e-ISSN: 3023-8609
RESEARCH ARTICLE
Vol.2 / No.2 / 2025 / pp.26-35
https://doi.org/10.69882/adba.cem.2025071

Deep Learning in Robot-Assisted Surgery: A Conceptual Framework for the da Vinci System

Hakan Özcan *,1
*Department of Computer Technology, Design Vocational School, Amasya University, Amasya, 05100, Türkiye.

ABSTRACT This study proposes a conceptual framework for integrating deep learning into the da Vinci Surgical System. The framework was developed after identifying common applications and challenges through a systematic review. It combines multiple data types, including visual, kinematic, and physiological signals, into a closed-loop system. This system includes four core components: data acquisition and fusion, deep learning-based analysis, adaptive control and feedback, and continuous skill assessment. These components interact to support real-time surgical guidance and personalized training, aiming to improve surgical outcomes. To inform the framework, a systematic search of the PubMed database was conducted, focusing on studies that combine deep learning with the da Vinci system. Two major application areas were identified: the first involves autonomous or semi-autonomous instrument control and image-guided navigation, and the second covers surgical skill assessment and workflow analysis. While existing studies demonstrate the potential of artificial intelligence in robotic surgery, they also reveal technical and practical limitations. By analyzing these common approaches and obstacles, this study provides a structured foundation for future research. The proposed framework offers a unified view that connects data processing, intelligent decision-making, and feedback, paving the way for smarter systems that assist both in real-time procedures and long-term training, and helping bridge the gap between advanced robotic hardware and cognitive support in surgical environments.

KEYWORDS
Artificial intelligence
Deep learning
Robotic surgery
Robot-assisted surgery
da Vinci System
Conceptual framework
Decision support system

INTRODUCTION

In recent years, advances in surgical technology have significantly transformed patient care and operative procedures. Robot-assisted surgery (RAS) now enables surgeons to perform complex interventions with enhanced precision, reduced invasiveness, and improved efficiency (Reddy *et al.* 2023). These innovations have contributed to better clinical outcomes, including shorter recovery times, reduced postoperative complications, and increased procedural consistency (Kinoshita *et al.* 2022; Saman 2024). As a result, robotic systems have become a central feature in many surgical specialties (Cannizzaro *et al.* 2024).

A particularly important development in this field is the integration of advanced decision-support technologies into RAS platforms (Iftikhar *et al.* 2024). Artificial intelligence (AI), and more specifically deep learning (DL) techniques, have emerged as critical components of this evolution (Panesar *et al.* 2019). These techniques are capable of processing complex, multi-modal datasets to support real-time intraoperative decision-making, analyze surgical performance, and predict patient-specific risks (Egert *et al.* 2020). Recent studies emphasize that the convergence of robotic systems with AI has the potential to enhance intraoperative safety, personalize surgical strategies, and improve outcomes by extending the surgeon's capabilities beyond mechanical precision alone (Iftikhar *et al.* 2024; Knudsen *et al.* 2024).

One of the most prominent platforms in RAS is the da Vinci Surgical System, developed by Intuitive Surgical in the late 1990s and first introduced into clinical practice in 1999 (Pugin *et al.* 2011). Now used in hospitals worldwide, the da Vinci system has become a cornerstone of minimally invasive surgery (Azizian *et al.* 2020; DiMaio *et al.* 2011). It integrates surgeon-controlled robotic arms with an ergonomic console and a high-definition (HD), 3-dimensional

Manuscript received: 10 May 2025,
Revised: 9 July 2025,
Accepted: 14 July 2025.

[1]hozcan@amasya.edu.tr (**Corresponding author**).

(3D) endoscopic imaging platform, enabling precise replication of the surgeon's hand movements with enhanced dexterity and tremor filtration (Cepolina and Razzoli 2022; Pugin *et al.* 2011). This setup allows for careful tissue handling and suturing through small incisions, improving surgical precision and supporting faster patient recovery (Cepolina and Razzoli 2022). The system's design provides a stable, immersive view of the operative field and facilitates complex procedures with greater control and better outcomes compared to conventional minimally invasive techniques (Steffens *et al.* 2023). Its 3D visualization significantly enhances spatial orientation and depth perception, which is essential for performing intricate procedures with minimal trauma to surrounding tissues (Koh *et al.* 2018). Over the past two decades, the da Vinci Surgical System has demonstrated substantial clinical impact across multiple specialties, particularly urology (Koukourikis and Rha 2021), gynecology (Li *et al.* 2022), and general surgery (Celotto *et al.* 2024). It has also served as a key research and training platform through initiatives such as the da Vinci Research Kit (dVRK), which has fostered innovation and collaboration in the field of surgical robotics (D'Ettorre *et al.* 2021).

Building on this established success, it becomes clear that the next frontier for RAS lies in augmenting these systems with advanced decision-support capabilities. Despite significant technological progress, a notable gap remains in the systematic and clinically validated integration of AI-driven systems into RAS workflows (Boal *et al.* 2024). While DL models have shown potential in interpreting large-scale visual, kinematic, and physiological data for intraoperative guidance (Knudsen *et al.* 2024), current research is often isolated to specific tasks, limited to controlled settings, or lacking cross-platform generalizability. Moreover, a unified architectural framework that integrates these diverse data modalities into a cohesive real-time support system is still lacking, despite the need for adaptable and validated solutions in evolving robotic surgery (Marcus *et al.* 2024). This fragmentation not only hampers clinical translation but also limits opportunities for scalable training, consistent evaluation, and adaptive surgical assistance.

To address this gap, the present study conducts a systematic literature review to examine thematic application domains where DL is employed to enhance various aspects of RAS with the da Vinci system. The review provides a detailed analysis of these applications and identifies persistent challenges hindering clinical integration. Building on these insights, the study proposes a conceptual framework for a multi-modal surgical decision support system (DSS). This framework integrates diverse data streams, including visual, kinematic, and physiological inputs, into a unified closed-loop system designed to support real-time decision-making and continuous surgical skill assessment.

By synthesizing recent advances and outlining future directions, this study aims to contribute to the development of intelligent, adaptive, and data-driven RAS platforms, ultimately facilitating more effective and personalized patient care.

MATERIAL AND METHODS

In order to identify the core components of a conceptual framework, a systematic literature review was conducted following the PRISMA guidelines (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (Page *et al.* 2021). Predefined inclusion and exclusion criteria were applied to select recent studies focused on DL applications in robotic surgery using the da Vinci Surgical System. The study design is illustrated in Figure 1. Selected studies were grouped by application domain, and key challenges were identified within each group. Based on these insights, a conceptual framework for a multi-modal surgical DSS was proposed to unify the identified components into a cohesive system.

