npj Digital Medicine

https://doi.org/10.1038/s41746-026-02660-z

Article in Press

Empowering surgeons with integrated synthetic data: solutions for mastering complex clinical scenarios

Received: 27 October 2025

Accepted: 12 April 2026

Cite this article as: Sakref, Y., Seenivasan, L., Ding, H. *et al.* Empowering surgeons with integrated synthetic data: solutions for mastering complex clinical scenarios. *npj Digit. Med.* (2026). https://doi.org/10.1038/s41746-026-02660-z

Yann Sakref, Lalithkumar Seenivasan, Hao Ding, Ruhika Iyer, Danush Kumar Venkatesh, Stefanie Speidel, Mathias Unberath, Jeffrey K. Jopling & Lisa Marie Knowlton

Empowering Surgeons with Integrated Synthetic Data: Solutions for Mastering Complex Clinical Scenarios*

Yann Sakref,[1] Lalithkumar Seenivasan,[2] Hao Ding,[2] Ruhika Iyer,[2] Danush Kumar Venkatesh,[3] Stefanie Speidel,[3] Mathias Unberath,[2] Jeffrey K. Jopling,[2] and Lisa M. Knowlton[1]

[1] *Stanford University, Department of General Surgery, Stanford, California*
[2] *Johns Hopkins University, Department of Computer Science, Baltimore, Maryland*
[3] *Department of Translational Surgical Oncology, NCT/UCC Dresden, a partnership between DKFZ, Faculty of Medicine and University Hospital Carl Gustav Carus, TUD Dresden, HZDR, Germany*

ABSTRACT

Synthetic data generation across domains can bridge gaps between visual training, skill development, and personalized surgical planning, ultimately transforming how surgeons and artificial intelligence (AI) systems prepare for the complexities of the operating room. In this Perspective, we explore applications of synthetic data to advance surgical education and AI across three key areas: visual data synthesis for training surgeons and AI systems, surgical simulation for skill development and robotics, and digital twins for patient-specific surgical planning and guidance. These domains have largely remained siloed, but their integration has potential to transform surgical training and AI development across the entire surgical workflow. To fully realize this potential, synthetic data must extend beyond routine surgical scenarios to anticipate atypical anatomy and intraoperative complications—the high-stakes clinical scenarios where enhanced training and AI support are most critical.

INTRODUCTION

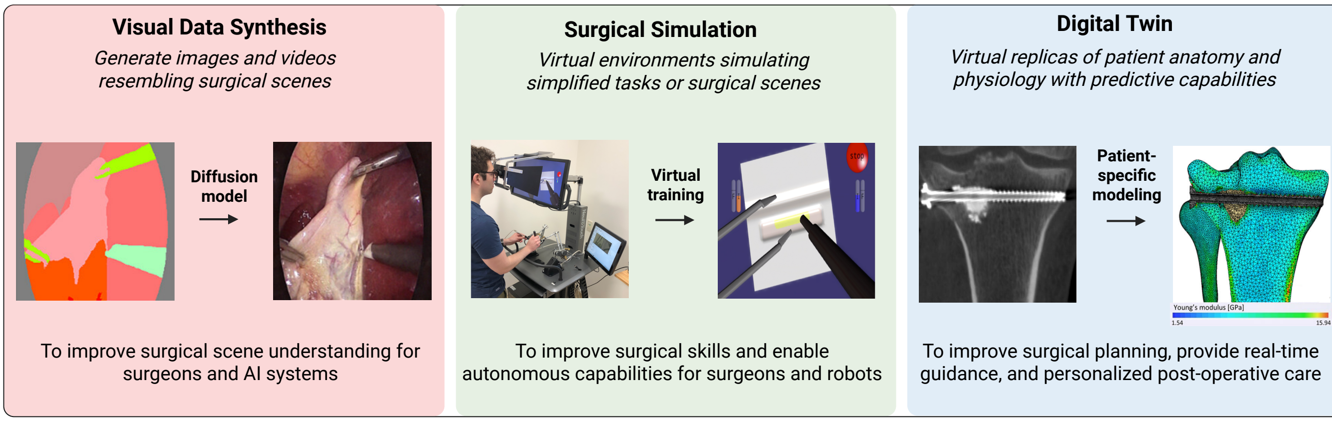
Surgery faces inherent limitations in providing comprehensive exposure to the full spectrum of surgical scenarios, constraining both trainees' education [1, 2] and artificial intelligence (AI) system development [3]. Synthetic data—algorithmically generated data that mimics the statistical, physical, and mechanistic properties of real-world data [4]—is vital to address this challenge and enhance the development of proficient surgeons [1, 5] and robust AI systems [4, 6]. Its transformative potential in surgery is evident across three key domains: visual data synthesis generates realistic surgical images and videos for AI training and surgical education [7–13]; surgical simulations create 3D virtual environments for safe and reproducible training of both human surgeons and robotic systems [14–19]; and digital twins provide patient-specific synthetic models for surgical planning and real-time intraoperative guidance [20–22].

Over the last decade, applications in each domain have progressed and expanded significantly. These domains, however, have mostly remained independent, both in current research that often focuses on specific downstream applications and in existing reviews [22]. This creates two significant issues: first, siloed approaches limit the synergies that could otherwise emerge between these domains, and second, they can perpetuate the perception that synthetic data serves only singular purposes—e.g., for training AI models or for digital twins. In this Perspective, we develop an integrated view of synthetic data for surgical applications. We highlight how a unified ecosystem can transform surgery across the entire workflow, providing unprecedented opportunities for advancing surgical training and developing AI and robotic systems to assist surgeons.

This integrated approach surfaces current critical limitations and future opportunities for the development of surgical synthetic data. While synthetic data predominantly captures routine surgical scenarios, more complex scenarios are largely unaddressed. These include time critical operations due to bleeding or hemodynamic instability, and technically complex cases involving aberrant anatomy, significant inflammation, or surgical planes altered by scar tissue. These are the exact intraoperative scenarios where surgical complications, including misidentification of anatomy or injury to critical structures, are most likely to occur, making it imperative for surgeons to develop skills to navigate these technical situations safely [2, 23, 24]. To date, the development of robust computer vision models and robots for intraoperative navigation and assistance during complex or time critical operations has also been limited, but this represents an important area of development where surgeons are most in need of support [25]. Synthetic data has the potential to generate such challenging cases—as scenes with massive bleeding [26, 27]—where real-world data is scarce and could significantly enhance surgical training, decision-making, and patient outcomes.

In this Perspective, we begin by presenting a few representative applications of synthetic data in surgery and a *vue d'ensemble* of these domains, highlighting their interconnectedness and impact across surgical care. We then discuss each sub-domain in more detail, exploring current approaches, their limitations, and their potential for

*Corresponding author: Yann Sakref (ysakref@stanford.edu)



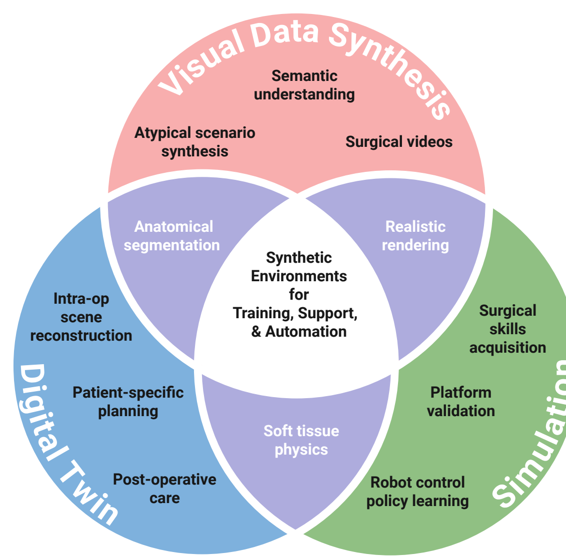
addressing complex scenarios that surgeons encounter in clinical practice. Finally, we examine how viewing these domains as an integrated ecosystem reveals both shared technical challenges and synergistic solutions, and how the generation of complex and atypical clinical scenarios is the current critical frontier that could unlock synthetic data's full potential.

I. CLINICAL APPLICATIONS OVERVIEW

We illustrate in Figure three applications of synthetic data to surgery. First, synthetic data can be used for visual data synthesis, i.e., to generate images and videos of surgical scenes. In the example of Figure (A), we show how diffusion models can generate synthetic surgical images from segmentation maps (mapping the position of organs and tools), creating photorealistic and automatically annotated scenes for laparoscopic procedures [28]. These are used for training AI computer vision algorithms in intra-operative critical structure recognition and delineation (semantic segmentation). Visual data synthesis also encompasses the generation of surgical videos and of atypical scenarios (Figure I).

