



# Towards Automated Tissue Classification for Markerless Orthopaedic Robotic Assistance

Stephen Laws<sup>1</sup>, Spyridon Souipas<sup>1</sup>, Brian Davies<sup>1</sup> and Ferdinando Rodriguez y Baena<sup>1</sup>

<sup>1</sup> Mechatronics in Medicine Laboratory, Imperial College London, London, U.K.  
S16114@ic.ac.uk, ss8413@ic.ac.uk, b.davies@imperial.ac.uk,  
f.rodriquez@imperial.ac.uk

## Abstract

A markerless computer aided orthopaedic platform will require a complex computer vision system to isolate and track rigid bodies used to localize a robot to a patient. Isolating rigid bodies such as bone requires accurate segmentation and this study explores using diffuse laser reflectivity to accurately classify tissue.

Lasers (Red, 650nm and IR, 850nm) intersected four material types; cartilage, ligament, muscle and metal surgical tools within a controlled cadaveric setup. Images were captured with an infrared CMOS sensor, pre-processed to isolate laser centres, and resized to test information requirements. Images for both laser types were scaled from 5x5 pixels to 30x30 pixels and trained on a convolutional neural network, GoogLeNet.

At sizes above 15x15 pixels the IR laser had a higher classification accuracy reaching 97.8% at 30x30 pixels, whereas the red laser peaked at 94.1%. It was shown as not possible to qualitatively identify materials that were not trained in the network based on their probability outputs. Further work will be done to classify multiple points in a single scene as a step toward segmenting entire surgical views for markerless CAOS systems.

## 1 Introduction

Orthopaedic robotic platforms accurately track markers rigidly attached to bones during joint replacement, allowing improved implant positioning [1, 2]. There is currently no CAOS platform that uses markerless techniques to track bones and tools intraoperatively, although research shows some success in markerless tool tracking [3]. One difficulty with markerless systems is differentiating between trackable rigid bodies and surrounding soft tissues, such that a robot can compute its position relative to the patient.

Techniques for segmenting tissue types during surgery are becoming increasingly popular, with image segmentation [4] and hyperspectral imaging [5]. This study builds on previous work [6] to study how laser diffuse reflectance imaging can be used to classify human tissue in real-time. Specifically,

our work compares how different wavelengths affect accuracy, quantify what resolution is required for accurate classification and if it is possible to qualitatively assess whether our method can identify if a sample material is part of the trained dataset.

## 2 Materials and Methods

The experiments comprised of acquiring training data for both a red and infrared laser shone on cadaveric tissue samples. The red laser builds on previous work [6] but is visible to the human eye, whilst the IR laser is a new wavelength to test the efficacy of invisible light. 12 human shoulder cadaveric samples were obtained from a licensed tissue bank. Ethical approval for their use was obtained from the Imperial College Healthcare Tissue Bank, under project R19040. For this study four material types were used for classification; cartilage, ligament, muscle and metal surgical tools, the latter having been included to test the efficacy of classifying non-organic material. The training data is an RGB image from a CMOS 8MP IR (up to 880nm) sensor (Pi-NoIR Camera V2, Raspberry Pi, UK). Each image contains a laser pointer (<1mW Class II 650nm±10nm & 850nm±10nm) intersecting the target material. Approximately 500 images of each material type were captured, of which 200 were hidden from the training set and used for validation.

Post processing was performed to crop these images to the area around the laser centre via local maxima greyscale thresholding. Each image was scaled to square images ranging from 5x5 to 30x30 pixels to explore how information content affected classification accuracy. The resized images were additionally augmented with random X-Y reflections and rotations to increase the diversity of the training data.

Each set of augmented training data, 52 in total (26 sizes and 2 laser types), was trained by reclassifying a pretrained convolutional neural network (CNN, GoogLeNet [7]). Networks were each trained for 100 Epochs on an Intel Core i7-7700K CPU (4.20GHz) with 16GB of RAM and a 4GB NVIDIA GeForce GTX 1050 Ti.

