

Toward reliable calcification detection: calibration of uncertainty in object detection from coronary optical coherence tomography images

Hongshan Liu,^a Xueshen Li,^a Abdul Latif Bamba^b, Xiaoyu Song,^c
Brigitta C. Brott^d, Silvio H. Litovsky,^d and Yu Gan^{a,*}

^aStevens Institute of Technology, Biomedical Engineering Department, Hoboken, New Jersey, United States

^bColumbia University, Department of Electrical Engineering, New York, United States

^cIcahn School of Medicine at Mount Sinai, New York, United States

^dUniversity of Alabama at Birmingham, School of Medicine, Birmingham, Alabama, United States

Abstract

Significance: Optical coherence tomography (OCT) has become increasingly essential in assisting the treatment of coronary artery disease (CAD). However, unidentified calcified regions within a narrowed artery could impair the outcome of the treatment. Fast and objective identification is paramount to automatically procuring accurate readings on calcifications within the artery.

Aim: We aim to rapidly identify calcification in coronary OCT images using a bounding box and reduce the prediction bias in automated prediction models.

Approach: We first adopt a deep learning-based object detection model to rapidly draw the calcified region from coronary OCT images using a bounding box. We measure the uncertainty of predictions based on the expected calibration errors, thus assessing the certainty level of detection results. To calibrate confidence scores of predictions, we implement dependent logistic calibration using each detection result's confidence and center coordinates.

Results: We implemented an object detection module to draw the boundary of the calcified region at a rate of 140 frames per second. With the calibrated confidence score of each prediction, we lower the uncertainty of predictions in calcification detection and eliminate the estimation bias from various object detection methods. The calibrated confidence of prediction results in a confidence error of ~ 0.13 , suggesting that the confidence calibration on calcification detection could provide a more trustworthy result.

Conclusions: Given the rapid detection and effective calibration of the proposed work, we expect that it can assist in clinical evaluation of treating the CAD during the imaging-guided procedure.

© The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JBO.28.3.036008](https://doi.org/10.1117/1.JBO.28.3.036008)]

Keywords: optical coherence tomography; coronary artery disease; deep learning; calibration.

Paper 220275GR received Nov. 21, 2022; accepted for publication Mar. 6, 2023; published online Mar. 27, 2023.

1 Introduction

Optical coherence tomography (OCT) can acquire high-resolution cross-sectional images of coronary arteries. The high-quality and detailed information from coronary OCT images facilitates the treatment of coronary artery disease (CAD). CAD causes 1 of every 5 deaths in Europe¹ and 1 of every 6 deaths in the United States,² and it remains one of the leading causes of

*Address all correspondence to Yu Gan, ygan5@stevens.edu

morbidity and mortality in developed countries.³ Coronary atherosclerosis is caused by the gradual buildup of plaque resulting from the depositing of calcium, lipids, and macrophages from the luminal blood into the arterial intima. Coronary atherosclerosis compounds and augments the risks of heart attack and heart failure. When treated improperly or left unattended, coronary atherosclerosis blocks the pathways to the heart's main arteries, known as the coronary arteries. The potential effects of plaque in CAD include chest pain, shortness of breath, heart failure, myocardial infarction, and sudden death.

A typical treatment for CAD is percutaneous coronary intervention (PCI), which is a non-surgical procedure used to treat the narrowing of the heart's coronary arteries. Unidentified calcified tissues within a narrowing artery often negatively impact the benefits of treatment. Approximately 700,000 PCIs are performed every year in the United States, and calcifications have been found in 17% to 35% of patients undergoing the procedure,⁴⁻⁶ highlighting a need to precisely locate the existence and extent of calcifications. Most PCI procedures involve using stents to open up obstructed coronary arteries.⁷ During the PCI procedure, a catheter with a tiny, folded balloon on its tip is inserted into the blood vessels until it arrives at the site where the plaque buildup is causing a blockage. At that point, the balloon is inflated to compress the plaque against the artery walls, therefore widening the passageway and restoring blood flow to the heart. After that, the balloon is deflated and removed. A stent implantation is performed in the plaque buildup area to keep the artery open after removing the balloon.⁸ Excess coronary calcification is highly related to the suboptimal deployment of the stent in the coronary during the PCI.⁹ Major calcifications are of great concern for two reasons.¹⁰ Calcifications can lead to stent underexpansion and strut malapposition. Malapposition of stent struts (e.g., an empty space between the strut and the adjacent vessel wall) might preclude healthy endothelial tissue growth. Even though stent deployment is generally effective in the short term, stent efficacy can be reduced and the risk can be increased by adverse clinical events, such as in-stent restenosis and thrombosis in the medium- and long-term.¹¹⁻¹⁷

Coronary imaging guidance during PCI is one of the key determinants of treatment outcomes. Imaging is integral to every stage of PCI, such as assessment of lesion severity, preprocedural planning, optimization, and management of immediate complications.^{18,19} OCT has significant advantages for characterizing coronary calcification that typically has a signal-poor area with sharply delineated borders.²⁰ A typical OCT system can achieve a high axial resolution at the micron level and a penetration depth of up to 2 mm, indicating superior imaging capability.^{21,22} The detection of calcified regions within coronary OCT images is critical for intervention.²³ On account of this, developing an object detection algorithm that is capable of detecting calcification in OCT images is essential.

