

Classification of laterality and mastectomy/lumpectomy for breast cancer patients for improved performance of deep learning auto segmentation

Helena Vivancos Bargalló^{a,b} , Line Bjerregaard Stick^b, Stine Sofia Korreman^{b,c}, Camilla Kronborg^{b,c}, Mathias M. Nielsen^b, Asbjørn Christian Borgen^b, Birgitte Vrou Offersen^{d,e}, Ole Nørrevang^b  and Jesper F. Kallehauge^{b,c} 

^aMedical Physics department, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain; ^bDanish Center for Particle Therapy, Aarhus University Hospital, Aarhus, Denmark; ^cDepartment of Clinical Medicine, Aarhus University, Aarhus, Denmark; ^dDepartment of Experimental Clinical Oncology, Aarhus University Hospital, Aarhus, Denmark; ^eDepartment of Experimental Clinical Oncology, Aarhus University Hospital, Aarhus, Denmark

ARTICLE HISTORY Received 25 May 2023; Accepted 3 August 2023

Background

In the last decade, there has been a growing interest in the use of auto-contouring in radiotherapy treatment planning [1–6]. Artificial intelligence and in particular U-net convolutional neural network [7] has been shown to be successful in image segmentation [8] and can be used for the segmentation of both normal tissue and target volumes of breast cancer patients [9–11]. However, due to the symmetry in breast cases, using a single model, regardless of laterality and type of surgery, for normal tissue and target structures in all patients can result in suboptimal target delineation. More specifically target volumes being contoured on the contralateral side or partial delineation of the target volume [10].

In order to address these shortcomings, it can be useful to train different models taking into account the laterality of the disease and the surgical extent (mastectomy/lumpectomy). One potential solution would be to train one model for the segmentation of normal tissue and four models (left mastectomy, left lumpectomy, right mastectomy, and right lumpectomy) for the target volumes [11,12].

When using different models, in order to be able to automate the whole segmentation process, classifiers (left/right, mastectomy/lumpectomy) need to be developed and implemented in the auto-contouring process. Even if the model could be chosen by the user, we believe that implementing an automatic method, without the need of human intervention, will minimise errors.

We have developed an automated pipeline based on the planning CT and two classifiers, which allow to classify the patients into the correct category and, therefore, get the contours automatically delineated by the appropriate model.

Material and methods

The pipeline from the CT to the final contoured volumes includes several steps using information from the previous steps (Figure 1):

1. Autocontouring of the normal tissue structures (Heart, Left Anterior Descending Coronary Artery, Lungs, Humeral Head, Trachea, Oesophagus and Liver) using the CT information and a trained nn-Unet [8] model.
2. Classification between left and right using the CT information.
3. Autocontouring of two sets of target volumes (CTVp, CTVn_IMN, CTVn_interpectoralis, CTVn_L1, CTVn_L2, CTVn_L3, and CTVn_L4 [13]) on the side determined in step 2: one using the lumpectomy model and one using the mastectomy model. In this step, each model also delineates contralateral breast as, like the targets, it also depends on the laterality and, therefore, is better delineated after laterality determination.
4. Classification between lumpectomy and mastectomy based on CT information and the volumes contoured in step 3.

Patient data

To train and validate the method all patients with breast cancer treated at DCPT (Danish Centre for Particle Therapy) before March 2023 excluding those with synchronous bilateral breast cancer were included in the study ($n = 129$) (The cohort characteristics are given in Table 1). All patients' targets were delineated according to the ESTRO consensus guidelines [13] while normal tissue was delineated according to the DBCG (Danish Breast Cancer Group) guidelines [14]. 113 breast cancer patients were stratified according to the laterality of the tumour or surgical extent (for the laterality classifier and the mastectomy/lumpectomy classifier respectively). Then they were used to train and validate (10-fold cross-validation) the two classifiers. A separate set of 16 patients was used to test the whole classification model.

