with minimal labeled data. This makes meta-learning a promising candidate for building scalable, robust, and adaptive semantic communication systems in real world environments.

2. Domain Adversarial Training:

Domain adversarial training is a widely adopted strategy for enhancing domain generalization by learning feature representations that are invariant across different domains [102]. Within semantic communication systems, this technique becomes highly relevant when models are expected to operate consistently in varying linguistic, acoustic, or environmental conditions. By jointly optimizing semantic accuracy and domain invariance, domain adversarial training helps models maintain stable performance in unseen contexts without requiring fine tuning or retraining.

The general architecture involves three key components: a feature extractor that converts input data (e.g., text, speech, or images) into latent representations; a task specific classifier that performs the semantic task (e.g., intent detection, translation, or sentiment classification); and a domain discriminator that attempts to infer the domain label from the same latent features. A minimax optimization is used, where the model minimizes task loss and simultaneously maximizes the domain discriminator's error. This competitive interaction compels the feature extractor to learn domain agnostic representations [103,104]. A central mechanism enabling this learning is the gradient reversal layer (GRL) [105]. During training, GRL inverts the gradient signals from the domain discriminator, pushing the feature extractor to produce embeddings that confuse the discriminator while still enabling accurate semantic task performance. This leads to representations that generalize well across domains an essential property for scalable semantic communication systems.

Consider, for instance, a multilingual voice assistant system integrated into smart home environments, where users communicate through speech across various accents, languages, room acoustics, and microphone qualities. In such a system, domain adversarial training ensures the semantic understanding module remains robust to these variations. Whether a command is issued in a quiet kitchen by a native speaker or in a noisy living room with accented speech, the model must consistently understand the intent and respond appropriately [106]. By training the feature extractor to ignore acoustic and linguistic domain differences, the system can generalize well across user and environmental variations, maintaining reliable semantic performance without reconfiguration.

Therefore, this technique is particularly beneficial in scenarios where explicit domain labels may not be available during inference, yet generalization is still critical. It effectively decouples semantic learning from domain specific nuances, ensuring models are not overfitted to source conditions. As a result, domain adversarial training is a core component of domain generalization strategies in semantic communication. It enables real time, cross condition adaptability and supports the deployment of intelligent communication systems in heterogeneous environments with minimal data preparation or retraining overhead.

3. Gradient Operation:

Gradient based optimization methods are essential tools for enhancing generalization in machine learning, particularly within semantic communication systems that operate in dynamic or unfamiliar environments. These techniques aim to stabilize training and guide models toward learning domain invariant representations by regulating how gradients are propagated and updated across tasks or domains. Among the widely used approaches are gradient clipping, gradient normalization, and algorithmic frameworks such as Representation Self Challenging (RSC) [107] and Fish [108], which are designed specifically for domain generalization.

Gradient clipping involves constraining the gradient values during backpropagation to avoid instability caused by excessively large updates. This is especially important in language centric tasks like semantic translation or summarization, where gradient explosion may hinder training convergence. In contrast, gradient normalization ensures that gradients across layers maintain consistent scales, addressing vanishing gradient problems and promoting uniform learning across network components [109]. These operations help maintain steady optimization dynamics, improving the model's robustness under domain shift conditions.

RSC introduces a self challenging mechanism where the model suppresses its most confident features during training, encouraging it to rely on more diverse and stable patterns across domains [107]. This forces the model to learn alternative semantic cues that generalize better to unseen conditions. Fish, on the other hand, adopts a gradient matching approach: it aligns gradients computed across multiple source domains during meta-training, promoting learning trajectories that are more consistent and transferable [108]. Together, these techniques contribute to building representations that are not only task relevant but also domain agnostic.

In semantic communication, such capabilities are particularly critical when transmitting or interpreting data across highly variable channels or user environments. Consider, for example, an industrial IoT scenario involving real time sensor data collection and communication across different factory floors, where environmental conditions (e.g., electromagnetic interference, machine noise, sensor brands) differ widely [110]. A semantic communication system trained on data from one location may perform poorly in another unless it can generalize across such conditions. Applying gradient based domain generalization techniques enables the system to extract semantically relevant information from sensor readings while ignoring environment specific noise. This leads to more consistent communication quality and decision making across industrial settings without retraining [110].

Overall, gradient operation techniques such as RSC and Fish serve as powerful strategies to equip semantic communication models with the ability to generalize across heterogeneous domains. By refining optimization trajectories and encouraging reliance on stable, transferable semantic patterns, these approaches significantly strengthen the reliability and scalability of semantic communication systems in real world deployments.

4. Distributionally Robust Optimization:

Distributionally Robust Optimization (DRO) is a powerful strategy in domain generalization aimed at training models to maintain performance across a range of possible data distributions. In semantic communication, where models must handle variations in context, language, channel conditions, and message structure, DRO offers a principled framework to enhance robustness and reliability [111]. Unlike standard empirical risk minimization, which assumes that training and test distributions are similar, DRO explicitly considers uncertainty in the data distribution by defining an ambiguity set typically formulated using metrics such as Wasserstein distance, moment constraints, or f-divergences. The objective is to minimize the worst case expected loss over all distributions within this ambiguity set, making it highly suited for unpredictable and dynamic environments [112].

The optimization process in DRO typically involves a nested minimax structure: the inner maximization identifies the worst case distribution in the ambiguity set under which model performance deteriorates the most, while the outer minimization updates the model parameters to reduce this worst case loss. Though computationally more intensive than conventional training, this framework encourages the model to learn representations that are not overly tailored to any specific domain, thus achieving better generalization across unseen or shifting conditions [113]. This approach is particularly relevant for semantic communication, where input data can exhibit significant variability across regions, user demographics, signal distortions, and linguistic constructs.

Several effective algorithms fall under the DRO paradigm. VRex (Variance Risk Extrapolation) minimizes variance in model predictions across domains, promoting more uniform generalization [114]. Just Train Twice (JTT) reweights training data to emphasize underrepresented or difficult examples, enhancing the model's resilience to rare or complex semantic structures [115]. GroupDRO takes a group based approach by minimizing the worst case loss across predefined or dynamically inferred groups, ensuring robustness even in the presence of latent domain heterogeneity [116]. These techniques allow semantic communication models to avoid overfitting to dominant data patterns and instead learn representations that support consistent message interpretation across diverse conditions.

A compelling use case for DRO in semantic communication is found in emergency communication networks deployed in low resource environments during natural disasters. In such scenarios, linguistic content can vary widely across users, and channel conditions may degrade unpredictably due to infrastructural damage or interference. A semantic communication system trained with DRO could remain functional across fluctuating noise levels, dialectical shifts, or compressed signal formats by focusing on the most challenging examples during training. GroupDRO, for instance, would ensure consistent performance across multiple user populations or message types, such as emergency alerts, weather updates, and rescue coordination messages each potentially originating from different data domains.

5. Self-Supervised Learning:

Self-supervised learning (SSL) has emerged as a promising approach for domain generalization by leveraging the inherent structure of data to learn meaningful representations without relying on labeled supervision [117]. In semantic communication, where the goal is to extract and transmit semantic meaning rather than surface level data, SSL enables models to generalize across diverse linguistic, contextual, and environmental conditions by creating surrogate tasks that force the network to learn high level abstractions.

SSL frameworks design pretext tasks in which the model learns to predict or reconstruct certain components of the input from others. These tasks encourage the model to capture relationships, structures, and dependencies within the data. In natural language processing (NLP), common pretext tasks include masked language modeling (MLM), next sentence prediction (NSP), and sentence permutation [118]. For instance, BERT [119] uses MLM, where input tokens are randomly masked and the model learns to predict the missing tokens using the surrounding context. Similarly, NSP trains the model to determine whether two sentences appear consecutively in a document, thereby reinforcing the understanding of semantic coherence [120]. These strategies are particularly valuable for semantic communication, where understanding context, semantics, and intent is essential for encoding and decoding messages across variable communication channels.

Several SSL techniques have shown strong domain generalization capabilities in broader machine learning tasks. JiGen (Jigsaw Puzzle Generation) combines image permutation with classification to help models learn robust representations across domains [121]. In the context of semantic communication, JiGen like techniques can be adapted for textual or multimodal inputs by segmenting data into logical units and forcing the model to learn their relationships. Another technique, SelfReg [122,123], introduces regularization losses during SSL training that promote representation consistency and reduce overfitting to specific source domains. These constraints enhance generalizability and have been successfully applied in cross domain adaptation scenarios.

One of the key advantages of SSL in semantic communication lies in its applicability to low resource settings where labeled data are scarce or infeasible to obtain. For example, consider a multilingual educational platform used globally by students in different regions, languages, and cultural contexts. Each user may access the platform with different semantic expectations, vocabulary styles, and interface languages. Training a semantic communication model using SSL enables the system to extract semantically meaningful structures from unlabeled interaction data such as forum posts, feedback, or spoken queries without requiring exhaustive manual annotation. This empowers the platform to personalize content delivery, answer queries accurately, and maintain consistent understanding across diverse user populations. Moreover, SSL is inherently scalable and efficient. Once trained, the encoder can be fine tuned on downstream tasks or directly integrated into end to end semantic transceivers for real time communication. Its capacity to generalize from unlabeled data makes it especially valuable in dynamic environments where domain shifts occur frequently, such as mobile applications, multilingual AI assistants, or adaptive learning systems.

6. Generative Adversarial Networks (GANs):

Generative Adversarial Networks (GANs) represent a powerful class of generative models that have been increasingly adopted to enhance domain generalization in semantic communication systems. Their core utility lies in generating high quality, diverse synthetic data samples that enrich training sets, enabling models to generalize across domains with varying characteristics [124,125]. GANs are composed of two neural networks a generator and a discriminator that are trained adversarially. The generator learns to produce synthetic data that mimic real samples, while the discriminator learns to distinguish real data from generated data. Through a minimax optimization framework, both networks improve iteratively and the generator becomes more proficient at producing indistinguishable data, while the discriminator sharpens its ability to detect subtle differences [126,127].

The GAN training process alternates between updating the discriminator and the generator. Initially, the discriminator is trained using both real and generated samples to distinguish between them. The generator is then trained to produce data that mislead the discriminator, minimizing its ability to correctly classify synthetic instances. This adversarial interplay results in a generator capable of producing highly realistic data, including text, speech, or semantic signals, depending on the application. Such synthetic data are critical for improving the generalization of semantic communication models, especially in domains with scarce labeled data or hard to simulate conditions [128].

From a domain generalization perspective, GANs contribute in two significant ways. First, they augment training datasets with samples from diverse distributions, exposing models to a broader spectrum of scenarios and reducing overfitting to narrow domains. Second, they help simulate edge cases or underrepresented conditions such as noisy channels, uncommon linguistic patterns, or sensor anomalies thereby preparing semantic models for real world deployment in highly variable environments [129,130]. Moreover, the adversarial learning process helps the model distinguish meaningful semantic content from domain specific noise, enhancing the transferability of learned features across tasks and domains [131].

Consider a practical application such as a semantic broadcast system designed to disseminate real time alerts and news updates across multilingual and multicultural regions. These systems must support domain robust semantic understanding and transmission across varying dialects, content styles, and reception formats. Using GANs, synthetic message datasets can be generated to simulate different broadcast styles, languages, and communication settings. This allows the semantic model to learn from both dominant and rare communication scenarios, improving its ability to generalize without needing exhaustive data collection across all locales. For example, GANs can generate synthetic versions of emergency alerts in multiple languages or simulate reception distortions caused by regional infrastructure, training the system to respond accurately under a wide range of conditions.

7. Ensemble Learning:

Ensemble learning is a powerful approach in machine learning that combines predictions from multiple models to improve robustness, accuracy, and generalization performance. Within the context of semantic communication, ensemble methods are especially valuable when systems need to operate across diverse and complex linguistic domains, user intents, or noisy environments [132]. The core idea is that while individual models may have limitations or biases toward certain domains, an ensemble of complementary models can compensate for each other's weaknesses, resulting in stronger overall performance.