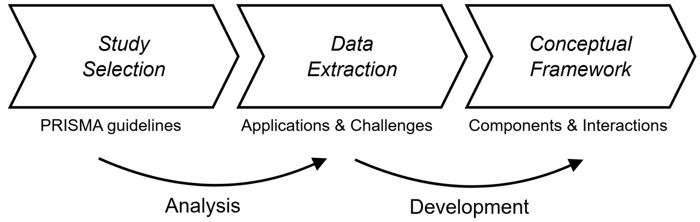

Figure 1 Study design

Study Selection

Initially, 23 publications related to the da Vinci Surgical System and DL applications were retrieved from the PubMed database using the query shown in Figure 2. Predefined inclusion and exclusion criteria guided the study selection. Specifically, studies focusing exclusively on RAS and DL or machine learning applications with the da Vinci Surgical System, published within the last five years and written in English, were included. During the abstract screening, three studies were excluded because they used da Vinci system outputs solely for classification model building or diagnostic support, without directly addressing robotic surgery. After full-text review, one additional study was excluded for not being directly related to robotic surgery or the da Vinci system. In total, 19 studies met the inclusion criteria and were incorporated into the review.

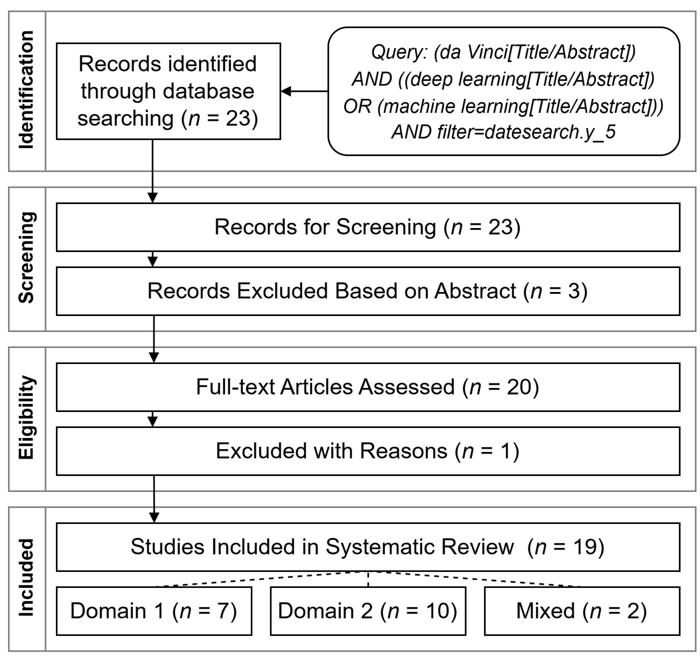

Figure 2 Study selection flow

Data Extraction and Analysis

Following PRISMA guidelines, the findings of each study were systematically summarized and analyzed. Methodological approaches, technologies utilized, and key outcomes were compared. The review was performed in two stages: first, the studies were grouped according to their application domains; then, each group was re-examined to identify specific challenges in applying DL, which were organized into subtopics.

Conceptual Framework Development

Finally, based on the insights gained from these studies, a conceptual framework for a multi-modal surgical DSS was proposed. This framework is designed to integrate multi-modal data streams from visual, kinematic, and physiological sources into a unified, closed-loop system that supports both surgical decision-making and continuous skill improvement. The framework is intended to serve as a foundation for future research and development in RAS DSSs.

RESULTS AND DISCUSSION

This study, guided by the PRISMA methodology outlined in the methods section, synthesizes current research on the integration of DL into the da Vinci Surgical System. Based on analysis of the selected studies, two primary domains of application have been identified. The first involves autonomous or semi-autonomous instrument control and image-guided navigation, where DL techniques are applied to enhance intraoperative decision-making, visual guidance, and robotic manipulation. The second focuses on surgical skill assessment and workflow analysis, highlighting how machine learning algorithms, combined with physiological and performance data, are used to evaluate surgical proficiency, monitor ergonomics, and support training. As illustrated in Figure 3, there has been a steady growth in the number of studies across both domains, with a noticeable increase in research addressing mixed implementation approaches in recent years. The following sections explore each of these domains in detail, emphasizing the transformative role of DL in RAS.

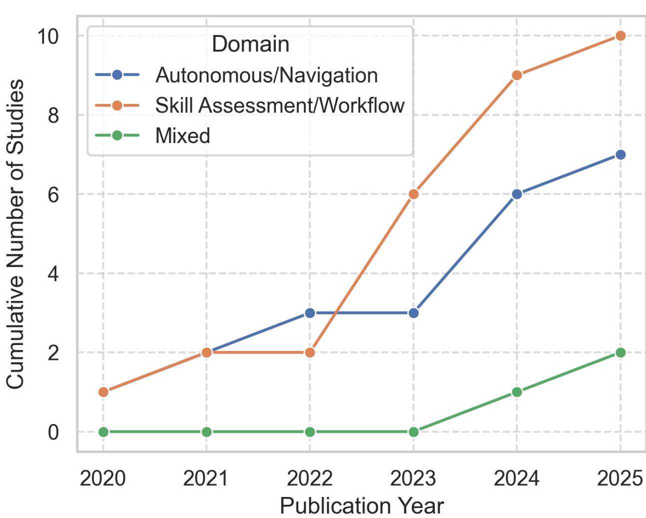

Figure 3 Cumulative deep learning–related *da Vinci* studies by domain (2020–2025)

Autonomous or Semi-Autonomous Instrument Control and Image-Guided Navigation

The integration of DL into the da Vinci Surgical System has led to significant advancements in autonomous and semi-autonomous robotic capabilities, particularly in the realms of real-time visual guidance, instrument control, and tactile perception. Relevant studies on autonomous and semi-autonomous instrument control and image-guided navigation are given in Table 1. These innovations are increasingly contributing to enhanced surgical precision, improved safety, and more efficient intraoperative workflows.

One of the most impactful developments involves real-time 3D model registration and augmented reality (AR)-based navigation. This was demonstrated in a study (Amparore *et al.* 2024) that leveraged the da Vinci system to enhance robotic partial nephrectomy by automatically overlaying patient-specific 3D kidney models onto the surgeon's live endoscopic view. Two complementary approaches were explored. The first utilized computer vision techniques, enhanced by indocyanine green (ICG) fluorescence imaging, to segment kidney structures and establish landmark-based registration. The second approach employed a convolutional neural network (CNN) to automatically detect and segment the kidney directly from live endoscopic images, without relying on fluorescence. Both methods were integrated into the da Vinci system via the TilePro multi-input display, enabling real-time AR guidance during surgery. The reported co-registration times ranged from 7 to 11 seconds, with only minimal manual adjustments required. This automatic overlay system improves intraoperative visualization, supports precise surgical navigation, and reduces cognitive load on the operating surgeon.

