



HAL
open science

Streaming DICOM Real-Time Video and Metadata Flows Outside The Operating Room

Saad El Jaouhari, Guillaume Pasquier, Amélie Serrand, Bernard Gibaud,
Patrick Hardy, Emmanuel Cordonnier

► **To cite this version:**

Saad El Jaouhari, Guillaume Pasquier, Amélie Serrand, Bernard Gibaud, Patrick Hardy, et al.. Streaming DICOM Real-Time Video and Metadata Flows Outside The Operating Room. IEEE International Conference on E-health Networking, Application & Services (Healthcom 2019), Oct 2019, Bogota, Colombia. hal-02431267

HAL Id: hal-02431267

<https://hal.science/hal-02431267>

Submitted on 7 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Streaming DICOM Real-Time Video and Metadata Flows Outside The Operating Room

Saad EL JAOUHARI* , Guillaume PASQUIER* , Amélie SERRAND* , Bernard GIBAUD† ,
Patrick HARDY‡ and Emmanuel CORDONNIER*

*B-com, Cesson Sévigné, France, †Univ Rennes, Inserm, LTSI - UMR 1099, F-35000 Rennes, France

‡Harmonic, Cesson Sévigné, France

Emails: *FirstName.LastName@b-com.com, †bernard.gibaud@univ-rennes1.fr, ‡patrick.hardy@harmonicinc.com

Abstract—With the current advancement in the medical world, surgeons are faced with the challenge of handling many sources of medical information in more and more complex and technological Operating Rooms (ORs). Obviously, in the next generation ones, there will be an increasing number of video flows during the surgery (e.g. endoscopes, cameras, ultrasounds, etc.), which can be also displayed all over the OR in order to facilitate the task for the surgeon and to avoid any adverse events or problems related to inadequate communication in the OR. Additionally, other information needs to be shared, pre/post/during an operation, such as the history of the digital images related to the patient in the PACS and the metadata coming from medical sensors. Moreover, these medical videos captured from the OR can be either displayed on a large screen in the OR in order to provide the surgeon with more visibility, in this case via DICOM-RTV, or streamed outside the OR via a P2P solution. The latter one can serve various purposes such as for teaching medical student in real-time or for remote-expertise with a remote senior surgeons. Hence, this paper addresses the challenges of streaming DICOM-RTV video and metadata flows live from the operating room, typically during an ongoing surgery, in real-time to the outside world. A Proof of Concept is also presented in order to demonstrate the feasibility of our solution.

Index Terms—DICOM, DICOM-RTV, WebRTC, Healthcare, Medical Metadata, Streaming, Real-Time, Education, Telemetering.

I. INTRODUCTION

DICOM (Digital Imaging and Communications in Medicine) [1] is currently the standard that defines how medical images must be managed inside a medical infrastructure. It basically specifies how the medical images should be acquired, transferred, processed and stored, and currently recognized as a standard by European Committee for Standardization (CEN), and International Organization for Standardization (ISO).

As an extension, a new supplement to DICOM standard, called DICOM-RTV (DICOM Real-Time Video) [2] is emerging. It allows the streaming of medical videos directly from the operating room, typically during a surgery, while synchronizing them with the medical metadata of the patient subject of surgery. This standard takes advantage of the multiplexing capability of IP, in order to transport multiple synchronized video/audio/data flows. The DICOM-RTV communication service is based on the SMPTE

ST 2110-10 architecture [3] (from the SMPTE ST 2110 family of standards, widely used in the professional video equipment(e.g., TV studios)).

Undoubtedly, in the next generation Operating Rooms (ORs), there will be an increasing number of video flows during the surgery (e.g. endoscopes, cameras, ultrasounds, etc.), which can be also displayed all over the OR in order to facilitate the surgeon's tasks and to better share information with the surgical staff. A factor that may help reducing adverse events. These medical videos captured in the OR can be either displayed on a large screen in the OR in order to provide the surgeon with more visibility, in this case via DICOM-RTV, or streamed outside the OR via a peer-to-peer (P2P) solution. The latter one can serve various purposes such as for teaching medical students in real-time or for remote-expertise with remote senior surgeons. Furthermore, other medical information needs to be shared, pre/post/during an operation, such as the history of the digital images related to the patient in the Picture Archiving and Communication System (PACS) and the medical metadata coming from medical sensors.

