Vision-Based Autonomous Surface Sampling in Healthcare Robotics: An Integrated Perception–Control Framework with Experimental Validation

Francesco Pio Briuolo[1]*, Alberto Maria Di Giacinto[2]*, Joseph Lovecchio[2], Oreste Riccardo Natale[3], Carmen Del Vecchio[1], Leandro Pecchia[2,4]

***Abstract*— Healthcare-Associated Infections (HAIs) remain a critical challenge, with surface contamination representing a key transmission vector. Current sampling procedures are performed manually, exposing operators to risks and introducing variability in data collection.**

The proposed pipeline combines YOLOv8-based target detection, surface localization, depth-derived geometric representation and fixed-base swabbing trajectory validation. Autonomous base navigation is treated as an upstream capability, while the proposed contribution focuses on perception-guided surface modeling and swabbing execution. The swabbing mission is divided into an arm reaching step and a trajectory execution step, with contact and tool-orientation constraints enforced only during trajectory execution.

The framework is evaluated through a fluorescence-based protocol that compares robotic and manual sampling. Human sampling achieves a higher mean raw fluorescence signal (mean $\mu = 26080.5$, standard deviation $\sigma = 8371.8$), whereas robotic sampling yields slightly lower values ($\mu = 24713.5$, $\sigma = 7794.1$). The coefficients of variation are similar, with 32.1% for manual sampling and 31.5% for robotic sampling. These findings support the feasibility of perception-guided robotic surface sampling under controlled conditions, without establishing superiority over expert human performance.

I. Introduction

Healthcare-Associated Infections (HAIs) represent a major global health concern, contributing to increased mortality, prolonged hospital stays and a significant economic burden on healthcare systems. According to international health agencies such as the World Health Organization (WHO) and the European Centre for Disease Prevention and Control (ECDC), HAIs affect millions of patients each year, highlighting the need for effective prevention and reliable monitoring strategies. Among the various transmission pathways, surface contamination has been identified as a critical vector, especially in high-risk clinical areas such as intensive care units. Several studies have demonstrated that pathogens can persist on inanimate surfaces and contribute significantly to cross-transmission dynamics [1]–[3].

Current environmental surface sampling procedures are predominantly manual and remain heterogeneous across studies and protocols. Differences in sampling tools (e.g., swabs, sponges, contact plates), sampled area definition, applied pressure, trajectory execution and surface coverage affect the reliability and comparability of collected measurements [4], [5]. Even under controlled conditions, swabbing performance can depend on operator execution and protocol details, influencing recovery efficiency and measurement variability [6]. These limitations motivate robotic solutions that can support more reproducible sampling procedures while reducing direct operator exposure to potentially contaminated surfaces.

In recent years, robotic systems have been increasingly investigated for environmental monitoring and contamination-sensitive operations, particularly in structured and industrial settings. Examples include robotic platforms for aseptic environmental monitoring, closed robotic work cells for contamination control and robotic swabbing systems in food safety applications [7]–[9]. Although these works confirm the feasibility of automating sampling or contamination-control procedures, they are typically task-specific, operate in highly structured environments, or are not integrated with general-purpose mobile manipulation platforms. As a result, the development of robotic pipelines that combine target detection, surface localization and contact-aware sampling execution remains an open research direction, especially for healthcare-oriented monitoring operations.

From a robotics perspective, surface sampling is a contact-intensive manipulation task, requiring persistent interaction between the robot end-effector, the tool and the target surface. This distinguishes it from standard pick-and-place or collision-free motion planning tasks and motivates constrained motion planning, force/contact-aware control and trajectory validation under task-specific admissibility constraints [10]–[12]. Recent constrained manipulation approaches provide useful formulations for enforcing task-space constraints [13], while reinforcement-learning-based control methodologies have also been investigated for autonomous decision and stabilization problems in uncertain dynamical systems [14]. However, such methods are rarely

*These authors contributed equally to this work

[1]Dipartimento di Ingegneria, Università del Sannio, Piazza Roma 21, Benevento 82100, Italy f.briuolo@studenti.unisannio.it, c.delvecchio@unisannio.it

[2]Università Campus Bio-Medico di Roma, Departmental Faculty of Engineering, Via Alvaro del Portillo 21, Rome, 00128, Italy a.digiacinto@unicampus.it, j.lovecchio@unicampus.it, leandro.pecchia@unicampus.it

[3]Mosaico Monitoraggio Integrato S.r.l., c/o Rete Poema - Zona Industriale ASI83040 Morra De Sanctis - Avellino, Italy o.r.natale@mosaico-mi.it

[4]Fondazione Policlinico Universitario Campus Bio-Medico di Roma, Via Alvaro del Portillo 200, Rome, 00183, Italy leandro.pecchia@unicampus.it

combined with perception-derived surface reconstruction and experimentally evaluated for healthcare-oriented surface sampling.

