Arthroscopic Tool Classification using Deep Learning

Bryce Palmer
Florida Polytechnic University
4700 Research Way
Lakeland, FL 33805
+1(863) 583-9050
bpalmer6967@floridapoly.edu

Gunnar Sundberg
Florida Polytechnic University
4700 Research Way
Lakeland, FL 33805
+1(863) 583-9050
gsundberg3513@floridapoly.edu

James Dials
Florida Polytechnic University
4700 Research Way
Lakeland, FL 33805
+1(863) 583-9050
jdials3259@floridapoly.edu

Bayazit Karaman
Florida Polytechnic University
4700 Research Way
Lakeland, FL 33805
+1(863) 583-9050
bkaraman@floridapoly.edu

Doga Demirel
Florida Polytechnic University
4700 Research Way
Lakeland, FL 33805
+1(863) 583-9050
ddemirel@floridapoly.edu

Muhammad Abid
Florida Polytechnic University
4700 Research Way
Lakeland, FL 33805
+1(863) 583-9050
mabid@floridapoly.edu

Tansel Halic
University of Central Arkansas
201 Donaghey Ave
Conway, AR 72035
+1(501) 450-5000
tanselh@uca.edu

Shahryar Ahmadi
University of Arkansas for Medical Sciences
4301 W Markham Street
Little Rock, AR 72205
+1(501) 686-7000
SAhmadi@uams.edu

ABSTRACT

Shoulder arthroscopy is a common surgery to diagnose and treat tears to improve patient's quality of life. Quality of cleaning the tear during shoulder arthroscopy significantly affects the outcome of the surgery. Appropriate cleaning is necessary to reduce healing time and avoid feature pain in the area. In this paper, we used convolutional neural networks to automatically differentiate between two tools-electrocautery and shaver tools- that are used during the cleaning phase of a shoulder arthroscopy. We captured images from the actual shoulder arthroscopy videos. We used 8,691 images that contain the shaver tool, 7,773 images that contain the electrocautery tool, and 4,834 images that contain no tools. Our results showed that average accuracy of our model is 99.1(+/- 0.49) %. For the electrocautery tool precision and sensitivity was calculated as 0.988 and 0.988, respectively. For the shaver tool precision and sensitivity was calculated as 0.993 and 0.988, respectively. For the no tool scenes precision and sensitivity was calculated as 1.0 and 1.0, respectively.

CCS Concepts

•Computing methodologies~Machine learning~Machine learning approaches~Neural networks

Keywords

Image Recognition; Deep Learning; Convolutional Neural Network

ICISDM 2020, May 15–17, 2020, Hawaii, HI, USA

ACM ISBN 978-1-4503-7765-2/20/05…$15.00
https://doi.org/10.1145/3404663.3404672

1. INTRODUCTION

Shoulder Arthroscopy is a surgical procedure that uses a small camera called an arthroscope to visualize, diagnose, and treat tears and other issues inside of the shoulder area[1]. This procedure is used as a minimally invasive option as opposed to open surgery. There are over 1.4 million arthroscopy surgeries performed in a year [2]. Rotator Cuff is one of the most commonly performed arthroscopic shoulder procedures with approximately 300,000 surgeries around the world every year[3], [4]. Rotator cuff is the group of muscles that connect the shoulder to the arm[5]. A tear in the rotator affects the stability and the rotational motion of the arm. During an arthroscopic rotator cuff procedure tear is diagnosed, cleaned and sutured back together.

A main portion of this surgery is the cleaning phase(bursectomy). The purpose of this cleaning phase is to create a clear view of the tear allowing for a better assessment and result. This portion of the surgery is very important given that if cleaning is not done properly the bone can either be weakened by shaving too much or the tear may not be fully visible if not enough shaving is carried out (Figure 1 shows adequate and inadequate cleaning).There are two tools that are used to clean the area of focus. Those tools are a shaver tool and an electrocautery tool. While, shaver is used to remove debris in the area to create a clear view of the tear, the electrocautery works by burning the area and sealing blood vessels that were shaved open by the shaver tool.

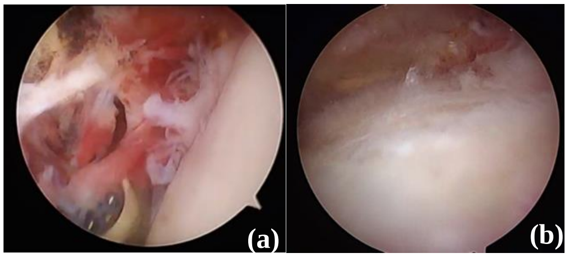

Figure 1. (a) Inadequate Cleaning
(b) Adequate Cleaning

Even though, arduous training is needed for arthroscopic rotator cuff surgery, there is no objective standard for performance assessment. The Arthroscopy Association of North America and American Board of Orthopedic surgery require orthopedic surgeons to carry out a number of surgery cases but there is no standard guideline.