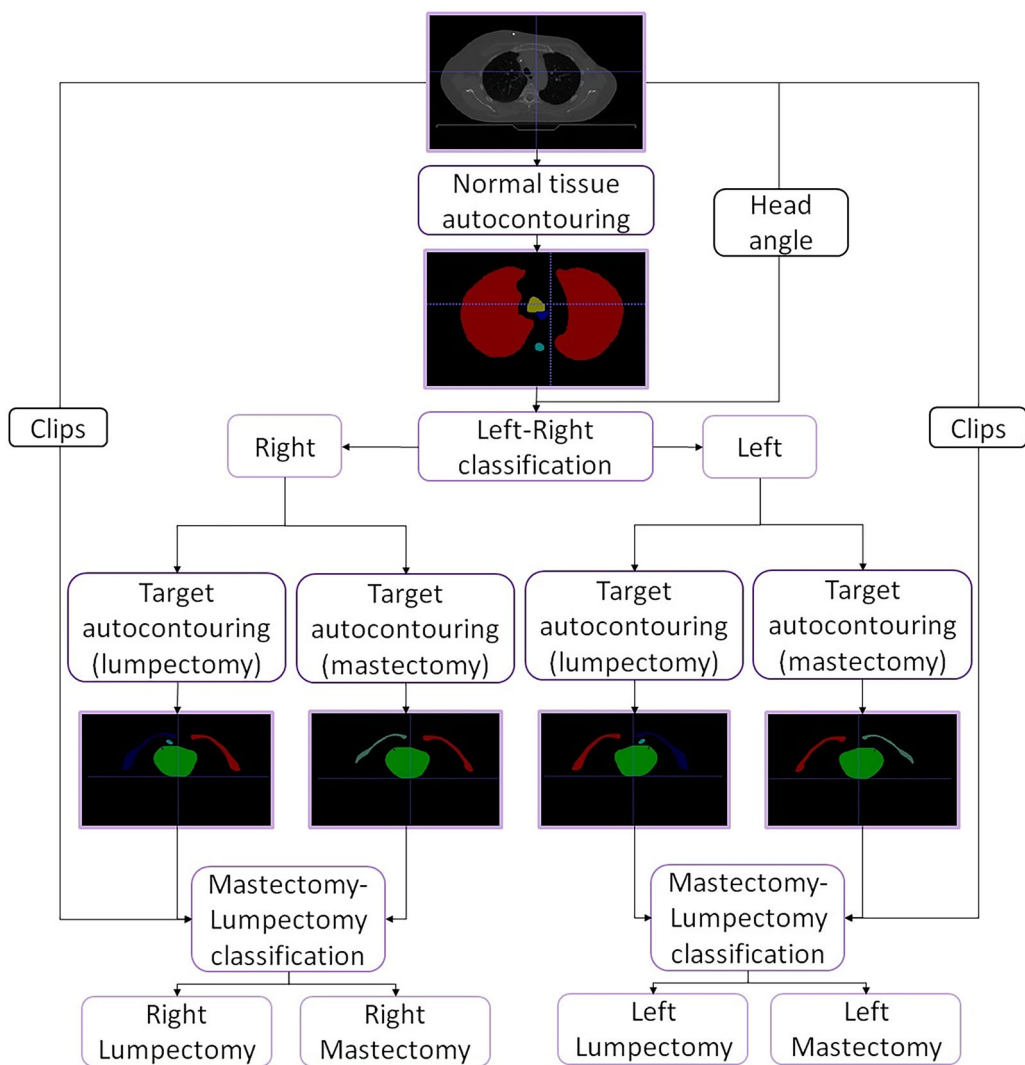


Figure 1. Pipeline design from the CT-scan to the contour of the final volumes. It includes normal tissue delineation, laterality classification, target delineation by the lumpectomy and mastectomy networks for the chosen tumour side and finally mastectomy/lumpectomy determination.

Table 1. Characteristics of the patients included in the study.

