NANCY

An Artificial Intelligent Aided Unified Network for Secure Beyond 5G Long Term Evolution
[GA: 101096456]

Deliverable 4.4

Semantic & goal-oriented communication schemes for beyond Shannon performance

Programme: HORIZON-JU-SNS-2022-STREAM-A-01-06

Start Date: 01 January 2023

Duration: 36 Months

Co-funded by the European Union

6GSNS

NANCY project has received funding from the Smart Networks and Services Joint Undertaking (SNS JU) under the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101096456.

Document Control Page

Deliverable Name	Semantic & goal-oriented communication schemes for beyond Shannon performance
Deliverable Number	D4.4 Semantic & goal-oriented communication schemes for beyond Shannon performance
Work Package	WP4
Associated Task	T4.4 Semantic & goal-oriented communication schemes for beyond Shannon performance
Dissemination Level	Public
Due Date	28 February 2025 (M26)
Completion Date	28 February 2025
Submission Date	28 February 2025
Deliverable Lead Partner	INNO
Deliverable Author(s)	Stylianos Trevlakis (INNO), Lamprini Mitsiou (INNO), Vasileios Kouvakis (INNO), Eirini Gkarnetidou (INNO), Theodoros Tsiftsis (INNO), Nikolaos Ntampakis (MINDS), Vasileios Gavresis (MINDS), Ioannis Makris (MINDS), Dimitrios Balompas (MINDS), Georgios Tsoumplekas (MINDS), Panagiotis Sarigiannidis (UOWM), Thomas Lagkas (UOWM), Dimitrios Pliatsios (UOWM), Athanasios Liatifis (UOWM), Sotirios Tegos (UOWM), Anna Triantafyllou (UOWM)
Version	1.0

Document History

Version	Date	Change History	Author(s)	Organisation
0.1	12/11/2024	ToC	Stylianos Trevlakis, Lamprini Mitsiou, Vasileios Kouvakis, Theodoros Tsiftsis	INNO
0.2	29/11/2024	Section 2 & 3	Stylianos Trevlakis, Lamprini Mitsiou, Vasileios Kouvakis, Theodoros Tsiftsis	INNO
0.3	14/12/2024	Section 4 & 5	Stylianos Trevlakis, Lamprini Mitsiou, Vasileios Kouvakis, Theodoros Tsiftsis	INNO

0.4	27/12/2024	Section 6	Stylianos Trevlakis, Lamprini Mitsiou, Vasileios Kouvakis, Theodoros Tsiftsis	INNO
0.5	18/01/2025	Section 7	Stylianos Trevlakis, Lamprini Mitsiou, Vasileios Kouvakis, Eirini Gkarnetidou, Theodoros Tsiftsis	INNO
0.5	24/01/2025	Section 1	Stylianos Trevlakis, Lamprini Mitsiou, Vasileios Kouvakis, Eirini Gkarnetidou, Theodoros Tsiftsis	INNO
0.6	31/01/2025	Section 8	Stylianos Trevlakis, Lamprini Mitsiou, Vasileios Kouvakis, Eirini Gkarnetidou, Theodoros Tsiftsis	INNO
0.7	05/02/2025	Editing and first complete draft	Stylianos Trevlakis, Lamprini Mitsiou, Vasileios Kouvakis, Eirini Gkarnetidou, Theodoros Tsiftsis	INNO
0.8	06/02/2025	Editing and general review	Ramon Sanchez-Iborra	UMU
	14/02/2025	Editing and review	Nikolaos Ntampakis, Vasileios Gavresis, Ioannis Makris, Dimitrios Balompas, Georgios Tsoumplekas	MINDS
	16/02/2025	Editing and review	Panagiotis Sarigiannidis, Thomas Lagkas, Dimitrios Pliatsios, Athanasios Liatifis, Sotirios Tegos, Anna Triantafyllou	UOWM
0.9	20/02/2025	Addressing internal reviewer comments	Stylianos Trevlakis, Lamprini Mitsiou, Vasileios Kouvakis, Eirini Gkarnetidou, Theodoros Tsiftsis	INNO
1.0	26/02/2025	Final quality revisions	Anna Triantafyllou, Dimitrios Pliatsios	UOWM

Internal Review History

Name	Organisation	Date
Dimitrios Pliatsios	UOWM	17 February 2025
Nikolaos Ntampakis	MINDS	14 February 2025

Quality Manager Revision

Name	Organisation	Date
Anna Triantafyllou, Dimitrios Pliatsios	UOWM	26 February 2025

Table of Contents

List of Figures

List of Tables

List of Acronyms

Acronym	Explanation
3D	Three-Dimensional
AI	Artificial Intelligence
AoI	Age of Information
AoII	Age of Incorrect Information
AoL	Age of Loop
ASL	American Sign Language
B5G	Beyond 5G
BER	Bit-Error-Rate
CNN	Convolutional Neural Network
DL	Deep learning
DNNs	Deep Neural Networks
DT	Digital Twin
E2E	End-to-End
HMMs	Hidden Markov Models
HTTP	Hypertext Transfer Protocol
IoT	Internet of Things
ISO	International Organization of Standardization
ITU	International Telecommunication Union
KNN	k-Nearest Neighbor
KPI	Key Performance Indicator
LSTM	Long Short-Term Memory
MAC	Medium Access Control
MIMO	Multiple-Input Multiple-Output
ML	Machine Learning
QAM	Quadrature Amplitude Modulation
RGB	Red-Green-Blue
RSU	Roadside Unit
SemCom	Semantic Communications
SER	Symbol-Error-Rate
SGD	Stochastic Gradient Descent
SLR	Sign Language Recognition
SNR	Signal-to-Noise Ratio
SSD	Single Shot Detector
THz	Terahertz
UML	Unified Modeling Language

UoI	Urgency of Information
URLLC	Ultra Reliable Low Latency Communications
V2I	Vehicle-to-Infrastructure
V2P	Vehicle-to-Pedestrian
V2V	Vehicle-to-Vehicle
V2x	Vehicle-to-Everything
VoI	Value of Information
WNCSs	Wireless Networked Control Systems
YOLO	You Only Look Once

Executive Summary

This deliverable explores the integration of Semantic Communications (SemCom) in NANCY and towards future 6G networks. It emphasizes the paradigm shift from traditional communications to a more context-aware and goal-driven approach, where the transmission of important information takes precedence over simple data transfer. It investigates the role of SemCom in supporting the complex, intelligent interactions envisioned for future network generations; therefore, enabling a transition from "connecting devices" to "connecting intelligence".

"D4.4 - Semantic & goal-oriented communication schemes for beyond Shannon performance" delves into the essential role of semantic knowledge in SemCom systems, exploring concepts such as shared knowledge bases, ontologies, and knowledge graphs. These elements enable machines to understand and interpret the semantics contained within the transmitted data; thus, laying the groundwork for efficient semantic communications. Additionally, the critical aspect of timeliness in SemCom is addressed, particularly for applications with stringent latency requirements. The document discusses various metrics like Age of Information (AoI), Value of Information (VoI), and Urgency of Information (UoI) to evaluate and optimize the time-sensitive aspects of semantic data transmission.

Furthermore, "D4.4 – Semantic & goal-oriented communication schemes for beyond Shannon performance" introduces a natively intelligent semantic architecture, encompassing both point-to-point and networking modalities. The point-to-point semantic architecture highlights the extraction of semantic information at the source and its accurate interpretation at the destination, utilizing shared knowledge as well as semantic coding techniques. The semantic networking architecture, building upon the point-to-point one, expands the scope to encompass multi-user, distributed scenarios, involving the edge-to-cloud continuum. The realization of both necessitates new knowledge extraction frameworks, theoretical models for managing time-varying traffic flows, and novel semantic-aware metrics for performance evaluation.

Finally, "D4.4 – Semantic & goal-oriented communication schemes for beyond Shannon performance" showcases goal-oriented communication schemes based on SemCom; thus, providing concrete examples of successful implementations across various domains. Two specific use cases are presented: (i) an American Sign Language (ASL) transmission system and (ii) an enhanced Vehicle-to-Everything (V2x) communications digital twin. These scenarios demonstrate the practical benefits of SemCom in NANCY as well as in future 6G networks, including improved accessibility for hearing-impaired individuals and optimized data transmission for intelligent transportation systems. This deliverable concludes by highlighting the transformative potential of SemCom in shaping the future of connected systems by driving efficiency and facilitating meaningful interactions between intelligent agents.

1. Introduction

1.1. Purpose of the Document

This document, "D4.4 – Semantic & goal-oriented communication schemes for beyond Shannon performance", has explored the emerging role of SemCom in shaping the future of 6G networks. SemCom represents a fundamental shift from traditional communication paradigms that focus solely on transmitting bits. Instead, SemCom prioritizes conveying only the necessary information and achieving specific communication goals, leading to more efficient and intelligent communication systems.

1.2. Relation to other Tasks and Deliverables

This deliverable builds upon previous research and findings within the NANCY project as follows:

- "D2.1 – NANCY Requirements Analysis" provides a strong foundation for the construction of the SemCom framework in terms of the NANCY's vision, use cases, usage scenarios, and requirements.
- "D3.1 – NANCY Architecture Design" describes the NANCY architecture along with details about its various components. A preliminary analysis of the SemCom approaches of NANCY is included in D3.1.
- "D4.1 – Computational Offloading and User-centric Caching" provides insights into the task offloading techniques that were developed in the NANCY project. Since SemCom interacts with the task offloading schemes, some alignment between the two components was necessary to establish the logic and flow of this interaction.
- Finally, "D6.1 – B-RAN and 5G End-to-end Facilities Setup" is focused on the integration between various NANCY components. This includes the integration point between SemCom and task offloading. To ensure smooth integration, it was necessary to align the content of deliverables D4.1, D4.4, and D6.1.

1.3. Structure of the Document

The rest of the document is structured as follows:

- **Section 2 – State of the Art Overview** provides a comprehensive analysis of the current state-of-the-art of SemCom. It covers recent advancements, key concepts, challenges, and the integration of SemCom within the context of 6G networks.
- **Section 3 – Technology Enablers for Semantic Communications** delves into the specific technologies that are driving the development of SemCom, focusing on how these technologies address the limitations of current 5G networks and pave the way for the realization of 6G systems. It highlights the role of artificial intelligence, edge computing, and other enabling technologies in shaping the future of SemCom.

- **Section 4 – Semantic Knowledge** explores the fundamental aspects of semantic knowledge representation and its significance in SemCom systems. It delves into the concepts of shared knowledge bases, ontologies, and knowledge graphs as tools for facilitating machine understanding and interpretation of transmitted data.
- **Section 5 –Timeliness Aspects** focuses on the crucial aspect of timeliness in SemCom, especially for applications with strict latency requirements. It examines various metrics like AoI, VoI, and UoI, providing insights into how these metrics can be used to evaluate and optimize time-sensitive aspects of semantic communication.
- **Section 6 – Natively Intelligent Semantic Architecture** introduces a novel semantic architecture designed to support intelligent communication in future networks. It presents both point-to-point and networking paradigms, highlighting the intelligent functionalities incorporated at both the transmitter and receiver sides.
- **Section 7 – Goal-oriented Communication Schemes** showcases practical applications and examples of SemCom systems, emphasizing their effectiveness in achieving specific communication goals. It explores case studies involving ASL transmission and DT creation for V2x communications, demonstrating the tangible benefits of adopting a SemCom approach.
- **Section 8 – Conclusion and Outlook** summarizes the key findings and insights of T4.4, providing concluding remarks on the role and potential of SemCom in future 6G networks. It also presents an outlook on future research directions and areas for further exploration.

2. State of the Art Overview

During the last decade, the development, standardization, and deployment efforts for 5G have been solidified and the research interest has been directed towards future communication networks. Specifically, several vision papers have been published that documented the goals and requirements of 6G networks, while research planning and standardization activities have been initiated [1, 2, 3, 4, 5, 6]. Among the core requirements is energy efficiency to extend the battery of mobile nodes and reduce the drastically increasing energy footprint. Moreover, the various services that are envisioned to be implemented in 6G networks, like extended reality (XR) applications, airborne networks, remote driving and more, are entangled with use-case-specific requirements that must be accounted for [7]. Thus, the requirement to move away from the classic Shannon's theory is ever more demanding [8].

In working toward this goal, we revisit the ideas of Shannon and Weaver, who identified three main communication layers: technical, semantic, and effectiveness [9]. The technical layer addresses the accurate transmission of symbols, the semantic layer ensures the message is precisely conveyed, and the effectiveness layer aims for maximum efficiency in delivering that message. Although Shannon's emphasis on the technical side led to the development of a probabilistic mathematical theory of communications, today's vision of intelligent networks is tightly connected to semantics, now viewed as a critical element of next-generation communications theory.

2.1. Semantic Communications

Over the last couple of years, the development of an applicable SemCom architecture has gained increasing attention. This research effort has identified various challenges that need to be resolved for this novel system-level evolution of networks to be applied in the real-world [10, 11]. Ideally, a SemCom system should include all three layers (technical, semantic and effectiveness) presented in Figure 1. The technical layer answers the question "How accurately can the symbols of communication be transmitted?". This question has been the main motivation of conventional communication research since the 1940s [9], [12]. As a matter of fact, even the question posed by the semantic layer, "How precisely do the transmitted symbols convey the desired meaning?", is not a completely new idea and can be traced back to the introduction of semantic information theory in 1952 [13]. Afterwards, the theoretical framework for compression of semantic data and reliable semantic communications was introduced [14], while, in the recent years, a plethora of applications for SemCom have been identified, including image [15, 16], sound [17] and text transmission [18, 19, 20]. Finally, the effectiveness layer aims to answer the question "How effectively does the received meaning affect conduct in the desired way?". This layer handles the orchestration of the previous two to fulfil the required key performance indicators (KPIs) as well as optimize the utilized resources. Therefore, the effectiveness layer can access the technical and semantic layers directly in order to ensure the intelligent allocation of computational and network resources even in applications where semantic data may be irrelevant.

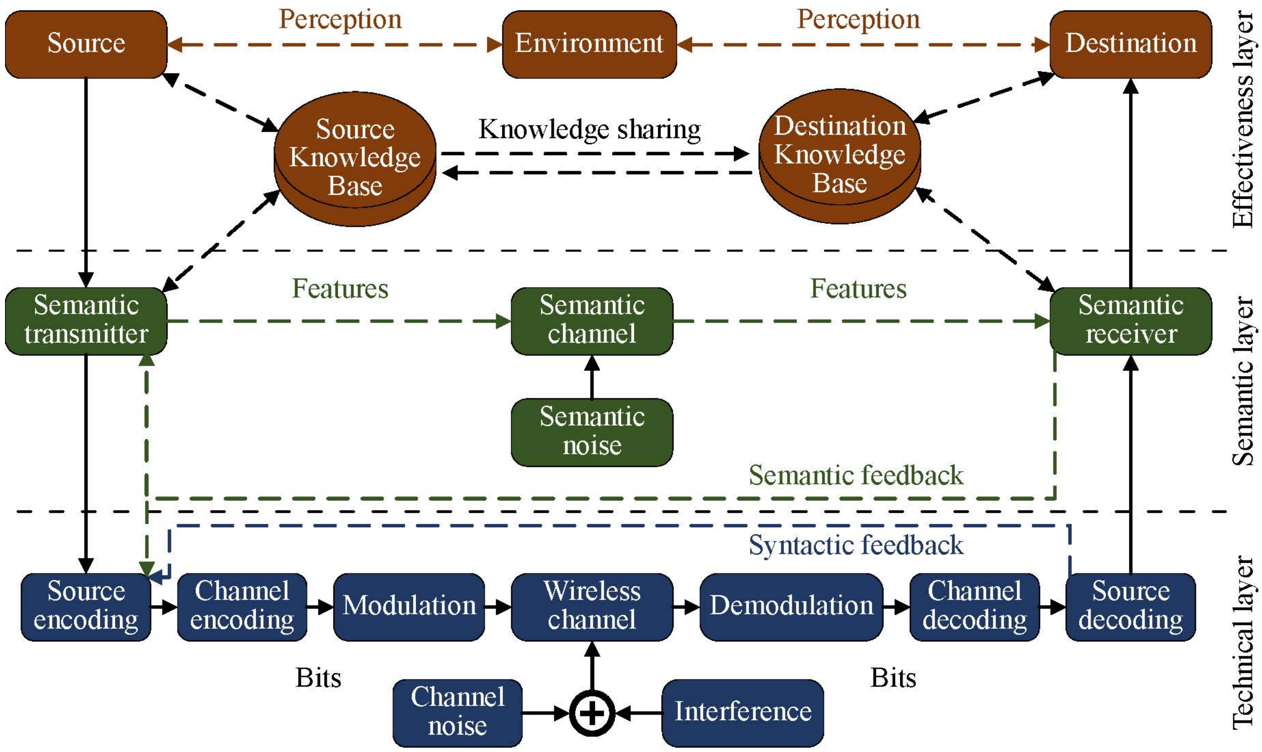

Figure 1. High-level SemCom architecture

The purpose of communication from an engineering point of view and that of SemCom are radically different. Specifically, as stated by Shannon, "*The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point*", while Weaver noted that "*The semantic problems are concerned with the interpretation of meaning by the receiver, as compared with the intended meaning of the sender*". By contrasting these two claims, we may say that the focus of SemCom is not on reproducing the transmitted messages word for word or even roughly, but on conveying the interpretations attached to those messages. Therefore, even though a message can be seen syntactically as a string of symbols, it is the expression that is constructed using those symbols. This significant difference between conventional and semantic communications is highlighted in the following commonplace example. Although the phrase "I have been recently elected Lecturer" can be transmitted without any syntactic errors to the destination, its interpretation may vary depending on the receiver's residency, since the term Lecturer is equivalent to that of an Assistant Professor in Europe, while in the US it represents a non-tenure position. Thus, it becomes evident that the semantic information source anticipates the destination to "understand" the transmitted messages. As illustrated in Figure 2, instead of blindly deciphering the message's syntax, the destination might use the information in the message and its own background knowledge to derive inferences.

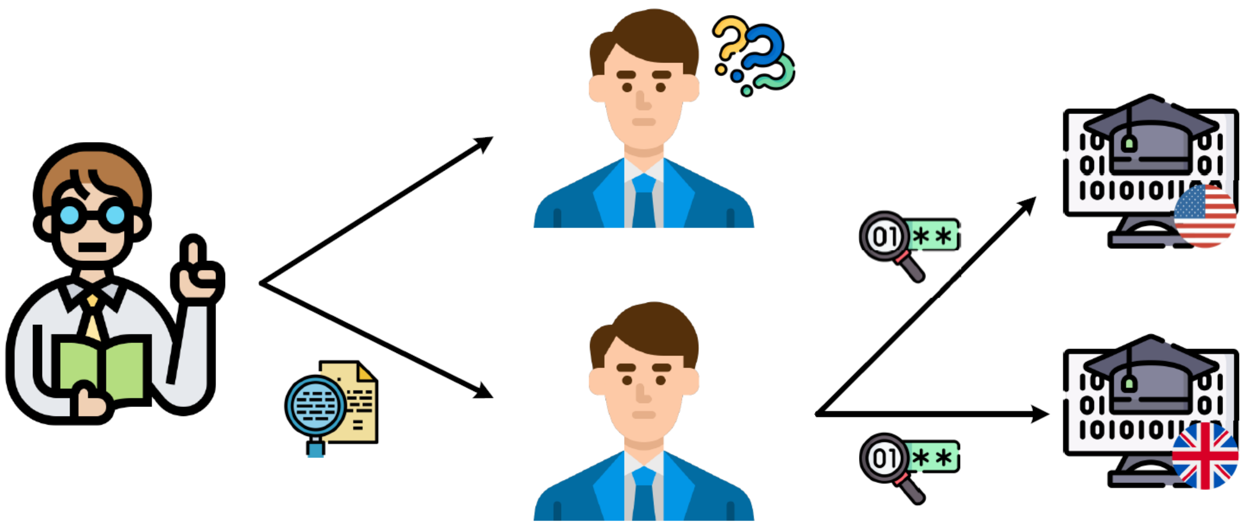

Figure 2. Illustration of the semantic interpretation problem

2.2. Related Work

By doing a thorough examination of the technical literature, one may identify several surveys and tutorials on semantic communications [10 – 31]. These works are presented in Table 1. Specifically, in [10], a survey of existing literature is presented covering the human-machine SemCom paradigm and highlighting the prospects of this field. It categorizes SemCom into three areas, namely human-to-machine, machine-to-machine, and KG-based ones. The authors of [11] discuss semantics-guided source and channel coding, a transmission paradigm in semantic communications that utilizes data semantics diversity and wireless channel diversity to enhance system performance, presenting the general system architecture and key techniques.

Table 1 Summary of recent semantic communications surveys and tutorials

Year	Reference	Data				Architecture		Knowledge extraction			Timeliness		Semantic information theory
		Text	Image	Sound	Multimodal	Point-to-point	Networking	Knowledge graphs	Ontologies	Domain adaptation	Timing	Multi-flow traffic	
2021	[10]	✓	✓			✓		✓					
2021	[11]		✓			✓							
2021	[21]		✓			✓					✓		
2021	[22]	✓	✓	✓		✓		✓					✓
2021	[23]	✓	✓	✓	✓	✓		✓					✓

2021	[24]	✓	✓	✓	✓	✓		✓					
2022	[25]	✓	✓			✓		✓			✓		✓
2022	[26]	✓	✓	✓		✓							
2022	[27]	✓	✓			✓		✓			✓		
2022	[28]	✓	✓	✓		✓							
2023	[29]	✓	✓		✓	✓		✓	✓		✓		
2023	[30]	✓	✓	✓		✓		✓			✓		✓
2023	[31]	✓	✓	✓	✓	✓		✓			✓		✓

Similarly, [21] introduces a new approach to real-time source reconstruction for remote actuation, utilizing semantics-empowered sampling and communication policies. This results in a significant reduction in reconstruction error, cost of actuation error, and uninformative sample generation. Also, [22] conducts a historical review of the basic semantic communication definitions and research efforts focusing on fundamental mathematical theory and illustrating the relation with beyond 5G (B5G) applications. Moreover, in [23], an in-depth analysis of semantic communications principles and challenges, focusing on the advancements in DL-enabled semantic communications for multimodal data transmission, including text, image, and audio, and outlines research challenges is provided. In [24], a novel architecture that uses federated edge intelligence for resource-efficient semantic-aware networking is introduced, enabling users to offload computationally intensive tasks to edge servers while protecting their proprietary model-related information through intermediate results.

The authors of [25] present a comprehensive vision for building next-generation reasoning-driven, AI-native semantic communication networks. They explore the transition from classical information theory to semantic information theory, focusing on the migration from bit-pipeline to semantic communications. The reasoning foundations for semantic communications are examined, focusing on knowledge-driven networks. The article proposes semantic KPIs for evaluating AI-native systems' performance and optimizing future semantic networks. It also discusses building scalable semantic communication networks and addressing computing, control, and networking challenges. Furthermore, [26] provides an overview of the most recent advancements in deep learning (DL) and end-to-end (E2E) communication for semantic communications. It also discusses the specific challenges that need to be addressed. The authors of [27] introduce an edge-driven training, maintenance, and execution of semantic communications for intelligent agents. In addition, this article investigates how SemCom can enhance edge intelligence by improving agent generalization capabilities at lower computation and communication overheads and presents a case study on semantic-aware resource optimization for wireless Internet of Things (IoT).

The authors of [28] discuss cross-layer interaction, goal-oriented communication applications, and semantic communications trends. This article highlights challenges and potential solutions for practical implementation in 6G wireless technology. Moreover, [29] discusses four approaches to semantic communication: classical semantic information, knowledge graphs, modern deep learning techniques, and focusing on the significance of information. The first approach extends information-theoretic results to include semantic information, while the second uses knowledge graphs for semantic communication. The third uses deep learning techniques for efficient communication, while the fourth emphasizes the significance of information. The authors of [30], discuss the latest SemCom efforts and their relation to other communication levels, including conventional ones. This article categorizes critical enabling techniques using explicit and implicit reasoning-based methods and explores their evolution and contribution to modern content and channel semantics-empowered communications. It also highlights advanced practical techniques for improving semantic accuracy, robustness, and large-scale scalability for future developments and industrial applications. Finally, [31] investigates three categories of SemCom systems, namely semantic-oriented, goal-oriented, and semantic-aware communications. Also, it divides semantic system architecture into semantic information extraction, transmission, and metrics. It discusses SemCom's potential uses in 6G and the new network architecture it may enable.

2.3. The Role of Semantic Communications in NANCY

The new vision of 6G networks that incorporates semantics and effectiveness aspects takes inspiration from, Nikola Tesla, who stated, in 1926: ‘‘*When wireless is perfectly applied, the whole Earth will be converted into a huge brain*’’ [32]. Following this approach, NANCY's strategy for 6G networks places semantics and effectiveness at the core of network design. By emphasizing the communication goal and semantic content, we can filter out only the data needed to convey the intended information or fulfill a particular goal. This means that ignoring irrelevant data becomes a key tactic to greatly reduce the amount of data transmitted and recovered, thereby improving data and energy efficiency. Under this perspective, goal-oriented and semantic communications are essential to uncovering the meaning behind bits and enabling cognitive, effective task execution among network nodes. This shift represents a major change in focus: the success of task execution at the destination (effectiveness) now takes precedence over error-free communications at the symbol level (technical). In practice, the NANCY semantic encoder (part of the NANCY SemCom component) identifies and extracts the semantic content from the source signal, removing or compressing irrelevant details. Meanwhile, the NANCY semantic decoder interprets the received data and reconstructs it in a form understandable to the user at the destination. It also evaluates user satisfaction and determines if semantic information was successfully received. Details about the SemCom component's placement are discussed in Section 6.3. Lastly, NANCY will address any semantic noise introduced during the communication process, which can lead to misunderstandings or incorrect reception of semantic information.

