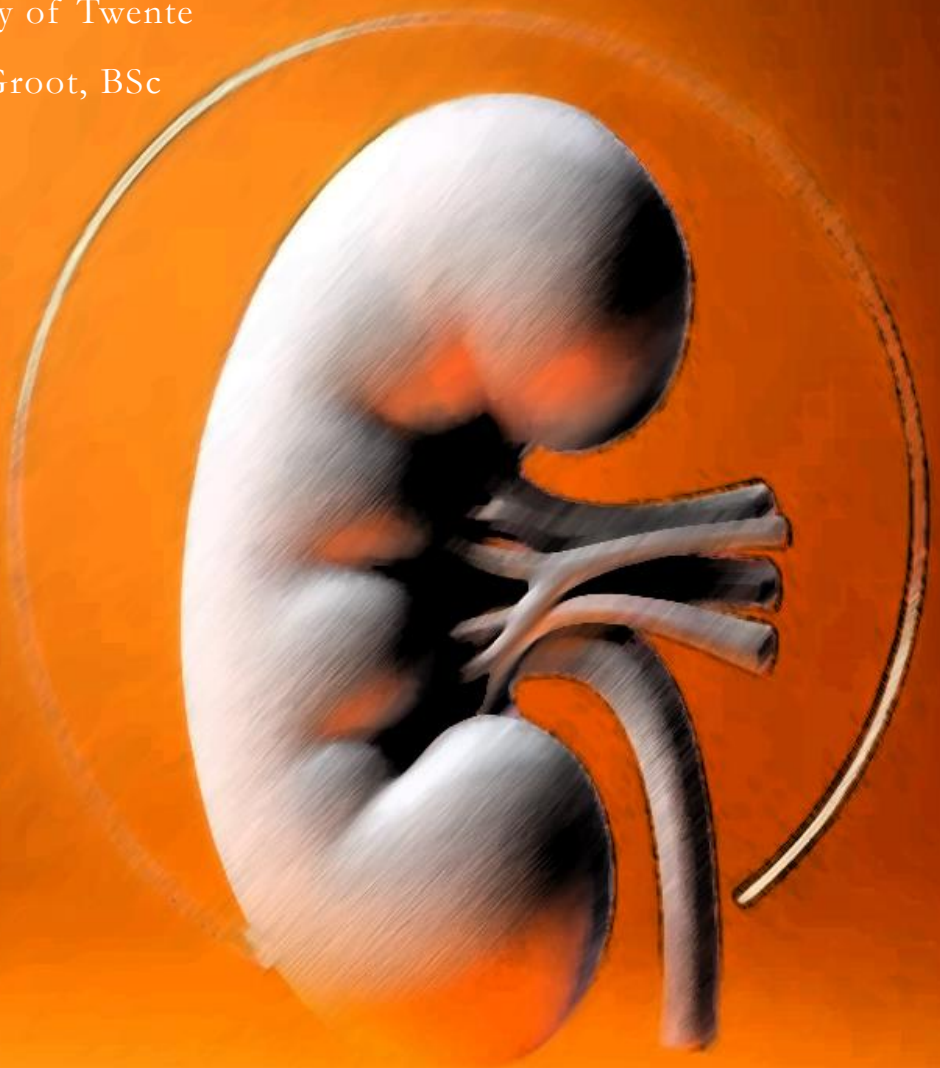
Feasibility of Infrared Reflective Marker Tracking for Nephron-Sparing Surgery Using Augmented Reality

Master Thesis
Technical Medicine
University of Twente

N.T. de Groot, BSc



princess máxima center
pediatric oncology

UNIVERSITY OF TWENTE.

Master thesis

Feasibility of Infrared Reflective Marker Tracking for Nephron-Sparing Surgery Using Augmented Reality

N.T. De Groot, BSc

A thesis submitted in partial fulfilment of the requirements of the degree of Master of Science

In

Technical Medicine
Track: Medical Imaging and Intervention
University of Twente

1 September 2024

UNIVERSITY OF TWENTE.

Graduation Committee

Chairman	Prof. dr. J.J. Fütterer Department of Medical Imaging Radboud University Medical Center Nijmegen, the Netherlands Robotics and Mechatronics University of Twente Enschede, the Netherlands
Medical Supervisor	Dr. A.F.W. van der Steeg Department of Pediatric Surgery Princess Máxima Center for Pediatric Oncology Utrecht, the Netherlands
Technical Supervisor	Dr. F.J. Siepel Robotics and Mechatronics University of Twente Enschede, the Netherlands
Process Supervisor	E. Walter, MSc Faculty of Science and Technology University of Twente Enschede, the Netherlands
Daily Supervisor	Dr. M. Fitski Department of Pediatric Surgery Princess Máxima Center for Pediatric Oncology Utrecht, the Netherlands
External member	Dr. ir. L. Alic Magnetic Detection and Imaging University of Twente Enschede, the Netherlands

Acknowledgments

In 2017 ben ik begonnen aan de opleiding Technische Geneeskunde waar, met het verdedigen van deze master thesis genaamd *'Feasibility of Infrared Reflective Marker Tracking for Nephron-Sparing Surgery Using Augmented Reality'*, (mogelijk) een einde is gekomen aan mijn studenten tijdperk. Een bewogen tijdperk met onwijs veel mooie herinneringen en ervaringen. Na het kiezen van mijn master-richting 'Medical Imaging and Intervention' is mijn interesse uitgegaan naar chirurgie en heb ik op diverse plekken, en in diverse domeinen zoals MRI, AI en chirurgische navigatie, kennis gemaakt met het vakgebied van de Technische Geneeskundige. Gedurende mijn M2 heb ik kennis leren maken met het Prinses Maxima Centrum, waarbij ik een systeem genaamd Augmented Reality Live UltraSound (ARLUS) heb ontworpen om intra-operatieve ultrasound beelden direct op het onderliggend weefsel af te beelden. Vanwege de eigen verantwoordelijkheid, vrijheid en zelfsturing, en tevens mijn wens om de kinderoncologische chirurgische behandeling te verbeteren, is mijn keuze gevallen om ook af te studeren in het Prinses Maxima Centrum. Een uniek centrum, met een unieke manier van werken waarbij iedere individu hard werkt om de kinderoncologische zorg te verbeteren. Afgelopen jaar heb ik mij ingezet om met behulp van Augmented Reality betere chirurgische navigatie te ontwerpen voor niersparende chirurgie bij kinderen met Wilms Tumoren. Een uitdagende opdracht waarbij veel verschillende wegen zijn bewandeld om bij het uiteindelijke resultaat te komen. Hierbij hoorde ook het leiden van een klinisch wetenschappelijk onderzoek naar het intra-operatief gebruik van AR tijdens totale nefrectomieën. Bij de inclusiegesprekken lieten de reacties van de kinderen en de ouders op de patiënt specifieke 3D modellen zien waarvoor je op de achtergrond bezig bent om de zorg te verbeteren. Door allerlei facetten mee te nemen in mijn klinische ontwikkeling wordt ook duidelijk waar ik, als Technische Geneeskunde, van meerwaarde ben en kan zijn. Ik ben onwijs blij met alle ontwikkelingen die ik heb doorgemaakt en de kansen die heb gekregen.

Dit was ook niet mogelijk zonder de hulp van alle betrokken die ik heb mogen ontvangen. Graag zou ik Lideke van der Steeg en Marc Wijnen onwijs willen bedanken voor alle ondersteuning en mogelijkheden die jullie mij hebben gegeven tijdens mijn M3. Zonder jullie, zou ik nooit zo mijn M3 hebben kunnen indelen en beleven als ik het nu heb ervaren. Daarnaast wil ik Matthijs Fitski bedanken voor de dagelijkse begeleiding. Je was altijd bereid om met mij mee te denken om problemen op te lossen. Tevens hebben wel veelvoudig samengewerkt om de klinische AR studie in goede banen te leiden, waarbij ik de samenwerking, zowel op persoonlijk als professioneel vlak, als zeer prettig en gezellig heb ervaren. Verder wil ik Françoise Siepel en Vincent Groenhuis bedanken voor de technische begeleiding waarmee ik mijn thesis naar een hoger niveau heb kunnen tillen. Omtrent de persoonlijke ontwikkeling die ik heb doorgemaakt wil ik Elyse Walter en mijn intervisiegroep bedanken. Door jullie scherpte en feedback gedurende de afgelopen twee jaar heb ik kritisch naar mijn eigen ontwikkeling kunnen kijken en waar nodig bij kunnen sturen. Dit heeft mij mede gevormd tot de professional die ik nu ben. Tevens wil ik alle kinderchirurgen, fellow's, kinderoncologen, verpleegkundig specialisten, anois'en en physician assistant's bedanken voor alle tijd die jullie in mijn hebben geïnvesteerd om mijn klinische ontwikkeling mogelijk te maken. Door al het extra geduld en uitleg heb ik onwijs veel kennis opgedaan over de kinderoncologische zorg. Daarnaast wil ik de onderzoeksgroep Wijnen bedanken alle geleverde input. Jasper, Kevin, Lorentz, Dominique, Myrthe, Rixt, Ceder en alle voorbijgekomen studenten, onwijs bedankt voor alle gezellige koffietjes, uitjes, borrels en de daar bijhorende goede gesprekken.

Als laatste wil ik alle personen bedanken die persoonlijke dicht bij mij staan. Marcel, Jolanda en Barbara, bedankt voor alle steun, gezelligheid en wijze lessen die ik van jullie heb mee mogen krijgen. Verder wil ik mijn vrienden en Manon bedanken omdat ik altijd bij jullie terecht kan en jullie mij veel afleiding hebben geboden in een drukke periode.

Dan rest mij niks anders dan vooruit te kijken naar de toekomst, en nogmaals aan iedereen, bedankt!

Abstract

Purpose: During Nephron-Sparing Surgery (NSS) for Wilms Tumors (WT), current clinical challenges involve preventing positive surgical margins and preserving as much healthy renal parenchyma as possible. Intraoperative Ultrasound (iUS) has multiple limitations, such as difficulties correlating two-dimensional (2D) images with 3D anatomy. Augmented Reality (AR) presents a potential solution by providing a 3D holographic model overlay to enhance intraoperative visualization and decision-making. This study tries to identify the clinical limitations of AR and propose a new AR system for surgical navigation during NSS for WTs.

Methods: AR is introduced in the clinical setting to support surgical navigation during NSS. The preliminary evaluation of AR holographic visualization during total nephrectomies revealed specific challenges related to registration accuracy due to tumor size, Quick Response(QR)-code, and continuous tracking (Chapter 3). An application was developed to enable continuous tracking of the kidney using infrared (IR) reflective markers (Chapter 4). Static and dynamic accuracy assessments of different IR markers using the DINO-CLINIC systems identified optimal configurations for clinical use (Chapter 5). The clinical usability and accuracy of the DINO-CLINIC system were further evaluated through phantom experiments, focusing on IR-marker placement surgical feasibility (Chapter 6).

Results: Successful AR registration was achieved in two out of six patients, with significant limitations related to QR-code registration and positional changes of the kidney affecting the holographic overlay (Chapter 3). The 6.4 mm and 12.7mm IR markers met technical accuracy requirements (<3.0 mm), with overall accuracies of -0.73 mm and -0.12, respectively, although an overestimation of distance was noted (Chapter 5). The DINO-CLINIC system exhibited a target registration error (TRE) of 11.82 mm and a positional localization error (PLE) of 8.74 mm, failing to meet clinical accuracy requirements (<5.0 mm). Surgeons rated the system's usability positively but identified significant tracking limitations (Chapter 6).

Conclusion: Inside-out IR-tracking with the HoloLens 2 shows potential for enhancing intraoperative guidance during NSS by providing continuous 3D visualization of the tumor and resection boundaries. However, further research and development must address limitations regarding holographic stability and accuracy before further clinical implementation.

Table of Contents

Lists of Figures

Lists of Tables

Lists of Abbreviations

2D	Two-Dimensional
3D	Three-Dimensional
ACT	Actinomycin D
AR	Augmented Reality
CT	Computed Tomography
EM	Electromagnetic
FPS	Frames per Second
HL2	HoloLens 2
IR	Infrared
iUS	Intraoperative Ultrasound
MRI	Magnetic Resonance Imaging
MX	Mixed Reality
NSS	Nephron-Sparing Surgery
PLE	Point Localization Error
QR	Quick Response
SIOP	International Society of Paediatric Oncology
SUS	System Usability Scale
SVD	Singular Value Decomposition
TDR	Tracking Detection Rate
TN	Total Nephrectomy
ToF	Time of Flight
TRE	Target Registration Error
US	UltraSound
VCR	Vincristine
VR	Virtual Reality
WT	Wilms Tumor

1 Introduction

1.1 Clinical background

This section highlights the clinical relevance of this thesis.

1.1.1 Epidemiology

Renal tumors comprise approximately 7% of all pediatric malignant tumors [1]. Among these, 95% are Wilms Tumors(WT), also known as nephroblastoma, which is often diagnosed before ten years of age [2][3]. The annual incidence of WT is around 7 per 1 million children under 15 years of age [3]. Solitary WTs are most prevalent, followed by multifocal loci (10%) and bilateral (5-7%) tumors. Due to advances in treatment strategies, the overall survival rate improved significantly from 20% in the 1960s to 90% today [4]. WT arises due to abnormal renal development, resulting in metanephric blastema without normal tubular and glomerular differentiation[5]. The underlying pathogenesis of WT is thought to involve the persistence of metanephric cells known as nephrogenic rests, which are present in most cases of WT. In addition, genetic factors play a role, such as alterations in WT1, CTNNB1, and WTX genes, which are found in 1/3 of all WTs [6]. Secondly, WTs are associated with multiple syndromes, including the WAGR-, Drash, and Beckwith-Wiedemann syndrome.

1.1.2 Diagnostics and staging

Children are often asymptomatic when presented in the clinic but have a palpable mass in their flanks that does not move during respiration[5]. Only 20% of the children have symptoms at the time of diagnosis, such as pain, hematuria, fever, hypertension, and urinary tract infections. The initial diagnostic imaging tool is ultrasound (US), which is used to confirm the presence of renal mass [7]. After the US, further diagnostic imaging is performed using contrast-enhanced Computed Tomography (CT) and, preferably, Magnetic Resonance Imaging (MRI) [8]. CT is beneficial for examining the lungs for metastases. MRI is primarily used to assess the kidneys, including the tumors, arteries, and veins. In 20% of WT, distant metastases are present at initial evaluation, where lung parenchyma is the most common location [8]. MRI is beneficial when nephron-sparing surgery (NSS) is considered due to the accurate imaging of the kidney and assessment of vascular involvement. Secondly, it can distinguish WT and nephroblastomatosis [5].

Based on the CT or MRI scan, the tumor is staged, following the UMBRELLA-protocol, as localized (stage I-III), metastatic (stage IV), or bilateral (stage V); see Appendix J. Current 5-year overall survival for I, II, III, and IV stage are respectively 97%, 95%, 92% and 82% [9]. The five-year survival rate for Stage V WT is 85%, based on the SIOP-9 trial [10] [11].

1.1.3 Treatment

Based on the UMBRELLA SIOP–RTSG 2016 guidelines, the current treatment of the WT depends on the staging of the tumor [12] [13]. Typically, the treatment protocol consists of preoperative chemotherapy, followed by surgical resection and additional postoperative chemotherapy [14]. International Society of Paediatric Oncology (SIOP) recommends preoperative chemotherapy because of a decreased risk of intraoperative rupture after chemotherapy, which would upstage the tumor and necessitate irradiation. Preoperative chemotherapy for localized WT (Stage I-III) consists of 4 weeks of vincristine (VCR) and actinomycin D (ACT) after reassessment of the tumor after week four. Metastatic disease (stage IV) receives

a combined treatment of VCR, ACT, and doxorubicin for six weeks, followed by local and metastatic disease assessment. Bilateral WT's receive six weeks of VCR and ACT. After treatment assessment, if the tumor remains inoperable, chemotherapy is extended up to a maximum of 12 weeks.

After neoadjuvant chemotherapy, surgical resection can be performed, which can be divided into either a total nephrectomy (TN) or NSS [12] [15]. TN involves complete removal of the kidney and should include complete removal of the tumor, including the adipose capsule and Gerota's fascia. On the other hand, NSS aims to preserve as much healthy renal parenchyma as possible while ensuring complete excision of malignant tissue [16]. Radical nephrectomy is the standard procedure for unilateral WTs. However, NSS is recommended in several cases for unilateral WT and especially for bilateral WT's. For unilateral WT, NSS is acceptable when specific criteria are met, see Figure 1 [17].

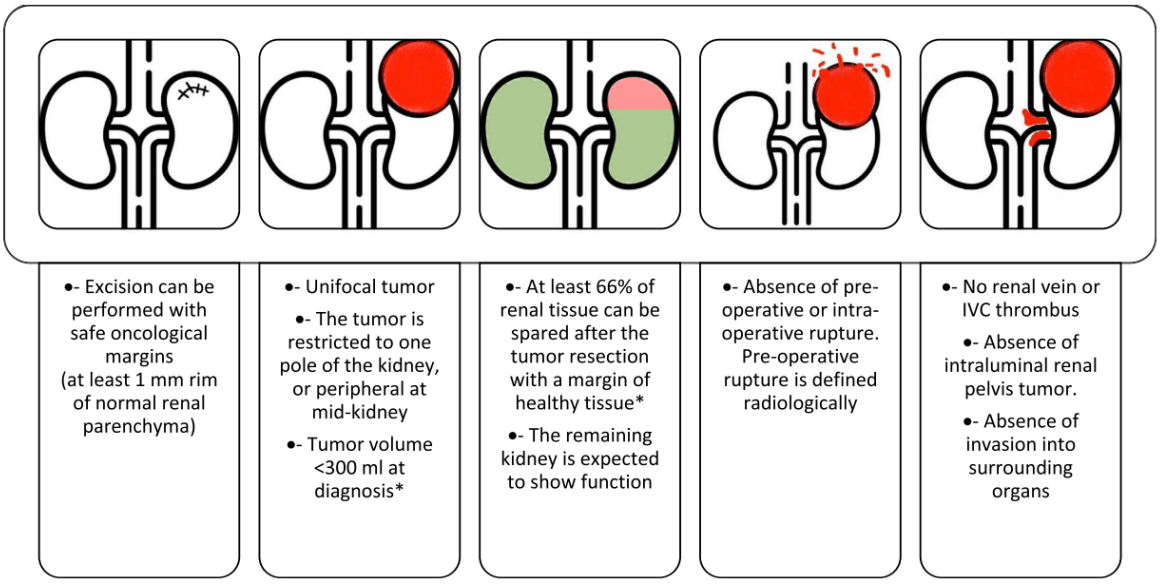


Figure 1: Illustration of criteria NSS for unilateral non-syndromic WT based on the UMBRELLA-protocol (2016). Adapted from Taghavi et al. [17].

1.1.4 Nephron sparing surgery

According to UMBRELLA Protocol SIOP 2016, NSS for WT is recommended under specific conditions (see Figure 1)[12]. WTs must be restricted to one pole of the kidney or located peripherally at the mid-kidney and have a volume of less than 300 ml at the time of diagnosis. Moreover, preoperative rupture of the tumor should be absent. Additionally, there should be no intraluminal tumor presence within the renal pelvis, no invasion of surrounding organs, no thrombus in the renal vein or vena cave, and no multifocal tumors. Most importantly, surgeons must be able to perform an excision with an oncological safe margin, and the remaining kidney should remain functional, with a total of 66% of healthy kidney parenchyma.

When a patient is eligible for NSS, the patient is placed in the supine position, and the choice between a bi-subcostal laparotomy incision or a retroperitoneal approach is made to optimize exposure of the affected kidneys and associated tumors [16]. Before the excision of the kidney parenchyma, complete mobilization of the kidneys is essential for a successful operation. Mobilization enables appreciation of the tumor for visual examination, palpation, and the use of intraoperative ultrasound (iUS). iUS is particularly valuable in identifying resection edges for small lesions that may not be visible from the capsular surface of the kidney. However, iUS is still limited since it requires specific skills and experience [18]. One limitation is the need to correlate two-dimensional (2D) US images to a patient's physical three-dimensional (3D) anatomy. Secondly, skilled hand-screen coordination is required, as the surgeon must simultaneously look directly at

the US screen while adjusting the position and orientation of the US probe. These limitations can impact the accuracy of the resection border and negatively impact intraoperative decision-making. Subsequently, the resection path is carefully marked to prevent disorientation and deviation during excision-induced bleeding, minimizing the risk of incomplete resection.

Two surgical techniques are feasible, namely partial nephrectomy and enucleation. Partial nephrectomy involves the resection of a tumor with a rim of normal parenchyma, where a safe rim is defined as 5 mm or more. Enucleation is a resection without a rim of normal renal parenchyma. However, it is mainly possible if the tumor has an intact capsule, which often results in a margin-negative resection. After resection, lymph node sampling is regarded as mandatory.

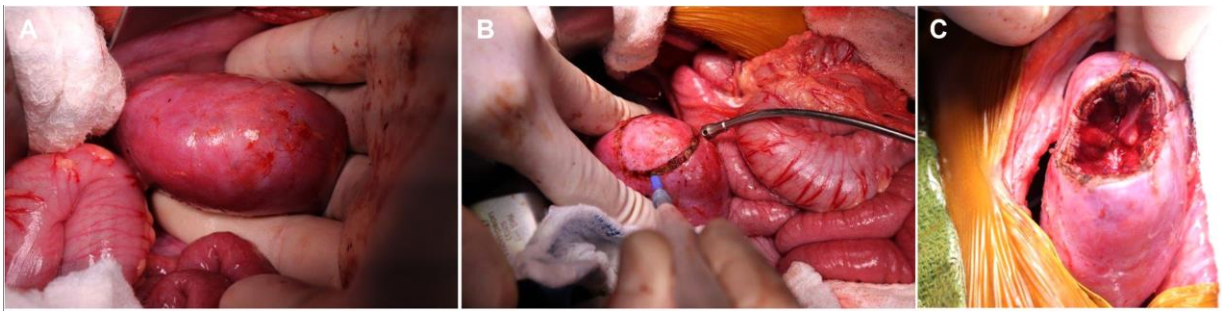


Figure 2: Overview of NSS A: The surgeon mobilizes the kidney and tumor to access the renal hilum and compress the vessels between their fingers. B: The surgeon marks the renal capsule around the tumor location to outline the expected dissection path, often using iUS. C: After marginally resecting the lesion, a kidney defect occurred. Adapted from Murphy et al. [16].

After NSS, the remaining kidney requires short and long-term follow-up within the UMBRELLA-2016 protocol. Doppler US should be performed two days post-operation to evaluate the spared renal tissue and scintigraphy six months later. Secondly, parameters such as creatinine clearance, blood pressure, and renal failure markers should be assessed every three months for the first two years, lasting three years. After three years, a patient is followed up annually.

1.1.5 Outcomes TN vs NSS

As mentioned earlier, long-term survival rates for WT are high, around 90%. The prolonged survival of WT patients gives new insight into the treatment-related morbidity and mortality after more than 25 years. Preserving renal parenchyma with NSS has multiple long-term benefits but worse short-term outcomes than TN. TN shows a lower complication rate (13%) in comparison with NSS (36.4%), of which the most common complication was prolonged urinary leakage [17] [19].

Oncological long-term outcomes showed similarities for unilateral WT with TN and NS. As described by Wilde et al., overall survival (OS) is similar for NSS 100% and TN 94.4%, and event-free survival is NSS 94.8% and TN 86.5% [20]. It is essential to mention that the NSS group comprised 65% stage I tumors in this cohort compared to 48% in the TN group. Secondly, the tumor volume at diagnosis is significantly larger in the TN group. However, positive surgical margins for NSS are often higher; reports show positive surgical margins ranging from 9% to 22.5% [21]. Prevention of positive surgical margins during NSS is critical because flank radiotherapy also affects the residual renal parenchyma [15]. As described in the systematic review of Vanden Berg et al., most studies reported similar long-term outcomes between NSS and TN for WT [22].

Further long-term outcomes are related to kidney function. A recent study following adults after childhood TN showed that after 30 years, 40% had kidney dysfunction, 22% had hypertension, and 23% had albuminuria [23]. For children undergoing TN and NSS, kidney function after a second decade was higher

for NSS than TN, with a kidney function of 110 ml/min/1.73 m^2 and 91.2 ml/min/1.73 m^2 [23]. Another study demonstrates that the patients undergoing TN have significantly higher odds of abnormal kidney function (OR 4.29, 95% Cl 1.02, 18.00) and lower estimated glomerular filtration rate (eGFR) (mean difference -8.99, 95% CI -16.4, -1.58) in comparison with patients undergoing NSS [24]. Moreover, for bilateral WT, patients with RN with contralateral kidney have, in comparison with bilateral NSS, higher changes of developing abnormal kidney function and hypertension (OR 5.81, 95% CI 1.31, 25.68).

Multiple indications suggest that NSS has potentially better long-term outcomes than TN for specific staged WT; however, this is currently under active debate within pediatric surgery and requires more long-term outcomes.

1.1.6 Clinical challenges

As described earlier, current clinical challenges are preventing positive surgical margins during NSS for WT and maintaining as much healthy renal parenchyma as possible. As visualized in Figure 2, WT are often not visible from the outside [25]. Therefore, iUS is used to find the resection edge. However, iUS has considerable limitations, such as it requires specific skills and experience, especially in a small pediatric abdomen, see Figure 3. Moreover, the correlation of 2D iUS images to the 3D physical anatomy of the patient remains challenging.

Augmented reality (AR) presents a promising solution to these limitations. By enabling a 3D holographic model overlay, AR can visualize in-depth information that overcomes the limitations of iUS [26], see Figure 3. AR has the potential to enhance the accuracy of intraoperative assessments and positively influence decision-making processes during surgery. Therefore, NSS could be performed more accurately, preventing positive margins and remaining as much healthy renal parenchyma as possible.

