DENTAL HEALTH MONITORING SYSTEM

By

Linjie Tong

Xin Wang

Yichen Shi

Zitai Kong

Final Report for ECE 445, Senior Design, Spring 2023

TA: Adeel Ahmed

23 May 2023

Project No. 37

Abstract

Point-of-care and remote diagnosis are hot points in recent years due to the COVID-19 epidemic. We designed a dental health monitoring system aiming to provide a deliverable, affordable, easy-to-use, and highly accurate method for remote dental diagnosis, which is necessary and promising. Our system consists of a mechanism, an APP, a segmentation software, and a 3D reconstruction software. The patient uses the hardware to record a video of his teeth scanning and send the video to the cloud via the smartphone, upload it to the cloud through an APP and some frames of the video will be selected to be segmented and 3D reconstructed. The dentist will use the video, the segmented images, and the 3D reconstructed model to make a diagnosis and send the result back to the patient through the app. This system has gained good performance and will be promising in the future. The mechanical system is highly adaptive, for adults and children with different mouth shapes, for phones with various dimensions, for misty gases coming out from the user's teeth when he coughs, and allows the shipping to be compact, which was achieved by careful design of the geometry. In the segmentation subsystem, a sample block is developed to sample video into images with different views of patients' teeth. A segmentation algorithm is developed to segment teeth in the palate and teeth in the mandible and it can outperform the benchmark. For the 3D reconstruction model, by introducing a segmentation module and larger RNN cores, and expanding the input capacity, the 3D reconstruction model can predict accurate voxel structure describing the surface of the teeth and the shape of the dental arch curve compared to the baseline.

Contents

1. Introduction 1
1.1 Background and Problem 1
1.2 Solution 1
1.3 Functionality 1
1.4 Subsystem Overview and Design Changes 2
1.4.1 Block Diagram 2
1.4.2 Description and Changes 2
2 Design 3
2.1 Design Procedure 3
2.1.1 Mechanical Subsystem 3
2.1.2 Interactive Application 4
2.1.3 Segmentation Subsystem 4
2.1.4 3D Reconstruction Subsystem 5
2.1.4.1 Alternatives and Issues 5
2.1.4.2 Equations 5
2.2 Design Details 6
2.2.1 Mechanical Subsystem 6
2.2.2 Interactive Application 9
2.2.2.1 Register and Login Subcomponent 9
2.2.2.2 Upload and Download Subcomponent (Patient) 9
2.2.2.3 Upload and Download Subcomponent (Dentist) 10
2.2.3 Segmentation Subsystem 11
2.2.3.1 Sample Block 11
2.2.2.1 Segmentation Block 11
2.2.4 3D Reconstruction Subsystem 12
2.2.4.1 Diagram Overview 12
2.2.4.2 Image Encoder 12
2.2.4.3 Combiner 13

2.2.4.4 Fine Volume Generator 13
2.2.4.5 Model Training Details 13
2.2.4.6 Visualized Model Results 14
3. Design Verification 14
3.1 Mechanical Subsystem 14
3.2 Interactive Application 19
3.3 Segmentation Subsystem 19
3.4 3D Reconstruction Subsystem 19
4. Cost and Schedule 20
4.1 Parts 20
4.2 Labor 20
4.3 Schedule 21
5. Conclusion 22
5.1 Accomplishments 22
5.2 Uncertainties 22
5.3 Ethical considerations 22
5.4 Future work 22
References 23
Appendix A Requirement and Verification Table 24

1. Introduction

1.1 Background and Problem

In the past era of COVID-19, point-of-care has become a hot topic in society and the medical field due to the high risk of infection when people go out to seek medical treatment. Nowadays, though the world is in a post-epidemic era, point-of-care remains a highly desirable and widely marketable field. One reason is that there is a common reluctance to seek medical treatment; another is that the cost of seeking medical treatment is high in both time and price, especially for those living in remote areas. Until now, point-of-care has developed in lots of medical application scenarios, like antigen detection, blood glucose detection, and so on.

With the fast development in people's daily life, nowadays, when people's living standards are rising rapidly, high-sugar, high-calorie diets are causing more and more oral problems. People are paying increasing attention to oral health and have more requirements for it. However, point-of-care for the diagnosis of oral diseases (e.g., caries analysis, orthodontics, oral health) is not yet well equipped. In the dental diagnosis process, dentists use professional dental tools, X-rays and CT scan machines to get geometric details of teeth, but these are difficult to obtain from non-medical facilities because professional equipment is rarely prepared in one's home.

In this case, developing a dental health monitoring system for remote dental diagnosis is necessary and promising.

1.2 Solution

Our team planned to design an oral point-of-care system for dental health monitoring to take good dental pictures in their homes with their phones, and send them to remote doctors for diagnosis, which will exempt the cost of offline medical treatment.

The solution consists of four major subsystems, the hardware, an interactive app, segmentation software, and 3D reconstruction software. The patient uses the hardware to record a video of his teeth scanning and send the video to the cloud via the smartphone, upload it to the cloud through an APP and some frames of the video will be selected to be segmented and 3D reconstructed. The dentist will use the video, the segmented images, and the 3D reconstructed model to make a diagnosis and send the result back to the patient through the app.

1.3 Functionality

The most direct part that interacts with the user is the non-deformable mechanical equipment. The mechanism is able to hold the smartphone in the correct position and open the patient's mouth to guarantee more parts of it can be shown. The user manually changes the phone camera orientation to record a video, which finishes the process of sample scanning. It will contain springs to reduce vibration and maintain stability. The friction between parts in the assembly should be reasonable so that the machine can be operated by hand but not too sensitive to force input.

For the segmentation, in the video sampling part, sampling on videos is in a certain number of frames and the number cooperates with predefined restrictions on how to record the video. In the segmentation part, an encoder-latent layer-decoder structure is applied to construct the segmentation algorithm. The encoder part receives the raw videos uploaded by the mechanical sample system, samples them into images, and encodes the images to latent features. The decoder part receives the features and uses them to do the pixel-wise classification and combine them with the original image to feed to the 3D reconstruction part. A novel loss function that can help the model learn the shape of the dental arch curve is designed and it improves the model's performance.

The 3D reconstruction subsystem takes teeth images, combines the information, and projects the images to a 96×96×96-voxel 3D model by using a neural network trained before, which can show the surface shape of the teeth. The result can provide a general picture of the teeth's relevant positions, shapes, and arch shape, which cannot be directly viewed through the 2D pictures. can help the dentist to make the diagnosis together with the video and segmented images.

For the APP, this subsystem can be divided into three subcomponents, which act together to make the whole subsystem work smoothly. For the “register and login subcomponent”, the patients' account can be registered at any time if the account id is unique. The dentist account can only be created by the administrator. It is related to the cloud database. The “upload and download subcomponent for the patient” is used for the upload and download process of the patients. With the help of this subcomponent, the patients can upload their videos and download the feedback by interacting with the cloud server. The “upload and download subcomponent for the dentist” subcomponent is used for the upload and download process of the dentist. With the help of this subcomponent, dentists can upload their feedback and download the video by interacting with the cloud server.

1.4 Subsystem Overview and Design Changes

1.4.1 Block Diagram

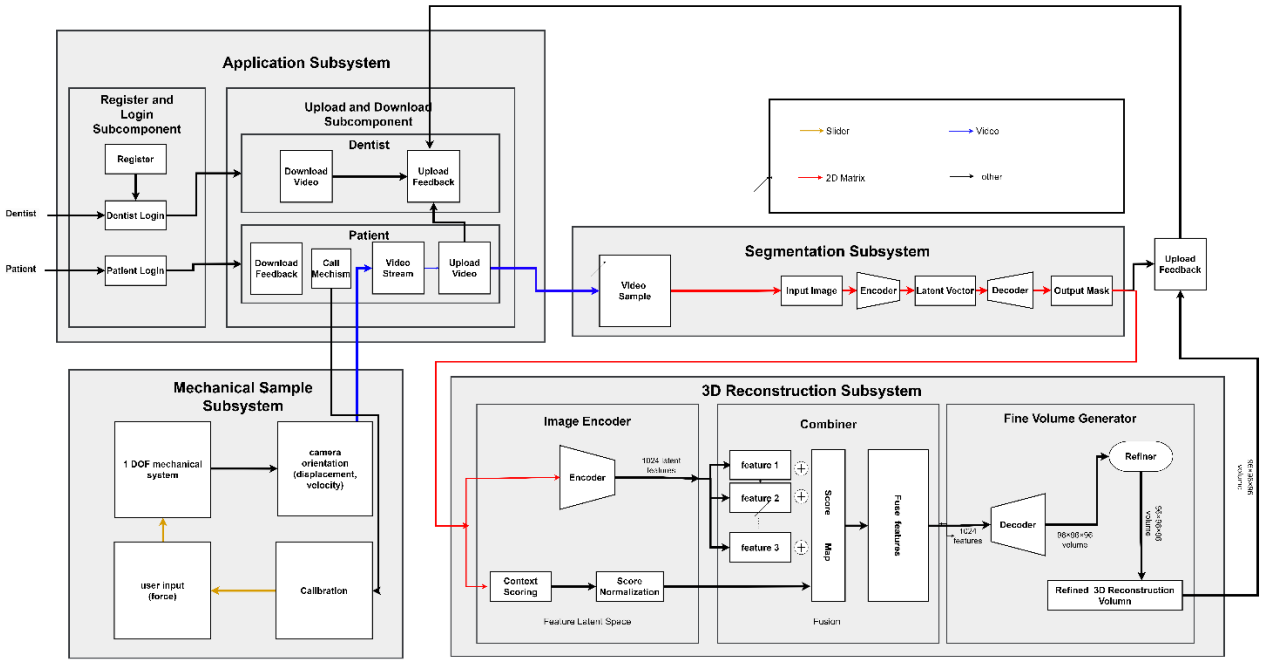


Figure 1 Block Diagram Overview.

