IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 27 February 2023, accepted 14 March 2023, date of publication 20 March 2023, date of current version 24 March 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3259539

RESEARCH ARTICLE

Robotic Manipulation System Design and Control for Non-Contact Remote Diagnosis in Otolaryngology: Digital Twin Approach

SHO-HYUN LEE[1], DONG-WOO LEE[1], HWA-SEOB SONG[2], SEONMIN JEONG[3], YONGBAE JI[3], JI-SUNG SONG[4], JIYOUNG KIM[5], AND BYUNG-JU YI[6], (Member, IEEE)

[1]Department of Electrical Engineering, Hanyang University, Seoul 15588, Republic of Korea
[2]Research Institute Engineering and Technology, Hanyang University, Seoul 15588, Republic of Korea
[3]Department of Otolaryngology-Head and Neck Surgery, College of Medicine, Hanyang University, Seoul 11923, Republic of Korea
[4]School of Communication Design, Hanyang University, Seoul 15588, Republic of Korea
[5]Department of Intelligent Robotics, Hanyang University, Seoul 15588, Republic of Korea
[6]School of Electrical Engineering, Hanyang University, Seoul 15588, Republic of Korea

Corresponding authors: Byung-Ju Yi (bj@hanyang.ac.kr) and Jiyoung Kim (kimjiyoung93@naver.com)

This research was supported in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education under Grant 2021R1I1A4A01051258, and in part by the BK21 Fostering Outstanding Universities for Research (FOUR) funded by the Ministry of Education (MOE, South Korea, and NRF).

ABSTRACT The COVID-19 pandemic has emphasized the need for non-contact medical robots to alleviate the heavy workload and emotional stress experienced by healthcare professionals while preventing infection. In response, we propose a non-contact robotic diagnostic system for otolaryngology clinics, utilizing a digital twin model for initial design optimization. The system employs a master-slave robot architecture, with the slave robot comprising a flexible endoscope manipulation robot and a parallel robot arm for controlling additional medical instruments. The novel 4 degrees of freedom (DOF) control mechanism enables the single robotic arm to handle the endoscope, facilitating the process compared to the traditional two-handed approach. Phantom experiments were conducted to evaluate the effectiveness of the proposed flexible endoscope manipulation system in terms of diagnosis completion time, NASA task load index (NASA-TLX), and subjective risk score. The results demonstrate the system's usability and its potential to alternate conventional diagnosis.

INDEX TERMS Digital twin, flexible endoscope, robot design, robotic diagnosis, telerobotic system, usability test.

I. INTRODUCTION

The COVID-19 pandemic has increased the stress and fatigue of otolaryngologists and patients. A cross-sectional study revealed a correlation between the emotional stress of medical workers and the COVID-19 outbreak [1]. Medical staff who directly or indirectly contacted patients reported anxiety and depression.

Therefore, in recent years, remote medical robot diagnosis systems gained prominence. Accordingly, many non-contact robotic systems for diagnosis, disinfection, surgery, and remote health care services have been developed, ever since the COVID-19 outbreak [2]. Among them, a non-contact swab sampling robot, that is, a swab robot, was developed primarily for diagnosis. Kim et al. [3] classified the non-contact sampling robot system into three types: a human-machine interface, a computer vision-based artificial intelligence system, and an instrument-based simple insertion system.

The human-machine interface allows medical staff to monitor and diagnose the patient by the master control device remotely. A computer-vision-based artificial intelligence system is a fully automatic system. The simple instrument-based insertion system places the subject in front of the instrument, and only movement after insertion into the body is required.

The associate editor coordinating the review of this manuscript and approving it for publication was Giovanni Merlino.

This classification is also applicable to other medical robotic diagnoses and treatment systems.

To apply the robotic method to medical diagnosis, it is necessary to analyse the target medical procedure and derive the degrees of freedom (DOF) required for the robot. This procedure can be explained using the swab robot development process.

The swab sampling procedure is briefly described as follows. COVID-19 is diagnosed by collecting a sample through a swab. The procedure involves inserting a swab for body odour through the oropharynx or nasopharynx, rotating it several times against the mucosa so that the secretions can be sufficiently absorbed, and then placing it into the collection tube [4]. It is more advantageous to obtain specimens from both the oropharynx and nasopharynx [5], [6]; however, in most triage rooms, the specimen is mainly obtained from the nasopharynx.

According to Kim et al. [3], the movement required for a robot in the sample-collection process requires at least six DOF. The movements involve 2 DOF insertion angle correction, swab insertion and retraction, planar motion horizontal to the patient's face, and rotation for absorption during insertion.

In addition, a recommendation by Kim et al. [3] for positioning diagnostic instruments before insertion is to use a universal serial robot, also called a collaborative robot (COBOT). Previous studies have used COBOTs in COVID-19 sampling robots [7], [8], [9].

As shown above, according to Kim et al. [3] and Shen et al. [2], many studies have focused on swab robots for COVID-19 screening used in triage rooms. However, the burden of preventing the transmission of COVID-19 is not limited to the triage room, so a new system that can be applied to the broader medical field, especially in hospitals that treat respiratory diseases, such as otolaryngology, is required.

Therefore, the ultimate goal of this study is to develop a non-contact robotic diagnostic system that can be applied to an otolaryngology clinic.

This study makes three major contributions to the literature.

1) First, a newly designed 4-DOF end-effector enables remote control with one hand through the graphic user interface (GUI) is proposed, making it possible to use the flexible endoscope and other medical instruments simultaneously.
2) Second, we conduct a phantom experiment and usability test of experts and novices to verify the physical implementation of the robot-held flexible endoscope system, including face recognition.
3) Third, we propose a digital twin (DT) model of two-arm manipulation for the diagnosis of otolaryngology.

Fig. 1-(a) demonstrates the typical diagnosis in a conventional otolaryngology clinic. Regarding COVID-19, both patients and doctors are exposed to viral infections. In contrast, the otolaryngology clinic that consists of the proposed diagnosis system is shown in Fig. 1–(b). The doctor and the patient are separated so that non-contact treatment is possible, thereby preventing infection.

Briefly summarising the system architecture, the proposed system is based on a master-slave robot system. The slave stage comprises a camera and two robot arms: a parallel robot to control the medical instruments, a 6-DOF robot arm to control the novel 4-DOF flexible endoscope control module, and a camera for face recognition. The doctor and master PC are located in the master stage, while a GUI on the master PC controls the slave stage.

Our work applies a computer vision-based artificial intelligence system to automatically position a diagnostic module to an appropriate location. Although this method has the advantage of being automated, it is challenging to prepare for emergencies. To solve this problem, we use a combination of computer vision-based and human-machine interface methods. In our system, the instruments are positioned automatically outside the body by face recognition, while during the instruments' insertion into the body, the doctor controls the module directly through monitoring.

In otolaryngology diagnostic procedure, an endoscope is generally used to secure the visual field of the body. The use of rigid and flexible endoscopes in conventional otolaryngology clinic are shown in Fig. 2. The flexible endoscope is used to insert deep into the larynx. However, a doctor must use two hands to handle a flexible endoscope, as shown in Fig. 2-(a). Therefore, this approach cannot be used along with other diagnostic instruments such as suction or forceps.

Alternatively, the doctor can use a rigid endoscope with one hand and another device with the other, as shown in Fig. 2-(b). The shortcoming of this approach is that the rigid endoscope cannot be inserted deeply into the body, so it cannot reach the laryngeal area.

The proposed system was developed to solve these problems. It aims to simultaneously use a flexible endoscope with other instruments using two robot arms.

To enable the simultaneous use of a flexible endoscope and other instruments, we analysed the flexible endoscopy diagnostic procedure as aforementioned example and then derived the DOF required for flexible endoscope control. Consequently, we designed a novel 4-DOF flexible endoscope control module. Moreover, we used a 6-DOF COBOT as a robot arm to properly position the module outside the body for diagnosis. An overall comparison between the two diagnostic methods for flexible endoscopy is summarised in Table. 1.

We also analysed the DOF of other instruments used in otolaryngology, such as forceps, suction, swabs, and ultrasonic probes. As another robot arm, we adopted a 4-DOF RRPaRR parallel robot with sufficient DOF to control instruments [10]. Considering the required task space for the diagnosis procedure and the size of the medical instrument, we located the parallel robot on the patient site's desk.

We performed physical implementation and performance experiments on a flexible endoscope manipulation system. As aforementioned, we used a GUI, unlike the physical control devices that are primarily used in other studies [7], [11]. Among various diagnostic instruments for otolaryngology, the physical implementation of flexible endoscope control was first performed in this study because the flexible endoscope requires the most DOF and is the most used instrument in the field. We used the NASA-TLX and completion time, which are mainly used for system performance evaluation [12], and a subjectively perceived risk score using a visual analogue scale (VAS) as evaluation methods.

The use of DTs has been increasingly highlighted in the medical field, where they have proven to be effective in providing real-time feedback to optimize physical systems. DTs are virtual models of physical systems that are connected to their physical counterparts and exchange data in real time. This allows DTs to optimize the physical system by providing and receiving feedback from the physical twin (PT) [13].

This feature is particularly useful in medical systems where real-time feedback is important. Previous research has shown that DTs have been used to develop virtual models of patients' bodies or physiological mechanisms for testing surgical procedures and pharmaceuticals, as well as for improving hospital processes [14], [15].

The integration of DT technology in the development of medical devices and robots has become a significant issue in the field of medical applications. The use of DTs allows for concurrent collaboration among multiple users and facilitates the testing of hardware through software development during the design and development phase of medical devices. This approach has been demonstrated to improve the efficiency and effectiveness of the design process [16].

In this study, we developed an independent digital space model of a diagnostic system using two robot arms and an otolaryngology clinic room environment to which this system was applied. It will be connected to the physical model and will share each lifecycle. Moreover, we simulated the diagnosis procedure with a two-robot arm system. Following the simulation in this digital space, we will commence manufacturing after gaining an understanding of the workspace and design conditions of the parallel robot, which handles diagnostic instruments. Thus, mistakes that may occur during the manufacturing phase can be reduced. Furthermore, it will help analyse and configure the trajectories of the two robot arms and end effectors.

In addition, in terms of user experience (UX) research, digital twins allow patients to experience the robot diagnosis environment in advance to increase understanding of the new diagnosis system, reduce objections, and improve the system by receiving feedback from actual users. Not only for patients but also it can be used for the training of diagnose operators like doctors.

The remainder of this paper is organised as follows. In Section II, DT is introduced, followed by the DT model and motion simulation in the digital environment. The components and development process of the physical implementation of the flexible endoscope manipulation system are explained in Section III. Section IV discusses usability test settings, procedures, and results. In Section V, we discuss the test results given in Section IV.

