Deep Learning for Automatic Segmentation of Background Parenchymal Enhancement in Breast MRI

S. Nowakowska[1] ABC@SAMPLE.EDU
[1] *Department of Diagnostic and Interventional Radiology, University Hospital Zürich, Rämistrasse 100, CH-8091 Zürich*
K.Borkowski[1] KAROL.BORKOWSKI@USZ.CH
C. Ruppert[1] CARLOTTA.RUPPERT@USZ.CH
P. Hejduk[1] PATRYK.HEJDUK@USZ.CH
A. Ciritsis[1] ALEXANDER.CIRITSIS@USZ.CH
A. Landsmann[1] ANNA.LANDSMANN@USZ.CH
M. Marcon[1] MAGDA.MACRON@USZ.CH
N. Berger[1] NICOLE.BERGER@USZ.CH
A. Boss[1] ANDREAS.BOSS@USZ.CH
C. Rossi[1] CRISTINA.ROSSI@USZ.CH

Abstract

Contrast-enhanced breast MRI plays a crucial role in the care of women at high risk of developing breast cancer. Contrast agent uptake in the breast tissue, i.e., Background Parenchymal Enhancement (BPE), may be an indicator of a higher risk of developing breast cancer and may limit the detectability of lesions. Not only the degree, but also the area of enhancement are elements of importance in the decision-making process in each case. However, they rely on the visual assessment of the reader and thus suffer from poor reliability and reproducibility. In this study, we have developed and evaluated a deep learning (DL) multiclass algorithm for segmentation of both: the BPE area and the non-enhancing tissue. For training, validation, and testing 3441 slices were used. The mean Dice Similarity Coefficient (DSCmean) for the test set amounted to 0.76. Our results show that accurate BPE segmentation is feasible with DL for all classes of enhancement. Such an algorithm may be implemented as part of a pipeline for precise BPE classification or may find direct clinical application in the management of high-risk patients in breast MRI.
Keywords: breast MRI, segmentation, 2D U-Net, BPE, BI-RADS, breast cancer

1. Introduction

MRI is a main modality in breast cancer diagnostics: due to its superior soft tissue contrast, it is used for diagnostics in high-risk patients, as well as for cancer treatment planning, monitoring and assessment. The European Society of Breast Imaging (EUSOBI) now also recommends this modality for screening of patients with dense breast tissue. Predominantly, Dynamic Contrast Enhanced (DCE) sequences are used, in which contrast uptake of the breast tissue, i.e., BPE, and lesions, in case those are present, is assessed. As high BPE is suggested to be correlated with the higher risk of breast cancer development, and it negatively influences tumor detectability, its assessment is crucial for the determination of further medical procedures for the individual patient. Hence, the American College of Radiology has included the BPE classification into the Breast Imaging Reporting & Data System (BI-RADS) atlas. The current approach relies on the visual assessment of enhanced images and their subsequent classification into four classes: “minimal”, “mild”, “moderate” and “marked”. However, such an assessment is prone to high inter-reader variability. (Grimm, 2015) Thus, standardized automatic BPE classification is of utmost importance. For this task, few approaches have been demonstrated in the past, including convolutional neural networks (Borkowski, 2020) or radiomic feature extraction followed by classification (Nam, 2021). To achieve highly reliable classification, precise segmentation of the BPE is crucial. Moreover, such segmentation may serve as input for further clinical studies concerned with e.g. factors influencing the BPE intensity and area. Thus, in this study, we have developed and evaluated a deep learning (DL) multiclass segmentation algorithm segmenting both, the enhancing tissue, i.e. the BPE, as well as the non-enhancing tissue directly from the first subtraction of the DCE sequence.

2. Materials and Methods

A subset of a dataset described in (Borkowski, 2020) was used: the first post-contrast subtraction data from breast MRI-DCE sequence for 38 patients from our institution. In total 3441 slices, not depicting tumor, were used for the deep learning algorithm training, validation and testing. BPE classification was performed slice-wise by two consenting radiologists with more than 5 years of experience in breast imaging. The ground truth masks were created in semi-automatic way using 3D Slicer v4.11.20210226 (Fedorov, 2012). Firstly, the breast was segmented with the use of the grow-from-seeds algorithm. In the second step, the skin was removed from the obtained mask. Lastly, the BPE was segmented by thresholding. The step of skin removal is important, as the skin has similar intensity as BPE and thus would be included in the BPE selection done by thresholding.