A second application of synthetic data is the creation of surgical simulations. In the example shown in Figure (B), a trainee performs a simplified subpial tumor resection using haptic-enabled instruments that provide realistic tactile feedback. The simulation allows precise tissue manipulation with assessment of surgical performance [29]. These virtual environments enable surgical trainees to practice procedures repeatedly in risk-free situations, allowing for skill acquisition without patient safety concerns. These platforms can also be used for robot training and surgical tool validation (Figure I).

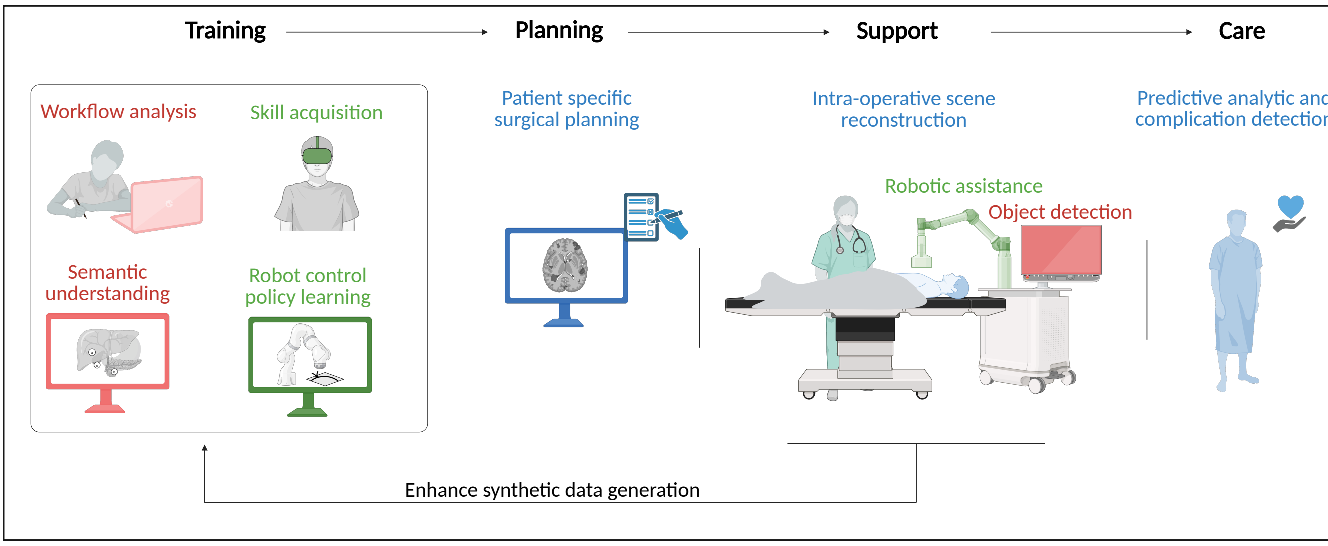
The third application of synthetic data is through digital twins, which are patient-specific synthetic models that replicate individual patient anatomy and physiology. In the example of Figure (C), we illustrate how 3D clinical imaging is transformed into patient-specific computational models with heterogeneous bone properties, enabling simulation of multiple surgical scenarios



to optimize treatment planning and predict healing outcomes [30]. This predictive modeling allows clinicians to assess different treatment approaches and customize surgical strategies based on individual patient anatomy and bone quality. Digital twins also support intra-operative guidance and post-operative care (Figure I).

As shown in Figure I, these domains should not be excessively compartmentalized, as the technical foundations underlying each one create natural synergies that amplify their collective impact. For example, advances in semantic segmentation algorithms through Visual Data Synthesis directly enhance Digital Twins' ability to create accurate patient-specific anatomical models, while Surgical Simulations' physics engines inform Visual Data Synthesis about realistic tissue deformation and instrument interaction patterns through realistic rendering techniques. Similarly, both Surgical Simulations and Digital Twins converge on the need for excellent soft tissue physics.

Moreover, as illustrated in Figure I, beyond these technical synergies, there exist clinical synergies, as synthetic



data is creating a transformation across the entire surgical workflow, from training to post-operative care. For instance, synthetic data can be used to train more robust AI systems—AI systems that will be used during the intra-operative phase for detecting critical anatomical structures and updating the digital twin. Similarly, there exists potential for powerful continuous feedback loops where intra-operative data from surgeries—including failure and complex cases—can inform better synthetic data generation and more comprehensive training in virtual environments. The result is an emerging ecosystem where synthetic data becomes fundamental to precise and robust surgical practice.

II. CURRENT APPROACHES ACROSS SURGICAL DOMAINS

A. Visual Data Synthesis

Semantic scene understanding, meaning the detection and delineation of anatomical structures and instruments, is a key application area for synthetic data. By generating large amounts of annotated surgical scenes, synthetic data eliminates the manual annotation burden while providing ready-to-use datasets that improve AI model precision. Two primary approaches have emerged. The first approach starts with 3D CT scans of anatomical structures (e.g. gallbladder, liver, pancreas) and reconstructs surgical environments by manually adding virtual meshes representing connective tissues. These scenes are then transformed into photorealistic surgical views using Generative Adversarial Networks (GANs) [7, 8] or diffusion models [9] that retexture the image based on real surgical footage. The second approach, shown in Figure (A), directly leverages diffusion models to generate photorealistic surgical scenes, guided by simple semantic masks that specify where anatomical structures and instruments should appear [11, 31, 32].

The addition of synthetic data to real datasets has been reported to improve segmentation accuracy by 4–16% in laparoscopic tasks [7, 9, 32, 33]. In a task-specific instance segmentation study on robotic suturing scenes, Leoncini *et al.* report that models trained on synthetic data alone achieved a Dice score of 0.122 on real images, whereas combining synthetic data with a limited number of real images achieved a Dice of 0.920 [34]. These results suggest that while synthetic augmentation can be beneficial, synthetic data alone remains insufficient if it does not adequately capture real-world variability. When diffusion-based methods simply replicate existing data, they may offer limited benefit to segmentation accuracy [11, 28]. Effective synthetic data generation therefore requires introducing meaningful surgical scene diversity rather than reproducing the statistical distribution of the training data. The most successful approaches employ structured variation—whether through systematic randomization of camera angles and tool positions as in Pfeiffer *et al.* [7] or compositional approaches that create statistically different scenes as in Venkatesh *et al.* [32].

These current methods demonstrate that visual data synthesis is rapidly becoming an effective tool for training more accurate surgical AI algorithms in semantic understanding. Key opportunities for advancement include studying how to systematically engineer the crucial diversity (through organ shape and texture, lighting, blood, etc.) needed to improve performance. Additionally, generating complex anatomical scenarios and edge cases represents an underexplored area that could significantly benefit robust surgical AI training.

Surgical video generation has shown great promise and increased realism. Recent studies demonstrate realistic video sequences representing distinct phases and actions in procedures such as laparoscopic cholecystectomy [10, 12, 13], endoscopy [35, 36], and robot-assisted radical prostatectomy [37]. A primary goal is surgical education, enabling trainees to create tailored video clips.

These videos are typically generated using video diffusion models trained on temporally annotated datasets, with labels at the phase or action level, or even action triplets—structured as "instrument, verb, target" (e.g., "hook dissect cystic duct") [12, 13]. Evaluation of the video's realism is often both qualitative, through expert surgeon review, and quantitative, through measuring the statistical similarity between real and generated videos. Recent studies show striking realism: expert surgeons could not distinguish synthetic from real clips in one-third of cases when generated from action triplets [12], with complete indistinguishability achieved when models use real surgical frames as starting points [36].

Video generation of surgical action and phase is rapidly evolving driven by significant advances in generative video models. Current challenges remain largely technical, such as extending video length, improving temporal and semantic consistency, and accurately simulating complex behaviors like multilayer dissection, though clinical validation also remains limited. Notably, video generation models produce less variation than real surgical footage, capturing dominant features at the cost of natural variability [10], hinting that a similar problem as in image generation lies here.

Atypical scenario synthesis, the generation of complex anatomical cases or adverse intra-operative conditions, is an application where synthetic data will have a critical clinical impact for both AI development and surgical training.

A few notable works have begun to explore these scenarios. Several studies focus on simulating intra-operative bleeding. For example, Mohamadipanah *et al.* [26] present a proof-of-concept in which image textures are modified using GANs to simulate massive bleeding during minimally invasive lobectomy—demonstrating that synthetic frames can realistically represent rare complications. Similarly, Venkatesh *et al.* [37] use synthetic video clips to significantly improve the prediction of staple line bleeding during prostatectomy. These works show synthetic data can modify existing images or generate full bleeding sequences—both clinically significant but underrepresented in datasets. Other approaches address visibility challenges from intra-operative smoke through "desmoking" methods that reconstruct clear views from obscured footage [38–40].