By analysing the current literature, it becomes evident that semantic information related to the purpose of the communication can be invaluable for ensuring that only necessary data are transmitted; thus, increasing bandwidth, delay, and energy efficiency of the system. Hence, by utilizing SemCom, the focus on transferring symbols and bits without errors is shifted towards the successful execution of the task at hand. This deliverable presents our vision for the role of SemCom in NANCY and in future networks. It also aims to initiate further research in SemCom for next-generation intelligent 6G wireless networks, contributing to the development of future research directions. To achieve the aforementioned objectives, in the upcoming sections, we provide a full examination of the integration of SemCom in the context of 6G by conducting a comprehensive analysis of current research and exploring the possible uses of SemCom in 6G. Additionally, we examine SemCom which falls under the category of conventional connection-oriented communication, together with goal-oriented communications (ASL transmission and V2x communications). To provide a clear plan for the implementation of SemCom, we examine and evaluate the primary components of a SemCom system across three dimensions: semantic knowledge, timeliness, and semantic information theory. In addition, we suggest a new SemCom semantic networking architecture that relies on point-to-point SemCom to provide semantic-aware networking with optimal resource utilization. We analyse the impact of various technologies and explain the lessons learnt on their respective SemCom categories, scenarios, advantages, and limits. Finally, a set of potential areas for further investigation are determined.

3. Technology Enablers for Semantic Communications

People tend to underestimate what can be accomplished in 10 years and overestimate what can be accomplished in 2 years. Many areas of our everyday lives will be enhanced by ultra-high-speed and ultra-reliable wireless connections, native artificial intelligence (AI), and sophisticated sensor technologies when new wireless communication systems, such as 6G, become more widely deployed. Although 5G is still in the process of being fully implemented, research efforts have already been focused towards developing beyond 5G (B5G) and 6G technologies to cover the needs of 2030 aiming to transition wireless networks from connecting devices to connecting intelligence [33]. Based on the latest preliminary proposal from the International Telecommunication Union (ITU) [34], the evolution of the fundamental services of wireless networks is summarized in Figure 3. With such a quick transition to B5G/6G development, one must question what limitations and shortcomings of 5G are pushing academia and industry to move towards the next generation of networks and in what way will its requirements be fulfilled.

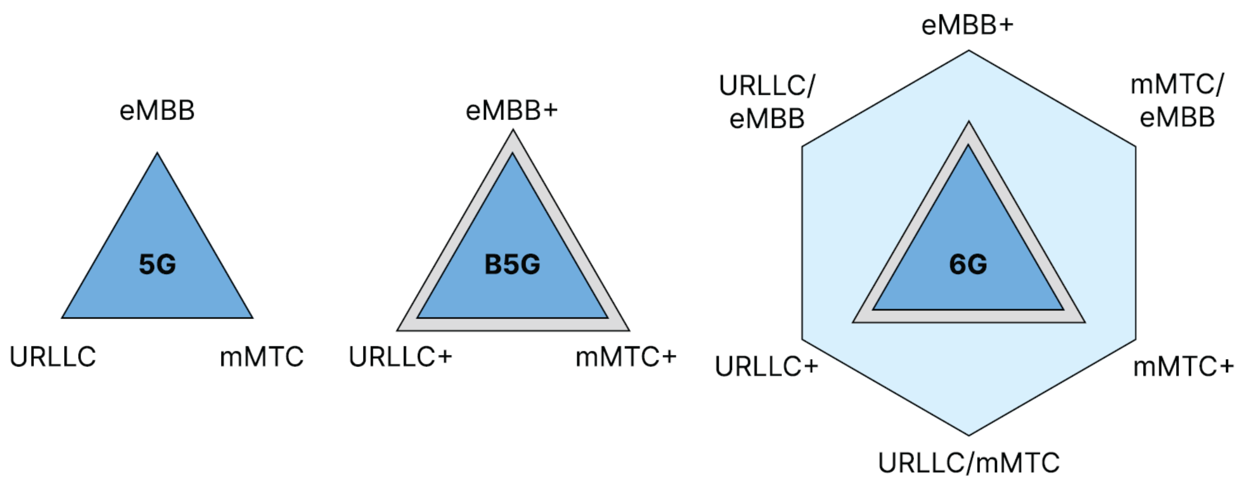

Figure 3. Evolution of fundamental services from 5G to 6G networks

Wireless networked control systems (WNCSs) include closing the feedback control loops using a wireless communication network [35]. Efficient network resource management strategies are necessary to handle the increasing number of WNCSs, since the impairments produced by network restrictions have a substantial impact on control performance. Transmitted data volume is a crucial performance metric in traditional communication networks. Conversely, in WNCSs, the network's effectiveness is evaluated based on its capacity to support control applications, but the pace of data transmission must be restricted to prevent network congestion.

Multiple studies demonstrate that explicitly considering the dynamics of the state and recognizing its significance for control objectives results in improved control performance [36, 37]. Studies have shown the effectiveness of semantic metrics, other than age, in wireless resource scheduling at the medium access control (MAC) layer. This has been supported by both theoretical research [38, 39, 40] and practical experiments [41]. Hardware-induced limits in control performance sometimes restrict real-life deployment situations due to cost constraints.

3.1. Native AI

AI advancements have significantly impacted 6G networks, with real-time AI expected to support distributed learning with high spectral efficiency and low latency. This will enable the interchange of data and model parameters across numerous intelligent agents, ensuring native trustworthiness and local data privacy. AI can alleviate the cost and manpower load of mobile networks, enabling zero-touch network orchestration through predictive network analytic services and end-to-end system automation. The production of software in 6G networks will evolve into a data coding paradigm, utilizing AI algorithms to construct deep neural networks (DNNs) for distributed learning and inference applications. This will enable real-time and large-scale inference requirements for society and vertical industries. Mobile communication networks will transform into platforms with integrated connectivity and computing capabilities, supporting AI services with fast learning convergence. Future mobile communication networks will generate, collect, and distribute vast amounts of data for orchestration and automation network services. To this end, SemCom will play a catalytic role in this transformation by reducing data transmission between multiple agents, resulting in higher task completion compared to content-blind data transmission.

WNCS design is fundamentally interdisciplinary, encompassing signal processing for accurately estimating the state of the plant, control theory for effectively regulating plant behaviour, and communication theory for reliably transmitting sensor and controller signals despite limited communication resources [42]. Given that estimate and control are dependent on the information provided by the communication system, the design of a WNCS to achieve optimum control performance should include the integration of estimating, control, and communication algorithms, which have a strong interdependence. It is desirable for such algorithms to be collaboratively constructed and optimized to maximize the control performance of WNCSs while operating within resource limitations. In this respect, most co-design challenges may be expressed as dynamic decision-making problems. However, when dealing with extensive state and action spaces, traditional approaches are not feasible owing to the curse-of-dimensionality. To address this problem, native AI methods have been used, using artificial neural networks to acquire knowledge on both the control and transmission scheduling signals [43]. It is important to mention that the DRL-based algorithms are independent of models and are used in real-world WNCS scenarios where precise understanding of the nonlinear system models is not necessary [44]. In contrast, traditional solutions rely only on model-based approaches.

3.2. Timeliness

6G wireless networks require instantaneous data transfers and stringent timing requirements for real-time applications. Recent studies have included various timing metrics, such as the AoI or age of loop (AoL), to better capture real-time activity. This requires a broader understanding of time and a more generalized framework that incorporates network services. Future networks must account for the interaction between the digital and physical worlds, as timing constraints are intrinsically analogue. The research community has developed metrics, like AoI and VoI and integrated learning- and consensus-based applications into timing characterization. SemCom can push future communications beyond the predetermined constraints of 5G, ensuring only the most timely, valuable, and relevant data is sent. This breakthrough network architecture is expected to increase network energy efficiency, computational efficiency, and resource utilization, benefiting future 6G applications, such as consumer robots, telemedicine, environmental monitoring, and autonomous vehicles.

The concept of cross-layer design of network algorithms, with respect to application objectives of WCNSs and the use of application-defined metrics, is now receiving attention in the semantic communication paradigm [45]. Semantic communication incorporates metrics that consider the importance of the conveyed data in relation to the objective of the application. The AoI is a commonly used semantic measure for real-time applications. AoI refers to the duration that has passed since the most recent update was generated and made accessible to the receiver. AoI strikes a balance between the sampling rate of status updates and the congestion caused by an excessive amount of generated information. AoI is a timing measure but in certain scenarios fails to capture the significance of generated and communicated data.

3.3. Ubiquitous Connectivity

6G networks aim to provide ubiquitous connectivity through three main pillars: terahertz (THz) communications, extreme multiple-input multiple-output (MIMO), and three-dimensional (3D) networks. THz communications are considered a potential contender for ultra-broadband in 6G applications. However, there are constraints on THz communication, such as direct line of sight transmission. SemCom can be used to compress data before transmission, improving network efficiency. Extreme MIMO, a refinement of massive MIMO, offers higher energy and spectral efficiency across a broader frequency range. However, it faces challenges such as heavy signal processing load, increased cost, large reference signal overhead, constraint fronthaul capacity, and inaccurate channel modelling. The addition of a third dimension, made possible by deploying network nodes in the sky and space, allows for the transition to 3D networks. The ground network includes wireless networks, wireless local area networks (WLANs), mobile ad-hoc, and more. The space network comprises ground infrastructure and satellites, offering advantages like high mobility, fast deployment, adaptable coverage, and low cost. SemCom is considered a key enabler of 3D communications due to the transmission of important semantic information in place of source data, improving energy efficiency, spectral efficiency, and cost.

To realize ubiquitous WNCSs, it is essential to have an infrastructure that can handle interconnected processing devices. More precisely, this infrastructure must have the capacity to support a large number of stationary and portable devices, often referred to as sensor nodes or motes, which communicate with each other via wireless transmission [46]. Similarly to the WNCS paradigm, sensor nodes of ubiquitous WNCSs have the task of monitoring and collecting parameters and they may either analyse the data locally or send it to one or more routers. Typically, these nodes are characterized by their small physical size, affordable cost, low power consumption, and low processing capabilities. To provide dependable, protected, and long-lasting communications, a diverse range of protocols are used to optimize resource utilization, precisely route sensor data, and successfully maintain wireless connections. When building ubiquitous WNCSs, it is important to consider several elements such as energy efficiency, topology, synchronization and scheduling, fault tolerance, coverage, and orchestration protocols.

3.4. Native Trustworthiness

6G networks must address privacy and security concerns by ensuring explainability and trustworthiness. With decentralized architecture, intelligent functionalities, and big data, new issues arise regarding trustworthiness, security, and privacy. For instance, trustworthiness is predicated upon both mutual identification and the integration of robust security and privacy preservation mechanisms inside the network's signalling and data flows. Conventional approaches must adapt to new paradigms like robot-to-robot and holographic communications. High-throughput data monitoring, encryption, and blockchain are essential features for security architecture. SemCom techniques can boost native trustworthiness by transmitting only the necessary data and incorporating autonomous security modalities like intelligent mechanisms and cryptography.

Native trustworthiness has the potential to be used for establishing, strengthening, and assessing trust in the WNCSs. Trustworthy encryption methods play a crucial role in guaranteeing security, privacy, and data integrity on the sensing layer. Given that WNCSs largely depend on a wide range of diverse IoT devices, the end devices are susceptible to many security risks, including data leakage, identity spoofing attacks, and manipulation [47]. Furthermore, asymmetric encryption facilitates the safe transport of data from end devices to servers for data processing. Reliable techniques provide a safe authentication procedure in the WNCS situation while minimizing the overheads seen in prior authentication protocols. Trust assessment techniques may be used to evaluate the reliability and credibility of each device, database, or server inside the network. The trust score of an entity enables network administrators to choose whether to let the entity to join the network or to bestow access permissions onto it. However, the reliability of trust assessment methodologies and device identification becomes vital since the trust connection between two entities is built based on their degree of trustworthiness. When there is user involvement, explainable

AI will provide a clear and understandable explanation of the decision-making process or the outcomes of medical data analysis.

4. Semantic Knowledge

The main distinguishing factor between conventional and SemCom systems is the existence of a shared background knowledge base for the source and destination [26]. This common knowledge base may contain text, images, video, or sounds and contains all the semantic information that is required by the system to function. Specifically, on the source's side, the semantic information is extracted from the message that needs to be transmitted and then it is conveyed to the destination over the semantic channel by the semantic transmitter. On the destination's side, the semantic receiver is responsible for the interpretation of the received messages based on the shared knowledge base. Thus, the generation of a common knowledge base plays an important role in realizing robust SemCom systems. In addition, the collection of the semantic information through environmental perception is a continuous self-learning process carried out by the system, much like the human learning process. However, it is not always possible for the source and the destination to have a shared knowledge base. One reason may be their distinct perception of the environment, while another their different capabilities.

To achieve robust SemCom between the source and the destination, it is necessary for the destination to not only decode the semantic information of the received messages, but also update its knowledge based on it [48]. Given that a variety of different representations can convey the same information, it is necessary to select the optimal knowledge representation framework to ensure the efficiency and performance of the system. However, this goal poses an insurmountable challenge when considering a single globally applied framework. Thus, each specific SemCom application must be tackled separately with its distinct knowledge representation framework, each with specific constraints, rules, facts, ontology, and reasoning engine. Although such a framework may be created specifically for a single application, it is impossible to cover 100% of the relations between different objects in its ontology. As a result, a trade-off is created between the reliability of the system and its time and complexity constraints. Finally, knowledge representation frameworks developed for SemCom are required to make decisions based only on the semantic knowledge of the transmitted data and should not be influenced by the structure of the messages. In other words, it is imperative that two distinct syntactic representations of the same message to be assigned to the same semantic information. For example, using different numerical representation systems (i.e., decimal, hexadecimal, or any other) should be equivalent in terms of semantics.

4.1. Semantic Heterogeneity and Ontologies

An open problem in SemCom stems from the heterogeneity of the information in communication systems and networks, where it is common for the source and the destination to use different "languages" to conceptualize and portray information. Such languages originate from a broad range of use cases and with different tasks at hand. Consequently, it is possible for the same message to have different interpretations depending on the application scenario. Because of this semantic diversity, the communication process is sometimes clouded by uncertainty. Even in scenarios where the same terminology is used, there may exist diverse interpretations of the information carried by a message [49]. That makes it difficult for machines and humans to share data with one another. To overcome this, a method that is capable of unambiguously and openly expressing the semantics can be provided by ontologies.

By defining the semantic interpretation of the symbols used in communications, ontologies aid distinct entities understand each other. The most adopted categorization of ontologies distinguishes them between logic and non-logic-based ones [50]. In a logic-based ontology, the semantics of its content are represented by names and relation symbols with their interactions being denoted by logical definitions and axioms. As a result, the semantics of the semantic information of the language are governed by the logical definitions and axioms that define the acceptability of certain interpretations.

Non-logical ontologies are often used to provide the semantics of terminology systems, such as semi-formal ontologies found in unified modelling language (UML) as well as natural languages included in the different international organization of standardization (ISO) standards. Non-logical ontologies do not limit the possible interpretations of words using logical axioms, which is how logical ontologies establish the semantics of systems. For example, standards are an important kind of ontology that is not logic-based. In more detail, standards define a vocabulary by requiring that all words be understood in the context of an unchanging and well-defined body of knowledge. By establishing a common vocabulary, humans eliminate the possibility of confusion over the interpretations of natural language terms; similarly, the use of diverse symbols can be utilized to eliminate ambiguity in communication terms.

The ontology stores all meta-information and knowledge about devices and their kinds, making it a crucial component in ensuring semantic interoperability. One frequent purpose in constructing ontologies is to provide a shared understanding of data across applications. The system may now define application-related notions without relying on a universally accepted theory [51]. When it comes to solving the challenge of how to get diverse actors to efficiently communicate, there are at least two distinct approaches that rely on ontologies [52]. In the first class of approaches, all actors use the same language throughout the exchange of information, i.e. the same, clear vocabulary (standard) will be used in their communications since it will define a common semantic terminology. Transformation mappings are required if an actor’s internal vocabulary deviates from the norm and the human factor is necessary for their development to resolve semantic heterogeneity based on the terminology as defined in the standard.

In the second type of ontology-based interoperability-preserving approach, applications make use of distinct terminological systems, which are in turn governed by a larger reference terminology [53]. Both types of terminologies, whose semantic information are given by means of logic-based ontologies, are linked with indirect connections between them. The mappings to and from the reference terminology are automatically calculated due to the semantics of the more particular terminologies being described using logic-based ontologies. The goal of the SemCom ontologies is to make it possible for all actors to automatically produce transformations between various terminological systems.

4.2. Knowledge Graphs

To facilitate visualization and explainability, the notion of the knowledge graph has arisen as a network of information. Knowledge graphs are representations of networks that depict the relationships between the various entities [54]. The data is often represented graphically as a network graph, which is stored in a database. The three essential parts of every knowledge graph are nodes, labels, and edges. Specifically, entities are represented by nodes while edges specify their relations. Hence, knowledge graphs are an integral part of semantic representation that has been applied in various applications, such as recommendation systems, data integration, and more, while they are envisioned to play an important role in SemCom systems.

One distinguished technique of knowledge graphs that is termed knowledge graph embedding has been utilized in many applications, such as real-time ranking, recommendation systems, and data integration [55, 56, 57]. This unique approach was introduced in [58] to translate and transmit semantics both from edges to nodes using nonlinear functions and from nodes to edges based on linear bias vectors. Specifically, in [55], an innovative ranking method that makes use of knowledge graphs is proposed to overcome the difficulty of deciphering the research terms in queries. This approach is based on an innovative knowledge extraction and visualization system called semantic knowledge graphs that create buffers between an edge and the connected nodes to generate new edges dynamically based on the underlying statistical characterization. Moreover, mutual information maximization has been applied to bring together the semantic spaces at the word and entity levels [56], while in [57], the authors develop a strategy for semantic data integration that makes use of the inherent keyword and structured search functionality of online sources.

Regarding SemCom, a paradigm for cognitive SemCom that makes use of the notion of a knowledge graph is introduced in [59]. By using triplets as semantic symbols, a straightforward and understandable technique for semantic information recognition is devised. This method enables the correction of symbolic errors at the destination. To achieve this, the pre-trained model is also improved to retrieve semantic information, overcoming the limitation of fixed bit length coding in variable length message transmission. Both semantic matching and translational distance models are the two most recognized categories of knowledge graphs. The former relies on comparisons of similarities, whereas the later employs measures of separation [60]. Specifically, semantic matching models, which is the most promising method for SemCom, use scoring methods to quantify the credibility of facts by comparing the vector space representations of entities and relations for latent semantic similarities.

4.3. Multi-modal Codebooks

Task execution [16], [61], [62] and data reconstruction [63, 64, 65, 66] are the two primary applications of the current efforts on SemCom. In the former, only task-specific semantic information is collected and encoded at the sender. In more detail, in [61], a model is presented for performing the image retrieval job, while reducing by both bandwidth and power requirements. In [16], a SemCom system focused on picture categorization is presented. Image categorization is another application that uses a SemCom system based on a vector quantization-variational autoencoder [62]. In data reconstruction applications, the semantic system uses the extracted global semantic information to rebuild the original data. Specifically, the authors of [63] presented a DeepSC method to encode text into sentences of variable length. In [64], the performance of a joint source channel coding image transmission system was investigated for various signal-to-noise ratio (SNR) values. In terms of video transmission, the cliff-effect may be avoided by utilizing DL methods, as described in [65]. Finally, in [66], the proposed SemCom technique encodes incoming voice messages as semantic representations at the receiver and then decodes them back into a waveform at the transmitter.

Although these methods have been proven effective in certain applications, they are limited in the sense that they use a single data modality. It is challenging for these models to fulfil several tasks using multi-modal sources as: i) changing the job necessitates updating the model, which, in turn, necessitates a large amount of gradient transmission for retraining the model; and ii) storing several models for use with various tasks may be impractical on edge devices due to space constraints. It becomes evident that the creation of a unified multi-task SemCom system is critical, since most devices nowadays need to be able to perform many functions simultaneously. To address this problem, a transformer-based paradigm was presented in [67], which, unfortunately, has not yet been unified for distinct tasks.

4.4. Domain Adaptation

Attempts on multi-task learning have been done in computer vision and natural language processing [68], [69]. The goal of multi-task learning is to make use of the domain-specific knowledge learned through training samples of similar tasks. As compared to models built for only one activity, multi-tasking models are characterized by i) drastically reduced storage requirements due to the multitasking ability in exchanging model parameters; and ii) if various linked tasks share semantic information, it becomes more efficient to train the model for numerous tasks at once and increase its performance. For example, although the codebook in [70] is trained using a multi-modal approach, its usefulness in multi-domain settings is severely constrained. This is mostly since data from various modalities and activities have distinct distributions, resulting in substantial variation in the encoded properties. Thus, a very large codebook is required to deal with these multiple tasks. To solve this problem, a single codebook for a variety of tasks by capitalizing on domain adaptability was developed in [71]. More specifically, a domain adaptation loss is enlisted for training to boost performance and decrease convergence time.

5. Timeliness Aspects

As we move toward 6G wireless systems, the requirement to merge the digital with the material world is tightened. Therefore, not only will real-time needs become tighter, but also the concepts of causality, presence, simultaneity, and time will become more nuanced and difficult to grasp. Latency alone is inadequate to characterize real-time needs, according to a growing corpus of research [72]. Information timeliness, as described by the AoI and its variants, is one such need. In general, new distributed paradigms like distributed consensus, learning, and sensing are not reflected in the metrics produced from a traditional black-box approach to communication network architecture.

5.1. Multi-flow Deadline Constrained Traffic

The emergence of real-time or near-real-time and interactive applications has shown that only reducing latency is insufficient to ensure the flawless functioning of these types of programs. Specifically, the network's operations should be time-bound [73]. In multi-flow deadline constrained traffic, once the transmitter creates a packet, the timer starts counting backward from the maximum delay, T. If a packet makes it before the timer hits zero, the transmission was successful. This is true regardless of the packet's delay. Time-sensitive traffic is important in industrial settings with ultra reliable low latency communications (URLLC)-related use cases leading the pack [74].

The messages that can be effectively successfully transmitted before the deadline are measured by a metric termed timely throughput [75]. Generally, there is often a trade-off between the throughput that can be achieved and the T 's tightness, since tighter T 's needs more resources for each packet [76]. The minimum latency under ideal circumstances serves as the hard floor for the achievable deadline. Although 6G systems aim to achieve under 1 ms requirements, technical obstacles may include computational and hardware component limitations, as well as the impact of the selected medium access techniques.

Due to the packets' potential importance, a first-come, first-serve policy may not be the best option if the utilized KPI is a future time reference. Therefore, achieving optimal or sub-optimal performance in the case of time-sensitive traffic requires to account for not only the latency but also the likelihood of deadline violations. This can be carried out by optimizing packet scheduling, interference control, resource provisioning, and media access [77, 78, 79, 80]. Scheduling optimization with regard to other transmission characteristics, including power management, as well as adaptive approaches that adjust scheduling choices to meet traffic patterns, such as randomization and Markov decision processes, have been investigated in recent literature [81, 82, 83].

End-to-end traffic deadlines may be achieved via the use of rate-adaptive transmission or probabilistic techniques to ensure that packets arrive on time [84], [85]. Multi-flow connections are often combined with packet-level coding, taking an application block's delay into consideration rather than a single packet's [86]. An intriguing example is provided by hypertext transfer protocol (HTTP) adaptive streaming, an application-layer protocol for streaming video [87]. In this case, there is no hard deadline because it does not signify interactivity but rather is contingent upon the capacity of the video buffer at the receiver, i.e. the transmission must be finalized before the currently playing video segment completes to avoid interruptions in the video's playback. Since the deadline for each data packet is conditional on both the packets' contents and the buffer's state, this is an illustrative example of a relative timing reference.

5.2. Timing Metrics

Although relative timing references, which consider complicated networking contexts, are increasingly being investigated, previously described metrics are still the most widely used form of time measurement in protocols and standards. This section unifies all these indicators under one umbrella and considers time and the associated metrics holistically.

5.2.1. Latency/delay

One of the earliest and most basic timing metrics is the latency (or delay), which counts the time required for a packet to travel between two points in the network. Latency is quantified based on a past time reference, which means that the timer begins when the packet leaves a certain layer in the protocol stack and ends when the packet reaches the corresponding layer at the receiver. Theoretical and applied investigations on latency have been conducted in a wide variety of system configurations [72]. In more detail, latency in networks has been analysed using queueing theory that is capable of scaling from basic $M/M/1$ systems to complicated access mechanisms with varying arrival patterns. Moreover, applications like video streaming are particularly sensitive to fluctuations in packet delays [88], while when sophisticated channel operations and correlations in arrival are taken into account, the complexity of the problem increases [89], [90].

