Maastricht University

Synthetic data of simulated microcalcification clusters to train and explain deep learning detection models in contrast-enhanced mammography

Citation for published version (APA):

Van Camp, A., Beuque, M., Cockmartin, L., Woodruff, H. C., Marshall, N. W., Lobbes, M., Lambin, P., & Bosmans, H. (2022). Synthetic data of simulated microcalcification clusters to train and explain deep learning detection models in contrast-enhanced mammography. In H. Bosmans, N. Marshall, & C. Van Ongeval (Eds.), *16th International Workshop on Breast Imaging: IWBI 2022* (Vol. 12286). Article 122860U SPIE. https://doi.org/10.1117/12.2621195

Document status and date:
Published: 01/01/2022

DOI:
10.1117/12.2621195

Document Version:
Publisher's PDF, also known as Version of record

Document license:
Taverne

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy
If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 13 Dec. 2023

PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Synthetic data of simulated microcalcification clusters to train and explain deep learning detection models in contrast-enhanced mammography

Astrid Van Camp, Manon Beuque, Lesley Cockmartin, Henry Woodruff, Nicholas Marshall, et al.

Astrid Van Camp, Manon Beuque, Lesley Cockmartin, Henry C. Woodruff, Nicholas W. Marshall, Marc Lobbes, Philippe Lambin, Hilde Bosmans, "Synthetic data of simulated microcalcification clusters to train and explain deep learning detection models in contrast-enhanced mammography," Proc. SPIE 12286, 16th International Workshop on Breast Imaging (IWBI2022), 122860U (13 July 2022); doi: 10.1117/12.2621195

SPIE.

Event: Sixteenth International Workshop on Breast Imaging, 2022, Leuven, Belgium

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 08 Sep 2022 Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Synthetic data of simulated microcalcification clusters to train and explain deep learning detection models in contrast-enhanced mammography

Astrid Van Camp[a], Manon Beuque[b], Lesley Cockmartin[c], Henry C. Woodruff[b,d], Nicholas W. Marshall[a,c], Marc Lobbes[b,d,e], Philippe Lambin[b.d], Hilde Bosmans[a,c]

[a]KU Leuven, Department of Imaging and Pathology, Division of Medical Physics and Quality Assessment, Leuven, Belgium
[b]The D-Lab, Department of Precision Medicine, GROW Research Institute, Maastricht University, Maastricht, the Netherlands
[c]UZ Leuven, Department of Radiology, Leuven, Belgium
[d]Department of Radiology and Nuclear Medicine, GROW Research Institute, Maastricht University Medical Centre+, Maastricht, the Netherlands
[e]Zuyderland Medical Center, Department of Medical Imaging, Zuyderland Medical Center, Sittard-Geleen, the Netherlands

ABSTRACT

Deep learning (DL) models can be trained on contrast-enhanced mammography (CEM) images to detect and classify lesions in the breast. As they often put more emphasis on the masses enhanced in the recombined image, they can fail in recognizing microcalcification clusters since these are hardly enhanced and are mainly visible in the (processed) low-energy image. Therefore, we developed a method to create synthetic data with simulated microcalcification clusters to be used for data augmentation and explainability studies when training DL models. At first 3-dimensional voxel models of simulated microcalcification clusters based on descriptors of the shape and structure were constructed. In a set of 500 simulated microcalcification clusters the range of the size and of the number of microcalcifications per cluster followed the distribution of real clusters. The insertion of these clusters in real images of non-delineated CEM cases was evaluated by radiologists. The realism score was acceptable for single view applications. Radiologists could more easily categorize synthetic clusters into benign versus malignant than real clusters. In a second phase of the work, the role of synthetic data for training and/or explaining DL models was explored. A Mask R-CNN model was trained with synthetic CEM images containing microcalcification clusters. After a training run of 100 epochs the model was found to overfit on a training set of 192 images. In an evaluation with multiple test sets, it was found that this high level of sensitivity was due to the model being capable of recognizing the image rather than the cluster. Synthetic data could be applied for more tests, such as the impact of particular features in both background and lesion models.