This paper focuses on this second part, namely the challenge of live-streaming the real-time DICOM video flows outside the OR. The video flow can be either a simple video flow coming from an endoscope in the surgery room, or it can be a perfectly synchronizing video combining several videos flows from the OR. Moreover, the paper addresses also the challenge of transferring DICOM metadata in real-time in parallel with the video flow. For this purpose, the Web Real-Time Communication (WebRTC) solution seems to be the appropriate solution since it allows real-time and peer-to-peer exchange of these flows using simply a compatible web browser, without the need to add or install additional plugins or software.

The remaining of this paper is structured as follows: in section II, a literature review of the existing scientific works and a quick background introduction of WebRTC are presented. Then, in section III, the architecture together with the related use cases of the proposed solution is introduced. Next, the

main results of the proposed solutions and discussions are presented in section IV. Finally, the section V concludes the paper with a summary and some perspectives.

II. RELATED WORK AND BACKGROUND

In the literature, none of the proposed solutions that use WebRTC deals with the live DICOM video and metadata flows coming directly from the operating room. However, some works use WebRTC in order to deal with DICOM images which shows the relevance of our proposed solution.

In [4], the authors describe a medical visualization system suitable for both diagnosis and medical training. The proposed system uses the patient DICOM images (e.g. Magnetic Resonance Imaging (MRI) scan), converts them into Portable Network Graphics (PNG) format, edits them using an online image editor, and reconstructs them into 3D model. The reconstructed images will be then combined with a web camera feed of the patient organ in question. The mixed image can be afterwards transmitted to a remote browser with WebRTC (e.g. for remote expertise). However, mixing the images and transmitting them using WebRTC is a future work that has not been developed. In [5], the authors present a web-based solution for collaborative DICOM medical image visualization for decision support in a telemedicine environment. They focus mainly on the bucomaxillofacial field (teledentistry). The solution allows participants (i.e. remote doctors) to visualize and manipulate DICOM images (extracted from the PACS) in a synchronized way, while visio-conferencing or just chatting using WebRTC. The authors in [6] and [7] also provide a web-based and real-time collaboration platform for on-line MTDMs (Multidisciplinary Team Meetings), by using WebRTC among other technologies and Application Programming Interfaces (APIs). It allows a secure access to DICOM medical images from the PACS, sending them via the WebRTC data channel and the image processing at the client side, while visio-conferencing with other peers. In [8], the authors give an overview of the DICOM standard and discuss the position of the P2P systems in the current development of the DICOM/PACS system. They mention that the P2P systems have been neglected in the DICOM/PACS domain, and highlight that WebRTC can be a solution toward P2P DICOM networks.

Web Real-Time Communication (WebRTC)

WebRTC [9] [10] provides real-time and peer-to-peer communication capability between multiple users with just their browsers. Moreover, many browsers such as Google Chrome or Mozilla Firefox natively support WebRTC. It is a free, open source project that provides browsers with real-time capability by implementing the Datagram Transport Layer Security (DTLS) protocol and by finally using the SRTP (Secure RTP (encryption embedded)) protocol in order to secure the exchanges. One of the major advantages of WebRTC is its capability to transport various types of data

using the WebRTC data channels. It utilizes SCTP (Stream Control Transmission Protocol), which is message-oriented, and transported via UDP while ensuring reliable, in-sequence transport of messages with congestion control like TCP. It comes also with the possibility of securing the exchanges via the implementation of DTLS.

WebRTC uses mainly three APIs: *MediaStream*, which is responsible for accessing the Input/Output multimedia devices and the creating the media stream. *RTCPeerConnection*, which is responsible for the creation of a secure media channel via SRTP; it also encompasses signaling, Network address translation (NAT) traversal and other important tasks. Finally, *RTCDataChannel*, which allows the creation of a secure data channel via SCTP encapsulated in DTLS. One of the advantages of WebRTC compared to the other real-time communication technologies is that the security was taken into account from the beginning, thus providing a strong security architecture.

