

Intelligent Sampling of Anterior Human Nasal Swabs using a Collaborative Robotic Arm

Roman Parak[✉], Martin Juricek[✉]

Institute of Automation and Computer Science, Brno University of Technology, Czech Republic
Roman.Parak@vutbr.cz[✉], Martin.Juricek1@vutbr.cz[✉]

Abstract

Advanced robotics does not always have to be associated with Industry 4.0, but can also be applied, for example, in the Smart Hospital concept. Developments in this field have been driven by the coronavirus disease (COVID-19), and any improvement in the work of medical staff is welcome. In this paper, an experimental robotic platform was designed and implemented whose main function is the swabbing samples from the nasal vestibule. The robotic platform represents a complete integration of software and hardware, where the operator has access to a web-based application and can control a number of functions. The increased safety and collaborative approach cannot be overlooked. The result of this work is a functional prototype of the robotic platform that can be further extended, for example, by using alternative technologies, extending patient safety, or clinical tests and studies. Code is available at https://github.com/Steigner/Robo_Medicinae_I.

Keywords: Robotics, Smart Hospital, Convolution Neural Network (CNN), U-Net, ASPOCRNet, Robot Operating System (ROS)

Received: 15 April 2022
Accepted: 22 June 2022
Online: 23 June 2022
Published: 30 June 2022

1 Introduction

The influx of COVID-19 has caused mortality and often painful diseases throughout the world. One of the main problems in health facilities has been the shortage and also the protection of medical personnel from the spread of the virus.

Due to the infection of the medical personnel, many departments have been forced to suspend scheduled examinations and possible operations for a period of time. Therefore, in these difficult times and with the prospect of another pandemic, the purpose of this proposal aims to innovate and provide a solution to nasal swab collection. Intensive research and development aimed at combating this disease is expanding into every possible field, including robotics and automation. Medical robotics as a scientific field has been evolving in recent years, and modern robots (da Vinci Surgical System [7], KUKA LBR Med [9]) working in the real world should autonomously learn new tasks and flexibly adapt to various changes. In this struggle, robotics plays the role of an assistant to medical personnel. For this reason, the Robo Medicinae I project was created to build on existing research on test samples analysis [17] at the Institute of Automation and Computer Science, Faculty of Mechanical Engineering, Brno University of Technology as part of the Industry 4.0 Cell (I4C) robotics laboratory [14, 13].

In this paper, we propose an experimental robotic platform whose main function is the swabbing samples of nasal vestibule using the collaborating robotic arm UR3 from Universal Robots [21]. In our case, the role of the collaborative robot is to help the medical per-



Figure 1: The robotic platform called Robo Medicinae I allows intelligent collection of anterior human nasal swabs using a collaborative robotic arm UR3. Video overview: <https://www.youtube.com/watch?v=mLLzvUdUReE>

sonnel to prevent the spread of the virus and to keep it as safe as possible from the risk of infection by the patient. The basic scene of our experiment in the initial position is shown in Figure 1.

We describe related work in the use of robotic systems for intelligent sampling of anterior swabs of the human nose (Section 2 Related work), and we also summarize the necessary methods needed to create our work (Section 3 Methods).

In the main part of the work, we focus on the design of an experimental robotic platform for the collection of samples from the nasal vestibule for possible subsequent antigen testing of COVID-19 disease (Section 4 Experiments and Results). A sub-part of the work was to solve the nostril detection problem from several possible methods that deal with image segmentation, using an artificial neural network (ANN), more precisely, a convolutional neural network (CNN). Our approach compares different segmentation models, the older U-Net model (2015) and the newer ASPOCRNet model (2021).

In the final part of the paper, we focus on the challenges we have encountered, the current limitations, and future extensions of our work (Section 5 Conclusion and Future Work).

2 Related Work

Around the world, several research teams have tried to contribute their inventions to the fight against COVID-19 disease, including those related to the sampling of anterior human nasal swabs. In the following section, we briefly discuss previous work on the use of robotic systems for intelligent sampling of anterior human nasal swabs using a modular robotic arm as well as a single-purpose manipulators.

A team from the University of Southern Denmark and the company Lifeline Robotics have devised the first automated robotic solution for COVID-19 sampling [4]. This was a solution using a robot from Universal Robots, while their gripper integrated a swab holder and a camera. The head of the patient is then placed in the bracket for the patients. The bracket prevents deflecting and endangering the patient. A similar solution was created by Brain Navi Biotechnology Co., Ltd. based in Taiwan's Hsinchu's Biomedical Science Park [12]. The robot automatically detects the patient's facial structure and nostril location independently of each other and gently samples to avoid close contact with the patients.

