International Journal of Computer Assisted Radiology and Surgery manuscript No.
(will be inserted by the editor)

Surgical Robotics Beyond Enhanced Dexterity Instrumentation

A survey of the machine learning techniques and their role in intelligent and autonomous surgical actions

Yohannes Kassahun · Bingbin Yu · Abraham Temesgen Tibebu · Danail Stoyanov · Stamatia Giannarou · Jan Hendrik Metzen · Emmanuel Vander Poorten

Received: date / Accepted: date

Abstract Advances in technology and computing play an increasingly important role in the evolution of modern surgical techniques and paradigms. In this article we review the current role of machine learning (ML) techniques in the context of surgery with a focus on surgical robotics and we provide a perspective on the future possibilities for enhancing the effectiveness of procedures by integrating ML in the operating room.

We focus our review on ML techniques directly applied to surgery, surgical robotics, and surgical training and assessment. The widespread use of ML methods in diagnosis and medical image computing is beyond the scope of the review. We performed searches on PubMed and IEEE Explore using combinations of the keywords: ML, surgery, robotics, surgical and medical robotics, skill learning, skill analysis, learning to perceive.

Studies making use of ML methods in the context of surgery are increasingly reported. In particular, there is focus on using ML for developing the tools to understand and model surgical skill, competence, and surgical workflow. Some initial works have begun integrating the increased understanding of the surgical process into the control of recent surgical robots and devices.

ML is an expanding field. It is widely used to efficiently process vast amounts of data and to interpret it for real-time decision making. Already widely used in imaging and diagnosis, we believe ML methods will also play an important role in surgery and interventional treatments. In particular ML could become a game changer into the conception of *cognitive* surgical robots. ML would allow extracting surgical skill, learned through demonstration by human experts, and transfer this to robotic skills as such offering intelligent surgical assistance. Such systems would significantly surpass the state of the art in surgical robotic technology, which, at present, merely plays the role of an instrument that enhances the surgeon's dexterity.

Keywords Surgical robotics · Skill learning · Skill analysis · Learning to perceive

Y. Kassahun · B. Yu · A. T. Tibebu · J. H. Metzen
Faculty 3 - Mathematics and Computer Science
University of Bremen,
Robert-Hooke-Str. 5, D-28359 Bremen, Germany
E-mail: bingbin.yu@uni-bremen.de

D. Stoyanov
Centre for Medical Image Computing and Department of Computer Science
University College London, UK

S. Giannarou
Hamlyn Centre for Robotic Surgery
Imperial College London, London, UK

E. Vander Poorten
Department of Mechanical Engineering
University of Leuven
Celestijnenlaan 300B, B-3001 Heverlee, Belgium

1 Motivation for Machine Learning in Surgical Robotics

To justify the cost of robotic surgery, technology providers and its users are searching for objective and measurable proof that robotic surgery possesses clinical advantages over existing manual techniques [1,2,3]. While such evidence remains sparse [4,5] or even discouraging at present [6,7,8], future systems that possess a certain degree of intelligence might show the clinical advantage people are looking for. Endowed with cognitive capabilities surgical robots could take over the simpler parts of a procedure and allow surgeons to focus on the more crucial and complex parts of the procedure. This could translate in increased reliability, performance or efficiency of robot-assisted compared to more traditional interventions. Indeed, some of the selling arguments of machine learning (ML) techniques are exactly that they allow *smoother trajectories*, more *accurate* or *faster* execution of repetitive and time-consuming tasks [9]. Through synthesis of technical or cognitive knowledge coming from a broad group of expert surgeons, the robots could play the role of a 'computerized assistant' that provides technical assistance during routine or even unusual interventions [2]. This should then translate towards increased reliability and improved performance of robot-assisted interventions compared to traditional interventions.

The aging population, a reduced workforce and an increasing workload on expert surgeons are arguments to automate certain surgical interventions and increase operating room efficiency. Especially in the early days developers aimed for fully automated procedure execution. For instance, systems like the Unimation Puma 200 from Kwoh *et al.* [10], the ROBODOC [11], MINERVA [12] or Cyberknife [13] operated mostly autonomously. All these systems work in an environment that shows a relatively large invariance with respect to the actual robotic action. However, such assumption is quite restrictive. It severely limits the range of procedures that can be considered or also the performance that can be achieved. Indeed the majority of surgical interventions do impact the environment, especially when interactions with soft, deformable structures are involved. This is also one of the challenges of and motivations behind the development of visual servoing techniques, namely to account for deformation and physiological motion. Typically visual servo techniques only focus on a specific detailed part of the surgical procedure. Servo techniques aiming for accurate control of forces or interaction become difficult if based on visual information only. Substantial efforts have been done to *explicitly* model in detail interactions or tissue deformation. Excellent works have appeared that model trajectories and interactions during surgical tasks, e.g., for knot tying [14], suturing [15], stitching [16], tissue retraction [17,18], puncturing [19], cochleostomy [20], anesthesia [21] or even diagnosis [22]. It should be noted that depending on particular choices of models and parameters the predictive power or applicability of a model can be rather limited. However, the derivation of valid models and identification of parameters can be a time-consuming and tedious task. Given the large variability between people, organs and tissues *explicit* modeling approaches have practical limitations.

In contrast, ML approaches that *implicitly* learn models directly from the real sensory data are appealing for the following reasons:

- they are generally applicable;
- they do not require deep understanding of the complex underlying physics;
- they are grounded in reality as they are trained on real data.

These properties explain why ML approaches recently receive more attention within the research community. Even for critical applications such as in surgery they are increasingly being considered.

2 Introduction to Machine Learning

ML is a multidisciplinary field that provides computers with the ability to learn without being explicitly programmed to perform specific tasks [23,24]. While ML techniques have been used extensively in a wide spectrum of robotic applications, it is only recently that ML methods have been considered for Surgical Robotics (SR). Figure 1 shows a schematic of an ML-enabled intelligent surgical robotic system for the case of catheter based surgery. The continuous interaction that takes place between the robot, the surgeon (domain expert) and the environment (human body) is an important feature of this scheme. The robot perceives the state of the environment through its sensors and executes an action based on this information. The environment gives the robot the next state based on the action executed by the robot. An action taken by the robot has an associated cost with it. The purpose of the robot is to learn a mapping function from perception $\boldsymbol{z}$ to action $\boldsymbol{a}$ that minimizes the total cost incurred. In SR, the mapping function from perception to action can be considered as the *surgical skill*. Such skill could be decomposed into two large parts. The first part is concerned with the state estimation, which maps the perception $\boldsymbol{z}$ to the estimate of the state of the environment $\hat{\boldsymbol{s}}$. The second part maps the estimated state $\hat{\boldsymbol{s}}$ to the action $\boldsymbol{a}$ that is to be taken. The cost quantifies the evaluation of the skill demonstrated by the robot or by the surgeon. It depends on the state $\boldsymbol{s}$ and action $\boldsymbol{a}$ taken by the robot.

