



## OPEN ACCESS

## EDITED BY

Philipp Taussky,  
Beth Israel Deaconess Medical Center and  
Harvard Medical School, United States

## REVIEWED BY

Prajwal Ghimire,  
King's College Hospital NHS Foundation Trust,  
United Kingdom  
Yaniv Zager,  
AdventHealth, United States

## \*CORRESPONDENCE

Mark C. Preul  
✉ [neuropub@barrowneuro.org](mailto:neuropub@barrowneuro.org)

RECEIVED 14 November 2024

ACCEPTED 31 January 2025

PUBLISHED 26 February 2025

## CITATION

Yangi K, On TJ, Xu Y, Gholami AS, Hong J,  
Reed AG, Puppalla P, Chen J, Tangsrivimol JA,  
Li B, Santello M, Lawton MT and Preul MC  
(2025) Artificial intelligence integration in  
surgery through hand and instrument tracking:  
a systematic literature review.  
*Front. Surg.* 12:1528362.  
doi: 10.3389/fsurg.2025.1528362

## COPYRIGHT

© 2025 Yangi, On, Xu, Gholami, Hong, Reed,  
Puppalla, Chen, Tangsrivimol, Li, Santello,  
Lawton and Preul. This is an open-access  
article distributed under the terms of the  
[Creative Commons Attribution License \(CC  
BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in  
other forums is permitted, provided the  
original author(s) and the copyright owner(s)  
are credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Artificial intelligence integration in surgery through hand and instrument tracking: a systematic literature review

Kivanc Yangi<sup>1</sup>, Thomas J. On<sup>1</sup>, Yuan Xu<sup>1</sup>, Arianna S. Gholami<sup>1</sup>, Jinpyo Hong<sup>1</sup>, Alexander G. Reed<sup>1</sup>, Pravarakhya Puppalla<sup>1</sup>, Jiuxu Chen<sup>1,2</sup>, Jonathan A. Tangsrivimol<sup>1</sup>, Baoxin Li<sup>2</sup>, Marco Santello<sup>3</sup>, Michael T. Lawton<sup>1</sup> and Mark C. Preul<sup>1\*</sup>

<sup>1</sup>The Loyal and Edith Davis Neurosurgical Research Laboratory, Department of Neurosurgery, Barrow Neurological Institute, St. Joseph's Hospital and Medical Center, Phoenix, AZ, United States, <sup>2</sup>School of Computing and Augmented Intelligence, Arizona State University, Tempe, AZ, United States, <sup>3</sup>School of Biological and Health Systems Engineering, Arizona State University, Tempe, AZ, United States

**Objective:** This systematic literature review of the integration of artificial intelligence (AI) applications in surgical practice through hand and instrument tracking provides an overview of recent advancements and analyzes current literature on the intersection of surgery with AI. Distinct AI algorithms and specific applications in surgical practice are also examined.

**Methods:** An advanced search using medical subject heading terms was conducted in Medline (via PubMed), SCOPUS, and Embase databases for articles published in English. A strict selection process was performed, adhering to PRISMA guidelines.

**Results:** A total of 225 articles were retrieved. After screening, 77 met inclusion criteria and were included in the review. Use of AI algorithms in surgical practice was uncommon during 2013–2017 but has gained significant popularity since 2018. Deep learning algorithms ( $n = 62$ ) are increasingly preferred over traditional machine learning algorithms ( $n = 15$ ). These technologies are used in surgical fields such as general surgery ( $n = 19$ ), neurosurgery ( $n = 10$ ), and ophthalmology ( $n = 9$ ). The most common functional sensors and systems used were prerecorded videos ( $n = 29$ ), cameras ( $n = 21$ ), and image datasets ( $n = 7$ ). The most common applications included laparoscopic ( $n = 13$ ), robotic-assisted ( $n = 13$ ), basic ( $n = 12$ ), and endoscopic ( $n = 8$ ) surgical skills training, as well as surgical simulation training ( $n = 8$ ).

**Conclusion:** AI technologies can be tailored to address distinct needs in surgical education and patient care. The use of AI in hand and instrument tracking improves surgical outcomes by optimizing surgical skills training. It is essential to acknowledge the current technical and social limitations of AI and work toward filling those gaps in future studies.

## KEYWORDS

artificial intelligence, deep learning, education, hand tracking, instrument tracking, machine learning, surgery, surgical motion

## 1 Introduction

The current paradigm shift in medicine involves technological advancements aimed at optimizing patient care tasks through research and collaboration. Artificial intelligence (AI) and machine learning (ML) algorithms are applicable in precision medicine, allowing integrative and personalized approaches to solve complex medical problems (1).

AI applications in surgical practice, mainly associated with hand tracking, have evolved significantly over time. Integrating AI technologies into surgical practice began attracting attention in the early 2010s; however, the use of AI algorithms for hand and instrument tracking is a more recent advancement (2). These technologies have been explored across various surgical specialties, including ophthalmology; plastic, reconstructive and aesthetic surgery; endocrine surgery; cardiac surgery; and general surgery (3–9). The main goals of using AI algorithms in surgical practice include enhancing surgical safety, decreasing the duration of time-consuming procedures, improving minimally invasive techniques, and facilitating the training of junior surgeons and students through surgical simulation education (10–12).

In recent years, ML technologies have also been used for disease detection and prevention, radiological interpretations, and improving surgical precision (13–16). The focus of this study was on the role of AI and ML in analyzing surgical precision. ML is often directed toward generating hand-tracking and motion-tracking solutions for the best possible postsurgical outcomes. AI and ML technologies can assess and improve human performance beyond conventional standards, particularly in surgical tasks.

A systematic review of the current literature on AI applications in surgical practice, with a particular focus on hand and instrument tracking, was conducted. AI technologies and their applications in various surgical procedures were explored. It also examined the numerous types of sensors employed in these processes. Finally, the analysis addressed the various AI algorithms, their application in distinct surgical areas via different types of sensors, and the use of AI for hand and instrument tracking in surgical practice, including the associated challenges and future directions.

## 2 Materials and methods

### 2.1 Search strategy

Systematic searches of the Medline (via PubMed), SCOPUS, and Embase databases were conducted in July 2024 without time restrictions, using the following keywords: [(Surgical motion) OR (Hand tracking) OR (instrument tracking) OR (surgical gesture)

OR (instrument detection) OR (tool movement) OR (instrument movement) OR (MoCap Motion Capture) OR (Motion Capture, MoCap) OR (Biomechanical Movement Capture) OR (Movement Capture, Biomechanical) OR (Optical Motion Capture) OR (Motion Capture, Optical) OR (Magnetic Motion Capture) OR (Motion Capture, Magnetic)] AND [(Machine Learning) OR (Learning, Machine) OR (Transfer Learning) OR (Learning, Transfer) OR (Deep Learning) OR (Learning, Deep) OR (Hierarchical Learning) OR (Learning, Hierarchical) OR (Intelligence, Artificial) OR (Computer Reasoning) OR (Reasoning, Computer) OR [AI (Artificial Intelligence)] OR (Machine Intelligence) OR (Intelligence, Machine) OR (Computational Intelligence) OR (Intelligence, Computational) OR (Computer Vision Systems) OR (Computer Vision System) OR (System, Computer Vision) OR (Systems, Computer Vision) OR (Vision System, Computer) OR (Vision Systems, Computer) OR [Knowledge Acquisition (Computer)] OR [Acquisition, Knowledge (Computer)] OR [Knowledge Representation (Computer)] OR [Knowledge Representations (Computer)] OR [Representation, Knowledge (Computer)] AND [(Surgery) OR (Surgical Procedures) OR (Procedures, Surgical) OR (Procedure, Surgical) OR (Surgical Procedure) OR (Operative Procedures) OR (Operative Procedure) OR (Procedure, Operative) OR (Procedures, Operative) OR (Operative Surgical Procedure) OR (Procedure, Operative Surgical) OR (Procedures, Operative Surgical) OR (Surgical Procedure, Operative) OR (Operative Procedures) OR (Operations) OR (Invasive Procedures) OR (Operative Therapy) OR (Preoperative Procedures) OR (Intraoperative Procedures) OR (Peroperative Procedures) OR (Perioperative Procedures)]. These medical subject heading terms were linked with Boolean operators “AND” and “OR” to maximize comprehensiveness.

### 2.2 Eligibility criteria

Our inclusion criteria focused on original research articles that used AI methods in surgical practice, specifically through the application of various sensors for hand or instrument tracking during surgical procedures. Articles were excluded if they lacked any of the 3 key components: AI systems, application in surgical procedures, or tracking of hands or instruments (Table 1). Additionally, review articles, editorials, errata, retracted articles, and studies without accessible full texts or those not available in English were excluded.

### 2.3 Study selection

Our search terms and articles were filtered by title or abstract. Only articles written in the English language were screened. Duplications were excluded. Three independent reviewers

#### Abbreviations

AI, artificial intelligence; CNN, convolutional neural network; EMG, electromyography; IMU, inertial measurement unit; LSTM, long short-term memory; ML, machine learning; SECTR, spectrally encoded coherence tomography and reflectometry; 3D, 3-dimensional; 4D, 4-dimensional.

TABLE 1 Inclusion and exclusion criteria.

| Inclusion criteria   | Exclusion criteria                                    |
|--|---|
| Original research articles   | Review articles, editorials, errata, retracted papers |
| Articles applying AI systems to surgical practice through hand or instrument tracking. | Full text unavailable                                 |
|  | Non-English language                                  |
|  | Not hand or instrument tracking                       |
|  | Not applying to surgical practice                     |

Articles that did not involve any of the 3 components (AI systems, application in surgical procedures, or tracking of hands or instruments) were excluded.

(A.S.G., J.H., K.Y.) screened the articles using the Rayaan platform (17). A strict selection process was performed, adhering to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines (18).

A manual search was also conducted to explore any additional relevant articles, and the reference lists of all included articles were examined by 2 independent reviewers (J.H. and A.S.G.), as recommended by systematic review manuals (19). After the screening, disagreements were discussed, and all reviewers (T.J.O., A.S.G., J.H., K.Y., P.P., A.G.R.) reached a consensus to include 77 articles in the study.

### 3 Results

Initially, 225 articles were retrieved. Thirty-five duplicated papers, 21 review articles and editorials, and 5 articles for which the full text was unavailable were excluded. Of the 164 remaining articles, 102 did not meet the inclusion criteria and were excluded. A reference-checking method was performed, and 10 additional articles were found that were within the scope of the study and included. A manual search method was also conducted, and 5 articles were found to be relevant and included in the study. Thus, after screening, a total of 77 articles were found to be eligible for the study and were included (Table 2) (2–12, 20–84). The selection process is documented in Figure 1 (85).

According to our study's findings, applying AI algorithms in hand and instrument tracking was an exciting but relatively little-studied topic between 2013 and 2017, whereas interest in this area has increased significantly since 2018, with a marked increase in research activity in the 2020s (Figure 2).

Various ML algorithms were used in these studies. For clarity, these algorithms can be categorized into deep learning and traditional ML (non-deep learning) algorithms (Figure 3). This review indicates that deep learning algorithms are significantly more popular than traditional ML algorithms. Specifically, 62 articles employed deep learning, whereas only 15 studies used traditional ML algorithms (Figure 2).

Examining the different functional sensors used in ML applications within surgical practice reveals a lack of uniformity, with a wide range of sensors being used. However, image datasets ( $n = 7$ ), videos ( $n = 29$ ), and camera recordings ( $n = 21$ )

are among the main systems used to collect information or train the algorithms (Table 2). ML algorithms are most extensively studied in general surgery ( $n = 19$ ), neurosurgery ( $n = 10$ ), and ophthalmology ( $n = 9$ ), but their applications are used in a broad spectrum of surgical specialties (Table 2; Figure 4).

ML algorithms have been used in a variety of surgical tasks, including assessment of basic surgical skills, such as suturing and knot tying ( $n = 12$ ); basic laparoscopic ( $n = 13$ ) and endoscopic training ( $n = 8$ ); robotic-assisted surgical training ( $n = 13$ ); and surgical simulation training ( $n = 8$ ) (Table 2). According to our findings, the primary use of the technologies is for educational purposes.

## 4 Discussion

### 4.1 AI and ML

With the advent of AI, the aim was to design a system that could operate intelligently and would be able to understand human language, perform mechanical tasks, solve complex computer-based problems, and return a human-like answer. Another skill humans possess is the ability to learn from our environment, as in language acquisition. When children are exposed to language and start to recognize patterns, they learn the rules of the language. They practice, evaluate their progress, receive feedback, and make adjustments. ML is a subset of AI that allows an AI system to accomplish a similar feat (Figure 3). The capacity of the system to learn is based on 3 factors: the data consumed, quantification of how incorrect the current behavior is compared to the ideal or model behavior, and a feedback mechanism to help guide and produce better future behavior (86). ML technology can predict stock market trends and serve as the foundation for self-driving cars, and its potential applications in medicine are equally, if not more, transformative. ML has been demonstrated to be effective in disease prediction (13), disease detection (14), radiologic interpretation (15), and enhancement of surgical precision (16), the last of which is the focus of this article.

One of the most important aspects of an ML system is its algorithms. The algorithm is effectively the rule book that determines how the data will be used. A single algorithm could have hundreds or even thousands of "rules" or parameters defining its decision-making. Some examples of the algorithms that are used in surgical applications include the following:

- Convolutional neural networks (CNNs) (87)
- Support vector machines (88)
- Random forests (89)
- K-nearest neighbors (90)
- Recurrent neural networks and long short-term memory (LSTM) networks (91)
- Clustering algorithms (92)

These algorithms can be used in distinct learning paradigms, including supervised learning (93), unsupervised learning (86), and reinforcement learning (94). These 3 paradigms are

TABLE 2 Summary of studies focused on applying artificial intelligence algorithms to surgical practice through hand and instrument tracking.

