This article has been accepted for publication in a future proceedings of this conference, but has not been fully edited. Content may change prior to final publication. Citation information: DOI: 10.1109/bdcat63179.2024.00066, 2024 IEEE/ACM International Conference on Big Data Computing, Applications and Technologies (BDCAT) - Workshop 2 - EIBBDESFT

Semantic Communications for Healthcare Applications: Opportunities and Challenges

Areej Athama[1], Kezhi Wang[1], Xiaomin Chen[2], and Yongmin Li[1]

[1]*Department of Computer Science*, Brunel University of London, United Kingdom

[2]*Department of Computer Science*, University of Reading, United Kingdom

Email: Areej.Athama@brunel.ac.uk, kezhi.wang@brunel.ac.uk, xiaomin.chen@reading.ac.uk, yongmin.li@brunel.ac.uk

***Abstract*—In this paper, we introduce the healthcare system where Semantic Communication (SC) technology is applied to improve the quality of service for healthcare and medical applications. We first show the concepts and possible architecture of SC. Then, we show different types of SC in the healthcare system. Next, some examples of SC-enhanced healthcare applications are discussed. Finally, we give research challenges and future research directions.**

***Index Terms*—Semantic communication, healthcare applications, medical applications, wireless communications.**

I. Introduction

The rapid development of wireless communication technologies has significantly enhanced the healthcare and medical sectors. For example, cellular networks (e.g., 4G/5G) have empowered many innovations, and there has been huge progress in how healthcare services are provided and in the administration of healthcare establishments. These communication technologies offer high connectivity for various medical devices through base stations or the Internet for applications involving telemedicine with the help of high-definition video conferencing and real-time patient monitoring systems.

One example is the development and application of Ultra-Reliable and Low-Latency Communications (URLLC) in healthcare sectors [1]. It enables the communication of critical short-packet information, e.g., control signals, to be delivered with high reliability. Also, it might help some healthcare-related tasks to be offloaded for processing in remote edge servers with low latency. Another example is tactile Internet [2], which can help build real-time interactive systems. This technology is useful for remote surgery in medical sectors, enabling doctors to operate robotic instruments performing physical examinations or operations with haptic feedback, thus simulating a tactile experience. Additionally, the Internet of Medical Things (IoMT) [3] has also been adopted recently in the healthcare sectors, which can help build an interconnected network of medical devices and sensors to help better monitor and manage patients.

As a result of the above-mentioned wireless communication technologies, many medical applications have been discussed. One example is remote surgery, where the URLLC and tactile Internet may be employed to help real-time, robot-assisted remote operation by allowing surgeons to perform complex procedures remotely with minimal latency. Another example is monitoring systems, where the IoMT may be applied to facilitate real-time patient monitoring through wearable devices, providing continuous data transmission of vital signs e.g., heart rate and oxygen levels for remote healthcare management. Also, considering real-time telemedicine, cellular networks may be used for remote consultations, diagnostics, and videoconferencing, enhancing patient access to healthcare professionals and diagnosis in remote areas.

Furthermore, we have several enabling technologies with the help of wireless communications in healthcare services. For example, Digital Twin has been popular in assisting in the creation of a digital replica of a physical entity, e.g., the patient, that can be used for real-time monitoring and simulation. This digital copy helps doctors make complex medical decisions with real-time access to information and data concerning the patients. Similarly, with Augmented Reality (AR) and Virtual Reality (VR) technologies, doctors can simulate, visualise, or even explore the various treatment possibilities in real-time. For example, doctors can use the Apple Vision Pro to view 2D/3D images or videos to understand better accurate real-time models, e.g., the heart Image 3D model and its shape when it is beating, thereby enhancing patient treatment.