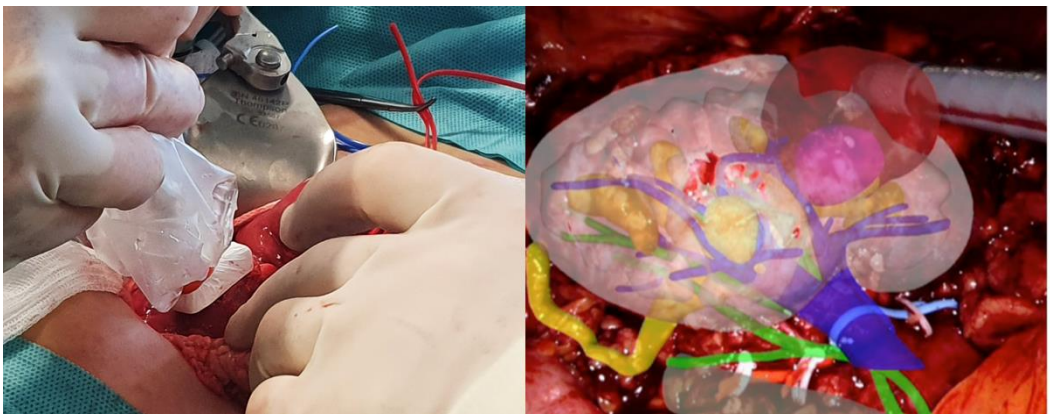


Figure 3: iUS and AR Left: Utilization of iUS during NSS. Right: Holographic overlay of the kidney during laparoscopic guided NSS, WT, arteries, and veins during TN. Adapted from Piana et al. [27]

1.2 Technical Background

1.2.1 Mixed Reality

Mixed Reality (MR) and Virtual Reality (VR) blend physical and digital environments. VR typically involves fully immersive digital experiences with minimal interaction with the real world. In contrast, MR overlays digital elements in the real world, allowing for interaction between physical and digital elements. AR is defined as a view of the real world enhanced by digital content [28]. It enhances the user experience of the real world by enabling users to interact with the real and virtual worlds simultaneously. AR has been introduced to health care and is widely studied in various clinical domains, such as orthopedic surgery, neurosurgery, laparoscopic surgery, and abdominal surgery [28]. Here, AR applications are developed for multiple users, such as physicians, students, and patients, and for multiple purposes, such as education, training, patient monitoring, displaying medical information, and even surgical navigation [29][30][31][32].

The use of AR for surgical navigation can potentially overcome the limitations of conventional surgical navigation systems such as iUS [33]. Conventional surgical navigation systems often display images outside of the surgical area. Therefore, surgeons must shift their attention to surgical navigation displays and the surgical area itself. Secondly, a surgeon must visualize a 2D image in a 3D real-life situation, which is mentally demanding and increases the chance of misinterpretation of 2D images. A study assessing the usefulness of 3D kidney models and AR before NSS showed that it improved anatomical understanding among surgeons [34].

The development of AR applications for surgical guidance involves 3D reconstruction, holographic visualization, registration, and tracking techniques [35], [36]. Holographic visualization is achieved through a head-mounted display (HMD), providing in situ visualization of 3D anatomical models, medical images, or surgical planning trajectories.

1.2.2 HoloLens 2

The HoloLens 2 (HL2) is the second generation of AR devices developed by Microsoft, and it was released in 2019 [37]. The HL2 is a head-mounted display that enables the visualization of holograms in the real world. The device is currently priced around $3500.

The HL2 has a Qualcomm Snapdragon 850 Compute Platform with a 4-GB LPDDR4x system DRAM memory card [38]. In addition, the HL2 is equipped with multiple sensors, such as four visible light cameras, two infrared (IR) cameras, an accelerometer, a gyroscope, a magnetometer, an RGB camera of 8-MP with a 1080p30 video resolution, and a 1 MP Time of Flight (ToF) depth sensor, see Figure 4. The ToF sensor has two depth modes: short throw (range 0-0.8m) and long throw (range 0.8m – 3.5m). Grayscale images representing IR light reflectivity can be used for both modes. These sensors enable the possibility of tracking and monitoring the environment around the user.

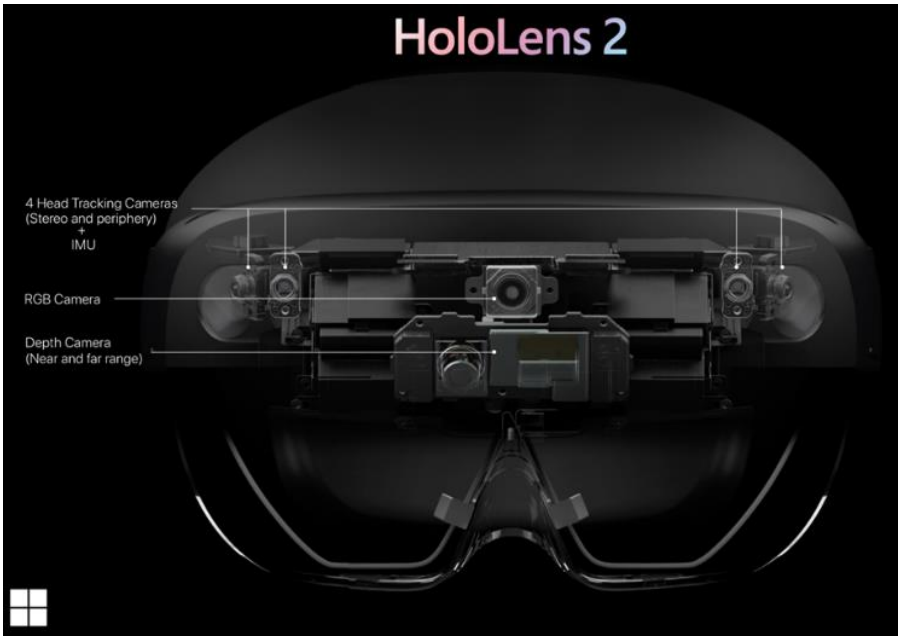


Figure 4: Integrated HL2 sensors consist of 4 Head Tracking Cameras, RGB camera and Depth Camera [37].

1.2.3 Registration

Registration for AR-guided surgery is necessary to transform the virtual 3D model to a real object in situ [39]. The HMD and virtual 3D model, as the real object in situ, have their own coordinate frames. The coordinate systems need to be aligned to align the virtual 3D model with the help of the HMD on the real object in situ. The following coordinates systems are defined; see Figure 5:

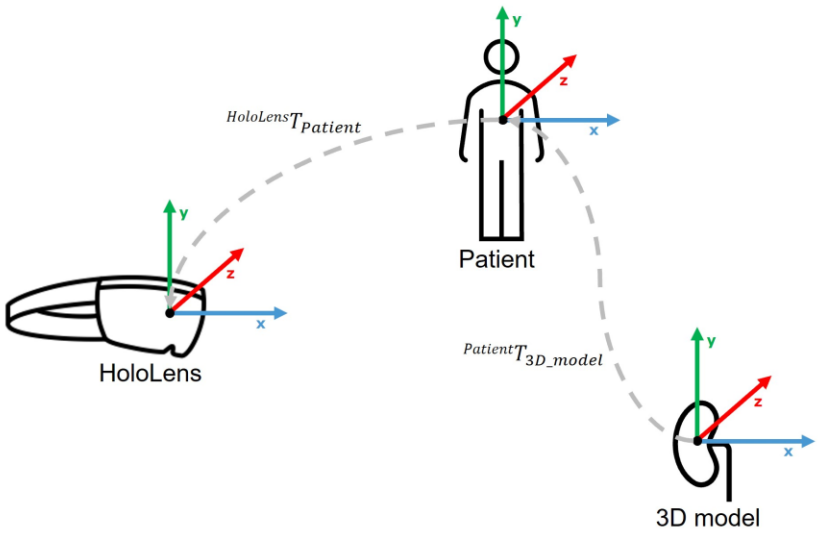


Figure 5: Coordinate systems from 3D model to HL2.

To align the multiple coordinate systems, such as coordinate systems A and B, spatial transformations like translation and rotation are necessary, represented by matrixes. A translation of coordinate system B to A can be defined as:

$$ {}^{A}t_{B} = \begin{bmatrix} {}^{A}t_{x} \\ {}^{A}t_{y} \\ {}^{A}t_{x} \end{bmatrix} \tag{1} $$

Considering two coordinate frames that are rotated from each other, both coordinate frames can be rotated, which are defined in vectors from B to A,

$$ x_A = r_{xx}x_B + r_{xy}y_B + r_{xz}z_B $$

$$ y_A = r_{yx}x_B + r_{yy}y_B + r_{yz}z_B $$

$$ z_A = r_{zx}x_B + r_{zy}y_B + r_{zz}z_B $$

Written in matrix-vector notation, this results in:

$$ \begin{bmatrix} x_A \\ y_A \\ z_A \end{bmatrix} = \begin{bmatrix} r_{xx} & r_{xy} & r_{xz} \\ r_{yx} & r_{yy} & r_{yz} \\ r_{zx} & r_{zy} & r_{zz} \end{bmatrix} \begin{bmatrix} x_B \\ y_B \\ z_B \end{bmatrix} = {}^{A}R_{B} \begin{bmatrix} x_B \\ y_B \\ z_B \end{bmatrix} \tag{2} $$

With both the translation and rotation matrices known, the transformation matrix can be defined as:

$$ {}^{A}p = {}^{A}T_{B}\ {}^{B}p $$

For two coordinate frames, a specific point within coordinate system B can be described in coordinate system A as:

$$ {}^{A}T_{B} = \begin{bmatrix} {}^{A}R_{B} & {}^{A}t_{B} \\ 0 \quad 0 \quad 0 & 1 \end{bmatrix} \tag{3} $$

Regarding the situation in Figure 5, three coordinate systems are defined, and the transformation matrix to define a specific point of the 3D model inside the HoloLens can be mathematically described as:

$$ {}^{HoloLens}p = {}^{HoloLens}T_{Patient}\ {}^{Patient}T_{3D_{model}}\ {}^{3D_{model}}p \tag{4} $$

Registration is necessary to find the relationship and the transformation matrix between the 3D model and the patient. Therefore, multiple registration techniques are available, such as manual, point-based, surface, and marker-based registration [36].

Manual registration in AR-guided surgery involves directly aligning the virtual model with the real object, eliminating the need for additional tracking calculations. During this process, the user views a hologram through the headset and uses software controls to adjust its position until it aligns spatially with the corresponding physical structure. This alignment process includes moving, rotating, and scaling the preoperative data to match accurately with the intraoperative data.

Point-based registration is based on virtual landmarks placed on the 3D model and the virtual points placed on the patient intraoperatively. The procedure involves preoperatively determining specific points on anatomical landmarks in the 3D model. Intraoperatively, similar points are placed on the patient in situ, which results in two similar point clouds. Then, a point-matching algorithm must be applied, which searches for the optimal configuration between two sets of points. The generalized Procrustes algorithm provides

the solution to this optimization problem. It employs isomorphic scaling, translation, and rotation in its rigorous examination of shapes to achieve the optimal alignment between two or more landmarked shapes [39]. Therefore, it takes advantage of the known one-to-one correspondence of landmarks. The precision of the point-based registration method depends on the accuracy of the manual extraction of fiducial points and the spatial distribution of the fiducial points [35]. Spijkerboer et al. presented an AR technique allowing fast registration using anatomical landmarks to localize chest wall tumors [40].

Surface-based registration allows for the alignment of the patient's anatomy surface with the virtual surface received from medical imaging. This process typically involves matching a surface mesh or a point cloud from the preoperative model to the intraoperative geometric shape. An Iterative Closest Point (ICP) algorithm is commonly used to align both point sets. Multiple studies have evaluated the use of surface-based registration in clinical settings. Zhang et al. combined stereo matching and 3D reconstruction with automatic segmentation and surface stitching to generate a dense intraoperative point cloud of the renal surface during laparoscopy partial nephrectomy [41]. This is achieved by initially an ICP algorithm followed by the coherent point drift algorithm. Further developed and implemented deep learning algorithms, such as those by Weber et al., which proposed using deep learning-based point cloud registration for AR using the HL2 and showed promising results [42]. In the case of pediatric surgery, van der Woude et al. explored the use of surface-based registration for the localization of chest wall tumors [43].

Marker-based registration uses a fiducial/marker which is fixated to an organ for intraoperative tracking and registration. Marker-based registration enables real-time tracking and registration of a moveable object since the markers move together with the organ or patient [35]. However, the accuracy of the registration depends on the visibility, lighting conditions and size of marker [44].

1.2.4 Tracking

Tracking is necessary to correctly project the hologram onto a position in the real world after registration. Multiple tracking techniques are used, such as self-localization, tracking with integrated sensors (inside-out), or external tracking (outside-in). A combination of self-localization and inside-out or outside-in tracking is called hybrid tracking [35]. Self-localization is a technology implemented within the HoloLens, where a camera monitors the changes within the surrounding environment so that the relative pose of the hologram and within the environment remains unchanged. However, the movement compensation of the HL2 is relatively low, leading to drift problems [35]. Outside-in tracking uses external devices to track markers within the operating room and to translate the position to the HL2. Often, optical trackers or EM-trackers are used for this application [45]. Inside-out tracking makes use of the integrated sensors within HoloLens. The HL2 has multiple camera sensors for environmental tracking, which are accessible in developer mode [38]. Currently, the HL2 has four Visible Light Environment Tracking Cameras and a depth camera, which operates in two modes: a high frequency near depth mode (45 frames per second (FPS)) and a low frequency (1-5 FPS) far-depth mode. These sensors can be used to track objects in different ways, such as Quick Response (QR) tracking, point-cloud tracking, or IR reflective sphere markers [46][47].

2

Research design and objectives

This chapter highlights the objectives of this thesis. Several objectives are set to be achieved by the end of the thesis. Multiple research questions have been formulated to correlate with these objectives. Finally, the thesis design is outlined.

2.1 Objectives

The primary objective of this thesis is to improve the AR system for use in surgical navigation during NSS for WT's. This involves the clinical validation of the conventional AR system through an in vivo study. Addressing the limitations of the current AR system, a new AR system has been developed that is capable of real-time kidney tracking to compensate for kidney movement during NSS. This system is designed to continuously update the hologram's position based on fixated IR reflective markers tracking the kidney's positional changes, ensuring usability throughout the surgery. Additionally, this study aims to validate the tracking accuracy and clinical feasibility of the developed IR-tracking AR application through a phantom experiment.

2.2 Research questions

The main research questions addressed in this thesis are:

- What is the clinical feasibility and accuracy of a point-based registration AR system in total nephrectomies?
- What is the technical performance of the IR-marker tracking application with the HL2?
- What is the intraoperative accuracy of manual IR-marker placement based on preoperative imaging with and without the HL2?
- What is the surgical feasibility and usability of the developed IR-marker inside-out tracking system using the HL2 for NSS?

2.3 Research design

Chapter 3 presents the preliminary results of intraoperative holographic visualization during total nephrectomies. The system Usability Scale (SUS) and intraoperative accuracy are determined.

Chapter 4 provides an in-depth analysis of DINO and the working principles of IR tracking within HL2 and details the necessary adaptations to the algorithm for intraoperative use. Additionally, a user interface is integrated into the DINO AR application to enhance the surgical workflow.

Chapter 5 is dedicated to validating the accuracy of the DINO-CLINIC system's IR tracking capabilities. Various geometric markers with different parameters are evaluated and compared to establish statistical measures of translational error. This chapter also assesses the update frequency of the IR-tracking algorithm and conducts a dynamic movement analysis to determine translational accuracy during marker movement.

Chapter 6 focuses on the intraoperative validation of IR tracking by simulating NSS using phantoms. This chapter measures the intraoperative accuracy of IR-marker placement and evaluates surgical feasibility using a Likert scale and a System Usability Scale (SUS) assessment.

3

Preliminary results: holographic surgical navigation for WT's

3.1 Introduction

WT's are renal tumors present in 95% of pediatric renal tumors [1]. WT's consists of a bilateral, multifocal, or unilateral disease. For patients with WT, surgical treatment consists of TN or NSS [3]. For bilateral and sometimes unilateral WT, NSS can be the preferred option over TN. Specific criteria are set to perform NSS in cases of unilateral WT; see Chapter 1.1.4. During NSS, the tumor is completely removed with a surgical margin. Therefore, NSS possibly leads to better long-term outcomes than TN, such as preventing renal failure [17], [24]. However, NSS is a complex procedure as the surgeon should identify the tumor, which is often not visible from the outside, and resect only the tumor with a small margin to maintain a healthy renal parenchyma [16]. iUS is used to visualize the resection edge during surgery. However, there are multiple clinical limitations with the use of iUS, as described in Chapter 1.1.6, such as iUS requires specific skills and experience, and it involves correlating 2D US images to the 3D physical 3D anatomy of the patient, which is challenging. AR could help to visualize the resection border in 3D [48] and improve understanding of critical anatomical structures during NSS.

Currently, AR is employed in NSS, primarily in laparoscopic procedures. Studies have shown that AR is beneficial for soft tissue navigation during laparoscopy. For instance, Teber et al. utilized custom-designed navigation, including a mobile C-arm used for cone-beam imaging, to evaluate the feasibility and reproducibility of an inside-out tracking application in a phantom model with an intraparenchymal tumor[49]. This application was also evaluated in clinical practice during laparoscopic partial nephrectomy. Therefore, point-based registrations with rigidly attached markers or anatomical landmarks for initial alignments in surface-based registrations are used.

In pediatric oncological surgery, resections of WT's are performed during open surgery. Translating AR to open surgery is essential as it can potentially improve the anatomical understanding and correlation of the tumor with the surrounding tissue. Multiple reviews show the extensive use of AR in open surgery for adolescents [50], [51]. A study presents a method using a 3D patient-specific guide attached with a visual pattern for automatic registration [52]. The patient-specific guide with the image target is fixated on the position of the bone. The system was validated on a phantom for extraosseous Ewing sarcoma. The placement of the patient-specific guided had a 1.87 mm accuracy, and the point localization error (PLE) had a mean of 2.9 mm. Van Doormaal et al. developed a neuronavigation system with an AR device by using a point-based registration and a pointer with an image target for neurosurgery [53]. They assessed the developed AR application in the operating room (OR) on patients before surgery and in a phantom experiment, which showed a fiducial registration error of 7.2 mm and 4.4 mm, respectively.

Currently, AR is being introduced to the field of pediatric oncological surgery. Van der Woude et al. demonstrated the potential of AR in pediatric chest wall tumor surgeries, where accurate surgical planning and tumor localization are crucial for achieving radical resections while sparing healthy tissue [43]. They used preoperative CT scans and intraoperative tracking with a QR code for a five-point registration method. Secondly, they used an additional QR code to track the patient's movement. The system improved the anatomical understanding and intraoperative tumor localization. Eyck et al developed a HL2 application for

pediatric NSS [54]. They used the integrated sensors within the HL2 to track a QR pointer and overlay 3D holograms onto the kidney based on a point registration algorithm. During a phantom experiment, this application achieved a matching accuracy of 3.16 ± 1.82 mm within the clinically acceptable error margin of <5 mm, demonstrating its potential for clinical use.

This study aims to assess and validate the HL2 technique developed by Eyck et al. inside the operating room on patients eligible for TN. Patients eligible for TN are selected instead of patients eligible for NSS, as the HoloLens does not influence clinical decision-making and patient outcomes. Therefore, no clinical risks are taken with the introduction of this technique. With this study, the initial steps are taken to introduce AR into the operating room by allowing the surgeon to get comfortable and gain experience with the HL2. Secondly, the holographic overlay can be assessed, and the initial limitations of the point registration method can be identified.

3.2 Materials and Methods

3.2.1 Preoperative planning

The study results were collected between January and July 2024 at the Princess Maxima Centrum (Utrecht, the Netherlands). Six patients, aged between 6 months and 18 years, with radiologically proven renal tumors and planned for TN were included. All patients have given informed consent to participate in this study. Preoperative MRI was performed, which consisted of T1-Thrive post-Gadolinium, T2 Fat Suppression Multivane-XD High Resolution 4 mm, and Balanced Triggered Angiography Non-Contrast Enhanced acquired with a 1.5 Tesla system (Ingenia, Philips Medical Systems, Best, The Netherlands). A 3D model was created for each patient based on the standard preoperative MRI data using Mimics Innovation Suite 25.0 (Materialise, Leuven, Belgium). The 3D models included the kidney, tumor, and renal arteries and veins. The preparation of the AR application consisted of defining multiple anatomical landmarks in Unity 2019.4.22f1 (Unity Software Inc., San Franciso, United States) in consultation with the surgeon. The landmarks included the cranial side of the kidney, the lateral side, the caudal side, the anterior side, the hilum, and the bifurcation of the renal arteries or veins. Several evaluation parameters were collected during preoperative planning: the time required to segment and prepare anatomical landmarks in Unity and the number of landmarks set preoperatively.

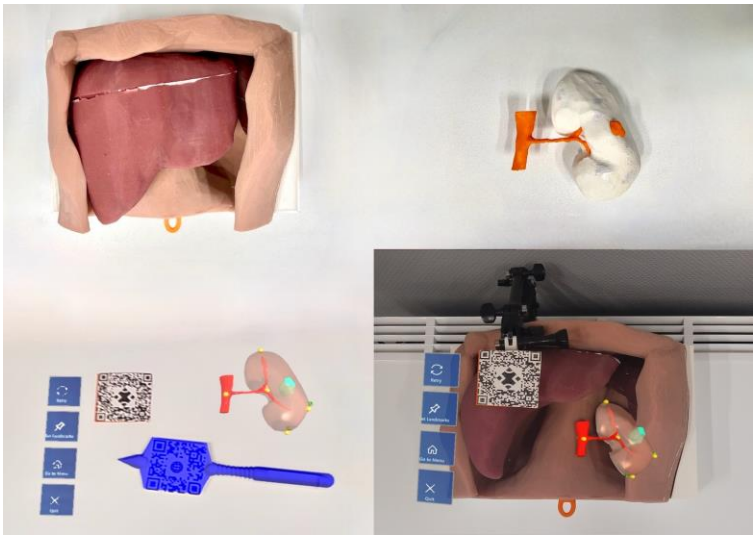


Figure 6: Overview of holographic application in a phantom study. Top-left: pediatric abdominal phantom. Top-right: 3D-printed phantom kidney. Bottom-left: Holographic application including reference QR-code and QR-pointer Bottom-right: utilization of the 5-point registration system in a phantom study. Adapted by Eyck et al. [54].

3.2.2 Intra-operative utilization AR

During the TN, the standard surgical protocol was followed to expose the kidney with the tumor from the surrounding tissue. After complete mobilization of the kidney and tumor, and when the clinical situation was deemed stable and safe, the HL2 (Microsoft Corporation Inc., Redmond, United States) device was positioned on the head of the surgeon by the technical physician, and the registration time measurement started. The surgeon selected the correct patient, and the holographic 3D model of the patient is shown, including the position of the preoperative landmarks. Manipulation of the holographic 3D model, by translation and rotation, enables an overall overview of all positions of the landmarks. Depending on the intraoperative situation, the surgeon decides which landmarks, at least four, are feasible to use for the point-registration. A reference QR code was positioned near the operation area, and using a sterilized QR-pointer, the surgeon pinpointed the digitally defined anatomical landmarks on the kidney (see Figure 7). Based on a Procrustes algorithm, the virtual points were transformed to the pinpointed physical points, allowing the hologram to be placed inside the patient, overlapping the kidney and tumor.

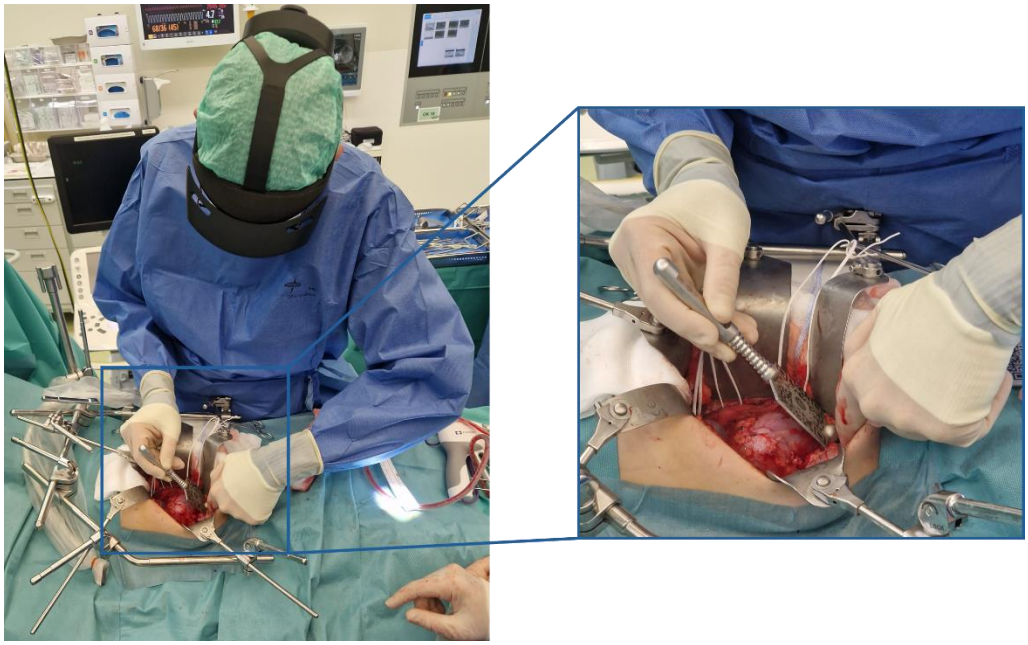


Figure 7: Point localization performed by the surgeon using the QR-pointer and HL2 during total nephrectomy.

The surgeon carefully observes the resulting holographic visualization and measures the distance between the hologram's and the kidney's boundaries with a ruler. Subsequently, the surgeon continues with the conventional surgical procedure and removes the kidney and tumor from the patient.

3.2.3 Qualitative assessment

Qualitative assessment aimed to determine the intraoperative system usability of the HL2 system. Therefore, after surgery, the surgeon filled in SUS and self-developed questionnaires. The SUS is determined, and the surgeons can Totally Disagree (1) and Totally Agree (5) with ten statements. The SUS is calculated using a formula [55]. In addition, to assess the observer's accuracy and comfort, a self-developed questionnaire is filled in which the surgeon scores questions based on a Likert scale of Totally Disagree (1) and Totally Agree (10).

3.3 Results

This section describes the preliminary results of the study. A total of six patients were included in the ongoing study, with a median tumor volume of 152 ml. The patient characteristics are visualized in Table 1.

Table 1: Patient characteristics

Case nr.	Age	Gender	Tumor Volume (ml)	Neoadjuvant chemo (weeks)	Tumor location	Successful registration?
Case 1	7	Male	33.14	VAD (8)	Right	Yes
Case 2	9	Female	827.92	VA (4)	Right	No
Case 3	3	Male	150.53	VAD (6)	Right	Yes
Case 4	4	Male	155.28	VA (4)	Right	No
Case 5	4	Female	89.28	VAD (6)	Left	No
Case 6	2	Male	348.62	VA (4)	Right	No

Successful registration was achieved in two patients. In the following three cases, the registrations were unsuccessful due to the size of the tumor, which hindered the surgeon from pinpointing the anatomical landmarks. For case 2, the tumor was across the midline with a tumor volume of 827ml. For case four, the tumor with a volume of 150ml protruded anteriorly, disabling the surgeon to move the kidney and pinpoint the anatomical landmarks. In case six, the tumor was too large, 348 ml, and it was impossible to find the anatomical landmarks. In addition, case five failed due to the attempt for reregistration to enhance the holographic accuracy. During the reregistration, the holographic overlay rotated with a 90-degree offset. Therefore, the registration and reregistration were defined as unsuccessful.