III. PROPOSED SOLUTION

A. Architecture

The architecture describes the main components that allow the live streaming of the videos from the operating rooms, together with the corresponding metadata. As shown in Fig. 1, these components can be divided into two parts, the components present in the OR (i.e. the DICOM-RTV source and the DICOM-RTV P2P Manager) and the components that allow the streaming of the video outside the OR represented by the WebRTC architecture.

DICOM-RTV source: is the source of two types of flows: i) a metadata flow transporting information about the patient being operated, ii) one or more RAW uncompressed DICOM video flows, produced by medical equipment such as video cameras, endoscopes, fluoroscopes and so forth. The format of the RAW video is explained later in section III-C and the structure of the DICOM metadata is presented in section III-D. Both flows can then be broadcasted, multicasted or simply unicasted depending on the needs via the DICOM-RTV source. The DICOM-RTV video essences are complying with the proposed DICOM-RTV standard [2] (approved and in its final steps for publication) based on SMPTE 2110-20 [11]. These video essences are then transported via IP on separate flows and in a synchronized way. The DICOM metadata is extracted directly from the hospital worklist, and can be easily extended to interact also with the PACS using the DICOM-DIMSE or the DICOMweb™ [1] protocols in order to extract the patient's corresponding medical images.

DICOM-RTV P2P Manager: is the component that adds the peer-to-peer capability to the DICOM-RTV standard. It has two main parts: 1) the video part, where the captured DICOM-RTV video flow is adapted in order to be usable by the WebRTC RTP protocol, and 2) the metadata part, where the DICOM metadata are decapsulated from the

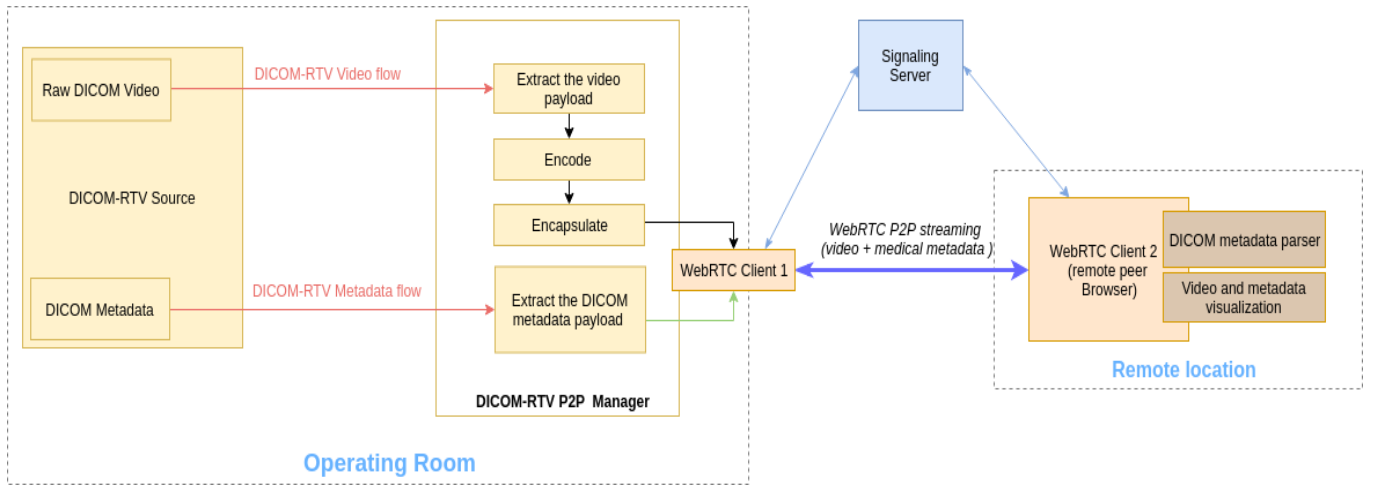


Fig. 1. General architecture

UDP/RTP packet and sent via the WebRTC data channel. Details regarding the video manipulation are presented later in section III-C. Both flows are then streamed via WebRTC.

