# Simulated image-specific microcalcification clusters and associated mass enhancement to enhance training of a deep learning model for cancer detection in contrast-enhanced mammography

# Simulated image-specific microcalcification clusters and associated mass enhancement to enhance training of a deep learning model for cancer detection in contrast-enhanced mammography

Astrid Van Camp*[a,b], Henry C. Woodruff[a,c], Lesley Cockmartin[d], Nicholas W. Marshall[b,d], Hilde Bosmans[b,d], Philippe Lambin[a,c]
[a]Department of Precision Medicine, GROW – Research Institute for Oncology and Reproduction, Maastricht University, Maastricht, The Netherlands
[b]KU Leuven, Department of Imaging and Pathology, Division of Medical Physics & Quality Assessment, Herestraat 49, 3000 Leuven, Belgium
[c]Department of Radiology and Nuclear Medicine, GROW – Research Institute for Oncology and Reproduction, Maastricht University Medical Centre+, Maastricht, The Netherlands
[d]UZ Leuven, Department of Radiology, Herestraat 49, 3000 Leuven, Belgium
*Contact author: a.vancamp@maastrichtuniversity.nl

## ABSTRACT

We present an automated method to generate synthetic contrast-enhanced mammography cases with simulated microcalcification clusters. This method accounts for existing textures in the breast, with the simulated clusters inserted in the low-energy image. In parallel, potential mass-like enhancement is modelled from real values in the recombined image.

The same deep learning model was trained with different amounts and ratios of real and synthetic data. When trained with real data only, malignant masses are more often correctly detected and classified than malignant microcalcification clusters. The addition of synthetic data with simulated clusters during training could increase detection sensitivity for all types of malignant lesions and maintained similar levels of AUC for classification. This enhanced performance was consistent on both internal and external test sets.

These findings demonstrate the potential applicability of synthetic data to enhance deep learning models, especially when real data are scarce or imbalanced.

**Keywords:** Microcalcification clusters, Mass-like enhancement, Contrast-enhanced mammography, Deep learning, Detection, Classification, Lesion simulation, Breast cancer

## 1. INTRODUCTION

Contrast-enhanced mammography (CEM) has become increasingly prevalent in many hospitals for its improved visualization of breast cancer in the recombined image, facilitated by the uptake of contrast agent[1]. Recently, a number of deep learning models have been developed to detect and classify breast lesions in CEM data. While these models achieve high levels of performance, they focus mainly on masses[2–4] or do not make a specific distinction between lesion types[5–9]. A significant challenge remains, however, in the detection of microcalcification clusters, as calcifications themselves do not take up contrast agent rendering them less conspicuous in the recombined image. These clusters are better visualized in the low-energy image, which is similar to digital mammography.
Limitations in detection sensitivity of microcalcification clusters, especially those with low levels of contrast enhancement, can be (partially) caused by data scarcity and the limited number of complex cases available. Therefore,

17th International Workshop on Breast Imaging (IWBI 2024), edited by Maryellen L. Giger, Heather M. Whitney, Karen Drukker, Hui Li, Proc. of SPIE Vol. 13174, 1317404
 · 0277-786X · doi: 10.1117/12.3026879

the aim of this work is to generate synthetic sets of CEM cases with simulated microcalcification clusters based on breast textures present in the images. Including these simulated clusters in the training of deep learning models can potentially enhance the model performance in terms of detection accuracy and reduced false positives/negative rates. The aim was to build a single deep learning model capable of detecting both masses and microcalcification clusters in a single CEM exam, including both the processed low-energy and recombined image in the evaluation. Models of the same architecture were trained with different numbers of real synthetic images; the ratio of real to synthetic training images was also varied. The performance of these models for detection and classification tasks was then compared.