However, integrating these communication technologies and facilitating systems into the healthcare setting are coupled with many challenges. Bandwidth constraints are among the first challenges. With the available limited wireless bandwidth, only a few devices can successfully communicate simultaneously without experiencing interference or noise. The IoMT and similar technologies might experience some problems when many sensors try to communicate at the same time. Additionally, each medical device may only transmit a limited amount of data due to the small available wireless bandwidth mentioned above. This may pose challenges to 3D video/image transmission as they generally include a large amount of data volumes. This may impact the Quality of Service (QoS) of AR, VR or digital twin applications in medical fields like remote surgery or real-time diagnostics, where data accuracy is important for precise representation.

To address the above issues, in this paper, we propose to apply SC techniques to healthcare applications to help the above-mentioned applications.

II. Brief introduction to semantic communication

Wireless technologies have evolved significantly during the past few decades, with nearly every ten years, a new generation

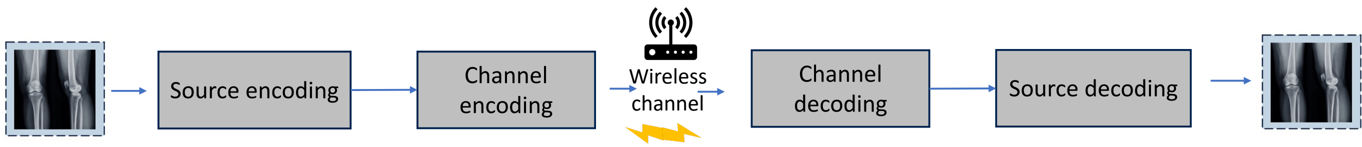

Fig. 1. Typical communication systems.

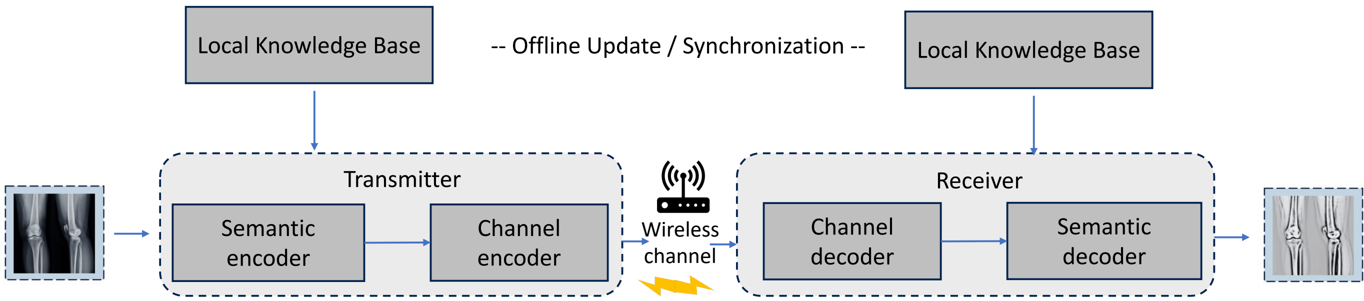

Fig. 2. Semantic communication systems.

of communication technology from 1G to now 5G. For typical communications, it usually has source encoding and decoding as well as channel encoding and decoding components, as shown in Fig. 1. In this system, we usually need to convert the data, e.g., the image, into bits, e.g., $'0101010111...'$ and then transmit them in the wireless physical channel. The communication target is to ensure the accuracy of the transmitted bits or symbols by overcoming the noise or interference that may lead to errors or loss of information. For example, in Fig. 1, we expect to receive the same X-ray image as transmitted. The other goal is to maximise the number of transmitted data bits in the wireless channel. In other words, we expect to transmit as much information as possible in the wireless channel with limited bandwidth.

On the other hand, Semantic Communication (SC), which was first proposed by Shannon and Weaver, described in their three levels of communication theory [4], focuses on conveying the meaning behind the data that ensures the sender and receiver understand each other and interpret the message in the same way. In other words, semantic communication aims to deliver the desired meaning with minimal data by only transmitting the key information or bits that express the contents or the parts they find most important or relevant while omitting redundant information which is irrelevant to both the sender and receiver [5], [6]. In this case, the exact bits are not required to be transmitted. There are two main advantages of using SC: 1) it saves bandwidth, and 2) it increases the transmission efficiency and saves energy consumption, as SC generally transmits fewer data bits instead of the whole information [7].

