Universitat Politècnica de Catalunya

Written by:
Name: Nil Crespo Peiró
Bachelor: Degree in Data Science and Engineering

Tutor:
Name: Pere-Pau Vazquez
Department: Computer Science

Final Report

Bachelor's Thesis: Surgical Video Analysis Tool

TFG: Surgical Video Analysis Tool

Simplifying Surgical Learning

Abstract - English

Part of the teaching of future medical surgeons is based on the viewing and analysis of videos of surgical operations. However, the current technology to visualize them has certain limitations.
The analysis of recorded videos of surgical operations that are currently present and available to the surgeons nowadays are very limited in number and functionalities. For instance, they don't allow to filter by the tools used in the operation nor the phases that it has. Because of that, I decided to create a tool that allows a more efficient viewing of annotated surgical operation videos. To do so, I developed a series of visualization tools and a user interface, in the form of a website to present the aforementioned tools.

The main goal of this thesis is to create a tool that can be used to solve some needs in the field of medical science. Specifically. to facilitate the learning process of surgeons and resident doctors. During the development of the project, the tool has been evaluated by six surgeons to back up the hypothesis of the usefulness of this tool to the medical surgeon profession.

Abstract - Spanish

Parte de la enseñanza de los futuros médicos cirujanos se basa en la visualización y el análisis de vídeos de operaciones quirúrgicas. Sin embargo, la tecnología actual para visualizarlos tiene ciertas limitaciones.
Los análisis de vídeos grabados de operaciones quirúrgicas que existen actualmente y están a disposición de los cirujanos son muy limitados en número y funcionalidades. Por ejemplo, no permiten filtrar por las herramientas utilizadas en la operación ni por las fases que ésta tiene. Por ello, decidí crear una herramienta que permitiera una visualización más eficiente de los vídeos de operaciones quirúrgicas comentadas. Para ello, he desarrollado una serie de herramientas de visualización y una interfaz de usuario, en forma de página web, para presentar dichas herramientas.

El objetivo principal de esta tesis es crear una herramienta que pueda ser utilizada para resolver algunas necesidades en el campo de la ciencia médica. En concreto, facilitar el proceso de aprendizaje de los cirujanos y los médicos residentes. Durante el desarrollo del proyecto, la herramienta ha sido evaluada por seis cirujanos para respaldar la hipótesis de la utilidad de esta herramienta para la profesión de médico cirujano.

Abstract - Catalan

Part de l'ensenyament dels futurs metges cirurgians es basa en la visualització i l'anàlisi de vídeos d'operacions quirúrgiques. No obstant això, la tecnologia actual per visualitzar-té certes limitacions.
Els anàlisis de vídeos gravats d'operacions quirúrgiques que existeixen actualment i estan a disposició dels cirurgians són molt limitats en nombre i funcionalitats. Per exemple, no permeten filtrar per les eines utilitzades en l'operació ni per les fases que aquesta té. Per això, vaig decidir crear una eina que permetés una visualització més eficient dels vídeos d'operacions quirúrgiques comentades. Per tal de conseguir-ho, vaig desenvolupar una sèrie d'eines de visualització i una interfície d'usuari, en forma d'aplicació web, per a presentar aquestes eines.

L'objectiu principal d'aquesta tesi és crear una eina que pugui ser utilitzada per resoldre algunes necessitats en el camp de la ciència mèdica. En concret, facilitar el procés d'aprenentatge dels cirurgians i els metges residents. Durant el desenvolupament de el projecte, l'eina ha estat avaluada per sis cirurgians per donar suport a la hipòtesi de la utilitat d'aquesta eina per a la professió de metge cirurgià.

Contents

List of Figures

List of Tables

Acknowledgments

Throughout the process of my thesis and the writing of this final report, I have received a great deal of support and assistance.

I would first like to thank my tutor, Associate Professor Pere-Pau Vázquez, whose expertise was really valuable when deciding on some of the main parts of the project. Your feedback pushed me to always find a way to improve my work and find better solutions.

This project began in Germany, and therefore I can't omit the big help they provided. I wanted to thank **Ph.D. Anirban Mukhopadhyay** and research assistant **Henry John Krumb**, for teaching me the fundamentals of *Deep Learning for Medical Imaging*, and my colleagues **Amin Ranem** and **Paco Rahn** for working along with me in that project.
On this same note, it has to be mentioned, that without the work of the **research group CAMMA** and the *University Hospital of Strasbourg/IRCAD* , who developed the dataset of annotation videos that I use in this project, none of this would be possible. And also the creators of the paper [Twi+16]*EndoNet: A Deep Architecture for Recognition Tasks on Laparoscopic Videos* which were of great help when designing the algorithm.

I would like to acknowledge the surgeons from the *Hospital Vall d'Hebrón* **Dr. Martí**, **Dr. Malagelada**, and **Dr. Pando, Dr. Blanco**, and **Dr. Dopazo**, and apprentices of surgeons **Dr. Martínez** and **Dr. Fernandes**, for all the help in the user study of the real life application of my thesis.

The workers at **RDLab**, have also been of great help in the process of using their GPU to improve the results of our model, even if it finally was not as good as expected.

Finally, I could not have completed my thesis without the support of my friends, **Marc Cortadellas**, **Xavier Rubiés**, **Guillem Borràs**, and **Roger Creus**, who provided stimulating discussions that made my final project even better as well as happy distractions and constant support.
In addition, I would like to thank my parents for their support throughout the process, and their wise counsel and sympathetic ear while writing this final report and preparing the oral defense.

1 Introduction

The project is carried out at the [UPC] *Universitat Politècnica de Catalunya* under the supervision of Associate Professor **Pere-Pau Vázquez**.

The purpose of this project is the creation of a tool that facilitates the **visual exploration** among a set of surgery videos dealing with cholecystectomy procedures.

In this section two main aspects of this bachelor thesis will be discussed:

1. The reason behind the creation of this project, with a detailed explanation of how we ended up with the idea.
2. The evaluation of other tools already developed, who give some similar functionalities to the final user, in the field of surgical learning.

Once the project has been introduced, the main goals of the project will be explained to give a general overview of what should be expected in this document.
Then, the technology used in this project will be explained and shortly after that the development of the thesis and the results will be presented.
And last but not least, we will show the user study that we have committed and the final conclusions to which we have arrived, after finalizing the whole process.

1.1 Project Background

The deadline to submit the idea for this bachelor thesis was in the beginnings of January, when I was currently performing my ERASMUS exchange semester in the [TUD] *Technische Universität Darmstadt*, a city near Frankfurt in Germany. Due to the fact that I had barely no contact with anyone from my home university I had problems regarding communication with teachers and tutors when needed information or assistance with the choice of my bachelor's thesis.

In this section, the project conducted in Germany will be detailed as well as the results that we obtained (myself with the help of my colleagues Amin Ranem and Paco Rahn), the challenges we faced and the ideas that came from it, which ended up creating the foundations for this project.

1.1.1 Definition of the project

It was the start of January and I was finishing a subject in Germany called **Deep Learning for Medical Imaging**. In that subject, we were given a dataset (the Cholec80 dataset), which contains 80 videos of cholecystectomy procedures carried out in the University Hospital of Strasbourg/IRCAD (Strasbourg, France)[Str]. The videos are annotated with labels that indicate the presence of tools in video frames and also in which phase of the operation are this frames.

With this in mind, we were asked to **develop a Deep Learning Model** to predict the tools on each frame of the videos. We based some of our work in the paper created by the authors of the dataset [Twi+16].

One of the main issues using the whole Cholec80 dataset was the size of each video. Although they had been downsampled to 1 frame per second (fps), each frame still needed to be resized to 224x224 pixels, in order to be used for transfer learning with PyTorch. However, once downsampled, the whole transformed dataset was still too big to train with. Based on this problematic, from each video 2000 slices were randomly selected, if the

video has more than 2000 frames, other-wise the whole video will be used for training. This was made using the random library from Python.
In the development of this thesis, we used two main models:

- A variation of AlexNet called **ToolNet**, which basically consists on taking AlexNet off the shelf and change the last layers in order to be able to predict the tools in each video frame.
- The second model was the **ResNet** model. The idea to test transfer learning using ResNet-50 is based on the Cholecystectomy Cataract lecture from the Deep Learning for Medical Imaging (DLMI) course where the ZIB-Net was introduced. The ZIB-Net used two different network architectures (AlexNet and ResNet). The results of the ZIB-Net showed that the ResNet-50 based model could achieve a higher accuracy then the AlexNet.

In the next section, we will conduct an analysis of the results for both models.

1.1.2 Results obtained

In this results section, the two tables corresponding to the results obtained when using the two models explained above will be seen:

First, this are the results of the AlexNet model that we used:

Model #	# videos	# epochs	learning_rate	batch_size	weight_decay	train_acc	train_loss	val_acc	val_loss	test_acc
1	37	30	1e-3	100	0.75	29.33%	0.003332	31.23%	0.003474	-
2	5	300	5e-3	200	0.5	21.45%	0.006631	6.85%	0.006808	-
3	5	60	1e-2	50	0.5	30.19%	0.01309	26.94%	0.01347	-
4	5	30	1e-3	50	0.9	17.81%	0.007252	6.486%	0.007888	-
5	5	60	1e-4	32	1e-3	98.89%	0.0002176	37.05%	0.01546	47%
6	20	60	1e-4	32	1e-3	96.76%	0.0004687	56.84%	0.01028	61%
7	37	10	1e-4	32	1e-3	89.30%	0.001431	72.99%	0.004042	61%
8	5	30	1e-4	32	5e-4	98.34%	0.0002409	45.77%	0.01377	46%
9	5	30	5e-5	32	5e-5	100%	5.323e-7	43.06%	0.03911	36%
10	5	30	5e-5	64	1e-5	100%	1.354e-7	46.89%	0.03258	59%
11	37	30	5e-5	64	1e-5	99.61%	3.064e-5	58.99%	0.007504	-

Table 1: Results of the AlexNet model that achieved 61% of test accuracy

And here next, the results for the ResNet model will be shown:

Model #	# videos	# epochs	learning_rate	batch_size	weight_decay	train_acc	val_acc	test_acc
1	37	30	1e-4	100	0.75	30.16%	38.63%	29% -
2	5	15	1e-4	30	0.01	58.61%	28.6%	30% -
3	5	15	1e-4	30	5e-3	60.2%	40.2%	43% -
4	5	10	1e-3	30	0.00025	67.22%	35.07%	40% - %
5	5	10	1e-3	30	5e-5	70.2%	45.24%	53%
6	5	10	5e-4	30	0.000125	72.2%	41.62%	47%
7	5	10	5e-3	30	0	68.7%	51.2%	27%
8	40	20	1e-3	62	5e-5	-%	72.2%	-%

Table 2: Results for the ResNet-50 model that achieved 53% of test accuracy

We achieved results around 60% of test accuracy which are acceptable, when taking into consideration the amount of data available (80 videos of about 25 to 40min, at 25fps) and the fact that we did not have our own

GPU, and had to use [Kag]Kaggle Kernels which only provides 13GB of RAM when using a GPU. This last drawback is relevant to mention, since the amount of power required to run the deep learning algorithm, that we have mentioned, is way larger than what a normal CPU could achieve.

