The creation of a large set of realistic synthetic microcalcification clusters for simulation in (contrast-enhanced) mammography images

The creation of a large set of realistic synthetic microcalcification clusters for simulation in (contrast-enhanced) mammography images

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

Astrid Van Camp, Lesley Cockmartin, Manon Beuque, Henry Woodruff, Nicholas Marshall, Philippe Lambin, Hilde Bosmans, "The creation of a large set of realistic synthetic microcalcification clusters for simulation in (contrast-enhanced) mammography images," Proc. SPIE 12031, Medical Imaging 2022: Physics of Medical Imaging, 120310V (4 April 2022); doi: 10.1117/12.2611393

SPIE.

Event: SPIE Medical Imaging, 2022, San Diego, California, United States

The creation of a large set of realistic synthetic microcalcification clusters for simulation in (contrast-enhanced) mammography images

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

[a]KU Leuven, Department of Imaging and Pathology, Division of Medical Physics and Quality Assessment, Leuven, Belgium
[b]UZ Leuven, Department of Radiology, Leuven, Belgium
[c]The D-Lab, Department of Precision Medicine, GROW – School for Oncology and Developmental Biology, Maastricht University, Maastricht, The Netherlands
[d]Department of Radiology and Nuclear Medicine, GROW – School for Oncology and Developmental Biology, Maastricht University Medical Centre+, Maastricht, The Netherlands

ABSTRACT

Characterization of microcalcification clusters in the breast and differentiation between benign and malignant structures on (contrast-enhanced) mammography (CEM) images is of great importance to determine cancerous lesions. Computer algorithms may help performing these tasks, but typically need large sets of data for model training. Therefore this paper develops a method to create synthetic microcalcification clusters that can later be used to overcome data sparsity problems. Starting from descriptors of the shape and size, both benign and malignant microcalcifications were created and then combined into 3-dimensional cluster models given realistic geometric properties. The distributions of the largest diameter and the number of microcalcifications per cluster in a set of 500 simulated clusters were set such that they agreed with those of real clusters. An existing simulation tool was then extended to insert the clusters into processed, low-energy CEM background images with appropriate contrast values. In a validation study comprised of 40 real and 40 synthetic cases, radiologists were asked to evaluate realism and malignancy. It was found that the shape and the structure of the individual microcalcifications as well as the complete clusters were realistic. Thus the descriptors were chosen correctly and enabled a good classification between benign and malignant cases. The realistic brightness and boundary smoothness proved the simulation tool can correctly insert the 3D clusters into real background images and is suitable of creating a large set of realistic microcalcification clusters simulated in existing (contrast-enhanced) mammography images. With improvements on the correspondence of insertion location in craniocaudal and mediolateral oblique view, which proved more challenging to simulate realistically, this promising method is expected to be applicable for modeling complete synthetic cases. Such a dataset can be used for data enrichment where data sources are limited and for development and training purposes.

Keywords: mammography, contrast-enhanced mammography, microcalcification clusters, simulation framework

1. INTRODUCTION

Contrast-enhanced mammography (CEM) is a promising technique that has been shown to be more sensitive than conventional full-field digital mammography (FFDM) in detecting and differentiating mass lesions.[1,2] The approach is to inject an iodine-based contrast agent that will enhance cancerous regions as it will be taken up in neovascular structures of the tumour. By producing low-energy and high-energy images which are then subtracted, a recombined image is obtained. In this image the background structures are largely eliminated and

Further author information: (Send correspondence to Astrid Van Camp)
Astrid Van Camp: E-mail: astrid.vancamp@uzleuven.be

Medical Imaging 2022: Physics of Medical Imaging, edited by Wei Zhao, Lifeng Yu, Proc. of SPIE
Vol. 12031, 120310V · © 2022 SPIE · 1605-7422 · doi: 10.1117/12.2611393

the cancerous structures are enhanced. This CEM technique has the same level of specificity and sensitivity as breast magnetic resonance imaging,[3,4] while it eliminates some of its disadvantages such as the time required or the higher cost.

To have an overall view, a radiologist will be given the combination of the recombined image as well as the processed low-energy image. Whereas CEM is mainly developed to enhance mass lesions, the low-energy image which is comparable to a FFDM image, is better suited to the visualization of calcifications.[5] This allows for a better detection of calcifications as well as differentiation between benign and malignant microcalcification clusters, which is crucial. The radiologists in our department prefer the combined reporting of masses and microcalcifications, a feature that is possible with CEM while it requires a multimodality evaluation in case of magnetic resonance imaging (MRI).

As CEM is a rather new technique, large studies and datasets are not yet widely available. To study the technique however, or to design computer-aided diagnosis tools, a large set of data for development and testing is needed. Therefore we present the tools to create synthetic CEM cases with microcalcification clusters. At first we discuss how to create individual microcalcifications and combine them into clusters. These can then be projected onto real background images of any (lesion free) mammogram creating a sufficiently large dataset of synthetic cases.