The process of evaluating the learned skills is referred to as skill analysis. A detailed review of work in SR on skill learning (Section 3.1 and Section 4) and skill analysis (Section 3.2) is provided. The robot can learn a surgical skills in multiple ways. First, it learns from its own interaction with the environment, by evaluating the appropriateness of its own actions in order to achieve some particular target states. The robot can

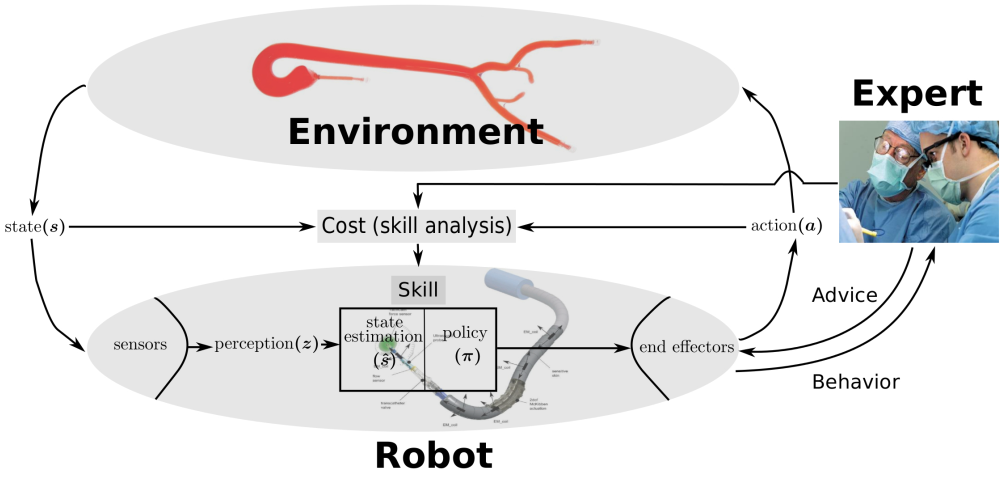

Fig. 1 Overview of a learning system in surgical robotics. The learning system is augmented with a process that allows a surgeon to watch the robot and provide advice based on the behavior of the robot. In the figure a catheter surgical robot and the aorta are depicted as examples of a surgical robot and environment, respectively.

also learn from human demonstration by observing experiments conducted by surgeons. From the demonstration, both the surgical skill and the cost function (Section 3.2.2) used to asses the quality of the skills can be learned. The cost function can also be defined explicitly by the domain expert/surgeon (which is described in Section 3.2.1). Through observation of robot actions, the surgeon can intervene and guide. The information provided by the surgeon (expert) is essential to further speed up and guide the learning process.

The surgeon's expert knowledge is also of use during the perception of the tasks. The surgeon can provide examples to the learning system for example to teach the robot how to detect natural landmarks and anatomies. This information can help a robot to choose relevant (optimal) actions when approaching difficult or risk prone areas. Some applications of ML in SR are introduced in Section 4. We now provide a brief introduction to three important areas of ML supervised learning, reinforcement learning, and unsupervised learning.

Supervised Learning

In *Supervised Learning* [24], training data is provided externally and consists of a set of known input vectors along with a set of known corresponding target vectors which might be discrete (classification) or continuous (regression). Supervised learning seeks to build a predictor model that predicts reasonable target vectors for new input vectors. The choice of the predictor model is typically up to the designer and often requires considerable ML expertise. Learning consists of finding optimal parameters value for the chosen model. Supervised learning is applied for instance in state estimation (see later sections).

Reinforcement Learning (RL)

RL deals with learning a policy, i.e., a mapping from states to actions. The most popular approaches in RL are value-function based approaches such as Q-learning [25]. In these approaches, the agent learns the optimal value function of a state action pair. Once the optimal value function is learned, it is possible to generate the optimal policy (skill) for a given task from the value function. Intuitively, a value of a state action pair shows how good it is for the robot (agent) to execute an action in a given state. The training data for RL is generated through direct interaction with the environment and autonomous generation of sequences of experience tuples. An experience tuple is an entry in training data at a particular time which consists of the current state, the current action, the next state and the reward received after executing the action. An important issue in RL is the trade-off between exploration and exploitation: in exploration, the agent tries actions which may be suboptimal according to its current knowledge but have the potential of resulting in better outcomes than expected. In exploitation, the agent always chooses the action which it considers to be optimal at the risk of missing other actions which turn out to be better in reality.

Unsupervised Learning

In *unsupervised learning*, the training data consists of a set of input vectors without a corresponding set of target vectors. The goal of such unsupervised learning is to discover structure and correlations in the data. Approaches in unsupervised learning include [24]:

- clustering for discovering groups of similar examples in the data;
- density estimation for determining the distribution of the data;
- dimensionality reduction for data compression, visualization, or accelerating subsequent learning.

3 ML-empowered Instrumentation for Assisted Surgery

3.1 Surgical Skill Learning from Expert Knowledge

Prior knowledge is of key importance in ML. For obvious reasons in the context of surgery expert knowledge is typically supplied by experienced surgeons. *Implicit imitation learning* is a form of supervised learning, which is usually concerned with accelerating RL through the observation of an expert mentor [26]. The agent observes the state transitions of the experts' actions and uses the information extracted from these observations to update its own states and actions. The mentor (surgeon) and the agent may have identical or different action capabilities, or identical or different reward structures. Several methods that have been developed for modeling human movement (see e.g. [27]) could be used to learn the state and actions of the expert. Human skill has been modeled from sets of recorded data using hidden Markov models [28,29], neural networks [30,31] and fuzzy nets [32,33].

The work in implicit imitation learning can be categorized into three groups. A first group tries to learn the mentor's policy; a second group learns the reward function of the mentor's behavior and optimizes its own behavior using the learned reward function. The third group employs a Bayesian framework for combining prior (explicit) knowledge and implicit imitation learning. In a series of works, trajectories recorded from human subjects are used to generate an initial policy.

Works on inverse RL [34,35] assume that the mentor has the same reward function as the observer and chooses from the same set of actions. The idea is then to infer the reward function of the mentor so as to produce the observed behavior. In other words, inverse RL accomplishes the task of learning both the reward function and the policy (apprenticeship learning).

Bayesian formulations of imitation learning are used to elegantly combine prior knowledge, model observations from the imitator's own experience and model observations derived from other agents. Works in this area developed algorithms for imitation learning that can handle knowledge transfer between agents with different reward structures, learning from multiple mentors, and selecting relevant portions of examples [26,36].

3.2 Skill Analysis in Robotic Surgery

Skill analysis is common to many disciplines and is known under different terms. In optimal control, a skill is the control policy to be designed. Such a control policy is evaluated using a cost function (reward function) [37].

In surgery, skill evaluation is performed in the context of training and competence evaluation. Training and competence evaluation are now widely recognized as critical for acquiring new clinical techniques or for operating complex devices that are used for patient monitoring or treatment. It is generally accepted that the skill level of clinicians is variable and can be enhanced with teaching, training and naturally through experience. Clinical outcomes have in the past been linked to clinical skill [38] and, as a result, effective surgical training and evaluation could have a significant impact on healthcare. However, despite advances in simulation, phantom models and task-based procedural trainers, typical training aimed at enhancing manual dexterity and instrument handling, still involves expert monitoring. This is time consuming and hence costly to the healthcare system. In addition it is also, to a certain extent, subjective in nature. It can also be inefficient when lacking real-time feedback of the task performance. It goes without saying that also for skill learning in SR, skill analysis plays a crucial role, as it can be used as a means to asses or evaluate the quality of the skills that were learned.

Different objective assessment techniques have been reported in the literature. Metrics can be based on task completion time, instrument speed, distances and more complex measures derived from positional information [39]. Such metrics can be derived for example from the information given by robotic encoders or from virtual reality simulation. Simulation environments are seen as a promising means to enhance our understanding and means of evaluating skills as they provide full geometric knowledge of the procedure [39,40]. This facilitates building links between instrument motion and motion of the surrounding tissue. Note that complex considerations are needed for acquiring such links. A distinction is made here into explicit and implicit types of SR skill analysis. Where possible illustrations are given for skill analysis in endovascular procedures.

3.2.1 Explicit Skill Analysis

In *explicit skill analysis* the form of the cost function is defined by the domain expert (surgeon).

Checklists and Rating Scales

Checklists and rating scales is a validated mode of assessment where experts appraise seven aspects of operative performance using a five point Likert scale [41]. The main challenge for rating scales is the vast amount of expert time required to analyse and score videos of trainee operations and unfortunately reducing the time by showing only selected parts of the procedure has not been shown to be as reliable [42]. As a result, the current measure of whether someone is competent to perform endovascular procedures is largely based on a classical view, counting the number of procedures performed and the time spent in training the respective skill.

Structured Assessment

The aim of structured assessment approaches is to attempt to standardize evaluation through rated checklists on a phantom bench-top model. Several successful examples of this assessment approach have been reported [43,44,45]. Objective Structured Assessment of Technical Skills (OSATS) is one of the first methods designed for objective medical skill assessment, which aims at quantifying medical skill evaluation without relying on expert evaluators. It consists of a global rating scale and a procedure-specific checklist. OSATS is one of the few methods that has been implemented in clinical practice [46,47].

