SIUJ Société Internationale d'Urologie Journal

Review

A Review of the Intraoperative Use of Artificial Intelligence in Urologic Surgery

Arjun Guduguntla [1,2], Abdullah Al-Khanaty [1], Catherine E. Davey [3], Oneel Patel [2], Anthony Ta [1,2] and Joseph Ischia [1,2,*]

[1] Department of Urology, Austin Health, Heidelberg, VIC 3084, Australia; anguduguntla@student.unimelb.edu.au (A.G.); abdullah.al-khanaty@austin.org.au (A.A.-K.); anthony.ta@austin.org.au (A.T.)
[2] Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Parkville, VIC 3052, Australia; patelo@unimelb.edu.au
[3] Faculty of Engineering and IT, University of Melbourne, Parkville, VIC 3052, Australia; catherine.davey@unimelb.edu.au
* Correspondence: jjischia@unimelb.edu.au; Tel.: +61438004100

Abstract: Introduction: Future evolutions of artificial intelligence (AI) will support autonomous surgery, conducted without the need for human decision making and implementation, but we have not yet achieved this level of technology. Presently, the predominant applications of AI in urological surgery are achieved using the tool of computer vision. This review aims to summarise potential intra-operative AI tools for urologists. **Method**: A systematic search was conducted through Scopus, PubMed, Embase, and Medline by two independent reviewers, with a third to resolve any conflicts. As a rule, only original articles describing the use or potential use of artificial intelligence intra-operatively in urologic surgery were included. A total of 60 articles were reviewed. **Key content and findings**: There is significant research investigating the ability to diagnose bladder tumours using AI assistance at the time of cystoscopy, with studies showing the ability to also grade tumour based on appearance and differentiate between carcinoma in situ and indeterminate lesions. With the aid of AI, kidney stones can accurately be identified and diagnosed morphologically intra-operatively. Various studies show the ability to overlay 2D and 3D anatomical models on a surgeon's screen, as well as correctly identify important anatomical landmarks and surgical instruments, with AI support. All types of intra-operative data can be analysed with AI to assess surgeon performance, predict post-operative outcomes such as continence post prostatectomy, and recognise complications such as bleeding and ischemia. **Conclusions**: AI holds great potential for urologists during surgery to improve safety, diagnostic accuracy, identification of anatomical structures and surgical instruments, assessment of the surgeon for self-evaluation, and prediction of post-operative outcomes. Before the use of AI as an aid during surgery becomes standard practice, more prospective studies are needed to evaluate its real-world application, feasibility, and costs.

Keywords: computer vision; artificial intelligence; machine learning; deep learning; artificial neural network

Academic Editor: Peter C. Black

Received: 15 October 2024
Revised: 9 December 2024
Accepted: 20 December 2024
Published: 12 February 2025

Citation: Guduguntla, A.; Al-Khanaty, A.; Davey, C.E.; Patel, O.; Ta, A.; Ischia, J. A Review of the Intraoperative Use of Artificial Intelligence in Urologic Surgery. *Soc. Int. Urol. J.* **2025**, *6*, 5. https://doi.org/10.3390/siuj6010005

1. Introduction

Artificial Intelligence (AI) technology holds the potential to underpin visual processing and decision making in autonomous robotic surgery. Autonomous robotic surgery is already in the animal testing phase, where intestinal anastomosis has successfully been performed on porcine subjects, assisted by AI-based algorithms that enable tissue motion

tracking and markerless tissue landmark detection [1]. In human surgery, currently, computer vision forms the mainstay for AI in assisting a surgeon intra-operatively by analysing images and videos to interpret and extract patterns [2].

In particular, there has been significant headway in general surgery, where AI-based software has been designed to identify tumours with high accuracy on endoscopy and colonoscopy, as well as software to accurately recognise relevant critical view structures during cholecystectomy [3–6]. Neurosurgeons have used AI for intra-operative diagnosis of brain tumours and recognition of vascular injury [7]. Intra-operative use of AI has been used in orthopaedics particularly during arthroscopy to assist with things such as anchor insertion at time of shoulder arthroscopy [8,9].

Prior to increased adoption of AI, urological surgeons utilised visual supports such as augmented reality (AR) and virtual reality (VR) [10–12]. AR has utility in overlaying artificial images or objects onto the surgical field and surgical screen, or in view of the operating theatre [13]. VR can display the same things, but in a fully immersed and simulated environment [13]. Two-dimensional and three-dimensional models have also been used to assist a surgeon in cognitively fusing their perception of a patient's anatomy, ensuring safe dissection and optimal excision [14]. Optical adjuncts, such as blue light cystoscopy, spectroscopy, and medical dyes, have all proven to be helpful in improving surgical safety and outcomes [15–17]. Whilst these traditional supports can aid urologists in real time, they usually require time consuming individualised analysis of patient data prior to surgery. They also can be susceptible to human error, as they require varying degrees of interpretation by the operating surgeon.

Convolutional neural networks (CNNs) are a type of artificial neural network (ANN) that are highly effective in interpreting visual data; the mathematical operation of convolution allows the network, designed to mimic the function of a human brain, to detect features in images [18]. Employing the use of CNNs or other forms of AI during surgery can diminish the drawbacks of conventional optical adjuncts by minimising surgeon cognitive load, and once AI-based programs have been established, will require less time and resources to implement and use [19]. AI may also completely negate the need for certain optical adjuncts, like medical dyes, as they can interpret visual data just as accurately without their use [20,21].

This review seeks to provide a concise overview of the existing literature regarding the increasing use, or potential application, of AI in assisting urological surgeons during operations. A glossary of terms and concepts is included in Appendix A.

2. Materials and Methods

A systematic search was conducted using Scopus, PubMed, Embase, and Medline by two independent reviewers (A.G. and A.A.K.), with a third reviewer used to adjudicate any conflicts (JI). The review was performed in accordance with the methodology outlined in the latest Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guideline (Figure 1) [22]. Key search words or themes to create MeSH terms included "Computer Vision", "Intra-operative", "Real-Time", "Detection", "Identification", "Recognition", "Image Segmentation", "Semantic Segmentation", "Image-Processing", "Computer-Assisted Surgery", "Augmented reality", "Virtual Reality", "Artificial Neural Network", "Convolutional Neural Network", "Artificial Intelligence", "Deep Learning", "Machine Learning", "Algorithm", and "Automation". As a rule, only original articles describing the use or potential use of artificial intelligence intra-operatively in urologic surgery, were included. We excluded case reports, conference abstracts, commentaries, animal studies, non-English literature, letters, video articles, and literature reviews. However,

pertinent content found in literature reviews were used to confirm the robustness of our search and search terms.

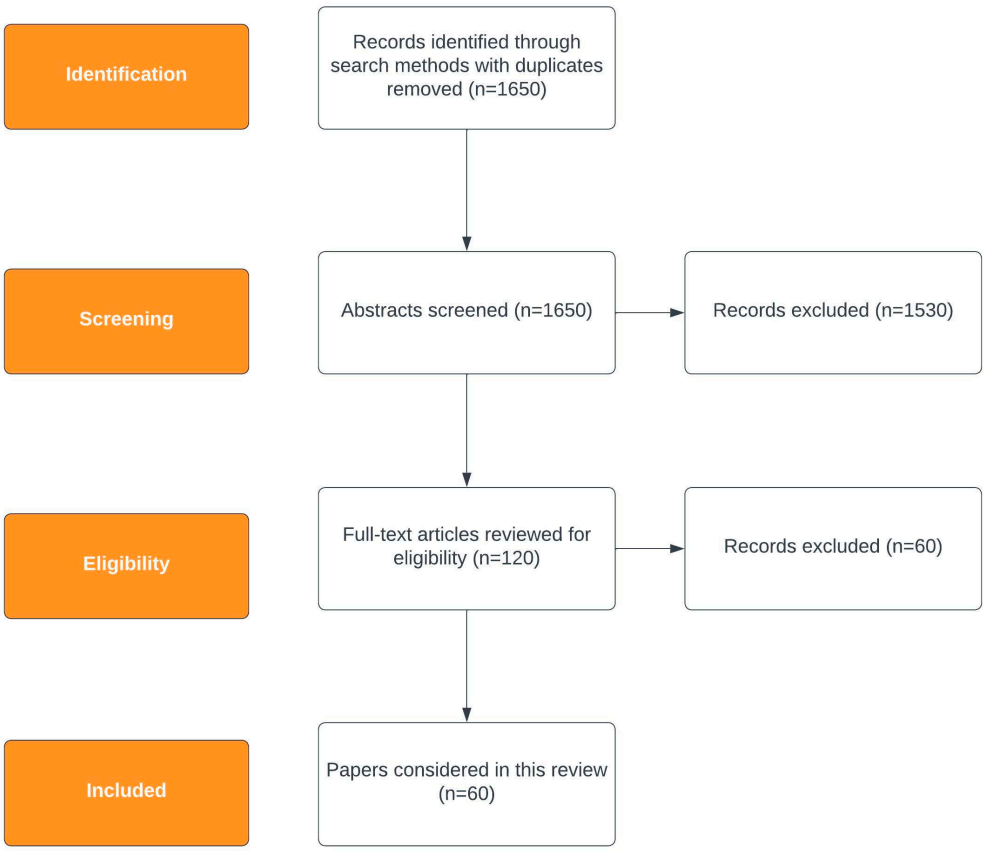

Figure 1. Flowchart detailing identification of articles reviewed for review using PRISMA guidelines.