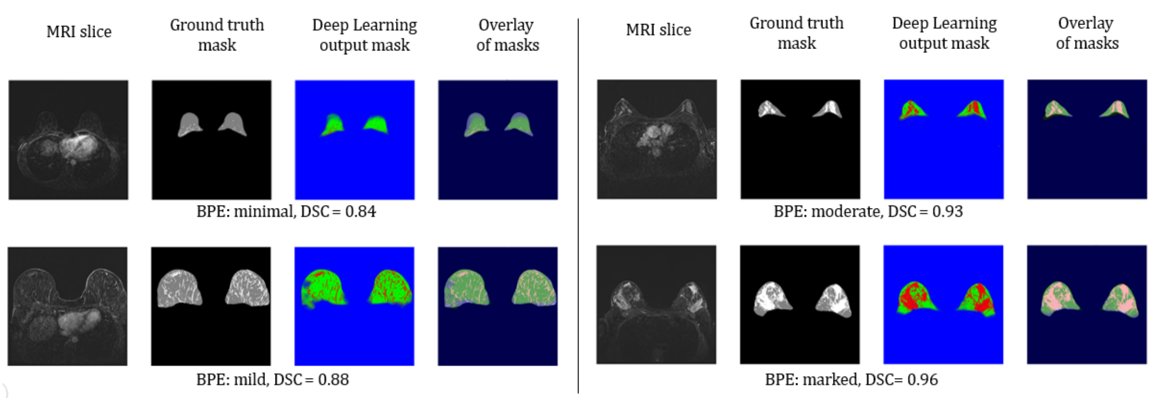

Figure 1: Comparison of the output of the deep learning segmentation algorithm and the ground truth presented for four MRI slices with different BPE classes. The DSC for each case is reported. Further hyperparameter tuning will be performed to optimize the model performance.

The dataset was split into training, validation and test sets containing 74%, 13% and 13% of the data, respectively. It was ensured that slices of a given patient only occurred in one of the sets and that each set contained slices from each BPE class. A hyperparameter tuning of a DL multi-class segmentation algorithm based on the U-Net architecture (Ronneberger, 2015) was performed. The resulting models were trained for 100 epochs. The model from an epoch, characterized by the lowest validation loss was chosen. Each model was evaluated by calculation of mean Dice Similarity Coefficient (DSCmean) score for the test set.

3. Results

In Figure 1, the performance of the model characterized by the overall DSCmean of 0.76 is visualized by comparison of the output mask with the ground truth mask for four chosen slices, each belonging to different BPE class. For each case, the corresponding DSC is reported. The overlays show that the higher the BPE class, the more precise the segmentation, which holds true for the whole test dataset: the DSCmean for slices classified as "minimal" amounts to 0.70 (177 slices), "mild" to 0.76 (124 slices), "moderate" to 0.80 (83 slices) and "marked" to 0.90 (66 slices).

4. Conclusion

Accurate segmentation of enhancing and non-enhancing tissue in breast MRI directly from subtraction images from DCE sequence can be achieved with the use of the DL algorithm. Such an algorithm constitutes an important building block of a pipeline for automatic assessment of breast MRI data, as its output can be further used for e.g., BPE classification task with the use of deep learning convolutional neural networks (CNNs) or for radiomic features extraction followed by classification.

Acknowledgments

This study was supported by the UZH CRPP AI in Oncological Imaging programm.

References

K. et al. Borkowski. Fully automatic classification of breast mri background parenchymal enhancement using a transfer learning approach. *Medicine (Baltimore)*, 99:e21243, 2020.

A. et al. Fedorov. 3d slicer as an image computing platform for the quantitative imaging network. *Magn. Reson. Imaging*, 30:1323–1341, 2012.

L. J. et al. Grimm. Interobserver variability between breast imagers using the fifth edition of the bi-rads mri. *Lexicon. Am. J. Roentgenol.*, 204:1120–1124, 2015.

Y. et al. Nam. Fully automatic assessment of background parenchymal enhancement on breast mri using machine-learning models. *J. Magn. Reson. Imaging*, 53:818–826, 2021.

O. et al. Ronneberger. U-net: Convolutional networks for biomedical image segmentation. in medical image computing and computer-assisted intervention. *MICCAI*, 9351:234–241, 2015.