| Study (year)                        | Machine learning method      | Sensor or system used   | Surgical field      | Procedure   |
|-------------------------------------|------------------------------|---|---------------------|---|
| AbuSneineh and Seales (2013) (20)   | Traditional machine learning | ECG monitor, eye trackers, cameras  | Nonspecific surgery | Minimally invasive surgery; pegboard ring transfer task                               |
| Ryu et al. (2013) (2)               | Traditional machine learning | Endoscopic vision-based tracking  | General surgery     | Laparoscopic surgery  |
| Cavallo et al. (2014) (21)          | Traditional machine learning | Optical movement detector, sensor instrument data logger module               | General surgery     | Laparoscopic procedures   |
| Partridge et al. (2014) (22)        | Traditional machine learning | Simulator camera and colored bands placed around the distal instrument shafts | General surgery     | Surgical simulation training (threading string through hoops)                         |
| Dayak and Cevik (2015) (12)         | Traditional machine learning | Logitech HD C525 webcam   | Nonspecific surgery | Minimally invasive surgery; surgical simulation training                              |
| Sahu et al. (2016) (23)             | Traditional machine learning | Laparoscopic images   | General surgery     | Robotized laparoscopic minimally invasive surgery                                     |
| Chen et al. (2017) (24)             | Deep learning                | Line segment detector   | Nonspecific surgery | Robotic minimally invasive surgery  |
| Mishra et al. (2017) (25)           | Deep learning                | Retinal microscopy instrument tracking dataset                                | Ophthalmology       | Retinal microsurgery  |
| Zisimopoulos et al. (2017) (26)     | Deep learning                | Video data and simulation   | Ophthalmology       | Cataract surgery  |
| Du et al. (2018) (27)               | Deep learning                | Endoscopic and microscopic datasets   | Nonspecific surgery | Articulated endoscopic surgical instrument tracking                                   |
| Kil et al. (2018) (28)              | Deep learning                | Cameras   | Nonspecific surgery | Assessment of suturing skill  |
| Richey et al. (2018) (29)           | Traditional machine learning | Bumblebee XB3 and Grasshopper stereo camera system; Polaris Vicra             | General surgery     | Lumpectomy  |
| Wang and Fey (2018) (30)            | Deep learning                | da Vinci surgical system  | Nonspecific surgery | Minimally invasive surgery; robotic surgery, suturing, needle passing, and knot tying |
| Wesierski and Jezierska (2018) (31) | Traditional machine learning | Cameras; videos   | Nonspecific surgery | Minimally invasive surgery; detection and pose estimation during robotic surgery      |
| Zhang et al. (2018) (32)            | Deep learning                | Video recordings  | Urology             | Prostatectomy   |
| Azari et al. (2019) (33)            | Traditional machine learning | GZ-E300 JVC camcorders  | Nonspecific surgery | Simple interrupted and runner subcuticular suturing                                   |
| Du et al. (2019) (34)               | Traditional machine learning | Videos  | General surgery     | Minimally invasive surgery  |
| El-Haddad et al. (2019) (35)        | Deep learning                | Volumetric OCT data   | Ophthalmology       | Ophthalmic surgery  |
| Kowalewski et al. (2019) (36)       | Traditional machine learning | Myo armband; NDI Polaris camera   | General surgery     | Laparoscopic suturing and knot-tying tasks  |
| Funke et al. (2019) (37)            | Deep learning                | Virtual camera programmed to mimic Intel RealSense D435                       | Nonspecific surgery | Minimally invasive surgery; suturing on da Vinci machine                              |
| Qiu et al. (2019) (38)              | Deep learning                | m2cai16-tool based dataset  | Nonspecific surgery | Minimally invasive robot-assisted surgery   |
| Tang et al. (2019) (39)             | Deep learning                | Custom 1,060 nm swept-source OCT engine to perform SPECT imaging              | Ophthalmology       | Intraoperative optical coherence tomography   |
| Winkler-Schwartz et al. (2019) (40) | Traditional machine learning | High-fidelity neurosurgical simulator   | Neurosurgery        | Tumor resection in virtual reality  |
| Zhao et al. (2019) (41)             | Deep learning                | Videos  | General surgery     | Robot-assisted minimally invasive surgery   |
| Zia et al. (2019) (42)              | Deep learning                | Videos  | Urology             | Robotic-assisted radical prostatectomies  |
| Khalid et al. (2020) (43)           | Deep learning                | Videos  | Nonspecific surgery | Knot tying; suturing; needle passing  |
| Lee et al. (2020) (7)               | Deep learning                | Video recordings  | General surgery     | Thyroidectomy; robotic surgery  |
| Leong-Hoi et al. (2020) (44)        | Deep learning                | Depth camera  | General surgery     | Laparoscopic mini-invasive surgery  |
| Morita et al. (2020) (45)           | Deep learning                | High-resolution video recordings  | Ophthalmology       | Cataract surgery  |
| Sivarasa and Jerew (2020) (46)      | Deep learning                | Videos  | Nonspecific surgery | Endoscopic procedures   |
| Tang et al. (2020) (47)             | Deep learning                | OCT and SER images  | Ophthalmology       | Intraoperative OCT  |
| Yu et al. (2020) (48)               | Deep learning                | Video dataset   | Nonspecific surgery | da Vinci surgical system doing 6 different surgical tasks                             |
| Zhang et al. (2020) (4)             | Deep learning                | Preselected video footage from YouTube  | General surgery     | Laparoscopic surgery; breast, GI, and head and neck surgery                           |
| Zhang and Gao (2020) (49)           | Deep learning                | Laparoscopic camera   | General surgery     | Laparoscopic surgery  |
| Cheng et al. (2021) (50)            | Deep learning                | Magnetic endoscope  | Thoracic surgery    | Video-assisted thoracoscopy surgery   |
| Deng et al. (2021) (51)             | Deep learning                | Image dataset   | Otolaryngology      | Identification of electrocautery surgical instruments from 20 open-neck procedures    |

(Continued)

TABLE 2 Continued

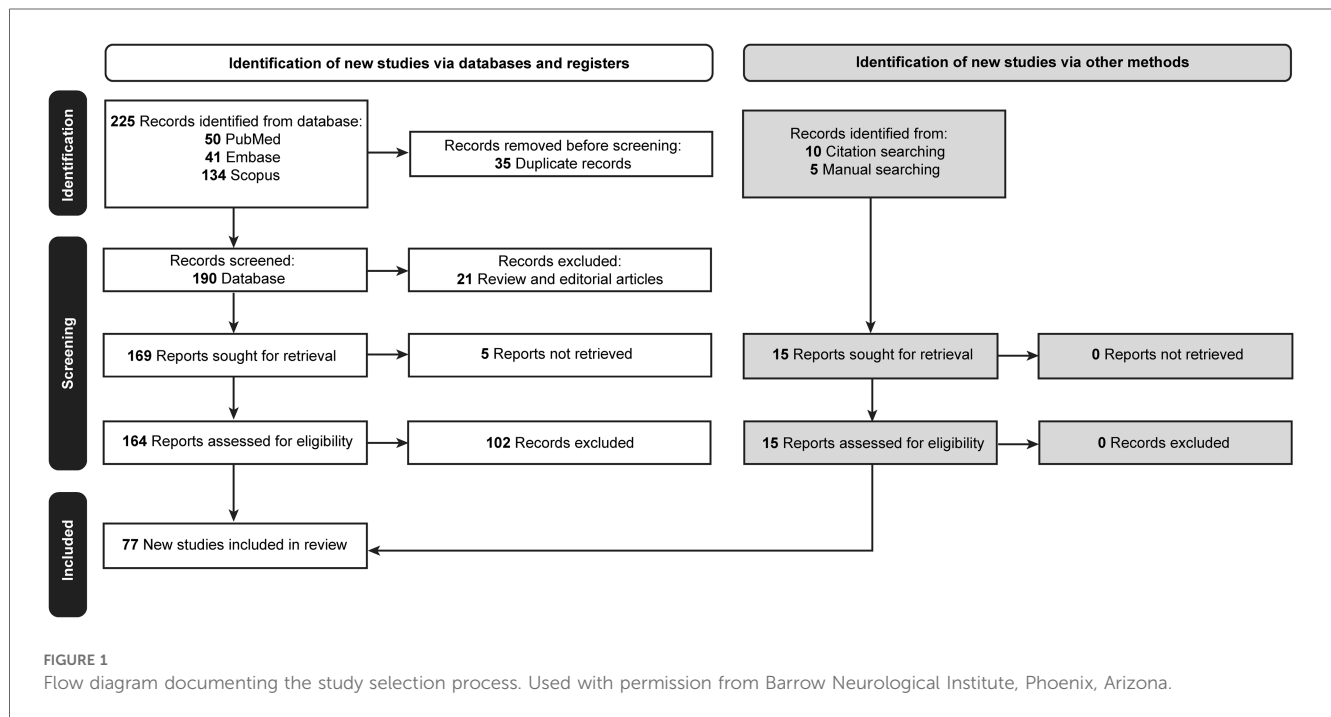
| Study (year)                         | Machine learning method      | Sensor or system used  | Surgical field      | Procedure  |
|--------------------------------------|------------------------------|--|---------------------|--|
| Garcia Nespolo et al. (2021) (52)    | Deep learning                | Video recordings   | Ophthalmology       | Cataract surgery   |
| Hein et al. (2021) (53)              | Deep learning                | RGB frames   | Neurosurgery        | Spine surgery  |
| Sani et al. (2021) (5)               | Deep learning                | Inertial measurement units, strain gauges, Polar Spectra optical motion capture                          | Cardiac surgery     | Mock heart microsurgery procedure  |
| Agarwal et al. (2022) (54)           | Deep learning                | Videos   | Nonspecific surgery | Suturing task  |
| De Backer et al. (2022) (55)         | Deep learning                | Annotated video frames from robotic surgeries  | Urology             | Robot-assisted partial nephrectomy and robot-assisted minimally invasive esophagectomy   |
| Despinoy et al. (2016) (56)          | Deep learning                | Videos   | Nonspecific surgery | Surgical hand gesture recognition  |
| Doughty and Ghugre (2022) (57)       | Deep learning                | Low-latency video  | Nonspecific surgery | Surgery involving rigid surgical drill   |
| Ebina et al. (2022) (58)             | Traditional machine learning | Motion capture system  | General surgery     | Automatic assessment of laparoscopic surgery skills  |
| Ebina et al. (2022) (59)             | Traditional machine learning | Motion capture system  | Urology             | Tissue dissection task and parenchymal suturing task   |
| Gazis et al. (2022) (60)             | Deep learning                | Video dataset  | General surgery     | Laparoscopy; 2 basic laparoscopic tasks (peg transfer and knot tying)  |
| Huang et al. (2022) (61)             | Deep learning                | Images; intelligent flexible endoscope system  | Nonspecific surgery | Minimally invasive surgery   |
| Kil et al. (2022) (62)               | Deep learning                | Cameras (Firefly MV USB 2.0, Point Grey Research Inc., British Columbia, Canada) and InertiaCube4 sensor | Nonspecific surgery | Assessment of open surgery suturing skill  |
| Müller et al. (2022) (63)            | Deep learning                | Intel RealSense D435i RGB camera   | General surgery     | Telestration for laparoscopic surgery  |
| Pangal et al. (2022) (64)            | Traditional machine learning | Intraoperative video recordings  | Neurosurgery        | Vascular injury control during endonasal surgery   |
| Soleymani et al. (2022) (65)         | Deep learning                | JIGSAWS dataset; a dual sparse coding algorithm  | Nonspecific surgery | Interpreting surgical trajectories   |
| Tang et al. (2022) (3)               | Deep learning                | SECTR system; OCT  | Ophthalmology       | Ophthalmic microsurgery  |
| Yamazaki et al. (2022) (66)          | Deep learning                | Video dataset  | General surgery     | Laparoscopic infrapyloric lymphadenectomy and suprapancreatic lymphadenectomy during laparoscopic gastrectomy for gastric cancer |
| Yibulayimu et al. (2022) (67)        | Traditional machine learning | Six-dimensional force sensor embedded in liposuction handle; OptiTrack camera                            | Plastic surgery     | Liposuction  |
| Baldi et al. (2023) (6)              | Deep learning                | Zeiss Lumera microscope  | Ophthalmology       | Vitreoretinal surgery  |
| Balu et al. (2023) (68)              | Deep learning                | Videos   | Neurosurgery        | Detection of surgical instruments in endoscopic endonasal carotid artery laceration repair in a perfused cadaveric simulator     |
| Burton et al. (2023) (69)            | Deep learning                | RGB camera   | Nonspecific surgery | Six-degree-of-freedom pose estimation of a representative surgical instrument in RGB scenes                                      |
| De Backer et al. (2023) (70)         | Deep learning                | Endoscopic video data  | Urology             | Robot-assisted partial nephrectomy and kidney transplant   |
| Gonzalez-Romo et al. (2023) (11)     | Deep learning                | Zeiss Pentero microscope; GoPro camera   | Neurosurgery        | Microvascular anastomosis simulation   |
| Kiyasseh et al. (2023) (71)          | Deep learning                | Videos   | Nonspecific surgery | Gesture recognition and skill assessment in robotic-assisted surgery   |
| Kögl et al. (2023) (8)               | Deep learning                | Single monocular RGB camera  | Neurosurgery        | Burr hole placement  |
| Lin et al. (2023) (72)               | Deep learning                | Cameras  | General surgery     | Robot-assisted laparoscopic surgery  |
| Louis et al. (2023) (73)             | Deep learning                | Video dataset  | Nonspecific surgery | Hand-pose estimation   |
| Luijten et al. (2023) (74)           | Deep learning                | Structured-light 3D scanners   | Oral surgery        | Oral and maxillofacial surgery   |
| Ran et al. (2023) (75)               | Deep learning                | Image dataset  | Orthopedic surgery  | Surgical instrument detection  |
| Yang et al. (2023) (76)              | Deep learning                | Videos   | General surgery     | Peritoneal closure   |
| Yuan et al. (2023) (9)               | Deep learning                | High-resolution microscope camera  | Otolaryngology      | Mastoidectomy and cochlear implant   |
| Zheng and Liu (2023) (77)            | Deep learning                | YOLOv7   | Nonspecific surgery | Minimally invasive surgery   |
| Badilla-Solórzano et al. (2024) (78) | Deep learning                | Annotated image dataset  | Oral surgery        | Wisdom teeth extraction surgery  |
| Balu et al. (2024) (79)              | Deep learning                | Videos   | Neurosurgery        | Minimally invasive spine surgery; CSF leak; durotomy repair  |

(Continued)

TABLE 2 Continued

| Study (year)                | Machine learning method | Sensor or system used                                | Surgical field      | Procedure  |
|-----------------------------|-------------------------|--|---------------------|--|
| Jurosch et al. (2024) (80)  | Deep learning           | Laparoscopic surgical videos recorded via 2D cameras | General surgery     | Surgical instrument location detection                             |
| Kil et al. (2024) (81)      | Deep learning           | Cameras  | Nonspecific surgery | Assessment of open surgery suturing skill                          |
| Kim et al. (2024) (82)      | Deep learning           | Image dataset  | Neurosurgery        | Estimation of accurate localization of spinal fixation instruments |
| Raymond et al. (2024) (83)  | Deep learning           | Surgical microscope                                  | Otolaryngology      | Mastoidectomy  |
| Simoens et al. (2024) (10)  | Deep learning           | Endoscopic camera                                    | Urology             | Robotic surgery, exercise training                                 |
| Sugiyama et al. (2024) (84) | Deep learning           | Videos   | Neurosurgery        | Microvascular anastomosis training                                 |
| Viviers et al. (2024) (85)  | Deep learning           | External optical camera                              | Neurosurgery        | Spine surgery  |

ECG, electrocardiogram; GI, gastrointestinal; HD, high definition; JIGSAWS, Johns Hopkins University–Intuitive Surgical Inc. Gesture and Skill Assessment Working Set; OCT, optic coherence tomography; RGB, red-green-blue; SECTR, spectrally encoded coherence tomography and reflectometry; SER, spectrally encoded reflectometry; SPECT, single-photon emission computed tomography.



the primary learning paradigms of ML, whereas deep learning is a distinct concept suitable for handling complex data structures (95).

### 4.1.1 Supervised learning

Supervised learning is a methodology that trains AI by using a dataset that is already labeled. This would be the equivalent of giving a child a stack of images in which some are apples and others are bananas. If the images are labeled, the child can learn which label is associated with the apple and which with the banana. When a new image appears without a label, they can accurately label it.

Supervised learning deals with 2 problems: classification and regression. Regression problems involve predicting a continuous output with a value that can exist at any point within a range; an example would be predicting Medical College Admission

Test scores based on hours studied. The classification problems ask the system to indicate a categorical output, such as benign or malignant (93). These systems are essential for problems that require accurate predictions and the ability to generalize new data. This learning can recognize and track instruments, hand movements, and anatomical landmarks in a surgical application.

#### 4.1.1.1 Application in using surgeon hand motions to predict surgical maneuvers

Azari et al. (33) asked 37 clinicians, ranging from medical students to retirees, to perform simple and running subcuticular suturing. The hand movements were recorded during the suturing tasks, and 3 ML techniques were applied: decision trees, random forests, and hidden Markov models. These techniques were used to classify surgical maneuvers every 2 s (60 frames) of video. The

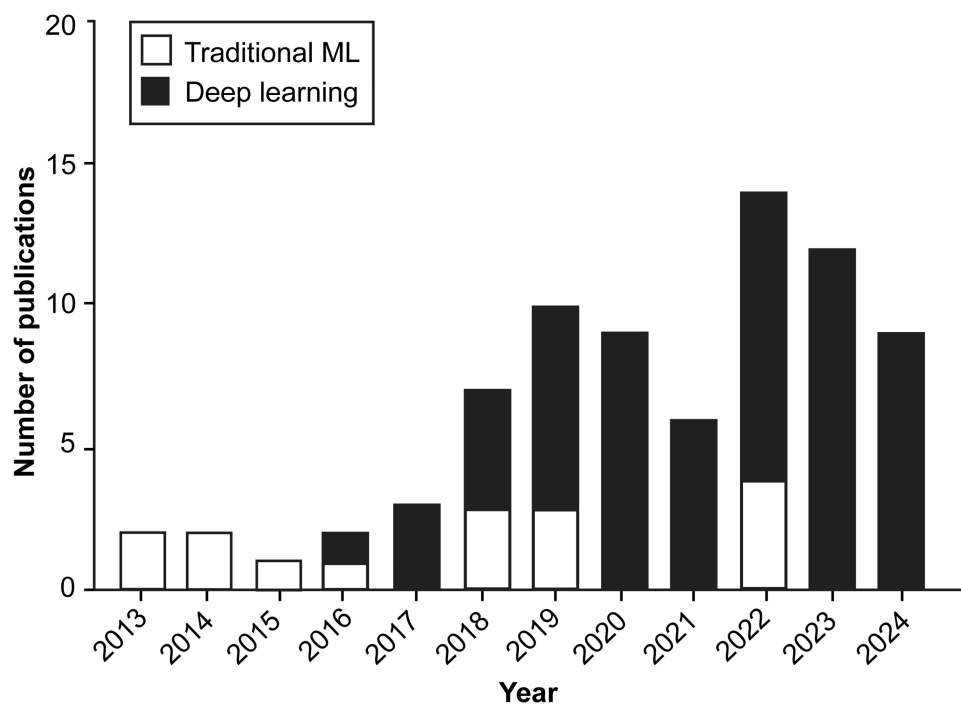


FIGURE 2

Bar chart illustrating the yearly distribution of publications exploring the application of machine learning (ML) and artificial intelligence algorithms in surgical practice through hand and instrument tracking. Since 2018, there has been a shift in the preferred types of ML algorithms used in these studies, with deep learning techniques gaining prominence over traditional machine learning methods. Although applications of these technologies to surgical practice through hand and instrument tracking were relatively sparse between 2013 and 2017, research interest in these technologies has increased substantially since 2018. Used with permission from Barrow Neurological Institute, Phoenix, Arizona.

random forest predictions correctly classified 74% of all video segments, whereas the hidden Markov model improved this to 79%. The results highlight the potential of ML for predicting surgical maneuvers using video data. This approach is promising for improving surgical training and objective assessment by providing a more efficient way to review surgical videos.

