M.A.R.V.I.N : Medical Augmented Reality Visualizer for Intuitive Navigation

May 14, 2023

Sponsors
Dr. Kader

Team Members
Nicolas Demello, Kenny Chan, Arsalan Sepahpour

Team Lead
Steve Morales

Executive Summary

The project at hand encompasses a wide range of objectives aimed at revolutionizing the field of surgery by addressing the limitations and challenges faced by surgeons when visualizing a patient's internal anatomy during surgical procedures. Traditional 2D displays have long been the standard, but they restrict surgeons' ability to perform surgeries with utmost precision and introduce a higher risk of complications. Therefore, the underlying motivation of this project is to harness the power of modern technology, specifically Microsoft Hololens and Augmented Reality (AR), to provide real-time data and three-dimensional organ modeling to surgeons in the midst of surgeries. By doing so, the project aims to significantly enhance surgical precision, minimize complications, and ultimately improve patient outcomes.

Furthermore, this endeavor seeks to go beyond the boundaries of physical location by enabling remote guidance and collaboration among surgeons worldwide. Through the implementation of advanced communication and visualization technologies, the project aims to break down barriers and increase access to high-quality healthcare. This will not only benefit patients who may otherwise face limited resources and expertise but also promote knowledge-sharing and skill development among medical professionals on a global scale.

To accomplish these ambitious goals, the project employs a combination of Unity, a powerful game development platform, and the Microsoft Mixed Reality Toolkit. By utilizing these tools, the development team is creating a custom application specifically designed for the Hololens2 device. This application allows surgeons to visualize a patient's internal anatomy within the augmented reality space generated by the Hololens2. To establish a seamless connection between the application and medical devices, such as ultrasounds, a communication framework based on WebRTC is being employed. This framework enables the transmission of live video footage from these medical devices to the Hololens2, where it is then overlaid with accurate and detailed organ models, providing surgeons with a comprehensive and intuitive visualization of the patient's anatomy during surgery. To facilitate secure communication and collaboration, the project also incorporates a signaling server, which ensures encrypted data transmission and enhances the ability of surgeons worldwide to work together effectively.

Throughout the project's development, several significant milestones have already been achieved. The team successfully tested the desktop application's capability to receive video input from a wide range of medical devices through the utilization of a capture card. Furthermore, they have also accomplished the conversion of two-dimensional CT scans into three-dimensional models that can be rendered within the Unity software. However, additional work is still required to establish a reliable and efficient transfer of these organ models from the desktop to the Hololens2 via remote connection using WebRTC, as certain challenges with the WebRTC library have arisen and need to be addressed.

Looking ahead, the project's future steps entail further enhancing the intuitive interface of the Hololens Unity software. This will involve the implementation of visual guidance elements and voice command support, which will empower surgeons to interact seamlessly with the AR environment. Additionally, the development team will finalize the user interface for both the desktop and Hololens2 applications, ensuring a seamless display of output from the DaVinci surgical system within the Hololens2 environment. Furthermore, the project will encompass the implementation of communication capabilities between multiple Hololens2 devices, thereby enabling surgeons to collaborate and assist each other in real-time during complex surgical procedures. By combining these advancements, the project aims to revolutionize the field of surgery and pave the way for safer, more precise, and globally connected healthcare practices.

Table of Contents

Introduction

Several problems arise in the field of surgical procedures. One of the primary problems plaguing surgeons is the difficulty visualizing a patient's internal anatomy during surgeries. Traditional visualization is only in the second dimension (2D displays) which limits a surgeon's ability to provide surgery with poor anatomical visualization. This causes procedures to be far more risky leading to additional complications for the patient. Medicine constantly evolves with technology, so as engineers, isn't there a way to integrate modern technology to improve these issues?

Throughout this project our team has to view two sides of the same coin, the patient and surgeon. This sets up our motivation and the context for our project. Not only do we want to design and build a technology that solves the problem at hand, but also ensure as potential patient ourselves, we would be more than comfortable to be assisted by it.

The societal and global context for this project is significant. Access to quality healthcare is a global concern, and the use of AR technology in surgical procedures can contribute to overcoming barriers and increasing access to innovative healthcare solutions. The sponsor is interested in this project due to the potential societal impact and the transformative nature of the technology. The project aligns with the sponsor's mission to advance healthcare delivery and improve patient outcomes whilst also making it easier on surgeons to perform their operations.

At a high-level the idea is to use Microsoft Hololens as our medium, such as a laptop, to provide surgeons with real-time data and 3D organ modeling from a patient during surgical procedures. By doing so, we aim to improve surgical precision, reduce the risk of complications, and ultimately improve patient outcomes. We do this by leveraging the capabilities of Augmented Reality (AR) technology. The Hololens is an AR device which can be used to enable surgeons to visualize a patient's internal anatomy in real-time without the need for traditional 2D displays. Additionally, the communication abilities will allow for remote guidance from other surgeons around the world increasing equitable access to quality healthcare while improving surgical outcomes for patients.

This technology has the potential to significantly improve healthcare delivery, increase access to quality healthcare, and inspire further innovation in the healthcare industry. Everyone needs to go to the doctor, this directly impacts you! By leveraging these technologies, the project addresses the limitations of traditional anatomical visualization and provides an immersive and interactive environment in real time enhancing surgical procedures and collaboration among surgeons worldwide and potentially saving lives. One of the biggest reasons why this project is important is because it's innovative in the healthcare industry leading to the development of new solutions which can benefit medical professionals and patients alike.