Beyond complications and visibility issues, generative models can simulate imperfect surgical views. Nwoye *et al.* [12] discuss that video diffusion models can be guided to reproduce sub-optimal camera angles—an important feature in surgical education, where training under difficult perspectives has been shown to reduce the learning curve [41].

Together, these studies highlight a possible shift in visual data synthesis: from generating generic scenes at scale to synthesizing rare or challenging surgical scenarios with high clinical relevance. Although still underexplored, this direction is likely to become increasingly important as generative models achieve finer control and deeper surgical understanding.

B. Surgical Simulation

Robot control policy learning for surgery aims to train robotic systems to perform actions ranging from sub-tasks like grasping and suturing to complete autonomous surgical procedures like clipping and cutting the cystic artery [42, 43] or autonomously navigating catheters and guidewires through vascular anatomy [44, 45]. The clinical motivation is multifaceted: reducing surgeon fatigue, facilitating remote surgery in underserved areas, and providing surgical capabilities where expert surgeons are unavailable. While real-world robot policy learning has shown success [42, 43], simulation offers compelling advantages: unlimited repetition, reduced costs, and critically, access to rare or critical cases absent from real training scenarios.

The main challenge of policy learning is transferring a robot's skills from simulation environments to the real world. This is called sim-to-real transfer, and requires simulations to be particularly faithful, both in terms of tissue physics and visual realism. In laparoscopic surgery, multiple studies have successfully demonstrated this transfer to physical dVRK systems [14–17, 46, 47], but with limited success, as state-of-the-art simulations achieve only 50% success rates for suture needle lift tasks [48]. Beyond laparoscopic robotics, sim-to-real transfer has also been explored in endovascular surgery. Open-source simulation platforms built on MuJoCo [44] and SOFA [45, 49] enable reinforcement learning for autonomous catheter and guidewire navigation. Simulation-trained controllers have transferred to physical phantom test benches for single-device tasks [45], and domain randomization has further supported deployment in real robotic systems [49]. However, measurable performance degradation remains after transfer, highlighting the persistent domain gap between simulation and real vascular environments [45, 49].

To improve sim-to-real transfer across surgical domains, more realistic and faster physics engines are being developed [18, 19, 48, 50, 51], evolving from oversimplified grasping rules to sophisticated soft tissue simulation with neural-augmented physics [50, 52] or GPU-accelerated physics [48]. Visual realism advances have been more modest, with examples like SuFIA-BC integrating CT-derived organ models to better prepare systems for clinical use [14].

Beyond improving sim-to-real transfer, current simulations still focus on subtasks rather than full procedures and operate under predictable, controlled conditions rather than the dynamic variability surgeons encounter.

Addressing these technical challenges represents a critical research priority, as simulation training holds tremendous potential to train robotic systems at scale on the adversarial and atypical cases that are essential for robust real-world performance.

Surgical skill acquisition is another exciting application of synthetic data, whereby surgeons can learn and master new skills with enhanced safety, repeatability, and reduced costs. Virtual simulation has already demonstrated remarkable efficacy, reducing operation time by half with successful skill transfer to real-world operation for laparoscopic salpingectomy [53, 54], and leading to a 38% reduction in posterior capsule rupture rates among trainees for cataract surgery [55]. Importantly, the benefits of simulation training extend globally, producing significant improvements in procedural understanding compared to textbook-based learning in resource-limited settings such as Rwanda [56].

The enthusiasm for virtual simulation extends beyond clinical outcomes to learner acceptance. A study at Yale University found that 91% of residents agreed that simulation training should be mandatory [57]. However, trainees noted limitations in graphics realism and haptic feedback, highlighting the importance of advancing VR synthetic data technology to address these concerns.

In that regard, contemporary platforms showcase more sophisticated haptic feedback capabilities that directly address these realism limitations. FIVR and AMBF+ for skull-base surgery exemplify these advancements through realistic drill vibrations that match actual surgical instruments and integrated multi-modal feedback systems [58, 59].

Despite these advances, two main challenges limit current virtual simulation effectiveness. First, many platforms focus on subtasks rather than comprehensive procedural training. Second, as for robot policy learning, current platforms operate within predictable and controlled conditions, highlighting an opportunity for synergistic technical solutions that could benefit both domains simultaneously.

Platform validation of surgical tools, techniques, and workflows through computer simulation has become increasingly important in medical technology development. These virtual testing environments enable extensive experimentation and refinement without endangering patients, while identifying potential problems with designs or procedures before they reach clinical settings.

The VTEST platform illustrates this by enabling development and testing of NOTES (Natural Orifice Transluminal Endoscopic Surgery), allowing surgeons to test this new technique and associated instruments in controlled, high-fidelity settings before moving to animal studies [60, 61]. For device validation, AR systems like LapAR™ enable rapid prototyping of new laparoscopic devices [62], while Cartucho et al. used simulation to validate a precise tracking system for laparoscopic instruments, testing performance across different conditions before confirming sub-millimeter tracking precision in ex vivo experiments [63].

Regulatory agencies increasingly recognize these methods' value, with computer-based trials gaining acceptance as supplements to conventional preclinical testing [64]. However, most simulation-based validations have yet to achieve full clinical implementation, and systematic comparisons with real-world outcomes remain limited. As these platforms advance in fidelity, they may increasingly function as legitimate frameworks supporting regulatory approval and clinical adoption.

C. Digital twin

Patient specific surgical planning represents a paradigm shift toward personalized medicine in surgery. In their simplest form, digital twins for surgical planning provide anatomical visualization. For example, in partial nephrectomies, detailed 3D models of the kidney, tumor, and vasculature enable surgeons to determine optimal cutting planes [21]. More advanced applications rely heavily on synthetic data through physical simulation, creating virtual scenarios and predictive outcomes that extend beyond medical images [22].

One key such application involves testing surgical approaches through simulations, enabling surgeons to compare different techniques and predict their outcomes in individual patients. Beyond the tibial plateau fracture example discussed in Figure (C), where different scenarios of tibial reconstruction can be tested [30], this approach extends across specialties: digital twins simulate fracture risks in cancer patients during vertebroplasty procedures [65], in endovascular surgery, simulation environments have been evaluated against physical phantoms for catheter shape and force sensing as well as tool–tissue interaction modeling [66], while cardiovascular surgery employs biomechanical models to simulate transcatheter aortic valve replacement, exploring different prostheses and implant depths [67].

Beyond simulation, digital twins can predict complications. CardioVision, for example, reconstructs three-dimensional models of the aortic root from CT imaging data to analyze aortic stenosis patients. By analyzing their calcification metrics, the system enables clinicians to predict adverse events and select optimal surgical approaches [68].

While digital twins for planning show significant promise, several challenges must be addressed for widespread clinical translation beyond anatomical visualization. Patient-specific parameter identification represents a fundamental hurdle, as individual variations in tissue properties are difficult to quantify non-invasively. Simply incorporating broad parameter ranges may address this variability, but reduce the model's precision and risk overlooking the precise

Synthetic Data Use Case	Surgical Applications	Illustrative Study	Key Challenge
Semantic understanding	Improving object detection and segmentation for intra-operative assistance	Generate thousands of fully annotated LapChole scene images [11]	Generating images significantly different from the training data ("out of its distribution")
Surgical video generation	Training residents on surgical procedure steps	Generate phase-action-triplets videos to teach residents specific endoscopic procedures [12]	Accurately generating longer and more complex actions
Atypical scenario synthesis	Improving vision models' robustness to adversarial surgical conditions	Add massive bleeding on images of minimally invasive lobectomy [26]	Require clinical expertise in designing and evaluating synthetic complications
Robot control policy learning	Training autonomous surgical systems for assistance	Train a dVRK robot in simulation to perform real suturing [15]	Improving simulation to real-world policy transfer
Surgical skill acquisition	Training residents to using surgical robots	Realistic VR environment for skull-based surgery with haptic feedback [52]	Improving scenarios' realism and diversity
Platform validation	Introducing new surgical techniques	Introduce surgeon to NOTES [54]	Improving fidelity and verify clinical validity
Patient specific surgical planning	Personalize surgical approaches for individual patients	Determine the best stabilization method following tibial plateau fractures [30]	Determining patient-specific parameters before incision
Intra-op scene reconstruction	Real-time surgical navigation and guidance	Offer real-time surgical guidance with optical tracking in skull based surgery [63]	Improving soft-tissue physics for real-time reconstruction and registration
Post-operative care	Monitor patient recovery and complications	Forecast complication after endovascular aneurysm repair [60]	Accurately distinguishing between normal recovery variations and clinically significant complications

patient-specific characteristics that define complex cases.