## 2. MATERIALS AND METHODS

### 2.1. AVAILABLE DATA

This study utilised a dataset of processed CEM images from MUMC+ (Maastricht, the Netherlands), obtained on a GE Senographe Essential system and annotated by an expert radiologist[6]. The annotations included clinical information as well as masks based on lesion delineations and information provided by the radiology and pathology reports. This comprehensive information on the lesion type allowed lesion types to be split between pure masses and lesions with a microcalcification cluster present. Figure 1 illustrates a representative a case with a microcalcification cluster present. The lesion images were used for training, validation, and internal testing of the deep learning models, whereas images of patients without any findings were used to create the synthetic cases. From the lesion images, 20% were withheld for internal testing. Of the remaining lesion images, 80% were used for training and 20% for validation.

In addition, a dataset from Gustave Roussy institute (GR) (Villejuif, France) was at hand, comprising processed images acquired on GE Senographe Essential and Senographe Pristina systems. This was only used as an external test set. Due to limited clinical information available, no distinction could be made between the presence of masses only or lesions with calcifications present.

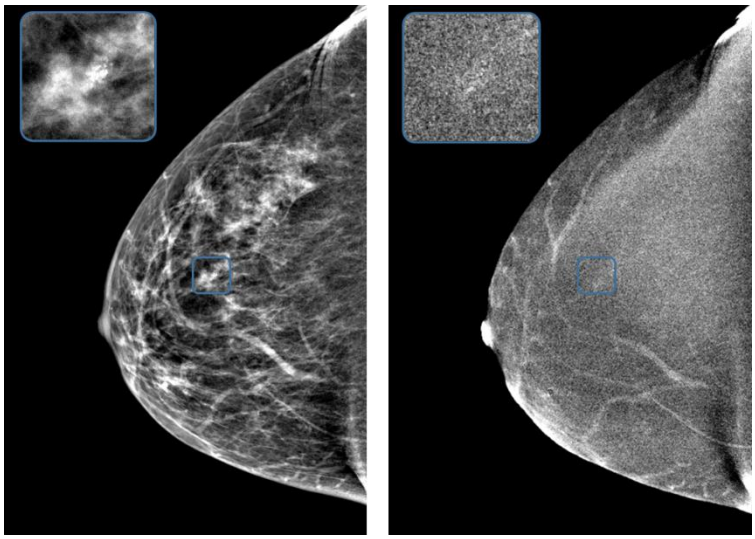

*Figure 1: Example of a microcalcification cluster from the MUMC+ dataset. The heterogeneous microcalcifications are apparent in the low-energy image on the left. The recombined image on the right shows no clear enhancement at the corresponding location.*

### 2.2. THE CREATION OF SYNTHETIC DATA

In an earlier study, we developed an automated workflow to simulate microcalcification clusters at a specific location and in relation to existing breast textures[10]. In short, this method obtains a plausible location for simulation by subdividing the breast region in a grid of cells. For each cell, a handcrafted radiomics analysis[11] evaluated texture and statistical features, which were used to rank the grid cells. The highest scoring cell was considered the most relevant

location for lesion insertion. Within this cell, the workflow detected locations of candidate calcifications based on local intensity, contrast and the output from a Frangi filter[12]. This approach presumes that the best integration of calcifications with the local environment occurs in regions of high contrast, high intensity or with increased structures. An iterative clustering algorithm then combined nearby candidate calcifications to create a two-dimensional (2D) binary lesion model of a new microcalcification cluster. For this work, parameters were tuned to either create a benign or malignant cluster, with the latter generally consisting of smaller microcalcifications of more heterogeneous shapes that are spread over a larger area[13].

The 2D binary cluster model could be inserted into the CEM images using an in-house developed simulation framework[14]. This framework was extended to work with 'For Presentation' images and to create an ideal template of the cluster model as if it was ray-traced based on image and system characteristics[10]. This insertion was only applied in the processed low-energy images, as it was found single calcifications do not take up contrast agent themselves and therefore do not show up in the recombined image[15].