Parameter	Train/validation set	Test set
Number of patients	113	16
Age	Median: 56 yr.Range: 26–82 yr.	Median: 63 yr.Range: 24–76 yr.
Dose prescription:		
50 Gy / 25 fx.	33 (29.20%)	1 (6.25%)
50 Gy / 25 fx. + SIB: 57 Gy / 25 fx. (EQD2: 50 Gy + 10 Gy)	3 (2.66%)	0 (0%)
51.52 Gy / 28 fx. + SIB: 63 Gy / 28 fx. (EQD2: 50 Gy + 16 Gy)	1 (0.89%)	0 (0%)
40 Gy / 15 fx. (EQD2: 45 Gy)	63 (55.75%)	13 (81.25%)
40 Gy / 15 fx. + SIB: 45.75 Gy / 15 fx. (EQD2: 45 Gy + 10 Gy)	5 (4.42%)	0 (0%)
42.3 Gy / 18 fx + SIB: 52.2 Gy / 18 fx. (EQD2: 45 Gy + 16 Gy)	8 (7.08%)	2 (12.5%)
Unilateral	113 (100%)	16 (100%)
Previous surgery or RT in the side being treated	0 (0%)	0 (0%)
Category		
Right mastectomy	16 (14.16%)	2 (12.50%)
Left mastectomy	42 (37.17%)	5 (31.25%)
Right lumpectomy	16 (14.16%)	5 (31.25%)
Left lumpectomy	39 (34.51%)	4 (25.00%)

Classification

For both classifiers simple and intuitive features were chosen to classify each patient into the correct category: right mastectomy, left mastectomy, right lumpectomy or left lumpectomy.

We tested two different types of classifiers for both laterality and mastectomy/lumpectomy classification:

1. Stochastic Gradient Descent (SGD) [15] with different loss functions. The SGD is an iterative method for optimising an objective function. The algorithm aims to

minimise the chosen loss function by calculating the gradient and changing the model parameters in the negative gradient direction.

2. Random Forest Classifier [16] which combines several tree predictors and for which the output is the class predicted by more trees.

We performed a grid search to determine the best loss function for Gradient Descent and to fine-tune parameters for both Gradient Descent and Random Forest (Supplementary Materials Table 1, 2, 3 and 4).

Using the results from the grid search we obtained accuracies values for the training set. For the test set a ROC (receiver operating characteristic) analysis [17] was performed, which represents the probability of the model being able to distinguish between the two classes. The AUC (Area Under the Curve) was also obtained, which is a single score representing how well the model performs (1 being the perfect classifier and 0.5 being a random classifier).

Laterality classifier

Institutional position guidelines state that patients with breast cancer should be positioned with both arms up and looking to the contralateral side to expose the lymph nodes on the treated side. The classifier uses a bone HU window on the 10 last slices of the CT on a bounding box including the head. This allows for detection of relevant bone structures, i.e. mandible and vertebrae which are then used to determine the head angle through simple linear regression (Figure 2a). The head angle feature is used to run a grid search with 10 fold cross validation and train a Stochastic Gradient Descent and a Random Forest Classifier with the optimal parameters using Python and scikit-learn [18] (Parameters used are specified in Supplementary Materials Table 1 and 2)

Mastectomy/lumpectomy classifier

From the CT information and the two CTVp model predictions (CTVp_breast for the lumpectomy model and CTVp_chestwall for the mastectomy model) three features are determined:

1. Number of surgical clips inside the CTVp (using the CTVp which includes more clips). To detect the clips, clusters of high HU are searched inside both CTVp's
2. Maximum relation between CTVp volume and contralateral breast volume (Figure 2b: CTVp_breast and 2c: CTVp_chestwall):

$$\frac{CTVp_{Volume}}{Contralateral\ breast\ volume}$$

(Using the CTVp for which this is relation is larger)

3. Relation between the two CTVp's:

$$\frac{CTVp_{breast}}{CTVp_{chestwall}}$$

These three features are used to run a grid search with 10 fold cross validation and train a Stochastic Gradient Descent and a