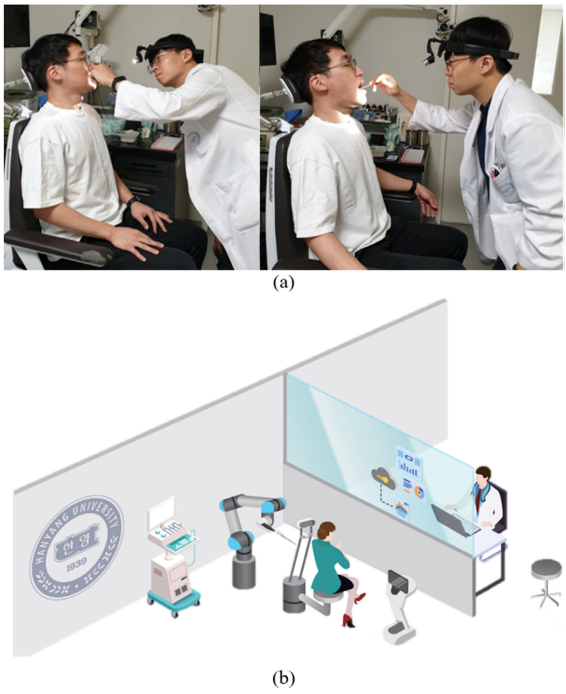


FIGURE 1. (a) A conventional otolaryngology clinic; (b) the clinical setup of the proposed system scheme.

FIGURE 2. (a) A doctor using a flexible endoscope in a conventional otolaryngology clinic; (b) a doctor using a rigid endoscope and other instruments in a conventional otolaryngology clinic.

TABLE 1. Comparing two diagnostic methods of flexible endoscope: the manual application and the proposed robotic application.

Property	Manual Application	Robotic Application
No. of hands	Two hands	One mechanism
Doctor's workload	High	Low
Remote control	Impossible	Possible
Viral infection	High risk	Low risk

II. DIGITAL TWIN OF THE REMOTE NON-CONTACT DIAGNOSIS SYSTEM

A DT is a digital representation of a physical entity, such as a system, object, or process, that is synchronized with the

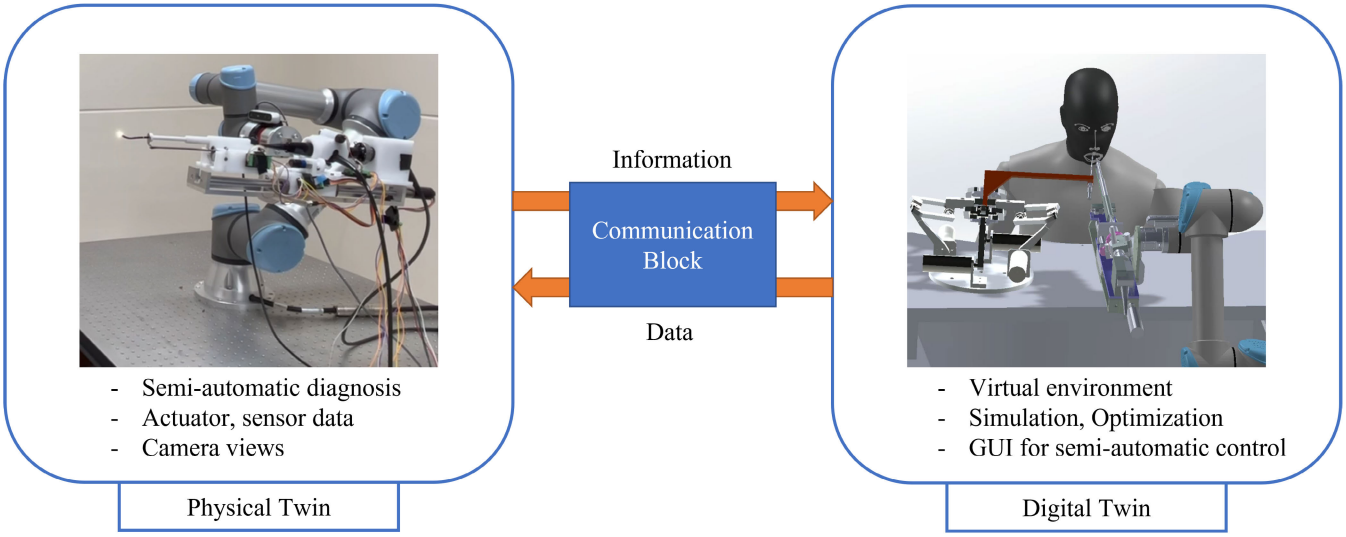


FIGURE 3. DT of two robot arm diagnostic system. The communication module manages information exchange between the virtual environment and the physical environment.

physical entity through real-time communication and has the same life cycle.

The life cycle of a digital twin refers to the stages from its creation to its eventual decommissioning and includes the design, development, operational, and maintenance phases. The type of life cycle can be divided into two cases: a DT that is created before the physical entity and a DT that is created for an already existing physical entity.

The configuration of a DT typically includes a digital and a physical space, as well as a communication module that connects them and a repository for storing data. The digital space includes the 3D CAD environment, objects, physical conditions, motion scenes of objects, and processes, and is sometimes referred to as the information model because it digitally describes the information of the physical model [17].

In this study, the digital space of the DT, which includes the 3D CAD environment and objects, is established in the early stage of the life cycle, specifically the design phase. Through simulations, the motion and workspace of the objects are identified and used to optimize the design of the end effector for driving the robotic arms and surgical instruments, as well as in motion planning for diagnostic procedures. Ultimately, the goal is to implement the digital twin in a physical form and establish a link between the digital and physical entities.

The structure of a DT is shown in Fig. 3. The PT provides information about its current state and environment by transmitting data such as the current configuration or joint values of the robotic arms, actuator and sensor data, and camera views. This information is received by the DT via the communication module. After performing motion simulation and optimization in the virtual environment using this information, the DT sends updated actuator values and robot joint information back to the PT.

During a diagnostic procedure, the doctor can observe the camera views provided by the PT and control the movement of the robotic arms through a GUI. These movements are transmitted to the DT and updated in real time, allowing for a seamless integration of the physical and digital entities.

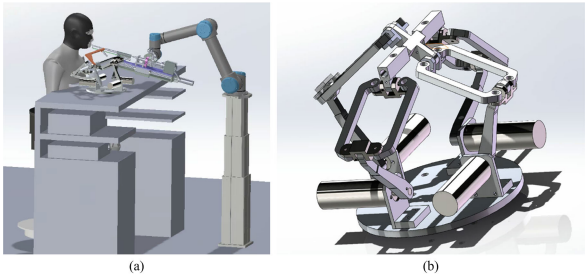


FIGURE 4. Representation of the digital space model of the slave stage. (a) An overall view of the digital space. (b) A close-up view of the parallel robot structure.

A. DIGITAL SPACE DESCRIPTION

The digital space created by us is shown in Fig. 4-(a). It consists of the patient area separated from the doctor area in the entire system, including the slave stage.

First, an electric chair is positioned for the patient to sit. The desk in front of the patient contains various medical instruments and by-products required for each instrument's diagnostic procedure. The flexible endoscope digital system is located to the patient's right.

We selected a 6-DOF serial robot arm to control the flexible endoscope module and a 4-DOF RRPaRR-type parallel robot to control another instrument. The required motion for a parallel robot was analysed as 4-DOF: one DOF for approaching the human face, one for inserting the instruments, one for adjusting the angle of insertion of the instruments into

the target, and one for the instruments' height adjustment because the height of the face feature varies for each person. The selected RRPaRR-type robot can create a pitch motion and X, Y, and Z linear motions, which is sufficient to create the required motions. Our research team developed this parallel robot in a previous study [10]. Using a DT environment, we optimised the parallel robot design and workspace. Because the parallel robot handles relatively small medical instruments, we designed it to be small enough to be placed on a desk. The digital space model of the selected parallel robot structure is shown in Fig. 4-(b). The detailed description of the parallel robot dimensions and workspace are provided in the Appendix C.

Among the two robots, the 6-DOF serial robot arm which controls the flexible endoscope driving module is mounted at the same height as the desk, and the base is located 295 mm far to the patient's left.

Another parallel robotic arm is installed on the 320 mm far right from the patient. An extending link is mounted on the top platform of the parallel robot for the appropriate positioning of the diagnostic instruments. This arrangement avoids collisions between the diagnostic instruments and the flexible endoscope module during the diagnosis process.

B. TWO ROBOT ARM DIAGNOSTIC SYSTEM SIMULATION

In this study, we developed a DT for a two-robot arm diagnostic system to demonstrate that the simultaneous use of two instruments is feasible.

A computational simulation of the system motion was performed to verify the workspace and motion required during the diagnosis procedure. There are two simulation targets: mouth and nose. The insertion angle of the flexible endoscope module was set to approximately 12°, similar to that in the experiments described in Section IV. The scenario is as follows:

First, the flexible endoscope module is positioned at a distance of approximately 2 cm from the target entrance. After the parallel robot moves the instruments near the nose, the instrument is inserted approximately 3 cm in vivo. Then, for medical treatment in a deeper area, a flexible endoscope module is inserted into the body approximately 3 cm to secure the field of view. Finally, a parallel robot inserts a diagnostic instrument of approximately 4 cm for treatment. The final state of the simulation process is shown in Fig. 5.

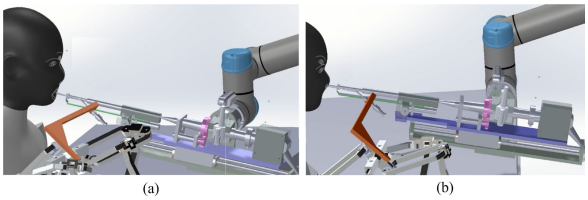


FIGURE 5. Final state of each target's simulation process captured in the side view. It shows the insertion of the two instruments into the target. (a) Side view of the target 1 (mouth) simulation. (b) Side view of the target 2 (nose) simulation.

A simulation video is provided in the Appendix B.

III. PHYSICAL IMPLEMENTATION OF FLEXIBLE ENDOSCOPE MANIPULATION SYSTEM

The flexible endoscope manipulation system was implemented physically. The overall configuration of the physical system is illustrated in Fig. 6.

The system is divided into a master stage on the doctor's side and a slave stage on the patient's side. At this time, the area on the patient's side and that on the doctor's side are physically separated to avoid virus propagation, as shown in Fig. 1(b).

In the master stage, the doctor can control the robot in the slave stage using the GUI. In addition, it is possible to monitor a patient's condition in real time using dual camera views transmitted from a face recognition camera attached to UR5e and a flexible endoscope's camera.

