available at www.sciencedirect.com
journal homepage: www.europeanurology.com

European Association of Urology

Case Series of the Month

Improving Augmented Reality Through Deep Learning: Real-time Instrument Delineation in Robotic Renal Surgery

Pieter De Backer[a,b,c,d,e,*], *Charles Van Praet*[c,d], *Jente Simoens*[a], *Maria Peraire Lores*[a], *Heleen Creemers*[c], *Kenzo Mestdagh*[c], *Charlotte Allaeys*[c,d], *Saar Vermijs*[b,c,e], *Pietro Piazza*[a,f], *Angelo Mottaran*[a,f], *Carlo A. Bravi*[a,g,h], *Marco Paciotti*[a,g,i], *Luca Sarchi*[a,g], *Rui Farinha*[a,g], *Stefano Puliatti*[a,j], *Francesco Cisternino*[k], *Federica Ferraguti*[k], *Charlotte Debbaut*[b,e], *Geert De Naeyer*[g], *Karel Decaestecker*[c,d,l,†], *Alexandre Mottrie*[a,g,†]

[a] *ORSI Academy, Melle, Belgium;* [b] *IBiTech-Biommeda, Department of Electronics and Information Systems, Faculty of Engineering and Architecture, Ghent University, Ghent, Belgium;* [c] *Department of Human Structure and Repair, Faculty of Medicine and Health Sciences, Ghent University, Ghent, Belgium;* [d] *Department of Urology, ERN eUROGEN accredited centre, Ghent University Hospital, Ghent, Belgium;* [e] *Cancer Research Institute Ghent, Ghent University, Ghent, Belgium;* [f] *Division of Urology, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy;* [g] *Department of Urology, Onze-Lieve-Vrouwziekenhuis Hospital, Aalst, Belgium;* [h] *Division of Oncology/Unit of Urology, Urological Research Institute, IRCCS Ospedale San Raffaele, Milan, Italy;* [i] *Department of Urology, Humanitas Clinical and Research Center, Rozzano, Milan, Italy;* [j] *Department of Urology, University of Modena and Reggio Emilia, Modena, Italy;* [k] *Department of Sciences and Methods for Engineering, University of Modena and Reggio Emilia, Modena, Italy;* [l] *Department of Urology, AZ Maria Middelares Hospital, Ghent, Belgium*

Article info

Article history:
Accepted February 13, 2023

Associate Editor:
Sarah P. Psutka

Keywords:
Three-dimensional models
Augmented reality
Instrument segmentation
Partial nephrectomy
Kidney transplantation
Robotic surgery
Renal cell carcinoma
Deep learning
Real time

Abstract

Several barriers prevent the integration and adoption of augmented reality (AR) in robotic renal surgery despite the increased availability of virtual three-dimensional (3D) models. Apart from correct model alignment and deformation, not all instruments are clearly visible in AR. Superimposition of a 3D model on top of the surgical stream, including the instruments, can result in a potentially hazardous surgical situation. We demonstrate real-time instrument detection during AR-guided robot-assisted partial nephrectomy and show the generalization of our algorithm to AR-guided robot-assisted kidney transplantation. We developed an algorithm using deep learning networks to detect all nonorganic items. This algorithm learned to extract this information for 65 927 manually labeled instruments on 15 100 frames. Our setup, which runs on a standalone laptop, was deployed in three different hospitals and used by four different surgeons. Instrument detection is a simple and feasible way to enhance the safety of AR-guided surgery. Future investigations should strive to optimize efficient video processing to minimize the 0.5-s delay

† Joint last authors.
* Corresponding author. ORSI Academy, Proefhoevestraat 12, 9090 Melle, Belgium. Tel. +32 472 394735.
E-mail address: pieter.de.backer@orsi.be (P. De Backer).

https://doi.org/10.1016/j.eururo.2023.02.024

currently experienced. General AR applications also need further optimization, including detection and tracking of organ deformation, for full clinical implementation.