1.4.2 Description and Changes

The hardware system is originally comprised of two blocks, the mechanical subsystem, and the sampling subsystem. The mechanical system holds the phone and gives the phone camera the right orientation to take pictures of a human mouth. The sampling subsystem is responsible to take picture sequences as the input of software blocks. The sampling subsystem is not robust because as the user uses the mechanical systems, he cannot see the view of the camera. As a result, the user is not conscious of the quality of the picture token. To improve the robustness of the sampling, we removed this block from the hardware block and include it in the software subsystems. As sampling by software is more stable and easier than by approaches using hardware.

For the segmentation subsystem, the video sampling part has been moved from the mechanical subsystem to the segmentation subsystem. The segmentation subsystem does 3-class classification. It will segment teeth in the palate, teeth in the mandible, and background. The latest segmentation subsystem consists of two parts: the video sampling part and the segmentation part. The video sampling part consists of a video sample block and the segmentation part consists of an encoder block and a decoder block.

The 3D reconstruction subsystem consists of three blocks. The coarse volume generator block will map the images to latent features, the combiner will combine the features to one that can represent the 3D model, and the fine volume generator will decode the feature to a 3D model and apply a refining process. The subsystem receives the segmented images from the segmentation subsystem and gives the 3D model to the doctor. The doctor will make the diagnosis and send the result through the app. Initially, the decoder block is put in the image encoder. In later design, I move it to the fine volume generator. The combiner block is separated from the fine volume generator.

For the APP subsystem, it is extra to make the project widely used by patients. The application has a single administrator with the highest privilege who can modify all data and files in the cloud database and server. Patients can register and log in using their own accounts, while dentists have special accounts created by the administrator for login. To maintain quality, the number of dentists is limited, and interested individuals can contact the administrator through the "about us" page. The subsystem comprises pages, a cloud database with user, dentist, request, and feedback tables, and a cloud server for storing uploaded videos, PDF files, and image resources. While storing images in the cloud server helps reduce the application package size, it may lead to slightly longer load times for background images.

2 Design

2.1 Design Procedure

2.1.1 Mechanical Subsystem

Stage 1, preliminary collection of information.

Make design objectives and come up with basic descriptions of parts needed to design. Make material list of all parts. Learning the geometry of human face and mouth.

Stage 2, first generation CAD model design, numerical simulations, and physical part verification.

Design the CAD models for parts. Make static force finite element analysis to test the strength of the parts and the whole assembly. Do multiphase flow simulation to test the distribution of water when people cough. Print all the main parts, observe to test the surface finish, the strength and how they fit each other and the performance of the parts which is directly interacting with the human face.

Stage 3, share information with other team members, make design changes to fulfill their specific requirements.

As the segmentation subsystem cannot distinguish the white color and other light colors, some specific parts need to be designed to have dark color and reduce the exposure of nature sunlight as much as possible.

Stage 4, make plans for parts that will not be 3D printed.

For parts that will be made by 304 stainless steel and HDPE and LDPE, I contacted the manufacturer to get their manufacture limitations, made changes in the design to fulfill the limitations and purchased the materials.

Stage 5, design for the methods to join all the parts.

To achieve the goal of not using nuts in the design, all the connections between parts are either 1) complementary; 2) buckle; 3) fixed by hot glue or a combination thereof. Because of the complexity of the substructure in the design is increased much, the requirement of accuracy of the 3D printing is raised. I tested the quality of the parts under different printing parameters and found the best printing parameter which will give me the best material strength and surface finish.

Stage 6, assemble the system and test with the other software subsystems.

Assemble the system and get a working prototype. Test the performance by placing a real phone in the mechanical system and take videos. And listen to the feedback from my teammates as they process the video in their algorithms.

Stage 7, strengthen part strength and replace parts by optimized new parts.

For some parts, because of the onset of large deformation in the testing, I changed the design to make the parts more rigid. Also, the mouth opener was re-designed to give a better view of teeth in the camera.

Stage 8, final change on the robustness of the whole system and user experience.

Considering the possible propagation of cracks when the diffuser collides with the edges of the calibrator, two structures of elastic materials is designed to reduce the impulse to the materials.

2.1.2 Interactive Application

Some problems that have ever occurred during the development process are listed. For how to exchange data from different pages, exchanging data between different pages is necessary since one needs to pass the user's login information from the login page to the user's home page after the user login successfully and has been redirected to his own account page. But both the data in the "data structure" on one page and variables "const, var, and let" can only be used locally. There are two ways to solve this problem. The first method is to use the wx.setStorageSync function to store the data as the (key, value) pair and use the wx.getStorageSync function to get the stored data by using the key you just set. The second method is to use the global variable declared in the app.js file to pass the information across pages.

For the data delay caused by the Sync or Async problems, sometimes the data get is not the data that are just input but the last one. This is caused by the Async programming. Async is sometimes more efficient than Sync but can cause this kind of problem if one does not design the algorithm and code structure well. For the application, the logic is comparably simple, in which case the efficiency is not the main challenge, the Sync programming is decided to be used later. Whether the function is Sync or Async is needed to be checked before one uses it.

The scope of the keyword "this", is the current function. If one wants to use the "this" keyword of the caller function, he needs to store it in a variable and pass it to the current function before he calls it.

2.1.3 Segmentation Subsystem

For the video sampling block, the video recorded by the mechanical subsystem is sampled into 24 images. In the video, patients record twelve different views about their teeth. And the more images we sample, the more likely we can get high-quality images about different views. However, with more images, the segmentation model needs more time to segment, and GPUs are also not large enough to reconstruct these images. Finally, sampling video into 24 images is chosen. Sampling video into 24 images means for each view, there are two images sampled which can increase the probability to get high-quality images from different views. And NVIDIA RTX 3090 GPU is capable of reconstructing 24 images into a 3D model. What's more, the segmentation model can segment these images within two minutes.

Table 1: Comparison of different segmentation methods

Methods	IoU
Unet	84.04
Unet++	84.25
Medical Transformer	79.31
Semask	84.31

For the segmentation block, a model that can segment teeth from the background is trained. After reading some articles about semantic segmentation, we compare four different methods. Two of them, U-Net [11] and U-Net++ [13] are U-net based. two of them, MedT[8], and Semask Transformer [7] are transformer-based methods. To guarantee the fairness of guaranteed experiments, all five methods are implemented by their original source code. Training images are resized to (512, 512) for Unet, Unet++, MedT. Training images are cropped into (512,512) for Semask due to implementation details about its original source code [7]. All the training hyperparameters are default values. At this time, the data for our project is not pre-proposed. To reproduce these codes and test their performance, Cone beam computed tomography images are used to train data. And Table 1 demonstrates the results. In the project, Intersection over Union (IoU) is used as the metric to measure the performance of the segmentation model.

$$IoU(A,B) = \frac{|A \cap B|}{|A \cup B|} \tag{1}$$

And the result shows that Semask performs the best in dental image segmentation, thus semantic prior proposed by Semask is used in our segmentation algorithm. Besides, to improve the performance of the 3D construction subsystem, the segmentation model is trained to do 3-class classification. Besides, for loss function, a novel loss function called dental arch curve shape loss is designed to help the model learn the shape of the dental arch curve and it improves the model's performance in our task.

2.1.4 3D Reconstruction Subsystem

2.1.4.1 Alternatives and Issues

There are two aspects of alternatives to be considered. Firstly, there are voxel-based methods like 3D-R2N2 [3] and point cloud-based methods like GraphX [2], PSG [5], RealPoint3D [12], and NeuralRecon [9]. Point cloud methods have the common problem of memory consumption, which can easily cause the problem of bugs, system compatibility, or insufficient memory on the server. Point clouds maybe with a smoother surface. But voxels can also be good if there are enough voxels. In this case, 3D-R2N2 is used as the baseline of the subsystem, and it passed tests successfully.

Secondly, the initial version is to generate the upper and lower teeth using two sets of models since the given datasets separate the upper teeth and lower teeth into two ground truths, and clinically dentists use either upper teeth or lower teeth individually. The current model combines two 3D reconstruction models and runs the upper and lower teeth at once. It can gain the same accuracy and it uses less memory and less time.