Deep learning has been increasingly explored in analyzing the diseased tissue in coronary OCT images.²⁴ In existing research works,²⁵⁻³⁰ extensive studies have been conducted to automatically identify plaque in coronary OCT images. A weighted majority voting from different convolutional neural networks (CNN)²⁶ was used to solve the multiclass classification problem of pathological formations in coronary artery tissues. A deep convolutional architecture named SegNet segmented calcification in coronary OCT images.¹⁰ A two-step deep learning approach²⁷ characterized plaques in coronary arteries in OCT images by first localizing the major calcification lesions using the CNN model and then applying the deep learning model (SegNet) to provide pixel-wise classifications of calcified plaques. A modified deep convolutional segmentation model UNet²⁸ was used to identify calcification in coronary OCT images. The segmentation module in MASK-RCNN was employed to identify the erosion region.³¹ Currently, the most popular way to perform automated analysis on OCT images is deep learning-based segmentation, which makes the pixel-wise classification and outputs the detailed shape and location of the tissue of interest. Demonstrably, the segmentation architecture results in large computational costs due to the burden of pixel-wise classification. By virtue of this, a more efficient way of enacting automated analysis of coronary images is through the use of object detection, which outputs the bounding box of the tissue region rather than the pixel-wise classification of the tissue region, to efficiently identify the diseased region in coronary images.

Although existing works also focus more on increasing the accuracy of deep learning models, the quality of predictions can be negatively impacted by overconfident deep learning models.³² The problem of overconfidence can be produced by deep learning models in the form of

providing high confidence scores for predictions.^{33–35} In general, recalibration methods of the well-trained model, such as Platt scaling,³⁶ histogram binning,³⁷ and temperature scaling,³⁸ can improve the calibration of the overconfident prediction results. In addition, model ensemble methods^{39,40} can also reduce overconfidence by aggregating the prediction results over multiple models. However, there are limited studies on correcting overconfident predictions in coronary OCT images. In OCT-related CAD treatment, overconfidence could be dangerous as confidence is often learned as the likelihood that the prediction is correct. Therefore, in safety-critical and risk-sensitive applications in clinical diagnosis, it is crucial to quantify and calibrate the uncertainty of predictions.

In this work, we aim to achieve reliable calcification detection for patients with CAD to boost the efficiency of clinical diagnosis. We summarize our contributions as follows.

1. We detect calcification in coronary OCT images via a deep learning-based object detection model. The object detection process delineates the bounding box of the calcified region within OCT images, providing a computationally efficient solution in comparison with conventional segmentation methods.
2. We propose calibrating the confidence of the coronary object detection task. We use a dependent logistic calibration method to reduce the bias in the prediction uncertainty.
3. We quantitatively and qualitatively evaluate the effectiveness of the proposed work on a human coronary dataset. The experimental results demonstrate the accuracy and speed of calcification detection and the effectiveness at reducing the bias of confidence among the three most popular object detection methods.

2 Methods

The workflow is shown in Fig. 1. The steps are as follows: (1) the coronary OCT data are first processed by a data augmentation module to create motion-blurred and horizontally flipped copies of each original OCT image. (2) The coronary OCT data after augmentation are trained by deep learning object detection models, and the detection results on test data are output. (3) Detections containing bounding box coordinates and confidence scores are processed through dependent logistic calibration, and a calibrated confidence score is output for each predicted bounding box.

2.1 Data Collection

Samples are imaged by the spectral domain OCT system (Thorlabs Ganymede, Newton, New Jersey, United States) with an axial resolution of $3\ \mu\text{m}$ and a lateral resolution of $4\ \mu\text{m}$ in air. Autopsy specimens of human heart vessels are collected and imaged through the same protocol given in Refs. 41 and 42. All images are acquired in the laboratory at the University of Alabama.

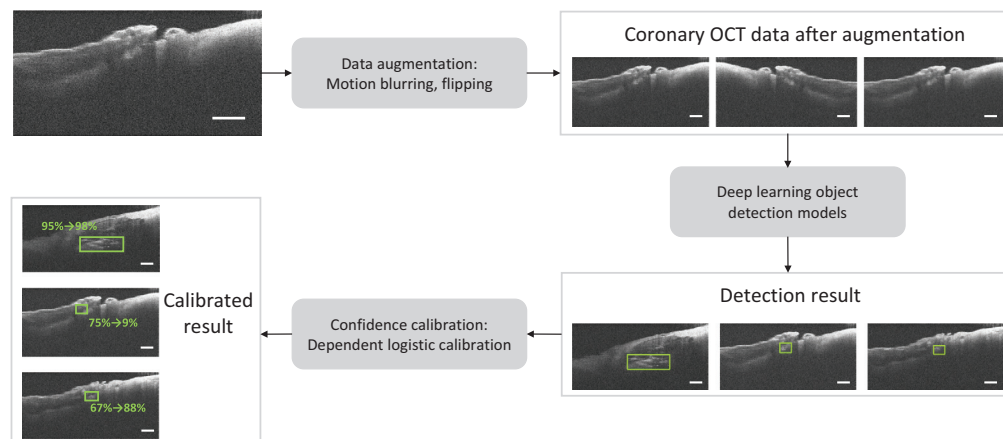


Fig. 1 Flowchart of the proposed work. Scale bar: $500\ \mu\text{m}$.

