# Uncertainty-Aware Pectoral Muscle Segmentation Based on Heteroscedastic Regression

Paul Zech<sup>1,2</sup> Christian Hümmer<sup>1</sup> Christopher Syben<sup>1</sup> Ludwig Ritschl<sup>1</sup> Sebastian Stober<sup>2</sup> P.ZECH@SIEMENS-HEALTHINEERS.COM CHRISTIAN.HUEMMER@SIEMENS-HEALTHINEERS.COM CHRISTOPHER.SYBEN@SIEMENS-HEALTHINEERS.COM LUDWIG.RITSCHL@SIEMENS-HEALTHINEERS.COM STOBER@OVGU.DE

<sup>1</sup> Siemens Healthineers AG, Forchheim, Germany

<sup>2</sup> Otto-von-Guericke University, Magdeburg, Germany

Editors: Accepted for publication at MIDL 2025

### Abstract

Accurate segmentation of the pectoral muscle is crucial for improving breast cancer diagnosis in mammograms. While modern deep learning models excel in segmentation, they often lack uncertainty quantification, which is essential for reliable clinical decisions. In this work, we propose a novel method for modeling uncertainty in pectoral muscle segmentation by combining the prediction of probabilistic heatmaps with heteroscedastic regression. For that, we investigate both an existing and a novel loss function derived from the heteroscedastic Laplace distribution, and show that our loss function is more robust for pectoral muscle segmentation in our setting. Further, we demonstrate that our method is capable of producing heatmaps with high-likelihood predictive distributions within a single model while outperforming an ensemble baseline in terms of accuracy.

Keywords: Pectoral muscle segmentation, uncertainty, heteroscedastic regression

## 1. Introduction

Breast cancer is one of the leading causes of mortality among all cancer types (Siegel et al., 2016). To improve early detection, digital mammography is the predominant method for screening patients due to its high cost-efficiency and effectiveness. In mammography image analysis, one important preprocessing step is the segmentation of the pectoral muscle (PM), for which various methods were introduced, ranging from traditional concepts that model the PM boundary (Ferrari et al., 2004) to modern deep learning based solutions that predict the PM region (Rampun et al., 2019), but they typically do not model uncertainty.

A common approach to modeling uncertainty is the assessment of variability in the predictive distribution generated by model ensembles. In the context of PM segmentation, these ensembles can be obtained using Monte Carlo dropout (Klanecek et al., 2023), different model states at various training stages (Tang et al., 2025) or training on different data subsets (Huemmer et al., 2024). However, ensembles primarily capture (epistemic) uncertainty in the model parameters rather than input-dependent heteroscedastic uncertainty (Cipolla et al., 2018; Chan et al., 2024), which is critical when the PM is partially undetectable due to dense glandular tissue and structural similarity to skin folds. In such ambiguous situations, modeling input-dependent uncertainty is crucial, as it is more informative than an overconfident yet inaccurate prediction. Hence, we propose a novel PM segmentation approach based on heteroscedastic uncertainty to capture local variations in noise and occlusions, requiring only a single model training.



Figure 1: Illustration of our method, where the orange line shows the row-wise mean and the uncertainty band represents the row-wise variance of the predicted heatmap.

#### 2. Methods

This paper builds on the work of Huemmer et al. (2024), who proposed to perform the PM segmentation as a regression task by leveraging a unique row-to-column index mapping to directly predict the PM boundary as a column-index (CI) vector. As a major difference, we employ an encoder-decoder architecture (U-Net) to perform heatmap regression through row-wise application of the soft-argmax operation (Luvizon et al., 2019), as shown in Figure 1. The integrated softmax operation enables a probabilistic interpretation and allows to compute relevant row-wise statistics from the heatmap

$$\hat{\mu}_i = \sum_j j \cdot h_{i,j}, \qquad \hat{\sigma}_i^2 = \sum_j h_{i,j} \cdot (j - \hat{\mu}_i)^2 \tag{1}$$

with mean  $\hat{\mu}_i$  as PM boundary prediction, variance  $\hat{\sigma}_i^2$  as measure for the uncertainty, and  $h_{i,j}$  as value of the softmax-activated heatmap at row *i* and column *j*. To train the uncertainty aware heatmap regression we utilize the  $\beta$ -NLL loss of Seitzer et al. (2022) which extends the standard heteroscedastic regression loss with a variance weighting for more stable convergence. It is derived from the negative log-likelihood (NLL) of the normal distribution and thus denoted as  $\mathcal{L}_{\beta-\mathcal{N}-\mathrm{NLL}}$ . Initial experiments empirically revealed that squared error terms, such as the mean-squared-error (MSE) can lead to poor mean predictions compared to the mean-absolute-error (MAE). Hence, we propose a similar loss function  $\mathcal{L}_{\beta-\mathrm{Lap-NLL}}$  that is derived from the Laplace distribution and leads to an absolute difference term in the heteroscedastic loss:

$$\mathcal{L}_{\beta-\text{Lap-NLL}} := \frac{1}{N} \sum_{i=1}^{N} \left\lfloor \hat{b}_{i}^{\beta} \right\rfloor \left( \frac{|y_{i} - \hat{\mu}_{i}|}{\hat{b}_{i}} + \log(2\hat{b}_{i}) \right) \quad \text{with } \hat{b}_{i} = \sqrt{\frac{\hat{\sigma}_{i}^{2}}{2}}.$$
 (2)