Common ensemble techniques include bagging, boosting, and stacking. Bagging (bootstrap aggregating) involves training multiple instances of the same model on different subsets of training data and aggregating their predictions to reduce variance and overfitting. This method is beneficial in tasks like speech recognition or text classification, where data heterogeneity can hinder a single model's performance [133,134]. Boosting builds a sequence of models where each subsequent model focuses on correcting the errors of the previous ones. This makes it effective for handling hard to classify semantic nuances, such as sentiment expressed with sarcasm or ambiguity. Stacking trains diverse base learners and then uses a meta-learner to synthesize their outputs, capturing a broader range of semantic features and improving decision accuracy in complex tasks like machine translation or dialogue generation [133,134].

In addition to these traditional methods, advanced ensemble based algorithms have been proposed to explicitly address domain generalization. Among them, D-SAM (Domain Specific Aggregation Module) [135], DAEL (Domain Adaptive Ensemble Learning) [136], and COPA (Collaborative Part based Aggregation) [137] are notable examples. D-SAM introduces domain specific attention modules within the ensemble, allowing each sub model to specialize in a particular domain while still contributing to a unified prediction [135]. DAEL promotes domain invariant feature learning while maintaining domain specific pathways, enhancing both adaptability and stability in variable environments [136]. COPA improves domain generalization through collaborative learning among domain specific classifiers, incorporating privacy preserving updates and normalization techniques to ensure scalability and resilience [137] .

A compelling use case for ensemble learning in semantic communication arises in multi agent robotic systems. In these environments, multiple autonomous agents (e.g., drones, delivery robots, or underwater vehicles) need to communicate semantically relevant information such as location updates, hazard signals, or task assignments under varying terrain, acoustic conditions, and operational settings. Each agent may operate in slightly different conditions and data domains. By employing ensemble learning, the communication framework can aggregate insights from multiple sub models trained on different operational contexts, ensuring that the system generalizes well across all agents. For instance, an ensemble of models specialized for different terrains (e.g., urban, indoor, rural) can improve the semantic encoding and decoding processes, resulting in more reliable coordination and fewer communication breakdowns.

8. Transfer Learning:

Transfer learning is a widely adopted technique in machine learning that enables models trained on one task or domain to be reused or adapted for a different, yet related, task. This method is particularly advantageous in semantic communication, where systems such as intelligent agents, translation tools, or multilingual chatbots must often operate in novel domains with minimal labeled data [138]. The key benefit of transfer learning lies in its ability to reduce training costs and data requirements while enhancing performance in previously unseen or underrepresented semantic environments. The typical workflow of transfer learning involves two phases i.e. pre-training and fine tuning. In the pre-training phase, a model is exposed to a large scale dataset from a broad, general purpose domain, allowing it to learn reusable representations such as grammatical structures, contextual embeddings, and high level semantics. In the fine tuning phase, the model is adapted to a more specific task or domain by retraining its weights either partially or fully on a smaller, domain specific dataset [139]. This dual phase training approach helps the model maintain general knowledge while becoming sensitive to new task specific patterns. In semantic communication, this framework enables the rapid adaptation of communication models to emerging application scenarios. For example, a model pre-trained on open domain conversational data can be fine tuned on a technical corpus such as customer support logs or medical dialogue, allowing it to interpret context rich queries more effectively. This approach is especially valuable in resource constrained environments where collecting and annotating new domain specific data is impractical or expensive.

Prominent pre-trained models such as BERT, GPT, RoBERTa, and T5 demonstrate the power of transfer learning. These models are typically trained on large corpora covering a range of topics and linguistic styles and can be fine tuned for specific semantic communication tasks such as question answering, paraphrase generation, and real time summarization [140]. By transferring learned knowledge of language and semantics, these systems generalize more efficiently to new domains, even when provided with minimal domain specific supervision.

A practical use case illustrating the power of transfer learning in semantic communication is its application in low resource cross lingual emergency response systems. During natural disasters or humanitarian crises, timely and accurate communication across multiple languages and dialects is critical. However, building and training separate models for each linguistic group is not always feasible due to time and resource constraints. By pre-training a semantic encoder on a high resource language corpus and then fine tuning it on limited examples of regional or dialect specific emergency texts, transfer learning enables effective communication across linguistic barriers. This allows for faster model deployment and ensures generalization across scenarios involving varied languages, writing styles, and urgency levels.

Across the techniques reviewed in this section, each family makes distinct trade-offs in generalization, compute, data, and stability summarized in Table 7. Meta-learning typically delivers the strongest out of domain performance and fastest few-shot adaptation, but incurs high

Table 7
Comparative trade-offs of domain generalization techniques.

DG technique	Key methods	Gen. performance	Complexity	Data Req.	Stability	Strengths	Limitations	Best for shifts
Meta-learning	MAML, MLDG, MetaReg	Strong, fast few-shot adaptation	High (bi-level opt.)	Moderate–high (diverse tasks)	Sensitive; unstable if poorly sampled	Rapid adaptation, versatile	Expensive training, episodic overhead	Source modality shifts, heterogeneous tasks
Domain-adversarial	DANN, GRL, ADDA	Good average-case robustness	Moderate (adds discriminator)	Low–moderate	Moderate; adversarial risk of collapse	Learns domain-invariant features, efficient	May fail on large unseen shifts	Channel shifts, heterogeneous envs.
Gradient-based	RSC, Fish, IRM	Stable, incremental improvements	Low–moderate (light ops)	Low	High stability	Simple, efficient, plug-and-play	Limited robustness on severe shifts	Source-level shifts (linguistic/visual)
DRO	GroupDRO, VREx, JTT	Strong worst-case guarantees	High (min–max opt.)	Moderate	Less stable (saddle-point issues)	Protects against worst-case domains	Computationally heavy, slower convergence	Severe channel + modality shifts
Self-supervised	SimCLR, MoCo, BYOL	Good via large pretraining	High (pretrain)/Low (fine-tune)	High (large unlabeled corpora)	Stable after pretraining	Uses unlabeled data; versatile	Expensive pretraining, task alignment needed	Broad shifts (linguistic, multimodal)
Generative	GANs, Diffusion, VAE	High fidelity reconstructions	High (large models)	High (paired/unpaired data)	GANs unstable; diffusion more stable but slow	Improves semantic fidelity, robust to corruption	Compute-heavy, latency issues	Channel corruption, noisy reconstructions
Ensemble	Bagging, Boosting, Snapshot Ensembles	Moderate boost via aggregation	High (parallel inference)	Low–moderate	Stable	Reliable, improves robustness	High inference/storage cost	General robustness across domains
Transfer learning	BERT, ResNet-pretrain, CLIP	Good when pretrained matches target	Low–moderate (fine-tuning)	Low (few-shot possible)	Stable; depends on pretrained quality	Strong baseline, efficient	Domain mismatch hurts performance	Cross-domain semantic shifts, low-resource settings

training overhead (bi-level optimization, episodic sampling) and can be brittle to task selection. By contrast, domain adversarial methods add only moderate complexity and deploy broadly, yet their domain invariant features can underperform under extreme or non-stationary shifts. Gradient based regularizers (e.g., self challenging, gradient matching) are lightweight and stable, offering an attractive efficiency robustness balance, though gains taper under severe distribution gaps. Distributionally robust optimization prioritizes worst case guarantees in unpredictable settings, at the cost of heavier min–max training and careful tuning. Self-supervised/contrastive pretraining lowers label dependence and scales well, but shifts cost to pretraining and demands alignment with downstream semantics. Generative/reconstruction driven models (GANs, diffusion, VAEs) improve semantic fidelity and corruption robustness, while increasing model size, latency, and (for GANs) training instability. Finally, ensembles and transfer learning are practical, scalable baselines, ensembles reliably boost robustness at parallel inference cost, and transfer learning excels when pretraining data match the target domain but degrades under semantic mismatch. In practice, method choice should reflect shift type channel corruption often favors generative or adversarial approaches, cross modality or semantic shifts favor self-supervised or transfer learning, and high uncertainty deployments justify DRO while hybrid designs balance robustness, latency, and compute budget.

In summary, the techniques discussed in this section represent significant advancements in enabling domain generalization for AI-enabled semantic communication. Each method ranging from meta-learning and domain adversarial training to gradient based optimization, distributionally robust optimization, self-supervised learning, generative modeling, ensemble strategies, and transfer learning offers unique mechanisms to improve adaptability, robustness, and scalability. When integrated into semantic communication architectures, these approaches empower models to operate reliably across dynamic and heterogeneous environments, reducing dependence on domain specific training and enhancing generalization to unseen conditions. As semantic communication systems are deployed in increasingly complex real world scenarios, the adoption of these domain generalization strategies will be essential for ensuring consistent performance, low latency adaptation, and reliable interpretation of meaning across varied linguistic, contextual, and operational domains. These capabilities are vital for next generation wireless technologies, supporting applications in autonomous systems, distributed AI agents, human machine interaction, and global scale communication networks.

5. Open issues, potential solutions, and future research directions in domain generalization for semantic communication

AI-enabled semantic communication represents a significant advancement in wireless communication, shifting the focus from traditional bit level transmission to the efficient exchange of meaningful information. While notable strides have been made in developing architectures and algorithms that support semantic inference, a number of critical research challenges remain particularly in the context of domain generalization. As semantic communication systems are deployed in increasingly diverse environments, the ability of AI models to generalize across different domains without retraining becomes a central requirement.

This section explores open issues and potential research directions related to domain generalization in semantic communication. We identify key challenges that impede the adaptability, robustness, and scalability of current systems and propose methodological strategies to address them. These challenges are categorized into four major themes:

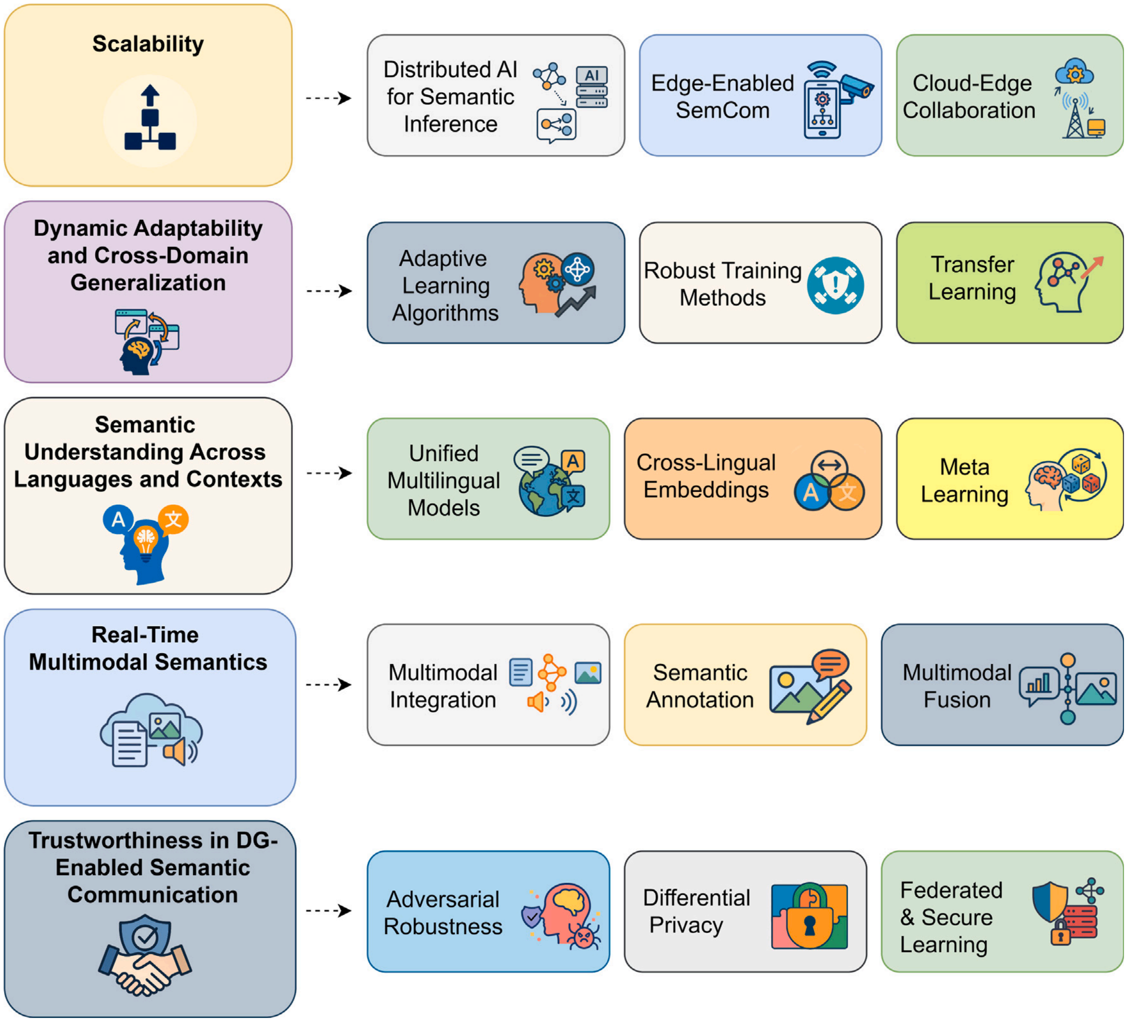

Fig. 8. Open Challenges and Solution Pathways for Domain Generalization in Semantic Communication System.