Our ultimate goal is to create an automated system that will differentiate between expert and novice arthroscopy surgeons which will be used as a teaching tool as well as a performance assessment. Automated tool differentiation is a step towards that goal.

2. RELATED WORKS

Arthroscopic surgery is a kind of minimally invasive surgical procedure that can be done on any joint. This treatment is performed by inserting arthroscopy tools into the joint after applying a small incision. This results in less joint pain and stiffness. Besides recovery is fast due to a very small incision than a large incision for traditional open surgery. This requires expertise in surgery field and minor mistake will convert small surgery into a complicated case. This is eminence as a research field due to its sensitivity and direct involvement of humans as the subject. There is not much research is done in this discipline.

Tyrshkin et al. [6] proposed a navigation system for shoulder arthroscopic surgery. They perform pre-surgery processing which involves the computation of the surface model of the shoulder by using tomography images. Later, the computed surface model was registered to the patient. They used to track freehand ultrasound images got from regions on the scapula. Here they displayed in real-time 3-dimensional models of the surgical instruments respective to the surface model.

Goncalves et al. [7] did research on computer-assisted surgery. They proposed a vision system for robotic ultrasound-guided orthopedic surgery. They developed algorithms to control a robotic manipulator that can be used in both real-time and simulation for a surgical procedure in hip resurfacing. The navigation has been done by acquiring 3D US bone surface from a sequence of US images during surgery. They estimated the bone location and alignment by registering the bone surface to the pre-operative bone model for a piece of accurate information.

Ren and Meng el al. [8] investigated the navigation and robotic system for computer-assisted orthopedic surgery. They did research on navigation and robotic system, which perform sensing and actuating tasks respectively. For navigation, they proposed a hybrid tracking method to integrate optical tracking and inertial sensing techniques. The authors proposed an OPT-aided inertial navigation system for computer-assisted pelvic-acetabular surgeries and demonstrated that the accumulated drift error from the inertial sensor unit can be corrected by the OPT system. There is a state-of-the-art computer vision technique are available for pattern recognition that can be used to recognize the surgical tools during surgery.

Abid et al.[9], [10] used hog for feature extraction, k-means++ for feature classification and bag-of-features for a final decision. Their research can be adapted for automated robot and doctor's communication and more specifically surgical tool detection and recognition.

3. METHODS

In order to achieve high-performance automated tool differentiation, a convolutional neural network (CNN), a class of deep neural networks, was employed. CNNs are most commonly applied to the analysis of visual imagery and were a natural fit for the image classification needed for tool differentiation. The performance of a CNN is heavily dependent on both the quality and quantity of the data used to train it. While surgical tool datasets do exist, none were of the specificity or scale necessary for the project, and as such, it was necessary to create a new dataset.

3.1 Data Acquisition

Data acquisition began with acquiring a set of footage (20 videos ~30 hours of footage) of arthroscopy rotator cuff procedures from partner surgeons. The footage selected was performed by expert surgeons, in this case expert being defined as surgeons who had undergone fellowship programs for rotator cuff repair procedures and had extensive surgical experience. Expert surgeons had performed the surgery more than two hundred times, and observed the surgery performed more than two hundred times in the last six months. The footage was then parsed for segments containing the target tools: the shaver tool and the electrocautery tool. From two full surgical videos, four segments containing tools were obtained. These segments ranged from one and a half to nearly eighteen minutes in duration, with two segments for the shaver tool and two segments for the electrocautery tool. In addition, fifteen video segments with no tools were created, all with a duration under one minute. From these video segments, one segment of the electrocautery tool was used, as well as both segments for the shaver tool and all fifteen segments with no tools. A Python script was implemented utilizing the OpenCV library to split the video frames into images.

Splitting the videos into frames yielded 17,572 images from the shaver tool video segments, 25,679 images from the electrocautery tool segment, and 4,834 images with no tools. The next step was to manually find frames containing the tools. The images with no tools required no manual processing. Frames with heavy motion blur and significant occlusions were dropped, although images with some motion blur and slight occlusion were kept, allowing for a resilient model that performs in real-life surgical conditions. From the 17,572 shaver tool images, 8,691 were selected for training. Of the 25,679 electrocautery tool images, 7,773 were selected for training. All 4,834 images with no tools were used for training. The dataset included images of both tools at a wide variety of poses and angles, with varying lighting as well (as seen in Figure 2).

