CHAPTER 8

Artificial intelligence assisted surgery

Elan Witkowski and Thomas Ward
Department of Surgery, Massachusetts General Hospital/Harvard Medical School, Boston, MA, United States

8.1 Introduction

The practice of surgery involves physical manipulation of tissue to treat disease. Over many centuries, surgical outcomes have improved with increasing human knowledge and the development of novel tools. Artificial intelligence (AI) has been defined as the study of algorithms that give machines the ability to reason and perform functions such as problem solving, object and word recognition, inference of world states, and decision-making [1]. Surgery is therefore a natural and yet complex application for AI technologies. Operations require surgeons to synthesize data from multiple sources to make decisions, identify anatomy, and carry out physical tasks in rapidly changing scenarios. Outside of the operating room, components of surgical care include diagnosis, preoperative assessment, postoperative care, assessment of outcomes, and training of surgeons (Fig. 8.1). AI promises to improve the quality and efficiency of perioperative care, improve surgical decision-making, augment the physical capabilities of human surgeons, and offer many exciting opportunities for future investigation—though not without potential pitfalls and challenges. In this chapter, past, present, and future applications of AI in surgery are reviewed.

8.2 Preoperative

Sun Tzu's declaration that "Every battle is won or lost before it is fought" rings especially true with surgery. Surgeons lay the operation's true foundation before the case itself with adequate preoperative patient selection, assessment, and optimization. Surgery is a controlled insult on the human physiology, and therefore the patient must have sufficient reserve to

Artificial Intelligence in Healthcare
DOI: https://doi.org/10.1016/B978-0-12-818438-7.00008-3

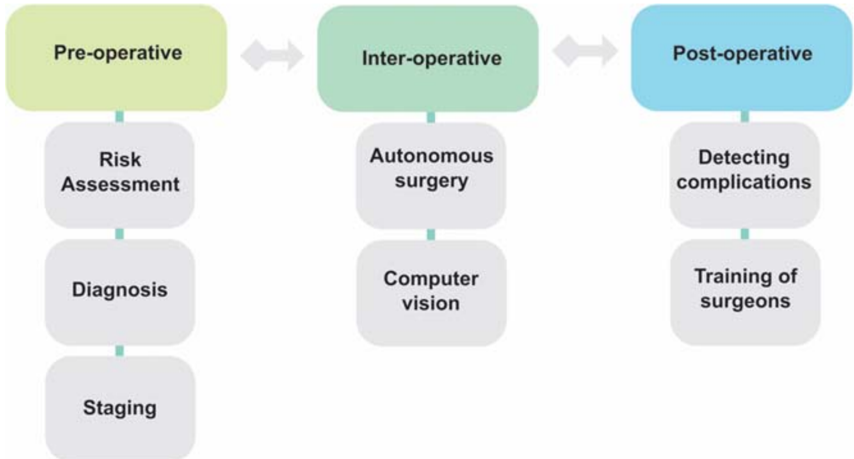

Figure 8.1 Applications of artificial intelligence in surgery and perioperative care.

tolerate the insult. The surgeon must also have adequate preoperative planning to correctly diagnose and assess the extent of the disease they will address intraoperatively. Lastly, the progress of medicine, in particular, neoadjuvant chemoradiotherapy prior to surgical resection of cancer, may obviate the need for curative resection and surgeons will need to identify these patients and prevent unnecessary operations.

8.2.1 Preoperative risk assessment

Before proceeding with an operation, the surgeon must determine if the patient can tolerate the procedure. Of paramount importance is adequate cardiac function. The American College of Cardiology and American Heart Association publish guidelines to guide physicians with risk-stratification and preoperative optimization [2]. These guidelines stratify patients into low-, medium-, and high-risk categories which correlate to probability of a perioperative major adverse cardiac event (MACE) of death or myocardial infarction. To evaluate a patient preoperatively, first physicians use risk-assessment calculators to estimate perioperative risk of MACE. These calculators include the Revised Cardiac Risk Index (RCRI) developed in the 1990s and the more recent Gupta Myocardial Infarction or Cardiac Arrest (MICA) risk model [3,4] After establishing that a patient is at elevated risk ($>1\%$), they then use subjective assessment of a patient's metabolic equivalents (METs) functional capacity. As long as the patient has adequate functional capacity, they are deemed to not require further expensive preoperative cardiac testing aimed at identifying intervenable coronary disease that could be optimized prior to surgery.

Recent studies call into question this algorithm [5,6]. The POISE trial showed 8000 patients who underwent noncardiac surgeries had a MACE rate of 6.9% [5]. The majority who had a MACE had either one or two points on the RCRI scale which should have correlated to 1.0%–2.4% MACE rate. Clearly, the scale underperformed. More recently, a group looked at patients' subjective assessment of their METs functional capacity. Patients' subjective assessments were only 19% sensitive to predict what METs they could obtain during formal cardiopulmonary exercise testing. This underestimation led to a significant proportion of high-risk patients placed incorrectly into a low-risk category. Obviously, preoperative cardiac evaluation methods need further refinement.

The American Society of Anesthesiologists (ASA) has also proposed a classification system to assess perioperative risk [7]. It look beyond cardiac status alone and rate patients on a scale from I to V based upon their overall physical status. Examining over 6000 patients, higher classes correlated to increased risk odds ratio (OR) of postoperative complications: ASA class III and IV with OR 2.2 and 4.2, respectively. Perhaps, the ASA scale's primary limitation is the inherit subjectivity (e.g., mild vs severe systemic disease vs unlikely to survive 24 hours). One survey of 304 board-certified anesthesiologists gave them ten patients to classify according to the ASA scale. Only six of the patients received consistent ratings [8]. Again, current methods of risk assessment are found lacking.

Clearly, current methods of preoperative risk assessment need improvement. Their shortcomings stem from one of two issues. The first issue is subjectivity. This subjectivity limits consistent application of models such as the ASA classification and METs functional capacity and can lead to striking underestimation of risk. The second issue is the assumption of models that risk factors are additive and consistently related. Often times, these models assess risk factors in a binary fashion (present vs absent). In addition to the cardiac risk calculators previously named, one such calculator is the American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP) calculator. This calculator uses a generalized linear mixed model to predict risk for multiple complications [9].

Medicine is obviously more nuanced, which reflects how these models consistently underperform. Application of machine learning (ML) methods could remove subjectivity and may allow for a better incorporation of variables in a nonlinear way that better reflects clinical reality. As ML models have gained popularity, many analyses and risk calculators that were previously created with traditional regression modeling techniques have been reconceived (using the same or similar data sets) with neural networks and other algorithmic approaches. A group from Massachusetts General Hospital recently created a ML-based score for emergency surgery patients [10]. Their algorithm (POTTER) used optimal classification trees to create a comprehensive decision tree to predict both mortality and morbidity [11]. Their score outperformed the ASA classification ACS-NSQIP calculator for both morbidity and mortality prediction, with a c-statistic of 0.9162 for over 382,960 patients. The strength of the POTTER approach was in its ability to more appropriately draw on data from representative patients to achieve risk prediction. Being one of the first ML-based calculators of its kind, the future should bring further scores and therefore refinement in our ability to stratify patients at risk for undergoing surgery and ultimately improve decision-making and counseling, assess appropriateness, and optimally prepare them for surgery—examples.