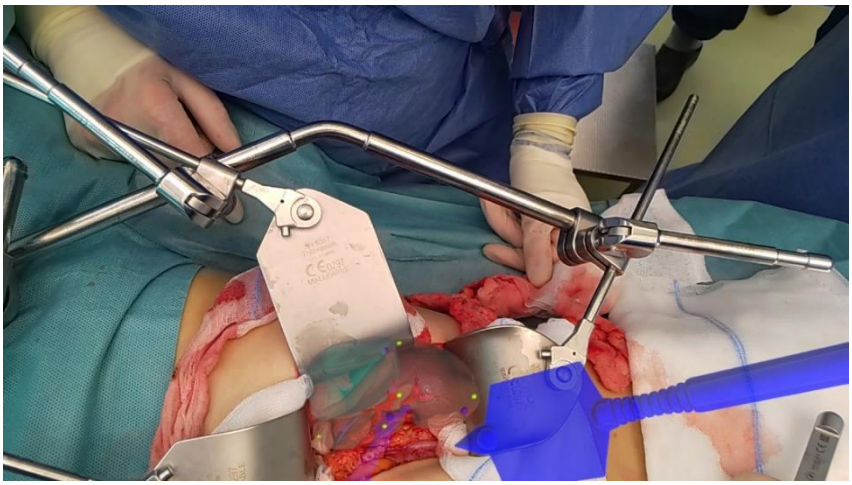


Figure 8: Intraoperative holographic visualization after registration of the kidney (brown), tumor (turquoise), arteries (red) and veins (blue). Yellow and purple dots are registration points. Large blue pointer is an overlay of the QR-pointer.

The preoperative and intraoperative measurements for the two successful cases are visualized in Table 2. The preparation time for cases 1 and 3 was 01:22 and 01:15. The surgeon used five intraoperative landmarks in both cases for registration, which lasted 11.5 minutes and 15 minutes in total. A 5 mm and 20 mm accuracy is reported for cases one and three. Consequently, the surgeons rated the subjective accuracy a six and a four.

Table 2: Preoperative, intraoperative and accuracy results

Pre-operative	**Case 1**	**Case 3**
Pre-operative number of landmarks	6	7
Preparation time (hh:mm)	01:22	01:15
Intra-operative		
Surgeon	4	2
Used landmarks for registration	5	5
Registration time (minutes)	11.5	15
Accuracy		
Accuracy (mm)	5	20
Subjective accuracy (1-10)	6	4

The surgeons filled in the SUS questionnaire, visualized in Table 3. In case 1, the surgeon finds the system acceptable based on the SUS categories. For case 3, the surgeon finds the system marginally user-friendly.

Table 3: Results of SUS for case 1 and case 3

No.	**Question**	**Case 1**	**Case 3**
1.1	I think that I would like to use this system frequently.	2	4
1.2	I found the system unnecessarily complex.	2	2
1.3	The system was easy to use.	4	2
1.4	I think that I would need the support of a technical person to be able to use this system.	3	2
1.5	I found the various functions in the product were well integrated	5	3
1.6	I thought there was too much inconsistency in this system	1	3
1.7	I would imagine that most people would learn to use this system very quickly.	5	2
1.8	I found the system very cumbersome to use.	2	2
1.9	I felt very confident using the system.	4	4
1.10	I needed to learn a lot of things before I could get going with this system.	2	4
	SUS	**75**	**55**

Based on the questionnaire, specific aspects of the application are evaluated, see Table 4. Overall, there are quite some similarities between both cases and surgeons, such as the interface of the software being straightforward, improved sense of depth, and improved anatomical relationship with the patient. However, there are differences between both cases. Most noticeable is the difference in determining the landmark's position on the kidney and if the hologram remained stable after positioning. Especially in case 3, these questions scored lower than case 1. The surgeon did not answer questions 1.6 and 1.7 in case 3 because the intraparenchymal vessels were not visualized.

Table 4: Results of a self-developed questionnaire for case 1 and case 3

No.	Question	Case 1	Case 3
1.1	Determining the position of the landmarks on the kidney of the patient was straightforward	7	2
1.2	The user interface of the HoloLens software was straightforward to use	8	8
1.3	The hologram remained stable on the kidney after positioning	10	4
1.4	The holographic overlay improves my sense of depth	7	8
1.5	The holographic overlay improves my understanding of the anatomical relationships of the patients	8	8
1.6	The holographic overlay visualized the position of the intraparenchymal vessels reliably	7	NA
1.7	The holographic overlay of the tumor and vessels is useful for NSS	8	NA
1.8	The holographic overlay is worth the additional time for NSS	8	5

3.4 Discussion

This current ongoing study is the first to use intraoperative holographic visualization of renal tumors in open surgery for pediatric patients. The study evaluated a point registration and holographic visualization system during total nephrectomies. For validation, intraoperative accuracy measurements were performed alongside a feasibility assessment based on the SUS and a self-developed questionnaire.

Six patients were included, of which successful registration was achieved in two patients. The qualitative analysis showed multiple limitations, primarily related to using the QR code and the system's instability. The size of the tumors, combined with the QR pointer, often made it challenging to pinpoint anatomical landmarks during total nephrectomies. This issue was particularly evident when comparing the results of case 3, which involved a larger tumor volume, to case 1. Case 3 showed difficulties locating the anatomical landmarks and worse overall accuracy and SUS. Tumor size was frequently the limiting factor for successfully implementing this holographic visualization technique in three cases.

The overall measured accuracy was low, with reported 5 mm and 20 mm accuracies. The subjective accuracy ratings were only 6 and 4, respectively. Consequently, the visualization was not sufficiently accurate for clinical guidance. In addition, surgeons mentioned that it was difficult to localize the position of the registration points intraoperatively, as the kidney does not have specific anatomical landmarks. Secondly, the position of the kidney is not tracked continuously. Therefore, small movements with the abdomen could displace the position of the kidney in correlation to the hologram, influencing the accuracy.

The implementation of holographic registration and visualization disrupted the surgical workflow due to the time required for registration, which was an average of 13.25 minutes. This disruption of the workflow is caused by the multiple steps that are necessary before registration. Selecting the patient, determining the landmarks used, and pinpointing the landmarks took a long time. A faster responsive device could enhance the registration time and, therefore, be better implemented in the existing workflow. The surgeons were concerned about whether this extra time was worth the additional time. To enhance the registration time and improve the clinical workflow, extra personnel were necessary to support the surgeons during the utilization of the AR device.

In the unsuccessful cases, the tumor size and QR code use were prominent limiting factors. There are multiple problems related to the intraoperative use of the QR code, such as size, detection, and tracking. Firstly, the QR-code size of 4.5 x 4.5 cm is often too large to accurately pinpoint the selected anatomical landmarks, especially in patients with a large tumor volume. Registration cannot be performed if the surgeon cannot pinpoint the landmarks. However, the application should be used for NSS, which has tumor volumes at diagnosis below 300ml. Therefore, the QR-pointer size could be a less pronounced problem for NSS. Secondly, the QR code was less detectable in the OR. This was due to the reflection of bright operating lights, the low contrast of the sterilized QR code due to the grey background, and the limited detection angle of the QR pointer. These limitations made it challenging to remain a Line of Sight with the QR code. These limitations could explain the low accuracy and low SUS.

It is challenging to compare our findings with existing literature due to the varying use of validation parameters, clinical applications, and usability questionnaires. However, compared to the study by Amini et al., which explored the use of AR for surgical planning of breast reconstruction post-single mastectomy, their AR system achieved a score of 71.5 on the SUS questionnaire, indicating decent usability [56]. Another study focused on AR-based 3D holographic navigation for brachial plexus tumor surgery. This study revealed that surgeons preferred the AR visualization of anatomical structures over conventional MRI images [57]. Additionally, a study by Thabit et al., more relevant to intraoperative surgery, demonstrated a combination of AR tracking with external electromagnetic tracking (EM) in a phantom experiment for craniectomy [58]. They reported a higher SUS score of 73 and a distance accuracy of 2.4 $\pm$ 1.2 from the annotated sutures to the planning reference. A comparison of these studies suggests that AR systems can achieve higher SUS scores and accuracy.

There are a few limitations in this study. First, the intraoperative accuracy measurements were limited. Accuracy was determined using a ruler to measure the displacement between the hologram and the kidney boundaries. This method is highly observer-dependent. Secondly, this method does not give insight into the rotational or translational accuracy across multiple axes. Therefore, it is hard to quantify errors and identify accuracy issues with the holographic overlay intraoperatively.

Further research should include more quantitative measurements, such as the Target Registration Error (TRE) [59]. In addition, extra validation methods could be used to validate the holographic accuracy for multiple points, as TRE does not provide information about the overall accuracy of the holographic visualization. Fick et al. showed which stage of the development and implementation of the AR system and which metrics are commonly used [59].

Another limitation is that the study's preliminary results only included six patients, which prevents conclusions about the SUS, overall accuracy, and self-developed questionnaire. However, early improvements within the application can already be identified. Further research should give clearer insights into the AR system's clinical feasibility and accuracy. In addition, a selection bias is present within the results of the SUS questionnaire, as only the successful registrations are included in the questionnaire.

Further research should focus on improving the tracking technique of the AR application to make the system faster, user-friendlier, and easier to use within the pediatric abdomen. The QR-pointer could be replaced by using IR reflective markers, which have the potential to be more accurate, increase usability, and lead to continuous tracking of the kidney. Secondly, a new validation method, including TRE, could improve this application's assessment and identify other areas of improvement.

3.5 Conclusion

This study reported the preliminary results of intraoperative holographic visualization during total nephrectomies. Out of six patients, there were two successful registrations. The primary issue for unsuccessful registration was the size of the tumor and the QR-pointer, which hindered the placement of the anatomical landmarks necessary for registration. Secondly, the algorithm is incapable of tracking the kidney, and therefore, a small positional change of the kidney results in an inaccurate holographic overlay. Based on the two successful cases, the intraoperative accuracy was 5 mm and 20 mm with a System Usability Scale (SUS) of 75 and 55. To conclude, the accuracy does not meet the clinical requirements; however, the AR system improved the depth perception and understanding of anatomical relationships in both cases. Further research should focus on improving the QR-marker and intraoperative holographic validation methods.

4
Software design

4.1 Introduction

IR-tracking presents a promising alternative for tracking using the HL2 compared to QR-code tracking. QR-tracking has multiple clinical and technical limitations, as identified in Chapter 3, such as reflectivity issues in the operating room, low tracking rate, size of QR-marker, and restricted detection angles. Especially in the case of NSS, QR-pointers are too large to be used in a restricted space within a pediatric abdomen, especially with large tumors. Secondly, QR codes are too small to be placed on an organ to allow for real-time organ tracking. Inside-out IR-tracking could potentially overcome the limitations of QR-tracking, as the IR-marker can be small and accurately detectable. Secondly, using an inside-out tracking system could theoretically limit Line of Sight problems since the surgeon wears the HL2 on the head, most of the time in Line of Sight with the IR markers. Recent studies have explored the use of inside-out IR-tracking with HL2 to overcome these limitations.

Kunz et al. were the first to describe the use of IR-marker tracking using the HL2 as an inside-out tracking device for neurosurgical interventions [46]. They developed and validated two different systems. The first approach used the short throw reflectivity (STR) frame and short throw depth (STD) frame. The second approach used the left and right environmental cameras of the HL2. However, in this approach, the IR-markers are not distinguishable from the background. Therefore, an additional illuminating IR-light source was fixated on the HL2 to visualize the IR-markers with the environmental sensors better. Accuracy validation is performed using a distance of circa 40-60 cm. For the first approach, utilizing STR and STD, the tracking accuracy is 0.76 mm. For the second approach, the accuracy was 0.69 mm. However, the additional illuminating IR-light source was mentioned to be a limitation of this approach, and therefore, they recommended exploring the utilization of the STR and STD approach.

Another study by Martin-Gomez et al. used a similar approach for k-wire insertion in orthopedic procedures [60]. They provide an algorithm for marker detection and tool tracking. They developed a framework that enables tracking multiple tools without prior knowledge of their geometries. This setup required a local network connection between the HL2 and a workstation. Secondly, they implemented a Kalman Filter to enhance the precision of the developed framework. They extensively validated and compared their framework to conventional optical navigation systems (Northern Digital Inc., Polaris Spectra) and ArUCo markers. The validation showed that a lateral accuracy is measured at 0.09 $\pm$ 0.06 mm and a longitudinal translation accuracy of 0.42 $\pm$ 0.32 mm with a rotational accuracy of 0.80 $\pm$ 0.39 degrees.

An open-source real-time IR-tracking and tool pose estimation algorithm for the HL2 was recently published on GitHub by the Department of Mechanical Engineering at Imperial College London [61]. The published algorithm, called DINO, which stands for Detection for Infrared Navigation with OST AR headsets, enables real-time tracking of IR markers and the visualization of a hologram. Hisham Iqbal et al. demonstrated that they could achieve an average dynamic tracking accuracy of 2.03 mm and 1.12° [62].

Initially, an algorithm was developed during this master thesis following a similar workflow as described by Kunz et al. An outline of the initial steps that succeeded is described in Appendix B. DINO is published as open-source C++ code and a Dynamic Link Library (DLL) on GitHub. This allows the source code to be adjusted to one's standards if necessary. Additionally, the algorithm is implemented in various Unity versions (2019, 2020, 2021), making it easy to implement on the HL2. Therefore, this thesis will further explore the

use of DINO for intraoperative organ tracking and as a replacement for the QR-code. The DINO framework will be described in more detail in the next paragraph.

4.2 Working principle

4.2.1 Input HL2

The basis of DINO lies in utilizing the ToF sensor for IR-tracking through the ResearchMode of the HL2. The short-throw mode of the ToF sensor can visualize IR-reflective spheres within the world. It has two modes, the depth (STD) and the IR reflectivity mode (STR), see Figure 9.

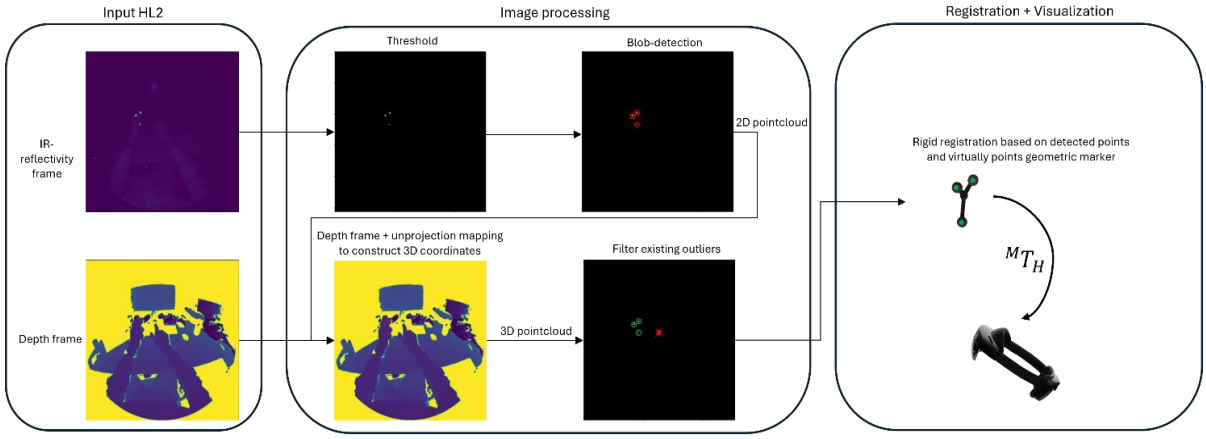


Figure 9: Schematic visualization of the DINO working principle. Input: IR-reflectivity frame and Depth Frame. Image Processing: consists of thresholding, blob detection algorithm, unprojection mapping, and filtering outliers. Registration: Rigid registration based on Singular Value Decompensation (SVD).

4.2.2 Image processing

Within DINO-DLL, a blob detection algorithm is used [61]. The algorithm identifies round bright spots on the Activity Brightness Image (IR-reflectivity frame) and produces a 2D point cloud. For this, OpenCV is integrated, utilizing 'findContours' [63]. First, the IR-reflectivity frame is binarized to enhance the contour detection. The detected contours are filtered based on the pixel area (A) ($5 < A < 16384$) and circularity ($C > 0.7$).

Based on the valid contours, the centroid is calculated, resulting in a 2D (X, Y) point cloud set. Then, the position of the IR markers is correlated to the depth camera (Short Throw Depth Frame). To find the point in 3D in the CameraUnitPlane, the (x,y) coordinates of the detected blobs are transformed to the CameraUnitPlane using a transformation (MapImagePointToCameraUnitPlan (u,v)). Following the blob detection process, the 2D positions of the IR markers within the STR frame are determined. An explanation of the methodology for deriving their 3D positions is provided by Labini et al., which will be briefly described here [64].

Suppose there is a point in a 3D space, namely [X,Y,Z], and the corresponding point in the projection plane is [x,y,1], see Figure 10 [65]. An un-projection map contains values that allow the translation of a 2D projection plane point to a 3D point. The un-projection map contains [u,v] values for each pixel in the projection plane such that [X,Y,Z]=Z*[u,v,1]. To retrieve the 3D coordinates based on the location within the projection plane, the mapping is applied over the depth frame for i-th row and j-th column width, thereby belonging u_ij and v_ij as un-projection map parameters. The depth position is [X_ij,Y_ij,Z_ij]=Z_ij*[u_ij,v_ij,1].

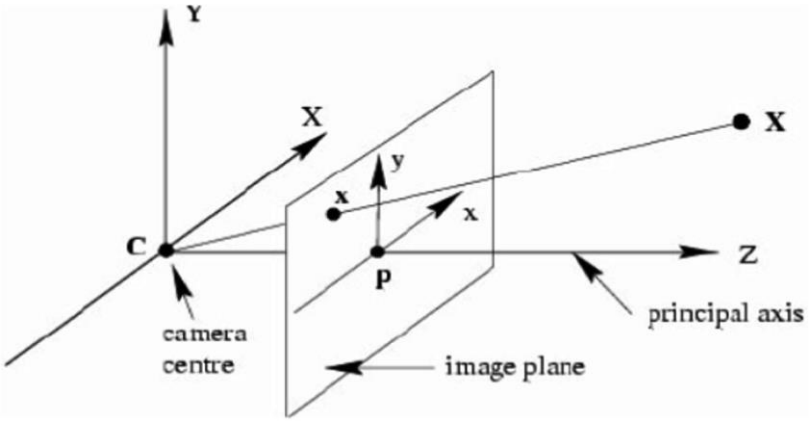


Figure 10: Schematic visualization of projection plan and 3D point. [X,Y,Z] = point in 3D, [x,y,1] = point on projection plane adapted from et al

After the 3D positions of the IR-markers are known, the 3D reconstruction is mapped to the App Coordinate System, and the position of the reconstruction is updated.

4.2.3 Registration and visualization

To determine the pose of geometric markers that can be customized, the detected pointset (p_{dps}) is aligned with the source tool geometries (p_{sps}), using a Singular Value Decomposition (SVD) algorithm [66]. SVD produces a rigid transformation between two-point sets. Therefore, the centroids of both point sets are calculated as follows:

$$c_{dps} = \frac{1}{N}\sum_{i=1}^{N} p_{dps} \tag{5}$$

$$c_{sps} = \frac{1}{N}\sum_{i=1}^{N} p_{sps} \tag{6}$$

Next, the point sets are centered by subtracting the centroids of the corresponding point sets:

$$D_i = p_{dps}^{i} - c_{dps} \tag{7}$$

$$S_i = p_{sps}^{i} - c_{sps} \tag{8}$$

Based on the matrices D and S, the cross-covariance matrix H is computed as follows:

$$H = S^T D \tag{9}$$

where S^T is the transpose of matrix S. The compute the rotation matrix, SVD is performed on H:

$$H = U \sum V^T \tag{10}$$

Here, U and V are orthogonal matrices, and Σ is a diagonal matrix containing singular values. Then, the rotation matrix can be determined by:

$$R = UV^T \tag{11}$$

Now, the translation vector is calculated by:

$$t = c_{dps} - Rc_{sps} \tag{12}$$

Finally, the 4x4 transformation matrix T can be constructed, which includes the translational and rotation:

$$T = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \tag{13}$$

The tool was considered invisible to the HL2 if there was no valid correspondence between both point sets.

4.3 DINO-CLINIC

An IR-tracking software for the HL2 has been developed based on DINO-DLL [61], named DINO-CLINIC, standing for Detection for Infrared Navigation with OST – Clinical Live Intraoperative guidance for NSS In Children. Developing and implementing a real-time kidney tracking and holographic visualization system should facilitate surgeons during NSS; therefore, multiple clinical requirements are specified. Particularly in pediatric surgery, the abdominal operating space is limited. Therefore, small sized IR-markers are required. The current DINO application uses ~1.0 cm flat IR-stickers and large shapes geometric tools. Therefore, we propose to use small IR-reflective spheres and geometric shapes. These small 3D IR-reflective spheres and smaller geometric tools reduce spatial interference within the abdomen, potentially enhancing the detection angle.

The complete clinical workflow consists of the preoperative phase, preparation of the DINO-CLINIC application, and the intraoperative use of DINO-CLINIC, see Figure 11. In the preoperative phase, a preoperative MRI is made a week before surgery [67]. Within the MRI protocol, two specific scans are used for segmentation: Balanced Triggered Angiography Non-Contract Enhanced (B-Trance) and the T2-weighted Fat Saturation MultiVane XD (T2 FS MVxd HR 4 mm), with 4 mm slice thickness. Tumor and healthy kidney parenchyma are segmented based on the T2 FS MVxd HR 4 mm scan. B-Trance scan is beneficial for segmenting the renal arteries and the vasculature of the kidney. Semi-automatic segmentation uses 3D segmentation software like 3D Slicer or Mimics Innovation Suite.

The DINO-CLINIC preparation phase consists of importing the patient-specific 3D model into the DINO-CLINIC application. In consultation with a pediatric surgeon, a choice is made regarding the IR-marker suitable for intraoperative use in this specific case. The IR-marker is placed on the tracked organ during the virtual planning within Unity 19.4.22.f1. Then, the 3D model tracking scene is linked to the patient selection scene, and the application is deployed on the HL2.

In the OR, the surgeon selects the correct patient, and then the patient-specific tracking scene is visualized. Based on visual assessment, the surgeon places the IR-markers onto the kidney surface and fixates it with tissue glue. The kidney is tracked based on manual registration, and the hologram is continuously updated based on the kidney's movement.

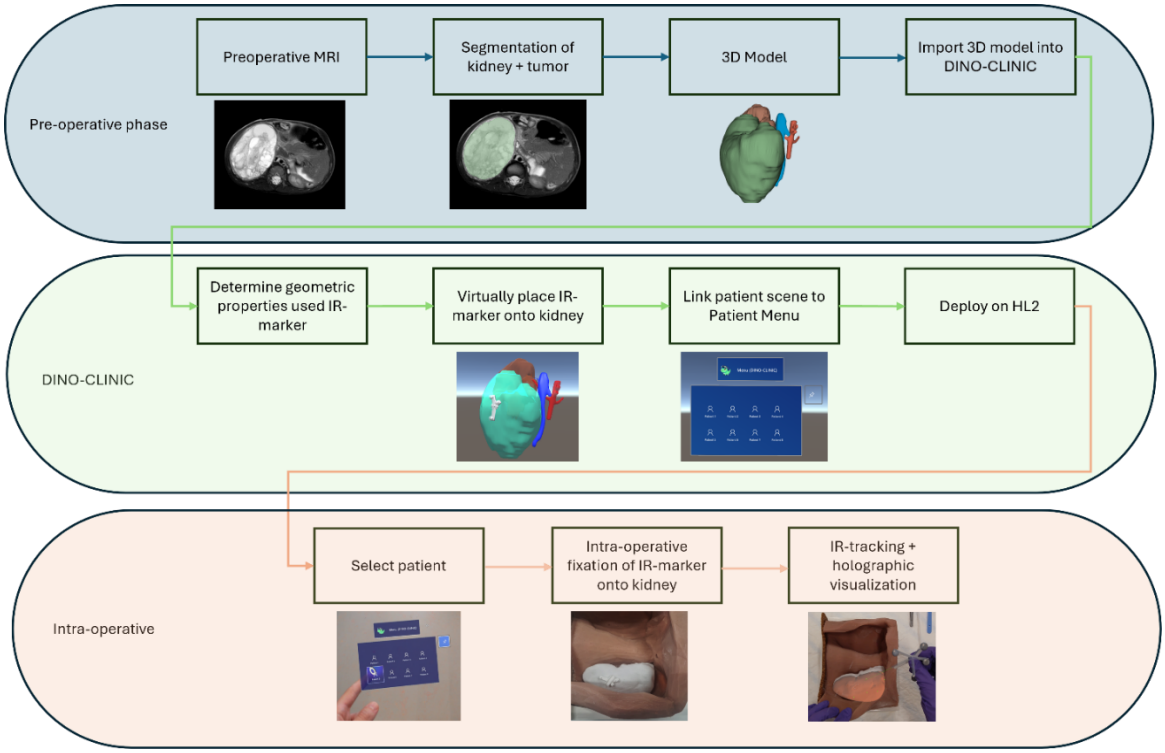


Figure 11: Schematic overview of the workflow of DINO-CLINIC, including the preoperative phase, DINO-CLINIC, and intraoperative phase. The pre-operative phase consists of segmentation-based preoperative imaging. DINO-CLINIC consists of virtually planning the IR-marker placement on the 3D model. In the intraoperative phase, the surgeons can select the correct patient and fixate the IR-marker for holographic visualization and continuous tracking.

4.3.1 Adaptation DINO-DLL

The threshold area size within the DINO-DLL algorithm needs adjustment to accommodate the smaller IR-markers. The area threshold is reduced ($1 < A < 16384$) to track IR-markers of 6.4 mm in size (Appendix C.1). Secondly, DINO-DLL initially uses flat IR-reflective stickers. However, IR-reflective spheres were used instead for an optimal viewing angle. The depth point is compensated by the radius (r) of the IR-markers to compensate for the diameter of the IR-markers, see Figure 12 (Appendix C.2).

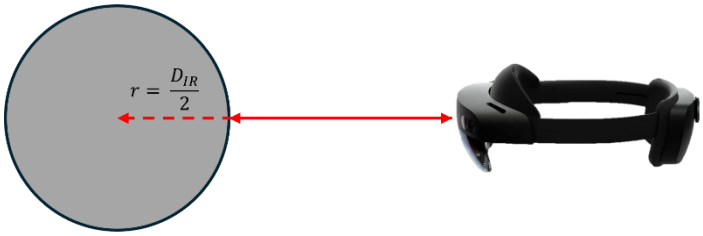


Figure 12: Compensation for diameter of IR-sphere based on radius.