Keywords: contrast-enhanced mammography, microcalcification clusters, simulation framework, synthetic data, deep learning, detection, explainability

1. INTRODUCTION

New techniques using artificial intelligence for computer aided diagnosis (CAD) in healthcare have been widely developed and evaluated. Among other applications, deep learning (DL) models can aid radiologists in interpreting results and

16th International Workshop on Breast Imaging (IWBI2022), edited by Hilde Bosmans, Nicholas Marshall, Chantal Van Ongeval, Proc. of SPIE Vol. 12286, 122860U
© 2022 SPIE · 0277-786X · doi: 10.1117/12.2621195

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 08 Sep 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

predicting outcomes in early screening mammography for breast cancer detection [1], [2]. For less widespread imaging modalities, however, the accuracy of such DL models can suffer from a lack of data. Unlike for the conventionally used full-field digital mammography (FFDM), large datasets of the more novel contrast-enhanced mammography (CEM) technique are not yet available. This can lead to poorer levels of sensitivity in detection or accuracy in classification of benign and malignant lesions.

The aim of CEM is to better enhance the neovascular structures of cancerous regions by injecting an iodine-based contrast agent [3], [4]. The subtraction of a high-energy and a low-energy image results in a recombined image in which background structures are largely eliminated whereas cancerous structures are enhanced. CEM has proven to induce the same level of specificity and sensitivity as breast magnetic resonance imaging (MRI) while maintaining a lower cost and faster examination [5], [6]. Thus, when FFDM screening images do not indicate a conclusive outcome or for high risk patients, CEM is a good alternative and a DL model can provide the radiologist with more information.

Whereas CEM is more optimal to enhance mass lesions, calcifications do not show up clearly in the recombined image and are difficult to detect [7]. In a preliminary DL model made available for detection of lesions in CEM, more emphasis was put on the masses in the recombined image while it had low level of sensitivity in the detection of microcalcification clusters. It was hypothesized the detection of such clusters could be further improved by using synthetic data. Such data aids in analyzing the model and increases its explainability by providing it with specific data. In this work we present a method to generate synthetic cases of CEM containing simulated microcalcification clusters. When feeding this data in a controlled manner to the DL model it is possible to investigate how specific changes influence the detection accuracy. Our work focused on how a DL model, trained for detection of microcalcification clusters, learns more from the lesion's features or from the image's features.

2. METHODS

2.1 Database

The CEM database available comprised 1917 cases acquired on a GE Essential system (GE Precision Healthcare, Buc, France) in Maastricht UMC+ between 2013 and 2018. All patients were either high-risk or recalled from mammography screening. For each patient the set of images consisted of the processed low-energy and recombined image in craniocaudal (CC) and mediolateral oblique (MLO) view of both breasts.

In 1062 cases lesions were delineated by radiologists and classified based on biopsy. Since the remaining 859 cases did not contain suspicious lesions, they could be utilized as background images for the insertion of simulated microcalcification clusters without the risk of feeding the DL model images with both real and synthetic lesions. Out of all real, delineated cases, 222 breasts contained a cluster of microcalcifications, 63 of which were benign and 159 malignant. None of the breasts contained more than one cluster.

2.2 Generating synthetic data

Starting from a set of individual microcalcifications, new clusters were created based on mathematical properties and findings of the shape, size and distribution, as is described in [8]. Benign clusters had less microcalcifications, with their individual shape being more spherical whereas malignant clusters generally had more microcalcifications and the microcalcifications were more irregular in shape [9], [10]. Taking these considerations into account the individual microcalcifications were combined in a 3-dimensional voxel model to create a cluster. Applying random choices on the size of the cluster, the number of microcalcifications and the location and rotation of the individual microcalcifications ensured every new cluster was unique.

After the creation of the voxel models of clusters, the next step consisted of inserting them into background images. When studying the relative contrast between calcifications and their surrounding background pixels, it was found that this was close to zero for the recombined images. This was a consequence of the limited uptake of contrast in calcifications as opposed to masses which show up more clearly in recombined images. Therefore, the simulation of microcalcification clusters was only performed in the processed low-energy images.

A simulation tool earlier developed and validated at KU Leuven [11], [12] was extended for realistically inserting voxel models into images acquired on a GE Essential system. The cluster and the background image as well as some necessary

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 08 Sep 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

parameters were given as inputs. These parameters include the system configuration, the spectrum, the attenuation coefficient of the simulated material and the breast thickness. The tool then created a template of the voxel model as it would be produced by the actual system. Based on findings of the relationship between 'FOR PROCESSING' and 'FOR PRESENTATION' images the template was adapted and then multiplied with the background image to create synthetic data.

