Consistent Learning-Based Breast Tumor Segmentation and Its Application in Sentinel Lymph Node Metastasis Prediction

Fengjun Zhao¹, Kaiming Huang¹, Zhipeng Sun¹, Xin Chen², Xiaowei He¹, Bin Wang¹* and Cao Xin¹*

Abstract— Accurate staging of lymph nodes provides crucial diagnostic information for breast cancer patients, where segmentation is of great importance by localizing and visualizing the breast tumor of interest. Nevertheless, current segmentation methods perform average when facing large span of tumor sizes, degraded image quality, blurred tumor boundaries, and resulting noise during manual annotation. Therefore, we develop a Multi-scale RepVGG-based Segmentation Network (MPSegNet) to segment breast tumor from MR images. In particular, we construct a consistent learning framework for the MPSegNet to alleviate the impact of noisy labels upon segmentation results. The rationale is that different views covering the same breast tumors are supposed to generate identical segmentation predictions. Then, we predict SLN metastasis given segmented breast tumors, where we evaluate the relationships between the predictive performance and tumor segmentation under different consistencies. The results show the superiority of our method over other state-of-the-art methods. A high consistency among multiple views can boost the segmentation performance during consistent learning. However, the optimal segmentation does not produce the best SLN metastatic prediction results, implying that the dependence of classification upon segmentation needs to be elaborately investigated further.

Clinical Relevance— This study facilitates more accurate segmentation of breast tumors with consistent learning, and provides an initial analysis between tumor segmentation and subsequent prediction of SLN metastasis, which has potential significance for the precise medical care of breast cancer patients.

I. INTRODUCTION

Breast cancer is the most common malignant tumor threatening women all over the world [1]. Accurate staging of lymph nodes provides important diagnostic and prognostic information in the management of patients with breast cancer [2]. Sentinel lymph node (SLN) represents the first affected drainage site in the event of tumor spread, which is clinically determined by biopsy. However, this invasive procedure not only has complications, such as pain, paresthesia and arm swelling, but also suffers from underestimation if not covering the tissue with metastatic tumor cells. Radiomics with noninvasive magnetic resonance imaging (MRI) facilitates the comprehensive analysis of breast tumors. In particular, various radiomic signatures have then been derived to predict breast cancer SLN metastasis in MRI [3-6], including T₁-weighted imaging, T₂-weighted imaging, diffusion-weighted imaging (DWI) and dynamic contrast-enhanced MRI. However, radiomics generally requires a complicated procedure consisting of tumor segmentation, feature extraction, feature selection and prediction model construction. Of note, segmentation can help localize and visualize the breast tumor of interest, which is essential for precise diagnosis and subsequent treatment.

Over the past decade, state-of-the-art deep learning (DL) models such as fully convolutional network (FCN) [7], encoder-decoder structure [8], attention-based network [9] have been widely used in medical image segmentation. Nevertheless, accurate segmentation of breast tumors from MRI remains challenging due to large span of tumor sizes, degraded image quality (e.g. motion artifacts), blurred tumor boundaries, and resulting noise during manual annotation. To cope with these challenges, modified DL models such as cascaded multi-scale encoder-decoder [10] and atrous spatial pyramid pooling (ASPP) [11] were performed to improve the segmentation for different sizes of tumors. Also, active contour models were integrated into DL models to focus on tumor boundary segmentation [12].

Despite showing promising performance, the above DL methods on one hand require complicated network architectures and thus consume many computational resources. Therefore, we propose a consistent learning-based model termed as Multi-scale RepVGG-based Segmentation Network (MPSegNet) to segment breast tumor from MR images. The conventional convolution in our MPSegNet is replaced by a lightweight RepVGG block [13], which prevents overfitting and improves the performance of the DL network. A convolutional block attention module (CBAM) [14] is inserted between two adjacent RepVGG blocks to highlight the salient features. On the other hand, it is difficult for DL models to eliminate the noise included in the human annotated labels. Shi et al. [15] have tried to distill supervision information from both pixel and image levels to solve low-quality labeling issue. Inspired by this study, we construct a consistent learning framework based on the developed MPSegNet to alleviate the impact of noisy labels upon segmentation results. The rationale is that different views covering the same breast tumors are supposed to generate identical segmentation predictions. Finally, we perform the prediction of SLN metastasis given the segmented breast tumors, where the relationships between the predictive performance and segmentations under different consistencies were well investigated.

This work was supported in part by Scientific Research Program Funded by Education Department of Shannxi Provincial Government (22IW0387), China Postdoctoral Science Foundation (2019M653717), and Shaanxi International S&T Cooperation Program (2021KW-55).

¹*Corresponding author: wbwn@nwu.edu.cn; caoxin@nwu.edu.cn.