An extra function is added to DINO-DLL for validation purposes, which saves all transformations of the detected IR-markers for 30 seconds, see Appendix C.3.

4.3.2 User-interface

To streamline the surgical workflow, the system includes a patient selection menu. This feature allows surgeons to quickly navigate to a patient's specific 3D model, ensuring that the right patient-specific tracking scene is available during surgery, see Figure 13. Therefore, a patient selection menu is designed using the Mixed Reality ToolKit (MRTK) [68]. The patient selection menu screen holds up to eight patient-specific

IR tracking scenes. The surgeon can position this menu anywhere in the room by pressing the pin button or manually moving the menu to the desired location. Secondly, the menu has a functionality that automatically follows the surgeon's movement across a room.

A hand palm user interface (UI) has been integrated into the system, providing surgeons with intuitive control over the holographic visualization when their left-hand palm is visual. Key features of this UI include a transparency slider embedded within the hand palm UI, enabling surgeons to adjust the transparency of the 3D hologram. This function allows for a clear view of the hologram or the underlying anatomical structures as desired by the surgeon's needs. Lastly, an exit button is also included in the hand palm UI, allowing surgeons to exit the application.

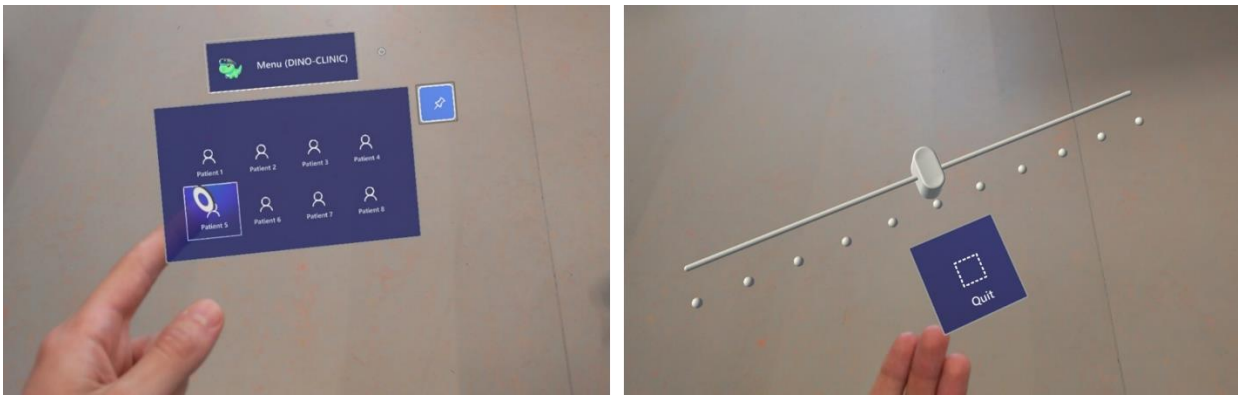


Figure 13: User interface DINO-CLINIC: Left: Patient selection menu. Right: Hand-palm UI with integrated slider. UI is activated based on the left hand-palm.

5
Validation of IR-tracking accuracy with HoloLens 2

5.1 Introduction

DINO-CLINIC is an in-house developed HL2 application, based on DINO-DLL [69], to intraoperatively track the position of the kidney in real-time for holographic visualization of WTs. As described in Chapter 4, DINO-CLINIC uses IR-reflective spheres fixated on a marker, which can be fixated to the kidney surface, to allow for intraoperative tracking of the kidney's position. Thereby, it could potentially overcome the limitations of QR tracking.

Multiple studies have evaluated the accuracy of IR tracking with the HL2, as described in Chapter 4.1. These studies evaluate the algorithms based on dynamic and rotational accuracy. Similar approaches reported tracking accuracy between 0.25 mm and 2.03 mm and a rotational accuracy of 0.8° – 1.12° in dynamic situations [46], [60], [62]. In these studies, the primary goal is to track surgical tools. Therefore, relatively large diameter IR-spheres enhance the IR-detection range over a large distance.

In the case of using IR-markers for intraoperative organ tracking, the IR-marker should be as small as possible to allow IR-tracking of the kidney as it is fixated onto the kidney. Large IR-markers will disrupt clinical workflow because they take up too much space within a pediatric abdomen's small, restricted space. Therefore, the question remains if the configuration of the IR-marker influences the accuracy of the IR-tracking algorithms. Brown et al. validated the influence of multiple configurations on the tracking accuracy of a conventional optical tracking system (NDI Polaris Vicra) [70]. Their geometric tools designs were limited to a number of four IR-markers, as the TRE significantly drops from three to four markers. There was no significant difference in tracking accuracy between the different designs. However, their geometric shapes did not vary in specific attributes determining the final IR-marker size, such as the number of IR-markers and physical size.

Therefore, the accuracy and technical performance of multiple possible IR-tracking shape configurations must be measured before proceeding to the pre-clinical validation. This section aims to determine the minimal size of the IR-marker. The physical size of the IR-marker is influenced by the diameter of the IR-reflective sphere, the size of the geometric shape, and the number of used IR-markers. Preferably, all aspects are as small as possible. Therefore, testing multiple configurations under variating view distances and angles is essential.

This study aims to validate the accuracy, FPS, and tracking detection rate (TDR) of the designed IR-markers under different conditions. The technical validation aims to result in a recommendation for IR-marker configuration, which are potentially clinically feasible for intraoperative organ tracking and can be used in a pre-clinical experiment.

5.2 Materials and Methods

The IR tracking performance was evaluated by three experiments based on the system's static accuracy, dynamic accuracy, and FPS. A static accuracy experiment, which aims to validate the performance of IR-markers, was performed first. Based on the initial results of the static experiment, specific markers were selected for the dynamic and FPS experiments.

Technical requirements were discussed with two pediatric surgeons and a technical physician with experience with intraoperative AR. The basis for the technical requirements lies within the expected clinical usability and the achieved accuracy. In previous internal studies, a 5 mm deviation of the hologram is considered clinically relevant. Within the proposed workflow, the system's accuracy depends on tracking, manual registration, and visualization accuracy. Therefore, in accordance with the reported accuracy of Hisqab et al., the limit of technical tracking accuracy is a translation error of 3.0 mm [69]. Secondly, at least an FPS of 10 should be achieved.

5.2.1 Design IR-markers

Multiple IR-marker designs were developed using Autodesk Fusion 360 (Autodesk, San Francisco, United States) to evaluate the accuracy of various intraoperative IR-markers (see Figure 14). These designs were categorized based on the number of IR-markers (three or four), the dimension of the geometric shapes (XS, S, M, L), and the diameter of the IR reflective spheres (6.4 mm or 12.7 mm), see Appendix A. The markers were 3D printed with the Bambu Lab X1 Carbon (Bambu Lab, Shenzhen, China). The scaling steps based on the smallest geometric shape (S) were 150% (M), 200% (L), and 250% (XL).

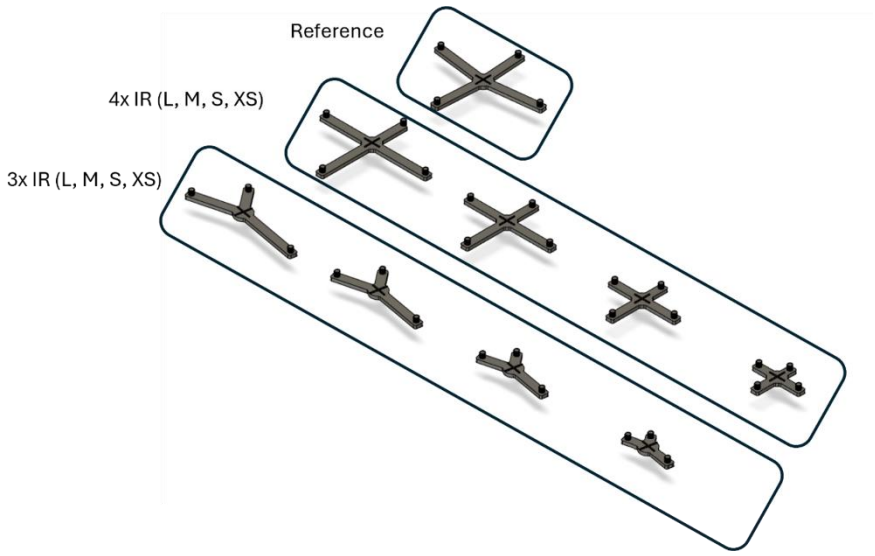


Figure 14: Designed geometric IR-markers. Three categories: 3x IR (XS, S, M, L), 4x IR (XS, S, M, L) and Reference IR-marker.

5.2.2 Hummel board design

A hummel board is a plate with fixed positions [71], often used to validate optical or EM tracking systems. This board allows the fixation of multiple geometrical shapes at a specified distance from each other. Within AutoDesk Fusion 360, a hummel board was designed with five-by-five holes, each separated by 50 mm. Within the center of the hummel board, extra pinpoint holes were designed at an angle of 45 degrees and 50 mm from the center point. To fixate the geometric shapes, custom-made 3D clamps were created that closely fit into the Hummel Board's pinhole and ensure that the IR-marker center point, corresponding to the center point calculated between all IR-markers, is in the middle of the pinhole. Lastly, to fixate the Hummel Board at multiple angles, multiple stands were created with an angle of 22.5, 45, 67,5, and 90 degrees. The hummel board, clamps for geometric shapes, and stands were 3D printed with a Bambu Lab X1 0.04 nozzle printer with 0.08 mm layer height using Polylactic Acid filament.

5.2.3 Static accuracy

The static accuracy experiment was divided into two phases. The first phase aims to determine the accuracy of the large IR-marker with sizes M and L and an IR-sphere diameter of 12.7 mm using the Optitrack Motion Capture Marker 12.7 mm M3 (OptiTrack, Corvallis, United States). Secondly, the goal was to

determine the influence of distance and viewing angle on the tracking accuracy. Therefore, two geometric shapes, consisting of one reference geometric shape and an intra-operative geometric shape, were fixated at a distance of 20 cm from each other on the Hummel Board. The HL2 was fixated on a tripod, and the distance from the HL2 (Microsoft, Redmond, Washington, United States) camera to the geometric shapes was changed to 30 cm, 60 cm, and 90 cm, see Figure 15. Secondly, the view angle was changed to 45 or 90 degrees for each distance, resulting in 24 configurations. For each configuration, 30 seconds of transformations ${}^{GeometricShape}T_{HoloLens}$ and ${}^{Reference}T_{HoloLens}$ were recorded, resulting in n~400 measurements.

The second phase determines the accuracy of the small IR-markers (XS, X, M) with IR diameter markers of 6.4 mm at 60 cm distance over multiple angles (22.5°, 45°, 66.7°, 90°) in a total of 24 configurations. A distance of 60 cm was regarded as the average viewing distance during an intra-operative situation.

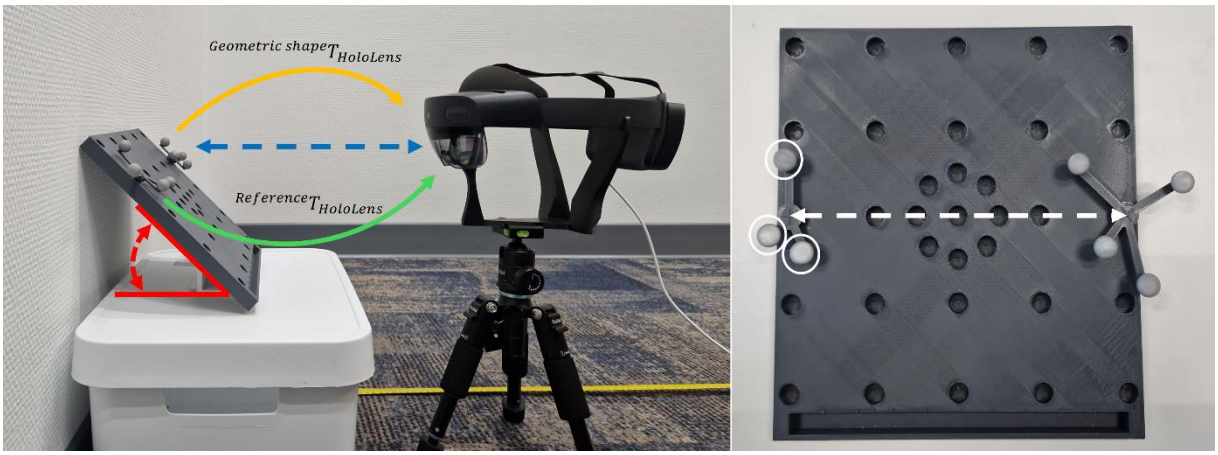


Figure 15: Overview of IR-accuracy validation setup using the HoloLens 2. Red: Viewing angle. Blue: Viewing distance. Yellow: Transformation of IR-marker to HoloLens 2. Green: Transformation of Reference to HoloLens 2. White line: fixed distance between IR-markers (200 mm). White circle: diameter of IR-spheres.

5.2.4 Dynamic accuracy

The dynamic experiment validated the dynamic tracking capabilities between the 12.7 mm IR-sphere and the 6.4 mm IR-sphere using Optitrack Motion Capture Marker. Therefore, a small IR-marker (3x IR, M, 6.4 mm) and a large IR-marker (3x IR, M, 12.7 mm) were used. The IR-markers were fixated next to the IR-reference marking with the corresponding IR-diameter. While wearing the HL2, the hummel board was moved across the x,y, and z axes (each axis circa 10 seconds) within a 30-second during measurement.

5.2.5 FPS and TDR

For the IR-marker (3x IR, M, 6.4 mm) and the IR-marker (3x IR, M, 12.7 mm), the average TDR and FPS were measured based on the results of the dynamic accuracy. The FPS is defined as the total number of updated frames divided by the duration of the dynamic measurement. The TDR indicates how often the IR-marker is detected during the dynamic experiment divided by the total duration of the measurement.

$$FPS = \frac{N_{total\ frames}}{T_{measurement}} \qquad \text{TDR} = \frac{N_{FramesFiltered}}{Tmeasurement} \tag{14}$$

5.2.6 Post-processing

Post-processing steps for the static and dynamic experiment consisted of converting XML files into an Excel (Microsoft, Redmond, Washington, United States) file, which is further processed within MATLAB 2023a (The Mathworks, Natick, Massachusetts, United States). The first filter steps within MATLAB consisted of

deleting all measurements not detected as an IR-marker together with the corresponding matrices. Based on the translational matrices of the reference marker and the geometric shape, the Euclidean distance between both markers at each recorded timestep was calculated using the following equation:

$$Euclidean\ distance = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \quad (15)$$

Then, the system's accuracy was defined in mm accuracy by subtracting the calculated Euclidean distance from the ground truth, which was 200 mm in all experiments.

$$Accuracy = 200 - Euclidean\ distance \quad (16)$$

A negative accuracy value indicates that the measured distance is larger than the actual distance, and, therefore, the distance is overestimated. A positive value indicates an underestimation of the actual distance.

5.2.7 Statistical analysis

The dataset of the static and dynamic experiments, with all groups, were tested with IBM SPSS Statistics 29.0.2.0 (IBM, Armonk, New York, United States) for normality using the Kolmogorov-Smrinov Test. Secondly, the medians of all groups were statistically tested for significance using a Kruskal-Wallis Test (one-way non-parametric ANOVO).

5.3 Results

This section presents the results of the static, dynamic, and FPS/TDR experiments.

5.3.1 Static accuracy

The performance of the IR-marker tracking algorithm was evaluated under various conditions for phase one (12.7 mm). In total, 24 measurements were performed, with deviating distance (30 cm/60 cm/90 cm), size (M/L), number of IR-markers (3/4), and angle (45/90). The results are visualized in the violin plots in Figure 16. A static accuracy of -0.12 ± 0.69 mm is measured for phase one (n = 11404). Specified for the size of the IR-marker, M, and L are the accuracy, respectively -0.07 ± 0.73 mm and -0.16 ± 0.66, for all measurements. For the number of IR-spheres, 3 and 4, the average accuracy is 0.02 ± 0.69 mm and -0.26 ± 0.68 mm.

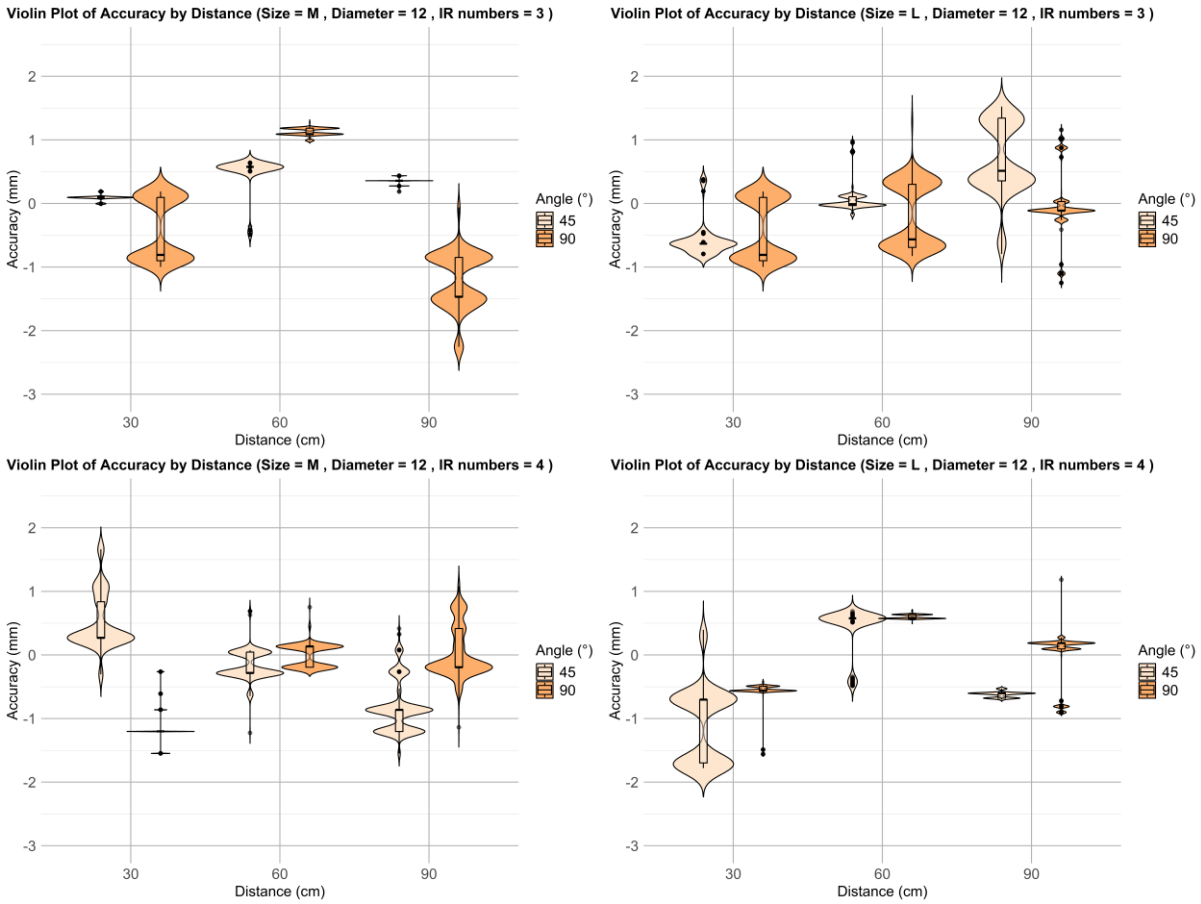


Figure 16: Results phase one: static accuracy results over multiple distances and angles. Top-left: 3X IR, Medium, 12.7 mm, Top-right: 3X, Large, 12.7 mm. Bottom-left: 4x IR, Medium, 12.7 mm. Bottom-right: 4x, Large, 12.7 mm

The phase two results evaluate the influence of the smaller IR-markers (6.4 mm) across a fixed distance of 60 cm for multiple angles (22.5°, 45°, 67.5°, 90°) and multiple sizes (XS, S, M). In Figure 17, the results are visualized. Overall, there is a mean of -0.73 ± 0.62 mm for all configurations (n = 8268). Specified for the size of the IR-markers (XS, S, M), there is an overall mean of -0.75 ± 0.55 mm, -0.79 ± 0.52 mm, and -0.67 ± 0.74 mm. Visual inspection of the violin plots shows that the accuracy is similar for all sizes and the number of IR-reflective spheres. However, an angle of 90° seems more accurate, as it is closer to zero.

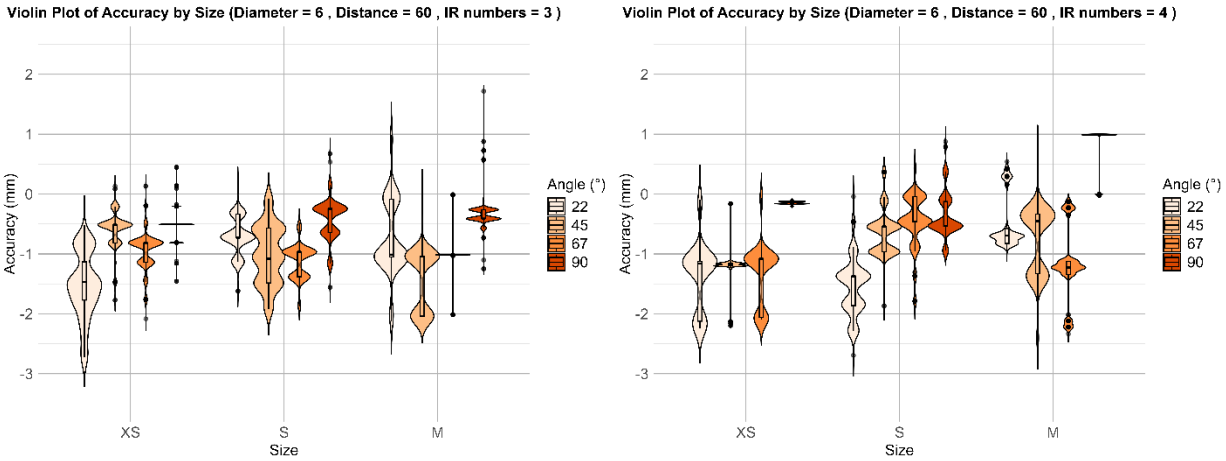


Figure 17: Results phase two: static accuracy for multiple angles and sizes. Left: 3x IR, 6.4 mm. Right: 4x IR, 6.4 mm.

5.3.2 Dynamic accuracy

Secondly, the results of the dynamic experiments are presented (Figure 18), which shows the impact of movement on accuracy measurements. The dynamic accuracy of the IR-marker (3x IR, M, 6.4 mm) is -0.80 ± 0.61 mm (n= 198). The accuracy for the IR-marker (3x IR, M, 12.7 mm) is -1.21 ± 0.44 mm (n = 348).

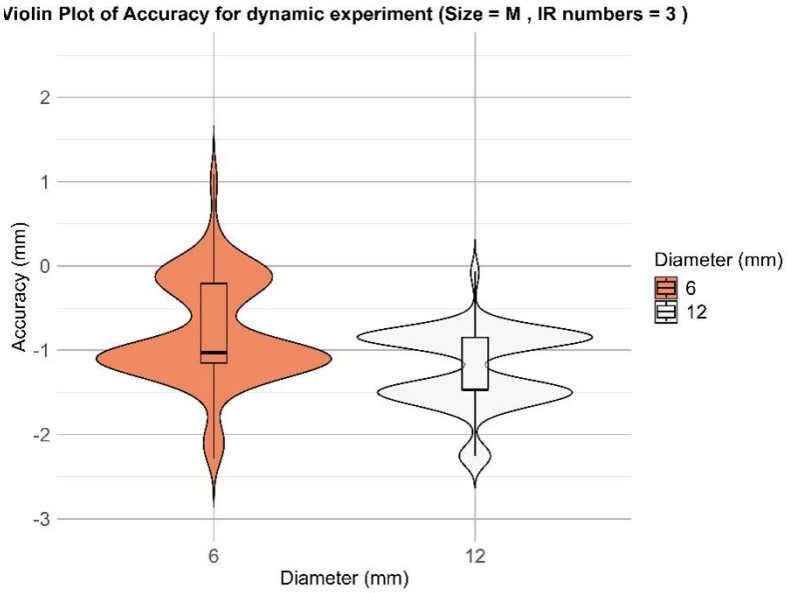


Figure 18: Results of dynamic accuracy experiment of 3x IR, Medium 6.4 mm and 12.7 mm IR-sphere.

5.3.3 FPS and TDR

Lastly, the FPS is determined for the 6.4 mm diameter marker and the 12.7 mm diameter marker. For the IR-marker (3x IR, M, 6.4 mm), the average FPS is 11.6. For the large IR-marker (3x IR, M, 12.7 mm), the average FPS is 12.56. TDR for 6.4 mm and 12.7 mm IR-markers are, respectively, 6.6 and 12.56. This result indicates that the detection rate for the 6.4 mm IR-marker is lower than the 12.7 mm; potentially, the HL2 finds it more difficult to detect the smaller 6.4 mm IR-markers.

5.4 Discussion

The study focused on comparing the performance of IR-markers with sizes of 6.4 mm and 12.7 mm for multiple configurations. The evaluation included static accuracy assessments under various conditions, dynamic accuracy validations during motion, and FPS and TDR measurements to evaluate the system's responsiveness.

In the first phase, the performance of the 12.7 mm marker was analyzed under various conditions, with an overall accuracy of -0.12 mm and a standard deviation of 0.69 mm. The findings revealed inconsistencies in variance across several configurations. Some configurations displayed overestimation, while others showed underestimation of the actual distance. Overall, there is an overestimation of the actual distance. Based on the violin plots, there does not appear to be a correlation between distance or angle on the accuracy of the IR-tracking.

Moreover, there is no clear difference between the number of IR-markers and the size of the IR-marker used. Statistical analysis identified significant differences in the median accuracy among various groups within the dataset.

In the second phase, the 6.4 mm marker was evaluated at a standard distance of 60 cm across multiple angles. The results were similar to those observed in the first phase, suggesting that the size of the IR-sphere

did not significantly influence IR-tracking accuracy. Overall, a mean accuracy of -0.73 ± 0.62 mm is measured. Notably, the viewing angle at 90 degrees seemed a more accurate representation of the distance between the IR markers than lower angles (22.5, 45, 67.6).

The dynamic accuracy experiments showed results consistent with the static experiments, with no significant differences observed between the 6.4 mm and 12.7 mm markers. However, the 12.7 mm marker showed a lower standard deviation, suggesting it might be less prone to measurement errors under dynamic conditions. Also, in the dynamic experiment, the actual distance is overestimated. Comparison of the 3x IR M 6.4 mm static experiment with the 3x IR M 6.4 mm dynamic experiment showed an overestimation of the distance during the static experiment; however, there was an underestimation of the actual distance during the dynamic experiments. The 6.4 mm diameter results appear consistent between static and dynamic accuracy measurements.