When inserting the clusters into the CEM background images, the choice of location is important. Even though microcalcifications can occur anywhere in the breast, indications by radiologists showed that regions with higher skewness were more typical. Therefore, the choice of insertion location was based on this parameter by computing the ten regions with the highest skewness. In order to eliminate problematic regions e.g. at the edge of the breast, the final location is hand-picked from the ten possibilities. To ensure the locations in both CC and MLO view corresponded, the distance to the nipple and angle with respect to the center of the pectoralis were taken into account as well [13]. Finally, a rotation adjusted the voxel model of the cluster to create a realistic projection in both views.

2.3 Validation of synthetic data

After inserting these simulated clusters into background processed low-energy CEM cases, a first set was used for evaluation by radiologists. In a validation study both real and synthetic cases were shown to radiologists experienced in mammography. Since the study did not include a query on the detectability, the cluster of importance was already denoted by an enlarged bounding box. As is explained in [8], radiologists would find these clusters to have a realistic shape, size and appearance. When showing one view only (CC or MLO) radiologists failed to distinguish real from synthetic cases. This denoted the realism of the generated synthetic cases with a faithful insertion in the real background image. At the same time, more realism would be needed if combined views are important.

To ensure a correct label (benign or malignant) was given to the synthetic case, radiologists were asked to rate this classification as well. The cluster was rated with a number in the range of 1 to 6 with 1 representing the most confidence in a benign cluster whereas 6 denoted that the radiologist was extremely confident the cluster was malignant. This was done first with showing both views at the same time. To investigate the importance of the correspondence of location, a second study was included as well. In this study half of the cases were shown to the radiologists with one view only (this was randomly either the CC or MLO view) whereas the other half still comprised both views. Apart from this, the setting was identical to the previous study with the cluster already indicated and a 6-level confidence scale given to rate the malignancy of the clusters shown.

2.4 Deep learning model

A deep learning model for the detection of microcalcification clusters in contrast-enhanced mammography was developed using PyTorch. The architecture was based on the Mask R-CNN model [14], which is capable of detecting regions by predicting bounding boxes. For each lesion it can perform segmentation as well and predict a class. This model had a MobileNet backbone and the weights were pre-trained on the COCO dataset [15].

Before providing the model with CEM data, it was first preprocessed to combine the low-energy and recombined images in the three channels of an RGB image. Binary masks were converted in such a way that, if multiple lesions were present, each lesion was assigned a different value ranging from 1 to the number of lesions to distinguish between unique occurrences of lesions. The same label was used for each occurrence of a specific class. Both the mask and the RGB image were cropped around the borders of the breast to reduce the size of the data to be processed. Images of the same breast in CC and MLO view were handled independently by the Mask R-CNN model.

As this work focused on studying the use of synthetic data for training and investigating a DL detection model, only images containing malignancies were used. This overcomes issues that could occur in the classification stage. The accuracy of the segmentation was not investigated at this point either, the focus was on evaluating the predicted bounding boxes. This model could be trained with real data only, synthetic data only and the combination of both. Specific test sets of synthetic data could be created to evaluate multiple criteria.

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 08 Sep 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

2.5 Explainability

Given the method to create synthetic data with the simulated microcalcification clusters and the available simulation tool, customized test sets were created. This allowed for specific tests on the DL model and report which changes in data would be detectable. Successive tests could be performed to validate or confirm the performance of a new model, and more importantly explain the AI. Opportunities included successively:

- Start from the same breast and same cluster, but change the location of insertion in the same view
- Start from the same breast and same cluster, but insert the same 2D template in the other view
- Start from the same patient and same cluster, but insert in the contralateral breast in the same view
- Start from the same patient and same cluster, but change the intensity of the inserted cluster
- Start from the same cluster, but insert it in the breast of a different patient at a similar location
- Start from the same breast and the same location, but insert a different cluster
- Start from an unseen cluster in the breast of an unseen patient

In this first work we investigated the sensitivity of a new detection model on two new sets. The first (test set A) contained 16 new CEM images without lesions in which simulated clusters were inserted that were already seen by the model in the training data. The new images were collected to have the same laterality, view and breast thickness as the original image. The location was chosen to be at the same coordinates as in the training image. When this choice of location was impossible due to different boundaries of the breast, a location as close as possible was chosen. In a second test set B 16 CEM images were used that were already passed to the model during training. Only in this case new simulated clusters were inserted at the same location as the previous. Finally, to compare the results on training data, the sensitivity on the partial training set (training set A) of the 16 images containing the original cluster in the original image was calculated.