1.1.3 GUI created

In order to make the algorithm more user-friendly we created a user interface, which could be used for prediction for the users. We wanted it to be very intuitive so this are some of the windows the user would have seen when running the following command `python3 SurgeryToolRecognition.py --use_gui`, the following Welcome window will be displayed, to welcome the user to the Surgical Tool Recognition System:

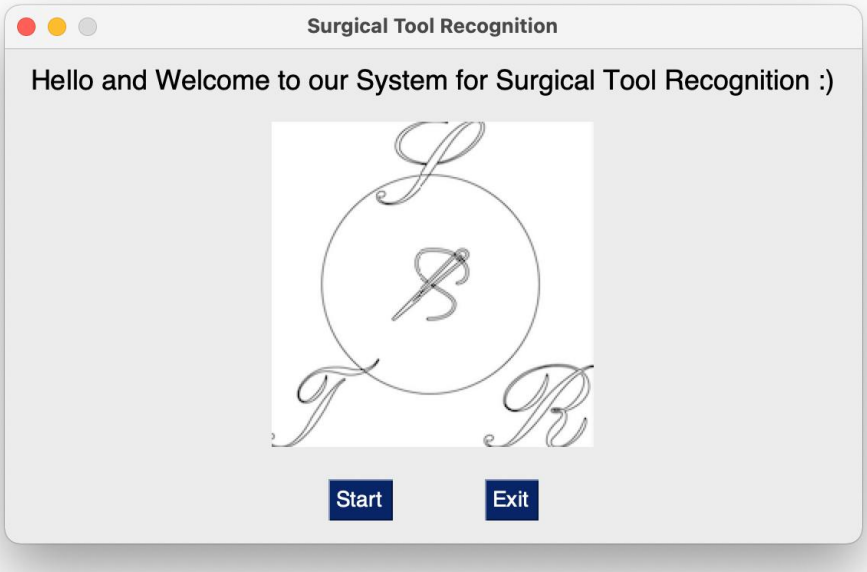

Figure 1: Welcome window

Once the user hit `Start`, a new window was displayed, were the user needed to specify the wanted video that had to be loaded in order to use the prediction tool in it:

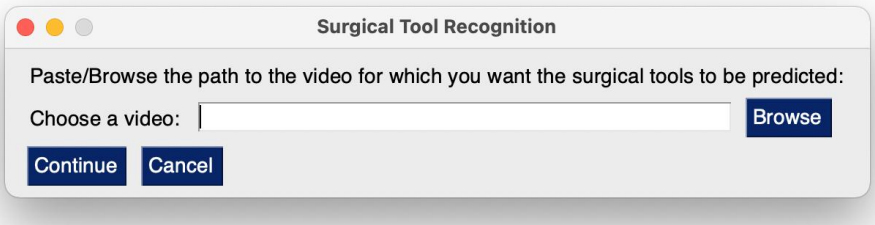

Figure 2: Browse window to select the video the user wants to input

After the chosen model (from between the two explained above: AlexNet/ResNet) , the parameters and specifications the selected model had been trained on will be displayed.

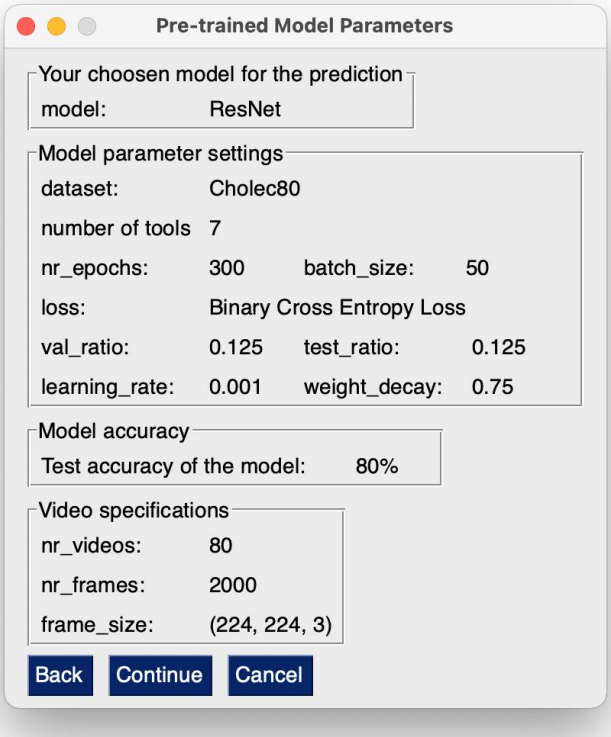

Figure 3: Window that shows the parameters the selected model has been trained on with its test accuracy

Once the user specified which model and device he/she wanted to predict on, the prediction would start. In this case a progress bar was shown to the user on which the progress can be seen.

Once the computation is finished, the final window would appear, showing the results. In this window the predicted tools for each frame were listed. The same output was going to be saved as a .json file after the user hit `Start again` to use a new video or `Finish` to exit the tool:

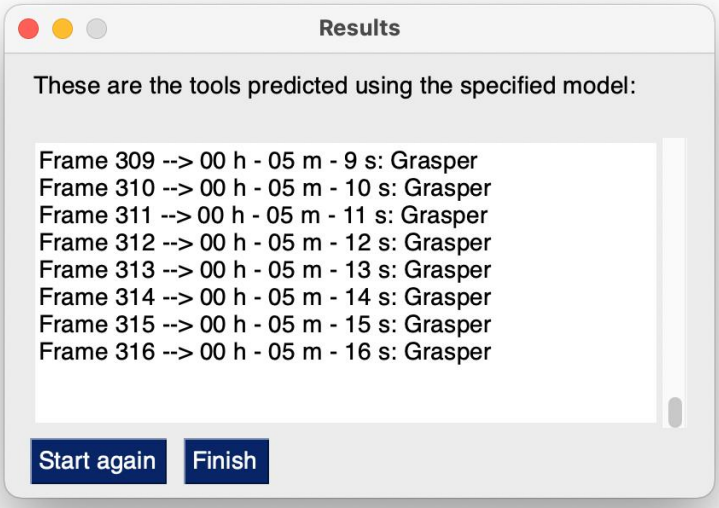

Figure 4: Results window with output

1.1.4 Challenges encountered and Suppositions

However, I knew that the dataset could be further explored and improved, so I originally thought of trying to improve the results we got. The first idea was to try to overcome the results achieved by the authors of the EndoNet system [Twi+16] (which was the one used as inspiration for our project and also the creators of the Cholec80 dataset) , but then realized that it was more interesting to create a tool that could help people in doing some tasks.

The software that was developed allowed classification of which tool appeared on each frame. However, it did not allow a visualization of the results in a friendly way, so it was hard to check if the results made sense without having to watch the entire video. So, in order to make the results useful, it would be interesting to create a tool to visually inspect the labeling results.

From this idea, the initial hypothesis/supposition was developed:

Part of the teaching process for future surgeons may consist in the inspection of videos of surgical procedures to understand how a certain operation must be carried out. This idea came from the fact that nowadays with the current Covid-19 situation very few people are allowed into the surgical rooms, and it is the main way of how surgeon apprentices or resident doctors best learned how to perform the surgical operations before. So thinking about other ways that these apprentices could learn, given the current and future technologies like

nanocams and 5G, videos of the operation from different angles and cameras could be recorded and then used for apprentices to understand how to proceed.
So, it would be very helpful if they could review a video of how an operation should be done, just before entering the surgical room, in order to refresh their memory on the key uses of certain tools or some key phases of the operation without having to actually visualize the whole video.

However, after conducting a series of informal interviews with some surgeons, that will be explained in section (6), we were able to realize that our original hypothesis was not accurate, even though it was not that far off: Even with the Covid-19 situation, apprentices of surgeons (or resident doctors, as they are called) can still attend surgical operations to learn the process, and they even do theirs (with assistance of their Tutor). However, the goal that was initially thought (to create a tool that could be used to visualize some key parts of the operation) it was very much appreciated and the interviewed surgeons mentioned the great utility it would have in the learning progress of resident doctors.

1.1.5 Thesis Idea

After discussing different possibilities with my tutor, we defined the goal as:

- ***Creating a visualization tool that facilitates the navigation in the videos by means of visual labeling of the timelines, as well as the development of searching widgets to facilitate the process of learning for surgeons.***

1.2 Related Work

When developing the idea, and after starting to work on the project, two questions arised:

1. *Was is it really solving a necessity in the field of surgical learning?*
2. *Was there a tool already developed with some or most of the functionalities that I wanted to create?*

The first question, was solved by conducting some informal interviews with surgical doctors from the Hospital de la Vall d'Hebrón, which will be further analyzed in section (6).

The other question was not that easy to solve since it was still not clear, at the moment of formulating these questions, all the functionalities that my tool was going to have.
However after talking to a surgeon, who said that she actually used a website called **WebSurg** [IRC00], to watch conferences and recordings of surgical operations, we finally found one website, who could give help to the surgeons in a similar way to what our tool offers.

WebSurg is a free (but you must register) website, property of the online university of IRCAD, which acronym stands for Research Institute against Digestive Cancer, (are the same who created the dataset used for the training of the deep learning algorithm). This website contains a huge compilation of videos: from **conferences**, to **online courses**, to **lectures**, to **recorded surgical operations**, which will be our focus.
As we can observe in the following image, this is an example of a recorded surgical operation:

Figure 5: Example of a Surgical Video in WebSurg

Some of the positive aspects of the surgical videos on WebSurg, compared to our tool, is the quantity of them. There is a wide range of surgical operations, and not only Laparscopical Cholecystectomies.
It can also be noted, that it has some features that our tool does **not have**:

- The ability to account the number of views of a video.
- The ability to share a video with other people.
- The option to save the video to a "watch later" list.
- The option to save the video to a playlist.

All of this features could be very useful to include in our web application/visual tool.
Some of the functionalities that there are in common between our tool and the video player of WebSurg, is the ability to skip the video directly to some parts, in the WebSurg case, to what they call "Keysteps", which are the list that one can see on the right of the video in Figure (5). In our case, however, instead of keysteps we have the defined phases of the operation.

On the other hand, there are also many features which **only the visual tool created in this project has**:

- There is no option in WebSurg that allows the user to select a tool, which would then make the video slow down when the tool is present and accelerate when the tool is not. This is helpful when wanting to remember, how one has to use this concrete tool, in an operation, without having to look for the tool in the whole video.
- Furthermore, there is no interactive visual graphic, that allows the user to see, at which instants of the video a tool appears, and how long does a phase lasts, or which tools are present in each phase, which our tool allows.
- Aside from this features from the main tool developed, after talking to the surgeons, two additional tools have been created, which are:

  * A video comparison tool, that allows to compare phases of two different videos so one can see the differences between their operation and a model operation.

  * A video annotation tool, that consists in a software that allows the annotation of videos during the visualization of a surgical video and then it allows the user to download this annotations with their corresponding timestamps, when they were recorded.

None of this is possible in the WebSurg website and because of that we can conclude that we are adding value to the market, when creating this tool.

1.3 Structure of the thesis

Once all the background and related work already created in the field of surgical learning has been detailed, a concise and short summary of how is this document structured will be shown:

- **Introduction**: In this section, the background of this project and the related work has been explained. Which consists on, where did the idea come from, how much work was made abroad, and which tools are already offering similar functionalities to what my web application offers.
- **Project Goals**: Indicates and details the goals of this project, if they have been fulfilled and the new goals that have been added mid-way the creation of this thesis.
- **State of the art of the technology used or applied in this thesis**: This section basically explains, all the technologies that have been used in each of three sections (website, visualization and code), in the development of this project, which problems have I encountered, if I had previous experience or had to learn a new technology from scratch, etc.
- **Project Development**: This section contains the main challenges faced during the development of the thesis will be exposed, as well as the methodology of work used during the whole project and how have I made each part of the web application, with detailed information about the process.
- **Results**: In this segment, the final structure of the web application is displayed as well as, the content of each tabs and a series of use cases, that can be solved with my tool (with detailed steps and images).
- **Evaluation**: Here, a user study is shown, where 6 surgeons of different fields and different experiences were ask to evaluate the usefulness of the web application and the visual tool and add give some insight of what could be improved (which some of it has actually been implemented).
- **Conclusions**: In this last section, the conclusions to which I have gotten after finishing this process have been explained and also there is a list of future developments that could be done in the near future.

2 Project Goals

As we had aforementioned, the initial goal of this thesis was to create a tool that would facilitate the visual exploration of several annotated surgical videos dealing with cholecystectomy procedures. This became a need when the result of the deep learning algorithm could not be visualized in order to check, if it was correctly detecting the tool at each frame.