FPS measurements indicated that for both IR-sphere diameters, an FPS of approximately 12 is reached. The TFS showed that the smaller 6.4 mm IR-sphere encountered detection difficulties at distances exceedingly approximately 70 cm, indicating a potential limitation in detection range for smaller markers. The inability to detect the IR-marker was noticeable during the experiment, and secondly, it showed that the IR-sphere with 6.4 is detected less.

Compared with the literature, this study's static and dynamic accuracy showed similar results. The average tracking accuracy reported in other studies is 0.76 mm[46]. The IR marker tracking accuracy in another study was 1.44 mm with a standard deviation of 0.33 mm [60]. Mean Absolute Errors (MAEs) for translation and rotation were 2.028 mm and 1.122 mm by the developers of DINO [62], respectively. Regarding FPS, Kunz et al. achieved a higher FPS, which reported a TDR of 22 FPS and an FPS of 50-60.

Several limitations were identified in this study for the static and dynamic experiments. The measurement accuracy of the HL2 is limited by its integrated algorithm, which visualizes 3D positions based on 1 mm increments. Consequently, minor deviations below 1 mm are not accurately measurable with the HL2, potentially explaining the low variance observed in some cases and the high frequencies of specific accuracy values. Moreover, the validation method used in this study has several limitations. Firstly, the accuracy of 3D printing could impact the actual distance between the IR-markers. The Muller board, geometric markers, and design clamps were printed with a Bambu X1 Carbon printer with a nozel diameter of approximately 0.4 mm [72]. Secondly, the design process within Autodesk Fusion 360 for the clamps and geometric markers might introduce small deviations in the calculation of the center point, potentially affecting the overall accuracy of the measurements. This study did not consider rotational accuracy, which significantly influences holographic visualization due to the lever effect. A slight deviation in the rotational position of the IR-markers can have a more pronounced effect at the outer boundaries compared to areas closer to the IR-marker. Future studies should consider using external optical or EM-tracking for rotational validation, as described in existing literature [60], [62], to enhance the accuracy of the validation method.

The FPS measurements were derived from dynamic experiments in which the Muller board was moved across multiple axes. The quantitative results indicated that the TFS for the smaller IR-spheres was lower than that for the larger IR-diameter. Based on an observation, the detection rate for the smaller IR-diameter was lower at larger distances. A limitation of this experiment is that if the IR-marker is outside of the detection window, the DTR is influenced. Therefore, incorrect movement of the IR-marker outside the detection view lowers TDR, which makes this measurement observer-dependent. Further research should focus on quantifying TFS for the 6.4 mm IR-marker at varying distances. In addition, software calculations could be performed on an external PC instead of the HoloLens 2, to enhance the detection rate.

Clinical requirements were established with a translational error threshold of 3.0 mm and a frame rate of 10 FPS on the HL2 device, with TFS not included as a technical requirement. Although the TFS for the 6.4

mm diameter fell below the 10 TFS, this discrepancy was partly due to the distance between the IR-marker and the HL2 during the experiment, which made this metric less significant. Given these clinical requirements, all configurations met the necessary standards. Therefore, from a clinical perspective, evaluating the feasibility of the smallest IR-markers, specifically the 3x XS IR Marker, is recommended as the smallest marker size is preferred.

5.5 Conclusion

This study assessed multiple IR-markers' static and dynamic accuracy using the DINO-CLINIC system. Results from the first phase indicated that 12.7mm IR-markers have an overall accuracy of -0.12 mm for multiple sizes, distances, and angles. The second phase validated the 6.4 mm IR-markers, which have a slightly higher mean accuracy of -0.73 mm. A 90-degree viewing angle indicates better accuracy for both markers. Dynamic experiments showed comparable results. Overall, there is an overestimation of the measured distance. To conclude, all IR marker configurations met the technical requirements for further clinical implementation. Based on the smallest IR-marker and the performance, the 6.4 mm 3x XS IR-marker is recommended for further clinical feasibility testing. Future studies should include rotational errors and explore methods to enhance the accuracy of IR-marker tracking.

6
A phantom study: clinical feasibility of IR-tracking for NSS

6.1 Introduction

DINO-CLINIC facilitates the tracking of organs by attaching an IR-marker to the organ's surface. This method is particularly useful in tracking soft tissues, such as the kidney, due to an organ's freedom of movement. In Chapter 5, DINO-CLINIC demonstrated the capability to meet the required technical performance requirements; therefore, it is suitable to explore the DINO-CLINIC application further in a pre-clinical phantom experiment that simulates an intraoperative NSS scenario.

The use of sensors or markers, which can be fixated on the organ's surface, for tracking the position of organs is also explored in other fields. In the field of EM-tracking, sensors are used to track the liver. Market et al. proposed a tracking system using sensors fixated for real-time tracking of the organ shape and vessel locations. The tracking system was integrated with a hepatectomy simulation software to virtually visualize the deformation of a plastic liver model [73]. Tracking is also already used intraoperatively. Smit et al. used a combination of iUS and EM-tracking for intraoperative localization of liver lesions [74]. Therefore, after laparotomy, an EM-sensor is attached to the surface of the liver, and a patient-specific 3D model is registered using iUS.

Optical markers are also used within the field of surgical guidance in the liver. A study demonstrates the use of reflective fiducials fixed on the surface of the liver to measure the position and deformation of the liver in an experiment of a pig's liver [75]. Similar methods are introduced to AR. Yinoussa et al. developed a marker-based navigation and registration method for open liver surgery [76]. Therefore, markers are manually placed on the liver's surface during surgery and tracked in real-time by IR-cameras.

These studies demonstrate the potential feasibility and utility of organ tracking through intraoperative studies and phantom experiments. Currently, there is no research on using IR-markers for organ tracking in open surgery for kidneys. Tracking organs intraoperatively poses several challenges, including restricted space using a compact IR-tracking marker. Secondly, line-of-sight (LoS) remains an issue. However, compared to standard surgical optical tracking, it is less of an issue since the IR camera is integrated into the HL2 worn by the surgeon.

Additionally, the manual placement of the IR-marker is inherently challenging. Mainly because there is an absence of clear anatomical landmarks on the kidney surface, errors in the holographic visualization of DINO-CLINIC can occur from the manual placement of the IR-marker during surgery and inaccuracies in the visualization of the hologram.

A manual registration experiment and a clinical feasibility study are conducted to evaluate the IR-tracking capabilities for kidney tracking on a phantom model. The manual registration experiment aims to assess the accuracy of IR-marker placement by multiple surgeons based on preoperative planning. Meanwhile, the clinical feasibility study investigates the practicality of using DINO-CLINIC for intraoperative IR-tracking. Through these studies, the potential of DINO-CLINIC in enhancing the precision and feasibility of intraoperative organ tracking is systematically explored.

6.2 Materials and Methods

This study evaluated the intraoperative accuracy and usability of the DINO-CLINIC software. This section is divided into two main experiments: the manual registration experiment and the clinical feasibility experiment. The manual registration experiment aimed to assess the accuracy of pediatric oncological surgeons' geometric marker placement on phantom kidneys compared to the preoperatively planning. The clinical feasibility study evaluated the intraoperative usability and holographic accuracy of the DINO system in a clinical setting.

6.2.1 Phantom

To create 3D holographic models and phantoms for this experiment, a pig kidney was used, which was scanned using a CT-scanner (Siemens Somatom Spiral CT scanner, Erlangen, Germany) with a 0.6 mm slice thickness at the Faculty for Veterinary Medicine (University of Utrecht, the Netherlands). The CT images were manually segmented using 3D Slicer 5.4.0 [77], [78]. The segmentations were smoothed with a smoothing factor of 1.5. On the surface of the kidney, five target points are selected, which are necessary for the accuracy of the experiment. After smoothing the 3D models, they were imported within Autodesk MeshMixer (Autodesk, San Francisco, California, United States) to create an object file, which was later imported into the DINO-CLINIC software using Unity 2019.4.22.f1 (Unity Technologies, San Francisco, United States). Secondly, within AutoDesk Fusion (Autodesk, San Francisco, United States), five pivot points were integrated within the phantom kidney to facilitate EM registration, which was used as a validation device. Moreover, a small cylinder was integrated into the kidney to fixate the reference sensor, see Figure 19. The 3D kidney model, without targets, was printed with a Bambu Lab X1 Carbon (Bambu Lab, Shenzhen, China), see Figure 19. To provide a minimal deformable model, the kidney was printed with thermoplastic polyurethane (TPU), which has deformable characteristics. A print density of 8% was used to enhance the deformability of the 3D model.

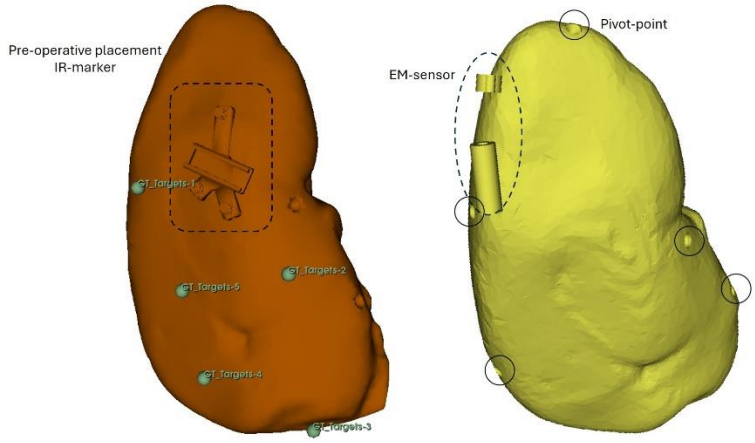


Figure 19: Overview of phantoms used for validation and landmark placement. Left: a 3D hologram of the kidney with the target points and virtual placement of the IR-marker. Right: 3D phantom with integrated EM sensor and pivot-points for registration.

6.2.2 Manual registration experiment

The primary objective of the manual registration experiment was to determine the accuracy of the intraoperative placement of the geometric shape in comparison with the preoperative planning. Inaccurate placement of the IR-geometric marker on the kidney could negatively affect the holographic accuracy.

A commercially available EM-tracking system by NDI (Northern Digital Inc., Polaris Vega ST, Waterloo, Ontario, Canada) was used to validate the manual registration experiment, see Figure 20. The field generator was placed directly beneath the abdominal phantom to enhance EM tracking accuracy. The IR-marker, integrated with an EM-sensor, was placed on the kidney along with a reference EM-sensor. Respectively, three and five-point registration was performed with EM-tracking to track both the geometric marker and the phantom kidney in EM coordinates. Participants in the experiment were pediatric oncological surgeons (n=3), the end-users of the DINO system. The experiment involved showing the surgeons preoperative 3D images and asking them to place the EM-tracked IR-marker on the kidney as accurately as possible without using HL2. Each surgeon performed the placement on the TPU phantom kidney. After this, the surgeons were asked to place the IR-marker with the help of the HL2 visualization. Therefore, the surgeons could overlay the kidney model based on visual inspection. Based on the surgeon's positioning of the IR marker, the translation and rotational difference can be determined compared with the preoperative planning. Secondly, based on the planning, an estimation of the TRE was calculated based on a new registration.

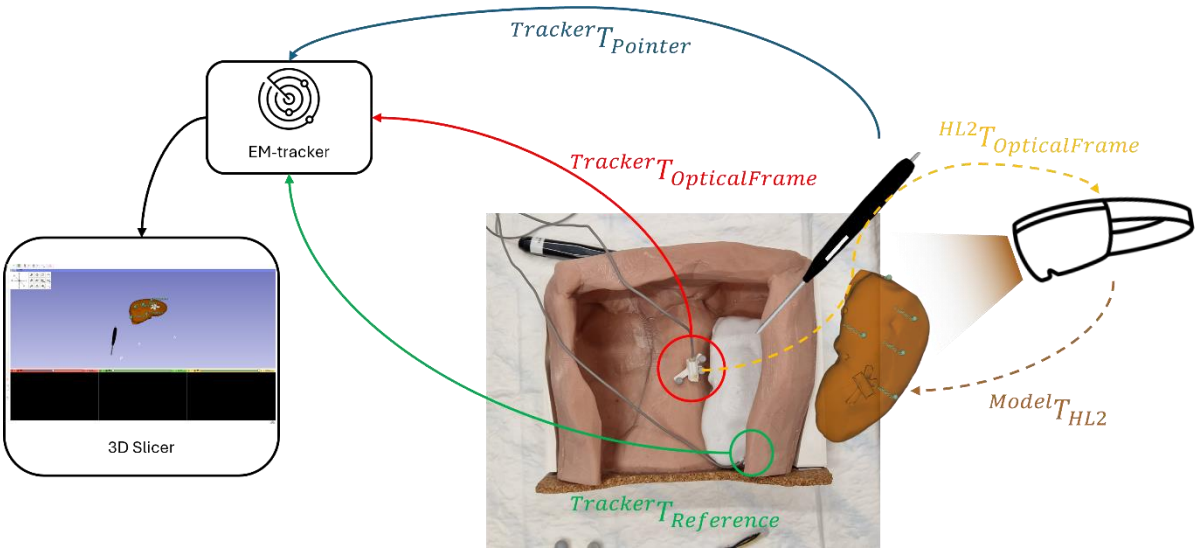


Figure 20: Experiment set-up landmark and clinical feasibility experiment. Green, Red, and Blue visualize the transformation of the necessary EM-tools for validation. Yellow and Green visualize the transformation regarding the HL2 (HL2).

6.2.3 Manual registration experiment post-processing

Within 3D Slicer, 3D points were placed within the pinholes IR-marker of the preoperative placement and the actual placement by the surgeon, resulting in two pointsets ($p_{planning}$ and $p_{placement}$). Then, within MATLAB 2023a (The Mathworks, Natick, Massachusetts, United States), the centroids of two-point sets were calculated, and the point sets were centered by subtracting these centroids. The translation error was determined by subtracting the centroid of $p_{planning}$ with the centroid of $p_{placement}$. A covariance matrix was computed from the centered points, and the SVD algorithm was applied to determine the rotation matrix. The rotation matrix was converted to Euler angles. Therefore, the output parameters included manual registration error, measured in terms of translation (in millimeters) and rotation (in Euler angles), and preoperative placement was compared with participant placement.

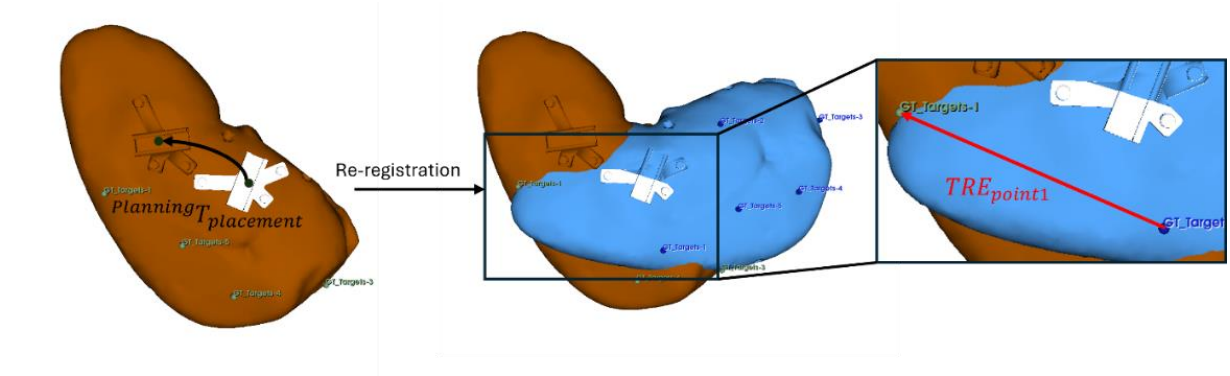


Figure 21: Reregistration within 3D slicer based on the IR-marker placement of surgeons + visualization of TRE.

The TRE for all five points was calculated based on the placement of the IR-marker, see Figure 21. Therefore in 3D slicer, for the initial placement (without HL2) and placement with HL2, a reregistration using a 3-point fiducial registration of a copy of the 3D model was performed based on the placement of the IR-marker. Then, the Euclidean distances between all five points were calculated.

6.2.4 Clinical feasibility experiment

This section aimed to evaluate the DINO-CLINIC system's clinical feasibility and the holographic visualization's accuracy. The experiment was conducted subsequent to the final placement in the manual registration experiment. In this experiment, pediatric surgeons (n = 3) were asked to identify five target points on the phantom kidney using the holographic visualization following a specified order (1 to 5). An EM-pointer, tracked in EM coordinates in correlation to the kidney's position, was used to point out the five targets, see Figure 22. If necessary, the surgeons were permitted to manipulate the kidney during the procedure. If the surgeon was satisfied with the localization of the EM-tracker tip on the target point based on the holographic visualization, the location of the target point was saved digitally within 3D Slicer. The PLE, similar to the TRE, was determined for all five target points [79].

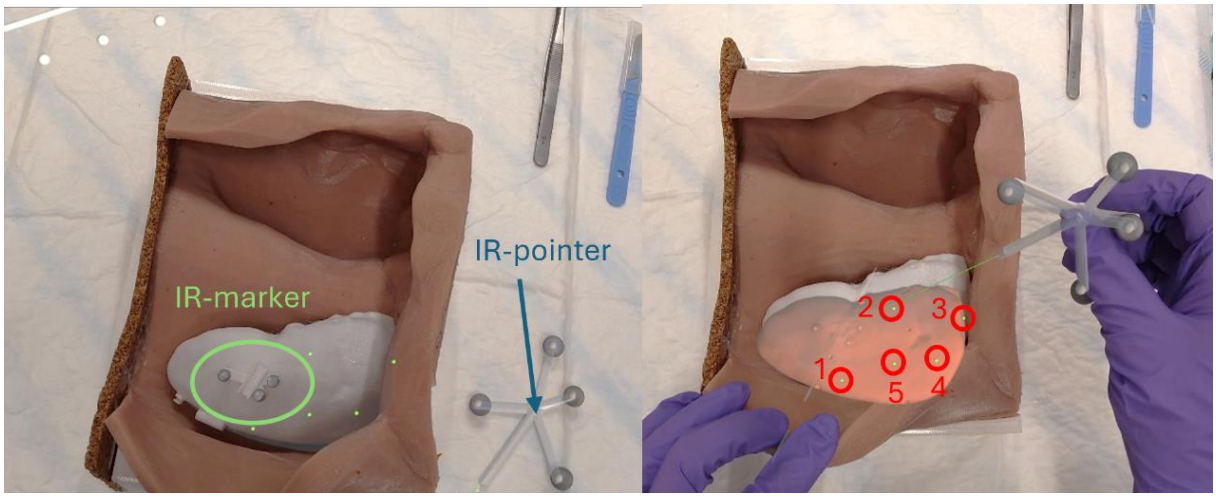


Figure 22: Holographic visualization of the experiment. Left: Placement of IR-marker on the kidney. Right: Holographic visualization of target points in the correct order (1 → 5).

Following the completion of the target identification task, the surgeons were asked to complete a System usability Scale (SUS) questionnaire rating multiple questions on a Strongly Disagree (1) and Strongly Agree (5) scale [55]. To calculate the final score, for the odd-numbered questions, one was subtracted from the score. Then, from the even-numbered questions, five were subtracted from the score. Both scores were summed and multiplied by 2.5 for the final SUS score. Additionally, the surgeons were asked to complete

another self-developed survey, which mainly focuses on the DINO-CLINIC application, on a Strongly Disagree (1) and Strongly Agree (5) scale.

6.2.5 Statistical analysis

The normality of the TRE dataset was assessed using the Kolmogorov-Smirnov test in IBM SPSS Statistics 29.0.2.0 (IBM, Armonk, New York, United States). To determine the significance of differences among the medians of the HL2, Initial and PLE groups, the Kruskal-Wallis test was applied.

6.3 Results

Within this section, the results of the manual registration experiment and the clinical feasibility experiment will be reported.

6.3.1 Manual registration experiment

Multiple surgeons (n=3) placed the intra-operative IR-marker on the phantom kidney, and the translation and rotational errors were measured. The rotational and translation errors for the initial and HL2 placement by three surgeons are visualized in Table 5.

Table 5: Rotational and translational errors of placement of IR-marker

	Surgeon 1		**Surgeon 2**		**Surgeon 3**	
	Initial	**HoloLens2**	**Initial**	**HoloLens2**	**Initial**	**HoloLens2**
Roll (°)	-31.01	-26.26	-0.44	-0.00	-5.44	-4.02
Pitch (°)	-2.37	-10.75	-2.48	-12.44	5.30	-2.95
Yaw (°)	55.38	55.94	1.70	4.47	6.54	5.52
X (mm)	14.77	12.22	-1.06	6.75	-1.77	0.61
Y (mm)	25.95	1.33	1.39	4.53	-0.45	3.46
Z (mm)	12.43	9.309	1.92	4.31	1.50	2.68

The visualization of the placement of the IR-marker is depicted in Figure 23. The 3D model is registered to the placement, illustrating the shift of the target points. For surgeon 1, there is a noticeable displacement of the IR-marker compared to the preoperative planning without using the HL2. When using the HL2 by surgeon 1, this displacement is corrected, although a considerable deviation remains. Surgeon 2 placed the geometric marker closer to the preoperative planning without the HL2 than with it, generally aligning the hologram well upon visual inspection. Surgeon 3 showed minimal visual differences when comparing the initial placement to the HL2 placement.

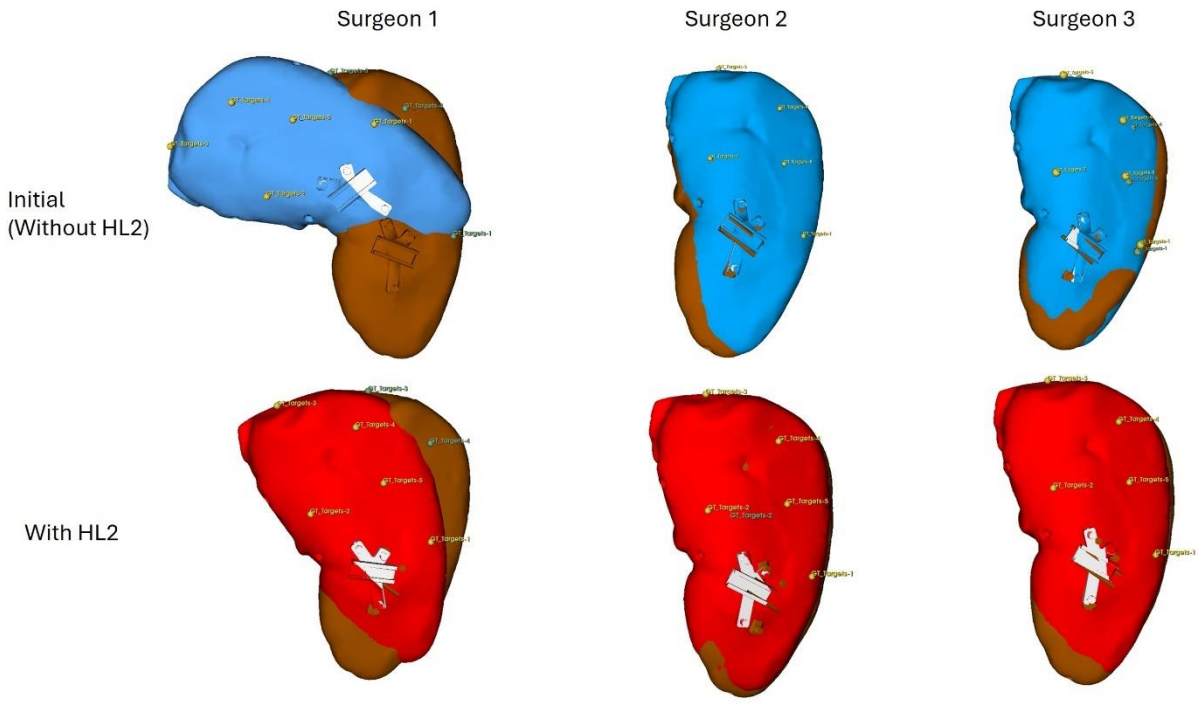


Figure 23: Visualization of results per surgeon of IR-marker placement without HL2 (blue), with HL2 (red) and ground truth (brown)

The TRE from the initial and the HL2 placements had statistical significance for normality in the Shapiro-Wilk test. Consequently, the median and interquartile range (IQR) are reported to represent the data accurately. The median TRE for the initial measurement is 5.65 mm [4.13 – 26.32]. In contrast, the median TRE for the HL2 placement is 11.82 mm [9.32 – 17.01]. Figure 24 visualizes each surgeon's boxplot of initial and HL2 TRE measurements. The Kruskal-Wallis is not significant ($p = 0.178$). Therefore, there is no significant difference in TRE between the initial and HL2 groups. Surgeon 1 has, for the initial and HL2 placement, a relatively large TRE in comparison with surgeons 2 and 3.

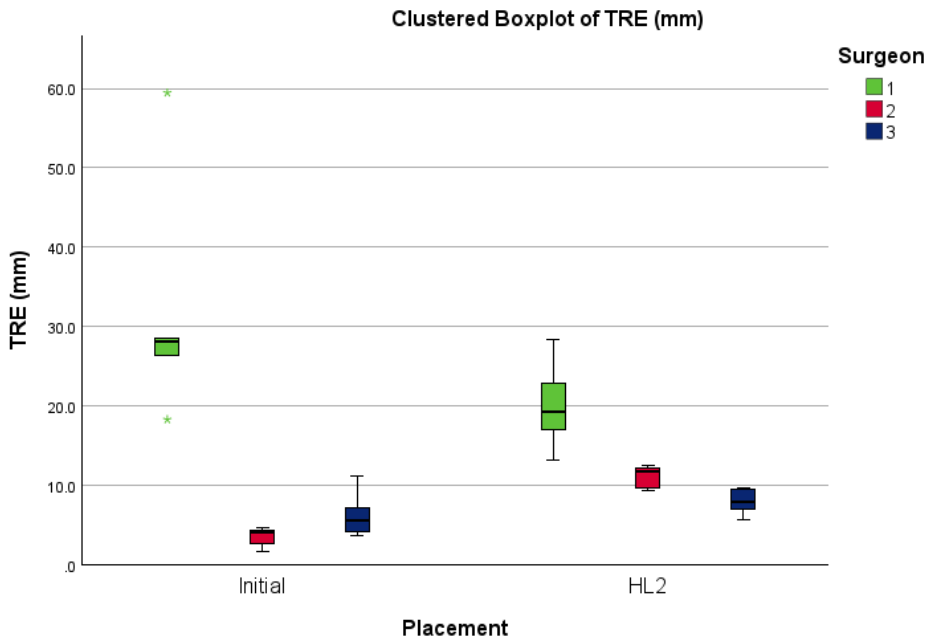


Figure 24: Boxplot visualizing TRE for each surgeon and placement (Initial or HL2).