3. RESULTS

3.1 Validation of synthetic data

The method described in the first section of 2.2 resulted in a dataset of 3-dimensional voxel models of 250 benign clusters and 250 malignant clusters. The size of all clusters varied from 2.25mm to 6.00mm with the number of individual microcalcifications per cluster within a range of 8 to 32. The malignant clusters contained on average more individual microcalcifications, 22.5 as compared to 11.8 for the benign cases. This is comparable to the differences discussed in [9] with a mean of 16.9 and 24.4 for the benign and malignant clusters respectively. The developed method can then create dataset of synthetic images with a unique cluster for each case. By randomly assigning a cluster to an image any cluster could be inserted in any image as in Figures 1-2.

For the first validation study with 40 real and 40 simulated clusters, Table 1 shows the statistical results. The high level of specificity for the synthetic data for both radiologists revealed a low number of false positives which denoted hardly any simulated benign clusters were considered to be malignant. For the simulated malignant clusters on the other hand, the sensitivity was 0.7 representing some misclassifications. Upon closer investigation, these appeared to be among the smaller clusters with fewer individual microcalcifications within the whole distribution of all simulated malignant clusters. Overall, the accuracy as well as the F1-score were higher for synthetic than for real data. This indicated that the simulated benign clusters as well as the larger simulated malignant clusters were created in such a way that they could be correctly classified.

Similar results were seen in the second validation study where half of the images were shown with one view only and the other half with both views. The values of specificity equal to 1 for the synthetic cases show all benign simulated clusters were correctly classified. The simulated malignant clusters were again harder for the radiologists to correctly categorize. The overall accuracy also increases for synthetic as compared to real cases. All F1-scores improved for synthetic cases as well except for the study with both views evaluated by radiologist B where a small drop can be noted.

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 08 Sep 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

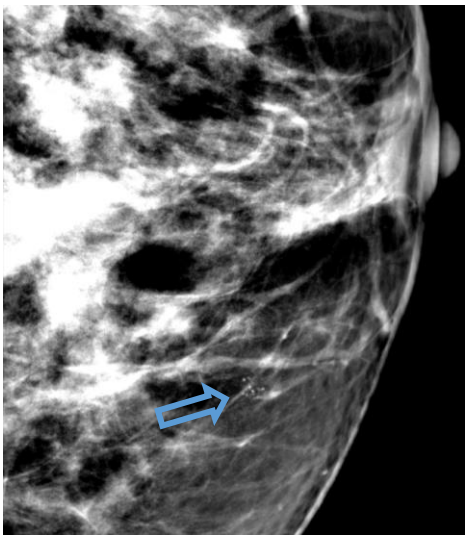

Figure 1: Simulated cluster in CC view

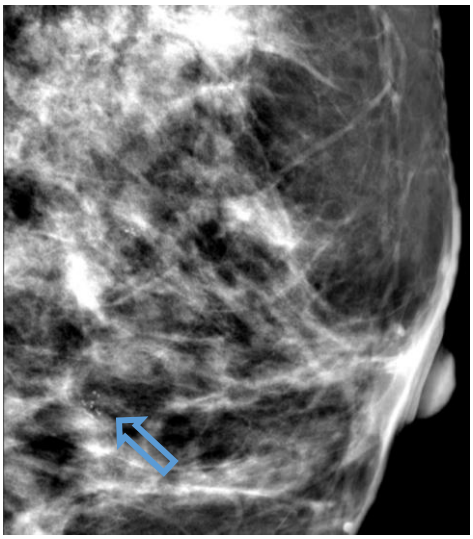

Figure 2: Simulated cluster in MLO view

Table 1: Results of classification by radiologists

		Sensitivity	Specificity	Accuracy	F1-score
Radiologist A	Real	0.90	0.55	0.73	0.77
	Synthetic	0.70	0.95	0.83	0.80
Radiologist B	Real	0.70	0.55	0.63	0.65
	Synthetic	0.70	1.00	0.85	0.82