2.2 Data Augmentation

Various data augmentation techniques have been proposed to improve the performance of deep learning models.⁴³ During imaging, the quality of OCT images may be impacted due to degradation caused by motion blur,^{44,45} which can be caused by sample and device movement.^{46–48} A motion blur filter is used to simulate this effect of real-world conditions.⁴⁹ Other common augmentation strategies, such as flipping, cropping, scaling, Gaussian noise, rotation, and shears, are routinely performed.⁵⁰ Noticeably, we do not prefer vertical flipping or rotation in OCT images because the light propagates in a fixed direction, and applying such methods will change the nature of OCT images. Therefore, in this work, two copies of each OCT image are created by applying a motion-blurring filter and flipping horizontally for training the deep learning model.

2.3 Object Detection

Object detection creates bounding box regions that identify an object’s position, size, and class within an image. We opt to use You-Only-Look-Once v5 (YOLO)⁵¹ to rapidly identify the bounding box and tissue types within an OCT image. Because of its lightweight and feature-reuse properties, the YOLO architecture is powerful at realizing fast and accurate detection. As the conceptual schematic shown in Fig. 2, to better predict objects of different sizes, YOLO enhances the detection performance by utilizing different scales of feature maps that are generated by applying filters to the input image or the feature map output of the prior layers.

The predictions have the outputs in two branches: confidence scores (confidence) and bounding boxes (x_{center} , y_{center} , width, and height). In the confidence score branch, the confidence score indicates a certain level that the prediction is true. In the bounding boxes branch, the values of the center coordinate, together with the width and height of the bounding box, depict the location of the predicted bounding box.

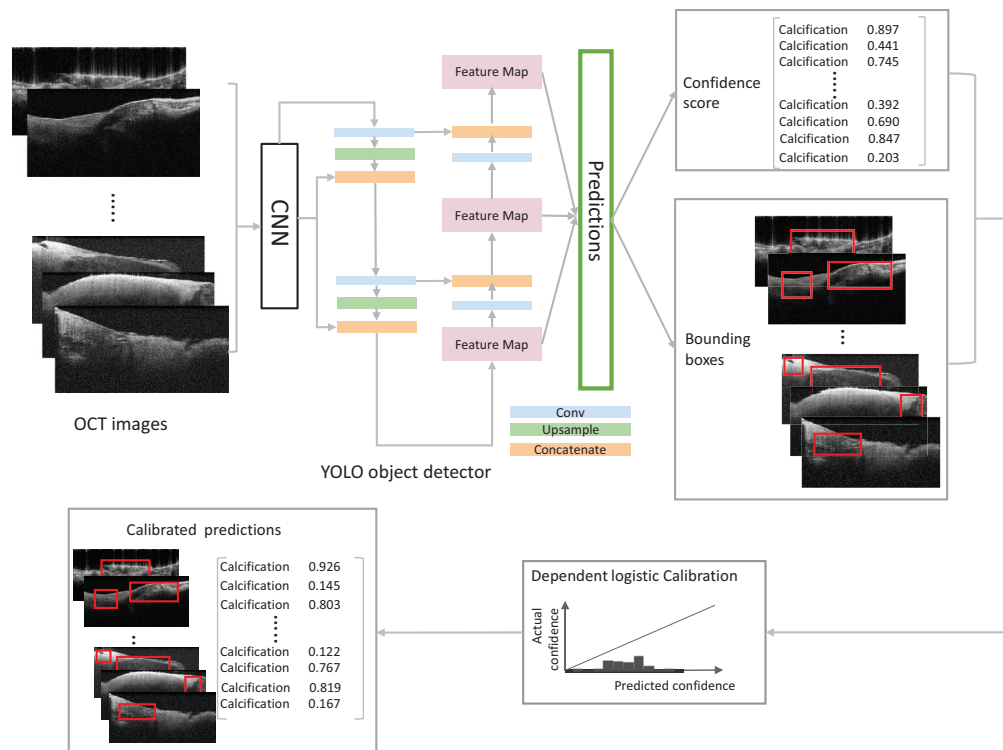


Fig. 2 Schematic of YOLO object detector and calibration.

2.4 Uncertainty Measurement and Confidence Calibration

The common use of calibration is for the classification task, in which only the confidence score is utilized for a given image. In object detection, one additional piece of information that can be included for calibration is the location and scale of the bounding box. Therefore, in the object detection task, the criterion of a calibrated model is defined as the precision of a prediction given the confidence, class category, and bounding box information,⁵² as in the following equation:

$$P(z = 1 | p \hat{=} \text{conf}, y \hat{=} y, r \hat{=} r) = \text{conf}, \forall \text{conf} \in [0,1], y \in Y, r \in [0,1]^K, \quad (1)$$

where $z = 1$ indicates that the prediction is correct, conf denotes the confidence of prediction, y is the predicted class in the set of all classes denoted by Y , and r is the bounding box information with k dimensions.

