Development of an AR tool using an Optical See-Trough Display for Surgical Applications

Alice Carcone
Medical Robotics and Technologies for Computer Assisted Surgery
Politecnico di Milano

Milano, Italy
alice.carcone@mail.polimi.it

Chiara Leoncavallo
Medical Robotics and Technologies for Computer Assisted Surgery
Politecnico di Milano

Milano, Italy
chiara.leoncavallo@mail.polimi.it

***Abstract* – Nephron-sparing surgery in pediatrics focuses on preserving kidney function by selectively removing malignant lesions, a crucial approach to minimize long-term complications in children. The precise localization of the tumor plays a fundamental role in ensuring the optimal outcome of the surgical intervention. To this end, this research explores the incorporation of Augmented Reality (AR) in pediatric spare nephrectomy procedures. Leveraging the Microsoft HoloLens 2, a mixed-reality headset, an immersive Augmented Reality (AR) application designed for intraoperative guidance is presented. The application offers the capability to adjust organ surface transparency levels and manually manipulate the organ of interest, allowing for seamless overlap with the real organ during surgical procedures. A quantitative evaluation, based on the Intersection over Union (IoU) metric between the virtual and real left kidney, was performed to assess system accuracy. Preliminary results indicate a high IoU in a brief duration, indicating the potential effectiveness of our AR application in pediatric kidney surgery.**

I. INTRODUCTION

Wilms tumors (WTs) are the most frequently occurring pediatric cancers of the kidney, with children being diagnosed at a median age of approximately 3.5 years [1]. The survival rate of children with WT is around 90%. The therapy generally consists of neo-adjuvant chemotherapy, followed by nephrectomy and adjuvant chemotherapy. Open surgery provides surgeons with a clearer view and better access in these situations, considering the dimension constraints associated with pediatric anatomy, allowing for precise manipulation and meticulous control during complex procedures. In contrast to a radical nephrectomy for local treatment, nephron-sparing surgery (NSS) can be used for nephrogenic preservation. This helps to protect the patient from long-term functional parenchymal loss.

NSS is a technically demanding procedure and it requires a thorough preoperative understanding of the patient-specific renal anatomy and intra-parenchymal vasculature. Therefore, a personalized planning and surgical strategy is essential. Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) are used for diagnosis and to differentiate between tumor and healthy renal tissue. Pediatric surgeons plan the surgery based on the 2D interpretation of these conventional imaging techniques. Data from MRI and CT can be used to reconstruct 3D virtual models that enable physicians to analyze complex anatomical structures in an interactive way and enhance their perspective of the surgical site, as shown in Figure 1. [2]-[4]

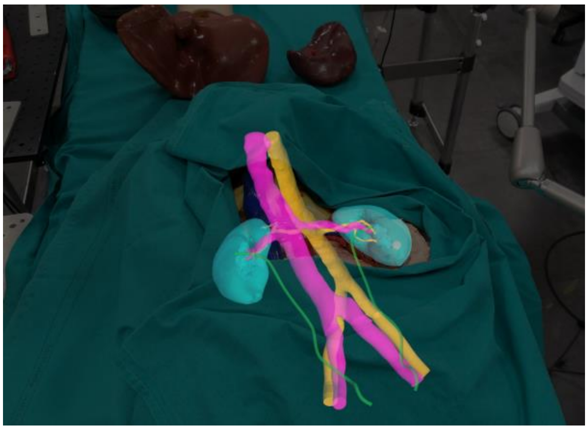

Fig. 1. Intra-operative visualization of patient-specific anatomy through HoloLens 2 see-through display.

Recent studies proposed the use of 3D visualizations to improve the understanding of the exact intrarenal tumor location and the assessment of relevant anatomical structures, such as arteries, veins, and urinary collection structures. [2]

Visualizing 3D models superimposed on actual anatomy during surgery would significantly support surgeons in increasing accuracy and safety in the intra-operative process.

Our contributions include a comprehensive review of the current state of the art in Augmented Reality surgical applications. Based on these studies, we developed an Augmented Reality (AR) tool in Unity and integrated it with an optical see-through display for surgical applications (HoloLens 2, Microsoft Corporation), as shown in Fig. 1. A quantitative evaluation was performed to assess the system accuracy, in order to demonstrate its feasibility. This study aims to contribute valuable insights into the application of AR technology in pediatric NSS, with the ultimate goal of improving accuracy, usability, and overall surgical outcomes.