1 Xi’an Key Lab of Radiomics and Intelligent Perception, School of Information Science and Technology, NorthWest University, Xi’an, Shaanxi 710069, China.
2 Department of Radiology, Second Affiliated Hospital of Xi’an Jiaotong University, Xi’an, Shaanxi 710007, China.
II. METHODS

A. Segmentation Model

As shown in Fig. 1, the proposed MPSegNet consists of an encoder, a decoder and a multi-scale fusion (MSF) module. The encoder has 22 RepVGG blocks that are grouped into 1, 2, 4, 14 and 1 corresponding 5 stages of feature abstraction. Each block encompasses two parallel convolutions with 3×3 and 1×1 kernel sizes, respectively, and one identity mapping. Note that the first block in each stage omits the identity mapping. The CBAM is inserted between two adjacent RepVGG blocks to improve the feature representation. The decoder has four upsampling and fusion steps. Each step upsamples the feature maps and concatenates them with the skip-connected feature maps from the encoder, followed by two 3×3 convolutions and ReLU activation. In this way, four fused feature maps denoted as D1, D2, D3 and D4 are sequentially generated. The MSF module generates four side output probability maps given D1-D4 by applying a 3×3 convolution and ReLU mapping. Then, we upsample these probability maps by different rates (8, 4, 2 and 1), and then concatenate them together. After performing 1×1 convolution and sigmoid activation, we finally obtain the output segmentation probability map with the same size as the input image.

B. Consistent Learning

We also perform consistent learning on our MPSegNet model to alleviate the inaccurate segmentation caused by noisy human annotations (labels). As shown in Fig. 2, the consistent learning framework includes two modules: a teacher network and a student network. Considering that breast tumors occupy only a small portion of an MR image, we first crop one central view and multiple adjacent views from the MR image containing entire tumor regions. When feeding the center view into the teacher network, the prediction should be identical with the predictions obtained from the student network by inputting adjacent views. Thus, the predictions of the student network can be used as an extra supervision to drive the teacher network to output consistent segmentation free of noise.

Specifically, one center view \(X\) and 8 adjacent views \(X_1, X_2, \ldots, X_8\) of size of 352×352 are generated from one MRI slice, where the latter are positioned top, bottom, left, right, upper left, lower left, upper right, lower right, respectively, 10 pixels away from the center. In each batch, the teacher and student networks output the predictions \(P\) and \(P_1, P_2, \ldots, P_8\) given views \(X\) and \(X_1, X_2, \ldots, X_8\), respectively. Then we take the average over \(P_1, P_2, \ldots, P_8\) to get the pseudo label \(\hat{P}\), whose entropy \(\beta\) was used to evaluate the uncertainty and determine the probability of correct pixel.

![Figure 1. Overview of the proposed MPSegNet model consisting of an encoder, a decoder and a multi-scale fusion (MSF) module.](image1)

![Figure 2. Consistent learning framework for our model, where X is the center view of slice, and X1-X8 are the adjacent views around the center view.](image2)
Authorized licensed use limited to: Taiyuan University of Technology. Downloaded on December 19, 2023 at 04:26:32 UTC from IEEE Xplore. Restrictions apply.
multiple views can boost the segmentation performance of our
thresholds. The results showed the superiority of our method
segmentation results. Finally, we evaluated the MPSegNet for
interference of noise labels on the model to obtain robust
we used a consistent learning framework to solve the
that can learn the multi-scale features of breast tumors. Then,
tumor segmentation model named MPSegNet and evaluated
precise medical care of breast cancer patients.

IV. CONCLUSION
We have developed a consistent learning-based breast
tumor segmentation model named MPSegNet and evaluated
its application in SLN metastasis prediction. First, we
proposed a new architecture based on RepVGG and CBAM
that can learn the multi-scale features of breast tumors. Then,
we used a consistent learning framework to solve the
interference of noise labels on the model to obtain robust
segmentation results. Finally, we evaluated the MPSegNet for
predicting SLN metastasis under different consistent thresholds. The results showed the superiority of our method
over other state-of-the-art methods. A high consistency among
multiple views can boost the segmentation performance of our
MPSegNet model during consistent learning. However, the
optimal segmentation does not produce the best SLN
metastatic prediction results, which implies that the
dependence of classification upon segmentation needs to be
elaborately investigated further. Overall, this study facilitates
more accurate segmentation of breast tumors and provides an
initial analysis between tumor segmentation and prediction of
SLN metastasis, which has potential significance for the

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Dice</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.7517</td>
<td>97.15</td>
<td>99.75</td>
</tr>
<tr>
<td>0.2</td>
<td>0.7717</td>
<td>96.30</td>
<td>99.76</td>
</tr>
<tr>
<td>0.3</td>
<td>0.7845</td>
<td>95.62</td>
<td>99.76</td>
</tr>
<tr>
<td>0.4</td>
<td>0.7927</td>
<td>94.77</td>
<td>99.77</td>
</tr>
<tr>
<td>0.5</td>
<td>0.8004</td>
<td>93.91</td>
<td>99.77</td>
</tr>
<tr>
<td>0.6</td>
<td>0.8089</td>
<td>92.77</td>
<td>99.79</td>
</tr>
<tr>
<td>0.7</td>
<td>0.8173</td>
<td>91.53</td>
<td>99.78</td>
</tr>
<tr>
<td>0.8</td>
<td>0.8213</td>
<td>90.23</td>
<td>99.79</td>
</tr>
<tr>
<td>0.9</td>
<td>0.8290</td>
<td>87.45</td>
<td>99.80</td>
</tr>
</tbody>
</table>

