IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

Autonomous System for Tumor Resection (ASTR) - Dual-Arm Robotic Midline Partial Glossectomy

Jiawei Ge, Michael Kam, Justin D. Opfermann, Hamed Saeidi, Simon Leonard, Leila J. Mady, Martin J. Schnermann, and Axel Krieger, *Senior Member, IEEE*

***Abstract*—Head and neck cancers are the seventh most common cancers worldwide, with squamous cell carcinoma being the most prevalent histologic subtype. Surgical resection is a primary treatment modality for many patients with head and neck squamous cell carcinoma, and accurately identifying tumor boundaries and ensuring sufficient resection margins are critical for optimizing oncologic outcomes. This letter presents an innovative autonomous system for tumor resection (ASTR) and conducts a feasibility study by performing supervised autonomous midline partial glossectomy for pseudotumor with millimeter accuracy. The proposed ASTR system consists of a dual-camera vision system, an electrosurgical instrument, a newly developed vacuum grasping instrument, two 6-DOF manipulators, and a novel autonomous control system. The letter introduces an ontology-based research framework for creating and implementing a complex autonomous surgical workflow, using the glossectomy as a case study. Porcine tongue tissues are used in this study, and marked using color inks and near-infrared fluorescent (NIRF) markers to indicate the pseudotumor. ASTR actively monitors the NIRF markers and gathers spatial and color data from the samples, enabling planning and execution of robot trajectories in accordance with the proposed glossectomy workflow. The system successfully performs six consecutive supervised autonomous pseudotumor resections on porcine specimens. The average surface and depth resection errors measure 0.73 ± 0.60 mm and 1.89 ± 0.54 mm, respectively, with no positive tumor margins detected in any of the six resections. The resection accuracy is demonstrated to be on par with manual pseudotumor glossectomy performed by an experienced otolaryngologist.**

Manuscript received 18 September 2023; accepted 20 November 2023. Date of publication 12 December 2023; date of current version 26 December 2023. This letter was recommended for publication by Associate Editor L. Mattos and Editor J. Burgner-Kahrs upon evaluation of the reviewers' comments. This work was supported in part by the National Institutes of Health under Grants 1R01EB020610 and R21EB024707, in part by the National Science Foundation's Foundational Research in Robotics CAREER Program under Grant 2144348, and in part by the Intramural Research Program of the National Institutes of Health, National Cancer Institute, Center for Cancer Research. *(Corresponding author: Jiawei Ge.)*

Jiawei Ge, Michael Kam, Justin D. Opfermann, and Axel Krieger are with the Department of Mechanical Engineering, Johns Hopkins University, Baltimore, MD 21211 USA (e-mail: jge9@jhu.edu; mkam2@jhu.edu; jopferm1@jhu.edu; axel@jhu.edu).

Hamed Saeidi is with the Department of Computer Science, University of North Carolina Wilmington, Wilmington, NC 28403 USA (e-mail: saeidih@uncw.edu).

Simon Leonard is with the Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD 21211 USA (e-mail: sleonard@jhu.edu).

Leila J. Mady is with the Department of Otolaryngology - Head and Neck Surgery, Johns Hopkins School of Medicine, Johns Hopkins University, Baltimore, MD 21287 USA (e-mail: lmady1@jh.edu).

Martin J. Schnermann is with the Chemical Biology Laboratory, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, MD 21702 USA (e-mail: martin.schnermann@nih.gov).

Digital Object Identifier 10.1109/LRA.2023.3341773

***Index Terms*—Medical robots and systems, software architecture for robotic and automation, control architectures and programming.**

I. Introduction

HEAD and neck cancer (HNC) arises from various anatomic sites in the upper aerodigestive tract, such as the oral cavity, larynx, pharynx, nasal cavity, and salivary glands, ranking as the seventh most common cancer worldwide [1]. Squamous cell carcinoma (SCC), as the most common histological subtype, comprises over 90% of HNC cases [1]. Early-stage head and neck SCC (HNSCC) is primarily treated through surgical resection, where the malignant tumor is excised from the surrounding healthy tissue. Late-stage HNSCC treatment typically involves adjuvant radiotherapy or chemoradiotherapy in combination with surgical techniques [2]. Surgeons localize HNSCC through physical exam augmented by diagnostic imaging such as ultrasonography, computed tomography (CT), and magnetic resonance imaging (MRI). Intraoperatively, surgeons delineate the tumor's extent and the corresponding tumor-free margin by palpation, followed by performing resection to achieve complete tumor removal. Surgical margin status significantly impacts HNSCC oncologic outcomes, with excessively narrow margins increasing recurrence or metastasis risk, and overly wide margins potentially compromising post-surgical function and quality of life [3].

Robotic-assisted minimally invasive surgery (RAMIS), introduced to otolaryngology in the early 2000 s for managing head and neck pathologies, is commonly termed trans-oral robotic surgery (TORS) due to the oral cavity being an access point [4]. Several commercially available robot systems are food and drug administration (FDA) approved for TORS, including da Vinci S, X, and SP (Intuitive Surgical, Sunnyvale, CA), and Flex (Medrobotics, Raynham, MA). They are teleoperation systems that provide surgeons superior visualization, dexterity, precision, and ergonomics [4]. However, two limitations exist despite the advantages. First, surgeon burnout has a ubiquitous presence in surgery [5]. The design of surgeon's console in TORS helps ease but not eliminate surgeon burnout/fatigue, which might lead to major medical errors [6]. Second, surgeons can easily lose track of tumor edges if the surgical site is obscured due to intraoperative bleeding. Although electrosurgical instruments are widely used in TORS of HNC patients to reduce bleeding compared with scalpel settings, they further obfuscate the surgical site by

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

causing tissue charring. In both aforementioned scenarios, any surgical error would negatively impact surgical outcomes.

Research on integrating autonomy into surgery aims to reduce surgeon workload and mitigate errors from burnout [6]. Several previous studies have explored autonomous tumor resections. Kehoe et al. used the Raven robot for tumor fragment removal, achieving autonomous dual-arm debridement of foam rubbers with grippers [7], while Hu et al. performed semi-autonomous debridement of jelly-like mixtures using a suction tool [8]. McKinley et al. employed the da Vinci Surgical Research Kit (dVRK) to autonomously palpate, expose, and remove a cylindrical rubber shape embedded in a silicone rubber pad [9]. However, these works are challenging to apply to real tissue surgeries due to their reliance on unrealistic phantom models. Saeidi et al. and Ge et al. developed supervised autonomous controllers for precise 3D incisions on porcine tissues using a Kuka robot and an electrosurgical tool [10], [11]. While effective, these techniques were not extended to complete tumor excision. In a subsequent study, Ge et al. used an electrosurgical robot and a customized vacuum grasping system, introducing an open-loop depth estimation technique for full tumor resection. This approach confirmed feasibility, but also highlighted accuracy challenges inherent to open-loop estimations [12].