(i) scalability in large and heterogeneous networks, (ii) dynamic adaptability across evolving domains, (iii) semantic understanding across languages and contexts, and (iv) real time processing of unstructured multimodal data.

Each challenge is paired with promising solution pathways, offering a structured roadmap for future research and system development. Fig. 8 provides a conceptual overview of the open problems and associated strategies. This section aims to guide researchers and practitioners toward building resilient and generalizable semantic communication systems capable of operating reliably across a wide range of real world scenarios.

1. Challenge: Scalability

Scalability is a core challenge in AI-enabled semantic communication, particularly when systems are expected to operate across vast, diverse, and dynamic environments. In the context of domain generalization, scalability refers not only to managing large data volumes but also to maintaining model robustness and adaptability across distributed domains without extensive retraining. With the emergence of 6G and intelligent networks, AI-based semantic systems must support massive IoT infrastructures, real time vehicular communication, AR/VR streaming, and mission critical services each generating multimodal, non stationary data that challenge both communication efficiency and semantic consistency [141,142].

As these networks scale, ensuring that semantic communication models generalize well across varied operational contexts languages, use cases, and devices without centralized retraining or performance degradation becomes increasingly difficult.

Potential Solutions:

- *Distributed AI for Semantic Inference:* Distributing semantic communication models across nodes within the network (e.g., user devices, sensors, base stations) allows for localized semantic encoding and decoding that is better aligned with domain specific data. This setup supports generalization by learning domain aware patterns in situ while maintaining coordination with the global model. For instance, federated or decentralized training frameworks can be employed for domain aware semantic learning across IoT devices without data centralization, thus scaling semantic communication while preserving privacy and local domain specificity [143,144].
- *Edge Enabled Semantic Communication:* AI models deployed at the edge (e.g., on mobile devices, smart cameras, or vehicular units) can perform near real time semantic encoding, reducing transmission overhead and adapting to the local environment. In AI-enabled semantic communication, edge AI enables rapid adaptation to context (e.g., accent, dialect, or sensor modality), improving generalization without requiring constant connectivity to centralized servers. For example, semantic decoders on autonomous vehicles can dynamically adapt to different environments or road signage styles across cities [145,146].
- *Cloud Edge Collaboration for Semantic Adaptation:* Combining edge responsiveness with cloud level training and knowledge management can address both generalization and scalability. Edge devices can handle lightweight semantic inference while the cloud

aggregates feedback from multiple domains to refine and synchronize global semantic representations. This architecture enables AI-enabled semantic communication systems to continuously adapt to unseen domains through periodic updates while maintaining responsiveness and scalability in real time deployments [147].

The system should achieve at least a 50% reduction in floating-point operations (FLOPs) or energy consumption per inference relative to baseline semantic encoders such as DeepSC, while maintaining high semantic fidelity with target thresholds of BLEU $\geqslant$ 0.75 or PSNR $\geqslant$ 30 dB under in-distribution to out-of-distribution (ID→ OOD) shifts. Moreover, the end to end inference latency at the network edge should be limited to 20–30 ms for interactive applications and further reduced to no more than 10 ms in ultra-reliable low-latency communication (URLLC) scenarios, ensuring both efficiency and responsiveness across diverse use cases [148,149].

2. Challenge: Dynamic Adaptability and Cross Domain Generalization

In AI-enabled semantic communication, dynamic adaptability and cross domain generalization are essential to maintaining robust performance across diverse operating environments. Communication networks are inherently dynamic, subject to rapid changes in user demand, environmental factors, network topology, and traffic patterns. For semantic communication systems to remain effective, AI models must adapt in real time to novel data distributions, evolving channel conditions, and domain shifts such as variations in language, context, or modality without retraining from scratch [150,151].

These requirements are especially critical in mission critical applications such as autonomous transportation, real time augmented reality (AR), and industrial IoT, where semantic fidelity must be preserved under continuous change. The ability to dynamically generalize across domains ensures that models maintain semantic accuracy and low latency in unseen scenarios, which is fundamental for deploying scalable and dependable semantic communication solutions.

Potential Solutions:

- *Adaptive Learning Algorithms:* Adaptive learning approaches allow semantic encoders and decoders to update their parameters in real time in response to distributional changes. These algorithms enhance cross domain generalization by enabling the system to learn from evolving patterns such as changes in user intent, context, or signal quality without compromising efficiency. Techniques like online meta-learning, experience replay, or few shot fine tuning help maintain semantic performance even under severe domain shifts. This adaptability is vital in systems where semantic communication must respond quickly to new interaction modalities or user behaviors [152,153].
- *Robust Training Methods:* Adopting robust training paradigms such as adversarial learning and distributionally robust optimization (DRO) prepares models to withstand noise, adversarial input, or worst case domain shifts. In semantic communication, adversarial training exposes models to perturbed inputs that simulate real world variation in encoding, transmission, or contextual ambiguity. DRO optimizes models not just for average performance but for worst case distribution performance, ensuring resilience across heterogeneous domains. These methods collectively increase the model's ability to generalize semantic representations under unpredictable and volatile conditions [154,155].
- *Transfer Learning:* Transfer learning enables semantic models to reuse knowledge from related domains and quickly adapt to new communication scenarios. This is particularly valuable when labeled data in the target domain is scarce or when rapid deployment is necessary. For example, a model pre-trained on general language or visual data can be fine tuned with limited task specific examples such as domain specific speech patterns, scene descriptors, or device logs while maintaining performance. This accelerates deployment and enhances generalization across diverse communication contexts such as multilingual virtual assistants or cross domain collaborative robotics [156].

Ideally, the adaptation cost should be limited to no more than five gradient steps with the use of at most 1% labeled target data, while the online adaptation time should not exceed one second. In addition, the relative performance degradation under ID→ OOD shifts should remain within 10% or less, measured using BLEU score or accuracy, and assessed across at least two previously unseen channel or content domains [157,158].

3. Challenge: Semantic Understanding Across Languages and Contexts

Semantic communication systems are designed to extract, encode, and transmit the meaning of messages rather than raw symbols. However, when deployed in multilingual and cross cultural environments, these systems face the complex challenge of preserving semantics across diverse linguistic structures, cultural norms, and contextual interpretations. Each language presents unique syntax, grammar, idiomatic expressions, and semantic nuances, making direct translation or semantic mapping non-trivial. Moreover, context specific meanings such as sarcasm, humor, or metaphor require AI models to perform fine grained semantic inference beyond surface level understanding [159].

The challenge becomes more acute in low resource languages, where labeled datasets are scarce and linguistic structures may differ significantly from those in high resource settings. Without effective domain generalization, semantic communication systems risk performance degradation or meaning distortion when operating across such linguistic gaps. Furthermore, as language evolves introducing new expressions, regional dialects, or user specific jargon models must dynamically adapt to ensure semantic consistency across shifting linguistic domains [160]. Evaluating semantic fidelity in this context is also difficult, as standard benchmarks may fail to capture the full spectrum of cultural and contextual diversity [161].

Potential Solutions:

- *Unified Multilingual Models:* Leveraging multilingual pre-trained models such as mBERT or XLM-RoBERTa offers a scalable solution for cross lingual semantic communication. These models are trained on diverse multilingual corpora and can capture generalizable linguistic patterns that support domain transfer from high resource to low resource languages. In semantic communication, such models enable robust semantic encoding and decoding across languages within a single architecture, reducing the need for language specific pipelines and improving system maintainability [162].
- *Cross Lingual Embeddings:* Embedding alignment methods map words from different languages into a shared semantic space, ensuring that semantically similar concepts are closely positioned regardless of their language of origin. This approach enhances the ability of semantic communication models to generalize across languages, enabling accurate message reconstruction and intent recognition in multilingual systems. Tools such as MUSE and VecMap support both supervised and unsupervised embedding alignment, benefiting tasks like cross lingual retrieval, multilingual dialogue generation, and translation [160,163].
- *Meta-Learning:* For low resource or emergent languages, meta-learning can provide a mechanism for efficient adaptation with minimal training data. Meta-learning frameworks train the model on a variety of linguistic tasks, allowing it to learn transferable semantic representations. When applied to semantic communication, these techniques enable rapid adaptation to new languages, dialects, or cultural expressions with limited supervision. This supports more inclusive and globally adaptable semantic communication systems, especially in multilingual environments such as global customer service or international disaster response networks [96,161].

In cross-lingual transfer, performance should retain at least 90% of the in-domain BLEU score when applied to OOD languages. For low resource adaptation, fine-tuning should be achievable with no more than 1000 parallel sentences. Moreover, semantic similarity should remain at or above 0.80 in zero or few-shot settings, evaluated across at least two distinct language families [164,165].

4. Challenge: Semantics for Unstructured Data in Real Time Multimodal Communication

One of the most complex challenges in achieving domain generalization within AI-enabled semantic communication systems lies in processing unstructured, multimodal data streams such as text, audio, images, and video in real time. Unlike structured tabular inputs, unstructured data lacks predefined formats and often carries nuanced semantic meaning embedded in variable length sequences, contextual dependencies, or multi sensor inputs. These data types dominate many real world communication settings, including video conferencing, intelligent surveillance, and human robot interaction, yet remain difficult to model due to their heterogeneity and semantic ambiguity [166].

Semantic communication systems are required not only to transmit this data efficiently but to extract and convey its underlying meaning accurately across domain shifts. These shifts may manifest in changes to visual context, acoustic background, user interaction styles, or data collection modalities. Therefore, models must generalize across diverse multimodal sources while remaining responsive to temporal dynamics and context specific semantic cues. Achieving this goal demands architectures that integrate, annotate, and fuse multimodal signals in a domain aware and latency sensitive manner [167].

Potential Solutions:

- *Multimodal Integration:* A key enabler of semantic generalization across unstructured data sources is the development of multimodal representation learning. Cross modal learning techniques allow semantic communication models to capture relationships between text, speech, and visual inputs, ensuring consistent meaning inference across domains. Transformer based architectures with multi head attention and shared embedding spaces can align modality specific semantics during training, enabling the model to adapt when one modality changes (e.g., different dialects, lighting conditions, or image resolutions). Training on diverse, mixed domain datasets also improves generalization under varying real world conditions [167,168].
- *Semantic Annotation:* Automated semantic annotation methods are critical for structuring unlabeled multimodal data into formats usable by semantic encoders and decoders. Annotating visual or audio signals with contextual tags (e.g., object types, speaker intent, emotion) supports domain generalization by reducing ambiguity in cross modal alignment. Techniques such as convolutional neural networks (CNNs) for image tagging, large language models (LLMs) for contextual text generation [169], and weak supervision methods for semi automated labeling help bootstrap domain robust annotations. These annotations allow models to extract generalizable patterns from limited or noisy data and maintain semantic coherence across different deployment environments [170,171].
- *Multimodal Fusion:* Real time semantic communication requires efficient fusion of multimodal features to generate coherent messages. Fusion strategies such as early fusion (combining raw data), late fusion (merging modality specific outputs), or hybrid fusion support the creation of unified semantic representations that retain the relevance of each input source. Attention based fusion mechanisms further improve domain generalization by dynamically weighting modality importance depending on context (e.g., prioritizing visual data in noisy environments or text in ambiguous audio). These fusion strategies enable semantic systems to adaptively interpret diverse inputs and maintain robustness under domain shifts [172].