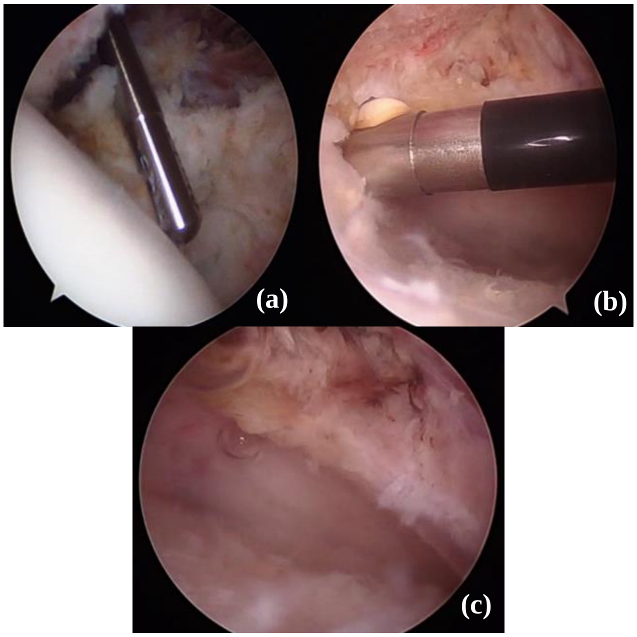

Figure 2. Dataset Examples (a) Electrocautery Tool (b) Shaver Tool (c) No Tool

3.2 Model

A CNN was implemented with a Rectified Liner Unit (ReLU) activation function on all layers but the output layer which used Softmax for multiclass classification. The goal of using a CNN was to find the features that differentiated whether an image contained an electrocautery tool, shaver tool, or no tools. Our CNN model used Max Pooling for down-sampling, as well as implemented L2 regularization and Dropout layers to prevent overfitting/underfitting.

When loading the images from the dataset to be used in the training of the CNN, they were resized to be 128x128x3 using OpenCV. The newly resized images were then loaded into a NumPy array and used in a K-Fold Cross Validation method of training the CNN implemented through use of scikit-learn's StratifiedKFold function. The CNN was trained with k=10 and was set to run for 200 epochs each fold with an early stopping condition set to end the current fold's training if the accuracy didn't improve within five epochs to prevent overtraining.

Overall, the CNN used a total of 26 layers. Figure 3 shows the layers of our CNN. There was a total of three convolutional blocks made up of seven layers each starting from 32 filters and ending at 128 filters to aid in finding much more detailed patterns. Figure 4 represents the layers of a singular convolutional block. The last five layers were made up of a Flatten layer to organize the data from the convolutional layers to be used in the dense layers, two dense layers with 128 nodes and 64 nodes respectively, a dropout layer to help prevent overfitting/underfitting, and a dense layer with three nodes for output.

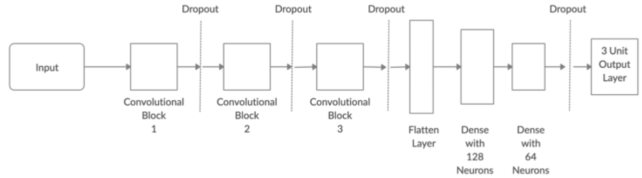

Figure 3. Graph showing the layers of our CNN

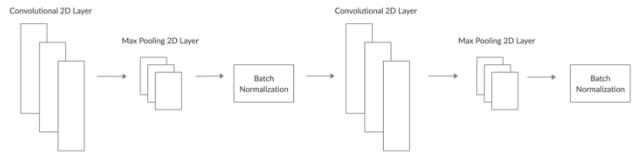

Figure 4. Layers of a singular convolutional block

4. RESULTS

Within the scope of this study, an automated tool differentiation for shoulder arthroscopy was modeled. Evaluation of the model was performed with 10-fold cross validation. The accuracies of each fold were shown in Figure 5. It is important to note that the average accuracy of the model was found **99.1(+/- 0.49) %**.