The expected calibration error (ECE) is used to measure the uncertainty of the prediction of the deep learning model. The ECE of object detection is calculated by binning the confidence \hat{p} into M equally spaced bins. Samples with different confidence scores fall into corresponding bin m . B_m is the number of samples in a bin, and N is the number of total samples. The Prec is the precision that represents the correct predictions among all predictions as defined in Eq. (4), and the conf denotes the average confidence score of the predictions. The ECE is given by

$$\text{ECE} = \sum_{m=1}^M \frac{|B_m|}{N} |\text{Prec}(m) - \text{conf}(m)|. \quad (2)$$

For confidence calibration, in this work, we take two additional bounding box pieces of information, the center- x and center- y positions, along with the confidence score to calibrate the prediction results using the dependent logistic calibration,⁵² with the multivariate probability density function being used to model the log-likelihood ratio (lr) of the combined input (confidence, bounding box). Taking the correlations between the confidence and bounding box into consideration, the calibration map is defined as g and is given as

$$g(\text{input}) \approx \frac{1}{1 + e^{-\text{lr}(\text{input})}}, \quad \text{lr}(s) = \frac{1}{2} [(s_-^T \Sigma_-^{-1} s_-) - (s_+^T \Sigma_+^{-1} s_+)] + c. \quad (3)$$

For the variables, $s_+ = s - \mu_+$ and $s_- = s - \mu_-$, and $c = \log \left| \frac{\Sigma_-}{\Sigma_+} \right|$, with μ_+ and μ_- as the mean vectors and Σ_- and Σ_+ as the covariance matrices for the incorrect and correct predictions, respectively. As shown in the calibrated predictions block in Fig. 2, a new confidence score for each prediction is obtained by mapping the input to the calibration map g . The ECEs of the prediction results before and after calibration are calculated to test the effect of calibration on model uncertainty.

3 Results

3.1 Experimental Setup

For model development, we use 943 OCT images from 14 OCT specimen segments for a three-fold cross validation. The OCT images were acquired from specimens that contain calcification regions, which include essential information for CAD treatments. Within each OCT volume, B-scans were sampled at an interval of 20 B-scans. Each B-scan has a size of 1024×1500 pixels, corresponding to a space of $1.98 \times 3 \text{ mm}^2$. In the confidence calibration stage, 60% predictions are used to fit the calibration model, with the remaining 40% predictions to be tested. The ground-truth annotations used for training and testing were made under the guidance of the pathologists.

The YOLO was built in Python 3.8, PyTorch 1.10, CUDA 11.1, and NVIDIA RTX 6000, and a pretrained weight⁵³ was used in this work, with a batch size of 8 and a learning rate of 0.001 using the Adam optimizer with a weight decay of 0.01. Two other popular object detection deep learning models were implemented to show the effectiveness of the calibration. A single-shot

multibox detector (SSD)⁵⁴ and faster region-based convolutional neural networks (Faster RCNN)⁵⁵ were built in Python 3.8, PyTorch 1.10, CUDA 11.1, and NVIDIA RTX 6000. The training process of the SSD was started by loading the pretrained weight,⁵⁶ with a batch size of 8 and a learning rate of 0.001 using the stochastic gradient descent optimizer with a momentum of 0.9. Faster RCNN used a pretrained weight⁵⁷ and was trained with a batch size of 8 and a learning rate of 0.0001 using the Adam optimizer with a weight decay of 0.001.

3.2 Object Detection

To evaluate the performance of calcification detection, three metrics, precision, recall, and *f1*-score, are calculated, as given in the following equations:

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (4)$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (5)$$

$$f1\text{-score} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}, \quad (6)$$

where the true positive (TP) means the model correctly predicts the region with calcification, the false negative (FN) is the wrong prediction for the region that has calcification, and the false positive (FP) is the wrong prediction for the region with no calcification. Precision indicates the number of correct predictions among all detections. Recall measures the fraction of correct predictions among ground truths. The *f1*-score is a measure of overall model performance determined by combining precision and recall.

Qualitatively, in Fig. 3, YOLO predicts all calcification in this coronary OCT image. The SSD and Faster RCNN fail to detect the calcification region in relatively lower contrast. The low recall of the SSD and Faster RCNN in Fig. 4 reveals higher FN predictions, which agrees with the observation in Fig. 3.

In addition, as shown in Table 1, the processing speed of YOLO is 140 frames per second (fps), showing that YOLO has great capability for real-time detection, which is especially desirable in the circumstance of processing a large volume of OCT images. The runtimes of the SSD and Faster RCNN are 68 and 35 fps, respectively. The runtime of OCT segmentation

