Multi-Class Detection of Laparoscopic Instruments for the Intelligent Box-Trainer System Using Faster R-CNN Architecture
Fatemeh Rashidi Fathabadi  
*Department of Electrical & Computer Engineering*  
*Western Michigan University*  
*Michigan, USA*

Janos L. Grantner  
*Department of Electrical & Computer Engineering*  
*Western Michigan University*  
*Michigan, USA*

Saad A. Shebrain  
*Department of Surgery, Homer Stryker M.D. School of Medicine*  
*Western Michigan University*  
*Michigan, USA*

Ikhlas Abdel-Qader  
*Department of Electrical & Computer Engineering*  
*Western Michigan University*  
*Michigan, USA*
Agenda

Introduction
  • Literature Review
  • Object detection

Methodology
  • Procedure for Laparoscopic Instrument Detection
  • Data Collection
  • Data Preparation
  • PyTorch
  • Choosing a Model
  • Training the Model
  • Predictions

Results
  • Detections of laparoscopic surgical instruments

Summary and Future Research

SAMI 2021
IEEE 19th World Symposium on Applied Machine Intelligence and Informatics
Minimally Invasive Surgery (MIS)

- Operating using smallest incisions
- Shortest recovery
- Providing the surgeon with the field view using cameras
- Reduces complications and health risks
- Decreases hospital stay
Virtual Reality (VR) Trainers

- Educational tools for the attainment of MIS surgery skills such as:
  - Strong eye-hand coordination while operating using visual information from the two-dimensional images
  - Control on the graspers and new challenging skills
- Developing skills without error risks
- Improve operating room (OR) performance
Laparoscopic Surgical Box-trainer

- Equipped with:
  - Cameras
  - Monitor
  - Light source
  - Needle holder, scissors, and grasper
- Allows the trainees to practice laparoscopic techniques in a low-cost environment
Related Work

Laparoscopic instrument detection and tool-tip tracking algorithms:

- Gradient-based features
- Color detection
- Haar wavelets
- Texture features
- Edge detection method
- Deep Learning methods
- Convolutional Neural Networks (CNN)
- Recurrent Neural Networks (RNN)
- Image segmentation
Object Detection

- A computer technology related to computer vision and image processing
- Detecting instances of semantic objects of a certain class (such as humans, buildings, or cars) in digital images and videos
- In this study we used deep learning algorithm
Procedure for Laparoscopic Instrument Detection

- Data collection
- Data preparation
- Choosing a model
- Training the model
- Making predictions
In the field of laparoscopic surgical box-trainer research there are no published datasets.

We have recorded many videos during pattern cutting tests procedure using our FLS box-trainer.

Our datasets consist of extracted frames of laparoscopic training videos.

We aim to release this dataset in public, so that other researchers can benefit from it.

We manually annotated each image using the Image Annotation Tool – LabelImg - an open-source image labeling tool.
Data Preparation

Partitioning of the Dataset

Typically, the ratio of this arrangement is 6:2:2

- 60% of the images for training the model
- 20% of the images for validation during training in each epoch
- The remaining 20% for testing
Creating a Label Map

- Each training algorithm requires a label map
- It maps each of the labels (classes) to an integer value
- It is used both by the training and the detection processes
Data Preparation

Data Augmentation

Geometric transformations for data augmentation:
- Translations
- Rotations
- Changes in scale
- Shearing
- Horizontal (and in some cases, vertical) flips

Details of the dataset per instruments before and after image augmentation

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Available size</th>
<th>Augmented size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scissors</td>
<td>300</td>
<td>640</td>
</tr>
<tr>
<td>Grasper</td>
<td>280</td>
<td>485</td>
</tr>
<tr>
<td>Circle</td>
<td>370</td>
<td>695</td>
</tr>
<tr>
<td>Total</td>
<td>950</td>
<td>1820</td>
</tr>
</tbody>
</table>
PyTorch is:

- An open-source machine learning library
- Based on the Torch library
- Used for applications such as computer vision and natural language processing, object detection
Faster R-CNN ResNet-50 FPN Architecture

• A: ResNet-50

• B: the RPN classifies scores and bounding boxes (Bbox) of the proposal regions by localizing each object

• C: the last conv feature map connects to a RoI Pooling Layer, connected to the proposal region

• D: Classifier, there are two output layers of Fast R-CNN, representing two vectors per proposal region: SoftMax probabilities and Bbox regression
Datasets are usually grouped into batches (especially when the amount of data is very large).

An epoch is a term used in machine learning and indicates the number of passes of the entire training dataset the machine learning algorithm has completed.

After the training data is prepared and the model is selected, we train the model to a specific number of epochs.
• The performance of a model is tested by making predictions on new images.

• The testing data is used to see if a model is capable of predicting the surgical instruments on images which haven’t been trained.

• The output of the detected object is then compared with the label map to evaluate whether the predicted instruments is correct or not.
Detections

Introduction
Methodology
Results
Summary

Circle: 0.82
Scissors: 0.71
Grasper: 0.92

Circle: 0.76
Scissors: 0.76
Grasper: 0.92

Circle: 0.70
Scissors: 0.72
Grasper: 0.76

SAM2021
IEEE 19th World Symposium on Applied Machine Intelligence and Informatics
Quantitative & Qualitative Discussion

Total Loss:

- Classification: determining the existence of an object in the image
- Localization: determining the location of a region of interest of the object in the image using the regression method

<table>
<thead>
<tr>
<th>Bounding Box Score</th>
<th>Instrument</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Scissors</td>
</tr>
<tr>
<td>Score after 15 epochs</td>
<td>0.6367</td>
</tr>
<tr>
<td>Score after 25 epochs</td>
<td>0.7658</td>
</tr>
<tr>
<td>Score after 35 epochs</td>
<td>0.7937</td>
</tr>
</tbody>
</table>
Summary and Future Research

• Importance of training experienced and skilled laparoscopic surgeons
• The performance of a laparoscopic surgeon with respect to the movements of the tool tips
• Preparing a dataset for future research in assessment of the residents’ performance
• A multi-class laparoscopic object detection method using a Faster R-CNN Architecture
• The module predicts each laparoscopic instrument
• In future research, we plan to assess the performance of surgeons using fuzzy logic assessment system.
Thank you for your attention!