The paper is structured as follows: Section II provides an overview of the state of the art about AR and its application in the medical field, with a focus on the usage of HoloLens 2. Section III details the materials and methods employed in constructing the virtual environment scene using the Mixed Reality Toolkit. along with the functionalities it provides. Section IV offers a comprehensive quantitative evaluation of registration accuracy, utilizing Intersection over Union (IoU) metrics. The results and Conclusion are presented in Sections V and VI, respectively.

II. STATE OF THE ART

A. *Augmented Reality*

Before delving into the subject, we believe it is pertinent to define what are immersive technologies and the concept of ‘Virtuality Continuum’ (Fig. 2). This continuum represents a spectrum that spans from a wholly real environment to an entirely virtual one.

Augmented Reality (AR) resides closer to the real-world end of the spectrum, where digital elements are merely superimposed onto our physical surroundings. In contrast, Virtual Reality (VR) sits at the opposite end, offering a fully digital environment that detaches users from the real world. Mixed Reality (MR) occupies a unique space in the middle, merging the real and virtual to such an extent that digital and physical objects can interact in real time. Extended Reality (XR) serves as an umbrella term that encompasses the entire spectrum, highlighting the seamless merge of real and virtual environments.

Both Augmented Reality (AR) and Mixed Reality (MR) are generating significant interest within the surgical community. These technologies hold the potential to revolutionize surgical practices by providing enhanced visualization, real-time data overlay, and interactive guidance, ultimately contributing to improved surgical precision and patient outcomes. Mixed Reality (MR) applications empower medical professionals to virtually recreate real-world images of anatomical structures. This is achieved through various digital display technologies, including [5]:

- Projection onto patients: the patient becomes the screen and anatomical images can be projected onto his abdomen.
- Optical see-through: projection of the augmentation on semi-transparent surfaces using monocular or binocular technologies.
- Video see-through: The surgical scene is captured in real-time by the tracked mobile display device's camera (VR headsets or mobile devices like smartphones or tablets) and displayed on the screen, simultaneously overlaying additional information or virtual data onto the real scene.
- Static video display: The surgical scene is displayed on a fixed screen without the ability to move or follow the surgeon.

In open surgery, using see-through technology or direct projections, surgeons do not have to take their eyes off the patient to look at different displays when they need to gather and interpret medical information. Moreover, the possibility

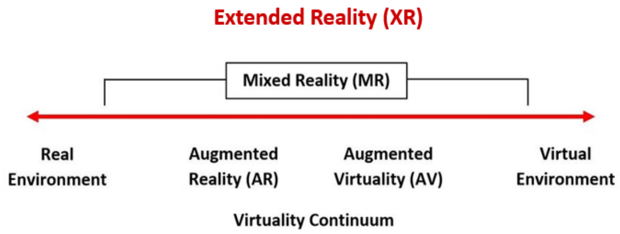

Fig. 2. AR and MR in Virtuality Continuum

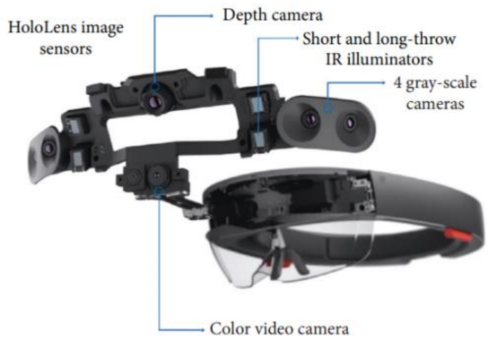

Fig. 3. Hololens hardware and details [6]

to use mixed reality features allows the surgeons to anchor virtual objects to the real world and interact with them. [6]

B. *Augmented Reality applications using Hololens 2*

One of the most popular mixed reality headsets is the Microsoft HoloLens 2, released by Microsoft Corporation in November 2019. They consist of a pair of mixed reality smart glasses, shown in Fig. 3, able to describe an environment in which real and virtual elements appear to coexist. This allows the protagonist to interact with the surrounding environment using holograms whilst engaging their senses throughout.