8.2.2 Preoperative diagnosis

Early diagnosis and detection of malignant lesions are essential. Almost universally, earlier detection leads to improved outcomes. One example is melanoma. Excised melanomas with a thickness under 1 millimeter (mm) have a 95% 5-year survival rate. More locally advanced melanomas with a thickness over 4 mm only have a 45% 5-year survival rate [12]. Therefore there has been a push for earlier and improved detection of melanomas. Local tools such as dermoscopy initially served as an extension of the examiner's eye when looking for malignant skin lesions. Dermoscopy uses a combination of a magnifying device with cross-polarized light to make the initial few layers of the skin translucent which affords analysis of deeper skin structures and improved diagnosis. The use of dermoscopy leads to an OR of 9.0−15.6 for detection of melanoma [13].

One trial sought to go beyond extending physicians' ability to visualize a lesion by extending a physician's judgment with ML. Research has previously shown that physicians are willing to change their clinical decisions a quarter of the time when presented with conflicting output from a clinical decision-support system [14]. The study enrolled expert clinicians (dermatologists) and nonexpert clinicians to examine 511 patients who had a

Figure 8.2 Example of a smart phone with a neural-network decision-support tool to analyze skin lesion images.

combined 3827 pigmented skin lesions [15]. The nonexpert physicians were allowed to use a neural-network decision-support tool to analyze skin lesion images they acquired with dermoscopy to increase their accuracy to the level of an expert clinician. The nonexperts, with the decision-support system's assistance, were actually able to exceed the specificity of an expert dermatologist (82% vs 72%).

Subsequently, a deep convolutional neural network was trained on 129,450 clinical images to classify images malignant versus benign categories [16]. Not only did it analyze dermoscopy images but also regular photos. When tested against 21 dermatologists, the system achieved superior sensitivity and specificity. The authors further added that the images could be acquired with only smartphones. Through a smartphone alone, the public could then have the ability to diagnosis malignant skin lesions (Fig. 8.2). This would provide almost universal access to specialist-level diagnostic care at a low cost through the power of AI and ML.

8.2.3 Preoperative staging

Once diagnosed, malignancies next need staging to determine resectability. The American Joint Committee on Cancer publishes a manual that is the gold standard and stages of tumors based on Tumor size/depth, Nodal status, and Metastases (TNM) [17]. Clinicians use history, physical exam, laboratory values, and radiographic imaging to determine a tumor's clinical stage and therefore resectability. If resectable, then the patient traditionally would go straight to surgery and then postoperative chemotherapy and radiation therapy (known as "adjuvant therapy"). More-and-more, there has been a move with certain tumors to perform chemoradiotherapy prior to surgical resection (neoadjuvant therapy) with improved survival and cure for patients. Examples include both esophageal and rectal cancer [18−20]. Interestingly, with the switch of chemoradiotherapy treatment prior to resection, surgeons found some resected specimens with no residual cancer. Esophageal and rectal cancer both had pathological complete responses (pCRs) with no residual tumor around 30% of the time [19,21]. With this observation, groups looked at the previously unimaginable: in patients with a complete clinical response opting to "watch and wait" rather than removing resectable disease [21,22]. Avoiding surgery limits the high morbidity that can accompany both rectal and esophageal resections. Initial results are promising, with overall 5-year survival and disease-free survival rates in rectal cancer of 100% and 92%, respectively, in one study [21] and no difference in 3-year overall survival in another [22]. In the latter study, 26% of patients were spared the morbidity of a permanent colostomy.

Of course, a complete clinical response does not necessarily correlate to a pCR which can only be known if the tumor is resected. As expected, local rectal cancer recurrences occurred in 34% of patients from assumed residual disease after chemoradiotherapy [22]. Interestingly, even in patients without complete clinical response, 7% had pCR in one cohort [21]. Groups have therefore looked at additional factors to predict pCR and limit the number of recurrences in patients initially manage nonoperatively. One group looked at pretreatment levels of carcinoembryonic antigen (CEA), a colorectal cancer serum tumor marker, and found that elevated pretreatment CEA was associated with a lower rate of pCR [23]. Another group performed multivariate analysis across over 300 rectal cancer patients and found that sex, age, body mass index (BMI), radiation dose, tumor differentiation, clinical stage, and CEA were not associated with pCR but that delaying surgery more than 8 weeks after chemoradiotherapy was associated with pCR [24]. With these confounding results, other researchers looked at other possible predictors including fluorodeoxyglucose-positron emission tomography (FDG PET) scans, which detect metabolically active tumors with radiolabeled glucose molecules. They found that the lower standardized uptake value posttreatment and pretherapy correlated with pCR [25]. Macomber et al. then applied ML techniques to various FDG PET features (mean uptake, max uptake, total lesion glycolysis, and volume) of esophageal cancers to predict pCR [26]. Their methods included k nearest neighbors, decision trees, support vector machines, and Naïve Bayes decision analysis (Fig. 8.3). The k nearest neighbors had the highest accuracy at 0.74. Future methods will work to incorporate not only imaging, but a variety of other clinical data with ML techniques to improve our ability to predict pCR. This improved prediction, surgery may transition from the first-line treatment to a rescue option; and with this transition, we may spare more-and-more patients the serious morbidity associated with these large operations.

8.3 Intraoperative

Some of the most exciting and promising applications of AI technology occurs in the operating room. An estimated 234 million surgeries happen across the globe annually and there exists a significant room for improvement: up to 20% of surgical patients experience complications [27,28]. AI technology (including computer vision and computer-assisted or eventually autonomous surgery) offers promising solutions to assist surgeons in the operating room with the hope of further decreasing complications.

8.3.1 Autonomous surgery

Bowel resection and subsequent reconnection (anastomosis) is one of the fundamental tasks for a general surgeon. Gastrointestinal anastomoses, across the board, leak 4.8% of the time. With each leak came a 10-times increased risk of return to the operating room, almost tripled length of hospital stay, tripled 30-day mortality, and doubled long-term mortality [29]. Reviewing 21,902 patients with a colon cancer resection and colorectal anastomoses, a meta-analysis also shockingly found an OR of 2 for local recurrence, OR 1.38 for distant recurrence, and OR 1.75 for long-term mortality [30]. As for quality of life, a colorectal anastomotic leak leads to more frequent stools, worse control of stools, and worse overall mental and physical component scores [31]. Surgeons traditionally sutured bowel anastomoses. However, leaks plagued the surgical community and innovators sought to ameliorate the problem with technology.