Nevertheless, even with structured methods, objective assessment of surgical skills is currently underdeveloped. Existing structured grading practices suffer from the need for well-structured task stations and the need for clinical experts to administer the assessment. Added cost and time, and also subjectivity, pose additional problems to the sustainability of structured approaches. Automated and analytical approaches are thus required and need to be researched and validated further.

Outcome-based Analysis

The outcome-based metrics, such as the number of complications, morbidity and mortality rates assuming that a strong relationship between skill level and patient outcomes exists, are measured [48]. This approach, however, suffers from limitations because patient outcomes are strongly dependent on patient characteristics, diagnostic information, theater staff and condition, and the difficulty of the procedure. For example, a less experienced surgeon could be selective and take low-risk cases than a more experienced surgeon and yet have better outcome-based skill measures. Therefore outcome-based metrics, while important and possibly a meaningful statistic to monitor, are not comprehensively meaningful and do not lend themselves to training and assessment during medical accreditation courses.

Motion Analysis

One of the most promising methodologies for task and manual dexterity evaluation is motion analysis. In motion analysis the surgeon's hand or tool motions are recorded and analyzed by different instruments such as Imperial College Surgical Assessment Device (ICSAD), the Advanced Dundee Psychomotor Tester (ADEPT), the ProMIS Augmented Reality Simulator, the Hiroshima University Endoscopic Surgical Assessment Device (HUESAD) and the TrEndo Tracking System [49,50,51,52,53,41]. The technique can provide a good assessment of dexterity and technical skill level, but it has not been used or investigated thus far for endovascular procedures [42]. Nevertheless this methodology has the most abundant literature references [42], especially with recent technological developments in robotics and particularly the da Vinci API [54,55]. Multiple studies have shown that skill metrics can be derived using statistical analysis (e.g. Hidden Markov modelling [56,57,28]) of instrument motion from this data [58].

Time Action Analysis

Time action analysis is a technique where the surgical procedure is broken down into several steps and the time to complete each one is measured usually by an expert watching a video recording of the training exercise. The limitation of this approach is that it is time consuming and does not report any measure of how well the particular surgical action was performed [59]. While not particularly informative about what the failings or technical limitations of a particular task are, time action analysis does offer a simple means of evaluation and can often be linked or correlated to clinical competence. It is logical that experienced and highly skilled surgeons would be able to perform tasks more quickly than novices. The problem lies in identifying when a performed task is done badly or with considerable potential risk to the patient or benchtop environment.

Virtual Reality

An emerging training modality is Virtual Reality (VR) and this potentially offers a vast amount of information for assessment and analysis of different surgical techniques [60]. The validity of VR in endovascular procedures is still under evaluation, though it seems logical that simulators will have a role in surgical training. Endovascular surgery simulators are available on the market, though they have not been integrated into any curriculum or formal accreditation course. With the current evolution of endovascular surgery training, however, and the merger between the disciplines or vascular surgery and interventional radiology, it is likely that VR will have a role in modern training.

Error Analysis

Error analysis, where the number of errors made during certain part of the procedure is scored, is an alternative and potentially more thorough skill evaluation technique. For endovascular surgery errors can be defined by for example the number, location or intensity of the contact with the vascular wall. Such parameters could be recorded by some simulation systems [49]. However, to the knowledge of the authors no in vivo or phantom study, taxonomy of errors or scores exists at this point including these parameters. It seems that this metric ought to be investigated since the wall vessels provide a geometric enclosure for the tool and therefore errors can be evaluated intuitively [61].

3.2.2 Implicit Skill Analysis

Implicit skill analysis uses a metric which is learned by an ML approach from a surgeon or group of surgeons. The learned metric can then be used to evaluate other surgeons (trainees) relative to the skills of a surgeon (surgeons) from which the metric is learned.

The Viterbi algorithm, a dynamic programming algorithm for computing most probable sequence of hidden states, and Hidden Markov models (HMMs) are often used to develop new models for prediction and analysis of sensory data recorded during task execution by telemanipulation [28]. HMMs are used to learn task levels, target levels, force and torque sensor signals. The Viterbi algorithm is used to analyse the force signals based on the model developed. By doing so it can perform excellent segmentation of a task into subtasks.

Reiley et al. [62] applied Vector Quantization (VQ), an unsupervised ML approach, and HMM to evaluate the skill from continuous velocity data of the da Vinci system. In Reiley's paper, HMMs based on skill are developed for three surgical levels such as novice, intermediate and expert. In the paper it is shown that HMMs are important methods to classify skill of unknown trial based on maximum likelihoods from trained skill models of novice, intermediate and expert surgeons.

VQ is employed in a laparoscopic porcine task of bowel suturing of 30 surgical subjects [63] for developing an objective evaluation of surgical skills. Each vector in the training data stream contains forces, torques, and velocities with respect to a coordinate system located at the port of each tool. It has been shown that the method can distinguish surgical skill level even if no prior human knowledge of the task has been used for training the algorithm.

An automatic method of parsing raw motion data from a surgical task from a labeled sequence of surgical gestures that would allow for the development of automatic evaluation of surgical skills is developed by Lin et al. [64]. The method has feature processing and classification steps where a Bayes classifier is used for the classification step. Results show that the method is able to correctly identify the different gestures for the case of a suturing task using the da Vinci surgical robot against benchtop models. It has been shown that based upon analysis of instrument motions it is possible to distinguish an expert surgeon from a surgeon having limited da Vinci experience. The method is further extended to handle data from live surgeries and for more number of users [65].

Allen et al. [66] compared three different methods (summed-ratios, z-score normalization and support vector machine [SVM]) in prediction of surgical competency within 696 trials performed by 30 participants (four experts and 26 novices). The SVM based analysis is proven to be more efficient to assess surgical skill based on motion data in these three methods.

Tao et al. [67] compared Mixture of Factor Analysis HMM (MFA-HMM) to KSVD-HMM for three different surgical tasks such as suturing, needle passing and knot tying. The dataset for these tasks consist of 39, 26 and 36 trials performed by 8 surgeons with expertise levels expert, intermediate and novice. In the paper it is shown that the proposed methods work with stable performance for different sparsity levels.

Rafii-Tari et al. [68] proposed a learning-from-demonstration framework for robot-assisted cardiac catheterization. The motion model of the catheterization procedure is trained by manipulations from experts and intermediate-level operators. The motion model is represented by a Gaussian Mixture Model (GMM) and clustered by the k-means algorithm. Then Gaussian Mixture Regression (GMR) is used to smooth the motion trajectory. For validation the same procedure is performed by novices assisted by a robotic catheter driver. A significant difference between skills of novices and with skills from experts and intermediate-skilled operators was observed.

Currently, a significant limitation of assessment methods based on analysis of surgical tool motion is that they do not consider the environment. Only some recent studies attempt to provide some context to the instrument motion data using the da Vinci simulation environment [39, 69]. The interaction with the environment can also be investigated within phantoms. These parameters would be especially powerful to evaluate the role/effect of guidance or of the use of novel instrumentation or control approaches.

3.3 Surgical Workflow Analysis and Episode Segmentation

A surgical procedure is in essence a concatenation of surgical acts, which when pertaining to the same specific surgical (sub)goal can be grouped into surgical (sub)tasks. *Workflow Analysis* can be conducted to identify the different surgical (sub)tasks that belong to a surgical intervention, the order in which (sub)tasks can follow each other and possible termination conditions that mark transients between distinct (sub)tasks. The analysis of the surgical workflow is essential to assist surgical navigation and enable the design of cognitive surgical systems that can adapt and operate in highly dynamic environments such as the cardiovascular system. In addition, the analysis of individual surgical tasks can provide quantitative evaluation of surgical skills during different procedural tasks.

