Convolutional neural networks for coronary plaque classification in Intravascular Optical Coherence Tomography (IVOCT) images

Chaitanya Kolluru†
Biomedical Engineering
CWRU
Cleveland OH USA
chaitanya.kolluru@case.edu

David Prabhu
Biomedical Engineering
CWRU
Cleveland OH USA
david.prabhu@case.edu

Yazan Gharaibeh
Biomedical Engineering
CWRU
Cleveland OH USA
yazan.gharaibeh@case.edu

Sanjaya Gajurel
RCCI [U]Tech
CWRU
Cleveland OH USA
sanjaya.gajurel@case.edu

ABSTRACT
Intravascular Optical Coherence Tomography (IVOCT) is a high resolution, high contrast imaging technique that can be used to characterize coronary artery plaques and guide stent interventions. Currently, IVOCT is the only imaging technique with the resolution necessary to identify vulnerable thin cap fibro-atheromas (TCFAs). IVOCT also has greater penetration depth in calcified plaques as compared to Intravascular Ultrasound (IVUS). Despite its advantages, IVOCT image interpretation is challenging and time consuming, especially among novice readers, with over 500 images generated in a single pullback. In recent years, convolutional neural networks (CNNs) have surpassed state of the art methods in visual recognition tasks such as image classification and semantic segmentation. In this paper, we propose a method to automatically classify A-lines in IVOCT images using a convolutional neural network. Conditional random fields were used to clean network predictions across frames. The neural network was trained using a dataset of over 4,500 image frames across 48 IVOCT pullbacks. Ten-fold cross validation with held-out pullbacks resulted in a classification accuracy of roughly 76% for fibrocalcific, 84% for fibrolipidic and 85% for other. The classification results across frames displayed in en face view matched closely to annotated counterparts.

CCS CONCEPTS
Machine learning algorithms

KEYWORDS
Intravascular Optical Coherence Tomography (IVOCT), Convolutional Neural Network (CNN)

1 Methods
IVOCT images are obtained in the polar domain. An automatic lumen segmentation algorithm based on a dynamic programming approach [1] is used to segment the lumen region and identify guidewire positions in the image. A-lines within the guidewire regions are discarded and are not considered for the classification task. The resulting A-lines are pixel shifted to reduce the effect of large lumen eccentricity. Speckle noise is reduced by taking a log transform and performing Gaussian filtering. A-line segments of length 1 mm (200 pixels) from the lumen boundary are fed as inputs to the CNN.

The CNN architecture used for this work is shown in Table I. Briefly, it consists of two convolutional layers, each followed by a reshaping operation and a max-pooling (MP) operation of length 2. Two fully connected (FC) layers with size 100 units and 3 units are used as the penultimate and final layers in the network. Softmax activation is used in the final layer to ensure that the network outputs probabilities for each class which sum to one. Overfitting is controlled by the addition of a dropout layer with a ratio of 0.5 after the first FC layer during training. Neural network results are cleaned in the en face (θ, z) view using a dense conditional random field as described in [2]. Parameters of the conditional random field algorithm, including the size and weight of the smoothness kernel are found in an ad hoc fashion.

2 Datasets
IVOCT images are annotated by an expert IVOCT reader by analyzing the image in cartesian domain. A second reader is presented with the same images and consensus is taken between the two readers to obtain the ground truth. A-lines are classified into three classes, namely, fibrocalcific, fibrolipidic and “other”. The “other” class contains all A-lines that are not of the previous two classes. A total of 4,469 image frames are analyzed covering a total of 48 pullbacks. Example annotations are shown in Fig 2.