WebRTC: composed of the two clients and the signaling server. The **WebRTC client 1** represents the extremity of the operating room and the exit point to the outside world. In order to allow a client 2 (e.g. a remote expert) to view the live streaming, a WebRTC session needs to be established between the two clients. Via the signaling server, the client 1 and 2 exchange Session Description Protocol (SDP) messages, containing the streaming media communication parameters such as the IP address (for the connectivity) and the supported codecs. Once done, the client 1 can stream the video flow via the RTP media channel and the DICOM metadata flow via the data channel directly with the client 2 in a peer-to-peer fashion. On the **WebRTC client 2** side, the video stream is displayed via HTML5/AngularJS, and the DICOM metadata is parsed at the client side via the *dicom-parser* tool [12] as explained in section III-D.

One of the main innovative features of the proposed solution is the exploitation of client-side computation capabilities using the WebRTC protocol, by transporting the live-streamed video flow (and audio if present), and the corresponding DICOM medical data in parallel channels, and by parsing and interpreting the DICOM metadata on the client side using suitable parsers. Secondly, it adopts state-of-the-art web technologies in order to implement the different components of the framework (e.g. AngularJS, HTML5, Dockerized components). The peer-to-peer nature of WebRTC eliminates any intermediate party overheads and their security threats (such as Man-In-The-Middle or replay attacks). A second feature is enabling the P2P connectivity so as to facilitate the access to the medical video and metadata from the OR by a remote peer by simply using their browser.

B. Use cases

In this paper, two main use cases are presented in order to illustrate the relevance of our proposed solution.

1) *Remote expertise (known also as telementoring):* live from the OR, and in almost real-time, a remote doctor can follow an ongoing surgery using only a compatible browser in a simple and transparent way. For instance a senior surgeon may supervise and even guide a junior surgeon in real-time from outside the OR. Furthermore, DICOM metadata containing all the necessary information related to the patient in question are sent periodically via WebRTC to the remote expert in order to provide him/her with a clearer view of the situation in the OR. The main purpose is to reduce the human error rate during a critical situation such as a surgery.

2) *Education:* the live streaming of an ongoing surgery from the OR can be a great learning tool for the medical students and especially future surgeons. For instance the trainees, medical students can witness the surgeons hands gestures during the operation, while having their professor explaining the operation phases and the reasons for each decision. The proposed solution is easy to use and does not require any additional installation. The student can simply use a WebRTC compatible browser in order to view the video stream and the DICOM metadata.

C. Video manipulation

The uncompressed active video used in our tests, which follows the SMPTE 2110-20 [11] standards, contains the video flow captured from an endoscope as shown in Fig. 2. The RAW video is then transported from the OR (DICOM-RTV source) to the DICOM-RTV P2P using an RTP format based on the RFC 4175 [13] via IP. In our tests the videos are being multicasted in the lab. The transported RAW video

has the following format:

Clock-rate	90000
Encoding-name	RAW
Sampling	YCbCr-4:2:2
Depth	10
Width	1280
Height	720
Payload type	96
Framerate	50 Hz

The DICOM-RTV P2P Manager subscribes to the multicast video flow and extracts the RAW video from the RTP packets using the *GStreamer* tool [14]. The second step is to encode the video, since WebRTC does not support sending RAW video directly. Hence, the video is encoded into a format supported by the browsers, namely VP8 or H264. In our case, the RAW video format is converted from the source format (any YUV format other than I420) to the I420 format since the Gstreamer encoders do not support the source video format. Finally, the result is transported also using an RTP encapsulation specified in RFC 3550 [15], as required by WebRTC in [16]. The RTP payload format depends on the used encoding. For H264 codec the RFC 6184 [17] is used, and for the VP8 codec the RFC 7741 [18] is used.