Thus far, the analysis of surgical workflow has been extensively studied for minimally invasive procedures. Approaches proposed in literature can be classified into methods for segmentation of high-level surgical tasks (surgical phases) and methods for the recognition of low-level tasks (surgical gestures) and into off-line and on-line approaches.

In [70, 71], laparoscopic cholecystectomy procedures are segmented into 14 different phases based on the presence of instruments in the surgical scene. AdaBoost is used in [70] to analyse the use of each surgical instrument in each phase of the surgery and weight them according to their discriminative power. For phase recognition, an average reference surgery is generated based on Dynamic Time Warping (DTW) and used to segment newly observed procedures. The standard DTW approach [71] segments different surgical phases with 92% accuracy and 5 sec tolerance while the adaptive DTW [70] improves the segmentation reliability to an error less than 5%.

A significant number of approaches to surgical gesture recognition focused on modeling kinematic data with HMMs using a variety of methods for modeling the observations such as vector-quantization of the observations into discrete symbols [72], Gaussian HMMs combined with Linear Discriminative Analysis (LDA) [73], Factor Analyzed HMMs (FA-HMMs) and Switched Linear Dynamical Systems (SLDSs) [74]. Sparse HMM have been used in [67] where the observations are modeled as sparse linear combinations of basic surgical motions. In [75], tool-tissue interactions of a knot-tying task in MIS have been modeled using Markov Models (MM) based on the kinematics (position and orientation) and the dynamics (force and torque) of the surgical tools.

From the above analysis, it becomes clear that an ideal surgical workflow analysis system should be based on a general methodology that can be used to describe almost any surgery. Depending on the application, on-line or off-line surgical phase and gesture recognition is required. One limitation of DTW-based approaches is that surgical task segmentation can only be performed after the whole video has been recorded. The majority of on-line segmentation approaches are based on HMMs, which are not able to capture the variability of complex gestures and may fail to recognise complex surgical actions.

4 Towards Autonomous Robotic Surgery

In section 3, different ML techniques which used to learn surgical skills from surgeons, to assess surgical skills and to analyze surgical workflow have been discussed. In this section the autonomous robotic surgery using ML is discussed.

When profound and up-to-date understanding of a surgical task is available and when a robotic system has demonstrated repeatedly its ability to correctly display an acceptable level of performance in executing the necessary surgical acts under similar surgical conditions, one might consider to let the robot perform these surgical gestures in an autonomous fashion. Setting up a system suitable for autonomous robotic surgery (ARS) is not to be taken lightly as many aspects need to be considered. The different technologies introduced in preceding sections could serve here as building blocks. These blocks could for example be plugged into the framework proposed by Muradore *et al.* [19] and briefly introduced in subsection 4.1. The framework by Muradore aims at streamlining the developments towards an ARS system. It further promotes modularity and exchangeability so that developments could potentially be reused in other applications. Many developments introducing automatic features and capabilities were already proposed in the past. A non-exhaustive overview of such developments is given in subsections 4.2, 4.3 and 4.4. From an analysis of these systems an overview of the main challenges and directions of future work in ARS is given in subsection 4.5.

4.1 A Verification Framework for Autonomous Robotic Surgery

Muradore and colleagues introduced recently a structured and formal method to approach autonomous surgery [19]. The basis of the proposed approach lies in the decomposition of the surgical act into smaller surgical subtasks with specific goals and well-described conditions for transitions between them. Through a detailed *workflow analysis* the procedure is broken down into several *episodes*. Automatic code-generation tools are employed to translate these into corresponding software blocks. To each episode a certain desired behavior or *surgical skill* is assigned. Equipped with knowledge about the surgical state, the state of the environment and of the robotic instruments, controllers are deployed to steer the system to closely display the desired behavior. A broad set of safety and error handling procedures are to be prepared, so that when a critical event is detected, appropriate procedures are triggered and an adequate response is given by the robotic system. Muradore *et al.* follow a model-based approach. This means that the entire procedure and its different components are explicitly modeled. An ML-based variant of this approach could also be envisaged. In such case models of the procedure, environment, instrument, etc. could be constructed and learned directly from the data. Table 1 summarizes the different aspects that could be covered in such case.

Table 1 Aspects of ARS where ML could play an enabling role

workflow analysis episode segmentation	surgical procedure broken into logical subtasks or episodes
environment modeling	reconstruction of environment, recognition of anatomical features and landmarks, rigid and flexible registration, mechanical and physiological modeling.
robot control	low-level modeling and robot-control
localization	localization of instrument/robot w.r.t. environment
skill analysis	analysis of surgical skill, derivation of performance metrics, cost functions for optimization
planning	high-level trajectory and interaction planning
critical event detection	detection of adverse events

4.2 Evolution in Autonomous Robotic Surgery

For an excellent review on works on autonomous and semi-autonomous robotic surgery we gladly refer to the work by Moustris *et al.* [76]. In this section we are mainly interested in high-lighting the evolution that is taking place in ARS and the role that ML plays in this evolution. Figure 2 gives a fair idea of the evolution in ARS (despite being based on a non-exhaustive set of ARS papers). It can be seen that the number of papers dealing with ARS is steadily growing over time. Furthermore, when looking at the share of ARS papers that employ ML techniques, one can appreciate that also this share rises steadily.

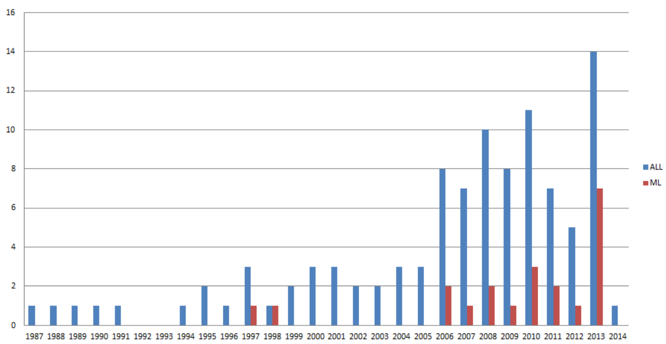

Fig. 2 Evolution of a reference number of publications regarding ARS over the years. The number and share of papers employing ML can be seen to rise over time.

4.3 Examples of ML used in Surgical Robotics Research

Intelligent Autonomous Endoscopic Guidance System for MIS

Modern laparoscopic surgery or MIS procedures make use of three or four access ports through which a plurality of instruments is inserted in the body. Typically this includes an endoscope that is used to visualize the patient's organs alongside instruments for grasping, cutting, ablating and so on.

The research has been conducted to automatically steer laparoscopes in such configuration by a.o. Casals *et al.* [77]. In order for such tracking system to behave in an automatic fashion, steering must be extremely reliable. This implies that such system should be capable of tracking all aspects of the procedure and this in a robust fashion. In contrast to the short-term prediction steps associated with typical control schemes that focus on the compensation of physiological motion such as heartbeat and breathing [78,79,80,81], Weede *et al.* [82] advocate the development of *long-term* prediction schemes that anticipate upon what the surgeon is going to do the next couple of minutes, so that the endoscope can always be moved to an appropriate position.

Autonomous Knot-Tying

Surgeons frequently have to tie knots to connect tissues or close openings. In MIS, where access and maneuverability is limited and haptic feel is typically poor if not absent, knot-tying is a tedious job. Whereas in open surgery a knot can be tightened within a few seconds, in MIS this can take up to three minutes per knot [83]. Research was also conducted to apply ML to solve suturing and knot-tying. For example Mayer *et al.* published a series of works to this end [83,84,85,86].

Knot-Tying with Neural Network

For instance, Mayer *et al.* [83] use recurrent neural networks (RNN) to tie knots autonomously. The system even speeds up the knot-tying, reducing the overall time of the surgical intervention. A sequence is presented to the network by a surgeon after which the sequence is learned. A neural network with a long-term storage [87] is used to learn this task. Only after a limited number of sequences, the network is capable of performing the basic steps.