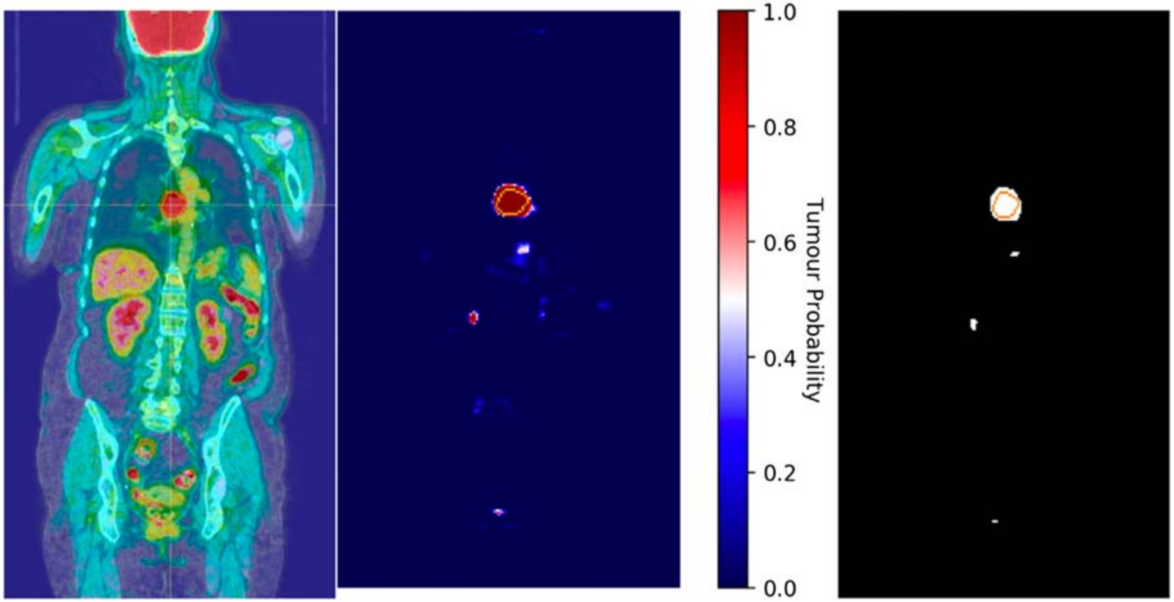

Figure 8.3 Application of deep learning for detection of esophageal lesions in positron emission tomographic-computed tomographic scans. *From I. Ackerley, R. Smith, J. Scuffham, M. Halling-Brown, E. Lewis, E. Spezi, V. Prakash, K. Wells, Using deep learning to detect oesophageal lesions in PET-CT. In Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging (Vol. 10953, p. 109530S), 2019, International Society for Optics and Photonics.*

Ravitch and Steichen popularized stapled (ST) anastomoses in the United States after earlier efforts by Russian surgeons [32]. The reliability and speed of uniformly applied staples versus a hand-sewn (HS) anastomosis theoretically promised faster and safer anastomoses. Unfortunately, this promise failed to bear fruit: most trials and meta-analyses show equivalence in outcomes between ST versus HS anastomoses [33,34]. Emergency general surgery is one area that may show an edge to HS anastomoses. One retrospective study showed almost double failure rate for ST compared to HS anastomoses [35]. The authors hypothesized that these relatively sicker patients had more likely bowel wall edema and size mismatch which a uniformly applied staple load could not properly address while a surgeon could customize for the discordance with each individual stitch. A later multicenter prospective trial enrolled patients and found that surgeons given this theoretical advantage preferred HS anastomosis for sicker patients (higher lactates, lower albumin, more vasopressor utilization, higher BMIs). Their leak rates were identical (~10%) despite the patients with HS anastomoses being significantly sicker [36]. Therefore HS anastomoses offer the potential promise for safer anastomoses in sicker patients, but still require an expert hand. With leak rates still approaching 5%–10%, the surgical community has room for improvement.

Autonomous surgery offers hope for safe surgery, particularly for gastrointestinal anastomoses. Researchers at Children's National Hospital developed a vision-guided robotics system that can perform autonomous suturing [37,38]. The Smart Tissue Anastomosis Robot (STAR), later named Smart Tissue Autonomous Robot, is a robotic arm with a suturing tool based on the Endo360° suturing device by EndoEvolution, LLC mounted to a light-weight robot. In its first iteration, it came with two modes: a manual mode and an automatic mode. The manual mode meant a surgeon selected each location for stitch placement. The automatic mode had a surgeon outline a linear incision and the STAR would take bites at a predefined interval. Comparing the STAR systems versus laparoscopic (both plain and robotic-assisted) surgery, the STAR system could finish a 5.5 cm incision closure in manual and automatic modes in 64.51 and 70.93 seconds, respectively. The laparoscopic surgeon took 9.34 minutes and the robotic-assisted surgeon took 5.71 minutes [37]. Not only was the STAR system more efficient, but it was more precise; stitches placed by surgeons had double the standard deviation for distance variation compared to the robot (Fig. 8.4). In addition, due to being able to

Figure 8.4 Image of smart tissue autonomous robot during preparation for intestinal anastomosis procedure. *From E. Svoboda,Your robot surgeon will see you now, Nature 573 (Sept), 2019, S110–S111.*

set defined tension for each suture bite, there was an order of magnitude less deflection of the stitches when they were pulled on tension compared to the much looser surgeon placed stitches (0.25 vs 2 mm).

Minimally invasive surgery, like autonomous surgery, aspires for safer and improved surgical practices. It was popularized as laparoscopic surgery matured in the 1980s, where a surgeon insufflates the abdomen then uses instruments through small incisions [39]. With the smaller incisions and decreased intraabdominal trauma, laparoscopic surgery compared to open comes with reduced postoperative pain, decreased wound infection rate, and quicker time to resumption of usual activities [40]. Laparoscopic surgery continues to present numerous technical challenges, however, with limited degrees of freedom for instrument articulation, diminished tactile feedback, and limited 2D vision, which makes technically challenging procedures more difficult except for the master laparoscopic surgeons. Recently, computer-assisted laparoscopic surgery offers augmented capabilities to address some of these issues. The da Vinci robotic system (Intuitive Surgical Inc) offers a "master-slave" configuration with the surgeon at a bedside console using inputs to control the robotic arms inside the patient [41]. Its display gives a three-dimensional image with instruments that have increased degrees of freedom to address the prior listed shortcomings of laparoscopic surgery. Despite these theoretical advantages, the use of robotic-assisted surgery has not created significant outcome differences. Short-term results are comparable in large reviews of colorectal, gynecologic, and urologic surgery [42–44]. The robotic-assisted surgery also negatively brought with increased cost ($2000 per case in the gynecologic operations) and length of procedure time. Despite the promise of laparoscopic and robotic-assisted surgery, they so far failed to deliver concrete improvements in patient outcomes.

Patients' surgical outcomes still rely on a human's technical ability which may inherently limit the impact of laparoscopy and robotic-assisted surgery. The creators of STAR built upon their prior work with a particular focus on bowel anastomoses [38]. Prior iterations could only suture on simple planar surfaces. To address the more realistic elastic surfaces one encounters throughout surgery (bowel, airways, blood vessels), the creators used both a 3D visual tracking system and a plenotopic 3D surface reconstruction system to plan suture placement. First, a surgeon placed a few near-infrared fluorescent imaging markers onto the corners of the bowel to give STAR the outlines of the planned anastomosis. Then, the plenotopic camera created a 3D point cloud in-between the markers to map the remaining bowel edges. The STAR then placed stitches based on an algorithm for optimal suture spacing and tension. Compared to open, laparoscopic, and robotic-assisted techniques, the STAR created equivalent if not superior anastomoses. With the consistent stitch placement and ideal tension application on the suture, the STAR anastomoses could withstand nearly double leak pressure compared to the other anastomoses. This technology offers great promise as it could enable any surgeon to instantly have the ability to perform a technically excellent anastomosis. The device adjusts tension with each stitch, therefore giving it unparalleled ability to adjust for varying tissue conditions. As discussed previously, HS anastomoses showed an edge over ST ones in patients with thick edematous bowels likely due to surgeons adjusting their tension and stitch placement for the tissue. Staplers do not adjust. The STAR therefore gives the benefit of an automated anastomosis that can adjust with each stitch and form the ideal connection to even the most novice of surgeons.