For real time multimodal pipelines, the end to end latency should remain below 10 ms for URLLC scenarios and within 20–30 ms for XR or other interactive tasks operating at a minimum of 30 frames per second. The system should sustain a PSNR of at least 30 dB for images and video, as well as a semantic similarity score of at least 0.80 under OOD channel conditions and noise. Additionally, the fusion overhead at the edge should contribute no more than 20% of the total latency [173,174].

5. Challenge: Trustworthiness in DG-Enabled Semantic Communication

A critical but often overlooked challenge in deploying domain generalization within semantic communication systems is ensuring trustworthiness, particularly against security risks and privacy concerns. While DG enhances robustness to distributional shifts, models remain vulnerable to adversarial attacks that exploit domain invariant features. Such perturbations can cause semantic misinterpretations or even malicious manipulation of transmitted meaning, posing risks in safety critical applications such as autonomous driving, telemedicine, and industrial automation. Additionally, reliance on shared knowledge bases (KBs) for grounding semantics raises privacy concerns, as sensitive contextual information may be inadvertently exposed or inferred, undermining user trust and data protection [175].

For DG-enabled semantic systems to achieve widespread adoption in real-world 6G deployments, they must not only generalize effectively but also ensure resilience against malicious actors and guarantee the confidentiality of semantic data. Addressing these dual concerns requires embedding security and privacy preservation as integral design principles rather than afterthoughts.

Potential Solutions:

- *Adversarial Robustness:* Apply adversarial training and certified robustness methods to harden DG models against perturbations targeting domain invariant features, ensuring semantic fidelity under malicious attacks [176].
- *Differential Privacy (DP):* Integrate DP into DG training and adaptation to provide formal guarantees against leakage of sensitive user data, enabling privacy preserving generalization without exposing knowledge base contents [177].
- *Federated and Secure Learning:* Employ federated DG frameworks with secure aggregation so that raw user data remains on-device, reducing central exposure while maintaining adaptability across heterogeneous domains. [178]

To ensure trustworthy deployment, DG-enabled semantic communication systems should demonstrate resilience against adversarial perturbations with less than 5% performance degradation, while maintaining differential privacy guarantees (e.g., $\epsilon < 5$) during training and adaptation. These safeguards should be achieved without exceeding the latency and energy constraints of URLLC or XR applications, thereby balancing generalization, efficiency, and trustworthiness [179].

The challenges outlined in this section including scalability, dynamic adaptability, semantic consistency across languages and contexts, the processing of unstructured multimodal data, and the overarching issue of trustworthiness constitute major barriers to realizing robust domain generalization in AI-enabled semantic communication. Overcoming these obstacles is vital for the design of systems capable of operating reliably across diverse real world conditions, heterogeneous user populations, and varied application domains. The potential solutions discussed such as distributed AI, adaptive learning, multilingual modeling, and multimodal fusion offer promising avenues toward building semantic communication frameworks that are not only accurate and context-aware but also resilient to environmental variability and domain shifts. As semantic communication advances, future research must continue to refine these strategies to ensure that AI models can generalize autonomously and meaningfully across the full spectrum of data modalities, linguistic diversity, and contextual dynamics that define next-generation global communication networks.

6. Conclusion

The integration of AI into semantic communication systems represents a foundational shift in how meaning is extracted, transmitted, and interpreted within next generation wireless networks. As networks become increasingly heterogeneous and data environments more diverse, the ability of AI models to generalize across unseen domains without retraining has emerged as a defining requirement for robust and scalable communication systems. This survey has provided a focused and indepth examination of domain generalization in the context of AI-enabled semantic communication, identifying its role as both a challenge and a catalyst for system wide adaptability. We discussed a broad set of techniques including meta-learning, domain adversarial training, distributionally robust optimization, self-supervised learning, transfer learning, and multimodal fusion that enable semantic communication systems to maintain performance across varying languages, modalities, environments, and user contexts. Furthermore, we identified key open issues related to scalability, dynamic adaptability, multilingual understanding, and real time unstructured data processing, offering a roadmap of potential solutions and research directions. By consolidating the current state of research in domain generalization for semantic communication, this work offers a reference framework for future studies aiming to develop context aware, resilient, and deployment ready AI communication models. As the demand for intelligent connectivity continues to grow across sectors ranging from IoT and autonomous systems to smart cities and global infrastructure ensuring that AI-enabled semantic communication systems can generalize across real world complexities will be central to their success. We hope this survey will inspire further interdisciplinary efforts to bridge the gap between generalization theory and practical communication system design, paving the way toward truly adaptive, intelligent, and globally interoperable wireless networks.

CRediT authorship contribution statement

Muhammad Furqan Zia: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Messaoud Ahmed Ouameur:** Supervision, Writing – review & editing, Validation, Project administration. **Miloud Bagaa:** Writing – review & editing. **Daniel Massicotte:** Writing – review & editing. **Adlen Ksentini:** Writing – review & editing.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this survey, generative AI tools (e.g., ChatGPT by OpenAI) were used to support drafting, formatting, and language refinement. All technical content, conceptual framing, and analytical insights were developed and validated solely by the authors. The authors accept full responsibility for the integrity and accuracy of the final manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work was supported under the discovery fund granted by the Natural Sciences and Engineering Research Council of Canada (NSERC).

Data availability

No data was used for the research described in the article.

References

[1] C.-X. Wang, X. You, X. Gao, X. Zhu, Z. Li, C. Zhang, H. Wang, Y. Huang, Y. Chen, H. Haas, J.S. Thompson, E.G. Larsson, M.D. Renzo, W. Tong, P. Zhu, X. Shen, H.V. Poor, L. Hanzo, On the road to 6G: Visions, requirements, key technologies, and testbeds, IEEE Commun. Surv. Tutor. 25 (2) (2023) 905–974, http://dx.doi.org/10.1109/COMST.2023.3249835.

[2] M. Banafaa, I. Shayea, J. Din, M. Hadri Azmi, A. Alashbi, Y. Ibrahim Daradkeh, A. Alhammadi, 6G mobile communication technology: Requirements, targets, applications, challenges, advantages, and opportunities, Alex. Eng. J. 64 (2023) 245–274, http://dx.doi.org/10.1016/j.aej.2022.08.017, URL https://www.sciencedirect.com/science/article/pii/S111001682200549X.

[3] N. Khan, S. Coleri, A. Abdallah, A. Çelik, A.M. Eltawil, Explainable and robust artificial intelligence for trustworthy resource management in 6G networks, IEEE Commun. Mag. 62 (4) (2023) 1–7, http://dx.doi.org/10.1109/MCOM.001.2300172.

[4] L. Zhao, Y. Zhao, S. Zhou, P. Yu, Q. Dong, L. Wang, J. Li, Discuss the key development direction of 6G key technology based on 5G technology, in: 2023 2nd International Conference on Artificial Intelligence and Computer Information Technology, AICIT, 2023, pp. 1–3, http://dx.doi.org/10.1109/AICIT59054.2023.10277958.

[5] S. Iyer, R. Khanai, D. Torse, et al., A survey on semantic communications for intelligent wireless networks, Wirel. Pers. Commun. 129 (2023) 569–611, http://dx.doi.org/10.1007/s11277-022-10111-7.

[6] R. Bruschi, J.F. Pajo, F. Davoli, C. Lombardo, Managing 5G network slicing and edge computing with the MATILDA telecom layer platform, Comput. Netw. 194 (2021) 108090, http://dx.doi.org/10.1016/j.comnet.2021.108090, URL https://www.sciencedirect.com/science/article/pii/S1389128621001766.

[7] W. Yang, Z.Q. Liew, W.Y.B. Lim, Z. Xiong, D. Niyato, X. Chi, X. Cao, K.B. Letaief, Semantic communication meets edge intelligence, IEEE Wirel. Commun. 29 (5) (2022) 28–35, http://dx.doi.org/10.1109/MWC.004.2200050.

[8] A. Haidine, F.Z. Salmam, A. Aqqal, A. Dahbi, Artificial intelligence and machine learning in 5G and beyond: A survey and perspectives, in: A. Haidine (Ed.), Moving Broadband Mobile Communications Forward, IntechOpen, Rijeka, 2021, http://dx.doi.org/10.5772/intechopen.98517, Ch. 3, URL https://doi.org/10.5772/intechopen.98517.

[9] Y. Sun, M. Peng, Y. Zhou, Y. Huang, S. Mao, Application of machine learning in wireless networks: Key techniques and open issues, IEEE Commun. Surv. Tutor. 21 (4) (2019) 3072–3108, http://dx.doi.org/10.1109/COMST.2019.2924243.

[10] K. Li, B.P.L. Lau, X. Yuan, W. Ni, M. Guizani, C. Yuen, Toward ubiquitous semantic metaverse: Challenges, approaches, and opportunities, IEEE Internet Things J. 10 (24) (2023) 21855–21872, http://dx.doi.org/10.1109/JIOT.2023.3302159.

[11] A. Celik, A.M. Eltawil, At the dawn of generative AI era: A tutorial-cum-survey on new frontiers in 6G wireless intelligence, IEEE Open J. Commun. Soc. (2024) http://dx.doi.org/10.1109/OJCOMS.2024.3362271, 1–1.

[12] L. Chen, S.T. Jose, I. Nikoloska, S. Park, T. Chen, O. Simeone, Learning with limited samples – meta-learning and applications to communication systems, 2022.

[13] N. Khan, S. Coleri, A. Abdallah, A. Celik, A.M. Eltawil, Explainable and robust artificial intelligence for trustworthy resource management in 6G networks, IEEE Commun. Mag. (2023) 1–7, http://dx.doi.org/10.1109/MCOM.001.2300172.

[14] S. Mukherjee, J.W. Rupe, J.S. Zhu, XAI for communication networks, in: 2022 IEEE International Symposium on Software Reliability Engineering Workshops, ISSREW, 2022, pp. 359–364, URL https://api.semanticscholar.org/CorpusID:255171438.

[15] M.F. Zia, H.M. Furqan, J.M. Hamamreh, Multi-cell, multi-user, and multi-carrier secure communication using non-orthogonal signals' superposition with dual-transmission for IoT in 6G and beyond, RS Open J. Innov. Commun. Technol. 2 (3) (2021) http://dx.doi.org/10.46470/03d8ffbd.08b7bd1d.

[16] M.F. Zia, Secrecy and resilience in next-gen Wi-Fi: Exploring a multi-user down-link non-orthogonal transmission framework, in: 2024 Intermountain Engineering, Technology and Computing, IETC, 2024, pp. 1–6, http://dx.doi.org/10.1109/IETC61393.2024.10564473.

[17] Z. Lu, R. Li, K. Lu, X. Chen, E. Hossain, Z. Zhao, H. Zhang, Semantics-empowered communications: A tutorial-cum-survey, IEEE Commun. Surv. Tutor. 26 (1) (2024) 41–79, http://dx.doi.org/10.1109/COMST.2023.3333342.

[18] W. Weaver, Recent contributions to the mathematical theory of communication, ETC: A Rev. Gen. Semant. 10 (4) (1953) 261–281, URL http://www.jstor.org/stable/42581364.

[19] J. Bao, P. Basu, M. Dean, C. Partridge, A. Swami, W. Leland, J.A. Hendler, Towards a theory of semantic communication, in: 2011 IEEE Network Science Workshop, 2011, pp. 110–117, http://dx.doi.org/10.1109/NSW.2011.6004632.

[20] W. Yang, H. Du, Z.Q. Liew, W.Y.B. Lim, Z. Xiong, D. Niyato, X. Chi, X. Shen, C. Miao, Semantic communications for future internet: Fundamentals, applications, and challenges, IEEE Commun. Surv. Tutor. 25 (1) (2023) 213–250, http://dx.doi.org/10.1109/COMST.2022.3223224.

[21] C. Liang, H. Du, Y. Sun, D. Niyato, J. Kang, D. Zhao, M.A. Imran, Generative AI-driven semantic communication networks: Architecture, technologies, and applications, IEEE Trans. Cogn. Commun. Netw. 11 (1) (2025) 27–47, http://dx.doi.org/10.1109/TCCN.2024.3435524.

[22] F. Jiang, L. Dong, Y. Peng, K. Wang, K. Yang, C. Pan, X. You, Large AI model empowered multimodal semantic communications, IEEE Commun. Mag. (2024).