Table 2: Results of classification by radiologists with one view only

		Sensitivity	Specificity	Accuracy	F1-score
Radiologist A	Real	1.00	0.40	0.70	0.77
	Synthetic	0.80	1.00	0.90	0.89
Radiologist B	Real	0.80	0.40	0.60	0.67
	Synthetic	0.60	1.00	0.80	0.75

Table 3: Results of classification by radiologists with two views

		Sensitivity	Specificity	Accuracy	F1-score
Radiologist A	Real	0.80	0.00	0.40	0.57
	Synthetic	0.80	1.00	0.80	0.80
Radiologist B	Real	0.90	0.40	0.65	0.72
	Synthetic	0.60	1.00	0.80	0.69

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 08 Sep 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

3.2 Evaluation of the deep learning detection model

In a first run the model was trained with 240 synthetic images available containing simulated malignant microcalcification clusters. This set was split into 80% (192 images) for training and 20% (48 images) for validation. A training run of 100 epochs was conducted. Intermediate results were saved as well at each tenth epoch. The first lines in Table 4 show the model started to overfit from epoch 50. In later epochs the difference in sensitivity between the training set and the validation set increases.

The sensitivity in the detection of this model was studied with multiple test sets. Sensitivity was calculated by computing the overlap between the predicted bounding box and the ground truth bounding box. If the overlapping area was larger than 10% of the ground truth area and the prediction score of the predicted bounding box was higher than 0.25, the cluster was considered to be detected.

As expected, the sensitivity for training set A increased (see Table 4) as the model was trained longer. For test set B the sensitivity also increased when the number of epochs was increased. This proves the model learned to recognize the images it had already seen. The dataset with previously seen clusters inserted into new images, on the other hand, only returns a sensitivity of 0.0 at each training point.

Table 4: Sensitivity of the trained DL detection model

	Epoch 30	Epoch 50	Epoch 80	Epoch 100
Validation set	0.00	0.06	0.17	0.17
Training set A	0.00	0.38	0.88	0.88
Test set A	0.00	0.00	0.00	0.0
Test set B	0.00	0.38	0.81	0.75

4. DISCUSSION

From the results in the previous section, one can note that the methods in section 2.2 can create a large set of realistic simulated microcalcification clusters. The shape of the microcalcifications, the distribution of the size of the microcalcifications and the number of microcalcifications per cluster follow properties found earlier for real clusters [10], [16]. With the extension of a previously developed simulation tool, new synthetic CEM cases including such clusters were created. In the evaluation studies performed by two radiologists experienced in mammography, it was found they could more confidently classify the simulated clusters in benign and malignant categories than real clusters. In all studies higher values of accuracy were found for synthetic cases. This may be due to the lower level of complexity as they follow more strictly set parameters and cannot have as wide a variety of appearance. The simulated benign clusters in particular were correctly classified every time, whereas malignant cases appeared to reduce sensitivity.

A Mask R-CCN model was trained on 192 images containing unique, simulated microcalcification clusters. When evaluating this model on test data it was found that the model could not recognize previously seen clusters in a new image. Newly simulated clusters inserted at the same location in already passed images, on the other hand, resulted in a higher level of detection sensitivity. The reason for this is that this particular model did not learn features of the clusters themselves, but recognized the background structure and the location. When training DL models for detection in medical images, it is clearly important to consider the number and variety of background images used.

These results have demonstrated the capability of synthetic data to explain the features learned by a DL model. When further evaluating DL models, a wider variety of synthetic test sets could be created to investigate the models behavior. This is the case for models trained on real data as well, especially in a medical imaging setting where less annotated data is available. For CEM where more cases contain a mass, microcalcification clusters were harder to detect. Our method to create synthetic cases containing such simulated clusters and to provide it as a test set could thus be used to improve a DL model for detection of all types of lesions in CEM.

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 08 Sep 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

5. CONCLUSION

The method described here can produce a new set of synthetic CEM image data using simulated microcalcification clusters. A dataset produced with this method was used to train a DL microcalcification detection model and to explain the model behavior. While new clusters in previously seen images could be detected, the model could not recognize known clusters in new images. This provides information on the features learnt by the DL detection model, where we found that not just the lesion itself is utilized by the model. In further work, the training could be extended to combine real and synthetic data. More test sets with specifics regarding the simulated lesion or the chosen image can be created in order to further understand the DL model and its training scheme.