In order to alleviate the need for manual annotations of all synthetic images, automated mask generation was included. This consisted of drawing a convex hull around all the inserted calcifications, which was then smoothed and enlarged with an algorithm similar to that proposed by Chaikin[16]. A final dilation step ensured the masks were more similar to delineations made by radiologists for the real cases, which also tended to include some pixels of the perilesional region.

Within this mask region, the recombined image could be altered as well to simulate a potential complementary mass-like enhancing component. To this end, we measured the relative contrast between the real lesion regions delineated by the radiologist, and the surrounding background area. This background was obtained by subtracting a once-dilated real lesion mask from a twice-dilated mask. The combination of relative contrast values from all real microcalcification clusters resulted in probability distributions for benign and malignant lesions presented in figure 2. For each synthetic case, the contrast value was sampled from the appropriate distribution, depending on the type of cluster simulated. After a multiplication of the original pixel values in the recombined image within the generated mask with this contrast value, Gaussian smoothing with a standard deviation of 5 was applied to give a more realistic blending in with the surrounding tissue.

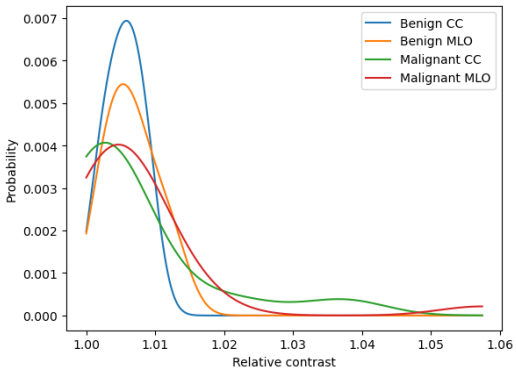


*Figure 2: Normalized probability distributions of the relative contrast between lesion and background measured for real cases.*

### 2.3. DETECTION AND CLASSIFICATION WITH A DEEP LEARNING MODEL

All data was first pre-processed[6] and subsequently compiled in a 3-channel PNG image, consisting of the filtered low-energy image, the original low-energy image and the filtered recombined image[17]. The mask R-CNN architecture utilised detects lesions via bounding boxes and then predicts a class for each box. We differentiated between masses and microcalcification clusters, and benign and malignant lesions to create four classes in total. In order to detect small structures such as calcifications, images were cropped to sub-images half the width and height of the original image. Predictions of all sub-images were combined to a single set for each full image, of which the five highest scoring regions were considered for further analysis.

In the real training dataset (total of 695 patients), 37 patients had a benign microcalcification cluster present and 94 patients had a malignant cluster. The deep learning model was first trained with this real set containing all lesion types. In the second and third training instances, the number of masses was preserved whereas the number of microcalcification clusters was doubled and tripled, respectively. Synthetic images with simulated clusters were added while maintaining the benign-malignant ratio of the real cluster cases.

The mask R-CNN model was initialised with weights pre-trained on the COCO dataset[18]. For all training instances, it was further trained for 30 epochs with a stochastic gradient descent optimizer, a learning rate of $5\times10^{-4}$ and a batch size of 2. For comparison, the optimal epoch was chosen at the lowest validation loss without any change in other hyperparameters. A lesion was considered detected (true positive) if the prediction score and intersection over union were both higher than 0.1, and if the lesion was correctly classified in terms of benign and malignant categories. For the classification stage, we focused on the area under the receiver operator characteristics curve (AUC) as the measure of performance.

## 3. RESULTS

### 3.1. THE CREATION OF SYNTHETIC DATA

The automated workflow allowed a large set of synthetic cases to be generated easily. For this work, simulated clusters were inserted in a number of lesion-free cases equal to twice the number of cluster cases in the real training set. In total, this resulted in 74 benign synthetic cases and 188 malignant synthetic cases. The simulated microcalcification clusters in the low-energy image had a realistic appearance[10]. Similarly, the enhancement levels in the recombined image adhered to the probability distribution observed in real cases. An example of a synthetic CEM with a simulated microcalcification cluster is shown in Figure 3.