1. Case series

In recent years, patient-specific three-dimensional (3D) models have been introduced in renal surgery as a tool to better define surgical strategies [1]. The adoption of robotic surgery has accelerated intraoperative integration as, 3D models can now be easily visualized on the robotic console. One further step is model integration in the live endoscopic surgical view via augmented reality (AR). During AR, preoperative images (in this case the 3D model) are superimposed on top of the surgical field in an attempt to provide additional information, such as important anatomical landmarks. Although the evidence is still preliminary, AR technologies show potential to further improve outcomes in renal surgery [2].

Figure 1A shows an endoscopic image next to its AR-enhanced equivalent during robot-assisted partial nephrectomy (RAPN) in Figure 1B. This state-of-the-art AR visualization for RAPN highlights two major barriers to current AR implementation. One major barrier is the lack of automated precise 3D model alignment throughout the surgery [3,4]. Figure 1B shows slight misalignment of the 3D model in comparison to the underlying kidney owing to the manual alignment process, the lack of adequate elastic deformation of the model, and motion artifacts such as breathing. Nevertheless, promising initial attempts are under way to tackle this via artificial intelligence, 3D modeling, 3D scene reconstruction, and computer vision algorithms [4–6].

A second major barrier to the implementation of safe AR-guided surgery is the occlusion of surgical instruments by the 3D model, as accentuated in green in Figure 1C. To the best of our knowledge, this is the first study to tackle this problem using real-time automated instrument delineation as depicted in Figure 1D. We present a proof-of-concept study that applies AR Instrument Delineation (AR-ID) technology in RAPN for tumor dissection [6] and transfers it to robot-assisted kidney transplantation (RAKT) for iliac vessel projection and atherosclerotic plaque visualization [7].

Our real-time instrument delineation method uses a binary segmentation algorithm to detect all instruments in the surgical scene, as depicted in Figure 2A. The algorithm was trained using only instruments in RAPN videos. The algorithm architecture is a deep learning convolutional neural network with a U-Net infrastructure, as depicted in Figure 2B.

A total of 15 100 surgical video frames were obtained by sampling 57 RAPN videos every 20 s [8]. Then 65 927 nonorganic items were manually delineated using the SuperAnnotate annotation platform (SuperAnnotate AI, Sunnyvale, CA, USA), resulting in 37 different instrument classes including all robotic/laparoscopic instruments, needles, and wires, among others [8]. The labeling effort duration was estimated at 1258 h [8]. The original images and the corresponding instrument delineation information were fed to a deep learning network.

The labeled data set of 15 100 images was split into three subsets: 10 573 images (40 procedures) were used to train the network, 3019 images (12 procedures) were used for network optimization/validation, and 1508 images (five procedures) were used for testing and reporting the final performance on unseen images. The three data sets contain different RAPN procedures so the algorithm has never seen these images beforehand during the training or optimization phase and hence cannot be biased. The algorithm is trained on a standalone computer with a dedicated computer graphics card (Nvidia RTX3090).

We quantified our results for the test set by reporting two technical metrics commonly used in the field of computer vision:

(1) Intersection over union (IoU) $=\frac{TP}{TP+FP+FN}$

and

(2) Dice score $=\frac{2\times TP}{2\times TP+FP+FN}$,

where TP = true-positive pixels, FP = false-positive pixels, and FN = false-negative pixels.

Our algorithm achieves state-of-the-art accuracy, with an IoU score of 94.4% and a Dice score of 97.10% (Fig. 3).

Once the training and optimization are finished, the algorithm is deployed intraoperatively. Figure 4 shows our

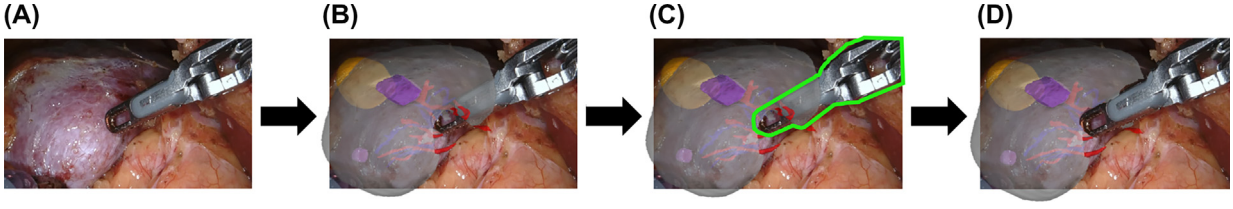

Fig. 1 – State of the art in augmented reality for robot-assisted partial nephrectomy. (A) Original intraoperative view. (B) Overlay of the three-dimensional (3D) constructed preoperatively. (C) Green delineation showing how the view of the force bipolar forceps is occluded by the 3D model. (d) Automated instrument detection during augmented reality solves this problem.