SC normally has five parts, as shown in Fig. 2: semantic encoder, channel encoder, channel decoder, semantic decoder and knowledge base [5], [6], [8]. The function of them is as follows:

- Semantic encoder is applied to extract key information from the source and encodes these features from the original data, with the target of increasing the possibility of effective and reliable transmission of key features over the wireless channel.
- Channel encoder is used to ensure that the selected semantic feature is suitable for transmission on the wireless physical channel. The aim is to mitigate the effects of noise and interference on the physical channel during the wireless transmission process.
- Channel decoder, as the reverse pair of channel encoder, is for demodulating and decoding the received signal and obtaining the transmitted semantic features to the maximum extent.
- Semantic decoder, as the reverse pair of the semantic encoder, is applied to recover the original message from the transmitted semantic features.
- Knowledge Base (KB) is considered the most important part of SC. The KB can help the semantic encoder and decoder understand and infer semantic information in a similar way. In healthcare and medical applications, we can use the specific KB to include the related medical models, recodes, knowledge and information. It can help improve the QoS of medical applications, support clinical decisions, and ensure that treatment plans are consistent with the clinical guidelines.

It can be seen from Fig. 2 that the received X-ray image may not be exactly the same as the transmitted image, however, their key information or concept should be the same, in order to meet the objective of the semantic communication.

III. Types of SC in Healthcare System

Recently, SC has drawn much attention from both academia and industry communities with the impetus of the development of Artificial Intelligence (AI) and Machine Learning (ML) techniques. For instance, researchers in [9] have

This article has been accepted for publication in a future proceedings of this conference, but has not been fully edited. Content may change prior to final publication. Citation information: DOI: 10.1109/bdcat63179.2024.00066, 2024 IEEE/ACM International Conference on Big Data Computing, Applications and Technologies (BDCAT) - Workshop 2 - EIBBDESFT

proposed a Joint Source Channel Coding (JSCC) scheme, where deep learning techniques are applied to the semantic encoder/decoder and the channel encoder/decoder to maximize the transmission efficiency through various network scenarios.

Generally, SC may be categorized into several types based on the nature of the data being sent and received [10], [11], e.g., text, audio, video, or any combination of these modalities, as illustrated in Fig.3.

A. Text

Traditional mobile health applications typically transmit detailed text-based medical information. With the development of Natural Language Processing (NLP) and SC, only essential semantic information needs to be transmitted, thus simplifying real-time data exchange and saving bandwidth usage, which further increases the operational efficiency of mobile medical devices.

SC normally, in the mobile healthcare scope, semantic textual information refers to text understanding and logical relations between different texts. This type of information exists in structured and unstructured forms: medical records, patient documents, or ontologies that define concepts and their interrelations. KB in SC can be used here to help patients understand medical terminology, identify the relations between medical entities and diagnoses, and get personalized treatment recommendations.

Additionally, KBs can store Electronic Health Records (EHRs), clinical notes, and relevant literature, and by doing so, retrieve and infer meaningful information from previously unstructured texts. Moreover, SC enables doctors and nurses to cross-check prescribed drugs against the patient's history, allergies, and other critical data which helps to decrease the chances of error and enhance patient safety. Given a well-designed KB, SC can be integrated with source and channel encoding/decoding strategies to adapt to changing network conditions in healthcare applications.

B. Image/Video

In modern medicine, images and videos are staples in the field of diagnosis, treatment, or research. They constitute a visual record that could be analyzed, shared, or archived for multiple purposes. Nowadays, many medical applications may depend on image and video processing for diagnostics and decision-making processes, especially in telerehabilitation or remote surgery applications.