Real-time surgical scene reconstruction and guidance during the intra-operative phase requires digital twins that continuously synchronize with rapidly changing surgical conditions. Unlike preoperative planning, which relies on offline digital twins constructed once from patient data, intra-operative applications demand online systems that update based on continuous data streams from surgical sensors, endoscopic video, and physiological monitors [69]. These systems must balance model accuracy with computational speed while capturing and predicting dynamic procedural conditions in real-time.

A promising application of online digital twins is enabling surgeons to maintain accurate spatial awareness as anatomy changes throughout procedures. Twin-S demonstrates this approach in phantom studies by providing real-time guidance during skull base surgery [70]. The system uses high-precision optical tracking to generate live digital replicas of the surgical field and in real-time the virtual patient anatomy model. Similar approaches enable holographic augmented reality guidance for liver tumor puncture by continuously updating virtual models to compensate for respiratory motion [71].

Beyond real-time tracking, the most sophisticated applications aim for predictive modeling and autonomous assistance that could anticipate surgical complications before they occur. Intelligent twins could dynamically map optimal tissue-resection pathways using biomechanical modeling to avoid inadvertent damage to critical structures, while patient-specific hemodynamic models simulate blood flow alterations to provide predictive alerts for anticipated complications [22, 72]. These predictive capabilities are being integrated into surgical robotics, where continuously updated digital twins enable autonomous assistance by adapting to changing conditions and incorporating surgeon feedback [73], with emerging systems supporting natural speech-based interaction for real-time surgical assistance [74].

Key clinical translation challenges for intra-operative assistance stem from two primary issues: linking pre-operative data to intra-operative scenes and limited validation beyond controlled conditions. The first challenge involves registering pre-operative scans to real-time surgical environments [20, 51]. While successful for rigid structures like the skull, soft-tissue deformation remains problematic due to unknown tissue-specific mechanical properties and computational complexity of accurate tissue physics modeling. The second challenge concerns validation scope; current systems remain restricted to phantom experiments or small animal models with artificial tumors [70, 71]. The gap between controlled experimental validation and the anatomical diversity, pathological complexity, and procedural variability of real clinical practice remains a key barrier to widespread adoption.

Post-operative care represents a significant opportunity for digital twin technology to transform patient recovery through continuous monitoring and personalized management.

Digital twins have been explored for tailored postoperative care in orthopedic procedures, where biomechanical modeling guides rehabilitation strategies [30], and for forecasting complications like implant failure after interventions such as endovascular aneurysm repair [67, 75]. These systems can also generate patient-specific surgical documentation for proactive complication identification [21]. Current challenges include ensuring robust data integration from multiple monitoring devices and developing algorithms that distinguish between normal recovery variations and clinically significant complications.

III. INTEGRATED CHALLENGES AND OPPORTUNITIES

Our exploration of synthetic data across surgical domains, illustrated in Table I, reveals significant progress alongside persistent challenges related to realism and diversity: (i) Visual Data Synthesis can now generate large volumes of annotated data at low marginal cost, yet ensuring meaningful and realistic diversity remains challenging [11, 37]; (ii) Surgical Simulation demonstrates increasingly sophisticated tools for surgeon and robot training, yet limitations persist in training diversity and in realism [48]; (iii) Digital Twins show significant clinical impact in pre-operative planning, but online intra-operative digital twins still face physics realism limitations, particularly for soft-tissue deformation where unknown tissue-specific mechanical properties prevent accurate modeling and registration [20, 21, 51]. A related risk is that simulation assumptions—such as idealized tissue physics or uniform lighting—can introduce systematic distributional bias into models trained on synthetic data, causing silent failures when these models encounter real operative conditions that violate those assumptions [4, 48].

These challenges span a broad spectrum of clinical maturity. At the most established end are patient-specific 3D anatomical models for pre-operative visualization and planning, which are routinely used in clinical workflows and supported by multiple FDA-cleared software platforms across specialties [21]. Simulation-based training is also clinically mature, with randomized trials demonstrating improved operative performance and reduced complications, including lower intraoperative and post-operative complication rates in laparoscopic inguinal hernia repair [76]. More advanced predictive uses of simulation are emerging. In the multicenter PRECISE-TAVI study, CT-derived simulations altered valve sizing or implantation strategy in 35% of anatomically complex cases, demonstrating prospective clinical influence [77]. By contrast, online intra-operative digital twins, visual data synthesis, and robot policy learning remain largely at the proof-of-concept or phantom stage, with sim-to-real transfer typically validated on bench models rather than live tissue. This spectrum—from clinically deployed planning tools to early-stage intra-operative systems—defines the translational path that surgical synthetic data is working to follow.

Viewing synthetic data for surgery altogether highlights that these challenges are interconnected, and that coordinated development across domains can yield multiplier effects—amplifying impact while minimizing resource investment. While most research efforts remain siloed, a small number of works illustrate what cross-domain integration can look like in practice. The AMBF ecosystem illustrates how a single technical foundation can support multiple clinical functions. At its core is AMBF+, a virtual surgical simulator. Patient CT scans are segmented and loaded into a VR environment where surgeons can rehearse drilling on patient-specific anatomy. Importantly, while the surgeon trains, the system automatically generates structured data—such as segmentation masks, depth maps, and stereo image pairs—that can be used to develop computer vision algorithms [59]. In other words, AMBF+ is primarily a simulation and data-generation platform.
FIVRS builds directly on this simulator but focuses on immersion and skill assessment. It integrates high-fidelity hardware (head-mounted display, haptics, foot pedals, eye tracking) to reproduce the operating room workflow and record multimodal performance metrics. While it retains data-generation capability, its primary contribution is as a fully immersive training and evaluation system [58].
In contrast, Twin-S is not a rehearsal simulator. It is a digital twin that runs in parallel with real surgery. Using optical tracking, it continuously updates a virtual model of the operative field to reflect real drilling in real time, achieving a mean spatial error of 1.39 mm in phantom experiments [70]. Its goal is intra-operative guidance rather than offline training.
Finally, AMBF-RL extends the same simulation backbone to robotic learning. Instead of training surgeons, it trains control policies for robotic manipulators in simulation before transferring them to physical systems [15].
Taken together, these components illustrate how cross-domain integration can emerge from a shared computational substrate: patient imaging supports digital twin construction; the digital twin enables simulation-based rehearsal; the simulation generates multimodal synthetic data; and these data support downstream computer vision development and skill assessment. Although each component has been evaluated separately and the integrated pipeline has not yet been validated in live patients, it represents a concrete near-term target for translational development.

Yu *et al.*'s platform (ORBIT-Surgical) provides a related example within the simulation domain, combining robot policy learning with structured synthetic data generation in a shared environment [48]. While not spanning the full digital twin–to–clinical workflow, it demonstrates how integration within a single technical substrate can simultaneously support training and data production. This integrated viewpoint also highlights that breakthrough advances in foundational technologies could unlock unprecedented progress across multiple domains. Improved soft tissue physics modeling—a widely recognized bottleneck for realistic simulation and accurate digital twin registration—represents a shared technical challenge where coordinated research investment could yield transformative improvement to surgical preparation [18–20, 48, 51]. Similarly, advances in real-time 3D scene reconstruction and registration methods critical for digital twins directly support simulation environment fidelity and visual data synthesis quality [20, 51]. From a clinical perspective, such an integrated approach of synthetic data into clinical workflow can fundamentally transform surgical preparation, leading to an ecosystem where surgeons could rehearse patient-specific procedures on digital twins while simultaneously generating training data for surgical education and AI-assisted decision support systems.

This approach becomes even more relevant when addressing the challenge apparent throughout all domains: generating complex, atypical, and complication scenarios. Current approaches struggle with this limitation in related ways—visual synthesis lacks diversity, simulations operate under predictable conditions rather than capturing procedural variability, and digital twins face difficulties modeling tissue property variations in complex cases. Integrated development could systematically address these limitations: patient-specific digital twins of anatomically complex cases could provide foundations for realistic simulations of rare scenarios, while these simulations could generate the diverse visual training data that current generative models struggle to create independently. The result would be a comprehensive synthetic data system that captures the full spectrum of surgical workflow—from routine procedures to the challenging cases where enhanced preparation would have the greatest clinical impact.

Finally, regulatory expectations differ substantially across use cases. Training-only applications—whether for surgeon education or for internal AI model development—generally face limited direct regulatory oversight. However, when AI systems are intended to support clinical decision-making or regulatory submissions, FDA guidance emphasizes clear context-of-use definition, transparency, lifecycle bias control, and ongoing performance monitoring [4, 78]. Although this guidance is not specific to synthetic data, these principles apply whenever synthetic data contribute to model development, validation, or performance claims. Intraoperative decision-support systems, by contrast, are regulated as medical device software and are therefore subject to substantially more stringent expectations regarding documentation, validation, transparency, and post-deployment surveillance [3, 78]. Established credibility frameworks exist for physics-based computational modeling in device submissions. However, these frameworks explicitly exclude standalone statistical or AI/ML models [79]. At present, no surgery-specific consensus standard defines what constitutes sufficient synthetic-to-real validation evidence for AI systems trained using simulation-derived data.