#### 4.1.2 Unsupervised learning

Supervised learning requires data to be pre-labeled, whereas unsupervised learning does not rely on any predefined labels for the data. Instead, it identifies intrinsic patterns within the data provided to the program. There are 2 different types of unsupervised learning: clustering and dimensionality reduction. Clustering involves the use of algorithms such as k-means and hierarchical clustering to group things together. Dimensionality reduction involves principal component analysis and T-distributed stochastic neighbor embedding, which are used to reduce the complexity of the data (86).

##### 4.1.2.1 Application in vision-based tracking of surgical instruments

Ryu et al. (2) demonstrated that they could identify the movement of surgical instruments and detect anomalies in their movements using an unsupervised learning method with k-means clustering. By combining this technique with a supervised CNN, they developed a collision warning system that enhanced surgical

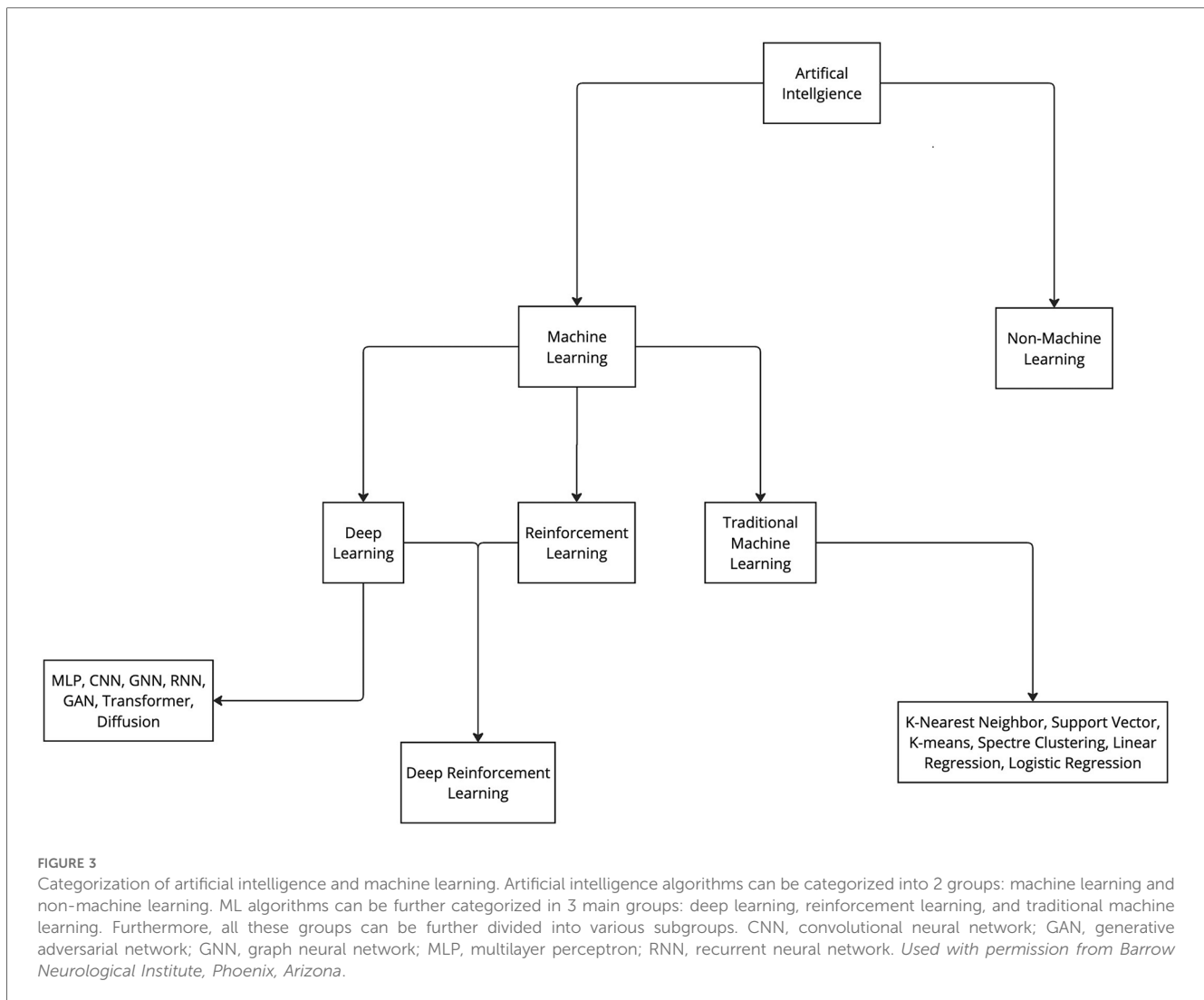
safety by reducing the risk of tissue perforation and instrument collisions.

#### 4.1.3 Reinforcement learning

Reinforcement learning is a type of ML that resembles unsupervised learning in that it does not use pre-labeled samples for training. However, it interacts continuously with the environment, which provides feedback, and through this mechanism, it can produce a desired behavior (86). This type of learning focuses on optimization through trial and error, which makes it suitable for dynamic and complex environments like surgery. There is high potential for its use, with possibilities ranging from skill acquisition for surgical robots to real-time decision support and personalized surgical assistance.

#### 4.1.4 Deep learning

Deep learning is another subset of ML that relies on artificial neural networks with multiple layers. This form of ML excels at learning from large amounts of data, which makes it suitable for applications like healthcare, visual recognition, text analytics, and cybersecurity (96). Within deep learning, there are techniques that can be classified as supervised learning. These include CNNs and recurrent neural networks. CNNs are helpful for image recognition and analysis, whereas recurrent neural networks are suitable for sequential data like time series or natural language processing.



#### 4.1.4.1 Application in estimation of the 6-degree-of-freedom poses of surgical instruments

CNNs were used by Burton et al. (69) to estimate the 6-degree-of-freedom poses of surgical instruments. The CNN was trained on a large dataset of simulated and real-world images. In simulated scenes, CNN achieved mean in-place translation errors of 13 mm and mean long-axis orientation errors of 5°. These errors increased to 29 mm and 8° in the real-world scenes. This demonstrated that the CNN can effectively estimate the 6-degree-of-freedom poses of surgical instruments in real-time, offering an alternative to marker-based tracking methods.

#### 4.1.4.2 Application in hand tracking for surgical telestration

A study by Müller et al. (63) aimed to develop a real-time system that could track hand movements to use in surgical telestration, potentially aiding in surgical training via the augmented reality visualization of surgeons' hands. A total of 14,000 annotated images were used as the dataset, and the model architecture included bounding box detection, skeleton tracking, and hand segmentation. The system achieved a mean detection accuracy of

98% and a Dice similarity coefficient for hand segmentation of 0.95, indicating high segmentation accuracy. Deep learning proved effective for developing real-time hand-tracking systems.

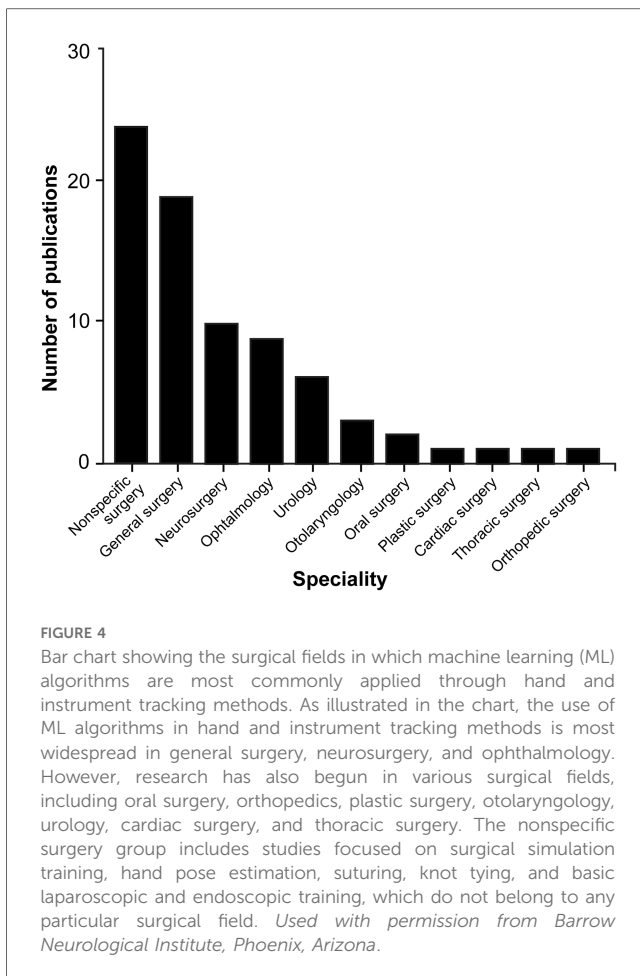
#### 4.1.4.3 Application in neuronavigation

Kögl et al. (8) used this ML methodology to develop a tool-free neuronavigation system using a deep learning model, a single red-green-blue (RGB) camera, and a laptop. This system achieved a high level of accuracy with a target registration error of 1.3 cm and provided real-time guidance for craniotomy planning and execution.

#### 4.1.4.4 Application in hand gesture recognition

Karrenbach et al. (97) created a dynamic hand gesture recognition system using electromyography (EMG) signals and a deep learning methodology known as bidirectional LSTM. This system works by classifying hand gestures and reached an accuracy of 79.5% through a recalibration training scheme. This study highlighted the potential of this type of ML to improve the accuracy and reliability of hand gesture recognition systems.





#### 4.1.4.5 Application in instrument tracking for ophthalmic surgical maneuvers

Tang et al. (39) initially developed a deep learning-based adaptively sampled spectrally encoded coherence tomography and reflectometry (SECTR) program for intraoperative optical coherence tomography for real-time imaging during ophthalmic surgical maneuvers. A deep learning model with a CNN was used to track surgical instrument positions and sample instrument tips. The study presented a novel real-time automated tracking of surgical instruments using deep learning.

Tang et al. (3) used a similar method to track surgical instruments, and the results demonstrated a resolution-limited tracking accuracy with a mean pixel error statistically significantly lower than the resolution limit of 2.95 pixels. This finding was also true for speed, depth, and orientation variations. This technology combines multimodal imaging with deep learning and shows high accuracy and speed. It shows promise for improving surgical outcomes by enhancing surgical visualization. Furthermore, tracking surgical instruments can help achieve greater precision during procedures, reduce errors, and make surgical interventions more effective.

#### 4.1.4.6 Application in automating hand tracking using computer vision

Zhang et al. (4) also developed a method to automate the detection and tracking of surgeon hand movements using

computer vision and a CNN. Data were collected from publicly available YouTube videos annotated with spatial bound boxes. The results showed an average mean precision of 70.4%, and 75.4% with fine-tuning of the dataset. The findings highlight the potential of computer vision for effective hand tracking. It provided a foundation for developing automated systems for surgical skill assessment, leading to improved surgical outcomes and training.

#### 4.1.4.7 Application in video-based surgical skill assessment using three-dimensional (3D) CNNs

Funke et al. (37) investigated a method for automatic, objective assessment of surgical skills using only video data. Their study, which used a 3D CNN and a temporal segment network trained on a kinetics dataset, demonstrated significant performance improvements. The approach achieved high classification accuracies for surgical skill assessment, with accuracy ranging from 95.1% to 100%. The findings reaffirm the reliability of the deep learning approach in evaluating surgical skills, paving the way for advancements in surgical training and outcomes.

#### 4.1.4.8 Application in instrument segmentation with TerausNet during temporal bone surgery

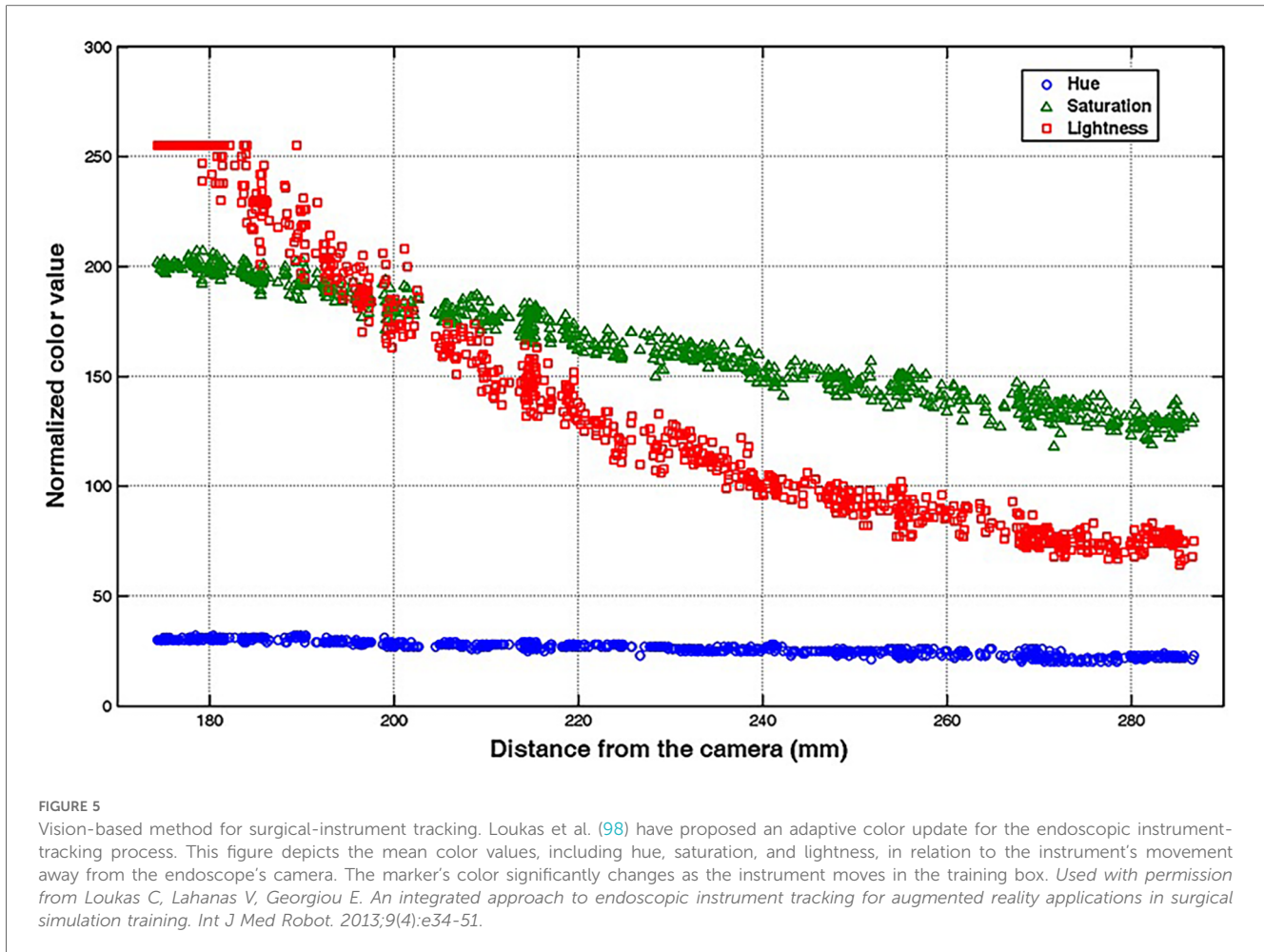
Yuan et al. (9) used an ML application known as TerausNet, a deep learning method, to achieve accurate segmentation of surgical instruments during temporal bone surgery. Video footage of 15 cochlear implant surgeries was collected, with a total of 744 annotated images. Results were measured as F1 dice scores, a single metric that combines precision and recall, evaluating the accuracy of a model. Binary segmentation achieved an F1 dice score of 95% during training and 87% during validation, and multiclass segmentation achieved an F1 score of 62% during training and 61% during validation. Intersection over union for binary segmentation was 91% during training and 77% during validation. Overall, the findings confirm the accuracy and practicality of ML-based segmentation.