2.1.4.2 Equations

Loss function and IoU are two metrics used to measure the performance of the subsystem. [3]

The Loss function is a useful method to evaluate the performance of a machine learning model by measuring the consistency between the training and validation set with their ground truth. It is defined as equation (3) in summed voxel-wise cross-entropy form. Let the final output at each voxel (i, j, k) be Bernoulli distribution [1 − p(i,j,k), p(i,j,k)], where the dependency on input X = {xt} t∈{1,...,T} is omitted, and let the corresponding ground truth occupancy be y(i,j,k) ∈ {0, 1}, then

$$L(X,y) = \sum_{i,j,k} y_{(i,j,k)}\, log(p_{(i,j,k)}) + \left(1 - y_{(i,j,k)}\right)\left(1 - lo\, g\left(p_{(i,j,k)}\right)\right) \tag{2}$$

The IoU is another machine learning method to measure whether the prediction (test set) is similar to its ground truth. It is defined as equation (3) with the same meaning for y and p with equation (2).

$$\mathrm{IoU} = \sum_{i,j,k}\left[\mathrm{I}(p_{(i,j,k)} > t)\mathrm{I}(y_{(i,j,k)})\right] / \sum_{i,j,k}\left[\mathrm{I}\left(\mathrm{I}(p_{(i,j,k)} > t) + \mathrm{I}(y_{(i,j,k)})\right)\right] \tag{3}$$

2.2 Design Details

2.2.1 Mechanical Subsystem

The explanation of fine structures in section view is shown in Figure 2.

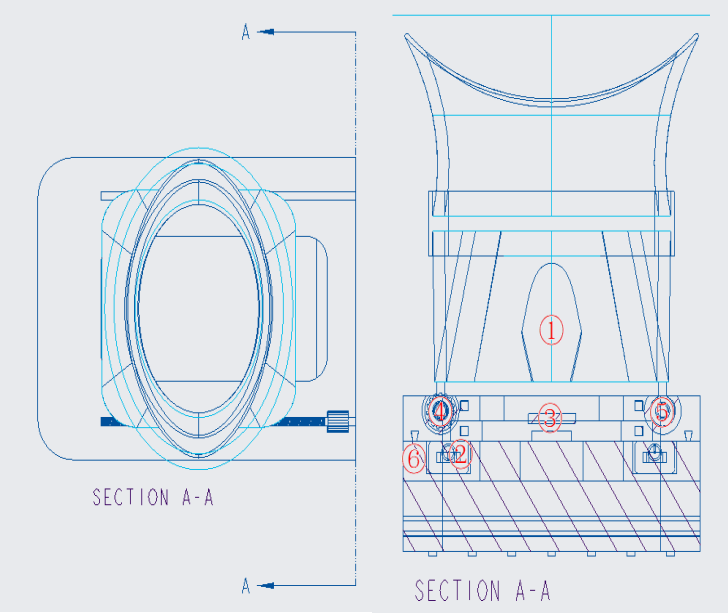


Figure 2. Section view of the mechanical system

For the modelling of airstream with water from the human mouth, the criteria of modelling the flow as incompressible and low-Mach number flow are the Mach number. Mach number is a nondimensional parameter of the ratio of the airspeed to the speed of sound.

$$Ma = \frac{U}{a} = \frac{U}{\sqrt{\gamma RT}} = \frac{50}{\sqrt{1.4 \times 287 \times 300}} = 0.14 < 0.3 \tag{4}$$

Where U is the airspeed, $\gamma = c_p / c_v = 1.4$ is the adiabatic coefficient, $R = \bar{R} / M = 8.314 / 0.029$ is the ideal gas constant, T is the temperature.

Given the simulation results in the following sections, it explained the construction of cavity ①.

For the restoration of the placement shown in Figure 3, one alternative is to use the piston.

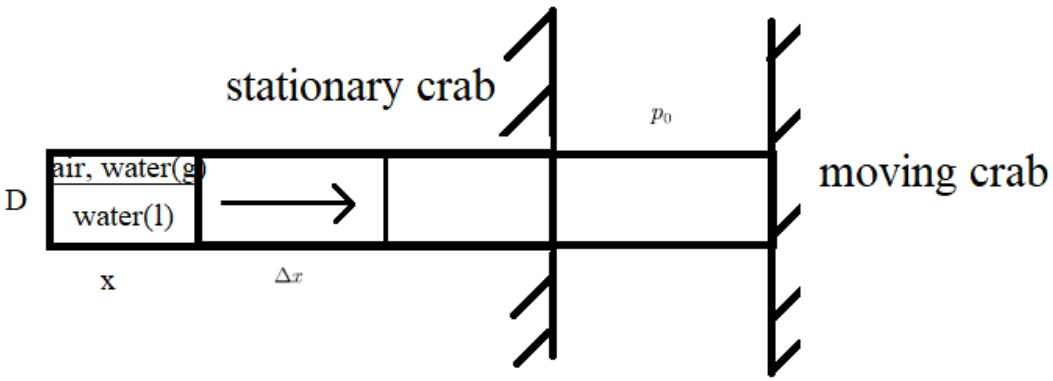


Figure 3 Retraction placement design as a piston.

If the portion of water is 0, the inner pressure is,

$$p/p_0 = \rho/\rho_0 = x/(x + \Delta x) \tag{5}$$

Given that $D = 3mm, p_0 = 101kPa, x = 2mm, \Delta x = 30cm$, The total force is,

$$F = (p_0 - p)\frac{\pi}{4}D^2 = \frac{\pi}{4}\frac{\Delta x}{x+\Delta x}p_0 D^2$$

D=3mm, p_0=101kPa, x=2mm,Δx=30cm

Given that

$D = 3mm, p_0 = 101kPa, x = 2mm, \Delta x = 30cm$,

$$F = (p_0 - p)\frac{\pi}{4}D^2 = \frac{\pi}{4}\frac{\Delta x}{x+\Delta x}p_0 D^2 = \frac{\pi}{4} \times \frac{30}{2+30} \times 101 \times 10^3 \times (3 \times 10^{-3})^2 = 0.669N \tag{7}$$

The mass of phone is around $m = 200g$, the minimum friction factor should be:

$$\mu = \frac{mg}{2F} = \frac{0.2 \times 9.8}{2 \times 0.669} = 1.465 \tag{8}$$

This friction factor is very hard to achieve by the materials I have, even though when I increase the roughness of the surface.

The other alternative is to use the spring.

The deflection of the spring is,

$$\delta = \frac{\partial E}{\partial F} = \frac{8FD^3N}{d^4G} \tag{9}$$

The elastic constant is,

$$k = F/\delta = \frac{d^4G}{8D^3N} \tag{10}$$

Where G is the shear modulus, for 304 steels, $G = 77GPa$

Consider an obtainable spring with dimension $d = 0.4mm, D = 3mm, l = 10mm$

From grounded ends,

$$N_t = L_s/d = 10/0.4 = 25 \tag{11}$$

$$k = \frac{d^4 G}{8D^3 N} = \frac{(4 \times 10^{-4})^4 \times (77 \times 10^9)}{8 \times (3 \times 10^{-3})^3 \times 25} = 365.037 N/m \tag{12}$$

If I install the 4 springs on the system in parallel and extend the spring to 10mm, the force will be

$$F = 4kl = 4 \times 365 \times 0.01 = 14.6N \tag{13}$$

The friction coefficient needed will be,

$$\mu = \frac{mg}{2F} = \frac{0.2 \times 9.8}{2 \times 14.6} = 0.067 \tag{14}$$

For the friction coefficient between PLA and PLA is greater than 0.3824[10], which means this set up will be able to hold the phone. This explained the construction of springs at ②.

To install the threads, we need to break the calibrator into two parts so that we can install the threads in the middle. Therefore, the design of complementary parts ③ is explained. On the one part, the shape is convex and on the other part it is concave, and when glues are put in the clearance of the two parts, it has been physically found that the parts can connect to each other firmly.

To adjust the location of the diffuser finely, the threaded components are needed. ④ are threaded components. On the diffuser there are inner threads and on the adjustment rotator there are outer threads. The sizes of the threads are M4, which is appropriate because it is large enough so that users do not need to rotate too many rounds to adjust the orientation of the diffuser while keeping the material strength of the diffuser as it penetrates the round parts of the diffuser.

⑤ is a locating hole. In the intermediate design, it was designed to place the 304 steel rods to guide the diffuser. However, as the final design changed the connection of diffuser from revolve to threaded, I did not remove it because it can guide the user to install the diffuser vertically and correctly, enabling the user to assemble this part by himself if he wants.

⑥ is the design of Mortise and tenon structure along with buckles. When installing the crab parts into the calibrator, this structure ensures the connection to be reliable. Also, the length of the cavities was specifically designed so that when users push the crab parts along the cavity to its end, the outer shapes will automatically be aligned. This explains the term buckle as the combined structure will be locked when the user pushes the crab parts to the end. Also, this kind of structure allows the installation to be fully reversible. If nuts are used, every time the parts are being assembled, some parts of the plastic will be peeled off. Also, when installing the nuts, people often give excessive forces to the nuts to make the connection firm. However, this kind of excessive force will result in the propagation of internal cracks of the plastic threads, where the nuts are being installed on. This will significantly increase the possibility of breaking the parts.