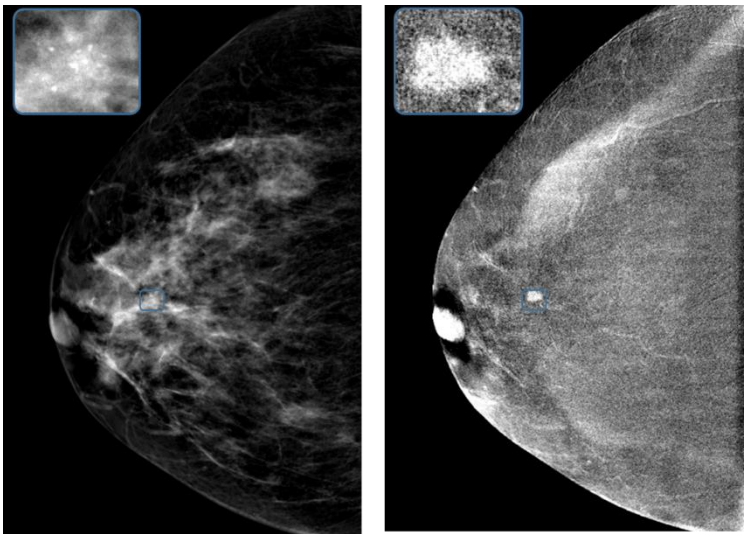

*Figure 3: Synthetic CEM case of the MUMC+ dataset with a simulated microcalcification cluster. The low-energy image on the left shows the microcalcifications of heterogeneous shape that blend in with the surrounding textures. On the right, one can see the associated mass-enhancement in the recombined image.*

### 3.2. DETECTION AND CLASSIFICATION WITH A DEEP LEARNING MODEL

When trained with real data only, the mask R-CNN model achieved a sensitivity of 82% for malignant masses on the internal test set. However, this sensitivity dropped to 51% for malignant microcalcification clusters, as depicted in Figure 4a. When the training set was augmented with synthetic data, sensitivity gradually increased not only for malignant clusters, but for malignant masses as well. Sensitivity was highest when the number of clusters was tripled, reaching 69% for malignant clusters and 92% for malignant masses. This increase in sensitivity was accompanied by a