Various AI-based latency reduction strategies have been suggested for 6G networks. For instance, overall service latency may be reduced by taking both communication and computation factors into account [91]. Also, latency in complicated processes has been decreased by strategically placing network functions [92]. Use cases with stringent timing limits and high throughput, like XR transmission, or quick mobility, like vehicle communications, present problems that must be dealt with on an individual basis. By providing additional training data and minimizing the influence of the training on the actual network, the digital counterpart of the network can also constitute learning-based optimization approaches more efficient. These methods facilitate resource allocation for sophisticated distributed systems and can be utilized along with federated learning and other methods [93].

The overall latency of a packet includes the significant initial access delay that might occur via wireless connections, which can be achieved by improving the service rate and reducing the rate at which packets join the network, both of which reduce the overall amount of delay. One of the most investigated issues is the trade-off between network throughput and latency. Latency increases significantly when congestion management measures exceed the available capacity. Congestion control techniques shape traffic through its latency-related practical function, which in turn influences the queue's current state and, by extension, the choices of all transmitters in the transmission chain [94]. Because of how closely these two variables are linked, studying how latency is used as a measure is an intriguing but difficult path to explore.

5.2.2. Timely Throughput

When considering packet routing and static arrivals, going beyond average delay and examining per-packet delay performance is a significantly more difficult issue with far fewer known outcomes. Finding the min-cost route between a source and a destination with an end-to-end delay restriction is NP-hard [95]. Considering varying arrival times is another challenge that must be addressed. To address this, the timely throughput metric has been introduced, which quantifies the number of messages that can be successfully sent before the deadline [96]. Tighter delays need more resources for each packet, hence there is usually a trade-off between throughput and delay tightness. In [97], an opportunistic scheduling method is presented that balances worst-case latency with timely throughput, and which, when applied to hop-count-limited transmissions, maintains the delay guarantee. Yet, it is dependent on a connection selection mechanism, which reduces its effectiveness in broad networks. In addition, throughput assurances are less rigorous under strict deadline restrictions, since throughput is measured relative to the worst-case delay rather than the deadline set by the application. Finally, it does not optimize the end-to-end latency, since packet scheduling on various lines is handled independently. In [98], the authors model the timely throughput maximization issue as a Markov decision process with exponential-size constraints and then develop an approximate solution by first finding the ideal path for a single packet to travel. Moreover, [99] deals with the more intricate wireless network configurations that include link interference. While the complexity drops from exponential to polynomial using this method, it may still become computationally costly since it needs to solve a dynamic programming problem for each packet in each time frame. As an added downside, none of these works prioritizes reducing operating expenses, which is crucial in today's elastic cloud settings. [98], [99].

5.2.3. Age of Information

During the last ten years, AoI research has made incredible strides [100]. AoI refers to the amount of time since the last successfully received packet reached the destination. Imagine a stream of packets, each of which is created at time τ_n by the source and received at time τ'_n by the destination. In this respect, the AoI can be expressed as

$$\Delta(t) = t - \xi(t), t \geq \tau'_1 , \quad (1)$$

with

$$\xi(t) = max\{\tau_n \mid \tau'_n \leq t\} , \quad (2)$$

denoting the creation time of the most recent message. The sawtooth shape associated with AoI is the result of a random process with sample pathways that rise linearly between message receptions. This is the result of the inherent unpredictability of packet production timings and transmission delay. Likewise, AoI sets the temporal anchor as the time at which the most recently received packet was produced. It is important to keep in mind that whenever a new packet arrives, the timing anchor is also changed, whereas, in the traditional latency, the anchor is refreshed each time the transmitter creates a new packet.

AoI is a typical metric for tasks involving remote control and monitoring since it quantifies the timeliness of the information at the destination. A great amount of research effort has been devoted to calculating the average of the sawtooth pattern of linear AoI in various systems, such as ALOHA [101], slotted ALOHA [102], and carrier sense multiple access [103]. The AoI has also produced several related metrics, such as peak AoI (PAoI), AoI at query (QAoI), age of synchronization (AoS), and age of loop (AoL), most of which account for contextual time considerations. In more detail, PAoI is the most recognized one, which takes a snapshot of the AoI just before a fresh packet arrives. The PAoI is able to capture the essence of ageing, and it offers great potential for evaluating the worst-case performance of a system as well as for analyzing the whole probability density function, tail-related metrics, or higher moments of age [104]. However, both AoI and PAoI assume that information should remain fresh throughout the entire communication process. Instead, in QAoI, a receiving application can only access the data at certain times [105]. This metric is quite like PAoI, with the exception that it does not sample the AoI unless the application specifically asks it. Due to this change, the environment shifts from a push-based system to a pull-based one, in which the application controls the flow of data during transmission. Moreover, to accurately reflect the timeliness of the local cache AoS is established in [106]. Specifically, it investigates the challenge of how a local server balances the need to update data from several sources while adhering to a certain limit on the total refresh rate. The final related metric is a sense–compute–actuate cycle with a maximum duration between events and actions. Node 1 acts as a sensor and actuator, while Node 2 controls the wireless network-controlled system. Node 1 delivers the system status to Node 2, which responds with a control command. AoL is the AoI over the two-way connection, which comprises the transfer of the system status from Node 1 to Node 2 and the instruction from Node 2 to Node 1 [107].

5.2.4. Non-linear AoI

In most cases, a linear function of time is utilized for the penalty function that characterizes AoI. Nevertheless, a more generic non-linear penalty function is possible due to the adaptability of the underlying data-exchanging mechanism; this function may adequately characterize the worth of information updates, which may lead to non-linear AoI behaviour [108]. The percentage of non-linearity in the AoI penalty function shows how much keeping updates as fresh as possible impacts performance.

Non-linear average AoI and PAoI performance under an $M/M/1$ queuing model was studied in [109]. When using source sensor nodes with energy-collecting capabilities, a comparable system model was also analysed in [110]. These studies, however, relied on a set packet error rate and/or service rate, assumed single-antenna transceivers, and paid little attention to the possibility of user mobility. Moreover, energy-AoI trade-offs have also been investigated in coded short-packet communication [111] and re-transmission policy scenarios [112]. The reach of the antenna array at the transceiver, the possibility for user mobility, and the nature of the wireless channel were not modelled in any of the contributions, which indicates that their outcomes are dependent on a particular packet error rate. Motivated by this, the authors of [113] studied the AoI and PAoI of point-to-point wireless communication systems, where the source node was a moving simple sensor monitoring device with an arbitrary moving speed profile. Finally, given the non-linear AoI, the linear one can be derived as a special case.

5.2.5. Value of Information

The metrics discussed so far only consider the timing aspect. However, to fulfil the goal of SemCom in transmitting only the necessary information at each moment, we need more comprehensive timing metrics that capture data quality aspects as well and determine their relevance for the system's end goal. To develop such metrics, research efforts focused on moving from statistical decisions to statistical estimations, which require a closer look at the task under evaluation. Such metrics work as a link between the significance and timeliness of the transmitted information, with the most recognized ones being age of incorrect information (AoII), value of information (VoI), and urgency of information (UoI). By extending the idea of timely updates to timely "informative" updates in statistical estimation, the AoII is the first of these estimation-based metrics to appear in the literature [114]. In AoII, the age grows as the quality of Node 2's guess about the process at Node 1 declines. VoI considers both the timeliness and the quality of the new information being transmitted [115]. Specifically, as the relative timing reference is based on statistical estimation, the goal of the VoI is to minimize the difference between the actual measured process and the one estimated by the receiver through the updates. Lastly, the UoI is a relatively new development that expands on VoI in this way, shifting the focus from only measuring the process accurately to ensuring that it can be controlled reliably as well [116]. If the controller must use the data sent over the wireless connection, the data that really matters is the data that modifies the controller's actions.

5.3 Semantic Information Theory

To begin exploring how to quantify semantic information, we must first establish what we mean by "semantic information", which has been the subject of a variety of research works since 1952 [13]. Based on classical information theory, a message's entropy is calculated based on the statistical likelihood of its symbols existing in the message, while, in semantic information theory, a statement's entropy is equivalent to the probability that it can be considered true in the observed setting. In [13], the authors introduce the initial semantic information theory, which differentiates between information and amount of information. Moreover, in the same work, the amount of information in a message is evaluated based on logical probabilities rather than the statistical ones employed by Shannon. For a given message, its logical probability represents the probability of it being true. Therefore, for two distinct messages, their intersection holds more information than each of them, while their union has less information. Thus, the semantic entropy of a single message can be expressed as

$$H_s(x) = -\log_2 (m(x)) , \quad (3)$$

where m denotes the logical probability of a message x and can be written as

$$m(x) = \frac{\sum_{w \in W, w \vDash x} \mu(w)}{\sum_{w \in W} \mu(w)} , \quad (4)$$

with w being a single situation from the set W, in which x is true.

The framework presented in [13] is considered a model-theoretic method for giving probabilities to logical statements. To achieve this, Bar-Hillel and Carnap used truth tables to determine the probability that a given propositional statement is true. Because of this, there is a strong correlation between the size of a message's statements and the amount of information it contains, with fewer statements in a message meaning that it is more "surprising" and informative. However, a paradox arises from this concept since contradictory messages are assigned the highest semantic entropy. This paradox is investigated in [117], [118], where, opposite to the already presented framework that quantifies semantic entropy based on all possible situations, the core concept is that the degree of semantic divergence from a specific situation selected as "truth" is a measure of the amount of information that the message contains. To achieve this, Floridi [118] defined the quantity of information in a message, x, based on the informativeness, I, and vacuity, V. In more detail, I quantifies the information provided by m, while V measures the vacuous information carried by it.

Thus, the semantic information can be expressed as

$$H_s(x) = I - V. \quad (5)$$

The semantic information theory framework presented by Floridi is insufficient for cases where all the situations are considered; thus, it is incapable of measuring the semantic entropy of some messages. To this end, the authors of [119] incorporated truth-likeness models into the previous framework to measure both the vacuity and inaccuracy of messages. This is achieved by computing the value aggregate of all models in a message against the situation that is considered true. The semantic entropy in the last two frameworks is relative to a particular reference situation and takes values in the range of [0, 1]. However, they are unable to measure semantic information without a reference and thus are characterized as semantic similarity approaches.

Several contributions have been presented aiming to provide a solution to the open issue of developing a universal semantic information theory framework [120, 121, 122]. However, the authors in [14] proposed a framework that considers the existence of a shared knowledge base between the source and the destination. In this case, semantic entropy becomes entangled with the background knowledge and is expressed based on the conditional logical probability as

$$H_s(x \mid K) = \log_2\ m(x \mid K), \qquad (6)$$

where m denotes the logical probability of a message x given a knowledge base K and is given by

$$m(x \mid K) = \frac{\sum_{w \in W, w \vDash K, x} \mu(w)}{\sum_{w \in W, w \vDash K} \mu(w)}. \qquad (7)$$

The existence of a knowledge base has a significant impact on semantic entropy in the sense that, when the source and the destination do not share a common background knowledge, the existence of knowledge diminishes semantic entropy. However, in the case of common background knowledge, its existence enables the semantic compression of the messages in the source, thus transmitting as much information as possible despite exchanging shorter messages. Rather than relying on task-specific definitions, like previously mentioned methods, the authors of [123] investigated an information-theoretic framework for quantifying semantic information across various tasks and sources. In their work, task complexity is defined as the smallest number of semantic inquiries with which it is possible to adequately forecast a certain task and can be expressed as

$$\begin{aligned} C_Q(x; y) &:= \min_E \mathbb{E}_x\left[\left|\text{Code}_Q^E(x)\right|\right] \\ &\text{s.t. } p(y \mid x) = p\left(y \mid \text{Code}_Q^E(x)\right) \forall x, y, \end{aligned}, \qquad (8)$$

with y, Q, E, and $CodeE(x)$ denoting the task, the inquiry set, the encoder, and the inquiry vector extracted from W by using E, respectively. Moreover, the authors also showed that when y is a function of W and Q includes all binary inquiries on W , the semantic entropy and task complexity are bound as

$$H_s(y) \le C_Q(x; y) < H_s(y) + 1. \qquad (9)$$

As a result, the challenge is transformed into finding the optimal semantic encoder to derive the semantic entropy from task complexity. However, this direction needs further research since it is still in its early phases.

In conclusion, published semantic information theory definitions range significantly from one another due to their dependence on the characteristics of the task in question. The final definition is transferable to other jobs, but it is just as difficult to discover the optimal semantic encoder as it is to get semantic entropy. Because of this, semantic information theory does not have a single, unified definition, and most formulations lack the practical applicability that Shannon's classical information theory has in a wide range of technical applications.

6. Natively Intelligent Semantic Architecture

6.1. Semantic Point-to-Point Architecture

To fully exploit the advantages of SemCom, we present a semantic networking architecture that can overcome the major limitations of state-of-the-art networks [124]. Figure 4 presents the point-to-point SemCom architecture accompanied by arrows that describe the flow of information between the source and the destination. The semantic encoder in the transmitter extracts semantic information from the message by utilizing knowledge extraction algorithms and sends it across a wireless channel to the receiver, where the semantic decoder is optimised to accurately recover the message's semantic information rather than relying on symbol-by-symbol reconstruction. This means that the first stage in SemCom design is to determine what semantic information will be communicated and then extract it from the source data/message.

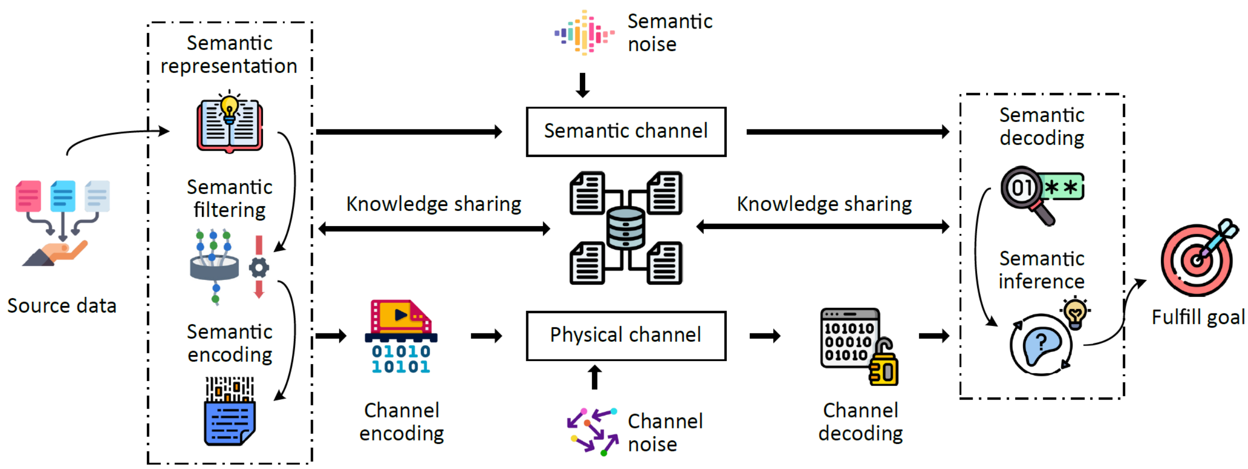

Figure 4. Semantic point-to-point architecture

In conventional communication systems, the source and channel coding that takes place aim at the mitigation of interference as well as ensuring the system's robust performance, while the semantic information carried by the messages is completely ignored. On the contrary, if we consider the semantic and effectiveness layers, the conventional communication system can be transformed into a SemCom system. The semantic transmitter and receiver perform additional intelligent operations that take into consideration transmitted semantic information. A semantic encoder uses the source knowledge base to perform the necessary semantic information extraction, while its output is then sent into the channel encoder, which prepares the data for transmission through the wireless channel. After the destination receives the transmitted signal, the channel decoder uses a semantic decoding operation and semantic inference to accomplish a desired action with respect to a communication objective or task and a destination knowledge base. Goal-specific action execution instructions that may ensure timing and timeliness metrics can be included in the output of the semantic inference module.

6.2. Semantic Networking Architecture

Building upon point-to-point SemCom systems, we present the semantic networking architecture of Figure 5 that accounts for goal-specific semantic extraction and filtering at the source prior to signal transmissions, as well as semantic decoding and post-processing at the destination. The differentiating factor of the networking architecture is the fact that the communication paradigm evolves into a multi-user, distributed, deep-edge-to-cloud networking architecture with multiple traffic flows. Hence, due to its complexity, the realization of such an architecture requires new semantic knowledge extraction and representation frameworks, theoretical models that can be used to predict and manage the multiple time-varying deadline/delay constrained traffic flows that will severely congest the network, as well as novel semantics-infused metrics capable of quantifying its performance while encapsulating its semantic nature.

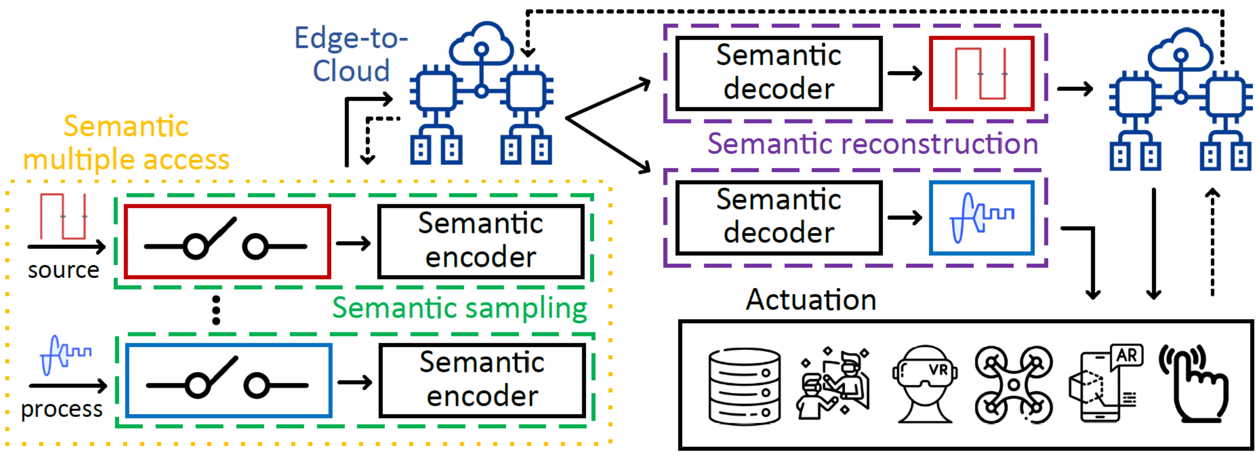

Figure 5. Semantic networking architecture

In SemCom, both the sender and the receiver function as "intelligent" entities that can perceive their environments and operate autonomously to achieve specific objectives, including data collection, system management, or information exchange. Traditionally, communication involves a sender generating a message, transmitting it through a noisy and unpredictable channel, and ensuring that the receiver can accurately reconstruct it. In contrast, SemCom enhances this process by embedding intelligence at both ends of the communication link, allowing the system to adapt to the context effectively and efficiently. Specifically, we assume that the source is tasked to convey the gathered semantic information to the destination. To achieve this, it must first extract the semantic knowledge and create the semantic representation of this information based on its knowledge base. Then, the message is translated into appropriate symbols to be transmitted through the physical channel. At the destination, the received symbols are decoded, and the extracted message is interpreted based on the common knowledge base. If the semantic interpretation extracted at the destination is the same as the transmitted one, then the system is characterized by semantic equivalence, which does not depend on the syntactic structure of the messages but only on the conveyed semantic information.

6.2.1. Semantic Channel

Due to the nature of SemCom, the physical channel represents only one part of the stochastic phenomena between the source and the destination. In more detail, the physical channel can model various stochastic and deterministic processes, such as interference, fading, multipath, etc., that may generate errors in the decoded message. However, SemCom has introduced the semantic noise, which resides in errors generated during the interpretation of the semantic information of the transmitted messages [19], [20]. The cause of these errors is usually case-specific. For example, text transmission errors may be generated due to the ambiguity of linguistics. Another cause of semantic errors can be the improper mapping between the source and destination knowledge, which leads to misinterpretation on the receiver's side.

Despite their differences, syntactic (due to the physical channel) and semantic errors can be intertwined. The reason is that the effectiveness and semantic layers of SemCom systems rely upon the technical layer in the sense that recovery of the semantic information of the message is impossible in the presence of excessive syntactic errors. However, under moderate syntactic noise, the accurate recovery of the transmitted semantic information is possible and even probable. Therefore, we can say that the semantic coding strategies can mitigate the effect of physical channel imperfection by exploiting the knowledge representation frameworks. The opposite may be valid as well. In other words, semantic noise could be present even without syntactic errors. This phenomenon is present when the utilized knowledge bases are characterized by inadequate semantic equivalence. Thus, even though the appropriate symbols are conveyed over the channel successfully, the message is misinterpreted at the semantic level.

In point-to-point SemCom systems, semantic noise is limited to the understanding between the sender and receiver resulting in complexity and fewer sources of disruption. Conversely, in the semantic networking architecture, which involves multiple nodes and indirect paths of communication, an intricate landscape unfolds. In such complex environments, semantic noise arises from varying interpretations among users, collaborative communication, and challenges of ensuring compatibility across devices and platforms. The likelihood of divergence over time is increased, while the dynamic nature of networked environments further contributes to the obstacles of managing semantic noise. Overall, while both types of systems experience semantic noise, its nature and complexity differ between the two.

6.2.2. Semantic Coding

Conventional source and channel coding strategies are not enough to counterbalance the combined effects of physical channel and semantic noise. To this end, semantic coding is required to compress the transmitted messages and extract their semantic features [14], [125], [126]. To achieve this, the semantic encoder handles the mapping between the source's semantic alphabet and the syntactic one. This mapping can be one-to-one, many-to-one, or one-to-many. Although the first case is the simplest one and needs no further analysis, in the many-to-one mapping, we observe the issue of semantic ambiguity, which is common in many natural language processing systems, while in the one-to-many case, we observe the semantic redundancy issue, where multiple symbols can be linked to the same semantic message.

The main goal of semantic coding is to achieve the most realistic representation of the perceived environment while transmitting the fewest possible messages. In conventional communications, the role of source coding is to reduce the redundancy of the transmitted information, while channel coding increases the communication reliability through defined message syntax. However, both coding schemes are concerned only with the syntactic nature of communications, ignoring the semantic one. At this point, the semantic coding schemes are introduced to preserve the semantics of the messages generated from the source. It has been proven that in low syntactic noise scenarios, semantic coding schemes can offer significant performance improvement due to the fact that received messages can be corrected by exploiting the shared knowledge between the source and the destination [14], [127].

In point-to-point semantic coding approaches the encoding process involves the sender encoding information for a receiver without the added complexity of multiple intermediaries [128]. The encoding protocols used in these scenarios are often customized to suit the needs of the two entities involved and they rely on models that are optimized for their direct exchange. On the other hand, networked communication systems introduce a distributed semantic coding approach that involves participants and nodes handling both the encoding and decoding steps. In such a scenario, interoperability becomes crucial, which means that standardized encoding formats and protocols are necessary to ensure interpretation across devices and platforms. Additionally, semantic networking systems may require mapping and support for coding to facilitate shared understanding among participants. It's important to note that networked environments are dynamic in nature. Therefore, adaptive semantic coding approaches are essential for communication systems ranging from simplicity and directness in point-to-point systems to complexity and interoperability challenges in networking environments.

6.3. Integration to the NANCY Architecture

This section aims to provide a clear placement of the SemCom component within the functional and deployment view of the NANCY architecture. The SemCom component is comprised of two main entities, namely the semantic encoder and the semantic decoder as presented in Figure 6. The former handles the extraction of the semantic information from the data collected in the end-user devices, processes the semantic information, encodes it, and forwards it to the transmitter. In turn, the transmitter forwards the encoded semantic information to the destination, which is equipped with the semantic decoder that decodes the received messages and converts the included semantic information into the appropriate format to fulfil the goal of the communication.