2.2.2 Interactive Application

2.2.2.1 Register and Login Subcomponent

The general workflow of this subcomponent can be completely shown through the figure below:

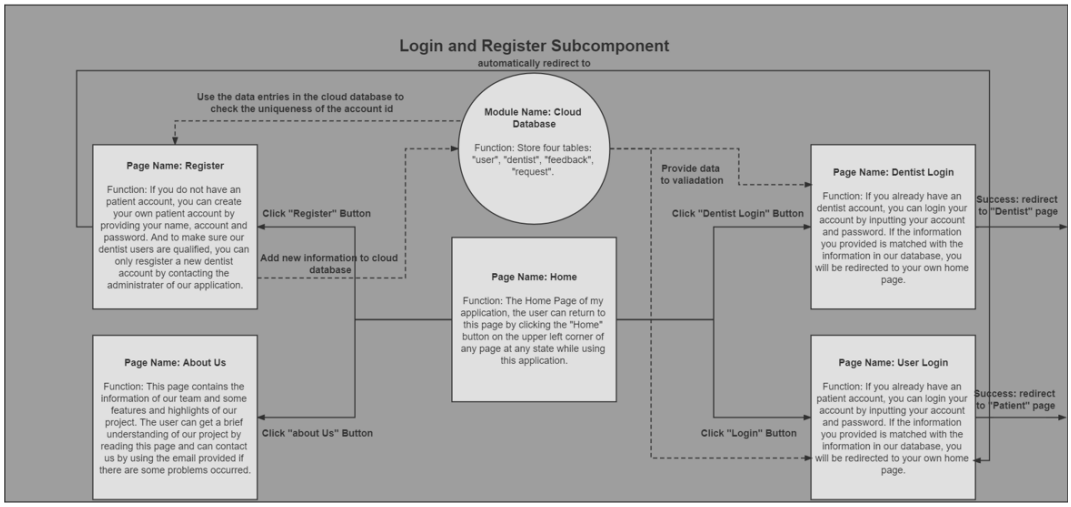


Figure 4 The flowchart of the login and register subcomponent.

Here are some detailed explanations of the involved pages and modules:

For the “About Us” page, the user can find the basic information of our team and the email of our team members in this page. The patient can use this information to contact the administrator and if anyone wants to be the dentist, he can contact us, and the administrator can add his information in the “dentist” table in the cloud database after checking his qualification. The user can also find some interesting highlights of our project. This will give the user a brief understanding of our dental health monitoring system. For the “Register” page, anyone can register a patient account by using the function of this page. The user needs to input the account id, password, and the name as the basic information of the account and the lengths of these information are constrained. If the account id is unique in our cloud database, the new account will be created successfully. Otherwise, the application will show prompt to indicate that the user needs to change the account id to satisfy our requirement. For the “Home” page, it is the home page of our application. The user can be redirected to this page at any time if he clicks the top left button. For the “User Login” page, the patient needs to login to their own individual home page by inputting correct account id and password on the “User Login” page. Like the “Register” page, the application will check the cloud database to determine whether the input information is qualified. For the “Dentist Login” Page, the dentist needs to login to their own individual home page by inputting correct account id and password on the “Dentist Login” page. Like the “Register” page, the application will check the cloud database to determine whether the input information is qualified. The cloud database contains four different tables which will be used when we need to match the data or fetch the data we stored previously.

2.2.2.2 Upload and Download Subcomponent (Patient)

The general workflow of this subcomponent can be completely shown through the figure below:

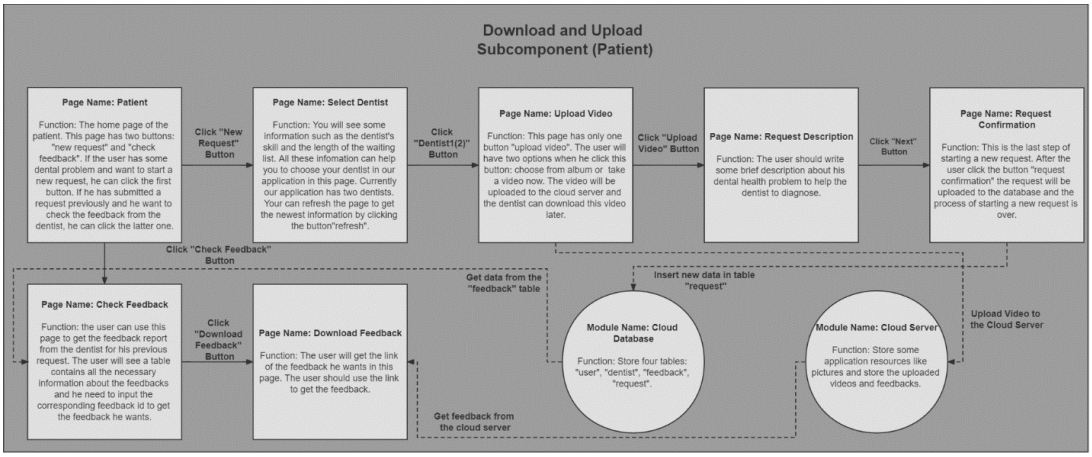


Figure 5 The flow chart of the download and upload subcomponent for patients.

Here are some detailed explanations of the involved pages and modules:

For the “Patient” page, it is the home page of the patient. The patient can return to this page while creating a new request or checking his feedback. For the “Select Dentist” page, when the patient wants to start a new request, he firstly needs to choose one dentist he prefers based on the information we provided on this page (the length of the waiting list and the research area of the dentist). For the “Upload Video” page, the patient needs to upload the video to the cloud server which will be used by the dentist to diagnose. The video can either be recorded in the application or be chosen from the album. For the “Request Description” page, the patient needs to write a brief explanation which needs to be less than one hundred words to help the dentist understand your request quickly. For the “Request Confirmation” page, this is the final request confirmation. The patient cannot cancel the request after this step. For the “Check Feedback” page, the first step to get feedback from the dentist. The patient needs to input the id of the feedback and he will navigate to the “Download Feedback” page. For the “Download Feedback” Page, the patient should use the link provided to get the pdf file of the feedback from the dentist. The cloud database contains four different tables which will be used when we need to match the data or fetch the data we stored previously. The cloud server stores uploaded videos, feedback, and the application resources like background and swiper images.

2.2.2.3 Upload and Download Subcomponent (Dentist)

The general workflow of this subcomponent can be completely shown through the figure below:

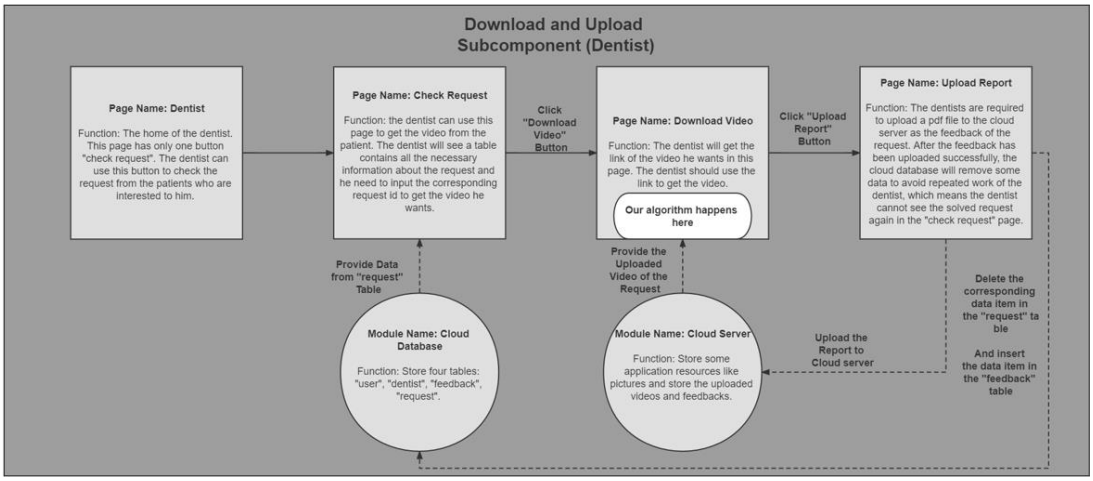


Figure 6 The flowchart of the download and upload subcomponent for dentists.

Here are some detailed explanations of the involved pages and modules:

For the “Dentist” page, it is the home page of the dentist. The dentist can return to this page at any time. For the “Upload Report” Page, the dentist needs to use this page to upload the feedback to the cloud server. For the “Check Request” page, the dentist needs to input the id of the request in this page, and he will be navigated to the “Download Video” page to get the video. For the “Download Video” page, the dentist should use the link provided to get video corresponding to the request from the patient. The cloud database contains four different tables which will be used when we need to match the data or fetch the data we stored previously. The cloud server stores uploaded videos, feedback, and the application resources like background and swiper images.

2.2.3 Segmentation Subsystem

2.2.3.1 Sample block

In the process of recording video, patients are required to record the left, middle, and right side of teeth with mouth closing, each for five seconds, and after that, patients are required to record up left, up middle, upright, left, middle, right, down left, down the middle, down right side of teeth with mouth opening. Sampling begins at a quarter of the sampling rate to avoid noisy images at the beginning and end of the video. And video sample part is developed with Python's mmcv package. It provides powerful tools to load videos and get the frame number of videos, which facilitates video processing.