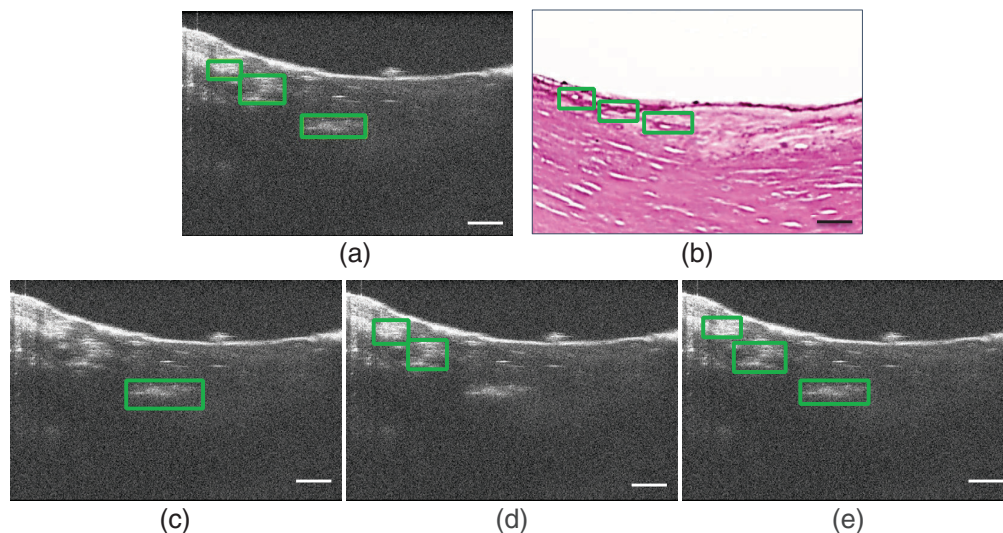


Fig. 3 Example of (a) ground-truth label, (b) corresponding histology, and object detection results from (c) Faster RCNN, (d) SSD, and (e) YOLO. Scale bar: 500 μm .

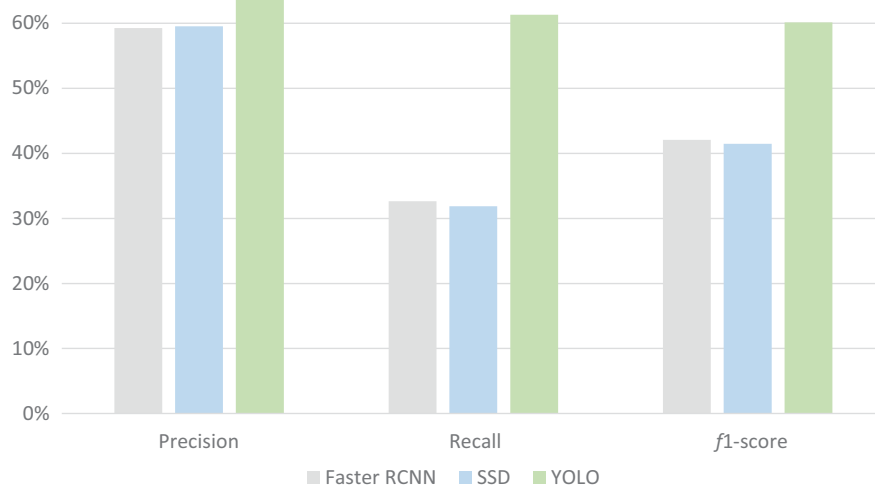


Fig. 4 Object detection results of deep learning models with a threshold of 0.4 in precision, recall, and f_1 score. The gray bars are the results of Faster RCNN, the blue bars are the results of the SSD, and the green bars are the results of YOLO.

Table 1 Runtime in fps for deep learning models detecting calcification in coronary OCT images in this work.

	Faster RCNN	SSD	YOLO
Runtime (fps)	35	68	140

of DeepRetina⁵⁸ and CNN-S⁵⁹ is ~ 5 fps, which indicates a larger computational burden than detection.

3.3 Uncertainty Measurement and Confidence Calibration

We evaluate the effectiveness of the calibration of predictions for the deep learning models. In Fig. 5, the adjustment of confidence scores is observed in the calibrated predictions. In Figs. 5(c) and 5(d), the predictions from Faster RCNN and the SSD in the red box show that the confidence score is slightly adjusted. In Fig. 5(e), the overconfident predictions shown in the yellow box

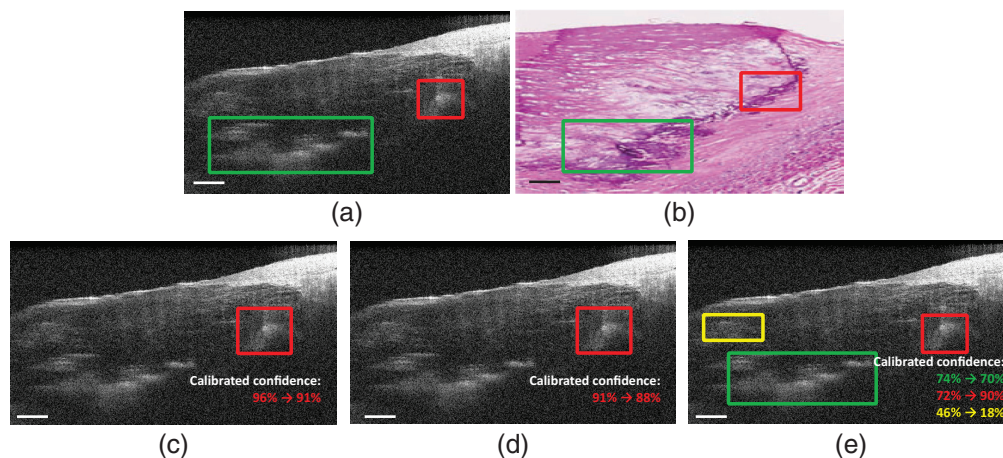


Fig. 5 Example result: (a) ground-truth label, (b) corresponding histology, and confidence calibration results of (c) Faster RCNN, (d) SSD, and (e) YOLO. Scale bar: 500 μm .