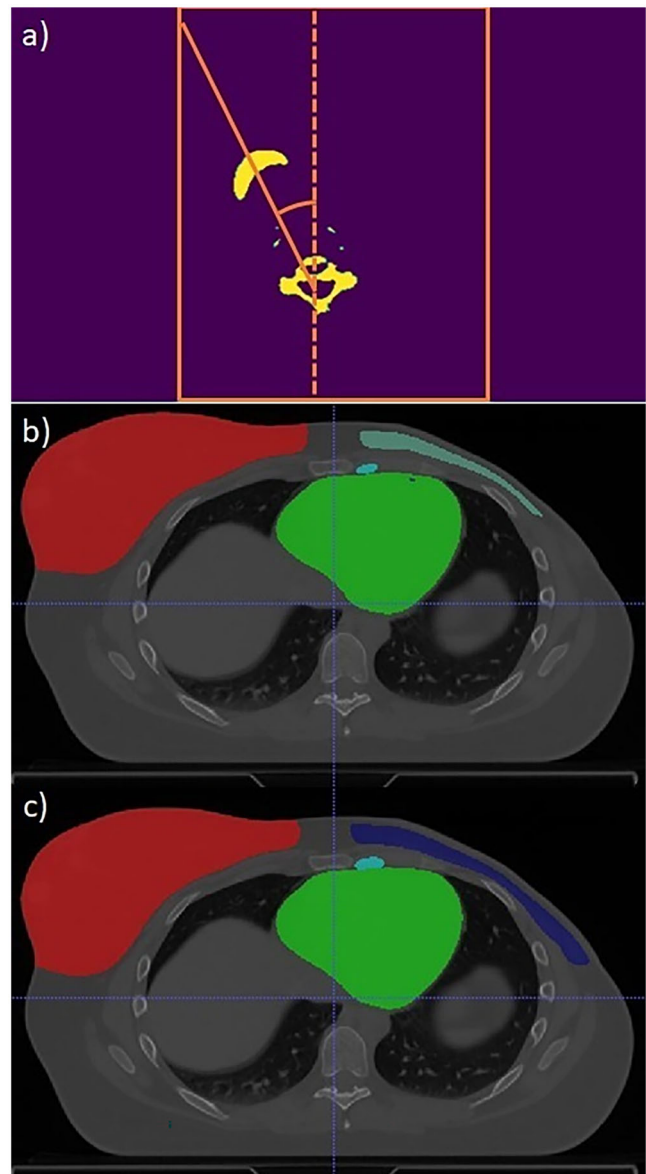


Figure 2. a: Head angle determination: performing a linear regression of the voxels (with bone HU window) of the last CT slices inside a bounding box that includes the head, b: Volumes segmented by the lumpectomy model and c: Volumes segmented by the mastectomy model.

Random Forest Classifier with the optimal parameters using Python (Parameters used are specified in Supplementary Materials Tables 3 and 4).

Results

Laterality classifier

The value of the angle allowed to distinguish between the two intended groups: patients with breast cancer in the left side and patients with breast cancer on the right side (Supplementary Materials Figure 1)

After running a grid search with 10 fold cross validation on the first set of 113 patients to determine the best parameters for each algorithm, the best accuracies regarding

laterality obtained were 100% (0% Standard Deviation (STD)) for both Stochastic Gradient Descent (for some specific loss functions) and Random Forest. The parameters used in the grid search and resulting accuracies and STD are specified in [Supplementary Materials Table 1 and 2](#), the final parameters chosen in the grid search are underlined. For the Stochastic Gradient Descent, the optimal loss function chosen was 'hinge', which gives a linear SVM (Support Vector Machine).

Even though the two models have very similar and high-level performance, the Random Forest was chosen as the best classifier as it provides an easier-to-understand feedback. In the separate test set 16/16 patients were classified correctly using the Random Forest classifier. A ROC analysis was performed for the Random Forest classifier with a resulting AUC of 1.0 ([Supplementary Material Figure 3](#))

Mastectomy/lumpectomy classifier

The features chosen allowed to distinguish between the two intended groups: patients who had a mastectomy and patients who had a lumpectomy ([Supplementary Materials Figure 2](#)).

After running a grid search on the first set of 113 patients to determine the best parameters for each algorithm, the best accuracies regarding surgical method of obtained were 93.8% (STD: 5.8%) and 98.3% (STD: 3.5%) for Stochastic Gradient Descent and Random Forest respectively. The parameters used in the grid search are specified in [Supplementary Materials Table 3 and 4](#), the final parameters chosen in the grid search are underlined.

In this case, the Random Forest was also chosen as the best classifier. From the separate test set 15/16 patients were classified correctly for the Random Forest. The AUC resulting from the ROC analysis of this classifier is of 0.93 ([Supplementary Material Figure 3](#)).

