Objective Surgical Skill Assessment of Computer-Aided Hysterectomy Procedures Documentation

Elif Bilgin (ebilgin1@jhu.edu)

Github: https://git.lcsr.jhu.edu/los-warmuphysterectomy/colpotomy_skill_assessment

May 9th, 2019

Introduction

This is the complete documentation of the code and data for this project. In the project, you will find the description of the directories, scripts and functions in detail.

Contents

1 Dependencies

The code utilizes Python 3, and requires the following packages:

- matplotlib
- numpy
- itertools
- os
- importlib
- scipy
- pandas
- math
- sklearn
- pickle

Code was run on the JHU compute server, thin6. All code was saved on the official project repository.

2 Input Data

The input data was from the warmup_data server, located at: `smb://10.162.34.158/warmup-data/OR/Data/processed/`.

The server contains the following data files:

- Label files: Video task labels, denoting what part of the hysterectomy procedure is being done at a specific timestamp.
- Motion files: Contains motion data files from the PSM1, PSM2, and PSM3 (if applicable), as well as event information based on timestamp. Event information gives insight into which tool is turned on/off in the system at a specific time. This file was used for energy usage calculations, as will be noted in the following sections.
- Video files: Contains video footage of the surgeries. This footage was used to verify user changes.

3 Scripts

3.1 main.py

This script calls some functions in the plotScript.py, and can be used to generate colpotomy specific cropped video footage,

energy usage data, and plots of the motion data extracted. The person running this script is allowed to specify which outputs they want by setting the corresponding flag to True or False.

3.2 plotScript.py

This script generates plots for all motion data points for the colpotomy step. It was expected to be indicative of skill of the surgeon, but was later on discarded as we found it was not as useful as expected. It calls plotVideo.py to generate a video of the motion data plots.

3.3 plotVideo.py

This is a script that takes in the specific motion data in XYZ coordinates and corresponding time stamps, and plots them based on time. It generates a GIF image that shows a video of the plot. This is then returned to main.py for output.

3.4 userChange.py

This script is used to go through surgery specific user change data, and extract the user changes that only correspond to colpotomy steps. These are then outputted as text files, as will be discussed in Section 4, Output Data.

3.5 extract_Colp.py

This script goes through each surgery's user change data that is outputted by the userChange.py script, and extracts all motion data corresponting to the colpotomy step, along with the timestamp. The motion data is in XYZ coordinates, and the timestamp is in seconds. It allows for the option to run the script on an individual specified surgery instance, or all surgeries in the data directory.

3.6 videoCrop.py

This script extracts video files of the colpotomy step from the complete surgery video file. If there are multiple instances of surgeons returning to the colpotomy step, it extracts them as seperate files, and numbers them accordingly.

3.7 master.py

This is the master script that goes through each surgery's user change data, and calculates the statistics (duration, path length,

energy usage) that are used as features for classification, and builds the dataset. Details of these features can be found in Section 5, Dataset.

3.8 prediction.py

This is the initial script used for classification. It splits the data into training and testing sets, trains the 5 different classification models we have on the training set, and makes classifications on the test set.

3.9 prediction_sweep.py

This is the final script used for classification. It performs a parameter sweep by splitting the data into three sets: training, testing and validation sets. It then trains the models on the training set, and then makes a classification using the testing and validation sets. It also calculated the % accuracy of these models.

This script was also used to determine how the performance of each model varies according to the number of folds.

4 Output Data

4.1 Animated_Numbered

This is the name of the directory containing the animated plots of motion data, numbered according to the instance of colpotomy steps. Each GIF is under the directory named after the surgery date/time.

4.2 Colp_Motion_PSM1

This is the name of the directory containing the motion data files solely of the colpotomy step, for each surgery, collected from PSM1. Each data file is found under the directory named after the surgery date/time. Each row denotes the timestamp and the XYZ coordinates of the tool.

4.3 Colp_Motion_PSM2

This is the name of the directory containing the motion data files solely of the colpotomy step, for each surgery, collected from PSM2. Each data file is found under the directory named after the surgery date/time. Each row denotes the timestamp and the XYZ coordinates of the tool.

4.4 Cropped_Numbered

This is the directory that contains the output of the video cropping script, which contains the video footage of each instance of the colpotomy step for each surgery. Each video file is found under the directory named after the surgery date/time, and is numbered according to the colpotomy instance within the surgery.

4.5 Energy_Usage

This is the directory that contains the output of main.py when run with the energy usage flag. Each surgery directory contains 5 files, corresponding to the 4 tools (mono_cut, mono_coag, bipolar_cut, bipolar_coag), and the total energy usage. Each row denotes a start and stop timestamp for energy activation.

4.6 User_Change

This is the directory that has the user change files for each surgery, for the duration of the colpotomy steps. Each row has a start and end timestamp, letter indicating the user (A-attending, R-resident, F-fellow) and a number, indicating which instance of colpotomy this is in reference to.

5 Dataset

The final dataset is a .csv file in which each row represents the data point number, surgery name (based on date/time of surgery), user (0 or 1 for expert vs novice), colpotomy instance within the surgery, total duration of the user, specific total durations for the 4 energy tools, total duration of energy usage, total Euclidean distance travelled in this interval for PSM1, specific path lengths for 4 energy tools for PSM1, total Euclidean distance travelled in this interval for PSM2, specific path lengths for 4 energy tools for PSM2, and the total counts for activation of each energy tool during this interval. Each row denotes a different user interval in a surgery, per colpotomy instance.