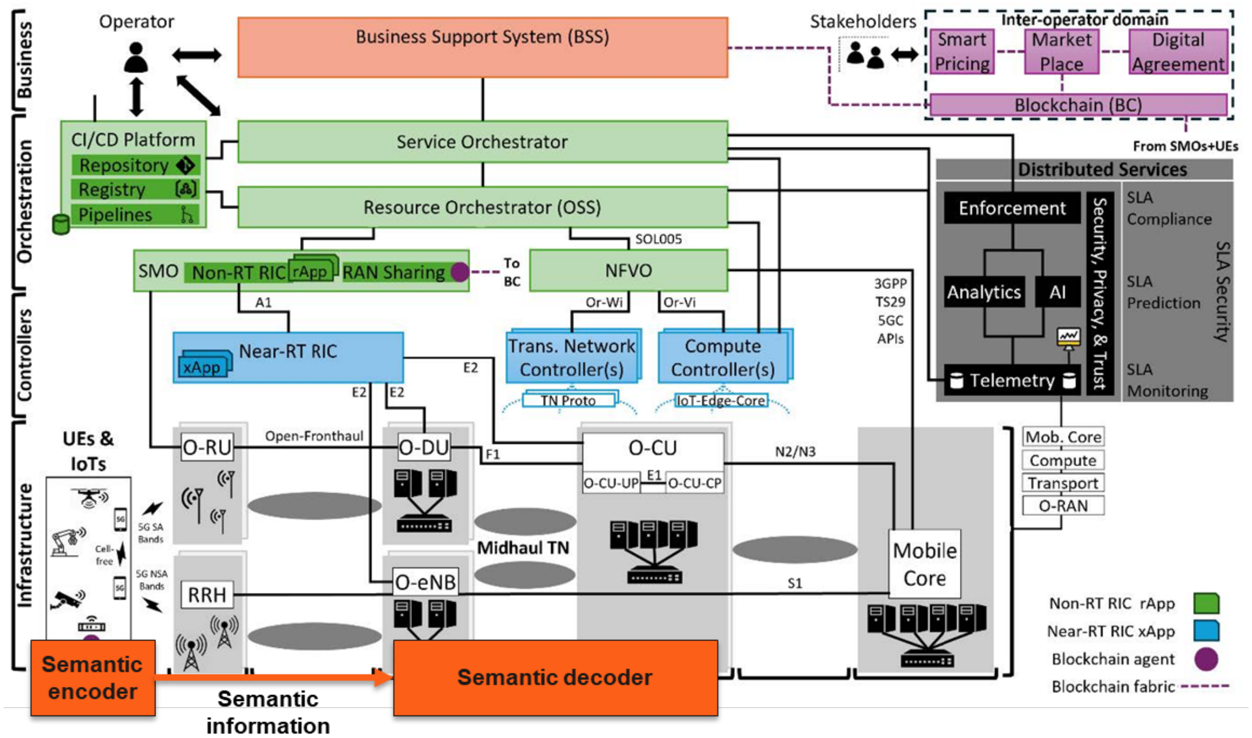

Figure 6. Functional and deployment view of the NANCY architecture

A functional diagram of the interactions between the SemCom component and NANCY's task offloading mechanisms is presented in Figure 7. The goal of this interaction is to showcase the energy and data efficiency achieved by semantic communications. In the simulated setup, the semantic encoder is deployed on the network nodes (cars, RSUs, drones, etc.) capturing video of the observed location. The extracted semantic information will be transmitted through the network to the edge server, where the semantic decoder will be deployed in order to create a DT. The specific workflow that is followed by the task offloading mechanism is analysed in D4.1 "Computational Offloading and User-centric Caching".

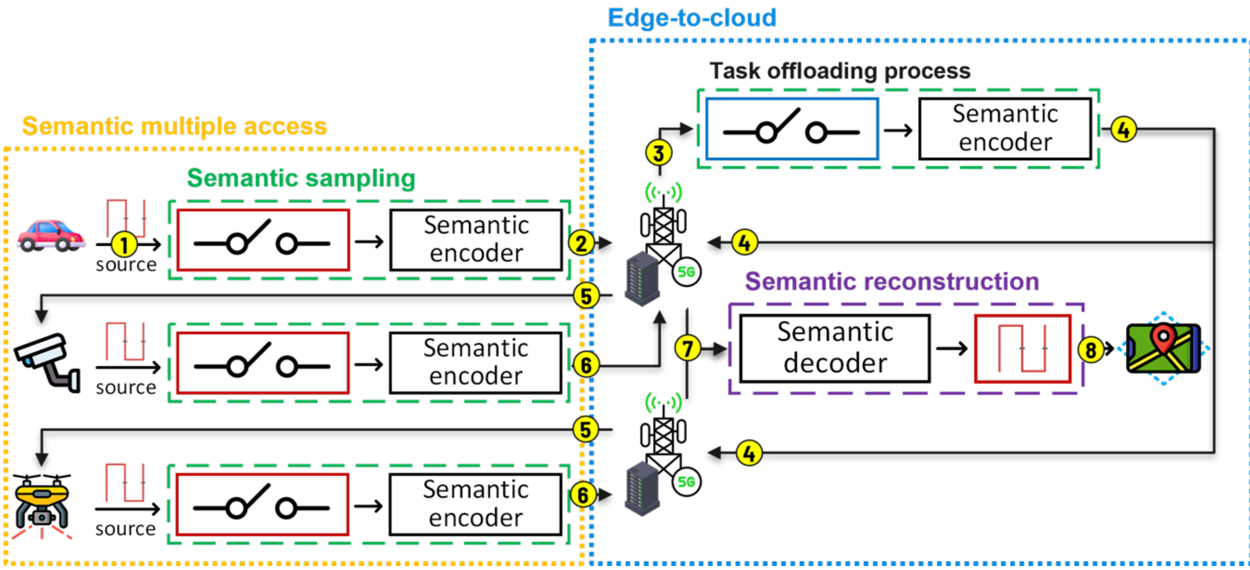

Figure 7. SemCom interactions with task offloading

7. Goal-oriented Communication Schemes

Building on the point-to-point semantic communications framework, extensive research has been carried out across various domains, including text [129, 130, 131], speech [132], images [133, 134, 135, 136, 15], and video [137]. In [129] researchers introduced DeepSC, a SemCom architecture based on a transformer model. The study in [132] extended this concept to speech transmission by incorporating convolutional neural networks (CNNs). Similarly, the work in [15] proposed a neural network that integrates source and channel coding to improve image compression, utilizing CNNs for better efficiency. Their method achieved a 3 dB increase in peak signal-to-noise ratio (PSNR) compared to traditional coding techniques under Rayleigh fading conditions. Further advancements include the image semantic coding model in [133], which applies a Laplacian pyramid to enhance compression. In [134], researchers introduced an adaptive image semantic coding technique using reinforcement learning to optimize the rate-perception-distortion metric for improved image reconstruction quality. These innovations in adaptive image coding, influenced the authors of [138] to compile a comprehensive survey on recent technological developments in semantic communications for intelligent wireless networks. This survey highlights key architectural aspects, cross-layer interactions, and real-world applications, while also addressing the challenges of implementing semantic communications in next-generation 6G wireless systems.

7.1. ASL Transmission

Sign Language Recognition (SLR), an interdisciplinary field combining computer vision and pattern recognition, has attracted considerable research attention. In [139], the authors provide a concise yet thorough review of the current state-of-the-art in SLR. They highlight that SLR commonly relies on Hidden Markov Models (HMMs) and neural networks. For example, studies in [140] and [141] use HMMs with motion sensors and modality fusion techniques for SLR. Additionally, much of the research in this area leverages neural networks, as seen in [142] and [143]. The former employs a 3D CNN, while the latter integrates Single Shot Detector (SSD), 2D CNN, 3D CNN, and Long Short-Term Memory (LSTM) models for sign language recognition. A shared focus on deep neural network-based SLR is evident in [144, 145, 146] where researchers enhance feature extraction and recognition by combining deep learning techniques, particularly CNNs and LSTMs. Similarly, in [147], an SLR model for ASL was developed by combining the k-Nearest Neighbor (KNN) method with LSTM recurrent neural networks. Data was collected using a Leap Motion Controller, and performance evaluations demonstrated improvements over support vector machines, recurrent neural networks, and standalone LSTM models. These advancements underscore the growing reliance on deep learning to improve the accuracy and efficiency of SLR systems.

Moreover, several studies have been conducted on the task-oriented image SemCom system, as documented in [133], [135]. The authors of [133] presented a SemComs system empowered by a deep neural network for the purpose of performing the visual question answering task, while in [134] a framework that utilizes reinforcement learning to enable image SemCom across various tasks is presented.

Finally, the authors of [136] introduced DeepWiVe, a video transmission joint source and channel coding system that leverages CNNs and non-local blocks to effectively capture the redundancies existing among frames.

Previous studies have represented the semantic information of images as feature vectors, which lack interpretability and are not easily understood or directly usable by humans. As a result, the receiver must reconstruct the original images—a process that is both complex and inefficient, requiring neural networks to decode the received data and transform it into meaningful information. To address this challenge, this section explores the role of semantic communications (SemComs) in sign language transmission. The key contributions of this section are as follows:

- A novel SemCom system model is proposed for image-based sign language transmission. This model is built on a modified quadrature amplitude modulation (QAM) scheme, which eliminates eight peripheral symbols from the 32-QAM constellation, resulting in a new variant called 24-QAM.
- A CNN-based semantic encoder is introduced to effectively utilize the ASL alphabet.
- A unique dataset is developed using RGB landmarks and key points to enhance hand pose representation.
- The proposed architecture's superior performance is demonstrated through simulation results.

7.1.1. System Model

As shown in Figure 1, the technical layer represents traditional semantic communication (SemCom) systems, which primarily consist of electronic components. In these systems, source and channel coding focus on reducing interference and ensuring reliable performance. However, these processes lack intelligent capabilities and do not consider the meaning or semantics of the transmitted information. In contrast, by incorporating the semantic and effectiveness layers, a conventional communication system can be transformed into a SemCom system. In this setup, the semantic transmitter and receiver perform additional intelligent operations that account for the meaning of the transmitted information. Specifically, the semantic source extracts the key semantic features from the original message and encodes them into the transmitted signal, while the semantic destination decodes this information using the common semantic knowledge base.

For example, consider an E2E communication system where a robot monitors a human using ASL. The human performs ASL signs at the transmitter side, and the robot interprets these commands to initiate specific actions at the receiver side. If the robot and the user are in the same environment, the robot captures and processes the images directly, extracting the necessary information to determine the appropriate actions. However, if the user and the robot are in different locations, the image data must be transmitted over a network to the robot, increasing the network load. In this case, the proposed semantic communication system can significantly improve data and energy efficiency by transmitting only the essential semantic information rather than the entire image data. This approach reduces the amount of data transferred, ensuring more efficient use of network resources while maintaining the system's effectiveness.

The semantic source and destination are intelligent agents, whether human, software, or hardware, capable of perceiving their environment and operating autonomously to achieve specific goals, such as sensing, controlling, or exchanging information. In traditional communication systems, the source generates a message that is transmitted over a stochastic channel to the destination. During this process, the source simply sends the message, and the destination receives it. However, with SemCom, this process evolves into a more intelligent framework. In this advanced approach, the source, S, is responsible for conveying gathered semantic information to the destination, D. To do this, S first creates a semantic representation, m, of the information based on its knowledge base, B_S. This representation captures the meaning of the information in a way that is both energy and data efficient. To achieve this, we developed a CNN that processes images of ASL gestures, d, and extracts the semantic information contained within them. The CNN then encodes this information into a compact and efficient representation, reducing the amount of data that needs to be transmitted while preserving the essential meaning. This approach not only improves communication efficiency but also enhances the system's ability to interpret and act on the transmitted information intelligently. The output of the CNN-based semantic encoder is one of the 24 letters included in the ASL alphabet and can be expressed as

$$m = S_e(d)\,, \tag{10}$$

where S_e denotes the semantic encoder.

After the message is converted into symbols s for transmission over the physical channel, a commonly used modulation scheme for 5-bit symbols is 32QAM, as shown in Figure 8 (left). However, since the semantic representation m corresponds to 24 distinct letters in the knowledge base B_S, transmitting these 24 symbols requires only 5 bits based on $\log_2(24) = 5$. This creates an issue: using 32QAM would result in 8 symbols remaining unused. To address this, we propose a modified QAM scheme that removes 8 peripheral symbols from the 32QAM constellation. The resulting configuration, now with 24 symbols, is named 24QAM. The proposed 24QAM constellation, along with its decision regions, is illustrated in Figure 8 (right). This adjustment ensures efficient use of all transmitted symbols, aligning the modulation scheme with the semantic requirements of the system.

After the transmitted signal passes through the stochastic physical channel, the received signal can be expressed as

$$y = h_i s + n\,, \tag{11}$$

with s and n denoting the modulated signal and the additive white Gaussian noise, respectively. Also, h_i represents the channel coefficient. Three distinct cases are investigated: a unitary, h_u, a Rayleigh channel response, h_r, and a Rician channel response, h_R. At the destination, the received symbols, y, are decoded and the extracted message, m' is interpreted based on its knowledge base, BS. In more detail, the received symbol is identified based on the 24-QAM decision regions that are illustrated in Figure 8 (right). The received symbol is then demodulated into corresponding bits, representing one of the 24 letters in the knowledge base. Finally, the identified letter is converted into an image of its ASL representation. If the extracted semantic meaning at the destination, matches the transmitted one, the system achieves semantic equivalence, relying only on the conveyed semantic information rather than the message's syntactic structure.

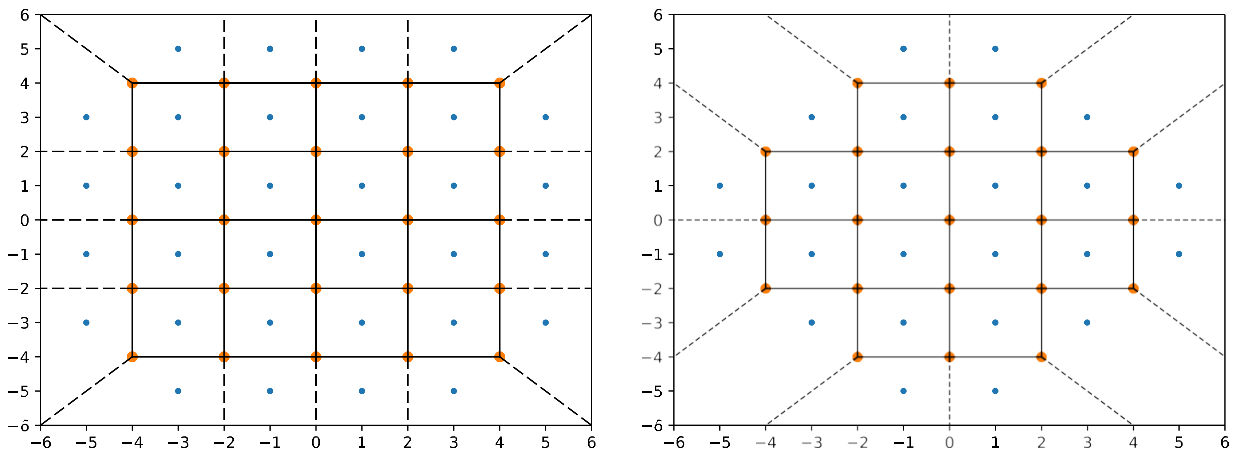

Figure 8. 32-QAM (left) and 24-QAM (right) constellations with decision regions

7.1.2. CNN Model

This section analyses the CNN model illustrated in Figure 9 [148], detailing its architecture, data collection, preprocessing techniques, and performance evaluation. The CNN serves as the semantic encoder within the semantic communication system. Evaluating the model's complexity requires analyzing its architectural design and the total number of parameters. The proposed architecture comprises five convolutional layers, each followed by a pooling layer, with a dense layer after the final convolutional layer. This structure increases the network's depth, contributing to its overall complexity. Additionally, the model's complexity is determined by the number of parameters, including weights and biases, required during training. In this case, the total number of parameters is 616.504.

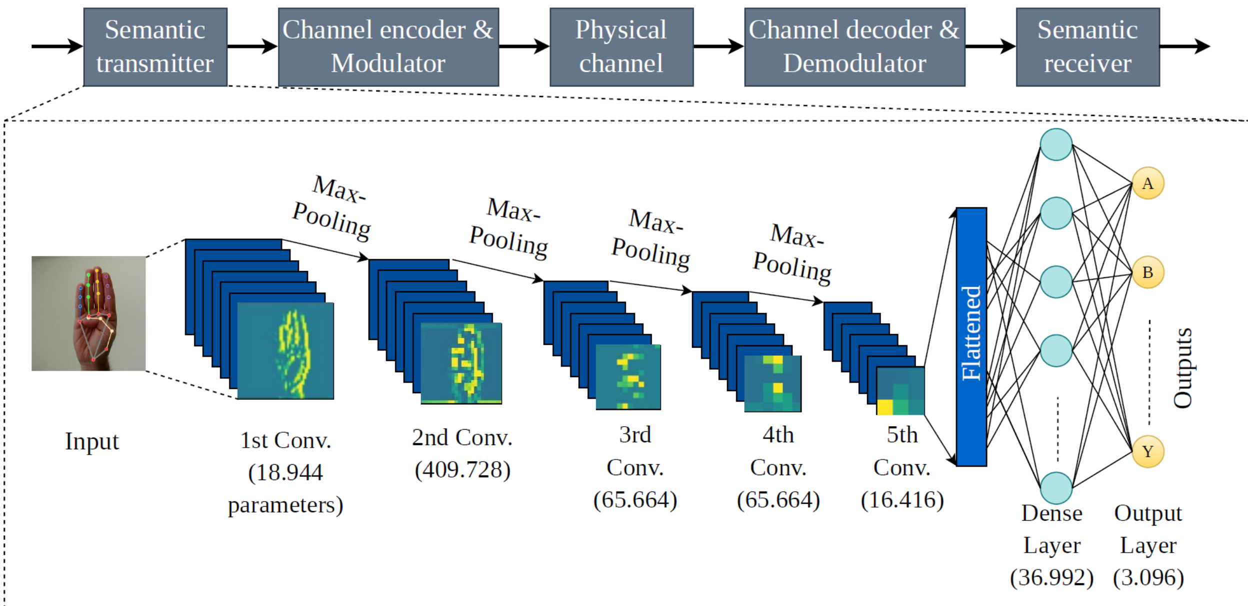

Figure 9. CNN model

7.1.2.1. Data collection and preprocessing

One of the biggest challenges in sign language recognition systems is maintaining high feature recognition efficiency while retaining only the most relevant features. For the ASL alphabet, a key factor is the precise positioning of each finger. According to [149], the meaning of a sign language image can be determined solely by the position of the fingers. Therefore, a dataset was created for this study, consisting exclusively of custom images with key points representing joint positions and landmarks for finger placement [150]. In addition, each finger is associated with a unique colour. This novel approach offers notable advantages in terms of feature extraction from the grayscale images and aims to improve the recognition process of the sign language images, while its implementation relies on the integration of libraries such as OpenCV [151] and MediaPipe [152]. Compared to grayscale datasets like MNIST, incorporating color information enhances the model's ability to distinguish features based on chromatic attributes. This advantage is particularly evident in our case, where each finger is assigned a different color, simplifying position representation. Notably, the model achieves high-performance metrics with relatively little training data, suggesting that using RGB images promotes a more efficient learning process, requiring fewer samples for superior performance. However, employing RGB images also adds complexity to the model architecture. This is primarily due to the increased number of parameters and the larger model size required for three-channel image representation. After data collection, preprocessing is a crucial step in the training process, as it directly impacts the quality and usability of the dataset. In particular, rescaling and normalization techniques are applied, as illustrated in Figure 10. First, rescaling standardizes image sizes to ensure consistency across the dataset. Specifically, image resolution is reduced from 400×400 to 100×100. Next, normalization is performed by dividing all pixel values by 255, scaling the CNN inputs to a range between 0 and 1. This approach enhances regularization, reduces the computational load during training, and helps prevent overfitting by preventing the model from relying too heavily on specific pixel values. Overall, these preprocessing techniques improve computational efficiency and consistency while reducing training time.

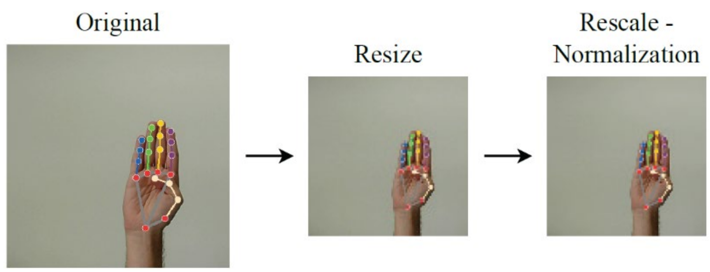

Figure 10. Data preprocessing procedure

7.1.2.2. Convolution layers

The selected ML model is based on a CNN, a deep learning architecture that reduces image size while preserving key features such as shapes, edges, and textures. The convolutional layers, the core components of a CNN, enable it to learn patterns and relationships within the data. As illustrated in Figure 11, the convolution process involves applying a small matrix of numbers, known as a kernel, to the image. This operation extracts essential features, such as edges, textures, and patterns, forming a feature map that retains only the most relevant information from each applied kernel.

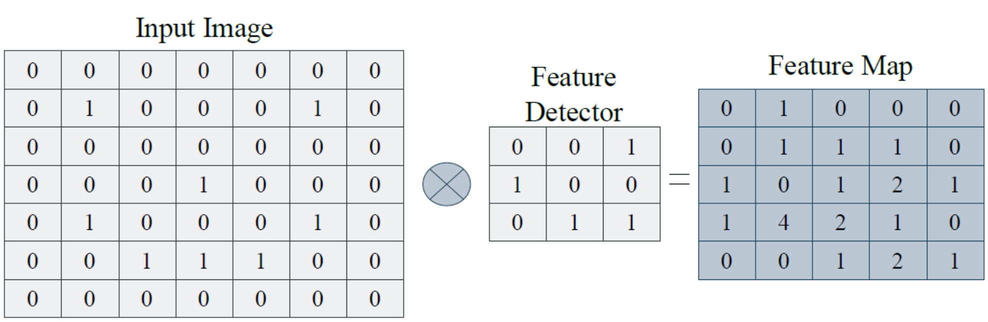

Figure 11. Kernel and Feature map

To calculate the pre-nonlinearity input for a specific unit of the convolution layer, it is necessary to aggregate the contributions from the cells in the previous layer, considering the weights assigned to each filter component [153].

Thus, the input is given by

$$x_{ij}^{\ell} = \sum_{a=0}^{m-1} \sum_{b=0}^{m-1} \omega_{ab} y_{(i+a)(j+b)}^{\ell-1} , \qquad (12)$$

with x and y denote the input and the output of the unit, respectively. Also, ω is the m-dimensional filter that is utilized. Thus, the non-linearity of the convolutional layer can be expressed as

$$y_{ij}^{\ell} = \sigma\left(x_{ij}^{\ell}\right) , \qquad (13)$$

in which σ represents the rectified linear unit (ReLU) activation function. The primary purpose of ReLU is to bring non-linear behaviour into the model, improving overall prediction accuracy. In simple terms, if the input to a neuron is greater than zero, the value is passed forward. Otherwise, the output is set to zero. This behaviour can be written as

$$\sigma(x) = x^{+} = max(0, x) , \qquad (14)$$

with its derivative given by

$$\sigma'(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x < 0 \end{cases} . \qquad (15)$$

The error we need to calculate for the previous layer is the partial derivative of the error function E with respect to each neuron's output. By applying the chain rule, we can then find the gradient component for each weight, which can be written as

$$\frac{\partial E}{\partial \omega_{ab}} = \sum_{i=0}^{N-m} \sum_{j=0}^{N-m} \frac{\partial E}{\partial x_{ij}^{\ell}} \frac{\partial x_{ij}^{\ell}}{\partial \omega_{ab}} = \sum_{i=0}^{N-m} \sum_{j=0}^{N-m} \frac{\partial E}{\partial x_{ij}^{\ell}} y_{(i+a}^{\ell-1} . \qquad (16)$$

In addition, contributions from every expression that includes the variable of interest must be combined. To find the gradient, we need the partial derivatives, which are expressed as

$$\frac{\partial E}{\partial x_{ij}^{\ell}} = \frac{\partial E}{\partial y_{ij}^{\ell}} \frac{\partial y_{ij}^{\ell}}{\partial x_{ij}^{\ell}} , \qquad (17)$$

or equivalent

$$\frac{\partial E}{\partial x_{ij}^{\ell}} = \frac{\partial E}{\partial y_{ij}^{\ell}} \frac{\partial}{\partial x_{ij}^{\ell}} \left(\sigma\left(x_{ij}^{\ell}\right)\right) , \qquad (18)$$

or

$$\frac{\partial E}{\partial x_{ij}^{\ell}} = \frac{\partial E}{\partial y_{ij}^{\ell}} \sigma'\left(x_{ij}^{\ell}\right) . \qquad (19)$$

From the earlier equation, it is clear that knowing the error at the current layer allows us to compute the partial derivatives by simply using the derivative of the activation function. Finally, to update the weights for each convolution layer, we must backpropagate the errors to the previous layer, which can be written as

$$\frac{\partial E}{\partial y_{ij}^{\ell-1}} = \sum_{a=0}^{m-1} \sum_{b=0}^{m-1} \frac{\partial E}{\partial x_{(i-a)(j-b)}^{\ell}} \frac{\partial x_{(i-a)(j-b)}^{\ell}}{\partial y_{ij}^{\ell-1}}, \quad (20)$$

or equivalently

$$\frac{\partial E}{\partial y_{ij}^{\ell-1}} = \sum_{a=0}^{m-1} \sum_{b=0}^{m-1} \frac{\partial E}{\partial x_{(i-a)(j-b)}^{\ell}} \omega_{ab} \,. \quad (21)$$

7.1.2.3. Pooling layers

Each convolutional layer is followed by a pooling layer, a standard practice in CNNs that reduces dimensionality while preserving the most important features of the image. Pooling layers help introduce translation invariance into the model, meaning the model can recognize objects in an image regardless of their position. Achieving translation invariance greatly improves the model's predictive performance, as it removes the need for objects to appear in specific locations within the images. This eliminates the requirement to train the model on a vast set of images showing the object in every possible position. The most commonly used pooling techniques are max pooling and average pooling, with max pooling being the method chosen for the proposed architecture. As illustrated in Figure 12, depending on the selected kernel size applied to the image, the pixel with the largest (most important) value is extracted.

