Paper Review : Uncertainty estimation and confidence calibration in semantic segmentation with deep convolutional neural networks

Why was this publication interesting to you?

This article interested me because it proposes innovative methods to improve the segmentation of medical images using deep neural networks. By focusing on confidence calibration and predictive uncertainty estimation, the authors aim to strengthen the reliability of model predictions, which is crucial in the medical field where segmentation errors could have serious consequences. . These advances could help improve the accuracy and confidence of results from AI systems used in medical diagnosis and treatment, paving the way for better clinical applications.

- ## Background

This article “Uncertainty estimation and confidence calibration in semantic segmentation with deep convolutional neural networks” constitutes an important contribution in the field of predictive uncertainty estimation and confidence calibration for semantic segmentation. It highlights the importance of the choice of the loss function, proposes a model assembly approach and presents a novel metric to assess the segmentation quality and detect non-distribution entries.
It explores methods to estimate predictive uncertainty for semantic segmentation using FCN . The authors analyze the choice of the loss function used in the formation of FCNs for semantic segmentation, specifically comparing cross-entropy loss and the loss of Dice . They examine the quality of segmentation obtained and the estimation of the predictive uncertainty for each loss function. To improve the confidence of the predictions and the estimation of the uncertainty, the authors propose a method of assembling models. They train multiple FCNs with random parameter initialization and a random shuffling of training data, which improves both the quality of segmentation and estimation of uncertainty. They also introduce a metric based on the average entropy on the predicted segmented object. This metric is used to predict the segmentation quality of foreground structures and to detect out-of-distribution inputs, i.e. data that differs significantly from training data and for which predictions may be less reliable.

However, the inherent uncertainty to the predictions of FCNs can pose problems in the medical context, where the taking of decision must be reliable and justifiable. Segmentation faults can lead to serious consequences for patients, such as misdiagnosis or inappropriate treatment.

It is therefore crucial to assess and quantify the uncertainty associated with model predictions segmentation of the medical image to ensure their reliability and clinical usefulness.

- **Abstract**

The article studies the estimation of the predictive uncertainty and the calibration of the confidence for the fully convolutional neural networks (FCN) used in the segmentation of medical images. The authors shed light on the problem of FCNs poorly calibrated, which produce overly reliable predictions for correct and erroneous classifications, making these models unreliable and difficult to interpret.
The main contributions of the article are as follows: The systematic comparison of the loss of cross-entropy and the loss of dice in terms of segmentation quality and uncertainty estimation for the FCNs. The proposal to use model assembly to calibrate the confidence of FCNs
trained with batch normalization and dice loss. Capacity assessment of NCFs calibrated to predict the quality of segmentation of structures and to detect test examples excluding distribution. Experiments carried out on three image segmentation applications (brain, heart and prostate) show that the models trained with the loss of dice outperform those trained with cross-entropy loss in performance of segmentation.
In addition, the assembly significantly improves the quality of calibration of the models trained with dice loss.

These results offer important perspectives on the estimation of predictive uncertainty and off-distribution detection in medical image segmentation. They also provide practical recommendations for calibrating confidence by demonstrating systematically that the assembly of models improves the calibration. Thus, this study contributes to improving the reliability and interpretability of FCNs for segmentation medical images, which can help practitioners make more clinical decisions informed and justifiable.