3. Key Content and Findings

3.1. Intra-Operative Diagnosis Aided by AI

3.1.1. Bladder Cancer Diagnosis

There are various studies that detail the use of different CNNs to assist with identifying urothelial carcinoma (UC) tumours during cystoscopy. There is also one known study investigating the ability of AI to help with diagnosis of upper tract UC during ureteroscopy [23]. These studies are summarised in Table 1.

Table 1. Summary of studies using AI to identify UC at time of cystoscopy and ureteroscopy, subdivided by types of images used to train and validate CNN (refer to Appendix A for definitions of the terms used).

Study Name [Year]	Authors	Endpoint Results [CNN Framework]	Frames Used to Train CNN	Comments
Trained and validated on cystoscopic images derived from cystoscopic videos (unless otherwise specified)				
Diagnostic Classification of Cystoscopic Images Using Convolutional Neural Networks (2018)	Eminaga et al. [24]	99.54% sensitivity (XCeption) 99.48% sensitivity (ResNet-50)	18,681 frames based on 479 images (15,327 non-tumour, 3354 with bladder tumours)	Digital images augmented to create significantly more frames. Model created classifies all cystoscopic findings from a digital atlas, not just bladder tumours. Significant limitation in that based on high-quality images.

Table 1. *Cont.*

Study Name [Year]	Authors	Endpoint Results [CNN Framework]	Frames Used to Train CNN	Comments
Augmented Bladder Tumour detection using Deep Learning (2019)	Shkolyar et al. [25]	90.9% per tumour sensitivity and 98.9% per tumour specificity (CystoNet based on ResNet-50)	2752 frames (2335 non-tumour, 417 with bladder tumours)	Was not trained on flat lesions but still detected them on validation.
Support System of Cystoscopic Diagnosis for Bladder Cancer Based on Artificial Intelligence (2020)	Ikeda et al. [26]	89.7% sensitivity, 94.0% specificity	2102 frames (1671 non-tumour, 431 with bladder tumours)	Low diagnostic accuracy when tumour occupying < 10% of frame space.
Cystoscopic Imaging for Bladder Cancer Detection Based on Stepwise Organic Transfer Learning with a Pretrained Convolutional Neural Network (2021)	Ikeda et al. [27]	95.4% sensitivity, 97.6% specificity	1.2 million general frames, 8728 gastroscopic frames and then the same 2102 frames as above (1671 non-tumour, 431 with bladder tumours)	Same framework as paper above by Ikeda et al. [26], but model pre-trained on gastroscopic images prior to cystoscopic. Ongoing issue with diagnostic accuracy tumours when <10% of frame.
A deep learning network-assisted bladder tumour recognition under cystoscopy based on Caffe deep learning framework and EasyDL platform (2021)	Du et al. [28]	82.9% accuracy (Caffe) 96.9% accuracy (EasyDL)	1736 frames (1002 non-tumour, 734 frames tumour)	EasyDL platform requires user to take photo of suspicious lesion and upload via phone app/website to get confidence value of whether lesion tumour.
A Comparative Study of Deep Neural Networks for Real-Time Semantic Segmentation during the Transurethral Resection of Bladder Tumours (2022)	Varnyu et al. [29]	97.39% sensitivity, 92.54 dice score, 0.9056 F-Score (U-Net)	2578 frames (these are filtered, sharpened, and enhanced, so 9280 frames overall used)	Investigates many different neural networks, with U-Net overall having best results.
Artificial Intelligence for Segmentation of Bladder Tumour Cystoscopic Images Performed by U-Net with Dilated Convolution (2022)	Mutaguchi et al. [30]	84.9% sensitivity, 88.5% specificity, 86.7% accuracy, 83.0 dice score (U-Net) 86.6% sensitivity, 89.8% specificity, 88.8% accuracy, 85.3 dice score (Dilated U-Net)	1790 frames (all with bladder tumour)	Low dice coefficient for flat tumours.
A comparative study of attention mechanism based deep learning methods for bladder tumour segmentation (2023)	Zhang et al. [31]	Dice score 82.0, 69% Mean Intersection over Union (mIoU)	737 frames	Attention-based mechanisms are used to enhance deep learning model so that there is better edge processing of lesions, leading to improved dice scores and better mIoU.
Tumour Detection under cystoscopy with transformer-augmented deep learning algorithm (2023)	Jia et al. [32]	97.3% sensitivity, 95.6% accuracy (CystoNet-T) 89.2–92.8% sensitivity, 88.4–93.6% accuracy (previous CystoNet models)	611 frames (all with bladder tumour, some with more than one tumour)	Builds on previous works of Chang et al. [33] and Shkolyar et al. [25], as above and below, on CystoNet. Utilises transformer-augmented pyramidal convolutional network architecture, as well as attention mechanisms, for better results.

Table 1. *Cont.*

Study Name [Year]	Authors	Endpoint Results [CNN Framework]	Frames Used to Train CNN	Comments
U-Net-Based Assistive Identification of Bladder Cancer: A Promising Approach for Improved Diagnosis (2024)	Guo et al. [34]	99% sensitivity, 98% accuracy, 0.98 F1-score (U-Net)	3500 frames from 100 patients, all with bladder tumour	Significant results based on U-Net architecture, better than other studies using U-Net in Table 1.
A lightweight bladder tumour segmentation method based on attention mechanism (2024)	Zhao et al. [35]	90.81% sensitivity, 91.52% accuracy, 0.9116 F1-score (NAFF-Net)	425 frames (211 papilloma, 214 bladder tumour)	Deep learning model with low computational requirements that incorporates an attention mechanism to still achieve similar results to other models in Table 1.
Selection of Convolutional Neural Network Model for Bladder Tumour Classification of Cystoscopy Images and Comparison with Humans (2024)	Lee et al. [36]	88% sensitivity, 74% specificity, 81% accuracy (EfficientNetB0)	3731 frames (2191 bladder tumour, 1540 without tumour)	17 CNN models trained for tumour classification, with EfficientNetB0 selected as appropriate out of them all. Comparison between CNN tumour diagnostic metrics with urologists as well as medical students performed by authors.
Trained and validated on cystoscopic images derived from cystoscopic videos; however, CNN trained to differentiate based on colour space only				
Automatic T1 bladder tumour detection by using wavelet analysis in cystoscopy images (2018)	Freitas et al. [37]	91% sensitivity, 92.9% specificity in Hue-Saturation-Value (HSV) colour space	353 frames (246 non-tumour, 107 with bladder tumours)	Focusses only on detection of T1 bladder tumours with best results achieved in the HSV colour space.
Deep Learning Diagnostics for Bladder Tumour Identification and Grade Prediction using RGB Method (2022)	Yoo et al. [38]	99.3% sensitivity, 98% specificity (WLI) 98.0% sensitivity, 91.2% specificity (WLI differentiating carcinoma in situ (CIS) and chronic inflammation) 67.0% sensitivity, 62.9% specificity (WLI in differentiating high-grade vs. non-high-grade lesions)	8244 frames (2742 non-tumour, 5502 bladder tumour) for training 1847 frames (1220 non-tumour, 627 bladder tumour) for validation	Red-green-blue (RGB) values from pixels in labelled tumour were used to train the AI model. White light imaging (WLI) and narrow band imaging (NBI) were used, and best results were found in the former. It is suspected this is because NBI limited the optical spectrum. The ability of the program to differentiate between high grade vs. non-high grade and CIS and chronic inflammation was also measured.
Trained and validated on blue light cystoscopic images derived from blue light cystoscopic videos				
Deep Learning-based classification of blue light cystoscopy imaging during transurethral resection of bladder tumours (2021)	Ali et al. [39]	95.77% sensitivity and 87.84% specificity in assessing malignancy in assessing malignancy (ResNet-50) 88% sensitivity and 96.56% specificity in assessing tumour invasiveness (ResNet-50)	216 blue light frames	Uses blue light cystoscopy as opposed to white light. Comparison between AI made with 2 experts. Secondary outcome assessed of tumour invasiveness prediction based on appearance.

Table 1. *Cont.*

Study Name [Year]	Authors	Endpoint Results [CNN Framework]	Frames Used to Train CNN	Comments
Trained on cystoscopic frames derived from cystoscopic videos, but validated with cystoscopic videos in real time (rather than frames derived from video)				
An Artificial Intelligence System for the Detection of Bladder Cancer via cystoscopy: A Multi-Centre Diagnostic Study (2021)	Wu et al. [40]	95% sensitivity, 93.9% accuracy, NPV 0.989–0.999 (PSPNet)	69,204 frames, but pre-trained on ImageNet prior	Software touted as Cystoscopy Artificial Intelligence Diagnostic System (CAIDS). <50 ms latency when used in real time. Has been integrated into practice at centres study conducted at.
Automatic recognition of bladder tumours using deep learning technology and its clinical application (2021)	Yang et al. [41]	91.23% accuracy (LeNet) 88.69% accuracy (AlexNet) 92.54% accuracy (GoogLeNet) 96.9% accuracy (EasyDL) 93.94% sensitivity (Easy DL) vs. 87.88% sensitivity (surgical expert) but $p > 0.05$	2350 frames (1150 non-tumour, 1200 tumour)	Similar to study above by Du et al. [28], but EasyDL compared with other CNNs and now system connected to cystoscope feed for live results. Secondary outcome assessed was classification of cancer between deep learning and surgical expert.
Efficient Augmented Intelligence Framework for Bladder Lesion Detection (2023)	Eminaga et al. [42]	81.4–88.1% sensitivity, 30–44.8% specificity between all models 100% sensitivity, 56–67% specificity between all models when frames analysed in frame-sequence blocks	272,799 frames (188,220 non-tumour, 84,579 with bladder tumour)	Validation and expansion on previous work by Eminaga et al., as above, based on digital atlas of cystoscopic images. Validated on videos at time of bladder tumour resection, with multiple different CNNs used. Assessed sensitivity and specificity in blocks of frames as well.
Real-Time Detection of Bladder Cancer using Augmented Cystoscopy with Deep Learning: A Pilot Study (2023)	Chang et al. [33]	41.55% per-frame tumour sensitivity, 95.1% per-frame tumour specificity for in office cystoscopy 52.9% per-frame tumour sensitivity, 95.4% per-frame tumour specificity for videos at time of tumour resection N.B per-frame sensitivity and specificity refers to the author's devised metrics when video is input, rather than frames	155,511 frames (129,478 non-tumour, 26,033 with bladder tumour)	This study is to assess the real-time integration of CystoNet developed by Shkolyar et al. [25], as above, when used in office cystoscopies or at time of bladder tumour resection. Patients with non-papillary or CIS only lesions were excluded from their study. Importantly, all positive cystoscopies were distinguished from negative ones.