Another key area of innovation is camera motion estimation in dynamic surgical environments, which is essential for developing intelligent camera systems. This challenge was addressed in a study (Huber *et al.* 2022) that proposed a DL approach for estimating laparoscopic camera motion using deep homography networks. The researchers created synthetic surgical sequences by introducing artificial camera motion into static da Vinci video frames and used these sequences to train a neural network based on ResNet-34. The proposed system was able to estimate camera motion accurately, even in the presence of moving instruments and tissue, significantly outperforming classical computer vision techniques such as SURF combined with RANSAC. The method achieved up to 41% higher precision and was approximately 43% faster on standard central processing units (CPUs). This capability lays the foundation for intelligent and adaptive camera control systems that can autonomously adjust viewpoint and maintain optimal visual context during complex procedures.

In the domain of skill transfer and task automation, a study (Gonzalez *et al.* 2021) used the dVRK as one of four robotic platforms in the development of the Dexterous Surgical Skill (DESK) dataset. The DESK dataset was designed to enable machine learning models to generalize across different robotic systems and environments, supporting the development of semi-autonomous capabilities in settings such as remote or battlefield surgery. Using the dVRK, the researchers performed a peg transfer task that was broken down into modular surgical gestures, referred to as "surgemes", such as grasping, transferring, and releasing pegs. Detailed kinematic data, including gripper state, position, and orientation, were recorded and used to train classification models. Remarkably, machine learning models trained exclusively on simulation data were able to achieve an accuracy of 93% when tested on real robotic executions using the da Vinci system. This result underscores the potential for leveraging virtual training environments and simulated data to develop generalizable models for real-world autonomous surgical applications.

Another important line of research involves enhancing the realism of biomechanical simulations used in training and preoperative planning. This was demonstrated in a study (Wu *et al.* 2020) that used the dVRK platform in combination with an RGB-depth

Table 1 Studies on autonomous and semi-autonomous instrument control & image-guided navigation

Study	Focus	Methodology/Algorithm	Modalities
(Wu *et al.* 2020)	Simulation correction in soft-tissue modeling	Modified 3D-UNet for correction factor prediction	RGB-D and kinematics
(Gonzalez *et al.* 2021)	Surgeme classification via transfer learning	Supervised learning (RF, SVM, MLP) with FFT features	Kinematics and annotated video images
(Huber *et al.* 2022)	Laparoscopic camera motion extraction	DNN-based homography estimation (using synthetic augmentation)	Laparoscopic surgery video images
(Amparore *et al.* 2024)	3D virtual model overlay in robotic partial nephrectomy	ICG-enhanced computer vision + CNN-based kidney segmentation	RGB endoscopic images (± NIRF) & 3D kidney models
(Ullah *et al.* 2024)	Fluorescence-guided tumor margin delineation	Comparison of visible vs. NIR probes; AI/ML for probe design	Visible and NIRF imaging
(Yilmaz *et al.* 2024)	Sensorless haptic feedback	Disturbance observer with deep learning dynamic compensation	Motor kinematics and force data
(Khan *et al.* 2025)	ML-based tension measurement	LSTM-based torque estimation	dVRK kinematics and force data

Abbreviations: 3D-UNet, Three-Dimensional U-Net; AI/ML, Artificial Intelligence / Machine Learning; AR, Augmented Reality; CNN, Convolutional Neural Network; dVRK, da Vinci Research Kit; DNN, Deep Neural Network; FFT, Fast Fourier Transform; ICG, Indocyanine Green; LSTM, Long Short-Term Memory; MLP, Multi-Layer Perceptron; NIRF, Near-Infrared Fluorescence; RF, Random Forest; RGB-D, Red Green Blue + Depth; SVM, Support Vector Machine.

camera sensor to capture visual and kinematic data while physically interacting with soft-tissue phantoms. They employed this real-time data to improve Finite Element Method (FEM)-based soft tissue models. Specifically, the authors train a neural network to learn the difference between the FEM's predicted tissue shape and the actual tissue shape, as measured by real-time camera and robot data (vision and kinematics). The network then generates a correction factor that adjusts the FEM output to more closely reflect the true, observed deformation of the soft-tissue phantom. The resulting hybrid model reduced simulation errors by 15–30%. These advances underscore the importance of increased realism in soft-tissue simulation, both for surgical training and for machine learning systems that depend on accurate virtual environments to learn complex surgical tasks.

The challenge of providing direct haptic feedback in the da Vinci system has been widely recognized, as tactile information is critical for safely manipulating delicate tissues. This issue was addressed in a study (Yilmaz *et al.* 2024) that implemented sensorless haptic feedback using internal dVRK signals such as joint positions, velocities, and motor torques. In this study, DL is used to estimate the robot's internal dynamics so the system can accurately distinguish external forces (e.g., contact with tissue or surgical objects) from the robot's own motion. Specifically, the authors train a Long Short-Term Memory (LSTM)-based neural network on the dVRK's joint positions and velocities to learn friction and inertia, then subtract this modeled disturbance from the measured motor torque to estimate external contact forces. This sensorless approach provides the basis for haptic feedback. Two experimental tasks were conducted. The first involved tumor detection via palpation, where sensorless feedback increased detection accuracy by approximately 30% compared to visual-only guidance and nearly matched the performance of systems using physical force sensors. The second task focused on the classic peg transfer evaluation, where sensorless feedback significantly reduced unwanted interaction forces by approximately threefold, improving safety and surgical control. Additionally, dynamic modeling contributed to a better user experience by reducing both physical strain and cognitive workload. Importantly, this technique required no additional hardware, highlighting its practical value for improving haptics in existing surgical systems.

In a related effort to enhance intraoperative safety, researchers (Khan *et al.* 2025) developed a machine learning-based method for measuring tissue tension during robotic surgery using the dVRK. Their approach involved using standard da Vinci instruments to apply pulling forces on porcine colon tissues. A LSTM network was trained on kinematic data (joint positions and velocities) and successfully predicted tissue tension with an average accuracy of 74%, reaching up to 88% in optimal cases. The correlation between predicted and ground-truth forces was consistently strong, with Spearman's correlations ≥ 0.81 in four of five experiments (mean ≈ 0.80). By enabling real-time monitoring of tissue stress, this method offers a way to prevent complications such as anastomotic leaks and opens new possibilities for quantitative intraoperative guidance.