Microsoft HoloLens is equipped with four gray-scale tracking cameras, a depth sensor, and an ambient light sensor to sense its environment and capture the gestures of the user. It also incorporates four microphones, built-in speakers to deliver sound, and a high-definition photo/video camera. The hardware also includes an inertial measurement unit (IMU) to detect head movements in three-dimensional space and a central processing unit (CPU). It is used in a variety of medical applications such as surgical navigation, human-computer interaction and AR-BCI systems integration, gait analysis, and rehabilitation, medical education and training (for example during the COVID-19 pandemic). [7]

Considering the technical characteristics, the Microsoft HoloLens 2 is the best head-mounted display headset on the market. It is characterized by improved comfort compared to alternatives: it is lighter to wear and presents a more balanced center of gravity. The 3D viewing is much more realistic, and the hand-tracking precision is higher. Moreover, no other commercially available system has undergone the rigorous validation process of the HoloLens 2. [8]

Despite the ergonomic and design improvements, there are still weaknesses. Some of the main limitations are the limited field of view, the half-obscured lenses that block 60% of the total light, the limited battery life (approximately 2-3 hours), the ergonomics and mechanical design as well as his weight, cybersickness after long exposure in an immersive environment. [6], [8]

III. MATERIALS AND METHODS

The workflow of the proposed Augmented Reality application is shown in Fig. 4: preoperative images are used to extract the patient-specific 3D organ models. These models are loaded in the Augmented Reality scene, which is deployed on the HoloLens 2. In the intra-operative phase, the surgeons can wear the see-through display, exploiting the AR support when needed.

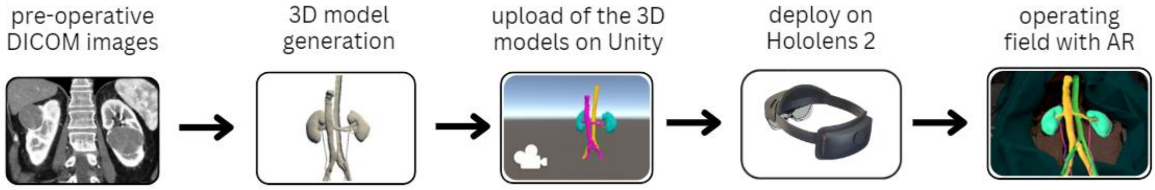

Fig. 4. Workflow of our system

The Augmented Reality scene was implemented as a virtual environment in Unity, optimally configured for the development of a Mixed Reality application. A virtual camera was added to the virtual environment to represent the surgeon's head point of view. The organs were extrapolated from DICOM pre-operative images and transformed into 3D models of the organs; then these 3D models were independently imported into the scene as .OBJ objects. For this application, the selected organs were: aorta, vena cava, left kidney, right kidney, left ureter, and right ureter. The left kidney is also characterized by a child object, the kidney tumor. Being a child object means that it inherits the transformation properties (position, rotation, and scale) of its parent object so that the tumor remains integral to the respective kidney.

Effective visualization of the Augmented Reality (AR) view in the operative field is crucial, and the texture and colors assigned to 3D models play a significant role in achieving this. To prevent color similarity with the operative field, we have enhanced the representativity of these models by employing bright and distinct colors.

We imported the open-source Mixed Reality Toolkit (MRTK), from the Microsoft Mixed Reality Feature Tool (v1.0.2209.0), which provides a set of components and features based on mixed reality functions. MRTK libraries allow us to implement essential functionalities such as hand tracking, eye gaze tracking, and detection of specific gestures.

The functionalities of our scene, represented in Fig. 5, are the following:

- **Organs Manual Manipulation:** The organs can be manually manipulated according to their 6 degrees of freedom (3 translations, 3 rotations). The manipulation is achieved through hand gesture recognition, managed by the HoloLens hand tracking library: we can both interact with them using a far pointer for long-distance or directly grabbing them thanks to the 'NearInteractionGrabbable' component.