In this context, the present work addresses the gap between the need for standardized healthcare surface sampling and the lack of robotic frameworks integrating perception-driven surface reconstruction with contact-consistent sampling execution.

The broader system includes perception, planning and control, but this work focuses on vision-based target detection, surface localization and fixed-base swabbing, while autonomous base navigation is treated as an upstream capability.

The added value lies in integrating target detection, depth-derived planar surface modeling, swabbing trajectory validation and fluorescence-based feasibility evaluation on a real mobile manipulator.

The framework is evaluated through a controlled fluorescence-based protocol comparing robotic and manual sampling. The results show that human sampling achieves higher mean collected fluorescence, whereas the robotic system exhibits comparable relative variability under standardized experimental conditions. Therefore, the experimental comparison is interpreted as a feasibility analysis of robotic surface sampling, rather than as evidence of superiority over human operators.

The main contributions of this work are:

- a perception-guided mobile manipulation pipeline integrating target detection, depth-derived surface modeling and fixed-base swabbing trajectory validation;
- a kinematic validation model for the swabbing phase under geometric contact consistency constraints;
- a custom target detection model and its quantitative evaluation on manually annotated scenes;
- a fluorescence-based feasibility comparison between robotic and human surface sampling.

II. Problem Formulation

Surface sampling is formulated as a classical constrained contact manipulation task. The overall pipeline includes an upstream *target reaching phase* and a fixed-base *swabbing phase*. During target reaching, the robot detects the target object, identifies the corresponding surface of interest and reaches a pre-sampling pose in front of the target surface. In this paper, the contribution to this phase is limited to target detection and surface localization; autonomous base navigation is assumed to be available and is not further investigated.

The swabbing phase starts once the robot is positioned in front of the target surface and the mobile base is stopped. We denote by $k = 0$ the initial discrete instant of the swabbing mission. The swabbing phase is divided into two sequential steps: an *arm reaching step*, in which the manipulator moves the swab to the first validated contact point without imposing surface contact; a *trajectory execution step*, in which the swab tip follows the validated surface trajectory. Let k_s denote the first instant of contact with the target surface and the transition index between the arm reaching step and the trajectory execution step, and let N denote the final discrete instant of the fixed-base swabbing mission. Contact and tool-orientation constraints are imposed only for $k = k_s, \ldots, N$, whereas the constant base pose, collision-avoidance, joint-limit and velocity-limit constraints are checked for $k = 0, \ldots, N$. If any of these constraints is violated, the swabbing mission is considered invalid.

Contact modeling assumptions. Physical contact is modeled geometrically with respect to the depth-derived estimate of the target surface. Force regulation, friction, compliance and contact dynamics are not modeled in this formulation and are left for future work. Therefore, the contact condition below should be interpreted as a geometric contact constraint, not as a force-based contact-control model.

A. Kinematic and Surface Representation

Let $n \in \mathbb{N}$ denote the number of arm joints and let $q_k \in \mathbb{R}^n$ be the arm joint configuration at discrete instant k. The candidate joint-space trajectory of the fixed-base swabbing mission is

$$Q = \operatorname{concat}\left(Q^{\text{arm}}, Q^{\text{contact}}\right) = \{q_0, \ldots, q_N\},$$

where $Q^{\text{arm}} = \{q_0, \ldots, q_{k_s-1}\}$ and $Q^{\text{contact}} = \{q_{k_s}, \ldots, q_N\}$ denote the arm reaching and trajectory execution segments, respectively.