Furthermore, emerging research is increasingly leveraging advanced fluorescence imaging modalities to support surgical decision-making. A recent review (Ullah *et al.* 2024) discussed the potential synergy between novel fluorescent probes, near-infrared imaging windows, and robotic platforms in fluorescence imaging-guided surgery. This work emphasized that future directions could include "fluorescence-guided da Vinci" systems, where DL potentially augments the fluorescence signal to provide automated tissue differentiation or tumor margin detection. Taken together with ICG-based AR, as demonstrated in the study (Amparore *et al.* 2024), these developments pave the way for even more sophisticated guidance strategies that seamlessly blend AI-driven analytics, robotics, and real-time imaging to enhance surgical precision.

Overall, these studies demonstrate how DL techniques, when integrated with the da Vinci Surgical System's visual and kinematic data, are pushing the boundaries of robotic surgery. From enhancing visual guidance through AR to predicting physical interactions with tissue, these developments are gradually transforming the da Vinci platform into a more autonomous and intelligent surgical assistant capable of supporting complex intraoperative tasks with greater safety, accuracy, and adaptability.

Surgical Skill Assessment and Workflow Analysis

The da Vinci Surgical System plays a central role in the development of objective, data-driven approaches to surgical skill assessment and intraoperative workflow analysis. Relevant studies focusing on surgical skill assessment and workflow analysis are presented in Table 2. Leveraging its precise motion tracking, real-time data capture capabilities, and compatibility with physiological monitoring systems, the da Vinci platform has enabled the integration of AI-based methods into both surgical training and evaluation environments.

A study (Egert *et al.* 2020) emphasize the foundational role of the da Vinci system in capturing automated performance metrics, such as instrument motion and camera positioning. These quantitative indicators of surgeon performance serve as inputs for machine learning algorithms that can assess surgical proficiency and even predict patient outcomes based on observed patterns. In addition to procedural logging, the da Vinci system supports training through playback features that allow trainees to review expert-performed tasks. These recordings can also be physically re-experienced via tactile robotic control, facilitating the development of procedural muscle memory and enhancing motor skill acquisition. These capabilities suggest that AI-integrated robotic platforms could provide standardized, objective feedback in surgical education.

Several studies have explored the integration of physiological and behavioral monitoring with the da Vinci system to assess surgical skill levels. One example of this approach is a study (Shafiei *et al.* 2024b) that predicted Robotic Anastomosis Competency Evaluation metrics during vesico-urethral anastomosis. The authors collected electroencephalography (EEG) and eye-tracking data from 23 participants performing procedures on plastic models and animal tissue using the da Vinci system. These data were used to train machine learning models, including random forest and gradient boosting regression, which could accurately predict subtask-level performance such as needle handling, tissue trauma, and suture placement. The ability to distinguish performance differences among surgeons of varying experience levels underscores the potential of using physiological and visual metrics for precise, objective skill assessment.

Complementary research (Takács *et al.* 2024) examined the stress and ergonomic impact of RAS with the da Vinci Xi while surgeons performed standardized tasks on a Sea Spikes phantom. Using digital sensors that captured heart-rate variability, posture, and hand movements, the authors trained machine learning models to classify self-reported workload levels and novice versus expert status from these physiological indicators. Preliminary findings suggest that, although the da Vinci's ergonomic design reduces physical strain relative to open surgery, remote tele-operation introduces distinct cognitive and stress-related challenges that differ between novice and experienced surgeons. This insight supports incorporating stress monitoring and ergonomics into future training and evaluation models.

The combined use of EEG and eye-tracking data was further explored in another study (Shafiei *et al.* 2023b) involving live animal surgeries using the da Vinci system. Surgeons performed cystectomy, hysterectomy, and nephrectomy procedures while their physiological responses were recorded during subtasks such as blunt dissection and tissue retraction. Skill levels were assigned using the modified Global Evaluative Assessment of Robotic Skills (GEARS) tool, and models such as multinomial logistic regression, random forest, and gradient boosting were trained to classify surgical expertise. The gradient boosting model achieved classification accuracies up to 93%, demonstrating the reliability of this multimodal approach for distinguishing skill levels in real-time surgical environments.

Building on this line of inquiry, an integrated EEG and eye-tracking study (Shafiei *et al.* 2024a) using eXtreme Gradient Boosting (XGBoost) evaluated mental workload across various simulator- and Fundamentals of Laparoscopic Surgery-based tasks in RAS. This study recruited 26 participants performing Matchboard and Ring Walk exercises on a da Vinci simulator, as well as Fundamentals of Laparoscopic Surgery pattern cut and suturing tasks. The XGBoost models achieved R^2 values above 0.80 for each task and identified key predictors such as pupil diameter, temporal-lobe functional connectivity, and task complexity. These findings further underscore the growing role of advanced physiological analytics in understanding surgeons' cognitive demands and optimizing training programs to improve overall performance.

Reliable model development also depends on high-quality, synchronised data, as demonstrated in a study (Hashemi *et al.* 2023) that outlined protocols for acquiring and processing da Vinci surgical data-combining stereoscopic endoscopic video with surgeon-movement recordings from depth cameras. The researchers further extracted event-level information such as clutch usage and instrument activation by means of computer-vision algorithms. Their open-source pipeline, together with a discussion of commercial recorders like dVLogger, provides a comprehensive framework for dataset development that supports reproducible and automated surgical performance evaluation.

The effectiveness of visual metrics was demonstrated in a study (Shafiei *et al.* 2023a), where participants performed robotic procedures on live pigs while wearing eye-tracking devices. Visual behavior during subtasks such as cold dissection and tissue retraction was analyzed and used to train a gradient boosting classifier, which achieved up to 96% accuracy across different tasks. Expert evaluations using GEARS served as the ground truth. These findings demonstrate that visual metrics collected during da Vinci-assisted procedures can reliably distinguish between varying skill levels, making eye-tracking a practical tool for real-time feedback and objective assessment.

The impact of automated feedback on learning was examined in a study (Brown and Kuchenbecker 2023) involving the da Vinci Standard system and a Smart Task Board. Trainees were divided into feedback and control groups, with the former receiving near-real-time GEARS score predictions generated by a regression-based machine learning model. While both groups improved in technical metrics such as force application and tool acceleration, the feedback group showed increased awareness of their performance gaps, suggesting that automated feedback may enhance reflective learning even if it does not immediately accelerate early psychomotor development.