Images and videos are conventionally processed and transmitted at the pixel or bit level, regardless of their importance. When considering the images or videos of future medical applications, e.g., 3D X-ray images, organ scans, or even skeletal models, the amount of data becomes huge, making transmission over wireless networks prone to interference or noise. Also, the large volume of data may cause a drop in QoS and introduce errors.

On the other hand, with the help of SC, image and video data transmission may be reorganized depending on the status of the network. It only transmits the key target or crucial content within the images or videos and their relationship [11]. In SC-based surgical applications, for instance, not all elements within the video stream are of equal importance. SC can recognise and send crucial information which enhances the experience and the precision for both doctors and patients, under different wireless channel conditions.

C. Audio

The analysis of voice and speech tone in a medical setting may be of high importance, especially in diagnostic scenarios where patients describe symptoms or emotional states. Compared with other data forms, e.g., text or images, semantic representation for voice and speech is more complicated [11] because of the characteristics of voice tone, background noise, and signal delay.

Future medical applications may use SC to process voice data into low-dimensional semantic representations, which are synchronized with the audio signals depending on the condition of a network. In telemedicine, SC can filter out non-essential acoustic features, focusing on capturing the patient's voice and contextual meaning. Further, spoken language may be transformed into semantically rich text, transmitting only the task-relevant features. These characteristics are subsequently explained in the context of medical applications, e.g., mental health monitoring or diagnostic interviews, which are nurtured by a tailored KB. An SC-enriched system not merely brings robustness but also reduces network traffic while maintaining performance.

D. Other forms

There are other models in medical applications where SC can help, e.g., smell, a highly complex sensory experience. Unlike sound and images, which can be broken down into simpler components (e.g., pixels for images or frequencies for sound), which we can now semantically convey through a network without focusing on the number of units, smells are made up of complex combinations of molecules. Human olfaction (i.e., our sense of smell) involves hundreds of different receptors that detect thousands of possible odours, and these combinations vary widely between different odours [12]. Given their chemical complexity, reproducing smells requires a device that can store and combine many different chemical compounds in precise proportions.

Therefore integrating SC in medical applications with the help of end-to-end networks [13] may help technically convey smells into forms which can be easily transmitted via wireless networks. The experience of perceiving smell is highly subjective and varies from person to person. Therefore, it is difficult to create a universally consistent experience [12]. By understanding the meaning of complex molecules and semantically classifying them to extract their basic structural components, wireless networks may be able to convey these molecules according to their structure as a concept rather than converting them into the exact bits, pixels or units. For example, in the COVID-19 case, we can help the patient diagnose using a digital sensing test remotely with the help

This article has been accepted for publication in a future proceedings of this conference, but has not been fully edited. Content may change prior to final publication. Citation information: DOI: 10.1109/bdcat63179.2024.00066, 2024 IEEE/ACM International Conference on Big Data Computing, Applications and Technologies (BDCAT) - Workshop 2 - EIBBDESFT