[23] M.U. Lokumarambage, V.S.S. Gowrisetty, H. Rezaei, T. Sivalingam, N. Rajatheva, A. Fernando, Wireless end-to-end image transmission system using semantic communications, IEEE Access 11 (2023) 37149–37163, http://dx.doi.org/10.1109/ACCESS.2023.3266656.

[24] D. Wheeler, B. Natarajan, Engineering semantic communication: A survey, IEEE Access 11 (2023) 13965–13995, http://dx.doi.org/10.1109/ACCESS.2023.3243065.

[25] Q. Wu, F. Liu, H. Xia, T. Zhang, Semantic transfer between different tasks in the semantic communication system, in: 2022 IEEE Wireless Communications and Networking Conference, WCNC, 2022, pp. 566–571, http://dx.doi.org/10.1109/WCNC51071.2022.9771768.

[26] Y. Liu, S. Jiang, Y. Zhang, K. Cao, L. Zhou, B.-C. Seet, H. Zhao, J. Wei, Extended context-based semantic communication system for text transmission, Digit. Commun. Netw. (2022) http://dx.doi.org/10.1016/j.dcan.2022.09.023, URL https://www.sciencedirect.com/science/article/pii/S2352864822001985.

[27] E. Beck, C. Bockelmann, A. Dekorsy, Semantic information recovery in wireless networks, Sensors 23 (14) (2023) 6347, http://dx.doi.org/10.3390/s23146347.

[28] P. Popovski, O. Simeone, F. Boccardi, D. Gündüz, O. Sahin, Semantic-Effectiveness filtering and control for Post-5G wireless connectivity, J. Indian Inst. Sci. 100 (2) (2020) 435–443.

[29] S. Pumma, A. Vishnu, Semantic-aware lossless data compression for deep learning recommendation model (DLRM), in: 2021 IEEE/ACM Workshop on Machine Learning in High Performance Computing Environments, MLHPC, 2021, pp. 1–8, http://dx.doi.org/10.1109/MLHPC54614.2021.00006.

[30] N. Alhakbani, M.M. Hassan, M. Ykhlef, G. Fortino, An efficient event matching system for semantic smart data in the Internet of Things (IoT) environment, Future Gener. Comput. Syst. 95 (2019) 163–174, http://dx.doi.org/10.1016/j.future.2018.12.064, URL https://www.sciencedirect.com/science/article/pii/S0167739X1831478X.

[31] L. Fallon, Y. Huang, D. O'Sullivan, Towards automated analysis and optimization of multimedia streaming services using clustering and semantic techniques, in: R. Brennan, J. Fleck, S. van der Meer (Eds.), Modelling Autonomic Communication Environments, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010, pp. 12–23.

[32] J. Wang, C. Lan, C. Liu, Y. Ouyang, T. Qin, W. Lu, Y. Chen, W. Zeng, P.S. Yu, Generalizing to unseen domains: A survey on domain generalization, IEEE Trans. Knowl. Data Eng. 35 (8) (2023) 8052–8072, http://dx.doi.org/10.1109/TKDE.2022.3178128.

[33] M. Akrout, A. Feriani, F. Bellili, A. Mezghani, E. Hossain, Domain generalization in machine learning models for wireless communications: Concepts, state-of-the-art, and open issues, IEEE Commun. Surv. Tutor. 25 (4) (2023) 3014–3037, http://dx.doi.org/10.1109/COMST.2023.3326399.

[34] C. Chaccour, W. Saad, M. Debbah, Z. Han, H. Vincent Poor, Less data, more knowledge: Building next-generation semantic communication networks, IEEE Commun. Surv. Tutor. 27 (1) (2025) 37–76, http://dx.doi.org/10.1109/COMST.2024.3412852.

[35] A. Arsénio, H. Serra, R. Francisco, F. Nabais, J. Andrade, E. Serrano, Internet of intelligent things: Bringing artificial intelligence into things and communication networks, in: F. Xhafa, N. Bessis (Eds.), Inter-Cooperative Collective Intelligence: Techniques and Applications, Springer Berlin Heidelberg, Berlin, Heidelberg, 2014, pp. 1–37, http://dx.doi.org/10.1007/978-3-642-35016-0_1.

[36] Z. Qin, T. Zhao, F. Li, X. Tao, Survey of research on multimodal semantic communication, J. Commun. 44 (5) (2023) 28–41.

[37] Z. Jiang, F.F. Xu, J. Araki, G. Neubig, How can we know what language models know? Trans. Assoc. Comput. Linguist. 8 (2020) 423–438, http://dx.doi.org/10.1162/tacl_a_00324.

[38] A.I. Maarala, X. Su, J. Riekki, Semantic reasoning for context-aware Internet of Things applications, IEEE Internet Things J. 4 (2) (2017) 461–473, http://dx.doi.org/10.1109/JIOT.2016.2587060.

[39] K.B. Letaief, W. Chen, Y. Shi, J. Zhang, Y.-J.A. Zhang, The roadmap to 6G: AI empowered wireless networks, IEEE Commun. Mag. 57 (8) (2019) 84–90, http://dx.doi.org/10.1109/MCOM.2019.1900271.

[40] H. Song, J. Bai, Y. Yi, J. Wu, L. Liu, Artificial intelligence enabled Internet of Things: Network architecture and spectrum access, IEEE Comput. Intell. Mag. 15 (1) (2020) 44–51, http://dx.doi.org/10.1109/MCI.2019.2954643.

[41] B. Zhao, J. Wu, Y. Ma, C. Yang, Meta-learning for wireless communications: A survey and a comparison to GNNs, IEEE Open J. Commun. Soc. 5 (2024) 1987–2015, http://dx.doi.org/10.1109/OJCOMS.2024.3380512.

[42] S. Imran, G. Charan, A. Alkhateeb, Environment semantic communication: Enabling distributed sensing aided networks, IEEE Open J. Commun. Soc. 5 (2024) 7767–7786, http://dx.doi.org/10.1109/OJCOMS.2024.3509453.

[43] Z. Qin, F. Gao, B. Lin, X. Tao, G. Liu, C. Pan, A generalized semantic communication system: From sources to channels, IEEE Wirel. Commun. 30 (3) (2023) 18–26, http://dx.doi.org/10.1109/MWC.013.2200553.

[44] A. Bensky, Chapter 14 - technologies and applications, in: A. Bensky (Ed.), Short-Range Wireless Communication, third ed., Newnes, 2019, pp. 387–430, http://dx.doi.org/10.1016/B978-0-12-815405-2.00014-2, URL https://www.sciencedirect.com/science/article/pii/B9780128154052000142.

[45] Y. Ding, L. Wang, B. Liang, S. Liang, Y. Wang, F. Chen, Domain generalization by learning and removing domain-specific features, in: S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, A. Oh (Eds.), in: Advances in Neural Information Processing Systems, vol. 35, Curran Associates, Inc., 2022, pp. 24226–24239, URL https://proceedings.neurips.cc/paper_files/paper/2022/file/9941833e8327910ef25daeb9005e4748-Paper-Conference.pdf.

[46] J. Wang, H. Li, S. Pan, X. Xie, A tutorial on domain generalization, in: Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining, WSDM '23, 2023, pp. 1236–1239, http://dx.doi.org/10.1145/3539597.3572722.

[47] B. Wang, M. Lapata, I. Titov, Meta-learning for domain generalization in semantic parsing, in: K. Toutanova, A. Rumshisky, L. Zettlemoyer, D. Hakkani-Tur, I. Beltagy, S. Bethard, R. Cotterell, T. Chakraborty, Y. Zhou (Eds.), Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Association for Computational Linguistics, Online, 2021, pp. 366–379, http://dx.doi.org/10.18653/v1/2021.naacl-main.33, URL https://aclanthology.org/2021.naacl-main.33/.

[48] K. Zhou, Z. Liu, Y. Qiao, T. Xiang, C.C. Loy, Domain generalization: A survey, IEEE Trans. Pattern Anal. Mach. Intell. 45 (4) (2023) 4396–4415, http://dx.doi.org/10.1109/TPAMI.2022.3195549.

[49] Q. Dou, D.C. Castro, K. Kamnitsas, B. Glocker, Domain generalization via model-agnostic learning of semantic features, in: Proceedings of the 33rd International Conference on Neural Information Processing Systems, Curran Associates Inc., Red Hook, NY, USA, 2019.

[50] K. Zhang, L. Li, W. Lin, Y. Yan, R. Li, W. Cheng, Z. Han, Semantic successive refinement: a generative AI-aided semantic communication framework, IEEE Trans. Cogn. Commun. Netw. (2025).

[51] H. Kim, Y. Kang, C. Oh, K.-J. Yoon, Single domain generalization for LiDAR semantic segmentation, in: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 2023, pp. 17587–17598, http://dx.doi.org/10.1109/CVPR52729.2023.01687.

[52] J. Ren, Y. Sun, H. Du, W. Yuan, C. Wang, X. Wang, Y. Zhou, Z. Zhu, F. Wang, S. Cui, Generative semantic communication: Architectures, technologies, and applications, Engineering (2025) http://dx.doi.org/10.1016/j.eng.2025.07.022, URL https://www.sciencedirect.com/science/article/pii/S2095809925004291.

[53] G. Zhang, Q. Hu, Z. Qin, Y. Cai, G. Yu, A unified multi-task semantic communication system with domain adaptation, in: GLOBECOM 2022 - 2022 IEEE Global Communications Conference, 2022, pp. 3971–3976, http://dx.doi.org/10.1109/GLOBECOM48099.2022.10000850.

[54] E.C. Strinati, P. Di Lorenzo, V. Sciancalepore, A. Aijaz, M. Kountouris, D. Gündüz, P. Popovski, M. Sana, P. Stavrou, B. Soret, N. Cordeschi, S. Scardapane, M. Merluzzi, L. Zanzi, M.B. Renato, T. Quek, N.D. Pietro, O. Forceville, F. Costanzo, P. Li, Goal-Oriented and Semantic Communication in 6G AI-Native Networks: The 6G-GOALS Approach, in: 2024 Joint European Conference on Networks and Communications & 6G Summit, EuCNC/6G Summit, IEEE, Antwerp, Belgium, 2024, pp. 1–6, http://dx.doi.org/10.1109/EuCNC/6GSummit60053.2024.10597087, URL https://cea.hal.science/cea-04760312.

[55] J. Gu, X. Zhang, Q. Cui, X. Tao, Semantic communication for multi-modal data transmission, in: 2023 International Conference on Wireless Communications and Signal Processing, WCSP, 2023, pp. 208–213, http://dx.doi.org/10.1109/WCSP58612.2023.10404655.

[56] F.-E. Yang, Y.-C. Cheng, Z.-Y. Shiau, Y.-C.F. Wang, Adversarial teacher-student representation learning for domain generalization, in: M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, J.W. Vaughan (Eds.), in: Advances in Neural Information Processing Systems, vol. 34, Curran Associates, Inc., 2021, pp. 19448–19460, URL https://proceedings.neurips.cc/paper_files/paper/2021/file/a2137a2ae8e39b5002a3f8909ecb88fe-Paper.pdf.

[57] J. Wang, IJCAI-ECAI 2022 tutorial on domain generalization, 2022, https://jd92.wang/assets/files/DGtutorial_ijcai22.pdf.

[58] Y. Liu, S. Jiang, Y. Zhang, K. Cao, L. Zhou, B.-C. Seet, H. Zhao, J. Wei, Extended context-based semantic communication system for text transmission, Digit. Commun. Netw. 10 (3) (2024) 568–576, http://dx.doi.org/10.1016/j.dcan.2022.09.023, URL https://www.sciencedirect.com/science/article/pii/S2352864822001985.

[59] A.G. Khoee, Y. Yu, R. Feldt, Domain generalization through meta-learning: a survey, Artif. Intell. Rev. 57 (2024) 285, http://dx.doi.org/10.1007/s10462-024-10922-z, Accepted: 22 August 2024, Published: 09 September 2024.

[60] J. Xu, T.-Y. Tung, B. Ai, W. Chen, Y. Sun, D. Gündüz, Deep joint source-channel coding for semantic communications, IEEE Commun. Mag. 61 (11) (2023) 42–48, http://dx.doi.org/10.1109/MCOM.004.2200819.