Table 2 Studies on surgical skill assessment & workflow analysis

Study	Methodology/Algorithm	Modalities
(Egert *et al.* 2020)	AI/ML (e.g., Computer Vision, ANN)	Instrument motion, eye tracking, force sensors, video
(Lyman *et al.* 2021)	CUSUM + ML (using ORI)	Kinematics (dVLogger from da Vinci system)
(Brown and Kuchenbecker 2023)	Automated GEARS-based scoring	Robotic instrument motion data (Smart Task Board)
(Hashemi *et al.* 2023)	AI-ready data acquisition framework	Video, 3D movement, event data
(Shafiei *et al.* 2023a)	GB	Eye gaze data
(Shafiei *et al.* 2023b)	GB, RF, MLR	EEG and eye gaze data
(Shafiei *et al.* 2024a)	XGBoost	EEG and eye-tracking (gaze, pupil, connectivity metrics)
(Shafiei *et al.* 2024b)	GB, RF	EEG and eye-tracking during VUA
(Takács *et al.* 2024)	Exploratory ML + statistical correlation analysis	Heart rate variability, posture, hand motion sensors (digital wearables)
(Hatcher *et al.* 2025)	Review of frameworks and assessments	Simulation tools, assessment checklists

Abbreviations: AI/ML, Artificial Intelligence / Machine Learning; ANN, Artificial Neural Network; CUSUM, Cumulative Sum Control Chart; dVLogger, da Vinci Logger; EEG, Electroencephalogram; GB, Gradient Boosting; GEARS, Global Evaluative Assessment of Robotic Skills; ML, Machine Learning; MLR, Multinomial Linear Regression; ORI, Operative Robotic Index; RF, Random Forest; VUA, Vesicourethral Anastomosis; XGBoost, Extreme Gradient Boosting.

Skill progression over time was the focus of a study (Lyman *et al.* 2021) that investigated the learning curve of novice surgeons performing hepaticojejunostomy using the da Vinci Surgical System. Kinematic data, including camera adjustments and tool movements, were recorded and analyzed with machine learning and stepwise "combination of an objective cumulative sum" (CUSUM) techniques. A new composite score, the Operative Robotic Index (ORI), was introduced to summarize performance across key metrics. The ORI effectively differentiated novice from intermediate skill levels and aligned closely with traditional CUSUM evaluations. These findings highlight how detailed kinematic tracking, and statistical modeling can be used to monitor skill development over time.

The evolution of robotic simulation education was the subject of a recent review (Hatcher *et al.* 2025) that emphasized the growing need for standardized training curricula tailored to RAS platforms like the da Vinci Xi. While early efforts such as the Fundamentals of Robotic Surgery program provide a baseline, more comprehensive tools and frameworks, such as GEARS, R-OSATS, and RO-SCORE, are needed to assess advanced robotic surgical skills. The authors also anticipate a future where emerging technologies such as AI, virtual reality, and AR are fully integrated into training platforms, making the da Vinci system a key player in next-generation surgical education.

Collectively, these studies illustrate the powerful role of the da Vinci Surgical System as both a surgical and educational platform. By capturing high-resolution data streams and integrating physiological, visual, and motion-based metrics, the system supports the development of automated models that can objectively assess skill levels, monitor stress responses, and improve training outcomes. These capabilities contribute to a future where surgical education is more standardized, feedback is more immediate and objective, and patient outcomes are continuously enhanced through data-driven learning.

Challenges in Adapting Deep Learning to Robot-Assisted Surgery

Despite rapid progress in applying DL and other machine learning methods to RAS, significant barriers remain before these techniques can be seamlessly adopted in real operating rooms. This section distills key challenges around data infrastructure, real-time performance, algorithm generalizability, interpretability, workflow integration, and regulatory/ethical concerns.

Data Availability, Quality, and Standardization: A fundamental challenge is assembling high-quality, large-scale datasets of robot-assisted procedures. The da Vinci Surgical System can record textual, kinematic, and video data, but there is no single, standardized protocol across institutions (Hashemi *et al.* 2023). In some contexts, researchers gather adverse event reports (Li *et al.* 2024) or produce specialized datasets for skill assessment (Lyman *et al.* 2021), but these remain fragmented. Meanwhile, anonymization requirements and the labor-intensive nature of manual annotation, particularly in advanced tasks like EEG- and eye-gaze–based skill assessment (Shafiei *et al.* 2023a,b, 2024b), discourage broad data-sharing. Even successful attempts to fuse multiple data modalities (e.g., vision, kinematics, and textual records) can be hampered by incompatible data formats, sensor noise, or incomplete metadata (Wu *et al.* 2020). The net effect is that many DL models are trained on small, nonrepresentative samples, limiting generalizability.

Real-Time Performance and Resource Constraints: DL algorithms used for tasks such as AR overlays, surgical-phase detection, or skill evaluation must produce sub-second updates. For instance, sensorless haptic feedback systems (Yilmaz *et al.* 2024) or real-time stress measurement (Takács *et al.* 2024) require tight latency bounds for closed-loop control. Yet many robotics labs still rely on external graphics processing units or slower CPUs not specifically tuned to handle real-time inference. Even modest delays can disrupt surgical flow. Moreover, models of this class, such as the deep-transformer pipeline "Bert-BiLSTM" used for adverse-event classification (Li *et al.* 2024), typically require considerable compute resources. Balancing model complexity with real-time feasibility is an ongoing struggle.

Transfer Learning and Algorithm Generalizability: Neural networks often degrade when applied to new surgical environments, different endoscopic hardware, or novel tissue types (Egert *et al.* 2020; Hatcher *et al.* 2025). Similarly, skill classifiers that excel in simulator tasks (Rivero-Moreno *et al.* 2023) may stumble in real-world surgery with unfamiliar lighting conditions or tool usage patterns. Bridging simulation to reality remains a major hurdle. For instance, the realism of biomechanical simulations was improved by integrating real vision and kinematics data (Wu *et al.* 2020), although differences between controlled phantom data and anatomical variance in patients still limit direct transfer. Approaches to domain adaptation (ranging from augmented training sets to semi-supervised learning) are essential if we want robust models that generalize well across procedures and platforms.

Interpretability and Surgeon Acceptance: Black-box deep models trained on massive data can be difficult to interpret, thereby reducing surgeon trust. The necessity of transparent and easily digestible metrics for surgical performance has been emphasized in prior work (Egert *et al.* 2020). Similarly, it has been shown that while EEG or eye-tracking data can accurately gauge skill level, explaining why a model produces a particular skill rating remains challenging (Shafiei *et al.* 2023a,b, 2024b). Surgeons or medical trainees often rely on step-by-step guidance and immediate feedback; purely numerical or "black-box" predictions might be undervalued if they cannot be explained in a clinically meaningful manner. Fostering trust may require "explainable AI" methods (Setchi *et al.* 2020) to highlight relevant video frames, tissue edges, or motion trajectories that drive each inference.