The slave stage consists of a robot arm, camera, and end-effector module for driving a flexible endoscope. For the 6-axis universal serial robot arm, we use UR5e from Universal Robots, which is used to position diagnostic instruments in place. We used a RealSense D415 from Intel for the face recognition task. We used an ENF-P4 endoscope and a CLV-S400 Light Source for the flexible endoscope system from Olympus. The end-effector module for driving the flexible endoscope was self-designed and mounted on the distal end of the robot arm using a tool changer. The design of the module is described in the following subsections.

The entire system was implemented on ROS Noetic. For communication, we used the TCP-IP protocol for UR5e robot control with a universal robot driver [18] and the CANopen communication protocol for the motors to operate the flexible endoscope end-effector module.

A. END-EFFECTOR MODULE DESIGN FOR CONTROLLING THE FLEXIBLE ENDOSCOPE

Several studies on the design and control of flexible endoscopes exist. Concerning the remote control of the flexible endoscope, large endoscopes, such as colonoscopy, include a kinematically supported and managed mechanism for the soft part at the distal end [19], [20], [21], [22]. Most of these methods use roller mechanisms. However, there is no soft part support mechanism for head and neck diagnosis in small endoscopes. In the case of a small endoscope for the head and neck, if there is no support, the insertion direction is fixed from top to bottom. Thus, the patient's posture is also limited to lying down or postural equivalent [23]. Alternatively, including a soft part as a continuum mechanism is challenging to control, has a slack problem, or has a limit in reducing the thickness [24]. Therefore, in this paper, we propose an end-effector module design that can handle all motions of a flexible endoscope, including the soft part support motion. It is clear that our mechanism is intended for use in diagnostic rooms and not operating rooms. According to Poon et al. [25], several robotic systems and mechanisms have already been developed for endoscopic head and neck surgery, but they are not available for use in diagnostic rooms.

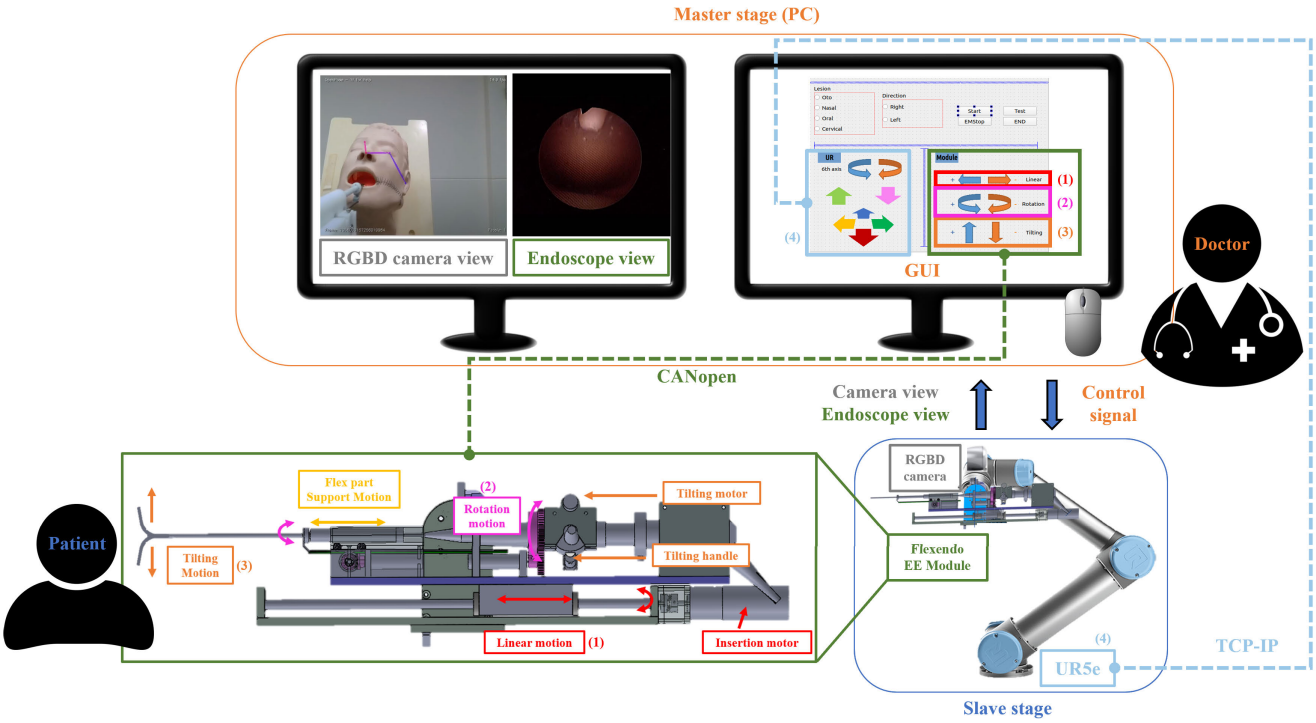


FIGURE 6. Proposed robot system configuration including the end-effector(EE) module for driving the flexible endoscope.

As aforementioned, a conventional flexible endoscope must be used with two hands. Mechanisms developed to use the endoscope with other instruments also exist, but they are predominantly developed for surgical purposes [26], [27], [28], [29]. Gafford et al. [30] developed an otoendoscope for diagnosis, and the endoscope could be controlled with one hand. However, it was developed for face-to-face treatment.

In this research, we designed a 4-DOF mechanism on which the flexible endoscope is mounted and drives all the motions required for diagnosis using the flexible endoscope: the linear motion required for insertion into the body, rotational motion, tilting motion of the distal end of the soft endoscope, and soft part support motion. The details of the flexible endoscope module design are shown in Fig. 7. It shows the mounting-side view of the actual module and the configuration for each DOF. It also shows how the motion is generated.

The control of this module is divided into two stages: insertion motion and remaining motion driving. A motor that generates the insertion motion is fixed to the base using a spool structure (Fig. 7–(a), (b)). The rotation of this motor is transmitted by the lead screw and guided by the linear bushing; subsequently, a linear insertion motion is generated.

The remaining motions are as follows. First, the actuator that generates the soft part tilting motion is attached directly to the tilting handle (Fig. 7–(b), (c)). Next, the actuator drive for the rotation of the entire body is transmitted by the gear. The gear ratio is 1:2. This requires the greatest torque among all the motions, except for the insertion motion.

The soft part support mechanism is of the telescope type, as shown in Fig. 7–(d), (e). A telescopic motion is realised using the rack-and-gear mechanism. The pulley acts as a gear, and the timing belt acts as a rack. The motion generation process is described as follows: First, the actuator is driven to rotate the coupled gears. Subsequently, the coupled pulley rotated accordingly. When the pulley rotates, the timing belt moves linearly; thus, a linear motion of the scope coupled with the belt is generated.

Owing to the soft part supporting mechanism, there is no sagging, even if the soft part faces upward at a 90° angle. Therefore, one advantage of performing a remote diagnosis using a flexible endoscope is that there are no restrictions on the patient's posture, so the patient is not required to lie down or the equivalent.

In the actual diagnosis, the soft part support motion occurs in the form of backward movement at the same time as the endoscope is inserted into the body. This backward motion prevents the collision of the supporting part with the body.

The important physical indices of the proposed end-effector module and endoscope is described in Table. 2 and Figure. 8. The LM 0 mm is the home position of the module. All actuators used were from Faulhaber.

B. FACE RECOGNITION

Face recognition aims to extract the position of a diagnosis site using a camera. The target examination areas are the nose, mouth, ears, and neck. To recognise facial features,

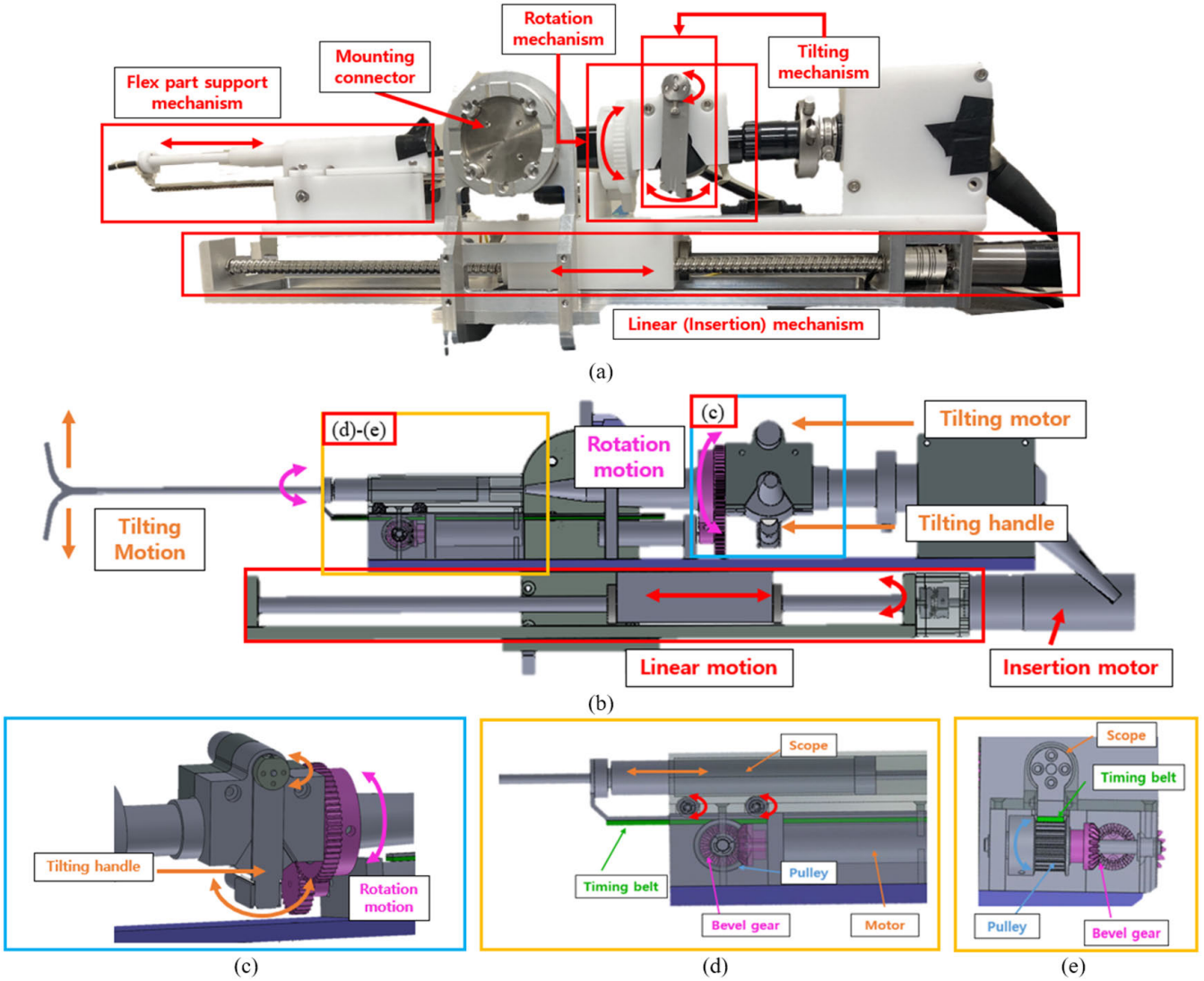


FIGURE 7. Proposed 4-DOF flexible endoscope end-effector module design. (a) Mounted-side view of the module; (b) Non-mounted-side view of the module's 3D CAD model; (c) Details of the tilting motion and rotation motion mechanism; (d) Side-view details of the flex part support motion mechanism; (e) Front-view of the flex part support motion mechanism.