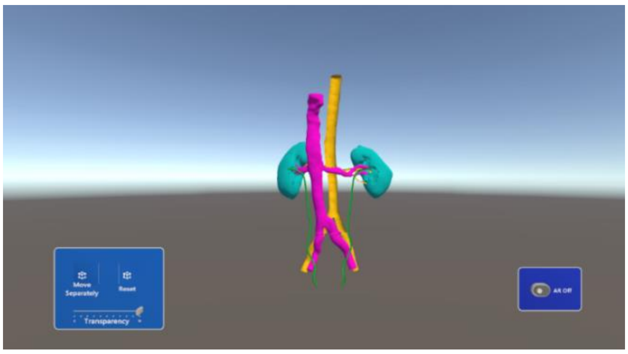

Fig. 5 Unity scene seen in 'Game modality', once Button AR is pressed

Since the models were obtained from pre-operative images, we kept the original position of such organs and the scaling of the models fixed.

- **AR view Menu:** 'AR Backplate', consists of a child element 'Toggleswitch AR' that activates the augmented reality when pressed, the Menu, the organs, and other visualization options superimpose to the real-world scenario.

- **Functionality Menu:** 'Menu Backplate', consists of 3 child elements: a 'Slider' thanks to which we can change the transparency of the virtual organs; a 'Reset' Button that allows us to reset the initial positions of the organs once moved in the scene, in order to facilitate their correct positioning; a 'Move Separately' Button, useful to move separately the single organs once pressed.

- **Streaming on external screen:** To make the surgical team of the operating room aware of the AR visualization, we can also replicate in real-time what the surgeon sees on other monitors: HoloLens 2 comes equipped with an integrated Mixed Reality Capture (MRC) feature that allows to broadcast the view to an external audience in real-time. We can configure the capture setting according to our requirements, for example including resolution, frame rate, and other important parameters for optimal streaming quality.

All buttons are aligned with the global AR reference system and equipped with a 'Follow' component to synchronize their movement with the head, minimizing potential eye discomfort resulting from constant subtle shifts.

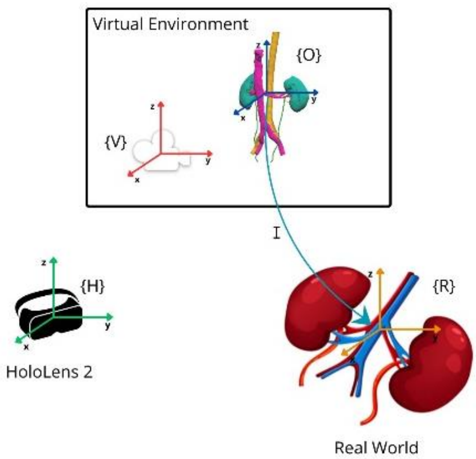

Fig. 6. Reference systems present in our set-up, with the homogeneous transformation matrix from {O} to {R} equal to identity matrix.

Alignment between reality and virtuality is a fundamental concept of AR, which is realized via registration. To maintain registration and synchronization of the viewpoint in the user's perspective, the position and orientation of the AR viewing camera with respect to the environment need to be tracked. This is obtained thanks to 'inside-out' (or intrinsic) tracking: the sensors are integrated within the AR device itself and thus the device can self-locate within the environment. [7]

Knowing all the reference systems (RFs) involved in our experimental setup is essential to guarantee that the virtual objects are correctly aligned with the real environment during the AR experience. Studying our experimental setup, we recognized different reference systems, represented in Fig. 6. In order to pass from one RF to another, we may calculate the homogeneous transformation matrices using registration methods. However, our objective was not to mathematically calculate these transformations, but to superimpose the RF of the virtual organs {O} to the RF of the real organs {R}, so that the homogeneous matrix describing that conversion becomes equal to the Identity matrix (I).

IV. EVALUATION

The proposed experimental setup for open surgery is illustrated in Fig. 7: the surgeon stands in front of a phantom of open abdomen surgery where the organs of interest are replicated. He wears some see-through holographic lenses (Microsoft HoloLens 2), able to display additional virtual elements to the real user's field of view.

The AR scene was built with Visual Studio 2022, through which we deployed the project on HoloLens 2. We first needed to establish a connection between HoloLens glasses and the host machine, which can be done through a wireless connection (Wi-Fi) or USB connection.