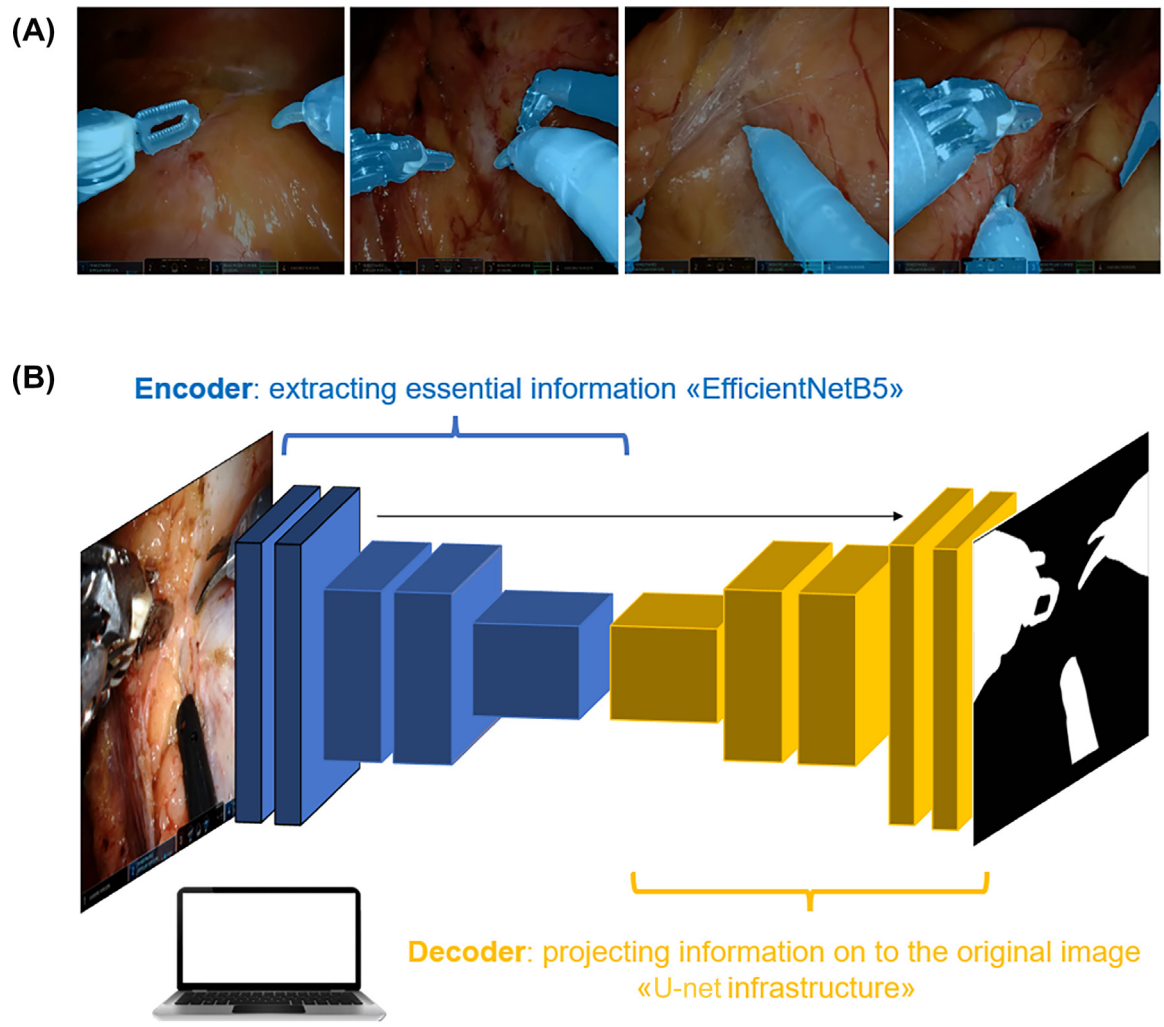

Fig. 2 – Real-time instrument segmentation via convolutional neural networks. (A) Results from our segmentation algorithm. All nonorganic items in the abdominal cavity are instantly detected by the algorithm, as depicted here in blue. This segmentation is called binary as it differentiates between two classes: nonorganic items and soft tissue. (B) Infrastructure of the deep learning convolutional neural network. The infrastructure used to achieve the segmentation in A is a classical infrastructure for deep learning. It consists of an encoder in blue (EfficientNetB5) that has learnt how to extract meaningful information for the task of instrument segmentation. The encoder is followed by a decoder in yellow (U-Net) that learns how to project this condensed meaningful information back onto the original surgical image it was presented with. This infrastructure can run on a laptop with a dedicated video graphics card and is written in PyTorch (https://pytorch.org/).

workflow, inspired by Schiavina et al [9]. (1) A laptop with a dedicated graphics card pulls in the endoscopic view by means of a video capture card. (2) The neural network deployed on the laptop delineates the instruments in real time. (3) A patient-specific 3D model is constructed preoperatively using Mimics (Materialise, Leuven, Belgium). (4) The 3D model is manually aligned and overlaid on the endoscopic view by a biomedical engineer (J.S.) using a Javascript 3D viewer (K3D Javascript, https://k3d.ivank.net/). (5) The laptop then merges the classical AR video with the instrument detection information using vMix software (StudioCoast, Robina, Australia) to generate an AR-ID video that is displayed inside the robotic console view.

The workflow was set up successfully in three tertiary referral centers for use by four surgeons (C.V.P., Ghent University Hospital; A.M. and G.D.N., OLV Aalst; and K.D., AZ Maria Middelares). In all cases, we overlaid a patient-specific 3D model on the corresponding anatomy. For oncologic cases, the tumor and possible cysts are included, while for transplantation cases the iliac arteries were projected to demonstrate feasibility in cases with atherosclerotic plaques.

Between May and October 2022, ten patients undergoing robotic renal surgery were enrolled in the study and signed informed consent. To be eligible for enrolment, an arterial computed tomography scan with a maximal slice thickness of 1 mm was required for adequate 3D model development.

Table 1 summarizes the individual patient characteristics. We performed eight transperitoneal RAPN procedures for localized renal cell carcinoma (including one case with intra-arterial cooling for a highly complex renal lesion), one RAKT with living donation for end-stage renal disease, and one robot-assisted kidney autotransplantation (RAKAT) for persistent loin-pain hematuria after a previous surgery. The median estimated blood loss for the ten surgeries was 200 ml (interquartile range 143.75–275). No intraoperative complications were recorded.