## 4.2 Types of sensors used in surgical motion and hand tracking

Technological advancements have enabled the development of hand-tracking tools to effectively analyze the motion of surgical instruments and hands, which can be used to train surgical residents and enhance their technique and motion efficiency. Several sensor systems for tracking surgical motions have been proposed and used, including optical, inertial, electromagnetic, ultrasonic, and hybrid systems that combine different tracking systems (Table 2). Here, we summarize the different types of sensors, their functions, and their applications in tracking surgical hand motions.

### 4.2.1 Vision-based methods

Early techniques for tracking surgical instruments typically relied on vision-based methods. One example is a 2014 study



by Partridge et al. (22), where colored bands on surgical instruments were used for motion tracking via computer vision processing. Similarly, a 2013 study by Loukas et al. (98) combined vision-based tracking using colored tips with the Hough-Kalman approach. Adaptive markers, capable of dynamically changing colors as the surgical instrument moved within the training box, were used (Figure 5) (98). More recently, computer vision processing has also been used to recognize surgical instruments to prevent surgical objects from being left behind during operations (99). Since then, newer and more precise methods of tracking instruments and hand motions have been proposed.

#### 4.2.2 Optical sensors

Optical sensors, such as the Polaris Vega and Stryker NAV3i systems, use infrared light and navigation markers to precisely track surgical motions. A fixed reference tool near the surgical field aids in localizing movements within the surgical volume, as the optical tracker detects infrared light reflected from the markers to determine the coordinates of the labeled instrument (100). Challenges like visual occlusion have been addressed by implementing multicamera modules, enabling instrument tracking from multiple angles (100).

#### 4.2.3 Simple video recordings

Simple video recordings of surgeries can also be used to effectively analyze and monitor the movement of surgical instruments and hands during operations. For example, motion-tracking systems can assign coordinates to instruments in selected videos, depicting the same surgical procedure with varying depth, surgical side, magnification, and resolution (10, 83). These datasets can be stored in various formats and analyzed using CNNs to enable precise motion tracking. In a study by Raymond et al., mastoidectomy surgical videos were used to track surgical instruments (83). The datasets were stored in COCO JSON format, and Detectron, an open-source Facebook AI research CNN, was employed to identify the surgical instruments. Subsequently, the motions of these instruments were tracked using the mastoidectomy instrument tracking module (83).

#### 4.2.4 Automated instrument tracking

Automated instrument tracking using SECTR has also been proposed. This method was used in 2 studies by Tang et al. (39, 47), in which the movements of the 25G ILM forceps, a surgical instrument, were tracked by spectrally encoded reflectometry. The forceps were tracked by training a CNN,

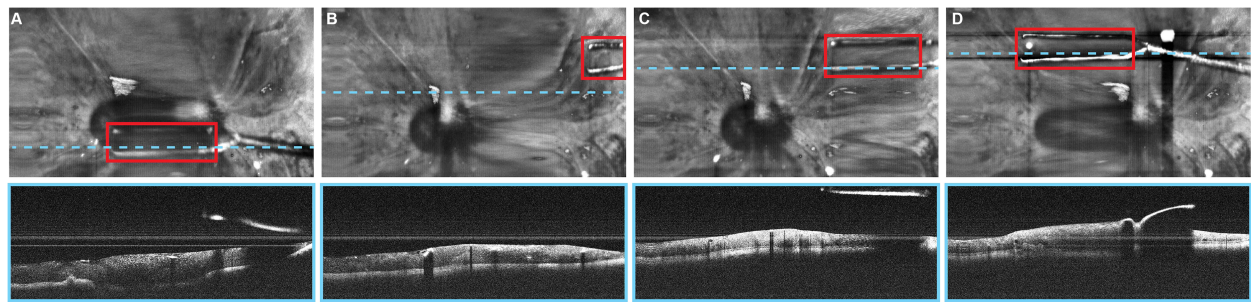


FIGURE 6

Automated instrument-tracking using SECTR. Tang et al. (39) trained a CNN to track 25G ILM forceps, a surgical instrument, to allow 4D visualization of the surgical motion during ophthalmologic surgeries. (A) Five-averaged SER image and representative OCT image demonstrating the tip of the forceps (red box) in an ex vivo bovine eye. Movement of the instrument out of the OCT plane (blue dashed line) and adaptive sampling can be seen on (B–D). CNN, convolutional neural network; OCT, optic coherence tomography; SECTR, spectrally encoded coherence tomography and reflectometry; SER, spectrally encoded reflectometry; 4D, four-dimensional. Used with permission from Tang E, El-Haddad M, Malone JD, et al. Automated instrument-tracking using deep-learning-based adaptively-sampled spectrally encoded coherence tomography and reflectometry (SECTR). *Investigative Ophthalmology & Visual Science*. 2019;60(9):1276-1276.

allowing 4-dimensional (4D) visualization of surgical motion during ophthalmological surgeries (Figure 6) (39, 47).

#### 4.2.5 Tool-free methods

Tool-free methods of tracking hand motion have been explored. In a study conducted by Kögl et al., a single RGB camera combined with ML-based hand pose estimation was used to track hand motion (8). The model was trained using 21 hand landmarks with both real and synthetic images. The Perspective- $n$ -Point problem was then solved, and 3D coordinates of the hand landmarks were obtained. Although the study aimed to facilitate tool-free neuronavigation by using a fingertip to select anatomical landmarks, it faced challenges caused by a high level of interference due to 3D landmark positions. To address this, 2 methods were implemented to filter out the noise.

Similarly, Gonzalez-Romo et al. (11) reported a sensor-free hand-tracking method using a motion-detector program created with the Python programming language (Python Software Foundation, <https://www.python.org/>) and a machine learning system developed by Google (Mediapipe) (<https://ai.google.dev/edge/mediapipe/solutions/guide>). The aim was to analyze hand movements during microanastomosis training. They also demonstrated the effectiveness of ML in assessing hand movement metrics without any physical sensors (Figure 7) (11).

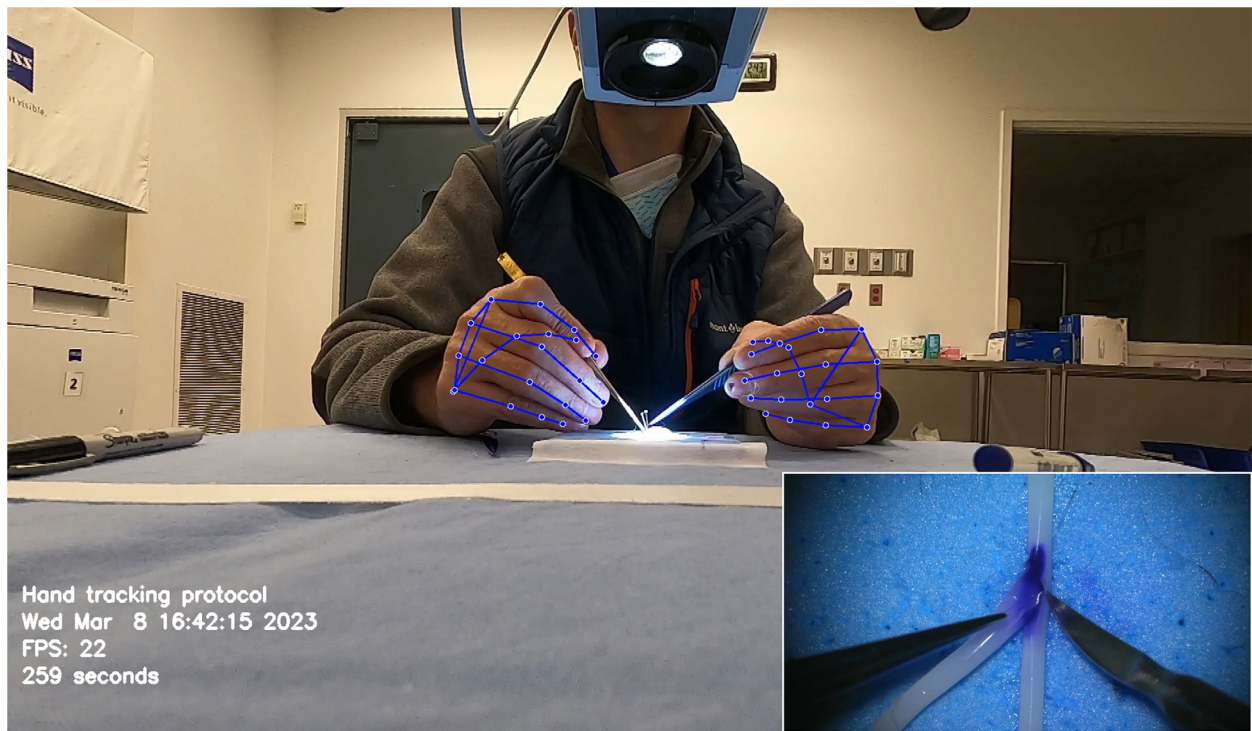
#### 4.2.6 Surgical instrument tracking and hand-eye coordination

Systems like the one proposed by AbuSneineh and Seales (20) measure hand motions along with heart rate and body movements, including those of the arms, head, and eyes, using dedicated cameras (20). Yang et al. explored hand-eye coordination, hypothesizing that a gaze-contingent framework using binocular eye-tracking could aid robotic surgeries (Figure 8) (101). The results highlighted that enhanced surgical instrument manipulation and hand-eye coordination could be achieved by analyzing surgeons' saccadic eye movements and ocular vergence (101).

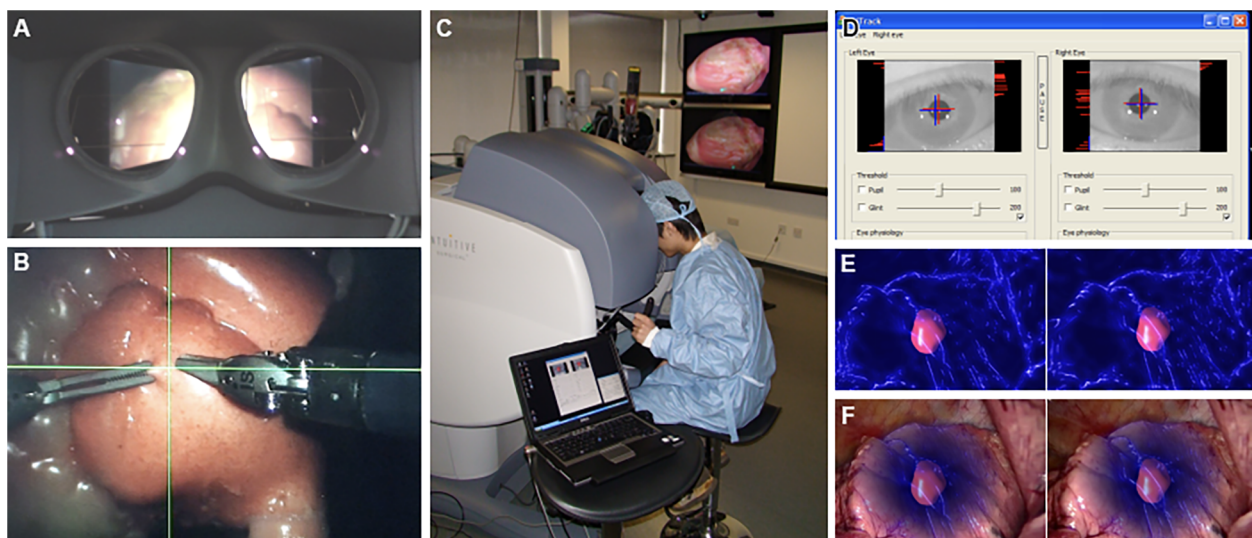
#### 4.2.7 Inertial sensors

Inertial sensors track motion based on the principle of inertia, which states that an object continues to move unless acted upon by an external force. These wearable sensors include accelerometers (tracking linear acceleration), gyroscopes (measuring angular velocity), and magnetometers (detecting magnetic field changes) (102). These different components are typically integrated into an inertial measurement unit (IMU) when used to detect human motion and posture to increase the accuracy of their measurements (102). The application of inertial sensors has been shown to be successful, with Sani et al. using a chain of 12 IMUs for surgical hand tracking (Figure 9) (5). These wearable IMUs focus on accurately tracking 4 main motions: joint angles of the digits and wrist, global hand position and orientation, global instrument position and orientation, and jaw angle of surgical instruments. The goal was to increase the degree of freedom to allow for smoother and more natural movements by the surgeon (5). Given that the index, middle finger, and thumb are the significant contributors to fine grasping and the use of the surgical instruments (in this case, Castroviejo needle-holder and forceps), each finger was tracked by 3 separate IMUs. The wrist joint was tracked via 2 IMUs, a reference point IMU on the forearm and one on the palm, with the rest of the IMUs being placed strategically to capture the best movements of the remaining joints of the hand.