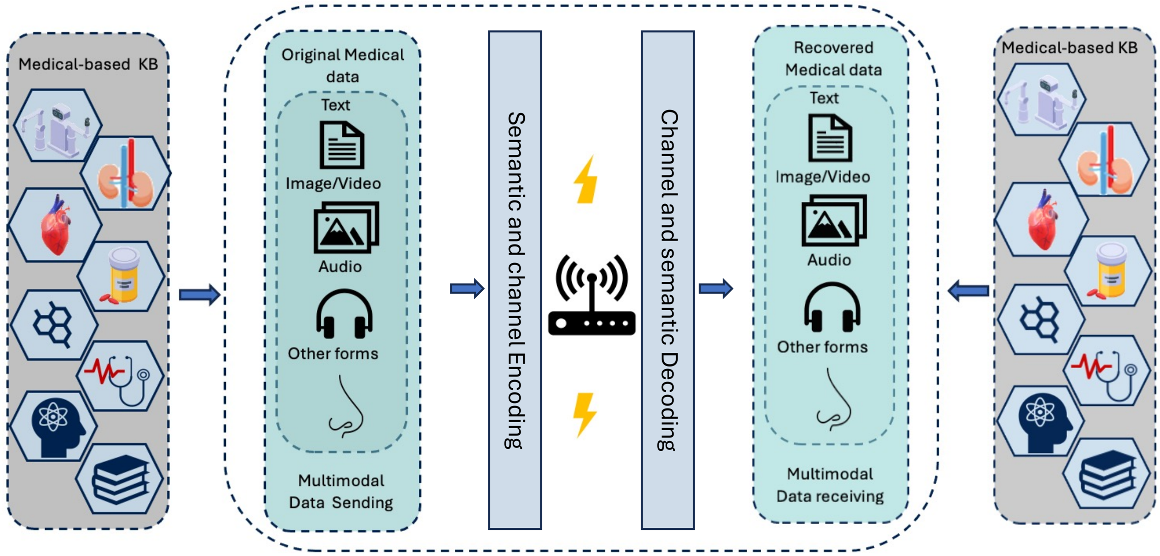

Fig. 3. Different types of semantic communication system.

of SC technologies. This technique can also be integrated into VR experiences which may require different types of sensory feelings.

E. Multimodal SC system

Multimodal SC systems are helpful for complex medical applications, e.g., real-time digital medical twin systems, which may require the combination of multidimensional medical data like text, image audio, and 3D video. It may be challenging to ensure low-latency and high-reliability communication if the raw data is transmitted due to large data volumes, especially in an emergency scenario like an earthquake, where some base stations may be destroyed, and wireless networks are affected. Robotic surgery can be an example, where 3D data may be transmitted from multiple sources, including medical imaging devices (CT/MRI), information from cameras, or haptic feedback. To reduce bandwidth usage and increase efficiency, different forms of modal data could be converted to one unified model, e.g., text modal, and then transmitted adaptively with the help of SC based on different wireless channel states and conditions. For example, in Fig. 4, we show that instead of transmitting the raw images, we may extract their key information in text and then transmit it over the wireless channel to save bandwidth. On the receiver side, we can get the key information using recovered semantics. Although the received and transmitted information may not be the same, we may still be able to recover the image via some techniques like [14] or AI-Generated Content (AIGC) [15], as its transmitted and received semantic information should be the same in general. Note that here unlike some medical image captioning techniques [16], which generate descriptive textual captions for medical images (e.g., X-rays, MRIs, and CT scans) using natural language, SC encodes information from medical images into a textual or symbolic format for efficient wireless transmission and accurate reconstruction or understanding on the receiver's end.

IV. SC Enhanced Healthcare System

In this section, we give some examples of different SC-enhanced healthcare systems.

A. SC-supported Telemedicine Healthcare Systems

Telemedicine systems allow patients to receive care from healthcare providers remotely, especially in areas with limited access to medical facilities. It may also be helpful for patients in remote areas or areas with poor network conditions. For example, healthcare professionals can remotely monitor patients using sensors, wearable devices or AR/VR equipment. In these scenarios, the key information instead of the whole data might be sent to the doctors/hospitals based on SC. In other words, depending on the network conditions, different semantic information with different data sizes will be sent in real-time. Additionally, in some emergency scenarios, if the base station is partially working, SC can be applied when the patient's voice is unclear, and the voice may be converted to an explanatory text summary and then displayed on the screen of smartphones in real time for clarity.

