BIOENGINEERING AND MEDICAL-SURGICAL SCIENCES

UNITO - Enhancing Robotic Surgery Autonomy through Computer Vision, AI, and Simulation

Funded By	UNIVERSITA' DEGLI STUDI DI TORINO [P.iva/CF:02099550010]

Supervisor	AREZZO ALBERTO - alberto.arezzo@polito.it

Contact	Federica Barontini DAGNINO GIULIO - giulio.dagnino@polito.it MARZOLA FRANCESCO - francesco.marzola@polito.it

Context of the research activity	Robotic surgery has the potential to enhance precision and clinical outcomes in minimally invasive procedures. Despite advantages like increased degrees of freedom, full teleoperation places significant cognitive demands on surgeons. Augmenting robotic systems with autonomous capabilities can improve surgical outcomes and safety, specifically in complex gestures such as suturing. This three-year PhD project focuses on advancing robotic surgery autonomy by integrating cognitive computer vision systems, simulation environments, and AI-driven methodologies.

	The project will pursue an integrated multimodal approach to support the long-term goal of achieving level 2 or 3 autonomy in robotic surgery, specifically automating the complex surgical gesture of suturing. Year one involves developing cognitive vision algorithms that interpret the surgical environment from endoscopic imaging, independent of robotic controls. Tasks will include real-time segmentation, object tracking, surgical action recognition, and depth estimation to characterize the surgical scenario accurately. In the second year, the candidate will focus on creating a realistic surgical simulation environment using tools such as Unity. These engines support high-quality 3D rendering and realistic lighting/shadows and can incorporate physics models for tissue deformation and tool interactions. This environment will support teleoperation for surgeon training and the development and training of autonomous agents. By leveraging such technology (as in modern surgical trainers), the candidate will build a hyper-realistic surgical training environment that includes anatomical models, surgical instruments, and even haptic feedback for realistic tool-tissue interaction. A known challenge is the sim-to-real gap by which policies or vision models trained in a virtual environment may not immediately generalize to real surgical scenes due to visual and dynamic differences. To overcome this, the project will incorporate domain adaptation techniques. Addressing the domain gap between

Objectives	simulated and real imagery through generative models will be a critical aspect, enabling more realistic and effective training conditions. In the final year, the emphasis shifts toward reconstructing real surgical scenarios within simulation environments. Using advanced 3D reconstruction methods and cognitive vision algorithms and by combining the 3D model of the patient's anatomy with the recorded instrument motion trajectories from the vision tracking system, the candidate will recreate the surgery in a virtual environment for review or analysis. This capability has a significant clinical and educational impact: surgeons and trainees could replay complex surgeries to study what transpired, analyze decision points, or identify causes of complications in a visually rich manner. Moreover, from a research perspective, realistic replays provide ground truth data to develop and validate new surgical AI algorithms by simulating them on past real scenarios. The candidate will collaborate closely with interdisciplinary teams, including surgeons, engineers, and AI specialists, fostering integration across clinical, computational, and robotic disciplines. The project will be mostly performed in the Minimally Invasive Therapy and Innovation Center (MITIC) at the Department of Surgical Sciences. This lab has a dynamic and collaborative environment. The candidate should be able to interface with different professional figures as well as be able to guide master's students who may be doing their final project at MITIC. The research will potentially result in significant contributions to surgical robotics literature, the development of novel methodologies, and direct clinical impact.

Skills and competencies for the development of the activity	Applicants must hold a Master's degree in Bioengineering with expertise in robotics and computer vision for endoluminal applications. Essential skills include proficiency in one or more areas: real-time segmentation algorithms (e.g., YOLO), simulation frameworks (Unity), endoscopic imaging, generative models, and AI/machine learning. Relevant scientific publications and conference participation are advantageous.