Let $\{W\}$ denote the world frame and $\{B\}$ the robot body frame. The homogeneous transformation ${}^{W}T_B(k) \in SE(3)$ denotes the pose of the robot body frame with respect to the world frame at discrete instant k. Since the mobile base is stationary during swabbing, this pose must remain constant:

$${}^{W}T_B(k) = {}^{W}T_B(0), \qquad k = 0, \ldots, N.$$

The swab-tip pose with respect to the robot body frame is

$${}^{B}T_{\text{tip}}(q_k) = \begin{bmatrix} {}^{B}R_{\text{tip}}(q_k) & {}^{B}p_{\text{tip}}(q_k) \\ 0 & 1 \end{bmatrix} \in SE(3),$$

where ${}^{B}p_{\text{tip}}(q_k) \in \mathbb{R}^3$ is the swab-tip position and ${}^{B}R_{\text{tip}}(q_k) \in SO(3)$ is the swab-tip orientation, both expressed in the robot body frame.

Let $\{S\}$ be the local frame attached to the estimated target surface, with pose ${}^{B}T_S$ with respect to $\{B\}$. Its rotation matrix is

$${}^{B}R_S = \begin{bmatrix} {}^{B}s_x & {}^{B}s_y & {}^{B}n_S \end{bmatrix},$$

where ${}^{B}s_x$ and ${}^{B}s_y$ span the tangent plane and ${}^{B}n_S$ is the estimated unit surface normal.

In this feasibility study, the target surface is assumed to be locally planar, consistent with the class of hospital surfaces considered in the experiments, such as bedside tables, trays, worktops and other high-touch planar surfaces [5]. The estimated planar patch is

$$\hat{\mathcal{S}}_B = \left\{ {}^{B}x_S(u,v) = {}^{B}o_S + {}^{B}R_S \begin{bmatrix} u \\ v \\ 0 \end{bmatrix}, \quad (u,v) \in \Omega_S \right\},$$

where ${}^{B}o_S$ is the origin of the local surface frame, (u, v) are local surface coordinates and $\Omega_S \subset \mathbb{R}^2$ is the bounded admissible sampling region. Extension to curved or non-planar surfaces is left for future work.

B. Validation Constraints

During trajectory execution, the swab tip is required to remain on the estimated surface patch $\hat{\mathcal{S}}_B$ within a tolerance. Let ${}^{B}p_{\mathrm{tip}}(q_k)$ denote the swab-tip position expressed in the robot body frame. The signed distance of the swab tip from the estimated tangent plane, measured along the estimated unit normal ${}^{B}n_S$, is defined as

$$d_S(q_k, u_k, v_k) = {}^{B}n_S^{\top} \left({}^{B}p_{\mathrm{tip}}(q_k) - {}^{B}x_S(u_k, v_k) \right),$$

where $(u_k, v_k) \in \Omega_S$ denotes the local surface coordinates of the planned contact point at discrete instant k. The geometric contact constraint requires this signed distance to remain within the admissible tolerance:

$$|d_S(q_k, u_k, v_k)| \leq \delta_c, \qquad k = k_s, \ldots, N,$$

where δ_c accounts for depth noise, calibration uncertainty and finite swab-tip size.

The tool orientation must also remain compatible with the estimated surface normal:

$${}^{B}n_S^{\top}\,{}^{B}z_{\mathrm{tip}}(q_k) \geq \cos(\epsilon), \qquad k = k_s, \ldots, N,$$

where ${}^{B}z_{\mathrm{tip}}(q_k) = {}^{B}R_{\mathrm{tip}}(q_k)e_z$ is the unit tool axis expressed in the robot body frame, with $e_z = [0, 0, 1]^{\top}$. This constraint bounds the angular deviation between the swab-tip axis and the estimated surface normal by ϵ.

Over the whole swabbing mission, the trajectory must satisfy joint limits, velocity limits and collision avoidance. In particular, the swab must avoid unintended collisions with detected obstacles, while the arm links must avoid collisions with both obstacles and the target surface. The only admissible target-surface contact is the intended swab-tip contact during trajectory execution.