Knot-Tying via Trajectory Transferring

Schulman et al. [88] developed recently a trajectory transfer method, which can tie knots in ropes by training the robots by human demonstration. During the procedure, a nonrigid transformation from training state to the testing state is registered. Based on the transformation and the training trajectory, the new trajectory for the testing task can be calculated. Five different types of knots were automated.

Algorithm for Superhuman Performance of Surgical Tasks

Van den Berg *et al.* [9] developed an algorithm that learns a task from multiple human demonstrations, and learns to execute the tasks with superhuman performance. The important parameters maximized during the learning process are smoothness and speed with which the tasks are performed. The approach is implemented on the Berkeley Surgical Robot and applied to two tasks: first drawing figures on a magnetic wireboards and second knot-tying.

Skill Transfer from a Surgeon Teleoperator to a Flexible Robot

Recently, Calinon *et al.* [89] developed a method based on inverse RL [34,35] for transferring skills from a surgeon teleoperator to a flexible robot. The flexible robot is a bio-inspired robot that mimics the way octopuses elegantly move through small openings and difficult environments. The method can handle the case where robots used for transfering skills have different morphological structures.

GMM/GMR based Learning from Demonstration

In recent papers [90,91,92], GMM algorithms are used to learn from demonstration by representing datasets stochastically using joint probability densities. Kassahun *et al.* [90] developed a method to learn the model of the interaction between catheter and aorta. GMM is used to model the joint probability densities of the multiple variables which are used to represent the catheter shape, touching states, entrance and tip points of the catheter. It has been shown that it is possible to predict the shape of the catheter only by knowing catheter entrance and tip points.

4.4 Potential Applications of ML in SR

ML can be used for different purposes in SR. In this section, we give some future applications of ML.

Automation of the Surgical Operation

The operating room (OR) is densely populated with different surgical equipments and the surgery team can be quite large. Therefore, the amount of information that is generated can be quite impressive. A surgeon's ability to process all the available information and at the same time establish and sustain an appropriate level of situation awareness is limited and also surgeon-dependent

[93,94]. The cognitive load could potentially be reduced by employing ML techniques. Based on knowledge of the procedure work flow such techniques could provide information and guidance, signaling critical events. Ultimately such techniques could take over repetitive and time consuming tasks. ML-techniques could steer surgical robots to safely, accurately and possibly at faster speed execute some specific surgical tasks. It has already been shown that the time taken by a surgical procedure can be reduced using a robotic scrub nurse [95]. Apart from a reduced operation time, enhanced surgery performance, reduced surgical team instability and miscommunication could be achieved.

Training Young Surgeons

In the current surgical practices, trainees are mainly under the supervision of senior surgeons and surgical skills are also evaluated based on the experiences of the supervisors. Therefore, the experiences of senior surgeons are used as the evaluation criteria for the skills learned by trainees. The evaluation criteria are, however, not accurate and have not been adequately quantified (see Section 3.2). ML approaches have the potential to learn a statistical model of surgical skills of experienced surgeons from a data collected in the OR [96]. The learned surgical skills can be used for quantitatively evaluating surgical skills learned by trainees. Moreover, ML techniques can be used to improve the existing trainers by accurately modeling the interaction among the surgeons, the patients and also the surgical instruments (robots).

Classification and Standardization of the Medical Practice

At present, it is difficult to compare and evaluate the different medical therapies, which are performed by different surgeons and in different hospitals. For reducing the costs and improving quality of health care, a standardization system for best medical practice is desired. A lot of research work has already been carried out over the last years to address this problem [97,98]. The main challenge is to classify the varieties of the skills for different surgeons. ML techniques are able to develop a statistical model, in which the procedure of medical practice can be separated into different steps, and the model learns the best medical practice from all of the surgeons for a given situation. The best practice that is learned can be continually updated or revised automatically.

Saving the Best Recipes of an Experienced Surgeon

ML can be used to learn the skills of an experienced surgeon and save it and use it latter in an OR or use it to train young surgeons. It can also be used to set an initial knowledge of newly introduced robots, so that they continue to refine the initial knowledge they received.

Discovering Novel Recipes

Where detailed and realistic digital or artificial models of the environment exist it becomes possible to combine ML techniques with these models and experiment upon these models. ML techniques such as decision trees and forests, artificial neural networks, Bayesian networks, Support Vector Machines and Gaussian processes [99] could discover and evaluate operating techniques that do not belong to, but potentially outperform current surgical practice.

Safe Interaction between Patients and Surgical Robots

ML-related research exists in developing techniques to model the environment (geometry) more explicitly or to identify some specific features such as anatomic landmarks, mechanical, or physiological properties of this environment. In robotically-assisted surgery, accurate perception of the surgical environment is essential to the surgical robotics control and decision making about how to interact safely within a fragile environment.

4.5 ML for SR - Challenges and Directions for Further Work

While ML is receiving more attention in surgery and robotic surgery in particular, its use in current surgical practice is still very limited. In the following a number of challenges that need to be faced by the research community are listed concisely.

High-quality Medical/Surgical Data

There is a need for large quantitaties of high-quality medical and surgical data to train ML techniques. Data is to be obtained following well-described protocols and stored in standardized formats to ensure interoperability and correct use. In case non-traditional imaging or data-capturing modalities are being used, i.e. requiring actions or sensing that deviates from current standard clinical practice, approval by an Ethical Commission might be required. Furthermore, measures should be installed to ensure protection of patient's privacy.

Modeling Challenges

The major challenge in modeling the surgical environment is the dynamic and deforming nature of the living body which restricts the use of pre-operatively estimated 3D maps and requires the analysis of intra-operative data. For that purpose, geometric, mechanical and physiological features of the environment should be considered and the catheter's proprioception should be combined with information captured from its exteroception to enable accurate perception of the surgical environment. However, the fusion of multiple sensors is not trivial as it involves theoretical and technical challenges such as sensor co-registration, synchronization and information fusion. On top of that, the modeling of the deformation of the environment due to respiratory motion and heart beat is a challenging task.

Learning and Defining Skill Analysis Metric

An important problem in learning skill analysis is to come up with a metric that adequately captures the characteristics of the best practice. Because of the variations during the procedures, the learned skill analysis would only be applicable to a certain group of surgeons. A major challenge for ML would be to learn a general skill analysis metric that can be applied across different groups of surgeons both at the national and international levels. Moreover, the definition of a metric that leads to a desired surgical procedure is difficult. Depending on the structure of the solution space, a given cost function may not lead to the optimal performance. Defining cost functions (objectives) that lead to desired behavior of a surgical robot remains a challenge, especially for complex surgical procedures.

Adaptation to Unknown or Previously Unobserved Situations

Any system deployed in the operating room and given decision making capabilities should be able to cope with uncertainty and unpredictable events just as the expert surgeon can adapt to such situations. The development of algorithms that are able to adapt the learned skill to novel (unexpected) situations is an important challenge for intelligent surgical robots. In this line, transfer learning aims at reducing the need of recollecting the training data, and improving learning in the new task by transferring the knowledge between different task domains.

Pipeline for Training and Deploying Autonomous Surgical Action

Given the large complexity and multidisciplinary nature of the surgical intervention and its automated counter-part, there is a need for a structured approach to efficiently transfer surgical skill towards automated execution. The envisioned framework would guide the skill transfer over all aspects of the surgical procedure, providing tools and guidance to:

- analyse the surgical workflow, query surgeons to identify procedures or parts of procedures for which automation would be of interest;
- set up the surgical scene for gathering data, providing documentation and directions to apply for approval at respective ethical regulatory bodies;
- gather, represent and store data in exchangeable and standardized formats;
- segment, filter and pre-process data for delivery to ML algorithms;
- extract surgical skills and associated reward functions from surgical data.
- train models and controllers to replicate or improve upon surgical skills. This can take place autonomously or through human demonstration and interaction with surgeons;
- evaluate robustness and transferability of learned skills;
- program robotic actions that display a targeted surgical skill
- analyse the scene and interaction to detect transitions or inconsistencies, triggering appropriate robotic actions, event or error handling methods.
- evaluate overall performance in an autonomous manner or by clinical experts.