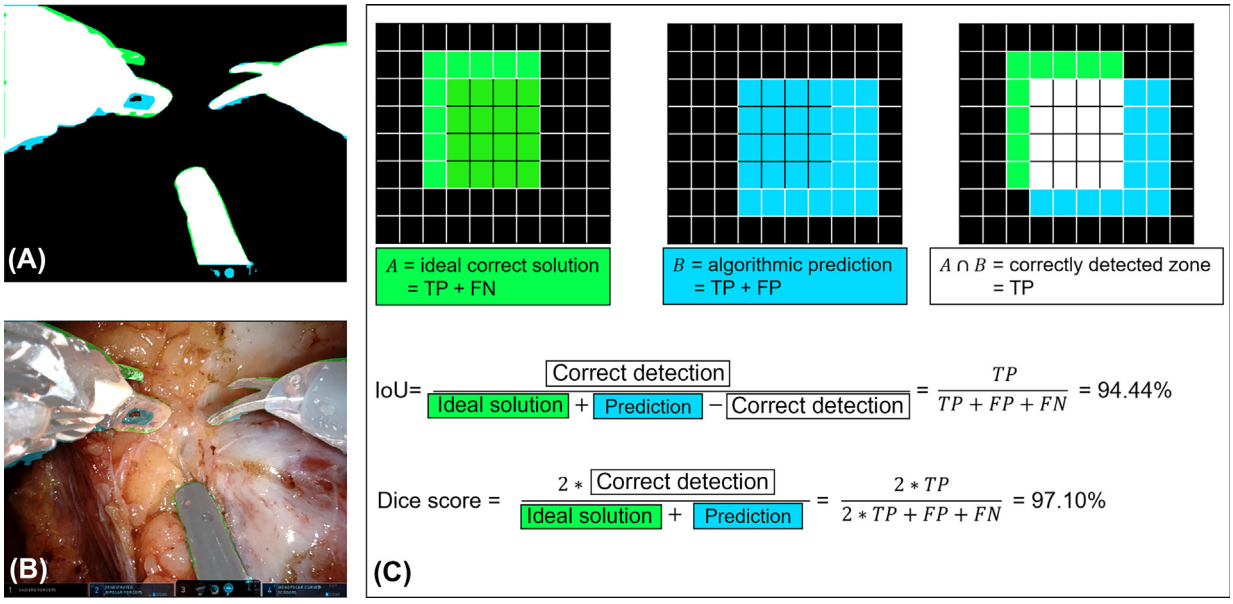

Fig. 3 – Quantification of algorithmic performance. (A) Simplified schematic of the quantification of algorithm performance. White pixels were correctly predicted as instruments (TP). Green pixels were missed by the algorithm but manually annotated by the authors as the perfect solution (FN). Blue pixels depict regions where the algorithm thinks instruments are present where they are not (FP). (B) Corresponding image in which the white, green, and blue colors are overlaid for clear error visualization. It is evident that the inside of the bipolar forceps is not correctly left out and the algorithm was not able to detect the top part of the forceps. (C) Schematic explaining the color coding for A and B. The ideal solution in green (TP + FN) and the full prediction in blue (TP + FP) are weighed against the part that was correctly predicted in white (TP) in terms of the intersection over union (IoU) parameter and the Dice score. The scores achieved scores both parameters are shown.TP = true positive; FN = false negative; FP = false positive.