The Polaris Spectra (Northern Digital Inc.) optical sensors were used to precisely track global hand position and orientation, with infrared markers attached directly to the Castroviejo surgical instruments and the surgeon's wrist (5). A significant disadvantage of the IMU system, however, is the accumulation of small errors in position estimation over time, leading to a reduced accuracy of tracking. These sensors typically drift over time and are sensitive to surrounding magnetic interference. To mitigate this issue, this study attempted to decrease drift by measuring relative angles between IMUs; in addition, they recalibrated the recordings every 5 min (5). IMU sensors have also been combined with



**FIGURE 7**  
Hand tracking using a machine learning system developed by Google (Mediapipe) (<https://ai.google.dev/edge/mediapipe/solutions/guide>). The neural network automatically detects hands without requiring preliminary registration. The hand-tracking process starts with activating a palm detector, followed by a landmark detector once the hands are located in the image. At the beginning of the hand-tracking session, the operator is asked to display both hands to the camera to ensure swift activation of the model. The effectiveness of hand detection is validated by displaying real-time tracking data during the simulation. This photograph was taken during microvascular anastomosis training at the Loyal and Edith Davis Neurosurgical Research Laboratory, Barrow Neurological Institute, in Phoenix, Arizona. *Used with permission from Barrow Neurological Institute, Phoenix, Arizona.*



**FIGURE 8**  
Gaze-contingent perceptual docking. Yang et al. (101) presented the general concept of perceptual docking for robotic control. The aim was to investigate the potential use of perceptual docking for knowledge acquisition in robotic-assisted minimally invasive surgery. This figure represents the different aspects of the gaze-contingent perceptual docking. (A) The viewpoint of the surgeon using this system. (B) The surgical field and surgical instruments are shown intraoperatively. (C) External view of the surgeon and the system. (D) The binocular tracking of the eyes. (E and F) The application of augmented reality to gaze-contingent eye tracking, showing (E) nonphotorealistic augmented reality rendering and (F) fused with the original video. *Used with permission from Yang G-Z, Mylonas GP, Kwok K-W, et al. Perceptual Docking for Robotic Control, in: T. Dohi, I. Sakuma and H. Liao (Eds.), Medical Imaging and Augmented Reality, Berlin, Heidelberg: Springer Berlin Heidelberg. 2008;21–30.*

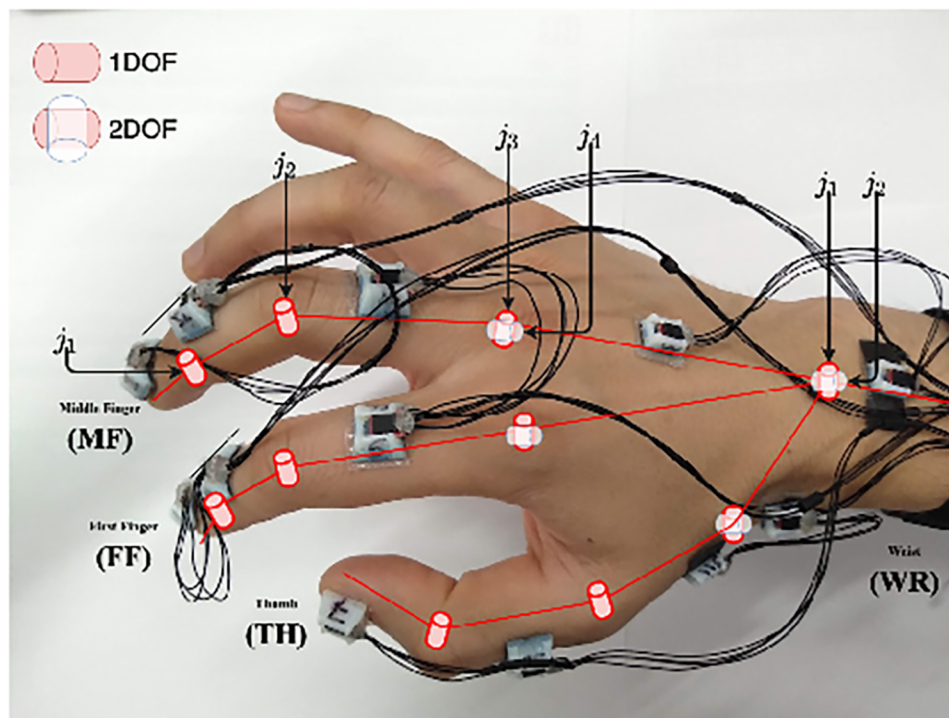


FIGURE 9

Inertial sensors. Sani et al. (5) have used inertial sensors for recording and mapping a surgeon's hand's gross and fine motions during cardiac microsurgery. This figure is a pictographic representation of the 12-unit-IMU sensor device used in this study. Each sensor is attached to a specific joint to allow precise measurement of position change. The DIP and PIP joints are represented by  $j_2$  and  $j_3$ , respectively. The axes of the MCP joint are represented by  $j_1$  and  $j_4$ . DIP, distal interphalangeal joint; IMU, inertial measurement unit; MCP, metacarpophalangeal joint; PIP, proximal interphalangeal joint. Used with permission from Sani MF, Ascione R, Dogramadzi S. Mapping Surgeon's Hand/Finger Motion During Conventional Microsurgery to Enhance Intuitive Surgical Robot Teleoperation. *ArXiv*. 2021;abs/2102.10585.

optical tracking methods to overcome the line-of-site barrier associated with optical tracking methods (103).

#### 4.2.8 Multimodal sensor systems

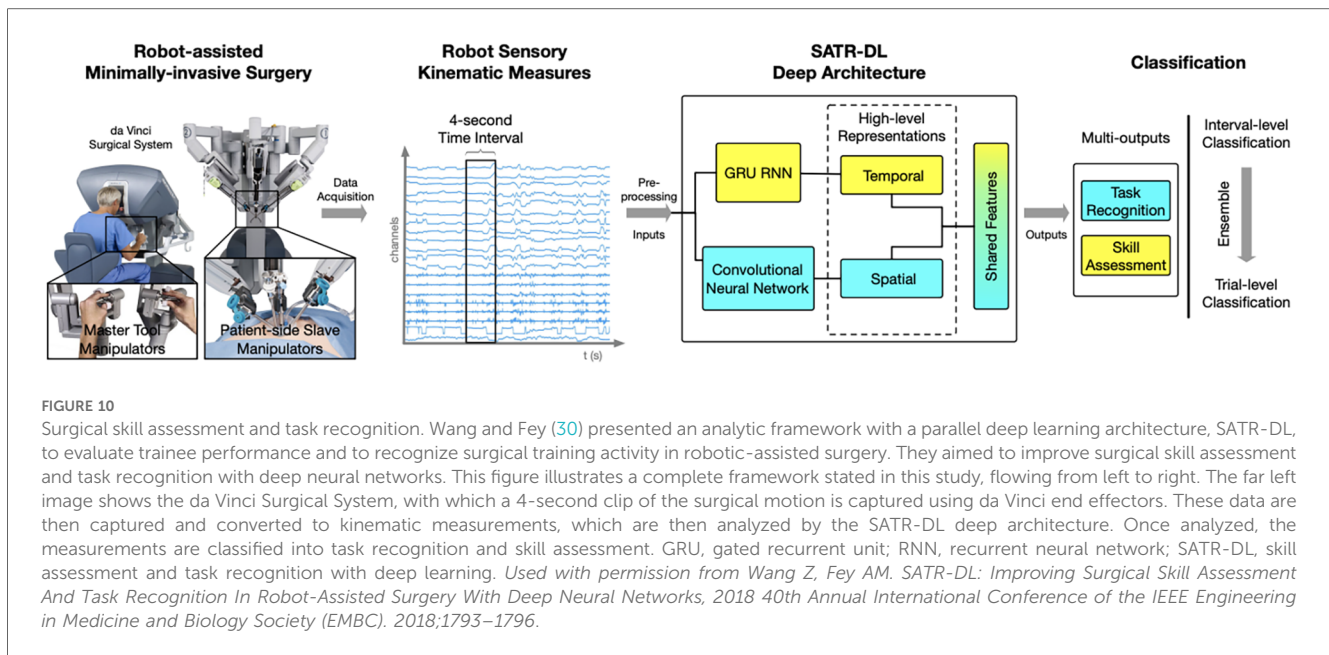
Multimodal systems integrate different sensor types, such as inertial and EMG technologies, to enhance tracking accuracy and address the limitations of single-sensor modalities. Kowalewski et al. published a study in which 28 participants, including 8 novices, 10 intermediates, and 10 experts, completed laparoscopic suturing and knot tying while wearing the Myo armband, a motion-sensing EMG technology combined with inertial sensors that can interpret and track motion (36). The inertial sensors within this technology can measure angular velocity, acceleration, and arm orientation in addition to Euler orientation (36).

EMG electrodes, available as dry or wet types, have been used to track hand motion (97). Dry electrodes are easy to use and attach directly to the skin, whereas wet electrodes require gel or paste but provide higher-quality electrical signals. These EMG electrodes come in both a wireless and a wired form. In a proof-of-concept study by Karrenbach et al. (97), hand gesture recognition using EMG electrodes was demonstrated using the MiSDIREKt (Multisession Single-subject Dynamic Interaction Recordings of EMG and Kinematics) dataset. This EMG hardware consisted of an armband with a wireless 8-channel surface EMG using dry

electrodes made of carbon-black infused thermoplastic polyurethane. It tracked various hand gestures, motions, and dynamic interactions (97).

#### 4.2.9 Electromagnetic sensors

Electromagnetic sensors measure the voltage change created by object movement in space. In the case of hand-tracking models, a coil is wrapped around a ring on a finger (104). This coil is fitted with a printed circuit board that generates a current through the coil, creating a magnetic field signal. This magnetic field signal is then detected by sensors attached to a wristband, which approximate the coil's pose on the finger. However, this is not an exact measurement; a calibration process is needed to help filter out parameters such as noise, gain, and bias. These systems use magnetic fields to track object movement without line-of-sight limitations, but their accuracy can be affected by device interference and surrounding magnetic noise (104). A study by Bordin et al. utilizing an augmented reality helmet, HoloLens 2, found promising results when applying this technology to hand tracking (105). This study used an electromagnetic measuring system to track hand motion (categorized as either no hand motion, simple hand motion, or hand motion while using a 3D-printed object) by wearing linked sensors on the index finger and thumb while wearing the HoloLens 2 helmet (105).



#### 4.2.10 da Vinci Surgical System

Hand motion has also been tracked using advanced systems like the da Vinci Surgical System. Wang and Fey used an end-to-end motion analysis framework with a multioutput deep learning architecture (SATR-DL) to track surgical trainee motions as raw kinematics during specific tasks (30). This method allowed for the representation of both spatial and temporal dynamics of hand motion during surgical tasks. Once the raw motion kinematics had been recorded, they could be input into SATR-DL to recognize the specific task and assess skill level (Figure 10) (30).

### 4.3 Applications of ML and hand and instrument tracking in surgery

ML and hand-tracking technologies are rapidly advancing and can potentially improve a range of healthcare technologies in areas such as preoperative planning, teaching modalities, intraoperative assistance, postoperative analysis, and rehabilitation. This section discusses recent and upcoming developments in ML and hand-tracking technologies and the potential for beneficial disruption in the surgical field. Some limitations of these technologies are addressed, and the developments necessary to ensure their widespread integration and use in the surgical field are discussed.

#### 4.3.1 Surgical instrument tracking

One of the most significant opportunities for ML and hand-tracking technologies to disrupt the surgical field is with intraoperative assistance. One way these technologies can be used is by using computer models to track and record the movement of surgical instruments in the hands of different surgeons.

Zhang and Gao proposed a marker-free surgical instrument tracking framework based on object extraction via deep learning

(49). They compared their deep learning framework with a marker-free tracking-by-detection framework. They found that the proposed deep learning framework improved accuracy by 37%, compared with a 23% improvement using the other model. The findings highlight the rapid development of accurate and efficient tracking technologies for use in various procedures (49).

Surgical instrument tracking has also been used with analysis by AI. Baldi et al. created a model of an eye and then recorded several videos in which surgical instruments were moved throughout the model eye. Several characteristics of the surgical tools were labeled in the frames of the videos, such as location, depth, and tool type (6). Two different deep learning models were trained to predict the tool characteristics, and their performances were recorded. They found that the accuracy of the classification model on the validation dataset was 83% for the x-y region, 96% for depth, 100% for instrument type, and 100% for laterality of insertion. The close-up detection model was performed at 67 frames per second with a mean average precision of 79.3%. The findings illustrate how deep learning models can enable software-based safety feedback mechanisms during surgery and document metrics to guide further research for optimizing surgical outcomes (6).

#### 4.3.2 Objective skill assessment and task recognition in surgery

Objective skill assessment and task recognition in surgery have also been areas of interest in using deep neural networks. Wang and Fey created an analytic framework with a parallel deep learning architecture, SATR-DL, to assess trainee expertise and recognize surgical training activity (30). According to the results, the deep learning model was able to successfully recognize trainee skills and surgical tasks, including suturing, needle passing, and knot tying. The reported accuracies of the deep learning model were 0.960 and 1.000 in skill assessment and task

recognition, respectively. Ultimately, the findings underscored the potential of these neural networks for efficient quantitative assessments in surgical applications, including robotic surgery (30).

#### 4.3.3 Pose estimation of the surgical tool

Sani et al. attempted to develop a system to train a deep neural network by mapping real-time data acquisition inputs of wrist, hand, and surgical tools during mock-up heart microsurgery procedures (5). The proposed network was trained to estimate the poses of the tools from refined hand joint angles, and the outputs were surgical tool orientation and jaw angle obtained by an optical motion capture system. It was found that the neural network could accurately estimate the surgical instrument's pose from the recorded hand movements of the surgeon, achieving a mean squared error of less than 0.3%. This system is suggested to potentially improve the remote teleoperation of similar surgical tools in microsurgery (5).

#### 4.3.4 Applications in robotic surgery

Deep learning models can also potentially be used to evaluate surgical skills in robotic surgery procedures. One of the difficulties with assessing these skills is that quantifying them requires having an accurate way to track the movement of surgical instruments (7). Lee et al. proposed a deep learning-based surgical instrument tracking algorithm to evaluate surgeons' skills in performing robotic surgery procedures. They overcame the drawbacks of occlusion and maintenance of the identity of the surgical instruments, and motion metrics were used to develop the deep learning models (7). It was found that the surgical skill prediction models had an accuracy of 83% with Objective Structured Assessment of Technical Skill and Global Evaluative Assessment of Robotic Surgery. Ultimately, it was argued that the performance of the deep learning model demonstrated how an automatic and quantitative method could be used to evaluate surgical skills, in contrast to previous methods (7).

#### 4.3.5 Endoscopic instrument tracking

Loukas et al. also adequately addressed the issue of occlusion handling by using 2 different tracking algorithms for endoscopic instrument tracking (98). The proposed method was validated based on several image sequences for tracking efficiency, pose estimation accuracy, and applicability in augmented reality-based training. The absolute average error of the tip position for the tool was 2.5 mm, and the average error of the instrument's angle to the camera plane was 2°. The approach showcased promising results for applying augmented reality technologies to laparoscopic skill training using a computer vision framework (98).

#### 4.3.6 Intraoperative real-time feedback

Although ML and hand-tracking technologies have been documented to analyze surgical techniques retrospectively, there has also been a significant interest in developing models that can provide real-time feedback in intraoperative settings. Yibulayimu et al. used ML to create a model that automatically evaluated the trainee's performance of liposuction surgery and provided visual

postoperative and real-time feedback (67). The model was assessed on liposuction surgery datasets, with an optimistic accuracy ranging from 89% to 94%. The findings highlight the potential for ML and other technologies to provide immediate feedback to surgical professionals (67).

#### 4.3.7 Surgical skill assessment during cataract surgery

Morita et al. sought to use the potential of ML in evaluating the surgical skills of ophthalmologists in cataract surgery (45). They proposed a method that used image classification and segmentation networks to detect surgical problems and the surgical instrument, respectively. The network could detect surgical problems with an area under the receiver operating characteristic curve of 0.97 and detect the tips of the forceps and incisional sites with 86.9% and 94.9% detection rates, respectively (45).

#### 4.3.8 Instrument tracking and adaptive sampling models for 4D imaging

Tang et al. proposed a deep learning-based automated instrument tracking and adaptive sampling model that could be used for 4D imaging of ophthalmic surgical maneuvers (3). The findings indicate that real-time localization of surgical instruments can be achieved using SECTR with the deep learning model. Moreover, adaptive sampling enhances the visualization of instrument-tissue interactions by increasing sampling density without sacrificing speed (3).

#### 4.3.9 ML-based surgical workflow analysis

Using surgical videos for AI-based analysis holds immense potential to identify defects and refine surgical techniques. Khan et al. (106) demonstrated this potential by employing Touch Surgery (Digital Surgery Ltd.) to develop and validate an ML-driven stage and procedure analysis tool. Expert surgeons labeled the stages and steps of 50 anonymized endoscopic transsphenoidal surgery videos, with 40 videos used to train a CNN. Based on 10 repetitive videos by Touch Surgery, the model's evaluation showed promising results, with the ML model's automatic detection of surgeons' stage and procedure recognition.

The results indicated that the longest phase was the sellar phase (median 28 min), followed by the nasal phase (median 22 min) and the closing phase (median 14 min). The longest step was tumor identification and resection (median 17 min). In comparison, posterior septectomy and sphenoid septation resection (median 14 min) and anterior sellar wall (median 10 min) varied considerably within the recorded stages.

Despite variations in video appearance, stage duration, and stage sequence, the model demonstrated high accuracy in recognizing surgical steps (91% accuracy, 90% F1 score) and procedures (76% accuracy, 75% F1 score). These findings underscored the reliability of ML-based surgical workflow analysis, which has numerous potential applications, such as education, operational efficiency, and patient outcomes (106).

#### 4.3.10 AI-based surgical phase recognition and analysis

Cheng et al. (107) presented an AI-based automated surgical phase recognition system in laparoscopic cholecystectomy. That study represented an excellent attempt to combine the capabilities of CNNs and LSTM networks to recognize the operative phases of a sequence of frames from a given laparoscopic cholecystectomy video. Although it is well-known that CNNs and LSTMs are highly effective techniques for analyzing medical data, this work leveraged CNNs to extract spatial information from adjacent pixels to generate frame-based visual features. Then, it passed them to an LSTM to capture the temporal transitions of the visual features across consecutive frames. This approach achieved very high accuracy in laparoscopic cholecystectomy operative phase recognition.