With HoloLens on, the surgeon is able to capture photos or videos of his current view. This can be accomplished in several ways: by pressing volume up and down buttons at the same time (similar to taking a screenshot), issuing voice commands like "Take a picture" or selecting the camera icon on the Start menu. The same options apply for recording videos: the surgeon can either press and hold volume buttons simultaneously until a 3-second countdown begins or say "Start/stop recording" or selecting the video icon on the Start menu. Exploiting these frames, a quantitative evaluation of the application is provided.

A. *Quantitative evaluation*

The quantitative evaluation of the AR application was based on the measurement of the registration accuracy between the real and virtual organ models. The metrics used is the Intersection over Union (IoU), which can be used to evaluate the overlap between real and virtual models in mixed reality scenarios, such as the HoloLens ones. Intersection over Union (IoU) is a commonly used metric in evaluating object detection or segmentation algorithms in computer vision applications. To calculate this metric, we need to obtain the 2D masks of virtual and real models, i.e. regions of interest (that represents the surface we are evaluating), calculate the intersection and union areas, and compute IoU as the ratio of the intersection area and the union area:

$$IoU = \frac{A \cap B}{A \cup B} \quad (1)$$

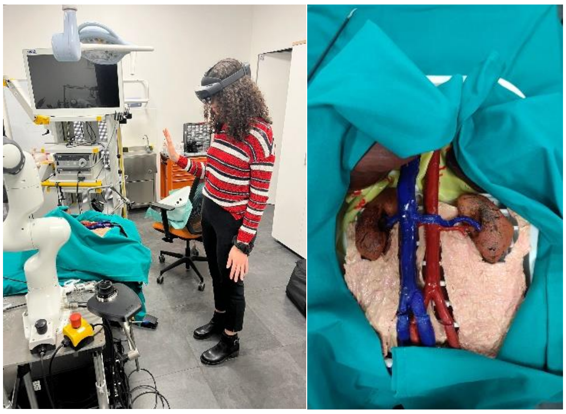

Fig. 7. Experimental setup (left), with the surgeon wearing Hololens 2 standing in front of the operating field (right)

Where A is the area of the real kidney and B is the area of the virtual kidney. An IoU closer to 1 indicates a high overlap between the real and virtual models, while an IoU closer to 0 indicates a low overlap.

To this end, we initially captured frames of the surgical field image, concentrating on the left kidney model. The task consisted of performing 12 times the manual registration of the kidney by superimposing the virtual one on the real one. Additionally, we recorded the task completion time invested for the 12 registrations using a chronometer, enabling the assessment of both mean and standard deviation for the registration process. To account for any intra-subject variations, the 12 measurements were conducted by two distinct subjects (6 by the first subject, 6 by the second subject).

To calculate the IoU, we wrote a Matlab code that calculates the binary mask of the real and virtual models. The masks are obtained by manually segmenting the RBG image saved using the 'roipoly' function, as illustrated in Fig. 8: roipoly returns the mask as a binary image, setting pixels inside the ROI to 1 and pixels outside to 0. We summed the number of pixels to obtain the area of interest and calculated the intersection union areas between the corresponding real and virtual models. Finally, we calculated the mean and standard deviation of our measures for statistical analysis.

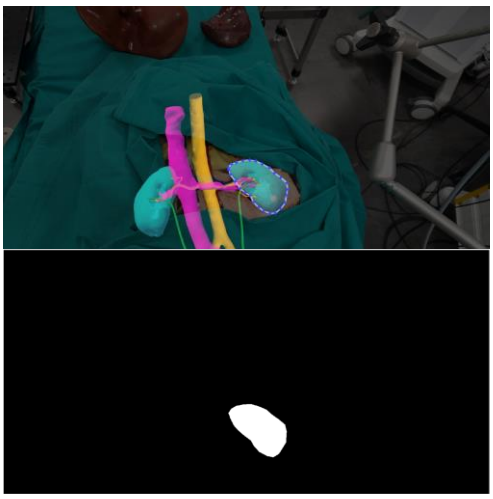

Fig. 8. Manual segmentation of virtual kidney using 'roipoly' function (top) and its corresponding binary mask (bottom)

V. RESULTS

The quantitative evaluation yielded optimal results, showing an IoU of 0.83563 ± 0.075499. This suggests that we were able to register the AR virtual models onto real models with an accuracy higher than 80%.