The team that also tackled the problem of oropharyngeal swab sampling for COVID-19 diagnosis was from the Chinese University of Hong Kong [6]. Their solution is also based on a collaborative robot from Universal Robots company, whose main invention is a complex gripper that grips a rigid-flexible coupled manipulator. The complete system thus consists of a collaborative robot, a linear manipulator with a servo motor for rotation, which holds a micro pneumatic actuator. Sensors such as the 3D camera, the endoscope, and the force detection system are also mounted on the gripper. A ring illuminator for front illumination on the camera axis is also an integral part. As for the software, the authors of this article addressed the detection of the oral cavity phantom. For this detection, they used one of the artificial intelligence methods, Mask Regions with Convolutional Neural Network Features. The result of the work is a robotic application that has already been validated by several volunteers and has

areas that can be improved and subsequently clinically tested.

There are several other types of applications that use single-purpose manipulators to solve the problem of a swab sampling system [10, 18, 22].

3 Methods

This section provides a brief introduction to the theory of machine perception, focusing primarily on machine vision, as well as artificial neural networks (ANNs).

3.1 Machine Perception

One of the most progressive and researched areas is machine perception. Machine perception is concerned with extracting information about a scene by analyzing signals from sensors. It is similar to the way a person perceives the world around him with his senses. The basic physical principle of the cameras is based on incident photons that eject electrons out of the valence layer in semiconductors, which are then trapped in a potential pit (electron trap). 3D cameras augmented the classical 2D sense with additional technology that simulates human binocular vision, and therefore captures three-dimensional images in point clouds form. Augmentation technologies in 3D cameras can be in many forms, from IR sensors to software 3D reconstruction from 2D images.

One of the integral tasks of machine vision is image segmentation. The main goal is to divide the image into parts that are related to objects or areas of the real world or to separate objects from the background. This task can therefore be classified as a lower level of processing, however, segmentation is widely used in further processing for higher-level processing tasks such as 3D measurement, classification, object detection, or dimensional measurement. The process itself is the division of an image into multiple segments that meet certain criteria. For example, if the pixels in a segmented part have the same gray level or brightness, this can lead to ambiguity.

Knowledge-based image segmentation is among the more advanced algorithms as they can make segmentation very easy because they use atlas models or templates of segmented objects that are generated automatically based on the training data. These advanced techniques can be also included in hybrid methods, these ones cannot be clearly classified but they are methods based on, for example, morphology, amplitude projection, or neural networks.

Hand-Eye Calibration:

In order to perform an accurate movement to the identified nostril from the 3D model reconstructed in the application, it is necessary to perform a so-called hand-eye calibration (Fig. 2 a). Alternatively, if the application solution used a camera statically positioned outside the robot and an eye-to-hand calibration (Fig. 2 b) would need to be performed.

The hand-eye calibration process can be defined as finding the relative position and orientation between the fixed camera mounted on the end-effector and the last joint of the robotic gripper [23]. Hand-eye calibration can be generally mathematically described as:

$$\mathbf{AX} = \mathbf{XB}, \quad (1)$$

where X is the unknown hand-eye transformation, A is the robot joint data and B corresponds to the camera motion.

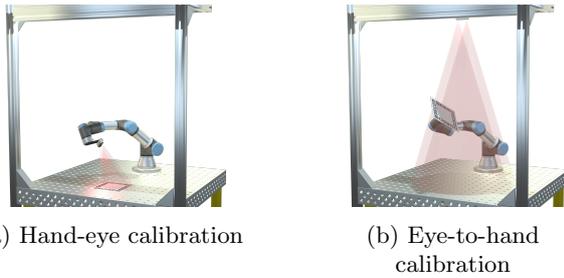


Figure 2: Hand-Eye calibration methods to determine the geometric relationship between the robot and the 3D camera.

3.2 Artificial Neural Networks

A major development in the field of artificial intelligence has been made possible by artificial neural networks in short ANNs. In its design, an artificial neural network has similarities to the biological computational model of the human brain, where signals are transmitted through a network of neurons. The first neural network can be considered the McCulloch-Pitts model (Fig. 3), which was published in 1943 [11]. Since then, the development in the field of artificial neural networks has undergone a great evolution, and since the theoretical model, artificial neural networks are currently used for a wide range of tasks.

A neural network is formed when multiple neurons are connected by inputs and outputs, but always consist of the same components, which can be expressed as the mathematical model of a neuron:

$$y(x) = f\left(\sum_{i=1}^n w_i \cdot x_i - b_i\right), \quad (2)$$

where y is the output signal, f is the activation function, x is the input signal, w are the weights and b is the bias.

Activation Functions:

Activation functions determine the output of a neuron depending on its inputs. To provide approximations of any continuous function, activation functions must be non-linear and continuously differentiable. However, not all known and used activation functions satisfy this condition.