2.2.3.2 Segmentation block

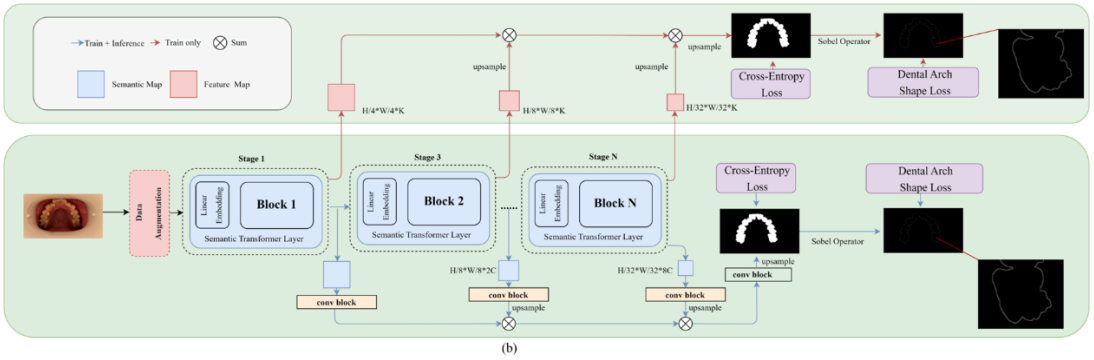


Figure 7. Pipeline of segmentation algorithm. It applies an encode-decoder structure, and it uses semantic prior and a novel dental arch loss to do segmentation.

The segmentation algorithm applies an encoder-decoder structure, and it uses semantic prior to segmenting small objects from the boundary of large objects. And a novel loss function is designed to help the algorithm learn the dental arch curve. And its details are as follows and Figure 6 is the pipeline of the segmentation algorithm.

After getting the prediction from the decoder, softargmax is applied to turn the prediction into one-hot form.

$$\operatorname{argmax}_{soft}(x, \beta) = \Sigma_c \frac{e^{x_c}}{\Sigma_{i,j} e^{x_{i,j}}} \cdot c \tag{15}$$

Then, the Sobel operator is applied to both the label and one-hot form of prediction, getting the edge of them.

To get the shape of the dental arch curve, we propose a Sobel-based Boundary detection Algorithm to extract the shape of the dentsl arch curve. The Sobel operator is used to extract the boundaries of the image in two dimensions by squaring and adding. The algorithm is as follow:

$$shape(img) = 255 \cdot \frac{sobel_x(img)^2 + sobel_y(img)^2}{max\{sobel_x(img)^2 + sobel_y(img)^2\}} \tag{16}$$

The novel dental arch curve loss uses a Sobel-based Boundary detection Algorithm to extract the shape of the dental arch curve in ground truth and prediction and compare them.

$$loss = \frac{\big(shape(GT) - shape(pred)\big)^2}{N_{pixel}} \tag{17}$$

N_{pixel} donates the number of pixels in that image.

This loss corporates with Cross-entropy loss function. Datasets are made for training of segmentation algorithm and 3D reconstruction algorithm. Labelme is used to label the raw data and a code is written to turn the output .json file from labelme into trainable datasets for the segmentation algorithm and 3D reconstruction algorithm.

2.2.4 3D Reconstruction Subsystem

2.2.4.1 Diagram Overview

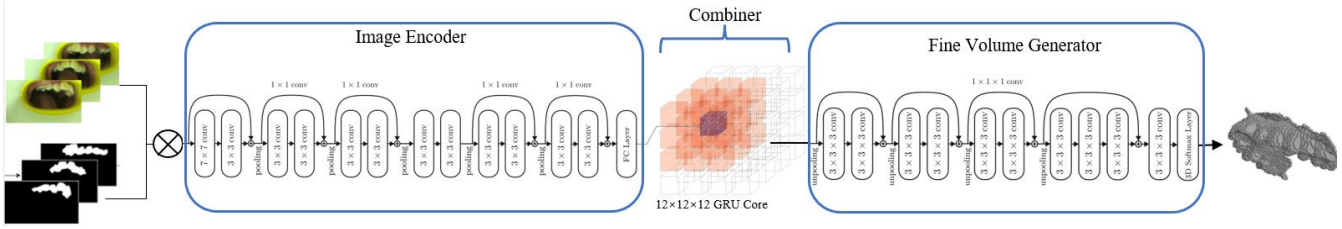


Figure 8 Overview diagram of 3D reconstruction subsystem

2.2.4.2 Image Encoder

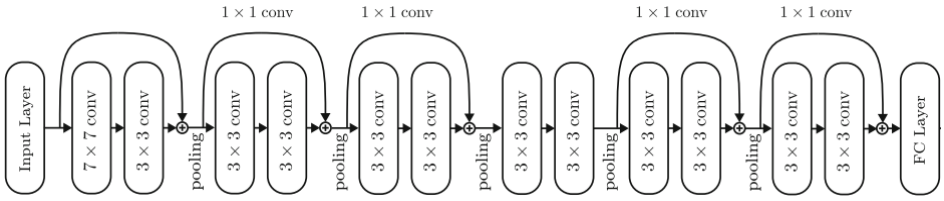


Figure 9 Block diagram of Image Encoder [3]

According to 3D-R2N2, the encoder consists of six sets of convolution layers (five of them are residual net), and each of the set consists of two convolution layers, shown in Figure 9. It can be seen as an encoder that takes 201×201 pixels RGB images, crops it by 1 in x and y dimensions, converts to grayscale images and generates the feature in 1024 format features by a full connection layer.

2.2.4.3 Combiner

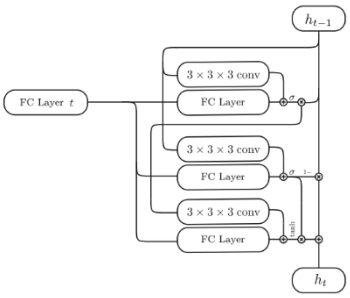


Figure 10 The structure of GRU units [3]

The combiner is the core structure to handle the multi-view images. The Recurrence structure GRU [4] is a classic model that can handle and combine a sequence of inputs. According to the 3D-R2N2 method [3], the network is made up of 12×12×12 GRU units with restricted connections to its spatial neighbors within the kernel, shown in the middle part of Figure 8. Each unit learns to reconstruct one part of the voxel space. This structure will combine the input images from all views of the teeth and generate the fine volume represented by a 1024 hidden state.

2.2.4.4 Fine Volume Generator

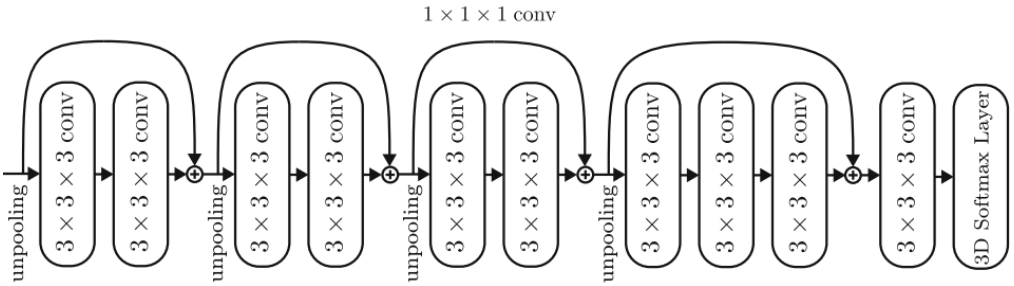


Figure 11 Block diagram of the fine volume generator [3]

Shown in Figure 11, the fine volume generator consists of four sets of residual nets. It takes the hidden states and transforms them into a final voxel occupancy map. This process also gives a refined step, that is, it converts the final activation $V \in R^{Nvox \times Nvox \times Nvox \times 2}$ to the occupancy probability p(i,j,k) of the voxel cell at (i, j, k) using voxel-wise softmax by the 3D softmax layer.

2.2.4.5 Model Training Details

For data augmentation, random crop, x padding by 1 pixel, y padding by 1 pixel, and random flip is applied.

For the training process, 7 patients' data from the datasets are used for training, 1 patient's data is for validation and the other one is for testing. The number of data workers is 5. If the user uses a blender, kill a worker after 100 iterations to clear the cache. The maximum number of training iterations is 80000, the number of renderings is 24, the maximum number of mini batches that can be put in a data queue is 15, the optimizer is Adam, with momentum 0.9, and makes a validation every 2000 iteration. The learning rate policy is starting with 0.0001 and at the 20000th iteration it will change to 0.00001 and at the 60000th iteration, it will change to 0.000001. The weight decay rate is set to 0.00005 and the model will be saved every 10000 iterations.

For testing and verification, the loss function and the IoU with threshold 0.4 are applied. Details are shown in the verification section.

2.2.4.6 Visualized Model Results

The visualization of 3D reconstruction model is in Figure 12. Compared to the baseline, our model opposes a better ability to predict the blur or invisible tissues of the teeth, which causes fewer blank parts. With raw data without segmentation, the noise gets worse, and the prediction on the inside of the teeth may go blank. This is due to the interference of the other half part of the teeth, tongue, and mouth wall that are not involved in the reconstruction. When reconstructing based on segmented images, the interference is got rid of and the surface of 3D reconstruction will be smoother.