During the registration process, we observed a challenge related to the optimal recognition of our pinch gestures by HoloLens 2. This affected the accurate superimposition of virtual organs onto real ones. There were instances where, upon releasing the pinch, the augmented reality organs did not precisely align with our intended location.

The mean time required for manual registration was 35.745 ± 37.9693 seconds, demonstrating a brief duration compared to the whole surgery. Furthermore, we observed a correlation between the speed of registration and its accuracy, with faster registration often providing worse results (for example, between the 12 frames, frame 7 provided the worst IoU result, equal to 0.6722, in only 14.7 seconds). This suggests that the accuracy of the measure is strictly related to surgeon's precision and involvement in performing the task. Additionally, we observed that the IoU value increases with practice, as shown in Fig. 9 related to subject 2. This trend underscores the positive impact of experience on the accuracy and effectiveness of the registration process, and strongly defends the thesis advocating the essential need to provide comprehensive training for surgeons on these new technologies before incorporating them into practice.

In summary, this preliminary analysis provides valuable insights into system feasibility and accuracy. To enhance the robustness of our findings, further evaluations involving a larger pool of surgeons and more registration processes are recommended. This extension will help identify which factors affect the measures and guide improvements for optimal performance.

VI. CONCLUSIONS

In conclusion, our paper proposes the integration of Augmented Reality (AR) with the Microsoft HoloLens 2 in pediatric nephron-sparing surgery. The aim of our study is to facilitate intraoperative guidance by enhancing the precision of tumor localization during surgery, ultimately minimizing long-term complications in children. The quantitative evaluation demonstrates a high accuracy of the AR application, but further improvements can be made to improve overall surgical outcomes.

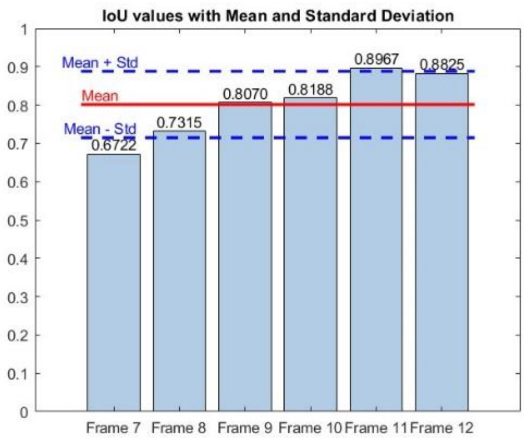

Fig. 9. Bar diagram showing IoU values measured on subject 2

The study acknowledges the challenges encountered during the manual registration process, as indicated by our findings. This overlapping procedure can be optimized through practice and further development could include incorporating voice commands to refine the superimposition of the two organ models. Additionally, in order to reach a fully automated model overlapping, we can pursue a computer vision strategy based on the identification of landmarks to link the virtual model (i.e. using the whole kidney as a marker). In the future, deep learning algorithms may also be integrated for training the software to recognize the kidney's features and texture, reaching a more precise and stable automatic tracking during the entire procedure. [9]

Moreover, we can also develop a ROS (Robot Operating System) package to enhance communication and real-time data streaming between HoloLens 2 and other applications/devices present in the operating scene (like exoscopes or surgical tools).. By using HoloLens as a video capture device, we can create an application that captures the video stream from its integrated camera. The integration of ROS support into this application, achieved through libraries available for Unity or C#, involves configuring a ROS node responsible for managing and transmitting video data from the camera (publishing node) and a node configured to receive the video data and transmit it to other devices or applications (subscriber node). The implementation of ROS within our application may be useful to add functionalities that require high computational costs, like integrating deep learning algorithms for automatic registration.

Our Unity scene has the advantage of being very versatile since it can also be deployed to other see-through devices, that may have different features than the proposed HoloLens 2.

To improve our AR scenario, we can change the illumination and the type of material for a better rendering of the organs. In particular, since surgeons will wear glasses for more hours it is very important to reduce the possible drawbacks and discomfort related to long exposure to cyber-sickness due to HoloLens. Moreover, in the future, we can simulate tissue deformations occurring during the dynamic phases of the intervention by including nonlinear parametric deformations to the 3D model meshes. In particular, for the described renal cancer surgery, bend and stretch deformers need to be selected. [10]

REFERENCES

[1] Breslow N, Olshan A, Beckwith JB, Green DM. Epidemiology of Wilms tumor. Med Pediatr Oncol. 1993;21(3):172-81. doi: 10.1002/mpo.2950210305. PMID: 7680412.