Table 1. *Cont.*

Study Name [Year]	Authors	Endpoint Results [CNN Framework]	Frames Used to Train CNN	Comments
Trained and validated on frames derived from endoscopic video during ureteroscopy, with further validation on endoscopic video during ureteroscopy				
Automated Upper Tract Urothelial Carcinoma Tumour Segmentation During Ureteroscopy Using Computer Vision Techniques (2024)	Lu et al. [23]	76% sensitivity, 99% specificity, 98% accuracy (U-Net)	3387 frames (from 20 videos of endoscopic treatment of upper tract UC)	Model had mild reduction in metrics when further validated on endoscopic video; however, it was able to produce automated overlay of predicted tumour at 15 frames per second. Mix of low-grade and high-grade upper tract UC cases used in this study.

Most of the studies in Table 1 employed similar methodologies, though they differ in the choice of CNN characteristics that the CNN focussed on analysing, as well as in terms of the types of images used to train the AI. For example, Freitas et al. used a discrete wavelet transform algorithm that relies on the fact that tumours are characterised by high texture variations to train their AI model in different colour spaces [37]. Similarly, Yoo et al. formulated their AI program using frames from the red-green-blue (RGB) colour space, but also secondarily investigated whether tumour grade or carcinoma in situ (CIS) and chronic inflammation could be differentiated using this ideology [38]. It is also evident that using transformer-augmented architecture and attention-based mechanisms increase all metrics in tumour identification for regional-based convolutional neural networks (r-CNNs) [31,32,35]. Ikeda et al. showed that pre-training their model with general and gastroscopic images before the cystoscopic ones improved sensitivity and specificity, suggesting similar visual tumour characteristics between the gastrointestinal and urological systems [26,27].

AI-assisted bladder tumour diagnostic software would ideally have a high sensitivity and low false positives or otherwise assist in reducing unnecessary biopsies for indeterminate lesions. Ali et al. show that, with blue light cystoscopy, they were able to differentiate between grades of tumour, as well as even differentiate CIS and benign lesions with a mean sensitivity of 78.1% [39]. Differentiating grade may be helpful in assessing candidacy for patients to receive intra-vesical therapy after tumour resection. Yoo et al. demonstrated an even better ability to differentiate CIS and chronic inflammation than Ali et al., and Wu et al. quantified that their software could avoid 62.5% of misdiagnosed cases of these lesions [38–40]. These three studies both show potential to reduce unnecessary biopsies for undefined red patches or flat lesions. Yang et al., Ali et al., Wu et al. and Lee et al. also compared tumour identification between AI and surgical experts; they all claimed that the metrics of their respective AI models are comparable to urologists [36,39–41]. Whilst these comparative evaluations seem promising, they are limited as their methodology lacks realistic representation of tumour identification in the clinical setting.

Applying AI clinically during urinary endoscopy introduces issues including variable magnification of an image, artefacts from light, variable image quality, distraction to the operator, and issues related to computational requirements and efficiency of the systems themselves. Chang et al. had significantly reduced sensitivity with respect to each specific frame when video was input into their system using CystoNet live during cystoscopy, rather than frames derived from the video [33]. Importantly, at the end of all the cystoscopies, all cases with tumour were still correctly identified [33]. Eminaga et al. likewise missed no cases of cancer when grouping frames from cystoscopic video into frame-sequence blocks [42]. Therefore, it is likely that even with increased complexities when using these

programs in practice, tumours are still identifiable. This suggests that end-points such as per frame sensitivity postulated by Chang et al. [33] may not be as clinically relevant as perhaps whether all tumours present are identified over the course of a whole cystoscopy; this would require standardising the method cystoscopy performed between operators but would be a more realistic metric. Eminaga et al. also showed that whilst some CNN models have higher sensitivity and specificity than others, they do not necessarily have low latency and low computational requirements [42]. This is significant as it complicates implementation of software into clinical practice, particularly because of the costs of systems requiring significant computational demands. Due to this, Zhao et al. created a lightweight model using attention-based mechanisms to increase its efficiency, with similar metrics to other studies in Table 1 [35]. Of all the literature, the CAIDS system by Wu et al. is the only software that we are aware of that has been implemented into practice, but they have had to report on their outcomes since implementation [40].

There is significant potential for AI to become a routine part of diagnostic cystoscopy and ureteroscopy for bladder cancer, specifically so that tumours are not missed, in order to differentiate indeterminate lesions and possibly to grade lesions based on appearance.

3.1.2. Kidney Stone Identification and Morphological Diagnosis

AI has mainly been used in endoscopic stone surgery to diagnose and detect stone composition and morphology based on visual appearance, as well as for intra-operative measurement and tracking of stone fragments. Initial studies investigating AI recognition of stone composition were limited as the images used to train the AI were from a simulated setting, or because the investigators used machine learning rather than CNNs, which have better sensitivity and specificity [43–47]. The findings of the different studies demonstrating the use of CNNs practically in stone surgery are summarised in Table 2. They show that prediction of stone composition can be performed with sensitivity and specificity over 90%, and tracking of fragments can be performed even during laser use. These studies suggest that, when applied clinically, it could potentially lead to shortened operative time and improved stone clearance rates.

Table 2. Summary of studies' practical use of CNNs in endoscopic stone surgery, sub-divided by whether the study assesses for stone composition, stone segmentation, or both (refer to Appendix A for the definition of terms used).

Study Name [Year]	Authors	Endpoint Results [CNN Framework]	Frames Used and Objective of CNN	Comments
Studies assessing stone composition				
Assessing deep learning methods for the identification of kidney stones in endoscopic images (2021)/On the in Vivo Recognition of Kidney Stones using Machine Learning (2024)	Lopez-Tiro et al. [48,49]	98% sensitivity, 97% precision (InceptionV3)	177 frames (90 surface, 87 cross-section) and then 181 frames	Two studies by the same author on the same dataset. Does not investigate fewer common classes of stones nor stones of mixed composition. Second study compares different underlying CNN architecture with InceptionV3 having best results.

Table 2. *Cont.*

Study Name [Year]	Authors	Endpoint Results [CNN Framework]	Frames Used and Objective of CNN	Comments
Towards Automatic Recognition of pure and mixed stones using intra-operative endoscopic images (2021)	Estrade et al. [50]	91–94% sensitivity, 90–93% specificity, 90–94% accuracy for calcium oxalate monohydrate 69–77% sensitivity, 95–97% specificity, 93–94% accuracy for calcium oxalate dihydrate 60–98% sensitivity, 97–99% specificity, 95–99% accuracy for uric acid (ResNet)	583 frames—to train CNN on stone composition	Recognition of 3 stone types (uric acid, calcium oxalate monohydrate, calcium oxalate dihydrate) in both pure and mixed forms during endoscopy, as these are most encountered types. Model perceived as limited as trained on clear and steady frames, which is very different to the clinical setting.
Deep morphological recognition of kidney stones using endoscopic digital videos (2022)	Estrade et al. [51]	80% sensitivity, 95% specificity, 88% accuracy, 0.78 F1-score averaged over both pure and mixed stone types (ResNet)	585 frames to train, but tested on 71 videos—to train CNN on stone composition	Application of previously created model by Estrade et al. [50] in clinical setting. Proved that did not require human intervention for stone delineation, or selection of good quality steady frames.
Boosting Kidney Stone Identification in Endoscopic Images Using Two-Step Transfer Learning (2023)	Lopez-Tiro et al. [52]	85.6% sensitivity, 85.6% accuracy, 0.854 F1-score on average for all stone types and stone images (ResNet)	Ex vivo dataset: 366 frames, 409 endoscopic frames	This study shows that pre-training images on ImageNet first, then ex-vivo (clear and detailed) images prior to endoscopic images, improved metrics in classifying types of stone.
Predicting stone composition via machine-learning models trained on intra-operative endoscopic digital images (2024)	Zhu et al. [53]	Overall sensitivity 99% Composition prediction rates: 99% calcium oxalate monohydrate, 100% calcium oxalate dihydrate, 100% apatite carbonate, 98% uric acid (ResNet)	1658 frames—to train CNN on stone composition	Like the second study by Estrade et al. [51], this work analysed intra-operative stone images in real time. This meant that the surgeon was capable of adjusting the laser frequency and energy according to the results provided.
Studies assessing tracking of stones and stone fragments				
Multi-class motion-based semantic segmentation for ureteroscopy and laser lithotripsy (2022)	Gupta et al. [54]	88.4/91.8% sensitivity, 0.87/0.92 dice coefficient for stone/laser device, (HybResUNet + DVFNet)	Images from animal studies performed at Boston Scientific and images from Oxford University Hospitals; 86 clips in vitro, 154 clips in vivo, 20 unseen clips—to train CNN on stone segmentation and localisation	Multi-class segmentation model (end-to-end CNN) that uses temporal information within 5 adjacent frames to improve segmentation results. Meticulous training, validation, and test sets in vitro, in vivo, and with previously unseen samples.