IV. CONCLUSION

Synthetic data has significant potential across surgery through three key domains: Visual Data Synthesis, Surgical Simulations, and Digital Twins. Individually, these technologies show promise to enhance surgical training and practice, but their siloed development limits the technical and clinical synergies that could emerge from their integration. Our analysis highlights that many overlapping challenges remain across domains, and that coordinated development will be necessary to advance their clinical translation. This is particularly evident in their shared limitation: the current difficulty of generating complex, atypical, and failure cases where enhanced training and AI support are most critically needed.

Moving forward, progress will depend not simply on scaling existing approaches within isolated domains, but on extending these integrated approaches to systematically address rare complications, anatomical variants, and equipment failures that are difficult to capture consistently in real-world training. Reaching this frontier will require targeted advances: in visual data synthesis, methods for generating clinically relevant out-of-distribution scenarios rather than replicating existing data patterns; in surgical simulation, improved physics models that can be calibrated and validated against physical substrates with known properties; and in digital twins, non-invasive inference of patient-specific tissue parameters for reliable predictive planning. Across all three domains, standardized benchmarks that explicitly evaluate performance on rare and complex cases will be important for advancing toward clinical translation [3]. By aligning synthetic data development with clinical workflow and validation requirements, safe and scalable integration of these technologies into surgical practice becomes achievable.

Search Strategy and Study Selection

This Perspective is based on a targeted literature search conducted across PubMed and Google Scholar to identify studies at the intersection of synthetic data and surgery, spanning visual data synthesis, surgical simulation, and digital twins. Search queries combined domain-specific terms—including "synthetic data," "data augmentation," "domain randomization," "domain adaptation," "surgical simulation," "robotic surgery," "sim-to-real," and "digital twin". These searches were supplemented by forward and backward citation tracking of key publications. Two authors independently screened titles and abstracts for relevance; full texts were reviewed for studies that demonstrated novel synthetic data generation methods, simulation frameworks, or digital twin applications with surgical relevance. Rather than aiming for exhaustive systematic coverage, we selected representative studies that best illustrate the state of the art, current limitations, and emerging opportunities within each domain. Thematic synthesis was conducted collaboratively to ensure balanced coverage across all three areas. Of the 79 cited works, 65 (82%) were published between 2020 and 2025, reflecting the field's recent growth. As this is a Perspective rather than a systematic review, the selection prioritizes conceptual breadth and illustrative depth over comprehensive enumeration.

Declarations

Competing Interests

The authors declare no competing financial or non-financial interests.

Author Contributions

Yann Sakref: Prepared the figures, contributed to the conception of the work, wrote the manuscript, and reviewed the manuscript. Lalithkumar Seenivasan: Contributed to the conception of the work, wrote the manuscript, and reviewed the manuscript. Hao Ding: Contributed to the conception of the work, wrote the manuscript, and reviewed the manuscript. Ruhika Iyer: Contributed to the conception of the work, wrote the manuscript, and reviewed the manuscript. Danush Kumar Venkatesh: Contributed to the conception of the work and reviewed the manuscript. Stefanie Speidel: Contributed to the conception of the work and reviewed the manuscript. Mathias Unberath: Contributed to the conception of the work and reviewed the manuscript. Jeffrey K. Jopling: Contributed to the conception of the work, wrote the manuscript, and reviewed the manuscript. Lisa M. Knowlton: Contributed to the conception of the work, wrote the manuscript, and reviewed the manuscript.

Acknowledgements

This research received no specific grant from any funding agency. The research activities of Hao Ding are supported by a collaborative research agreement with the MultiScale Medical Robotics Center at The Chinese University of Hong Kong. The research activities of Danush Kumar Venkatesh are partly supported by the Federal Ministry of Research, Technology and Space in DAAD project 57616814 (SECAI, School of Embedded Composite AI, https://secai.org/).

Ethics Approval and Consent to Participate

This study did not involve human participants, human data, or human tissue. Ethical approval was not required as the work consisted of a review and analysis of existing literature.

[1] R. S. Haluck and T. M. Krummel, "Computers and Virtual Reality for Surgical Education in the 21st Century," *Archives of Surgery*, vol. 135, pp. 786–792, July 2000.

[2] L. E. Kearse, A. Zeineddin, I. S. Schmiederer, J. R. Korndorffer, and J. N. Lau, "A 20-year review of surgical training case logs: Is general surgery still general?," *Surgery*, vol. 170, pp. 1347–1352, Nov. 2021.

[3] L. Maier-Hein, M. Eisenmann, D. Sarikaya, K. März, T. Collins, A. Malpani, J. Fallert, H. Feussner, S. Giannarou, P. Mascagni, H. Nakawala, A. Park, C. Pugh, D. Stoyanov, S. S. Vedula, K. Cleary, G. Fichtinger, G. Forestier, B. Gibaud, T. Grantcharov, M. Hashizume, D. Heckmann-Nötzel, H. G. Kenngott, R. Kikinis, L. Mündermann, N. Navab, S. Onogur, T. Roß, R. Sznitman, R. H. Taylor, M. D. Tizabi, M. Wagner, G. D. Hager, T. Neumuth, N. Padoy, J. Collins, I. Gockel, J. Goedeke, D. A. Hashimoto, L. Joyeux, K. Lam, D. R. Leff, A. Madani, H. J. Marcus, O. Meireles, A. Seitel, D. Teber, F. Ückert, B. P. Müller-Stich, P. Jannin, and S. Speidel, "Surgical data science – from concepts toward clinical translation," *Medical Image Analysis*, vol. 76, p. 102306, Feb. 2022.

[4] R. J. Chen, M. Y. Lu, T. Y. Chen, D. F. K. Williamson, and F. Mahmood, "Synthetic data in machine learning for medicine and healthcare," *Nature Biomedical Engineering*, vol. 5, pp. 493–497, June 2021.

[5] S. A. Cardoso, J. Suyambu, J. Iqbal, D. C. Cortes Jaimes, A. Amin, J. T. Sikto, M. Valderrama, S. S. Aulakh, V. Ramana, B. Shaukat, and T. Patel, "Exploring the Role of Simulation Training in Improving Surgical Skills Among Residents: A Narrative Review," *Cureus*, Sept. 2023.

[6] E. Schonfeld and A. Veeravagu, "Demonstrating the successful application of synthetic learning in spine surgery for training multi–center models with increased patient privacy," *Scientific Reports*, vol. 13, p. 12481, Aug. 2023.

[7] M. Pfeiffer, I. Funke, M. R. Robu, S. Bodenstedt, L. Strenger, S. Engelhardt, T. Roß, M. J. Clarkson, K. Gurusamy, B. R. Davidson, L. Maier-Hein, C. Riediger, T. Welsch, J. Weitz, and S. Speidel, "Generating Large Labeled Data Sets for Laparoscopic Image Processing Tasks Using Unpaired Image-to-Image Translation," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019* (D. Shen, T. Liu, T. M. Peters, L. H. Staib, C. Essert, S. Zhou, P.-T. Yap, and A. Khan, eds.), (Cham), pp. 119–127, Springer International Publishing, 2019.

[8] J. Yoon, S. Hong, S. Hong, J. Lee, S. Shin, B. Park, N. Sung, H. Yu, S. Kim, S. Park, W. J. Hyung, and M.-K. Choi, "Surgical Scene Segmentation Using Semantic Image Synthesis with a Virtual Surgery Environment," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022* (L. Wang, Q. Dou, P. T. Fletcher, S. Speidel, and S. Li, eds.), (Cham), pp. 551–561, Springer Nature Switzerland, 2022.

[9] D. K. Venkatesh, D. Rivoir, M. Pfeiffer, and S. Speidel, "SurgicaL-CD: Generating Surgical Images via Unpaired Image Translation with Latent Consistency Diffusion Models," in *Computer Vision – ECCV 2024 Workshops* (A. Del Bue, C. Canton, J. Pont-Tuset, and T. Tommasi, eds.), (Cham), pp. 218–235, Springer Nature Switzerland, 2025.

[10] J. Cho, S. Schmidgall, C. Zakka, M. Mathur, D. Kaur, R. Shad, and W. Hiesinger, "SurGen: Text-Guided Diffusion Model for Surgical Video Generation," Sept. 2024. arXiv:2408.14028 [cs].

[11] A. Bhat, R. Bose, C. I. Nwoye, and N. Padoy, "SimGen: A Diffusion-Based Framework for Simultaneous Surgical Image and Segmentation Mask Generation," Jan. 2025. arXiv:2501.09008 [cs].