Here,  $\lfloor . \rfloor$  denotes a gradient-stop operation and  $y_i$  refers to the target at row *i*. Equivalent to  $\mathcal{L}_{\beta-\mathcal{N}-\mathrm{NLL}}$  (Seitzer et al., 2022),  $\hat{b}_i^{\beta}$  is multiplied to the NLL to interpolate between uniform sample weighting and the NLL loss by selecting  $\beta$  between 0 and 1.

Furthermore, we extend the loss function regularization in (2), inspired by the static variance regularization of Nibali et al. (2018), by using

$$\mathcal{L}_{\text{reg}} = \frac{1}{N} \sum_{i=1}^{N} \mathcal{D}_{\text{JS}}(\mathbf{h}_i \| \mathcal{Q}_i) \quad \text{with} \quad \mathcal{Q}_i = \begin{cases} \mathcal{N}(\hat{\mu}_i, \hat{\sigma}_i^2) & (\text{Gaussian}) \\ \text{Lap}(\hat{\mu}_i, \hat{b}_i) & (\text{Laplace}) \end{cases}$$
(3)

where  $\mathcal{D}_{\text{JS}}(\mathbf{h}_i \parallel \mathcal{Q}_i)$  denotes the Jensen-Shannon divergence between the softmax-activated heatmap  $\mathbf{h}_i$  (row index *i*) and a template distribution  $\mathcal{Q}_i$ , constructed from the predicted heatmaps' statistics in (1). The regularization term of (3) is weighted by a constant factor  $\lambda = 100$  and added to the NLL loss with the goal to produce meaningful distributions that align with the probabilistic model of the underlying loss function.

#### 3. Experiments and results

We evaluate our method on 8861 mediolateral oblique (MLO) mammograms (both left and right laterality) from the MBTST dataset (Dahlblom et al., 2019). Ground truth labels were provided by clinical experts. For independent testing, 15% of the images were extracted while keeping patient boundaries. Mammograms were resized to  $128 \times 128$ , normalized to range of [0,1], and edge-padded by 1/4 of their spatial dimensions for boundary context. Left-lateral images were horizontally flipped for consistent right-laterality. We trained our method using the  $\mathcal{L}_{\beta-\mathcal{N}-\text{NLL}}$  and  $\mathcal{L}_{\beta-\text{Lap-NLL}}$  with various  $\beta$  parametrizations, and compared it against the CI vector regression ensemble of DenseNets by Huemmer et al. (2024). All architectures were adapted to have  $\sim 450$ k parameters. The results shown in Table 1 are averaged over five runs with different data splits, maintaining the same test set. Aligning with prior assumptions, our method using  $\mathcal{L}_{\beta-\text{Lap-NLL}}$  produces more accurate mean fits, particularly for higher  $\beta$ -values, where  $\beta$  controls the tradeoff between accuracy and loglikelihood (LL). Additionally, it outperforms the ensemble baseline in terms of accuracy while providing high-likelihood predictive distributions within a single model. Upon visual inspection of the test set, we observed an increase in row-wise variance in regions where the muscle outline is obscured, indicating valuable information about the models' uncertainty in difficult situations. This is further supported by the Pearson coefficient, which measures image-wise correlation between absolute error and variance.

|                 |  | 1                   |  | 0  |  |  |
|-----------------|--|---------------------|--|--|--|--|
| Method          | Loss   | $\beta$             | $\mathrm{MAE}\downarrow$   | $\mathrm{RMSE}\downarrow$  | Pearson $\uparrow$   | $LL\uparrow$   |
| Ensemble        | SAE  | _                   | $0.53\pm0.00$  | $0.98\pm0.01$  | $0.70\pm0.01$  | $-87.63 \pm 10.60$   |
| U-Net<br>(Ours) | $\mathcal{L}_{eta	ext{-N-NLL}} \ + \lambda  \mathcal{L}_{	ext{reg}}$   | $0.0 \\ 0.5 \\ 1.0$ | $\begin{array}{c} 0.52 \pm 0.02 \\ 0.48 \pm 0.01 \\ 0.53 \pm 0.04 \end{array}$ | $\begin{array}{c} 1.09 \pm 0.06 \\ 1.01 \pm 0.02 \\ 1.09 \pm 0.05 \end{array}$ | $\begin{array}{c} 0.69 \pm 0.00 \\ 0.69 \pm 0.02 \\ 0.68 \pm 0.03 \end{array}$ | $\begin{array}{c} 51.83 \pm 16.89 \\ \text{-}20.13 \pm 55.95 \\ \text{-}75.79 \pm 48.96 \end{array}$ |
|                 | $\mathcal{L}_{eta	ext{-Lap-NLL}} \ + \lambda  \mathcal{L}_{	ext{reg}}$ | $0.0 \\ 0.5 \\ 1.0$ | $\begin{array}{c} 0.51 \pm 0.02 \\ 0.47 \pm 0.00 \\ 0.46 \pm 0.01 \end{array}$ | $\begin{array}{c} 1.12 \pm 0.03 \\ 1.01 \pm 0.02 \\ 0.98 \pm 0.01 \end{array}$ | $\begin{array}{c} 0.72 \pm 0.01 \\ 0.71 \pm 0.00 \\ 0.72 \pm 0.00 \end{array}$ | $\begin{array}{c} 111.96 \pm 11.94 \\ 104.71 \pm 6.01 \\ 87.18 \pm 10.64 \end{array}$                |