B. SC-supported Digital Medical Twin System

Medical digital twin systems can help doctors visualize and operate virtual/digital models or the simulation of patients' physical bodies or organs. Conventional medical digital twin systems may need real-time data from various sources, which may require high-speed wireless data communications. Also, it may be difficult to ensure accurate synchronization between the physical and virtual worlds, due to the limited amount of wireless bandwidth. With the support of SC, only the

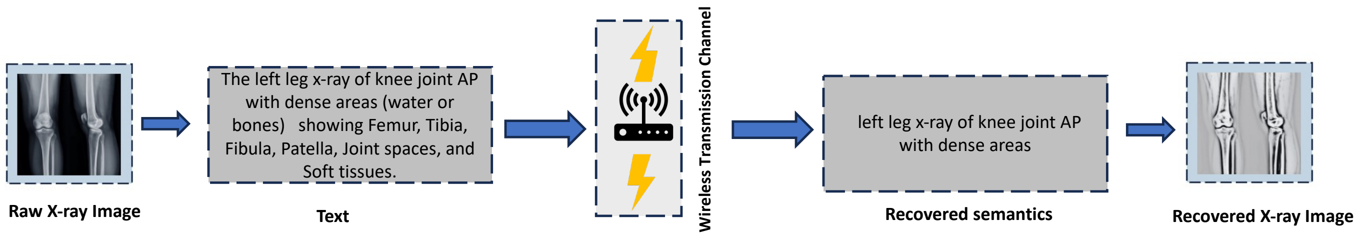

Fig. 4. The example of SC-enhanced text-based medical image application.

most relevant data or information with high priority will be transmitted and shared, e.g., key physiological metrics or essential alerts. For example, in a remote operation scenario, the doctor can do the surgery by using a remote console and controller, via a 3D VR digital system. For this system, the selection of some essential parts, or critical physiological metrics and the changing components might be synchronized through the wireless channel under different conditions.

C. *SC-supported Rehabilitation System*

Following the COVID-19 pandemic, many at-risk patients have shown a preference for telemedicine consultations with doctors whenever face-to-face consultations are not necessary. This saves time and decreases the chances of contracting other diseases. For example, some telerehabilitation applications, have been proposed for patients, including those with musculoskeletal disorders, to hold discussions about their health with doctors from the comfort of their homes thus avoiding the need for frequent visits to the hospital. Through these applications, doctors can communicate with patients via video conferencing and guide them through a series of physical activities.

Traditionally, the above-mentioned rehabilitation systems relied on a marker-based system, whereby patients have to wear sensors and perform specific movements in front of some special equipment. Recently, markerless systems [17] have been discussed, which may not rely on some specific setup, therefore patients may use their phone cameras to conduct remote consultations. In this case, doctors may monitor patients' movement in real-time using video conferencing. SC might be utilized here for improving the QoS, where the patients can do some physical therapy exercises as prescribed, and the key semantic information, including skeletal details, can be extracted, encoded and transmitted based on the different wireless network conditions, as shown in Fig. 5. Some techniques to extract the skeleton can be seen here [18], [19]. From the doctor's point of view, the patient's full-body movement may also be reconstructed in real-time if necessary based on the skeletal information obtained with the help of some AIGC technologies. This application ensures the effective transmission of data while maintaining the required quality and accuracy for effective remote rehabilitation.

V. Research Challenges and Future Directions

A. *Latency Requirement*

Real-time applications, e.g., robotic surgery or AR/VR-assisted rehabilitation systems, typically require a low-latency response. Despite the promises of ultra-low latency from 5G, real deployments are tough. Although SC can significantly reduce the transmission bandwidth, it may require some processing that involves additional computation to extract semantic information from raw data. However, these medical mobile devices or wearable sensors usually have limited power capacity, due to their small size, which may not be able to conduct complex semantic extraction and transmission. We might explore local energy-efficient or lightweight AI/ML models that run on low-power devices. Additionally, we may consider offloading the computational overhead to the edge or cloud to help with this process.