[12] C. I. Nwoye, R. Bose, K. Elgohary, L. Arboit, G. Carlino, J. L. Lavanchy, P. Mascagni, and N. Padoy, "Surgical Text-to-Image Generation," *Pattern Recognition Letters*, vol. 190, pp. 73–80, Apr. 2025. arXiv:2407.09230 [cs].

[13] D. Biagini, N. Navab, and A. Farshad, "HieraSurg: Hierarchy-Aware Diffusion Model for Surgical Video Generation," June 2025. arXiv:2506.21287 [cs].

[14] M. Moghani, N. Nelson, M. Ghanem, A. Diaz-Pinto, K. Hari, M. Azizian, K. Goldberg, S. Huver, and A. Garg, "SuFIA-BC: Generating High Quality Demonstration Data for Visuomotor Policy Learning in Surgical Subtasks," Apr. 2025. arXiv:2504.14857 [cs].

[15] V. M. Varier, D. K. Rajamani, F. Tavakkolmoghaddam, A. Munawar, and G. S. Fischer, "AMBF-RL: A realtime simulation based Reinforcement Learning toolkit for Medical Robotics," in *2022 International Symposium on Medical Robotics (ISMR)*, (GA, USA), pp. 1–8, IEEE, Apr. 2022.

[16] P. M. Scheikl, B. Gyenes, R. Younis, C. Haas, G. Neumann, M. Wagner, and F. Mathis-Ullrich, "LapGym - An Open Source Framework for Reinforcement Learning in Robot-Assisted Laparoscopic Surgery," *Journal of Machine Learning Research*, vol. 24, pp. 1–42, Dec. 2023.

[17] Y. Ou and M. Tavakoli, "Sim-to-Real Surgical Robot Learning and Autonomous Planning for Internal Tissue Points Manipulation Using Reinforcement Learning," *IEEE Robotics and Automation Letters*, vol. 8, pp. 2502–2509, May 2023.

[18] J. Xu, B. Li, B. Lu, Y.-H. Liu, Q. Dou, and P.-A. Heng, "SurRoL: An Open-source Reinforcement Learning Centered and dVRK Compatible Platform for Surgical Robot Learning," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, (Prague, Czech Republic), pp. 1821–1828, IEEE, Sept. 2021.

[19] A. Munawar, J. Y. Wu, G. S. Fischer, R. H. Taylor, and P. Kazanzides, "Open Simulation Environment for Learning and Practice of Robot-Assisted Surgical Suturing," *IEEE Robotics and Automation Letters*, vol. 7, pp. 3843–3850, Apr. 2022.

[20] H. Ding, L. Seenivasan, B. D. Killeen, S. M. Cho, and M. Unberath, "Digital twins as a unifying framework for surgical data science: the enabling role of geometric scene understanding," *Artificial Intelligence Surgery*, vol. 4, pp. 109–38, July 2024.

[21] L. Asciak, J. Kyeremeh, X. Luo, A. Kazakidi, P. Connolly, F. Picard, K. O'Neill, S. A. Tsaftaris, G. D. Stewart, and W. Shu, "Digital twin assisted surgery, concept, opportunities, and challenges," *npj Digital Medicine*, vol. 8, p. 32, Jan. 2025.

[22] Y. M. Mekki, G. Luijten, E. Hagert, S. Belkhair, C. Varghese, J. Qadir, B. Solaiman, M. Bilal, J. Dhanda, J. Egger, J. Deng, V. Khanduja, A. F. Frangi, S. M. Zughaier, and M. A. Stotland, "Digital twins for the era of personalized surgery," *npj Digital Medicine*, vol. 8, p. 283, May 2025.

[23] F. Moghul and S. Kashyap, "Bile Duct Injury," in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2025.

[24] J. A. Proaño-Zamudio, I. C. Nzenwa, M. Abiad, D. Argandykov, A.-S. C. Romijn, E. Lagazzi, W. Rafaqat, C. N. Paranjape, G. C. Velmahos, H. M. Kaafarani, and J. O. Hwabejire, "Impact of intraoperative adverse events in general and gastrointestinal surgery: A nationwide study," *The American Journal of Surgery*, vol. 240, p. 116125, Feb. 2025.

[25] H. Ding, Y. Zhang, T. Lu, R. Liang, H. Shu, L. Seenivasan, Y. Long, Q. Dou, C. Gao, Y. Leng, S. B. Yoo, E.-J. Lee, N. Ghamsarian, K. Schoeffmann, R. Sznitman, Z. Wu, Y. Chen, S. E. Salcudean, S. Irshad, S. Albarqouni, S. T. Kim, Y. Sun, A. Wang, L. Bai, H. Ren, I. Ullah, H.-G. Ha, A. Khan, H. Lee, S. Kondo, S. Kasai, K. Hirasawa, S. Tailor, R. Sanchez-Matilla, I. Luengo, T. Fu, J. Ma, B. Wang, M. Fernández-Rodríguez, E. Lima, J. L. Vilaça, and M. Unberath, "SegSTRONG-C: Segmenting Surgical Tools Robustly On Non-adversarial Generated Corruptions – An EndoVis'24 Challenge," Apr. 2025. arXiv:2407.11906 [cs].

[26] H. Mohamadipanah, L. Kearse, B. Wise, L. Backhus, and C. Pugh, "Generating Rare Surgical Events Using CycleGAN: Addressing Lack of Data for Artificial Intelligence Event Recognition," *Journal of Surgical Research*, vol. 283, pp. 594–605, Mar. 2023.

[27] D. K. Venkatesh, I. Funke, M. Pfeiffer, F. Kolbinger, H. M. Schmeiser, J. Weitz, M. Distler, and S. Speidel, "Mission Balance: Generating Under-represented Class Samples using Video Diffusion Models," May 2025. arXiv:2505.09858 [cs].

[28] Y. Zhou, R. Towning, Z. Awad, and S. Giannarou, "Image synthesis with class-aware semantic diffusion models for surgical scene segmentation," *Healthcare Technology Letters*, vol. 12, no. 1, p. e70003, 2025.

[29] N. Mirchi, V. Bissonnette, R. Yilmaz, N. Ledwos, A. Winkler-Schwartz, and R. F. D. Maestro, "The Virtual Operative Assistant: An explainable artificial intelligence tool for simulation-based training in surgery and medicine," *PLOS ONE*, vol. 15, p. e0229596, Feb. 2020.

[30] K. Aubert, A. Germaneau, M. Rochette, W. Ye, M. Severyns, M. Billot, P. Rigoard, and T. Vendeuvre, "Development of Digital Twins to Optimize Trauma Surgery and Postoperative Management. A Case Study Focusing on Tibial Plateau Fracture," *Frontiers in Bioengineering and Biotechnology*, vol. 9, p. 722275, Oct. 2021.

[31] A. Marzullo, S. Moccia, M. Catellani, F. Calimeri, and E. D. Momi, "Towards realistic laparoscopic image generation using image-domain translation," *Computer Methods and Programs in Biomedicine*, vol. 200, p. 105834, Mar. 2021.

[32] D. K. Venkatesh, D. Rivoir, M. Pfeiffer, F. Kolbinger, and S. Speidel, "Data Augmentation for Surgical Scene Segmentation with Anatomy-Aware Diffusion Models," in *2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 2280–2290, Feb. 2025.

[33] J. Sleeman, L. Krames, and W. Nahm, "Towards Liver Segmentation in Laparoscopic Images by Training U-Net With Synthetic Data," *Current Directions in Biomedical Engineering*, vol. 10, pp. 600–603, Dec. 2024.

[34] P. Leoncini, F. Marzola, M. Pescio, M. Casadio, A. Arezzo, and G. Dagnino, "A reproducible framework for synthetic data generation and instance segmentation

in robotic suturing," *International Journal of Computer Assisted Radiology and Surgery*, vol. 20, pp. 1567–1576, Aug. 2025.

[35] C. Li, H. Liu, Y. Liu, B. Y. Feng, W. Li, X. Liu, Z. Chen, J. Shao, and Y. Yuan, "Endora: Video Generation Models as Endoscopy Simulators," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2024* (M. G. Linguraru, Q. Dou, A. Feragen, S. Giannarou, B. Glocker, K. Lekadir, and J. A. Schnabel, eds.), (Cham), pp. 230–240, Springer Nature Switzerland, 2024.

[36] T. Chen, S. Yang, J. Wang, L. Bai, H. Ren, and L. Zhou, "SurgSora: Object-Aware Diffusion Model for Controllable Surgical Video Generation," June 2025. arXiv:2412.14018 [cs].