Table 1: Comparison of our method against an ensemble baseline.

# 4. Conclusion

In conclusion, we introduce a novel approach to PM segmentation that models inputdependent uncertainty. Preliminary results are promising and suggest to further validate the methods ability to provide meaningful and reliable uncertainty measures.

**Disclaimer:** The presented methods in this paper are not commercially available and their future availability cannot be guaranteed.

## References

- Matthew A. Chan, Maria J. Molina, and Christopher A. Metzler. Estimating epistemic and aleatoric uncertainty with a single model. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang, editors, Advances in Neural Information Processing Systems, volume 37, pages 109845–109870. Curran Associates, Inc., 2024. URL https://proceedings.neurips.cc/paper\_files/paper/2024/file/ c693c3ff83259aebcd55a41ab19a5d84-Paper-Conference.pdf.
- Roberto Cipolla, Yarin Gal, and Alex Kendall. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. In 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 7482–7491, 2018. doi: 10.1109/CVPR.2018. 00781.
- Victor Dahlblom, Magnus Dustler, and Sophia Zackrisson. Mammograms from cancer screening. Available: AIDA Data Hub, 2019. URL https://datahub.aida.scilifelab.se/10.23698/aida/mbtst-dm.
- R.J. Ferrari, A.F. Frere, R.M. Rangayyan, J. Desautels, and R.A. Borges. Identification of the breast boundary in mammograms using active contour models. *Medical and Biological Engineering and Computing*, 42:201–208, 2004.
- Christian Huemmer, Ramyar Biniazan, Manasi Datar, Martin Kraus, Andreas Fieselmann, and Steffen Kappler. Improved Pectoral Muscle Segmentation in Mammograms through Regression-Based Deep Learning and Knowledge Distillation. In 2024 IEEE International Symposium on Biomedical Imaging (ISBI), pages 1–5, Athens, Greece, May 2024. IEEE. ISBN 979-8-3503-1333-8. doi: 10.1109/ISBI56570.2024.10635448.
- Zan Klanecek, Tobias Wagner, Yao-Kuan Wang, Lesley Cockmartin, Nicholas Marshall, Brayden Schott, Ali Deatsch, Andrej Studen, Kristijana Hertl, Katja Jarm, Mateja Krajc, Miloš Vrhovec, Hilde Bosmans, and Robert Jeraj. Uncertainty estimation for deep learning-based pectoral muscle segmentation via Monte Carlo dropout. *Physics* in Medicine & Biology, 68(11):115007, June 2023. ISSN 0031-9155, 1361-6560. doi: 10.1088/1361-6560/acd221.
- Diogo C. Luvizon, Hedi Tabia, and David Picard. Human pose regression by combining indirect part detection and contextual information. *Computers & Graphics*, 85:15–22, December 2019. ISSN 00978493. doi: 10.1016/j.cag.2019.09.002.
- Aiden Nibali, Zhen He, Stuart Morgan, and Luke Prendergast. Numerical coordinate regression with convolutional neural networks. arXiv preprint arXiv:1801.07372, 2018.
- A. Rampun, K. López-Linares, P.J. Morrow, B.W. Scotney, H. Wang, I.G. Ocaña, G. Maclair, R. Zwiggelaar, M.A.G. Ballester, and I. Macía. Breast pectoral muscle segmentation in mammograms using a modified holistically-nested edge detection network. *Medical image analysis*, 57:1–17, 2019.

- Maximilian Seitzer, Arash Tavakoli, Dimitrije Antic, and Georg Martius. On the pitfalls of heteroscedastic uncertainty estimation with probabilistic neural networks. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=aP0pXlnV1T.
- R.L. Siegel, K.D. Miller, and A. Jemal. Cancer statistics, 2016. CA: a cancer journal for clinicians, 66(1):7–30, 2016.
- Yutao Tang, Yongze Guo, Huayu Wang, Ting Song, and Yao Lu. Uncertainty-Aware Semi-Supervised Method for Pectoral Muscle Segmentation. *Bioengineering*, 12(1):36, January 2025. ISSN 2306-5354. doi: 10.3390/bioengineering12010036.