REFERENCES

[1] S. M. McKinney *et al.*, "International evaluation of an AI system for breast cancer screening," *Nature*, vol. 577, no. 7788, pp. 89–94, 2020, doi: 10.1038/s41586-019-1799-6.

[2] A. Rodriguez-Ruiz *et al.*, "Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists," *J. Natl. Cancer Inst.*, vol. 111, no. 9, pp. 916–922, 2019, doi: 10.1093/jnci/djy222.

[3] M. B. I. Lobbes, M. L. Smidt, J. Houwers, V. C. Tjan-heijnen, and J. E. Wildberger, "Contrast enhanced mammography : Techniques , current results , and potential indications," *Clin. Radiol.*, vol. 68, no. 9, pp. 935–944, 2013, doi: 10.1016/j.crad.2013.04.009.

[4] J. J. James and S. L. Tennant, "Contrast-enhanced spectral mammography (CESM)," *Clin. Radiol.*, vol. 73, no. 8, pp. 715–723, 2018, doi: 10.1016/j.crad.2018.05.005.

[5] L. Li *et al.*, "Contrast-enhanced spectral mammography (CESM) versus breast magnetic resonance imaging (MRI): A retrospective comparison in 66 breast lesions," *Diagn. Interv. Imaging*, vol. 98, no. 2, pp. 113–123, 2017, doi: 10.1016/j.diii.2016.08.013.

[6] J. M. Lewin and M. J. Yaffe, "A history of Contrast-Enhanced Mammography," in *Contrast-Enhanced Mammography*, Springer International Publishing, 2019, pp. 1–21.

[7] U. C. Lalji *et al.*, "Evaluation of low-energy contrast-enhanced spectral mammography images by comparing them to full-field digital mammography using EUREF image quality criteria," *Eur. Radiol.*, vol. 25, no. 10, pp. 2813–2820, 2015, doi: 10.1007/s00330-015-3695-2.

[8] A. Van Camp *et al.*, "The creation of a large set of realistic synthetic microcalcification clusters for simulation in (contrast-enhanced) mammography images," 2022, no. April, p. 51, doi: 10.1117/12.2611393.

[9] E. S. Burnside *et al.*, "Use of microcalcification descriptors in BI-RADS 4th edition to stratify risk of malignancy," *Radiology*, vol. 242, no. 2, pp. 388–395, 2007, doi: 10.1148/radiol.2422052130.

[10] E. Fondrinier, G. Lorimier, V. Guerin-Boblet, A. F. Bertrand, C. Mayras, and N. Dauver, "Breast microcalcifications: Multivariate analysis of radiologic and clinical factors for carcinoma," *World J. Surg.*, vol. 26, no. 3, pp. 290–296, 2002, doi: 10.1007/s00268-001-0220-3.

[11] E. Shaheen, F. De Keyzer, H. Bosmans, D. R. Dance, K. C. Young, and C. Van Ongeval, "The simulation of 3D mass models in 2D digital mammography and breast tomosynthesis," *Med. Phys.*, vol. 41, no. 8, pp. 10–12, 2014, doi: 10.1118/1.4890590.

[12] L. Vancoillie, N. Marshall, L. Cockmartin, J. Vignero, G. Zhang, and H. Bosmans, "Verification of the accuracy of a hybrid breast imaging simulation framework for virtual clinical trial applications," *J. Med. Imaging*, vol. 7, no. 04, p. 1, 2020, doi: 10.1117/1.jmi.7.4.042804.

[13] S. Van Engeland, S. Timp, and N. Karssemeijer, "Finding corresponding regions of interest in mediolateral oblique and craniocaudal mammographic views," *Med. Phys.*, vol. 33, no. 9, pp. 3203–3212, 2006, doi: 10.1118/1.2230359.

[14] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2017-October, pp. 2980–2988, 2017, doi: 10.1109/ICCV.2017.322.

[15] T. Y. Lin *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)*, 2014, vol. 8693 LNCS, no.

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 08 Sep 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

PART 5, pp. 740–755, doi: 10.1007/978-3-319-10602-1_48.
[16] L. M. Warren, A. Mackenzie, D. R. Dance, and K. C. Young, “Comparison of the x-ray attenuation properties of breast calcifications, aluminium, hydroxyapatite and calcium oxalate,” *Phys. Med. Biol.*, vol. 58, no. 7, 2013, doi: 10.1088/0031-9155/58/7/N103.

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 08 Sep 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use