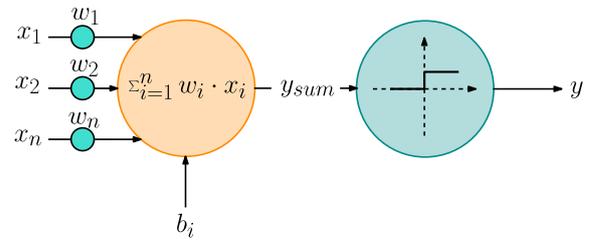


Figure 3: McCulloch-Pitts model

One of the most popular activation functions is ReLu (Rectified Linear Unit). The ReLu function can be mathematically described by the Eq. (3). Compared to others, it provides faster learning and less computational complexity [5].

$$f(x) = \max(0, x) \quad (3)$$

An activation function that finds its appropriate application in the output layer of the classification network is the Softmax (normalized exponential function). This activation function can be defined according to the Eq. (4), where the individual values determine the degree of classification.

$$f(x_i) = \frac{e^{x_i}}{\sum_{j=0}^n e^{x_j}} \quad (4)$$

Loss Functions:

The loss function is a function used to measure the prediction error rate against the training data. The loss function is chosen according to the type of problem to be solved. One of the most widely used loss functions for image segmentation using ANN is the cross-entropy loss function. Cross-entropy loss functions have many modifications [5]. Basic variants include binary cross-entropy. This method is used for two-class classification. While categorical cross-entropy is mainly used for classification into three or more classes.

The function is defined by the average difference between the actual probability distribution (y_i) and the predicted probability distribution (\hat{y}_i). The output size is the number of scalar values (N) in the output of the model.

$$L_{CE} = - \sum_{i=1}^N y_i \log(\hat{y}_i) \quad (5)$$

Neural Networks for Image Segmentation:

As mentioned above, a method that uses artificial neural networks for segmentation is classified as a hybrid method. Convolutional neural networks are based on the use of convolutional, pooling, and fully connected layers, where if an ANN has at least one convolutional neural layer, we refer to them as convolutional neural networks in the abbreviation CNN. Mathematically, it can be expressed as:

$$g(x, y) = h(x, y) * k(x, y) = \sum_{(i,j) \in O} h(x - i, y - j) \cdot k(i, j), \quad (6)$$

where g is the result image and h is the source image, k is the convolution kernel and O is the small neighborhood of the representative pixel (currently processed).

Another very important layer for image segmentation is the pooling layer. The pooling layer, like convolution, is also used for data reduction to preserve the information. This fact plays a significant role in reducing the dimensionality of feature maps and reducing computational complexity. The objective of this layer is to perform a transformation, such that one pixel in the new image is represented as a group of neighboring pixels, from the original one. The difference to the convolutional layer is that the transformation is known in advance. However, the weights and biases are unknown in the convolutional layer, this is determined from the training data. The pooling layers can be divided into max-pooling and average-pooling layers, which can be described as:

$$g(x, y) = \max_{(i,j) \in O} h(i, j), \quad (7)$$

$$g(x, y) = \sum_{(i,j) \in O} \frac{1}{mn} h(i, j), \quad (8)$$

where m is the width of the window and n is the height of the window.

In recent years, several typical segmentation models have emerged. These models have their specific architectures but may use different backbones. Backbone is not a universal technical term, which is used in *DeepLab* to refer to the feature extractor network [3]. Segmentation models from their original designs are modifiable in the form of these backbone networks which have their specific uses, examples of these backbone networks are ResNet, DenseNet, MobileNet and others.

4 Experiments and Results

4.1 Setting up the Robotic Environment

This robotic platform integrates the UR3 cobot with designed gripper on which the Intel RealSense D435i 3D camera, the OnRobot RG2 two-finger gripper and the OnRobot HEX-E force-torque sensor are mounted (Figure 4). The UR3 collaborative robotic arm from Universal Robots, more precisely a 6-axis robotic arm with a working radius of 500 mm/19.7 inches, a payload of 3 kg/6.6 pounds, and a repeatability of ± 0.1 mm [21].

The paper also developed a unique human-machine interface (HMI) (Fig. 8) through which the operator can control processes such as access to the patients

database, moving the robot, streaming the camera image, etc. The entire system is designed for simulation with the RVIZ visualization tool [20, 19] and the Gazebo 3D simulation tool [1, 8], as well as for use in real applications with the official ROS driver for Universal Robots [15, 2].

Several different areas of technology are connected in the application, such as pick-and-place, digital twin, virtualization, databases, machine vision, and 3D data processing. All of this is assembled on a complex but modular platform that creates a great foundation for tuning for real-world deployments.