Though exciting, truly autonomous surgical systems still are several years from clinical application. In the interim, the current surgical landscape still has much room for improvement [38]. The STAR takes 50 minutes to perform a bowel anastomosis compared to just 8 minutes for an open surgeon. It also currently operates in a "supervisory autonomous mode," which allows a surgeon to make minor adjustments prior to each stitch placement and is necessary for 42% of stitches. Obviously in its infancy, the technology will continue to grow. Already the creators expanded the system to semiautonomously resect tumors with electrocautery [45]. Then once viable the technology will need to go through the clinical approval process. While we wait as others develop an autonomous future, AI and ML techniques for computer vision offer a more immediate impact in the intraoperative phase of surgical care as an expansion of a surgeon's collective knowledge and pattern recognition capability.

8.3.2 Computer vision

What ultimately separates a good from an excellent surgeon is not technical ability but intraoperative judgment. Skills such as suturing, knot tying, and stapling form surgical fundamentals necessary for performance of an operation. To dissect out vital structures and reach the stage to use these fundamentals, however, requires a thoughtful surgeon who uses pattern recognition to progress an operation. Surgical trainees and junior surgeons often lack that refined ability to see planes and quickly but safely operate. Birkmeyer et al. showed with an analysis of almost half-a-million patients that for more complicated procedures (esophagectomy, lung resection, aortic valve replacement, abdominal aortic aneurysm repair, pancreatic resection, coronary artery bypass grafting, and cystectomy) that a surgeon's case volume had an inverse relationship to patient mortality. The OR for operative death between a low-volume and high-volume surgeon varied from 1.24 for lung resections to 3.61 for pancreatic resections [46]. Similarly, 20 bariatric surgeons in Michigan submitted videos of their complex laparoscopic bariatric surgeries. Independent surgeons then rated each video. Surgeons in the top quartile had lower rates of reoperation, readmission, emergency department visits, surgical complications, and medical complications [47]. Performing an operation frequently likely gave those surgeons a better appreciation for the nuances of the case to more safely and expediently perform them and minimize postoperative complications.

Currently, the only way to improve is to do more surgery. For pancreatic resections, surgeons become dramatically more proficient after 60 cases, with reduced blood loss, operative time, and rate of removed specimen's positive cancer margins [48]. In laparoscopic colonic resections, Tekkis et al. showed that completion of a case without need to convert to open also correlated to experience, with about 60 cases (55 for right-sided resection, 62 cases for left-sided resections) needed for competency

[49]. While surgeons acquire experience, they subject patients to objectively worse outcomes. Fortunately, AI and ML offer a potential solution: computer vision.

Computer vision gives a machine understanding of images and videos. This chapter mentioned some examples of computer vision earlier, with diagnosis of dermatologic processes and STAR's soft-tissue mapping capabilities [16,38]. In the 1990s and 2000s, computer vision had few applications due to the limited capabilities of earlier AI approaches. Convolutional neural networks (ConvNets) led to a revolution. ConvNets were designed to process multiple arrays, such as data representation of images which have multiple arrays of images' pixel values. They, in fact, ultimately take on a networked structure quite similar to the visual cortex pathway in primates [50]. Krizhevsky et al. applied ConvNets with great success on the ImageNet challenge, an image recognition test across 1000 image categories, to obtain an error rate half that of all other competitors [51]. This advance, combined with increased processing ability of computers, led to an explosion of computer vision applications throughout the industry, including recent advances for intraoperative real-time guidance.

Early applications of computer vision to intraoperative recognition of surgical phases and structures were limited. As expected, trying to identify a constantly changing operative field brings challenges. Some groups identified that instruments were easily identifiable and tried to use the appearance of certain instruments (e.g., the clips used in gallbladder surgery to ligate the ducts and arteries) to predict the current stage of an operation [52,53]. A typical operation typically uses instruments at multiple stages and some surgeons use different instruments at different times, often in quite unique ways, which limited this approach. Other groups, similar to the STAR approach described earlier, used near-infrared markers (NIFs). At a case's start, surgeons would apply biodegradable inert markers at key locations in the operative field to allow the system to then identify structures and predict an operation's course—essentially augmented reality [54]. The surgeon's need for placement at an operation's start and the inherent limitations of a constantly changing operative field with instruments and organs intermittently concealing these markers restricted NIFs' ultimate ceiling.

More recently, a group developed a method that used a standard high-definition video of laparoscopic surgery to identify surgical phases in real time that did not rely on instrument detection nor surgeon-placed markers [55]. One issue with processing live video is that 1 minute of data contains 25 times the amount of data in a high-resolution computer tomography scan image [56]. To solve this, images were made into a representative histogram for analysis based off visual cues determined from discussions with surgeons. These cues included color, position, shape, and texture. Next, the histograms were processed into a compressed amount of data called a k-segment coreset stream which stripped out all information not useful to determine real-time operative phase identification. These compressive steps reduced the size of the data by over 90% with no accuracy loss [55]. With a useable amount of information, they could then apply support vector machine and hidden Markov models to classify and predict surgical phases. After training the system on only 10 hours of video, they could, with 92.8% accuracy, predict the stage of a laparoscopic sleeve gastrectomy in real time using just a standard laparoscopic surgery high-definition video feed.

Real-time operative computer vision analysis promises exciting possibilities. Upon review of surgical videos, we already can reliably detect negative intraoperative events. A Canadian group reviewed 54 recordings of bariatric laparoscopic surgery. In 70% of the recordings, they detected near-miss events (events that if not detected and addressed would lead to adverse outcomes and injury). Two-thirds of the events needed additional interventions such as further suturing or attempts at hemostasis. Fortunately, in these select videos, the surgeon detected all events intraoperatively [57]. Computer vision techniques could detect these events in real time and alert the surgeon, hopefully with even earlier detection than the surgeon whose focus may be on a separate area of the visual field. Computer vision could possibly even identify before a surgeon performs a near-miss event and alert them to prevent it from happening in the first place. In this sense, any surgeon, no matter how novice, would have the collective surgical experience from every surgeon whose videos trained the machine. Each operation will have a "surgical fingerprint" that once deviated from, the machine can alert the surgeon and help guide them back to a safer course. No longer would patients' outcomes suffer as the surgeon climbed the previously described learning curves. Instead, they would have the real-time decision support of multiple senior surgeons to guide them through their operations. As we continue to face a shortage of surgeons, particularly in rural communities, general surgeons will need to safely perform a myriad of operations, for which many they will not have had the experience or volume to have yet surmounted the learning curve [58]. Decision support from computer vision will level the playing field and shrink healthcare inequality to allow for safe delivery of care to patients, no matter their access to high-volume centers.