In the domain of autonomous robotic surgery, especially for complex procedures like tumor resection, variations in procedural definitions have emerged. For instance, while Hu et al. interpreted *resection* as the removal of jelly-like mixtures via suction [8], Ge et al. depicted it as tissue incision [11], diverging from standard clinical definitions. Such inconsistencies underscored the need for a common conceptualization, specifically ontology, in surgical autonomy to ensure aligning between engineering research and clinical practices. While IEEE championed the Core Ontology for Robotics and Automation (CORA) for a broader autonomous robotics domain [13], there were efforts to craft ontologies specifically for surgical workflows, such as the Surgical Ontology for Computer-Assisted Surgery (SOCAS), the Ontology for Data Integration in Surgery (ODIS), and the Ontology for Surgical Process Modeling (OntoSPM) [14]. Notably, OntoSPM stood out due to its collaborative development across European institutions, sustained dedication to ontology refinement, practical applications in ontology-based autonomous robotic surgery, and holistic perspective to surgical subtask automation [14], [15]. Its emphasis on ontology-based subtask granularity allows for encapsulating non-procedure-specific actions, laying the foundation for diverse autonomous surgical procedures [15].

Addressing challenges posed by intraoperative obfuscation from blood and charred tissue, advancements in HNSCC imaging techniques emerge. By conjugating antibodies with near-infrared fluorescent (NIRF) dyes, selective binding to HNSCC could be achieved, thereby increasing the tumor-to-background ratios (TBRs) [16]. Dot-shaped and indocyanine green (ICG)-based NIRF markers were developed to replace traditional tumor edge marking methods, such as electrocautery or ink pens, providing superior signal-to-noise ratios (SNRs) regardless of blood and tissue charring [17].

In this study, we developed an autonomous system for tumor resection (ASTR), and conducted autonomous pseudotumor resections on porcine tongue tissues under human supervision. Our first contribution is developing the ASTR, a dual-arm robot system designed for autonomous HNSCC resections. Specifically, the current scope of ASTR is primarily focused on early-stage, midline tongue tumors. The ASTR integrates a novel autonomous control strategy, soft tissue resection planners, and a new laparoscopic vacuum grasping robot, enhancing our previously developed laparoscopic electrosurgical robot and a dual-camera vision system [12]. The ASTR presents multiple advancements: 1) deformed tissue tracking for precise planning; 2) closed-loop vacuum robot control; 3) integration of industry-grade equipment for reliability; 4) integration of the spherical linear interpolation (SLERP) algorithm for robust orientation control [18]; 5) soft real-time control for timely response; and 6) a three-layer supervisory control ensuring procedural safety. The second contribution is establishing a novel ontology-based research framework for complex autonomous robotic surgeries. In essence, a surgical workflow should be formed using consistent terminologies (e.g. OntoSPM) in surgical process model (SPM) representations at subtask granularity [19], and paired with the finite state machine (FSM) state diagram for further procedural elaboration and assured implementation [20], as applied in our autonomous glossectomy case study. This approach provides autonomous robot surgery data for SPM, broadening its scope to encompass both human and robotic operators, and potentially influence management practices across operating rooms (ORs), hospitals, and overall patient care in the future. For our third contribution, we not only prove the feasibility of employing the ASTR within the ontology-based research framework, but also report the first study that accomplishes accurate supervised autonomous tumor resection on animal tissues. In a surgical site setup mimicking clinical glossectomy, the autonomous surgeries succeed in all six consecutive experiments under human supervision, creating no positive margins. The resection accuracy exhibits great improvement in contrast to our prior study [12]. Comparing with three manual resections performed by an experienced otolaryngologist, both the autonomous and manual surface incisions display sub-millimeter accuracy. Autonomous resections show slightly better depth dissection accuracy than manual studies, but the difference isn't statistically significant.

II. Procedure of Glossectomy

The HNSCC resection workflows are influenced by the tumor's location, size, and depth. This study focuses on the midline partial glossectomy, a procedure employed to address early-stage, small-sized SCC that primarily arises on the midline superior surface of the tongue. Such a glossectomy is classified as a type 1, or mucosectomy [21]. A clinical report on an SCC (T1N0M0) glossectomy offers valuable insights into this procedure [22]. This procedure essentially resembles the excavation of a shallow mass growing from the tissue surface, and its progression into deeper tissue layers.

The choice of glossectomy is influenced by three factors. First, it fits the HNSCC resection introduction, which covers