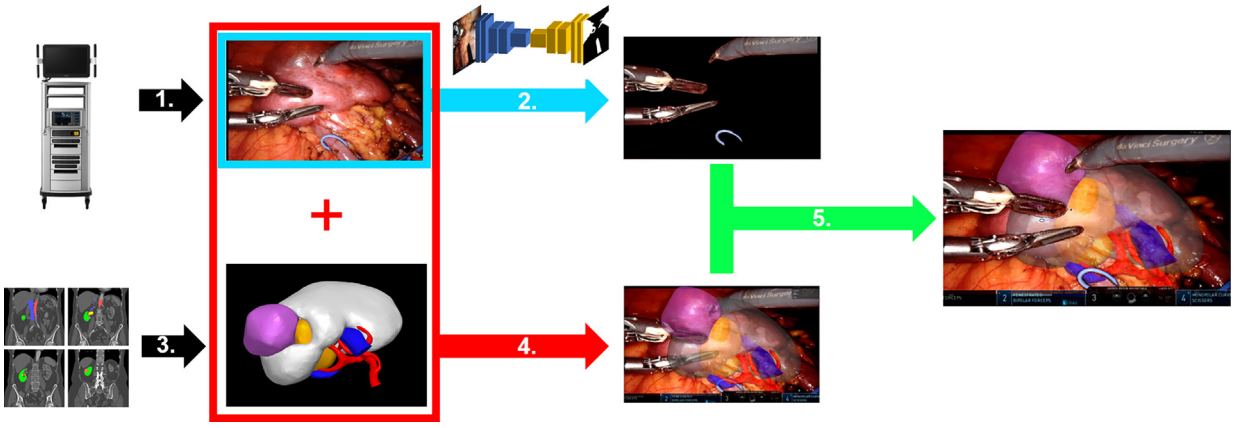

Fig. 4 – Workflow in the operating room allowing real-time instrument segmentation during augmented reality (AR) application. (1) A laptop with a dedicated graphics card (NVIDIA RTX A2000) provides the endoscopic view via a video capture card. (2) The instrument segmentation neural network runs on the laptop and extracts images of the instruments from the incoming video in real time. (3) The patient-specific three-dimensional (3D) model is constructed preoperatively using Mimics (Materialise, Leuven, Belgium). (4) The 3D model is manually aligned and overlaid on the endoscopic view on the laptop, with model rotation and alignment performed in Javascript (https://k3d.ivank.net/). (5) The laptop merges the classical AR video with the real-time instrument detection using vMix software (StudioCoast, Robina, Australia). This generates the instrument detection AR video, which is displayed on the robotic console via an Intuitive TilePro input.

Qualitative visual assessment of our real-time binary segmentation algorithm shows precise real-time delineation of all instruments during both RAPN and RAK(A)T. We did measure a 0.5-s delay when comparing the direct endoscopic view with the TilePro window (Intuitive Surgical, Sunnyvale, CA, USA). This is because of the serial infrastructure for both the capture card and the laptop, whereby each video frame is reprocessed multiple times.

Figure 5A shows the setup as visualized on the robotic console during echographic tumor demarcation. Figure 5B presents different examples of the current state of the art by comparing images with AR-ID detection off and on. It is evident that AR-guidance facilitates tumor demarcation in RAPN, while visualization of the iliac vessels is enhanced during RAKT. Figure 5C provides a QR video link to view the difference between AR with and without instrument detection.

Table 1 – Characteristics of the patients in whom augmented reality instrument detection was applied

	RAPN 1	RAPN 2	RAPN 3	RAPN 4	RAPN 5	RAPN 6	RAPN 7	RAPN 8 [a]	RAKAT	RAKT
Sex	Female	Male	Female	Female	Male	Female	Male	Female	Female	Male
Age (yr)	44	67	56	61	60	76	77	54	30	61
BMI (kg/m^2)	36	23.5	22.5	27.7	31.5	21	33.9	26	20.5	27.9
Side	Right	Left	Right	Right	Right	Left	Left	Right	Left	Right
LS (mm)	19	24	16	55	35	19	24	67	NA	NA
Padua score	7	8	3	9	6	6	6	13	NA	NA
WIT (min)	0	19	16	0	16.5	11.5	12.85	96 [b]/3	60 [b]/3	123 [b]/48
Histotype	OCC	ccRCC	chRCC	ccRCC	ccRCC	ccRCC	OCC	ccRCC	NA	NA
Surgical margins	Free	Free	Free	Free	Free	Free	Free	Free	NA	NA
$eGFR_{pr}$ (ml/min)	82	63.1	85	83	88	80	72	90	90	14.7
CRT_{pr} (mg/dl)	0.86	1.18	0.77	0.77	0.94	0.73	0.96	0.68	0.55	4.09
$eGFR_{po}$ (ml/min)	90	78	90	81	90	63	73	90	90	34
CRT_{po} (mg/dl)	0.68	0.99	0.64	0.79	0.83	0.9	0.99	0.6	0.59	2.04
CT (min)	150	150	120	70	180	75	80	216	260	207
POCs	None	None	None	None	None	C1: HPG	C2: CDC	None	None	C2: RTI

BMI = body mass index; LS = lesion size; WIT = warm ischemia time, $eGFR_{pr}$ = preoperative estimated glomerular filtration rate; $eGFR_{po}$ = postoperative eGFR; CRT_{pr} = preoperative creatinine, CRT_{po} = postoperative CRT; CT = console time; RAPN = robot-assisted partial nephrectomy; RAKAT = robot-assisted kidney autotransplantation; RAKT = robot-assisted kidney transplantation; NA = not applicable; OCC = oncocytoma; ccRCC = clear cell renal cell carcinoma; chRCC = chromophobic RCC; POCs = postoperative complications rated using the Clavien-Dindo classification [10]; C1 = class 1 POC; C2 = class 2 POC; HPG = hypoglycemia; RTI = respiratory tract infection, CDC = cardiac decompensation.