8.3.2 Clinical feasibility experiment

During the clinical feasibility experiment, surgeons pointed to five targets based on the holographic visualization of the DINO-CLINIC application, and the PLE was measured. Additionally, the surgeons completed a System Usability Scale and an additional questionnaire regarding the DINO-CLINIC application. The PLE is normally distributed; the mean and standard deviation are reported. The median PLE is 8.74 mm [6.38 – 10.85]. Surgeon 1 has a median PLE of 9.96 mm [6.98 – 11.73], surgeon 2 has a median of 10.35 [8.85 – 12.03], and surgeon 3 has the lowest median PLE of 6.38 mm [3.13 – 7.85]. Figure 25 shows the results of the clinical feasibility experiment (PLE) and the TRE of the manual registration experiment. Compared with the initial and HL2 TRE/PLE, there are no significant differences within PLE/TRE ($p = 0.136$).

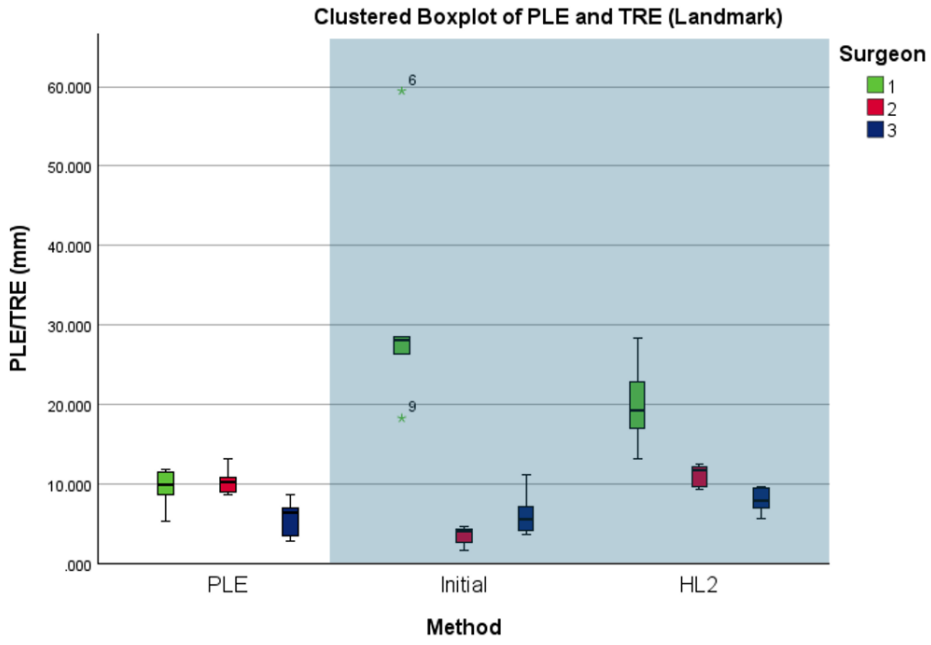


Figure 25: Boxplot visualizing the PLE and results manual registration experiment (blue)

The results of the System Usability Scale (SUS), which assesses the usability of the DINO-CLINIC application for intraoperative IR-tracking of the kidney, are presented in Table 6. Surgeon 1 and Surgeon 3 achieved a SUS score of 77.5, graded as B, indicating they found the DINO-CLINIC application to be 'good'. In contrast, Surgeon 2 reported a lower score of 60, reflecting a below-average rating.

A notable concern among the surgeons was the system's inconsistency, particularly highlighted by Surgeon 2. Despite this, all surgeons agreed that the system is easy to use, as evidenced by their responses on questions 1.2 and 1.3, indicating that the system was straightforward and not unnecessarily complex. However, there were discrepancies regarding whether the surgeons would choose to use the system again in future procedures.

Table 6: Results SUS for each surgeon

No	Question	Surgeon 1	Surgeon 2	Surgeon 3
1.1	I think that I would like to use this system frequently.	5	2	4
1.2	I found the system unnecessarily complex.	1	1	1
1.3	The system was easy to use.	4	5	4
1.4	I think that I would need the support of a technical person to be able to use this system.	3	4	2
1.5	I found the various functions in the product were well integrated	4	4	4

1.6	I thought there was too much inconsistency in this system	3	5	4
1.7	I would imagine that most people would learn to use this system very quickly.	4	4	4
1.8	I found the system very cumbersome to use.	2	3	1
1.9	I felt very confident using the system.	4	4	4
1.10	I needed to learn a lot of things before I could get going with this system.	1	2	1
SUS		**77.5**	**60**	**77.5**

The results of the self-developed questionnaire are visualized in Table 7. There were discrepancies between surgeons regarding whether the positioning of the IR marker on the kidney was straightforward. Surgeon 2 disagreed that the placement of the IR marker was straightforward, in contradiction to Surgeon 1 and 3.

The user interface was generally found to be straightforward. Based on the sub-questions, multiple surgeons considered the slider a good addition. Only surgeon 1 did not find it especially useful. This preference is caused by the surgical instruments not being easily visible when the hologram was bright and overlayed the instrument. Manual adjustment of the transparency allows for better focus on the kidney surface instead of the hologram.

One major issue the surgeons identified was that the hologram did not remain stable on the kidney surface, as all surgeons disagreed with question 2.3. This instability is primarily due to continuous tracking; even minor deviations in tracking positions result in noticeable movement of the hologram, even when the kidney is stationary. However, most surgeons found the real-time tracking of the kidney to be useful. Specifically, regarding the size of the IR-diameter, surgeons found that the small IR-marker did not provide sufficient and fast enough tracking for this purpose. The tracking was perceived as faster and more sufficient for the large IR-marker, though still imperfect.

All surgeons agreed that the holographic overlay justifies the additional time required for NSS surgery. Additionally, there was consensus that the system could benefit other clinical use cases.

Table 7: Results of a self-developed questionnaire for each surgeon

No.	Question	Surgeon 1	Surgeon 2	Surgeon 3
2.1	Determining the position of the IR-marker on the kidney of the patient was straightforward.	4	2	4
2.2	The user interface of the HoloLens software was straightforward to use.	4	4	5
2.2.1	*The hand-palm user interface was straightforward to use.*	4	4	5
2.2.2	*The integrated slider was helpful to visualize the kidney correctly.*	2	5	5
2.3	The hologram remained stable on the kidney after positioning.	1	2	2
2.4	Real-time tracking of the kidney is useful	3	5	4
2.4.1	*The infrared-marker tracking is sufficient and fast enough for the small infrared-markers.*	2	2	3
2.4.2	*The infrared marker tracking is sufficient and fast enough for the large infrared-markers.*	4	3	4
2.5	The holographic overlay is worth the additional time for NSS.	4	5	4
2.6	The system has the potential to be used for other clinical use cases	4	5	4

6.4 Discussion

6.4.1 Manual registration

The manual registration experiment evaluated the accuracy of IR-marker placement with and without using the HL2. It focused on the translational and rotational error and the TRE. The translational and rotational errors showed a large deviation between the preoperative placement, initial placement, and HL2 placement, especially for surgeon 1. The TRE for the initial measurement was 5.65 mm [4.13 – 26.32], while the median TRE for HL2 was 11.82 mm [9.32 – 17.01]. These results indicate that IR-marker placement accuracy was worse with HL2 compared to preoperative planning. This effect could be caused by surgeons' ability to visually inspect and correct the IR marker's position by aligning the hologram's borders with the kidney. Moreover, based on experience with HL2, the position of the surgeon's eyes also influences the visualization, potentially leading surgeons to compensate for this variation in their adjustments, explaining the deviation between the preoperative and intraoperative placements. In addition, surgeons 2 and 3 achieved more accurate placements without HL2, possibly because they perceived the overlay's accuracy better. Theoretically, placement similar to preoperative planning should give the best results. Another effect visible with the TRE results is the lever effect, where a slight rotation caused a significant TRE error over large distances, suggesting that the IR marker should be positioned close to the center of the trackable organ or target tissue. In contrast, the results of Surgeon 1 deviated from the others, with improved placement using HL2. This could be attributed to the difficulty in interpreting the preoperative 3D image due to a lack of visible landmarks for orientation.

Other studies showed the challenges and comparable results of manual registration. Scheider et al. reported an accuracy of 15.8 ± 14.2 mm [80]. Fotouhi et al. evaluated AR navigation using a multi-planar phantom with fiducial markers for validation [81]. Their results showed an accuracy of 13.2 ± 2.89 mm. Both studies showed similar results as indicated in this experiment.

The manual registration experiment had several limitations. The semi-deformable phantom used did not replicate the deformability of an actual kidney, making it challenging to secure the IR-marker. Additionally, the IR-marker's shape did not conform to the kidney's curvature, which could have enhanced its placement. Lastly, using EM sensors with cables further hindered the surgeons, as the cables crowded the small pediatric abdomen and added tension to the geometric marker, hindering the IR-marker stable placement.

Future research should concentrate on integrating a five-point registration system with IR-marker tracking. Manual registration is heavily user-dependent, particularly in this context. By manually fixing the IR-marker, the error in the hologram can be compensated through point-based registration, potentially enhancing accuracy. However, this approach may lead to increased registration time.

6.4.2 Clinical feasibility

The clinical feasibility experiment validated the end-point accuracy of the DINO-CLINIC system based on a phantom study and assessed its usability. There was no significant difference in TRE between initial and HL2-assisted placements, although the number of measurements per surgeon was limited. The mean PLE was 8.74 mm [6.38 – 10.85]. The determined acceptable clinical accuracy is 5 mm. Therefore, the DINO-CLINIC system currently exceeds the acceptable clinical accuracy of 5 mm. The PLE results were comparable to the TRE placement with the HL2. The surgeons performed better than the TRE based on the HL2 placement.

The results of the SUS indicated that surgeons were generally satisfied with the system's usability, ranging scores between 60 and 77.5. However, the system stability was a significant issue, particularly in tracking the small IR-marker. Surgeons noted flickering during the experiments, as the HL2 continuously updated the position even when the kidney was stationary. Flickering of the hologram is caused by the presence of noise

within the positional measurements. Moreover, a small rotation of the IR code resulted in a large visual rotation at the boundaries of the kidney. Therefore, further research should implement a Kalman filter for the tool pose estimation, which could filter out noise [82] [60]. Another limitation of the SUS questionnaire was the inclusion of surgeons experienced with AR. Including experienced pediatric surgeons could positively affect the results of the SUS as the surgeons are used to control AR applications. Including inexperienced surgeons could give a more detailed image of the system's usability, even for inexperienced users. In addition, future research should include more surgeons.

Analyzing both the questionnaire and the qualitative measurements, it was noted that Surgeon 1 found determining the position of the IR-marker easy. However, this was not reflected in the quantitative PLE measurements. This discrepancy might be due to differences in how the hologram visualization aligned with preoperative planning. It was also noted that the placement of the IR-marker did not significantly influence the PLE, suggesting other factors might be affecting overall accuracy.

Based on the self-developed questionnaire, surgeons expressed greater satisfaction with the larger 12.7 mm IR markers, suggesting these could replace QR codes due to their resistance to light disruption. Two surgeons deemed the integrated slider a valuable addition. The slider increases the visibility of the kidney and surgical instruments. All surgeons agreed that AR could be useful in NSS.

6.5 Conclusion

The study evaluated the clinical usability and accuracy of the system through a phantom experiment. The placement of the IR-marker proved challenging, with an initial TRE of 5.65 mm, which increased to 11.82 mm when using the HL2. The SUS score indicated that surgeons found the system generally usable. However, significant limitations are present within the tracking capabilities of the DINO-CLINIC, particularly for the 6.4 mm marker. This IR-Marker was rated poorly in both the SUS and self-developed questionnaire. This issue was less pronounced with the 12.7 marker, suggesting that this marker could potentially replace the QR code. In addition, the PLE (8.74 mm) indicated that the DINO-CLINIC system does not meet the clinical accuracy requirements (<5.0 mm). Therefore, further research should focus on implementing a Kalman Filter within DINO-CLINIC to address flickering. Additionally, exploring the possibility of integrating a registration method, such as point-based registration, with DINO-CLINIC is recommended. To conclude, the use of IR-markers for tracking the kidney surface using the HL2 shows potential. However, improvements are necessary before further clinical implementation.

7
General discussion

The main objective of this thesis was to apply the current AR system used in the clinical setting and to develop and validate a new application based on the clinical findings. This newly developed application enabled real-time intraoperative organ tracking to compensate for kidney movement using the HL2. The results of this thesis showed the feasibility of the AR system intraoperatively, highlighting its current limitations within the clinical context, primarily related to QR tracking and limited abdominal space to pinpoint anatomical landmarks accurately. To address these issues, a new tracking system has been developed to overcome QR-tracking and point registration limitations.

Initially, the integration of 3D holographic guidance in total nephrectomies shows promising results in terms of depth perception and correlation between tumors and critical anatomical structures. However, the software application appears unstable in the OR environment, resulting in incomplete or inaccurate registration. As previously mentioned, a large tumor and the limited space within the abdomen necessitate a markerless registration device or a small marker that minimally infers with the abdominal space. The instability of the software must be addressed, including QR tracking, registration time, and overall user feasibility.

The IR-tracking system developed showed promising tracking accuracy with sub-millimeter precision, meeting the technical requirements for further exploration within a phantom study. In comparing the QR tracking results evaluated by Eyck et al.[54], the same application used in Chapter 3, with the IR-tracking results, notable differences were observed [54]. In Eyck's study, to ensure a stable HoloLens position, the device was placed on a 3D-printed stand fixed to a table, with a QR marker positioned on a Lego tower at a specific height and distance (50 cm) from the HoloLens, simulating the distance during surgery. This method is less accurate than the muller board design, as a 3D print has very high accuracy. The median mean error using a 4.5 cm x 4.5 cm QR code was determined to be 4.5 mm. In contrast, the IR-tracking system reported an accuracy of 0.80 mm for the 6.4 mm IR-marker, sized M, in a dynamic situation. Therefore, IR-tracking proposes a significant improvement in accuracy over the QR-tracking system. In addition, similar detection angles are reported. Therefore, it was feasible to explore the use of IR-tracking as a replacement for the QR-code for holographic visualization and to implement IR-tracking of the kidney.

Clinical validation in a phantom experiment of the IR-tracking system, which consists of manual registration and continuous IR-tracking of the kidney, did not achieve the clinically acceptable TRE of 5 mm. The results of the manual registration experiment suggest that this method introduces substantial errors in the hologram visualization. Additionally, a significant limitation was the instability of the hologram. The technical results could not quantify the instability due to the lack of a rotational accuracy metric. Consequently, it is imperative to incorporate this measurement in future research to evaluate the rotational accuracy and standard deviation. A comparison of the results between the clinical AR application using the five-point registration with a QR code (see Chapter 3) and the manual registration for organ tracking with IR markers (see Chapter 6) showed that both systems exceeded the clinical criteria of 5 mm accuracy despite the limitations of intraoperative measurement methods. During the comparison, it should be considered that the intraoperative environment is more stressful and has more issues due to the restricted abdominal operating space compared to the phantom experiment. The IR-tracking system slightly outperformed overall SUS scores, particularly in ease of use, due to the simplicity of placing the IR marker in one location without needing point registration. A significant limitation of the IR-tracking system compared to the five-point registration system is its inconsistency in hologram visualization, with continuous tracking resulting in

hologram instability due to constant updating of the hologram's position. The 5-point registration is a one-time registration, and the tracking is purely based on the fixated QR code. Therefore, the hologram remains more stable. Both studies showed the potential use of AR for NSS. However, neither application meets the clinical accuracy requirements. Secondly, it should be noted that the surgeon had discrepancies in the usability of both systems. Therefore, while IR-tracking enables live tracking, it does not currently improve the overall usability of the AR system.

7.1 Clinical implementation

Additional considerations are required for further clinical implementation of the DINO system, particularly for organ tracking in open surgery.

Sterilized IR-sphere markers, such as the NDI 11.8 and 13 mm markers, are readily available in clinical settings [83]. However, smaller IR reflective spheres, such as the 6.4 mm described in this thesis, are not readily available in a sterilized version. While IR reflective spheres are sterilizable and have the potential for internal use, they are not certified for such applications yet. The 3D-printed IR markers should be sterilized before use. Secondly, it proposed to design multiple IR markers with curved surfaces to better adhere to the kidney surface. The flat IR markers in this thesis do not fully contact the kidney surface. Therefore, the IR markers are challenging to fixate correctly on the surface of the kidney.

Furthermore, the intraoperative adhesive glue that can be used to secure the IR marker and how this affects the clinical workflow should be examined. The IR-marker attachment during NSS should be removable after complete visualization or resection. The developed system can also be applied in other fields of pediatric surgical oncology, such as the clinical use in (osteo)sarcomas. It can also be utilized for other clinical treatments and diagnostics, such as AR-guided biopsies, where the trajectory of a tool can be visualized to guide the surgical instrument to the correct in vivo position, possibly enhancing the biopsy accuracy [84].

7.2 Technological improvements

Technological improvements to DINO-CLINIC are necessary to meet current clinical needs. Implementing a filter algorithm like a Kalman filter could help reduce noise and improve the stability of the hologram. Currently, the system relies on manual registration, which depends on the user. Integrating a registration technique, such as point registration, with IR tracking could potentially enhance the accuracy of the IR tracking system. However, as discussed in Chapter 3, point registration has several challenges, including difficulty identifying anatomical landmarks during surgery. Moreover, the current design of the pointer is large; however, the use of IR-markers and the placement of the markers further away from the tip of the tool could enhance the ability to pinpoint anatomical landmarks.

7.3 Future perspectives

Future research and development should focus on implementing an alternative registration system with the HL2 as a visualization device. Outside-in tracking could prove beneficial. Navigation techniques such as EM tracking show potential for improving registration and accuracy. The DINO-CLINIC system could be integrated with EM tracking, allowing registration and tracking with EM and 3D holographic visualization with HL2. A similar approach, as described by Thabit et al., is promising [85]. They use a QR code to register the HL2 to the EM system. However, the QR code could be replaced by an optical reference star and the DINO-CLINIC system, thereby improving registration accuracy and potentially enhancing detection in the OR. This method could also be suitable for kidney tracking, similar to the approach used for the liver, where an EM sensor can be attached to the kidney's surface.

DINO-CLINIC can track the kidney but does not account for organ deformability, which is present during NSS. The phantom study simulates surgery but does not accurately replicate the intraoperative situation due

to the lack of kidney deformability. Deformability remains an issue in further improving the accuracy of 3D navigation. Recent studies have shown that multiple markers are used to simulate deformation. Haouchine et al. demonstrated using a real-time biomechanical model of the liver and arteries to compute volumetric displacements based on liver surface motion [86]. Adagolodjo et al. combined marker-based tracking with deformation during open liver surgery [87]. A similar approach can be used with the DINO-CLINIC system. Also, in the field of partial nephrectomy, accurate volumetric deformation and robust navigation methods are explored [88].

As discussed, both systems exhibit marginal usability, according to the SUS. Therefore, discussing whether the HL2, as a standalone device, is sufficient for clinical purposes is relevant. While external trackers may enhance accuracy, the integration could introduce additional complexity, requiring more knowledge and system-specific process steps. Consequently, exploring alternative AR/MX devices, which could work as standalone devices, could yield better outcomes. Recent advancements in MX suggest promising developments and possibilities as more advanced sensors, processors, and graphic cards are introduced. Devices such as the Meta Quest 3 and Apple Vision Pro [89], [90], [91] may offer solutions for intraoperative visualization and automatic registration of 3D models, possibly enhancing the system's usability. Therefore, MX could provide more solutions similar to surgical navigation systems in laparoscopic guided surgery due to the availability of high-definition stereoscopic RGB cameras and light detection and ranging, also known as LiDAR, sensors. These methods mainly rely on point cloud registration methods, which still pose problems such as susceptibility to local minima and extended time required for registration [92]. However, multiple studies show promising results [93]. Zhang et al. demonstrated a markerless automatic deformable registration method for partial nephrectomy [94]. Their system used stereoscopic endoscopic images for 3D point cloud stitching and course to fine point cloud registration with preoperative models based on CT/MRI, which could be implemented on an MX device. Using deep learning algorithms for point cloud estimation could enhance registration accuracy and potentially overcome the limitations of standard point cloud registration algorithms [95]. Deep learning algorithms are widely used in multiple high-tech sectors, such as autonomous driving and robotics [96] [92]. The introduction of deep learning algorithms for registration in clinical applications is being explored [92], [97]. This technology holds immense potential for further advancements.

8
Conclusion

This thesis aimed to improve the AR system for surgical navigation during NSS for WT's. Moreover, we are the first to validate the use of AR intraoperatively during total nephrectomies, which showed a successful registration in two out of the six patients. Despite improving the depth perception and understanding of anatomical relationships, the five-point registration showed intraoperative limitations related to QR-code and one-time registration. To overcome the limitations of the current AR system, we developed and validated an inside-out IR-tracking AR application using the HL2, enabling continuous intraoperative kidney tracking. IR-tracking validation showed a high accuracy, with an accuracy of -0.70 mm for small IR spheres. Manual registration and fixation of a small IR-marker was challenging, which showed a TRE of 11.82 and a PLE of 8.74 mm. The instability of IR-tracking and the lack of a proper registration method results in low PLE and is, therefore, not yet clinically feasible (<5 mm).

The participating surgeons underscored that real-time tracking of the kidney and a holographic overlay are clinically useful during NSS despite considerable limitations in the stability of the hologram. Therefore, it can potentially improve the surgical clinical outcomes of pediatric patients with Wilms Tumors by supporting the surgeon with continuous 3D visualization of the tumor and resection boundaries. Secondly, the surgeons indicated that the system is potentially helpful for other clinical use cases supporting the broader utilization of IR tracking and AR within surgical pediatric oncology.

To conclude, inside-out IR-tracking with the HL2 has significant limitations but is potentially helpful for intraoperative surgical guidance during NSS and could improve the current AR system by replacing the QR code and implementing continuous tracking of the kidney. Further research is required to optimize the stability of the hologram during continuous tracking and implement alternative registration methods.

9
References

[1] J. S. Lee, T. R. Sanchez, and S. Wootton-Gorges, "Malignant renal tumors in children," *J. Kidney Cancer VHL*, vol. 2, no. 3, p. 84, May 2015.

[2] S. W. Leslie, H. Sajjad, and P. B. Murphy, "Wilms Tumor," *StatPearls*, May 2023.

[3] "Presentation, diagnosis, and staging of Wilms tumor - UpToDate." [Online]. Available: https://www.uptodate.com/contents/presentation-diagnosis-and-staging-of-wilms-tumor. [Accessed: 07-Mar-2023].

[4] "Treatment and prognosis of Wilms tumor - UpToDate." [Online]. Available: https://www.uptodate.com/contents/treatment-and-prognosis-of-wilms-tumor?sectionName=Anaplastic histology&topicRef=6249&anchor=H12305109&source=see_link#H12305109. [Accessed: 07-Mar-2023].

[5] S. W. Leslie, H. Sajjad, and P. B. Murphy, "Wilms Tumor," *StatPearls*, May 2023.

[6] E. C. Ruteshouser, S. M. Robinson, and V. Huff, "Wilms Tumor Genetics: Mutations in WT1, WTX, and CTNNB1 Account for Only About One-Third of Tumors," *Genes. Chromosomes Cancer*, vol. 47, no. 6, p. 461, Jun. 2008.

[7] B. C *et al.*, "Current Role of Imaging in the Management of Children with Wilms Tumor According to the New UMBRELLA Protocol," *Transl. Med.*, vol. 9, no. 1, pp. 1–16, May 2019.

[8] S. E. Servaes, F. A. Hoffer, E. A. Smith, and G. Khanna, "Imaging of Wilms tumor: an update," *Pediatr. Radiol.*, vol. 49, no. 11, pp. 1441–1452, Oct. 2019.

[9] J. S. Dome, E. J. Perlman, and N. Graf, "Risk Stratification for Wilms Tumor: Current Approach and Future Directions," *Am. Soc. Clin. Oncol. Educ. B.*, no. 34, pp. 215–223, May 2014.

[10] A. J. W. Millar, S. Cox, and A. Davidson, "Management of Bilateral Wilms Tumours," *Wilms Tumor*, pp. 61–76, Mar. 2016.

[11] A. Weirich *et al.*, "Survival in nephroblastoma treated according to the trial and study SIOP-9/GPOH with respect to relapse and morbidity," *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.*, vol. 15, no. 5, pp. 808–820, 2004.

[12] M. M. Van Den Heuvel-Eibrink *et al.*, "Position paper: Rationale for the treatment of Wilms tumour in the UMBRELLA SIOP-RTSG 2016 protocol," *Nat. Rev. Urol.*, vol. 14, no. 12, pp. 743–752, Dec. 2017.

[13] G. M. Vujanić *et al.*, "The UMBRELLA SIOP–RTSG 2016 Wilms tumour pathology and molecular biology protocol," *Nat. Rev. Urol. 2018 1511*, vol. 15, no. 11, pp. 693–701, Oct. 2018.

[14] J. Wang, M. Li, D. Tang, W. Gu, J. Mao, and Q. Shu, "Current treatment for Wilms tumor: COG and SIOP standards," *World J. Pediatr. Surg.*, vol. 2, no. 3, pp. 11–14, 2019.

[15] K. Milford, K. DeCotiis, and A. Lorenzo, "Wilms tumor: a review of current surgical controversies," *Transl. Androl. Urol.*, vol. 9, no. 5, pp. 2382392–2382392, Oct. 2020.

[16] A. J. Murphy and A. M. Davidoff, "Nephron-sparing surgery for Wilms tumor," *Front. Pediatr.*, vol. 11, p. 10, Jan. 2023.

[17] K. Taghavi, S. Sarnacki, T. Blanc, O. Boyer, and Y. Heloury, "The rationale for nephron-sparing surgery in unilateral non-syndromic Wilms tumour," *Pediatr. Nephrol.*, vol. 1, pp. 1–10, Aug. 2023.

[18] F. Von Haxthausen, C. Rüger, M. M. Sieren, R. Kloeckner, and F. Ernst, "Augmenting Image-Guided Procedures through In Situ Visualization of 3D Ultrasound via a Head-Mounted Display," pp. 1–16, 2023.

[19] H. R. Spiegl *et al.*, "Complications Following Nephron-Sparing Surgery for Wilms Tumor," *J. Pediatr. Surg.*, vol. 55, no. 1, pp. 126–129, Jan. 2020.

[20] J. C. H. Wilde *et al.*, "Nephron sparing surgery (NSS) for unilateral wilms tumor (UWT): the SIOP 2001 experience," *Pediatr. Blood Cancer*, vol. 61, no. 12, pp. 2175–2179, Dec. 2014.