Table 2 Uncertainty measurements of confidence calibration in ECE. The rows of before/after calibration shows the changes in ECE during the calibration.

	Faster RCNN	SSD	YOLO
ECE	0.429	0.731	0.233
Calibrated ECE	0.151	0.134	0.146
Before/after calibration	0.278	0.585	0.099

reduces the confidence score from 46% to 18% after calibration, whereas the other confidence scores of predicted boxes are slightly adjusted.

For quantitative evaluation, we use the ECE to measure the uncertainty using the mean value of all test targets. In Table 2, before calibration, YOLO has a lower level of uncertainty in ECE, indicating that YOLO produces more reliable predictions. For the three deep learning models, the calibration errors are lowered to the same level around ~ 0.14 after calibration, which shows the effectiveness of the calibration process that helps rectify the overconfident predictions.

4 Discussion and Conclusion

In this work, we reported calcification detection in coronary OCT images using deep learning models with uncertainty measurements and confidence calibration to reduce the bias in deep learning models. Although tissue detection and segmentation in OCT images have been studied, to our best knowledge, this work is the first to implement uncertainty measurement and confidence calibration for deep learning-based calcification detection in coronary OCT images. We investigated the calcification detection performance of deep learning object detection models and evaluated the reliability of predictions by detection accuracy and uncertainty measures. With an exceptional runtime of 140 fps, YOLO had the potential to become the real-time detector for predicting calcification in coronary OCT images. This work also implemented confidence calibration by integrating the bounding box information with the confidence score. The quantitative and qualitative results showed the effectiveness of the calibration, indicating its practical value in safe-critical and risk-sensitive applications, for example, the calcification detection in coronary OCT images during PCI.

In the future, we will implement other calibration methods on the predicted confidence score and seek to ensemble multiple models to produce more robust and reliable predictions for calcification detection in OCT images. Furthermore, by providing additional information critical to diagnosis, the calibrated confidence and uncertainty measures can be used in future clinical practice.

Disclosures

All authors declare that they have no conflicts of interest.

Acknowledgements

The authors would like to thank Dr. Dezhi Wang from the University of Alabama, Birmingham, for histology service. This work was supported in part by the National Science Foundation (Grant No. CR11-1948540), New Jersey Health Foundation, and National Center for Advancing Translational Research of the National Institutes of Health (Award No. UL1TR003096).

Data Availability Statement

The datasets generated and analyzed in this work are available from the corresponding author upon reasonable request.

References

1. M. Nichols et al., "Cardiovascular disease in Europe: epidemiological update," *Eur. Heart J.* **34**(39), 3028–3034 (2013).
2. A. S. Go et al., "Heart disease and stroke statistics—2014 update: a report from the American Heart Association," *Circulation* **129**(3), e28–e292 (2014).
3. K. Okrainec, D. K. Banerjee, and M. J. Eisenberg, "Coronary artery disease in the developing world," *Am. Heart J.* **148**(1), 7–15 (2004).
4. I. Moussa et al., "Impact of coronary culprit lesion calcium in patients undergoing paclitaxel-eluting stent implantation (a TAXUS-IV sub study)," *Am. J. Cardiol.* **96**(9), 1242–1247 (2005).
5. M. Farag et al., "Treatment of calcified coronary artery lesions," *Expert Rev. Cardiovasc. Ther.* **14**(6), 683–690 (2016).
6. R. Kawaguchi et al., "Impact of lesion calcification on clinical and angiographic outcome after sirolimus-eluting stent implantation in real-world patients," *Cardiovasc. Revasc. Med.* **9**(1), 2–8 (2008).
7. L. K. Kim et al., "Rate of percutaneous coronary intervention for the management of acute coronary syndromes and stable coronary artery disease in the United States (2007 to 2011)," *Am. J. Cardiol.* **114**(7), 1003–1010 (2014).
8. F. Migliavacca et al., "Mechanical behavior of coronary stents investigated through the finite element method," *J. Biomech.* **35**(6), 803–811 (2002).
9. V. Sinitsyn et al., "Relationships between coronary calcification detected at electron beam computed tomography and percutaneous transluminal coronary angioplasty results in coronary artery disease patients," *Eur. Radiol.* **13**(1), 62–67 (2003).
10. Y. Gharaibeh et al., "Coronary calcification segmentation in intravascular OCT images using deep learning: application to calcification scoring," *J. Med. Imaging* **6**(4), 045002 (2019).
11. E. R. Edelman and C. Rogers, "Pathobiologic responses to stenting," *Am. J. Cardiol.* **81**(7), 4E–6E (1998).
12. G. D. Dangas et al., "In-stent restenosis in the drug-eluting stent era," *J. Am. Coll. Cardiol.* **56**(23), 1897–1907 (2010).
13. K. Fujii et al., "Stent underexpansion and residual reference segment stenosis are related to stent thrombosis after sirolimus-eluting stent implantation: an intravascular ultrasound study," *J. Am. Coll. Cardiol.* **45**(7), 995–998 (2005).
14. G. F. Attizzani et al., "Mechanisms, pathophysiology, and clinical aspects of incomplete stent apposition," *J. Am. Coll. Cardiol.* **63**(14), 1355–1367 (2014).
15. H. Doi et al., "Impact of post-intervention minimal stent area on 9-month follow-up patency of paclitaxel-eluting stents: an integrated intravascular ultrasound analysis from the TAXUS IV, V, and VI and TAXUS ATLAS workhorse, long lesion, and direct stent trials," *JACC: Cardiovasc. Interv.* **2**(12), 1269–1275 (2009).
16. M.-K. Hong et al., "Intravascular ultrasound predictors of angiographic restenosis after sirolimus-eluting stent implantation," *Eur. Heart J.* **27**(11), 1305–1310 (2006).
17. N. G. Uren et al., "Predictors and outcomes of stent thrombosis. An intravascular ultrasound registry," *Eur. Heart J.* **23**(2), 124–132 (2002).
18. S. Mehrotra, S. Mishra, and G. Paramasivam, "Imaging during percutaneous coronary intervention for optimizing outcomes," *Indian Heart J.* **70**, S456–S465 (2018).
19. G. Richards and T. Johnson, "A vision of percutaneous coronary revascularisation in 2021: how to take advantage of intra-coronary imaging to perform more effective PCI," *JRSM Cardiovasc. Dis.* **10**, 204800402110499 (2021).
20. A. Fujino et al., "A new optical coherence tomography-based calcium scoring system to predict stent underexpansion," *EuroIntervention* **13**(18), 2182–2189 (2018).
21. F. Prati et al., "Expert review document part 2: methodology, terminology and clinical applications of optical coherence tomography for the assessment of interventional procedures," *Eur. Heart J.* **33**(20), 2513–2520 (2012).
22. H. G. Bezerra et al., "Intracoronary optical coherence tomography: a comprehensive review: clinical and research applications," *JACC: Cardiovasc. Interv.* **2**(11), 1035–1046 (2009).