5 Conclusion

In this paper, we have reported a detailed review of ML methods designed for and used in surgery and SR forming the learning system shown in Fig.1. Synthesizing and exploiting the knowledge and experience of the surgeon requires a thorough understanding and analysis of skill, training and evaluation. By developing the methods and knowledge of the surgical process it may be possible to compute the mapping from perception to action (imitation learning) for various surgical tasks and meanwhile qualitatively analyze a learned skill. By subdividing surgical procedures into individual surgical tasks through episode segmentation a detailed surgical workflow analysis pipeline can be constructed. In each episode, the desired behavior can be learned as skill within the surgical robot control loop and decision making mechanism and thus surgical robots could be developed to operate semi-autonomously or fully autonomously in the future.

ML will play a crucial role in the development of surgical robots. A thorough understanding of skill learning and analysis methods will enable us to develop intelligent surgical robots that learn from human experts efficiently and that can improve existing surgical procedures or discover new ones.

The challenges to overcome with ML technologies for surgical robotics are high-quality medical/surgical data, modeling challenges, learning and defining skill analysis metric, adaptation to unknown or previously unobserved situations and pipeline for training and deploying autonomous surgical action (see section 4.5). Beside those technical challenges getting the acceptance and trust of the physicians and patients is the most crucial challenge.

Conflict of interest

Yohannes Kassahun, Bingbin Yu, Abraham Temesgen Tibebu, Danail Stoyanov, Stamatia Giannarou, Jan Hendrik Metzen and Emmanuel Vander Poorten declare that they have no conflict of interest.

References

1. J. Troccaz. Medical robotics: where we come from, where we are and where we could go. *Industrial Robot: An International Journal*, 35(4), 2008.
2. D. Herron and M. Marohn. A consensus document on robotic surgery. *Surgical Endoscopy*, 22(2):313–325, 2008.
3. C. Camberlin, A. Senn, M. Leys, and C. De Laet. *Robot-assisted Surgery: health technology assessment*. KCE reports 104A. Federaal Kenniscentrum voor de Gezondheidszorg, 2009.
4. K. S. ElSahwi, C. Hooper, M. C. D. Leon, T. N. Gallo, E. Ratner, D.-A. Silasi, A. D. Santin, P. E. Schwartz, T. J. Rutherford, and M. Azodi. Comparison between 155 cases of robotic vs. 150 cases of open surgical staging for endometrial cancer. *Gynecologic Oncology*, 124(2):260 – 264, 2012.
5. L. F. Brandao, R. Autorino, H. Laydner, G.-P. Haber, I. Ouzaid, M. D. Sio, S. Perdonà, R. J. Stein, F. Porpiglia, and J. H. Kaouk. Robotic versus laparoscopic adrenalectomy: A systematic review and meta-analysis. *European Urology*, 65(6):1154 – 1161, 2014.
6. M. Morino, L. Pellegrino, C. Giaccone, C. Garrone, and F. Rebecchi. Randomized clinical trial of robot-assisted versus laparoscopic nissen fundoplication. *British Journal of Surgery*, 93(5):553–558, 2006.
7. I. Nakadi, C. Mélot, J. Closset, V. Moor, K. Bétroune, P. Feron, P. Lingier, and M. Gelin. Evaluation of da Vinci nissen fundoplication clinical results and cost minimization. *World Journal of Surgery*, 30(6):1050–1054, 2006.
8. J. S. Park, G.-S. Choi, S. Y. Park, H. J. Kim, and J. P. Ryuk. Randomized clinical trial of robot-assisted versus standard laparoscopic right colectomy. *British Journal of Surgery*, 99(9):1219–1226, 2012.
9. J. van den Berg, S. Miller, D. Duckworth, H. Hu, A. Wan, X. Fu, K. Goldberg, and P. Abbeel. Superhuman performance of surgical tasks by robots using iterative learning from human-guided demonstrations. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, pages 2074–2081, 2010.
10. Y. Kwoh, J. Hou, E. Jonckheere, and S. Hayati. A robot with improved absolute positioning accuracy for CT guided stereotactic brain surgery. *IEEE Transactions on Biomedical Engineering*, 35(2):153 –160, 1988.
11. H. Paul, W. Bargar, B. Mittlestadt, B. Musits, R. Taylor, P. Kazanzides, J. Zuhars, B. Williamson, and W. Hanson. Development of a surgical robot for cementless total hip arthroplasty. *Clinical Orthopaedics and Related Research ®*, 285:57–66, 1992.
12. C. W. Burckhardt, P. Flury, and D. Glauser. Stereotactic brain surgery. *Engineering in Medicine and Biology Magazine, IEEE*, 14(3):314–317, May 1995.
13. J. Adler, S. Chang, M. Murphy, J. Doty, P. Geis, and S. Hancock. The Cyberknife: a frameless robotic system for radiosurgery. *Stereotactic and Functinonal Neurosurgery*, 69(1-4):124–128, 1997.
14. H. Kang. *Robotic assisted suturing in minimally invasive surgery*. PhD thesis, Rensselaer Polytechnic Institute , Troy, New York, 2002.
15. R. Jackson and M. Cavusoglu. Needle path planning for autonomous robotic surgical suturing. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 1669–1675, May 2013.
16. F. Nageotte, P. Zanne, C. Doignon, and M. de Mathelin. Stitching planning in laparoscopic surgery : Towards robot-assisted suturing. *The International Journal of Robotics Research*, 2009.
17. R. Jansen, K. Hauser, N. Chentanez, F. van der Stappen, and K. Goldberg. Surgical retraction of non-uniform deformable layers of tissue: 2d robot grasping and path planning. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4092–4097, Oct 2009.
18. S. Patil and R. Alterovitz. Toward automated tissue retraction in robot-assisted surgery. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 2088–2094, May 2010.
19. R. Muradore, D. Bresolin, L. Geretti, P. Fiorini, and T. Villa. Robotic surgery - formal verification of plans. *Robotics Automation Magazine, IEEE*, 18(3):24–32, 2011.
20. P. Brett, R. Taylor, D. Proops, C. Coulson, A. Reid, and M. Griffiths. A surgical robot for cochleostomy. In *Proceedings of 29th Annual International conference of the IEEE Engineering in Medicine and Biology Society (EMBS)*, pages 1229 –1232, 2007.
21. J. Larsson and B. Hayes-Roth. Guardian: intelligent autonomous agent for medical monitoring and diagnosis. *Intelligent Systems and their Applications, IEEE*, 13(1):58–64, Jan 1998.
22. C. Frei, A. Gentilini, M. Derighetti, A. Glattfelder, M. Morari, T. Schnider, and A. Zbinden. Automation in anesthesia. In *Proceedings of the American Control Conference (ACC)*, pages 1258–1263, 1999.
23. T. M. Mitchell. *Machine Learning*. McGraw-Hill, 1997.
24. C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer Verlag, 2006.