8.4 Postoperative

8.4.1 Detecting complications

"When the abdomen is open you control it, when closed it controls you."

Moshe Schein [59].

Much to the chagrin of surgeons, intraoperative excellence only partly determines an operation's ultimate success. Recent literature discusses that the majority of perioperative mortality differences originates from "failure to rescue," where institutions are unable to prevent mortality after an adverse complication [60]. Complication rates across the top and bottom 20 percentile institutions over 270,000 patients are nearly identical. Despite this, the mortality rate is 2.5 times higher for a bottom-tier institution. The difference is even more pronounced for pancreatic surgery with a 13-times higher mortality rate [61]. Some of the treatment gap occurs from delayed recognition of operative complications. There currently exist a number of scoring systems to try and identify patients at higher risk for inpatient complications. The National Early Warning Score (NEWS) from the United Kingdom incorporates clinical parameters to try and predict events in the next 24 hours including cardiac arrest, unanticipated ICU admission, and death. NEWS is only partly effective though, with area under receiver operating characteristic curve (AUROC) values for those respective complications of 0.72, 0.56, and 0.94 [62].

The linear nature of input variables limits current systems such as NEWS. NEWS, for example, only uses seven parameters, a patient's vital signs and level of consciousness, to make predictions about generalized morbidity and mortality. These limited inputs neglect the remaining wealth of information in the electronic medical record (EMR) including notes, laboratory values, and radiology tests. AI is not limited like these linear and logistic regression models, and the fine-grained nature of AI, particularly with nature language processing (NLP), allows for improved and more-refined predictions. NLP allows a computer to understand human language and therefore perform large-scaled analysis of EMR data [63]. Groups have already developed accurate deep learning of EMR data with AUROC values for mortality and various morbidities outperforming traditional logistic regression models [64,65]. These studies however only looked at patient data from an entire admission rather than snapshots early in the admission to predict later hospital-course complications. Rajkomar et al. advanced this work to develop deep learning models that incorporated as many as 175,639 data points per patient and could predict hospital-stay mortality and even discharge diagnoses after only 24 hours of admission [66].

NLP and deep learning offer promise for earlier detection of postoperative complications and hopefully a resultant decrease in "failure to rescue" (Fig. 8.5). To predict postoperative complications, often the answer lies in examination of the operation itself. Ideally, one could read (or use as an input for ML) the narrative operative report from the surgeon to have a clear sense of each step of the operation. The quality though of an operative report varies drastically from surgeon to surgeon. Almost a third of complications during a laparoscopic cholecystectomy are not dictated into the operative note [67]. Similarly, when surgeons examined systematic video recordings versus narrative operative notes, they found that the notes only adequately described 52.5% of the steps [68]. Accurate assessment of potential intraoperative difficulties leading to complications therefore lies upon analysis of intraoperative video and instrument manipulation.

Hung et al. developed objective metrics to rate surgeons from in robotic-assisted radical prostatectomy. They used surgical manipulation

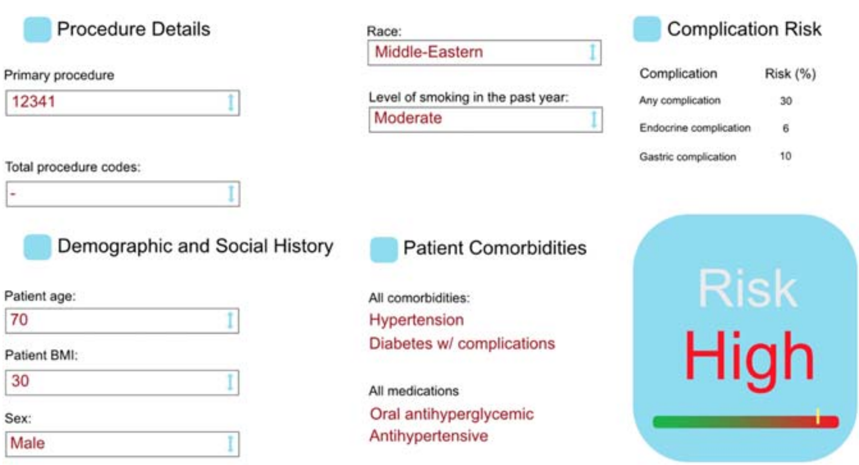

Figure 8.5 Example of surgery classification system for assessing postoperative outcomes including relevant scoring and process functions.

data from the da Vinci Surgical system over 100 cases for ten expert and ten novice surgeons. The data included task completion time, instrument path length, and dominant instrument moving velocity. Through analysis, they were able to objectively classify surgeons into expert and novice categories [69]. Using this objective scale, they then developed a ML algorithm that based on automated performance metrics predicted perioperative outcomes, in particular, prolonged hospital stay over 2 days with 87.2% accuracy [70]. With further development, they hope to broaden predictions to a wider variety of perioperative complications.

Computer vision also offers promise to predict postoperative complications. For example, laparoscopic cholecystectomy, one of the most common surgical procedures, is often an outpatient procedure with few complications. It does, however, have one dreaded complication: common bile duct injury. This injury occurs at a rate of 0.4%–0.6% cases. In one series, surgeons only recognized this complication 32.4% of the time, leading to an increased risk of uncontrolled sepsis and mortality [71]. Furthermore, in another series, 96% of bile duct injuries were found to have occurred as a result of errors in human visual perception [72]. Computer vision offers a chance in real time to prevent such an injury, but also could analyze videos postoperatively to detect the injury for earlier treatment and for future education to prevent the surgeon from making the same mistake again.

8.4.2 Training and certification of surgeons

Continuing education curriculum and recertification processes for surgeons are undergoing significant changes. Historically, to keep one's board-certification active, the American Board of Medical Specialties (ABMS) required board-certified physicians to take a four-part assessment every 10 years. Physicians called this practice into question. The test was "one-size-fits-all" and not tailored to one's clinical practice, so a clinician with a subspecialized practice, such as Breast surgeon, would need to know all of general surgery. The tests additional were quite costly, in terms of both time and cost. Lastly, they only forced a refreshment of knowledge every 10 years [73]. To address the latter concern, the ABMS has transitioned toward promoting lifelong learning with a continuous certification system. The American Board of Surgery (ABS), for example, transitioned the recertification process to a continuous certification which entails a 2-week online assessment process performed every-other-year

[74]. This change unfortunately fails to address that surgery not only requires clinical knowledge, but operative ability, which is not assessed at any point during the recertification process.

Greenberg et al. recognize the importance of operative ability in clinical practice. They note the need for deliberate practice, performance reflection, and intentional adjustments for a surgeon to continue to grow operative skills. Coaches could optimize this growth just as they do in sports. To address the incompleteness of surgical recertification, they advocate that the ABS incorporate active coaching into the certification process. This coaching would provide personalized coaching to a surgeon's practice and also assess and more importantly hopefully improve their surgical technique and outcomes [75]. Obvious limitations include that there are currently almost 23,000 general surgeons across the country [58]. Such a one-on-one coaching experience would be difficult to scale such a population. Computer vision could address this issue with surgical fingerprints. To maintain certification, a surgeon could submit surgical fingerprints obtained from their operative videos. As long as they fall within a level of acceptable technique, their operative ability would then pass recertification [76]. Extrapolating this further, when new techniques (e.g., laparoscopic surgery, robotic-assisted surgery, natural orifice transluminal endoscopic surgery) are invented, surgeons could perform first in a simulated setting then graduate to observed-practice one their surgical fingerprint is acceptable. The computer vision could then in real time guide the surgeons in the learning processes then once they reach acceptable fingerprints, they would become "certified" to perform that procedure independently.