Discussion and conclusion

For both classifiers, the Random Forest Classifier performed better than or as well as the Stochastic Gradient Descent classifiers trained with different loss functions. Furthermore, even though both the laterality and the mastectomy/lumpectomy classifiers are very simple (one and three features respectively) both have excellent accuracy regarding laterality and surgical extent respectively.

The high accuracy in classifying the patients into the correct category will allow the correct network to contour the volumes instead of being segmented by one more general model. This approach of classification instead of relying in the neural network to choose the side where the targets should be delineated will result in fewer cases with regions delineated on the opposite side [10]

The laterality classifier is potentially dependent on the institutional positioning guidelines of the patients with breast cancer and in the current single centre implementation, the patients are positioned looking to the healthy side with both arms up, with the aim to expose the lymph nodes

on the treatment side. It is still to be investigated how well this generalises to other centres with different patient positioning, i.e. looking straight with one arm up. For differences in patient positioning more features may have to be taken into account which is easily incorporated in the more general approach of using Random Forest classifier instead of traditional rule-based programming. Moreover, the Random Forest allows easily understandable feedback related to the importance of each feature used for the classification.

Furthermore, the laterality classifier also relies on the detection of the bone structures on the head area. When the CT-scan length allows for the detection of both mandible and vertebrae, the head angle feature is easily and reliably determined. On the other hand, if in some centres shorter scans are used, the angle obtained based only on the vertebrae information may not be as accurate. In these cases, more features may also need to be added to the model.

Regarding the mastectomy/lumpectomy classifier, further testing should be performed to include more patients with unusual anatomy such as patients with implants, with seromas or with previous contralateral breast surgery. For these outlier cases the classifier may not performed as well as expected. Moreover, it should be taken into consideration that for this study CT-scans without metal marking of the scar or contralateral breast (mastectomy patients only) were used, as in the local centre these markings were included in a separate low dose CT-scan. If metal markings are present in the CT-scan used for classification special care should be taken in the clip detection method.

The information about laterality and surgical extension can also be given as manually input before performing auto-segmentation. We opted for the classifier approach due to workflow considerations. In the clinical setting, once the CT scan has been performed it is automatically and immediately exported to the segmentation pipeline with no obvious manual interaction possibilities. Both manual interaction and imperfect classifiers are error-prone, and may be improved in the case of automatic classification by expanding the training set to include more extreme cases.

As already commented, potential drawbacks of the automatic classification may arise when changes to departmental patient positioning guidelines or post-surgery clip insertion are introduced. This would require retraining of the algorithm unless a more general set of features are identified covering a wider set of patient scenarios. Finally, the classification of the lumpectomy/mastectomy classification currently uses the prediction from two nn-Unet models which one is not used after this classification step. An implementation that combines this two prediction may reduce the time used for the entire pipeline.

In conclusion, we have shown a stable scheme for the classification of breast cancer patients based on laterality and whether they are lumpectomised or mastectomized.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Helena Vivancos Bargalló  <http://orcid.org/0000-0003-0394-6152>

Ole Nørrevang  <http://orcid.org/0000-0002-2100-0208>

Jesper F. Kallehauge  <http://orcid.org/0000-0003-3705-5390>

Data availability statement

Due to the nature of the research, due to legal reasons the supporting data is not available.