With this objective in mind, we had to define some initial tasks or main goals which were:

1. **Creation of a User Interface** to place the Video Analysis Tool
2. **Creation of a Surgical Video Analysis Tool** that allowed the user to see at which instants a tool appeared, and filter the video for tool and phase

Another complementary goal that was set, but more like an additional goal aside from the two main goals above described was the following:

3. **Improvement of the current Deep Learning Algorithm** to detect Tool and ability to also detect Phase of the video

After showing the initial tool to the experts of the field, in this case surgeons, some ideas were proposed, so two more tasks were created:

4. Creation of a tool that would allow the surgeons and resident doctors to **compare side by side** their video of a surgical operation with an operation of an expert or a theoretically perfect operation.
5. Creation of a tool that would allow the surgeons, who are tutors of resident doctors, to **evaluate** the surgeries of resident doctors by means of creating annotations.

2.1 Creation of a User Interface to place the Video Analysis Tool

Since a video analysis tool was the main initial structure I planned on developing, I needed a user interface where the tool could be placed.

I initially thought, since I had already created one, of creating a GUI like the one in figure 1. However I did not end up satisfied with the design and overall looks of it, since it looked a little too simple and plain.
Then, I considered creating a mobile application, but discarded this option as visualizing videos in a mobile device, was not appealing and not the best solution. Added to this, the goal was also that users would not miss details, and thus, the small size of screens in mobile devices would probably made them not suitable for this problem.

In the end, I went with a web application, that could contain this Visual Tool and other features that I would think of, further along the process.
Using a web application, has also other advantages, such as the fact that the user does not need to install anything on their computer. This is a great feature, because it is difficult to install third-party software in computers belonging to hospitals due to security issues.

2.2 Surgical Video Analysis Tool

As it has been explained in the introduction section, due to the lack of visual feedback the results of the deep learning algorithm gave us (figure 4) the goal was to develop a visual tool that would allow the user to see at

which frames did a tool appear and in which frames it did not.

The main goals that this tool had to fulfill where the following:

1. It should be interactive, as when hovering over the heatmap, it would indicate which tool was the user hovering over, at what time of the video was the cursor on, also which phase and most importantly: if the tool was present at that instant.
2. The user should also be able to filter by tool and then clearly visualize only the instants where the tool appears.
3. It should be able to have the same feature set but filtering by phase of the video instead of tool.

2.3 Improvement of the Deep Learning Algorithm

When the project carried out in Germany came to an end, we were a little bit disappointed in the results because, we felt that we could have done more if we had had a strong, non-limited GPU available.

Because of this, one of the goals I set for this project was to improve the results that I had gotten in that project, so they could be visualized in the visual tool (instead of the handmade annotated ones).

Aside from the improvement, another task that was set was to add another feature to the algorithm so it could learn also the phases of the surgical operation. This was explained in Twinanda et al [Twi+16], so I thought it was feasible, and therefore it was also a goal of this project.

2.4 Video Comparison Tool

After performing the user study, I realized that creating a tool that would allow the surgeons to compare their videos with more complicated surgeries, or surgeries made by experts in a concrete surgery would be very beneficial and useful to them so I decided to mark it as an additional goal.

The requirements of this tools would be:

1. Have the ability to place two videos side-by-side.
2. This tool would need to have a button to play and pause both videos at the same time.
3. Additionally, it should also allow the user to stop and pause any of the videos individually in case one of them was slower or faster than the other.
4. Lastly, it would be very useful to have the ability to go directly to a phase of the operation where the user knew that most of the differences were.

2.5 Video Annotation Tool

As it was the case with the video comparison tool, this goal was also added a little after all the main goals, due to the conversations with the surgeons of the user study that had been conducted.

The goal would be to create a tool that would allow the surgeons, who act as tutors for resident doctors to evaluate the work of this residents, by annotating messages, where mistakes have been made in their surgical operations.

The initial set of identified requirements would be:

- The ability to record an annotation, and automatically detect at which instant the annotation has been made
- The ability to go directly to a phase of a video and annotate mistakes in that phase.
- The development of a component to store all of the annotations made by the tutor, so he/she can see which annotations have been made and when have errors been committed, while the annotation process is still ongoing.
- The ability to download a file with the annotations and their corresponding timestamp.

3 State of the art

To develop this project several software packages and technologies were used. I already had some experience with Python and Altair, but others were the first time I had even heard of them. As a result, some of this tasks were not as fast as they could have been, since it took me time to master a level of expertise sufficient to develop the tasks that I wanted to perform in each of the new technologies that I had to use.

The technologies used in this Bachelor's Thesis have been divided in three main sections:

- **Web Application**: All the software that has been used in order to create all the files, components, pages and interactions of the website.
- **Visualization**: This section will include how have been made all of the visualizations and interactive graphics in the website and in this file.
- **Code**: And the last section will consist on the libraries and languages used to create the deep learning algorithm and the creation of the processed videos.

With all or most of them I encountered some challenges and they were solved through help from friends who gave me support and ideas and also with the aid of Google and more specifically the website of StackOverflow [Stab] which most of the times contained some information that was very helpful in order to solve the problems that I was facing.

3.1 Website Technologies

Up until this moment all the websites I had built were through a web application which had templates and one could edit them. However, one thing that I wish I learned in this career is to be able to develop apps or webs in order to publish the data science projects we make in this bachelor of Data Science and Engineering that I'm about to finish. Because of this and also the need of creating an interface where I could lay down the video analysis tool I decided to create a website.

However since I had no experience with HTML, CSS nor JavaSCript, I had to look for some information and tutorials of each of them so I could learn the basics and progress from there.

- **HTML + CSS = Bulma**: In order to facilitate the task of learning how to code a website from scratch I found out that there was a framework called Bulma, which allowed me to not have lo learn CSS and only focus on HTML and JavaScript. Bulma is an open source framework that provides ready-to-use frontend components that can be easily combined with HTML to build responsive web interfaces. So by only learning which Bulma components I wanted to use, I did not have to define with CSS all the styling and properties of each component. In order to learn how to program a web using HTML and Bulma I completed the following tutorial/course by the web NinjaNet [Nin].
- **JavaScript**: The JavaScript part of the project is where most of the problems came from since, for every feature I wanted to add, I encountered a problem and had to look ways to solve it and this was a recurring problem throughout the project, mainly due to the lack of knowledge in this language used to create the visual analysis tool website.

3.2 Visualization Technologies

In this subsection, the technologies that helped me to create the visualizations that I have used in either the website or this document will be exposed and detailed.

- **Altair**: Altair is a declarative statistical visualization library for Python, based on Vega and Vega-Lite. This python library has been the main visualization tool that has been used in the thesis, and it is present on the main tool of the website as an visual video exploration tool. It consists of a heatmap with interaction and also contains legends with the ability to be toggled, to preview only some parts of the data (by tool and phase of the video).

  I had some previous experience with Altair as it was the main language used in the course of Information Visualization, coursed in the second year of the degree in the UPC [UPC].

- **Infogram, Flourish, Datawrapper**: These websites are data visualization websites, that without the need of programming allow the user to create visualization reports that are visually appealing, and have been used to show to the surgeons and also to visualize the results in the Evaluation section (6).

  I had also previous experience with this technology as I have used it in many projects up to day.

- **D3**: Since the implementation of Altair had some drawbacks that we will explain further on, I decided to have a look at D3 since it offered some solutions to that problem. The idea was to use it to create the interactive heatmap in the main tool, and another one in the annotation tool. I had zero experience with D3 so I decided to look into a tutorial from a youtube account called [fre] freeCodeCamp.org. However, in the end it did not get implemented because the visualization that had been made in Altair had also lots of strengths that I wanted to keep and in the end it was left in the section of future developments to integrate it in the video annotation page.

3.3 Code Technologies

Aside from the coding that had to be done in order to design and create the website, there has also been programming with a different coding language, which has been the well-known Python.

- **Python**: This coding language has been used in our bachelor specially in the past 2 years, where we have started to program lots of data science projects, and Python is really useful for it. However in this project it has been used mainly in two parts. To modify the deep learning algorithm and trying to improve it, and also to create, process and edit the videos corresponding to the surgical operations where only one phase is shown or where a video is fast-forwarded when a certain tool is present and slowed-down when not.

  PyTorch: In order to accomplish, the first goal the PyTorch library has been used, which is an open source machine learning framework that facilitates the creation of deep learning models and architectures.

  MoviePy: In the other hand, there is MoviePy, which is a python library used for movie editing. The functionalities for which I have used this library have been mentioned above, but this library would also be used to develop a feature that is written in the section of future developements (7.2], which is the addition of subtitles to the video that indicated which phase is the video on and the names of the tools that appear in each frame.

3.4 Apps/Websites used

In order to use all the technologies, that have been explained in the past three subsections, there have been four apps or websites have been used, which should also be mentioned:

- **Google Colaboratory**[Goo]: Has been the main website where all the python-code related parts have been developed as well as the process of cleaning and adapting the dataset to the appropriate form in

order to create the Altair visualizations. It has also been used to edit, transform, cut and process the clips of the videos filtered by phase, tool and phase and tool together.

- **Kaggle**[Kag]: After struggling to find a working GPU, Kaggle was once more used again a little bit to see if an improvement of the model could be made in a small scale while waiting for an actual high-power GPU.
- **Visual Studio Code**[Cod]: Has been used to program all the website contents, from the CSS files, to the HTML and JavaScript ones.
- **Overleaf**[Ove]: Finally, Overleaf has been used to write all the files that are present in the website window of Documentation, which is detailed in section (5.2.2). All the help I needed to solve the design problems I needed to solve were found in the website of [Staa]StackExchange and the overleaf documentation.

4 Project Development

4.1 Methodology

In this project I decided to implement an agile methodology, due to the fact that I knew I was going to face challenges, which I could not predict and plan ahead of time and that I should be able to adapt and overcome them through a methodology to allowed me to constantly vary the goals based on the current limitations that I had or problems that I was facing.

The Agile methodology is a way to manage a project by breaking it up into several phases. It involved a constant collaboration with my tutor, in my case, and continuous improvement at every stage. Once the work began, I cycled through a process of planning, executing, and evaluating each step. Continuous discussion and evaluation is vital, that is the reason behind the decision to hold weekly meetings with my tutor.

In order to implement this methodology I worked with a series of frameworks/software that allowed me to pursue this way of organizing myself.

- **Sprints:** Sprints are fundamental, when implementing an agile methodology. A sprint is a short period of time (a week approximately in my case), when a team, which in this case was composed of my tutor and myself, works to complete a set amount of work. By using sprints in the right way, it helped me organize better, while solving more challenges without losing productivity.

  In this project there were two kinds of sprints defined:

  The sprints with the tutor were once a week, were we discussed the work improvements and talked about the tasks that were going to be developed in this next sprint/week, while also solving some problems when I thought I could get some advice from his expertise.

  The other kind of sprints were with myself also once or twice per week, where I reviewed my Kanban board and mini Gantt to be able to plan, adapt and solve the tasks that I needed to get done in this sprint.

- **Mini Gantt Charts:** A Gantt chart, is one of the most popular and useful ways of showing tasks displayed against time. On the left of the chart is a list of the activities and along the top is a suitable time scale. Each activity is represented by a bar; the position and length of the bar reflects the start date, duration and end date of the activity. This allows you to see at a glance:

  - What the various tasks are.
  - When each task is supposed to begin and end.
  - Where activities overlap with other activities, and by how much

  Even though, Gantt charts are not really suitable for an agile methodology, I had to present a Gantt chart in each of the two first reports that I was asked to submit during the process of the creation of this thesis. However, I also developed mini Gantt charts in order to organize and schedule the work I needed to do in the following two weeks.

  One would think that using Gantt charts is the polar opposite to using an agile methodology but that is not true, simply because the goal of this mini Gantt charts was not to perfectly follow the schedule that

I had set but rather have an approximation of time that should be dedicated to each component.

By doing this I tried to avoid getting stuck with a problem and rather go to the next task and set time in the next week to solve this problem. Also when the problem was urgent, another strategy that was used is to adapt to the consequences of spending more time than the planned, which consisted on delaying the next tasks and find alternative solutions that could be done, while that problem was still not fixed.