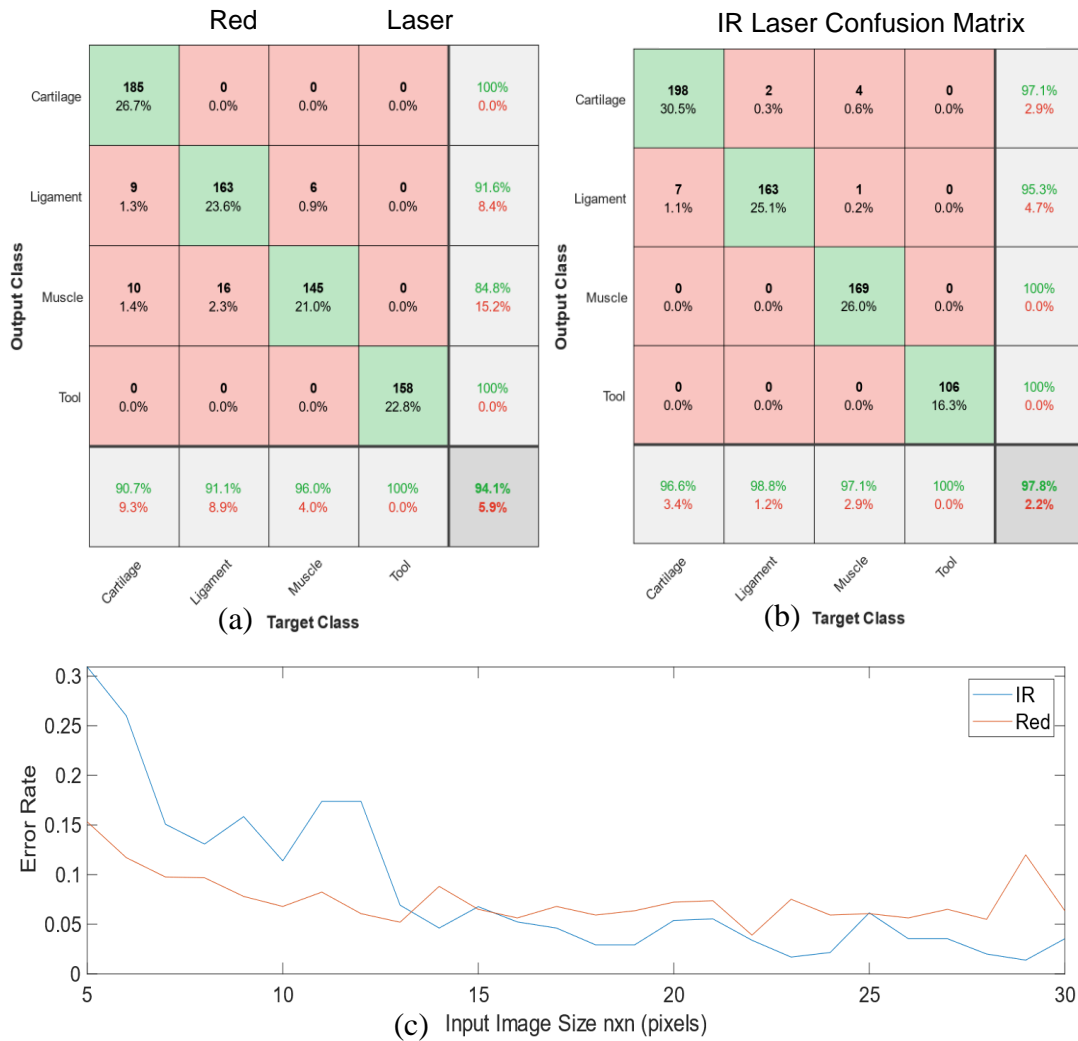
There are two methodologies for classifying ‘unknown’ objects with a CNN. The first method includes the unknown class in the training data so the algorithm can isolate features unique to outliers; this however can weaken the overall classification accuracy of target classes. The method used in this study uses the classification probabilities to classify samples as ‘other’ if they are below a probability threshold.

## 3 Results

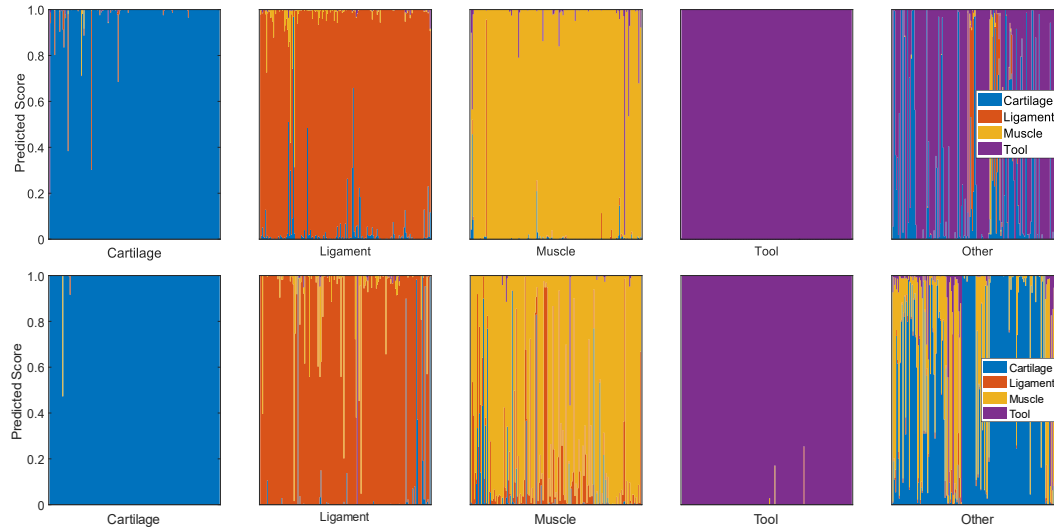
Validation of the trained networks involved classifying the unseen images of each class to evaluate the error rate for each of the sizes and different laser types, as shown in Figure 1. Confusion plots for the 30x30 pixel inputs of the Red and IR Lasers can be seen in Figure 1a and 1b, respectively. The right-hand column of the confusion plots indicates the precision (false discovery rate) of each class and the bottom row indicates the recall (false negative rate). Figure 1c plots the error rate for IR and Red lasers as information content (size) increases.