[2] Wellens LM, Meulstee J, van de Ven CP, Terwisscha van Scheltinga CEJ, Littooij AS, van den Heuvel-Eibrink MM, Fiocco M, Rios AC, Maal T, Wijnen MHWA. Comparison of 3-Dimensional and Augmented Reality Kidney Models With Conventional Imaging Data in the Preoperative Assessment of Children With Wilms Tumors. JAMA Netw Open. 2019 Apr 5;2(4):e192633. doi: 10.1001/jamanetworkopen.2019.2633. PMID: 31002326; PMCID: PMC6481457.

[3] Porpiglia F, Checcucci E, Amparore D, Piramide F, Volpi G, Granato S, Verri P, Manfredi M, Bellin A, Piazzolla P, Autorino R, Morra I, Fiori C, Mottrie A. Three-dimensional Augmented Reality Robot-assisted Partial Nephrectomy in Case of Complex Tumours (PADUA ≥10): A New Intraoperative Tool Overcoming the Ultrasound Guidance. Eur Urol. 2020 Aug;78(2):229-238. doi: 10.1016/j.eururo.2019.11.024. Epub 2019 Dec 30. PMID: 31898992.

[4] Matthijs Fitski, Jene W. Meulstee, Annemieke S. Littooij, Cornelis P. van de Ven, Alida F. W. van der Steeg, Marc H. W. A. Wijnen, "MRI-Based 3-Dimensional Visualization Workflow for the Preoperative

Planning of Nephron-Sparing Surgery in Wilms' Tumor Surgery: A Pilot Study", Journal of Healthcare Engineering, vol. 2020, Article ID 8899049, 6 pages, 2020. https://doi.org/10.1155/2020/8899049

[5] Sylvain Bernhardt, Stéphane A. Nicolau, Luc Soler, Christophe Doignon, The status of augmented reality in laparoscopic surgery as of 2016, Medical Image Analysis, Volume 37, 2017, Pages 66-90, ISSN 1361-8415, https://doi.org/10.1016/j.media.2017.01.007.

[6] Galati R, Simone M, Barile G, De Luca R, Cartanese C, Grassi G. Experimental Setup Employed in the Operating Room Based on Virtual and Mixed Reality: Analysis of Pros and Cons in Open Abdomen Surgery. J Healthc Eng. 2020 Aug 6;2020:8851964. doi: 10.1155/2020/8851964. PMID: 32832048; PMCID: PMC7428968.

[7] Christina Gsaxner, Jianning Li, Antonio Pepe, Yuan Jin, Jens Kleesiek, Dieter Schmalstieg, Jan Egger, The HoloLens in medicine: A systematic review and taxonomy, Medical Image Analysis, Volume 85, 2023, 102757, ISSN 1361-8415, https://doi.org/10.1016/j.media.2023.102757.

[8] Palumbo A. Microsoft HoloLens 2 in Medical and Healthcare Context: State of the Art and Future Prospects. Sensors (Basel). 2022 Oct 11;22(20):7709. doi: 10.3390/s22207709. PMID: 36298059; PMCID: PMC9611914.

[9] Kunz, Christian, Maierhofer, Pascal, Gyenes, Balázs, Franke, Nikolai, Younis, Rayan, Müller-Stich, Beat-Peter, Wagner, Martin and Mathis-Ullrich, Franziska. "Augmented Reality-based Robot Control for Laparoscopic Surgery" Current Directions in Biomedical Engineering, vol. 8, no. 1, 2022, pp. 54-57. https://doi.org/10.1515/cdbme-2022-00

[10] Amparore D, Checcucci E, Piazzolla P, Piramide F, De Cillis S, Piana A, Verri P, Manfredi M, Fiori C, Vezzetti E, Porpiglia F. Indocyanine Green Drives Computer Vision Based 3D Augmented Reality Robot Assisted Partial Nephrectomy: The Beginning of "Automatic" Overlapping Era. Urology. 2022 Jun;164:e312-e316. doi: 10.1016/j.urology.2021.10.053. Epub 2022 Jan 19. PMID: 35063460.