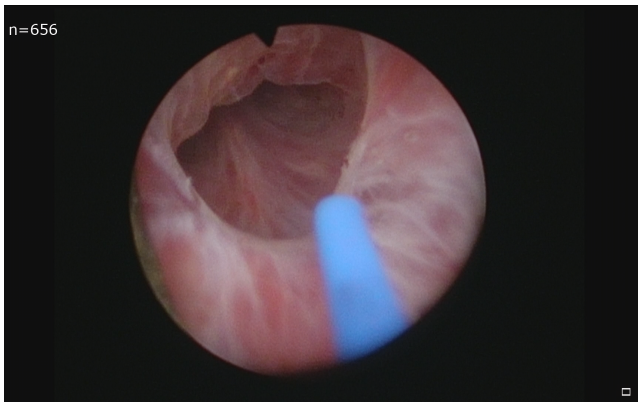


Fig. 2. Endoscope video from the OR streamed via WebRTC

D. DICOM metadata

The DICOM-RTV P2P subscribes to the metadata multicast flow using the WebSocket protocol, and extracts the RAW DICOM metadata from the RTP packets. The received packets are structured following the format shown in Fig. 3 and explained in detail in [2]. The DICOM-RTV P2P removes all the header and extracts the DICOM payload from the received packets, i.e. the static and the dynamic parts of the DICOM dataset. These data are then sent directly via the WebRTC data channel in parallel with the real-time video flow. Unfortunately, the current WebRTC does not provide a mechanism for synchronizing the metadata with the streamed video. However, the DICOM-RTV metadata is timestamped for future synchronization purpose.

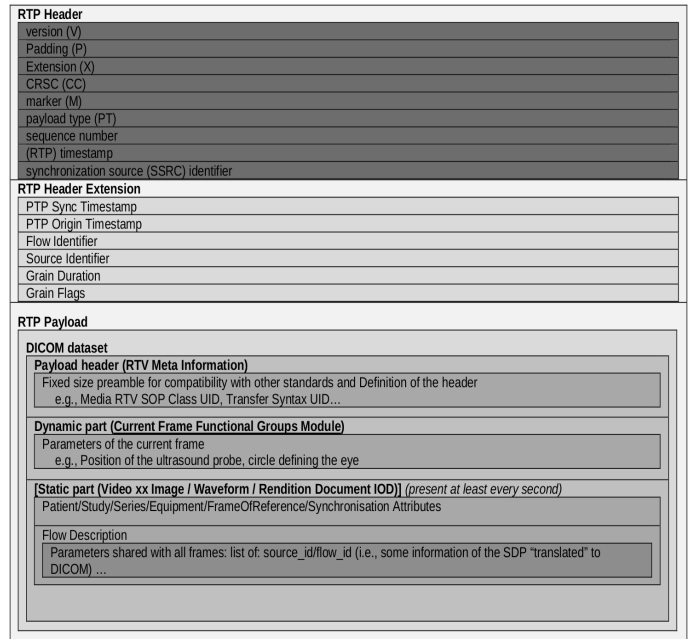


Fig. 3. DICOM dataset encapsulation within RTP [2]