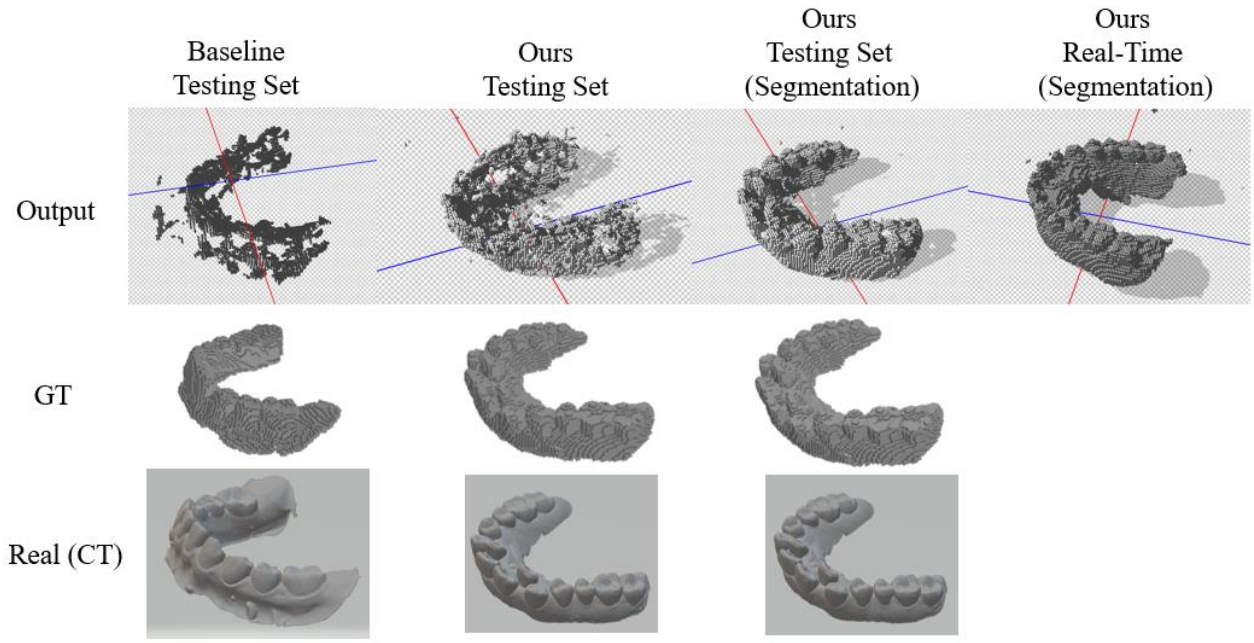


Figure 12 Visualized 3D Reconstruction Results

3. Design Verification

3.1 Mechanical Subsystem

1. stress strain analysis from finite element analysis.

First, the stress strain distributions for a load of 10N is tested. The simulation results are as Figure 13. The simulations show that the maximum tolerance by material deformation is smaller than 0.4mm, which is appropriate for practical use.

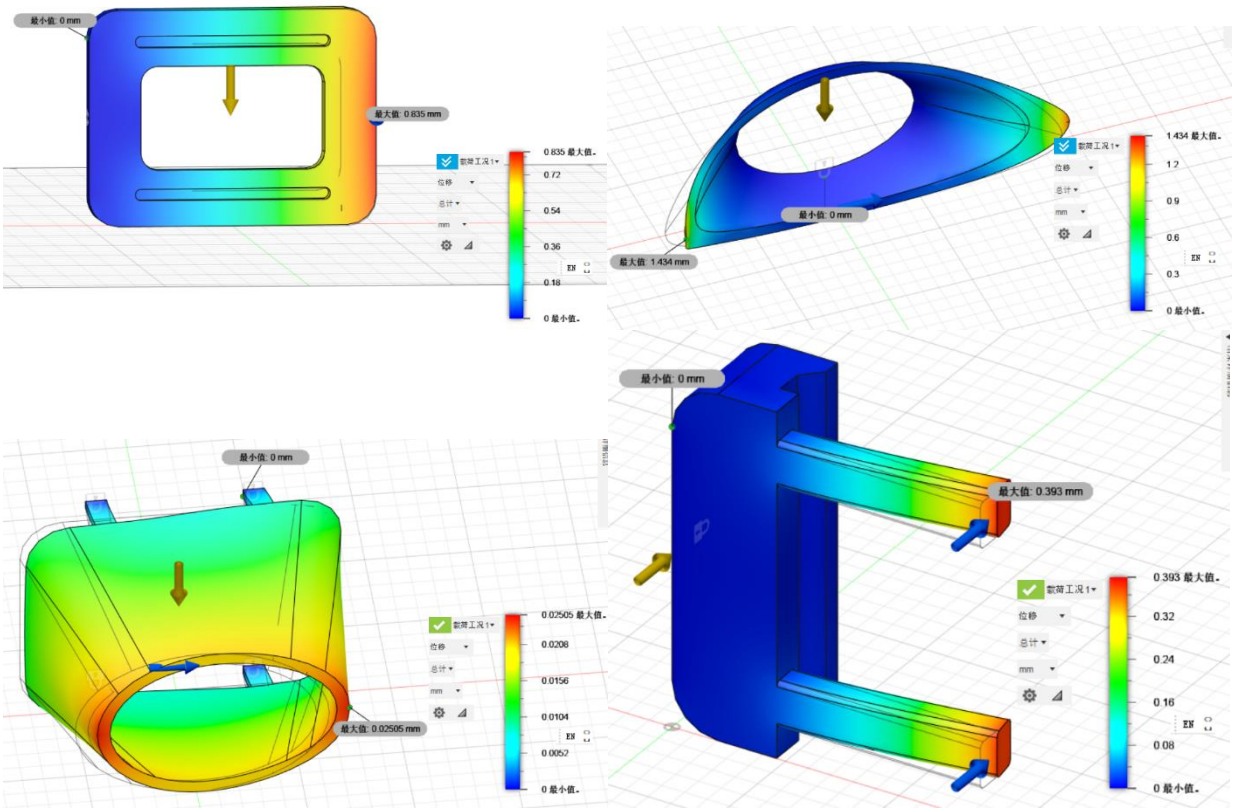


Figure 13 Static force analysis for some parts

Then the whole mechanical system for the stress concentrations on the phone is tested, shown in Figure 14.

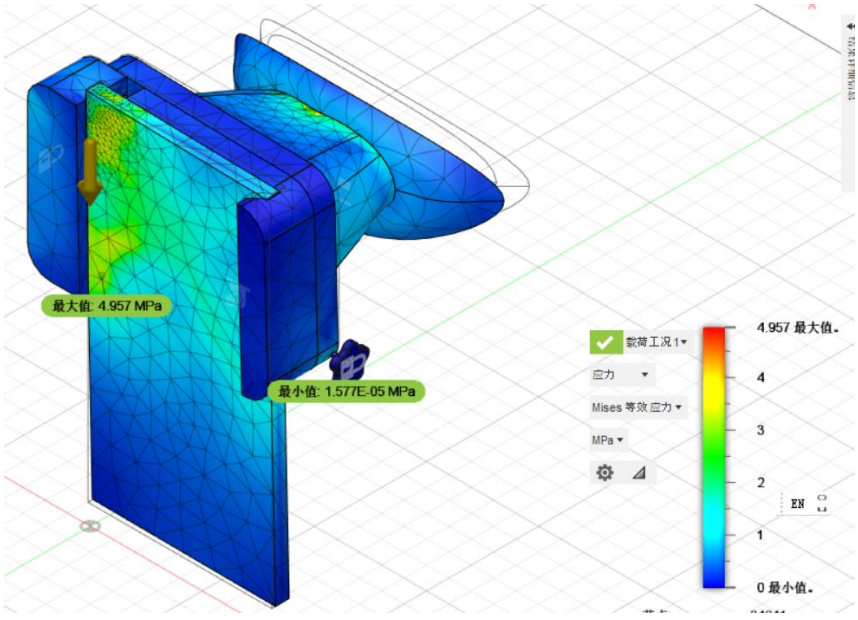
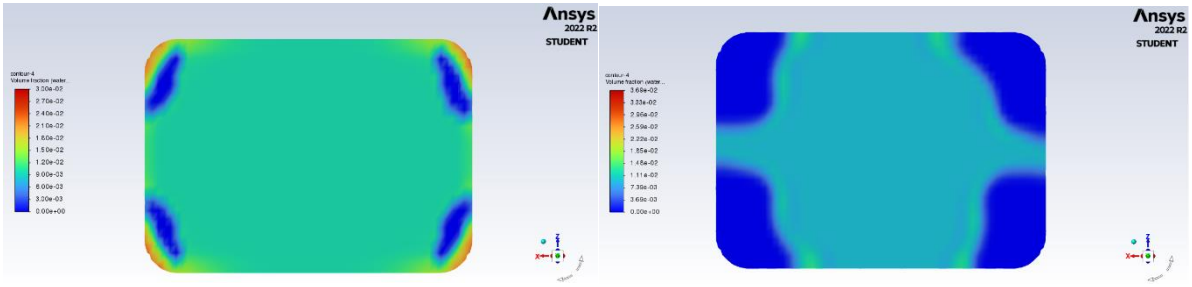


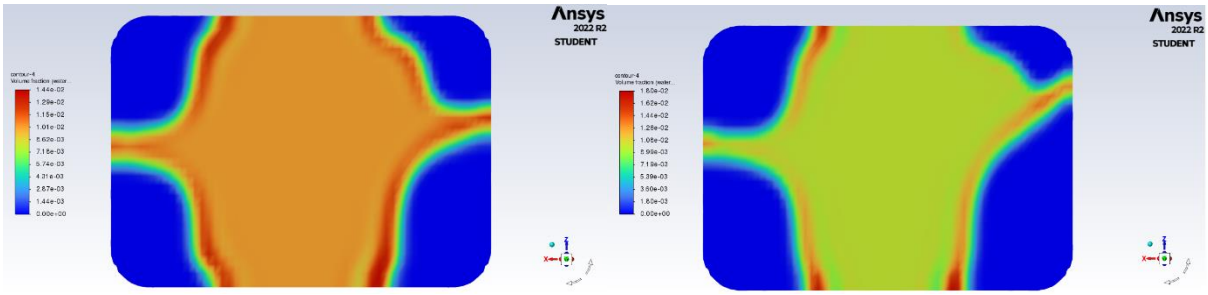
Figure 14 Static force analysis for the overall mechanical system

The maximum von mises stress is 5 MPa, which is far below the yielding stress for steel. The mechanical system is appropriate and will not make damage to the phone.