TABLE 2. Important physical indices of the proposed end-effector module and endoscope.

Parameter name	Value
Width(MW) x Height(MH) of module	464 x 151.9 mm
Range of the linear motion (LM)	-156 to 156 mm
Range of the Scope expansion (SE)	106.05 to 282.15 mm
Length of the flexible endoscope's flex part (LF)	300 mm
Weight of the module (WM)	4.02 kg
Weight of the mounting part (WP)	0.65 kg

we used OpenPose, a real-time system that extracts body, foot, hand, and facial feature keypoints from an image [31]. In addition, we used the Ros_openpose library [32] to convert the position of the keypoints in the 2D-image coordinate system to the position in the Cartesian coordinate system. We always set the target point slightly farther from the face for safety.

1) COORDINATES REGISTRATION

The slave stage system consists of five coordinate systems: the global coordinate system {0}, the robot's distal end frame {6}, the end-effector frame of the flexible endoscope module $\{e\}$, the camera frame $\{c\}$, the face frame $\{f\}$ and the target frame $\{t\}$.

The homogeneous transformation ${}^{0}_{6}T$ represents the position and orientation at the distal end of the robotic arm with respect to the global coordinate system [33]. The relationship between the 4-DOF end effector module and the robotic arm's distal end frame is represented by ${}^{6}_{e}T$. The kinematics of the whole slave system comprises a 4-DOF end effector and the

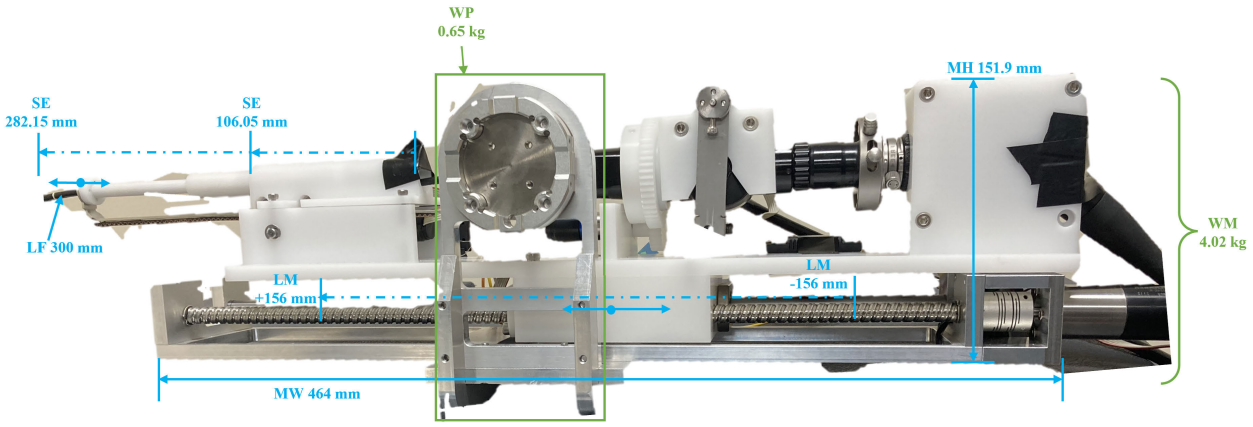


FIGURE 8. Physical indices of the designed end-effector module.

robot arm can be represented as follows.

$$ {}^{0}_{e}T = {}^{0}_{6}T\,{}^{6}_{e}T \tag{1} $$

Specifically, ${}^{6}_{e}T$ can be expressed as the product of T_f, R_f, and Tl_f, where T_f is the transformation matrix that represents the relation between the robot arm's distal end and the end effector's distal end when the end effector is in the home position, R_f is a 4×4 rotation matrix that represents the rotation DOF of the end effector, and Tl_f is a 4×4 linear translation matrix that represents the linear motion DOF of the module. θ in R_f is rotation angle and a_f in Tl_f refers to the distance due to the end effector's linear motion DOF.

$$ {}^{6}_{e}T = T_f\,R_f\,Tl_f \tag{2} $$

$$ T_f = \begin{bmatrix} 1 & 0 & 0 & 0.4083 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & -0.142 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3} $$

$$ R_f = \begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4} $$

$$ Tl_f = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & a_f \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5} $$

The remaining DOFs of 4-DOF, such as flex part support and tilting motion, are not included as independent DOF because they depend on other DOFs or are fine-tuned during the manual process in the patient's body.

Because the camera is mounted on the 6$^{\text{th}}$ joint of the robot arm, the relationship between the robot's distal-end frame and the camera frame is denoted as ${}^{6}_{c}T$. The face frame relative to the camera frame is denoted by ${}^{c}_{f}T$. The target point is set according to the examination area. ${}^{t}_{f}T$ represents the relationship between the target and face frames. All the frames are shown in Fig. 9.

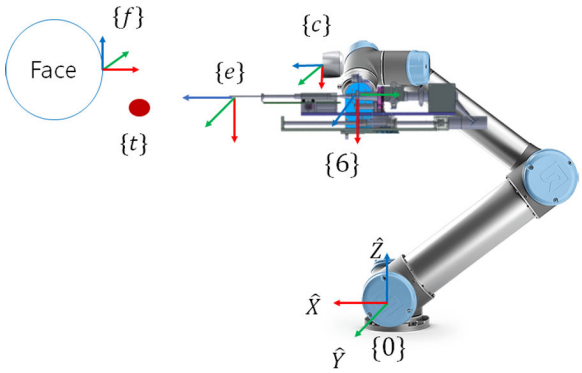


FIGURE 9. Coordinate systems of the slave stage.

2) AUTOMATIC POSITIONING

A positioning task is performed when the end-effector module is located at the target point [34]. It is represented by

$$ {}^{e}_{t}T = {}^{0}_{e}T^{-1}\,{}^{0}_{c}T\,{}^{c}_{f}T\,{}^{f}_{t}T = I[4 \times 4]. \tag{6} $$

The positioning process is as follows: First, the examination area is selected. The face is then recognised, and the target point is determined. We set the target point near the target feature outside the body. Next, the position of the target point is calculated relative to the global frame by calculating the forward kinematics as follows:

$$ {}^{0}_{t}T = {}^{0}_{c}T\,{}^{c}_{f}T\,{}^{f}_{t}T. \tag{7} $$

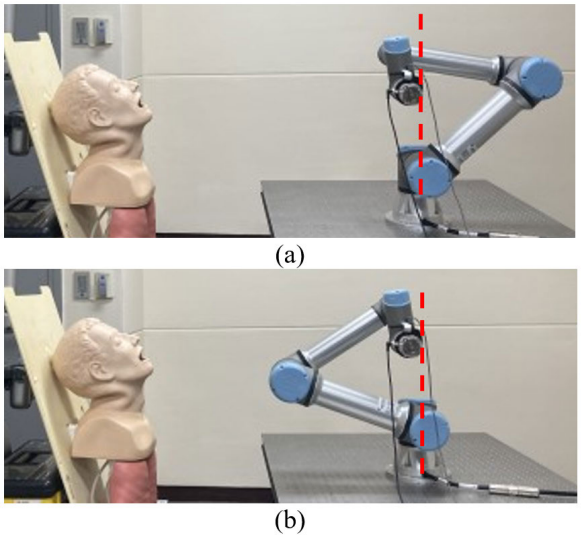
Then, the position of the target point relative to the global frame is set as the new end-effector position. The distal end position of the new robot arm is solved by calculating ${}^{0}_{6}T = {}^{0}_{e}T\,{}^{6}_{e}T^{-1}$. Finally, the inverse kinematics is solved, and the robot arm is moved to a new position to locate the end effector at the target.

C. SELECT THE INVERSE KINEMATICS SOLUTION OF THE ROBOT ARM

We assumed that the idea of a robotic arm being close to the patient during the diagnostic process would stress the patient. Therefore, when we selected a solution for the inverse kinematics of a robot arm, two poses were compared: the elbow-up and the elbow-down poses. According to [33], elbow joint parameter θ_3 can be obtained using Equation (8). ${}^1P_{4xz}$ is the origin of the 4th joint frame with respect to the 1st frame, which is described on the xzplane of the 1st frame, and a_2 and a_3 are the DH parameters. The elbow up and down poses are represented by the positive and negative sign of Equation (8), respectively. The two poses are shown in Fig. 10.

$$\theta_3 = \pm \arccos \frac{|{}^1P_{4xz}| - {a_2}^2 - {a_3}^2}{2a_2a_3}. \tag{8}$$

Among them, the elbow-up pose with a small volume protruding toward the patient was selected. Other solutions were not considered owing to limitations in terms of the experimental environment, a meaningful posture for diagnosis, and the convenience of mounting the flexible endoscope end-effector module.



(a)

(b)

FIGURE 10. Comparison of (a) elbow-up pose and (b) elbow-down pose of the robot arm. There are fewer protruding parts of the robot toward the phantom direction in the case of (a). The center of the robot arm's base is indicated by a dashed red line.

IV. EXPERIMENTS AND RESULTS

The design of the GUI used for the usability test is illustrated in Fig. 11. The start button was created for the initialisation sequence of the module system. The initialisation sequence involved placing the module near the diagnostic target feature of the face after the face recognition sequence. These sequences were integrated for user convenience.

The UR5e control area has eight buttons to control the X, Y, and Z positions and the rotation of the 6th axis of

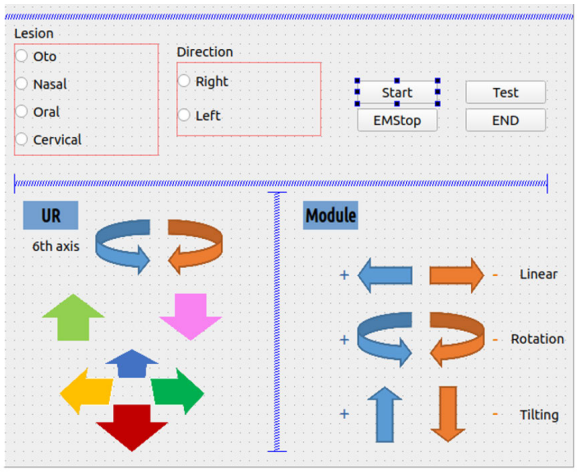


FIGURE 11. GUI used for the system implementation and experiment.

the UR5e end-effector. The UR5e side design refers to the position control of the UR5e's own control panel provided by Universal Robots.