8.5 Conclusion

AI technology offers innumerable promises to advance surgery. In the preoperative phase, it helps surgeons diagnose, adequately stage, and properly risk-stratify patients well-beyond our past capabilities. Intraoperatively, it offers the promise of flawless execution of technical skills through autonomous surgery and with computer vision the augmentation of a surgeon's decision and pattern-recognition abilities with the collective surgical consciousness. Postoperatively, it promises faster recognition of complications to decrease "failure-to-rescue" and integration into the recertification process to more adequately train and evaluate surgeons. Even though we are in the infancy of AI in surgery, it has already achieved startling advancements and promises numerous more in the future. We look forward to the coming years to see how far we can advance surgery and increase the safety of patients everywhere.

References

[1] Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial intelligence in surgery: promises and perils. Ann Surg 2018;268(1):70−6. Available from: https://doi.org/10.1097/SLA.0000000000002693.

[2] Fleisher LA, Fleischmann KE, Auerbach AD, et al. 2014 ACC/AHA guideline on perioperative cardiovascular evaluation and management of patients undergoing non-cardiac surgery: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Circulation 2014;130(24):2215−45. Available from: https://doi.org/10.1161/CIR.0000000000000105.

[3] Lee TH, Marcantonio ER, Mangione CM, et al. Derivation and prospective validation of a simple index for prediction of cardiac risk of major noncardiac surgery. Circulation 1999;100(10):1043−9.

[4] Gupta PK, Gupta H, Sundaram A, et al. Development and validation of a risk calculator for prediction of cardiac risk after surgery. Circulation 2011;124(4):381−7. Available from: https://doi.org/10.1161/CIRCULATIONAHA.110.015701.

[5] POISE Study Group, Devereaux PJ, Yang H, et al. Effects of extended-release metoprolol succinate in patients undergoing non-cardiac surgery (POISE trial): a randomised controlled trial. Lancet Lond Engl 2008;371(9627):1839−47. Available from: https://doi.org/10.1016/S0140-6736(08)60601-7.

[6] Wijeysundera DN, Pearse RM, Shulman MA, et al. Assessment of functional capacity before major non-cardiac surgery: an international, prospective cohort study. Lancet Lond Engl 2018;391(10140):2631−40. Available from: https://doi.org/10.1016/S0140-6736(18)31131-0.

[7] Wolters U, Wolf T, Stützer H, Schröder T. ASA classification and perioperative variables as predictors of postoperative outcome. Br J Anaesth 1996;77(2):217−22. Available from: https://doi.org/10.1093/bja/77.2.217.

[8] Owens WD, Felts JA, Spitznagel EL. ASA physical status classifications: a study of consistency of ratings. Anesthesiology 1978;49(4):239−43.

[9] Bilimoria KY, Liu Y, Paruch JL, et al. Development and evaluation of the universal ACS NSQIP surgical risk calculator: a decision aid and informed consent tool for patients and surgeons. J Am Coll Surg 2013;217(5):833−842.e1−3. Available from: https://doi.org/10.1016/j.jamcollsurg.2013.07.385.

[10] Bertsimas D, Dunn J, Velmahos GC, Kaafarani HMA. Surgical risk is not linear: derivation and validation of a novel, user-friendly, and machine-learning-based Predictive OpTimal Trees in Emergency Surgery Risk (POTTER) Calculator. Ann Surg 2018;268(4):574−83. Available from: https://doi.org/10.1097/SLA.0000000000002956.

[11] Bertsimas D, Dunn J. Optimal classification trees. Mach Learn 2017;106(7):1039−82. Available from: https://doi.org/10.1007/s10994-017-5633-9.

[12] Balch CM, Buzaid AC, Soong SJ, et al. Final version of the American Joint Committee on Cancer staging system for cutaneous melanoma. J Clin Oncol 2001;19(16):3635−48. Available from: https://doi.org/10.1200/JCO.2001.19.16.3635.

[13] Vestergaard ME, Macaskill P, Holt PE, Menzies SW. Dermoscopy compared with naked eye examination for the diagnosis of primary melanoma: a meta-analysis of studies performed in a clinical setting. Br J Dermatol 2008;159(3):669−76. Available from: https://doi.org/10.1111/j.1365-2133.2008.08713.x.

[14] Dreiseitl S, Binder M. Do physicians value decision support? A look at the effect of decision support systems on physician opinion. Artif Intell Med 2005;33(1):25−30. Available from: https://doi.org/10.1016/j.artmed.2004.07.007.

[15] Dreiseitl S, Binder M, Hable K, Kittler H. Computer versus human diagnosis of melanoma: evaluation of the feasibility of an automated diagnostic system in a prospective clinical trial. Melanoma Res 2009;19(3):180−4. Available from: https://doi.org/10.1097/CMR.0b013e32832a1e41.

[16] Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nat Lond 2017;542(7639):115−118G 10.1038/nature21056.

[17] Brierley J, Gospodarowicz MK, Wittekind C, editors. TNM classification of malignant tumours. 8th ed. Chichester, West Sussex, UK ; Hoboken, NJ: John Wiley & Sons, Inc; 2017.

[18] Benson AB, Venook AP, Al-Hawary MM, et al. Rectal Cancer, Version 2.2018, NCCN clinical practice guidelines in oncology. J Natl Compr Cancer Netw 2018;16(7):874−901. Available from: https://doi.org/10.6004/jnccn.2018.0061.

[19] van Hagen P, Hulshof MCCM, van Lanschot JJB, et al. Preoperative chemoradiotherapy for esophageal or junctional cancer. N Engl J Med 2012;366 (22):2074−84. Available from: https://doi.org/10.1056/NEJMoa1112088.

[20] Shapiro J, van Lanschot JJB, Hulshof MCCM, et al. Neoadjuvant chemoradiotherapy plus surgery versus surgery alone for oesophageal or junctional cancer (CROSS): long-term results of a randomised controlled trial. Lancet Oncol 2015;16(9):1090−8. Available from: https://doi.org/10.1016/S1470-2045(15)00040-6.

[21] Habr-Gama A, Perez RO, Nadalin W, et al. Operative versus nonoperative treatment for stage 0 distal rectal cancer following chemoradiation therapy. Ann Surg 2004;240(4):711−18. Available from: https://doi.org/10.1097/01.sla.0000141194.27992.32.

[22] Renehan AG, Malcomson L, Emsley R, et al. Watch-and-wait approach versus surgical resection after chemoradiotherapy for patients with rectal cancer (the OnCoRe project): a propensity-score matched cohort analysis. Lancet Oncol 2016;17(2):174−83. Available from: https://doi.org/10.1016/S1470-2045(15)00467-2.