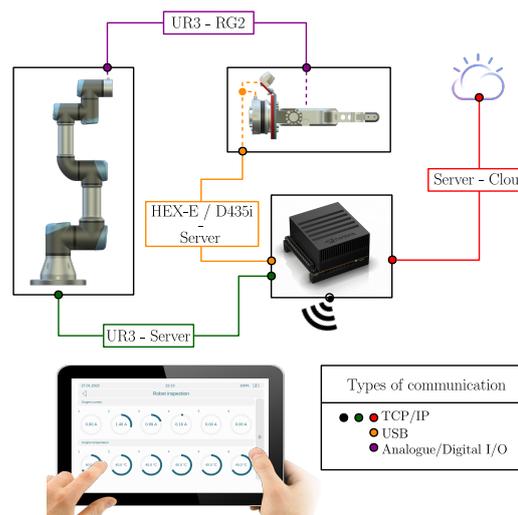


Figure 4: System integration of the robotic platform Robo Medicinae I.

4.2 Definition of the Experiment

In our problem, we chose a task focused on collecting samples using a nasal swab. The goal of the task is to create a complex facial scanning function that would take care of the detection of the center of the nostril and the actual 3D reconstruction in the form of a point cloud.

Assuming that the operator makes the device into operation, it can be defined in several basic steps:

1. The patient is positioned in a defined examination area.
2. Search the patient database based on a Personal Identification Number (PIN) or a QR code scanned with a PIN from a color image.
3. The operator activates the faceID function and patient identification is performed.
4. A face detection process is then performed to ensure that the robotic arm rotates in the correct position.
5. Identification center of nostril - segmentation model CNN to detect nostril, on the basis of this centroid calculation.

6. Face Scan - scan align color to depth image, on the basis of this reconstruct 3D point cloud.
7. In the last stage, a trajectory is autonomously generated from a initial position to a target position for sampling anterior swabs from the human nose.

Motion planning in our problem is solved using algorithms from the Open Motion Planning Library (OMPL) within the ROS Melodic Morenia distribution [15]. OMPL has also implemented collision detection and obstacle avoidance features.

As part of the solution to this task, a working prototype was created in which, however, it is necessary to perform these processes semi-automatically due to the safety of real-world testing.

4.3 Learning Process

An integral part of the whole application is the facial scanning process. This is a function in which the patient's head is scanned by using a 3D camera, followed by detection of the nostrils using a convolutional neural network. A 3D reconstruction of the image is then processed from the depth and color images. To start this process, the operator must first perform faceID and facial validation to ensure that the correct patient is scanned and in a good position.

One of the main problems encountered in the solution was the detection of nostrils. Since there is no pre-trained model for image segmentation with nostril detection, it was necessary to design a customized solution. The solution to this problem was based on a complex artificial neural network tool, more precisely a convolutional neural network segmentation model.

Two convolutional neural network segmentation models, namely U-Net [16] and ASPOCRNet [24], were selected to solve the nostril detection. However, before these models can be learned, it is necessary to collect the dataset and perform so-called masking. To solve the problem of nostril detection, regions of interest have been labeled using a polygon.

The data were collected using a 3D camera, where it was necessary to capture a colour image aligned with the depth image. This alignment is justified by the 3D reconstruction. The resolution of the captured colour image is 640×480 pixels with suppression of the environmental scene where the disparity offset has a fixed value. A total of 100 images were captured, in three different environments, and three persons were captured.

Afterwards, the images were labeled using the open-source tool *Make-Sense*. This was followed by a masking process in which three classes of masks were created. The labeled left nostril has an image value of 1, the labeled right nostril has a value of 3 and the rest of the scene has a value of 2. The data was then split into a ratio of 80% training and 20% validation. The simplified process represents in Fig. 5.

Among the final parts of the nostril, the detection process is the learning of the neural network. In this

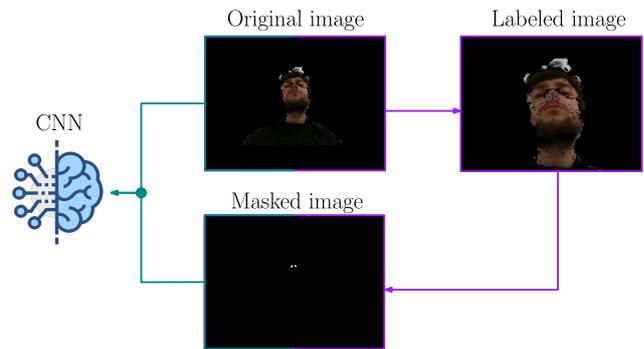


Figure 5: Pre-process: Original image - Labeled image - Masked image

part, all parameters such as the number of epochs, callbacks, and others are configured. Both convolutional neural networks models were trained on a desktop computer with the *TensorFlow* framework optimized to transfer the computation to an Nvidia GTX 1050 Ti graphics card. The hyper-parameters given in the table (Table 2) were used to compare these models.