In the endoscope module area, six buttons can control the flexible endoscope control module's insertion and reverse, left- and right-direction rotation motions, and up- and down-tilting motions (Fig. 6 (1) (4)).

If the buttons are pressed and held manually, motion occurs at the constant low speed set for the patient's safety. As for the speed, expert operators determined an appropriate speed through several experiments.

We measured the motor response time when a GUI button was clicked or released. We conducted at least 10 measurements for each button. The maximum delay was found to be less than 20ms, and each has an average of approximately 12 to 15 ms.

We performed a phantom experiment to test the implementation of the system. For the experimental settings, six expert operators and three novice operators performed the scenarios five times for each target. All expert operators were from the Department of Otolaryngology at Hanyang University. All the novice operators were independent of the medical field. Operator 1 had more than 20 years of experience, and operators 2, 3, 4, 5, and 6 were in residency.

There were a total of two targets: the first target was the vocal cords, and the second target was the nasopharynx.

This scenario comprised four sequences. During the operation, the phantom's location and posture, module's home position, and robot arm's home position were constant.

The 1st sequence is the process of face recognition and moving UR5e from its home position to the ex vivo target point. The operator selected the examination area on the GUI and pushed the start button. When the start button is pressed, the camera mounted on UR5e recognises the patient's face and detects the position information of the examination area. UR5e then places the module near the examination location and is set as our ex vivo target point. In our case, the ex vivo

target point for vocal cord diagnosis was the recognised mouth position by the RGBD camera, and the ex vivo target point for the nasopharynx was 5 cm away from the recognised nose position by the RGBD camera. This is illustrated in Fig. 12.

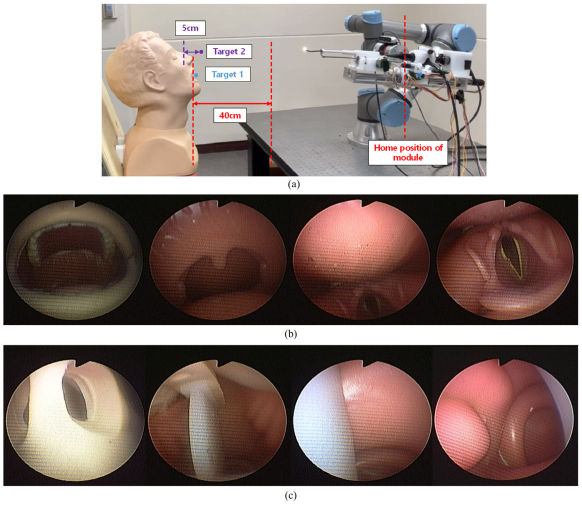


FIGURE 12. (a) Patient's location and posture, module's home position, and the ex-vivo targets that are set after the face recognition. Target 1 is for vocal cord diagnosis and Target 2 is for nasopharynx diagnosis. (b) and (c) Flexible endoscope's view during the diagnosis process, starting from the left. (b) is for Target 1 and (c) is for Target 2. The rightmost pictures represent each target.

The 2nd sequence moves from the ex vivo target to the in vivo target. An operator controls the UR5e and the endoscope end-effector module with the GUI to process the in vivo flexible endoscopic diagnosis. This sequence includes determining a proper pose for insertion by moving UR5e and the endoscope module outside the body and capturing the endoscopic image by moving toward the body target.

The 3rd sequence returns to its home position. After the operator confirms the in vivo target with the endoscope, the module is moved from the treatment position to the module's home position outside the body.

Finally, the 4th sequence returns UR5e to its home position. The full usability test scenario is illustrated as a flowchart in Fig. 13. The usability test physical implementation video is provided in Appendix A.

Three evaluation methods were used to analyse the experimental results. The first is the time taken for the 2nd and 3rd sequences, and their sum. The results are expressed as a learning curve. The second is the NASA-TLX index, which represents the workload with six subscales (mental, physical, temporal demands, frustration, effort, and performance) and is generally used to evaluate usability [12]. The third is the risk score for infection and other factors subjectively perceived during diagnosis using VAS from 0 to 10. This refers to the index used by [1].

The results are presented in Figures 14–19 and Tables 3–5. In Figures 14–19, the x--axis label or legend OP indicates the operator.

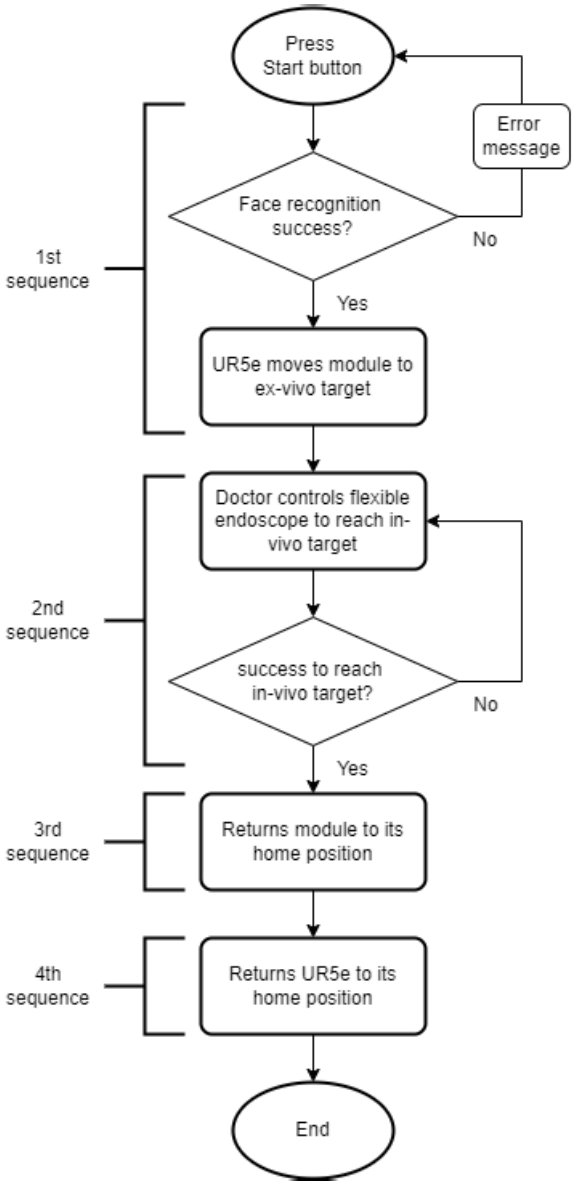


FIGURE 13. Flowchart of the usability test scenario. Each sequence is represented with annotation.

The analysis of the results is presented in the next section.

TABLE 3. NASA-TLX score of each operator.

Operator	Target	NASA-TLX
1	1	30.4 ± 10.43
	2	14.8 ± 2.17
2	1	15.8 ± 3.63
	2	10.6 ± 1.52
3	1	34.6 ± 7.37
	2	16.4 ± 7.44
4	1	19 ± 2.92
	2	19 ± 2.45
5	1	22.6 ± 3.85
	2	17 ± 2.92
6	1	18.8 ± 4.38
	2	26 ± 8.8

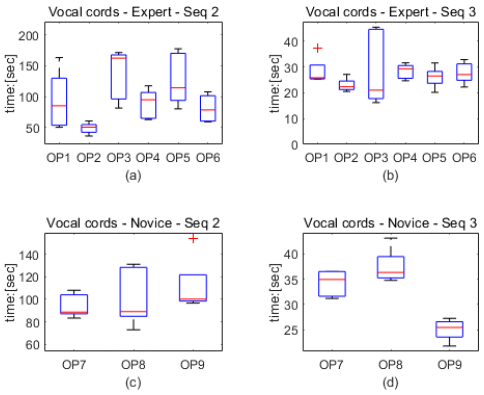


FIGURE 14. Operation time of sequence 2, 3 for Target 1 by operators(Vocal cords). (a) Operation time for sequence 2 by experts. (b) Operation time for sequence 3 by experts. (c) Operation time for sequence 2 by novices. (d) Operation time for sequence 3 by novices.

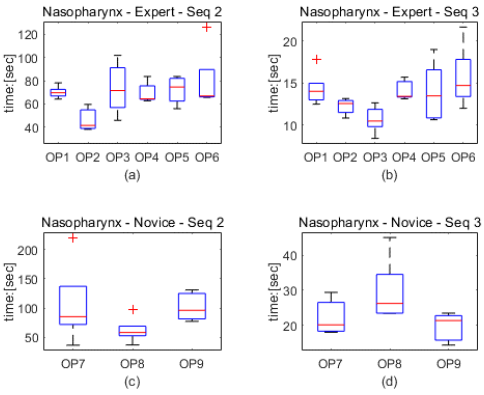


FIGURE 15. Operation time of each sequence for Target 2 (Nasopharynx) by operators. (a) Operation time for sequence 2 by experts. (b) Operation time for sequence 3 by experts. (c) Operation time for sequence 2 by novices. (d) Operation time for sequence 3 by novices.

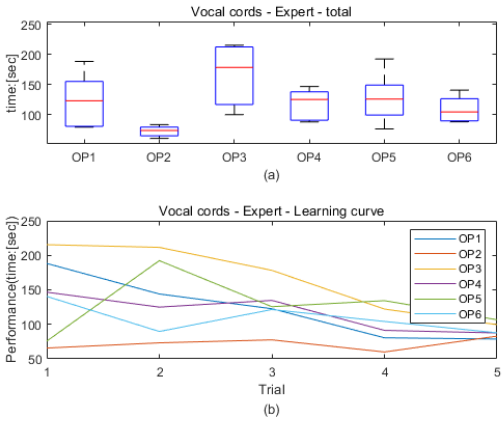


FIGURE 16. Each expert operator's total operation time and learning curve for Target 1 (Vocal cords). (a) Total operation time. (b) Learning curve.

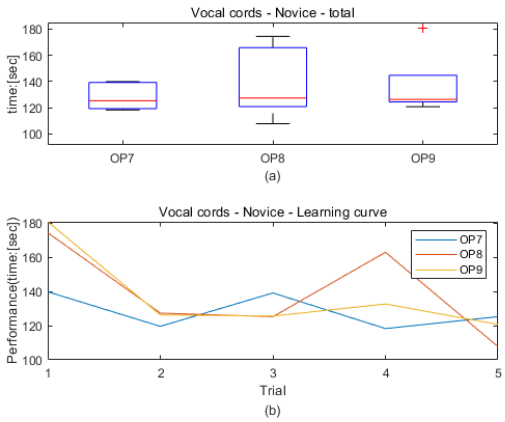


FIGURE 17. Each novice operator's total operation time and learning curve for Target 1 (Vocal cords). (a) Total operation time. (b) Learning curve.