25. C. J. C. H. Watkins and P. Dayan. Q-Learning. *Machine Learning*, 8(3-4):279–292, May 1992.
26. B. Price and C. Boutilier. A Bayesian approach to imitation in reinforcement learning. In *Proceedings of the 18th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 712–720, 2003.
27. C. E. Birkhimer. *Extracting human strategies for use in robotic assembly*. PhD thesis, Case Western Reserve University, Cleveland, OH, USA, 2005. AAI3159443.
28. B. Hannaford and P. Lee. Hidden Markov model analysis of force/torque information in telemanipulation. *The International Journal of Robotics Research*, 10(5):528–539, 1991.
29. K. Itabashi, K. Hayakawa, T. Suzuki, S. Okuma, and F. Fujiwara. Modeling of the peg-in-hole task based on impedance parameters and HMM. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, volume 1, pages 451–457, 1997.
30. H. Asada and S. Liu. Transfer of human skills to neural net robot controllers. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, volume 3, pages 2442–2448, 1991.
31. M. Kaiser and R. Dillmann. Building elementary robot skills from human demonstration. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, volume 3, pages 2700–2705, 1996.
32. D. Myers. An approach to automated programming of industrial robots. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, volume 4, pages 3109–3114, 1999.
33. M. Skubic and R. Volz. Acquiring robust, force-based assembly skills from human demonstration. *IEEE Transactions on Robotics and Automation*, 16(6):772–781, 2000.
34. A. Y. Ng and S. J. Russell. Algorithms for Inverse Reinforcement Learning. In *Proceedings of the International Conference on Machine Learning (ICML)*, pages 663–670, 2000.
35. P. Abbeel and A. Y. Ng. Apprenticeship learning via inverse reinforcement learning. In *Proceedings of the International Conference on Machine Learning (ICML)*, 2004.
36. D. Ramachandran and E. Amir. Bayesian inverse reinforcement learning. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, pages 2586–2591, 2007.
37. D. E. Kirk. *Optimal Control Theory: An Introduction*. Prentice-Hall, 1970.
38. R. K. Reznick and H. MacRae. Teaching surgical skills - changes in the wind. *The New England Journal of Medicine*, 355(25):2664–2669, 2006.
39. A. Jog, B. Itkowitz, M. Liu, S. DiMaio, G. Hager, M. Curet, and R. Kumar. Towards integrating task information in skills assessment for dexterous tasks in surgery and simulation. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, pages 5273–5278. IEEE, 2011.
40. P. A. Kenney, M. F. Wszolek, J. J. Gould, J. A. Libertino, and A. Moinzadeh. Face, content, and construct validity of dV-trainer, a novel virtual reality simulator for robotic surgery. *Urology*, 73(6):1288–1292, 2009.
41. M. M. Tedesco, J. J. Pak, E. J. Harris Jr, T. M. Krummel, R. L. Dalman, and J. T. Lee. Simulation-based endovascular skills assessment: the future of credentialing? *Journal of Vascular Surgery*, 47(5):1008 – 1014, 2008.
42. C. E. Reiley, H. C. Lin, D. D. Yuh, and G. D. Hager. Review of methods for objective surgical skill evaluation. *Surgical Endoscopy*, 25(2):356–366, 2011.
43. K. Moorthy, Y. Munz, S. K. Sarker, and A. Darzi. Objective assessment of technical skills in surgery. *The British Medical Journal (BMJ)*, 327(7422):1032–1037, 2003.
44. R. Reznick, G. Regehr, H. MacRae, J. Martin, and W. McCulloch. Testing technical skill via an innovative bench station examination. *The American Journal of Surgery*, 173(3):226–230, 1997.
45. J. D. Doyle, E. M. Webber, and R. S. Sidhu. A universal global rating scale for the evaluation of technical skills in the operating room. *The American Journal of Surgery*, 193(5):551–555, 2007.
46. P. D. Van Hove, G. J. M. Tuijthof, E. G. G. Verdaasdonk, L. P. S. Stassen, and J. Dankelman. Objective assessment of technical surgical skills. *British Journal of Surgery*, 97:972–987, 2010.
47. H. Niitsu, N. Hirabayashi, M. Yoshimitsu, T. Mimura, J. Taomoto, Y. Sugiyama, S. Murakami, S. Saeki, H. Mukaida, and W. Takiyama. Using the objective structured assessment of technical skills (OSATS) global rating scale to evaluate the skills of surgical trainees in the operating room. *Surgery Today*, 43(3):271–275, 2013.
48. B. Bridgewater, A. D. Grayson, M. Jackson, N. Brooks, G. J. Grotte, D. J. M. Keenan, R. Millner, B. M. Fabri, and J. Mark. Surgeon specific mortality in adult cardiac surgery: comparison between crude and risk stratified data. *The British Medical Journal (BMJ)*, 327(7405):13–17, 2003.
49. A. Dosis, R. Aggarwal, F. Bello, K. Moorthy, Y. Munz, D. Gillies, and A. Darzi. Synchronized video and motion analysis for the assessment of procedures in the operating theater. *Archives of Surgery*, 140(3):293–299, 2005.
50. M. P. Schijven, J. Jakimowicz, and C. Schot. The advanced dundee endoscopic psychomotor tester (ADEPT) objectifying subjective psychomotor test performance. *Surgical Endoscopy*, 16(6):943–948, 2002.
51. S. M. B. I. Botden, S. N. Buzink, M. P. Schijven, and J. J. Jakimowicz. ProMIS augmented reality training of laparoscopic procedures face validity. *The Journal of the Society for Simulation in Healthcare*, 3(2):97–102, 2008.
52. M. Hattori, H. Egi, M. Tokunaga, T. Suzuki, H. Ohdan, and T. Kawahara. The integrated deviation in the HUESAD (Hiroshima University Endoscopic Surgical Assessment Device) represents the surgeon's visual-spatial ability. In *Proceedings of The 2012 ICME International Conference on Complex Medical Engineering (CME)*, pages 316–320, July 2012.
53. P. J. van Empel, L. B. van Rijssen, J. P. Commandeur, M. G. E. Verdam, J. A. Huirne, F. Scheele, H. J. Bonjer, and W. J. Meijerink. Validation of a new box trainer-related tracking device: the TrEndo. *Surgical Endoscopy*, 26:2346–2352, 2012.
54. S. P. DiMaio and C. J. Hasser. The da Vinci research interface. *MICCAI Workshop on Systems and Arch. for Computer Assisted Interventions, MIDAS Journal*, 2008.
55. P. Kazanzides, S. DiMaio, A. Deguet, B. Vagvolgyi, M. Balicki, C. Schneider, R. Kumar, A. Jog, B. Itkowitz, C. Hasser, and R. Taylor. The Surgical Assistant Workstation (SAW) in minimally-invasive surgery and microsurgery. In *Proceedings of MICCAI Workshop on Systems and Architecture for Computer Assisted Interventions (SACAI)*, 2010.
56. J. Rosen, J. D. Brown, L. Chang, M. N. Sinanan, and B. Hannaford. Generalized approach for modeling minimally invasive surgery as a stochastic process using a discrete Markov model. *IEEE Transactions on Biomedical Engineering*, 53(3):399–413, 2006.