Table 2. *Cont.*

Study Name [Year]	Authors	Endpoint Results [CNN Framework]	Frames Used and Objective of CNN	Comments
Computer Vision Enabled Segmentation of Kidney Stones During Ureteroscopy and Laser Lithotripsy (2023)	Setia et al. [55]	94% sensitivity, 92% accuracy digital scope 52% sensitivity, 73% accuracy during fragmentation 41% sensitivity, 80% accuracy during dusting (U-Net++)	578 frames—to train CNN on stone segmentation and localisation	Demonstrated good segmentation performance during stone treatment and ability to track stone fragments less than 270 micrometres in diameter.
Study that assesses both stone composition and segmentation of stones				
Development of UroSAM: A Machine Learning Model to Automatically Identify Kidney Stone Composition from Endoscopic Video (2024)	Leng et al. [56]	94.77% precision, 0.9135 dice co-efficient for segmentation of images 60% correct classification of stone composition on average, with highest of 84.4% for calcium oxalate monohydrate (UNet)	1677 frames—to train CNN on stone composition and segmentation/ localisation	Software named UroSAM, that is unlike any of the other studies in Table 2; both segments and classifies composition of stones.

3.2. Intra-Operative Identification of Anatomy or Surgical Instruments Aided by AI

3.2.1. Cystoscopy

The majority of literature using AI during cystoscopy has focussed on diagnosis of bladder cancer, as stated earlier. The study by Eminaga et al. did show, however, it is possible to classify other cystoscopic findings with AI assistance, such as bladder stones [24].

Another area of interest is being able to identify the ureteral orifices (UOs) at the time of cystoscopy. Liu et al. developed a system that identifies and tracks UOs during cystoscopy with up to 92.2% sensitivity and 90.2% accuracy [57]. Liang et al. developed a similar system with a sensitivity of 79.8% and 90.2% precision [58]. These systems can be used to train novice urologists and act as a safety measure to preserve the UOs during endoscopic surgery, as well as being a first step towards machine-guided ureteric stent insertion.

3.2.2. Ureteroscopy

During ureteroscopy, it is important to correctly identify the lumen of the ureter to safely guide the surgeon during retrograde access of the upper tract. Accordingly, Lazo et al. developed a CNN that could recognise the ureteric lumen with a dice coefficient of 0.80 [59]. Importantly, because the model is spatio-temporal based and used multi-frame information, segmentation was effective in the presence of poor visibility, occasional bleeding, or specular reflections [59].

3.2.3. Radical Prostatectomy

Identification of the neurovascular bundles, as well other posteriorly located anatomical structures, is critical during radical prostatectomy. AI has significant potential to help with identifying these structures during robotic-assisted radical prostatectomy (RARP), as anatomical images or models can be displayed onto the surgical console easily.

Tanzi et al. were able to overlay 3D AR anatomical images onto the surgeon's console at the time of RARP, primarily through AI-based identification of the urinary catheter, so that they could accurately excise tissue for additional margins following removal of the

prostate [60]. The same investigators then used this technology to assess positive surgical margin rates when they attempted to nerve spare suspected pT3 prostate cancer cases, which had extra-capsular extension on MRI [61]. They showed that 14/16 suspected pT3 cases correlated to true pathological pT3 prostate cancer, but importantly, this translated to only 7.1% positive surgical margin rates (Gleason 3 + 3 highest score). Therefore, with the use of AI, they were able to nerve spare radiological pT3 prostate cancer without compromising oncological safety.

Takeshita et al. created a model that accurately recognises the seminal vesicles and vas deferens when the posterior approach for RARP is performed. They report that this could reduce complications and assist surgeons, particularly novice ones [62]. Park et al. showed that it is possible to not only identify the seminal vesicles and vas deferens but also identify surgical instruments, bladder, and prostate during RARP [63].

These studies reveal that intraoperative AI assistance during RARP may improve surgeons' ability to delineate anatomy and surgical instruments. This may translate to improved surgical margins and reduce complications by avoiding iatrogenic injury to important structures.

3.2.4. Partial Nephrectomy

The majority of partial nephrectomies (PNs) are performed laparoscopically, either with or without robotic assistance. Therefore, similarly to RARP, anatomical models or images can be displayed on the surgical screen or surgical console.

Zhang et al. created a model that enhances a surgeon's intra-operative perception during PN by overlaying 3D models onto the video screen [64]. They created the model from a pre-operative CT and used a CNN, trained on previous surgeries, to segment the renal surfaces in real time so that the model can be displayed accurately. This has similarly been performed by Piana et al. and Amparore et al., importantly showing that it is possible to overlay 3D models without the use of exogenous contrast or expensive tracking devices [20,21]. Padovan et al. were able to use the same AI-based program to overlay 3D anatomical models during PN, as well as RARP [65]. De Backer et al. used AI to prevent their 3D anatomical model overlay from obstructing the view of surgical instruments on the robotic console; however, they were limited by a 0.5 s delay, due to computer processing time (Figure 2) [66]. Given that temporal consistency is quite important with a dynamic operating view, Grammatikopoulou et al. successfully segmented anatomical surfaces during PN using AI based on a spatio-temporal network, but their overlay is 2D, unlike the other studies previously mentioned [67]. Zhang et al., in a further study, employed CNNs to automate and enhance the accuracy of segmenting kidneys, tumours, and renal artery trees from medical images [68]. These segmentations served as crucial inputs for the subsequent construction of a 3D model, which surgeons could use to improve their precision in segmental renal artery clamping during PN [68].

Given the time pressure experienced by surgeons during PN, the ability to better clarify renal anatomy with the help of AI could lead to improved surgical margins and possibly allow for surgeons to operate faster; these potential benefits can be assessed in future studies.

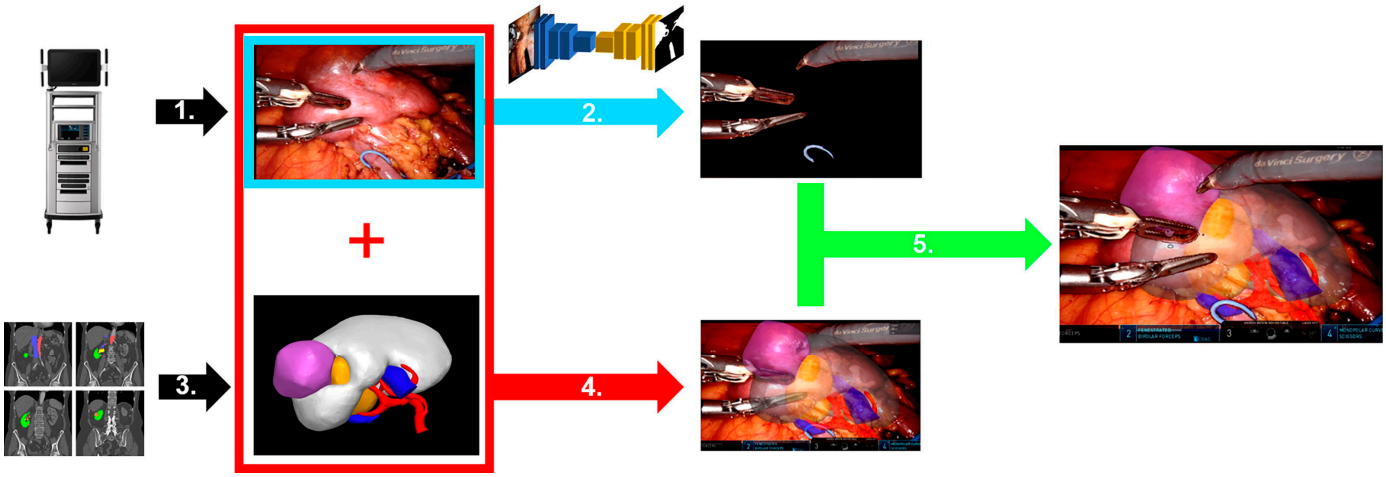

Figure 2. Overlay of 3D anatomical model during PN with delineation of robotic arm instruments as described by De Backer et al. [66] (reproduced with permission).

3.3. Intra-Operative Evaluation of Surgeon Performance and Prediction of Post-Operative Outcomes with the Aid of AI

Surgeons can be assessed through analysis of data derived from surgical instruments. Hung et al. were the first to show that measurement of automated performance metrics (APMs) during prostatectomy predicted functional outcomes, specifically urinary continence [69]. APM data were derived from the robotic surgery console and included instrument and endoscopic camera tracking, as well as events data, such as energy usage [69]. As suturing technique for vesico-urethral anastomosis is integral for continence, Luongo et al. were able to train AI to recognise classifications of suturing gestures during RARP [70]. Cui et al. went further by correlating suturing sub-skills to a skill assessment with included continence outcomes [71].