Workflow Integration and Educational Gaps: Even when DL outputs are accurate, they can disrupt the robotic workflow if delivered through clunky interfaces or if they add cognitive load. Objective skill metrics such as the dVLogger-based ORI are presently reviewed post-hoc (Lyman *et al.* 2021); if such feedback were ever pushed into real time, an unfiltered stream of notifications could overwhelm surgeons. Likewise, the "Bert-BiLSTM" classifier developed to label da Vinci adverse-event reports remains a bench prototype (Li *et al.* 2024), and its clinical deployment would require careful alert management. At the same time, most residency programmes still lack dedicated AI or data-science instruction, creating a literacy gap that slows adoption (Hatcher *et al.* 2025; Egert *et al.* 2020). Reviews consequently stress the need for intuitive user-interface design and formal curricula that teach clinicians how to interpret and respond to algorithmic recommendations (Rivero-Moreno *et al.* 2023). Without such human–machine-interface refinements, even high-performing models may remain sidelined.

Regulatory, Ethical, and Liability Considerations: Regulatory frameworks for AI-driven surgical systems are still evolving (Egert *et al.* 2020). Tools that automate or partially automate tasks such as trocar placement or tissue retraction face extra scrutiny, since errors can have severe ramifications (Attanasio *et al.* 2020). Liability is another gray area (Marcus *et al.* 2024): if a trained model or algorithm misclassifies an adverse event or suggests an ill-advised maneuver, it is unclear whether the responsibility lies with the surgeon, the device manufacturer, or the hospital system. Data privacy is yet another consideration particularly for advanced analytics involving EEG, eye-gaze, or physiologic signals (Hashemi *et al.* 2023; Shafiei *et al.* 2024b). Without robust, standardized governance, it is challenging to deploy advanced AI functionalities at scale.

Proposed Conceptual Framework for a Multi-Modal Surgical Decision Support System

In light of previous studies (as discussed in the Results section) that focused on isolated aspects of robotic surgery, this study proposes a conceptual framework (Figure 4) that integrates these components into a unified, closed-loop system. The need for such a framework arises from the growing complexity of robotic surgical procedures and the demand for real-time, comprehensive decision support that enhances both surgical performance and training.

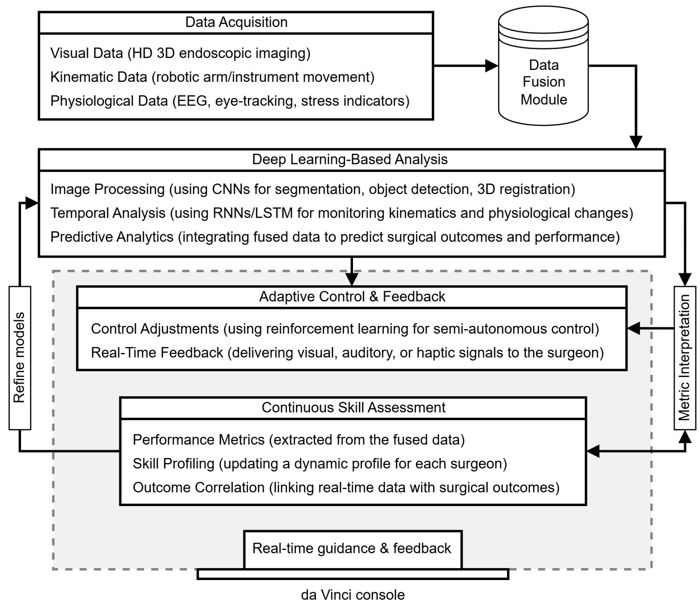

Figure 4 Proposed conceptual framework for the *da Vinci* system. This figure illustrates the interaction between the four core components: data acquisition and fusion, deep learning-based analysis, adaptive control and feedback, and continuous skill assessment. 3D, Three-Dimensional; CNN, Convolutional Neural Network; EEG, Electroencephalography; HD, High-Definition; LSTM, Long Short-Term Memory; RNN, Recurrent Neural Network.

Framework Components and Their Functions: The framework is composed of several interconnected components designed to capture, analyze, and act upon multi-modal data. Each component plays a vital role in ensuring that information is processed efficiently and utilized effectively to support surgical decision-making and continuous skill improvement.

The data acquisition and fusion component is critical for collecting and synchronizing diverse data sources to provide a comprehensive view of the surgical environment:

- **Visual Data:** HD 3D endoscopic imaging is captured by the da Vinci system. This data is used for AR overlays and real-time segmentation and registration of surgical sites.
- **Kinematic Data:** Detailed movement data from robotic arms and instruments, recorded via tools such as the dVLogger,

provides insights into instrument trajectories and motion patterns.

- **Physiological Data:** Wearable sensors collect metrics such as EEG, eye-tracking information, and stress indicators. This information reflects the surgeon's cognitive and emotional state during procedures.
- **Fusion Mechanism:** A central data fusion module synchronizes these diverse data streams into a unified dataset, ensuring that all modalities contribute to real-time analysis.

The deep learning-based analysis component leverages advanced DL models to interpret the fused data and extract actionable insights:

- **Image Processing:** CNNs are used to process visual data. These networks perform tasks such as object detection, segmentation, and 3D model alignment, thereby enhancing the visual context provided to the surgeon.
- **Temporal Analysis:** Recurrent neural network- or LSTM-based models analyze the kinematic and physiological data over time. This analysis supports dynamic skill assessment and monitors the surgeon's state throughout the procedure.
- **Predictive Analytics:** Machine learning algorithms integrate the fused data to predict surgical outcomes, provide alerts, and recommend adjustments based on historical performance and real-time trends.

The adaptive control and feedback component enables the system to dynamically adjust surgical parameters and provide real-time guidance based on continuous data analysis:

- **Control Adjustments:** Reinforcement learning algorithms enable the system to make semi-autonomous control adjustments. These adjustments can optimize instrument movements and camera positioning during surgery.
- **Real-Time Feedback:** The framework provides continuous feedback to the surgeon via the da Vinci console. This feedback may be visual, auditory, or haptic in nature, ensuring that the surgeon is informed of any deviations or potential improvements. Additionally, model interpretation (MI), such as through Shapley value-based approaches (Dean Pelegrina and Siraj 2024), helps the system deliver intelligible feedback (e.g., highlighting why a control adjustment is recommended).

The continuous skill assessment component focuses on evaluating and tracking surgeon performance over time to support targeted training and improved outcomes:

- **Performance Metrics:** Automated performance metrics are extracted from the fused data, allowing for an objective evaluation of surgical proficiency.
- **Skill Profiling:** This component continuously updates a skill profile for each surgeon. The profile is used to identify strengths, weaknesses, and trends in performance, which in turn supports targeted training interventions.
- **Outcome Correlation:** The framework correlates real-time performance data with surgical outcomes, contributing to a feedback loop that helps refine both the control system and the training protocols. Additionally, MI helps visually and numerically explain how specific data (e.g., instrument trajectories, EEG patterns) influence performance assessments. The two-sided connection provides an adaptive relationship between the component and MI (e.g., emphasizing temporal motion irregularities if fine motor control is a known issue). This transparency enables surgeons and educators to digest complex analytics in an intuitive and actionable manner.