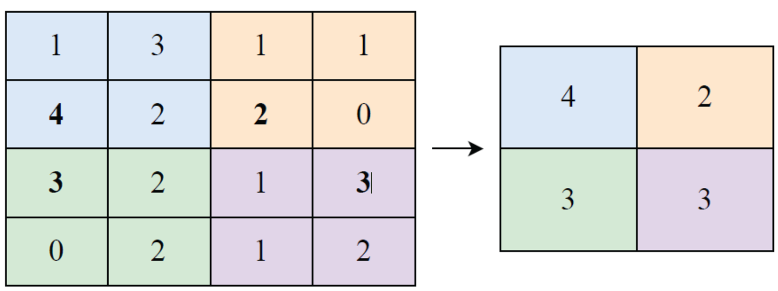

Figure 12. Max pooling process

If an object is within any examined region, it will be detected unless other elements with higher pixel values are present. However, if objects consistently appear in a small area, max pooling may not effectively achieve translation invariance. Conversely, when the dataset includes sufficient variability in object positions, max pooling can greatly enhance model performance. This was considered during the creation of the RGB dataset to ensure diverse object placements, improving the model's generalization ability.

7.1.2.4. Fully connected layers

After the final pooling layer, a fully connected (dense) layer follows, where each neuron is linked to every neuron in the previous and next layers. The primary role of this dense layer is to generalize relationships between extracted features and recognize global patterns. To achieve this, the SoftMax activation function is used to transform the model's outputs into a probability vector, ensuring each output represents the likelihood of the input belonging to a specific class. The SoftMax activation function is given by

$$\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}, \text{ for } i = 1,2,\ldots,K \,, \qquad (22)$$

where z_i represents the output of i neuron, and K denotes the total number of neurons in the layers.

7.1.2.5. Model training

For the model's optimizer, the Adam optimizer is utilized, which is a stochastic gradient descent (SGD) method based on adaptive estimation of first and second order moments. This technique has been proved to outperform other optimizers, such as SGD and adaptive gradient in DNN, while requiring less memory [154]. The Adam optimization algorithm adjusts individual learning rates for each parameter, accelerating convergence and improving the final solution. It performs well in scenarios with noisy gradients, such as deep learning training with mini-batches, while maintaining memory efficiency by using only two additional variables per parameter. Additionally, Adam is relatively insensitive to hyperparameter selection, making it a convenient choice for the developed CNN architecture.

Algorithm 1 Adam optimizer

$m_0 \leftarrow 0$
$v_0 \leftarrow 0$
$t \leftarrow 0$
while θ_t not converged **do**
  $t \leftarrow t + 1$
  $g_t \leftarrow \nabla_\theta L_{CE}\left(\theta_{t-1}\right)$
  $m_t \leftarrow \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t$ (
  $v_t \leftarrow \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2$
  $\widehat{m}_t \leftarrow m_t / \left(1 - \beta_1^t\right)$
  $\widehat{v}_t \leftarrow v_t / \left(1 - \beta_2^t\right)$
  $\theta_t \leftarrow \theta_{t-1} - \alpha \cdot \widehat{m}_t / \left(\sqrt{\widehat{v}_t} + \epsilon\right)$
  return θ_t
end while

The implementation of Adam's optimizer that was used is presented in Algorithm 1, where θ_t represents the parameter vector at time t, a is the step size, while β_1 and β_2 are the exponential decay rates for the moment estimates. In addition, L_{CE} denotes the loss function, which in our case is the sparse categorical cross-entropy. This function is used to measure the difference between the predicted probabilities and the category labels, making it suitable for multi-class classification tasks. A key benefit of sparse categorical cross-entropy is that it handles imbalanced datasets effectively and is compatible with the chosen Softmax activation function in the fully connected layer. The loss function can be expressed as

$$L_{CE} = -\sum_{i=1}^{n} \psi_i \log\left(p_i\right) , \qquad (23)$$

with n is the number of classes, ψ_i represents the truth label, and pi denotes the Softmax probability for the ith class. During its execution, initially g_t, m_t, and v_t are computed, which represent the gradients regarding the stochastic objective as well as the updated first and second moment estimations, respectively. Afterward, the bias-corrected first and second moments are calculated to get the updated parameter vector, which serves as the output. By thoroughly examining each step in the training process, we can determine the total time complexity of the Adam optimizer. In practice, the time complexity is often approximated by

$$C = E(G + P + R)\,, \qquad (24)$$

where, E represents the number of training epochs, while G, P, and R, denote the gradient computation, parameter updates, and running averages and moments, respectively.

7.1.3. Numerical Results

This section presents an evaluation of the CNN architecture and the overall E2E SemComs system. Specifically, the analysis utilizes three datasets to provide numerical results on CNN training and testing, as well as the communication channel's quality and efficiency in terms of data and energy usage. Additionally, a detailed breakdown of the model's outputs is provided, including accuracy scores, precision-recall metrics, and F1-scores. Finally, the results are complemented by insightful discussions that highlight the strengths and limitations of the proposed system.

7.1.3.1. Datasets

Table 2. Datasets characteristics

Dataset	Training images	Testing images	Image size
MNIST	27.455	7.172	28x28
LEXSET	21.600	2.400	100x100
RGB	10.490	1.800	100x100

As presented in Table 2, three different datasets were used for training the CNN architecture. Specifically:

- MNIST: The MNIST dataset has been selected as a standard for ASL-based tasks [155]. It contains 24 categories with each representing a letter of the alphabet1. The images are 28x28, while their total number is 34.627, which was split into 27.455 for training and 7.172 testing.
- LEXSET: The LEXSET dataset contains 27 classes, 24 of which were used to ensure compatibility with the other utilized datasets [156]. It includes 21.600 training images and 2.400 testing images, each with dimensions of 512×512. To achieve better performance, the images were downscaled to 100×100.
- RGB: This dataset was develop using coloured landmarks and key-points to provide a better presentation of the hand's structure. The size of the captured images is 400x400. For the training

process, they were resized to 100x100. There are 12.290 images in total, split into 10.490 for training and 1.800 for testing.

Table 3. Training, validation, and testing accuracy

Dataset	Training	Validation	Testing
MNIST	99.61%	100%	97.12%
LEXSET	98.39%	98.52%	98.95%
RGB	99.21%	99.81%	99.72%

7.1.3.2. Training performance

In this section, we present the accuracy of the CNN model across the training phase. Figure 13 illustrates how both training and validation accuracy evolve over multiple epochs for the three different datasets. The accuracy metric, defined as the ratio of correct predictions to the total number of predictions, can be expressed mathematically as

$$A = \frac{T_p+T_n}{T_p+T_n+F_n+F_n}, \qquad (25)$$

where T and F indicate whether predictions are true or false, respectively, while the subscripts n and p represent the negative or positive nature of the prediction. Additionally, solid lines show the training accuracy, and dotted lines represent the validation accuracy. From this figure, we can see that all three datasets share a steady learning pattern, showing gradual improvements without any major fluctuations. In addition, it can be observed that 90% accuracy is achieved after 8 epochs, while beyond the 25th epoch, all three datasets consistently surpass the 97% accuracy threshold for both training and validation. However, the plot reveals notable differences between the datasets as well. Specifically, as illustrated in Table 3, using the MNIST dataset we have a training accuracy of 99.61% and validation of 100% after 90 epochs. It is important to highlight that the MNIST dataset demonstrated the quickest improvement in validation accuracy but required the most epochs to reach its optimal performance on the evaluation dataset. The LEXSET dataset achieved a training accuracy of 98.39% and a validation accuracy of 98.52% after 60 epochs. For the RGB dataset, the model reached a training accuracy of 99.21% and a validation accuracy of 99.81% after 80 epochs. Overall, since the validation and training accuracies are very close, with the validation accuracy never exceeding the training accuracy, there is no evidence of overfitting. This suggests that the models do not simply memorize the training data but are capable of generalizing well to unseen data.

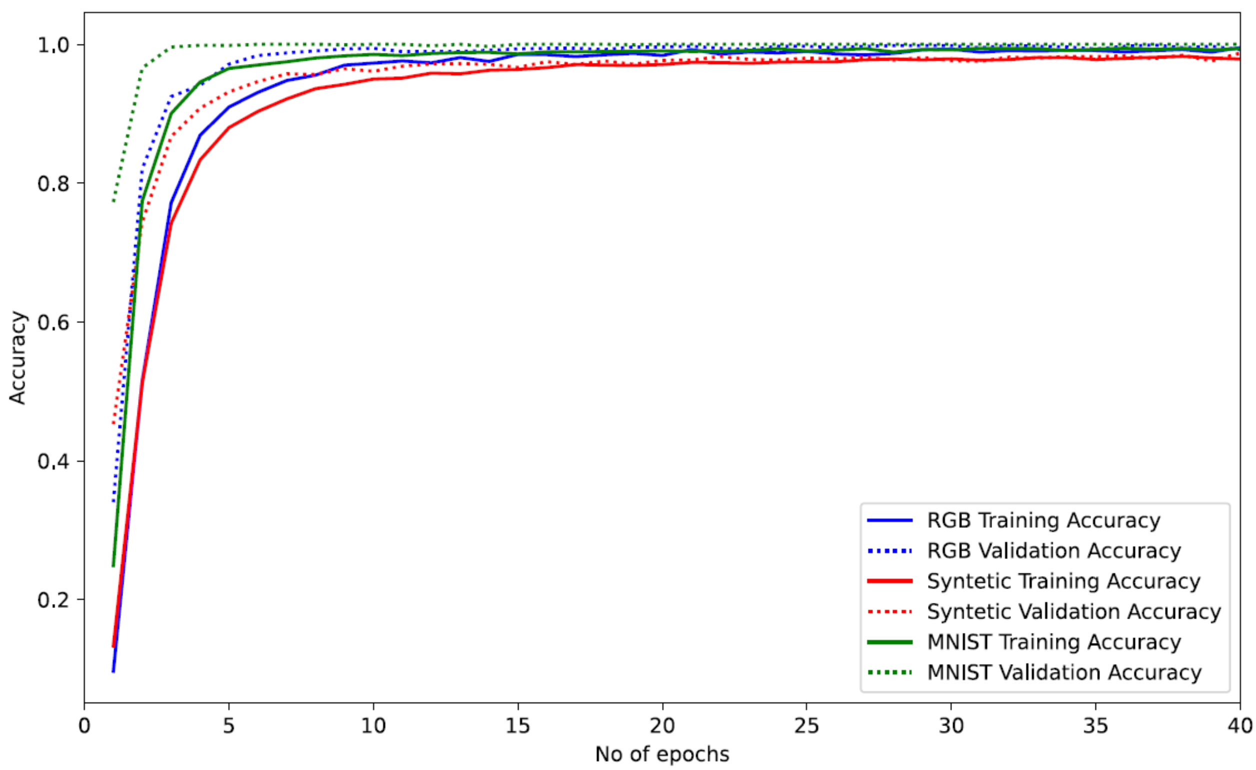

Figure 13. Training and validation accuracy over epochs

The same takeaways can be seen in Figure 14, which shows how the training and validation losses changed over time (epochs) for the three datasets. The MNIST dataset had a training loss of 0.0152 and a validation loss of $1.3588\text{x}10^{-8}$, the LEXSET dataset had a training loss of 0.0533 and a validation loss of 0.0551, and the RGB dataset had a training loss of 0.0354 and a validation loss of 0.00046. What's important is that for all three datasets, the gap between training and validation losses is very small. This means the models are good at generalizing to new, unseen data.

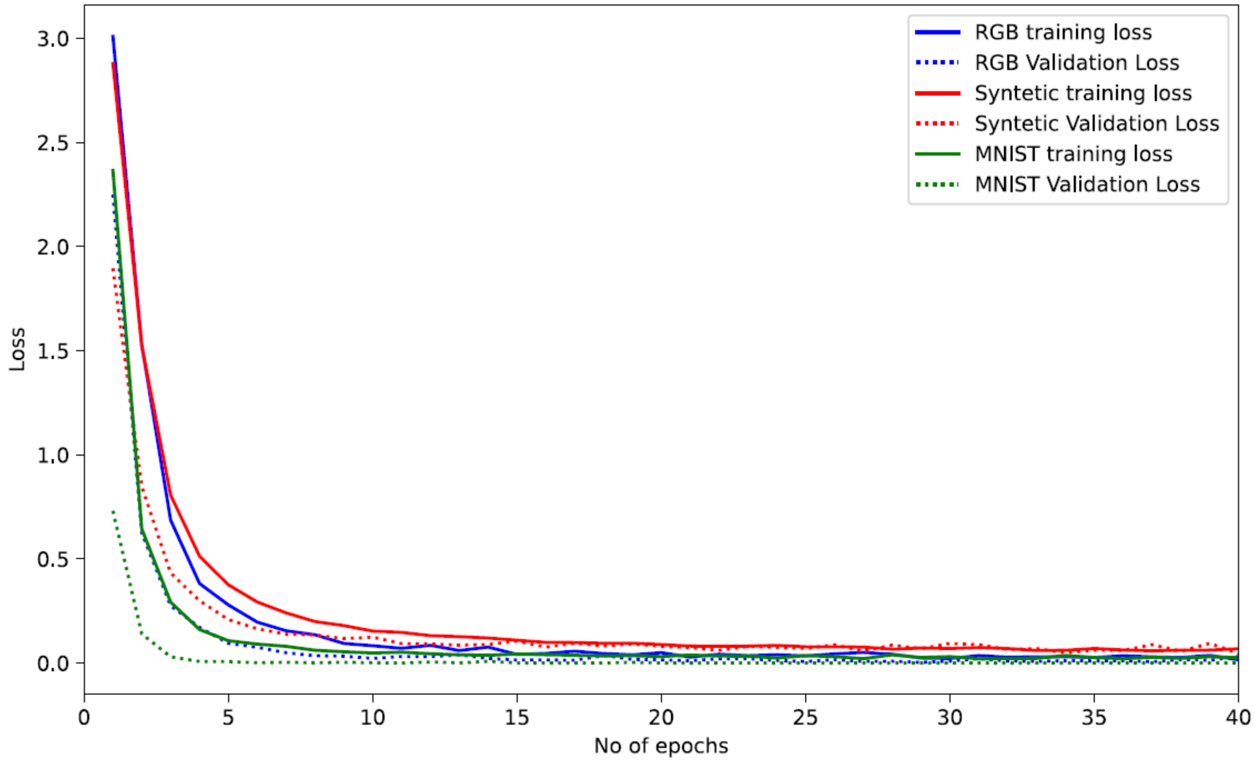

Figure 14. Training and validation loss over epochs

7.1.3.3. Testing performance

Figure 15 shows the confusion matrices for each dataset used in the study. These matrices display the true labels compared to the predicted labels in a grid format, where each cell indicates the number of images classified under a specific label. All three trained models perform well in extracting semantic information from the images, but the RGB model stands out, delivering higher prediction accuracy and requiring less data. For example, the RGB model achieves 99% testing accuracy with around 10.490 images, while the other models need more than 20.000 images to reach slightly lower accuracy levels. Additionally, the RGB model has the fewest errors, with 20 out of 24 classes achieving perfect prediction rates. The LEXSET model performs similarly but with slightly lower accuracy, as only 8 out of 24 categories have 100% correct predictions. In contrast, the MNIST model shows the most significant discrepancies between predicted and actual labels.

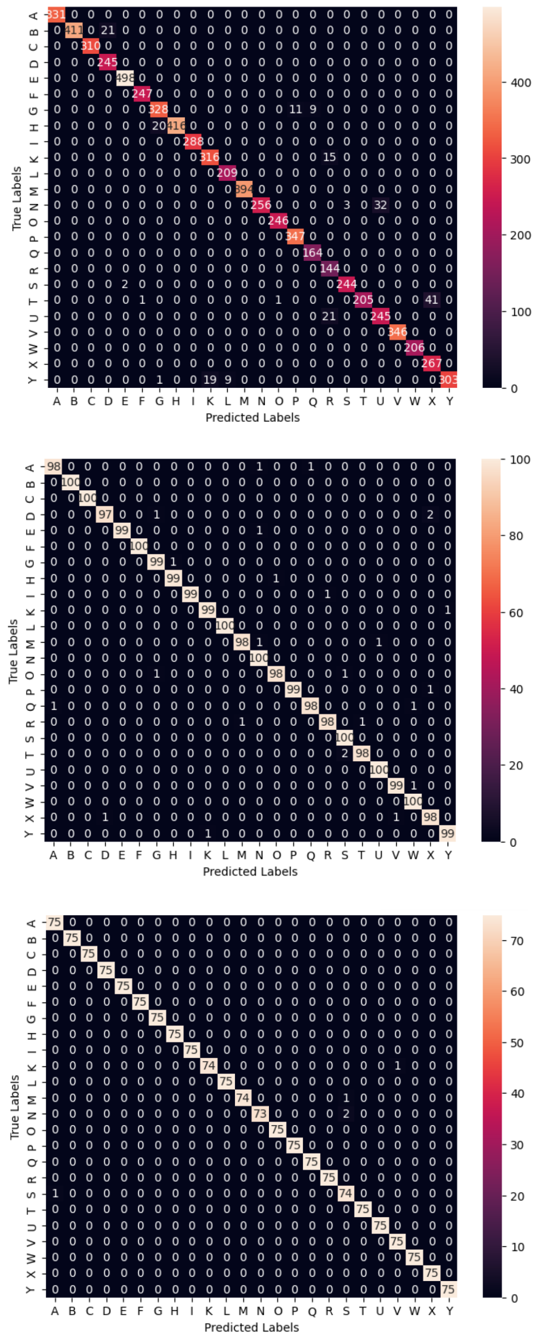

Figure 15. Confusion matrices of the models trained with the MNIST, LEXSET, and RGB datasets

Table 4. Precision, recall, and F1 for the models trained with the datasets

	(a) MNIST			(b) LEXSET			(c) RGB		
Class	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
0	100.00%	100.00%	100.00%	98.99%	98.00%	98.49%	98.68%	100.00%	99.34%
1	100.00%	95.14%	97.51%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
2	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
3	92.11%	100.00%	95.89%	98.98%	97.00%	97.98%	100.00%	100.00%	100.00%
4	99.60%	100.00%	99.80%	100.00%	99.00%	99.50%	100.00%	100.00%	100.00%
5	99.60%	100.00%	99.80%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
6	93.98%	94.25%	94.12%	98.02%	99.00%	98.51%	100.00%	100.00%	100.00%
7	100.00%	95.41%	97.65%	99.00%	99.00%	99.00%	100.00%	100.00%	100.00%
8	100.00%	100.00%	100.00%	100.00%	99.00%	99.50%	100.00%	100.00%	100.00%
9	94.33%	95.47%	94.89%	99.00%	99.00%	99.00%	100.00%	98.67%	99.33%
10	95.87%	100.00%	97.89%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
11	100.00%	100.00%	100.00%	98.99%	98.00%	98.49%	100.00%	98.67%	99.33%
12	100.00%	87.97%	93.60%	97.09%	100.00%	98.52%	100.00%	97.33%	98.65%
13	99.60%	100.00%	99.80%	98.99%	98.00%	98.49%	100.00%	100.00%	100.00%
14	96.93%	100.00%	98.44%	100.00%	99.00%	99.50%	100.00%	100.00%	100.00%
15	94.80%	100.00%	97.33%	98.99%	98.00%	98.49%	100.00%	100.00%	100.00%
16	80.00%	100.00%	88.89%	98.99%	98.00%	98.49%	100.00%	100.00%	100.00%
17	98.79%	99.19%	98.99%	97.09%	100.00%	98.52%	96.10%	98.67%	97.37%
18	100.00%	82.66%	90.51%	98.99%	98.00%	98.49%	100.00%	100.00%	100.00%
19	88.45%	92.11%	90.24%	99.01%	100.00%	99.50%	100.00%	100.00%	100.00%
20	100.00%	100.00%	100.00%	99.00%	99.00%	99.00%	98.68%	100.00%	99.34%
21	100.00%	100.00%	100.00%	98.04%	100.00%	99.01%	100.00%	100.00%	100.00%
22	86.69%	100.00%	92.87%	97.03%	98.00%	97.51%	100.00%	100.00%	100.00%
23	100.00%	91.27%	95.43%	99.00%	99.00%	99.00%	100.00%	100.00%	100.00%

Table 4 shows the precision, recall, and F1 scores for each class, comparing models trained on the MNIST, LEXSET, and RGB datasets. These scores measure how well the models performed during testing. Specifically, the utilised metrics are

$$P = \frac{T_p}{T_p + F_p}, \quad (26)$$

$$R = \frac{T_p}{T_p + F_n}, \quad (27)$$

and

$$F1 = \frac{2PR}{P+R} = \frac{2 * T_p}{2 * T_p + F_p + F_n}. \quad (28)$$

It is clear that the model trained on the RGB dataset performs the best. It hits perfect scores across nearly all classes for every metric, showing it not only classifies samples correctly but also nails true positives without letting any false positives slip through. The LEXSET-trained model doesn't do as well, especially in precision and F1 scores, though it still outperforms the MNIST-trained model. The MNIST model comes in last among the three but still manages to deliver decent, reliable predictions. Overall, all three models perform well enough to meet the bar for successful classification.

7.1.3.4. Communication performance

Figure 16 illustrates the normalized power and bit requirements as the number of transmitted letters increases for both conventional and SemCom systems. In conventional systems, transmitting a single letter involves sending an entire image, which amounts to roughly $3x10^6$ bits. In contrast, the SemCom system uses a semantic encoder to identify the letter in the image, reducing the data to just 5 bits for transmission over the physical channel. As the number of letters increases, both systems see a rise in the number of bits and power consumed. However, the conventional system demands far more data and power compared to the SemCom system for the same task, showcasing significant improvements in data and power efficiency. For example, the SemCom system requires only 10 normalized power units to transmit the 5 bits representing a single letter, while the conventional system needs 10^7 normalized power units to transmit $3x10^6$ bits. Overall, when efficiency is measured as the ratio of semantic information to the required data and power, the proposed SemCom system achieves $6x10^7$ times greater efficiency in both data and power usage.

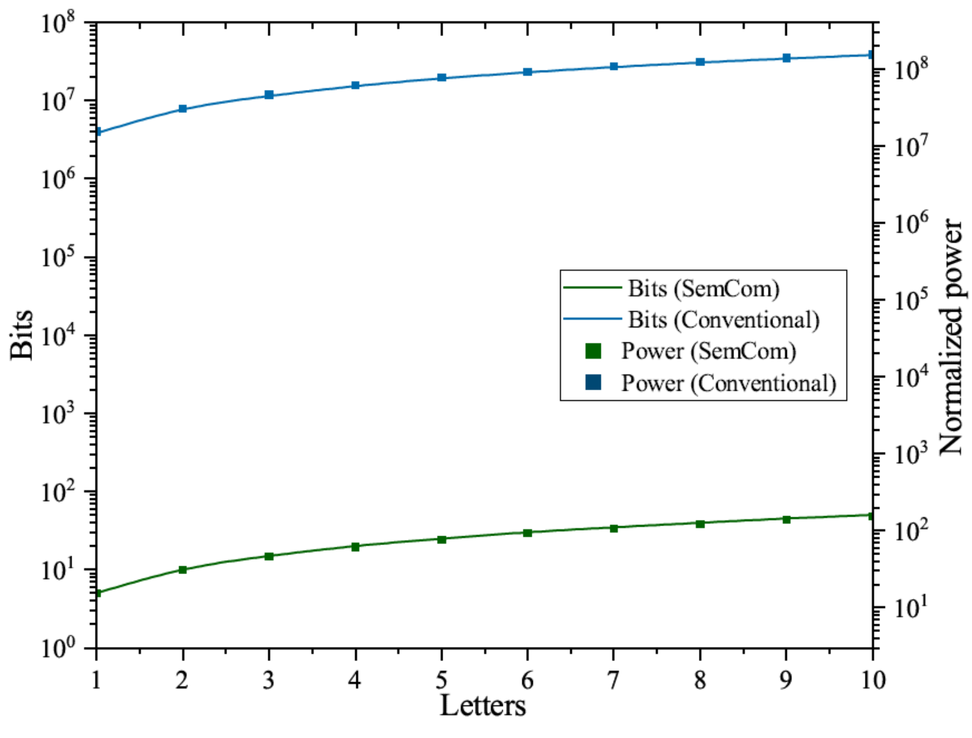

Figure 16. Data and power relative to semantic information

Figure 17 shows the BER and SER as a function of the SNR for conventional systems using 32QAM and 24QAM modulation schemes, compared to the E2E SemCom system. The first two systems rely on conventional communication channels that do not extract semantic information from the data, while the E2E SemCom system uses the proposed architecture to extract and transmit only the semantic information through the physical channel. In the figure, solid lines represent BER, and dotted lines represent SER. As expected, both BER and SER decrease as SNR increases. The results also show that the 24QAM modulation scheme consistently outperforms 32QAM in terms of both BER and SER across all SNR values, with the performance gap widening as SNR increases. For example, at an SNR of 16, switching from 32QAM to 24QAM reduces BER by 55%, from 2×10^{-2} to 9×10^{-3}.

Another key observation is that the SemCom system significantly outperforms conventional systems in low SNR conditions. However, in high SNR scenarios, a balance emerges between error performance and power/data efficiency. For SNR values above 20, the BER and SER of the SemCom system stabilize at around 10^{-3} and 2×10^{-3}, respectively, while conventional systems continue to improve their error rates. This is due to the semantic nature of the proposed system, where errors in the semantic encoder can lead to incorrect symbol transmission. These semantic errors propagate through the system, causing additional errors at the bit and symbol levels. This is supported by the fact that the CNN-based semantic encoder has an accuracy of approximately 99.7%, meaning its error rate is on the order of 10^{-3}.