[61] J. Huang, K. Yuan, C. Huang, K. Huang, D^2-JSCC: Digital deep joint source-channel coding for semantic communications, IEEE J. Sel. Areas Commun. 43 (4) (2025) 1246–1261, http://dx.doi.org/10.1109/JSAC.2025.3531546.

[62] X. Han, Y. Wu, Z. Gao, B. Feng, Y. Shi, D. Gündüz, W. Zhang, SCSC: A novel standards-compatible semantic communication framework for image transmission, IEEE Trans. Commun. (2025).

[63] J. Wu, C. Wu, Y. Lin, T. Yoshinaga, L. Zhong, X. Chen, Y. Ji, Semantic segmentation-based semantic communication system for image transmission, Digit. Commun. Netw. 10 (3) (2024) 519–527, http://dx.doi.org/10.1016/j.dcan.2023.02.006, URL https://www.sciencedirect.com/science/article/pii/S235286482300041X.

[64] Y. Shao, Q. Cao, D. Gündüz, A theory of semantic communication, IEEE Trans. Mob. Comput. 23 (12) (2024) 12211–12228, http://dx.doi.org/10.1109/TMC.2024.3406375.

[65] C. Liu, X. Sun, J. Wang, H. Tang, T. Li, T. Qin, W. Chen, T.-Y. Liu, Learning causal semantic representation for out-of-distribution prediction, in: M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, J.W. Vaughan (Eds.), in: Advances in Neural Information Processing Systems, vol. 34, Curran Associates, Inc., 2021, pp. 6155–6170, URL https://proceedings.neurips.cc/paper_files/paper/2021/file/310614fca8fb8e5491295336298c340f-Paper.pdf.

[66] S.R. Pokhrel, J. Choi, Understand-before-talk (UBT): A semantic communication approach to 6G networks, IEEE Trans. Veh. Technol. 72 (3) (2023) 3544–3556, http://dx.doi.org/10.1109/TVT.2022.3219363.

[67] D. Viji, S. Revathy, Semantic similarity detection from text document using XLNet with a DKM-clustered Bi-LSTM model, in: 2022 13th International Conference on Computing Communication and Networking Technologies, ICCCNT, 2022, pp. 1–8, http://dx.doi.org/10.1109/ICCCNT54827.2022.9984333.

[68] A. Ashari, M. Riasetiawan, Document summarization using TextRank and semantic network, Int. J. Inf. Secur. Appl. 11 (4) (2017) 26, http://dx.doi.org/10.5815/ijisa.2017.11.04.

[69] K. Lu, Q. Zhou, R. Li, Z. Zhao, X. Chen, J. Wu, H. Zhang, Rethinking modern communication from semantic coding to semantic communication, IEEE Wirel. Commun. 30 (1) (2023) 158–164, http://dx.doi.org/10.1109/MWC.013.2100642.

[70] Y. Fu, W. Cheng, W. Zhang, J. Wang, Scalable extraction based semantic communication for 6G wireless networks, IEEE Commun. Mag. (2023) 1–7, http://dx.doi.org/10.1109/MCOM.021.2300269.

[71] T.M. Getu, G. Kaddoum, M. Bennis, Making sense of meaning: A survey on metrics for semantic and goal-oriented communication, 2023, arXiv:2303.10867.

[72] A. Miaschi, D. Brunato, F. Dell'Orletta, G. Venturi, What makes my model perplexed? A linguistic investigation on neural language models perplexity, in: E. Agirre, M. Apidianaki, I. Vulić (Eds.), Proceedings of Deep Learning Inside Out (DeeLIO): The 2nd Workshop on Knowledge Extraction and Integration for Deep Learning Architectures, Association for Computational Linguistics, 2021, pp. 40–47, http://dx.doi.org/10.18653/v1/2021.deelio-1.5, Online, URL https://aclanthology.org/2021.deelio-1.5.

[73] H. Hu, X. Zhu, F. Zhou, W. Wu, R.Q. Hu, H. Zhu, One-to-many semantic communication systems: Design, implementation, performance evaluation, IEEE Commun. Lett. 26 (12) (2022) 2959–2963, http://dx.doi.org/10.1109/LCOMM.2022.3203984.

[74] M. Colombo, Semantic similarity measures, in: Phenotropic Interaction: Improving Interfaces with Computing with Words and Perceptions, Springer Nature Switzerland, Cham, 2024, pp. 49–69, http://dx.doi.org/10.1007/978-3-031-42819-7_4.

[75] T.M. Getu, G. Kaddoum, M. Bennis, Semantic communication: A survey on research landscape, challenges, and future directions, Proc. IEEE 112 (11) (2024) 1649–1685, http://dx.doi.org/10.1109/JPROC.2024.3520707.

[76] T.M. Getu, G. Kaddoum, M. Bennis, A survey on goal-oriented semantic communication: Techniques, challenges, and future directions, IEEE Access 12 (2024) 51223–51274, http://dx.doi.org/10.1109/ACCESS.2024.3381967.

[77] L. Xia, Y. Sun, C. Liang, L. Zhang, M.A. Imran, D. Niyato, Generative AI for semantic communication: Architecture, challenges, and outlook, IEEE Wirel. Commun. 32 (1) (2025) 132–140, http://dx.doi.org/10.1109/MWC.003.2300351.

[78] D. Wheeler, B. Natarajan, Engineering semantic communication: A survey, IEEE Access 11 (2023) 13965–13995, http://dx.doi.org/10.1109/access.2023.3243065.

[79] S. Guo, Y. Wang, N. Zhang, Z. Su, T.H. Luan, Z. Tian, X. Shen, A survey on semantic communication networks: Architecture, security, and privacy, IEEE Commun. Surv. Tutor. (2024) http://dx.doi.org/10.1109/COMST.2024.3516819, 1–1.

[80] Y. Wang, H. Han, Y. Feng, J. Zheng, B. Zhang, Semantic communication empowered 6G networks: Techniques, applications, and challenges, IEEE Access 13 (2025) 28293–28314, http://dx.doi.org/10.1109/ACCESS.2025.3532797.

[81] T.M. Getu, G. Kaddoum, M. Bennis, Making sense of meaning: A survey on metrics for semantic and goal-oriented communication, IEEE Access 11 (2023) 45456–45492, http://dx.doi.org/10.1109/ACCESS.2023.3269848.

[82] T.M. Getu, G. Kaddoum, M. Bennis, A survey on goal-oriented semantic communication: Techniques, challenges, and future directions, IEEE Access 12 (2024) 51223–51274.

[83] K. Zhou, Z. Liu, Y. Qiao, T. Xiang, C.C. Loy, Domain generalization in vision: A survey, 2021, URL https://api.semanticscholar.org/CorpusID:236090250.

[84] H. Xie, Z. Qin, G.Y. Li, B.-H. Juang, Deep learning enabled semantic communication systems, IEEE Trans. Signal Process. 69 (2020) 2663–2675, URL https://api.semanticscholar.org/CorpusID:219792180.

[85] Y. Hua, J. Du, Deep semantic correlation with adversarial learning for cross-modal retrieval, in: 2019 IEEE 9th International Conference on Electronics Information and Emergency Communication, ICEIEC, 2019, pp. 256–259, http://dx.doi.org/10.1109/ICEIEC.2019.8784597.

[86] H. Zhao, H. Li, D. Xu, S. Song, K.B. Letaief, Multi-modal self-supervised semantic communication, 2025, arXiv:2503.13940, URL https://arxiv.org/abs/2503.13940.

[87] H. Chen, S. Chen, C. He, H. Li, A scalable semantic communication system based on meta-learning, in: 2024 International Wireless Communications and Mobile Computing, IWCMC, 2024, pp. 561–566, http://dx.doi.org/10.1109/IWCMC61514.2024.10592533.

[88] Y. Bian, X. Zhang, G. Luosang, D. Renzeng, D. Renqing, X. Ding, Federated learning and semantic communication for the metaverse: Challenges and potential solutions, Electronics 14 (5) (2025) http://dx.doi.org/10.3390/electronics14050868, URL https://www.mdpi.com/2079-9292/14/5/868.

[89] L.X. Nguyen, K. Kim, Y. Lin Tun, S. Salman Hassan, Y. Kyaw Tun, Z. Han, C. Seon Hong, Optimizing multi-user semantic communication via transfer learning and knowledge distillation, IEEE Commun. Lett. 29 (1) (2025) 90–94, http://dx.doi.org/10.1109/lcomm.2024.3499956.

[90] J. Mao, K. Xiong, M. Liu, Z. Qin, W. Chen, P. Fan, K.B. Letaief, A GAN-based semantic communication for text without CSI, IEEE Trans. Wirel. Commun. 23 (10) (2024) 14498–14514, http://dx.doi.org/10.1109/TWC.2024.3415363.

[91] K. Zhang, L. Li, W. Lin, Y. Yan, R. Li, W. Cheng, Z. Han, Semantic successive refinement: A generative AI-aided semantic communication framework, IEEE Trans. Cogn. Commun. Netw. 11 (2) (2025) 687–699, http://dx.doi.org/10.1109/TCCN.2025.3526839.

[92] F. Jiang, C. Tang, L. Dong, K. Wang, K. Yang, C. Pan, Visual language model-based cross-modal semantic communication systems, IEEE Trans. Wirel. Commun. 24 (5) (2025) 3937–3948, http://dx.doi.org/10.1109/TWC.2025.3539526.

[93] C. Liang, D. Li, Joint source-channel noise adding with adaptive denoising for diffusion-based semantic communications, IEEE Internet Things J. (2025) http://dx.doi.org/10.1109/JIOT.2025.3589398, 1–1.

[94] C. Liang, D. Li, Image generation with supervised selection based on multimodal features for semantic communications, 2025, arXiv:2411.17428, URL https://arxiv.org/abs/2411.17428.

[95] C. Liang, D. Li, Z. Lin, H. Cao, Selection-based image generation for semantic communication systems, IEEE Commun. Lett. 28 (1) (2024) 34–38, http://dx.doi.org/10.1109/LCOMM.2023.3339534.

[96] A. Vettoruzzo, M.-R. Bouguelia, J. Vanschoren, T. Rognvaldsson, K. Santosh, Advances and challenges in meta-learning: A technical review, IEEE Trans. Pattern Anal. Mach. Intell. (2024) 1–20, http://dx.doi.org/10.1109/TPAMI.2024.3357847.

[97] C. Finn, P. Abbeel, S. Levine, Model-agnostic meta-learning for fast adaptation of deep networks, in: D. Precup, Y.W. Teh (Eds.), Proceedings of the 34th International Conference on Machine Learning, in: Proceedings of Machine Learning Research, vol. 70, PMLR, 2017, pp. 1126–1135, URL https://proceedings.mlr.press/v70/finn17a.html.

[98] A. Nichol, J. Achiam, J. Schulman, On first-order meta-learning algorithms, 2018, arXiv:1803.02999.

[99] D. Li, Y. Yang, Y.-Z. Song, T. Hospedales, Learning to generalize: Meta-learning for domain generalization, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 32, (1) 2018.

[100] Y. Balaji, S. Sankaranarayanan, R. Chellappa, Metareg: Towards domain generalization using meta-regularization, Adv. Neural Inf. Process. Syst. 31 (2018).

[101] T. Hospedales, A. Antoniou, P. Micaelli, A. Storkey, Meta-learning in neural networks: A survey, IEEE Trans. Pattern Anal. Mach. Intell. 44 (9) (2022) 5149–5169, http://dx.doi.org/10.1109/TPAMI.2021.3079209.

[102] H. Li, S.J. Pan, S. Wang, A.C. Kot, Domain generalization with adversarial feature learning, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2018.

[103] X. Jin, C. Lan, W. Zeng, Z. Chen, Re-energizing domain discriminator with sample relabeling for adversarial domain adaptation, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, ICCV, 2021, pp. 9174–9183.

[104] Y. Li, Y. Wu, Vision transformer-based adversarial domain adaptation, 2024, arXiv:2404.15817.

[105] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, V. Lempitsky, Domain-adversarial training of neural networks, in: Advances in Neural Information Processing Systems, 2016, pp. 2097–2105, URL https://papers.nips.cc/paper/2016/file/ab88b15790c24f6733e7d4d8c7b2a576-Paper.pdf.