References

- [1] Harrison K, Pullen H, Welsh C, et al. Machine learning for auto-segmentation in radiotherapy planning. *Clin Oncol (R Coll Radiol)*. 2022;; 34(2):74–88. doi: [10.1016/j.clon.2021.12.003](https://doi.org/10.1016/j.clon.2021.12.003).
- [2] Cardenas CE, Yang J, Anderson BM, et al. Advances in auto-segmentation. *Semin Radiat Oncol*. 2019; 29(3):185–197. doi: [10.1016/j.semradonc.2019.02.001](https://doi.org/10.1016/j.semradonc.2019.02.001).
- [3] Vandewinckele L, Claessens M, Dinkla A, et al. Overview of artificial intelligence-based applications in radiotherapy: recommendations for implementation and quality assurance. *Radiother Oncol*. 2020;153:55–66. doi: [10.1016/j.radonc.2020.09.008](https://doi.org/10.1016/j.radonc.2020.09.008).
- [4] Samarasinghe G, Jameson M, Vinod S, et al. Deep learning for segmentation in radiation therapy planning: a review. *J Med Imaging Radiat Oncol*. 2021; 65(5):578–595. doi: [10.1111/1754-9485.13286](https://doi.org/10.1111/1754-9485.13286).
- [5] Kalantar R, Lin G, Winfield JM, et al. Automatic segmentation of pelvic cancers using deep learning: state-of-the-art approaches and challenges. *Diagnostics*. 2021; 11(11):1964. doi: [10.3390/diagnostics11111964](https://doi.org/10.3390/diagnostics11111964).
- [6] Kosmin M, Ledsam J, Romera-Paredes B, et al. Rapid advances in auto-segmentation of organs at risk and target volumes in head and neck cancer. *Radiother Oncol*. 2019; 135:130–140. doi: [10.1016/j.radonc.2019.03.004](https://doi.org/10.1016/j.radonc.2019.03.004).
- [7] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (editors). *Medical image computing and computer-assisted intervention – MICCAI 2015*. MICCAI 2015. Lecture notes in computer science. Vol 9351, Cham: Springer; 2015
- [8] Isensee F, Jaeger PF, Kohl SAA, et al. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat Methods*. 2021;18(2):203–211. doi: [10.1038/s41592-020-01008-z](https://doi.org/10.1038/s41592-020-01008-z).
- [9] Byun HK, Chang JS, Choi MS, et al. Evaluation of deep learning-based autosegmentation in breast cancer radiotherapy. *Radiat Oncol*. 2021; 16(1):203. doi: [10.1186/s13014-021-01923-1](https://doi.org/10.1186/s13014-021-01923-1).
- [10] Oya M, Sugimoto S, Sasai K, et al. Investigation of clinical target volume segmentation for whole breast irradiation using three-dimensional convolutional neural networks with gradient-weighted class activation mapping. *Radiol Phys Technol*. 2021; 14(3):238–247. doi: [10.1007/s12194-021-00620-8](https://doi.org/10.1007/s12194-021-00620-8).
- [11] Men K, Zhang T, Chen X, et al. Fully automatic and robust segmentation of the clinical target volume for radiotherapy of breast cancer using big data and deep learning. *Phys Med*. 2018;50:13–19. doi: [10.1016/j.ejmp.2018.05.006](https://doi.org/10.1016/j.ejmp.2018.05.006).
- [12] Qi X, Hu J, Zhang L, et al. Automated segmentation of the clinical target volume in the planning CT for breast cancer using deep neural networks. *IEEE Trans Cybern*. 2022;52(5):3446–3456. doi: [10.1109/TCYB.2020.3012186](https://doi.org/10.1109/TCYB.2020.3012186).
- [13] Offersen BV, Boersma LJ, Kirkove C, et al. ESTRO consensus guideline on target volume delineation for elective radiation therapy of early stage breast cancer. *Radiother Oncol*. 2015;114(1):3–10. doi: [10.1016/j.radonc.2014.11.030](https://doi.org/10.1016/j.radonc.2014.11.030).
- [14] Nielsen MH, Berg M, Pedersen AN, et al. Delineation of target volumes and organs at risk in adjuvant radiotherapy of early breast cancer: National guidelines and contouring atlas by the Danish Breast Cancer Cooperative Group. *Acta Oncol*. 2013.
- [15] Bottou L. Stochastic gradient learning in neural networks. *Proceedings of Neuro-Nimes 91, EC2*. Nimes, France; 1991.
- [16] Breiman L. Random forests. *Mach Learn*. 2001;45(1):5–32. doi: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324).
- [17] Fawcett T. An introduction to ROC analysis. *Pattern Recog Lett*. 2006;27(8):861–874. doi: [10.1016/j.patrec.2005.10.010](https://doi.org/10.1016/j.patrec.2005.10.010).
- [18] Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in python. *J. Mach. Learn. Res*. 2011;12:2825–2830.