[23] Probst CP, Becerra AZ, Tejani MA, et al. Watch and wait? - Elevated pretreatment CEA is associated with decreased pathological complete response in rectal cancer. Gastroenterology 2015;148(4, Suppl. 1):S-1105. Available from: https://doi.org/10.1016/S0016-5085(15)33766-5.

[24] Kalady MF, de Campos-Lobato LF, Stocchi L, et al. Predictive factors of pathologic complete response after neoadjuvant chemoradiation for rectal cancer. Trans Meet Am Surg Assoc 2009;127:213−20. Available from: https://doi.org/10.1097/SLA.0b013e3181b91e63.

[25] Shanmugan S, Arrangoiz R, Nitzkorski JR, et al. Predicting pathologic response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer using FDG PET/CT. J Clin Oncol 2011;29(4_Suppl). Available from: https://doi.org/10.1200/jco.2011.29.4_suppl.505 505-505.

[26] Macomber MW, Samareh A, Chaovalitwongse WA, et al. Prediction of pathologic complete response to neoadjuvant chemoradiation in the treatment of esophageal cancer using machine learning. Int J Radiat Oncol 2016;96(2, Suppl.):E699. Available from: https://doi.org/10.1016/j.ijrobp.2016.06.2379.

[27] Weiser TG, Regenbogen SE, Thompson KD, et al. An estimation of the global volume of surgery: a modelling strategy based on available data. Lancet 2008;372 (9633):139−44. Available from: https://doi.org/10.1016/S0140-6736(08)60878-8.

[28] Healey MA, Shackford SR, Osler TM, Rogers FB, Burns E. Complications in surgical patients. Arch Surg 2002;137(5):611−18. Available from: https://doi.org/10.1001/archsurg.137.5.611.

[29] Turrentine FE, Denlinger CE, Simpson VB, et al. Morbidity, mortality, cost, and survival estimates of gastrointestinal anastomotic leaks. J Am Coll Surg 2015;220 (2):195−206. Available from: https://doi.org/10.1016/j.jamcollsurg.2014.11.002.

[30] Mirnezami A, Mirnezami R, Chandrakumaran K, Sasapu K, Sagar P, Finan P. Increased local recurrence and reduced survival from colorectal cancer following anastomotic leak: systematic review and meta-analysis. Ann Surg 2011;253(5):890−9. Available from: https://doi.org/10.1097/SLA.0b013e3182128929.

[31] Ashburn JH, Stocchi L, Kiran RP, Dietz DW, Remzi FH. Consequences of anastomotic leak after restorative proctectomy for cancer: effect on long-term function and quality of life. Dis Colon Rectum 2013;56(3):275−80. Available from: https://doi.org/10.1097/DCR.0b013e318277e8a5.

[32] Ravitch MM, Steichen FM. A stapling instrument for end-to-end inverting anastomoses in the gastrointestinal tract. Ann Surg 1979;189(6):791−7.

[33] Neutzling CB, Lustosa SA, Proenca IM, da Silva EM, Matos D. Stapled versus handsewn methods for colorectal anastomosis surgery. Cochrane Database Syst Rev 2012;(2). Available from: https://doi.org/10.1002/14651858.CD003144.pub2.

[34] Honda M, Kuriyama A, Noma H, Nunobe S, Furukawa TA. Hand-sewn versus mechanical esophagogastric anastomosis after esophagectomy: a systematic review and meta-analysis. Ann Surg 2013;257(2):238−48. Available from: https://doi.org/10.1097/SLA.0b013e31826d4723.

[35] Farrah JP, Lauer CW, Bray MS, et al. Stapled versus hand-sewn anastomoses in emergency general surgery: a retrospective review of outcomes in a unique patient population. J Trauma Acute Care Surg 2013;74(5):1187−92. Available from: https://doi.org/10.1097/TA.0b013e31828cc9c4 discussion 1192-1194.

[36] Bruns BR, Morris DS, Zielinski M, et al. Stapled versus hand-sewn: a prospective emergency surgery study. An American Association for the Surgery of Trauma multi-institutional study. J Trauma Acute Care Surg 2017;82(3):435−43. Available from: https://doi.org/10.1097/TA.0000000000001354.

[37] Leonard S, Wu KL, Kim Y, Krieger A, Kim PCW. Smart Tissue Anastomosis Robot (STAR): a vision-guided robotics system for laparoscopic suturing. IEEE Trans Biomed Eng 2014;61(4):1305−17. Available from: https://doi.org/10.1109/TBME.2014.2302385.

[38] Shademan A, Decker RS, Opfermann JD, Leonard S, Krieger A, Kim PCW. Supervised autonomous robotic soft tissue surgery. Sci Transl Med 2016;8(337). Available from: https://doi.org/10.1126/scitranslmed.aad9398 337ra64.

[39] Jones SB, Jones DB. Surgical aspects and future developments of laparoscopy. Anesthesiol Clin N Am 2001;19(1):107−24.

[40] Jaschinski T, Mosch CG, Eikermann M, Neugebauer EA, Sauerland S. Laparoscopic versus open surgery for suspected appendicitis. Cochrane Database Syst Rev 2018;(11). Available from: https://doi.org/10.1002/14651858.CD001546.pub4.

[41] Lanfranco AR, Castellanos AE, Desai JP, Meyers WC. Robotic surgery. Ann Surg 2004;239(1):14−21. Available from: https://doi.org/10.1097/01.sla.0000103020.19595.7d.

[42] Kim CW, Kim CH, Baik SH. Outcomes of robotic-assisted colorectal surgery compared with laparoscopic and open surgery: a systematic review. J Gastrointest Surg 2014;18(4):816−30. Available from: https://doi.org/10.1007/s11605-014-2469-5.