B. *Strict QoS Requirement*

In critical medical applications, errors or losses of life-critical data are intolerable, especially in surgery, where real-time video feeds or haptic feedback should be as accurate and reliable as possible. SC involves compressing or discarding less relevant data, which may lead to the risk of losing essential medical information. Hence, ensuring robust error-checking mechanisms and reliable QoS guarantees while implementing semantic compression is paramount. One may develop some reliable SC frameworks that use error detection and correction algorithms specifically for healthcare applications.

C. *Key Semantic Information Extraction*

In medical applications, healthcare data may come from different sources, e.g., MRI, CT, or motion sensors. Identifying the key information suitable for wireless transmission is a big challenge. In other words, we could prioritize the relevant parts of the original information and ensure the delivery of critical clinical information through the wireless networks. Moreover, we may convert or unify different models (e.g., visual, textual, or tactile) into low-complex models (e.g., text or bits) for efficient transmission.

D. *SC-related Healthcare Standard*

Different healthcare applications may have different requirements, and it is very challenging to implement a unified framework. They may vary depending on devices, data formats,

This article has been accepted for publication in a future proceedings of this conference, but has not been fully edited. Content may change prior to final publication. Citation information: DOI: 10.1109/bdcat63179.2024.00066, 2024 IEEE/ACM International Conference on Big Data Computing, Applications and Technologies (BDCAT) - Workshop 2 - EIBBDESFT

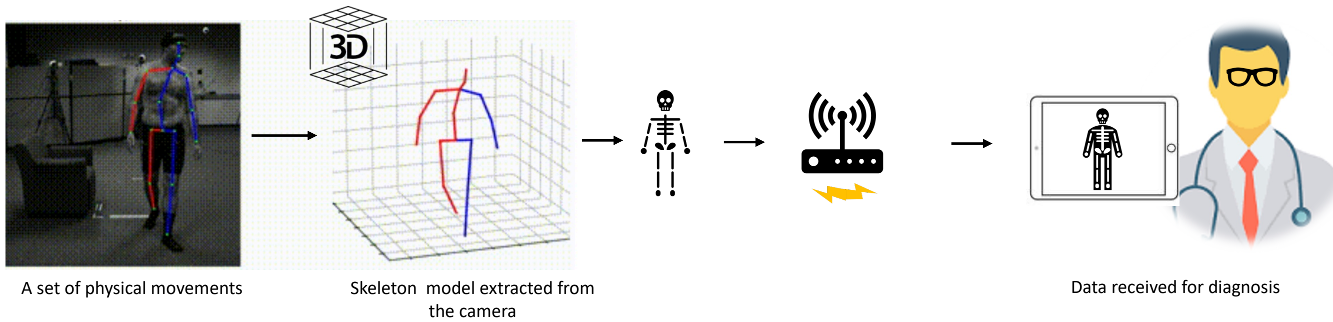

Fig. 5. The example of SC-based rehabilitation system

and modality [20]. With the introduction of SC, healthcare professionals may integrate SC models with existing healthcare systems and infrastructures. Interoperable protocols and standards that handle data from diverse devices and systems may be highly required.

E. Security and Privacy Concerns

Healthcare data is susceptible and subject to strict regulatory frameworks. Some sensitive user data, such as personal information, raises privacy concerns and requires robust security measures. When we send critical information using SC, it may be more vulnerable to privacy or security violations due to key information being transmitted. For example, the patient's name, critical health condition, and other sensitive information may be leaked. Privacy-preserving or security-enhanced SC technologies, e.g., federated learning or homomorphic encryption, which allow data to be processed without revealing sensitive information, may be explored.

VI. Conclusion

In this paper, we have introduced SC for healthcare application systems. First, we have given a brief introduction to semantic communications. Then, we have discussed different types of SC in healthcare applications, e.g., text, video, audio and images. We have also introduced some possible use cases for potential medical applications. Finally, we have discussed some challenges and future research directions.