Kitaguchi et al. (108) aimed to develop a deep learning model for automatic surgical phase recognition during video laparoscopic sigmoidectomy. This model can be used for real-time phase recognition and to evaluate the accuracy of recognizing surgical steps and actions using image data via a CNN. The automatic surgical phase recognition achieved an accuracy of 91.9%, whereas the recognition of extracorporeal actions and irrigation reached accuracies of 89.4% and 82.5%, respectively. Additionally, the system performed real-time phase recognition at 32 frames per second, with the CNN facilitating the recognition of surgical steps and actions based on manually annotated data. The results highlight the system's high accuracy in automatically recognizing surgical steps and specific actions. Furthermore, it confirmed the feasibility of a real-time automatic surgical phase recognition system with a high frame rate (108).

#### 4.3.11 Future directions

Ultimately, expanding ML and hand-tracking technologies pose considerable opportunities for disruption in the surgical field. From improving the ability of trainees to learn valuable techniques through retrospective analysis to real-time feedback in intraoperative settings, there are myriad applications for these technologies. One limitation of these technologies is the lack of more extensive validation studies to more comprehensively evaluate the accuracy of these ML and hand-tracking models. In addition, future research needs to be done across a more diverse set of procedures to showcase these technologies' potential for beneficial disruption. Nonetheless, these technologies are rapidly expanding and improving, and surgical professionals should strongly consider how they can improve various areas within the field.

### 4.4 Applications of ML and hand and instrument tracking in neurosurgery

AI and ML have been applied to neurosurgery in several ways. Deep learning approaches have allowed neurosurgeons to track the movements of instruments intraoperatively. For example, in a paper by Raymond et al., a deep learning network was initially trained with images from surgery to recognize and track surgical

instruments. The Residual Network 101, a CNN consisting of 101 deep layers, was used. These types of networks work by deconstructing aspects of the image and analyzing the different spectral bands in the image, allowing the network to identify specific components within the image. From here, the images were deconstructed in both low and high levels of resolution, eventually allowing for every detail of the image to be identified. This process is crucial for properly identifying surgical instruments within a video. After the deconstruction of these images, they were analyzed by a region proposal network, which allows for determining regions of interest. Once regions of interest have been identified, they are tested by the CNN to determine the accuracy of instrument recognition and prediction accuracy. Bounding boxes are used as a pictographic representation of these predictions. A computer vision model was used to track instruments during mastoidectomy surgeries. These deep learning technologies allowed the authors to track the exact coordinates of the surgical instruments during the procedure. Images from the procedure were analyzed as described above to determine the precise coordinates of these instruments and the bounding boxes. The model then measured outcomes such as the intersection over union ratio, accuracy, precision, and average precision (Figure 11) (83).

Other papers have also proposed ML methods to track objects or hands' movements in surgery. Kögl et al. published a study in 2022 describing the use of ML to track the movements of hands via hand topology (8). This tool-free neuronavigational method focused on 21 landmarks assigned to a hand and used real and synthetic images to train 2 deep learning networks to estimate hand pose. Real images were used to understand the 2-dimensional coordinates of the hand, whereas synthetic images were used to train the model on relative depth. The Google MediaPipe ML framework was implemented to determine the 2.5-dimensional hand pose. However, the aim was to estimate the 3D pose. Therefore, once these images had been used to train the neural networks, solving the Perspective-n-Point problem was then necessary. This was done by determining where the hand was in space relative to the frame, allowing the determination of the 3D hand pose. Once the 3D hand pose had been estimated, these coordinates were inputted into the 3D Slicer using the OpenIGTLink. The overall aim was to have tool-free neuronavigation that would allow the estimation of hand pose to allow neurosurgeons to use their fingertips to plan craniotomies instead of conventional pointers (Figure 12) (8).

Similar methods of hand pose estimation have been used in neurosurgical procedures such as microanastomosis. Gonzalez-Romo et al. published a study that used 21 hand landmarks in conjunction with ML to identify gross and fine movements during microanastomosis procedures for evaluating surgeons' performance (Figure 7) (11). The Python programming language was used to develop a motion detector incorporating Google Mediapipe (<https://ai.google.dev/edge/mediapipe/solutions/guide>). This motion detector was able to analyze videos of anastomosis procedures to track the motion of hands in a physical sensor-free manner.

In a recent study by Kim et al. (82), a deep learning-based object detection model (YOLOv8 OBB Framework) was used to



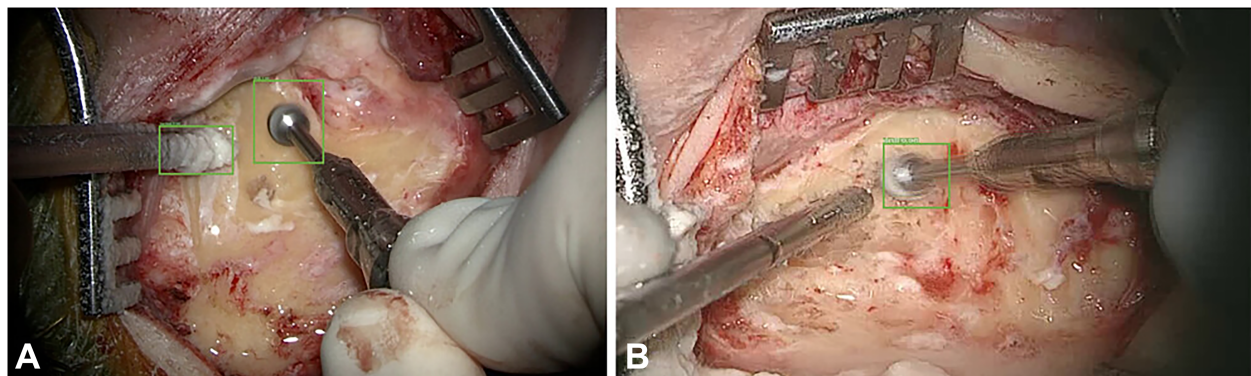


FIGURE 11

Machine learning (ML) applications in neurosurgery. Raymond et al. (83) tried to develop a convolutional neural network–based computer vision model to track 2 surgical instruments used in mastoidectomy procedures (drill and suction irrigator). Intraoperative video recordings were used to teach the model. Then, the model measured the outcomes, such as accuracy, precision, and average precision. The study stated 2 extremes regarding detection accuracy. (A) Detection accuracy for both the drill and the suction-irrigator was 100%. (B) Detection accuracy for the drill was 98%, but absent for the suction irrigator. Green bounding boxes were generated by the ML model to enhance visualization. *Used with permission from Raymond MJ, Biswal B, Pipaliya RM, Rowley MA, Meyer TA. Convolutional neural network-based deep learning engine for mastoidectomy instrument recognition and movement tracking. Otolaryngol Head Neck Surg. 2024;170(6):1555–60. doi: 10.1002/ohn.733.*

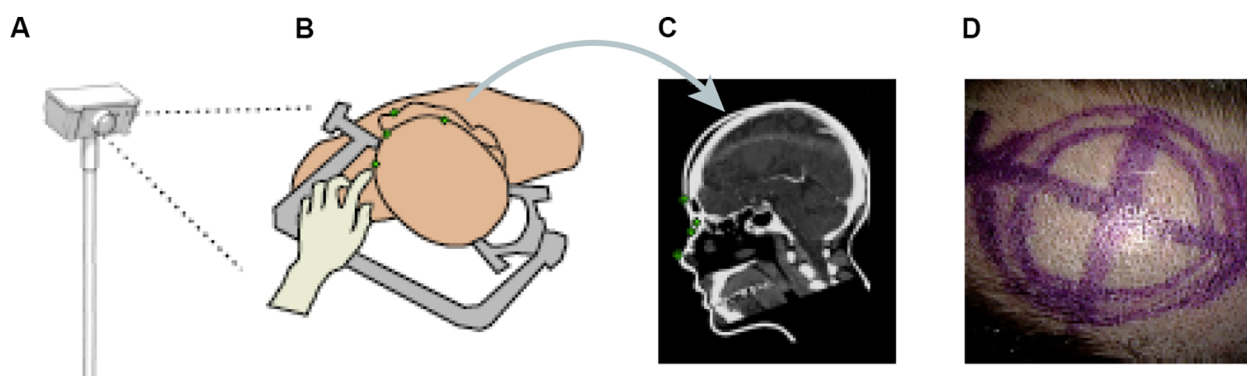


FIGURE 12

Machine learning (ML) application in neurosurgery tool-free methods. In an ML-based hand pose estimation study, Kögl et al. (8) presented a new tool-free neuronavigation method using an RGB camera during burr hole placement. This figure shows the stepwise process used to track hand motion for landmark selection. (A) Illustration of the monocular RGB camera used for this experiment to capture the motion of the hand. (B) Illustration of the hand being used to select landmarks for burr hole placement. (C) Computed tomography showing predefined anatomical landmarks. (D) Photograph showing the final position of the burr hole placement for a craniotomy using this system. *Used with permission from Kögl FV, Léger É, Haouchine N, et al. A Tool-free Neuronavigation Method based on Single-view Hand Tracking. Comput Methods Biomech Biomed Eng Imaging Vis. 2023;11(4):1307–1315. doi:10.1080/21681163.2022.2163428.*

estimate the accurate localization of spinal fixation instruments during spinal surgery. The model was trained using various datasets of images from different manufacturers. Their model's F1 score is 0.86 (1.00 precision and 0.80 recall). Although the model had high precision, the study stated that it had some limitations in detecting all the instruments present.

Applications of ML algorithms in neurosurgery are on the rise and have a significant potential for development. As improvements in ML technologies progress, their integration into neurosurgery is becoming more widespread. This trend highlights the opportunities for leveraging ML algorithms to enhance

neurosurgical procedures such as diagnosis, preoperative surgical planning, and outcome prediction.

#### 4.5 Challenges, ethical considerations, and future directions

Despite the rapid advancement of ML and hand-tracking technologies in recent years, their implementation in clinical settings still faces several limitations. Technical challenges such as data quality and integration from various sensors, as well as

real-time processing and computational limitations, are at the forefront of these limitations. However, the eventual expansion of platforms such as video acquisition can help address some of these computational concerns. Filicori et al. surveyed currently available video recording technology companies and found that those on the market already allow for in-depth performance analysis through manual and automated review (109). In addition, these devices have the potential to be integrated into future robotic surgical platforms.

From another perspective, the integration of AI algorithms into surgical practice may heavily rely on human effort. For instance, successful surgical instrument detection requires surgical videos to be annotated, and even frame by frame labeling or categorization may be necessary. Although this process is time-consuming, it is crucial for providing the necessary visualization to AI algorithms (110). The use of these models in surgical practice has not always been successful. Although these models have shown success in laparoscopic procedures, they have been less effective in open surgical procedures. This may be because of the complex anatomy encountered during open surgeries and the greater difficulty in annotating video frames of open surgical procedures compared to laparoscopic ones. More studies are needed to enhance the potential applications of AI in open surgical procedures (111).

Another concern may be the cost and infrastructure needs associated with integrating these technologies. However, Dayak and Cevik (12) proposed a computer vision-based simulator using a single camera and planar mirror that improved accuracy by 60% over relevant methods in the literature, even at low resolutions and low processing time. The authors also argued that their method was a low-cost alternative for computer vision-based minimally invasive surgery training tools (12). Other toolkits, such as multimodal learning analytics, are being researched to potentially complement traditional measures of learning by providing high-frequency data on the behavior and cognition of trainees (112). Moreover, the integration of these technologies into daily surgical practice may negatively impact the skill development of surgical trainees. This issue should not be overlooked when leveraging AI's facilitative and enhancing effects on surgical education, and the surgical skill development of novice surgeons must be closely monitored (113, 114).

The trustworthiness of AI can be questionable in certain clinical settings (115). Although AI algorithms can rapidly analyze vast amounts of data, their reliability depends on factors such as the quality of the data, differences in patient populations, and other variables (110). Although AI systems in diagnostic settings may reduce interobserver variability, the training data must adequately represent the broader population (110, 116). To address this issue, AI systems should be trained using diverse and comprehensive datasets that accurately reflect the variability and characteristics of the broader population. On the other hand, there is a need for qualified personnel to implement these algorithms in surgical practice, because clinicians and healthcare professionals must be properly trained to use these systems effectively. Regular and continuous training programs should be

developed to ensure that healthcare professionals receive adequate education on the use of rapidly evolving AI models (110).

Another set of challenges in implementing these technologies are ethical and legal considerations. These include the potential effects on data privacy and security as well as the challenges for regulating these ML-based surgical systems. The data used to train these algorithms often contain sensitive patient information, raising ethical concerns about privacy and security (110, 117, 118). A potential data leak could compromise patient confidentiality. To address these concerns, surgical and technological professionals must establish standards defining how and in which situations these technologies will be used. In addition, regulatory guidelines need to be developed and implemented simultaneously because these technologies continue to advance rapidly.

Another ethical consideration in using AI technologies in surgical settings is potential bias. Algorithms trained on biased datasets can reinforce healthcare inequalities related to race, sex, and socioeconomic status. Automation bias poses additional risks, particularly for under-resourced populations. To mitigate these issues, AI models should be developed and implemented with a focus on fairness and equity (118).

In addition to privacy and bias, a significant concern is the potential effect of AI's widespread use in modern medical practice, particularly in fields like radiology and pathology, for which it provides highly accurate diagnostic support (119). This potential raises concerns about job opportunities and employment in these areas, but it is important to remember that AI is a supportive tool in medical practice, not a replacement for human expertise (110).

Although challenges will surface when ML and hand-tracking technologies are developed and implemented, the potential for beneficial disruption in the surgical field is apparent. One exciting future trend will be integrating augmented reality and virtual reality technologies, which can potentially improve visualization during surgical skills training and intraoperative situations. In addition, advancements in sensor technologies and data fusion techniques will likely continue to improve upon the demonstrated accuracy of ML and hand-tracking models in both retrospective and real-time formats.

Ultimately, future research endeavors and the continued advancement of these technologies will be needed to address the current challenges. Nonetheless, ML and hand-tracking technologies can potentially have many exciting applications in the surgical field that should be considered. Future studies should focus on reducing the costs of applying AI systems in surgical practice, developing training programs to meet the demand for qualified personnel, and training AI algorithms with diverse datasets to enhance their applicability across various surgical disciplines (110). Emerging ML technologies like federated learning allow multiple devices or institutions to collaboratively train models while keeping their data private. The use of these recent AI models should be expanded comprehensively in future studies, with efforts focused on eliminating potential algorithmic biases and securing robust data privacy (120).

## 4.6 Limitations

This study aimed to collect the current literature on using diverse AI applications across various surgical settings involving hand or instrument tracking. Our primary goal was to create a resource for future studies rather than compare the effectiveness of different AI applications or ML paradigms. The heterogeneous nature of the data across the included studies represents a limitation, because differences in surgical instruments, settings, and methodologies make robust statistical comparisons challenging, and a risk of bias analysis was not conducted for the included studies.

## 5 Conclusions

In summary, ML and its subcategories, including supervised learning, unsupervised learning, reinforced learning, and deep learning, as well as the various sensor types such as optical, inertial, electromagnetic, ultrasonic, and hybrid sensors, offer unique strengths that could be combined and leveraged in different surgical training or patient care settings. The versatility of these technologies allows for the generation of adaptable models that cater to the specific surgical education and patient care questions that need to be addressed.

The hand-tracking technology and the applicability of ML have tremendous potential to grow, which leads to a direct, tangible impact in a wide range of sectors within surgery, such as surgical education of trainees and patient outcome management ranging from preoperative to postoperative stages. Acknowledging and tackling the current limitations of the development and use of ML will provide grounds for its successful optimization and application in the future.

Future studies that qualitatively and quantitatively depict the improvements generated by ML in surgical education and operative outcomes are warranted. A multifaceted approach to solving complex surgical problems in and out of the operating room should be the prioritized task to create novel solutions that complement and advance the delivery of patient-centered care.

## Author contributions

KY: Conceptualization, Writing – review & editing, Data curation, Methodology, Writing – original draft. TO: Data curation, Writing – original draft. YX: Conceptualization, Data curation, Writing – original draft. AG: Data curation, Formal analysis, Methodology, Writing – original draft. JH: Data curation, Formal analysis, Writing – original draft. AR: Data

curation, Formal analysis, Writing – original draft. PP: Methodology, Writing – original draft. JC: Methodology, Validation, Writing – original draft. JT: Methodology, Writing – original draft. BL: Conceptualization, Validation, Writing – review & editing. MS: Conceptualization, Validation, Writing – review & editing. ML: Writing – review & editing. MP: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was supported by funds from the Newsome Chair of Neurosurgery Research, held by Dr. Preul.