Table 1: Hyperparameters used to detect nostrils using Convolutional Neural Network (CNN)

Hyper-parameter	Value
Number of epochs	15
Number of classes	3
Optimizer	Adam
Learning rate (α)	0.001
Steps per train epoch	train images / 4
Steps per validation epoch	test images / 4
Loss function (E)	Sparse categorical crossentropy
Activation function (ϕ)	ReLU, Softmax
Backbones	Mobilenet, Resnet, SE-ResNet, Xception
Pretrained weights	Imagenet dataset

The learning process was then monitored using the machine learning tool *Weights & Biases*. Thanks to this tool, it is possible to monitor the current loss function, the accuracy of the model, or possibly the GPU temperature or the GPU utilization after each epoch. The training termination criterion is set to reach the maximum number of epochs.

In order to be able to use the trained model output in a comprehensive way, the *ONNX (Open Neural Network Exchange)* tool was used to convert the *TensorFlow.pb* model format to the general *.onnx* format. The size of the trained model itself also plays an important role, when exporting to *.onnx* the size was reduced by about 33%.

4.4 Results of the Experiment

To verify the functionality of the entire robotic platform, wiring and real-world testing was carried out. All of the above-mentioned devices were used in the solution, from the UR3 collaborative robot to the HEX-E sensor. The control unit was represented by an Nvidia Jetson Xavier single board computer in development kit configuration.

Once the center of the nostril is detected, this point is transferred to the original colour image. Then the RGBD image creation process is performed. The conversion function of the *Open3D* library [25] creates an RGBD image from the pair of the color image and the depth image. Where the depth image is stored in a data type representing the depth value in metres.

The *Open3D* library also provides a function for the subsequent calculation of point clouds by conversion of the RGBD image. To perform the conversion, it is necessary to specify the parameters of the 3D camera. These parameters includes depth scale d_s , the focal length (f_x, f_y) and the optical center (c_x, c_y) . The parameters of the RGBD image are the resolution (a, b) and the depth d [25]. The calculation of the point cloud is then performed based on Eqs. (9) - (11).

$$z = \frac{d}{d_s} \quad (9)$$

$$x = \frac{(a - c_x) \cdot z}{f_x} \quad (10)$$

$$y = \frac{(b - c_y) \cdot z}{f_y} \quad (11)$$

The point cloud result of 3D face reconstruction plays a crucial role in calculating the trajectory of the robot's movement towards the detected center of the patient's nostril.

In the context of the application, the center of the nostril is then calculated from the classified pixels. The relationship for the calculation can be defined as follows:

$$c_{x,y} = \frac{\sum_{n=1}^N p_{x,y}}{l_{x,y}}, \quad (12)$$

where $c_{x,y}$ represent center on the x, y -axis, $p_{x,y}$ pixels of interest on the x, y -axis, and $l_{x,y}$ is number of pixels of interest on the x, y -axis.

Table 2: Results of testing 5 subjects after 15 attempts in a real-world (75 attempts in total).

Task	Success Rate
Nostrils Detection	96.00%
Motion Simulation	93.33%
Real-World Movement	38.33%