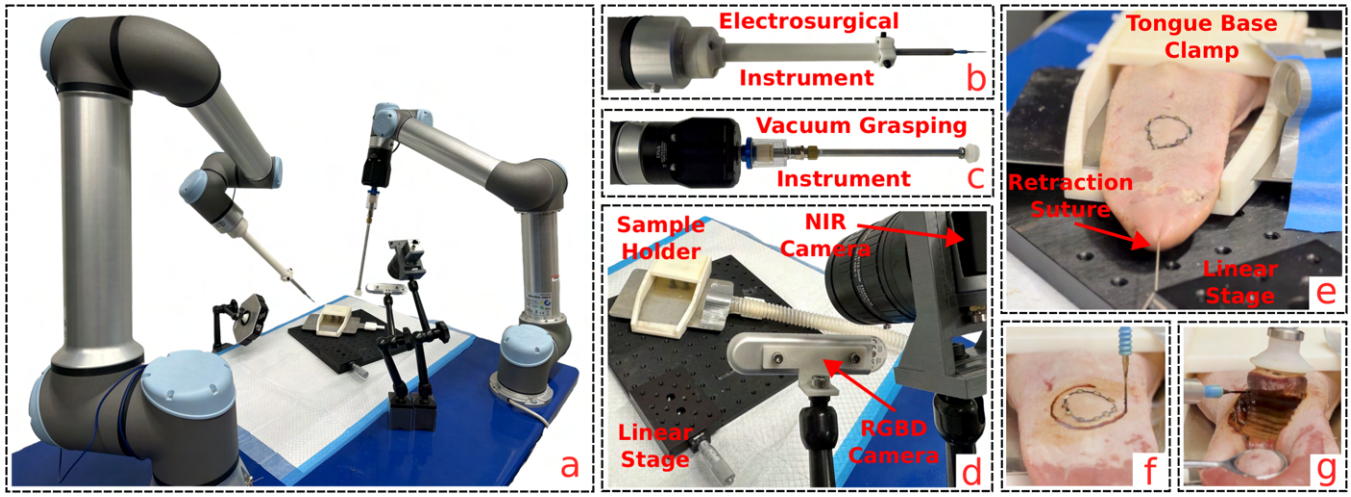

Fig. 1. Images of the (a) ASTR, (b) monopolar electrosurgical instrument, (c) vacuum grasping instrument, d) dual-camera vision system, sample holder, grounding pad, smoke evacuation tube, linear motion stage, and (e) simulated clinical setting featuring a porcine tongue specimen stretched using retraction sutures. The close views during the (f) surface incision, and (g) deep margin dissection for a pseudotumor on a porcine tongue tissue.

tongue SCC resection. Second, midline partial glossectomy shares characteristics with other HNSCC surgeries, like those targeting the hard palate, tongue base, tonsil, and oropharynx. Third, its transoral approach offers a more spacious workspace than surgeries in deeper head and neck regions, making it an appropriate initial step toward autonomous HNSCC resection. In practice, surgeons use oral gags to open the mouth, and traction sutures at the tongue tip to pull the anterior tongue out of the oral cavity [23], [24]. The standard method involves a hand-held electrocautery pen and surgical forceps to remove lesions [23], [24], approximating an open surgical setting.

In this study, we designed a pseudotumor on a porcine tongue to simulate an early-stage SCC that invades the mucosa, akin to a *tumor chip* [21]. SCC surface boundaries are derived from clinical reports, and marked on the porcine tongue's midline surface using black ink. The pseudotumor has a surface dimension of 2 cm and an invasion depth of $\ll$ 1 mm [21], [24]. Taking clinical resection margins into account, we aim to extend beyond the pseudotumor by a margin of 5 mm both laterally and in depth [25].

III. Methods

A. Autonomous System for Tumor Resection (ASTR)

The ASTR, an autonomous, dual-arm, vision-guided robotic system, adopts certain technical insights related to low-level control from the STAR system [26]. While STAR focuses on anastomosis, ASTR is specifically tailored for tumor resection, representing a unique application.

1) ASTR Hardware Architecture: Fig. 1(a)–(d) demonstrates the ASTR. To facilitate translation to clinical settings, we primarily integrate products from industry-leading manufacturers. For example, UR manipulators (Universal Robots, Odense, Denmark) are used in several FDA-approved surgical robots like the TMS-Cobot (Axilum Robotics, Strasbourg, France).

The primary instrument for tumor resection, the electrosurgical instrument, is used to cut tissues. It consists of a standard 25 mm length and 1 mm diameter monopolar electrode (Bovie, Clearwater, FL), a customized 35 cm laparoscopic extension, a grounding pad, and an electrosurgical generator (ASG-300ESU, DRE Veterinary, Louisville, KY). A portable smoke evacuator (Smoke Shark, Bovie) is also included to reduce electrosurgical smoke. The vacuum grasping instrument serves as a secondary tool for maneuvering tissues and exposing the area of interest for easier cutting, with its design detailed in Section III-A2. The electrosurgical and vacuum grasping instruments are mounted on the flanges of a UR10e and a UR5 robotic manipulators, respectively. These manipulators carry and direct the instruments to reach various positions and orientations within the surgical sites. The vision system incorporates an RGBD camera (D415, Intel, Santa Clara, CA), a 2D near-infrared (NIR) camera (acA2040-90umNIR, Basler AG, Ahrensburg, Germany), a 845 ± 27.5 nm band-pass filter (Chroma Technology, Bellows Falls, VT), and a 760 nm high power light-emitting diode (North Coast Technical, Chesterland, OH). The NIR camera and light source are chosen to achieve high SNRs of ICG, the most frequently used FDA-approved NIRF dye [17]. Images from both cameras are fused, allowing the retrieval of 3D positions of color and NIR signals to create surgical plans and guide the robots. The development details of the electrosurgical tool and vision system are explained in our previous work [10], [11], [12].

2) Vacuum Grasping Instrument: The vacuum grasping is an alternative to forceps grasping for soft tissue manipulation, reducing potential tissue damage [27]. For tumor resections, this approach reduces the risk of inadvertently breaking tumor tissues, leaving undetected fragments, and potentially causing cancer recurrence. Furthermore, since SCC originates from the epithelium, using vacuum grasping on the tumor surface simultaneously immobilizes the tumor and retracts the tissue.

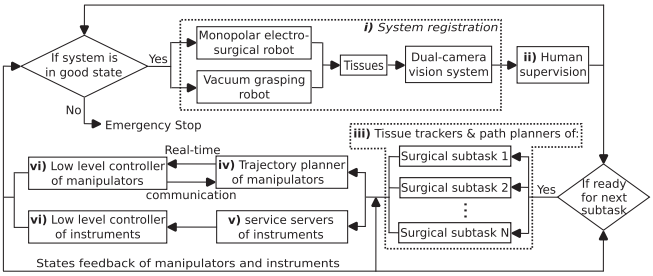

Fig. 2. Control diagram of the ASTR.

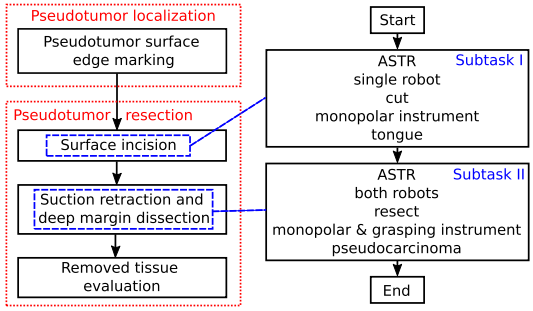

Fig. 3. Simplified workflow of the midline partial glossectomy (left), and the workflow using SPM representations at subtask granularity (right).