23. W. Wu et al., “Automatic detection of coronary artery stenosis by convolutional neural network with temporal constraint,” *Comput. Biol. Med.* **118**, 103657 (2020).
24. Y. Tong et al., “Application of machine learning in ophthalmic imaging modalities,” *Eye Vision* **7**(1), 1–15 (2020).
25. L. S. Athanasiou et al., “A deep learning approach to classify atherosclerosis using intracoronary optical coherence tomography,” *Proc. SPIE* **10950**, 109500N (2019).
26. A. Abdolmanafi et al., “Characterization of coronary artery pathological formations from OCT imaging using deep learning,” *Biomed. Opt. Express* **9**(10), 4936–4960 (2018).
27. J. Lee et al., “Segmentation of coronary calcified plaque in intravascular OCT images using a two-step deep learning approach,” *IEEE Access* **8**, 225581–225593 (2020).
28. Y. Avital et al., “Identification of coronary calcifications in optical coherence tomography imaging using deep learning,” *Sci. Rep.* **11**, 11269 (2021).
29. H. Shibutani et al., “Automated classification of coronary atherosclerotic plaque in optical frequency domain imaging based on deep learning,” *Atherosclerosis* **328**, 100–105 (2021).
30. J. Lee et al., “Automated plaque characterization using deep learning on coronary intravascular optical coherence tomographic images,” *Biomed. Opt. Express* **10**(12), 6497–6515 (2019).
31. H. Sun et al., “In vivo detection of plaque erosion by intravascular optical coherence tomography using artificial intelligence,” *Biomed. Opt. Express* **13**, 3922–3938 (2022).
32. A. Nguyen, J. Yosinski, and J. Clune, “Deep neural networks are easily fooled: High confidence predictions for unrecognizable images,” in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 427–436 (2015).
33. H. Wei et al., “Mitigating neural network overconfidence with logit normalization,” in *Int. Conf. Mach. Learn.*, pp. 23631–23644, PMLR (2022).
34. A. A. Mishra et al., “Uncertainty quantification for deep learning in particle accelerator applications,” *Phys. Rev. Accel. Beams* **24**(11), 114601 (2021).
35. A. Nguyen, J. Yosinski, and J. Clune, “Deep neural networks are easily fooled: high confidence predictions for unrecognizable images,” in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 427–436 (2015).
36. J. Platt et al., “Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods,” *Adv. Large Margin Classifiers* **10**(3), 61–74 (1999).
37. B. Zadrozny and C. Elkan, “Obtaining calibrated probability estimates from decision trees and naive Bayesian classifiers,” in *Int. Conf. Mach. Learn.*, vol. 1, pp. 609–616, Citeseer (2001).
38. C. Guo et al., “On calibration of modern neural networks,” in *Int. Conf. Mach. Learn.*, pp. 1321–1330, PMLR (2017).
39. M. Wyatt et al., “Using ensemble methods to improve the robustness of deep learning for image classification in marine environments,” *Methods Ecol. Evol.* **13**(6), 1317–1328 (2022).
40. C. Ju, A. Bibaut, and M. van der Laan, “The relative performance of ensemble methods with deep convolutional neural networks for image classification,” *J. Appl. Stat.* **45**(15), 2800–2818 (2018).
41. S. Cao et al., “Super-resolution technology to simultaneously improve optical & digital resolution of optical coherence tomography via deep learning,” in *2020 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 1879–1882, IEEE (2020).
42. X. Li et al., “Multi-scale reconstruction of undersampled spectral-spatial OCT data for coronary imaging using deep learning,” *IEEE Trans. Biomed. Eng.* **69**(12), 3667–3677 (2022).
43. Z. Hussain et al., “Differential data augmentation techniques for medical imaging classification tasks,” in *AMIA Annu. Symp. Proc.*, Vol. **2017**, p. 979, American Medical Informatics Association (2017).
44. M. F. Kraus et al., “Motion correction in optical coherence tomography volumes on a per A-scan basis using orthogonal scan patterns,” *Biomed. Opt. Express* **3**(6), 1182–1199 (2012).

