SimuScope: Realistic Endoscopic Synthetic Dataset Generation through Surgical Simulation and Diffusion Models
-Supplementary Material-

Sabina Martyniak[1] Joanna Kaleta[1,2] Diego Dall'Alba[3] Michał Naskręt[1]
Szymon Płotka[1,4] Przemysław Korzeniowski[1]

Sano Centre for Computational Medicine, Poland[1] Warsaw University of Technology, Poland[2]
University of Verona, Italy[3] University of Warsaw, Poland[4]

A. Publicly available dataset

To foster research in the field, we release our dataset for a community. Here, we present more generated surgical images provided by our proposed method called SimuScope. In total, the simulator dataset consists of 13,064 images. During the generation of realistic images (see Figure 1), artifacts appeared on the generated images (see Figure 2) which have also been attached. The number of images with artifacts is 1673 out of 13064.

Figure 1 presents a wide range of sample images generated by SimuScope, showcasing various perspectives and anatomical regions during cholecystectomy. These examples emphasize the simulator's capability to produce high-fidelity visualizations from different angles, providing a detailed representation of the surgical environment.

Figure 2 presents artifacts that emerged during the SimuScope image generation process, which are a natural consequence of the simulator's high realism. The complex lighting effects and diverse textures used in the simulator contribute to these artifacts, reflecting the intricate nature of realistic surgical environments. While some of the artifacts, such as instrument placement or color saturation inconsistencies, may appear, they result from the simulator's effort to accurately replicate real-world scenarios, highlighting areas for further refinement in the simulation.

Figure 3 illustrates the Da Vinci surgical tools in the dataset used for the cholecystectomy simulator. The dataset includes a diverse range of robotic instruments, enabling realistic modeling of surgical procedures. These tools are essential for replicating the precise and controlled actions of robotic-assisted surgery, making the dataset a valuable resource for developing and evaluating simulation-based methodologies. The variety of instruments allows the simulator to closely mirror real-world surgical techniques, enhancing its utility for both training and research applications.

B. Visual evaluation and parameter configuration for LoRA adapters

Table 1 presents the selected experimental parameter values for various LoRA (Low-Rank Adaptation) configurations applied to the models. It highlights key parameters such as Denoising Strength, CFG (Classifier-Free Guidance), SoftEdge Strength, Depth Control Strength, and Reference Control Strength. All experiments consistently utilize the DPM++2M Karras noise scheduler to maintain uniform noise handling across simulations, thereby improving the quality and reliability of the generated outputs. For the specific case of the CholectD45 model, the parameter values are: Denois at 0.65, CFG at 7.0, SoftEdge at 0.45, Depth at 0.65, and Reference at 0.65.

This set of images (see Figure 4) illustrates the outcomes of experiments using different LoRA adapter with varying parameter settings. The figure shows a series of visual comparisons at different LoRA weights (0.45, 0.6, and 0.9) across multiple configurations, highlighting how the integration of LoRAs affects the realism of simulated endoscopic images. At lower weights, the results are less realistic, with visible artifacts and a synthetic appearance. At higher weights, improvements in tissue and instrument rendering are observed, although color artifacts appear. The combination of multiple LoRAs provides the most realistic and faithful simulation.

Table 1. Selected experimentaly LoRA's parameter values for each model: denoising strength, CFG, SoftEdge, Depth, and Reference control strength. All models utilize the noise scheduler DPM++2M Karras, ensuring consistent noise management across simulations and enhancing the fidelity of generated outputs.