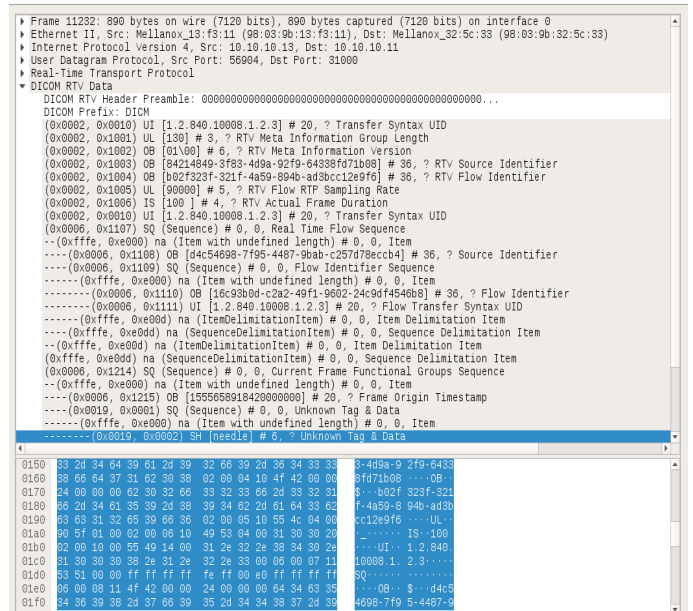


Fig. 4. DICOM parsed data using Wireshark

An example of the received and parsed DICOM metadata (only the RTP payload), captured via Wireshark, is shown in Fig. 4. In this example, the 3D position of the needle inside the patient being operated is transmitted in real-time in the DICOM-RTV metadata.

IV. RESULTS AND DISCUSSION

In order to demonstrate the feasibility of our approach, a Proof of Concept (PoC) of the proposed solution is implemented in our lab. The different components that exist inside the OR

surgeons, clarified that clearly the latency is an important challenge that needs to be as low as possible. It can be crucial in many cases, in particular for surgeries that require about a dozen of milliseconds, such as for the surgical precision and task completion during the future robotic-assisted remote telepresence surgery. For short terms, the obtained latency still satisfies the surgeons when it comes to human interaction with a remote expert, such as providing an advice or guidance during an ongoing surgery that do not require instantaneous reaction, however, in future work, it needs to be optimized and kept below 200 ms, as studied in [20].

As for the security considerations, and as a first step, all along the process, the data remains within the private network of the medical infrastructure, which can be an infrastructure with multiple delocalized hospitals, and where the communication between the different hospitals is established via a VPN (Virtual Private Network). Therefore, in this case, the remote expert or medical students (for education goal) must be physically inside one of the hospitals, and connected to WebRTC via the hospital's VPN to the remote real-time stream of the ongoing surgery. Moreover, the signaling server is also implemented inside the technical room of the hospital with the purpose of avoiding any data leak and to keep the private sensitive medical data of the patient within the medical infrastructure network. Additionally, WebRTC offers multiple mechanisms in order to secure the media and data flows with the remote peer by using Datagram Transport Layer Security (DTLS) over Stream Control Transmission Protocol (SCTP) to transport the metadata and Secure Real-time Transport Protocol (SRTP) as a means to transport the media flow. Furthermore, future works are planned to implement additional security mechanisms so as to:

- Identify and authenticate the peer that wants to access the platform.
- To control the access, so that only the authorized peers can visualize the real-time surgery and the corresponding metadata, via access control mechanisms such as Role-Based Access Control (RBAC) [21] or Attribute-based access control (ABAC) [22].
- To de-identify (i.e. anonymize) the patient data when needed. For instance an authorized remote expert can access the medical data of the patient, but medical students who are watching the real-time flow only see de-identified metadata.

V. CONCLUSION

In this paper, a solution to the challenges of streaming video flows from the operating room, typically during an ongoing surgery, is presented. The proposed solution leans on the WebRTC technology in order to stream the video and the corresponding metadata in real-time to the remote peer. The proposed solution was illustrated via two use cases. A PoC was also presented in order to demonstrate our proposal. As future work, we are interested in enabling real-time feedbacks

from the remote expert. Solutions, including the use of Augmented Reality (AR), in order to allow the remote expert to guide/advise the (junior) surgeon while he/she is operating the patient, can be a starting point toward solving this challenge.