## Acknowledgments

We thank the staff of Neuroscience Publications at Barrow Neurological Institute for assistance with manuscript preparation.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of *Frontiers*, at the time of submission. This had no impact on the peer review process and the final decision.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- MacEachern SJ, Forkert ND. Machine learning for precision medicine. *Genome*. (2021) 64(4):416–25. doi: 10.1139/gen-2020-0131
- Ryu J, Choi J, Kim HC. Endoscopic vision-based tracking of multiple surgical instruments during robot-assisted surgery. *Artif Organs*. (2013) 37(1):107–12. doi: 10.1111/j.1525-1594.2012.01543.x
- Tang EM, El-Haddad MT, Patel SN, Tao YK. Automated instrument-tracking for 4D video-rate imaging of ophthalmic surgical maneuvers. *Biomed Opt Express*. (2022) 13(3):1471–84. doi: 10.1364/BOE.450814
- Zhang M, Cheng X, Copeland D, Desai A, Guan MY, Brat GA, et al. Using computer vision to automate hand detection and tracking of surgeon movements in videos of open surgery. *AMIA Annu Symp Proc*. (2020) 2020:1373–82. doi: 10.48550/arXiv.2012.06948
- Sani MF, Ascione R, Dogramadzi S. Mapping Surgeon's Hand/Finger Motion During Conventional Microsurgery to Enhance Intuitive Surgical Robot Teleoperation. *ArXiv: abs/2102.10585* (2021).
- Baldi PF, Abdelkarim S, Liu J, To JK, Ibarra MD, Browne AW. Vitreoretinal surgical instrument tracking in three dimensions using deep learning. *Transl Vis Sci Technol*. (2023) 12(1):20. doi: 10.1167/tvst.12.1.20
- Lee D, Yu HW, Kwon H, Kong HJ, Lee KE, Kim HC. Evaluation of surgical skills during robotic surgery by deep learning-based multiple surgical instrument tracking in training and actual operations. *J Clin Med*. (2020) 9(6). doi: 10.3390/jcm9061964
- Kögl FV, Leger E, Haouchine N, Torio E, Juvekar P, Navab N, et al. A tool-free neuronavigation method based on single-view hand tracking. *Comput Methods Biomech Biomed Eng Imaging Vis*. (2023) 11(4):1307–15. doi: 10.1080/21681163.2022.2163428
- Yuan C, Shin D, Le TN, Lin VY, Chen JM, Kahrs LA, et al. Instrument segmentation with TeranusNet during temporal bone surgery. *Int J Comput Assist Radiol Surg*. (2023) 18:S49–50. doi: 10.1007/s11548-023-02878-2
- Simoens J, Mottrie A, Hofman J, Besi G, Ferraguti F, De Wel O, et al. A.I. instrument tracking during proficiency-based progression training: does instrument velocity really matter? *Eur Urol*. (2024) 85:S845–S6. doi: 10.1016/s0302-2838(24)00704-8
- Gonzalez-Romo NI, Hanalioglu S, Mignucci-Jimenez G, Koskay G, Abramov I, Xu Y, et al. Quantification of motion during microvascular anastomosis simulation using machine learning hand detection. *Neurosurg Focus*. (2023) 54(6):E2. doi: 10.3171/2023.3.FOCUS2380
- Dayak E, Cevik U. Endoscopic instrument tracking for surgical simulation training in a controlled environment via a camera and a planar mirror. *Comput Biol Med*. (2015) 67:161–71. doi: 10.1016/j.compbiomed.2015.10.012
- Nilashi M, Ibrahim OB, Ahmadi H, Shahmoradi L. An analytical method for diseases prediction using machine learning techniques. *Comput Chem Eng*. (2017) 106:212–23. doi: 10.1016/j.compchemeng.2017.06.011
- Rahman SMA, Ibtisum S, Bazgir E, Barai T. The Significance of Machine Learning in Clinical Disease Diagnosis: A Review. *ArXiv: abs/2310.16978* (2023).
- Wang S, Summers RM. Machine learning and radiology. *Med Image Anal*. (2012) 16(5):933–51. doi: 10.1016/j.media.2012.02.005
- Egert M, Steward JE, Sundaram CP. Machine learning and artificial intelligence in surgical fields. *Indian J Surg Oncol*. (2020) 11(4):573–7. doi: 10.1007/s13193-020-01166-8
- Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. *Syst Rev*. (2016) 5(1):210. doi: 10.1186/s13643-016-0384-4
- Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. (2009) 6(7):e1000097. doi: 10.1371/journal.pmed.1000097
- Horsley T, Dingwall O, Sampson M. Checking reference lists to find additional studies for systematic reviews. *Cochrane Database Syst Rev*. (2011) 2011(8):MR000026. doi: 10.1002/14651858.MR000026.pub2
- AbuSneineh S, Seales B. 2013 Scientific Session of the Society of American Gastrointestinal and Endoscopic Surgeons (SAGES). Baltimore, Maryland, USA, 17–20 April 2013. *Surg Endosc*. (2013) 27(S1):304–503. doi: 10.1007/s00464-013-2881-z
- Cavallo F, Sinigaglia S, Megali G, Pietrabissa A, Dario P, Mosca F, et al. Biomechanics-machine learning system for surgical gesture analysis and development of technologies for minimal access surgery. *Surg Innov*. (2014) 21(5):504–12. doi: 10.1177/1553350613510612
- Partridge RW, Hughes MA, Brennan PM, Hennessey IA. Accessible laparoscopic instrument tracking (“InsTrac”): construct validity in a take-home box simulator. *J Laparosc Adv Surg Tech A*. (2014) 24(8):578–83. doi: 10.1089/lap.2014.0015
- Sahu M, Moerman D, Mewes P, Mountney P, Rose G. Instrument state recognition and tracking for effective control of robotized laparoscopic systems. *Int J Mech Eng Robot Res*. (2016) 5(1):33–8. doi: 10.18178/ijmerr.5.1.33-38
- Chen Z, Zhao Z, Cheng X, editors. Surgical instruments tracking based on deep learning with lines detection and spatio-temporal context. *2017 Chinese Automation Congress (CAC)*; 2017 20–22 Oct. (2017).
- Mishra K, Sathish R, Sheet D. Tracking of retinal microsurgery tools using late fusion of responses from convolutional neural network over pyramidally decomposed frames. In: Mukherjee S, Mukherjee S, Mukherjee SP, Sivaswamy J, Awate S, Setlur S, editors. *Computer Vision, Graphics, and Image Processing*. Cham: Springer International Publishing (2017). p. 358–66.
- Zisimopoulos O, Flouty E, Stacey M, Muscroft S, Giataganas P, Nehme J, et al. Can surgical simulation be used to train detection and classification of neural networks? *Health Technol Lett*. (2017) 4(5):216–22. doi: 10.1049/htl.2017.0064
- Du X, Kurmann T, Chang PL, Allan M, Ourselin S, Sznitman R, et al. Articulated multi-instrument 2-D pose estimation using fully convolutional networks. *IEEE Trans Med Imaging*. (2018) 37(5):1276–87. doi: 10.1109/TMI.2017.2787672
- Kil I, Groff RE, Singapogu RB. Surgical suturing with depth constraints: image-based metrics to assess skill. *Annu Int Conf IEEE Eng Med Biol Soc*. (2018) 2018:4146–9. doi: 10.1109/EMBC.2018.8513266
- Richey WL, Luo M, Goodale SE, Clements LW, Meszoely IM, Miga MI. A system for automatic monitoring of surgical instruments and dynamic, non-rigid surface deformations in breast cancer surgery. *Proc SPIE Int Soc Opt Eng*. (2018) 10576. doi: 10.1117/12.2295221
- Wang Z, Fey AM. SATR-DL: improving surgical skill assessment and task recognition in robot-assisted surgery with deep neural networks. *Annu Int Conf IEEE Eng Med Biol Soc*. (2018) 2018:1793–6. doi: 10.1109/EMBC.2018.8512575
- Wesierski D, Jezierska A. Instrument detection and pose estimation with rigid part mixtures model in video-assisted surgeries. *Med Image Anal*. (2018) 46:244–65. doi: 10.1016/j.media.2018.03.012
- Zhang Y, Law H, Kim T-K, Miller D, Montie J, Deng J, et al. PD58-12 surgeon technical skill assessment using computer vision-based analysis. *J Urol*. (2018) 199(4S):e1138. doi: 10.1016/j.juro.2018.02.2800
- Azari DP, Hu YH, Miller BL, Le BV, Radwin RG. Using surgeon hand motions to predict surgical maneuvers. *Hum Factors*. (2019) 61(8):1326–39. doi: 10.1177/0018720819838901
- Du X, Allan M, Bodenstedt S, Maier-Hein L, Speidel S, Dore A, et al. Patch-based adaptive weighting with segmentation and scale (PAWSS) for visual tracking in surgical video. *Med Image Anal*. (2019) 57:120–35. doi: 10.1016/j.media.2019.07.002
- El-Haddad M, Malone J, Hoang N, Tao Y. *Deep-Learning Based Automated Instrument Tracking and Adaptive-Sampling of Intraoperative Oct for Video-Rate Volumetric Imaging of Ophthalmic Surgical Maneuvers*. Bellingham, WA: SPIE (2019).
- Kowalewski KF, Garrow CR, Schmidt MW, Benner L, Muller-Stich BP, Nickel F. Sensor-based machine learning for workflow detection and as key to detect expert level in laparoscopic suturing and knot-tying. *Surg Endosc*. (2019) 33(11):3732–40. doi: 10.1007/s00464-019-06667-4
- Funke I, Mees ST, Weitz J, Speidel S. Video-based surgical skill assessment using 3D convolutional neural networks. *Int J Comput Assist Radiol Surg*. (2019) 14(7):1217–25. doi: 10.1007/s11548-019-01995-1
- Qiu L, Li C, Ren H. Real-time surgical instrument tracking in robot-assisted surgery using multi-domain convolutional neural network. *Health Technol Lett*. (2019) 6(6):159–64. doi: 10.1049/htl.2019.0068
- Tang E, El-Haddad M, Malone JD, Tao Y. Automated instrument-tracking using deep-learning-based adaptively-sampled spectrally encoded coherence tomography and reflectometry (SECTR). *Invest Ophthalmol Vis Sci*. (2019) 60(9):1276.
- Winkler-Schwartz A, Yilmaz R, Mirchi N, Bissonnette V, Ledwos N, Siyar S, et al. Machine learning identification of surgical and operative factors associated with surgical expertise in virtual reality simulation. *JAMA Netw Open*. (2019) 2(8):e198363. doi: 10.1001/jamanetworkopen.2019.8363
- Zhao Z, Cai T, Chang F, Cheng X. Real-time surgical instrument detection in robot-assisted surgery using a convolutional neural network cascade. *Health Technol Lett*. (2019) 6(6):275–9. doi: 10.1049/htl.2019.0064
- Zia A, Guo L, Zhou L, Essa I, Jarc A. Novel evaluation of surgical activity recognition models using task-based efficiency metrics. *Int J Comput Assist Radiol Surg*. (2019) 14(12):2155–63. doi: 10.1007/s11548-019-02025-w
- Khalid S, Goldenberg M, Grantcharov T, Taati B, Rudzicz F. Evaluation of deep learning models for identifying surgical actions and measuring performance. *JAMA Netw Open*. (2020) 3(3):e201664. doi: 10.1001/jamanetworkopen.2020.1664
- Leong-Hoi A, Murat S, Acker T, Sattler Q, Radoux J-P. *Touchless Interface Interaction by Hand Tracking with a Depth Camera and a Convolutional Neural Network*. Bellingham, WA: SPIE (2020).
- Morita S, Tabuchi H, Masumoto H, Tanabe H, Kamiura N. Real-time surgical problem detection and instrument tracking in cataract surgery. *J Clin Med*. (2020) 9(12):3896. doi: 10.3390/jcm9123896