Both the IR and Red lasers proved good at classifying the 4 tissues types accurately as the image sizes approached 30x30 pixels. The IR images had a consistently lower error rate than the red laser after a size of 15x15 pixels was reached. However, the red laser was twice as accurate at classifying tissues at 5x5 pixels than the IR laser, Figure 1c. The IR classifier attained an overall accuracy of 97.8% with high recall for all tissue types, Figure 1b. Whereas the red laser was less accurate at classifying both cartilage and ligaments with relatively low recalls for both, Figure 1a.

Figure 2 shows the stacked prediction scores for both the IR and red lasers for 100 test images (30x30 pixels input size) of each class. Each sub-column indicates the probability prediction for each image in that class, indicating how confidently the network could accurately categorize each image. The horizontal axis indicates the true class in each instance, which includes using the trained network to classify images of ‘unknown’ class. Attempts to isolate a probability feature to classify ‘unknown’ images were unsuccessful, primarily due to the diverse range of images that the network has not been trained on for both the IR and Red lasers. This is evidenced by the high-confidence predictions of a single class (majority single colour sub-columns) which is not possible to isolate from true classifications.



**Figure 1:** [(a): Red, (b):IR] Confusion Matrix of 30x30 Pixel Input for Trained Network. (c) Error Rate Against Information Content (size) for IR and Red Laser



**Figure 2:** Comparison of Predicted Scores on 30x30 pixel inputs for trained classes and injected ‘Other’ class. Top: IR Laser. Bottom: Red Laser

## 4 Discussion

For markerless tracking to be viable it should be invisible to the surgeon, thus the IR laser is preferable for segmenting the tissue. The laser has proved accurate with laser spots at least 15x15 pixels across. Inference of a single spot currently takes 20ms, which can be accelerated by orders of magnitude with inference hardware [8], and a network optimized for 15x15 inputs, 200 times smaller than the GoogLeNet input size (224x224).

This study has shown that laser reflected diffuse classification is accurate even at relatively low resolutions and is suitable for online identification as part of a markerless tracking system. Further research is needed in a few areas. Firstly, the study showed that it is necessary for all possible materials in a surgical scene to be part of the training data as to limit the number of ‘unknowns’ within the tracked scene. Secondly, refracting multiple points onto the surgical site and classifying each is necessary to fully segment a scene. An accurately classified scene could then differentiate trackable rigid bodies when coupled with a depth camera system.

## References

- [1] E. L. Hampp *et al.*, “Robotic-arm assisted total knee arthroplasty demonstrated greater accuracy and precision to plan compared with manual techniques,” *J. Knee Surg.*, vol. 32, no. 03, pp. 239–250, 2019.
- [2] J. A. Lawson, A. T. Garber, J. D. Stimac, R. Ramakrishnan, L. S. Smith, and A. L. Malkani, “Does Robotic-Assisted Total Hip Arthroplasty Improve Accuracy of Cup Positioning?,” *J. Hip*

- Surg.*, vol. 3, no. 04, pp. 176–180, 2019.
- [3] S. C. Lee *et al.*, “Multi-modal imaging, model-based tracking, and mixed reality visualisation for orthopaedic surgery,” *Healthc. Technol. Lett.*, vol. 4, no. 5, pp. 168–173, 2017.
  - [4] T. de Moor, A. Rodriguez-Ruiz, A. G. Mérida, R. Mann, and J. Teuwen, “Automated soft tissue lesion detection and segmentation in digital mammography using a u-net deep learning network,” *arXiv Prepr. arXiv1802.06865*, 2018.
  - [5] E. L. Wisotzky *et al.*, “Validation of two techniques for intraoperative hyperspectral human tissue determination,” no. March 2019, p. 71, 2019.
  - [6] S. Laws, B. L. Davies, and F. Rodriguez y Baena, “Automated Soft Tissue Classification using Laser Pointer Diffuse Reflection Imaging,” in *International Society of Computer Aided Orthopaedic Surgery*, 2019.
  - [7] C. Szegedy *et al.*, “Going deeper with convolutions,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, vol. 57, pp. 1–9.
  - [8] V. Sze, Y. H. Chen, J. Emer, A. Suleiman, and Z. Zhang, “Hardware for machine learning: Challenges and opportunities,” *2018 IEEE Cust. Integr. Circuits Conf. CICC 2018*, pp. 1–8, 2018.