From an economic perspective, the project offers several factors that motivate its development. Firstly, by improving surgical precision and reducing complications, the project has the potential to decrease the costs associated with surgical procedures. This includes shorter hospital stays, faster recovery times, and a potential reduction in the need for expensive simulation labs or cadavers. Additionally, the implementation of AR technology can lead to more efficient and quicker surgical procedures, resulting in time and cost savings.

While there are other solutions out there that are similar, they are often limited in their scope or effectiveness. For example, traditional surgical navigation systems may require the use of cumbersome equipment or have limited accuracy. Hololens technology, on the other hand, provides a hands-free and intuitive interface that can provide doctors with the information they need in real-time, without interfering with their workflow.

Other technologies using virtual reality (VR) and AR have been used in the healthcare industry to enhance medical education, training, and patient care. However, similar products are often cumbersome and unintuitive for surgical procedures as they can be bulky, require extensive training, and may not provide real-time feedback and communication between medical professionals.

As such, our project provides a more comprehensive and accurate solution that can address the limitations of existing technologies in the healthcare industry while also bridging the gap between medical education globally.

Technical Background

We use the following technologies: Unity, Augmented Reality (AR), Electron and Node.js, WebRTC (Web Real-Time Communication). We do not write anything too in-depth and primarily write at a high level to explain the technicalities.

Unity is a popular cross-platform game engine used for developing interactive 3D applications. In this project, Unity is utilized to create the application for Microsoft Hololens2, enabling the visualization of a patient's internal anatomy in an Augmented Reality (AR) space. Unity provides a robust framework for rendering 3D models, handling user interactions, and integrating various other UI components of the project.

AR technology enhances the real-world environment by overlaying virtual objects in the real world. In this project, AR technology is leveraged through Microsoft Hololens2 to provide surgeons with real-time data and 3D organ modeling during surgical procedures. By blending virtual information with the surgeon's view, AR improves anatomical visualization and surgical precision.

Electron is an open-source framework that allows for the development of desktop applications using web technologies such as HTML, CSS, and JavaScript. Node.js is a runtime environment that enables server-side execution of JavaScript code. In this project, Electron and Node.js are used to develop the desktop application that facilitates the transmission of data and DICOM files from the desktop computer to the Hololens2. Electron provides an intuitive user interface, while Node.js enables communication with the Hololens2 and other components of the system.

WebRTC is a communication protocol that enables real-time audio and video communication directly between web browsers and applications. It facilitates the transmission of live video footage from medical devices, such as ultrasounds, to the Hololens2 in real-time. WebRTC ensures seamless and secure communication between the desktop application and the Hololens2, enhancing the visualization capabilities of the project.

System Design

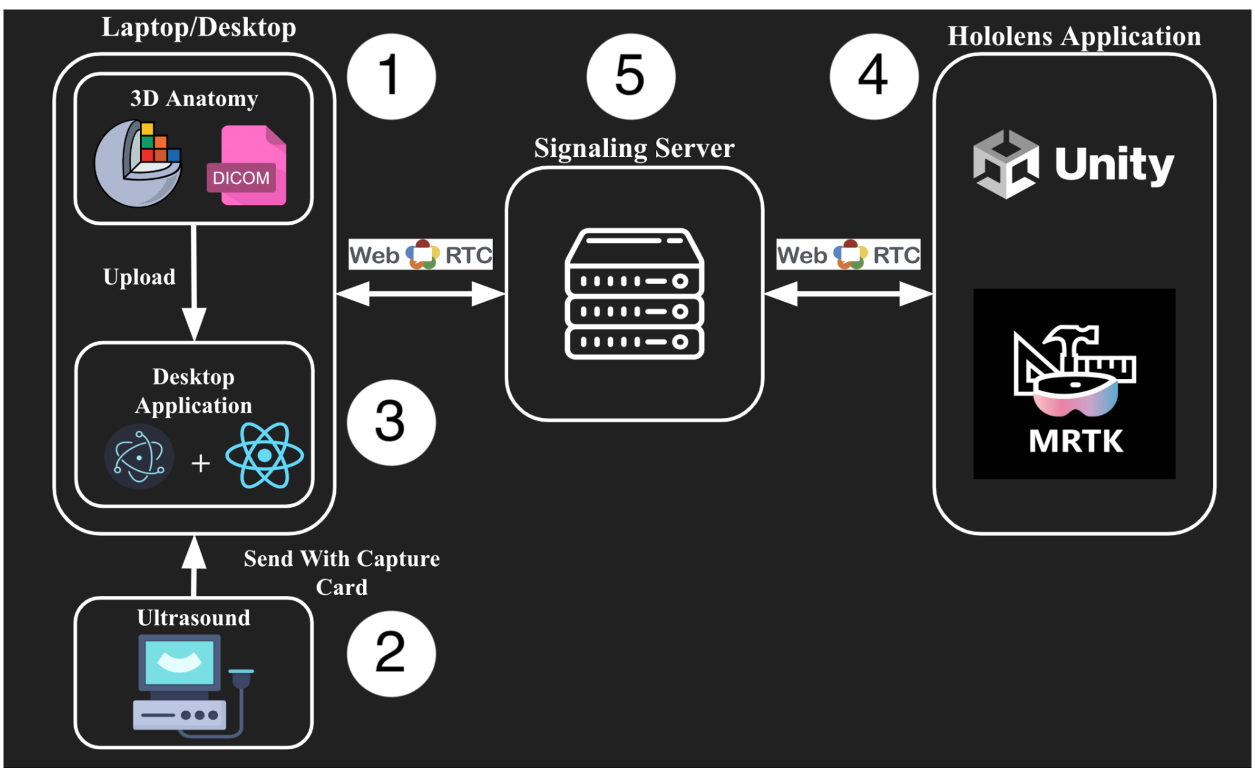
This is typically the longest section, giving a description of your design. Include, as appropriate, block diagrams, circuit diagrams, etc. as well as English explanations of the elements of the design. Explain the challenges you faced and corresponding solutions.