[43] Wright JD, Ananth CV, Lewin SN, et al. Robotically assisted vs laparoscopic hysterectomy among women with benign gynecologic disease. JAMA. 2013;309(7):689−98. Available from: https://doi.org/10.1001/jama.2013.186.
[44] Jeong IG, Khandwala YS, Kim JH, et al. Association of robotic-assisted vs laparoscopic radical nephrectomy with perioperative outcomes and health care costs, 2003 to 2015. JAMA. 2017;318(16):1561−8. Available from: https://doi.org/10.1001/jama.2017.14586.
[45] Opfermann JD, Leonard S, Decker RS, et al. Semi-autonomous electrosurgery for tumor resection using a multi-degree of freedom electrosurgical tool and visual servoing. In: 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE; 2017. p. 3653−60. Available from: <https://doi.org/10.1109/IROS.2017.8206210>.
[46] Birkmeyer JD, Stukel TA, Siewers AE, Goodney PP, Wennberg DE, Lucas FL. Surgeon volume and operative mortality in the United States. N Engl J Med 2003;349(22):2117−27. Available from: https://doi.org/10.1056/NEJMsa035205.
[47] Birkmeyer JD, Finks JF, O'Reilly A, et al. Surgical skill and complication rates after bariatric surgery. N Engl J Med 2013;369(15):1434−42. Available from: https://doi.org/10.1056/NEJMsa1300625.
[48] Tseng JF, Pisters PWT, Lee JE, et al. The learning curve in pancreatic surgery. Surgery. 2007;141(4):456−63. Available from: https://doi.org/10.1016/j.surg.2006.09.013.
[49] Tekkis PP, Senagore AJ, Delaney CP, Fazio VW. Evaluation of the learning curve in laparoscopic colorectal surgery. Ann Surg 2005;242(1):83−91. Available from: https://doi.org/10.1097/01.sla.0000167857.14690.68.
[50] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521(7553):436−44. Available from: https://doi.org/10.1038/nature14539.
[51] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Adv Neural Inform Process Syst 2012;25:1097−105.
[52] Bouget D, Benenson R, Omran M, Riffaud L, Schiele B, Jannin P. Detecting surgical tools by modelling local appearance and global shape. IEEE Trans Med Imaging 2015;34(12):2603−17. Available from: https://doi.org/10.1109/TMI.2015.2450831.
[53] Padoy N, Blum T, Feussner H, Marie-Odile B, Navab N. On-line recognition of surgical activity for monitoring in the operating room. In: Proceedings of the National Conference on Artificial Intelligence, vol. 3; 2008. p. 1718−24.
[54] Wild E, Teber D, Schmid D, et al. Robust augmented reality guidance with fluorescent markers in laparoscopic surgery. Int J Comput Assist Radiol Surg 2016;11(6):899−907. Available from: https://doi.org/10.1007/s11548-016-1385-4.
[55] Volkov M, Hashimoto DA, Rosman G, Meireles OR, Rus D. Machine learning and coresets for automated real-time video segmentation of laparoscopic and robot-assisted surgery. In: 2017 IEEE International Conference on Robotics and Automation (ICRA). IEEE. 2017. p. 754−9. Available from: <https://doi.org/10.1109/ICRA.2017.7989093>.
[56] Natarajan P, Frenzel JC, Smaltz DH. Demystifying big data and machine learning for healthcare. Boca Raton, FL: CRC Press Taylor & Francis Group; 2017.
[57] Bonrath EM, Gordon LE, Grantcharov TP. Characterising 'near miss' events in complex laparoscopic surgery through video analysis. BMJ Qual Saf 2015;24(8):516−21. Available from: https://doi.org/10.1136/bmjqs-2014-003816.
[58] Ellison EC, Pawlik TM, Way DP, Satiani B, Williams TE. Ten-year reassessment of the shortage of general surgeons: increases in graduation numbers of general surgery residents are insufficient to meet the future demand for general surgeons. Surgery 2018;164(4):726−32. Available from: https://doi.org/10.1016/j.surg.2018.04.042.

[59] Schein M, Rogers PN, editors. Schein's common sense emergency abdominal surgery: 21 tables. 2nd ed. Berlin: Springer; 2005.

[60] Silber JH, Williams SV, Krakauer H, Schwartz JS. Hospital and patient characteristics associated with death after surgery. A study of adverse occurrence and failure to rescue. Med Care 1992;30(7):615–29.

[61] Ghaferi AA, Birkmeyer JD, Dimick JB. Complications, failure to rescue, and mortality with major inpatient surgery in medicare patients. Ann Surg 2009;250(6):1029–34. Available from: https://doi.org/10.1097/SLA.0b013e3181bef697.

[62] Smith GB, Prytherch DR, Meredith P, Schmidt PE, Featherstone PI. The ability of the National Early Warning Score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. Resuscitation 2013;84(4):465–70. Available from: https://doi.org/10.1016/j.resuscitation.2012.12.016.

[63] Cambria E, White B. Jumping NLP curves: a review of natural language processing research [review article]. IEEE Comput Intell Mag 2014;9(2):48–57. Available from: https://doi.org/10.1109/MCI.2014.2307227.

[64] Murff HJ, FitzHenry F, Matheny ME, et al. Automated identification of postoperative complications within an electronic medical record using natural language processing. JAMA 2011;306(8):848–55. Available from: https://doi.org/10.1001/jama.2011.1204.

[65] FitzHenry F, Murff HJ, Matheny ME, et al. Exploring the frontier of electronic health record surveillance: the case of post-operative complications. Med Care 2013;51(6):509–16. Available from: https://doi.org/10.1097/MLR.0b013e31828d1210.

[66] Rajkomar A, Oren E, Chen K, et al. Scalable and accurate deep learning with electronic health records. NPJ Digit Med 2018;1(1):18. Available from: https://doi.org/10.1038/s41746-018-0029-1.

[67] Wauben LSGL, van Grevenstein WMU, Goossens RHM, van der Meulen FH, Lange JF. Operative notes do not reflect reality in laparoscopic cholecystectomy. BJS 2011;98(10):1431–6. Available from: https://doi.org/10.1002/bjs.7576.

[68] van de Graaf FW, Lange MM, Spakman JI, et al. Comparison of systematic video documentation with narrative operative report in colorectal cancer surgery. JAMA Surg 2019;154:381–9. Available from: https://doi.org/10.1001/jamasurg.2018.5246.

[69] Hung AJ, Chen J, Jarc A, Hatcher D, Djaladat H, Gill IS. Development and validation of objective performance metrics for robot-assisted radical prostatectomy: a pilot study. J Urol 2018;199(1):296–304. Available from: https://doi.org/10.1016/j.juro.2017.07.081.

[70] Hung AJ, Chen J, Gill IS. Automated performance metrics and machine learning algorithms to measure surgeon performance and anticipate clinical outcomes in robotic surgery. JAMA Surg 2018;153(8):770. Available from: https://doi.org/10.1001/jamasurg.2018.1512.

[71] Sicklick JK, Camp MS, Lillemoe KD, et al. Surgical management of bile duct injuries sustained during laparoscopic cholecystectomy. Ann Surg 2005;241(5):786–95. Available from: https://doi.org/10.1097/01.sla.0000161029.27410.71.

[72] Way LW, Stewart L, Gantert W, et al. Causes and prevention of laparoscopic bile duct injuries: analysis of 252 cases from a human factors and cognitive psychology perspective. Ann Surg 2003;237(4):460–9. Available from: https://doi.org/10.1097/01.SLA.0000060680.92690.E9.

[73] Weinberger SE. Can maintenance of certification pass the test? JAMA 2019;321:641–2. Available from: https://doi.org/10.1001/jama.2019.0084.

[74] American Board of Surgery. Continuous certification, <http://www.absurgery.org/default.jsp?exam-moc> [accessed 17.02.19].

[75] Greenberg CC, Ghousseini HN, Pavuluri Quamme SR, Beasley HL, Wiegmann DA. Surgical coaching for individual performance improvement. Ann Surg 2015;261 (1):32−4. Available from: https://doi.org/10.1097/SLA.0000000000000776.

[76] Iglehart JK, Baron RB. Ensuring physicians' competence—is maintenance of certification the answer? N. Engl. J. Med. 2012;367(26):2543−9. Available from: https://doi.org/10.1056/NEJMhpr1211043.

Further reading

Iglehart JK, Baron RB. Ensuring physicians' competence—is maintenance of certification the answer? N Engl J Med 2012;367(26):2543−9. Available from: https://doi.org/10.1056/NEJMhpr1211043.