Towards Validated Surgical Skills Assessment: Automated monitoring of Tracheoesophageal Fistula Repair
Siddharth Jain, BE1, Katherine A Barsness, MD, MS2, Ellie O'Brien, BS3, Brenna D Argall, PhD1
1Northwestern University, Rehabilitation Institute of Chicago, 2Ann and Robert H. Lurie Children’s Hospital of Chicago, Northwestern Univ. Dept. of Medical Education, 3Northwestern University Department of Medical Education

1st Author: Siddarath Jain, BE | Presenter: Siddarath Jain, BE

Topic: Simulation-Based Education

Number of Reviewers: 4
Total Score: 30
Mean Score: 7.5
Would like to be considered IRCAD Award: NO
Would like to be considered Basic Science Award: NO

<table>
<thead>
<tr>
<th>Score</th>
<th>Reviewer</th>
<th>Recommendation</th>
<th>Overall Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Celeste Hollands</td>
<td>Oral/Podium</td>
<td>well designed simulator study</td>
</tr>
<tr>
<td>6</td>
<td>Tadashi Iwanaka</td>
<td>Poster</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Todd A. Ponsky</td>
<td>Oral/Podium</td>
<td>although in its infancy this may become a real tool</td>
</tr>
<tr>
<td>9</td>
<td>Atul J. Sabharwal</td>
<td>Plenary</td>
<td>Excellent</td>
</tr>
</tbody>
</table>

Background: Validated assessments of technical skill are critical to the successful implementation of simulation-based educational curricula. Unfortunately, the evaluation of surgical skill is often time consuming, and cumbersome to work into the natural flow of a training or testing event. Computer vision and machine learning, as a component of surgical skills assessment, may alleviate many of the inefficiencies of current assessment strategies. The purposes of this work were to develop, and evaluate for accuracy, an algorithm to autonomously detect correct and incorrect procedural steps occurring during simulated thoracoscopic ligation of a tracheoesophageal fistula (TEF).

Methods: Using a previously described synthetic thoracoscopic TEF repair model, two separate video-taped performances of correct and incorrect procedural steps for TEF ligation were collected. The operative telescope (OT) video was used for surgeon visualization and to establish ground truth for correct and incorrect procedural steps. The tracheal telescope (TT) video was used for subsequent development and testing of the algorithm.

Image frames from the TT video data were processed for the automated detection of correct and incorrect procedural steps, classified as – (i) Initial state, (ii) Bronchial ligation, (iii) Partial ligation, (iv) Correct ligation and (v) Tracheal compression. The detection algorithm proceeded as follows. The color image frame was converted to a high contrast grayscale image, and a binary image was obtained using thresholding, and the complement was applied. The image consisted of regions of interest, which were identified by connecting components. The geometric parameters, namely—centroid, area and equivalent diameter were then calculated and compared to the initial image parameters (Fig. 1). Tracheal compression, without a defined geometric pattern, was more difficult to detect using computer vision techniques. Therefore, machine learning was employed. A Support Vector Machine (SVM) classifier using histogram of gradient features was trained to associate sets of features with known instances of Tracheal compressions versus not, and to classify unknown instances using the previously learned model.

Results: The automated algorithm examined two TT videotaped performances of thoracoscopic fistula ligation. The output was compared with the recorded ground truth and this comparison was used as the performance measure. The algorithm performed very well, as the output on the correct and incorrect procedural steps imitated the ground truth for the videotaped evaluations (Fig. 2). SVM classification achieved acceptable accuracy (86%) at discriminating the Tracheal compression cases on unseen test data.

Discussion: The combination of computer vision techniques and a machine learning classifier was successfully used for the automated determination of correct and incorrect procedural steps during simulated thoracoscopic ligation of TEF. Experimental validation of the algorithm demonstrated the practical feasibility of online detection during the procedure. These strategies allowed for efficient and objective assessment of surgical skills. Future steps involve evaluation of operative videos from surgeons with variable skill levels, and assessment of the algorithm performance on a broader range of surgical skills.

Figure 1: (i) Initial state, (ii) Bronchial ligation, (iii) Partial ligation, (iv) Correct ligation

http://www.ipeg.org/printabsSCORE.php
Figure 2: Illustration of performance comparison