The vacuum grasping instrument is comprised of an electric vacuum gripper (EPick, Robotiq, Levis, Canada), a moisture trap, a vacuum pad, and a stainless steel extension tube and fittings (Fig. 1(c)). The tube has a length of 35 cm and an outer diameter (OD) of 10 mm. The soft Bellows vacuum pad has an 12 mm inner diameter (ID) and a 24 mm OD, with food grade FDA compliant replacement options for future use. An in-line moisture trap connects the tube and EPick using pipe fittings, preventing blood, water, and tissue fluid from being sucked into the EPick and causing damage or corrosion to the internal electric panels. Custom-programmed robot operating system (ROS) service servers, developed in Python, control EPick vacuum between continuous on and off modes. The negative pressure is preset at -70 kPa while the vacuum mode is on, with the theoretical vacuum force on the 12 mm ID Bellows pad equaling 7.92 N.

3) Autonomous Control Strategy: The system control software is developed using the ROS and Open Robot Control Software (OROCOS) frameworks, and operates on RT_PREEMPT Linux for real-time functionality. This architecture is also employed in dVRK, a pioneering surgical robotics project with documented bench-to-bedside initiatives [28].

The ASTR's control strategy encompasses six components (Fig. 2). Initially, a checkerboard is placed near the target tissue serving as a shared *world* coordinate frame. Following standard hand-eye calibrations for all robots and cameras, the frames of each component are integrated, aligning them with this *world* frame for a unified system registration. Consequently, the dual-camera fusion is achieved. For safe autonomous surgery executions, a human supervisor oversees the ASTR system in three main ways: i) They monitor live image streams, determining if the system is prepared to execute the next subtask; ii) They review and approve initial plans before the robot executes a subtask, requesting changes if necessary; iii) They can immediately halt all robot operations if any irregularities are detected. Once the initial checks are done, the ASTR's tracking and planning modules take over the control. These modules are tailored to their respective surgical workflows. Two subtask planners are developed in this study (explained in Section III-B2), but more can be developed and integrated into the ASTR. Commands generated by these planners can be position and orientation targets forwarded to the robot's trajectory planner, or requests for specific instrument functions directed to ROS service servers. The trajectory planner incorporates robot kinematics, and controls speed modes like fast, accurate, or idle to cater to different needs. This ensures smooth robotic movements towards targets, adapting in real-time, even if a new target is received before the current one is reached. The ROS services for vacuum functions can be called and activated, and a confirmation signal is received from the EPick low-level controller regarding the vacuum's functional status. Upon completion of a surgical subtask, the human supervisor decides whether to proceed with the next subtask. Iterations continue until the final surgical subtask is completed.

The trajectory planner is implemented using OROCOS Real-Time Toolkit (RTT) for real-time capabilities, OROCOS Kinematics and Dynamics Library (KDL) for kinematics calculations, and Reflexxes Motion Library (RML) for instantaneous robot joint trajectory interpolations. The SLERP algorithm is integrated to enhance orientation control, utilizing quaternion representations. This approach not only ensures the shortest path between two orientations, but also avoids challenges associated with Euler angle interpolation, such as gimbal lock and ambiguities where multiple Euler angle sets produce the same orientation [18]. In terms of motion parameters, The maximum linear velocity along each Cartesian axis for the fast and accurate speed modes is set to 6 cm/sec and 2 mm/sec, respectively, while the maximum angular velocity is set to 15°/sec and 5°/sec, respectively.

B. A Research Framework for Autonomous Robotic Surgeries

In this paper, we propose a two-step research framework to design and implement an autonomous robotic surgery workflow. First, we use ontologies to formalize the workflow using SPM representations at subtask level. Second, we utilize FSM techniques to further elaborate the procedure and achieve implementation [20]. OntoSPM ontology and midline partial glossectomy are chosen to illustrate this approach.

1) Ontology-Based Surgical Workflow Using SPM Representations At Subtask Granularity: The clinical workflow of midline partial glossectomy can be decomposed into three tasks: diagnostic imaging, tumor resection, and surgical site reconstruction. This resection study focuses on the first two tasks, and simplifies the workflow at subtask level for the pseudotumor model (Fig. 3) [12]. To clarify, TBRs of a pseudotumor equal one, making medical imaging and pathological evaluation of tumor resection infeasible. The workflow is then formalized

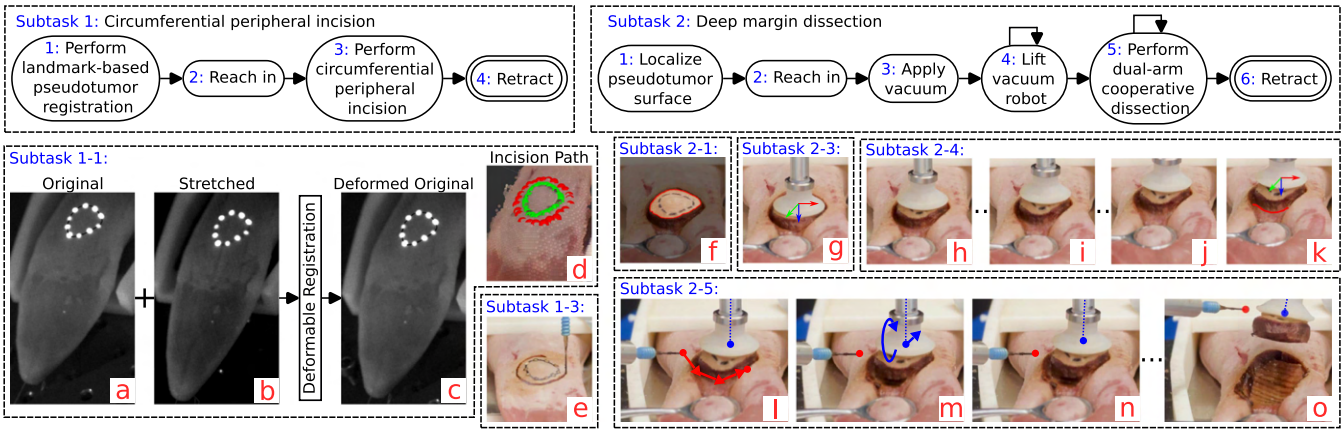

Fig. 4. FSM state diagram of pseudotumor surface incision (upper left), and deep margin dissection (upper right). NIR images of (a) a porcine tongue featuring drawn pseudotumor contours and ten NIR markers, (b) an elongated tongue, (c) an estimated tumor bed after landmark-based deformable registration, and (d) a point cloud image displaying 3D overlay of pseudotumor contours in green and robotic incision path in red after applying a 5 mm surface offset. Color images of (e) surface incision, (f) detected pseudotumor contours in red after incision, g) vacuum grasping instrument application, (g-k) iterations of autonomous pseudotumor deep margin exposure via lifting vacuum grasping instrument, k) the front edge of pseudotumor's deep margin highlighted in red, (l-m) deep margin dissection movement pattern highlighted in red, and m-n) tissue retraction motion pattern highlighted in blue during the first iteration of dual-arm cooperative dissection, o) post complete pseudotumor removal. All images are from the autonomous resection #1.