Surgeons can also be assessed through the analysis of surgical images and videos, which AI can assist with interpreting. Baghdadi et al. have developed an automated system to assess surgical performance of pelvic lymph node dissection by using computer vision to assess objective findings such as vessels and nerves, cleared area, and change in median colour map [72]. This data were then used to train a logistic regression model and were compared to the standardised pelvic lymphadenectomy scoring system, achieving 83% accuracy. It is also significant as the principles can be extrapolated to all robotic surgery, as well as when similar lymph node dissection is performed at the time of prostatectomy. Similarly, Bakker et al. and Nakamura et al. both found that using machine learning to estimate urethral length during prostatectomy correlated with continence outcomes post-prostatectomy [73,74]. Both of these studies are limited in that they do not have long-term follow-up and lack variance of surgeons, but they show that this measurement has the potential to be a marker of performance and help counsel patients. Deol et al. took a different approach to the aforementioned studies in that they used AI to recognise the surgical steps during bladder tumour resection for training and documentation purposes [75].

The surgical field is not the only thing that can be examined intra-operatively. Kumar et al. used AI to assess postural ergonomics during laparoscopic urologic procedures through the analysis of images of the surgeons themselves [76]. This shows the potential for AI to prevent musculo-skeletal injuries of surgeons.

3.4. Intra-Operative Detection of Complications with the Aid of AI

Having systems that automatically detect impending intra-operative complications is achievable automatically and in real time, with the assistance of AI. Jeong et al. used

machine learning to create a warning system with over 95% accuracy to warn a surgeon of tissue damage from the laser [77]. Checucci et al. developed BLAIR, a system to predict bleeding occurrence during RARP with over 90% accuracy [78,79]. Ayala et al. used spectral imaging as a non-invasive method to monitor ischemia during PN [80]. The system utilises a deep-learning-based algorithm to detect ischemia by identifying spectral outliers, eliminating the need for contrast agents and enabling continuous monitoring. As similar technologies such as these are developed, surgeons will be able to recognise intra-operative complications earlier, or recognise complications previously not perceived.

4. Future Considerations

It is likely that urologic surgeries will continue to have a significant amount of research investigating the use of AI to help improve patient outcomes, decrease surgeon cognitive load, and reduce human error, without significant use of resources for each patient. There are some studies in this review where there is a perceived benefit in these metrics, but it was not actually measured. There should be greater focus on ensuring there is an actual benefit from adopting the assistance of AI-based systems. It is equally important that appropriate ethical measures including privacy and consent are maintained by surgeons when using such systems. Figure 3 encapsulates the current potential intra-operative uses of AI in urologic surgery. We believe that there is potential for further research into adopting AI-based systems into percutaneous nephrolithotomy and robotic endoscopic stone surgery.

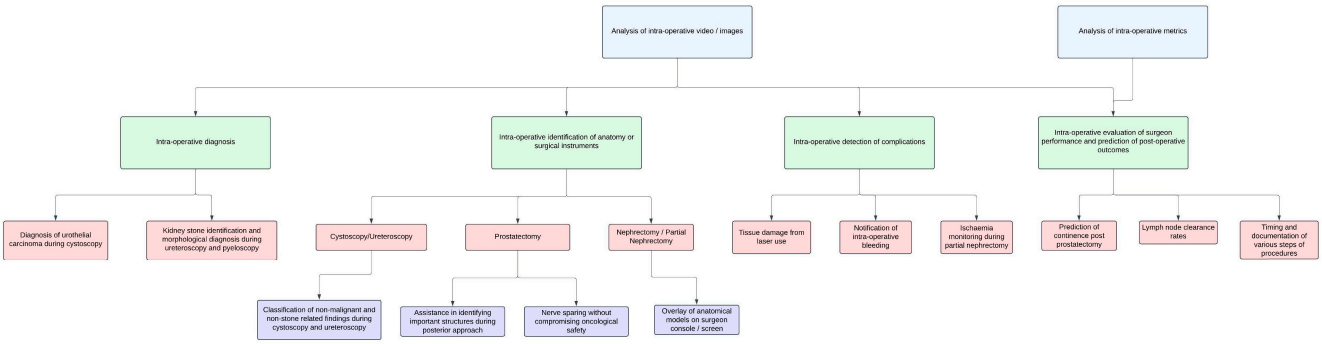

Figure 3. Summary of current or potential intra-operative aids to urologists.

5. Conclusions

AI holds great potential for urologists during surgery to enhance safety, accuracy in diagnosis, identification of anatomical structures and surgical instruments, and assessment of the surgeon for self-evaluation, as well as to predict post-operative outcomes. There needs to be more prospective studies to evaluate its real-world application, feasibility, and costs before it can be widely adopted during surgery. Future programs might improve by pre-training models on relevant large datasets, using attention mechanisms for efficiency purposes, using spatio-temporal information to improve metrics, and selecting the right CNN framework for optimal results.

Author Contributions: A.G.—methodology, data curation/investigation (major), formal analysis, writing—original draft, writing—review and editing, project administration; A.A.-K.—data curation/investigation (minor), writing—review and editing; C.E.D.—writing—review and editing; O.P.—writing—review and editing; A.T.—writing—review and editing; J.I.—supervision, conceptualisation, methodology, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors acknowledge open access provided by University of Melbourne and the Austin Health Library.

Conflicts of Interest: The authors declare no conflicts of interest.

List of Abbreviations

AI	artificial intelligence
AR	augmented reality
VR	virtual reality
ANN	artificial neural network
UC	urothelial carcinoma
CNN	convolutional neural network
r-CNN	regional-based convolutional neural network
RGB	red-green-blue
HSV	hue saturation value
WLI	white light imaging
NBI	narrow band imaging
MIOU	mean intersection over union
RARP	robotic-assisted radical prostatectomy
APM	automated performance metric
PN	partial nephrectomy

Appendix A. [81,82]

Note: All definitions are in the context of this review

Accuracy—Proportion or percentage of correct classifications.

Artificial neural network (ANN)—Networks that consist of inter-connected nodes that are supposed to simulate the function of human neurons. Each connection, imitating a synapse, has a weight, and the network learns by adjusting these weights to recognise patterns in data.

Attention-based mechanisms—Feature that can be applied to CNNs such that the network selectively emphasises certain aspects of an image. In the context of bladder cancer, attention-based mechanisms could be applied to place more weight on parts of the image with papillary architecture, or similarly place more emphasis in identifying tumours in areas with unusual colourations. This also usually means software can be created with less computational requirements to achieve the same result.

Class—Refers to a category or label assigned to parts of an image or frame. A class could include anatomical structures, surgical tools, or even intra-operative events. Classes can be further sub-divided into sub-classes. For example, bladder tumour could be further sub-divided into a high-grade or low-grade sub-class.

Computer vision—Machine learning (ML) and artificial intelligence (AI) that focusses on replicating the ability of human vision. Therefore, with respect to medicine, this encompasses tasks such as image processing, as well as the identification of anatomy and detection of instruments.

Convolutional neural network (CNN)—A type of ANN that is highly effective in interpreting visual data. The mathematical operation of convolution allows the network to detect features in images. Ground truth masks are used to train the CNN during the training phase to predict classes and generate class-specific masks during testing, validation, and when deployed.

Deep learning—Type of machine learning that uses artificial neural networks to detect complex and intricate patterns.

Dice coefficient—Used to calculate how much overlap there is between two datasets, with values ranging between 0 (no overlap) and 1 (complete overlap). The formula used is two times the area of overlap divided by the combined area of prediction and area of ground truth (refer to Figure 1 below for visual representation of these). This could be used to compare AI's prediction of a bladder tumour on an image, or to the bladder tumour labelled by the surgeon on an image. Please refer to mIoU below for the difference between the dice coefficient and mean intersection over union.

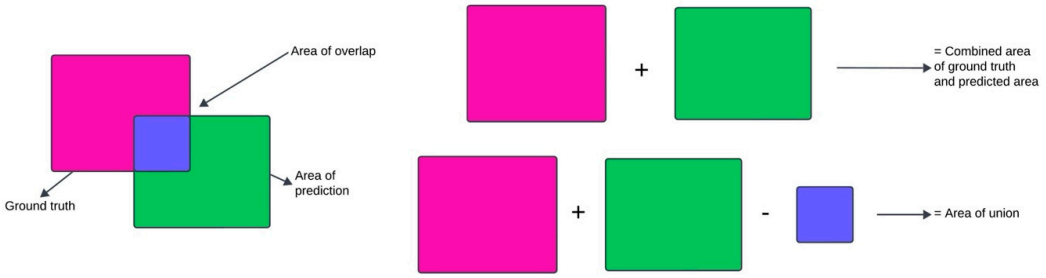

Figure A1. Visual representation of combined area of ground truth and predicted truth, as well as area of union.

Frame—Single still image (2D) extracted from a video sequence.

Ground truth—Refers to the accurate annotation of intra-operative images or frames performed by a human.

Machine learning—Branch of AI that focusses on using data and algorithms to make future predictions, based on learning from previous data. There are many types of machine learning.

Mask—Image where each pixel within the image corresponds to a particular class. In the context of this review, a mask will be based on a single frame. For example, a mask based on a cystoscopic frame would have all pixels labelled as either benign findings (e.g., squamous metaplasia), normal mucosa, or as bladder tumour. Masks can be labelled with a bounding box, which is a rectangular box over the area of interest. Otherwise, it can be labelled with a polygon, which defines the exact shape of the area of interest. The latter generally provides better results but requires more manual effort and increased computational requirements.