[37] Venkatesh Danush Kumar, Funke Isabel, Pfeiffer Micha, Kolbinger Fiona, Schmeiser Hanna Maria, Distler Marius, Weitz Jürgen, and Speidel Stefanie, "Mission Balance: Generating Under-represented Class Samples using Video Diffusion Models," in *proceedings of Medical Image Computing and Computer Assisted Intervention*, vol. LNCS 15970, Springer Nature Switzerland, Sept. 2025.

[38] S. Salehin, A. J. Islam, S. U. Alam, S. Paul, K. Islam, and P. Paul, "Desmoking and Perceptual Quality Restoration of Laparoscopic Videos Using CycleGAN," in *2024 IEEE International Conference on Biomedical Engineering, Computer and Information Technology for Health (BECITHCON)*, pp. 31–34, Nov. 2024.

[39] W. Wang, Q. Yuan, P. Huang, X. Wang, and H. Song, "Desmoke-VCU: Improved unpaired image-to-image translation for removing smoke from laparoscopic images," *Digital Signal Processing*, vol. 162, p. 105177, July 2025.

[40] S. Salazar-Colores, H. M. Jiménez, C. J. Ortiz-Echeverri, and G. Flores, "Desmoking Laparoscopy Surgery Images Using an Image-to-Image Translation Guided by an Embedded Dark Channel," *IEEE Access*, vol. 8, pp. 208898–208909, 2020.

[41] B. Kengen, H. Van Goor, and J.-M. Luursema, "Laparoscopic simulator performance and learning curves under different optical angles," *BMC Medical Education*, vol. 23, p. 613, Aug. 2023.

[42] J. W. Kim, T. Z. Zhao, S. Schmidgall, A. Deguet, M. Kobilarov, C. Finn, and A. Krieger, "Surgical Robot Transformer (SRT): Imitation Learning for Surgical Tasks," July 2024. arXiv:2407.12998 [cs].

[43] J. Woong, T. Chen, P. Hansen, L. X. Shi, A. Goldenberg, S. Schmidgall, P. M. Scheikl, A. Deguet, B. M. White, C. Finn, and A. Krieger, "SRT-H: A hierarchical framework for autonomous surgery via language-conditioned imitation learning," *Science Robotics*, vol. 10, Aug. 2025.

[44] T. Jianu, B. Huang, M. Nhat Vu, M. E. M. K. Abdelaziz, S. Fichera, C.-Y. Lee, P. Berthet-Rayne, F. Rodriguez y Baena, and A. Nguyen, "CathSim: An Open-Source Simulator for Endovascular Intervention," *IEEE Transactions on Medical Robotics and Bionics*, vol. 6, pp. 971–979, Aug. 2024.

[45] L. Karstensen, H. Robertshaw, J. Hatzl, B. Jackson, J. Langejürgen, K. Breininger, C. Uhl, S. M. H. Sadati, T. Booth, C. Bergeles, and F. Mathis-Ullrich, "Learning-Based Autonomous Navigation, Benchmark Environments and Simulation Framework for Endovascular Interventions," Oct. 2024. arXiv:2410.01956 [cs].

[46] P. M. Scheikl, E. Tagliabue, B. Gyenes, M. Wagner, D. Dall'Alba, P. Fiorini, and F. Mathis-Ullrich, "Sim-To-Real Transfer for Visual Reinforcement Learning of Deformable Object Manipulation for Robot-Assisted Surgery," *IEEE Robotics and Automation Letters*, vol. 8, pp. 560–567, Feb. 2023. arXiv:2406.06092 [cs].

[47] Y.-J. Ho, Z.-Y. Chiu, Y. Zhi, and M. C. Yip, "SurgIRL: Towards Life-Long Learning for Surgical Automation by Incremental Reinforcement Learning," Sept. 2024. arXiv:2409.15651 [cs].

[48] Q. Yu, M. Moghani, K. Dharmarajan, V. Schorp, W. C.-H. Panitch, J. Liu, K. Hari, H. Huang, M. Mittal, K. Goldberg, and A. Garg, "Orbit-Surgical: An Open-Simulation Framework for Learning Surgical Augmented Dexterity," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, (Yokohama, Japan), pp. 15509–15516, IEEE, May 2024.

[49] T. Yao, H. Wang, B. Lu, J. Ge, Z. Pei, M. Kowarschik, L. Sun, L. Seneviratne, and P. Qi, "Sim2Real Learning With Domain Randomization for Autonomous Guidewire Navigation in Robotic-Assisted Endovascular Procedures," *IEEE Transactions on Automation Science and Engineering*, vol. 22, pp. 13842–13854, 2025.

[50] J. Zhang, Y. Zhong, and C. Gu, "Deformable Models for Surgical Simulation: A Survey," *IEEE Reviews in Biomedical Engineering*, vol. 11, pp. 143–164, 2018.

[51] X. Liang, F. Liu, Y. Zhang, Y. Li, S. Lin, and M. Yip, "Real-to-Sim Deformable Object Manipulation: Optimizing Physics Models with Residual Mappings for Robotic Surgery," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 15471–15477, May 2024.

[52] A. Shahbazi, K. Pereira, J. S. Heiselman, E. Akbari, A. C. Benson, S. Seifi, X. Liu, G. L. Johnston, E. Terpstra, A. Draaisma, J.-J. Severes, J. Y. Wu, N. Simaan, M. L. Miga, and S. Kolouri, "Neural-Augmented Kelvinlet: Real-Time Soft Tissue Deformation with Multiple Graspers," June 2025. arXiv:2506.08043 [cs].

[53] C. R. Larsen, J. L. Soerensen, T. P. Grantcharov, T. Dalsgaard, L. Schouenborg, C. Ottosen, T. V. Schroeder, and B. S. Ottesen, "Effect of virtual reality training on laparoscopic surgery: randomised controlled trial," *BMJ*, May 2009.

[54] T. P. Grantcharov, V. B. Kristiansen, J. Bendix, L. Bardram, J. Rosenberg, and P. Funch-Jensen, "Randomized clinical trial of virtual reality simulation for laparoscopic skills training," *British Journal of Surgery*, vol. 91, pp. 146–150, Feb. 2004.

[55] J. D. Ferris, P. H. Donachie, R. L. Johnston, B. Barnes, M. Olaitan, and J. M. Sparrow, "Royal College of Ophthalmologists' National Ophthalmology Database study of cataract surgery: report 6. The impact of EyeSi virtual reality training on complications rates of cataract surgery performed by first and second year trainees," *British Journal of Ophthalmology*, vol. 104, pp. 324–329, Mar. 2020.

[56] G. J. Bunogerane, K. Taylor, Y. Lin, and A. Costas-Chavarri, "Using Touch Surgery to Improve Surgical Education in Low- and Middle-Income Settings: A Randomized Control Trial," *Journal of Surgical Education*, vol. 75, pp. 231–237, Jan. 2018.

[57] S. Shetty, B. Zevin, T. P. Grantcharov, K. E. Roberts, and A. J. Duffy, "Perceptions, Training Experiences, and Preferences of Surgical Residents Toward Laparoscopic Simulation Training: A Resident Survey," *Journal of*

Surgical Education, vol. 71, pp. 727–733, Sept. 2014.

[58] A. Munawar, Z. Li, N. Nagururu, D. Trakimas, P. Kazanzides, R. H. Taylor, and F. X. Creighton, "Fully immersive virtual reality for skull-base surgery: surgical training and beyond," *International Journal of Computer Assisted Radiology and Surgery*, vol. 19, pp. 51–59, Jan. 2024.

[59] A. Munawar, Z. Li, P. Kunjam, N. Nagururu, A. S. Ding, P. Kazanzides, T. Looi, F. X. Creighton, R. H. Taylor, and M. Unberath, "Virtual reality for synergistic surgical training and data generation," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 10, pp. 366–374, July 2022.

[60] M. A. Gromski, W. Ahn, K. Matthes, and S. De, "Preclinical Training for New Notes Procedures," *Gastrointestinal Endoscopy Clinics of North America*, vol. 26, pp. 401–412, Apr. 2016.

[61] G. Sankaranarayanan, K. Matthes, A. Nemani, W. Ahn, M. Kato, D. B. Jones, S. Schwaitzberg, and S. De, "Needs analysis for developing a virtual-reality NOTES simulator," *Surgical Endoscopy*, vol. 27, pp. 1607–1616, May 2013.

[62] S. Colman, M. El-Bahnasawi, N. Abdulkader, Z. Aloul, J. Brown, P. Luthra, and D. Rawaf, "The 'LapAR' augmented reality training device in surgical simulation: a multi-center pilot study," *Global Surgical Education - Journal of the Association for Surgical Education*, vol. 4, p. 32, Jan. 2025.

[63] J. Cartucho, C. Wang, B. Huang, D. S. Elson, A. Darzi, and S. Giannarou, "An enhanced marker pattern that achieves improved accuracy in surgical tool tracking," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 10, pp. 400–408, July 2022.