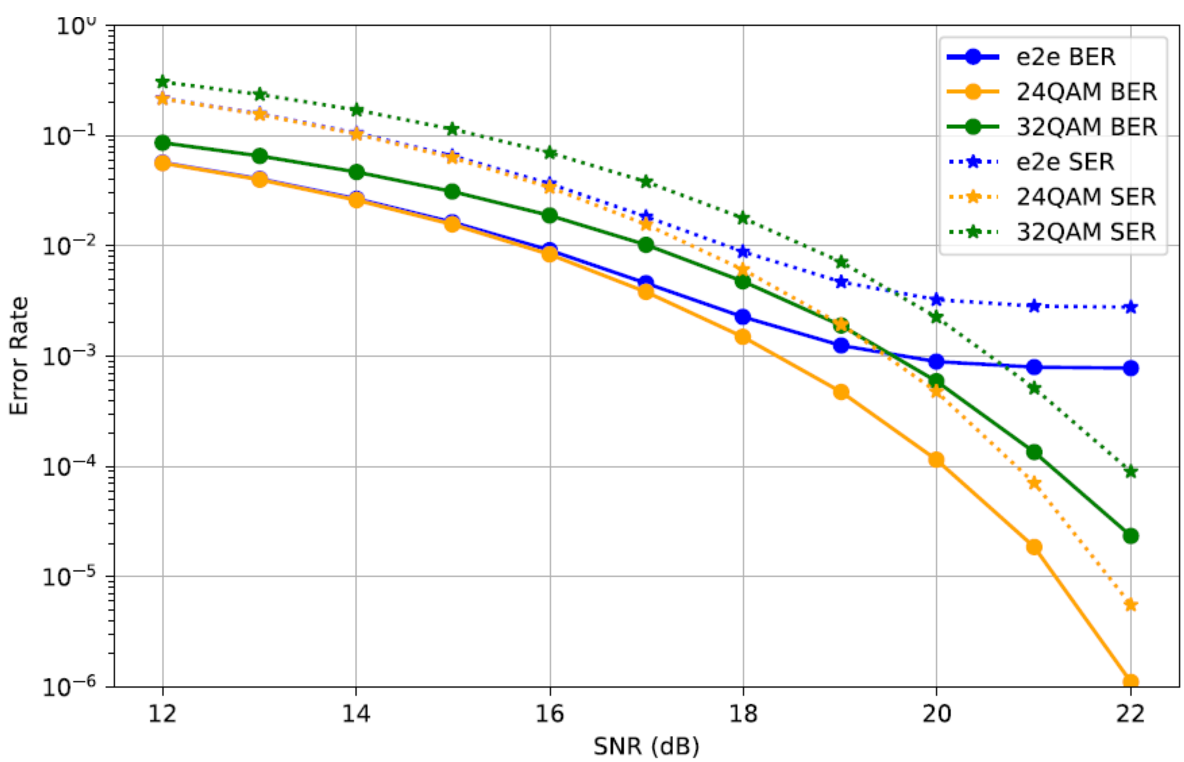

Figure 17. BER and SER vs SNR

Figure 18 shows the relationship BER and SNR under different channel conditions. The figure includes multiple curves, each representing BER performance for three channel models (Rayleigh, Rician, and AWGN) across three scenarios: 32QAM, 24QAM, and the E2E semantic communication system. At an SNR of 12 dB, all channels and scenarios start with a similar BER of around 10^{-1}. As the SNR increases, the BER decreases consistently across all channels and scenarios, demonstrating how stronger signal strength reduces errors. However, the Rayleigh channel shows the smallest improvement in BER as SNR rises. Even at 22 dB, its BER barely drops below 10^{-2}, unlike the other channels. In contrast, the Rician channel consistently reduces BER to approximately 10^{-3} across all scenarios. The AWGN channel performs best, with the 24QAM configuration achieving the lowest BER, nearly reaching 10^{-6} at 22 dB. Meanwhile, the E2E scenario, which incorporates a CNN model, stops improving at a BER of 10^{-3}, reflecting the limitations introduced by the CNN.

These results highlight the varying performance of different channel models under different QAM scenarios and SNR conditions. The unique characteristics of each channel provide valuable insights into their behaviour in communication systems operating under diverse conditions. A key trend emerges: as we move from the idealized AWGN channel to more realistic scenarios like Rayleigh and Rician, the BER remains relatively high, typically around 10^{-3}. This observation supports the claim that the proposed SemCom framework, with its integrated CNN model, maintains consistent performance in practical settings without significantly limiting the system's overall effectiveness.

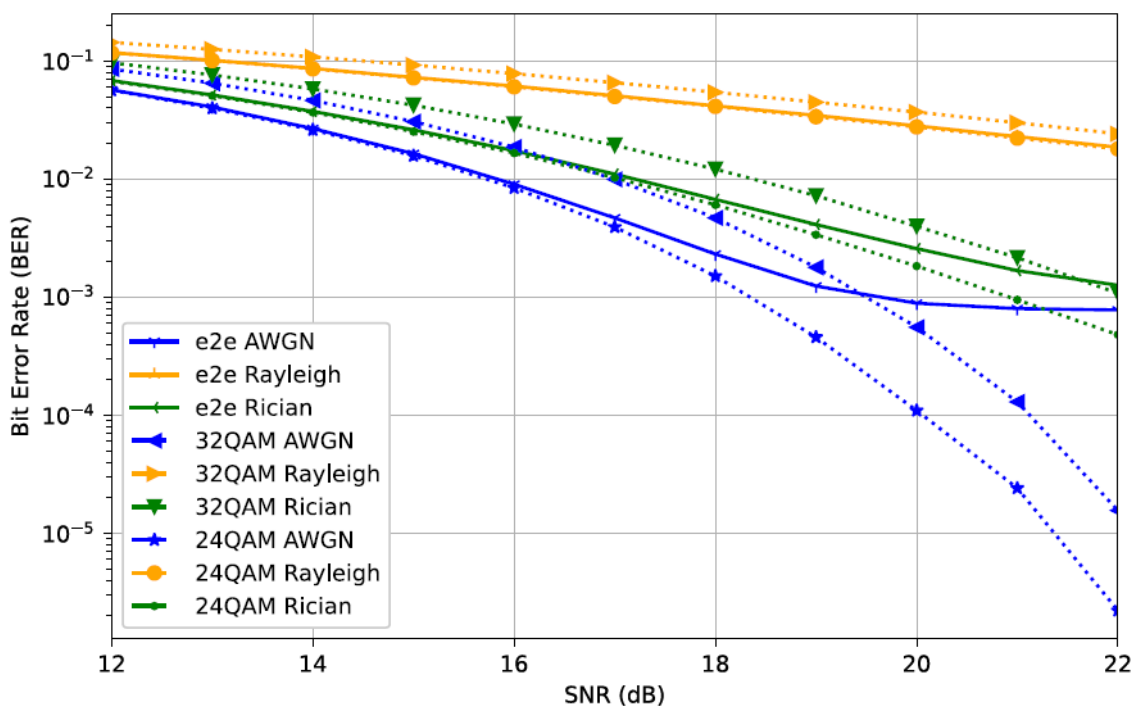

Figure 18. BER vs SNR under different channel conditions

7.1.4. ASL Transmission Conclusions

This section introduced a SemCom architecture tailored for transmitting ASL. It proposed a new system model for image-based SemCom, utilizing a modified version of the QAM technique called 24QAM. This was achieved by removing 8 peripheral symbols from the standard 32QAM constellation. Additionally, a novel semantic encoding approach was presented, using a CNN optimized for the ASL alphabet. To enhance the system, an original dataset was created by improving hand posture representation through RGB landmarks and key-points. The training, testing, and communication performance of the proposed system were evaluated using numerical results, highlighting its advantages and revealing a balance between conventional communication systems and the proposed SemCom framework.

The semantic communication approach discussed in this section focuses on understanding the meaning behind information and aims to maximize the benefits of integrating information creation, transmission, and usage. By adopting such methods, it becomes possible to prioritize and transmit only the most relevant and critical data between the source and destination. This ensures that the end user receives timely, useful, and actionable information necessary to achieve their goals. Leveraging semantics in communication is expected to significantly reduce energy consumption and optimize the use of network resources. Ultimately, semantic networking will support the scalability of future intelligent, interconnected systems.

The transformative potential of semantic communications is highlighted by the following points that also outline future directions that could revolutionize communication and promote inclusivity across various fields:

- The integration of ASL-based semantic communication into smart home devices will empower hearing-impaired individuals to interact with and control their environments more efficiently.
- Advances in ASL-based semantic communication technology are expected to greatly enhance communication for hearing-impaired individuals, offering real-time translation and improving interactions with those who are not fluent in ASL.
- The development of ASL-based semantic communication is set to transform educational environments by improving communication between hearing-impaired students and educators, fostering a more inclusive learning experience.
- Integrating this technology into video conferencing tools will enable individuals with hearing impairments to participate more effectively in virtual meetings, webinars, and online courses, enhancing accessibility.
- This technology will ensure that hearing-impaired individuals can fully participate in professional development sessions, training workshops, and similar activities.
- Future advancements in ASL-based semantic communication will improve customer service interactions, making them more accessible for hearing-impaired customers.
- Progress in this technology will ensure faster and more accurate communication between emergency responders and individuals using ASL during critical situations.

- ASL-based semantic communication represents an innovative tool designed to improve communication for individuals with hearing impairments. This discussion explores its potential impact across various sectors, highlighting its role in promoting inclusivity and breaking down communication barriers.
- Future developments in ASL-based semantic communication will facilitate better communication in legal settings, such as court proceedings, police interactions, and other legal matters involving sign language users.
- Beyond traditional subtitles, this technology will provide an additional layer of accessibility for live broadcasts and situations where pre-recorded captions may not suffice.

7.2. V2x Communications

The evolution of V2x technology in the last few years has changed the transportation industry, offering the opportunity for better, more efficient, and sustainable transportation systems. The V2x technology includes many communications systems, such as vehicle-to-vehicle (V2V) [157], vehicle-to-infrastructure (V2I) [158], and vehicle-to-pedestrian (V2P) [159], creating an ecosystem that enables data transfer between different systems. These areas of V2x have paved the way for intelligent transportation systems (ITS), capable of facing critical challenges in this area, like traffic congestion [160] and energy optimization [161], [162]. In recent years, the concept of a digital twin (DT) has shown promising signs as a solution to enhance and improve V2x systems. A DT model can transfer and replicate a real-world system into a digital version. This transformation gives the opportunity to monitor, optimize and create a realistic simulation of the real-world system. In the context of V2x, DT provides a holistic framework for the simulation of the object's behaviour and a perfect recreation of the environment, making it an ideal method for studying the behaviour of objects and experimentation with various parameters of the system to improve it, without compromising the security and functionality of the original system. Recent research has explored the application of DT in V2x systems in order to study various contexts such as traffic and potentially hazardous situations [163], and the integration of it, in a 5G network [164].

The integration of SemCom between the V2x and DT is one more technology that presents a significant enhancement in their evolution and efficiency. The distinctive difference between traditional communication systems and SemCom is the fact that SemCom focuses on the transmission of important data and the semantic information within the message and not the whole message. This inherent feature of SemCom is particularly useful in V2x models, where the rapid and accurate transmission of only useful data is crucial for their operation. The application of SemCom in V2x offers the opportunity for less bandwidth usage, improvement of the transmitted data and a more accurate representation of the message. Recent studies have shown that semantic communication has the potential to improve vehicular networking in a variety of ways, as in the case of [165] where the authors have presented a SemCom model that enhances the data transmission speed, by taking into account the nature of the transmitted data and tailored communication techniques. Despite significant progress in the integration of DTs and SemCom in

V2x systems, challenges remain in achieving seamless interoperability, real-time scalability, and context-aware adaptability. Moreover, the collaborative use of these technologies is still in its infancy, with limited exploration of their combined capabilities for transforming V2x applications. In NANCY, we address these gaps by presenting a semantic DT model for V2x, which aims to enhance the overall operation. This is accomplished by extracting the essential semantic information from messages using AI.

7.2.1. System Model

The constructed DT, as already discussed in Section 6.3, considers a traffic scene with three cameras, one car, one roadside unit (RSU), and a drone. In this scenario, we assume that a vehicle enters the campus, which is the area of interest, and the car as the rest of the camera devices have access to the DT service. A novel aspect of this research is the utilization of SemCom between camera outputs and DT input. With this approach, we can transmit only the coordinates of detected objects to the DT, instead of sending entire video frames. This efficient use of semantic information reduces the data volume transmitted, ensuring faster processing and better utilization of communication resources. Additionally, by utilizing SemCom, the system can integrate data from multiple sources and create more complex systems.

It is worth noting that all 3 devices have a deactivated SemCom encoder module available. The activation of these encoders is triggered by the corresponding request that the devices will send to the first base station to which they will be connected through task offloading. After the extraction of only the semantic information from the encoders the data is transferred to the edge server for the creation of the DT. In more detail, the blue box highlights the components that are deployed in the edge-to-cloud continuum. The SemCom decoder receives the semantic information from all selected nodes and recreates the DT scene. Through this novel approach, we achieve a higher data efficiency and scalability of the system that enables more advanced and complex V2x models along with the integration of DT.

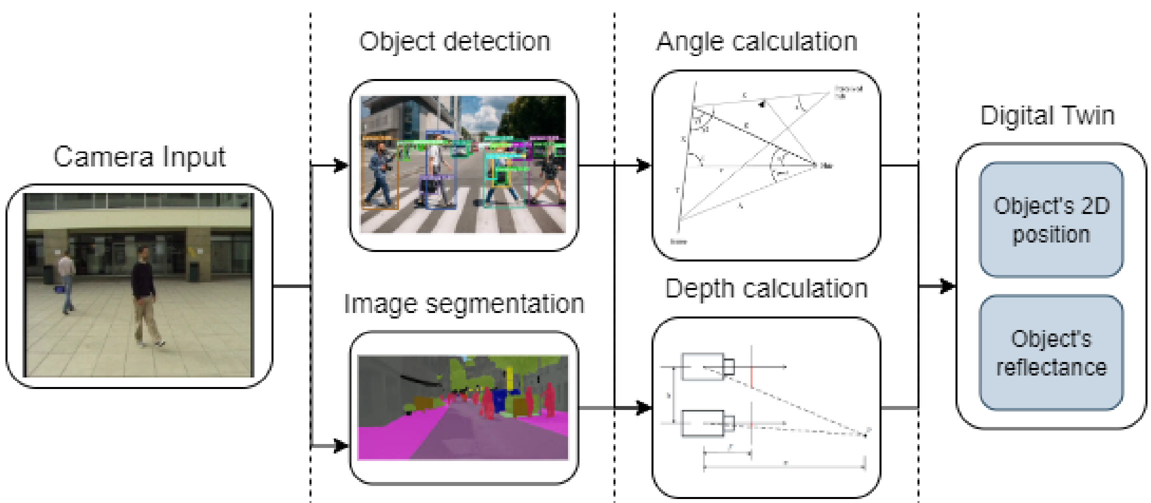

Figure 19. DT creation workflow

7.2.2. Open Dataset

The first version of the V2x SemCom component was created based on an open dataset of two cameras that observed the same area from different angles [166]. As illustrated in Figure 19, for the creation of the DT, four different techniques from the field of computer vision were used. Initially, based on the original image (frame) from each camera, the YOLO model was applied to perform the classification and segmentation of the desired objects (people). Then, the angle and distance of the objects from the centre of the frame were calculated based on the camera's specifications. By transmitting only semantic information (the coordinates of the objects, the viewing angle, and the distance from the camera) and converting the above information into real coordinates, the DT was created. Afterwards, using AI methods, we applied object detection locally in each camera and extracted the necessary semantic information. This process contains the analysis of the frames, the identification and classification of the objects of interest and the extraction of the necessary semantic information from them. Then, only the semantic information will be transferred through the network to the semantic decoder to create the digital twin. In the current scenario, the DT consists of a top view map which includes the two cameras together with the objects they see. In more detail, the objects by the first camera are shown in red while the objects that identified by the second camera are represented in blue. The objects that are jointly identified by the DT service are presented in purple as it can be seen in Figure 20. Finally, a demo of the SemCom DT creation has been uploaded on YouTube [167].

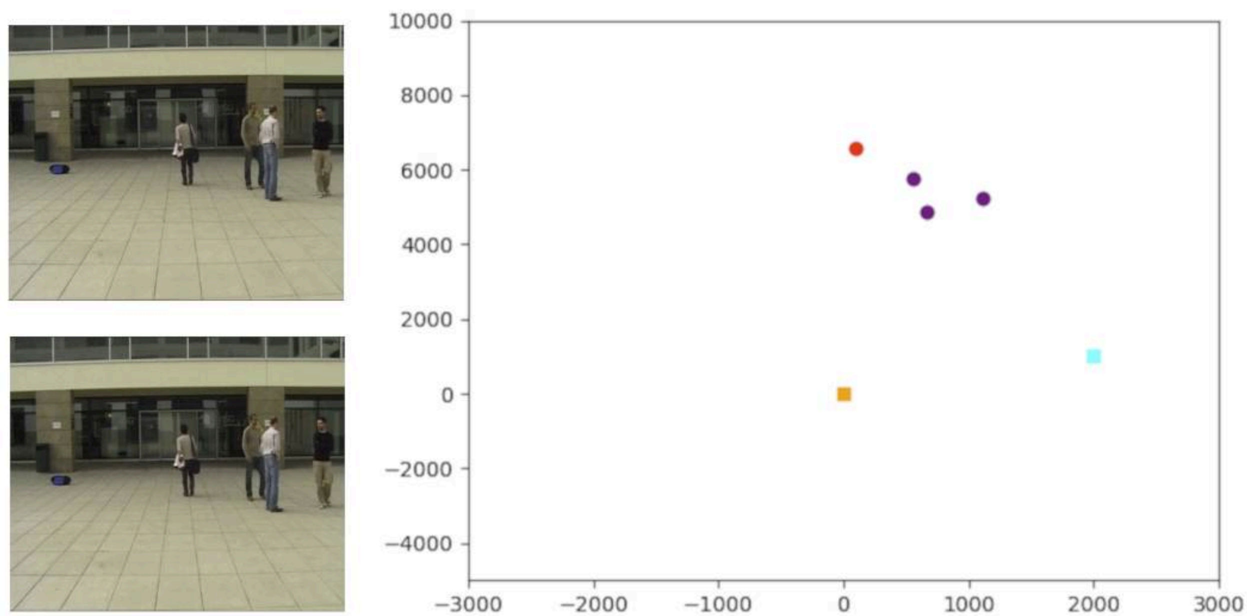

Figure 20. Schematic diagram of SemCom DT creation from the open dataset

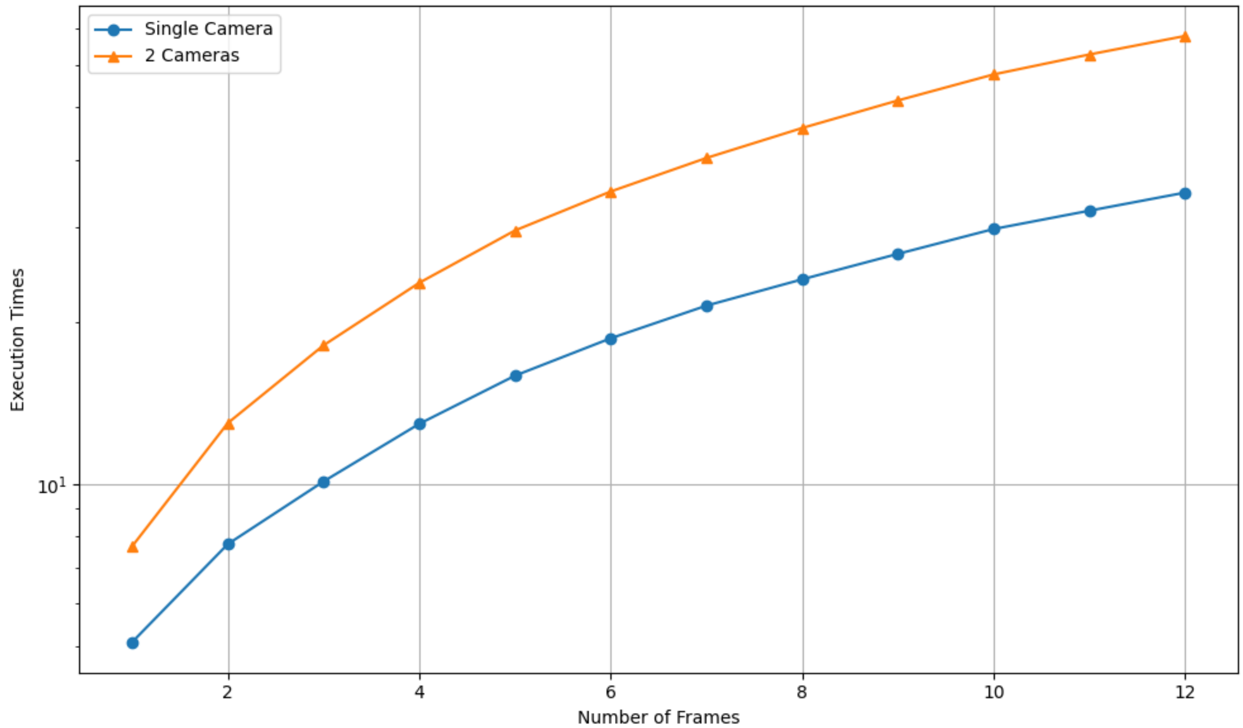

Figure 21. Scalability of SemCom DT from the open dataset

Figure 21 shows the execution time of a Semantic DT in relation to the number of frames for 2 different scenarios. In the first scenario, the DT is generated based on one camera while in the second one, it is generated from two cameras. Both scenarios show similar behavior as a function of the number of frames. Despite their similarities, it is obvious that in the single camera scenario, the execution times are lower than in the two-cameras scenario. It is worth noting that at the beginning of the simulations, both scenarios have similar execution times but as the number of frames increases the distance between the times increases as well.

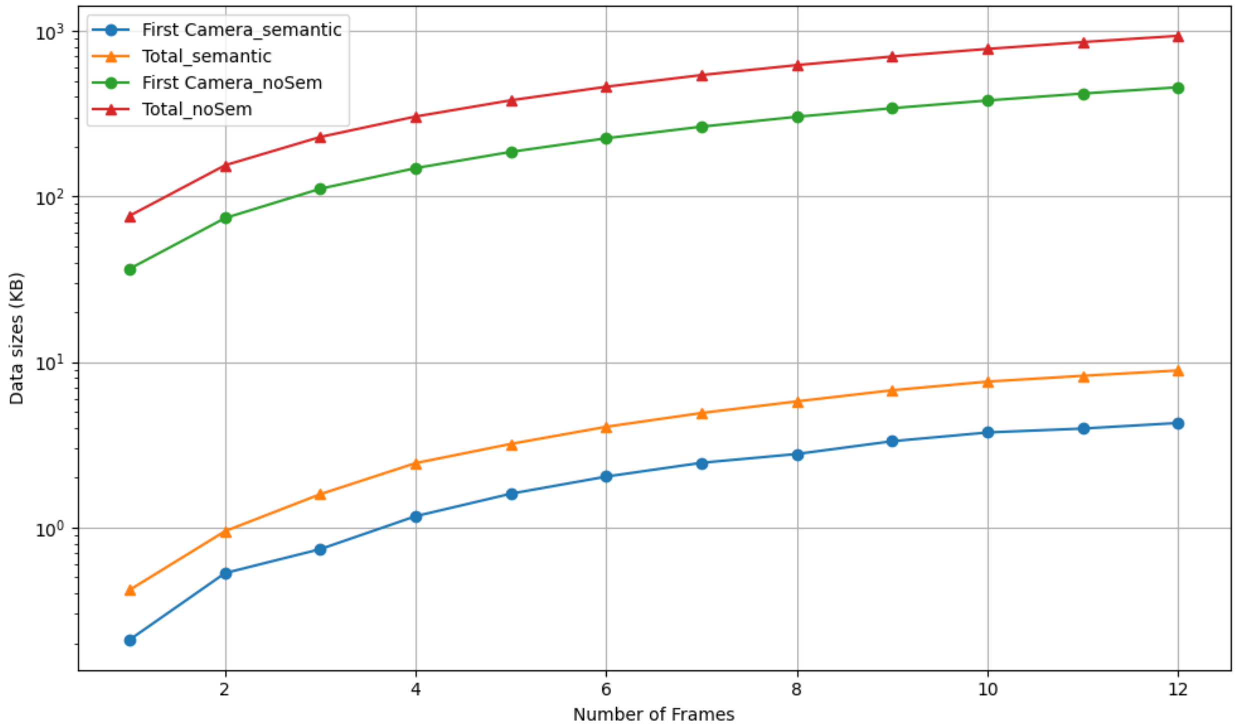

Figure 22. Data efficiency of SemCom DT from the open dataset

Figure 22 shows the amount of data needed by DT in two different scenarios (single camera and two cameras) under two different methods (with and without the use of semantic encoder). Although both methods show the same behaviour, it is obvious that using only semantic information sends significantly less information compared to the method that sends the full image. A typical example is the two-camera scenario where in the semantic version of the model it requires approximately 10 KB to send 12 frames while the corresponding scenario in the classical generation of the DT requires 10^3 KB. This highlights the increased data efficiency of SemCom.