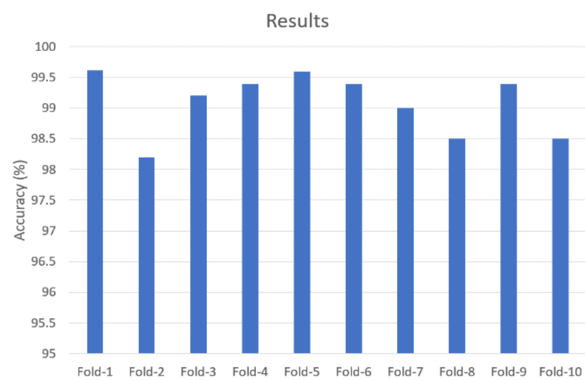

Figure 5. Accuracies of each fold for the model

Two statistical measures which are precision, and sensitivity were used to analyze the results (Table 1). Precision is the proportion of correctly classified samples over all samples that are classified as this class. Moreover, sensitivity is the proportion of correctly classified samples over all actual samples. The precision and sensitivity results from Table 1 shows that shaver tool images were classified slightly better than electrocautery tool images. Furthermore, the images with no tool were classified with 100% with our model.

Table 1. Performance evaluation of the classification method

	Precision	Sensitivity	Number of Images
Electrocautery Tool	0.988	0.988	777
Shaver Tool	0.993	0.988	868
No Tool	1.00	1.00	483

5. CONCLUSIONS

Shoulder Arthroscopy is an operation for the treatment of tears on the shoulder. While this surgery is very common, there is no objective assessment for classifying expert surgeons from novice surgeons. Furthermore, cleaning of the tear's region is the most important phase for the surgery. Therefore, our study focused on two tools which are used to perform this phase. We believe that how long these tools are used depends on whether the surgeon is expert or novice. Thus, we created a model that classifies the images according to the tools in the image using CNN. Our results show that the model is able to discriminate these tools with 99% accuracy.

As future work, we want to automate the process for determining the tool used during the cleaning phase and incorporate it to help assess the surgeons as novice and expert.

6. ACKNOWLEDGMENTS

This project was made possible by the Arkansas INBRE program, supported by a grant from the National Institute of General Medical Sciences, (NIGMS), P20 GM103429 from the National Institutes of Health (NIH). This project was also supported by NIH/NIAMS R44AR075481-01, NIH/NCI 5R01CA197491 and NIH/NHLBI NIH/NIBIB 1R01EB025241, R56EB026490.

7. REFERENCES

[1] R. Treuting, “Minimally invasive orthopedic surgery: arthroscopy,” *Ochsner J.*, vol. 2, no. 3, pp. 158–163, 2000.

[2] “Shoulder Surgery - OrthoInfo - AAOS.” https://www.orthoinfo.org/en/treatment/shoulder-surgery/ (accessed Feb. 07, 2020).

[3] H. S. Mahon, J. E. Christensen, and S. F. Brockmeier, “Shoulder rotator cuff pathology: common problems and solutions,” *Clin. Sports Med.*, vol. 37, no. 2, pp. 179–196, 2018.

[4] M. D. McElvany, E. McGoldrick, A. O. Gee, M. B. Neradilek, and F. A. Matsen III, “Rotator cuff repair: published evidence on factors associated with repair integrity and clinical outcome,” *Am. J. Sports Med.*, vol. 43, no. 2, pp. 491–500, 2015.

[5] D. Demirel *et al.*, “A hierarchical task analysis of shoulder arthroscopy for a virtual arthroscopic tear diagnosis and evaluation platform (VATDEP),” *Int. J. Med. Robot.*, 2016, Accessed: Mar. 06, 2017. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1002/rcs.1799/full.

[6] K. Tyryshkin, P. Mousavi, M. Beek, R. E. Ellis, D. R. Pichora, and P. Abolmaesumi, “A navigation system for shoulder arthroscopic surgery,” *Proc. Inst. Mech. Eng. [H]*, vol. 221, no. 7, pp. 801–812, 2007.

[7] P. J. S. Gonçalves, P. M. Torres, F. Santos, R. António, N. Catarino, and J. M. M. Martins, “A vision system for robotic ultrasound guided orthopaedic surgery,” *J. Intell. Robot. Syst.*, vol. 77, no. 2, pp. 327–339, 2015.

[8] H. Ren and M. Q.-H. Meng, “Investigation of navigation and robotic system for computer assisted orthopedic surgery: State-of-art and preliminary results,” *Int. J. Inf. Acquis.*, vol. 6, no. 03, pp. 171–179, 2009.

[9] M. R. Abid, P. E. Meszaros, R. F. d Silva, and E. M. Petriu, “Dynamic hand gesture recognition for human-robot and inter-robot communication,” in *2014 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, 2014, pp. 12–17.

[10] M. R. Abid, E. M. Petriu, and E. Amjadian, “Dynamic sign language recognition for smart home interactive application using stochastic linear formal grammar,” *IEEE Trans. Instrum. Meas.*, vol. 64, no. 3, pp. 596–605, 2014.