[21] K. Milford, K. DeCotiis, and A. Lorenzo, "Wilms tumor: a review of current surgical controversies," *Transl. Androl. Urol.*, vol. 9, no. 5, p. 2382, Oct. 2020.

[22] R. N. W. Vanden Berg, E. N. Bierman, M. Van Noord, H. E. Rice, and J. C. Routh, "Nephron-sparing surgery for Wilms tumor: A systematic review," *Urol. Oncol.*, vol. 34, no. 1, pp. 24–32, Jan. 2016.

[23] D. A. Cozzi, S. Ceccanti, and F. Cozzi, "Renal function up to the 5th decade of life after nephrectomy in childhood: A literature review," *Nephrology (Carlton).*, vol. 23, no. 5, pp. 397–404, May 2018.

[24] A. Khondker, A. Jain, M. L. Groff, J. Brzezinski, A. J. Lorenzo, and M. Zappitelli, "Late Kidney Effects of Nephron-Sparing vs Radical Nephrectomy for Wilms Tumor: A Systematic Review and Meta-Analysis," *J. Urol.*, vol. 207, no. 3, pp. 513–523, Mar. 2022.

[25] J. M. Gleason, A. J. Lorenzo, P. R. Bowlin, and M. A. Koyle, "Innovations in the management of Wilms' tumor," *Ther. Adv. Urol.*, vol. 6, no. 4, pp. 165–176, 2014.

[26] A. Piana *et al.*, "Automatic 3D Augmented-Reality Robot-Assisted Partial Nephrectomy Using Machine Learning: Our Pioneer Experience," *Cancers 2024, Vol. 16, Page 1047*, vol. 16, no. 5, p. 1047, Mar. 2024.

[27] A. Piana *et al.*, "Automatic 3D Augmented-Reality Robot-Assisted Partial Nephrectomy Using Machine Learning: Our Pioneer Experience," *Cancers 2024, Vol. 16, Page 1047*, vol. 16, no. 5, p. 1047, Mar. 2024.

[28] C. Gsaxner *et al.*, "The HoloLens in medicine: A systematic review and taxonomy," *Med. Image Anal.*, vol. 85, 2023.

[29] D. Suresh, A. Aydin, S. James, K. Ahmed, and P. Dasgupta, "The Role of Augmented Reality in Surgical Training: A Systematic Review," *Surg. Innov.*, vol. 30, no. 3, p. 366, Jun. 2023.

[30] E. Bollen, L. Awad, B. Langridge, and P. E. M. Butler, "The intraoperative use of augmented and mixed reality technology to improve surgical outcomes: A systematic review," *Int. J. Med. Robot. Comput. Assist. Surg.*, vol. 18, no. 6, p. e2450, Dec. 2022.

[31] J. W. Yoon *et al.*, "Augmented reality for the surgeon: Systematic review," *Int. J. Med. Robot. Comput. Assist. Surg.*, vol. 14, no. 4, p. e1914, Aug. 2018.

[32] A. Palumbo, "Microsoft HoloLens 2 in Medical and Healthcare Context: State of the Art and Future Prospects," *Sensors (Basel).*, vol. 22, no. 20, Oct. 2022.

[33] L. Ma, T. Huang, J. Wang, and H. Liao, "Visualization, registration and tracking techniques for augmented reality guided surgery: a review," *Phys. Med. Biol.*, vol. 68, no. 4, 2023.

[34] L. M. Wellens *et al.*, "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*, vol. 2, no. 4, Apr. 2019.

[35] L. Ma, T. Huang, J. Wang, and H. Liao, "Visualization, registration and tracking techniques for augmented reality guided surgery: a review," *Phys. Med. Biol.*, vol. 68, no. 4, p. 04TR02, Feb. 2023.

[36] C. M. Andrews, A. B. Henry, I. M. Soriano, M. K. Southworth, and J. R. Silva, "Registration Techniques for Clinical Applications of Three-Dimensional Augmented Reality Devices," *IEEE J. Transl. Eng. Heal. Med.*, vol. 9, 2021.

[37] "HoloLens 2—Overview, Features, and Specs | Microsoft HoloLens." [Online]. Available: https://www.microsoft.com/en-us/hololens/hardware#document-experiences. [Accessed: 05-Aug-2024].

[38] "HoloLens 2 hardware | Microsoft Learn." [Online]. Available: https://learn.microsoft.com/en-us/hololens/hololens2-hardware. [Accessed: 17-Nov-2023].

[39] F. Van der Heijden, "Snt 1 – Point Registration," p. 14, 2020.

[40] K. G. P. Spijkerboer, M. Fitski, F. J. Siepel, C. P. van de Ven, and A. F. W. van der Steeg, "Augmented reality-guided localization of a chest wall tumor in a pediatric patient," *Eur. J. Cancer*, vol. 170, pp. 103–105, Jul. 2022.

[41] X. Zhang *et al.*, "A markerless automatic deformable registration framework for augmented reality navigation of laparoscopy partial nephrectomy," *Int. J. Comput. Assist. Radiol. Surg. 2019 148*, vol. 14, no. 8, pp. 1285–1294, Apr. 2019.

[42] M. Weber, D. Wild, J. Kleesiek, J. Egger, and C. Gsaxner, "Deep Learning-based Point Cloud Registration for Augmented Reality-guided Surgery," May 2024.

[43] R. van der Woude *et al.*, "Clinical Application and Further Development of Augmented Reality Guidance for the Surgical Localization of Pediatric Chest Wall Tumors," *J. Pediatr. Surg.*, vol. 59, no. 8, pp. 1549–1555, Aug. 2024.

[44] M. Weber, D. Wild, J. Kleesiek, J. Egger, and C. Gsaxner, "Deep Learning-based Point Cloud Registration for Augmented Reality-guided Surgery," May 2024.

[45] H. Alexander Jaeger, K. O'Donoghue, and P. Cantillon-Murphy, "Electromagnetic tracking in guided medical interventions," *2022 3rd URSI Atl. Asia Pacific Radio Sci. Meet. AT-AP-RASC 2022*, no. April, 2022.

[46] C. Kunz *et al.*, "Infrared marker tracking with the HoloLens for neurosurgical interventions," *Curr. Dir. Biomed. Eng.*, vol. 6, no. 1, May 2020.

[47] M. Doughty, N. R. Ghugre, and G. A. Wright, "Augmenting Performance: A Systematic Review of Optical See-Through Head-Mounted Displays in Surgery," *J. Imaging*, vol. 8, no. 7, Jul. 2022.

[48] J. M. van der Zee *et al.*, "Virtual Resection: A New Tool for Preparing for Nephron-Sparing Surgery in Wilms Tumor Patients," *Curr. Oncol.*, vol. 29, no. 2, pp. 777–784, Feb. 2022.

[49] D. Teber *et al.*, "Augmented reality: a new tool to improve surgical accuracy during laparoscopic partial nephrectomy? Preliminary in vitro and in vivo results," *Eur. Urol.*, vol. 56, no. 2, pp. 332–338, Aug. 2009.

[50] B. Fida, F. Cutolo, G. di Franco, M. Ferrari, and V. Ferrari, "Augmented reality in open surgery," *Updat. Surg. 2018 703*, vol. 70, no. 3, pp. 389–400, Jul. 2018.

[51] A. Ragnhildstveit *et al.*, "Intra-operative applications of augmented reality in glioma surgery: a systematic review," *Front. Surg.*, vol. 10, p. 1245851, Aug. 2023.

[52] R. Moreta-Martinez, D. García-Mato, M. García-Sevilla, R. Pérez-Mañanes, J. Calvo-Haro, and J. Pascau, "Augmented reality in computer-assisted interventions based on patient-specific 3D printed reference," *Healthc. Technol. Lett.*, vol. 5, no. 5, pp. 162–166, 2018.

[53] T. P. C. Van Doormaal, J. A. M. Van Doormaal, and T. Mensink, "Clinical Accuracy of Holographic Navigation Using Point-Based Registration on Augmented-Reality Glasses," *Oper. Neurosurg.*, vol. 17, no. 6, p. 588, Dec. 2019.

[54] Q. D. Eyck, "Towards holographic navigation for paediatric nephron-sparing Wilms tumour surgery : a feasibility study," 2021.

[55] "(PDF) SUS: A quick and dirty usability scale." [Online]. Available: https://www.researchgate.net/publication/228593520_SUS_A_quick_and_dirty_usability_scale. [Accessed: 26-Aug-2024].

[56] S. Amini and M. Kersten-Oertel, "Augmented reality mastectomy surgical planning prototype using the HoloLens template for healthcare technology letters," *Healthc. Technol. Lett.*, vol. 6, no. 6, p. 261, 2019.

[57] X. Zhao, H. Zhao, W. Zheng, A. Gohritz, Y. Shen, and W. Xu, "Clinical evaluation of augmented reality-based 3D navigation system for brachial plexus tumor surgery," *World J. Surg. Oncol.*, vol. 22, no. 1, p. 20, Dec. 2024.

[58] A. Thabit, M. Benmahdjoub, M. L. C. van Veelen, W. J. Niessen, E. B. Wolvius, and T. van Walsum, "Augmented reality navigation for minimally invasive craniosynostosis surgery: a phantom study," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 17, no. 8, p. 1453, Aug. 2022.

[59] T. Fick, J. A. M. van Doormaal, E. W. Hoving, P. W. A. Willems, and T. P. C. van Doormaal, "Current Accuracy of Augmented Reality Neuronavigation Systems: Systematic Review and Meta-Analysis," *World Neurosurg.*, vol. 146, pp. 179–188, Feb. 2021.

[60] A. Martin-Gomez *et al.*, "STTAR: Surgical Tool Tracking using Off-the-Shelf Augmented Reality Head-Mounted Displays," *IEEE Trans. Vis. Comput. Graph.*, vol. 30, no. 7, p. 3578, Jul. 2024.

[61] "HL2-DINO/DINO-DLL: A C++ HoloLens 2 DLL used for detecting and tracking infrared markers with the Research Mode API." [Online]. Available: https://github.com/HL2-DINO/DINO-DLL. [Accessed: 06-Mar-2024].

[62] H. Iqbal and F. Rodriguez Y Baena, "Semi-Automatic Infrared Calibration for Augmented Reality Systems in Surgery," *IEEE Int. Conf. Intell. Robot. Syst.*, vol. 2022-October, pp. 4957–4964, 2022.

[63] "OpenCV - Open Computer Vision Library." [Online]. Available: https://opencv.org/. [Accessed: 05-Aug-2024].

[64] M. Sylos Labini, C. Gsaxner, A. Pepe, J. Wallner, J. Egger, and V. Bevilacqua, "Depth-Awareness in a System for Mixed-Reality Aided Surgical Procedures," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11645 LNAI, pp. 716–726, 2019.

[65] M. Sylos Labini, C. Gsaxner, A. Pepe, J. Wallner, J. Egger, and V. Bevilacqua, "Depth-Awareness in a System for Mixed-Reality Aided Surgical Procedures," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11645 LNAI, pp. 716–726, 2019.

[66] S. L. Brunton and J. N. Kutz, "Singular Value Decomposition (SVD)," *Data-Driven Sci. Eng.*, pp. 3–46, Feb. 2019.

[67] M. Fitski, J. W. Meulstee, A. S. Littooij, C. P. Van De Ven, A. F. W. Van Der Steeg, and M. 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," *J. Healthc. Eng.*, vol. 2020, no. 1, p. 8899049, Jan. 2020.

[68] "Mixed Reality Toolkit 3 Developer Documentation - MRTK3 | Microsoft Learn." [Online]. Available: https://learn.microsoft.com/en-us/windows/mixed-reality/mrtk-unity/mrtk3-overview/. [Accessed: 05-Aug-2024].

[69] H. Iqbal and F. Rodriguez Y Baena, "Semi-Automatic Infrared Calibration for Augmented Reality Systems in Surgery," *IEEE Int. Conf. Intell. Robot. Syst.*, vol. 2022-Octob, pp. 4957–4964, 2022.

[70] B. A, U. A, S. T, M. A, and S. JH, "Design and validation of an open-source library of dynamic reference frames for research and education in optical tracking," *J. Med. imaging (Bellingham, Wash.)*, vol. 5, no. 2, p. 1, Feb. 2018.

[71] J. B. Hummel *et al.*, "Design and application of an assessment protocol for electromagnetic tracking systems.," *Med. Phys.*, vol. 32, no. 7Part1, pp. 2371–2379, Jul. 2005.

[72] "Bambu Lab X1 Series | Desktop 3D Printer | X1-Carbon - Bambu Lab." [Online]. Available: https://bambulab.com/en-eu/x1. [Accessed: 07-Aug-2024].

[73] J. Lu, L. Zhang, and R. Maeda, "Real-time tracking of organ-shape and vessel-locations for surgical navigation using MEMS tri-axis magnetic sensors," *Med. Eng. Phys.*, vol. 93, pp. 42–48, Jul. 2021.

[74] J. N. Smit *et al.*, "Ultrasound-based navigation for open liver surgery using active liver tracking," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 17, no. 10, pp. 1765–1773, Oct. 2022.

[75] M. Markert, A. Koschany, and T. Lueth, "Tracking of the liver for navigation in open surgery," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 5, no. 3, pp. 229–235, Aug. 2010.

[76] Y. Adagolodjo, N. Golse, E. Vibert, M. de Mathelin, S. Cotin, and H. Courtecuisse, "Marker-based Registration for Large Deformations -Application to Open Liver Surgery," May 2018.

[77] A. Fedorov *et al.*, "3D Slicer as an image computing platform for the Quantitative Imaging Network," *Magn. Reson. Imaging*, vol. 30, no. 9, pp. 1323–1341, Nov. 2012.

[78] "3D Slicer image computing platform | 3D Slicer." [Online]. Available: https://www.slicer.org/. [Accessed: 07-Aug-2024].

[79] J. M. Fitzpatrick and J. B. West, "The distribution of target registration error in rigid-body point-based registration," *IEEE Trans. Med. Imaging*, vol. 20, no. 9, pp. 917–927, Sep. 2001.

[80] C. Schneider *et al.*, "Comparison of manual and semi-automatic registration in augmented reality image-guided liver surgery: a clinical feasibility study," *Surg. Endosc.*, vol. 34, no. 10, pp. 4702–4711, Oct. 2020.

[81] J. Fotouhi *et al.*, "Interactive Flying Frustums (IFFs): spatially aware surgical data visualization," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 14, no. 6, p. 913, Jun. 2019.

[82] Q. Li, R. Li, K. Ji, and W. Dai, "Kalman filter and its application," *Proc. - 8th Int. Conf. Intell. Networks Intell. Syst. ICINIS 2015*, pp. 74–77, Aug. 2016.

[83] "Optical Navigation Markers - NDI." [Online]. Available: https://www.ndigital.com/optical-navigation-technology/navigation-markers/. [Accessed: 26-Aug-2024].

[84] H. Li *et al.*, "Navigate biopsy with ultrasound under augmented reality device: Towards higher system performance," *Comput. Biol. Med.*, vol. 174, p. 108453, May 2024.

[85] A. Thabit, M. Benmahdjoub, M. L. C. van Veelen, W. J. Niessen, E. B. Wolvius, and T. van Walsum, "Augmented reality navigation for minimally invasive craniosynostosis surgery: a phantom study," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 17, no. 8, pp. 1453–1460, Aug. 2022.

[86] N. Haouchine, J. Dequidt, I. Peterlik, E. Kerrien, M. O. Berger, and S. Cotin, "Image-guided simulation of heterogeneous tissue deformation for augmented reality during hepatic surgery," *2013 IEEE Int. Symp. Mix. Augment. Reality, ISMAR 2013*, pp. 199–208, 2013.

[87] Y. Adagolodjo, N. Golse, E. Vibert, M. De Mathelin, S. Cotin, and H. Courtecuisse, "Marker-Based Registration for Large Deformations - Application to Open Liver Surgery," *Proc. - IEEE Int. Conf. Robot. Autom.*, pp. 4007–4012, 2018.

[88] X. Zhang, T. Wang, X. Zhang, Y. Zhang, and J. Wang, "Assessment and application of the coherent point drift algorithm to augmented reality surgical navigation for laparoscopic partial nephrectomy," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 15, no. 6, pp. 989–999, Jun. 2020.

[89] "Meta Quest 3: nieuwe mixed reality VR-headset - Shoppen | Meta Store | Meta Store." [Online]. Available: https://www.meta.com/nl/quest/quest-3/. [Accessed: 27-Aug-2024].

[90] "Apple Vision Pro - Apple." [Online]. Available: https://www.apple.com/apple-vision-pro/.

[Accessed: 27-Aug-2024].

[91] E. Waisberg *et al.*, "Apple Vision Pro: the future of surgery with advances in virtual and augmented reality," *Irish Journal of Medical Science*, vol. 193, no. 1. Springer Science and Business Media Deutschland GmbH, pp. 345–346, 01-Feb-2024.

[92] M. Weber, D. Wild, J. Kleesiek, J. Egger, and C. Gsaxner, "Deep Learning-based Point Cloud Registration for Augmented Reality-guided Surgery," May 2024.

[93] M. A. Audette, F. P. Ferrie, and T. M. Peters, "An algorithmic overview of surface registration techniques for medical imaging," *Med. Image Anal.*, vol. 4, no. 3, pp. 201–217, Sep. 2000.

[94] X. Zhang *et al.*, "A markerless automatic deformable registration framework for augmented reality navigation of laparoscopy partial nephrectomy," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 14, no. 8, pp. 1285–1294, Aug. 2019.

[95] L. Chen, C. Feng, Y. Ma, Y. Zhao, and C. Wang, "A review of rigid point cloud registration based on deep learning," *Front. Neurorobot.*, vol. 17, p. 1281332, Jan. 2023.

[96] G. M. Netto and M. M. Oliveira, "Robust point-cloud registration based on dense point matching and probabilistic modeling," *Vis. Comput.*, vol. 38, no. 9–10, pp. 3217–3230, Sep. 2022.

[97] Y. Adagolodjo, R. Trivisonne, N. Haouchine, S. Cotin, and H. Courtecuisse, "Silhouette-based pose estimation for deformable organs application to surgical augmented reality," *IEEE Int. Conf. Intell. Robot. Syst.*, vol. 2017-September, pp. 539–544, Dec. 2017.

Appendix A: Dimensions of IR-markers

Table 8: Dimensions of all designed IR-markers

Name	**No. IR**	**Diameter**	**Dimensions [X, Y, Z]**			
			Point 1	Point 2	Point 3	Point 4
3_IR_ExtraSmall	3	6.4	[-2.81, -15.36, 0]	[2.81, 11.03, 0]	[5.63, 4.32, 0]	x
3_IR_Small	3	6.4	[-4.21, -23.04, 0]	[-4.22, 16.55, 0]	[8.43, 6.48, 0]	x
3_IR_Medium	3	6.4 / 12.7	[-5.62, -30.72, 0]	[-5.63, 22.08, 0]	[11.25, 8.65, 0]	x
3_IR_Large	3	6.4 / 12.7	[-7.03, -38.41, 0]	[-7.04, 27.59, 0]	[14.06, 10.81, 0]	x
4_IR_ExtraSmall	4	6.4	[1.60, -13.54, 0]	[-11.6, 0.68, 0]	[1.60, 12.86, 0]	[10.00, 0.68, 0]
4_IR_Small	4	6.4	[2.40, -20.31, 0]	[-17.4, 1.02, 0]	[2.4, 19.29, 0]	[15.00, 1.02, 0]
4_IR_Medium	4	6.4 / 12.7	[3.2, -27.08, 0]	[-23.20, 1.36, 0]	[3.2, 25.72, 0]	[20.00, 1.36, 0]
4_IR_Large	4	6.4 / 12.7	[4.00, -33.85, 0]	[-29.00, 1.70, 0]	[4.00, 32.15, 0]	[25.00, 1.70, 0]
4_IR_Reference	4	12.7	[6.07, -40.83, 0]	[-34.17, 4.96, 0]	[17.43, 35.96, 0]	[23.59, 4.96, 0]

Appendix B: Outline initial steps

The following steps were taken in the initial phase to develop a own IR-tracking algorithm. The following steps are integrated with the ResearchModeVideoStream.cs from petergu684 (https://github.com/petergu684/HoloLens2-ResearchMode-Unity).

1. **Create simpleBlobDetector:**

```
using OpenCVForUnity.CoreModule;
using OpenCVForUnity.ImgprocModule;
using OpenCVForUnity.Features2dModule;
using OpenCVForUnity.UnityUtils;

private MatOfKeyPoint keypoints;
private SimpleBlobDetector blobDetector;

SimpleBlobDetector_Params param = new SimpleBlobDetector_Params();
param.set_thresholdStep(1.0f);
param.set_minThreshold(252.0f);
param.set_maxThreshold(255.0f);
param.set_filterByColor(false);
param.set_filterByArea(true);
param.set_minArea(1);
param.set_maxArea(8000);
param.set_filterByCircularity(true);
param.set_minCircularity(0.75f);
param.set_maxCircularity(100000);
param.set_filterByInertia(true);
param.set_minInertiaRatio(0.75f);
param.set_maxInertiaRatio(100000);
param.set_filterByConvexity(true);
param.set_minConvexity(0.75f);
param.set_maxConvexity(100000);

blobDetector = SimpleBlobDetector.create(param);
```

2. **Create function to detect blob on STD frame and correlate 2D points to depth frame.**

```
#if ENABLE_WINMD_SUPPORT
private Mat DetectBlobs(byte[] binaryFrame)
  {
    // Convert binary frame to Mat
    Mat binaryMat = new Mat(512, 512, CvType.CV_8UC1);
    resultDisplayTexture.LoadRawTextureData(binaryFrame);
    Utils.texture2DToMat(resultDisplayTexture, binaryMat);

    // Detect blobs
    keypoints = new MatOfKeyPoint();
    blobDetector.detect(binaryMat, keypoints);

    Debug.Log("Number of keypoints: " + keypoints.toList().Count);

    // Count Keypoints and visualize coordinates

    KeyPoint[] keypointArray = keypoints.toArray();
    for (int i = 0; i < keypointArray.Length; i++)
    {
      // Get the pixel value from depthFrameData using blob center coordinates
      int centerX = (int)keypointArray[i].pt.x;
      int centerY = (int)keypointArray[i].pt.y;

      if (centerX >= 0 && centerX < 512 && centerY >= 0 && centerY < 512)
      {
        // Assuming depthFrameData is a linear array of bytes, calculate the index
        int index = centerY * 512 + centerX;

        // Get the pixel value from depthFrameData at the calculated index
        byte pixelValue = depthFrameData[index];
```

```
        Debug.Log("Keypoint " + i + ": X =" + keypointArray[i].pt.x + ", Y =" + keypointArray[i].pt.y +
",Z =" + pixelValue);
      }
    }
    Mat outImgMat = new Mat();

    // Draw keypoints on the outImgMat
    Features2d.drawKeypoints(binaryMat, keypoints, outImgMat, new Scalar(0, 0, 0));

    return outImgMat;
  }

#endif
```

3. Visualize detected blobs in HoloLens 2

```
shortAbImageMediaTexture.LoadRawTextureData(shortAbImageFrameData);
        shortAbImageMediaTexture.Apply();

        // Detect keypoints [j,i]
        Mat outImgMat = DetectBlobs(shortAbImageFrameData);

        // MaptoImageUnitPlane



        if (resultDisplayPlane != null)
        {
          resultDisplayMaterial = resultDisplayPlane.GetComponent<MeshRenderer>().material;
          Utils.matToTexture2D(outImgMat, resultDisplayTexture);
          resultDisplayTexture.Apply();
        }
```

Appendix C: Adaptations DINO-DLL

C.1 Threshold

Orginial code from: https://github.com/HL2-DINO/DINO-DLL. Adaptations are visualized in red.

```
void DetectBlobs2DBasic(cv::Mat& processed_image, std::vector<cv::Point2f>& outPixelLocations)
 {
  PROFILE_BLOCK(DetectBlobsBasic);
  if (outPixelLocations.size() > 0) outPixelLocations.clear();

  // binarisation to speed up contour detect
  cv::threshold(processed_image, processed_image, BINARY_THRESH_8BIT, 255, cv::THRESH_BINARY);

  // external contours of interest in this scenario
  std::vector<std::vector<cv::Point>> contours;
  cv::findContours(processed_image, contours, cv::RETR_EXTERNAL, cv::CHAIN_APPROX_SIMPLE);

  double area, perimeter, circ;
  float blobX, blobY;

  for (const std::vector<cv::Point> contour : contours)
  {
   area = cv::contourArea(contour);
   perimeter = cv::arcLength(contour, true);

   /* filter by area. max: 1/16th of total image */
   if (area < 1 || area > 16384) continue;

   /* see https://en.wikipedia.org/wiki/Roundness */
   /* for definition of roundedness formula used */
   circ = (4 * PI * area) / (perimeter * perimeter);
   if (circ < 0.7f) continue; // filter by circularity

   // see https://en.wikipedia.org/wiki/Image_moment
   // for defining a centroid from an image moment
   auto M = cv::moments(contour);
   blobX = M.m10 / M.m00;
   blobY = M.m01 / M.m00;
   outPixelLocations.emplace_back(blobX, blobY);
  }
 }
```

C.2 Compensation sphere markers

Orginal code from: https://github.com/HL2-DINO/DINO-DLL. Adaptations are visualized in red.

```
pointInDepth.normalize(); // turn it into a unit vector

   pointInDepth *= (static_cast<double>(depthVal+3.2) / 1000.0); // convert into metres

   pointInWorld = transform * pointInDepth.homogeneous();

pointInDepth.normalize(); // turn it into a unit vector

   pointInDepth *= (static_cast<double>(depthVal+6.3) / 1000.0); // convert into metres

   pointInWorld = transform * pointInDepth.homogeneous();
```