3 Results and Discussion
We employ a ten-fold cross validation procedure to assess classifier performance. Specifically, each loop of the cross-validation procedure contained five pullbacks for testing and the remaining 43 pullbacks for training and validation. For each fold, five pullbacks
TABLE I. CNN ARCHITECTURE USED TO CLASSIFY A-LINES

<table>
<thead>
<tr>
<th>Layer</th>
<th>Output shape</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>(200, 1)</td>
<td>Individual A-lines</td>
</tr>
<tr>
<td>Padding layer</td>
<td>(210, 1)</td>
<td>Pixels replicated at corners</td>
</tr>
<tr>
<td>Convolutional layer</td>
<td>(200, 1, 32)</td>
<td>Kernel size of (11, 1)</td>
</tr>
<tr>
<td>Max-pooling layer</td>
<td>(100, 1, 32)</td>
<td>Pool size of (2,2) and stride 2</td>
</tr>
<tr>
<td>Convolutional layer</td>
<td>(100, 1, 64)</td>
<td>Kernel size of (9, 32)</td>
</tr>
<tr>
<td>Max-pooling layer</td>
<td>(50, 1, 64)</td>
<td>Pool size of (2,2) and stride 2</td>
</tr>
<tr>
<td>Fully connected layer</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Fully connected layer</td>
<td>3</td>
<td>Output layer, softmax</td>
</tr>
</tbody>
</table>

Fig 2. Manual A-line segmentation of plaques in XY images (labels indicated in a ring around image). Fibrocalcific A-lines are indicated in red color and fibrolipidic plaques are shown in green, other A-lines shown in blue and A-lines within guidewire shadow are colored black.

were randomly selected from the training set and were used to form the validation dataset. The classifier performance after noise cleaning by the conditional random field was aggregated over all the folds and a confusion matrix is created (Table II). Results from the automatic classification procedure are shown for a few example frames in Figure 3. Although the overall classification accuracy is around 82%, we argue that this learning method produces clinically useful results. During an intervention, a cardiologist is interested in locating plaque deposits that are much larger than single A-lines. For example, a cardiologist would want to place a longer stent if he suspects the presence of a large lipid pool. Similarly, large calcifications hamper the expansion of a stent. Both scenarios show that high-resolution accuracy is not necessary.

This work has been carried out in Case Western Reserve University (CWRU) High Performance Computing (HPC) cluster where singularity [3] is used to pull the tensorflow image from the docker container.

TABLE II. CONFUSION MATRIX FOR 10-FOLD CROSS VALIDATION

<table>
<thead>
<tr>
<th></th>
<th>Predicted Fibrocalcific</th>
<th>Predicted Fibrolipid</th>
<th>Predicted Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Fibrocalcific</td>
<td>243821 (76%)</td>
<td>33753 (10%)</td>
<td>43753 (13%)</td>
</tr>
<tr>
<td>True Fibrolipid</td>
<td>19535 (4%)</td>
<td>362332 (84%)</td>
<td>48144 (11%)</td>
</tr>
<tr>
<td>True Other</td>
<td>67455 (5%)</td>
<td>126298 (10%)</td>
<td>1101586 (85%)</td>
</tr>
</tbody>
</table>

The dependency packages such as keras, h5py, cython, graphviz, pydot, opencv-python, and libtiff are then installed in the singularity image. The neural networks were trained in a reasonable amount of time (under 20 minutes per fold) on a single GPU node (Intel(R) Xeon(R) CPU E5-2640 v4 @ 2.40GHz, 2x GPU Tesla P100 cards). The GitHub pages [4][5] are available for its reproducibility. We will consider utilizing multiple GPU nodes in the HPC to scale up the performance on large training datasets in the future.

Fig 3. Results of A-line classification after noise cleaning and ground truth overlaid on OCT frames within a test pullback. Inner ring indicates classification results; outer ring is the ground truth. Color scheme: Blue indicates “other”, red indicates fibrocalcific and green indicates fibrolipidic A-lines.

Since we intend to deploy this software to surgical suites, it is necessary that the classification software run as a stand-alone system and present results in near real-time. Our initial experiments show that this is possible, as we could run the trained network on a notebook GPU (NVIDIA GeForce GTX 950M) and could pre-process, predict and display classification result on a test image within 0.3 seconds.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health through grant number R01HL114406. This work made use of the High Performance Computing Resource in the Core Facility for Advanced Research Computing at Case Western Reserve University (CWRU).

REFERENCES