Mean intersection over union (mIoU)—Similar to the dice coefficient, also used to define the area of overlap between ground truth and area of prediction ranging between 0 (no overlap) and 1 (complete overlap). The formula to calculate the intersection over union (IoU) is the area of overlap divided by the area of union between the ground truth and area of prediction (refer to Figure 1 above for visual representation). The mean IoU for all classes in a frame is the mIoU. Thus, the denominator used in the formula to calculate IoU and dice coefficient differ. It is important to note because of this difference, the dice coefficient provides more weight to the overlapping region, and mIoU can be affected when the area of prediction is larger or smaller than the ground truth.

Predicted segmentation—Prediction of objects or areas in an image by AI, based on previous learning and training.

Precision—How often the correct class is positively identified.

Recall—Same measure as sensitivity.

Regional-based convolutional neural network (r-CNN)—Combines the power of a CNN with a regional proposal method. A region where an object may be located is initially identified, and then features of this region are extracted and input into a pretrained CNN, which then classifies and further localises the object.

References

1. Saeidi, H.; Opfermann, J.D.; Kam, M.; Wei, S.; Leonard, S.; Hsieh, M.H.; Kang, J.U.; Krieger, A. Autonomous robotic laparoscopic surgery for intestinal anastomosis. *Sci. Robot.* **2022**, *7*, eabj2908. [CrossRef] [PubMed]
2. Guni, A.; Varma, P.; Zhang, J.; Fehervari, M.; Ashrafian, H. Artificial Intelligence in Surgery: The Future is Now. *Eur. Surg. Res.* 2024, *in press*. [CrossRef] [PubMed]
3. Kamitani, Y.; Nonaka, K.; Isomoto, H. Current Status and Future Perspectives of Artificial Intelligence in Colonoscopy. *J. Clin. Med.* **2022**, *11*, 2923. [CrossRef] [PubMed]
4. Fernicola, A.; Palomba, G.; Capuano, M.; De Palma, G.D.; Aprea, G. Artificial intelligence applied to laparoscopic cholecystectomy: What is the next step? A narrative review. *Updates Surg.* **2024**, *76*, 1655–1667. [CrossRef]
5. Wang, P.; Xiao, X.; Glissen Brown, J.R.; Berzin, T.M.; Tu, M.; Xiong, F.; Hu, X.; Liu, P.; Song, Y.; Zhang, D.; et al. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. *Nat. Biomed. Eng.* **2018**, *2*, 741–748. [CrossRef]
6. Luo, H.; Xu, G.; Li, C.; He, L.; Luo, L.; Wang, Z.; Jing, B.; Deng, Y.; Jin, Y.; Li, Y.; et al. Real-time artificial intelligence for detection of upper gastrointestinal cancer by endoscopy: A multicentre, case-control, diagnostic study. *Lancet Oncol.* **2019**, *20*, 1645–1654. [CrossRef]
7. Tangsrivimol, J.A.; Schonfeld, E.; Zhang, M.; Veeravagu, A.; Smith, T.R.; Härtl, R.; Lawton, M.T.; El-Sherbini, A.H.; Prevedello, D.M.; Glicksberg, B.S.; et al. Artificial Intelligence in Neurosurgery: A State-of-the-Art Review from Past to Future. *Diagnostics* **2023**, *13*, 2429. [CrossRef]
8. Chen, H. Application progress of artificial intelligence and augmented reality in orthopaedic arthroscopy surgery. *J. Orthop. Surg. Res.* **2023**, *18*, 775. [CrossRef]
9. Farhadi, F.; Barnes, M.R.; Sugito, H.R.; Sin, J.M.; Henderson, E.R.; Levy, J.J. Applications of artificial intelligence in orthopaedic surgery. *Front. Med. Technol.* **2022**, *4*, 995526. [CrossRef]
10. Kalia, M.; Avinash, A.; Navab, N.; Salcudean, S. Preclinical evaluation of a markerless, real-time, augmented reality guidance system for robot-assisted radical prostatectomy. *Int. J. Comput. Assist. Radiol. Surg.* **2021**, *16*, 1181–1188. [CrossRef]
11. Noël, J.; Moschovas, M.C.; Patel, E.; Rogers, T.; Marquinez, J.; Rocco, B.; Mottrie, A.; Patel, V. Step-by-step optimisation of robotic-assisted radical prostatectomy using augmented reality. *Int. Braz. J. Urol.* **2022**, *48*, 600–601. [CrossRef] [PubMed]
12. Checcucci, E.; Amparore, D.; Volpi, G.; De Cillis, S.; Piramide, F.; Verri, P.; Piana, A.; Sica, M.; Gatti, C.; Alessio, P.; et al. Metaverse Surgical Planning with Three-dimensional Virtual Models for Minimally Invasive Partial Nephrectomy. *Eur. Urol.* **2024**, *85*, 320–325. [CrossRef] [PubMed]
13. Khor, W.S.; Baker, B.; Amin, K.; Chan, A.; Patel, K.; Wong, J. Augmented and virtual reality in surgery-the digital surgical environment: Applications, limitations and legal pitfalls. *Ann. Transl. Med.* **2016**, *4*, 454. [CrossRef] [PubMed]
14. Amparore, D.; Piramide, F.; Checcucci, E.; Verri, P.; De Cillis, S.; Piana, A.; Volpi, G.; Busacca, G.; Colombo, M.; Fiori, C.; et al. Three-dimensional Virtual Models of the Kidney with Colored Perfusion Regions: A New Algorithm-based Tool for Optimizing the Clamping Strategy During Robot-assisted Partial Nephrectomy. *Eur. Urol.* **2023**, *84*, 418–425. [CrossRef]
15. Michael, P.; Kevin, C.Z.; Jean-Philippe, T.; Joannie, D.; Frédérick, D.; Kelly, A.; Eric, T.M.; Chris, K.; Frederic, L.; Dominique, T.; et al. Integration of a Raman spectroscopy system to a robotic-assisted surgical system for real-time tissue characterization during radical prostatectomy procedures. *J. Biomed. Opt.* **2019**, *24*, 025001. [CrossRef]
16. Amara, N.; Al Youssef, T.; Massa, J.; Fidjel, A.; Khoury, E.E.; Patel, B.; Flais, M.; Deswarte, C. Intraoperative angiography of the neurovascular bundle using indocyanine green and near-infrared fluorescence improves anatomical dissection during robot-assisted radical prostatectomy: Initial clinical experience. *J. Robot. Surg.* **2023**, *17*, 687–694. [CrossRef]
17. Cahill, E.M.; Chua, K.; Doppalapudi, S.K.; Ghodoussipour, S. The use of blue-light cystoscopy in the detection and surveillance of nonmuscle invasive bladder cancer. *Curr. Urol.* **2022**, *16*, 121–126. [CrossRef]
18. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [CrossRef]
19. Navarrete-Welton, A.J.; Hashimoto, D.A. Current applications of artificial intelligence for intraoperative decision support in surgery. *Front. Med.* **2020**, *14*, 369–381. [CrossRef]
20. Piana, A.; Amparore, D.; Sica, M.; Volpi, G.; Checcucci, E.; Piramide, F.; De Cillis, S.; Busacca, G.; Scarpelli, G.; Sidoti, F.; et al. Automatic 3D Augmented-Reality Robot-Assisted Partial Nephrectomy Using Machine Learning: Our Pioneer Experience. *Cancers* **2024**, *16*, 1047. [CrossRef]
21. Amparore, D.; Sica, M.; Verri, P.; Piramide, F.; Checcucci, E.; De Cillis, S.; Piana, A.; Campobasso, D.; Burgio, M.; Cisero, E.; et al. Computer Vision and Machine-Learning Techniques for Automatic 3D Virtual Images Overlapping During Augmented Reality Guided Robotic Partial Nephrectomy. *Technol. Cancer Res. Treat.* **2024**, *23*. [CrossRef] [PubMed]
22. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Syst. Rev.* **2021**, *10*, 89. [CrossRef] [PubMed]