In putting forth a description of your design, we do not want a chronological description (as in first we did this, and then we did that, and it didn't work, and so next we tried this...). Rather you should provide a functional description of whatever is the final design you adopted.

On the other hand, if there are things you tried which didn't work, and it would be instructive to the sponsor or to a future student to know that, then you can describe that (but do it in the context of a separate subsection presenting alternative design choices which didn't work as well, not in the context of a chronological narrative).

Overview

A diagram or figure showing the main modules/submodules of the system, how they interface with each other, and the workflow. A there are a total of 5 modules to our whole system.

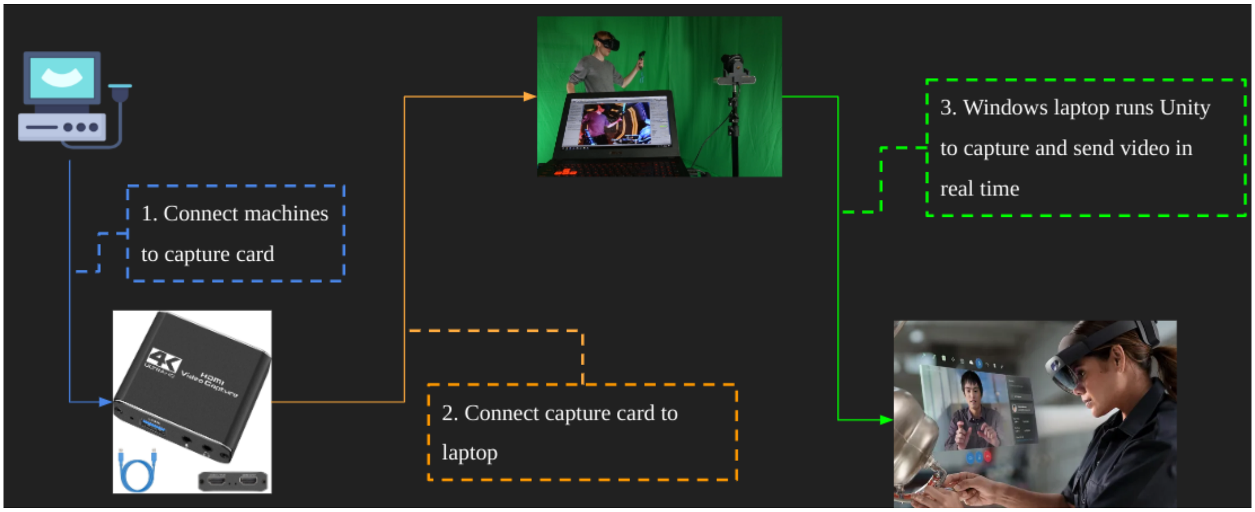


1. 3D Anatomy

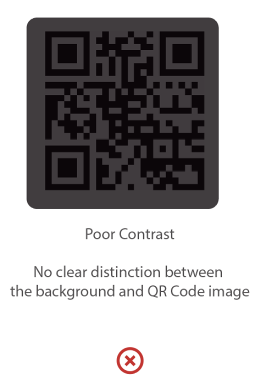
For turning DICOM files into 3D models, we used a software called [Slicer] to auto segment the 2D images and create 3D models of the organs. For the implementation of the auto-segmenetor we installed both the ML model, and its dataset within the application, a detailed tutorial can be found in this [github] repo. It may also be beneficial to explain DICOM files are an industry standard to encapsulate a patient's anatomy from scans or x-rays. This standardized format makes it easier to exchange files and work with other doctors.

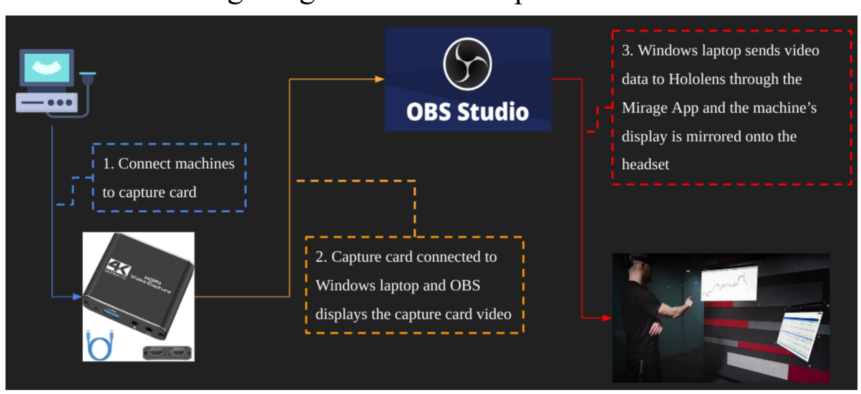
2. Video Capture

To get video input into the user's laptop, we used a video capture card as a middle man, where the medical device gets connected to it and then it is inputted into the laptop. This in turn allows the user to select the video input via the desktop application.