D. Prediction of SLN Metastasis
The predictive results for SLN metastasis under different
consistent thresholds are shown in Table 3. It is unexpected
that the highest Dice coefficient of 0.8213 (at 0.8 threshold)
generates only suboptimal predictive performance with the
area under the curve (AUC) and accuracy of 0.6920 and
73.91%, respectively. The optimal predictive performance is
obtained when the threshold equals 0.3-0.4, achieving an AUC
and accuracy of 0.7321-0.7634 and 73.21-78.26%,
respectively. We notice that corresponding Dice coefficients
are only 0.7845-0.7927, i.e. the interval of over-segmentation.
Overall, the optimal segmentation may not be directly related to
the accurate prediction of SLN metastasis, which is
noteworthy in future breast cancer diagnosis.

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Dice</th>
<th>AUC</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.7517</td>
<td>0.6380</td>
<td>60.87</td>
</tr>
<tr>
<td>0.2</td>
<td>0.7717</td>
<td>0.6294</td>
<td>65.22</td>
</tr>
<tr>
<td>0.3</td>
<td>0.7845</td>
<td>0.7634</td>
<td>78.26</td>
</tr>
<tr>
<td>0.4</td>
<td>0.7927</td>
<td>0.7321</td>
<td>73.21</td>
</tr>
<tr>
<td>0.5</td>
<td>0.8004</td>
<td>0.6607</td>
<td>69.57</td>
</tr>
<tr>
<td>0.6</td>
<td>0.8089</td>
<td>0.5893</td>
<td>65.22</td>
</tr>
<tr>
<td>0.7</td>
<td>0.8173</td>
<td>0.6607</td>
<td>69.57</td>
</tr>
<tr>
<td>0.8</td>
<td>0.8213</td>
<td>0.6920</td>
<td>73.91</td>
</tr>
<tr>
<td>0.9</td>
<td>0.8290</td>
<td>0.5893</td>
<td>65.92</td>
</tr>
</tbody>
</table>

REFERENCES
Sastre-Gara, et al., "Prediction of tumour involvement in remaining
auxiliary lymph nodes when the sentinel node in a woman with breast
"Preoperative prediction of sentinel lymph node metastasis in breast
cancer based on radiomics of T2-weighted fat-suppression and diffusion-
nomogram for prediction of axillary lymph node metastasis in breast
"Preoperative prediction of sentinel lymph node metastasis in breast
cancer by radiomic signatures from dynamic contrast-enhanced MRI,
auxiliary lymph node metastasis in invasive breast cancer patients: A
comparison of radiomic signatures from multiparametric breast MR
sequences," Journal of Magnetic Resonance Imaging, vol. 50, pp. 1125-
Networks for Semantic Segmentation," in IEEE Conference on
Computer Vision and Pattern Recognition (CVPR), Boston, MA, 2015,
pp. 3431-3440.
Networks for Biomedical Image Segmentation," in 18th International
Conference on Medical Image Computing and Computer-Assisted
Intervention (MICCAI), Munich, GERMANY, 2015, pp. 234-241.
attention-guided dense-upsampling networks for breast mass
segmentation in whole mammograms," Physics in Medicine & Biology,
vol. 65, pp. 55005, Feb 2020.
Cochezet, et al., "Cascaded multi-scale convolutional encoder-decoders
for breast mass segmentation in high-resolution mammograms," in 41st
Annual International Conference of the IEEE Engineering in Medicine
and Biology Society (EMBC), Berlin, GERMANY, 2019, pp. 6738-6741.
Context Enhanced Residual Attention Mechanism for Segmentation of
Breast Tumors in Magnetic Resonance Image. in IEEE International
Conference on Bioinformatics and Biomedicine (BIBM). 2019. San
Diego, CA.
[12] Fang, Z., et al., "Combining a Fully Convolutional Network and an Active
Contour Model for Automatic 2D Breast Tumor Segmentation from
Ultrasound Images. Journal Of Medical Imaging And Health Informatics,
2019, 9(7); p. 1510-1515.
[13] X. Ding, X. Zhang, N. Ma, J. Han, G. Ding and J. Sun, "RepVGG:
Making VGG-style ConvNets Great Again," in 2021 IEEE-CVF
Conference on Computer Vision and Pattern Recognition (CVPR),
Block Attention Module. In: Ferr ari, V., Hebert, M., Sminchises cu, C.,
13728-13737.
Medical Image Segmentation with Noisy Labels. In: et al. Medical
MICCAI 2021. Lecture Notes in Computer Science, vol 12901. Springer,