C. Problem Statement

The surface sampling task is formulated as a standard trajectory validation problem. Given a candidate swabbing trajectory Q, the objective is to verify that the following constraints are satisfied over their corresponding index ranges:

$$\begin{aligned}
&{}^{W}T_B(k) = {}^{W}T_B(0), && k = 0, \ldots, N,\\
&|d_S(q_k, u_k, v_k)| \leq \delta_c, && k = k_s, \ldots, N,\\
&(u_k, v_k) \in \Omega_S, && k = k_s, \ldots, N,\\
&{}^{B}n_S^{\top}\,{}^{B}z_{\mathrm{tip}}(q_k) \geq \cos(\epsilon), && k = k_s, \ldots, N,\\
&\mathcal{G}_{\mathrm{swab}}(q_k) \cap \mathcal{O} = \emptyset, && k = 0, \ldots, N,\\
&\mathcal{L}_i^{\mathrm{arm}}(q_k) \cap (\mathcal{O} \cup \hat{\mathcal{S}}_B) = \emptyset, && i = 1, \ldots, n_\ell,\ k = 0, \ldots, N,\\
&q_{\min} \leq q_k \leq q_{\max}, && k = 0, \ldots, N,\\
&|\dot{q}_k| \leq \dot{q}_{\max}, && k = 0, \ldots, N.
\end{aligned} \tag{1}$$

where $\mathcal{G}_{\mathrm{swab}}(q_k)$ is the occupied volume of the swab at configuration q_k, $\mathcal{O}$ is the set of detected obstacles excluding the target surface and $\mathcal{L}_i^{\mathrm{arm}}(q_k)$ is the collision geometry of the i-th arm link, with n_ℓ denoting the number of arm links considered in the collision model. The vectors $q_{\min}$ and $q_{\max}$ denote the lower and upper joint limits, respectively. The vector $\dot{q}_k$ denotes the discrete joint velocity at instant k, while $\dot{q}_{\max}$ is the maximum admissible joint-velocity vector. The inequalities involving q_k and $\dot{q}_k$ are intended component-wise.

A swabbing trajectory satisfying all constraints in Eq. (1) is defined as valid and can be executed for surface sampling.

III. System Overview

The proposed method is developed within a perception-guided robotic pipeline for surface sampling. The overall system includes target detection, surface localization, trajectory generation, validation and execution. The contribution to the target reaching phase concerns vision-based detection and surface localization, whereas autonomous base navigation is treated as an upstream capability. Accordingly, this section summarizes the computational pipeline used to transform perception data into a body-frame surface representation and to generate and validate fixed-base swabbing trajectories.

As summarized in Fig. 1, the pipeline is structured into the following main components:

- **Sensor Acquisition:** the robot acquires an RGB image I_{RGB} and the corresponding depth/LiDAR measurement I_D. The transformation from the depth sensor frame to the robot body frame is denoted by ${}^{B}T_D$.
- **Perception Mapping:** RGB and depth/LiDAR data are mapped into a body-frame representation of the target surface:

$$\Phi_{\mathrm{perc}} : (I_{\mathrm{RGB}}, I_D, {}^{B}T_D) \mapsto (\mathcal{P}_B, \mathcal{M}_B, \hat{\mathcal{S}}_B, {}^{B}T_S),$$

where $\mathcal{P}_B$ is the target point cloud, $\mathcal{M}_B$ is the reconstructed local surface mesh or planar model, $\hat{\mathcal{S}}_B$ is the admissible surface patch and ${}^{B}T_S$ is the local surface frame. The target object is detected using a fine-tuned YOLOv8 model and the resulting region of interest is combined with depth/LiDAR data to reconstruct the local surface geometry.
- **Trajectory Planning:** given $\hat{\mathcal{S}}_B$, the planner generates N_c desired swab-tip contact poses:

$$\Phi_{\mathrm{plan}} : \hat{\mathcal{S}}_B \mapsto \mathcal{T}_{\mathrm{tip}}^{d} = \left\{ {}^{B}T_{\mathrm{tip},j}^{d} \right\}_{j=1}^{N_c}.$$

Here, N_c denotes the number of planned contact poses and the superscript d denotes desired quantities. Each pose specifies a contact point on the surface and a tool orientation compatible with the estimated surface normal. The corresponding joint-space trajectory is obtained through sequential inverse kinematics:

$$\Phi_{\mathrm{IK}} : \mathcal{T}_{\mathrm{tip}}^{d} \mapsto Q = \mathrm{concat}\left(Q^{\mathrm{arm}}, Q^{\mathrm{contact}}\right).$$

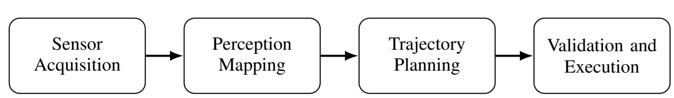


Fig. 1. Architecture of the proposed perception-guided surface sampling pipeline.