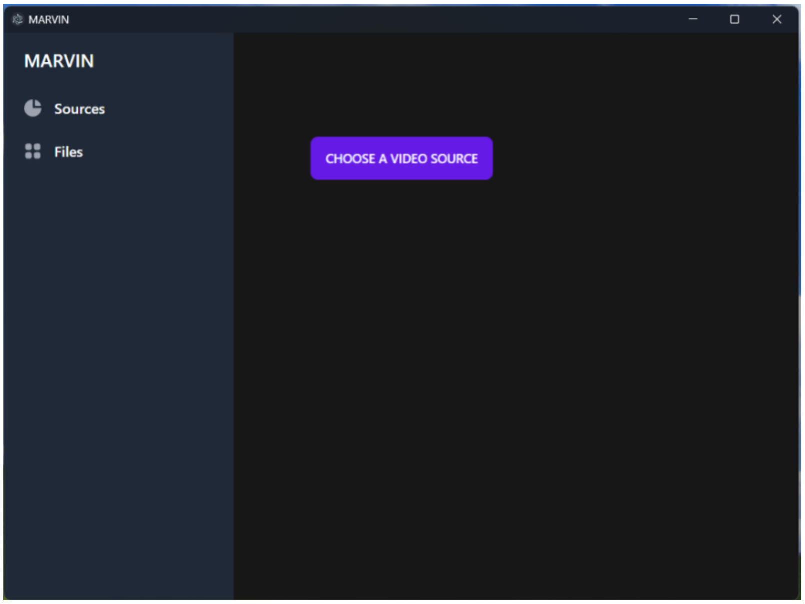
One of the previous solutions was very similar except that we used a software called Mirage to display the video output to the Hololens. Unfortunately we had connectivity issues as the QR code the Hololens needs in order to connect to other devices remotely was very difficult to scan. Specifically the background of the QR code was too dark so the contrast was basically causing it look like a black box. The following images are from the previous solution.





3. Desktop Application

We created a desktop application using React and Electron where the user can select their video input and upload their 3D models. We chose these tools because we noticed the operating system used among doctors and hospitals vary from place to place, using Electron to create the application enables us to create one application for all operating systems. However, there is a slight concern that the application size might become quite large and consume too much processing power.
The following is the desktop application we made.



Previously we were looking into making it with Unity, but it does not offer the best toolset to allow for proper UI.

4. Hololens Applications

Unity is a powerful and versatile game engine that can also be used to create applications for various platforms, including the Microsoft HoloLens 2. The HoloLens 2 is an augmented reality headset developed by Microsoft, which allows users to interact with virtual objects in the real world.

To create applications for the HoloLens 2 using Unity, we use the Mixed Reality Toolkit (MRTK). MRTK is an open-source framework provided by Microsoft that simplifies the development process and provides a set of pre-built components and tools specifically designed for mixed reality experiences.

MRTK serves as a bridge between Unity and the HoloLens 2 hardware, providing a collection of scripts, prefabs, and assets that enable developers to build applications with features like hand tracking, spatial mapping, voice recognition, and gesture recognition. It includes a wide range of building blocks and interactions that can be easily integrated into Unity projects, saving developers significant time and effort.

Unity's versatility is a key factor in its popularity for creating HoloLens 2 applications. Unity is a cross-platform game engine that supports multiple operating systems and devices, including Windows, iOS, Android, and more. This flexibility allows us to create applications for the Hololens.