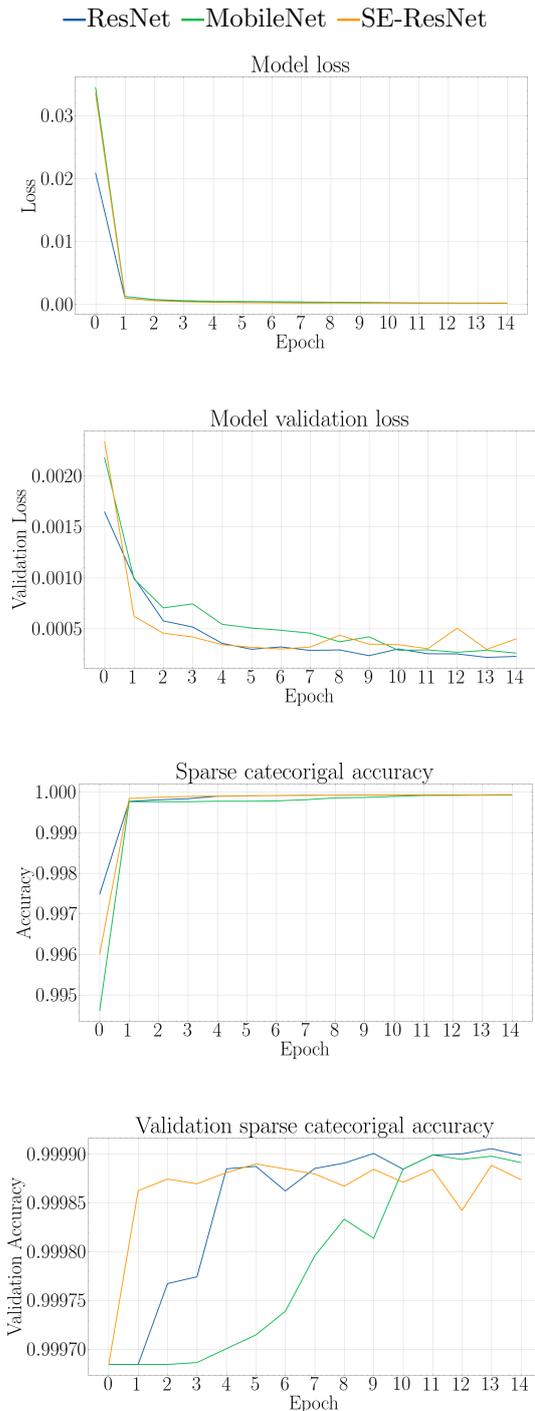


Figure 6: Training results of the U-Net for nostril detection.

5 Conclusion and Future Work

In this work, we provided a design for an experimental robotic platform. In this context, a robotic platform with the name Robo Medicinæ I was developed. The main function of this platform is the collection of samples from the nasal vestibule for possible subsequent antigen testing of COVID-19 disease. The design of this experimental robotic platform extends the previous results of scientific and technical teams that have

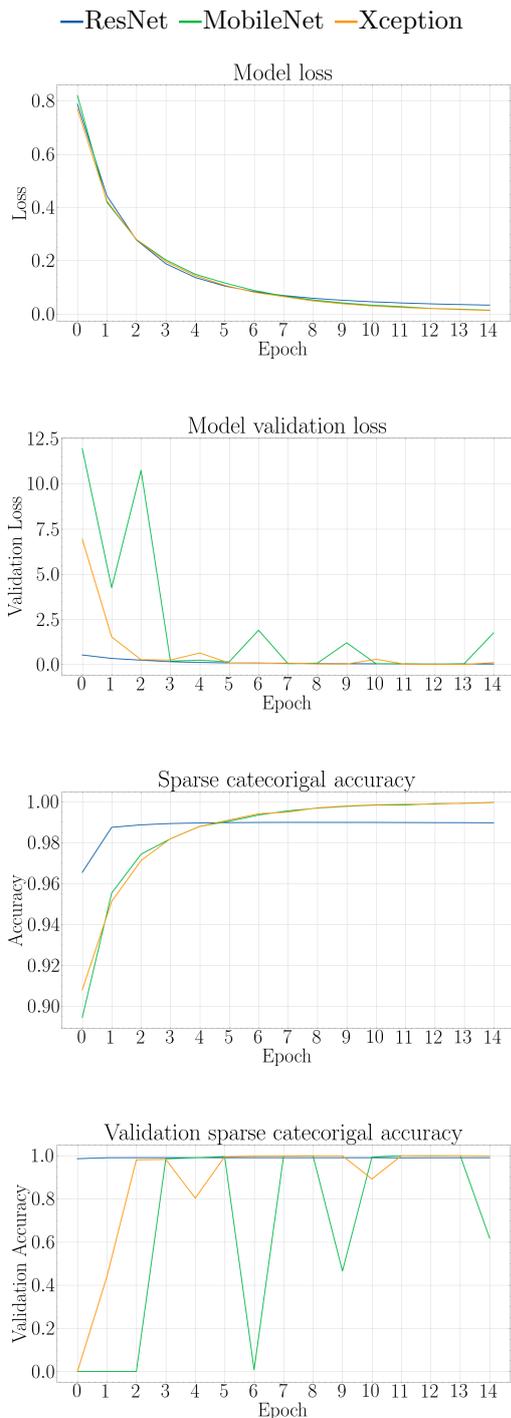


Figure 7: Training results of the ASPOCRNet for nostril detection.

been involved in the collection of samples for the purpose of diagnosing the COVID-19 disease.

The nostril detection function was created from several possible methods that deal with image segmentation, using an artificial neural network, more precisely a convolutional neural network. Two segmentation models, the older U-Net model (2015) and the newer ASPOCRNet model (2021), were compared in the solution. The result of this comparison is that to solve the

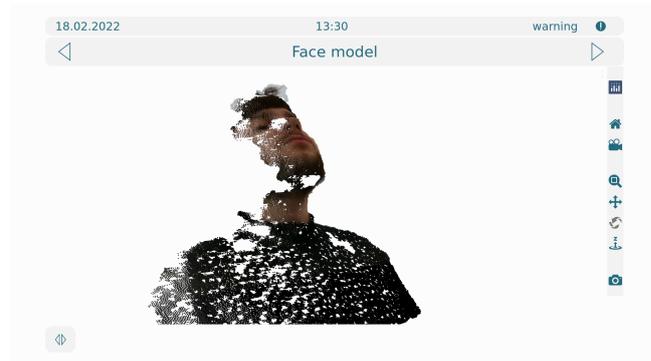


Figure 8: Reconstruction of a face model visualized in a Human-Machine Interface (HMI) as a point cloud.