In addition, the numerical simulation is done to the condition when the user cough. The mixture coming from the human mouth is modeled as multiphase with 5% volume of fraction of water. The contour of water volume fraction is shown in Figure 15.



Contour of water volume fraction (0.002s) Contour of water volume fraction (0.0055s)



Contour of water volume fraction (0.007s) Contour of water volume fraction (0.02s)

Figure 15 Multiphase flow simulation for airflow with 5% volume of fraction of water

Figure 15 shows that if the phone is installed properly. The camera will be on the side of the square, which will be likely be confronted with high volume of fraction of water droplets.

When the assembly was made, OPPO phone was tested for the performance, shown in Figure 16. When moving the slider, as the diffuser is symmetric with its openings in the center, at some orientations, the view of the camera get blocked, and some teeth cannot be seen clearly. The phone camera is also too far away from the mouth, causing that a lot of portions of the camera view are wasted. What's more, sometimes the buttons on the edges of the phone are being pressed down by the crabs.

Figure 16 Video chips in the OPPO's camera installed in the intermediate design prototypes.

Different phones are used to test the performance. For the iPhone12, shown in Figure 17, it cannot be fitted into the system because there is a protection skin on the phone, which is very thick that do not match with the crabs.

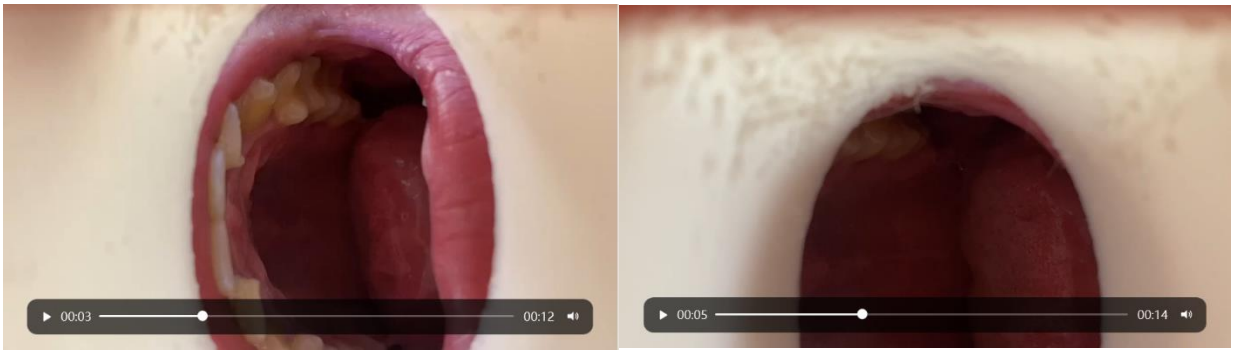


Figure 17 Video chips in the iPhone12's camera installed in the intermediate design prototypes.

For HUAWEI, shown in Figure 18, the mechanical system works fine, and clear figures of the upper teeth can be seen. Some teeth are buried under the video, which is not a problem. As the camera moves, these teeth can be seen by the camera from a different orientation.

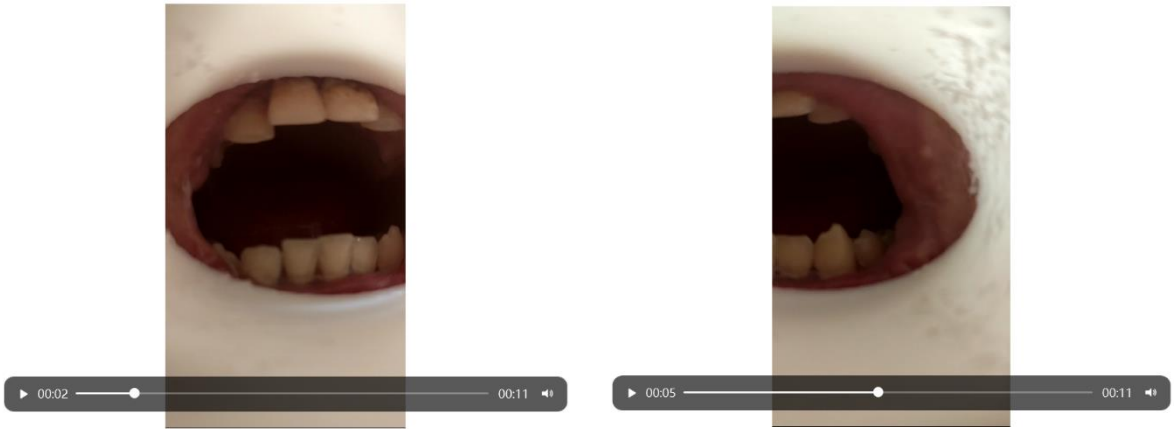


Figure 18 Video chips in the HUAWEI's camera installed in the intermediate design prototypes.

After the design improvement to the mouth opener, we see a much clearer view from the camera, shown in Figure 19.

Figure 19 Video chips in the camera installed in the intermediate design prototypes.

The outer green rim is influencing the performance of segmentation algorithm. Surrounding the outer rim by electric tapes will perfectly reduce this problem. The outer rim is the 3D printed part by elastic material TPU. This material is chosen because it is elastic and can give a good match between the diffuser and the mouth opener.

Figure 20 Final mechanical system

The dimension of the green rim is the same as the diffuser, i.e., no clearance. And the parts are connected firmly when I install the rim to the two adjacent parts.

After installing the black tapes, a much clearer view from the camera can be seen, shown in Figure 210.

Figure 21 Video chips in the camera installed in the final design prototypes.

3.2 Interactive Application

The application has been uploaded to the application market named “Your Private Dentist” and no bad feedback received. All functionality performs well during our test stage.

3.3 Segmentation Subsystem

To verify the video sample part, firstly, at the code level, it should be executed without errors. At the algorithm level, with video meeting recording requirement, the code should output 24 images and these images should include the left, middle, and right side of teeth with teeth closing and up left, up middle, upright, left, middle, right, down left, down the middle, down right side of teeth with teeth opening. The code of the video sampling part should be executed without errors, and the output of the video sampling part is visualized to showcase that it can fulfill the requirement at the algorithm level.

To verify the segmentation part, at the code level, it should be executed without errors. At the algorithm level, it can segment a tooth with IoU larger than 0.5 when tested in the test set and successfully combine the mask with the original image. The segmentation algorithm is compared to the benchmark, showing it outperforms the benchmark. Table 2 showcases the results.

Table 2 Comparison between our method and benchmark

	Background		Teeth In Mandible		Teeth In Palate	
	IoU	F1 Score	IoU	F1 Score	IoU	F1 Score
Benchmark	97.78	99.47	93.2	94.77	93.35	94.86
Our Method	97.86	98.74	93.28	96.74	94.03	97.69

3.4 3D Reconstruction Subsystem

Since the machine learning model applied finally is a black box, the intermediate features are meaningless for the explanation. Therefore, the previous verification methods cannot be used and only the system testing will be applied.

The requirements are listed in Appendix A. In verification, a dataset will be separated into a training set, a validation set, and a test set in the proportion of 8:1:1. The training and validation set will be put into the model to execute the machine learning training process and give an averaged train loss plot and an averaged validation loss back. For both losses, a value $\leq 0.1 \pm 0.05$ will be considered a good result. After this process, the test set will be put into the trained model to gain a test loss and an averaged IoU. The IoU$\geq$0.5 with a threshold = 0.4 will be considered a good result.

The data obtained are listed in Table 3. The data in the first line come from our baseline, the ShapeNet trained 3D-R2N2. The dataset is directly applied on it. And the data in the second line comes from our model. Our model meets all the requirements and gains better results than the baseline.

Table 3 Requirements Indices of 3D Reconstruction Subsystem

	Train Loss	**Validation Loss**	**Test Loss**	**Test IoU**
Baseline	0.000195	0.128109	0.174293	0.296464
Ours without segmentation	0.000179	0.040246	0.048032	0.587891
Ours with segmentation	0.000158	0.044999	**0.046675**	**0.676007**

4. Cost and Schedule

4.1 Parts

The parts and manufacturing prototype costs are listed in Table 4.