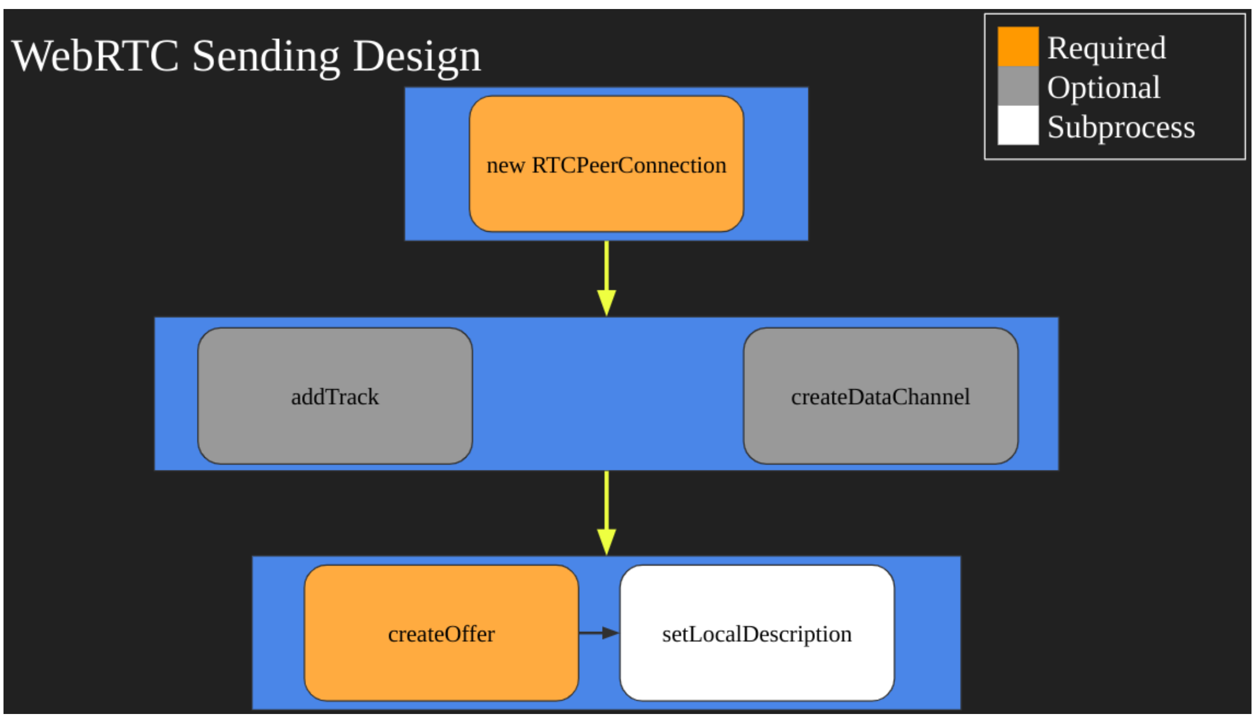
5. WebRTC + Signaling Server

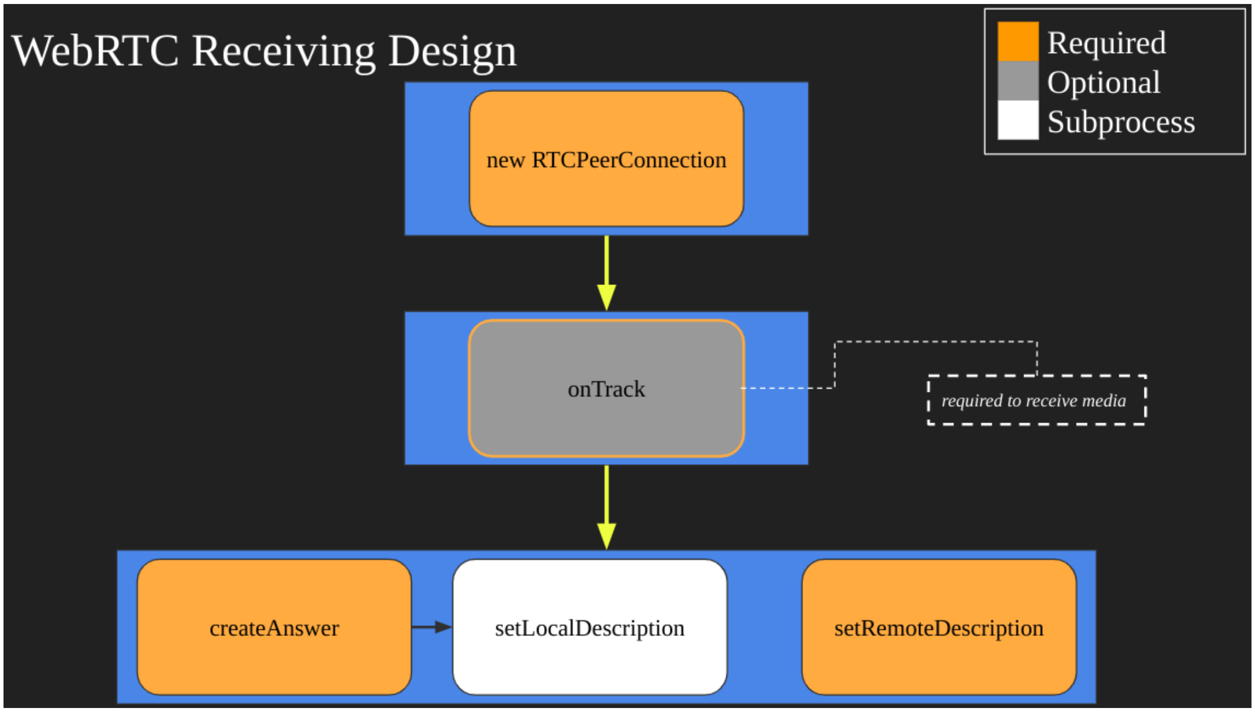
In order to create a communication system to send and receive video we choose to use the WebRTC technology. WebRTC enables real-time communication between web browsers or applications without the need for plugins. The process involves two main steps: signaling and peer connection.

During signaling (explained later), the peers (devices) exchange information through a server to establish a connection. This information includes network addresses, encryption keys, and media settings. Once the signaling is complete, the peers can establish a direct peer connection.

To send media, the sending peer captures the audio or video stream through the user's microphone or camera. This stream is encoded using supported codecs and packaged for transmission. The encoded media is then sent over the network using standard protocols, with ICE facilitating connectivity between peers. On the receiving end, the media is received, decoded, and rendered for playback. WebRTC also provides a Data Channel for exchanging additional application-specific data. With the simplicity of its API, WebRTC enables real-time communication and collaboration directly within web browsers or applications.

The following are the high level designs for video sharing from each side.





The bidirectional design of WebRTC allows for sending video using a set of API method names. Here's a simplified explanation of the process:

1. Getting User Media: The first step is to access the user's media input, such as their camera and microphone. This is achieved using the getUserMedia() method. It prompts the user for permission to access their media devices and returns a stream of audio and video tracks.