REFERENCES

- [1] O. S. Pinykh, *Digital imaging and communications in medicine (DICOM): a practical introduction and survival guide*. Springer Science & Business Media, 2009.
- [2] "Digital Imaging and Communications in Medicine (DICOM) Supplement 202: Real-Time Video," Jan. 2019.
- [3] "ST 2110-10:2017 - SMPTE Standard - Professional Media Over Managed IP Networks: System Timing and Definitions," *ST 2110-10:2017*, pp. 1–17, Nov 2017.
- [4] I. Virag, L. Stoicu-Tivadar, and E. Amrici, "Browser-based medical visualization system," in *2014 IEEE 9th IEEE International Symposium on Applied Computational Intelligence and Informatics (SACI)*, May 2014, pp. 355–359.
- [5] A. Puel, A. v. Wangenheim, M. I. Meurer, and D. D. J. d. Macedo, "BUCOMAX: Collaborative Multimedia Platform for Real Time Manipulation and Visualization of Bucomaxillofacial Diagnostic Images," in *2014 IEEE 27th International Symposium on Computer-Based Medical Systems*, May 2014, pp. 392–395.
- [6] C. Andrikos, G. Rassias, P. Tsanakas, and I. Maglogiannis, "Real-time medical collaboration services over the web," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2015, pp. 1393–1396.
- [7] —, "An Enhanced Device-Transparent Real-Time Teleconsultation Environment for Radiologists," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 1, pp. 374–386, Jan 2019.
- [8] I. Dmasin and M. Grgi, "Is There Still a Place for Peer-to-Peer in the DICOM World?" in *2018 25th International Conference on Systems, Signals and Image Processing (IWSSIP)*, June 2018, pp. 1–5.
- [9] A. Sergiienko, *WebRTC Blueprints*. Packt Publishing Ltd, 2014.
- [10] C. Jennings, T. Hardie, and M. Westerlund, "Real-time communications for the web," *IEEE Communications Magazine*, vol. 51, no. 4, pp. 20–26, 2013.
- [11] "ST 2110-20:2017 - SMPTE Standard - Professional Media Over Managed IP Networks: Uncompressed Active Video," *ST 2110-20:2017*, pp. 1–22, Nov 2017.
- [12] C. Hafey, "dicom-parser," 2016, last visited: 15/03/2019. [Online]. Available: <https://www.npmjs.com/package/dicom-parser>
- [13] L. Gharai and C. Perkins, "RTP Payload Format for Uncompressed Video," IETF, RFC 4175, Sep. 2005. [Online]. Available: <http://tools.ietf.org/rfc/rfc4175.txt>
- [14] "GStreamer- open source multimedia framework," last visited: 15/03/2019. [Online]. Available: <https://gstreamer.freedesktop.org>
- [15] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications," IETF, RFC 3550, Jul. 2003. [Online]. Available: <http://tools.ietf.org/rfc/rfc3550.txt>
- [16] P. C., M. Westerlund, and J. Ott, "Web Real-Time Communication (WebRTC): Media Transport and Use of RTP," IETF, Internet-Draft, 2016. [Online]. Available: <https://tools.ietf.org/html/draft-ietf-rtcweb-rtcp-usage-26>
- [17] Y.-K. Wang, R. Even, T. Kristensen, and R. Jesup, "RTP Payload Format for H.264 Video," IETF, RFC 6184, May 2011. [Online]. Available: <http://tools.ietf.org/rfc/rfc6184.txt>
- [18] P. Westin, H. Lundin, M. Glover, J. Uberti, and F. Galligan, "RTP Payload Format for VP8 Video," IETF, RFC 7741, Mar. 2016. [Online]. Available: <http://tools.ietf.org/rfc/rfc7741.txt>
- [19] H. Alvestrand and V. Singh, "Identifiers for webrtc statistics api," *W3C, W3C Working Draft*, 2018.
- [20] Perez, Manuela and Xu, Song and Chauhan, Sanket and Tanaka, Alyssa and Simpson, Khara and Abdul-Muhsin, Haidar and Smith, Roger, "Impact of delay on telesurgical performance: study on the robotic simulator dV-Trainer," *International Journal of Computer Assisted Radiology and Surgery*, vol. 11, no. 4, pp. 581–587, Apr 2016. [Online]. Available: <https://doi.org/10.1007/s11548-015-1306-y>
- [21] R. S. Sandhu, E. J. Coyne, H. L. Feinstein, and C. E. Youman, "Role-based access control models." IEEE Computer Society Press, 1996.
- [22] M. A. Al-Kahtani and R. Sandhu, "A model for attribute-based user-role assignment," in *Computer Security Applications Conference, 2002. Proceedings. 18th Annual.* IEEE, 2002, pp. 353–362.