2Pa5-10

Quantitative Uncertainty Assessment of Lumen Detection in Intravascular Ultrasound Images Using Deep Ensembles

Deep Ensembles を用いた血管内超音波画像の血管内腔検出における不確実性の定量評価

Naoya Kanno[1 ‡], Takashi Orihara[1], Hiroyuki Yagami[2], Masanori Kawasaki[3], Munenori Okubo[3], Takuro Ishii[1,4], Hitoshi Matsuo[3], and Yoshifumi Saijo[1] ([1]Grad. School of Biomed. Eng., Tohoku Univ.; [2]Terumo Corporation; [3]Gifu Heart Center; [4]Frontier Research Institute for Interdisciplinary Sciences, Tohoku Univ.)
菅野尚哉 [1 ‡]，折原隆志 [1]，矢上弘之 [2]，川崎雅規 [3]，大久保宗則 [3]，石井琢郎 [1,4]，松尾仁司 [3]，西條芳文 [1]（[1] 東北大院 医工, [2] テルモ株式会社，[3] 岐阜ハートセンター, [4] 東北大 学際科学フロンティア研究所）

1. Introduction

Intravascular ultrasound (IVUS) and the fractional flow reserve (FFR) tests are necessary examinations that assess detailed structures and hemodynamics of the pathological coronary artery and provide useful information to determine the necessity of the invasive percutaneous coronary intervention (PCI) treatment[1)]. Since the FFR test requires inserting pressure sensors into the stenosed artery and administrating drugs to induce cardiac stress, alternative methods that can estimate FFR index without those cumbersome techniques have been studied. One of which is using IVUS images to extract changes in diameter around the pathological artery and estimating the FFR index from those geometrical features[2)]. The major problem of the method is that physicians must manually trace the arterial wall from consequent images as it is demanding and sometimes impossible procedure because the boundary is often unclear or invisible in IVUS images. As a solution, Deep Learning models to automatically trace the vessel wall has been reported. However, the Deep Learning techniques have a problem called overconfidence, where the estimated probability in classification problems is biased towards 0 and 1[3)], and the estimated probability value does not represent its uncertainty. Therefore, the reliability of each extracted boundary has not been considered. This study aimed to develop a machine-learning based method for both detecting the lumen of the coronary artery and quantifying its uncertainty using Deep Ensembles algorithm.

2. Materials and Methods

2.1 Uncertainty Estimation

In order to estimate the uncertainty of a trained deep learning network, a method combining multiple convolutional neural network (CNN) models with the same structure, so-called Deep Ensembles, was employed. **Fig. 1** shows the schematic of Deep Ensembles. Essentially, each trained CNN model has its unique prediction result depending on the initial values. If a feature in the test data set was not sufficiently included in the training data set, extracted boundaries around the feature from multiple models trained with different initial values would show high variability. Deep Ensembles algorithm uses the variability of the predicted probability as the uncertainty map of the extracted boundaries.

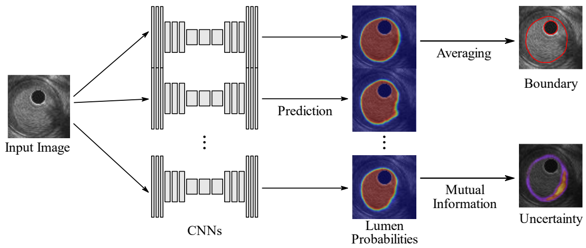

Fig. 1 Schematic of boundary and uncertainty estimation from a IVUS image using Deep Ensembles algorithm.

In this paper, our model was a two-label classification problem, identifying intra- and extra-lumen of the blood vessel to extract the boundary. Since the predicted probability of the lumen is distributed from 0 to 1, high uncertainty means that the probability is close to 0.5. To quantify the characteristics of the uncertainty, mutual information (MI)[4)] was used in this study and expressed as

$$MI = [-\bar{p}\log\bar{p} - (1-\bar{p})\log(1-\bar{p})]$$
$$-\frac{1}{L}\sum_{l=1}^{L}[-p_l\log p_l - (1-p_l)\log(1-p_l)], \quad (1)$$

where $\bar{p}$ was the average of the prediction probabilities, p_l was the l-th lumen prediction