Formally, Φ_{IK} denotes a sequential numerical inverse-kinematics map based on damped least-squares task-space error minimization, continuity regularization and joint-limit enforcement [15].
The IK objective minimizes a weighted task-space pose error regularized toward the previous valid configuration, subject to joint limits. In the current implementation, the candidate joint-space trajectory is generated through sequential numerical inverse kinematics, while low-level trajectory execution is handled through the robot control stack. Candidate trajectories are then checked against the admissibility constraints introduced in Section II. The use of sequential decision and control strategies is consistent with previous learning-based control formulations proposed for uncertain dynamical systems [16].

- **Validation and Execution:** validation and execution are treated jointly. A candidate trajectory is evaluated through the validation map

$$\mathcal{V} : (Q, \hat{\mathcal{S}}_B, \mathcal{O}) \mapsto (\nu, \eta),$$

where $\nu \in \{0, 1\}$ is a binary validity flag and η denotes validation and execution metrics. The trajectory is accepted only if the constant base pose, joint-limit, velocity-limit, collision-avoidance, contact and orientation constraints are satisfied over the index ranges specified in Section II. Since the prototype does not include force or tactile sensing at the swab tip, execution success is summarized by

$$H_{\mathrm{exec}} = h_{\mathrm{contact}} h_{\mathrm{collision}} h_{\mathrm{completion}},$$

where the three binary terms indicate visible contact maintenance, absence of unintended collisions and completion of the trajectory, respectively.

The following section details the proposed method for perception-to-geometry mapping, contact-aware trajectory generation and execution-level validation, operationalizing the admissibility conditions defined in Section II.

IV. Proposed Sampling Framework

This section describes the implementation of the proposed perception-guided robotic surface sampling framework. The implemented workflow consists of RGB-depth acquisition, YOLOv8-based target detection, depth projection into the robot body frame, planar surface estimation, generation of N_c contact waypoints, sequential inverse kinematics, constraint validation and execution-level feedback.

A. Perception-to-Geometry Mapping

Given the inputs I_{RGB} and I_D defined in Section III, the target object is localized using a YOLOv8 object detector fine-tuned on a custom dataset of bedside-table scenes acquired from different viewpoints and illumination conditions, including daytime and nighttime acquisitions. Images were initially annotated manually in Label Studio and then refined through a human-in-the-loop workflow based on a Label Studio ML backend [17], with model-assisted predictions visually checked before inclusion in the final dataset. Top-surface annotations include both obstacle-free and partially occupied surfaces, allowing the perception module to identify the subset of the target surface that remains available for sampling. Accordingly, $\hat{\mathcal{S}}_B$ represents the admissible sampling region, obtained by excluding occupied or obstructed areas from the detected top surface.

The detected region is then combined with the depth/LiDAR measurement to reconstruct the local 3D geometry of the target surface. Valid depth samples inside the region of interest are back-projected into 3D and transformed into the robot body frame using the frame transformations provided by the Boston Dynamics perception and robot-state modules. The resulting body-frame point cloud is filtered to remove outliers and points outside the target region.

Since the surfaces considered in this study are assumed to be locally planar, the filtered point cloud is approximated by a planar surface model. The estimated surface centroid and normal define the local surface frame $\{S\}$ and the admissible surface patch $\hat{\mathcal{S}}_B$ introduced in Section II. This representation provides the geometric input for the generation of contact waypoints and the validation of the trajectory.

B. Trajectory Generation and Execution-Level Validation

The trajectory generation module operates after the target reaching phase has been completed. At this stage, the robot is positioned in front of the target surface and the mobile base is stationary. The fixed-base swabbing mission is composed of an arm reaching segment and a trajectory execution segment. The arm reaching segment moves the swab from its initial configuration to the first validated contact waypoint without imposing surface contact, while the trajectory execution segment follows the validated contact path on $\hat{\mathcal{S}}_B$.

Given the estimated surface patch, the trajectory execution segment is constructed as a sequence of N_c contact waypoints with positions on $\hat{\mathcal{S}}_B$ and tool orientations compatible with the estimated surface normal. The corresponding joint-space trajectory is generated by the sequential IK procedure introduced in Section III and executed through the robot control stack.