2. Creating Peer Connection: To establish a connection between two peers, the RTCPeerConnection object is created using the new RTCPeerConnection() method. This object represents a local peer and handles the encoding, decoding, and transmission of media.

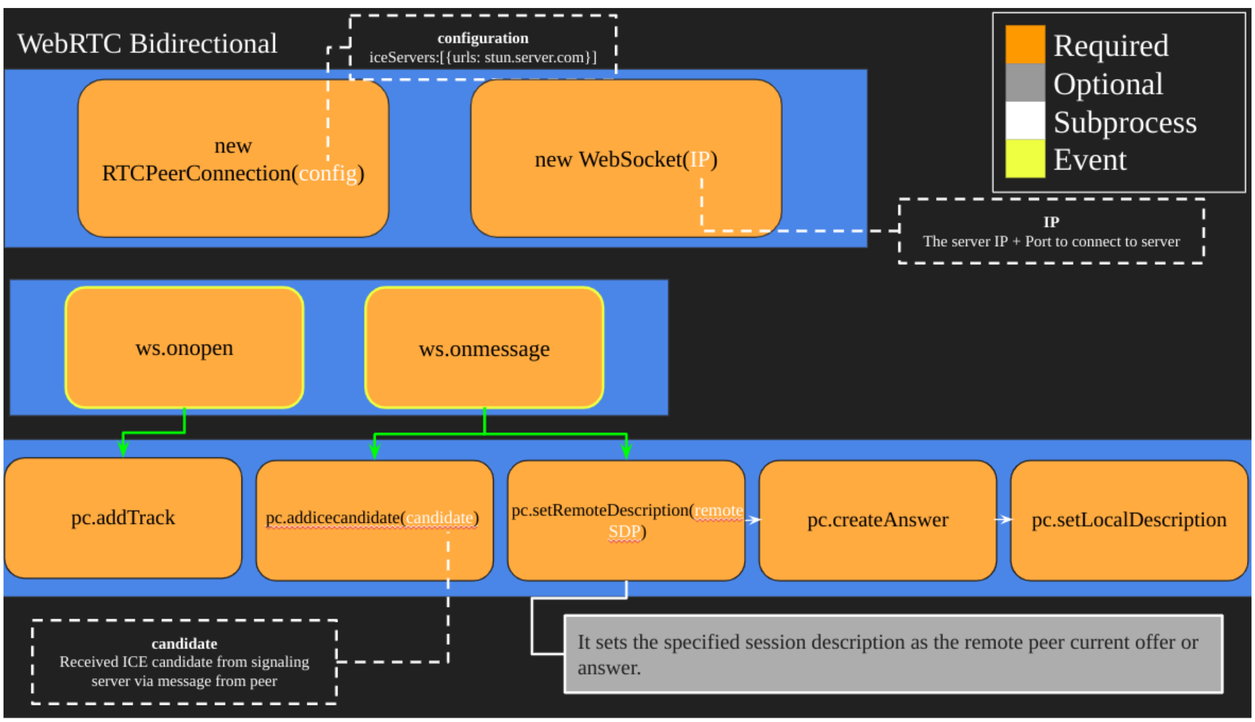
3. Adding Local Media Stream: The local media stream, obtained from getUserMedia(), is added to the peer connection using the addTrack() method. This makes the video stream available for sending to the remote peer.

4. Signaling and Establishing Connection: The signaling process begins to exchange session information with the remote peer. The local and remote peers exchange signaling messages containing network information and other metadata. The specific method names for signaling may vary depending on the signaling implementation chosen.

5. Receiving Remote Media Stream: As signaling progresses, the remote peer's media stream is received. When the remote signaling message is received, the addTrack() method is used to add the incoming video track to the peer connection. This allows the video to be rendered locally.

6. Rendering Video: The video stream is rendered on the user's screen using HTML5 video elements or a canvas. The specific method for rendering depends on the chosen UI framework or custom implementation.

7. Sending and Receiving Video: With the connection established and the video streams added, the media is automatically sent between peers. WebRTC handles the encoding, packetization, and transmission of video data over the network. The bidirectional nature of WebRTC allows for both peers to send and receive video simultaneously.



Design and Implementation

One major challenge we had was getting communication working between the desktop and Hololens as well as communication between Hololens. This is a challenge as there is no official software for interacting with the Firebase server using Google's SDK so we must create our own CRUD operations to the server using HTTP requests. This communication with the server would be used to communicate between Hololens and other Hololens devices as well as the desktop computer which acts as the medium to send data to Hololens devices such as video streams of monitoring systems and 3D model scans of organs for overlaying onto the patient.

Machine Learning segmentation faces challenges such as poor data quality, overfitting, class imbalance, complexity, interpretability, and generalization. These challenges can result in biased models, inaccurate predictions, poor generalization, and difficulty in clinical decision-making. Techniques such as regularization, oversampling, undersampling, class weighting, optimization, parallelization, and transfer learning can be used to mitigate these issues. However, tweaking these parameters is cumbersome and generally takes time.

In order to create 3D objects by using CT and MRI scans, we are using an open source Machine Learning model that is able to segment the medical scan and isolate the organ we are interested in displaying. This is crucial and saves many hours. Initially without this model, to produce each 3D object from DICOM files, we would need to sit down with a medical professional that could understand the scan and tell us what should be isolated. This process was long and impractical.