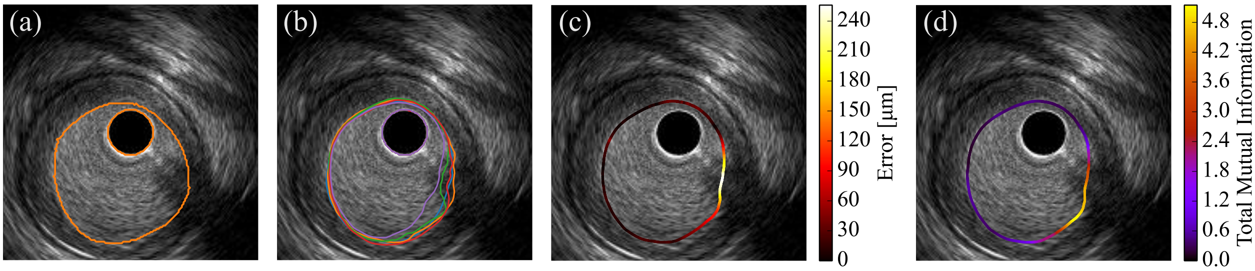

Fig. 2 An example of detecting the lumen and quantifying its uncertainty using Deep ensemble; (a) True boundary, (b) Predicted boundaries from five models trained independently, (c) Error between true and mean predicted boundaries, and (d) Mutual Information as the uncertainty map.

probability, and L was the number of CNN models for Deep Ensembles.

2.2 Experimental Setup

Fifty-three cases of IVUS images consisting of 3,027 images taken in Gifu heart center were used to train and test the deep ensembles model. IVUS imaging was performed using a mechanical IVUS system (VISICUBE for Ueda Japan Radio Corp., Japan) with 60MHz (AltaView) imaging catheters (Terumo Corp., Japan). As preprocessing, the IVUS images were resized from 512×512 to 256×256 pixels and converted to 8-bit grayscale. The true boundaries of the arterial wall in all the images were manually drawn by a skilled technician and then the label images were generated as the inside and the outside of the boundary are to be 1 and 0, respectively. The 53 sequences of IVUS images of the coronary artery of each patient were divided into 40 cases (2,323 images) for the training and 13 cases (704 images) for the testing.

For the deep learning model, we used U-Net, which is a CNN model for semantic segmentation[5]. For Deep Ensembles, five U-Net models were trained independently with a batch size of 16 and a training epoch of 50. The predicted probability was the average of the probabilities of the lumen to be output for each pixel from the models. The boundary between the lumen and the wall was defined as the line where the predicted probability was 0.5.

2.3 Evaluation Criteria

To verify the global accuracy of the predictions, the Intersection over Union (IoU) that evaluates the overlap between the target region and the predicted region was used. In addition, to analyze the local accuracy of the predictions, the difference between the target and prediction boundary was calculated for each angle in the polar coordinates. The target and the prediction boundary between the lumen and the wall were transformed into the polar coordinates, then the error was calculated for each angle. Furthermore, the MI value was transformed into the polar coordinates, and was summed for each angle to analyze the uncertainty for each angle.

naoya.kannno.s8@dc.tohoku.ac.jp

3. Results and Discussion

The mean of the IoU achieved in 5 trained models was 0.824 ± 0.152, indicating that all the models were comparably trained well.

Fig. 2 (a) shows the input image and the boundary. The input IVUS image contained artifacts derived from the guidewire due to multiple reflections and shadows, and Fig. 2 (b) shows the boundaries that the models predicted. The boundaries show the line where the probability predicted by each model was 0.5 and varied in the area where the artifacts occurred. Fig. 2 (c) shows the output boundary from the models and the error from the target boundary. The estimation error was larger in the artifact part. Fig. 2 (d) shows the MI values at each angle. The MI values were high in the areas where artifacts occurred. This was because of the higher mutual information in the areas where the output probability of each model varied. This observation suggested that it would be possible to estimate the uncertainty of the model predictions by looking at the variability of the model estimation results using the mutual information values.

4. Conclusion

In this paper, we proposed a method to estimate the uncertainty of IVUS images using Deep Ensembles. The differences in the prediction results from multiple neural network models suggested that the increased uncertainty due to artifacts can be visualized.

References

1. N. H. J. Pijls *et al.*: Circ. **92** (1995) 3183.
2. F. Seike *et al.*: Circ. J. **82** (2018) 815.
3. C. Guo *et al.*: Proc. 34th Int. Conf. Mac. Learn. PMLR 2017 70 (2017) 1321.
4. L. Smith and Y. Gal: 4th Conf. Uncertain. Artif. Intell. 2018 UAI 2018 **2** (2018) 560.
5. O. Ronneberger *et al.*: Med. Image Comput. Comput. Assist. Interv. MICCAI 2015 **9351** (2015) 234.