A trajectory is accepted only if the constant base pose constraint, collision-avoidance constraints, joint-limit constraints and velocity-limit constraints are satisfied over the whole swabbing mission, and if the contact and orientation constraints are satisfied during trajectory execution. If any of these checks fails, the swabbing mission is rejected.

Since the current prototype does not include force or tactile sensing at the swab tip, execution-level feasibility is assessed through the human-evaluated score H_{exec} defined in Section III. The system should therefore be interpreted as a perception-guided, geometrically validated execution pipeline rather than as a closed-loop contact-control system. Closed-loop sampling with online re-planning, force feedback, tactile sensing and adaptive contact control is left for future work.

V. Experimental Feasibility Evaluation

This section presents a controlled experimental feasibility evaluation of the proposed robotic sampling procedure. The objective is not to demonstrate superiority over manual sampling, but to assess whether robotic sampling can produce measurable fluorescence recovery with variability comparable to manual execution under standardized conditions.

A. Experimental Setup

The robotic platform used in the experiments is the Boston Dynamics Spot legged robot equipped with the Spot Arm, which provides six degrees of freedom and an integrated gripper. For robotic sampling, a rack with six sterile collection tubes was manually mounted on the robot body (see Fig. 2a). Each tube houses one swab, allowing the arm to extract a swab before sampling and reinsert it after task completion. At the end of the mission, the tubes can be manually removed and transported to the laboratory for analysis.

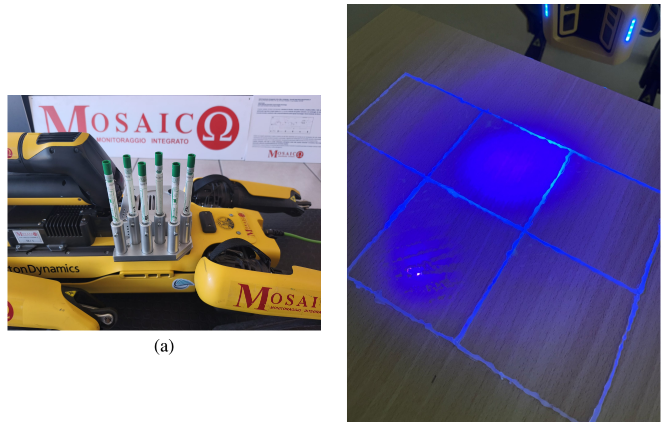
The evaluation is conducted using a fluorescence-based protocol. A known quantity of fluorescein is uniformly applied to a planar target surface and subsequently collected using a standardized swab (Copan 952C, Copan Diagnostics) in both manual and robotic trials. Sampling is performed over a structured 6×2 grid, corresponding to $N_c = 12$ contact waypoints in the experimental protocol, to promote consistent spatial coverage across trials following ISO 18593 guidelines. The same sampling material, surface preparation procedure and fluorescence analysis protocol are adopted for both experimental conditions (see Fig. 2b).

In the experiments, δ_c is set to 0.0015 m and is used as the admissible bound on the signed distance between the swab tip and the estimated sampling plane. This parameter accounts for residual uncertainty due to depth sensing, surface fitting, frame transformations and the finite size of the swab tip. The candidate arm trajectory is computed through a sequential numerical inverse-kinematics solver based on damped least-squares updates. At each contact waypoint, the previous valid arm configuration is used to initialize the solver, so as to promote joint-space continuity and reduce abrupt configuration changes. The resulting joint trajectory is accepted only if the geometric contact tolerance, tool-orientation constraint, joint limits, velocity limits and collision-avoidance checks are satisfied.

A total of 72 trials are performed with the robotic system and 72 with manual operators under the same protocol. Each trial consists of surface preparation, sampling execution, sample collection and fluorescence measurement using a spectrofluorometer (Infinite 200 Pro, Tecan Group Ltd.). Background-corrected fluorescence is computed by subtracting the baseline signal obtained from blank samples. Fluorescence values are reported in arbitrary units (a.u.).