using OntoSPM terminologies, resulting in a SPM activity diagram (Fig. 3). Semantic adjustments are made for clarity, such as substituting *grasping instrument* for *suction instrument*, because the latter is defined as: a sucking instrument has the function to suck and is used to remove blood, fluid or debris from operative sites. Since traditional SPM nodes exclusively represent human operations, three non-ontology terms (i.e. ASTR, robot, single) are used to indicate robot operators without distinguishing between left or right hand. When integrated with the surgery time from experimental results for SPM nodes, the reported autonomous robotic surgeries can be used for SPM.

2) Subtasks Implementation and Detailed Workflow: Fig. 4 shows the FSM state diagrams of the two subtasks, and representative images for better illustrating the workflow.

Subtask 1 is a four-state FSM that employs both cameras and the electrosurgical robot. Preoperatively, a NIR image is taken to record the original tumor bed (Fig. 4(a)). The porcine tongue is then stretched from the tongue tip by 2 cm to imitate the clinical workflow of exposing tumor, and another NIR image is taken (Fig. 4(b)). A landmark-based deformable image registration technique, using Insight Segmentation and Registration Toolkit (ITK) libraries, is performed to recover the 2D deformed pseudotumor surface boundary (Fig. 4(c)) [11]. The 3D positions of the deformed pseudotumor boundary are generated from the dual-camera image fusion design, and downsampled to waypoints with an average distance of 2 mm (green dots in Fig. 4(d)). Then, a point 10 mm above the center of waypoints is planed as the reach-in goal for the electrosurgical instrument. The surface incision path is planned by applying a 5 mm lateral offset (red dots in Fig. 4(d)), and a 5 mm incision depth (Fig. 4(e)) to all waypoints. The reach-in point and incision path are executed sequentially, with the electrosurgical instrument maintaining a vertical orientation (Fig. 4(e)). The electrosurgical robot then retracts to its initial configuration to prepare the next subtask.

Subtask 2 is a six-state FSM that employs the RGBD camera and both robots. Once the human supervisor evaluates the performance of subtask 1 and deems it satisfactory, a standard Collins tongue forceps is manually applied on porcine tongue to expose the pseudotumor deep margin (Fig. 4(f)). A color-based region growing segmentation technique, utilizing Point Cloud Library (PCL) libraries, is employed to detect the current pseudotumor surface contours (Fig. 4(f)). Both robots then move near the sample concurrently in preparation for deep margin dissection. The pseudotumor surface center point and a vertical orientation are used as the target pose of the vacuum grasping instrument. Once the target is reached, the vacuum is turned on by executing the customized Robotiq vacuum ROS service (Fig. 4(g)). Subsequently, the vacuum grasping instrument is raised in 2 mm increments to reveal the front edge of the pseudotumor's deep margin to the RGBD camera (Fig. 4(g)–(k)). The exposed area of pseudotumor's side margin is calculated at each step using a simple HSV thresholding of the dark brown color beneath the vacuum instrument. If the area stops increasing after one step during tissue retraction, the side margin is considered fully exposed (i.e. Fig. 4(j)–(k). The front edge of the pseudotumor's deep margin is acquired by performing a vertical slicing of 1 mm thickness to the side margin point cloud in the grasping instrument frame, and keeping the lowest slice (Fig. 4(k)). As the end effector frames of the instruments can be retrieved from the robotic kinematics, the complete pseudotumor deep margin contours are generated by projecting the pseudotumor surface contours (Fig. 4(f)) onto the xy-plane of the grasping instrument frame (Fig. 4(g)), and offsetting to the depth of the aforementioned lowest slice (Fig. 4(k)). The pseudotumor deep margin contours are further downsampled with an average distance of 2 mm along the y axis of the grasping instrument frame, and grouped into left and right pairs. The electrosurgical instrument is guided by the paired poses with a 20° downward-tilting orientation to perform the deep margin dissection in iterations

Fig. 5. Pseudotumor surface shapes for six autonomous resections (#1 to #6) derived from three clinically reported tumor surface shapes. The rulers included in the images display 1 mm increments.

(Fig. 4(l)–(m)). The vacuum instrument is then lifted by 1 mm, with horizontal offset from the electrosurgical instrument by 1 mm, and tilted towards the RGBD camera by 2° for improved tissue retraction (Fig. 4(m) and (n)). This repetitive dual-arm cooperative dissection process continues until all paired poses are traversed, and the pseudotumor is removed from the sample (Fig. 4(o)). Finally, both robots retract concurrently to their initial configurations.

IV. Experiments and Results

A. Sample and Surgical Site Preparation

Three distinct oral tongue SCC surface shapes are identified from clinical case reports, and scaled to fit in a 2 cm diameter bounding circle, corresponding to T1/T2 stage SCC dimensions [24], [29], [30], [31]. Nine porcine tongues ($N = 9$), obtained from a grocery store (H Mart, Lyndhurst, NJ), are used as samples due to their anatomical resemblance to human tongues. Each surface contour is manually drawn on the anterior midline of two porcine tongues ($N = 2 \cdot 3 = 6$) for autonomous resections using a black ink pen and laser-cut pattern films, with ten NIRF markers evenly distributed along each contour (Fig. 5) [17]. Each surface contour is manually drawn on one porcine tongue ($N = 1 \cdot 3 = 3$) for manual resections as comparison studies using a black ink pen.

A clinically-mimicked glossectomy setup is used for both autonomous and manual resections (Fig. 1(e)) [23], [24]. We 3D printed a sample holder to replicate the jawbone and mouth floor, and added a clamp on top to stabilize posterior tongue motion. Before each test, tongue tips are sutured and extended by 2 cm using a linear stage. In autonomous resections, after subtask 1 execution, a Collin tongue forceps is manually clamped on the anterior tongue's edge (Fig. 4(f)). This technique exposes the front edge of the pseudotumor deep margin during subtask 2, while minimizing deformation from vacuum tissue retraction. In manual resections, the otolaryngologist uses a handheld Bovie electrosurgical pen, a surgical forceps, and a ruler, bypassing the ASTR components [23], [24]. The ruler is actively used during resection to provide surgeons with dimensional feedback in millimeters.