23. Lu, D.; Reed, A.; Pace, N.; Luckenbaugh, A.N.; Pallauf, M.; Singla, N.; Oguz, I.; Kavoussi, N. Automated Upper Tract Urothelial Carcinoma Tumor Segmentation During Ureteroscopy Using Computer Vision Techniques. *J. Endourol.* **2024**, *38*, 836–842. [CrossRef] [PubMed]
24. Eminaga, O.; Eminaga, N.; Semjonow, A.; Breil, B. Diagnostic classification of cystoscopic images using deep convolutional neural networks. *JCO Clin. Cancer Inform.* **2018**, *2018*, 1–8. [CrossRef]
25. Shkolyar, E.; Jia, X.; Chang, T.C.; Trivedi, D.; Mach, K.E.; Meng, M.Q.; Xing, L.; Liao, J.C. Augmented Bladder Tumor Detection Using Deep Learning. *Eur. Urol.* **2019**, *76*, 714–718. [CrossRef]
26. Ikeda, A.; Nosato, H.; Kochi, Y.; Kojima, T.; Kawai, K.; Sakanashi, H.; Murakawa, M.; Nishiyama, H. Support System of Cystoscopic Diagnosis for Bladder Cancer Based on Artificial Intelligence. *J. Endourol.* **2020**, *34*, 352–358. [CrossRef]
27. Ikeda, A.; Nosato, H.; Kochi, Y.; Negoro, H.; Kojima, T.; Sakanashi, H.; Murakawa, M.; Nishiyama, H. Cystoscopic Imaging for Bladder Cancer Detection Based on Stepwise Organic Transfer Learning with a Pretrained Convolutional Neural Network. *J. Endourol.* **2020**, *35*, 1030–1035. [CrossRef]
28. Du, Y.; Yang, R.; Chen, Z.; Wang, L.; Weng, X.; Liu, X. A deep learning network-assisted bladder tumour recognition under cystoscopy based on Caffe deep learning framework and EasyDL platform. *Int. J. Med. Robot.* **2021**, *17*, 1–8. [CrossRef]
29. Varnyu, D.; Szirmay-Kalos, L. A Comparative Study of Deep Neural Networks for Real-Time Semantic Segmentation during the Transurethral Resection of Bladder Tumors. *Diagnostics* **2022**, *12*, 2849. [CrossRef]
30. Mutaguchi, J.; Morooka, K.; Kobayashi, S.; Umehara, A.; Miyauchi, S.; Kinoshita, F.; Inokuchi, J.; Oda, Y.; Kurazume, R.; Eto, M. Artificial Intelligence for Segmentation of Bladder Tumor Cystoscopic Images Performed by U-Net with Dilated Convolution. *J. Endourol.* **2022**, *36*, 827–834. [CrossRef]
31. Zhang, Q.; Liang, Y.; Zhang, Y.; Tao, Z.; Li, R.; Bi, H. A comparative study of attention mechanism based deep learning methods for bladder tumor segmentation. *Int. J. Med. Inf.* **2023**, *171*, 104984. [CrossRef] [PubMed]
32. Jia, X.; Shkolyar, E.; Laurie, M.A.; Eminaga, O.; Liao, J.C.; Xing, L. Tumor detection under cystoscopy with transformer-augmented deep learning algorithm. *Phys. Med. Biol.* **2023**, *68*. [CrossRef] [PubMed]
33. Chang, T.C.; Shkolyar, E.; Del Giudice, F.; Eminaga, O.; Lee, T.; Laurie, M.; Seufert, C.; Jia, X.; Mach, K.E.; Xing, L.; et al. Real-time Detection of Bladder Cancer Using Augmented Cystoscopy with Deep Learning: A Pilot Study. *J. Endourol.* 2023, *in press*. [CrossRef] [PubMed]
34. Guo, Y.; Li, C.; Zhang, S.; Zhu, G.; Sun, L.; Jin, T.; Wang, Z.; Li, S.; Zhou, F. U-Net-Based Assistive Identification of Bladder Cancer: A Promising Approach for Improved Diagnosis. *Urol. Int.* **2024**, *108*, 100–107. [CrossRef]
35. Zhao, X.; Lai, L.; Li, Y.; Zhou, X.; Cheng, X.; Chen, Y.; Huang, H.; Guo, J.; Wang, G. A lightweight bladder tumor segmentation method based on attention mechanism. *Med. Biol. Eng. Comput.* **2024**, *62*, 1519–1534. [CrossRef]
36. Lee, J.Y.; Lee, Y.S.; Tae, J.H.; Chang, I.H.; Kim, T.-H.; Myung, S.C.; Nguyen, T.T.; Lee, J.H.; Choi, J.; Kim, J.H.; et al. Selection of Convolutional Neural Network Model for Bladder Tumor Classification of Cystoscopy Images and Comparison with Humans. *J. Endourol.* **2024**, *38*, 1036–1043. [CrossRef]
37. Freitas, N.R.; Vieira, P.M.; Lima, E.; Lima, C.S. Automatic T1 bladder tumor detection by using wavelet analysis in cystoscopy images. *Phys. Med. Biol.* **2018**, *63*, 035031. [CrossRef]
38. Yoo, J.W.; Koo, K.C.; Chung, B.H.; Baek, S.Y.; Lee, S.J.; Park, K.H.; Lee, K.S. Deep learning diagnostics for bladder tumor identification and grade prediction using RGB method. *Sci. Rep.* **2022**, *12*, 17699. [CrossRef]
39. Ali, N.; Bolenz, C.; Todenhofer, T.; Stenzel, A.; Deetmar, P.; Kriegmair, M.; Knoll, T.; Porubsky, S.; Hartmann, A.; Popp, J.; et al. Deep learning-based classification of blue light cystoscopy imaging during transurethral resection of bladder tumors. *Sci. Rep.* **2021**, *11*, 11629. [CrossRef]
40. Wu, S.; Chen, X.; Pan, J.; Dong, W.; Diao, X.; Zhang, R.; Zhang, Y.; Zhang, Y.; Qian, G.; Chen, H.; et al. An Artificial Intelligence System for the Detection of Bladder Cancer via Cystoscopy: A Multicenter Diagnostic Study. *J. Natl. Cancer Inst.* **2022**, *114*, 220–227. [CrossRef]
41. Yang, R.; Du, Y.; Weng, X.; Chen, Z.; Wang, S.; Liu, X. Automatic recognition of bladder tumours using deep learning technology and its clinical application. *Int. J. Med. Robot. Comput. Assist. Surg.* **2021**, *17*, e2194. [CrossRef] [PubMed]
42. Eminaga, O.; Lee, T.J.; Laurie, M.; Ge, T.J.; La, V.; Long, J.; Semjonow, A.; Bogemann, M.; Lau, H.; Shkolyar, E.; et al. Efficient Augmented Intelligence Framework for Bladder Lesion Detection. *JCO Clin. Cancer Inform.* **2023**, *7*, e2300031. [CrossRef] [PubMed]
43. Serrat, J.; Lumbreras, F.; Blanco, F.; Valiente, M.; López-Mesas, M. myStone: A system for automatic kidney stone classification. *Expert Syst. Appl.* **2017**, *89*, 41–51. [CrossRef]
44. Black, K.M.; Law, H.; Aldoukhi, A.; Deng, J.; Ghani, K.R. Deep learning computer vision algorithm for detecting kidney stone composition. *BJU Int.* **2020**, *125*, 920–924. [CrossRef]
45. Kim, U.S.; Kwon, H.S.; Yang, W.; Lee, W.; Choi, C.; Kim, J.K.; Lee, S.H.; Rim, D.; Han, J.H. Prediction of the composition of urinary stones using deep learning. *Investig. Clin. Urol.* **2022**, *63*, 441–447. [CrossRef]