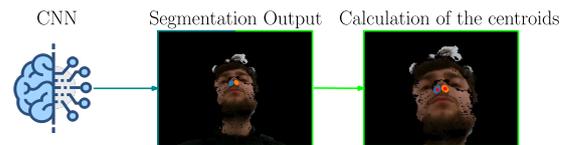


Figure 9: Post-process: Segmentation image - Centroid of nostrils image



(a) Patient identification using a QR code.



(b) The process of moving to the detected centre of the nostril.

Figure 10: Visualization of real-world testing of intelligent anterior human nose sampling using a collaborative robotic arm.

defined problem it is not necessary to apply the complex segmentation model represented by ASPOCRNet, which finds its use mainly for solving robotic segmentation scenes with a larger number of classes. Whereas U-Net has demonstrated good learning properties for this relatively simple segmentation task. An important difference compared to U-net is that the ASPOCRNet model offers the advantage of a significantly smaller size due to its architecture. In comparison, ASPOCRNet was half the size of U-net.

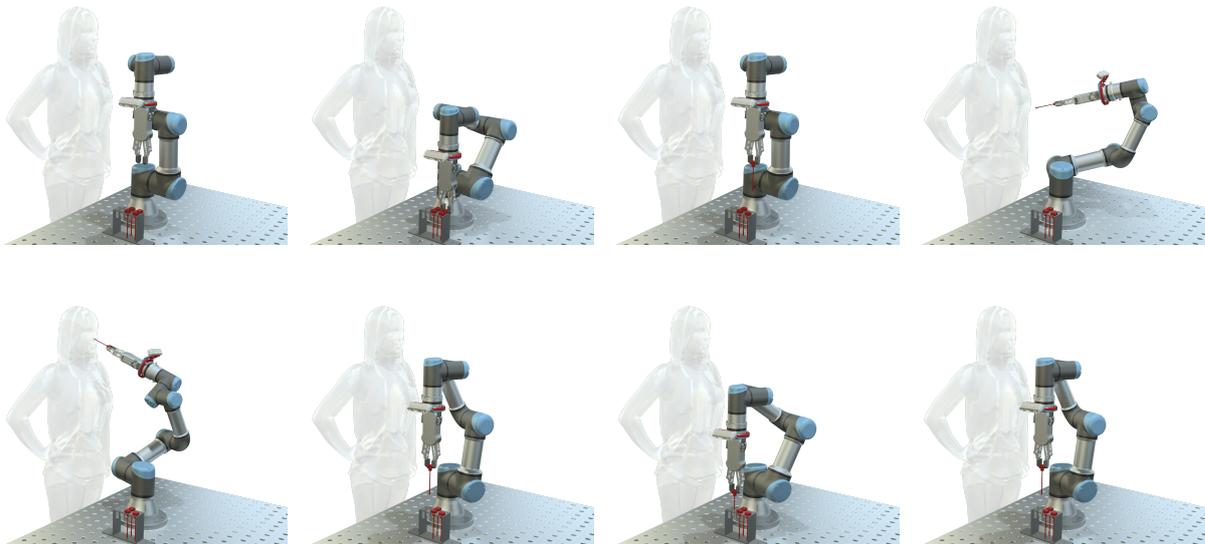


Figure 11: Visualisation of the nostril sampling process. A more detailed description of the process is described in Subsection 4.1 Setting up the robotic environment.

The result is a working prototype that emphasizes patient freedom and safety, but the human factor played a crucial role. Other factors that influenced the actual functionality are the calibration of the robot, the camera, and the hand-eye itself.

This work can provide a foundation for future research in medical robotics, as well as a suitable basis for intelligent guidance of robots in three-dimensional space.

Acknowledgement: This work was supported by Internal grant agency of BUT: FME-S-20-6538 “Industry 4.0 and AI methods”.