Interactions Between Components: The proposed framework functions as an interconnected system in which each component reinforces and informs the others. In this system, the data acquisition module collects multi-modal information (including visual, kinematic, and physiological data) and feeds it into the DL-based analysis module, which interprets these data and subsequently informs the adaptive control system that provides real-time feedback to the surgeon. Meanwhile, the continuous skill assessment module monitors performance metrics and updates the surgeon's skill profile, which in turn is used to refine control strategies and training protocols. By integrating these diverse data streams into a unified framework, this approach bridges gaps observed in previous research, enhancing intraoperative decision-making and fostering a continuous learning environment that ultimately improves surgical outcomes and training efficiency. Additionally, the system integrates MI-based explanations that help visualize how specific inputs influence assessment outcomes. These explanations support clearer decision-making, promote user understanding of AI-driven suggestions, and contribute to more informed adjustments during both training and live procedures.

CONCLUSION

This study systematically reviewed recent advances in integrating DL with the da Vinci Surgical System, highlighting key innovations in autonomous instrument control, real-time image-guided navigation, skill assessment, and sensorless haptic feedback. These developments show great promise, yet challenges persist in data standardization, real-time performance, domain adaptation, and model interpretability. To address these challenges, the review organized recent DL applications into a comprehensive conceptual framework that supports real-time feedback, adaptive control, and continuous skill assessment, providing a foundation for future development and validation in RAS. As DL advances alongside sensing technologies, user-interface research, and evolving regulatory guidance, the framework can steer innovation toward safer, more autonomous, and intelligent surgical systems that improve both clinical practice and surgical education.

Although this preliminary framework is an important step toward a multi-modal DSS for RAS, further research and rigorous validation remain essential to assess real-world feasibility, optimize each component, and fully unlock the benefits of this integrated approach. The included studies vary methodologically, and many were conducted in simulated or controlled settings rather than operating rooms. Future work should prioritize clinical validation through well-designed prospective trials and enhance model transparency to foster surgeon trust and ensure robustness under real-world conditions. Ongoing updates to the da Vinci platform must be tracked to keep the framework compatible and clinically relevant. Progress will require close collaboration among engineers, surgeons, data scientists, and regulators to translate laboratory prototypes into safe, widely adopted clinical tools.

Acknowledgments

This study received no funding, grants, or other external support. An abstract based on the literature review component of this study was presented as an oral presentation at the 2nd *Biyobilim Buluşmaları*, held on April 17–18, 2025.

Availability of data and material

The analyses presented in this study are based on previously published literature, all of which is fully cited within the article. No new datasets were generated or used.

Conflicts of interest

The author declares that there is no conflict of interest regarding the publication of this paper.

Ethical standard

The author has no relevant financial or non-financial interests to disclose. This study did not involve human participants or animal experimentation.

LITERATURE CITED

Amparore, D., M. Sica, P. Verri, F. Piramide, E. Checcucci, *et al.*, 2024 Computer Vision and Machine-Learning Techniques for Automatic 3D Virtual Images Overlapping During Augmented Reality Guided Robotic Partial Nephrectomy. Technology in Cancer Research & Treatment **23**: 15330338241229368.

Attanasio, A., B. Scaglioni, M. Leonetti, A. F. Frangi, W. Cross, *et al.*, 2020 Autonomous Tissue Retraction in Robotic Assisted Minimally Invasive Surgery – A Feasibility Study. IEEE Robotics and Automation Letters **5**: 6528–6535.

Azizian, M., M. Liu, I. Khalaji, J. Sorger, D. Oh, *et al.*, 2020 The da Vinci Surgical System. In *Handbook of Robotic and Image-Guided Surgery*, pp. 39–55, Elsevier.

Boal, M. W. E., D. Anastasiou, F. Tesfai, W. Ghamrawi, E. Mazomenos, *et al.*, 2024 Evaluation of objective tools and artificial intelligence in robotic surgery technical skills assessment: a systematic review. British Journal of Surgery **111**: znad331.

Brown, J. D. and K. J. Kuchenbecker, 2023 Effects of automated skill assessment on robotic surgery training. The International Journal of Medical Robotics and Computer Assisted Surgery **19**: e2492.

Cannizzaro, D., M. Scalise, C. Zancanella, S. Paulli, S. Peron, *et al.*, 2024 Comparative Evaluation of Major Robotic Systems in Microanastomosis Procedures: A Systematic Review of Current Capabilities and Future Potential. Brain Sciences **14**: 1235.

Celotto, F., N. Ramacciotti, A. Mangano, G. Danieli, F. Pinto, *et al.*, 2024 Da Vinci single-port robotic system current application and future perspective in general surgery: A scoping review. Surgical Endoscopy **38**: 4814–4830.

Cepolina, F. and R. P. Razzoli, 2022 An introductory review of robotically assisted surgical systems. The International Journal of Medical Robotics and Computer Assisted Surgery **18**: e2409.

Dean Pelegrina, G. and S. Siraj, 2024 Shapley Value-Based Approaches to Explain the Quality of Predictions by Classifiers. IEEE Transactions on Artificial Intelligence **5**: 4217–4231.

D'Ettorre, C., A. Mariani, A. Stilli, F. Rodriguez Y Baena, P. Valdastri, *et al.*, 2021 Accelerating Surgical Robotics Research: A Review of 10 Years With the da Vinci Research Kit. IEEE Robotics & Automation Magazine **28**: 56–78.

DiMaio, S., M. Hanuschik, and U. Kreaden, 2011 The da Vinci Surgical System. In *Surgical Robotics*, edited by J. Rosen, B. Hannaford, and R. M. Satava, pp. 199–217, Springer US, Boston, MA.

Egert, M., J. E. Steward, and C. P. Sundaram, 2020 Machine Learning and Artificial Intelligence in Surgical Fields. Indian Journal of Surgical Oncology **11**: 573–577.

Gonzalez, G. T., U. Kaur, M. Rahman, V. Venkatesh, N. Sanchez, *et al.*, 2021 From the Dexterous Surgical Skill to the Battlefield—A Robotics Exploratory Study. Military Medicine **186**: 288–294.

Hashemi, N., M. B. S. Svendsen, F. Bjerrum, S. Rasmussen, M. G. Tolsgaard, *et al.*, 2023 Acquisition and usage of robotic surgical data for machine learning analysis. Surgical Endoscopy **37**: 6588–6601.

Hatcher, A. J., B. T. Beneville, and M. M. Awad, 2025 The evolution of surgical skills simulation education: Robotic skills. Surgery **181**: 109173.