[a] With intracorporeal cooling.

[b] Cold ischemia time.

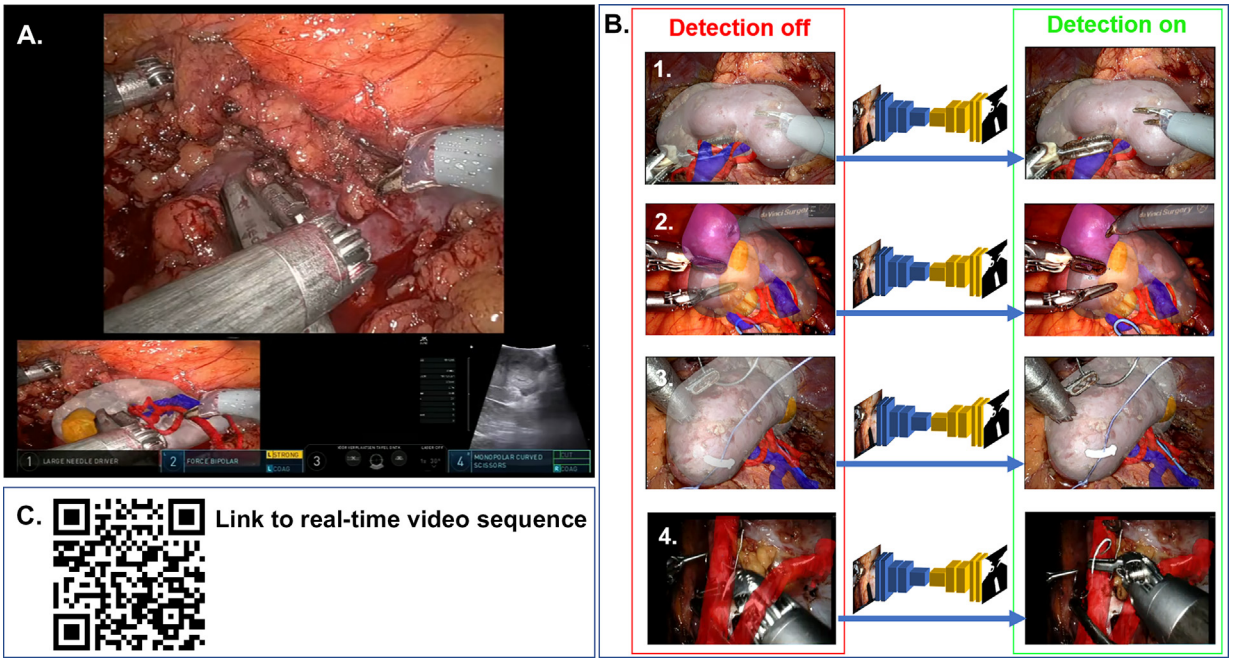

Fig. 5 – Real-time examples of augmented reality instrument detection (AR-ID). (A) The robotic console during surgery shows delineation of the monopolar curved scissors, force bipolar forceps, and ultrasound probe on top of the 3D model projection in the left TilePro window, while the right TilePro window shows the echography image during scanning across the tumor. (B) Four different cases of the greater visualization when detection is toggled on, namely, (1) precision of instrument detection in robot-assisted partial nephrectomy; (2) palpation of cysts and visualization of the underlying tumor is improved in AR, with detection of the vessel loop; (3) detection of wires, needles, Hem-o-lok clips, and vessel loops; and (4) integral transfers of the technology to robot-assisted kidney transplantation during arterial anastomosis, with detection of the robotic instruments and the sutures and bulldog clamps applied. (C) A QR code with a link to a video showing the on-off toggling of the AR-ID technology in different cases.