B. Evaluation Metrics

Upon pseudotumor removal, images of the top, left, and right views are captured for all nine studies. Using a ruler in these images for pixel-to-millimeter scaling, we measure a resolution of approximately 0.02 mm. A custom Python script,

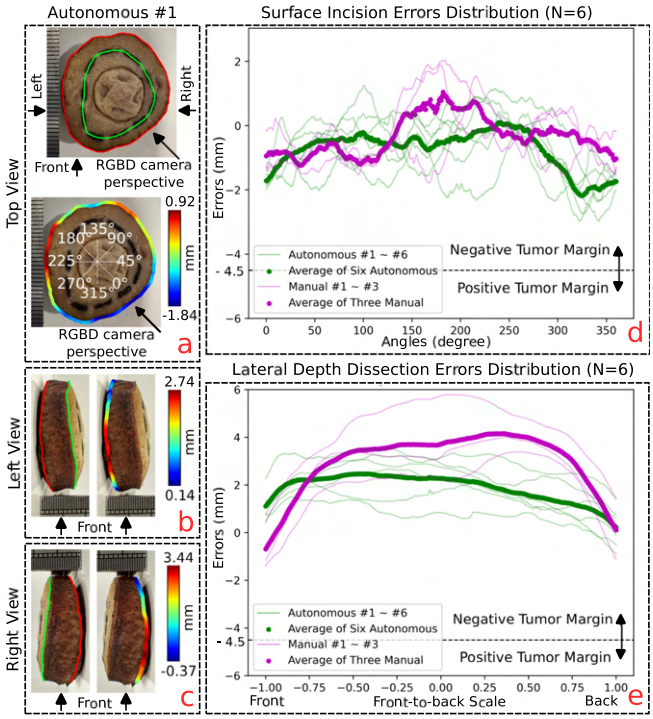

Fig. 6. Color images of the a) top, b) left, and c) right views of removed pseudotumor from the autonomous resection #1. The actual cut and reference contours are highlighted in red and green colors, respectively. Jet color mapped errors are also highlighted. Plots of the d) surface incision errors distribution, and e) lateral depth dissection errors distribution for both autonomous and manual resections.

developed using OpenCV libraries, processes the top view images to calculate surface incision errors [10]. The incision error for each point on the actual incision contour (**ic**, red in Fig. 6(a)) is determined by the distance to closest point on the pre-drawn pseudotumor contour (**pdc**, green in Fig. 6(a)), subtracting the desired surface margin of 4.5 mm (due to a 5 mm offset and a 0.5 mm radius electrosurgical electrode). This is mathematically represented as $(\min_j \|\mathbf{ic}_i - \mathbf{pdc}_j\|_2 - 4.5)$ mm for each point i on the incision contour. Similarly, a custom Python script calculates depth dissection errors using the left and right view images [10]. The dissection error for each point on the actual dissection contour (**dc**, red in Fig. 6(b), (c)) is measured by the distance to closest point on the surface contour (**sc**, green in Fig. 6(b), (c)), subtracting the 4.5 mm depth margin. This is represented as $(\min_j \|\mathbf{dc}_i - \mathbf{sc}_j\|_2 - 4.5)$ mm for each point i on the dissection contour. The average, standard deviation, maximum, and minimum of surface incision and depth dissection errors for each resection experiment will be calculated and reported. The two-tailed t-test will assess differences in incision and dissection errors between autonomous and manual resections, with a significance level of 0.05.

Clinically, successful tumor resection requires complete tumor removal with negative tumor margins. In this study, a glossectomy is deemed successful if: i) the entire pseudotumor is removed, ii) its surface contour remains outside the pre-drawn contour (i.e. minimum surface incision error > -4.5 mm), and iii) its depth contour doesn't intersect the surface contour (i.e.

TABLE I
TABLE OF RESULTS (A FOR AUTONOMOUS RESECTIONS, AND M FOR MANUAL RESECTIONS)

			A1	A2	A3	A4	A5	A6	**A Average**	M1	M2	M3	**M Average**
Subtask 1: Circumferential Peripheral Incision	Errors (mm)	Mean	-0.51	-0.67	-1.10	-0.64	-1.04	-0.31	**-0.73**	-0.22	-0.21	-0.57	**-0.32**
		Std	0.77	1.14	0.84	0.64	0.73	0.93	**0.60**	0.86	0.98	0.66	**0.60**
		Max Margin	0.92	1.26	0.37	0.44	0.17	1.17	**0.10**	1.90	2.03	1.56	**1.04**
		Min Margin	-1.84	-2.97	-2.75	-1.92	-2.97	-2.40	**-2.18**	-1.54	-1.51	-1.78	**-1.24**
	Procedure Duration (s)		80.52	81.53	78.87	82.88	80.67	81.18	**80.94**	98	90	135	**107.67**
	Number of Replannings		3	1	4	1	3	1	**2.17**	-	-	-	-
	Computation Time for Replannings (s)		0.96	0.67	1.88	0.84	2.70	0.67	**1.29**	-	-	-	-
Subtask 2: Deep Margin Dissection	Errors (mm)	Mean	2.00	2.75	1.52	0.76	1.62	2.68	**1.89**	1.97	4.31	2.79	**3.02**
		Std	0.87	0.49	0.36	0.47	0.82	0.69	**0.54**	1.18	1.33	1.20	**1.19**
		Max Margin	3.08	3.34	2.19	1.51	2.66	3.55	**2.47**	3.40	5.79	4.01	**4.15**
		Min Margin	-0.11	1.31	0.64	-0.90	-0.50	0.76	**0.22**	-1.39	0.46	-1.16	**-0.69**
	Procedure Duration (s)		496.27	517.68	476.55	486.34	471.92	491.00	**489.96**	118	210	63	**130.33**
	Number of Replannings		3	3	12	4	5	14	**6.83**	-	-	-	-
	Computation Time for Replannings (s)		6.99	11.00	5.10	2.00	12.61	9.98	**7.95**	-	-	-	-

The bold values emphasize the average results for both autonomous and manual resections.

minimum depth dissection error > -4.5 mm). We will report the procedure duration for each subtask within each experiment. We will also document the number of supervisor requested replannings, and the computation time for replannings at subtasks 1-1 and 2-1.