46. Sivarasa A, Jerew OD, editors. Deep learning for minimally invasive computer assisted surgery. *2020 5th International Conference on Innovative Technologies in Intelligent Systems and Industrial Applications (CITISIA)*; 2020 25-27 Nov (2020).
47. Tang E, El-Haddad MT, Malone JD, Tao Y. Deep-learning based automated instrument-tracking and adaptive-sampling for 4D imaging of ophthalmic surgical maneuvers. *Invest Ophthalmol Vis Sci.* (2020) 61(7):2544.
48. Yu L, Wang P, Yan Y, Xia Y, Cao W. Massd: multi-scale attention single shot detector for surgical instruments. *Comput Biol Med.* (2020) 123:103867. doi: 10.1016/j.combiomed.2020.103867
49. Zhang J, Gao X. Object extraction via deep learning-based marker-free tracking framework of surgical instruments for laparoscope-holder robots. *Int J Comput Assist Radiol Surg.* (2020) 15(8):1335–45. doi: 10.1007/s11548-020-02214-y
50. Cheng T, Li W, Ng WY, Huang Y, Li J, Ng SH, et al. Deep learning assisted robotic magnetic anchored and guided endoscope for real-time instrument tracking. *IEEE Robot Autom Lett.* (2021) 6(2):3979–86. doi: 10.1109/LRA.2021.3066834
51. Deng T, Gulati S, Rodriguez W, Dawant BM, Langerman A. Automated detection of electrocautery instrument in videos of open neck procedures using YOLOv3. *Annu Int Conf IEEE Eng Med Biol Soc.* (2021) 2021:2071–4. doi: 10.1109/EMBC46164.2021.9630961
52. Garcia Nespola R, Leiderman YI, Luciano C, Valikodath N, Cole E. An automated algorithm for tool tracking and turbulence detection during phacoemulsification cataract surgery using computer vision and deep learning neural networks. *Invest Ophthalmol Vis Sci.* (2021) 62(11):30.
53. Hein J, Seibold M, Bogo F, Farshad M, Pollefeys M, Furnstahl P, et al. Towards markerless surgical tool and hand pose estimation. *Int J Comput Assist Radiol Surg.* (2021) 16(5):799–808. doi: 10.1007/s11548-021-02369-2
54. Agarwal S, Pradeep CS, Sinha N. Temporal surgical gesture segmentation and classification in multi-gesture robotic surgery using fine-tuned features and calibrated MS-TCN. *2022 IEEE International Conference on Signal Processing and Communications (SPCOM)*; 2022 11-15 July (2022).
55. De Backer P, Eckhoff JA, Simoens J, Muller DT, Allaey C, Creemers H, et al. Multicentric exploration of tool annotation in robotic surgery: lessons learned when starting a surgical artificial intelligence project. *Surg Endosc.* (2022) 36(11):8533–48. doi: 10.1007/s00464-022-09487-1
56. Despinoy F, Bouget D, Forestier G, Penet C, Zemiti N, Pognet P, et al. Unsupervised trajectory segmentation for surgical gesture recognition in robotic training. *IEEE Trans Biomed Eng.* (2016) 63(6):1280–91. doi: 10.1109/TBME.2015.2493100
57. Doughty M, Ghugre NR. HMD-EgoPose: head-mounted display-based egocentric marker-less tool and hand pose estimation for augmented surgical guidance. *Int J Comput Assist Radiol Surg.* (2022) 17(12):2253–62. doi: 10.1007/s11548-022-02688-y
58. Ebina K, Abe T, Hotta K, Higuchi M, Furumido J, Iwahara N, et al. Automatic assessment of laparoscopic surgical skill competence based on motion metrics. *PLoS One.* (2022) 17(11):e0277105. doi: 10.1371/journal.pone.0277105
59. Ebina K, Abe T, Hotta K, Higuchi M, Furumido J, Iwahara N, et al. Objective evaluation of laparoscopic surgical skills in wet lab training based on motion analysis and machine learning. *Langenbecks Arch Surg.* (2022) 407(5):2123–32. doi: 10.1007/s00423-022-02505-9
60. Gazis A, Karaiskos P, Loukas C. Surgical gesture recognition in laparoscopic tasks based on the transformer network and self-supervised learning. *Bioengineering (Basel).* (2022) 9(12). doi: 10.3390/bioengineering9120737
61. Huang Y, Li J, Zhang X, Xie K, Li J, Liu Y, et al. A surgeon preference-guided autonomous instrument tracking method with a robotic flexible endoscope based on dVRK platform. *IEEE Robotics and Automation Letters.* (2022) 7(2):2250–7. doi: 10.1109/LRA.2022.3143305
62. Kil I, Eidt JF, Groff RE, Singapogu RB. Assessment of open surgery suturing skill: simulator platform, force-based, and motion-based metrics. *Front Med (Lausanne).* (2022) 9:897219. doi: 10.3389/fmed.2022.897219
63. Müller LR, Petersen J, Yamlahi A, Wise P, Adler TJ, Seitel A, et al. Robust hand tracking for surgical telestration. *Int J Comput Assist Radiol Surg.* (2022) 17(8):1477–86. doi: 10.1007/s11548-022-02637-9
64. Pangal DJ, Kugener G, Cardinal T, Lechtholz-Zey E, Collet C, Lasky S, et al. Use of surgical video-based automated performance metrics to predict blood loss and success of simulated vascular injury control in neurosurgery: a pilot study. *J Neurosurg.* (2022) 137(3):840–9. doi: 10.3171/2021.10.JNS211064
65. Soleymani A, Li X, Tavakoli M. Surgical procedure understanding, evaluation, and interpretation: a dictionary factorization approach. *IEEE Trans Med Robot Bionics.* (2022) 4(2):423–35. doi: 10.1109/TMRB.2022.3170210
66. Yamazaki Y, Kanaji S, Kudo T, Takiguchi G, Urakawa N, Hasegawa H, et al. Quantitative comparison of surgical device usage in laparoscopic gastrectomy between surgeons' skill levels: an automated analysis using a neural network. *J Gastrointest Surg.* (2022) 26(5):1006–14. doi: 10.1007/s11605-021-05161-4
67. Yibulayimu S, Wang Y, Liu Y, Sun Z, Wang Y, Jiang H, et al. An explainable machine learning method for assessing surgical skill in liposuction surgery. *Int J Comput Assist Radiol Surg.* (2022) 17(12):2325–36. doi: 10.1007/s11548-022-02739-4
68. Balu A, Pangal DJ, Kugener G, Donoho DA. Pilot analysis of surgeon instrument utilization signatures based on shannon entropy and deep learning for surgeon performance assessment in a cadaveric carotid artery injury control simulation. *Oper Neurosurg (Hagerstown).* (2023) 25(6):e330–e7. doi: 10.1227/ons.0000000000000888
69. Burton W, Myers C, Rutherford M, Rullkoetter P. Evaluation of single-stage vision models for pose estimation of surgical instruments. *Int J Comput Assist Radiol Surg.* (2023) 18(12):2125–42. doi: 10.1007/s11548-023-02890-6
70. De Backer P, Van Praet C, Simoens J, Peraire Loes M, Creemers H, Mestdagh K, et al. Improving augmented reality through deep learning: real-time instrument delineation in robotic renal surgery. *Eur Urol.* (2023) 84(1):86–91. doi: 10.1016/j.euro.2023.02.024
71. Kiyasseh D, Ma R, Haque TF, Miles BJ, Wagner C, Donoho DA, et al. A vision transformer for decoding surgeon activity from surgical videos. *Nat Biomed Eng.* (2023) 7(6):780–96. doi: 10.1038/s41551-023-01010-8
72. Lin W, Shi J, Ma H, Wang L, Hu J, Zhou C. Vision-based real-time tracking of surgical instruments in robot-assisted laparoscopic surgery. *IECON 2023- 49th Annual Conference of the IEEE Industrial Electronics Society*; 2023 16-19 Oct (2023).
73. Louis N, Zhou L, Yule SJ, Dias RD, Manojlovich M, Pagani FD, et al. Temporally guided articulated hand pose tracking in surgical videos. *Int J Comput Assist Radiol Surg.* (2023) 18(1):117–25. doi: 10.1007/s11548-022-02761-6
74. Luijten G, Gsaxner C, Li J, Pepe A, Ambigapathy N, Kim M, et al. 3D surgical instrument collection for computer vision and extended reality. *Sci Data.* (2023) 10(1):796. doi: 10.1038/s41597-023-02684-0
75. Ran B, Huang B, Liang S, Hou Y. Surgical instrument detection algorithm based on improved YOLOv7x. *Sensors (Basel).* (2023) 23(11). doi: 10.3390/s23115037
76. Yang JH, Goodman ED, Dawes AJ, Gahagan JV, Esquivel MM, Liebert CA, et al. Using ai and computer vision to analyze technical proficiency in robotic surgery. *Surg Endosc.* (2023) 37(4):3010–7. doi: 10.1007/s00464-022-09781-y
77. Zheng L, Liu Z. Real time surgical instrument object detection using YOLOv7. In: Stanimirović PS, Zhang Y, Xiao D, Cao X, editors. *6th EAI International Conference on Robotic Sensor Networks (ROSENET 2022) EAI/Springer Innovations in Communication and Computing. 6th EAI International Conference on Robotic Sensor Networks.* Cham: Springer International Publishing (2023). p. 81–90.
78. Badilla-Solórzano J, Gellrich N-C, Seel T, Ihler S. Modular, label-efficient dataset generation for instrument detection for robotic scrub nurses. In: Xue Y, Chen, Chen C., Zuo C., Liu L., Y., editors. *Data Augmentation, Labelling, and Imperfections Miccai 2023 Lecture Notes in Computer Science. Data Augmentation, Labelling, and Imperfections.* Cham: Springer Nature Switzerland (2024). p. 95–105.
79. Balu A, Kugener G, Pangal DJ, Lee H, Lasky S, Han J, et al. Simulated outcomes for durotomy repair in minimally invasive spine surgery. *Sci Data.* (2024) 11(1):62. doi: 10.1038/s41597-023-02744-5
80. Jurosch F, Wagner L, Jell A, Islerst E, Wilhelm D, Berlet M. Extra-abdominal trocar and instrument detection for enhanced surgical workflow understanding. *Int J Comput Assist Radiol Surg.* (2024) 19(10):1939–45. doi: 10.1007/s11548-024-03220-0
81. Kil I, Eidt JF, Singapogu RB, Groff RE. Assessment of open surgery suturing skill: image-based metrics using computer vision. *J Surg Educ.* (2024) 81(7):983–93. doi: 10.1016/j.jsurg.2024.03.020
82. Kim KH, Koo HW, Lee BJ. Deep learning-based localization and orientation estimation of pedicle screws in spinal fusion surgery. *Korean J Neurotrauma.* (2024) 20(2):90–100. doi: 10.13004/kjnt.2024.20.e17
83. Raymond MJ, Biswal B, Pipaliya RM, Rowley MA, Meyer TA. Convolutional neural network-based deep learning engine for mastoidectomy instrument recognition and movement tracking. *Otolaryngol Head Neck Surg.* (2024) 170(6):1555–60. doi: 10.1002/ohn.733
84. Sugiyama T, Sugimori H, Tang M, Ito Y, Gekka M, Uchino H, et al. Deep learning-based video-analysis of instrument motion in microvascular anastomosis training. *Acta Neurochir (Wien).* (2024) 166(1):6. doi: 10.1007/s00701-024-05896-4
85. Viviers CGA, Filatova L, Termeer M, de With PHN, van der Sommen F. Advancing 6-DoF instrument pose estimation in variable x-ray imaging geometries. *IEEE Trans Image Process.* (2024) 33:2462–76. doi: 10.1109/TIP.2024.3378469
86. Haddaway NR, Page MJ, Pritchard CC, McGuinness LA. PRISMA2020: an R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis. *Campbell Syst Rev.* (2022) 18(2):e1230. doi: 10.1002/cl2.1230
87. Joshi. *Machine Learning and Artificial Intelligence.* 2nd ed. Cham, Switzerland: Springer Nature (2023).
88. O'Shea K, Nash R. An Introduction to Convolutional Neural Networks. *ArXiv: abs/1511.08458* (2015).
89. Salcedo-Sanz S, Rojo-Álvarez JL, Martínez-Ramón M, Camps-Valls G. Support vector machines in engineering: an overview. *Wiley Interdiscip Rev Data Min Knowl Discov.* (2014) 4(3):234–67. doi: 10.1002/widm.1125

90. Breiman L. Random forests. *Mach Learn.* (2001) 45(1):5–32. doi: 10.1023/A:1010933404324
91. Taunk K, De S, Verma S, Swetapadma A. A brief review of nearest neighbor algorithm for learning and classification. *2019 International Conference on Intelligent Computing and Control Systems (ICCS)* (2019). p. 1255–60
92. Schmidt RM. Recurrent Neural Networks (RNNs): A Gentle Introduction and Overview. *ArXiv: abs/1912.05911* (2019).
93. Bindra K, Misra DA. A detailed study of clustering algorithms. *2017 6th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO);* (2017). p. 371–6.
94. Jiang T, Gradus JL, Rosellini AJ. Supervised machine learning: a brief primer. *Behav Ther.* (2020) 51(5):675–87. doi: 10.1016/j.beth.2020.05.002
95. AlMahamid F, Grolinger K. Reinforcement learning algorithms: an overview and classification. *2021 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE);* (2021). p. 1–7.
96. Montesinos López OA, Montesinos López A, Crossa J. Fundamentals of artificial neural networks and deep learning. In: Montesinos López OA, Montesinos López A, Crossa J, editors. *Multivariate Statistical Machine Learning Methods for Genomic Prediction.* Cham: Springer International Publishing (2022). p. 379–425.
97. Sarker IH. Deep learning: a comprehensive overview on techniques, taxonomy. *Applications and Research Directions. SN Comput Sci.* (2021) 2(6):420. doi: 10.1007/s42979-021-00815-1
98. Karrenbach M, Preechayasomboon P, Sauer P, Boe D, Rombokas E. Deep learning and session-specific rapid recalibration for dynamic hand gesture recognition from EMG. *Front Bioeng Biotechnol.* (2022) 10:1034672. doi: 10.3389/fbioe.2022.1034672
99. Loukas C, Lahanas V, Georgiou E. An integrated approach to endoscopic instrument tracking for augmented reality applications in surgical simulation training. *Int J Med Robot.* (2013) 9(4):e34–51. doi: 10.1002/rcs.1485
100. Arda MS, Guirnaldo SA, Permites ID, Salaaan CJO, editors. Object detection as a technological adjunct to the manual counting protocol during surgery. *2021 IEEE 13th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM);* 2021 28–30 Nov. (2021).
101. Chen L, Ma L, Zhang F, Yang X, Sun L. An intelligent tracking system for surgical instruments in complex surgical environment. *Expert Syst Appl.* (2023) 230:120743. doi: 10.1016/j.eswa.2023.120743
102. Perceptual docking for robotic control. In: Yang G-Z, Mylonas GP, Kwok K-W, Chung A, editors. *Medical Imaging and Augmented Reality.* Berlin, Heidelberg: Springer Berlin Heidelberg (2008). p. 21–30.
103. Huang X, Xue Y, Ren S, Wang F. Sensor-based wearable systems for monitoring human motion and posture: a review. *Sensors (Basel).* (2023) 23(22). doi: 10.3390/s23229047
104. Hartwig R, Dorigan A, Ostler D, Wilhelm D. CARS 2020-Computer Assisted Radiology and Surgery Proceedings of the 34th International Congress and Exhibition, Munich, Germany, June 23–27, 2020. *Int J Comput Assist Radiol Surg.* (2020) 15(Suppl 1):1–214. doi: 10.1007/s11548-020-02171-6
105. Chen D, Wang M, He C, Luo Q, Iravantchi Y, Sample A, et al. *Magx. Proceedings of the 27th Annual International Conference on Mobile Computing and Networking;* New Orleans, Louisiana: Association for Computing Machinery (2021). p. 269–82.
106. Bordin A, Tibamoso-Pedraza G, Miró J, Duong L. CARS 2023-Computer Assisted Radiology and Surgery Proceedings of the 37th International Congress and Exhibition, Munich, Germany, June 20–23, 2023. *Int J Comput Assist Radiol Surg.* (2023) 18(S1):1–123. doi: 10.1007/s11548-023-02878-2
107. Khan DZ, Luengo I, Barbarisi S, Addis C, Culshaw L, Dorward NL, et al. Automated operative workflow analysis of endoscopic pituitary surgery using machine learning: development and preclinical evaluation (ideal stage 0). *J Neurosurg.* (2022) 137(1):51–8. doi: 10.3171/2021.6.JNS21923
108. Cheng K, You J, Wu S, Chen Z, Zhou Z, Guan J, et al. Artificial intelligence-based automated laparoscopic cholecystectomy surgical phase recognition and analysis. *Surg Endosc.* (2022) 36(5):3160–8. doi: 10.1007/s00464-021-08619-3
109. Kitaguchi D, Takeshita N, Matsuzaki H, Takano H, Owada Y, Enomoto T, et al. Real-time automatic surgical phase recognition in laparoscopic sigmoidectomy using the convolutional neural network-based deep learning approach. *Surg Endosc.* (2020) 34(11):4924–31. doi: 10.1007/s00464-019-07281-0
110. Filicori F, Bitner DP, Fuchs HF, Anvari M, Sankaranarayanan G, Bloom MB, et al. SAGES video acquisition framework-analysis of available OR recording technologies by the SAGES AI task force. *Surg Endosc.* (2023) 37(6):4321–7. doi: 10.1007/s00464-022-09825-3
111. Yilmaz R, Browd S, Donoho DA. Controversies in artificial intelligence in neurosurgery. *Neurosurg Clin N Am.* (2025) 36(1):91–100. doi: 10.1016/j.nec.2024.08.008
112. Goodman ED, Patel KK, Zhang Y, Locke W, Kennedy CJ, Mehrotra R, et al. Analyzing surgical technique in diverse open surgical videos with multitask machine learning. *JAMA Surg.* (2024) 159(2):185–92. doi: 10.1001/jamasurg.2023.6262
113. Hassan J, Leong J, Schneider B. Multimodal data collection made easy: the EZ-MMLA toolkit. *LAK21: 11th International Learning Analytics and Knowledge Conference;* Irvine, CA, USA: Association for Computing Machinery (2021). p. 579–85.
114. Satapathy P, Hermis AH, Rustagi S, Pradhan KB, Padhi BK, Sah R. Artificial intelligence in surgical education and training: opportunities, challenges, and ethical considerations - correspondence. *Int J Surg.* (2023) 109(5):1543–4. doi: 10.1097/JS9.0000000000000387
115. Adegbesan A, Akingbola A, Aremu O, Adewole O, Amamdikwa JC, Shagaya U. From scalpels to algorithms: the risk of dependence on artificial intelligence in surgery. *J Med Surg Public Health.* (2024) 3:100140. doi: 10.1016/j.jglmedi.2024.100140
116. Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial intelligence in surgery: promises and perils. *Ann Surg.* (2018) 268(1):70–6. doi: 10.1097/SLA.0000000000002693
117. Tizhoosh HR, Diamandis P, Campbell CJV, Safarpour A, Kalra S, Maleki D, et al. Searching images for consensus: can AI remove observer variability in pathology? *Am J Pathol.* (2021) 191(10):1702–8. doi: 10.1016/j.ajpath.2021.01.015
118. Othman D, Kaleem A. The intraoperative role of artificial intelligence within general surgery: a systematic review. *Cureus.* (2024) 16(11):e73006. doi: 10.7759/cureus.73006
119. Jiang L, Wu Z, Xu X, Zhan Y, Jin X, Wang L, et al. Opportunities and challenges of artificial intelligence in the medical field: current application, emerging problems, and problem-solving strategies. *J Int Med Res.* (2021) 49(3):3000605211000157. doi: 10.1177/03000605211000157
120. Liu PR, Lu L, Zhang JY, Huo TT, Liu SX, Ye ZW. Application of artificial intelligence in medicine: an overview. *Curr Med Sci.* (2021) 41(6):1105–15. doi: 10.1007/s11596-021-2474-3
121. Wen J, Zhang Z, Lan Y, Cui Z, Cai J, Zhang W. A survey on federated learning: challenges and applications. *Int J Mach Learn Cybern.* (2023) 14(2):513–35. doi: 10.1007/s13042-022-01647-y