Experimentation

We have tested that the desktop app is able to receive capture card video input from various medical devices successfully. The Unity software is also able to render a 3D model as well but we still need to implement a reliable way to transfer this 3D model from desktop to hololens through the said remote connection using WebRTC and Firebase.

Next Steps

The Hololens Unity software as it currently stands is not the most intuitive interface. We would like to implement visual guidance for the application and voice command support to facilitate the use of the headset to create a more intuitive interface. We are also aiming to complete the user interface for both the desktop and HoloLens2 application to display the output from the DaVinci surgical system in the HoloLens2. Communication between multiple HoloLens2 will be implemented to allow surgeons to assist each other in surgery.

Impacts & Concerns

MARVIN has the potential to impact several dimensions including economic, social/cultural, global, environmental, public health, safety, and welfare, as well as ethics and professional responsibility. From an economic standpoint, MARVIN can potentially reduce costs associated with surgical procedures, including reducing complications, faster recovery times, shorter hospital stays, and potentially reducing the need for expensive simulation labs or cadavers. Additionally, the implementation of MARVIN can lead to more efficient and quicker surgical procedures, saving both time and money. Long-distance remote guidance communication facilitated by MARVIN can also save on travel costs.

From a social/cultural perspective, MARVIN can facilitate collaboration between medical professionals from different countries and cultures, sharing knowledge and expertise across borders. This can help address the global shortage of skilled medical professionals and create equity by increasing access to cutting-edge medical technology and expertise. The growth of the medical technology sector through projects like MARVIN can create new jobs and revenue streams.

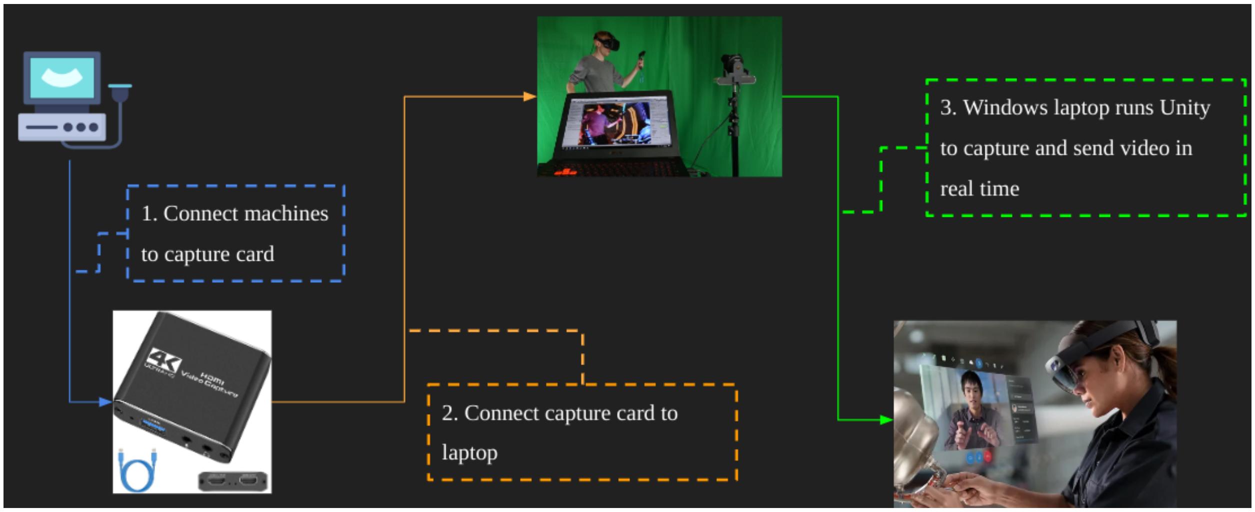
In terms of public health, safety, and welfare, MARVIN can improve the safety of surgical procedures by removing distractions like monitors and improving focus on the patient with centralized patient information. The improved comfortability and convenience for the surgeon by having a better view of monitoring systems can reduce the probability of surgical error. MARVIN can also improve the overall quality of healthcare by enabling equitable access, modern data analysis, and accurate diagnoses and treatment plans.

One of the concerns we have come from an ethical and responsible standpoint, MARVIN must prioritize patient safety, and doctors should not heavily rely on technology. There must be a balance between experience and technology. Data privacy is also crucial, and patient medical data must be protected. Software encryption when transferring data between devices and disclosing benefits, risks, and potential impacts to the patient is essential. MARVIN can address these ethical and responsible considerations by prioritizing patient safety and privacy while enabling modern medical technology.