VII. Acknowledgment

We would like to acknowledge the support in part by the EU Eureka 5G4PHealth project (via Innovate UK, Grant No. 10093679), as well as the Horizon Europe HarmonicAI project Grant No. 101131117 (via UKRI Grant No. EP/Y03743X/1).

References

[1] Z. Zhou, Z. Wang, H. Yu, H. Liao, S. Mumtaz, L. Oliveira, and V. Frascolla, "Learning-based URLLC-aware task offloading for Internet of Health Things," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 2, pp. 396–410, 2021.

[2] M. Simsek, A. Aijaz, M. Dohler, J. Sachs, and G. Fettweis, "5G-enabled tactile internet," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 3, pp. 460–473, 2016.

[3] S. Razdan and S. Sharma, "Internet of Medical Things (IoMT): Overview, emerging technologies, and case studies," *IETE Technical Review*, vol. 39, no. 4, pp. 775–788, 2022.

[4] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*. The University of Illinois Press, 1949.

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

[6] 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.

[7] Q. Lan, D. Wen, Z. Zhang, Q. Zeng, X. Chen, P. Popovski, and K. Huang, "What is semantic communication? a view on conveying meaning in the era of machine intelligence," *Journal of Communications and Information Networks*, vol. 6, no. 4, pp. 336–371, 2021.

[8] F. Jiang, Y. Peng, L. Dong, K. Wang, K. Yang, C. Pan, and X. You, "Large AI model-based semantic communications," *IEEE Wireless Communications*, vol. 31, no. 3, pp. 68–75, 2024.

[9] E. Bourtsoulatze, D. B. Kurka, and D. Gündüz, "Deep joint source-channel coding for wireless image transmission," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 4774–4778.

[10] F. Jiang, L. Dong, Y. Peng, K. Wang, K. Yang, C. Pan, and X. You, "Large AI model empowered multimodal semantic communications," *IEEE Communications Magazine*, pp. 1–7, 2024.

[11] Z. Qin, F. Gao, B. Lin, X. Tao, G. Liu, and C. Pan, "A generalized semantic communication system: From sources to channels," *IEEE Wireless Communications*, vol. 30, no. 3, pp. 18–26, 2023.

[12] Y. Liu, C. K. Yiu, and et al, "Soft, miniaturized, wireless olfactory interface for virtual reality," *Nature Communications*, vol. 14, 2023.

[13] H. Ye, L. Liang, G. Y. Li, and B.-H. Juang, "Deep learning-based end-to-end wireless communication systems with conditional gans as unknown channels," *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 3133–3143, 2020.

[14] M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, and M. J. Black, *SMPL: A Skinned Multi-Person Linear Model*, 1st ed. New York, NY, USA: Association for Computing Machinery, 2023.

[15] Y. Cao, S. Li, Y. Liu, Z. Yan, Y. Dai, P. S. Yu, and L. Sun, "A comprehensive survey of AI-generated content (AIGC): A history of generative AI from GAN to ChatGPT," *arXiv preprint arXiv:2303.04226*, 2023.

[16] M. Saraswat, C. V. Reddy, and G. Y. Singh, "Image captioning using NLP," in *2024 First International Conference on Technological Innovations and Advance Computing (TIACOMP)*, 2024, pp. 549–553.

[17] S. L. Colyer, M. Evans, D. P. Cosker, and A. I. Salo, "A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system," *Sports medicine-open*, vol. 4, pp. 1–15, 2018.

[18] S. Ghorbani, K. Mahdaviani, A. Thaler, K. Kording, D. J. Cook, G. Blohm, and N. F. Troje, "MoVi: A Large Multipurpose Motion and Video Dataset," 2020.

[19] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "Openpose: Realtime multi-person 2D pose estimation using part affinity fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 1, pp. 172–186, 2021.

[20] S. Sachdeva and S. Bhalla, "Semantic interoperability in standardized electronic health record databases," *ACM J. Data Inf. Qual.*, vol. 3, pp. 1:1–1:37, 2012.