7.2.3. UMU Dataset

A second SemCom model was developed using a custom dataset created by the University of Murcia [168]. This dataset includes video inputs from three different devices operating within the same area, as can be seen in Figure 23. The first camera was mounted on a road vehicle (car) that can provide a dynamic perspective of the scenario, by collecting data from its onboarded camera while moving around the test area. The second camera was placed on an RSU, which offers a stable vantage point ensuring continuous monitoring of a specific area, while at the same time it helps in the creation of a comprehensive and localized digital twin by providing consistent and reliable data from a fixed perspective. Finally, the third camera was mounted on a drone, that provides a flexible angle and a panoramic view, which offers better spatial information about the area while it can help fill the gaps that the other two cameras may have in an area. Afterwards, we use YOLO to recognize objects locally in each camera and extract the essential semantic information, as presented in Figure 23. This technique includes frame analysis, object identification and categorization and semantic information extraction. As can be seen, in this case, the target is to recognize individuals crossing a crosswalk, under regular traffic conditions.

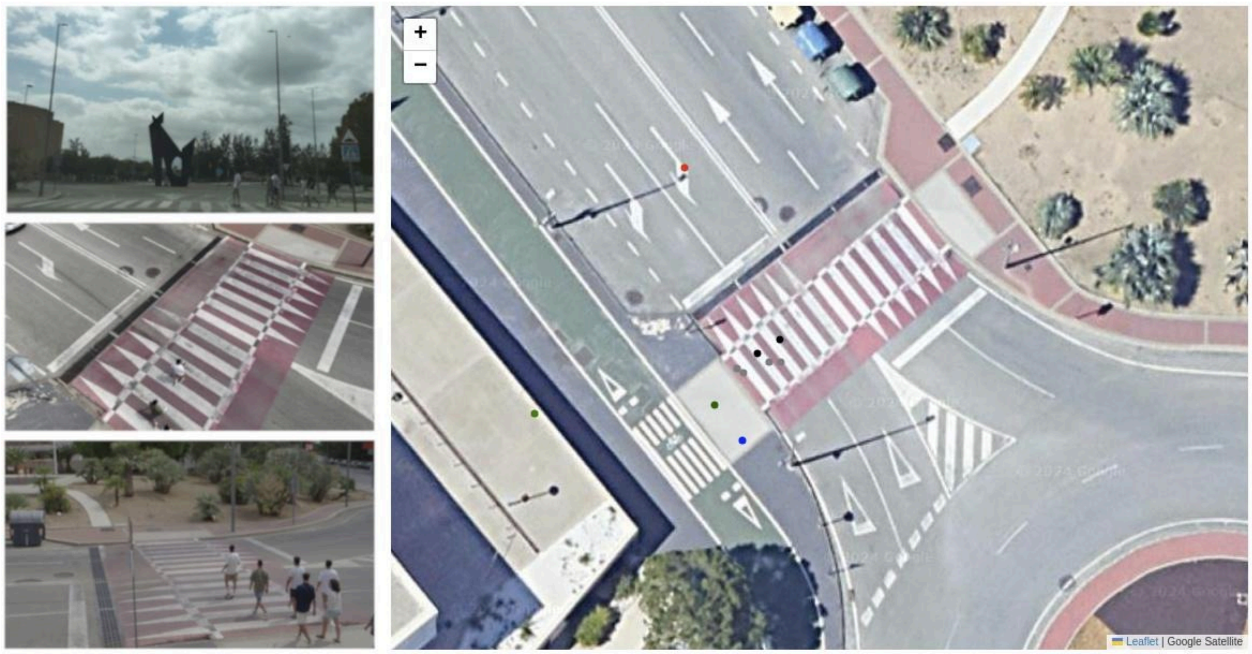

Figure 23. Schematic diagram of SemCom DT creation from the UMU dataset

Then, only the semantic information is transferred through the network to the semantic decoder to create the DT. In the current scenario, the DT consists of a top view real-world map that includes the three cameras together with the objects they see. In more detail, the objects captured by the first camera are shown in red, the objects from the drone in blue, while the objects that are identified by the RSU camera are represented in green. Finally, objects identified by two out of the three cameras are displayed in gray, while objects jointly identified by all three cameras in the DT model are displayed in black.

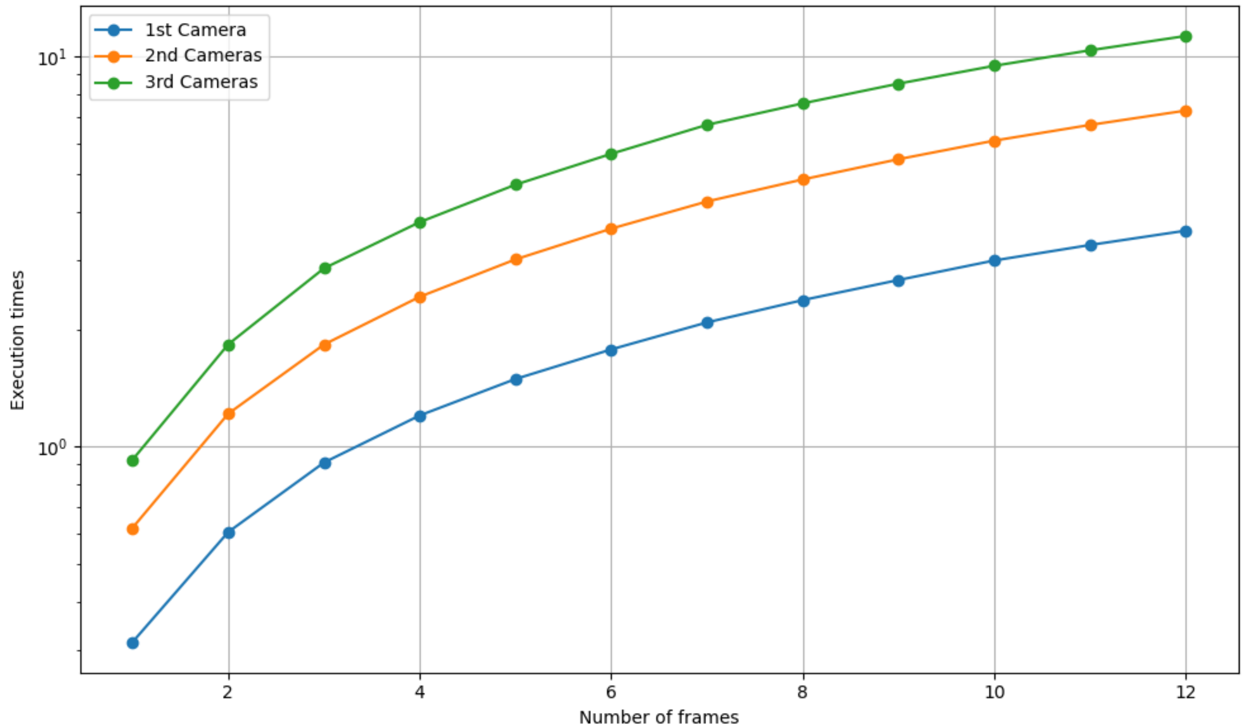

Figure 24. Scalability of SemCom DT from the UMU dataset

Figure 24 depicts the runtime for creating a DT using only semantic information, with a focus on the scalability of the model's complexity through different numbers of input nodes equipped with cameras. Specifically, it distinguishes between 3 different scenarios of creating a semantic DT using 1, 2 and 3 cameras, respectively. From all three scenarios it is obvious that as the number of frames increases, the execution time of the DT increases as well. The scenario with one camera has the shortest execution time while three cameras have the longest. Finally, the more cameras we enlist the more detailed and stable DT output we have but on the counterpart the DT generation time increases. On the contrary, using fewer cameras reduces processing time but may limit the depth and accuracy of the resulting model.

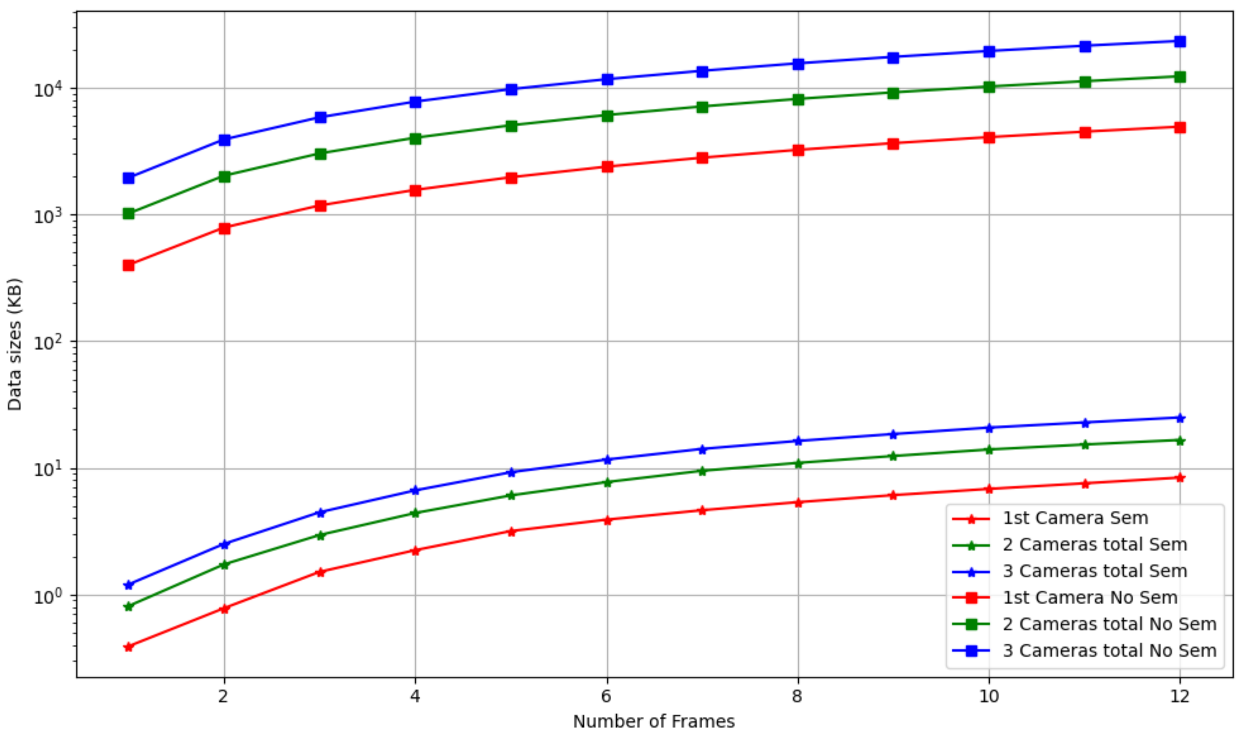

Figure 25. Data efficiency of SemCom DT from the UMU dataset

In Figure 25, we present the relationship between the transmitted data against the number of frames, for two methods. In the first method, as part of the creation of the DT, there is a semantic encoder that extracts the semantic information of each frame. On the contrary, in the second method, the whole image is transmitted for the composition of the DT. For each method, 3 different scenarios were used (one camera, two cameras, and three cameras, as input for the DT) to examine in more detail and in depth the two different approaches as well as the limits in scalability that they may have. In the non-SemCom case, all scenarios present a larger amount of data. It is worth noting that in the case where the DT is composed of one camera, we observe that the smallest value starts from 10^3 kB and as the number of frames increases there is a corresponding increase in the volume of sent data, a behavior that is observed in all three scenarios. In contrast to the case where only semantic information is sent to the DT, where a significant drop in data volume is extended to all three scenarios. A typical example is the more complex case with three cameras in which even the maximum value in the case of most frames does not exceed 10 kB. This efficiency is observed in all scenarios with the SemCom version outperforming the non-SemCom one. While the non-SemCom approach requires significantly more bandwidth, especially as the number of cameras and frames increases, the SemCom approach remains lightweight and scalable.

8. Conclusion and Outlook

This deliverable presents our perspective on the significance of SemCom in the 6G era. A high-level analysis is presented that covers the fundamentals of SemCom alongside enabling technologies, basic building blocks, and design considerations. The added contribution of semantics is highlighted through its ability to separately analyse the syntactic and semantic layers. Conventional systems have been designed with only the technical layer in mind and, therefore, revolve around the error-free transmission of symbols, while they disregard semantics. Contrary to that, the design principle of SemCom systems establishes semantics as the primary goal. Thus, the system transmits only the information that is relevant and useful to the accomplishment of this goal, while achieving higher efficiency and ensuring the satisfaction of constraints. Any errors in the syntactic level that are not detrimental for the semantic reconstruction of the message can be considered irrelevant for the communication. However, the efforts so far neglect the advantages of the effectiveness layer, which can orchestrate the functionalities and resources of the other two layers in the pursuit of the goal of communication. These novel interaction mechanisms will aid future communication systems to overcome limitations that reside in their purely syntactic nature.

When developing SemCom approaches for point-to-point communication and networking alike the creation of structured methods for exchanging information that carries rich semantic information is important. In this direction, important guidelines include:

- Clearly define the models that will be used to encode and decode information. This involves specifying the terms, concepts and relationships within the communication domain. Clearly defining terms and concepts will help minimize ambiguity and reduce the chances of misinterpretation.
- Establish proper semantic knowledge representation to ensure consistency across entities within the communication network. This entails defining a shared set of rules, formats, and structures for encoding information.
- Integrate mechanisms for filtering to eliminate or minimize semantic noise. This involves identifying and excluding information thereby improving the signal-to-noise ratio.
- Incorporate redundancy in communication to enhance error detection and correction capabilities. Expressing concepts in ways that helps the system tolerate and identify instances of semantic noise more effectively.
- Utilize existing standards and ontologies as a foundation to build upon established conventions. This can enhance compatibility and interoperability with systems adhering to standards.
- Consider the range of devices, platforms, and systems that may be involved in the network. Develop semantic coding methods that promote interoperability between entities with varying models enabling seamless communication.

- To address security concerns related to coding it is crucial to protect against semantic attacks or unauthorized manipulation of semantic information. This can be achieved by implementing encryption and authentication measures that safeguard the integrity and confidentiality of the data.
- Design dynamic processes for coding that can adapt to changes in the communication environment over time. This includes accommodating evolving semantics incorporating modified concepts and adjusting to variations in the network.
- Implement context processing by considering the surrounding context of communication. Understanding the context aids in filtering out interpretations. Enhances the system's ability to grasp the intended importance.

References

[1] K. David and H. Berndt, “6G vision and requirements: Is there any need for beyond 5G?” IEEE Veh. Technol. Mag., vol. 13, no. 3, pp. 72–80, Sep. 2018.

[2] A. Yastrebova, R. Kirichek, Y. Koucheryavy, A. Borodin, and A. Koucheryavy, “Future networks 2030: Architecture & requirements,” in Proc. 10th Int. Congr. Ultra Modern Telecommun. Control Syst. Workshops (ICUMT). 2018, pp. 1–8.

[3] V. Raghavan and J. Li, “Evolution of physical-layer communications research in the post-5G era,” IEEE Access, vol. 7, pp. 10392–10401, 2019.

[4] E. C. Strinati et al., “6G: The next frontier: From holographic messaging to artificial intelligence using subterahertz and visible light communication,” IEEE Veh. Technol. Mag., vol. 14, no. 3, pp. 42–50, Sep. 2019..

[5] F. Tariq, M. R. A. Khandaker, K.-K. Wong, M. A. Imran, M. Bennis, and M. Debbah, “A speculative study on 6G,” IEEE Wireless Commun., vol. 27, no. 4, pp. 118–125, Aug. 2020..

[6] W. Saad, M. Bennis, and M. Chen, “A vision of 6G wireless systems: Applications, trends, technologies, and open research problems,” IEEE Netw., vol. 34, no. 3, pp. 134–142, May 2020.

[7] S. E. Trevlakis et al., “Localization as a key enabler of 6G wireless systems: A comprehensive survey and an outlook,” IEEE Open J. Intell. Comun. Soc., vol. 4, pp. 2733–2801, Jan. 2023.

[8] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjoland, and F. Tufvesson, “6G wireless systems: Vision, requirements, challenges, insights, and opportunities,” Proc. IEEE, vol. 109, no. 7, pp. 1166–1199, Jul. 2021..

[9] C. E. Shannon, “A mathematical theory of communication,” Bell Syst. Tech. J., vol. 27, no. 3, pp. 379–423, Jul. 1948.

[10] Q. Lan et al., “What is semantic communication? A view on conveying meaning in the era of machine intelligence,” J. Commun. Inf. Netw., vol. 6, no. 4, pp. 336–371, Dec. 2021.

[11] J. Dai, P. Zhang, K. Niu, S. Wang, Z. Si, and X. Qin, “Communication beyond transmitting bits: Semantics-guided source and channel coding,” IEEE Wireless Commun., vol. 30, no. 4, pp. 170–177, Aug. 2023.

[12] W. Weaver, "Recent contributions to the mathematical theory of communication," ETC: A Rev. Gener. Semant., vol. 10, no. 4, pp. 261–281, 1953.

[13] R. Carnap and Y. Bar-Hillel, "An outline of a theory of semantic information," Res. Lab. Electron., Massachusetts Inst. Technol., Cambridge, MA, USA, Rep. TR. 247, Oct. 1952.

[14] J. Bao et al., "Towards a theory of semantic communication," in Proc. IEEE Netw. Sci. Workshop, 2011, pp. 110–117.

[15] E. Bourtsoulatze, D. B. Kurka, and D. Gunduz, "Deep joint source-channel coding for wireless image transmission," in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), 2019, pp. 4774–4778..

[16] C.-H. Lee, J.-W. Lin, P.-H. Chen, and Y.-C. Chang, "Deep learningconstructed joint transmission-recognition for Internet of Things," IEEE Access, vol. 7, pp. 76547–76561, 2019.

[17] Z. Weng, Z. Qin, and G. Y. Li, "Semantic communications for speech signals," 2020, arXiv:2012.05369v2.

[18] N. Farsad, M. Rao, and A. Goldsmith, "Deep learning for joint source-channel coding of text," in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), 2018, pp. 1–5.

[19] H. Xie, Z. Qin, G. Y. Li, and B.-H. Juang, "Deep learning based semantic communications: An initial investigation," in Proc. IEEE Glob. Commun. Conf., 2020, pp. 1–6.

[20] H. Xie and Z. Qin, "A lite distributed semantic communication system for Internet of Things," IEEE J. Sel. Areas Commun., vol. 39, no. 1, pp. 142–153, Jan. 2021.

[21] M. Kountouris and N. Pappas, "Semantics-empowered communication for networked intelligent systems," IEEE Commun. Mag., vol. 59, no. 6, pp. 96–102, Jun. 2021.

[22] E. C. Strinati and S. Barbarossa, "6G networks: Beyond Shannon towards semantic and goal-oriented communications," Comput. Netw., vol. 190, May 2021, Art. no. 107930.

[23] Z. Qin, X. Tao, J. Lu, W. Tong, and G. Y. Li, "Semantic communications: Principles and challenges," 2021, arXiv:2201.0138.

[24] G. Shi, Y. Xiao, Y. Li, and X. Xie, "From semantic communication to semantic-aware networking: Model, architecture, and open problems," IEEE Commun. Mag., vol. 59, no. 8, pp. 44–50, Aug. 2021.

[25] C. Chaccour, W. Saad, M. Debbah, Z. Han, and H. V. Poor, “Less data, more knowledge: Building next generation semantic communication networks,” 2022, arXiv:2211.14343.

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

[27] W. Yang et al., “Semantic communication meets edge intelligence,” IEEE Wireless Commun., vol. 29, no. 5, pp. 28–35, Oct. 2022.

[28] S. Iyer et al., “A survey on semantic communications for intelligent wireless networks,” Wireless Pers. Commun., vol. 129, no. 1, pp. 569–611, 2023.

[29] D. Wheeler and B. Natarajan, “Engineering semantic communication: A survey,” IEEE Access, vol. 11, pp. 13965–13995, 2023.

[30] Z. Lu et al., “Semantics-empowered communications: A tutorialcum-survey,” IEEE Commun. Surveys Tuts., vol. 26, no. 1, pp. 41–79, 1st Quart., 2024.

[31] W. Yang et al., “Semantic communications for future Internet: Fundamentals, applications, and challenges,” IEEE Commun. Surveys Tuts., vol. 25, no. 1, pp. 213–250, 1st Quart., 2023.

[32] P. P. et al., “Semantic effectiveness filtering and control for post-5G wireless connectivity,” *J. Indian Inst. Sci.,* vol. 100, p. 435–443, 2020.

[33] M. Moussaoui, E. Bertin, and N. Crespi, “5G shortcomings and beyond-5G/6G requirements,” in Proc. 1st Int. Conf. 6G Netw. (6GNet), 2022, pp. 1–8.

[34] ITU-R WP 5D: Future Technology Trends of Terrestrial International Mobile Telecommunications Systems Towards 2030 and Beyond, ITU-Rec. M.2516-0 , Int. Telecommun. Union, Geneva, Switzerland, 2022.

[35] P. Kutsevol, O. Ayan, N. Pappas, and W. Kellerer, “Experimental study of transport layer protocols for wireless networked control systems,” in Proc. 20th Annu. IEEE Int. Conf. Sens., Commun., Netw. (SECON), 2023, pp. 438–446.

[36] B. W. Carabelli, R. Blind, F. Dürr, and K. Rothermel, “Statedependent priority scheduling for networked control systems,” in Proc. Amer. Control Conf. (ACC), 2017, pp. 1003–1010.

[37] M. H. Mamduhi, J. P. Champati, J. Gross, and K. H. Johansson, "Where freshness matters in the control loop: Mixed age-ofinformation and event-based co-design for multi-loop networked control systems," J. Sensor Actuator Netw., vol. 9, no. 3, p. 43, 2020.

[38] H. Ma, L. Wang, and S. Zhou, "Scheduling to minimize control cost in multi-loop wireless networked control with imperfect sensors," in Proc. IEEE 95th Veh. Technol. Conf., 2022, pp. 1–7.

[39] T. Soleymani, J. S. Baras, and S. Hirche, "Value of information in feedback control: Quantification," IEEE Trans. Autom. Control, vol. 67, no. 7, pp. 3730–3737, Jul. 2022.

[40] A. Nikkhah, A. Ephremides, and N. Pappas, "Age of actuation in a wireless power transfer system," in Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS), 2023, pp. 1–6.

[41] O. Ayan, P. Kutsevol, H. Y. Özkan, and W. Kellerer, "Taskoriented scheduling for networked control systems: An age of information-aware implementation on software-defined radios," 2022, arXiv:2202.09189.

[42] Z. Zhao, W. Liu, D. E. Quevedo, Y. Li, and B. Vucetic, "Deep learning for wireless-networked systems: A joint estimation-control– scheduling approach," IEEE Internet Things J., vol. 11, no. 3, pp. 4535–4550, Feb. 2024.

[43] D. Baumann, J.-J. Zhu, G. Martius, and S. Trimpe, "Deep reinforcement learning for event-triggered control," in Proc. IEEE Conf. Decis. Control (CDC), 2018, pp. 943–950.

[44] V. Lima, M. Eisen, K. Gatsis, and A. Ribeiro, "Model-free design of control systems over wireless fading channels," Signal Process., vol. 197, Aug. 2022, Art. no. 108540.

[45] E. Uysal et al., "Semantic communications in networked systems: A data significance perspective," IEEE Netw., vol. 36, no. 4, pp. 233–240, Jul./Aug. 2022.

[46] M. Sharif and A. Sadeghi-Niaraki, "Ubiquitous sensor network simulation and emulation environments: A survey," J. Netw. Comput. Appl., vol. 93, pp. 150–181, Sep. 2017.

[47] Y. Wang, X. Kang, T. Li, H. Wang, C. Cheng, and Z. Lei, "SIXtrust for 6G: Toward a secure and trustworthy future network," IEEE Access, vol. 11, pp. 107657–107668, 2023.

[48] M. Chein and M.-L. Mugnier, Graph-Based Knowledge Representation: Computational Foundations of Conceptual Graphs. London, U.K.: Springer, 2008.

[49] V. R. S. Kumar et al., "Ontologies for industry 4.0," Knowl. Eng. Rev., vol. 34, p. e17, Nov. 2019.

[50] T. Bittner, M. Donnelly, and S. Winter, "Ontology and semantic interoperability," in Large-Scale 3D Data Integration. Boca Raton, FL, USA: CRC Press, 2005, pp. 139–160.

[51] H. Rahman and M. I. Hussain, "A comprehensive survey on semantic interoperability for Internet of Things: State-of-the-art and research challenges," Trans. Emerg. Telecommun. Technol., vol. 31, no. 12, 2020, Art. no. e3902.

[52] H. Stuckenschmidt, F. Van Harmelen, L. Serafini, P. Bouquet, and F. Giunchiglia, "Using C-OWL for the alignment and merging of medical ontologies," Dept. Inf. Commun. Technol., Univ. Trento, Trento, Italy, Rep. DIT-04-010, 2004.

[53] Y. He, Z. Xiang, J. Zheng, Y. Lin, J. A. Overton, and E. Ong, "The eXtensible ontology development (XOD) principles and tool implementation to support ontology interoperability," J. Biomed. Semant., vol. 9, pp. 1–10, Jan. 2018.

[54] D. Gündüz et al., "Beyond transmitting bits: Context, semantics, and task-oriented communications," IEEE J. Sel. Areas Commun., vol. 41, no. 1, pp. 5–41, Jan. 2023.

[55] C. Xiong, R. Power, and J. Callan, "Explicit semantic ranking for academic search via knowledge graph embedding," in Proc. 26th Int. Conf. world wide web, 2017, pp. 1271–1279.