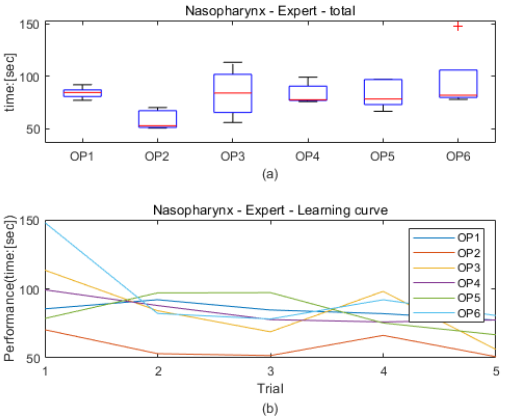


FIGURE 18. Each expert operator's total operation time and learning curve for Target 2 (Nasopharynx). (a) Total operation time. (b) Learning curve.

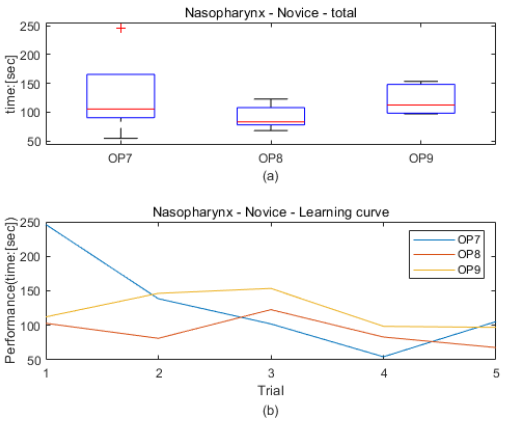


FIGURE 19. Each novice operator's total operation time and learning curve for Target 2 (Nasopharynx). (a) Total operation time. (b) Learning curve.

V. DISCUSSION

In Figures 14–19, the time required for each target of every operator can be determined. When targeting the vocal cords, the endoscope is inserted approximately 10–13 cm into the oral cavity, and when targeting the nasopharynx, the endoscope is inserted approximately 6–8 cm into the nasal

TABLE 4. NASA-TLX score of each novice operator.

Operator	Target	NASA-TLX
7	1	36.6 ± 2.19
	2	31.4 ± 5.18
8	1	28.6 ± 6.35
	2	20.8 ± 7.09
9	1	32.2 ± 2.39
	2	33.2 ± 2.95

cavity. Therefore, diagnosing vocal cords takes longer in sequence 3, as the soft part goes deeper into the body. Similarly, in sequence 2, targeting the nasopharynx takes less time then vocal cords. The main reason is that the examination difficulty varies between the two targets; however, it is also influenced by the precedent experiment on the vocal cords.

The results varied depending on the difficulty of each target, the operator's operation order, and the skill or experience of each operator. According to the learning curve in Fig. 16–19, after several trials, the completion time generally decreased, implying that the learning speed was fast. It was found that the operator's career was not a factor, because the robotic approach was new to all operators. However, skill, according to the experience of dealing with mechanical systems in daily life, may play an important role. It was found to be significant in the performance of the second operator, who was familiar with games. She demonstrated breakneck speed from the first trial, which was continuously maintained without significant changes. In addition, in the case of the novice operator, the overall operation time tended to decrease over several trials. Operators unfamiliar with the existing diagnostic procedures can quickly learn how to use the system. Like in the case of experts, similar factors also have an effect. However, the operation time was longer on average compared to expert operators because they were unfamiliar with the human body structure or the use of flexible endoscopes. Additionally, if the target is missed in the middle of the operation, it may take longer than in the previous trial.

The NASA-TLX is an index that evaluates system usability. The averages and standard deviations of the total scores for each target are listed in Tables 3–4. In general, Target 1 had a long completion time owing to great examination difficulty, and proportionally, the NASA-TLX score was also high in the case of Target 1. Only the 6^{th} and 9^{th} operators demonstrated higher scores for the second target. Furthermore, as can be seen from the expert learning curve in Fig. 16 and 18, because the completion time changed more in target 1 than in target 2, it can be confirmed that the standard deviation is generally more significant in target 1. Therefore, this shows that the system can learn quickly and easily.

Novice learning curves are shown in Fig. 17 and 19. Although both tended to decrease, there is no significant difference in the rate of the decrease. Thus, in the case of the novice operator, the standard deviation was generally significant in the second target, which showed an opposite tendency to the expert operator case. Because there was no significant difference in the decreasing rate, and considering that the NASA-TLX is the result of a sensory survey, it may have appeared different from the expert's evaluation.

Because the phantom posture or position did not change, the initial face recognition position did not differ significantly. However, there was a slight difference depending on the performance of the camera and the experimental environment. As a result, the operator could perceive a difference in usability in the process of moving the robot arm and module from the ex-vivo target point to a position where it is easy to insert into the body. In Fig. 18(b), in the case of a low-difficulty target, because the path through the human body is simple, making a proper insertion pose outside the body is as important as gaining experience through multiple attempts.

However, in the case of the novice, it was not easy to obtain an appropriate initial insertion position ex-vivo; therefore, it was not a factor that significantly influenced the operation time.

Several clinical trials are required to automatically create a good insertion pose after the face recognition process. Additionally, the ex-vivo target point must be closer to the subject. As a result, the diagnosis process is expected to be further simplified to reduce the time and workload of the doctor.

The decrease in psychological anxiety caused by the non-contact system is shown in Table. 5. Doctors are extremely anxious when using the conventional diagnostic method with contact; however, if it is performed remotely, there is little risk of infection because operators are completely separated from the patient unless it is an emergency or an exceptional situation. In the case of Novicer, a similar trend was observed. To make the system completely remote, sterilisation of the entire robotic system must be performed automatically, which is currently under study.

In summary, it was confirmed that the proposed non-contact robot system reduces the psychological risk felt by the operator and is easy to use.

In the 1^{st} sequence of the experimental scenario, we set the target position 5 cm from the patient's face for safety reasons. However, position errors can occur in the recognition process because the facial recognition algorithm may measure the position slightly differently due to the curvature of the nose or mouth target. Although using an external sensor is one way to eliminate the error, in our system, we employ the master-slave system architecture, which can adjust for such prior position errors. Therefore, we excluded the external measurement processes.

We tested the positional stability and motion speed of the end effector frame (6-th joint's frame origin) of a UR5e robot arm with and without mounting a 4 DOF end effector module. We measured the values displayed on the UR5e control panel. The results showed that the motion speed was only affected by the UR5e's own speed setting, not by adding the 4 DOF module. Thus, the magnitude of positional changes in the end effector's location was found to be within $\pm$0.4 0.6mm in each x, y, and z direction, regardless of the presence of the 4 DOF module. In conclusion, the addition of the end

TABLE 5. Subjective risk score of each operator. The C in the table means contact case and R in the table means remote case.

Operator	1	2	3	4	5	6	7	8	9
C	6	7	6	7	6	5	7	7	7
R	1	1	0	2	2	0	1	0	7

effector did not significantly impact the stability or motion speed of the UR5e robot arm.

We only applied the visual feedback through the camera attached to the UR5e. This choice is based on the survey of expert operators 78% of which prefer visual feedback rather than haptic feedback.

However, in future clinical trials, adding haptic feedback and warning with sound or light will be considered when the threshold is exceeded.

In the simulation of the two-robot arm diagnostic system, we employed a parallel robot to control other medical instruments. This is because a parallel robot generally has a high payload and high precision. In addition, it can be designed with a compact size such that it can be placed on a diagnostic station.

Motion simulation is aimed at a scenario in which a doctor's sight is guaranteed with a flexible endoscope, followed by a simple treatment or diagnosis performed with instruments mounted on a parallel robot.

Consequently, we successfully implemented this simulation scenario. However, it is necessary to modify the end of the flexible endoscope module more compactly so that the instruments and flexible endoscope end do not interfere with each other when inserted through the same side of the nasal cavity.

VI. CONCLUSION

In this study, we propose a two-robot arm diagnosis system that can be applied in an otolaryngology clinic room. The novel system includes a 6-DOF serial arm, a newly designed 4-DOF end-effector module for driving a flexible endoscope, and a parallel robot for controlling other medical instruments.

The first contribution of this work is the novel 4-DOF module that enables control of the flexible endoscope with one hand so that the flexible endoscope can be used with other instruments simultaneously.

Physical implementation was performed using the novel 4-DOF module. The architecture is a master-slave structure, and the slave stage is controlled by the GUI of the master stage. The slave stage comprises a robot arm for positioning medical instruments, a face recognition camera mounted on the end of the robot arm, and an end-effector module for driving a flexible endoscope, which was patented by our research team [35]. The robot arm, which was mounted with the flexible endoscope module, was set to the configuration in which the patient was least anxious during diagnosis, based on the experiences of doctors and patients. For face recognition, we registered the camera, robot arm, and patient frame.

As the second contribution, usability tests were conducted. The experiment was conducted by expert and non-expert groups. As a result, the operators quickly learned how to use the proposed system. The physical fatigue experienced by the doctor was less when diagnosing with the new system than with the manual method, in which two hands were used to support the flexible endoscope.

As the third contribution, we propose a DT architecture for our system and design a digital space. We conducted a 3D simulation of the two-robot arm diagnostic system in a virtual environment to optimise the entire robotic system.

The goal of our future work is to improve the remote robot system and develop a suitable process such that diagnosis can be carried out at a similar speed or faster than the manual process. We will also develop a physical end-effector module for other diagnostic instruments that can be applicable to our solution.

Furthermore, we are planning to physically implement the entire two-robot arm diagnostic system, which is comprised of a proposed endoscope diagnosis system, a parallel robot, and a series of new end-effector modules for other diagnostic tools. Consequently, an advanced DT will be built and the corresponding physical system will be established. We will conduct a UX study by incorporating patients in the novel robot diagnosis system in this virtual environment.

In addition, we will conduct a clinical trial in patients to validate the effectiveness of a non-contact otolaryngology diagnosis and treatment. Moreover, we plan to develop an otolaryngology robot system with autonomous diagnosis by adding a reinforcement-learning-based robot control algorithm.

APPENDIX A
PHANTOM EXPERIMENT VIDEO

This appendix is a phantom experimental video using the flexible endoscope manipulation system introduced in this paper. It is a total of 3 minute 14 seconds of video. In the video, the first diagnostic target is the vocal cords, and the second target is the nasopharynx.