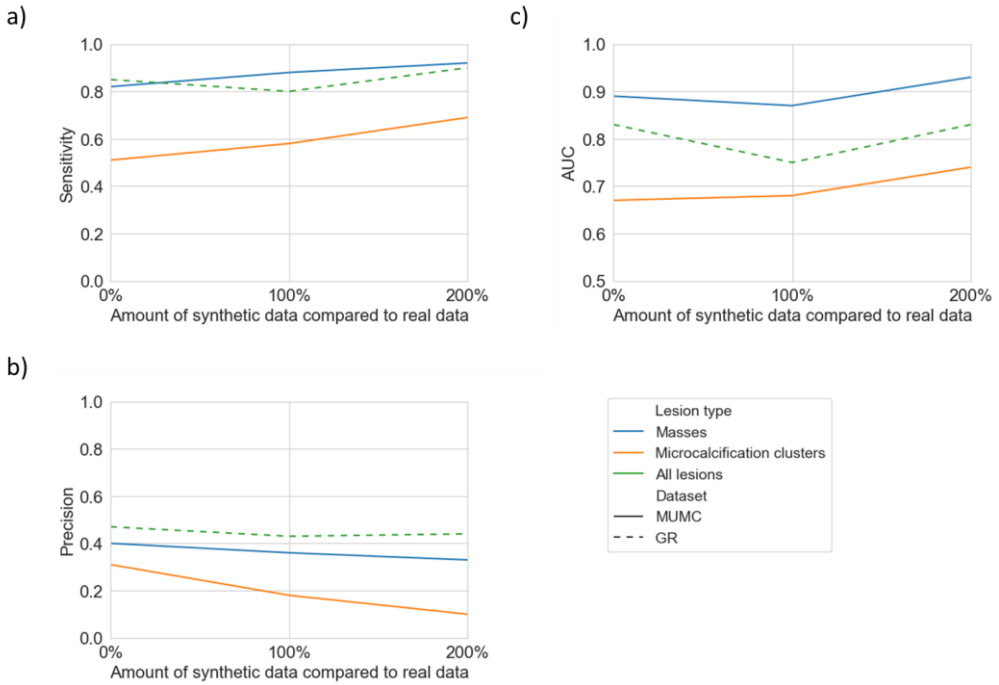


*Figure 4: Levels of a) sensitivity and b) precision for the detection of malignant lesions, and c) AUC for classification benign-malignant, for the trained mask R-CNN models. The full and dotted lines represent the internal (MUMC) and external (GR) test sets, respectively. All training instances included the complete real training set, and either no synthetic data was included (0%), or the number of clusters was doubled or tripled by adding a number of synthetic clusters equal to 100% or 200% of the real clusters, respectively.*

decrease in precision, illustrated in Figure 4b, which was most significant for the malignant microcalcification clusters. For the external test set the same trends were found. The impact of adding synthetic training data was smaller with an increase in sensitivity from 85% to 90% for all malignant lesions.

Figure 4c shows that the AUC remained relatively stable for most experiments. A discernible decrease was observed for the malignant lesions of the GR set, upon doubling the number of clusters. In contrast, when the number of clusters was tripled, the AUC increased for all categories of malignant lesions in the test sets.

## 4. DISCUSSION

We have introduced a workflow specifically designed to simulate microcalcification clusters in CEM cases. By leveraging a radiomics analysis of existing breast textures, individual calcifications and cluster architectures were constructed with a realistic appearance that integrated well with the surrounding tissue. This novel approach ensured a unique microcalcification cluster was modelled for each image, resulting in a versatile dataset suitable for data augmentation in deep learning applications. The additional method automatically generated a mask around the inserted cluster. We then modelled potential mass-like enhancement in the recombined image based on real, measured contrast levels. To our knowledge, this is the first attempt to model such lesions in complete 'For Presentation' CEM cases, utilizing both the processed low-energy and the recombined image.

Different setups were compared when training a deep learning model in CEM for breast cancer detection. The study focused on the detection of malignant lesions as these were of higher concern than benign lesions. The initial training, performed solely with real data, showed a substantial difference in the detection sensitivity of malignant microcalcification clusters as compared to malignant masses. For this reason, the impact of adding synthetically

generated images with simulated microcalcification clusters during training was investigated. As a result, more malignant clusters were detected but the method has not yet reached the sensitivity level attained for mass detection.

In future work, the remaining false negative malignant cases and the drop in precision when synthetic data is included during training, will be investigated. The deep learning model can be tuned to the different training instances and different amounts of real data can be included for training. Customizing the model optimization can be beneficial for different lesion types. In addition, the automated workflow allows for tuning the microcalcification cluster simulation based on user or model requirements. One option is to closely model the (malignant) lesions that are currently not detected by the deep learning model.

Apart from deep learning applications, the synthetic data with simulated models can be beneficial for a variety of applications such as medical training or in optimising image processing. The developed approach can improve diagnostic accuracy and ultimately patient outcomes in the context of breast cancer detection.

## 5. CONCLUSION

This work discusses a deep learning model to detect all lesion types in CEM images. A discrepancy between the performance for masses and microcalcification clusters was observed, with the latter posing a greater challenge for accurate characterization. To address this, synthetic cases were created using physics-based simulations of microcalcification clusters and mass-like enhancement. The addition of synthetic data for training contributed to an increase in the detection rates of malignant lesions, while maintaining similar levels of AUC in classification but causing a drop in precision. The impact increased as more synthetic data were included.

Our findings demonstrate that synthetic data can serve as a tool for enhancing the efficacy of deep learning models in medical imaging. This is especially relevant for areas where real data are scarce or imbalanced.