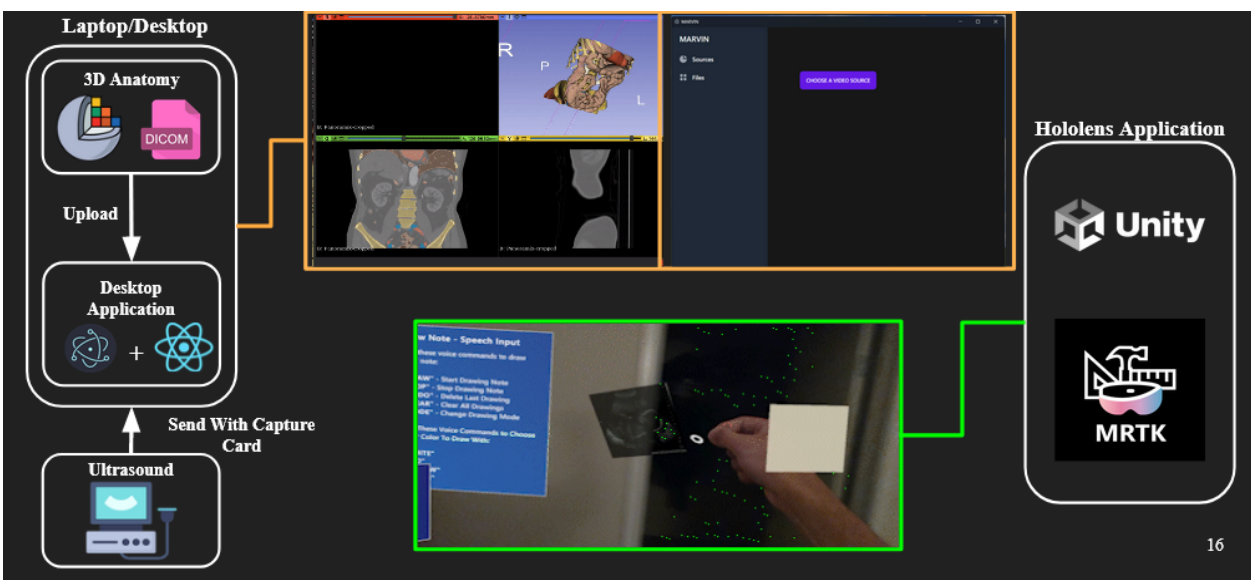
Overall, the MARVIN project has the potential to impact several critical dimensions, including economic, social/cultural, global, public health, safety, and welfare, and ethics and professional responsibility. The successful implementation of the project can create new jobs, improve medical procedures, enable equitable access to medical technology, and improve patient safety and privacy.

Results

We successfully achieved the majority of our initial goals and requirements. Early on we figured out how to capture video using capture cards (shown in the image). We also used Uity to be able to display the video output to the Hololens.



Initially, our primary objectives were to develop 3D models of organs, capture video footage from medical devices, and enable visualization on the Hololens device. We accomplished the tasks related to video footage and 3D modeling without significant hurdles. However, when it came to transferring the video and models to the Hololens, we initially implemented a solution using FastAPI, which initially worked but was later deemed inadequate due to lag issues identified during testing. To address this challenge and achieve faster transfer rates, we opted to utilize WebRTC. Unfortunately, we encountered a setback as the required software package for Hololens was unavailable, preventing us from establishing seamless communication between the database and the Hololens.



In the orange highlighted (left), you can see DICOM segmentation using the slicer app which is a UI which allows you to turn 2D DICOM files into segmented 3D models. The right of the orange highlighted shows our UI which has a tab for the user to switch between the file sharing tab and the source sharing tab. In the source sharing tab the user is able to select USB and video devices such as webcam and capture card. When medical equipment is connected to the laptop, it displays on the sources tab. In the green segment you can see the FastAPI implementation which displays virtual monitors, 3D models, and has a drawing capability.

Conclusion

Throughout the course of this project, significant accomplishments have been achieved in addressing the limitations of surgical procedures through the integration of modern technology. By leveraging Augmented Reality (AR) technology and Microsoft Hololens2, the project has successfully developed an application using Unity, enabled real-time data visualization, and improved surgical precision. The project has also implemented communication capabilities through WebRTC and Firebase, facilitating remote guidance and collaboration among surgeons worldwide.

In a societal and global context, the impact of this project is substantial. By improving surgical procedures, the project contributes to the enhancement of patient outcomes, reducing complications, and increasing access to quality healthcare. The implementation of AR technology has the potential to transcend geographical boundaries and bridge the gap in medical education globally. Surgeons from different countries and cultures can collaborate, share knowledge, and provide expertise, thereby addressing the global shortage of skilled medical professionals. This project empowers the medical community to leverage cutting-edge technology and overcome barriers to equitable healthcare.

Beyond the primary use of enhancing surgical procedures, this project opens up possibilities for various other applications. The use of AR technology can be extended to medical education and training, allowing students and aspiring surgeons to learn and practice in immersive and interactive environments. Additionally, AR-based telemedicine can be explored, enabling remote consultations and diagnoses, particularly in underserved regions. The potential for further innovation and expansion of this project is vast, with the capacity to revolutionize healthcare delivery on a global scale.

For the continuation of this project by another team, the focus should be on refining and enhancing the user interface and experience. Implementing visual guidance and voice command support for a more intuitive interface would further improve usability. The team should aim to complete the user interface for both the desktop and Hololens2 applications, enabling seamless integration with the DaVinci surgical system and facilitating the display of output in the Hololens2. Moreover, the team should explore the implementation of communication between multiple Hololens2 devices, allowing surgeons to assist each other in real-time during surgeries. Continuation of the project by another team holds great potential to further advance the technology, expand its capabilities, and optimize its impact on healthcare.

References

Unity: https://docs.unity3d.com/Manual/index.html
Microsoft Mixed Reality Toolkit:
https://learn.microsoft.com/en-us/windows/mixed-reality/mrtk-unity/mrtk2/?view=mrtkunity-2022-05
WebRTC: https://webrtc.org/getting-started/overview
DICOM: https://www.dicomstandard.org/current/
Firebase: https://firebase.google.com/docs
Electron: https://www.electronjs.org/docs