Style	Denois	CFG	SoftEdge	Depth	Reference
CholectD45	0.65	7.0	0.45	0.65	0.65

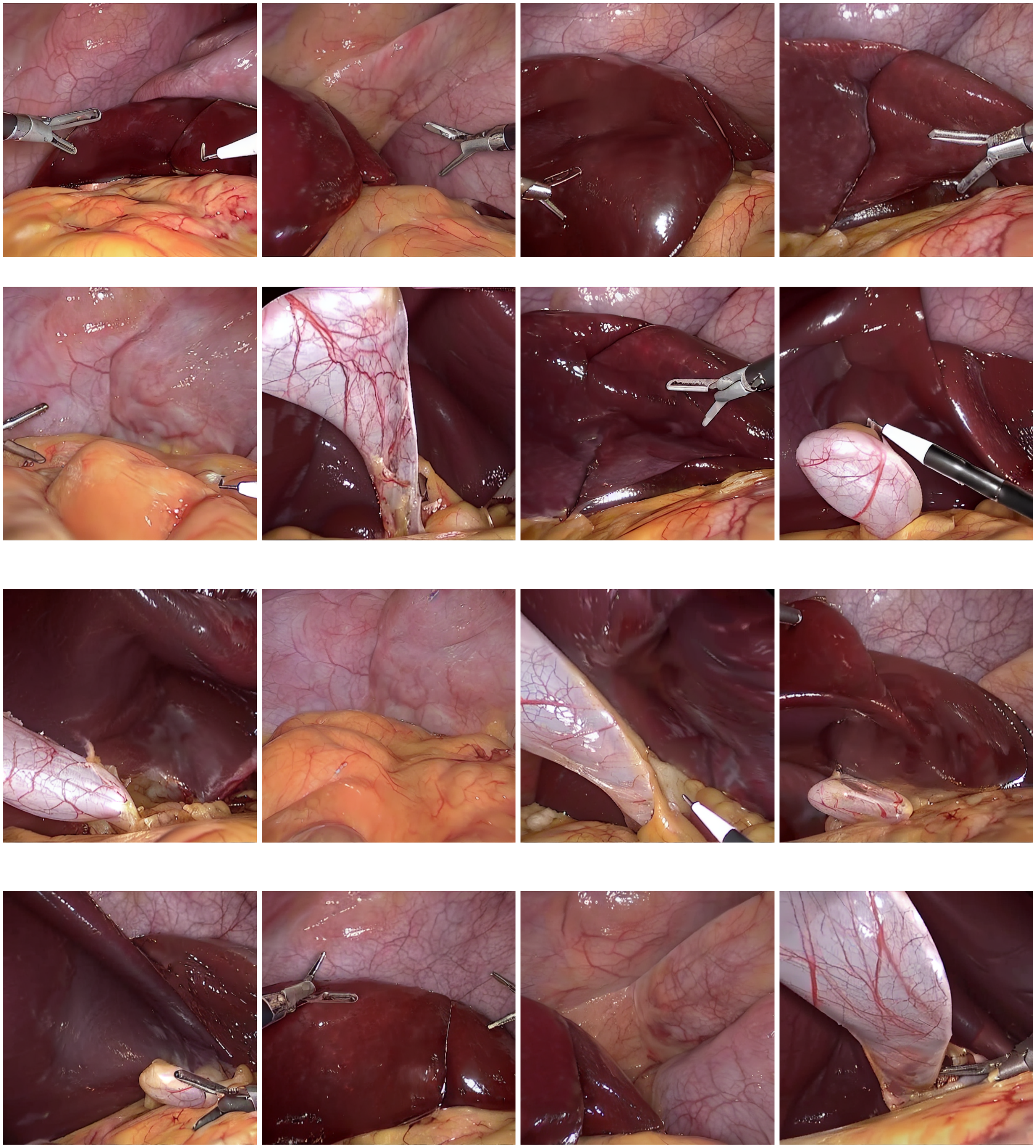

Figure 1. SimuScope generated images. A collection of sample generated images from our simulator is presented here, showcasing a diverse range of perspectives. These examples highlight the simulator's capability to produce detailed visualizations from different angles and orientations. The images encompass various anatomical regions and demonstrate the versatility of the simulation in replicating realistic scenarios.