C.3 Save transformations

A developed algorithm to save transformation matrices of detected tools for 30 seconds in a text file and save it locally on the HoloLens 2.

```
private List<(float, Dictionary<int, ToolTrackingUtils.TrackedTool>)> recordedToolDictionaries = new
List<(float, Dictionary<int, ToolTrackingUtils.TrackedTool>)>();

#if ENABLE_WINMD_SUPPORT
 async
#endif

 public void FetchTransform()
 {
#if ENABLE_WINMD_SUPPORT
 if (ToolManagerScript == null) return;

 // Clear the recordedToolDictionaries list before starting a new recording
 recordedToolDictionaries.Clear();

 Thread.Sleep(5000);

 // Start a coroutine to record transform data for 5 seconds
 StartCoroutine(RecordTransformDataForDuration(30f));
#endif
 }

private IEnumerator RecordTransformDataForDuration(float duration)
 {
#if ENABLE_WINMD_SUPPORT
  float elapsedTime = 0f;

  while (elapsedTime < duration)
  {
   // Get the ToolDictionary
   IReadOnlyDictionary<int, ToolTrackingUtils.TrackedTool> transformToPrint =
ToolManagerScript.GetToolDictionary();

   // Create a new dictionary to store the copy of transformToPrint
   Dictionary<int, ToolTrackingUtils.TrackedTool> transformCopy = new Dictionary<int,
ToolTrackingUtils.TrackedTool>();

   // Deep copy each key-value pair from transformToPrint
   foreach (var pair in transformToPrint)
   {
    // Create a new TrackedTool instance with only Tool_HoloFrame_LH copied
    var copiedTool = new ToolTrackingUtils.TrackedTool();
    copiedTool.Tool_HoloFrame_LH = pair.Value.Tool_HoloFrame_LH;

    // Add the copied instance to the transformCopy dictionary
    transformCopy.Add(pair.Key, copiedTool);
   }

   // Copy
   recordedToolDictionaries.Add((elapsedTime, transformCopy));

   string debugInfo = $"Elapsed Time: {elapsedTime.ToString("F2")} / {duration}\n";
   ConsoleDebugTextMesh.text = debugInfo;

   // Increment the elapsed time
   elapsedTime += time.deltaTime;

   // Wait for the next frame
   yield return null;
  }

  // Once recording is complete, save the recorded transform data
  SaveRecordedToolDictionary();

 }

 private void SaveRecordedToolDictionary()
 {
  // Combine all recorded ToolDictionaries into a single string
  string allToolDictionaries = "";

  foreach (var toolTuple in recordedToolDictionaries)
  {
   foreach (var pair in toolTuple.Item2)
```

```
    {
    allToolDictionaries += $"Elapsed Time: {toolTuple.Item1.ToString("F2")}\n";
    allToolDictionaries += $"Key: {pair.Key}\n";
    allToolDictionaries += $"{pair.Value.Tool_HoloFrame_LH.ToString("F3")}\n";
    }
   }

   // Save all recorded ToolDictionaries to a text file using SaveProfilerString
   SaveProfilerString(allToolDictionaries, $"DINO-
AR_AllToolDictionary_Unity_v2{Application.unityVersion}.txt");

   // Clear the recorded ToolDictionaries list for the next recording
   recordedToolDictionaries.Clear();
#else
   yield break;
#endif }
```

Appendix D: Post-processing MATLAB

```
%% Nick de Groot | s1914472 | M3 | PMC | Post processing IR tracking
clear all
close all
clc
disp('Nick de Groot | s1914472 | M3 | PMC | Data Analysis')
disp("LOAD DATA MANUALLY NOW --> Run to End")
%%   Manual load xlsx data
% load data manually
D = Dynamic12;
tic
%%   Pre-process: Delete NaN values + separate transformations of tools
D = table2array(D);

% Delete NaN
D = rmmissing(D);
D_size = size(D,1);

% Split tables (Tool 1 and Tool 2)
Tool1 = zeros(D_size/2,4);
Tool2 = zeros(D_size/2,4);

% Select tool 1 transforms
for i = 1:D_size/8
  x = 1;
  y = 4;
  z = x + 8 * i-8;
  p = y + 8 * i-8;
  Tool1((i-1)*4+1:i*4,:) = D(z:p,:);
end

% Select tool 2 transforms
for i = 1:D_size/8
  x = 5;
  y = 8;
  z = x + 8 * (i-1);
  p = y + 8 * (i-1);
  Tool2((i-1)*4+1:i*4,:) = D(z:p,:);
end
clear x y z p i D_size

%%   Pre-process: Separate translation and rotation matrices
Trans_1 = Tool1(:,4);
Trans_2 = Tool2(:,4);
Trans_xyz_1 = Tool1(:,4);
Trans_xyz_2 = Tool2(:,4);
Trans_xyz_1(4:4:end,:) = [];
Trans_xyz_2(4:4:end,:) = [];
Rot_1 = Tool1(:,1:3);
Rot_2 = Tool2(:,1:3);

%%   Pre-process: XYZ seperate values
Trans_x_1 = Trans_xyz_1(1:3:end);              % Select all x-axis values
Trans_y_1 = Trans_xyz_1(2:3:end);              % Select all y-axis values
Trans_z_1 = Trans_xyz_1(3:3:end);
Trans_x_2 = Trans_xyz_2(1:3:end);              % Select all x-axis values
Trans_y_2 = Trans_xyz_2(2:3:end);              % Select all y-axis values
Trans_z_2 = Trans_xyz_2(3:3:end);

Trans_1 = horzcat(Trans_x_1,Trans_y_1,Trans_z_1);
Trans_2 = horzcat(Trans_x_2,Trans_y_2,Trans_z_2);



%%   Pre-process: XYZ filter out invalid values
Trans_1_2 = horzcat(Trans_1, Trans_2);
Trans_1_2 = Trans_1_2(all(Trans_1_2,2),:);
```

```
Trans_1 = Trans_1_2(:,1:3);
Trans_2 = Trans_1_2(:,4:6);

%%   Pre-process: Rotation delete scaling
Rot_1( ~any(Rot_1,2), : ) = []; %rows
Rot_2( ~any(Rot_2, 2),: ) = [];

%%   Pre-process: Rotation convert to 3D matrix
A = size(Rot_1,1)/3;

for i = 1:A
Rot_1_3D(:,:,i) = Rot_1(-1+(2*i):1+(2*i),:);
end

for i = 1:A
Rot_2_3D(:,:,i) = Rot_2(-1+(2*i):1+(2*i),:);
end

%%   Pre-process: Rotation to Euler Angles
eul_angle_1 = rotm2eul(Rot_1_3D);
eul_angle_2 = rotm2eul(Rot_2_3D);

%%   Metric: Calculate euclidean distance
d = sqrt((Trans_2(:,1)-Trans_1(:,1)).^2 + (Trans_2(:,2)- ...
  Trans_1(:,2)).^2 + (Trans_2(:,3)-Trans_1(:,3)).^2);     % Calculate euclidean distance (m)

d_mm = d * 1000;                        % Convert to mm
d_real = 200;                          % Define real distance between two geometric shapes = 20cm -->
200mm
euclidean = -1*(d_mm - d_real);                % Calculate difference between measured and real
distance

%%   Metric: Calculate X Y Z difference
diff_x_tool = (Trans_2(:,1) - Trans_1(:,1))*1000;
diff_y_tool = (Trans_2(:,2) - Trans_1(:,2))*1000;
diff_z_tool = (Trans_2(:,3) - Trans_1(:,3))*1000;
diff_x_tool = diff_x_tool - 200;
diff_xyz = horzcat(diff_x_tool,diff_y_tool,diff_z_tool);
clear diff_x_tool diff_y_tool diff_z_tool

%%   Metric: Mean, Median, STD euclidean distance
Euc_Mean = mean(euclidean)
Euc_Median = median(euclidean)
Euc_STD = std(euclidean)
Euc_Max = max(euclidean)
Euc_Min = min(euclidean)

%%   Metric: Euler-angle mean
Euler_xyz_mean_1 = mean(eul_angle_1,1)
Euler_xyz_mean_2 = mean(eul_angle_2,1)
```

Appendix E: Clinical feasibility questionnaires

E.1 System Usability Scale (SUS)

Table 9: System Usability Scale questionnaire

No.	Question	Strongly Disagree				Strongly Agree
1.1	I think that I would like to use this system frequently.	1	2	3	4	5
1.2	I found the system unnecessarily complex.	1	2	3	4	5
1.3	The system was easy to use.	1	2	3	4	5
1.4	I think that I would need the support of a technical person to be able to use this system.	1	2	3	4	5
1.5	I found the various functions in the product were well integrated	1	2	3	4	5
1.6	I thought there was too much inconsistency in this system	1	2	3	4	5
1.7	I would imagine that most people would learn to use this system very quickly.	1	2	3	4	5
1.8	I found the system very cumbersome to use.	1	2	3	4	5
1.9	I felt very confident using the system.	1	2	3	4	5
1.10	I needed to learn a lot of things before I could get going with this system.	1	2	3	4	5

Comments:

E.2 Self developed questionnaire

Table 10: Self developed questionnaire for the validation of IR-tracking

No.	Question	Strongly Disagree				Strongly Agree
2.1	Determining the position of the IR-marker on the kidney of the patient was straightforward.	1	2	3	4	5
2.2	The user interface of the HoloLens software was straightforward to use.	1	2	3	4	5
2.2.1	*The hand-palm user interface was straightforward to use.*	1	2	3	4	5
2.2.2	*The integrated slider was helpful to visualize the kidney correctly.*	1	2	3	4	5
2.3	The hologram remained stable on the kidney after positioning.	1	2	3	4	5
2.4	Real-time tracking of the kidney is useful	1	2	3	4	5
2.4.1	*The infrared-marker tracking is sufficient and fast enough for the **small** infrared-markers.*	1	2	3	4	5
2.4.2	*The infrared marker tracking is sufficient and fast enough for the **large** infrared-markers.*	1	2	3	4	5
2.5	The holographic overlay is worth the additional time for NSS.	1	2	3	4	5
2.6	The system has the potential to be used for other clinical use cases	1	2	3	4	5

Comments:

Appendix F: Results phase one (12.7 mm)

Table 11: All results of IR-tracking accuracy of phase one (12.7mm)

Size	Number of IR markers	Angle	Distance (cm)	Mean ± std (mm)	Median (mm)	Max (mm)	Min (mm)	N
M	3	45	30	0.09 ± 0.04	0.10	0.19	-0.01	470
			60	0.51 ± 0.25	0.58	0.63	-0.48	483
			90	0.35 ± 0.05	0.36	0.44	0.19	447
		90	30	0.48 ± 0.31	0.58	0.19	-1.00	487
			60	1.13 ± 0.52	1.09	1.27	0.99	474
			90	-1.21 ± 0.44	-1.46	-0.06	-2.25	348
L	3	45	30	-0.58 ± 0.27	-0.63	0.38	-0.79	469
			60	0.06 ± 0.21	-0.01	0.98	-0.17	477
			90	0.68 ± 0.58	0.51	1.52	-0.79	479
		90	30	-0.41 ± 0.49	-0.81	0.18	-0.10	487
			60	-0.18 ± 0.52	-0.56	1.29	0.83	465
			90	-0.03 ± 0.50	-0.11	1.16	-1.25	471
M	4	45	30	0.51 ± 0.45	0.27	1.66	-0.30	374
			60	-0.17 ± 0.21	-0.28	0.69	-1.23	444
			90	-0.84 ± 0.39	-0.86	0.42	-1.54	441
		90	30	-1.14 ± 0.31	-1.20	-0.26	-1.55	443
			60	-0.01 ± 0.18	0.14	0.75	-0.19	403
			90	0.04 ± 0.40	-0.19	1.08	-1.14	440
L	4	45	30	-1.10 ± 0.60	-0.70	0.40	-1.78	603
			60	0.45 ± 0.34	0.58	0.69	-0.48	595
			90	-0.62 ± 0.05	-0.60	-0.53	-0.70	509
		90	30	-0.55 ± 0.12	-0.56	-0.42	-1.56	617
			60	0.60 ± 0.03	0.58	0.69	0.51	486
			90	-0.03 ± 0.40	0.19	1.18	-0.90	492

Appendix G: Results phase two (6.4 mm)

Table 12: All results of IR-tracking accuracy of phase two (6.4mm)

Size	Number of IR markers	Distance (cm)	Angle (°)	Mean ± std (mm)	Median (mm)	Max (mm)	Min (mm)	N
XS	3	60	22.5	-1.54 ± 0.55	-1.47	-0.51	-2.72	131
			45	-0.60 ± 0.29	-0.51	0.13	-1.77	474
			67.5	-0.95 ± 0.30	-0.81	0.13	-2.08	475
			90	-0.52 ± 0.24	-0.50	0.45	-1.46	467
S	3	60	22.5	-0.69 ± 0.31	-0.73	0.18	-1.62	267
			45	-1.01 ± 0.48	-1.08	-0.07	-1.91	260
			67.5	-1.12 ± 0.29	-0.97	-0.47	-1.88	483
			90	-0.42 ± 0.34	-0.24	0.67	-1.56	306
M	3	60	22.5	-0.74 ± 0.63	-1.01	0.99	-2.12	298
			45	-1.42 ± 0.52	-1.04	-0.02	-2.05	330
			67.5	-1.01 ± 0.18	-1.01	-0.01	-2.01	483
			90	-0.33 ± 0.22	-0.41	1.71	-1.24	472
XS	4	60	22.5	-1.34 ± 0.57	-1.16	-0.09	-2.25	132
			45	-1.22 ± 0.36	-1.18	-0.16	-2.20	98
			67.5	-1.31 ± 0.52	-1.08	-0.06	-2.10	438
			90	-0.14 ± 0.02	-0.16	-0.09	-0.20	456
S	4	60	22.5	-1.56 ± 0.45	-1.37	-0.04	-2.69	196
			45	-0.63 ± 0.36	-0.54	0.37	-1.87	503
			67.5	-0.43 ±0.42	-0.46	0.45	-1.79	172
			90	-0.36 ± 0.37	-0.53	0.88	-0.94	415
M	4	60	22.5	-0.50 ± 0.45	-0.69	0.54	-0.97	458
			45	-0.78 ± 0.58	-0.45	0.66	-2.44	315
			67.5	-1.18 ± 0.60	-1.22	-0.13	-2.33	365
			90	-0.02 ± 0.37	0.99	1.00	-0.02	274

Appendix H: Results phantom experiment

H.1 Results manual registration

Table 13: Results (translation and rotation) of manual registration

	Surgeon 1		Surgeon 2		Surgeon 3	
	Initial	**HL2**	**Initial**	**HL2**	**Initial**	**HL2**
Roll (°)	-31.01	-26.26	-0.44	-0.00	-5.44	-4.02
Pitch (°)	-2.37	-10.75	-2.48	-12.44	5.30	-2.95
Yaw (°)	55.38	55.94	1.70	4.47	6.54	5.52
X (mm)	14.77	12.22	-1.06	6.75	-1.77	0.61
Y (mm)	25.95	1.33	1.39	4.53	-0.45	3.46
Z (mm)	12.43	9.309	1.92	4.31	1.50	2.68

H.2 Results TRE and PLE

Table 14: Results TRE (all placements) and PLE

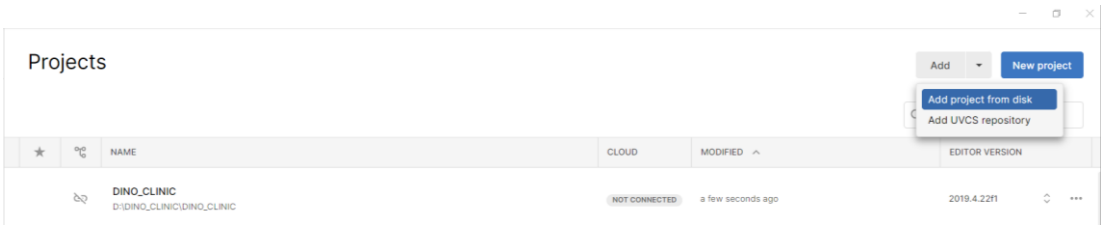
Target	**Surgeon 1**			**Surgeon 2**			**Surgeon 3**		
	Initial (TRE)	With HL2 (TRE)	**PLE**	Initial (TRE)	With HL2 (TRE)	**PLE**	Initial (TRE)	With HL2 (TRE)	**PLE**
Target 1	59.46	19.27	**9.96**	4.66	12.5	**10.85**	3.6	6.97	**6.38**
Target 2	26.32	13.19	**11.91**	1.68	9.66	**8.96**	5.65	5.64	**3.48**
Target 3	28.06	22.93	**11.54**	2.66	9.32	**13.2**	11.21	9.59	**2.78**
Target 4	18.3	28.31	**5.27**	4.35	12.13	**8.74**	7.24	9.7	**6.99**
Target 5	28.48	17.01	**8.69**	4.13	11.82	**10.35**	4.17	7.87	**8.7**

Appendix I: DINO protocol

This protocol highlights the steps to employ the DINO-CLINIC system on the HoloLens 2.

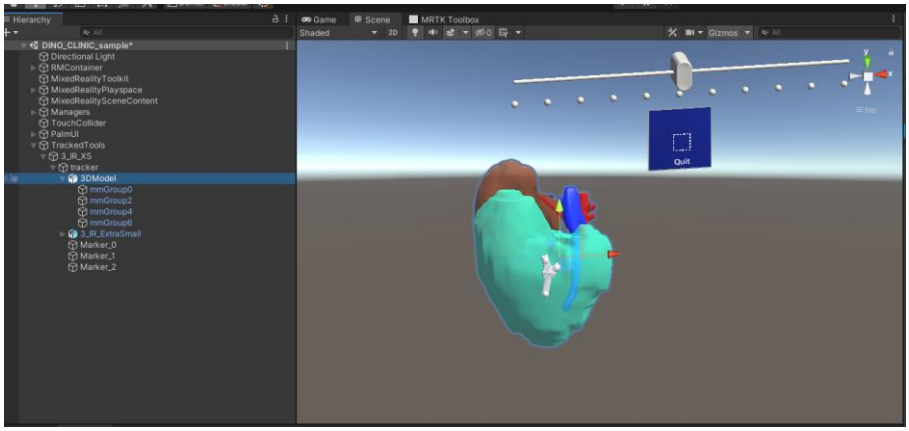
DINO-CLINIC installation

1. Install the required software
    a. Unity Hub
    b. Unity 2019.4.22f1
    c. Visual Studio 2019
2. Import the 'DINO-CLINIC' file in Unity Hub (Add → Add project from Disk



DINO-CLINIC

3. Open the 'DINO_CLINIC_sample' scene (Assets → Scences → Cases → DINO_CLINIC_sample.unity).
4. Import patient-specific 3D model (.obj) (Assets → Models → Patients).
5. Translate/rotate 3D model for preoperative planning of the position of the IR-marker.
    a. Shortcut: W = translate, E = rotate.

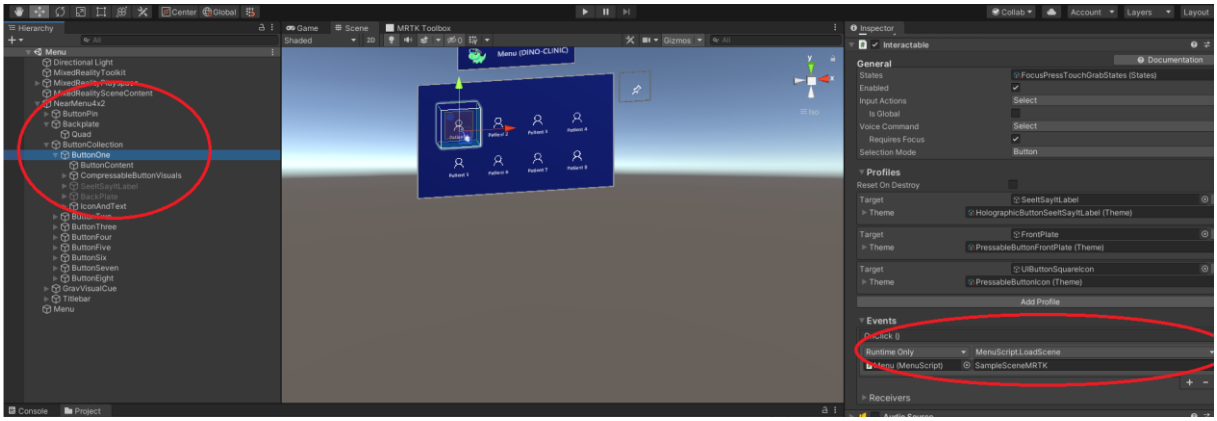


6. Colour 3D model
    a. In the hierarchy window, unfold 3D model of patients.
    b. In Assets → Materials, drop material to group.
7. Open Menu Scene (Assets → Scenes -> Menu.Unity).
8. Connect Menu button to correct the patient's scene

a. In Hierarchy, unfold NearMenu4x2 → ButtonCollection.

b. Select a button (1 to 8).

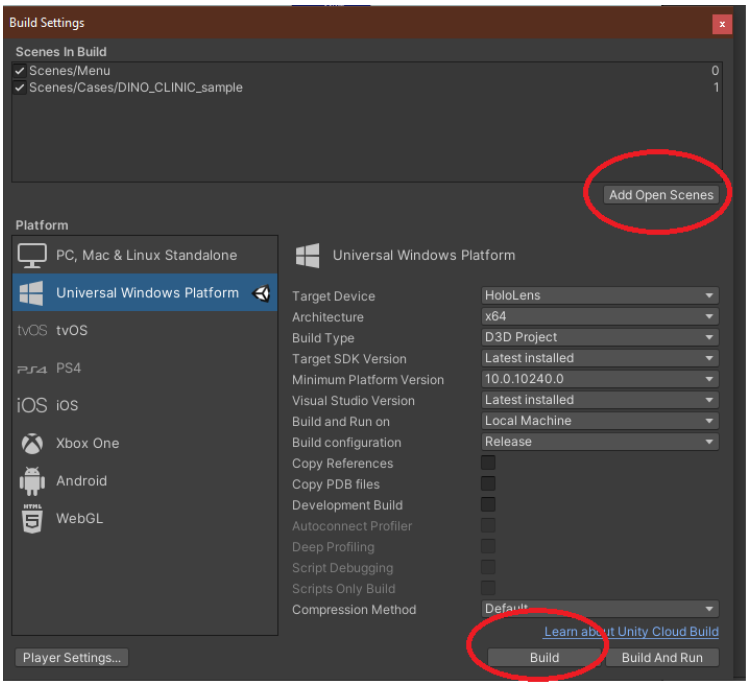
c. In Inspector, go to Events → OnClick () → Replace title with scene title of patient scene.



9. Build DINO-CLINIC

a. Go to File (left top corner) → Build Settings

b. Click on Add Open Scene (Patient scene and Menu scene). Menu scene should be at the top of the list.



c. Then press build

Adaptation Windows Explorer

10. Open build folder → DINO Clinic → Open Package.APPXMANIFEST File with text editor

a. Copy and paste the following text:

```
<?xml version="1.0" encoding="utf-8"?>
<Package xmlns:mp="http://schemas.microsoft.com/appx/2014/phone/manifest"
xmlns:uap="http://schemas.microsoft.com/appx/manifest/uap/windows10"
xmlns:uap2="http://schemas.microsoft.com/appx/manifest/uap/windows10/2"
xmlns:uap3="http://schemas.microsoft.com/appx/manifest/uap/windows10/3"
xmlns:uap4="http://schemas.microsoft.com/appx/manifest/uap/windows10/4"
xmlns:iot="http://schemas.microsoft.com/appx/manifest/iot/windows10"
xmlns:mobile="http://schemas.microsoft.com/appx/manifest/mobile/windows10"
xmlns:rescap="http://schemas.microsoft.com/appx/manifest/foundation/windows10/restrictedcapabilitie
s" IgnorableNamespaces="uap uap2 uap3 uap4 mp mobile iot rescap"
xmlns="http://schemas.microsoft.com/appx/manifest/foundation/windows10">
 <Identity Name="DINO-Unity19" Publisher="CN=MiM" Version="1.0.0.0" />
 <mp:PhoneIdentity PhoneProductId="3542521e-1b13-44f5-8d0d-ee5c20c8a2ab"
PhonePublisherId="00000000-0000-0000-0000-000000000000" />
 <Properties>
  <DisplayName>DINO-Unity19</DisplayName>
  <PublisherDisplayName>MiM</PublisherDisplayName>
  <Logo>Assets\StoreLogo.png</Logo>
 </Properties>
 <Dependencies>
  <TargetDeviceFamily Name="Windows.Holographic" MinVersion="10.0.10240.0"
MaxVersionTested="10.0.22621.0" />
 </Dependencies>
 <Resources>
  <Resource Language="x-generate" />
 </Resources>
 <Applications>
  <Application Id="App" Executable="$targetnametoken$.exe" EntryPoint="DINO-Unity19.App">
   <uap:VisualElements DisplayName="DINO-Unity19" Square150x150Logo="Assets\Square150x150Logo.png"
Square44x44Logo="Assets\Square44x44Logo.png" Description="Sample Unity app for HL2-DINO"
BackgroundColor="transparent">
    <uap:DefaultTile ShortName="DINO-Unity19" Wide310x150Logo="Assets\Wide310x150Logo.png" />
    <uap:SplashScreen Image="Assets\SplashScreen.png" BackgroundColor="#FFFFFF" />
    <uap:InitialRotationPreference>
     <uap:Rotation Preference="landscape" />
     <uap:Rotation Preference="landscapeFlipped" />
     <uap:Rotation Preference="portrait" />
     <uap:Rotation Preference="portraitFlipped" />
    </uap:InitialRotationPreference>
   </uap:VisualElements>
  </Application>
 </Applications>
 <Capabilities>
  <rescap:Capability Name="perceptionSensorsExperimental" />
  <Capability Name="internetClient" />
  <Capability Name="internetClientServer" />
  <Capability Name="privateNetworkClientServer" />
  <uap2:Capability Name="spatialPerception" />
  <DeviceCapability Name="webcam" />
 </Capabilities>
</Package>
```

Deploy DINO Clinic using VS2019

11. Open DINO Clinic.sln
12. Adjust settings:
    a. Switch (top bar) from ARM to ARM64
    b. Right-click on DINO Clinic in Solution Explorer (right side) → Properties
    c. Machine Name: [HoloLens IP]
    d. Apply settings
13. Deploy app on HoloLens 2: Debug → Start without Debugging
    a. Make sure that the HoloLens 2 is correctly connected to the PC.

Appendix J: SIOP staging

Table 15: SIOP staging for Wilms Tumors

Stages	Description
Stage 1	(a) Tumor is limited to kidney and is completely resected (b) The tumor may be protruding into the pelvic system and "dipping" into the ureter (but it is not infiltrating their walls) (c) The vessels of the renal sinus are not involved (d) Intrarenal vessel involvement may be present
Stage 2	(a) The tumor extends beyond kidneys or penetrates through the renal capsule and/or fibrous pseudocapsule into perirenal fat but is completely resected (b) The tumor infiltrates the renal sinus and/or invades blood and lympahtic vessels outside the renal parenchayma but is completely resected (c) The tumor infiltrates adjacent organs or vena cava but is completely resected
Stage 3	(a) Incomplete excision of the tumor, which extends beyond the resection margins (b) Any abdominal lymph nodes are involved (c) Tumor rupture before or intraoperatively (d) The tumor has penetrated through the peritoneal surface (e) Tumor thrombi present at resection margins of vessels or ureter, transected or removed piecemeal by surgeon (f) The tumor has been surgically biopsied (wedge bipsy) prior to preoperative chemotherapy or surgery
Stage 4	Hematogenous metastases (lung, liver, bone brain, etc.) or lymph node metastases outside the abdominopelvic region
Stage 5	Bilateral renal tumors at diagnosis