Table 4 Parts Costs

Part	Quantity	Manufacturer	Vendor	Cost/unit	Total cost
Mouth opener for adults	2	3D printer	/	$10	$20
Mouth opener for children	1	3D printer	/	$10	$10
diffuser	2	3D printer	/	$10	$20
Threaded 304 stainless steel rod	4	Provided by the innovation lab	/	$5	$20
rotator	2	3D printer	/	$1	$2
calibrator	1	3D printer	/	$20	$20
Movable crab	1	3D printer	/	$15	$15
Fixed crab	1	3D printer	/	$35	$35
TPU filament to print Elastic connector parts	1	Sanlv inc.	/	$15	$15
HDPE black sheet (not used in final design)	2	Yueji inc.	/	$8	$16
LDPE blue disposable roll (not used in final design)	1	Mowang inc.	/	$6	$6
Illuminating module (resistors, switches, batteries, wires, not used in final design)	1	Risym flagship store	/	$6	$6
Rent the Server	1	Ali cloud	/	$20	$20
Total					$162

4.2 Labor

The fixed development costs are estimated to be 40¥/hour, 10 hours/week, in total 8 weeks for four people.

$$4 \cdot \frac{40¥}{hr} \cdot \frac{10hr}{wk} \cdot 8wk = 12800¥$$

4.3 Schedule

Table 5 Schedule

Week	Zitai Kong	Yichen Shi	Xin Wang	Linjie Tong
3/13/23	Survey on 3D reconstruction Algorithms	Build basic CAD models. Start on mechanical design	Learn some basic algorithms of 3D reconstruction	Learn some state-of-art methods about image semantic segmentation
3/20/23	Study the principle and code of voxel method. Run the code successfully.	Build prototypes for some parts. Do FEA analysis on primary design	Understand these algorithms including the advantages and disadvantages of these methods and test the performance of these algorithms.	Compare some state-of-art methods about semantic segmentation and choose the one that is suitable for our projects.
3/27/23	Study the principle and code of point cloud method. Run the code successfully.	Nonstandard parts delivered from manufacturer. Test reliability of these parts and make subsequent design	Modify these codes and use these algorithms to solve our design problem.	Prepare dataset for our projects.
4/3/23	Label and preprocess the dataset	Build prototypes for all parts. Fabrication and primary evaluation	Learn some basic knowledge about application development.	Execute chosen state-of-art and modify it for our project.
4/10/23	Execute the code with the dataset of our project	Design optimization and reshape nonstandard parts	Achieve the register and login part of the application.	Find out methods that can improve performance of chosen algorithm.
4/17/23	Modify the code and try to improve the algorithm	Refine standard parts, assemble all the parts, remake broken parts.	Achieve the upload and download function for patient.	Conducting experiments about our methods.
4/24/23	Refine the code and combine it with the segmentation part	Finalize all parts and the assembly, test the performance	Achieve the upload and download function for dentist.	Refine the interface of our code and successfully assemble it with the 3D reconstruction partition code.

5/1/23	Finalize our 3D reconstruction algorithm	Small changes based on user experiences	Integrate with other parts.	Test our whole system and do some refining.

5. Conclusion

5.1 Accomplishments

The mechanical assembly was made and satisfy all requirements for the design. The highlight for this design is listed below. The methods to manufacture the parts are very easy, as the production only requires some stainless steel and PLA and TPU filaments. All installation processes are reversible so that users can have the options to get the systems in parts and they can assemble them on their own, which is beneficial to reduce the shipping cost and increase profit of companies. All materials are bio-friendly, and the structure is strong. As a result, the system will not break, and it is very safe for children to use. In the segmentation subsystem, video sample part can sample video into 30 images that include twelve views of patients' teeth. The segmentation algorithm can learn the shape of dental arche curve. It can outperformance state-of-the-art method, semask transformer. The 3D reconstruction part can output a 3D model about basic shape of the surface of teeth and dental arch curve. An extra application that can connect patients with doctors was developed.

5.2 Uncertainties

For the mechanical subsystem, we are not sure about the performance of the system for people with strongly asymmetric shapes. For the 3D reconstruction model, the robustness of the model is relatively poor, and the difference in the quality and information content of the input images will lead to the reconstruction results with a large gap. The prediction ability of the model for missing teeth and interdental space is weak, so it will choose to fill these vacant positions. This is due to the small amount of the dataset, which makes the model unable to learn a lot of tooth morphology, and the generalization ability of the model will be poor. The resolution of the model is not high enough to support clinical application, which is due to the current lack of computational power.

5.3 Ethical considerations

Our team pledges to follow the IEEE Ethics guidelines [6] and the ACM Ethics guidelines [1] as closely as possible. We will do the mechanical design by ourselves. We will not refer to the CAD files on the internet. If we must use some specific design features by other people, we will cite them [1]. We will make the mechanical design based on the user experiences. We will not use cheap but toxic materials for our design. We will not reduce the smoothing processes to reduce our workload, which will make the mechanism more likely to harm the user [1, 6]. During using the data of users, we will hold paramount the safety, health, and welfare of the patients, follow medical rigor and the Nightingale oath, and avoid any forms of data discrimination and user information leakage [1, 6].

5.4 Future work

In the future, for the mechanical system, we will give some warm prompts on the system so it will be easier for the beginner users. Also, instead of using raw data provided by the sponsor, we will use our mechanical system to take images of patients and use Cone Beam CT images to construct more 3D models to build a larger dataset. By doing so, the accuracy of the segmentation model and 3D reconstruction model will improve. Besides, we will try to get GPU with higher computation capability like NVIDIA A100, by doing so, we can train a 3D reconstruction model that can output a more precise 3D model. For the application part, we want to refine the user interface design and add a charge function to the application.

References

[1] ACM Code of Ethics and Professional Conduct, web page. Available at: https://www.acm.org/code-of-ethics. Accessed May 2023.

[2] A.D. Nguyen, et al. "GraphX-convolution for point cloud deformation in 2D-to-3D conversion." *Proceedings of the ieee/cvf international conference on computer vision*, 2019, pp. 8628-8637.

[3] C. B. Choy, D. Xu, J. Gwak, K. Chen and S. Savarese. "3D-R2N2: A Unified Approach for Single and Multi-view 3D Object Reconstruction", *CoRR* vol. abs/1604.0, 2016.

[4] D. Rahul, and F. M. Salem. "Gate-variants of gated recurrent unit (GRU) neural networks." *2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS), IEEE,* 2017, pp. 1597-1600.

[5] H. Fan, H. Su, and L. J. Guibas. "A point set generation network for 3d object reconstruction from a single image." *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 605-613.

[6] IEEE Code of Ethics, web page. Available at: https://www.ieee.org/about/corporate/governance/p7-8.html. Accessed May 2023.

[7] J. Jain, A. Singh, N. Orlov, et al., "Semask: Semantically masked transformers for semantic segmentation," *arXiv preprint* arXiv:2112.12782, 2021

[8] J. M. J. Valanarasu, P. Oza, I. Hacihaliloglu, and V. M. Patel, "Medical transformer: Gated axial-attention for medical image segmentation," in *Medical Image Computing and Computer Assisted Intervention*, 2021, Proceedings, Part I 24, Springer, 2021, pp. 36–46.

[9] J. Sun et al. "Neural 3D reconstruction in the wild." *ACM SIGGRAPH 2022 Conference Proceedings*, Vancouver, BC, Canada, 2022, pp. 1-9.

[10] Muthui Z. W, et. al., Polylactic acid (PLA) viscoelastic properties and their degradation compared with those of polyethylene, *International Journal of Physical Sciences,*2015

[11] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention*, 2015, Proceedings, Part III 18, Springer, 2015, pp. 234–241.

[12] Y. Zhang, et al. "RealPoint3D: An efficient generation network for 3D object reconstruction from a single image." *IEEE Access*, 7, 2019, pp. 57539-57549.

[13] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," in *Deep learning in medical image analysis and multimodal learning for clinical decision support*, *Springer*, 2018, pp. 3–11.

Appendix A Requirement and Verification Table

Table 6 System Requirements and Verifications

Requirement	Verification	Verification status (Y or N)
1. Precise orientation of the mechanical system a. 1 degree of freedom b. Correct location of the placement of the system	1. Verification a. From certain orientations, rotate the threads 20 times clockwise and 20 times counterclockwise, test if the final orientation is the same as the initial orientation b. For 4 members in the group, we install the matcher into our face, and use centimeters to test the accuracy of the placement of the system. A desirable situation is that the centerline of the system is within 1mm the centerline of human face.	Y
2. Stability of the mechanical system a. Self-locked system orientation b. Invulnerable to shaking	2. Verification a. As we use threads to control the orientation of the system to the mouth. An easy way to test is to first make the threads vertical, and then install a nut on the threads. If the nut does not slip down, the requirement is satisfied. b. 1Nm*s impulse on the free end of the beam, which will typically cause the beam to vibrate. When the maximum amplitude of the response is within 1mm, the requirement is satisfied. c. Also, when the phone is installed on the crab, the phone is locked by springs. We will put the system on a soft material to prevent falling, and then start shaking the mechanical system. The objective is that we want to make sure the phone will not fall due to shaking of human hands.	Y
3. 1. Sample block should successfully sample video into images with different views of teeth. 2.Segmentation model should make pixelwise prediction with high accuracy. 3. Segmentation subsystem should successfully combine mask with original image to extract teeth from background.	3. a. i. code must be executed without bugs. ii. Visualization of output should contain different views of teeth. b. i. code must be executed without bugs. ii. when tested by test set, segmentation model should achieve an accuracy above 0.5 mIoU. c. i. Visualization of output should demonstrate segmentation combine mask with original image successfully.	Y
4. The 3D reconstruction subsystem generate the final refined 3D model of its origin 2D feature with IoU>0.5 (at threshold=0.4), and final loss < 0.05	4. i. Input the test dataset images. ii. Calculate the averaged 3D IoU of outputs. iii. Calculate the loss function. iv. Observe the shape of loss curve and the final level value of it.	Y