- **Kanban:** A Kanban board is a method widely used when implementing an agile methodology, which consists in a board that contains tasks divided into three columns (normally):

  1. **TO DO:** In here go all the set tasks and goals that have to be done in this sprint (in my case lasted a week or two). They can be marked with different colors depending on the urgency of the task.
  2. **DOING:** This column serves as a landing spot for all tasks that are currently being developed. They can also be color coded depending on the progress made in the task (red if they are just starting to be done, orange if they are halfway done, and red if there is barely work left to do).
  3. **DONE**: In this column one can keep track of all the tasks that have been finished.

  In order to implement this board I used a website called [Tre]**Trello** that allows one to create a Kanban board.

4.2 Development and Challenges

In this section, the process of creation of each part of the web application will be explained.
Which tasks were defined, how I ended up creating each component, if I came across any setback or challenge that delayed the development of the component or if it even made me avoid that task and find another way around, etc.

4.2.1 Documentation

In the documentation web page, I had barely any problems. This web page will be explained more into details in the corresponding section of results, but it primarily consists on a page that contains all the documents that have been created during the process of developing this Bachelor's Thesis.

The only setback I had to deal with, was the fact that since I had to start this thesis in the beginning of the semester, in early February, and I did not end my semester abroad until March, because I still had exams, I couldn't start on time and had to hurry to deliver the Project Proposal and Initial Work Plan in time. Along with that, the goals of the project were not clearly defined and it was hard to plan that much further ahead. It was because of that, that I decided to follow an agile methodology in order to plan a week at a time and be able to manage better the tasks that I wanted to complete.

4.2.2 Surgical Procedure tool

The surgical Procedure Tool was initially the main and only feature that this thesis was going to have and therefore, lots of effort and time has been put into it. However, it has to be mentioned, that since there were lots of features and components that wanted to be added, with them came also a variety of challenges that I had to face and overcome.

- **Altair Visualization:** This was the original visual tool that I originally thought of when deciding to do this project. Since I had no experience in any other coding language that allowed to create interactive

visualizations, I used Altair, which I had already used in the subject of Information Visualization, taught in my second year of the degree.

The use of Altair had some positive aspects, which were that I had already programmed some complex interactive visualizations, but on another note, there is a limited amount of answers to problems since it is not a really popular choice and one depends on the documentation examples and a couple of answers in some forums in order to solve all problems that may arise.

Some of the problems I encountered during the process are the following:

- In order to embed the visualization into the website, I had to use a third party website, which is called [Dat]Datapane, and embed the visualization as an iFrame.
- The fact that the visualization was an iFrame, made that HTML and JavaScript treated the visualization as a whole component and there was no way to be able to interact with the elements of the Altair visualization, and that JavaScript would recognize it.
- The fact that the heatmap ended up being an iFrame, made this component be bigger than the width of the text that can be written on the web and it appeared a little off.
- Find a color scheme for the visualization and the legend which would be appropriate, according to the color rules of visualizations.
- Make the axis labels the same color as the legend options, so that the users can establish a visual connection.
- Re-order the legend labels without an alphabetical order
- Implement a functionality so that the opacity would be higher when the interception of both legend filters (by phase and tool) was selected (view figure (14).

- **Video Filtering Buttons**: The initial idea was that by clicking the legend of the visualization, the user would then be able to visualize the video of a concrete phase, if he/she clicked the phase legend or a tool likewise. However, since it was not possible, as it has been explained above, I needed to find an alternative way.

  That was when I decided I would create two additional buttons (one to change the phase of the video, and one to explore how is a tool used in an operation), which would change the video that was being visualized, depending on the choice selected.

  A user could select a tool the tool button and by doing so, the video would change to a video which is formed of parts where the video is fast-forwarded (which indicates that the tool is not present at that time of the video) and parts where the video is slowed-down (in order to better observe what are the expected movements and tasks that have to be completed with this tool.

  If the user would want to use the other button (without selecting an option in the tool button), when choosing a phase from this button's options, the video would show only that phase of the operation, in case the focus of the user is only to look at a concrete phase of the operation.

  Furthermore, when the user selects a tool in the tool button and a phase in the corresponding button, then, the video that is being visualized is the one that shows how to use a tool, but only in the concrete phase that has been chosen, in case that this tool is used in that phase.

  When clicking any of those options, it appears a message above the video that informs the user of which video is he/she visualizing and in case two options from both buttons have been selected, it also informs

the user in case a tool is not used in an operation.

There have not been many problems during the development of this component, except for the fact that I had to create the different videos many times, due to different requirements that varied along the process and I had to fulfill. Also the JavaScript function to link a choice to a change of video was not trivial for a person who barely had experience with JavaScript.

- **Video**: The process involving the development of the video included 3 steps which all caused some trouble in a way.
  - The first and main step was to create the videos that would appear after the user clicked an option in either the tool button or the phase button. To do so, I developed a function that creates a new video given the following parameters: the objective (to visualize a tool or to visualize a phase), the option (a tool if it was the tool button that was clicked or a phase otherwise), and the video from which the processed video will be created.

    In this step I encountered some problems, since the original video was really heavy and the function took also a lot of time to create a single processed video, which made this process to go awfully slow, if we take into account the number of tools, number of phases and number of tools used in each phase.

  - The next challenge I had to face was the size of the video, ideally I wanted to have the video as large as possible without having to scroll up and down to be able to visualize it all. Added to that, I also wanted it to be really close to the heatmap so one could really follow the heatmap while watching the video in order to identify in which part of the heatmap was currently the video and the user would then know which tools were appearing on it.

    It was hard to fulfill all of these requests so in the end I ended up making it big enough so visualizing all the steps and details of the video is clear, but small enough so one can fit the video in the screen without the need to scroll.

  - Lastly, I had to battle with the alignment of the video with the heatmap. The reason behind this was the same as the argument of closeness explained above, to be able to follow the progress of the video in the heatmap. Since I was no expert on CSS and HTML this was a little bit more challenging that what it should have been.

- **Progress Bar**: The idea of creating a progress bar came from the fact that the video could not be perfectly aligned nor perfectly joined with the heatmap and it was difficult to follow the advance of the video on the heatmap.

  The heatmap then served as a component to make that function. It has been placed just above the heatmap and it follows the progress of the video filling a bar up with red, as the video advances. By looking at the red progress bar, the user can then explore the part of the heatmap that the video is currently on, as well as the tools that are present and in which phase the video is at that moment.

4.2.3 Video Comparison

In this secondary tool, there were also a few things that needed to be handled in order to made this page of the web application fully functional.

- **New video**: This tool consists on a component that will allow users to compare their videos with videos of complex operations or surgical operations performed by excellent surgeons in a specific surgery. Since

I had no way of obtaining a video of an expert in cholecystectomy and the dataset I was using contained only this kind of operations, I had to use another video from the dataset to exemplify how it would look like when having the appropriate data videos.

With this new video that I decided to use as an expert video example, I had to process all the phase videos so the video could also be filtered by phase like the one we already processed. This took also quite a bit of time to complete.

- **Buttons and Functionalities**: Another important aspect of the development of this tool, was the ability to go directly to a phase. As mentioned before, the new video was processed so it could be filtered by phase. This was important due to the fact that, a surgeon might want to skip directly to a phase where he knows most of the differences will be, or maybe the user interest is to skip some phases because are simply not relevant to see.

  The other functionality that was added to this tool was the ability to pause and play both videos at the same time, so the user could start both videos without the delay of having to click play in the two different videos or stop them at the same time to check a difference that has been made between his video and the one made by an expert. However, both videos still have controls in the video, that allow the user to stop and pause an individual video in case one of them goes too fast or too slow compared to the other one.

4.2.4 Video Annotation

This is the last tool that was developed and it is also the one that brought more complications from the two that were added later on the project. The main goal of this tool is to be able to correct or to write certain comments or annotations at certain instants of time where an error has been made. This way, resident doctors can have his surgical operations corrected with a level of detail that is not available nowadays.

- **Record Annotations**: The first feature that I wanted to add was the ability to record annotations of errors made in operations while watching the videos.
  - There should be a button to create annotations, that by clicking it, a box where the user would write the mistake made would appear and once the annotation has been made, it should disappear. It can also serve as a tool to annotate general or important comments in the videos. After researching a bit, I managed to solve this by defining the visibility of that component to hidden by default and to make it visible once the button "Annotate Video" was clicked.
  - Another feature I wanted to add is the automatic tracking of times in videos for the annotations, so it detects the minute and seconds of the video when the button is clicked and then the video stops and records the time of the annotation in the corresponding input slot in the annotation box.
  - Finally, a confirmation feature was created which allowed the user to not record an annotation if an error in the process of creating the annotation was made. In order to do this, once the annotation has been made, a pop-up window appears, that indicates the user to please confirm that this annotation is correct. In case the user has made an error, the annotation will not get recorded, and otherwise, it will be written in the annotation box below the video.
- **Save Annotations**: As it has just been mentioned, all the annotations get written down in a box below the video so the tutor can follow the progress of which annotations have been made. In this step is when most of the complications arose: Annotations were saved in a JSON-format, however I wanted to print it in a more appealing way and to do so I had to really look for a way around. I ended up trying to print them in a nice table, but in the end, decided to print only the text of each cell since the design of a table

inside a box was not appealing.

The other challenge came when wanting to create a button to download this annotations. After battling to download the annotations in a JSON-format I ended up finding a better solution, which was to convert the JSON data to CSV and to download a CSV that could be easily open with a program like excel or numbers. This way, when downloading the data, the user would receive a table which contained, the time where the annotation had been made and also the message associated with it.

- Lastly, there was another functionality that was added later on which brought little to no complication which was the ability to add a button which would filter the video by phase. This way, the user could skip to certain sections which could be more prone to errors or just more interesting to examine.

  The only problem I encountered with the development of this feature was the fact that the recorded instant of the video had to be the time of the whole video, not the time based on a phase of the video. In other words, the time that had to be recorded was the sum of the instant where the annotation was made in that phase, plus the time the phase started in the whole video.

4.2.5 Deep Learning Algorithm

- **Definition**: The goal of this deep learning algorithm was to allow the computer to recognize the tools that take part in the surgical operation and also to identify the different phases of it. All of the base code was already created in the project carried out in Germany, so the steps I wanted to follow in order to complete the goal I set at the beginning of the project was, first to improve the results of the tool detection and if I saw significant improvement, to extend the capabilities of the algorithm to be able to predict phases of an operation like the algorithm in the paper already mentioned above did [Kag].

- **Results**: Due to some challenges I will explain in the next item of this list, I did not have access to a powerful GPU until an advanced part of the process and therefore I tried to improve the algorithm by using the Kaggle Kernels [Kag] that had been used also while developing the initial code, back in Germany .

  It has to be noted that the the results that will be seen now are only made with a small portion of the dataset and therefore are not representative.

  The goal of this training phase is to tune the parameters in order to obtain a high training accuracy in a dataset with very few data (5 videos), and therefore overfit on a small dataset and then increment the size of the dataset so the percentage of accuracy doesn't go too much down (these results can be observed in Table 4.2.5). As one can see in the table, models 1-3 were trained with respect to this goal. Even though the model seems to achieve better results it's not regularizing well enough on large datasets.

  The results of this experiment that was made are the following:

Model #	# videos	# epochs	learning_rate	batch_size	weight_decay	train_acc	train_loss	val_acc	val_loss	test_acc
1	5	30	1e-4	32	5e-4	98.34%	0.0002409	45.77%	0.01377	46%
2	5	30	5e-5	32	5e-5	100%	5.323e-7	43.06%	0.03911	36%
3	5	30	5e-5	64	1e-5	100%	1.354e-7	46.89%	0.03258	59%
4	37	30	5e-5	64	1e-5	99.61%	3.064e-5	58.99%	0.007504	

Table 3: Models performed when trying to improve the model

- **Challenges**: The first challenge that I had to face was again, like it happened in Germany, that I did not have access to a personal strong GPU to use. I knew that some colleagues had access to a GPU from the telecommunications department, so I planned on using that to improve the results.