[106] C. Franzreb, T. Polzehl, Domain adversarial training for german accented speech recognition, in: DAGA 2023 - 49. Jahrestagung FÜr Akustik. Deutsche Jahrestagung FÜr Akustik (DAGA-2023), 49., March 6-9, Hamburg, Germany, DEGA e.V, 2023, pp. 1413–1416.

[107] Z. Huang, H. Wang, E.P. Xing, D. Huang, Self-challenging improves cross-domain generalization, in: Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part II 16, Springer, 2020, pp. 124–140.

[108] Y. Shi, J. Seely, P. Torr, S. N., A. Hannun, N. Usunier, G. Synnaeve, Gradient matching for domain generalization, in: International Conference on Learning Representations, 2022, URL https://openreview.net/forum?id=vDwBW49HmO.

[109] A. Koloskova, H. Hendrikx, S.U. Stich, Revisiting gradient clipping: Stochastic bias and tight convergence guarantees, in: A. Krause, E. Brunskill, K. Cho, B. Engelhardt, S. Sabato, J. Scarlett (Eds.), Proceedings of the 40th International Conference on Machine Learning, in: Proceedings of Machine Learning Research, vol. 202, PMLR, 2023, pp. 17343–17363, URL https://proceedings.mlr.press/v202/koloskova23a.html.

[110] E. Hosseini, L. Reinhardt, D.B. Rawat, Optimizing gradient methods for IoT applications, IEEE Internet Things J. 9 (15) (2022) 13694–13704, http://dx.doi.org/10.1109/JIOT.2022.3142200.

[111] R. Chen, I.C. Paschalidis, Distributionally robust learning, Found. Trends® Optim. 4 (1–2) (2020) 1–243, http://dx.doi.org/10.1561/2400000026.

[112] R. Gao, A. Kleywegt, Distributionally robust stochastic optimization with wasserstein distance, Math. Oper. Res. 48 (2) (2022) 603–655, http://dx.doi.org/10.1287/moor.2022.1275.

[113] J. Liu, Z. Shen, P. Cui, L. Zhou, K. Kuang, B. Li, Distributionally robust learning with stable adversarial training, IEEE Trans. Knowl. Data Eng. 35 (11) (2023) 11288–11300, http://dx.doi.org/10.1109/TKDE.2022.3224056.

[114] D. Krueger, E. Caballero, J.-H. Jacobsen, A. Zhang, J. Binas, D. Zhang, R. Le Priol, A. Courville, Out-of-distribution generalization via risk extrapolation (rex), in: International Conference on Machine Learning, PMLR, 2021, pp. 5815–5826.

[115] E.Z. Liu, B. Haghgoo, A.S. Chen, A. Raghunathan, P.W. Koh, S. Sagawa, P. Liang, C. Finn, Just train twice: Improving group robustness without training group information, in: International Conference on Machine Learning, PMLR, 2021, pp. 6781–6792.

[116] S. Sagawa*, P.W. Koh*, T.B. Hashimoto, P. Liang, Distributionally robust neural networks, in: International Conference on Learning Representations, 2020, URL https://openreview.net/forum?id=ryxGuJrFvS.

[117] X. Liu, F. Zhang, Z. Hou, L. Mian, Z. Wang, J. Zhang, J. Tang, Self-supervised learning: Generative or contrastive, IEEE Trans. Knowl. Data Eng. 35 (1) (2023) 857–876, http://dx.doi.org/10.1109/TKDE.2021.3090866.

[118] K.R. Chowdhary, Natural language processing, in: Fundamentals of Artificial Intelligence, Springer India, New Delhi, 2020, pp. 603–649, http://dx.doi.org/10.1007/978-81-322-3972-7_19.

[119] L. Kryeziu, V. Shehu, Pre-training MLM using bert for the albanian language, SEEU Rev. 18 (1) (2023) 52–62.

[120] W. Shi, V. Demberg, Next sentence prediction helps implicit discourse relation classification within and across domains, in: K. Inui, J. Jiang, V. Ng, X. Wan (Eds.), Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP, Association for Computational Linguistics, Hong Kong, China, 2019, pp. 5790–5796, http://dx.doi.org/10.18653/v1/D19-1586, URL https://aclanthology.org/D19-1586.

[121] F.M. Carlucci, A. D'Innocente, S. Bucci, B. Caputo, T. Tommasi, Domain generalization by solving jigsaw puzzles, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 2229–2238.

[122] M. Noroozi, A. Vinjimoor, P. Favaro, H. Pirsiavash, Boosting self-supervised learning via knowledge transfer, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 9359–9367.

[123] D. Kim, Y. Yoo, S. Park, J. Kim, J. Lee, Selfreg: Self-supervised contrastive regularization for domain generalization, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 9619–9628.

[124] N. Park, M. Mohammadi, K. Gorde, S. Jajodia, H. Park, Y. Kim, Data synthesis based on generative adversarial networks, Proc. VLDB Endow. 11 (10) (2018) 1071–1083, http://dx.doi.org/10.14778/3231751.3231757.

[125] M. Sabuhi, M. Zhou, C.-P. Bezemer, P. Musilek, Applications of generative adversarial networks in anomaly detection: a systematic literature review, IEEE Access 9 (2021) 161003–161029.

[126] B. Zhu, J. Jiao, D. Tse, Deconstructing generative adversarial networks, IEEE Trans. Inform. Theory 66 (11) (2020) 7155–7179, http://dx.doi.org/10.1109/TIT.2020.2983698.

[127] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, X. Chen, Improved techniques for training GANs, in: Proceedings of the 30th International Conference on Neural Information Processing Systems, NIPS '16, Curran Associates Inc., Red Hook, NY, USA, 2016, pp. 2234–2242.

[128] A. Radford, J.W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, et al., Learning transferable visual models from natural language supervision, in: International Conference on Machine Learning, PMLR, 2021, pp. 8748–8763.

[129] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, T. Aila, Analyzing and improving the image quality of stylegan, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 8110–8119.

[130] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, H. Greenspan, GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification, Neurocomputing 321 (2018) 321–331.

[131] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, A.A. Bharath, Generative adversarial networks: An overview, IEEE Signal Process. Mag. 35 (1) (2018) 53–65.

[132] G. Kunapuli, Ensemble Methods for Machine Learning, Simon and Schuster, 2023.

[133] T.G. Dietterich, Ensemble methods in machine learning, in: Multiple Classifier Systems, Springer Berlin Heidelberg, Berlin, Heidelberg, 2000, pp. 1–15.

[134] Y. Zhang, J. Liu, W. Shen, A review of ensemble learning algorithms used in remote sensing applications, Appl. Sci. 12 (17) (2022) http://dx.doi.org/10.3390/app12178654, URL https://www.mdpi.com/2076-3417/12/17/8654.

[135] A. D'Innocente, B. Caputo, Domain generalization with domain-specific aggregation modules, in: Pattern Recognition: 40th German Conference, GCPR 2018, Stuttgart, Germany, October 9-12, 2018, Proceedings 40, Springer, 2019, pp. 187–198.

[136] K. Zhou, Y. Yang, Y. Qiao, T. Xiang, Domain adaptive ensemble learning, IEEE Trans. Image Process. 30 (2021) 8008–8018.

[137] G. Wu, S. Gong, Collaborative optimization and aggregation for decentralized domain generalization and adaptation, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 6484–6493.

[138] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, Q. He, A comprehensive survey on transfer learning, Proc. IEEE 109 (1) (2021) 43–76, http://dx.doi.org/10.1109/JPROC.2020.3004555.

[139] J. Howard, S. Ruder, Universal language model fine-tuning for text classification, 2018, arXiv:1801.06146.

[140] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, 2018, arXiv preprint arXiv:1810.04805.

[141] C. De Alwis, P. Kumar, Q.-V. Pham, K. Dev, A. Kalla, M. Liyanage, W.-J. Hwang, Towards 6G: Key technological directions, ICT Express 9 (4) (2023) 525–533, http://dx.doi.org/10.1016/j.icte.2022.10.005, URL https://www.sciencedirect.com/science/article/pii/S2405959522001485.

[142] P. Soldati, E. Ghadimi, B. Demirel, Y. Wang, R. Gaigalas, M. Sintorn, Design principles for model generalization and scalable AI integration in radio access networks, IEEE Commun. Mag. 63 (1) (2025) 84–91, http://dx.doi.org/10.1109/MCOM.001.2300812.

[143] N. Janbi, I. Katib, R. Mehmood, Distributed artificial intelligence: Taxonomy, review, framework, and reference architecture, Intell. Syst. Appl. 18 (2023) 200231, http://dx.doi.org/10.1016/j.iswa.2023.200231, URL https://www.sciencedirect.com/science/article/pii/S266730532300056X.

[144] C. Mwase, Y. Jin, T. Westerlund, H. Tenhunen, Z. Zou, Communication-efficient distributed AI strategies for the IoT edge, Future Gener. Comput. Syst. 131 (2022) 292–308, http://dx.doi.org/10.1016/j.future.2022.01.013, URL https://www.sciencedirect.com/science/article/pii/S0167739X22000218.

[145] X. Kong, Y. Wu, H. Wang, F. Xia, Edge computing for internet of everything: A survey, IEEE Internet Things J. PP (2022) http://dx.doi.org/10.1109/JIOT.2022.3200431, 1–1.

[146] H. Wu, X. Li, Y. Deng, Deep learning-driven wireless communication for edge-cloud computing: opportunities and challenges, J. Cloud Comput. 9 (21) (2020) http://dx.doi.org/10.1186/s13677-020-00168-9.

[147] R. Roman, J. Lopez, M. Mambo, Mobile edge computing, fog et al.: A survey and analysis of security threats and challenges, Future Gener. Comput. Syst. 78 (2018) 680–698.

[148] S.H. Wanasekara, V.-D. Nguyen, K.-S. Wong, M.-D. Nguyen, S. Chatzinotas, O.A. Dobre, SC-GIR: Goal-oriented semantic communication via invariant representation learning for image transmission, IEEE Trans. Mob. Comput. (2025) 1–15, http://dx.doi.org/10.1109/tmc.2025.3600434.

[149] Z. Yang, M. Chen, Z. Zhang, C. Huang, Energy efficient semantic communication over wireless networks with rate splitting, IEEE J. Sel. Areas Commun. 41 (5) (2023) 1484–1495, http://dx.doi.org/10.1109/JSAC.2023.3240713.

[150] 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 J. Sel. Areas Commun. 41 (1) (2023) 170–185, http://dx.doi.org/10.1109/JSAC.2022.3221991.

[151] Z. Sun, Z. Shen, L. Lin, Y. Yu, Z. Yang, S. Yang, W. Chen, Dynamic domain generalization, in: Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, (IJCAI-22, IJCAI, Vienna, Austria, 2022, pp. 2449–2455.

[152] H. Xie, Z. Qin, A lite distributed semantic communication system for Internet of Things, IEEE J. Sel. Areas Commun. 39 (1) (2020) 142–153.

[153] 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 J. Sel. Areas Commun. 41 (1) (2022) 170–185.

[154] L. Zhang, Adversarial Learning for Wireless Communications, Loughborough University Repository, 2023, http://dx.doi.org/10.26174/thesis.lboro.22122518.v1, URL https://repository.lboro.ac.uk/articles/thesis/Adversarial_learning_for_wireless_communications/22122518.

[155] X. Luo, Z. Chen, M. Tao, F. Yang, Encrypted semantic communication using adversarial training for privacy preserving, IEEE Commun. Lett. 27 (6) (2023) 1486–1490.

[156] L.X. Nguyen, K. Kim, Y. Lin Tun, S. Salman Hassan, Y. Kyaw Tun, Z. Han, C. Seon Hong, Optimizing multi-user semantic communication via transfer learning and knowledge distillation, IEEE Commun. Lett. 29 (1) (2025) 90–94, http://dx.doi.org/10.1109/LCOMM.2024.3499956.

[157] J. Pei, C. Feng, P. Wang, H. Tabassum, D. Shi, Latent diffusion model-enabled low-latency semantic communication in the presence of semantic ambiguities and wireless channel noises, IEEE Trans. Wirel. Commun. 24 (5) (2025) 4055–4072, http://dx.doi.org/10.1109/TWC.2025.3535714.