References

- [1] AGUERO, C., KOENIG, N., CHEN, I., BOYER, H., PETERS, S., HSU, J., GERKEY, B., PAEPCKE, S., RIVERO, J., MANZO, J., KROTKOV, E., AND PRATT, G. Inside the virtual robotics challenge: Simulating real-time robotic disaster response. *Automation Science and Engineering, IEEE Transactions on* 12, 2 (April 2015), 494–506.
- [2] ANDERSEN, T. T. Optimizing the universal robots ros driver. Tech. rep., Technical University of Denmark, Department of Electrical Engineering, 2015.
- [3] CHEN, L.-C., ZHU, Y., PAPANDREOU, G., SCHROFF, F., AND ADAM, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Computer Vision – ECCV 2018* (Cham, 2018), Springer International Publishing, pp. 833–851.
- [4] CROWE, S. Danish startup develops throat swabbing robot for covid-19 testing. Online, 2020.
- [5] GOODFELLOW, I., BENGIO, Y., AND COURVILLE, A. *Deep Learning*. MIT Press, 2016.
- [6] HU, Y., LI, J., CHEN, Y., WANG, Q., CHI, C., ZHANG, H., GAO, Q., LAN, Y., LI, Z., MU, Z., SUN, Z., AND KNOLL, A. Design and control of a highly redundant rigid-flexible coupling robot to assist the covid-19 oropharyngeal-swab sampling. *IEEE Robotics and Automation Letters* 7, 2 (2022), 1856–1863.
- [7] INTUITIVE. Da vinci surgical system: User manual, 2021. Online. Available at: <https://intuitivesurgical.com>.
- [8] KOENIG, N., AND HOWARD, A. Design and use paradigms for gazebo, an open-source multi-robot simulator. In *IEEE/RSJ International Conference on Intelligent Robots and Systems* (Sendai, Japan, Sep 2004), pp. 2149–2154.
- [9] KUKA. Medical robotics, lbr med, 2017. Online. Available at: <https://www.kuka.com>.
- [10] LI, C., GU, X., XIAO, X., LIM, C. M., DUAN, X., AND REN, H. A flexible transoral robot towards covid-19 swab sampling. *Frontiers in Robotics and AI* 8 (2021).
- [11] MCCULLOCH, W., AND PITTS, W. A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 5 (1943), 127–147.
- [12] NAVI, B. Brain navi develops new robot performing nasal swab tests to prevent cross infections. Online, 2020.
- [13] PARÁK, R., AND MATOUŠEK, R. Comparison of multiple reinforcement learning and deep reinforcement learning methods for the task aimed at achieving the goal. *MENDEL Journal* 27, 1 (June 2021), 1–8.

- [14] PARÁK, R., MATOUŠEK, R., AND LACKO, B. I4c – robotická buňka podle konceptu průmyslu 4.0. *AUTOMA 27*, 1 (January 2021), 10–12.
- [15] QUIGLEY, M., CONLEY, K., GERKEY, B., FAUST, J., FOOTE, T., LEIBS, J., WHEELER, R., AND NG, A. Ros: an open-source robot operating system. *ICRA Workshop on Open Source Software 3* (03 2009).
- [16] RONNEBERGER, O., FISCHER, P., AND BROX, T. U-net: Convolutional networks for biomedical image segmentation. *CoRR abs/1505.04597* (2015).
- [17] RŮŽIČKOVÁ, K. The robot from but will save the work of laboratory technicians. it could help with the covid-19 samples at the university hospital brno. Online, 2020.
- [18] SEO, J., SHIM, S., PARK, H., BAEK, J., CHO, J. H., AND KIM, N.-H. Development of robot-assisted untact swab sampling system for upper respiratory disease. *Applied Sciences 10*, 21 (2020).
- [19] David Coleman, Ioan A. Sucan, Sachin Chitta, Nikolaus Correll. Reducing the barrier to entry of complex robotic software: a moveit!case study. *Journal of Software Engineering for Robotics 5* (2014), 3–16.
- [20] Ioan A. Sucan and Sachin Chitta. Moveit. Online. Available at: moveit.ros.org.
- [21] Universal Robots. Ur3. Online. Available at: <https://www.universal-robots.com>.
- [22] WANG, S., WANG, K., TANG, R., QIAO, J., LIU, H., AND HOU, Z.-G. Design of a low-cost miniature robot to assist the COVID-19 nasopharyngeal swab sampling. *IEEE Transactions on Medical Robotics and Bionics 3*, 1 (feb 2021), 289–293.
- [23] YANG, L., CAO, Q., LIN, M., ZHANG, H., AND MA, Z. Robotic hand-eye calibration with depth camera: A sphere model approach. In *2018 4th International Conference on Control, Automation and Robotics (ICCAR)* (2018), pp. 104–110.
- [24] YUAN, Y., CHEN, X., AND WANG, J. Object-contextual representations for semantic segmentation. *CoRR abs/1909.11065* (2019).
- [25] ZHOU, Q.-Y., PARK, J., AND KOLTUN, V. Open3D: A modern library for 3D data processing. *arXiv:1801.09847* (2018).