2. Discussion

Our case series shows that it is feasible to remove the safety hazard of not visualizing instruments during AR surgery and that our approach adds a sense of depth to 3D model interaction. Our real-time AR-ID method is robust and transfers well from RAPN to RAK(A)T. The setup requires a laptop with a graphics card (€1000–1500), a capture card (€100–400), and an additional basic computer monitor (€100–200). It is easily applicable in different operating rooms and hospitals. Nevertheless, this setup has a 0.5-s delay, so it is not currently unacceptable to perform surgery based solely on the TilePro window. Thus, our future research is

focused on reducing the time delay via parallel integration of all components in a single computing device, which is expected to increase the hardware costs at least threefold. Second, the current application detects all instruments at once without a quantitative comparison of the detection level for individual instruments, which would require complete refactoring and retraining of the system. However, this approach might enhance our proposed solution for cases in which detection is imprecise. Third, manual alignment of static 3D models to a constantly changing operative view remains bothersome. The implementation of AR-guided surgery requires assurance regarding perfect alignment. In this context, deep learning might help in solving the localization problem [4], while numerical methods could assist in generating realistic model deformation sequences. All three barriers identified require dedicated and interconnected computational resources and represent a future challenge for successful and smooth clinical integration.

Conflicts of interest: The authors have nothing to disclose.

Acknowledgments: This research was supported by Ipsen NV (grant A20/TT/1655), the special research fund of Ghent University (BOF starting grant BOFSTA201909015), and Flanders Innovation & Entrepreneurship (VLAIO; Baekeland grant HBC.2020.2252 to Ghent University [reference A20/TT/0337] and ORSI Academy). The sponsors provided unconditional grants, supporting research for kidney sparing treatments, without having any say in the design, study conduct, data aspects, nor were they involved at any part in the manuscript writing, review or approval. The authors would like to thank Rania Matthys for support in three-dimensional model making, Tim Oosterlinck, Julie Lippens, and Amélie Hallemeesch for support in instrument annotation, Luigi Nocera and Adele Piro for support in video recording, and Camille Berquin and Ruben De Groote for clinical feedback.

Peer Review Summary and Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eururo.2023.02.024.

References

[1] Piramide F, Kowalewski KF, Cacciamani G, et al. Three-dimensional model-assisted minimally invasive partial nephrectomy: a systematic review with meta-analysis of comparative studies. Eur Urol Oncol 2022;5:640–50.
[2] Roberts S, Desai A, Checcucci E, et al. 'Augmented reality' applications in urology: a systematic review. Minerva Urol Nephrol 2022;74:528–37.
[3] Piramide F, Amparore D, Pecoraro A, et al. Augmented reality 3D robot-assisted partial nephrectomy: tips and tricks to improve surgical strategies and outcomes. Urol Video J 2022;13:100137.
[4] Khaddad A, Bernhard JC, Margue G, et al. A survey of augmented reality methods to guide minimally invasive partial nephrectomy. World J Urol 2022. https://doi.org/10.1007/s00345-022-04078-0, In press.
[5] Amparore D, Checcucci E, Piazzolla P, et al. Indocyanine green drives computer vision based 3D augmented reality robot assisted partial nephrectomy: the beginning of "automatic" overlapping era. Urology 2022;164:e312–6.
[6] Porpiglia F, Checcucci E, Amparore D, et al. Three-dimensional augmented reality robot-assisted partial nephrectomy in case of complex tumours (PADUA ≥10): a new intraoperative tool overcoming the ultrasound guidance. Eur Urol 2020;78:229–38.
[7] Piana A, Gallioli A, Amparore D, et al. Three-dimensional augmented reality–guided robotic-assisted kidney transplantation: breaking the limit of atheromatic plaques. Eur Urol 2022;82:419–26.
[8] De Backer P, Eckhoff JA, Simoens J, et al. Multicentric exploration of tool annotation in robotic surgery: lessons learned when starting a surgical artificial intelligence project. Surg Endosc 2022;36:8533–48.
[9] Schiavina R, Bianchi L, Chessa F, et al. Augmented reality to guide selective clamping and tumor dissection during robot-assisted partial nephrectomy: a preliminary experience. Clin Genitourin Cancer 2021;19:e149–55.
[10] Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. Ann Surg 2004;240:205–13.