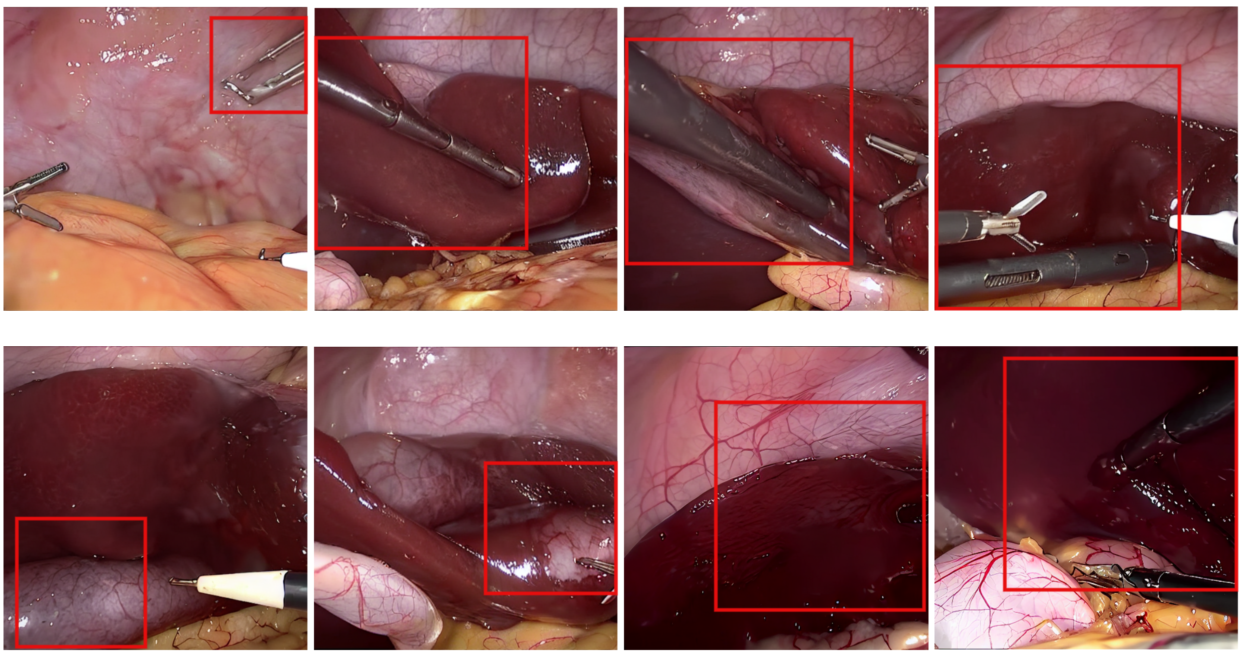

Figure 2. SimuScope artifacts. During the generation of images, various artifacts appeared. As shown in the attached images, these mainly include the addition of instruments in parts of the liver or abdominal wall, which can be observed in the first four images from the first row. These artifacts may interfere with accurate interpretation by introducing extraneous elements that do not belong to the actual anatomical structures. The next four images from the second row exhibit artifacts related to color saturation, where abnormal intensities and hues may distort the visual information. These color saturation artifacts can obscure important details and mislead diagnostic assessments by creating false impressions of tissue characteristics.