Huber, M., S. Ourselin, C. Bergeles, and T. Vercauteren, 2022 Deep homography estimation in dynamic surgical scenes for laparoscopic camera motion extraction. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization **10**: 321–329.

Iftikhar, M., M. Saqib, M. Zareen, and H. Mumtaz, 2024 Artificial intelligence: revolutionizing robotic surgery: review. Annals of Medicine & Surgery **86**: 5401–5409.

Khan, A., H. Yang, D. R. S. Habib, D. Ali, and J. Y. Wu, 2025 Development of a machine learning-based tension measurement method in robotic surgery. Surgical Endoscopy .

Kinoshita, T., R. Sato, E. Akimoto, Y. Tanaka, T. Okayama, *et al.*, 2022 Reduction in postoperative complications by robotic surgery: a case–control study of robotic versus conventional laparoscopic surgery for gastric cancer. Surgical Endoscopy **36**: 1989–1998.

Knudsen, J. E., U. Ghaffar, R. Ma, and A. J. Hung, 2024 Clinical applications of artificial intelligence in robotic surgery. Journal of Robotic Surgery **18**: 102.

Koh, D. H., W. S. Jang, J. W. Park, W. S. Ham, W. K. Han, *et al.*, 2018 Efficacy and Safety of Robotic Procedures Performed Using the da Vinci Robotic Surgical System at a Single Institute in Korea: Experience with 10000 Cases. Yonsei Medical Journal **59**: 975.

Koukourikis, P. and K. H. Rha, 2021 Robotic surgical systems in urology: What is currently available? Investigative and Clinical Urology **62**: 14.

Li, J., X. Gong, P. Li, L. Xiao, X. Chang, *et al.*, 2022 Application of Da Vinci robotic surgery system in cervical cancer: A single institution experience of 557 cases. Asian Journal of Surgery **45**: 707–711.

Li, T., W. Zhu, W. Xia, L. Wang, W. Li, *et al.*, 2024 Research on adverse event classification algorithm of da Vinci surgical robot based on Bert-BiLSTM model. Frontiers in Computational Neuroscience **18**: 1476164.

Lyman, W. B., M. J. Passeri, K. Murphy, I. A. Siddiqui, A. S. Khan, *et al.*, 2021 An objective approach to evaluate novice robotic surgeons using a combination of kinematics and stepwise cumulative sum (CUSUM) analyses. Surgical Endoscopy **35**: 2765–2772.

Marcus, H. J., P. T. Ramirez, D. Z. Khan, H. Layard Horsfall, J. G. Hanrahan, *et al.*, 2024 The IDEAL framework for surgical robotics: development, comparative evaluation and long-term monitoring. Nature Medicine **30**: 61–75.

Page, M. J., J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, *et al.*, 2021 The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ p. n71.

Panesar, S., Y. Cagle, D. Chander, J. Morey, J. Fernandez-Miranda, *et al.*, 2019 Artificial Intelligence and the Future of Surgical Robotics. Annals of Surgery **270**: 223–226.

Pugin, F., P. Bucher, and P. Morel, 2011 History of robotic surgery : From AESOP® and ZEUS® to da Vinci®. Journal of Visceral Surgery **148**: e3–e8.

Reddy, K., P. Gharde, H. Tayade, M. Patil, L. S. Reddy, *et al.*, 2023 Advancements in Robotic Surgery: A Comprehensive Overview of Current Utilizations and Upcoming Frontiers. Cureus .

Rivero-Moreno, Y., S. Echevarria, C. Vidal-Valderrama, L. Stefano-Pianetti, J. Cordova-Guilarte, *et al.*, 2023 Robotic Surgery: A Comprehensive Review of the Literature and Current Trends. Cureus .

Saman, A., 2024 Next-Gen Surgery: AI Robots Leading the Way in

Healthcare. Archives of Community Medicine and Public Health **10**: 013–017.

Setchi, R., M. B. Dehkordi, and J. S. Khan, 2020 Explainable Robotics in Human-Robot Interactions. Procedia Computer Science **176**: 3057–3066.

Shafiei, S. B., S. Shadpour, and J. L. Mohler, 2024a An Integrated Electroencephalography and Eye-Tracking Analysis Using eXtreme Gradient Boosting for Mental Workload Evaluation in Surgery. Human Factors: The Journal of the Human Factors and Ergonomics Society p. 00187208241285513.

Shafiei, S. B., S. Shadpour, J. L. Mohler, K. Attwood, Q. Liu, *et al.*, 2023a Developing Surgical Skill Level Classification Model Using Visual Metrics and a Gradient Boosting Algorithm. Annals of Surgery Open **4**: e292.

Shafiei, S. B., S. Shadpour, J. L. Mohler, P. Rashidi, M. S. Toussi, *et al.*, 2024b Prediction of Robotic Anastomosis Competency Evaluation (RACE) metrics during vesico-urethral anastomosis using electroencephalography, eye-tracking, and machine learning. Scientific Reports **14**: 14611.

Shafiei, S. B., S. Shadpour, J. L. Mohler, F. Sasangohar, C. Gutierrez, *et al.*, 2023b Surgical skill level classification model development using EEG and eye-gaze data and machine learning algorithms. Journal of Robotic Surgery **17**: 2963–2971.

Steffens, D., K. E. McBride, N. Hirst, M. J. Solomon, T. Anderson, *et al.*, 2023 Surgical outcomes and cost analysis of a multi-specialty robotic-assisted surgery caseload in the Australian public health system. Journal of Robotic Surgery **17**: 2237–2245.

Takács, K., E. Lukács, R. Levendovics, D. Pekli, A. Szijártó, *et al.*, 2024 Assessment of Surgeons' Stress Levels with Digital Sensors during Robot-Assisted Surgery: An Experimental Study. Sensors **24**: 2915.

Ullah, Z., S. Roy, S. Muhammad, C. Yu, H. Huang, *et al.*, 2024 Fluorescence imaging-guided surgery: current status and future directions. Biomaterials Science **12**: 3765–3804.

Wu, J. Y., P. Kazanzides, and M. Unberath, 2020 Leveraging vision and kinematics data to improve realism of biomechanic soft tissue simulation for robotic surgery. International Journal of Computer Assisted Radiology and Surgery **15**: 811–818.

Yilmaz, N., B. Burkhart, A. Deguet, P. Kazanzides, and U. Tumerdem, 2024 Enhancing robotic telesurgery with sensorless haptic feedback. International Journal of Computer Assisted Radiology and Surgery **19**: 1147–1155.

How to cite this article: Özcan, H. Deep Learning in Robot-Assisted Surgery: A Conceptual Framework for the da Vinci System. *Computers and Electronics in Medicine*, 2(2), 26-35, 2025.