Improving Mammography Malignancy Segmentation by Designing the Training Process

Abstract

We work on the breast imaging malignancy segmentation task while focusing on the training process instead of network complexity. We designed a training process based on a modified U-Net, increasing the overall segmentation performances by using both, benign and malignant data for training. Our approach makes use of only a small amount of annotated data and relies on transfer learning from a self-supervised reconstruction task, and favors explainability.

Keywords: Mammography, Segmentation, Malignancy Detection, Explainability

1. Introduction

Breast cancer is one of the most frequent cancers, and thus, an important public health issue. X-ray imaging (mammography) is the most common screening step for early breast cancer detection. While a big amount of image data is generated, its use for learning-based CAD solutions is not obvious due to the lack of detailed annotations around the malignant regions. In this work, we propose a method to train a segmentation network from a small amount of annotated images.

Typically, to perform image segmentation with a supervised deep learning approach, the network needs an explicit segmentation mask as a ground truth. For medical imaging, this means that each image shall have delimited regions of interest provided by an expert. In clinical practice, such ground truth is rarely available since its collection is a tedious and time-consuming process. Besides, only malignant images can be labeled pixel-wise, which are only a portion of the available data. However, even with 1/8 of women population being affected by breast cancer (Siegel et al., 2019), a great part of mammography imaging is benign and can not be used for supervised segmentation training.

We address the problem with a two-step training process. We follow-up on the work of (Zhou et al., 2019), which proposes pre-training the neural network for reconstruction in a self-supervised manner, before training it for a same or different application. Such an approach enables taking advantage of large amounts of benign images to initialize the model with prior knowledge, improving the overall performance. As in (Zhou et al., 2019), we train a reconstruction network in self-supervised manner, but adapt it to full-images instead of patches (see Section 2). In addition, we wrapped the backbone U-Net (see Section 2) with a subtraction layer on top of the U-Net (see Fig. 1), taking the output from the input, so the network is still trained for reconstruction, rather than being trained to yield the probability of pixel belonging to a malignant region.

© 2020 A.n. withheld.
2. Methods

Our training process (see Fig. 1) is composed of two steps: i) image-wise self-supervised training on full-sized benign mammograms for the reconstruction task followed by ii) image-wise discriminative reconstruction training on malignant mammograms.

**U-Net for full-size mammograms** We trained the network on the full images (mammograms rescaled to $1536 \times 1536$), instead of patches. Hence, we modified traditional U-Net (Ronneberger et al., 2015) as follows: i) we used separable convolutions (Chen et al., 2018; Qi et al., 2019), ii) we adjusted the skip connections, removing the top long skip connections, and introducing short skip connections in each block (Drozdzal et al., 2016; Szegedy et al., 2016), iii) we used instance normalization and leaky ReLU activations (Isensee et al., 2019).

**Self-supervised training** We adopted the self-supervised training approach (Zhou et al., 2019), with few modifications. However, unlike Zhou et al. who focused on patch-wise training of CT images, we worked with image-wise training on mammograms, so we introduced operations that better fit the task, namely: i) instead of non-linear intensity transformation we used gamma-based one, ii) we excluded local-shuffling as too harmful for mammograms, and iii) we extended in-painting with a range of non-rectangle shapes.

At this stage, we use only benign mammograms, filtered by the patient-wise labels: that is, only confirmed benign patient-files were included. We filtered the dataset according to the patient-wise labels of ACR cancer probability. We kept only the cases classified as ACR1 and ACR2 and remove all the others (ACR3-6) as uncertain or malign.

**Image-wise segmentation** Once the U-Net network was initialized with the reconstruction task on benign mammograms, we trained it on the malignant images (ACR4-6) in a supervised manner with pixel-wise ground truth. However, instead of changing the top activation of the U-Net, we added a layer on top of the U-Net (see Fig. 1), subtracting the output from the input. Thus, we continue to train the backbone network for a reconstruction task, while the whole model was designed to yield malignancy regions segmentation (see Fig. 2).

3. Results

We show the performances on the malignant images of INBreast dataset (Moreira et al., 2012) in Table 1 (80% train, 20% test). We note, that unlike common U-Net implementations (Sun et al., 2020), we combine both, masses and calcification masks.

<table>
<thead>
<tr>
<th>Method</th>
<th>$DICE_{Max}$</th>
<th>$DICE_{Avg}$ (on 10 eps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fully Supervised segmentation (<em>from scratch</em>) (baseline)</td>
<td>0.58</td>
<td>0.44</td>
</tr>
<tr>
<td>Fully Supervised seg. (<em>pre-trained</em>) (Zhou et al., 2019)</td>
<td>0.58</td>
<td>0.57</td>
</tr>
<tr>
<td>Discriminative segmentation (<em>from scratch</em>) (ours)</td>
<td><strong>0.61</strong></td>
<td><strong>0.52</strong></td>
</tr>
<tr>
<td>Discriminative segmentation (<em>with pre-training</em>) (ours)</td>
<td><strong>0.61</strong></td>
<td><strong>0.59</strong></td>
</tr>
</tbody>
</table>
4. Discussion

Our network can be trained on both, benign and malignant images. Such an approach yields better segmentation performances (i.e. higher $DICE_{\text{Max}}$) and acts as a natural regularizer (i.e. higher $DICE_{\text{Avg}}$) since the transition to malignant images is closer to fine-tuning than to transfer learning. Our results show that such duality is beneficial for the segmentation results (see Tab. 1). Working on high-resolution mammograms, our discriminative network is sensitive to both, masses and calcifications (see Fig. 2).

Our approach contributes to the explainability of the output: our network is designed to discriminate malignant regions on the full-sized mammograms instead of yielding a probability of pixel to belong to a malignant region. Our network is trained with full mammograms, so, the spatial information is kept by design. The proposed process is scalable, meaning it can take advantage of both benign and malignant imaging. Furthermore, such a process more naturally integrates with the federated or distributed learning scheme thanks to its efficient data use, thus its ability to switch easier between vendors.

The obtained results may be helpful to the radiologists when performing diagnosis. It also can help radiographers making additional examinations (such as spot localizations or magnification), before interpretation by a radiologist.
Acknowledgments

Acknowledgments withheld.

References