[158] H. Wei, W. Wang, W. Ni, W. Xu, Y. Huang, D. Niyato, P. Zhang, Task-agnostic semantic communications relying on information bottleneck and federated meta-learning, 2025, arXiv:2504.21723, URL https://arxiv.org/abs/2504.21723.

[159] W. Zhu, H. Liu, Q. Dong, J. Xu, L. Kong, J. Chen, L. Li, S. Huang, Multilingual machine translation with large language models: Empirical results and analysis, 2023, Journal Name Volume Number (Issue Number), Page Numbers.

[160] S. Ruder, A. Søgaard, I. Vulić, Unsupervised cross-lingual representation learning, in: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: Tutorial Abstracts, 2019, pp. 31–38.

[161] C. Finn, P. Abbeel, S. Levine, Model-agnostic meta-learning for fast adaptation of deep networks, in: International Conference on Machine Learning, PMLR, 2017, pp. 1126–1135.

[162] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, BERT: Pre-training of deep bidirectional transformers for language understanding, in: North American Chapter of the Association for Computational Linguistics, 2019, URL https://api.semanticscholar.org/CorpusID:52967399.

[163] M. Artetxe, H. Schwenk, Massively multilingual sentence embeddings for zero-shot cross-lingual transfer and beyond, Trans. Assoc. Comput. Linguist. 7 (2019) 597–610.

[164] T. Sherborne, T. Hosking, M. Lapata, Optimal transport posterior alignment for cross-lingual semantic parsing, Trans. Assoc. Comput. Linguist. 11 (2023) 1432–1450, http://dx.doi.org/10.1162/tacl_a_00611, arXiv:https://direct.mit.edu/tacl/article-pdf/doi/10.1162/tacl_a_00611/2184055/tacl_a_00611.pdf, URL https://doi.org/10.1162/tacl_a_00611.

[165] L. Zhang, M. Das, Y. Sun, D. Niyato, X. Yuan, Prompt-based transceiver cooperation for semantic communications with domain-incremental background knowledge, in: GLOBECOM 2023 - 2023 IEEE Global Communications Conference, 2023, pp. 2087–2092, http://dx.doi.org/10.1109/GLOBECOM54140.2023.10437903.

[166] N. Lin, W. Zhao, S. Liang, M. Zhong, Real-time segmentation of unstructured environments by combining domain generalization and attention mechanisms, Sensors 23 (13) (2023) 6008.

[167] G. Gupta, R. Kapila, K. Gupta, R. Raskar, Domain generalization in robust invariant representation, 2023, arXiv preprint arXiv:2304.03431.

[168] A. Dayal, V. K.B., L.R. Cenkeramaddi, C. Mohan, A. Kumar, V. N. Balasubramanian, MADG: margin-based adversarial learning for domain generalization, Adv. Neural Inf. Process. Syst. 36 (2024).

[169] Y. Wang, Y. Pan, Z. Su, Y. Deng, Q. Zhao, L. Du, T.H. Luan, J. Kang, D. Niyato, Large model based agents: State-of-the-art, cooperation paradigms, security and privacy, and future trends, IEEE Commun. Surv. Tutor. (2025) http://dx.doi.org/10.1109/COMST.2025.3576176, 1–1.

[170] J. Niemeijer, M. Schwonberg, J.-A. Termöhlen, N.M. Schmidt, T. Fingscheidt, Generalization by adaptation: Diffusion-based domain extension for domain-generalized semantic segmentation, in: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2024, pp. 2830–2840.

[171] C. Xu, X. Tian, Semantic-aware mixup for domain generalization, in: 2023 International Joint Conference on Neural Networks, IJCNN, IEEE, 2023, pp. 1–8.

[172] T. Baltrušaitis, C. Ahuja, L.-P. Morency, Multimodal machine learning: A survey and taxonomy, IEEE Trans. Pattern Anal. Mach. Intell. 41 (2) (2018) 423–443.

[173] Y. Wang, W. Yang, Z. Xiong, Y. Zhao, S. Mao, T.Q.S. Quek, H.V. Poor, FAST-GSC: Fast and adaptive semantic transmission for generative semantic communication, 2024, arXiv:2407.15395, URL https://arxiv.org/abs/2407.15395.

[174] M. Zhang, M. Abdi, V.R. Dasari, F. Restuccia, Semantic edge computing and semantic communications in 6G networks: A unifying survey and research challenges, Comput. Netw. 270 (2025) 111531, http://dx.doi.org/10.1016/j.comnet.2025.111531.

[175] 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 Commun. Surv. Tutor. (2024).

[176] V.-T. Hoang, V.-L. Nguyen, R.-G. Chang, P.-C. Lin, R.-H. Hwang, T.Q. Duong, Adversarial attacks against shared knowledge interpretation in semantic communications, IEEE Trans. Cogn. Commun. Netw. (2025).

[177] H. Liu, B. Wang, R. Meng, S. Han, X. Xu, Adaptive privacy budget-based differential privacy co-training for wireless semantic communication, in: 2024 IEEE Wireless Communications and Networking Conference, WCNC, IEEE, 2024, pp. 1–6.

[178] M.A.P. Putra, N. Karna, S.N. Hertiana, A. Zainudin, R.N. Alief, G.A. Sampedro, D.-S. Kim, J.-M. Lee, SECC: A secure and efficient communication channel for federated learning system, in: 2024 15th International Conference on Information and Communication Technology Convergence, ICTC, IEEE, 2024, pp. 2099–2104.

[179] Y.E. Sagduyu, T. Erpek, S. Ulukus, A. Yener, Is semantic communication secure? A tale of multi-domain adversarial attacks, IEEE Commun. Mag. 61 (11) (2023) 50–55, http://dx.doi.org/10.1109/MCOM.006.2200878.

Muhammad Furqan Zia received his Bachelors of Engineering degree in Electrical Engineering from DHA Suffa University, Karachi, Pakistan, in 2017, and his Master of Science degree in Electrical and Computer Engineering from Antalya Bilim University, Turkey, in 2021. He is currently pursuing a Ph.D. in Electrical Engineering at the University of Quebec at Trois-Riviéres (UQTR), Canada, after transferring his doctoral studies from Koç University, Turkey, in the summer of 2023. From 2018 to 2021, he served as a Research Assistant at the Wireless Intelligent Systems Laboratory, Antalya Bilim University, and subsequently as a Graduate Research and Teaching Assistant at Koç University from 2021 to 2023.

During his academic career, he has received several prestigious grants and scholarships, including a master's scholarship and an industrial Ph.D. scholarship from the Scientific and Technological Research Council of Turkey (TÜBİTAK), in collaboration with VESTEL, Turkey. He was also awarded an Outstanding Graduate Research Performance Scholarship under the TÜBİTAK 2250 program at Koç University in 2022–2023. His current Ph.D. research at UQTR is funded by the Natural Sciences and Engineering Research Council of Canada (NSERC).

He has authored several peer-reviewed articles and patent applications, contributing significantly to his field. His research interests include trustworthy and robust AI for intelligent wireless communication, multiple-input multiple-output (MIMO), AI applications in open radio access networks (Open RAN), semantic communication, energy-efficient network design, and explainable AI and domain generalization, with a focus on enhancing the transparency and adaptability of wireless technologies across diverse and unseen environments.

Messaoud Ahmed Ouameur (Senior Member, IEEE) received a bachelors degree in electrical engineering from the Institute national d'électronique et d'électricité (INELEC), Boumerdes, Algeria, in 1998, the M.B.A. degree from the Graduate School of International Studies, Ajou University, Suwon, South Korea, in 2000, and the masters and Ph.D. degrees (Hons.) in electrical engineering from the Université du Québec à Trois-Rivières (UQTR), QC, Canada, in 2002 and 2006, respectively. Before joining UQTR in 2018, Dr. Ahmed Ouameur was an radio system expert at NUTAQ innovation, Quebec, Canada. Therein, he was a major system architect and design engineer for several software defined radio development systems including Titan MIMO, a 23-by-32 massive MIMO text bed. As a professor at UQTR, his research interests include embedded real-time systems, parallel and distributed processing with applications to distributed Massive MIMO, reconfigurable intelligent surfaces, deep learning and machine learning for communication systems, and the Internet of Things with an emphasis on embedded AI using approximate computing.

Miloud Bagaa (Senior Member, IEEE) received the Engineering, master's, and Ph.D. degrees from the University of Science and Technology Houari Boumediene, Algiers, Algeria, in 2005, 2008, and 2014, respectively. From 2009 to 2015, he was a Researcher with the Research Center on Scientific and Technical Information, Algiers. From 2015 to 2016, he was a Postdoctoral Researcher with the Norwegian University of Science and Technology, Trondheim, Norway. From 2016 to 2019, he was a Postdoctoral Researcher with Aalto University, where he was a Senior Researcher, from 2019 to 2020. From October 2020 to November 2022, he was a Visiting Researcher with Aalto University and a Senior System Cloud Specialist with CSC. He is currently a Professor with the Department of Electrical and Computer Engineering, Université du Québec à Trois-Rivières (UQTR).

Daniel Massicotte (S'91–M'94–SM'08) (Senior Member, IEEE) received the B.Sc.A. and M.Sc.A. degrees in electrical engineering and industrial electronics in 1987 and 1990 respectively from the Université du Québec à Trois-Rivières (UQTR), Qc, Canada. He obtained the Ph.D. degree in electrical engineering in 1995 at the École Polytechnique de Montréal, Qc, Canada. In 1994, he joined the Department of Electrical and Computer Engineering, Université du Québec à Trois-Rivières, where he is currently a Full Professor. He is Founder of the Laboratory of Signal and Systems Integration. Since 2001, he has been Founding President and Chief Technology Officer of Axiocom Inc. He has been Head of the Industrial Electronic Research Group from 2011 to 2018 and Head of the Electrical and Computer Engineering Department from 2014 to 2020. Head of the Research Chair in Signals and Intelligence of High Performance Systems, Trois-Rivières, since 2018. He has proposed many methods based on modern signal processing, such as machine learning, transform domain, and metaheuristics. He is the author/coauthor of more than 200 technical papers in international conferences and journals and 9 inventions. His research interests include advanced VLSI implementation, digital signal processing for wireless communications, measurement, medical, real-time simulation, and control problems for linear/nonlinear complex systems. He has proposed many methods based on modern signal and biosignal processing,such as machine learning, transform domain, and metaheuristics.

Dr. Massicotte is also a member of the Ordre des Ingénieurs du Québec and the Microsystems Strategic Alliance of Quebec (ReSMiQ). He received the Douglas R. Colton Medal for Research Excellence awarded by the Canadian Microelectronics Corporation, the PMC-Sierra High Speed Networking and Communication Award, and the Second Place at the Complex Multimedia/Telecom IP Design Contest from Europractice. He was the General Chair of IEEE New Circuits and Systems (NEWCAS) 2014.

Adlen Ksentini (Senior Member, IEEE) is a professor in the Communication Systems Department of EURECOM. He is leading the Network softwarization group activities related to Network softwarization, 5G/6G, and Edge Computing. Adlen Ksentini's research interests are Network Sofwerization and Network Cloudification, focusing on topics related to network virtualization, Software Defined Networking (SDN), and Edge Computing for 5G and 6G networks. He has been participating to several H2020 and Horizon Europe projects on 5G and beyond, such as 5G!Pagoda, 5GTransformer, 5G!Drones, MonB5G, ImagineB5G, 6GBricks, 6G-Intense, Sunrise-6G and AC3. He is the technical manager of 6GIntense and AC3, on zero-touch management of 6G resources and applications, and Cloud Edge Continuum, respectively. He is interested in the system and architectural issues but also in algorithm problems related to those topics, using Markov Chains, Optimization algorithms, and Machine Learning (ML). Adlen Ksentini has given several tutorials in IEEE international conferences, IEEE Globecom 2015, IEEEE CCNC 2017/2018/2023, IEEE ICC 2017, IEEE/IFIP p=IM 2017, IEEE School 2019. Adlen Ksentini is a member of the OAI board of directors, where he is in charge of OAI 5G Core Network and ORAN management (O1, E2) for OAI RAN activity.