  However, when I asked about the availability of it, I was told I had no access, so I had to find another way around. Talking to my tutor, he put me in contact with the people at **RDLab**, who had a GPU but they were not sure if it would be useful to me since it was a little old and I needed lots of computational power. They offered me all kinds of help in setting up the necessary requirements to make it work.

  When I finally obtained the possibility to have a GPU from RDLab, I was in an advanced stage of the development of the whole thesis and was in the middle of completing the user study. After talking to the surgeons about my tool and showing them the functionalities it would have, I received really good feedback. After all this arguments already explained I had to make a decision:

  - I could spend lots of time hyperparameter tuning the algorithm to improve the tool detection algorithm to about 5% or even 10%, making it reach an accuracy of about 70% and then also implement a phase detection algorithm which would also be a little below the percentage that the creators of the original paper from whom we based our algorithm [Twi+16] . Even if this could be achieved, we would then be using the results of it to plot the heatmap (that can be seen in figure 11) in the procedure visual tool of the web application, but this results would contain an error of about 25 or 30% and therefore it would not serve the purpose of being able to have a global view of the tools and phases of the video.
  - The other option, which was the one I had ended up deciding for after hearing all the praises we got, from the initial prototype of surgical video analysis tool that we showed to the surgeons that participated in the user study. This option consisted in spend also lots of time, improving the tool we had showed to the surgeons so it would fit their needs.

    Aside from that, after collecting all of the proposals that they also had, regarding unsolved problems that the surgeons had nowadays in their day-to-day life, I could also spend some time implementing a prototype of how would a tool that solved this problems looked like.

    Based on this thoughts, I decided to develop the already explained Surgical Video Comparison Tool and the Annotation Tool.

  Looking back, I think I ended up having a **more complete and varied web application** which contains lots of features that could be really helpful to solve the real-life problems of surgeons, since I confirmed this in the completion of the user study that had been made.

5 Results

5.1 Use Cases

All of the following examples of use case examples, will be detailed in the Annex, since there was no more space in this document.

Each of these items will have a detailed description of a step-by-step list on how to perform each of these tasks with the web application that I have developed in this project. Also they will come accompanied by pictures making it really illustrative so one can really see how does this web application works in practice.

- **Get information about the surgery performed in the videos**
- **Get information on the dataset used: Cholec80**
- **Learn how to use each visual tool present in this web application**
- **How to contact us**
- **Explore how a tool is used in an operation**
- **Explore how to perform a concrete phase of an operation**
- **Explore how a tool is used inside a concrete phase of an operation**
- **Contextualize when and how much a tool is used in an operation**
- **Contextualize how long does a concrete phase of an operation lasts and which tools are used in it**
- **Research if a tool is used in a concrete phase of an operation**
- **Annotate mistakes made in an operation**
- **Annotate mistakes made in an operation in a concrete phase**
- **Download file with mistakes and send it to resident doctor**
- **Explore differences between a video of a model operation or a surgery performed by an expert surgeon and the user's**
- **Explore differences between two video in a concrete phase**
- **Read and get information about the evolution of the thesis**

5.2 Structure

The whole project has been created inside a hand-made website, which has been improving every time since day one. However, now that the final structure has been set, it will be shown in the following diagram:

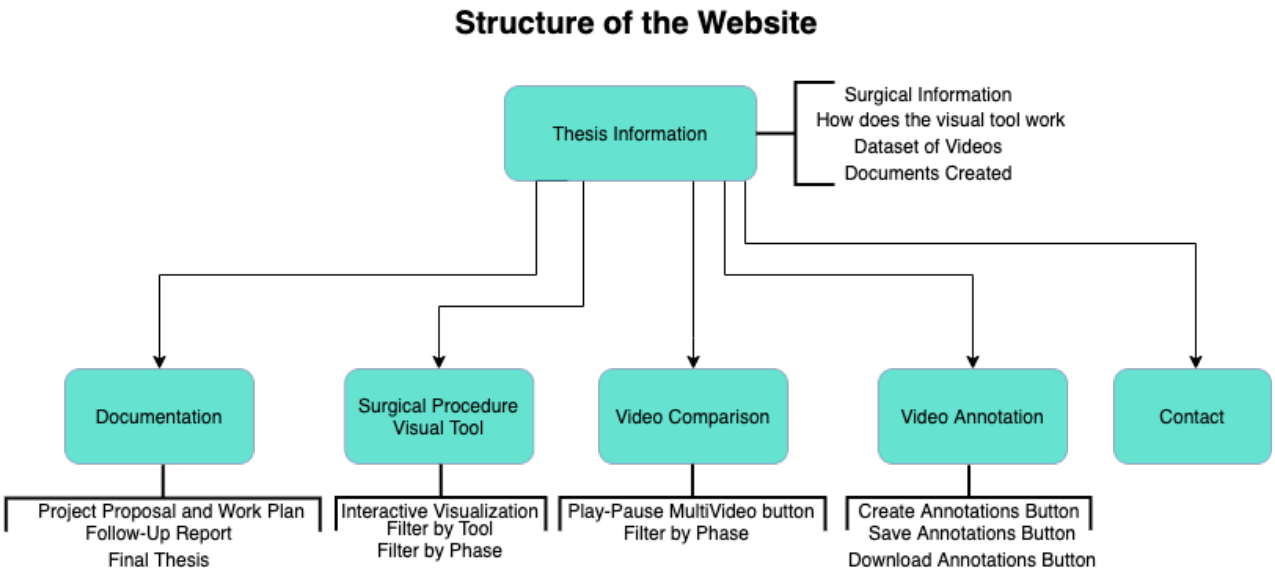

Figure 6: Structure of the thesis website

As it can be seen in the figure (6), there is a main or home page, which contains an overview of the necessary information that a user would need to know to fully understand all the capabilities of the tools in the website. And then there are different pages for each tool or other functionalities.

5.2.1 Home Page - Thesis Information

Like we have already mentioned above, the purpose of this web application page is to give all the information in order to let a user without previous knowledge, run all the tools present in the web and get an idea of how can this be extrapolated to real-life scenarios.

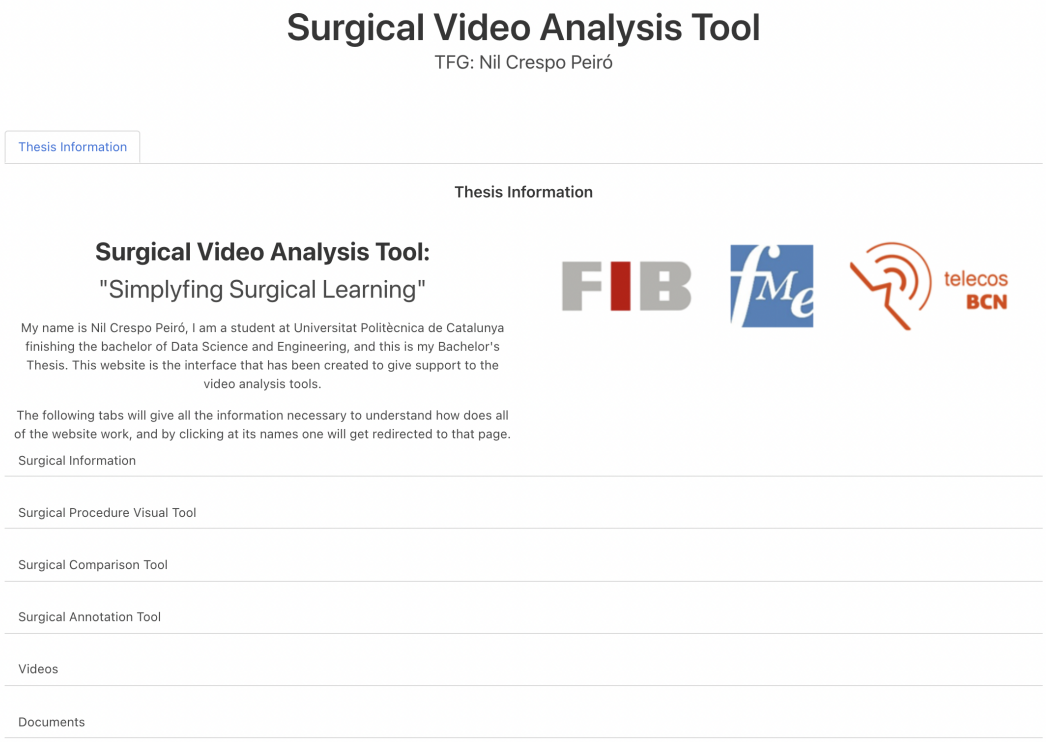

Figure 7: View of the main page of the web application

As it can be seen in the image above (figure 7), all this information is stored in tabs that can be minimized or maximized upon clicking at them, so it does not take too much space in the initial home page.

- The first tab contains an **introduction** of the idea that this web application has been created in order to have a user interface to display the visual tools that were going to assist the process of surgical learning. It indicates, who has created this Bachelor Thesis, where has been this thesis created(under which tutoring, faculties, university, etc.).
- The next tab serves as an **information tab** that gives all the necessary information to the user, regarding the surgical specifications: which operation will be seen in the videos, which method has been used to perform this surgical operation, how does this method work, if it is common, how has the video been recorded, etc.
- Next, there is **information in the following tab about the main tool developed, and the two posterior tools** that were created after conducting the user study. For each tool there is an explanation of how to use the tool and what functionality does each element of that page have. All of them are accompanied by images, supporting the description.
- Then, another important thing is shown, which is the explanation of how have I gotten this **dataset of videos**, who has created this dataset, under which legislation, who are the authors of the paper who first mentions this dataset, etc.
- Finally, there is a last tab that mentions that during the process of creating this thesis, some **documents** have been created, which show the progress that has been made from the original work plan and goals,

to the follow-up report that consisted on evaluating these goals and modifying them based on how have each of them developed, and also mentioning the goals that were left to do until the presentation of this document. Also, this document will be available there, in case there is interest in seeing it.

5.2.2 Documentation Page

Initially, this page was created for me to have all the documents in one place, so I could re-visit the goals that I had set for the project and how was I doing time-related. This is important, since as it has been explained in section (4.1), I established an agile method/procedure and therefore, it was critical to be able to evolve and adapt the goals depending on the problems that I encountered.
However, then I found it was a good idea to have all this documents public in a page, so it could be seen the progress that has been made during the process of four months that this project has lasted.

Figure 8: View of the documentation page in the web application

By looking at the figure above, it can be seen that this web application page consists on three documents:

1. **Project Proposal and Work Plan:** This document served as an initial outline of the work that had

to be done during the process of writing this thesis.

- Firstly, a short **summary of the project** had been written, followed up by the main goals of the thesis.
- Then a **temporal schedule** was written (of how the tasks and milestones were going to be distributed along the months that the project took place) using a GANTT Diagram and also a **schedule for the meetings** with my tutor, **Pere-Pau Vázquez,** who is in charge of this thesis.
- Lastly, the **generic skills** that will be used in this project will be checked in the appropriate table

2. **Follow-Up Report:** In this document it was necessary to record how the project evolved, based on the project proposal and initial work plan and to update the work plan to the current goals at that moment.
   - Firstly, the **initial goals** that were initially proposed were **analyzed and redefined** according to the problems encountered during the process of trying to achieve the set goals.
   - Then an **updated temporal schedule** was written bsaed on what has been explained in the last item.
   - Lastly, there was an enumeration of the **work that was left to do** before the final review.
3. **Final Thesis:** The goal of this document is to give a clear overview of the project, and to explain every part of work that has been done. Like all thesis it consists of a skeleton formed by what has been explained in section (1.3).

5.2.3 Surgical Procedure Visual Tool

This is the main page of the website, it contains the initial visual tool that was created and it has all the functionalities that were defined in the initial work plan.