[64] T. Smit, N. Aage, D. Haschtmann, S. J. Ferguson, and B. Helgason, "In silico medical device testing of anatomically and mechanically conforming patient-specific spinal fusion cages designed by full-scale topology optimisation," *Frontiers in Bioengineering and Biotechnology*, vol. 12, p. 1347961, Sept. 2024.

[65] H. Ahmadian, P. Mageswaran, B. A. Walter, D. M. Blakaj, E. C. Bourekas, E. Mendel, W. S. Marras, and S. Soghrati, "A digital twin for simulating the vertebroplasty procedure and its impact on mechanical stability of vertebra in cancer patients," *International Journal for Numerical Methods in Biomedical Engineering*, vol. 38, no. 6, p. e3600, 2022. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cnm.3600.

[66] M. Pescio, C. Li, D. Kundrat, M. Casadio, and G. Dagnino, "A simulation environment for robot-assisted endovascular interventions," *International Journal of Computer Assisted Radiology and Surgery*, vol. 20, pp. 2259–2267, Nov. 2025.

[67] D. Kim, D. Narayanan, S.-H. Sung, H.-M. Cheng, C.-H. Chen, C.-S. Kim, R. Mukkamala, and J.-O. Hahn, "Transmission line model as a digital twin for abdominal aortic aneurysm patients," *npj Digital Medicine*, vol. 7, p. 301, Oct. 2024.

[68] A. Rouhollahi, J. N. Willi, S. Haltmeier, A. Mehrtash, R. Straughan, H. Javadikasgari, J. Brown, A. Itoh, K. I. De La Cruz, E. Aikawa, E. R. Edelman, and F. R. Nezami, "CardioVision: A fully automated deep learning package for medical image segmentation and reconstruction generating digital twins for patients with aortic stenosis," *Computerized Medical Imaging and Graphics*, vol. 109, p. 102289, Oct. 2023.

[69] Y. Shen, H. Ding, L. Seenivasan, T. Shu, and M. Unberath, "Position: Foundation Models Need Digital Twin Representations," May 2025. arXiv:2505.03798 [cs].

[70] H. Shu, R. Liang, Z. Li, A. Goodridge, X. Zhang, H. Ding, N. Nagururu, M. Sahu, F. X. Creighton, R. H. Taylor, A. Munawar, and M. Unberath, "Twin-S: a digital twin for skull base surgery," *International Journal of Computer Assisted Radiology and Surgery*, vol. 18, pp. 1077–1084, May 2023.

[71] Y. Shi, X. Deng, Y. Tong, R. Li, Y. Zhang, L. Ren, and W. Si, "Synergistic Digital Twin and Holographic Augmented-Reality-Guided Percutaneous Puncture of Respiratory Liver Tumor," *IEEE Transactions on Human-Machine Systems*, vol. 52, pp. 1364–1374, Dec. 2022.

[72] J. Cowley, J. Kyeremeh, G. D. Stewart, X. Luo, W. Shu, and A. Kazakidi, "Near Real-Time Estimation of Blood Loss and Flow–Pressure Redistribution during Unilateral Nephrectomy," *Fluids*, vol. 9, p. 214, Sept. 2024.

[73] H. Ding, Y. Zhang, W. Cheng, X. Wang, X. Lian, C. Yu, H. Shu, J. W. Kim, A. Krieger, and M. Unberath, "Towards Robust Algorithms for Surgical Phase Recognition via Digital Twin Representation," Mar. 2025. arXiv:2410.20026 [cs].

[74] L. Seenivasan, J. Xu, R. D. S. Mukul, H. Ding, G. Byrd, Y.-C. Ku, J. L. Porras, M. Ishii, and M. Unberath, "Beyond Rigid AI: Towards Natural Human-Machine Symbiosis for Interoperative Surgical Assistance," July 2025. arXiv:2507.23088 [cs].

[75] A. M. DiGioia and B. Jaramaz, "Computer-assisted tools and interventional technologies," *Lancet (London, England)*, vol. 354 Suppl, p. SIV46, Dec. 1999.

[76] B. Zendejas, D. A. Cook, J. Bingener, M. Huebner, W. F. Dunn, M. G. Sarr, and D. R. Farley, "Simulation-Based Mastery Learning Improves Patient Outcomes in Laparoscopic Inguinal Hernia Repair: A Randomized Controlled Trial," *Annals of Surgery*, vol. 254, p. 502, Sept. 2011.

[77] T. W. Hokken, H. Wienemann, J. Dargan, D.-J. v. Ginkel, C. Dowling, A. Unbehaun, J. Bosmans, A. Bader-Wolfe, R. Gooley, M. Swaans, S. J. Brecker, M. Adam, and N. M. Van Mieghem, "Clinical value of CT-derived simulations of transcatheter-aortic-valve-implantation in challenging anatomies the PRECISE-TAVI trial," *Catheterization and Cardiovascular Interventions*, vol. 102, no. 6, pp. 1140–1148, 2023. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ccd.30816.

[78] U. F. a. D. Administration, "Artificial Intelligence-Enabled Device Software Functions: Lifecycle Management and Marketing Submission Recommendations," Jan. 2025.

[79] U. F. a. D. Administration, "Assessing the Credibility of Computational Modeling and Simulation in Medical Device Submissions," Nov. 2023.

Figure Legends

Figure 1: Three Domains of Synthetic Data in Surgery. Each domain addresses distinct clinical needs through different synthetic data approaches. **(A) Visual Data Synthesis** generates synthetic annotated surgical images for AI training and surgical education, shown with diffusion models creating realistic surgical scenes from segmentation maps for anatomical structure recognition (adapted from Zhou et al. [28]). **(B) Surgical Simulation** creates virtual environments for skills training, illustrated with a trainee performing virtual tumor resection using haptic-enabled instruments in a risk-free setting (adapted from Mirchi *et al.* [29]). **(C) Digital Twin** transforms 3D X-ray clinical imaging into patient-specific computational models with heterogeneous bone properties, enabling simulation of multiple surgical scenarios to optimize treatment planning and predict healing outcomes (from Aubert et al. [30]). Panel (A) adapted under CC BY 4.0 from Zhou et al. [28]. Panel (B) adapted under CC BY 4.0 from Mirchi et al. [29]. Panel (C) adapted under CC BY 4.0 from Aubert et al. [30]. Created in BioRender. Sakref, Y. (2025) https://BioRender.com/0oybj1h.

Figure 2: The Three Interconnected Domains of Synthetic Data in Surgery. Visual Data Synthesis is the generation of annotated surgical images and videos for AI system development and surgical education. Surgical Simulation corresponds to virtual environments for robot policy learning, surgeon skill acquisition, and platform validation. Digital Twin involves patient-specific models for personalized surgical planning, real-time guidance, and post-operative care. The overlapping regions highlight key technical synergies between the three domains: realistic rendering between Visual Data Synthesis and Surgical Simulation, anatomical segmentation shared between Visual Data Synthesis and Digital Twin, and soft tissue physics common to Surgical Simulation and Digital Twin, with the ultimate convergence of the domains toward synthetic environments enabling training, support, and automation. Key takeaway: these three domains share deep technical interdependencies, and the most impactful advances will target their shared foundations to drive convergence. Created in BioRender. Sakref, Y. (2025) https://BioRender.com/0oybj1h

Figure 3: Synthetic Data Integration Across the Surgical Workflow. Synthetic data is becoming integral to every phase of surgical care, from pre-operative training and planning through intra-operative guidance and post-operative care. Illustrated applications include workflow analysis and semantic understanding (with diverse synthetic images and on-demand video generation), skill acquisition and robot control policy learning through virtual environments, patient-specific surgical planning with digital twins, intra-operative scene reconstruction and guidance, and post-operative care with predictive analytics. The workflow also shows how operations (normal as well as complex cases) can enhance synthetic data generation leading to more comprehensive training. Synthetic data applications are categorized as follows: visual data synthesis in red, simulation in green, and digital twin in blue. Key takeaway: synthetic data is emerging across the surgical workflow and enables bidirectional feedback between training, intraoperative deployment, and post-operative learning—particularly through incorporation of complex and failure cases. Created in BioRender. Sakref, Y. (2025) https://BioRender.com/0oybj1h.

Table 1: Synthetic Data Use Cases in Surgery by Domain. Synthetic data applications in surgery categorized by visual data synthesis (red), simulation (green), and digital twin (blue) approaches. The listed studies are illustrative examples of representative research efforts and reflect varying stages of technical and clinical maturity within each subfield. Challenges highlight current methodological limitations, particularly regarding diversity, complexity, and atypical case generation. Created in BioRender. Sakref, Y. (2025) https://BioRender.com/0oybj1h