45. J. Lian et al., “Deblurring retinal optical coherence tomography via a convolutional neural network with anisotropic and double convolution layer,” *IET Comput. Vision* **12**(6), 900–907 (2018).
46. G. Tearney et al., “Rapid acquisition of in vivo biological images by use of optical coherence tomography,” *Opt. Lett.* **21**(17), 1408–1410 (1996).
47. Y. Cui et al., “Imaging artifacts and segmentation errors with wide-field swept-source optical coherence tomography angiography in diabetic retinopathy,” *Transl. Vis. Sci. Technol.* **8**(6), 18 (2019).
48. F. Schwarzhans et al., “Generating large field of view en-face projection images from intra-acquisition motion compensated volumetric optical coherence tomography data,” *Biomed. Opt. Express* **11**(12), 6881–6904 (2020).
49. O. O. Abayomi-Alli et al., “Cassava disease recognition from low-quality images using enhanced data augmentation model and deep learning,” *Expert Syst.* **38**(7), e12746 (2021).
50. A. Sharif Razavian et al., “CNN features off-the-shelf: an astounding baseline for recognition,” in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit. Workshops*, pp. 806–813 (2014).
51. G. Jocher, “YOLOv5 by ultralytics,” <https://github.com/ultralytics/yolov5> (2020).
52. F. Kuppens et al., “Multivariate confidence calibration for object detection,” in *Proc. IEEE/CVF Conf. Comput. Vision and Pattern Recognit. Workshops*, pp. 326–327 (2020).
53. G. Jocher, “YOLO5s pre-trained weight using coco dataset,” <https://github.com/ultralytics/yolov5/releases/download/v7.0/yolov5s.pt> (2020).
54. W. Liu et al., “SSD: single shot multibox detector,” in *Eur. Conf. Comput. Vision*, pp. 21–37, Springer (2016).
55. S. Ren et al., “Faster R-CNN: towards real-time object detection with region proposal networks,” *Adv. Neural Inf. Process. Syst.* **28**, 91–99 (2015).
56. H. Yu et al., “TensorFlow model garden,” <https://github.com/tensorflow/models> (2020).
57. TorchVision maintainers and contributors, “TorchVision: PyTorch’s computer vision library,” <https://github.com/pytorch/vision> (2016).
58. Q. Li et al., “DeepRetina: layer segmentation of retina in OCT images using deep learning,” *Transl. Vision Sci. Technol.* **9**(2), 61 (2020).
59. A. Shah et al., “Multiple surface segmentation using convolution neural nets: application to retinal layer segmentation in OCT images,” *Biomed. Opt. Express* **9**(9), 4509–4526 (2018).

Hongshan Liu received her BS degree in physics from Zhejiang University and her MS degree in electrical engineering from the University of Michigan-Ann Arbor. She is a doctoral student in the Biomedical Engineering Department at Stevens Institute of Technology. Her research focuses on deep learning-based image processing in clinical applications of optical coherence tomography.

Xueshen Li received his MS degree in biomedical engineering from Eindhoven University of Technology. He is a doctoral student in the Biomedical Engineering Department at Stevens Institute of Technology. His research focuses on biomedical image processing.

Abdul Latif Bamba is an undergraduate student from Columbia University. His research is in image processing and microelectronic and solid-state devices.

Xiaoyu Song received her DrPH in biostatistics from Columbia University. She is an assistant professor at Icahn School of Medicine at Mount Sinai. Her research interest is in biostatistics and statistical genomics.

Brigitta C. Brott obtained her cardiology and interventional cardiology training at Duke University Medical Center. She is an interventional cardiologist with a background in materials science and engineering. She is a professor of medicine and biomedical engineering at the University of Alabama at Birmingham. Her research interests include innovative coatings to improve healing after device implantation and optimization of imaging and physiology assessments to guide cardiac interventional procedures.

Silvio H. Litovsky received his MD degree from the University of Buenos Aires, Argentina. He is a professor of pathology at the University of Alabama at Birmingham. His research interest includes multiple areas of cardiovascular pathology including high-risk atherosclerotic plaques.

Yu Gan received his PhD in electrical engineering from Columbia University. He is an assistant professor in the Biomedical Engineering Department at Stevens Institute of Technology. His research interest is in optical coherence tomography, biomedical image processing, and computer vision.