[56] T. Pugazhenthi and H. Liang, "Improving conversational recommender systems via knowledge graph-based semantic fusion with historical interaction data," in Proc. IEEE Int. Conf. Big Data (Big Data), 2022, pp. 4303–4312.

[57] D. Collarana, M. Galkin, I. Traverso-Ribón, C. Lange, M.-E. Vidal, and S. Auer, "Semantic data integration for knowledge graph construction at query time," in Proc. IEEE 11th Int. Conf. Semantic Comput. (ICSC), 2017, pp. 109–116.

[58] S. Yang et al., "TransMS: Knowledge graph embedding for complex relations by multidirectional semantics," in Proc. IJCAI, 2019, pp. 1935–1942.

[59] F. Zhou, Y. Li, X. Zhang, Q. Wu, X. Lei, and R. Q. Hu, "Cognitive semantic communication systems driven by knowledge graph," in Proc. IEEE Int. Conf. Commun., 2022, pp. 4860–4865.

[60] M. Nickel, V. Tresp, and H.-P. Kriegel, "A three-way model for collective learning on multi-relational data," in Proc. 28th ICML, vol. 11, 2011, pp. 3104482–3104584.

[61] M. Jankowski, D. Gündüz, and K. Mikolajczyk, "Wireless image retrieval at the edge," IEEE J. Sel. Areas Commun., vol. 39, no. 1, pp. 89–100, Jan. 2021.

[62] Q. Hu, G. Zhang, Z. Qin, Y. Cai, G. Yu, and G. Y. Li, “Robust semantic communications with masked VQ-VAE enabled codebook,” 2022, arXiv:2206.04011.

[63] H. Xie, Z. Qin, G. Y. Li, and B.-H. Juang, “Deep learning enabled semantic communication systems,” IEEE Trans. Signal Process., vol. 69, pp. 2663–2675, Apr. 2021.

[64] J. Xu, B. Ai, W. Chen, A. Yang, P. Sun, and M. Rodrigues, “Wireless image transmission using deep source channel coding with attention modules,” IEEE Trans. Circuits Syst. Video Technol., vol. 32, no. 4, pp. 2315–2328, Apr. 2022.

[65] T.-Y. Tung and D. Gündüz, “DeepWiVe: Deep-learning-aided wireless video transmission,” IEEE J. Sel. Areas Commun., vol. 40, no. 9, pp. 2570–2583, Sep. 2022.

[66] Z. Weng and Z. Qin, “Semantic communication systems for speech transmission,” IEEE J. Sel. Areas Commun., vol. 39, no. 8, pp. 2434–2444, Aug. 2021.

[67] H. Xie, Z. Qin, X. Tao, and K. B. Letaief, “Task-oriented multi-user semantic communications,” IEEE J. Sel. Areas Commun., vol. 40, no. 9, pp. 2584–2597, Sep. 2022.

[68] S. Vandenhende, S. Georgoulis, W. Van Gansbeke, M. Proesmans, D. Dai, and L. Van Gool, “Multi-task learning for dense prediction tasks: A survey,” IEEE Trans. pattern Anal. Mach. Intell., vol. 44, no. 7, pp. 3614–3633, Jan. 2021.

[69] M. Crawshaw, “Multi-task learning with deep neural networks: A survey,” 2020, arXiv:2009.09796.

[70] G. Zhang et al., “A unified multi-task semantic communication system for multimodal data,” 2022, arXiv:2209.07689.

[71] M. Wang and W. Deng, “Deep visual domain adaptation: A survey,” Neurocomputing, vol. 312, pp. 135–153, Oct. 2018.

[72] P. Popovski et al., “A perspective on time toward wireless 6G,” Proc. IEEE, vol. 110, no. 8, pp. 1116–1146, Aug. 2022.

[73] P. Popovski, O. Simeone, F. Boccardi, D. Gündüz, and O. Sahin, “Semantic-effectiveness filtering and control for post-5G wireless connectivity,” J. Indian Inst. Sci., vol. 100, no. 2, pp. 435–443, Apr. 2020.

[74] P. Popovski, K. F. Trillingsgaard, O. Simeone, and G. Durisi, “5G wireless network slicing for eMBB, URLLC, and mMTC: A communication-theoretic view,” IEEE Access, vol. 6, pp. 55765–55779, 2018.

[75] S. Lashgari and A. S. Avestimehr, “Timely throughput of heterogeneous wireless networks: Fundamental limits and algorithms,” IEEE Trans. Inf. Theory, vol. 59, no. 12, pp. 8414–8433, Dec. 2013.

[76] B. Soret, P. Mogensen, K. I. Pedersen, and M. C. Aguayo-Torres, “Fundamental tradeoffs among reliability, latency and throughput in cellular networks,” in Proc. IEEE Globecom Workshops (GC Wkshps), 2014, pp. 1391–1396.

[77] A. Destounis, G. S. Paschos, J. Arnau, and M. Kountouris, “Scheduling URLLC users with reliable latency guarantees,” in Proc. 16th Int. Symp. Model. Optim. Mobile, Ad Hoc, Wireless Netw. (WiOpt), 2018, pp. 1–8.

[78] B. Soret, A. De Domenico, S. Bazzi, N. H. Mahmood, and K. I. Pedersen, “Interference coordination for 5G new radio,” IEEE Wireless Commun., vol. 25, no. 3, pp. 131–137, Jun. 2018.

[79] G. Pocovi, B. Soret, K. I. Pedersen, and P. Mogensen, “MAC layer enhancements for ultra-reliable low-latency communications in cellular networks,” in Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops), 2017, pp. 1005–1010.

[80] A. Anand and G. de Veciana, “Resource allocation and HARQ optimization for URLLC traffic in 5G wireless networks,” IEEE J. Sel. Areas Commun., vol. 36, no. 11, pp. 2411–2421, Nov. 2018.

[81] E. Fountoulakis, N. Pappas, Q. Liao, A. Ephremides, and V. Angelakis, “Dynamic power control for packets with deadlines,” in Proc. IEEE Glob. Commun. Conf. (GLOBECOM), 2018, pp. 1–6.

[82] C. Tsanikidis and J. Ghaderi, “On the power of randomization for scheduling real-time traffic in wireless networks,” IEEE/ACM Trans. Netw., vol. 29, no. 4, pp. 1703–1716, Apr. 2021.

[83] A. E. Ewaisha and C. Tepedelenlioˇglu, “Optimal power control and scheduling for real-time and non-real-time data,” IEEE Trans. Veh. Technol., vol. 67, no. 3, pp. 2727–2740, Mar. 2018.

[84] F. Chiariotti, A. Zanella, S. Kucera, K. Fahmi, and H. Claussen, “The HOP protocol: Reliable latency-bounded end-to-end multipath communication,” IEEE/ACM Trans. Netw., vol. 29, no. 5, pp. 2281–2295, Oct. 2021.

[85] F. Chiariotti, S. Kucera, A. Zanella, and H. Claussen, “Analysis and design of a latency control protocol for multi-path data delivery with pre-defined QoS guarantees,” IEEE/ACM Trans. Netw., vol. 27, no. 3, pp. 1165–1178, Jun. 2019.

[86] Y. E. Guo, A. Nikravesh, Z. M. Mao, F. Qian, and S. Sen, “DEMS: Decoupled multipath scheduler for accelerating multipath transport,” in Proc. 23rd Annu. Int. Conf. Mobile Comput. Netw., 2017, pp. 477–479.

[87] A. Bentaleb, B. Taani, A. C. Begen, C. Timmerer, and R. Zimmermann, “A survey on bitrate adaptation schemes for streaming media over HTTP,” IEEE Commun. Surveys Tuts., vol. 21, no. 1, pp. 562–585, 1st Quart., 2018.

[88] M. Uitto and A. Heikkinen, “Evaluation of live video streaming performance for low latency use cases in 5G,” in Proc. Joint Eur. Conf. Netw. Commun. 6G Summit (EuCNC/6G Summit), 2021, pp. 431–436.

[89] S. K. Sharma, I. Woungang, A. Anpalagan, and S. Chatzinotas, “Toward tactile Internet in beyond 5G era: Recent advances, current issues, and future directions,” IEEE Access, vol. 8, pp. 56948–56991, 2020.

[90] N. Pappas, M. Kountouris, A. Ephremides, and A. Traganitis, “Relayassisted multiple access with full-duplex multi-packet reception,” IEEE Trans. Wireless Commun., vol. 14, no. 7, pp. 3544–3558, Jul. 2015.

[91] Z. Liao et al., “Distributed probabilistic offloading in edge computing for 6G-enabled massive Internet of Things,” IEEE Internet Things J., vol. 8, no. 7, pp. 5298–5308, Apr. 2021.

[92] K. Kaur, F. Guillemin, V. Q. Rodriguez, and F. Sailhan, “Latency and network aware placement for cloud-native 5G/6G services,” in Proc. IEEE 19th Annu. Consum. Commun. Netw. Conf. (CCNC), 2022, pp. 114–119.

[93] Y. Lu, X. Huang, K. Zhang, S. Maharjan, and Y. Zhang, “Lowlatency federated learning and blockchain for edge association in digital twin empowered 6G networks,” IEEE Trans. Ind. Informat., vol. 17, no. 7, pp. 5098–5107, Jul. 2021.

[94] M. Polese, F. Chiariotti, E. Bonetto, F. Rigotto, A. Zanella, and M. Zorzi, “A survey on recent advances in transport layer protocols,” IEEE Commun. Surveys Tuts., vol. 21, no. 4, pp. 3584–3608, 4th Quart., 2019.

[95] Y. Cai, J. Llorca, A. M. Tulino, and A. F. Molisch, “Ultrareliable distributed cloud network control with end-to-end latency constraints,” IEEE/ACM Trans. Netw., vol. 30, no. 6, pp. 2505–2520, Dec. 2022.

[96] E. Fountoulakis, T. Charalambous, A. Ephremides, and N. Pappas, “Scheduling policies for AoI minimization with timely throughput constraints,” IEEE Trans. Commun., vol. 71, no. 7, pp. 3905–3917, Jul. 2023.

[97] Q. Wang, C. He, K. Jaffrès-Runser, J. Huang, and Y. Xu, “Timely-throughput optimal scheduling for wireless flows with deep reinforcement learning,” in Proc. IEEE/ACM 30th Int. Symp. Qual. Service (IWQoS), 2022, pp. 1–11.

[98] R. Singh and P. Kumar, “Throughput optimal decentralized scheduling of multihop networks with end-to-end deadline constraints: Unreliable links,” IEEE Trans. Autom. Control, vol. 64, no. 1, pp. 127–142, Jan. 2018.

[99] R. Singh and P. Kumar, “Adaptive CSMA for decentralized scheduling of multi-hop networks with end-to-end deadline constraints,” IEEE/ACM Trans. Netw., vol. 29, no. 3, pp. 1224–1237, Jul. 2021.

[100] A. Kosta, N. Pappas, and V. Angelakis, “Age of information: A new concept, metric, and tool,” Found. Trends Netw., vol. 12, no. 3, pp. 162–259, 2017.

[101] R. D. Yates and S. K. Kaul, “Age of information in uncoordinated unslotted updating,” in Proc. IEEE Int. Symp. Inf. Theory (ISIT), 2020, pp. 1759–1764.

[102] X. Chen, K. Gatsis, H. Hassani, and S. S. Bidokhti, “Age of information in random access channels,” IEEE Trans. Inf. Theory, vol. 68, no. 10, pp. 6548–6568, 2022.

[103] A. Maatouk, M. Assaad, and A. Ephremides, “On the age of information in a CSMA environment,” IEEE/ACM Trans. Netw., vol. 28, no. 2, pp. 818–831, Apr. 2020.

[104] E. Fountoulakis, T. Charalambous, N. Nomikos, A. Ephremides, and N. Pappas, “Information freshness and packet drop rate interplay in a two-user multi-access channel,” J. Commun. Netw., vol. 24, no. 3, pp. 357–364, 2022.

[105] F. Chiariotti et al., “Query age of information: Freshness in pullbased communication,” IEEE Trans. Commun., vol. 70, no. 3, pp. 1606–1622, Mar. 2022.

[106] J. Zhong, R. D. Yates, and E. Soljanin, “Two freshness metrics for local cache refresh,” in Proc. IEEE Int. Symp. Inf. Theory (ISIT), 2018, pp. 1924–1928.

[107] P. M. De Sant Ana, N. Marchenko, P. Popovski, and B. Soret, “Age of loop for wireless networked control systems optimization,” in Proc. IEEE 32nd Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC), 2021, pp. 1–7.

[108] A. Kosta, N. Pappas, A. Ephremides, and V. Angelakis, “Age and value of information: Non-linear age case,” in Proc. IEEE Int. Symp. Inf. Theory (ISIT), 2017, pp. 326–330.

[109] A. Kosta, N. Pappas, A. Ephremides, and V. Angelakis, “The cost of delay in status updates and their value: Non-linear ageing,” IEEE Trans. Commun., vol. 68, no. 8, pp. 4905–4918, 2020.

[110] X. Zheng, S. Zhou, Z. Jiang, and Z. Niu, “Closed-form analysis of non-linear age of information in status updates with an energy harvesting transmitter,” IEEE Trans. Wireless Commun., vol. 18, no. 8, pp. 4129–4142, Aug. 2019.

[111] M. Xie, J. Gong, and X. Ma, “Age-energy tradeoff of short packet based transmissions in multicast networks with ARQ,” in Proc. IEEE 91st Veh. Technol. Conf. (VTC), 2020, pp. 1–5.

[112] J. Huang, J. Gong, and W. Zhan, “Energy-age tradeoff in status update communication systems based on HARQ,” in Proc. 13th Int. Conf. Wireless Commun. Signal Process. (WCSP), 2021, pp. 1–5.

[113] N. I. Miridakis, T. A. Tsiftsis, and G. Yang, “Non-linear age of information: An energy efficient receiver-centric approach,” IEEE Wireless Commun. Lett., vol. 11, no. 3, pp. 655–659, Mar. 2022.

[114] A. Maatouk, S. Kriouile, M. Assaad, and A. Ephremides, “The age of incorrect information: A new performance metric for status updates,” IEEE/ACM Trans. Netw., vol. 28, no. 5, pp. 2215–2228, Oct. 2020.

[115] O. Ayan, M. Vilgelm, M. Klügel, S. Hirche, and W. Kellerer, “Age-of-information vs. value-of-information scheduling for cellular networked control systems,” in Proc. 10th ACM/IEEE Int. Conf. Cyber-Phys. Syst., 2019, pp. 109–117.

[116] X. Zheng, S. Zhou, and Z. Niu, “Urgency of information for contextaware timely status updates in remote control systems,” IEEE Trans. Wireless Commun., vol. 19, no. 11, pp. 7237–7250, Nov. 2020.

[117] L. Floridi, “Outline of a theory of strongly semantic information,” Minds Mach., vol. 14, pp. 197–221, May 2004.

[118] L. Floridi, “Philosophical conceptions of information,” in Formal Theories of Information. Berlin, Germany: Springer, 2009.

[119] S. D’Alfonso, “On quantifying semantic information,” Information, vol. 2, no. 1, pp. 61–101, 2011.

[120] B. Juba and M. Sudan, “Universal semantic communication I,” in Proc. Fortieth Annu. ACM Symp. Theory Comput., 2008, pp. 123–132.

[121] B. Juba and M. Sudan, “Universal semantic communication II: A theory of goal-oriented communication,” in Electron. Colloq. Comput. Complex. (ECCC), vol. 15, 2008, pp. 1–29.

[122] G. Sommaruga, Formal Theories of Information. Berlin, Germany: Springer, 2009.

[123] A. Chattopadhyay, B. D. Haeffele, D. Geman, and R. Vidal, “Quantifying task complexity through generalized information measures,” in Proc. ICLR, 2021, pp. 1–21.

[124] S. E. Trevlakis, N. Pappas and A. -A. A. Boulogeorgos, "Toward Natively Intelligent Semantic Communications and Networking," in IEEE Open Journal of the Communications Society, vol. 5, pp. 1486-1503, 2024.

[125] P. Agheli, N. Pappas, and M. Kountouris, “Semantics-aware source coding in status update systems,” in IEEE International Conference on Communications Workshops, 2022.

[126] https://www.google.comP. Agheli, N. Pappas, and M. Kountouris, “Semantic source coding for two users with heterogeneous goals,” in Proc. IEEE Glob. Commun. Conf. (GLOBECOM), 2022, pp. 4983–4988.

[127] P. Basu, J. Bao, M. Dean, and J. Hendler, “Preserving quality of information by using semantic relationships,” Pervasive Mobile Comput., vol. 11, pp. 188–202, Apr. 2014.

[128] V. Kouvakis, S. Trevlakis, and A.-A. A. Boulogeorgos, “Semantic communications for image-based sign language transmission,” TechRxiv, Sep. 2023.

[129] H. Xie, Z. Qin, G. Y. Li, and B.-H. Juang, “Deep learning enabled semantic communication systems,” IEEE Transactions on Signal Processing, vol. 69, pp. 2663–2675, 2021.

[130] N. Farsad, M. Rao, and A. Goldsmith, “Deep Learning for Joint Source-Channel Coding of Text,” in 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, Apr 2018.

[131] H. Xie and Z. Qin, “A Lite Distributed Semantic Communication System for Internet of Things,” IEEE Journal on Selected Areas in Communications, vol. 39, no. 1, pp. 142–153, jan 2021.

[132] Z. Weng and Z. Qin, “Semantic communication systems for speech transmission,” IEEE Journal on Selected Areas in Communications, vol. 39, no. 8, pp. 2434–2444, 2021.

[133] J. Wang, Y. Duan, X. Tao, M. Xu, and J. Lu, “Semantic perceptual image compression with a laplacian pyramid of convolutional networks,” IEEE Transactions on Image Processing, vol. 30, pp. 4225–4237, 2021.

[134] D. Huang, X. Tao, F. Gao, and J. Lu, “Deep learning-based image semantic coding for semantic communications,” in 2021 IEEE Global Communications Conference (GLOBECOM), 2021, pp. 1–6.

[135] X. Li, J. Shi, and Z. Chen, “Task-driven semantic coding via reinforcement learning,” IEEE Transactions on Image Processing, vol. 30, pp. 6307–6320, 2021.

[136] D. B. Kurka and D. Gu¨ndu¨z, “Deepjscc-f: Deep joint source-channel coding of images with feedback,” 2020.

[137] T.-Y. Tung and D. Gu¨ndu¨z, “Deepwive: Deep-learning-aided wireless video transmission,” 2021.

[138] S. Iyer, R. Khanai, D. Torse, R. J. Pandya, K. M. Rabie, K. Pai, W. U. Khan, and Z. Fadlullah, “A survey on semantic communications for intelligent wireless networks,” Wireless Personal Communications, vol. 129, no. 1, pp. 569–611, 2023.

[139] D. M. Madhiarasan, P. Roy, and P. Pratim, “A comprehensive review of sign language recognition: Different types, modalities, and datasets,” arXiv preprint arXiv:2204.03328, 2022.

[140] R. Fatmi, S. Rashad, R. Integlia, and G. Hutchison, “American sign language recognition using hidden markov models and wearable motion sensors,” Trans. Mach. Learn. Data Min., vol. 10, pp. 41–55, 2017.

[141] J. Forster, C. Oberdo¨rfer, O. Koller, and H. Ney, “Modality combination techniques for continuous sign language recognition,” 06 2013.

[142] M. Al-Hammadi, G. Muhammad, W. Abdul, M. Alsulaiman, M. A. Bencherif, and M. A. Mekhtiche, “Hand gesture recognition for sign language using 3dcnn,” IEEE Access, vol. 8, pp. 79 491–79 509, 2020.

[143] R. Rastgoo, K. Kiani, and S. Escalera, “Hand sign language recognition using multi-view hand skeleton,” Expert Syst. Appl., vol. 150, p. 113336, 2020.

[144] O. M. Sincan and H. Y. Keles, “Autsl: A large scale multi-modal turkish sign language dataset and baseline methods,” IEEE Access, vol. 8, pp. 181 340–181 355, 2020.

[145] G. Yuan, X. Liu, Q. Yan, S. Qiao, Z. Wang, and L. Yuan, “Hand gesture recognition using deep feature fusion network based on wearable sensors,” IEEE Sensors Journal, vol. PP, pp. 1–1, 08 2020.

[146] S. K. H. Aly and W. Aly, “Deeparslr: A novel signer-independent deep learning framework for isolated arabic sign language gestures recognition,” IEEE Access, vol. 8, pp. 83 199–83 212, 2020.

[147] C. Lee, K. K. Ng, C.-H. Chen, H. Lau, S. Chung, and T. Tsoi, “American sign language recognition and training method with recurrent neural network,” Expert Systems with Applications, vol. 167, p. 114403, 04 2021.

[148] V. Kouvakis, S. E. Trevlakis and A. -A. A. Boulogeorgos, "Semantic Communications for Image-Based Sign Language Transmission," in IEEE Open Journal of the Communications Society, vol. 5, pp. 1088-1100, 2024.

[149] M. Rinalduzzi, A. De Angelis, F. Santoni, E. Buchicchio, A. Moschitta, P. Carbone, P. Bellitti, and M. Serpelloni, “Gesture recognition of sign language alphabet using a magnetic positioning system,” Applied Sciences, vol. 11, no. 12, Jun 2021.

[150] Kouvakis, V., Mitsiou, L., Trevlakis, S. E., Boulogeorgos, A.-A. A., & Tsiftsis, T. (2025). American Sign Language dataset for semantic communications [Data set]. IEEEDataPort & Zenodo. https://doi.org/10.21227/2c1z-8j21.

[151] OpenCV, “Opencv,” last accessed 28 Aug 2023. [Online]. Available: https://opencv.org/.

[152] MediaPipe, “Mediapipe,” last accessed 28 Aug 2023. [Online]. Available: https://developers.google.com/mediapipe/framework.

[153] T. Wiatowski and H. Bo¨lcskei, “A mathematical theory of deep convolutional neural networks for feature extraction,” IEEE Transactions on Information Theory, vol. 64, no. 3, pp. 1845–1866, Nov 2017.

[154] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” arXiv preprint arXiv:1412.6980, Dec 2014.

[155] MNIST, “Sign language mnist,” 2018. [Online]. Available: https://www.kaggle.com/datasets/datamunge/sign-language-mnist.

[156] LEXSET, “Synthetic asl alphabet,” 2022. [Online]. Available: https://www.kaggle.com/datasets/lexset/synthetic-asl-alphabet.

[157] A. Demba and D. P. F. Möller, "Vehicle-to-Vehicle Communication Technology," 2018 IEEE International Conference on Electro/Information Technology (EIT), Rochester, MI, USA, 2018, pp. 0459-0464.

[158] J. Jansons, E. Petersons and N. Bogdanovs, "Vehicle-to-infrastructure communication based on 802.11n wireless local area network technology," 2012 2nd Baltic Congress on Future Internet Communications, Vilnius, Lithuania, 2012, pp. 26-31.

[159] J. J. Anaya, P. Merdrignac, O. Shagdar, F. Nashashibi and J. E. Naranjo, "Vehicle to pedestrian communications for protection of vulnerable road users," 2014 IEEE Intelligent Vehicles Symposium Proceedings, Dearborn, MI, USA, 2014, pp. 1037-1042.

[160] J. W. Wedel, B. Schünemann and I. Radusch, "V2X-Based Traffic Congestion Recognition and Avoidance," 2009 10th International Symposium on Pervasive Systems, Algorithms, and Networks, Kaoshiung, Taiwan, 2009, pp. 637-641.

[161] Xu, W.; Tian, J. Energy Efficiency Optimization for a V2X Network with NOMA-CoMP Enabled. Electronics 2023, 12, 4590.

[162] Zhang, Zheng, et al. "DRL-Based Optimization for AoI and Energy Consumption in C-V2X Enabled IoV." arXiv preprint arXiv:2411.13104 (2024).

[163] M. Helmy, Y. Essam, S. Khaled, O. Khaled and O. Moataz, "Design and Implementation of Digital Twin V2X System of Broken Vehicle," 2023 International Conference on Computer and Applications (ICCA), Cairo, Egypt, 2023, pp. 1-5.

[164] Haitham Y. Adarbah, Mehdi Sookhak, and Mohammed Atiquzzaman. 2023. A Digital Twin Environment for 5G Vehicle-to-Everything: Architecture and Open Issues.

[165] J. M. Gimenez-Guzman, I. Leyva-Mayorga and P. Popovski, "Semantic V2X Communications for Image Transmission in 6G Systems," in IEEE Network, vol. 38, no. 6, pp. 48-54, Nov. 2024.

[166] "EPFL" data set: Multi-camera Pedestrian Videos, https://www.epfl.ch/labs/cvlab/data/data-pom-index-php/.

[167] Semantic communications for Multi-View Digital Twin creation, https://www.youtube.com/watch?v=wXzT3vcAh2E&pp=ygUNaW5ub2N1YmUgcC5jLg%3D%3D.

[168] Ramon Sanchez-Iborra, Rodrigo Asensio-Garriga, Gonzalo Alarcon-Hellin, Luis Bernal-Escobedo, Stylianos Trevlakis, Theodoros Tsiftsis, Antonio Skarmeta, January 20, 2025, "Multi-perspective Traffic Video Recording", IEEE Dataport, doi: https://dx.doi.org/1.