57. S. Speidel, T. Zentek, G. Sudra, T. Gehrig, B. P. Müller-Stich, C. Gutt, and R. Dillmann. Recognition of surgical skills using hidden markov models. In *Proceedings of SPIE Medical Imaging*, volume 7261, page 25, 2009.
58. H. C. Lin, I. Shafran, T. E. Murphy, A. M. Okamura, D. D. Yuh, and G. D. Hager. Automatic detection and segmentation of robot-assisted surgical motions. In *Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 802–810. Springer, 2005.
59. K. Ahmed, D. Miskovic, A. Darzi, T. Athanasiou, and G. B. Hanna. Observational tools for assessment of procedural skills: a systematic review. *The American Journal of Surgery*, 202(4):469–480, 2011.
60. C. V. Shah, C. V. Riga, D. Stoyanov, G.-Z. Yang, N. J. W. Cheshire, and C. D. Bicknell. Video motion analysis for approaching objective assessment of catheter-based endovascular interventions. In *Proceedings of The Hamlyn Symposium on Medical Robotics*, pages 93–94, 2011.
61. S. K. Neequaye, R. Aggarwal, I. Van Herzeele, A. Darzi, and N. J. Cheshire. Endovascular skills training and assessment. *Journal of Vascular Surgery*, 46(5):1055–1064, 2007.
62. C. E. Reiley and G. D. Hager. Task versus subtask surgical skill evaluation of robotic minimally invasive surgery. In *Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI)* , pages 435–442. Springer, 2009.
63. T. Mkowalewski, J. Rosen, L. Chang, M. Sinanan, and B. Hannaford. Optimization of a vector quantization codebook for objective evaluation of surgical skill. *Studies in Health Technology and Informatics*, 98:174–179, 2004.
64. H. Lin, I. Shafran, D. Yuh, and G. Hager. Towards automatic skill evaluation: detection and segmentation of robot-assisted surgical motions. *Computer Aided Surgery*, 11(5):220–230, 2006.
65. C. Reiley, H. Lin, B. Varadarajan, B. Vagvolgyi, S. Khudanpur, D. Yuh, and Hager. Automatic recognition of surgical motions using statistical modeling for capturing variability. *Studies in Health Technology and Informatics*, 132:396–401, 2008.
66. B. Allen, V. Nistor, E. Dutson, G. Carman, C. Lewis, and P. Faloutsos. Support vector machines improve the accuracy of evaluation for the performance of laparoscopic training tasks. *Surgical Endoscopy*, 24:170, 2009.
67. L. Tao, E. Elhamifar, S. Khudanpur, G. D. Hager, and R. Vidal. Sparse hidden markov models for surgical gesture classification and skill evaluation. In *Proceedings of Information Processing in Computer-Assisted Interventions (IPCAI)*, pages 167–177. Springer, 2012.
68. H. Rafii-Tari, J. Liu, S.-L. Lee, C. Bicknell, and G.-Z. Yang. Learning-based modeling of endovascular navigation for collaborative robotic catheterization. In *Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 8150, pages 369–377. Springer, 2013.
69. R. Kumar, A. Jog, B. Vagvolgyi, H. Nguyen, G. Hager, C. C. G. Chen, and D. Yuh. Objective measures for longitudinal assessment of robotic surgery training. *The Journal of Thoracic and Cardiovascular Surgery*, 143(3):528–534, 2012.
70. N. Padoy, T. Blum, I. Essa, H. Feussner, M.-O. Berger, and N. Navab. A boosted segmentation method for surgical workflow analysis. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 4791 of *Lecture Notes in Computer Science*, pages 102–109. Springer Berlin Heidelberg, 2007.
71. S.-A. Ahmadi, T. Sielhorst, R. Stauder, M. Horn, H. Feussner, and N. Navab. Recovery of surgical workflow without explicit models. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 4190, pages 420–428. Springer, 2006.
72. C. Reiley and G. Hager. Task versus subtask surgical skill evaluation of robotic minimally invasive surgery. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 5761, pages 435–442. Springer, 2009.
73. B. Varadarajan, C. Reiley, H. Lin, S. Khudanpur, and G. Hager. Data-derived models for segmentation with application to surgical assessment and training. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 5761, pages 426–434. Springer, 2009.
74. B. Varadarajan. *Learning and Inference Algorithms for Dynamical System Models of Dextrous Motion*. PhD thesis, The Johns Hopkins University, 2011.
75. J. Rosen, J. Brown, L. Chang, M. Sinanan, and B. Hannaford. Generalized approach for modeling minimally invasive surgery as a stochastic process using a discrete markov model. *IEEE Transactions on Biomedical Engineering*, 53(3):399–413, 2006.
76. G. P. Moustris, S. C. Hiridis, K. M. Deliparaschos, and K. M. Konstantinidis. Evolution of autonomous and semi-autonomous robotic surgical systems: a review of the literature. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 7(4):375–392, 2011.
77. A. Casals, J. Amat, and E. Laporte. Automatic guidance of an assistant robot in laparoscopic surgery. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, volume 1, pages 895–900, 1996.
78. A. Krupa, J. Gangloff, C. Doignon, M. de Mathelin, G. Morel, J. Leroy, L. Soler, and J. Marescaux. Autonomous 3-D positioning of surgical instruments in robotized laparoscopic surgery using visual servoing. *IEEE Transactions on Robotics and Automation*, 19(5):842–853, 2003.
79. P. Hynes, G. Dodds, and A. J. Wilkinson. Uncalibrated Visual-Servoing of a Dual-Arm Robot for MIS Suturing. In *Proceedings of the first IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, pages 420–425, 2006.
80. T. Osa, C. Staub, and A. Knoll. Framework of automatic robot surgery system using Visual servoing. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1837–1842, 2010.
81. R. Ginhoux, J. Gangloff, M. De Mathelin, L. Soler, M. Sanchez, and J. Marescaux. Active filtering of physiological motion in robotized surgery using predictive control. *IEEE Transactions on Robotics*, 21(1):67–79, 2005.
82. O. Weede, H. Mönnich, B. Müller, and H. Wörn. An intelligent and autonomous endoscopic guidance system for minimally invasive surgery. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, pages 5762–5768, 2011.
83. H. Mayer, F. Gomez, D. Wierstra, I. Nagy, A. Knoll, and J. Schmidhuber. A System for Robotic Heart Surgery that Learns to Tie Knots Using Recurrent Neural Networks. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 543–548, 2006.

84. H. Mayer, D. Wierstra, F. Gomez, I. Nagy, A. Knoll, and J. Schmidhuber. A System for Robotic Heart Surgery that Learns to Tie Knots Using Recurrent Neural Networks. *Advanced Robotics*, 22(13-14):1521–1537, 2008.
85. H. Mayer, I. Nagy, D. Burschka, A. Knoll, E. Braum, R. Lange, and R. Bauernschmitt. Automation of Manual Tasks for Minimally Invasive Surgery. In *Proceedings of International Conference on Autonomic and Autonomous Systems (ICAS)*, pages 260–265, 2008.
86. H. Mayer, I. Nagy, A. Knoll, E. Braun, R. Lange, and R. Bauernschmitt. Adaptive control for human-robot skilltransfer: Trajectory planning based on fluid dynamics. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, pages 1800–1807, April 2007.
87. S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.
88. J. Schulman, J. Ho, C. Lee, and P. Abbeel. Learning from Demonstrations through the Use of Non-Rigid Registration. In *Proceedings of the 16th International Symposium on Robotics Research (ISRR)*, 2013.
89. S. Calinon, D. Bruno, M. S. Malekzadeh, T. Nanayakkara, and D. G. Caldwell. Human-robot skills transfer interfaces for a flexible surgical robot. *Computer Methods and Programs in Biomedicine*, 2014.
90. Y. Kassahun, B. Yu, and E. Vander Poorten. Learning catheter-aorta interaction model using joint probability densities. In *Proceedings of the 3rd Joint Workshop on New Technologies for Computer/Robot Assisted Surgery*, 2013.
91. G. P. Mylonas, P. Giataganas, M. Chaudery, V. Vitiello, A. Darzi, and G.-Z. Yang. Autonomous eFAST ultrasound scanning by a robotic manipulator using learning from demonstrations. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3251–3256. IEEE, 2013.
92. P. Giataganas, V. Vitiello, V. Simaiaki, E. Lopez, and G.-Z. Yang. Cooperative in situ microscopic scanning and simultaneous tissue surface reconstruction using a compliant robotic manipulator. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, pages 5378–5383. IEEE, 2013.
93. J. K. Maron and M. A. Zenati. Framework for managing cognitive load in a data-rich robotic operating room. In *Proceedings of the 3rd Joint Workshop on New Technologies for Computer/Robot Assisted Surgery (CRAS)*, pages 150–153, 2013.
94. M. R. Endsley. Toward a Theory of Situation Awareness in Dynamic System. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1):32–64, 1995.
95. M. G. Jacob, Y.-T. Li, G. A. Akingba, and J. P. Wachs. Collaboration with a robotic scrub nurse. *Communications of the ACM*, 56(5):68–75, 2013.
96. C. Reiley, E. Plaku, and G. Hager. Motion generation of robotic surgical tasks: Learning from expert demonstrations. In *Proceedings of Annual International Conference of the IEEE on Engineering in Medicine and Biology Society (EMBC)*, pages 967–970, 31 2010-sept. 4 2010.
97. ACOG Committee on Patient Safety and Quality Improvement. ACOG committee opinion no. 526: Standardization of practice to improve outcomes. *Obstetrics & Gynaecology*, 119(5):1081–1082, 2012.
98. PSQH: Advancing the Practice of Evidence-Based Medicine with Standardized Order Sets. [Online; accessed 14-April-2014].
99. G. Meyfroidt, F. Güizab, J. Ramonb, and M. Bruynooghe. Machine learning techniques to examine large patient databases. *Best Practice & Research: Clinical Anaesthesiology*, 23(1):127–143, 2009.