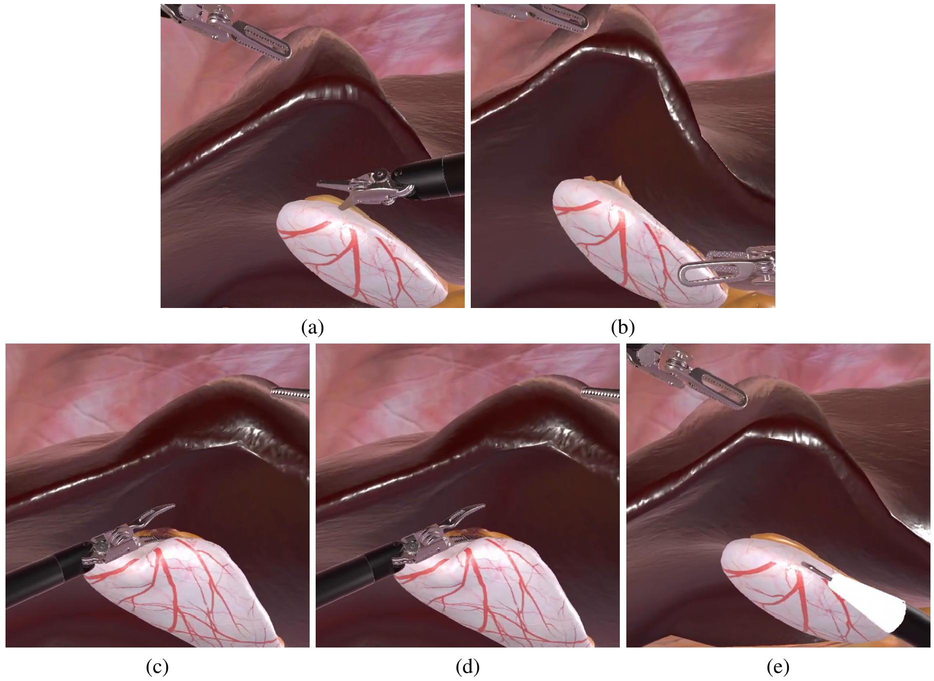

Figure 3. Images showing DaVinci tools used in the cholecystectomy simulator: a) Round Tip Scissors - versatile scissors for delicate tissue manipulation, providing precise cuts and control during the procedure, b) Cadiere Forceps - essential for grasping and holding tissues securely, allowing for firm yet gentle manipulation of tissues, c) Maryland Bipolar Forceps - used for precise tissue dissection and coagulation, offering both cutting and hemostatic capabilities, d) Monopolar Curved Scissors - precision tool for cutting and dissecting tissues, designed to facilitate curved and angled incisions, and e) Hook - versatile instrument for hooking and retracting tissues, aiding in the exposure and manipulation of the surgical site.

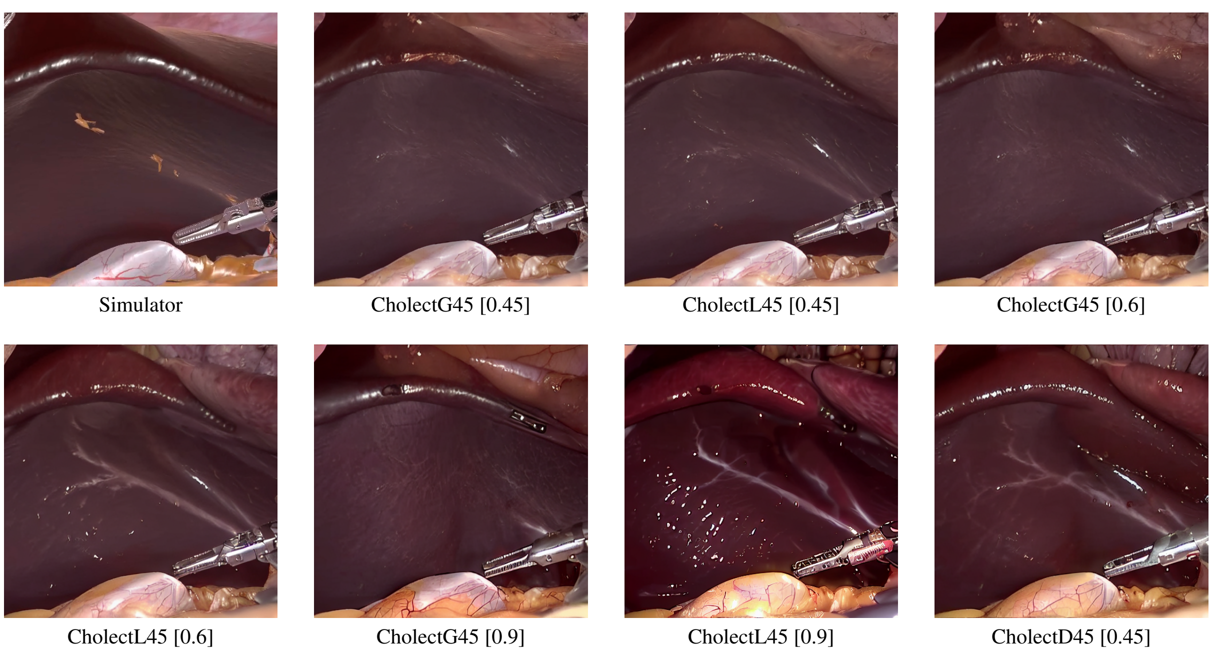

Figure 4. Visual comparisons reveal that using LoRAs individually degrades results. At 0.45 weight, CholectD45 shows minimal improvement. At 0.6, tissue and instrument appearance improves, though results remain synthetic. At 0.9, color artifacts appear. Combining two LoRAs yields the best results, closely resembling real endoscopic images with enhanced realism and fidelity.