B. Metrics and Results

Detection performance is evaluated on manually reviewed validation and test scenes using precision P, recall R, F1 score, mAP@0.50 and mAP@0.50:0.95. The YOLOv8m detector is fine-tuned from pretrained weights on 1550 training images and 173 validation images, with 20 additional public images used for external inference. Training uses 640-pixel inputs, batch size 16, SGD, initial learning rate 10^{-5}, dropout 0.2 and early stopping with patience 5, reaching convergence after approximately 30 epochs.



(a) (b)

Fig. 2. Experimental setup: (a) robot-mounted rack with sterile tubes and swabs; (b) fluorescein-marked planar sampling grid.

Table I shows high precision and mAP@0.50, indicating reliable target localization under the validation conditions. The lower mAP@0.50:0.95 reflects the stricter localization requirement across multiple IoU thresholds, while the precision–recall balance supports the detector as an upstream module for surface-based trajectory generation.

Sampling performance was evaluated using the mean raw fluorescence signal, standard deviation, coefficient of variation ($\text{CV} = 100\sigma/\mu$) and background-corrected mean fluorescence. Execution-level feasibility was also monitored using the human-evaluated score H_{exec} defined in Section III, which summarizes visible contact maintenance, absence of unintended collisions and task completion.

As shown in Table II, manual sampling achieved a higher mean collected fluorescence signal in both raw and background-corrected measurements. Robotic sampling yielded slightly lower mean values, while the coefficients of variation were similar across the two methods. These results indicate comparable relative dispersion under controlled conditions, but do not establish superiority of robotic sampling over manual execution. All robotic trials included in the analysis completed the planned trajectory without visible loss of contact or unintended collisions, resulting in $H_{\text{exec}} = 1$.

TABLE I
Quantitative detection performance of the fine-tuned YOLOv8m model used for target localization.

Prec.	Rec.	F1	mAP@0.50	mAP@0.50:0.95
0.9562	0.8760	0.9144	0.9572	0.7186

TABLE II

COMPARISON BETWEEN MANUAL AND ROBOTIC SAMPLING. FLUORESCENCE MEASUREMENTS ARE EXPRESSED IN ARBITRARY UNITS (A.U.). "BC" DENOTES BACKGROUND-CORRECTED MEASUREMENTS.

Method	**Mean raw [a.u.]**	**Std raw [a.u.]**	**CV [%]**	**Mean bc [a.u.]**
Manual	26080.5	8371.8	32.1	21492.7
Robot	24713.5	7794.1	31.5	20627.3

C. Discussion

The results should be interpreted as a feasibility comparison rather than as evidence of robotic superiority. Manual sampling recovered a higher mean fluorescence signal, whereas robotic sampling showed comparable relative variability. The lower mean value obtained by the robot may be related to differences in contact execution and to the current mechanical setup, including swab handling, air exposure during transfer from the target surface to the body-mounted rack and reinsertion into the collection tube.

The inclusion of partially occupied top-surface annotations further supports sampling-area selection under moderate surface occlusion, although systematic stress testing under severe clutter, reflective surfaces and highly variable lighting remains future work.

The current prototype validates contact geometrically and through human-evaluated feedback, without force or tactile measurements. Therefore, force consistency remains outside the scope of the present validation and will be addressed through dedicated sensing. The fluorescence-based protocol provides a controlled and repeatable proxy for sampling performance, but does not fully capture microbial adhesion, survival, or air-exposure effects. Validation with clinically relevant microbial samples is left for future work.

VI. CONCLUSION

This work presents a feasibility study of a perception-guided robotic framework for healthcare-oriented surface sampling. The proposed approach integrates YOLOv8-based target detection, depth-derived surface localization and constraint-based validation of fixed-base swabbing trajectories.

The experimental evaluation shows that manual sampling achieves a higher mean fluorescence signal, while robotic sampling exhibits comparable relative variability under controlled conditions. These findings support the feasibility of robotic surface sampling as a step toward more standardized monitoring workflows, without establishing superiority over expert manual sampling.

Future work will extend the evaluation to non-planar surfaces, cluttered environments and clinically relevant microbial contamination. In addition, the current perception-guided execution pipeline will be extended toward a closed-loop sampling framework by integrating force and tactile sensing, quantitative contact-force measurement, online trajectory adaptation and mechanical optimization of the swab-handling rack.

Acknowledgment— This work was supported by Mosaico Monitoraggio Integrato S.r.l.