The video shows a slave stage side and two views streaming to the master PC. The slave stage has a robot arm, a camera, and a flexible endoscope control module. As the Master PC view, the RGBD camera and endoscope view are displayed to recognize facial features and secure the doctor's field of view.

APPENDIX B
TWO-ARM ROBOT SIMULATION VIDEO IN DIGITAL SPACE

This appendix is a video of two-arm robot simulation scenario from Section. II-B in digital space. It is a total of 1 minute 6 seconds of video. In the video, the first diagnostic target is the vocal cords, and second target is nasopharynx.

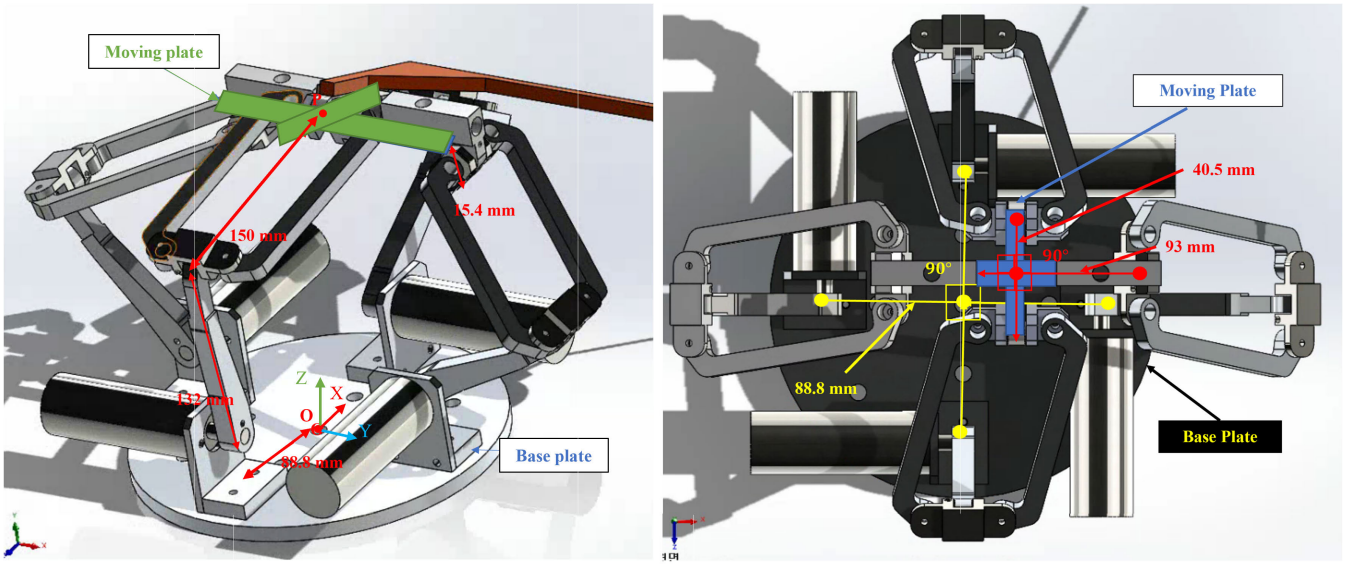


FIGURE 20. (a) 4 DOF Parallel manipulator configuration with 3D cad model. (b) Moving plate and base plate's offset angles.

The video shows a side view of slave stage. The slave stage has a serial robot arm which controls flexible endoscope module and a parallel robot which controls medical instrument.

APPENDIX C
PARALLEL ROBOT

This appendix includes the optimal design parameters, workspace analysation, and actuator sizing of the parallel robot.

We used a 4 DOF RRPaRR-type parallel manipulator previously developed by our team at [10] for medical instrument control. The optimal design parameters of the 4-DOF manipulator are shown in Fig. 20 and Table. 6.

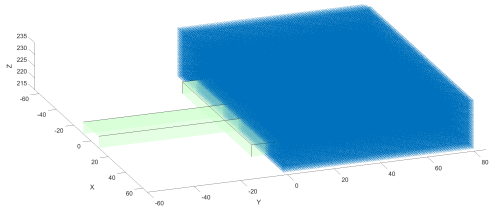


FIGURE 21. 4 DOF parallel manipulator required workspace visualisation (The green cross represents the moving plate of the parallel robot and the blue box denotes the available workspace of the parallel robot.).

The optimization of the design parameters of the parallel robot was carried out through the following process.

We place the parallel robot on the left side of the desk in order to avoid collision with flexible endoscope module. Therefore, the center of the base plate of the parallel robot was set to be located 320 mm to the left of the desk with respect to the sagittal plane of the patient's body and 192 mm away from the nose in front of the patient.

TABLE 6. Optimal kinematic design parameters.

Body	Optimal design parameters
Radius of the base plate (R)	88.8 mm
Radii of the moving plate RRPaR side (r_1, r_3)	40.5 mm
Radii of the moving plate RRPaRR side (r_2, r_4)	93 mm
Offset angles on the base plate (γ_{ti})	$\gamma_{ti} = 0°, 90°, 180°, 270°$ for $i = 1, 2, 3, 4$
Offset angles on the moving plate (γ_{bi})	$\gamma_{bi} = 0°, 90°, 180°, 270°$ for $i = 1, 2, 3, 4$
Lower link length (l_{i1})	132 mm for $i = 1, 2, 3, 4$
Upper link length (l_{i2})	150 mm for $i = 1, 2, 3, 4$
Offset distance (Z_{offset})	15.4 mm

Fig. 20 shows the base frame and frame's origin O on the base plate and the ouput frame's origin P on the moving plate. The x-axis is the instrument insertion direction, the positive y-axis is direction toward the patient, the z-axis is the height adjustment direction, and the pitch angle is the instrument's angle. The home position of the parallel robot was set to be (0,0) for the x and y and 220 mm for the z coordinate.

We employed a extended link to hold the diagnostic instrument as shown in Figure. 7. The length of the extended link was set to be 290 mm in width and 120 mm in height. With the suction attached to the extension link, the distance between the center of the patient's nose and the distal end of the suction tip is 2 cm on the x-axis, 7 cm on the y-axis, and 1.5 cm on the z-axis. Thus, the minimum dexterous workspace required for diagnosis was set in consideration of the following reasons. According to the scenario of Section. II-B, a maximum of 4 cm insertion and retraction are required in the x direction, and adjustment in the z direction is necessary in consideration of the patient's height. In addition, it should be possible to adjust the pitch during the diagnosis process. Accordingly, the minimum dexterous workspace required for diagnosis is set to ±60 mm in the x-axis, 0 mm to 80 mm in the y-axis,

and 215 mm to 235 mm in the z-axis. Finally, the size of the parallel robot is decided by the kinematic optimization and the parameters of Table. 6 were obtained.

The actuator sizing of the designed parallel robot is based of [10]. As a result, the required torque τ_i is about 30 Nm, and the required speed $\dot{q}_i$ is about 7 RPM. Therefore, we decide to use DCX35L GB KL 24V motor of Maxon Motor company was selected with some margin to the calculated value, and the gear ratio was set to 326:1.

ACKNOWLEDGMENT

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2021R1I1A4A01051258).

REFERENCES

[1] C. Park, J.-M. Hwang, S. Jo, S. J. Bae, and J. Sakong, "COVID-19 outbreak and its association with healthcare workers' emotional stress: A cross-sectional study," *J. Korean Med. Sci.*, vol. 35, no. 41, pp. 1–10, 2020.

[2] Y. Shen, D. Guo, F. Long, L. A. Mateos, H. Ding, Z. Xiu, R. B. Hellman, A. King, S. Chen, C. Zhang, and H. Tan, "Robots under COVID-19 pandemic: A comprehensive survey," *IEEE Access*, vol. 9, pp. 1590–1615, 2021.

[3] K. Kim, S. Choi, and J. Seo, "Non-face-to-face sampling robot system to increase safety and convenience for medical staff," *Robot Hum.*, vol. 18, no. 2, pp. 11–16, 2021.

[4] F. M. Marty, K. Chen, and K. A. Verrill, "How to obtain a nasopharyngeal swab specimen," *New England J. Med.*, vol. 382, no. 22, p. e76, May 2020.

[5] L. L. Hammitt, S. Kazungu, S. Welch, A. Bett, C. O. Onyango, R. N. Gunson, J. A. G. Scott, and D. J. Nokes, "Added value of an oropharyngeal swab in detection of viruses in children hospitalized with lower respiratory tract infection," *J. Clin. Microbiol.*, vol. 49, no. 6, pp. 2318–2320, Jun. 2011.

[6] C. Kim, J. A. Ahmed, R. B. Eidex, R. Nyoka, L. W. Waiboci, D. Erdman, A. Tepo, A. S. Mahamud, W. Kabura, M. Nguhi, P. Muthoka, W. Burton, R. F. Breiman, M. K. Njenga, and M. A. Katz, "Comparison of nasopharyngeal and oropharyngeal swabs for the diagnosis of eight respiratory viruses by real-time reverse transcription-PCR assays," *PLoS ONE*, vol. 6, no. 6, Jun. 2011, Art. no. e21610.

[7] J. Lee, "Robot thru and which automatically tests corona and to be presented at Hoseo University AI Week.(dizzo.com)," Digitalchosun News, Dec. 26, 2022.

[8] *Automated Robot Takes Swabs for Safe COVID-19 Testing*, Medgadget Editors, Eugene, OR, USA, Jun. 2020.

[9] *Robotic Clinicians for Taking Nasal Swabs During COVID Pandemic*, Medgadget Editors, Eugene, OR, USA, Aug. 2020.

[10] B.-J. Yi, S. M. Kim, H. K. Kwak, and W. Kim, "Multi-task oriented design of an asymmetric 3T1R type 4-DOF parallel mechanism," *Proc. Inst. Mech. Eng., C, J. Mech. Eng. Sci.*, vol. 227, no. 10, pp. 2236–2255, Oct. 2013.

[11] J. Seo, S. Shim, H. Park, J. Baek, J. H. Cho, and N.-H. Kim, "Development of robot-assisted untact swab sampling system for upper respiratory disease," *Appl. Sci.*, vol. 10, no. 21, p. 7707, Oct. 2020.

[12] S. G. Hart, "NASA-task load index (NASA-TLX); 20 years later," *Proc. Hum. Factors Ergonom. Soc. Annu. Meeting*, vol. 50, no. 9, pp. 904–908, Oct. 2006.

[13] B. R. Barricelli, E. Casiraghi, and D. Fogli, "A survey on digital twin: Definitions, characteristics, applications, and design implications," *IEEE Access*, vol. 7, pp. 167653–167671, 2019.