It consists on three components that are linked:

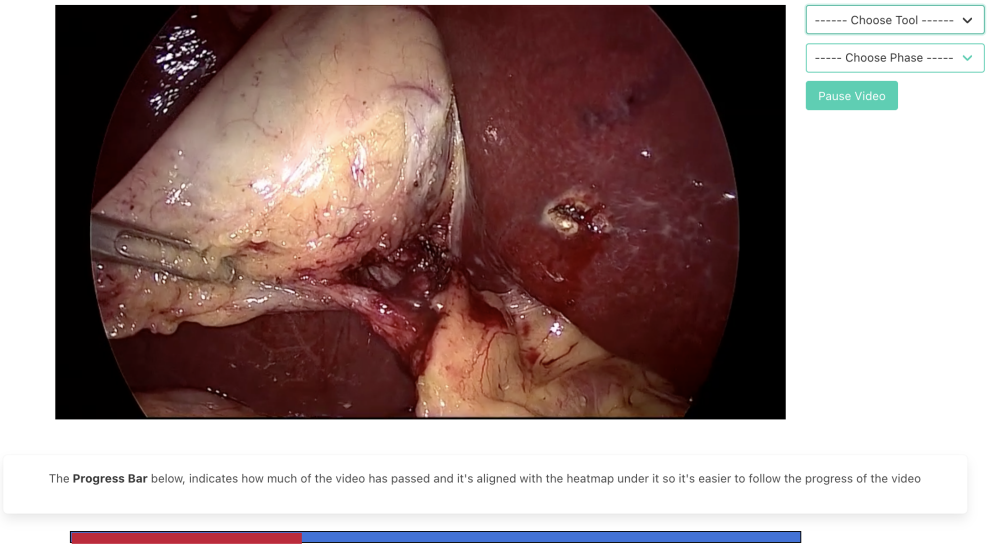

Figure 9: View of the Visual Procedure Tool

- First, there is a large video corresponding to one of the Cholec80 dataset videos [CAM]. This video can be reproduced by either a button on the right of the video that says: **"Play Video"** or the controls at the video itself. One can observe that when clicking this button, the **progress bar** below it fills up, showing the progress of the video in a more intuitive way and it is also aligned with the heatmap under it so it's easier to follow the advance of the video. This way, the user can visualize the whole surgical operation from start to finish.

  Another feature that this progress bar contains is the ability to be clicked in any part of the bar, and by doing this, the video gets skipped to the corresponding moment of the video. This is very useful in the sense that the user can then explore with the heatmap below, which parts could be of interest and then click the bar to go to that instant and not having to search for it in the video, which would be a more tedious process.

  On the right of the video there are also two additional buttons, aside from the one that allows to play and pause the video. This two buttons are the main feature of this tool, and is what makes it unique from other software with similar functionalities.

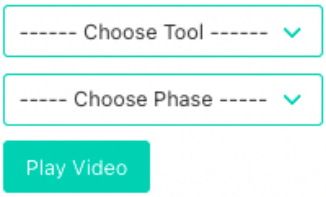

Figure 10: Buttons from the Visual Procedure Tool

The first button that is named **"Choose Tool"**, allows the user to select a tool (as its name indicates) that he/she has interest on remembering his use in this surgical operation. When choosing a tool, the video that was on screen before, which was the whole video, will disappear and in its place a new video will appear, which will consists on a video that is composed of parts where the chosen tool is not present (where the video will be fast-forwarded), and parts where it is present (and the video will be slowed-down).

When choosing a tool, a message above the video will appear, indicating which video are we watching if it is slowed-down or fast-forwarded depending on some tool or not, for example.

The button below, where **"Choose Phase"** can be seen, it follows the same structure as the one aforementioned, but in this case, the user can select a phase that he/she is interested in and only visualize this part of the video. It also appears a comment above which indicates the phase that the user is visualizing.

Lastly, when choosing a tool in the above button and a phase in the button under it, the user can then visualize how is this tool used in a specific phase (in case that the tool is used in that phase, otherwise a message will pop up which asks the user to change either the tool or the phase and choose so using the heatmap below, to decide which tool or phase should be changed).

There will also appear a message above the video that the user is viewing a fast-fowarded video when the selected tool does not appear in the selected phase of the video and slowed-down when it is.

- The other main component of this page is the **heatmap**. This visualization serves as an exploration tool of the surgical video.

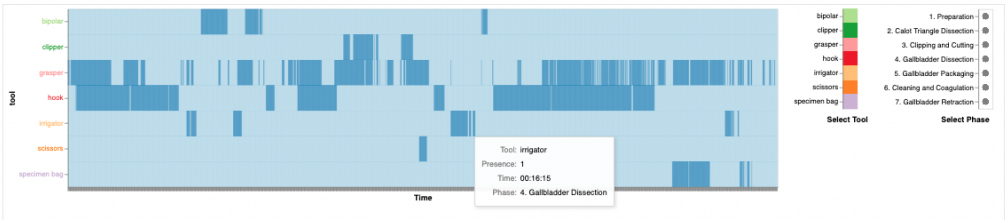

Figure 11: Heatmap in the Visual Procedure Tool

In general traits, the heatmap has a **dark blue color if that tool appears in that instance of the video and light blue if it does not appear**.

This heatmap is **interactive**, which means that the user can hover over the visualization and a message will appear that shows the name of the tool, phase and if it is present at that time of the video or not.

This is clear, when looking at figure (11).

It also has interactive legends, which allow the users to select a tool (figure 12) and preview only when does that tool appear in the video.

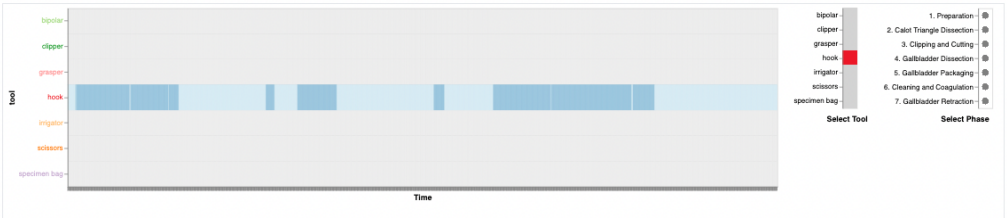

Figure 12: Heatmap, when filtered by tool

It also allows selection of the phase (figure 13) to understand how much it lasts and which tools are present in each phase.

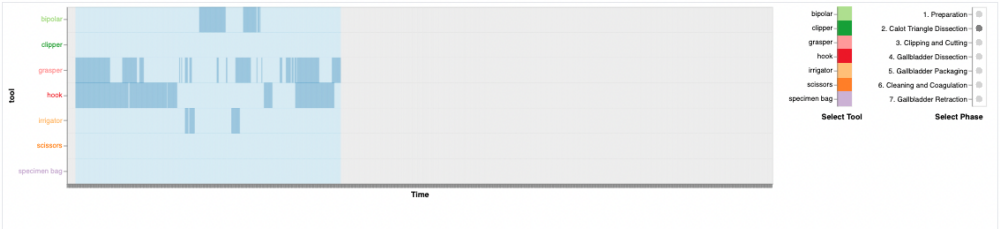

Figure 13: Heatmap, when filtered by phase

If the user would want to select both tool and phase it would be visible with more opacity the intersection

of both filters (if a tool appears in a phase). It can be seen in figure (14

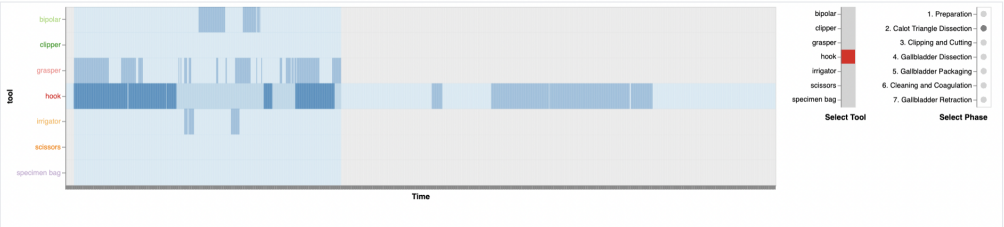

Figure 14: Heatmap, when filtered by phase and tool together

To sum up, the purpose of the main tool is to facilitate the learning process by providing enhanced navigation methods for surgery videos. To this end, we provide some interactive visualization widgets, like this heatmap which allow the user to have a **global view of the video** that is playing above; how long does a phase last in comparison of the whole video, which parts are being slowed-down or fast-forwarded, in which phases does each tool appear, etc.

- Lastly, there is a small component in the bottom which are these **"cards"** (figure 15), which has the look of a video selection tool. However, this feature is not currently implemented.

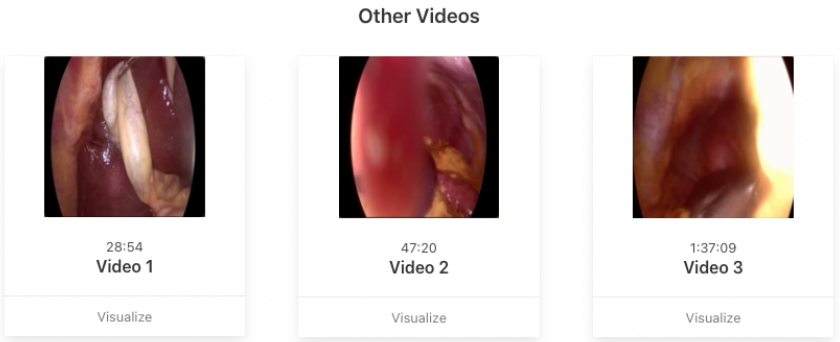

Figure 15: Cards, a video selection feature

5.2.4 Video Comparison Tool

This tool, as well as the (5.2.5) Video Annotation Tool, have been developed after doing the user study. We found that there two other needs, described by the study participants, that could complement the main goal of our tool.
These next two functionalities aim to solve these needs, that surgeon apprentices, resident doctors and surgeons have nowadays in real life.
During the user study, we found out that the Cholecystectomy operation is a very common and basic surgical operation and therefore, the experienced surgeons would not use this tool because they had no need to review parts of the surgery because they have done this kind of operation multiple times and they already know how to perform a good Cholecystectomy.

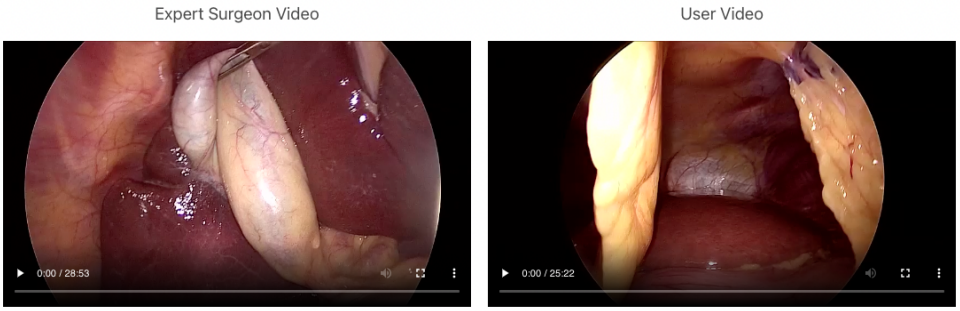

Figure 16: View of the Video Comparison Tool, from the web application

However several of the participant surgeons commented that this tool would be of use if another more complex operation was being visualized. From this idea, I decided to create this tool, which consists on two videos side by side that have a play/pause button that controls both videos at the same time, but the user can also stop and play any of the two videos individually by using the video controls. This is useful when one of the two videos is taking longer in some parts, so this way one can stop the other video and wait for the slow one to "catch up".
There is only a feature (in figure 17) which allows to select a phase and this way both videos skip to that phase, which could be the phase where it is more technically difficult or the user has special interest on.