REFERENCES

[1] D. J. Weber, W. A. Rutala, M. B. Miller, K. Huslage, and E. Sickbert-Bennett, "Role of hospital surfaces in the transmission of emerging health care-associated pathogens: norovirus, clostridium difficile, and acinetobacter species," *American journal of infection control*, vol. 38, no. 5, pp. S25–S33, 2010.

[2] J. A. Otter, S. Yezli, J. A. Salkeld, and G. L. French, "Evidence that contaminated surfaces contribute to the transmission of hospital pathogens and an overview of strategies to address contaminated surfaces in hospital settings," *American journal of infection control*, vol. 41, no. 5, pp. S6–S11, 2013.

[3] G. Suleyman, G. Alangaden, and A. C. Bardossy, "The role of environmental contamination in the transmission of nosocomial pathogens and healthcare-associated infections," *Current infectious disease reports*, vol. 20, no. 6, p. 12, 2018.

[4] S. Rawlinson, L. Ciric, and E. Cloutman-Green, "Covid-19 pandemic–let's not forget surfaces," *Journal of hospital infection*, vol. 105, no. 4, pp. 790–791, 2020.

[5] C. Willis, K. Nye, H. Aird, D. Lamph, and A. Fox, "Examining food, water and environmental samples from healthcare environments. microbiological guideline," *Pub Health England Pp*, pp. 1–48, 2013.

[6] A. Babiker, J. Van Riel, S. Lohsen, A. Page, A. Strudwick, E. Wilber, M. Woodworth, and S. Satola, "Evaluation of swabbing methods for culture and non-culture-based recovery of multidrug-resistant organisms from environmental surfaces," *Infection Control & Hospital Epidemiology*, vol. 46, no. 8, pp. 837–844, 2025.

[7] T. Deguchi, H. Maruyama, H. Matsuoka, and J. Akers, "Application of robotics in aseptic environmental surface monitoring," *Pharmaceutical Technology*, 2010.

[8] J. McCall, N. Barnard, K. Gadient, C. Kasireddy, A. Kurtz, Y. Li, T. Page, T. Putnam, and A. Brennan, "Environmental monitoring for closed robotic workcells used in aseptic processing: Data to support advanced environmental monitoring strategies," *AAPS PharmSciTech*, vol. 23, no. 6, p. 215, 2022.

[9] S. Mahmoudi, C. Griscom, P. Sohrabipour, Y. Tian, C. Pallerla, P. Crandall, and D. Wang, "Evaluation of robotic swabbing and fluorescent sensing to monitor the hygiene of food contact surfaces," *Foods*, vol. 14, no. 19, p. 3311, 2025.

[10] D. Berenson, S. Srinivasa, and J. Kuffner, "Task space regions: A framework for pose-constrained manipulation planning," *The International Journal of Robotics Research*, vol. 30, no. 12, pp. 1435–1460, 2011.

[11] J. Cortés and T. Siméon, "Sampling-based motion planning under kinematic loop-closure constraints," in *Algorithmic Foundations of Robotics VI*. Springer, 2005, pp. 75–90.

[12] M. H. Raibert and J. J. Craig, "Hybrid position/force control of manipulators," *Journal of Dynamic Systems, Measurement, and Control*, vol. 103, no. 2, pp. 126–133, 06 1981. [Online]. Available: https://doi.org/10.1115/1.3139652

[13] J. Panthi, F. Alambeigi, and M. Pryor, "A closed-chain approach to generating affordance joint trajectories for robotic manipulators," *IEEE Transactions on Robotics*, 2025.

[14] A. Acernese, A. Yerudkar, L. Glielmo, and C. Del Vecchio, "Reinforcement learning approach to feedback stabilization problem of probabilistic boolean control networks," *IEEE Control Systems Letters*, vol. 5, no. 1, pp. 337–342, 2020.

[15] B. Siciliano, L. Sciavicco, L. Villani, and G. Oriolo, *Robotics: modelling, planning and control*. Springer, 2009.

[16] A. Acernese, A. Yerudkar, L. Glielmo, and C. Del Vecchio, "Model-free self-triggered control co-design for probabilistic boolean control networks," *IEEE Control Systems Letters*, vol. 5, no. 5, pp. 1639–1644, 2020.

[17] HumanSignal, "Label Studio ML Backend," https://github.com/HumanSignal/label-studio-ml-backend, 2026, accessed: 2026-05-29.