## REFERENCES


1. Jochelson, M. S. & Lobbes, M. B. I. Contrast-enhanced Mammography: State of the art. *Radiology* vol. 299 36–48 (2021).
2. Chen, Y. *et al.* Detection and classification of breast lesions using multiple information on contrast-enhanced mammography by a multiprocess deep-learning system: A multicenter study. *Chinese J. Cancer Res.* **35**, 408–423 (2023).
3. Mao, N. *et al.* Attention-based deep learning for breast lesions classification on contrast enhanced spectral mammography: a multicentre study. *Br. J. Cancer* 1–12 (2022) doi:10.1038/s41416-022-02092-y.
4. Zheng, T. *et al.* Deep learning-enabled fully automated pipeline system for segmentation and classification of single-mass breast lesions using contrast-enhanced mammography: a prospective, multicentre study. *eClinicalMedicine* **58**, 1–14 (2023).
5. Khaled, R. *et al.* Categorized contrast enhanced mammography dataset for diagnostic and artificial intelligence research. *Sci. Data* **9**, 1–10 (2022).
6. Beuque, M. P. L. *et al.* Combining deep learning and handcrafted radiomics for classification of suspicious lesions on contrast-enhanced mammograms. *Radiology* **307**, e221843 (2023).
7. Jailin, C. *et al.* AI-Based Cancer Detection Model for Contrast-Enhanced Mammography. *Bioengineering* **10**, 1–17 (2023).
8. Dominique, C. *et al.* Deep learning analysis of contrast-enhanced spectral mammography to determine histoprognostic factors of malignant breast tumours. *Eur. Radiol.* **32**, 4834–4844 (2022).
9. Helal, M. *et al.* Validation of artificial intelligence contrast mammography in diagnosis of breast cancer: Relationship to histopathological results. *Eur. J. Radiol.* **173**, 111392 (2024).
10. Van Camp, A. *et al. An Automated Toolbox for Microcalcification Cluster Modelling in Mammographic Imaging*. (2024).[Manuscript in preparation]
11. Lambin, P. *et al.* Radiomics: Extracting more information from medical images using advanced feature analysis. *Eur. J. Cancer* **48**, 441–446 (2012).

12. Frangi, A. F., Niessen, W. J., Vincken, K. L. & Viergever, M. A. Multiscale Vessel Enhancement Filtering. in *Medical Image Computing and Computer-Assisted Intervention — MICCAI'98* 130–137 (1998).
13. Burnside, E. S. *et al.* Use of microcalcification descriptors in BI-RADS 4th edition to stratify risk of malignancy. *Radiology* **242**, 388–395 (2007).
14. Vancoillie, L. *et al.* Verification of the accuracy of a hybrid breast imaging simulation framework for virtual clinical trial applications. *J. Med. Imaging* **7**, 1 (2020).
15. Van Camp, A. *et al.* The creation of a large set of realistic synthetic microcalcification clusters for simulation in (contrast-enhanced) mammography images. in *Proceedings SPIE Medical Imaging 2022: Physics of Medical Imaging* 120310V (2022). doi:10.1117/12.2611393.
16. Chaikin, G. M. An algorithm for high-speed curve generation. *Comput. Graph. Image Process.* **3**, 346–349 (1974).
17. Van Camp, A. *et al.* A deep learning model optimised to detect both masses and microcalcification clusters in contrast-enhanced mammography. in *ECR 2023* C–11661 (2023). doi:10.26044/ecr2023/C-11661.
18. Lin, T. Y. *et al.* Microsoft COCO: Common objects in context. in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* vol. 8693 LNCS 740–755 (2014).

## ACKNOWLEDGEMENTS

Authors acknowledge the global partnership agreement 2020 KU Leuven—Maastricht University. Authors acknowledge financial support from the European Union's Horizon research and innovation program under Grant Agreement: CHAIMELEON n°952172, EuCanImage n°952103, EUCAIM n°101100633, REALM n° 101095435, RADIOVAL n° 101057699, AIDAVA n°101057062 and IMI-OPTIMA n°101034347.