C. Results

The results of all nine experiments are presented in Fig. 6 and Table I. In six autonomous experiments, the average surface incision and depth dissection errors were -0.73 ± 0.60 and 1.89 ± 0.54 mm, respectively. The average autonomous surgery duration for subtask 1 and 2 were around 1 min 21 sec and 8 min 10 sec, respectively. No emergency stop of the ASTR was triggered by the human supervisor. In three manual experiments, the average surface incision and depth dissection errors were -0.32 ± 0.60, and 3.02 ± 1.19 mm, respectively. The average manual surgery duration for subtask 1 and 2 were around 1 min 48 sec and 2 min 10 sec, respectively. All six autonomous and three manual glossectomies successfully removed the pseudotumors, with no positive tumor margin identified in any of the nine samples. Both autonomous and manual resections achieved sub-millimeter accuracy in the surface incision, with no significant difference ($p = 0.098$). While autonomous resections showed better deep margin dissection accuracy than manual resections (1.89 ± 0.54 mm vs. 3.02 ± 1.19 mm), the difference wasn't significant ($p = 0.118$). Our current autonomous resection results showed accuracy improvements over the previous study, which reported averages of -1.19 mm for surface incision, and -1.83 mm for depth dissection errors [12]. Two-tailed t-tests indicated near significance for surface incision ($p = 0.052$), and definitive significance for depth dissection improvements ($p < 0.001$).

V. DISCUSSION

In this study, we have identified several key observations from the resection results. First, both autonomous and manual resections of pseudotumors exhibit a trend of narrower average surface margins (-0.73 mm and -0.32 mm, respectively) and thicker depth margins (1.89 mm and 3.02 mm, respectively). The former could be influenced by the electrosurgical electrode vaporizing more tissue than its radius due to heat effects, while the latter might result from the stretching of dead tissues during the procedure, causing removed pseudotumors to be thicker than initially intended. Second, a spike in autonomous surface resection errors near the cameras' perspectives suggests the need for improved camera setup, calibration, and dual-camera fusion in future work (Fig. 6(a)). Third, large autonomous depth resection inconsistencies occur at the front and back of the removed pseudotumor (Fig. 6(e)). The former may result from the initial deep margin dissection process using a 20° downward-tilting electrode angle, while the latter could be due to decreasing retraction force in the latter half of the dissection iterations. Fourth, subtask 1-3 in manual resection is a repetitive task. Surgeons lack robot precision and vision feedback in units, and require repetitive ruler measurements for accurate margin control. As a result, autonomous incisions averaged 1 min 21 sec, compared to the manual incisions which took 1 min 48 sec. Fifth, the hand motion in manual resections is significantly faster than the ASTR's preset 2 mm/sec speed for precise movements. Consequently, the average time for deep margin resection was reported 8 min 10 sec in autonomous operations, compared to 2 min 10 sec for manual procedures. Future work should evaluate and optimize the ASTR's speed settings to potentially decrease resection duration, while achieving a comparable or improved resection accuracy.

Automating a complex surgery entails the autonomous and continuous execution of a detailed, lengthy sequence of predefined surgical workflows. Errors can accumulate throughout the procedure, as the image processing and robot control algorithms in each sequence interact. Consequently, a failure in one part of the algorithmic sequence can jeopardize the entire procedure. Ensuring near 100% success rates and improved accuracy for each algorithm is crucial for both safety and effectiveness. Another challenge arises from the depth camera's inaccuracy at incision sites, where deep gaps and small openings hinder the camera from providing millimeter-accurate depth information. Robot kinematics can be employed alongside precise camera data to address this issue. For instance, we use the accurate depth of the pseudotumor surface at subtask 2-1, precise color projection of the pseudotumor deep margin at subtask 2-4, and accurate kinematics of the vacuum grasping robot during the six states of subtask 2 to determine the accurate depth of the pseudotumor deep margin. The ASTR and experimental design exhibit two primary limitations. First, the pseudotumor model, while useful, is an oversimplified representation of HNSCC, resembling a

tumor chip. Consequently, our evaluation metrics, tailored to a pseudotumor with a TBR of 1, might not translate directly to clinical scenarios, particularly for pathological evaluations. Second, by focusing on the midline partial glossectomy, which mirrors an open surgical environment without simulating oral access constraints, we haven't fully address challenges related to HNSCC resection in deeper regions of the head and neck.

VI. Conclusion

This letter reports the first accurate supervised autonomous tumor resection using animal tissues. Our primary contributions include the development of the ASTR for autonomous HNSCC resections, where we integrate vacuum grasping as an alternative to traditional forceps, aiming to reduce potential tissue damage and subsequent cancer recurrence. We also develop a safe and reliable control strategy for the autonomous robot system under human supervisory. Furthermore, we establish an ontology-based research framework for complex autonomous robotic surgeries, which bridges medical and programming knowledge, offering a structured and accurate representation of procedures and ensuring their robotic implementation. This framework provides autonomous robotic surgery data for SPM, extending its reach to include both human and robot operators, and holds the potential to reshape future OR management practices. Lastly, our results from supervised autonomous tumor resections on porcine tissues demonstrate a precision that competes with, and at some aspects exceeds, manual resections performed by an experienced otolaryngologist. This also underscores the feasibility of the proposed ontology-based research framework.

For future work, the ASTR will be utilized to perform HNSCC resections in more realistic and challenging scenarios, such as an autonomous tongue base tumor resection with blood obfuscation and spatial constraints. The ASTR will be integrated with advanced deep learning-based vision techniques to improve the system's perception capabilities. We will also develop more subtask planners, and employ a more complex control scheme to provide decision-making capabilities, such as pausing resections to clear blood. Force sensors will be integrated to the instrument tips to measure the cutting force, allowing path planning adjustments to increase safety. These future directions will enhance the efficacy of the proposed method in real-world surgical settings.

References

[1] H. Sung et al., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer J. Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.

[2] N. Sadeghi, N.-W. Li, M. R. Taheri, S. Easley, and R. S. Siegel, "Neoadjuvant chemotherapy and transoral surgery as a definitive treatment for oropharyngeal cancer: A feasible novel approach," *Head Neck*, vol. 38, no. 12, pp. 1837–1846, Dec. 2016.