46. El Beze, J.; Mazeaud, C.; Daul, C.; Ochoa-Ruiz, G.; Daudon, M.; Eschwège, P.; Hubert, J. Evaluation and understanding of automated urinary stone recognition methods. *BJU Int.* **2022**, *130*, 786–798. [CrossRef]
47. Martínez, A.; Trinh, D.-H.; El Beze, J.; Hubert, J.; Eschwege, P.; Estrade, V.; Aguilar, L.; Daul, C.; Ochoa, G. Towards an automated classification method for ureteroscopic kidney stone images using ensemble learning. In Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, QC, Canada, 20–24 July 2020; pp. 1936–1939.
48. Lopez-Tiro, F.; Varelo, A.; Hinojosa, O.; Mendez Ruiz, M.; Trinh, H.; ElBeze, Y.; Hubert, J.; Estrade, V.; Gonzalez-Mendoza, M.; Ochoa-Ruiz, G.; et al. Assessing deep learning methods for the identification of kidney stones in endoscopic images. In Proceedings of the 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society, EMBC 2021, Virtual, 1–5 November 2021; Volume 2021, pp. 2778–2781.
49. Lopez-Tiro, F.; Estrade, V.; Hubert, J.; Flores-Araiza, D.; Gonzalez-Mendoza, M.; Ochoa, G.; Daul, C. On the In Vivo Recognition of Kidney Stones Using Machine Learning. *IEEE Access* **2024**, *12*, 10736–10759. [CrossRef]
50. Estrade, V.; Daudon, M.; Richard, E.; Bernhard, J.C.; Bladou, F.; Robert, G.; Denis de Senneville, B. Towards automatic recognition of pure and mixed stones using intra-operative endoscopic digital images. *BJU Int.* **2022**, *129*, 234–242. [CrossRef]
51. Estrade, V.; Daudon, M.; Richard, E.; Bernhard, J.C.; Bladou, F.; Robert, G.; Facq, L.; Denis de Senneville, B. Deep morphological recognition of kidney stones using intra-operative endoscopic digital videos. *Phys. Med. Biol.* **2022**, *67*, 16. [CrossRef]
52. Lopez-Tiro, F.; Flores-Araiza, D.; Betancur-Rengifo, J.P.; Reyes-Amezcua, I.; Hubert, J.; Ochoa-Ruiz, G.; Daul, C. *Boosting Kidney Stone Identification in Endoscopic Images Using Two-Step Transfer Learning*; Springer Nature: Cham, Switzerland, 2024; pp. 131–141.
53. Zhu, G.; Li, C.; Guo, Y.; Sun, L.; Jin, T.; Wang, Z.; Li, S.; Zhou, F. Predicting stone composition via machine-learning models trained on intra-operative endoscopic digital images. *BMC Urol.* **2024**, *24*, 5. [CrossRef]
54. Gupta, S.; Ali, S.; Goldsmith, L.; Turney, B.; Rittscher, J. Multi-class motion-based semantic segmentation for ureteroscopy and laser lithotripsy. *Comput. Med. Imaging Graph.* **2022**, *101*, 102112. [CrossRef] [PubMed]
55. Setia, S.A.; Stoebner, Z.A.; Floyd, C.; Lu, D.; Oguz, I.; Kavoussi, N.L. Computer Vision Enabled Segmentation of Kidney Stones During Ureteroscopy and Laser Lithotripsy. *J. Endourol.* **2023**, *37*, 495–501. [CrossRef] [PubMed]
56. Leng, J.; Liu, J.; Cheng, G.; Wang, H.; Quarrier, S.; Luo, J.; Jain, R. Development of UroSAM: A Machine Learning Model to Automatically Identify Kidney Stone Composition from Endoscopic Video. *J. Endourol.* **2024**, *31*, 31. [CrossRef]
57. Liu, D.; Peng, X.; Liu, X.; Li, Y.; Bao, Y.; Xu, J.; Bian, X.; Xue, W.; Qian, D. A real-time system using deep learning to detect and track ureteral orifices during urinary endoscopy. *Comput. Biol. Med.* **2021**, *128*, 104104. [CrossRef]
58. Liang, L.; Yuanjun, W. UO-YOLO: Ureteral Orifice Detection Network Based on YOLO and Biformer Attention Mechanism. *Appl. Sci.* **2024**, *14*, 5124. [CrossRef]
59. Lazo, J.F.; Marzullo, A.; Moccia, S.; Catellani, M.; Rosa, B.; de Mathelin, M.; De Momi, E. Using spatial-temporal ensembles of convolutional neural networks for lumen segmentation in ureteroscopy. *Int. J. Comput. Assist. Radiol. Surg.* **2021**, *16*, 915–922. [CrossRef]
60. Tanzi, L.; Piazzolla, P.; Porpiglia, F.; Vezzetti, E. Real-time deep learning semantic segmentation during intra-operative surgery for 3D augmented reality assistance. *Int. J. Comput. Assist. Radiol. Surg.* **2021**, *16*, 1435–1445. [CrossRef]
61. Checcucci, E.; Piana, A.; Volpi, G.; Piazzolla, P.; Amparore, D.; De Cillis, S.; Piramide, F.; Gatti, C.; Stura, I.; Bollito, E.; et al. Three-dimensional automatic artificial intelligence driven augmented-reality selective biopsy during nerve-sparing robot-assisted radical prostatectomy: A feasibility and accuracy study. *Asian J. Urol.* **2023**, *10*, 407–415. [CrossRef]
62. Takeshita, N.; Sakamoto, S.; Kitaguchi, D.; Takeshita, N.; Yajima, S.; Koike, T.; Ishikawa, Y.; Matsuzaki, H.; Mori, K.; Masuda, H.; et al. Deep Learning-Based Seminal Vesicle and Vas Deferens Recognition in the Posterior Approach of Robot-Assisted Radical Prostatectomy. *Urology* **2023**, *173*, 98–103. [CrossRef]
63. Gon Park, S.; Park, J.; Rock Choi, H.; Ho Lee, J.; Tae Cho, S.; Goo Lee, Y.; Ahn, H.; Pak, S. Deep Learning Model for Real-time Semantic Segmentation During Intraoperative Robotic Prostatectomy. *Eur. Urol. Open Sci.* **2024**, *62*, 47–53. [CrossRef]
64. Zhang, X.; Wang, J.; Wang, T.; Ji, X.; Shen, Y.; Sun, Z.; Zhang, X. A markerless automatic deformable registration framework for augmented reality navigation of laparoscopy partial nephrectomy. *Int. J. Comput. Assist. Radiol. Surg.* **2019**, *14*, 1285–1294. [CrossRef] [PubMed]
65. Padovan, E.; Marullo, G.; Tanzi, L.; Piazzolla, P.; Moos, S.; Porpiglia, F.; Vezzetti, E. A deep learning framework for real-time 3D model registration in robot-assisted laparoscopic surgery. *Int. J. Med. Robot. Comput. Assist. Surg.* **2022**, *18*, e2387. [CrossRef] [PubMed]
66. De Backer, P.; Van Praet, C.; Simoens, J.; Peraire Lores, M.; Creemers, H.; Mestdagh, K.; Allaeys, C.; Vermijs, S.; Piazza, P.; Mottaran, A.; et al. Improving Augmented Reality Through Deep Learning: Real-time Instrument Delineation in Robotic Renal Surgery. *Eur. Urol.* **2023**, *84*, 86–91. [CrossRef] [PubMed]
67. Grammatikopoulou, M.; Sanchez-Matilla, R.; Bragman, F.; Owen, D.; Culshaw, L.; Kerr, K.; Stoyanov, D.; Luengo, I. A spatio-temporal network for video semantic segmentation in surgical videos. *Int. J. Comput. Assist. Radiol. Surg.* **2024**, *19*, 375–382. [CrossRef] [PubMed]

68. Zhang, S.; Yang, G.; Qian, J.; Zhu, X.; Li, J.; Li, P.; He, Y.; Xu, Y.; Shao, P.; Wang, Z. A novel 3D deep learning model to automatically demonstrate renal artery segmentation and its validation in nephron-sparing surgery. *Front. Oncol.* **2022**, *12*, 997911. [CrossRef]
69. Hung, A.J.; Chen, J.; Ghodoussipour, S.; Oh, P.J.; Liu, Z.; Nguyen, J.; Purushotham, S.; Gill, I.S.; Liu, Y. A deep-learning model using automated performance metrics and clinical features to predict urinary continence recovery after robot-assisted radical prostatectomy. *BJU Int.* **2019**, *124*, 487–495. [CrossRef]
70. Luongo, F.; Hakim, R.; Nguyen, J.H.; Anandkumar, A.; Hung, A.J. Deep learning-based computer vision to recognize and classify suturing gestures in robot-assisted surgery. *Surgery* **2021**, *169*, 1240–1244. [CrossRef]
71. Cui, Z.; Ma, R.; Yang, C.H.; Malpani, A.; Chu, T.N.; Ghazi, A.; Davis, J.W.; Miles, B.J.; Lau, C.; Liu, Y.; et al. Capturing relationships between suturing sub-skills to improve automatic suturing assessment. *NPJ Digit. Med.* **2024**, *7*, 152. [CrossRef]
72. Baghdadi, A.; Hussein, A.A.; Ahmed, Y.; Cavuoto, L.A.; Guru, K.A. A computer vision technique for automated assessment of surgical performance using surgeons' console-feed videos. *Int. J. Comput. Assist. Radiol. Surg.* **2019**, *14*, 697–707. [CrossRef]
73. Bakker, A.; de Nijs, J.V.; Jaspers, T.J.M.; de With, P.H.N.; Beulens, A.J.W.; van der Poel, H.G.; van der Sommen, F.; Brinkman, W.M. Estimating Surgical Urethral Length on Intraoperative Robot-Assisted Prostatectomy Images Using Artificial Intelligence Anatomy Recognition. *J. Endourol.* **2024**, *38*, 690–696. [CrossRef]
74. Nakamura, W.; Sumitomo, M.; Zennami, K.; Takenaka, M.; Ichino, M.; Takahara, K.; Teramoto, A.; Shiroki, R. Combination of deep learning and ensemble machine learning using intraoperative video images strongly predicts recovery of urinary continence after robot-assisted radical prostatectomy. *Cancer Rep.* **2023**, *6*, e1861. [CrossRef] [PubMed]
75. Deol, E.S.; Tollefson, M.K.; Antolin, A.; Zohar, M.; Bar, O.; Ben-Ayoun, D.; Mynderse, L.A.; Lomas, D.J.; Avant, R.A.; Miller, A.R.; et al. Automated surgical step recognition in transurethral bladder tumor resection using artificial intelligence: Transfer learning across surgical modalities. *Front. Artif. Intell.* **2024**, *7*, 1375482. [CrossRef] [PubMed]
76. Kumar, A.P.; Sindhu, S.; Manangi, M.; Chikkanayakanahalli, S.S.; Venkatappa, S.K.; Naik, M.G.; Shivaprakash, N. An Adaptation of Computer Vision of Artificial Intelligence for the Assessment of Postural Ergonomics in Laparoscopic Surgery. *World J. Laparosc. Surg.* **2023**, *16*, 119–124. [CrossRef]
77. Jeong, J.; Chang, K.; Lee, J.; Choi, J. A warning system for urolithiasis via retrograde intrarenal surgery using machine learning: An experimental study. *BMC Urol.* **2022**, *22*, 80. [CrossRef] [PubMed]
78. Checcucci, E.; Piazzolla, P.; Marullo, G.; Innocente, C.; Salerno, F.; Ulrich, L.; Moos, S.; Quara, A.; Volpi, G.; Amparore, D.; et al. Development of Bleeding Artificial Intelligence Detector (BLAIR) System for Robotic Radical Prostatectomy. *J. Clin. Med.* **2023**, *12*, 7355. [CrossRef]
79. Checcucci, E.; De Cillis, S.; Amparore, D.; Gabriele, V.; Piramide, F.; Piana, A.; Fiori, C.; Piazzolla, P.; Porpiglia, F. Artificial Intelligence alert systems during robotic surgery: A new potential tool to improve the safety of the intervention. *Urol. Video J.* **2023**, *18*, 100221. [CrossRef]
80. Ayala, L.; Adler, T.J.; Seidlitz, S.; Wirkert, S.; Engels, C.; Seitel, A.; Sellner, J.; Aksenov, A.; Bodenbach, M.; Bader, P.; et al. Spectral imaging enables contrast agent-free real-time ischemia monitoring in laparoscopic surgery. *Sci. Adv.* **2022**, *9*, eadd6778. [CrossRef]
81. Russell, S.; Norvig, P. *Artificial Intelligence: A Modern Approach*; Pearson: London, UK, 2016.
82. Elgendy, M. *Deep Learning for Vision Systems*; Manning Publications: Shelter Island, NY, USA, 2020.