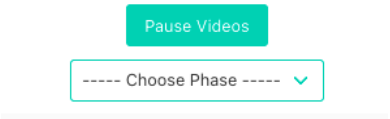

Figure 17: Main feature in the Video Comparison Tool

The main goal of this tool is to compare the differences between the video the user will input and a video of a complex surgery or a video created by an excellent professional in a really concrete surgical operation. This video comparison tool can be of use to 2 different groups of persons:

- For experienced surgeons, as it has been previously mentioned, if on the left video, a video of a difficult surgical operation or an operation from an expert surgeon in a field would be visualized, while on the right the user could place his operation it would be very helpful to see what has been done differently, and what techniques this user can apply to his surgeries in order to improve his performance.
- This tool can also be used for resident doctors, with the goal of making the learning progress of this apprentices more visual and easy. In order to be useful to them, on the left hand side of the screen, there should be a video which could be a model made by the hospital of how to perform a Cholecystectomy in this case. This way, all resident doctors would have a model operation from which they would want to learn and compare their work with it.

5.2.5 Video Annotation Tool

As it has been mentioned above, this is the third tool created, after gathering ideas/problems that the surgeons have nowadays.

After discussing the methods of evaluation of resident doctors with a surgeon from the user study, we got into the conclusion that nowadays, the evaluation of resident doctors is complicated. To illustrate one of the problems is that, for example, the tutors can not be present to every operation that the resident doctors do. Based on this, an idea to try to solve this challenge could be the following:

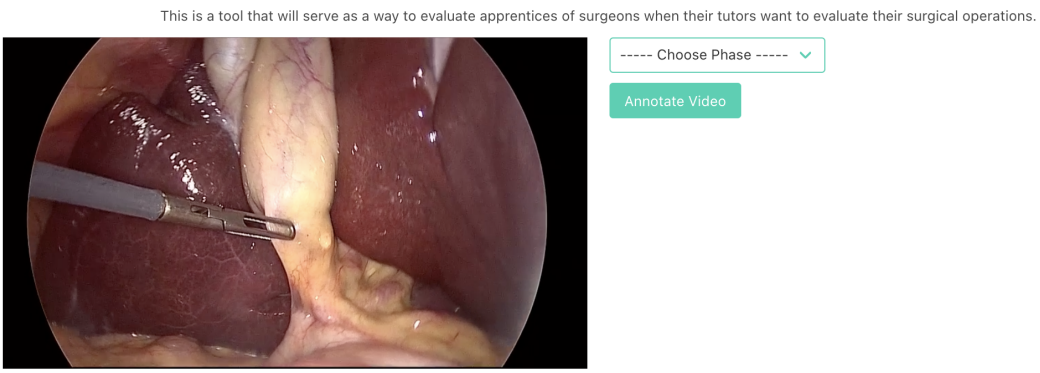

Figure 18: View of the Video Annotation Tool, from the web application

If the resident doctor could record their operation and send it to his/her tutor, the tutor could evaluate it later on, without needing to be present. This would be done by annotating at which instants a mistake has been made, download this file and then send it to the resident doctor so he/she will be able to know which errors has he/she committed and when.
Based on this, the tool has been developed and works in the following way:

- The video is visualized on the left by the tutor and he/she can choose whether to watch the whole video or skip directly to a concrete phase that could be more interesting or more prone to errors.
- On the right of the video there is a button that says: "Annotate Video" (figure 19). When clicking this button, the video stops and it automatically records the timestamp of the video and let's the tutor add an a comment, explaining what mistake has been done. Once finished, the tutor can click the button "Save Annotation", to save the annotations and keep watching the video. When doing this, a pop-up window appears, asking the user to confirm or cancel this annotation.

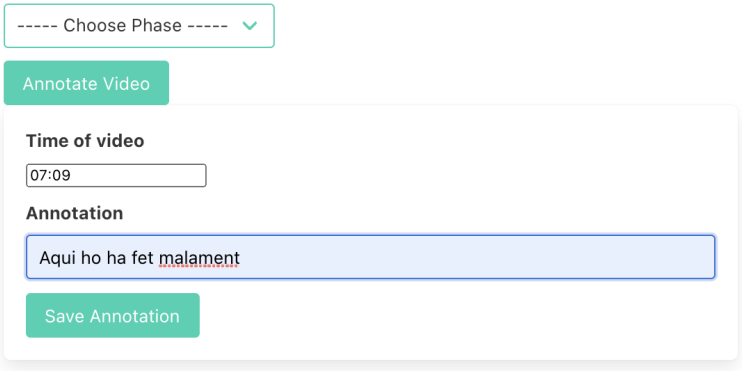

Figure 19: Buttons that allow the creation of annotations for the visualized video

- Under the video, it appears a box (figure 20), where all the annotations will be written after the user has confirmed each annotation and then there is also a button, that allows the user to download this annotations in a CSV-format, that can be sent to the resident doctor, so he/she can analyze where mistakes have been made an understand the mark he/she has received from the tutor in a totally objective way.

Figure 20: View of how are the annotations saved

5.2.6 Contact

This page can only be accessed through the home page, and it is simply a form that could be used in the future to receive comments from people who want to learn more about this project, who want to keep developing it, or has ideas to improve it or complains about it.

It is a form composed of different parts:

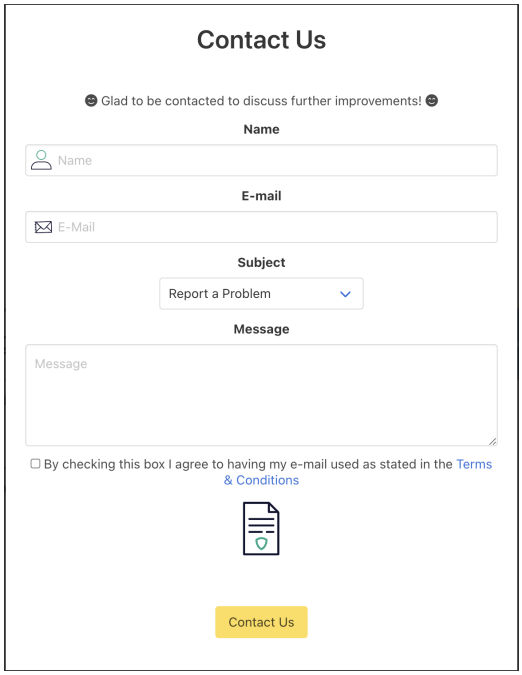

Figure 21: View of the Contact Pop-Up page, from the web application

The initial objective was, that when a user would fill up this form a notification or an e-mail would be sent to me, with all the information that this form contained.
However this was not possible due to the fact that in order to do so, I would have needed to have knowledge on backend components like PHP, and thought it was not worth the trouble.

Even if that was not implemented, what it has been successfully developed is the ability to open the user's mail with all of the data inputted in the form so the user can send me a mail with all the corresponding information.

This workaround ended up being really convenient as it solved the intended task without the need of spending time learning a whole new technology that I had no experience with.

6 Evaluation - User study

While doing the project, a disturbing thought came into mind when I was reflecting about the real application in life.
If it would be useful to someone, how could this tool help the final user, if I was really solving a real-life problem or not, etc..

So, with this in mind, I decided to get in contact with several surgeons, who gave me their insight on how this tool could be used, if it would help them and how could I improve the tool to make it even more useful.

As we will see, the evaluations we got from the participants were highly useful and helped us to design a set of extensions that solve some further problems.

6.1 User Study

Throughout the development of the project, we have performed an informal user study with several surgeons in order to analyze the features of the application together with the usability.
This informal study has been done in a continuous way, in the meaning that different versions were shown to the participants, although the main features were already present in the first demonstration.

6.2 Development

Initially, since I had no contact nor relationship with any surgeon or doctor of any hospital, I decided to reach out to **Prof. Jordi Vitrià**, who is a professor in the University of Barcelona, more concretely in the Mathematics and Statistics Department. Since he leads the Data Science group there, I hoped he could redirect me to some professionals in the field of Surgical Operations.

He indeed put me in contact with **Dr. Carolina Malagelada**, who worked in the Vall d'Hebron Hospital in Barcelona. She listened to my demonstration of the surgical tool and told me that even though she found it really interesting she admitted that she did not perform this kind of operations and that she would provide me of the contact of a surgeon that could really answer my questions.

Through Dr. Malagelada I contacted **Dr. Marc Martí** who, after evaluating the tool, gave us some interesting insights and also provided us with the contacts of **Dr. Elizabeth Pando, Dr. Laia Blanco, and Dr. Cristina Dopazo** who were more specialized in cholecystectomy Operations.
Since the physicians agreed that the surgery students could be a good target audience for our application, we also obtained the contacts of 2 resident doctors (**María Martinez** and **Naïr Fernandes**), who also evaluated our tool.

The contact limitations in Catalonia caused by the COVID pandemic, we had to perform all the studies in remote.

6.3 User Profile

So as I mentioned, 6 users participated in the evaluation (1 male, 5 females). From them, four were senior surgeons (between 10-17 years of experience) and 2 were resident doctors (in initial years of specialization).

6.4 Procedure

The study was carried out the following way: after describing the objectives, we introduced the tool with a brief demonstration and asked the users whether they had understood how the tool works, and whether they had any doubts.

After that, we issued a small questionnaire with 7 questions that the users had to answer in a 7-point Likert scale, where 1 means completely disagree and 7 means completely agree. The questions were:

1. **Do you think the tool can be used to help in the learning/teaching process?**
2. **Do you think the tool can be useful for the work of a surgeon?**
3. **Do you think the tool is easy to use?**
4. **Do you think one can easily learn how to use the tool?**
5. **Do you think the tool covers a need that was not yet covered for any software?**
6. **Do you think the tool is more useful that the current software?**
7. **Do you think you would use this system frequently?**

We also asked them the following open-ended questions:

- **What would you add or remove from the tool to make it more useful to surgeons and students?**
- **Do you think this tool could serve any other purpose that I have not mentioned?**

The evaluation lasted from 30 minutes to 1 hour.

6.5 Results

The table below indicates the results of each surgeon and apprentice of surgeon we interviewed and their corresponding results to the questionnaire, we've already explained.

Table 4: Results of Questionnaire

	Dr. Martí	**Dra. Pando**	**Dra. Blanco**	**Dra. Dopazo**	**Maria Martínez**	**Naïr Fernandes**	**Average**
Q #1	6	5	5	7	7	7	**6,167**
Q #2	5	6	4	4	5	5	**4,833**
Q #3	7	5	7	7	7	7	**6,667**
Q #4	6	5	7	7	7	7	**6,500**
Q #5	5	4	3	7	7	7	**5,500**
Q #6	6	2	6	-	7	7	**5,600**
Q #7	2	3	5	7	6	7	**5**

From this table, we can visualize it more clearly with the following visualization:

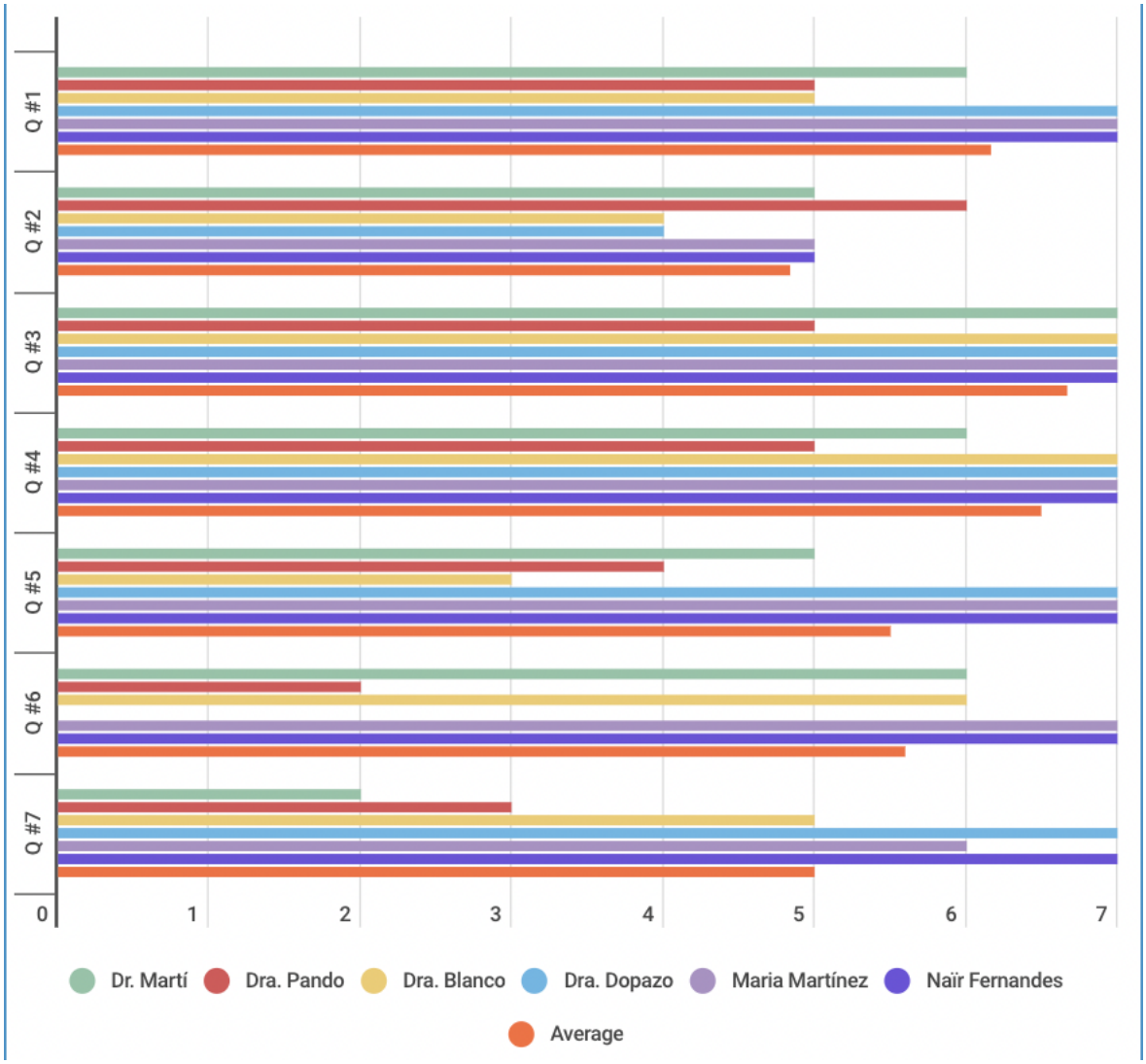

Figure 22: Individual responses to the questionnaire

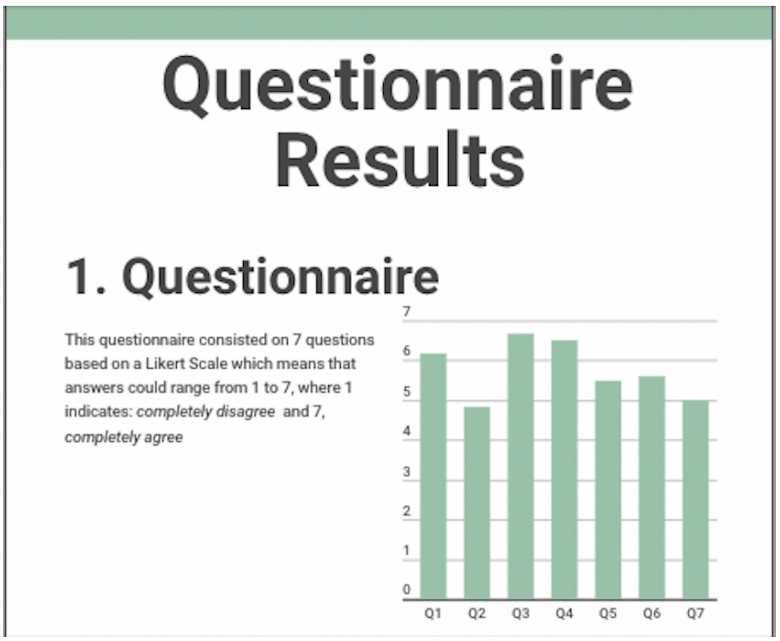

Figure 23: Analysis of Questionnaire results

We can clearly see that most of the answers were over 5 out of 7 in the Likert Scale, which indicate a good response from the surgeons.

Also, the open-ended questions were really helpful and will be discussed in the discussion section.

6.6 Discussion

The discussion with the physicians also gave us very interesting ideas for further development.

One of the things that we discovered from these interviews is that the core idea of our application could be used in three scenarios. In order to do so, we need to add some features to our initial tool.

For example, one surgeon suggested that for experienced surgeons our original tool was not useful, but if it could do the same with other more complex operations, or operations carried out by an expert of a field, then it could be interesting to see the differences between their operation and the one made by this expert surgeon. This **has been implemented** by putting two videos side by side with a button that allows the playing and stopping at the same time, and also the filtering by phase, so it is easier for the user to compare an specific part of the operation.

Another doctor suggested that a variation of the tool, with the adequate elements to evaluate surgeries carried out by apprentices/resident doctors could be very valuable. It would work as follows:

- First, the system should allow the students to upload a small set of videos (around 5 according to the interviewed surgeons). The tool might also offer some editing features, to label the important parts of the operation.
- Then, in a second step, the tutors would be able to inspect and grade the videos. This inspection would work very similarly to what we have presented up to now.

This has **also been implemented**, by creating a button next to the video that the tutor will visualize, which allows the annotation of messages/comments that indicate where the student has made a mistake. This comments get recorded in the format (Time: Annotation Message), so that when the annotation process is over, one can see in a visualization at which times an annotation has been made and go directly to that time of the video.

The goal of this tool would be to facilitate the evaluation of resident surgeons in a more objective and numeric way that it is done now. As we commented before, this extension of our tool has also been carried out.

Besides that, other potential lines for future development were also pointed out, such as:

- Adding some extra labels/text that indicate meta data, such as the position the patient is in (French way, American way, Spread out legs...).es un
- Indication of the hand the surgeon is grabbing the tool with, which would be really helpful for teaching/learning.
- The inclusion of a filtering method that would allow the user to filter the video by wrong moves performed in an operation and that the user could go directly to those instants of the video.
- Be able to combine this with a virtual simulator of cholecystectomy for the apprentices of surgeons.
- Adding overlaid information regarding the anatomical regions, to facilitate a quick identification.

7 Conclusions and Future Developments

7.1 Conclusions

In this section, the goals of the project will be evaluated and then a final comment about the overall process of conducting this thesis will also be made.

7.1.1 Evaluation of Goals

As one can observe in section 2, two main goals were defined, aside from a complementary one, and two other that were created further along the project.

The first goal was to be able to **create a user interface** which allowed the placement of the video analysis tool that I was going to develop.
This goal was clearly met, since without it, none of the project would have been able to be created. The user interface that had been created was a web application, coded from scratch, which contains all the features that were asked in the initial goals (which were that this interface should support the placement of the video analysis tool).

The other main goal, was to actually **develop the visual tool**. This goal has also been successfully completed through the union of three components: The video of the surgical operation, the tools that allow the filtering of the video by tool and phase, and lastly the heatmap developed with Altair, that allows the user to have a global vision of the operation.

After evaluating the first two goals, there was a complementary goal that was defined which consisted on **improving the deep learning algorithm** in order to get better results and extend the features of the tool, so it could also detect the phases of the surgical video.
As it has been explained in section 4.2.5, due to some reasons it was not possible to fully complete this goal. One of this reasons was that it was difficult to get access to a GPU-based cluster. However, this setback is independent of the work done and has not affected the development of any of the other goals or tasks in the project.

After completing the user study two other goals were defined, which were the **creation of a surgical video comparison and annotation tool**. This two goals were two goals that directly solved problems that some surgeons or resident doctors have nowadays in their day-to-day life, so it's safe to say, that the successfully development of this two tools could have a major impact on some of this surgeons or resident doctors.
These two tools have been successfully developed and integrated to the web application with all the features that were initially thought of in section 2.

7.1.2 Final Remarks

After finalizing this whole process of creation a Bachelor Thesis, I can gladly say that it has been a fullfilling experience, due to the fact that I have ended up developing a software, which could be of use to so many people in the field of surgical learning.

In this project I have learned to organize myself according to an initial plan and also be able to adapt and overcome all the challenges that I faced during the creation of every component of the web application and the thesis in general.

In the end I ended up very satisfied with the work done, the final software created and I hope this can lead to major improvements in the field of surgical learning, in Catalonia and more specifically in the Vall D'Hebrón hospital, from where I got all the help needed to conduct the user study.

7.2 Future Developments

Even though I finally ended up implementing many features to the web application and also a user study has been made, some tasks were not completed due to a variety of reasons; from lack of time to lack of resources, aside from other aspects.

However, I decided to keep track of them, so it could be shown the amount of work that can still be done in this direction:

- Improve the deep learning algorithm with a powerful GPU and try to improve the results gotten with the Twinanda et al. paper[Twi+16] add the functionality to predict the phases of the videos.
- Visually compare the results of the deep learning algorithm to visualize where has the algorithm made mistakes in the prediction of tools or phases.
- Improve the video comparison tool, so it allows a fast forward method or the ability to annotate, differences between both videos. By doing this, the two tools that are created would be merged together (however in this case it has not been done only for the sake of clarity).
- The addition of a feature which would create subtitles for the video from the predictions made by the deep learning algorithm, which would indicates which tool appears or in which phase is the video currently in, so it's easier to follow.
- Develop an improved video annotation tool, so it allows the user to upload their video and once all annotations are made it shows a heatmap made in d3 (since its components can be detected in JavaScript) which allows the user to go directly to the instants of the video where an error has been committed.
- Upload the website to a web hosting service so it is available for everyone who is interested in it.

It is also important to note all of the further developments that were extracted from the user study, which have been mentioned in section 6.6

8 References

Papers

[Twi+16] Andru P. Twinanda et al. "EndoNet: A Deep Architecture for Recognition Tasks on Laparoscopic Videos". In: (2016). URL: https://arxiv.org/pdf/1602.03012.pdf.

Universities

[Str] CHRU - Strasbourg. *Universiy Hospital of Strasbourg*. URL: http://www.chru-strasbourg.fr/.

[TUD] TUD. *Technische Universität Darmstadt*. URL: https://www.tu-darmstadt.de/.

[UPC] UPC. *Universitat Politècnica de Catalunya*. URL: https://www.upc.edu/ca.

Websites and Tutorials

[IRC00] IRCAD. *Websurg, Online University*. 2000. URL: https://www.websurg.com/en/dashboard/.

[CAM] CAMMA. *Cholec80 Dataset*. URL: http://www.camma.u-strasbg.fr/datasets.

[Cod] Visual Studio Code. *Visual Studio Code*. URL: https://www.code.visualstudio.com/.

[Dat] Datapane. *Connect data science insights with end users*. URL: https://www.datapane.com/.

[fre] freeCodeCamp.org. *Data Visualization with D3.js - Full Tutorial Course*. URL: https://www.youtube.com/watch?v=_8V5o2UHG0E&t=232s.

[Goo] Google. *Google Collaboratory*. URL: https://www.colab.research.google.com/notebooks/.

[Kag] Kaggle. *Kaggle*. URL: https://www.kaggle.com/.

[Nin] The Net Ninja. *Bulma CSS Tutorial*. URL: https://www.youtube.com/playlist?list=PL4cUxeGkcC9iXItWKbaQxc

[Ove] Overleaf. *LaTeX, Evolved*. URL: https://www.overleaf.com/.

[Staa] StackExchange. *LaTeX Stack Exchange is a question and answer site for users of TeX, LaTeX, ConTeXt, and related typesetting systems*. URL: https://www.tex.stackexchange.com/.

[Stab] StackOverflow. *Stack Overflow is a question and answer site for professional and enthusiast programmers*. URL: https://www.stackoverflow.com/.

[Tre] Trello. *Trello helps teams move work forward*. URL: https://www.trello.com/.