[14] T. Erol, A. F. Mendi, and D. Dogan, "The digital twin revolution in healthcare," in *Proc. 4th Int. Symp. Multidisciplinary Stud. Innov. Technol. (ISMSIT)*, Oct. 2020, pp. 1–7.

[15] G. Coorey, G. A. Figtree, D. F. Fletcher, V. J. Snelson, S. T. Vernon, D. Winlaw, S. M. Grieve, A. McEwan, J. Y. H. Yang, P. Qian, K. O'Brien, J. Orchard, J. Kim, S. Patel, and J. Redfern, "The health digital twin to tackle cardiovascular disease—A review of an emerging interdisciplinary field," *NPJ Digit. Med.*, vol. 5, no. 1, p. 126, Aug. 2022.

[16] *Digital Twins and Robots in Medical Device Development and Testing Auriga. Outsourcing Software Development*, Airat Sadykov, Kazan, Russia, Nov. 2019.

[17] A. A. Malik and A. Brem, "Digital twins for collaborative robots: A case study in human-robot interaction," *Robot. Comput.-Integr. Manuf.*, vol. 68, Apr. 2021, Art. no. 102092.

[18] *UniversalRobots/Universal_Robots_Ros_Driver: Universal Robots ROS Driver Supporting CB3 and e-Series*, GitHub, San Francisco, CA, USA.

[19] J. Woo, J. H. Choi, J. T. Seo, T. I. Kim, and B.-J. Yi, "Development of a robotic colonoscopic manipulation system, using haptic feedback algorithm," *Yonsei Med. J.*, vol. 58, no. 1, pp. 139–143, Jan. 2017.

[20] S. G. Lim, "The development of robotic flexible endoscopic platforms," *Int. J. Gastrointestinal Intervent*, vol. 9, no. 1, pp. 9–12, Jan. 2020.

[21] J. G. Ruiter, G. M. Bonnema, M. C. van der Voort, and I. A. M. J. Broeders, "Robotic control of a traditional flexible endoscope for therapy," *J. Robotic Surg.*, vol. 7, no. 3, pp. 227–234, Sep. 2013.

[22] Z. Li and P. W.-Y. Chiu, "Robotic endoscopy," *Visceral Med.*, vol. 34, no. 1, pp. 45–51, Feb. 2018.

[23] P. Li, H. M. Yip, D. Navarro-Alarcon, Y. Liu, C. F. M. Tong, and I. Leung, "Development of a robotic endoscope holder for nasal surgery," in *Proc. IEEE Int. Conf. Inf. Autom. (ICIA)*, Yinchuan, China, Aug. 2013, pp. 1194–1199.

[24] C. Song, X. Ma, X. Xia, P. W. Y. Chiu, C. C. N. Chong, and Z. Li, "A robotic flexible endoscope with shared autonomy: A study of mockup cholecystectomy," *Surgical Endoscopy*, vol. 34, no. 6, pp. 2730–2741, Jun. 2020.

[25] H. Poon, C. Li, W. Gao, H. Ren, and C. M. Lim, "Evolution of robotic systems for transoral head and neck surgery," *Oral Oncol.*, vol. 87, pp. 82–88, Dec. 2018.

[26] D. Lomanto, S. Wijerathne, L. K. Y. Ho, and L. S. J. Phee, "Flexible endoscopic robot," *Minimally Invasive Therapy & Allied Technol.*, vol. 24, no. 1, pp. 37–44, Jan. 2015, doi: 10.3109/13645706.2014.996163.

[27] H.-S. Yoon, J. H. Jeong, and B.-J. Yi, "Image-guided dual master–slave robotic system for maxillary sinus surgery," *IEEE Trans. Robot.*, vol. 34, no. 4, pp. 1098–1111, Aug. 2018.

[28] T. A. D. C. Visconti, J. P. Otoch, and E. L. D. A. Artifon, "Robotic endoscopy. A review of the literature," *Acta Cirúrgica Brasileira*, vol. 35, no. 2, 2020, Art. no. e202000206.

[29] X. Wang, J. Yan, X. Ma, J. Y. K. Chan, R. H. Taylor, S. S. Cheng, and K. W. S. Au, "Hybrid-structure hand-held robotic endoscope for sinus surgery with enhanced distal dexterity," *IEEE/ASME Trans. Mechatronics*, vol. 27, no. 4, pp. 1863–1872, Aug. 2022.

[30] J. Gafford, M. Freeman, L. Fichera, J. Noble, R. Labadie, and R. J. Webster, "Eyes in ears: A miniature steerable digital endoscope for trans-nasal diagnosis of middle ear disease," *Ann. Biomed. Eng.*, vol. 49, no. 1, pp. 219–232, Jan. 2021.

[31] Z. Cao, G. Hidalgo, T. Simon, S. E. Wei, and Y. Sheikh, "OpenPose: Realtime multi-person 2D pose estimation using part affinity fields," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 1, pp. 172–186, Jan. 2021.

[32] R. Joshi, *Ros_Openpose*, Kyutech, Japan, Dec. 2022.

[33] R. S. Andersen, "Kinematics of a UR5," Aalborg Univ., Denmark, Tech. Rep., 2018.

[34] S. Hutchinson, G. D. Hager, and P. I. Corke, "A tutorial on visual servo control," *IEEE Trans. Robot. Autom.*, vol. 12, no. 5, pp. 651–670, Oct. 1996.

[35] S. Lee and B.-J. Yi, "Insertion drive device for telemedicine and operating method thereof," Korea Patent 10 2021 0 192 999, 2021.

SHO-HYUN LEE received the B.S. degree from the Department of Electrical and Electronic Engineering, Hanyang University, Ansan, South Korea, in 2020, where she is currently pursuing the integrated Ph.D. degree. Her interests include medical robots and artificial intelligence.

DONG-WOO LEE received the B.S. degree from the Department of Electrical and Electronic Engineering, Hanyang University, Ansan, South Korea, in 2021, where he is currently pursuing the master's degree. His research interests include mobile robots and artificial intelligence.

HWA-SEOB SONG received the B.S. degree in electronics and communication engineering from Hanyang University, Ansan, South Korea, in 2016, and the Ph.D. degree in electrical and electronic engineering from Hanyang University, Seoul, South Korea, in 2022. He is currently a Postdoctoral Fellow with the Research Institute of Engineering and Technology, Hanyang University. His research interests include medical robots and artificial intelligence.

SEONMIN JEONG received the M.D. and Ph.D. degrees in medicine from Hanyang University, Seoul, South Korea, in 2016 and 2022, respectively. She has been holding a fellowship with the Department of Otolaryngology-Head and Neck Surgery, Hanyang University Guri Hospital, Guri-si, Gyeonggi-do, since March 2021.

YONGBAE JI received the M.D. and Ph.D. degrees in medicine from Hanyang University, Seoul, South Korea, in 2001 and 2006, respectively. From March 2013 to February 2017, he was an Assistant Professor with the Department of Otolaryngology-Head and Neck Surgery, Hanyang University Guri Hospital, Guri-si, Gyeonggi-do. From September 2017 to July 2018, he was a Visiting Scholar with The Johns Hopkins Hospital, Baltimore, MD, USA. He was a Visiting Professor with The Johns Hopkins University, in 2017. He is currently a Professor with the Department of Otolaryngology-Head and Neck Surgery, Hanyang University. His research interests include head and neck oncology, robotics, and endoscopic surgery in the head and neck area. He is a member of the Korean Society of Head and Neck Surgery and an Editorial Board Member of *Clinical and Experimental Otorhinolaryngology* and *Korean Journal of Otorhinolaryngology-Head and Neck Surgery*. He is also the Vice Secretary General of the Asia-Pacific Society of Thyroid Surgery and the Director of foreign relations of the Korean Society of Head and Neck Surgery.

JI-SUNG SONG received the bachelor's degree in applied arts, the master's degree in industrial design, and the Ph.D. degree in design from Seoul National University, in February 1986, February 1997, and February 2011, respectively. He was appointed as a Professor with the College of Design, Hanyang University, in March 1999. He is currently a Professor with the College of Design, Hanyang University. He majored in vision design (identity design + UX service design) and RX (Relationship eXperience) design. He is also the Director of the AI UX Design Center, Hanyang University. He is a Professor-in-Charge of 125 patent applications and the Director of the Korean Design and Culture Association-COPE (design, engineering, humanities convergence). He is also engaged in many activities, one of which is being the Dean of the College of Design, Hanyang University. He is a Professor-in-Charge of signing the design MOU of the Ministry of Patriots and Veterans Affairs and a Design Consultant for Hanssem Company Ltd. He is also a Design Adviser with the Bank of Korea. He was with the Seoul Design Foundation DDP Operation Advisory Committee and a member of the Saemangeum CI Advisory Committee and the Jeollabuk-do Global Brand Marketing Steering Committee. He is an Advisor of the National Branding Commission. He is a Consulting Committee Member of the Ministry of Commerce, Industry and Energy on the world's best product design brand as well as a member of the Brand Development Steering Committee of the Ministry of Planning and Budget and the Global Brand Steering Committee of the Small and Medium Business Administration.

JIYOUNG KIM received the Bachelor of Arts, Master of Education, and Doctor of Science degrees from Hanyang University, Seoul, South Korea, in 1997, 2010, and 2016, respectively. Her Ph.D. thesis was the study of UX design of realistic content applying a psychological model. She is currently an Assistant Professor of intelligent robotics with the College of Engineering, Hanyang University. Her research interests include convergence education and robot UX design.

BYUNG-JU YI (Member, IEEE) received the B.S. degree in mechanical engineering from Hanyang University, Seoul, South Korea, in 1984, and the M.S. and Ph.D. degrees in mechanical engineering from The University of Texas at Austin, Austin, TX, USA, in 1986 and 1991, respectively. From January 1991 to August 1992, he was a Postdoctoral Fellow with the Robotics Group, The University of Texas at Austin. From September 1992 to February 1995, he was an Assistant Professor with the Department of Mechanical and Control Engineering, Korea Institute of Technology and Education, Cheonan, Chungnam, South Korea. In March 1995, he joined the Department of Control and Instrumentation Engineering, Hanyang University. He was a Visiting Professor with The Johns Hopkins University, Baltimore, MA, USA, in 2004, and a JSPS Fellow with Kyushu University, Japan, in 2011. He is currently a Professor with the Department of Electronic Systems Engineering, Hanyang University. His research interests include general robot mechanics with application to surgical robotic systems (ENT, neurosurgical, and needle insertion areas), deep-learning-based robotic manipulation, and ubiquitous sensor network-based robotics. He is a member of the IEEE Robotics and Automation Society. He was the President of the Korean Society of Medical Robotics, from 2015 to 2018. He was the President of the Korean Robotics Society, in 2019. He was an Associate Editor of the IEEE TRANSACTIONS ON ROBOTICS from 2005 to 2008.