[3] M. M. Li, S. V. Puram, D. A. Silverman, M. O. Old, J. W. Rocco, and S. Y. Kang, "Margin analysis in head and neck cancer: State of the art and future directions," *Ann. Surg. Oncol.*, vol. 26, no. 12, pp. 4070–4080, 2019.

[4] A. Tamaki, J. W. Rocco, and E. Ozer, "The future of robotic surgery in otolaryngology–head and neck surgery," *Oral Oncol.*, vol. 101, 2020, Art. no. 104510.

[5] T. D. Shanafelt et al., "Burnout and career satisfaction among american surgeons," *Ann. Surg.*, vol. 250, no. 3, pp. 463–471, 2009.

[6] T. D. Shanafelt et al., "Burnout and medical errors among american surgeons," *Ann. Surg.*, vol. 251, no. 6, pp. 995–1000, 2010.

[7] B. Kehoe et al., "Autonomous multilateral debridement with the raven surgical robot," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2014, pp. 1432–1439.

[8] D. Hu, Y. Gong, E. J. Seibel, L. N. Sekhar, and B. Hannaford, "Semi-autonomous image-guided brain tumour resection using an integrated robotic system: A bench-top study," *Int. J. Med. Robot. Comput. Assist. Surg.*, vol. 14, no. 1, 2018, Art. no. e1872.

[9] S. McKinley et al., "An interchangeable surgical instrument system with application to supervised automation of multilateral tumor resection," in *Proc. IEEE Int. Conf. Automat. Sci. Eng.*, 2016, pp. 821–826.

[10] H. Saeidi et al., "Supervised autonomous electrosurgery via biocompatible near-infrared tissue tracking techniques," *IEEE Trans. Med. Robot. Bionics*, vol. 1, no. 4, pp. 228–236, Nov. 2019.

[11] J. Ge, H. Saeidi, J. D. Opfermann, A. S. Joshi, and A. Krieger, "Landmark-guided deformable image registration for supervised autonomous robotic tumor resection," *Med. Image Comput. Comput. Assist. Interv.*, vol. 11764, pp. 320–328, 2019.

[12] J. Ge, H. Saeidi, M. Kam, J. Opfermann, and A. Krieger, "Supervised autonomous electrosurgery for soft tissue resection," in *Proc. IEEE 21st Int. Conf. Bioinf. Bioeng.*, 2021, pp. 1–7.

[13] P. J. Gonçalves et al., "IEEE standard for autonomous robotics ontology [Standards]," *IEEE Robot. Automat. Mag.*, vol. 28, no. 3, pp. 171–173, Sep. 2021.

[14] B. Gibaud et al., "Toward a standard ontology of surgical process models," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 13, no. 9, pp. 1397–1408, 2018.

[15] D. Á. Nagy, T. D. Nagy, R. Elek, I. J. Rudas, and T. Haidegger, "Ontology-based surgical subtask automation, automating blunt dissection," *J. Med. Robot. Res.*, vol. 3, no. 03n04, 2018, Art. no. 1841005.

[16] S. van Keulen et al., "The clinical application of fluorescence-guided surgery in head and neck cancer," *J. Nucl. Med.*, vol. 60, no. 6, pp. 758–763, 2019.

[17] J. Ge et al., "A novel indocyanine green-based fluorescent marker for guiding surgical tumor resection," *J. Innov. Opt. Health Sci.*, vol. 14, no. 03, May 2021, Art. no. 2150013.

[18] E. G. Hemingway and O. M. O'Reilly, "Perspectives on euler angle singularities, gimbal lock, and the orthogonality of applied forces and applied moments," *Multibody Syst. Dyn.*, vol. 44, no. 1, pp. 31–56, 2018.

[19] T. Neumuth, "Surgical process modeling," *Innov. Surg. Sci.*, vol. 2, no. 3, pp. 123–137, 2017.

[20] R. Balogh and D. Obdržálek, "Using finite state machines in introductory robotics," in *Proc. Int. Conf. Robot. Educ.*, 2017, pp. 85–91.

[21] M. Ansarin et al., "Classification of glossectomies: Proposal for tongue cancer resections," *Head Neck*, vol. 41, no. 3, pp. 821–827, 2019.

[22] T. Toh, T. Akita, M. Saito, and T. Shigetomi, "Squamous cell carcinoma initially arising in the midline of the dorsum of the tongue in a young adult: A case report and review of the literature," *J. Oral Maxillofac. Surg., Med. Pathol.*, vol. 31, no. 3, pp. 185–188, 2019.

[23] B. J. Bailey, *Atlas of Head & Neck Surgery–Otolaryngology*. Philadelphia, PA, USA: Lippincott Williams & Wilkins, 2001.

[24] J.-L. M. Bigcas and O. T. Okuyemi, "Glossectomy," in *StatPearls [Internet]*. St. Petersburg, FL, USA: StatPearls Publishing, Dec. 2023. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK560636/

[25] M. Alicandri-Ciufelli et al., "Surgical margins in head and neck squamous cell carcinoma: What is 'close?," *Eur. Arch. Oto-Rhino- Laryngol.*, vol. 270, pp. 2603–2609, 2013.

[26] H. Saeidi et al., "Autonomous robotic laparoscopic surgery for intestinal anastomosis," *Sci. Robot.*, vol. 7, no. 62, 2022, Art. no. eabj2908.

[27] D. Vonck, R. Goossens, D. Van Eijk, I. De Hingh, and J. Jakimowicz, "Vacuum grasping as a manipulation technique for minimally invasive surgery," *Surg. Endoscopy*, vol. 24, no. 10, pp. 2418–2423, 2010.

[28] F. Richter, E. K. Funk, W. S. Park, R. K. Orosco, and M. C. Yip, "From bench to bedside: The first live robotic surgery on the dVRK to enable remote telesurgery with motion scaling," in *Proc. IEEE Int. Symp. Med. Robot.*, 2021, pp. 1–7.

[29] T. Suga et al., "Case report: Hidden oral squamous cell carcinoma in oral somatic symptom disorder," *Front. Psychiatry*, vol. 12, 2021, Art. no. 651871.

[30] P. Stanko et al., "Squamous cell carcinoma and piercing of the tongue–a case report," *J. Cranio- Maxillofac. Surg.*, vol. 40, no. 4, pp. 329–331, 2012.

[31] S. K. Hirota, D. A. Migliari, and N. N. Sugaya, "Oral squamous cell carcinoma in a young patient: Case report and literature review," *Anais Brasileiros de Dermatologia*, vol. 81, pp. 251–254, 2006.