2. METHODS

2.1 Database

A database of CEM cases, acquired on a GE Essential system (GE Precision Healthcare, Buc, France) between 2013 and 2018, was available. It consisted of 1,917 cases of high-risk patients or patients recalled from mammography screening of which 1,058 included lesions that had been delineated by radiologists and classified based on biopsy. The remaining 859 non-delineated cases were normal and could be used as background images to simulate new structures without the need to take into account another potential lesion already in place.

All cases included both the craniocaudal (CC) and the mediolateral oblique (MLO) view of both breasts. For each case the processed low-energy and the recombined image were saved, resulting in a complete dataset comprising 15,336 images.

2.2 Generating Realistic Microcalcification Clusters

The creation of a set of clusters started from a set of individual microcalcifications. For the malignant clusters, these were obtained from segmenting real clusters. A dataset of real biopsied clusters which had been imaged by a micro-CT and then processed to obtain a 3-dimensional binary voxel model, was available.[6] The segmentation was done based on finding connected regions in the 3-dimensional model and then selecting each connected region as a single microcalcification. A lower limit on the size of 0.0001 times the volume of the cluster was included to avoid adding small noisy structures of only a few pixels to the dataset of microcalcifications.

For benign clusters on the other hand, microcalcification clusters were not biopsied and therefore another approach was needed. Given that benign microcalcifications were typically found to have a more regular shape such as annular or round,[7,8] these were created starting from the basic binary structure of a sphere. The diameter ranged from 0.09mm to 0.6mm with a higher probability for the smaller values as the probability weights scaled linearly from 1/0.09 to 1/0.6. This followed the distribution of diameters in the dataset of malignant microcalcifications and of those described by Warren et al.[9] Afterwards noise uniformly distributed between -1 and 1 was added and a threshold of 0.5 was applied on the resulting intensity values to give the microcalcification a more realistic, less regular shape. A final opening operation could then remove the background noise whereas a closing operation filled the voxels of the remaining microcalcification. By choosing the diameter of the sphere randomly and by adding random noise, each newly created microcalcification was unique.

The individual microcalcifications were then combined to create new clusters as in Figure 1. The initial cluster was an empty cubic grid with the dimension randomly chosen on a uniform scale between 2.25 mm and 6 mm. This corresponded to the sizes found for the real clusters that were available and had an average size of 4x4x4 mm. The first step consisted of randomly choosing the number of microcalcifications to be added to this cluster

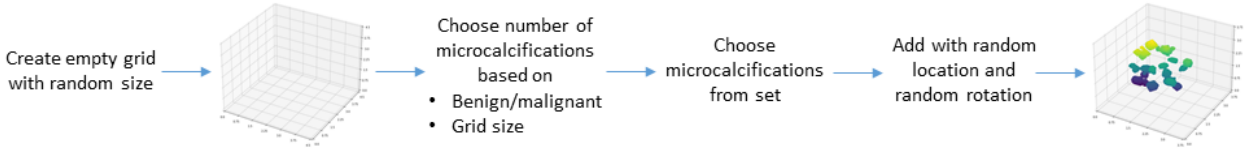


Figure 1: Pipeline of generating a realistic microcalcification cluster.

from a pre-determined range. Since real malignant clusters on average contained more microcalcifications,[10] the range in between which the number of microcalcifications could be chosen, was higher than for the benign case with lower limits of 18 and 8 respectively. Next to this, more microcalcifications were allowed when the cluster is larger. Therefore the upper limit linearly increased with the cluster size to a maximum of 25 for the benign and 35 for the malignant clusters. After choosing the microcalcifications randomly from the dataset, each of them was added to the grid with the only restriction that no overlap of multiple microcalcifications was allowed. Variation of appearance after 2-dimensional projection of microcalcifications was ensured by including a 3-dimensional rotation with angles randomly chosen between 0° and 360°. This method could be repeated for as many clusters as needed by generating new random choices thus enabling the creation of a large set of unique 3-dimensional voxel models of microcalcification clusters.

2.3 Simulation into (Contrast-Enhanced) Mammography Images

With 3-dimensional models of clusters created, these could be inserted in background images of the CEM dataset. This insertion was built upon a simulation framework that had previously been designed and validated[11,12] and was now extended to include the characteristics of the GE Essential system. This framework produced a template of the cluster to be inserted in the 2-dimensional projection of either the CC or the MLO view. As the CEM dataset consisted of the processed low-energy and recombined images of the breasts only and did not contain the 'for processing' images, the templates obtained by the simulation framework for simulation into raw images had to be investigated. Computations however, showed that in the recombined image the relative contrast between the microcalcifications and a nearby background region was close to zero. This indicated the microcalcifications did not show in the recombined image, due to the fact that they did not take up contrast, unlike masses. Therefore when creating new synthetic cases with inserted microcalcification clusters, these were only simulated into the low-energy image.

To do so, the relationship between intensities of raw and processed images was studied. Values of relative contrast for some microcalcifications in patient images as well as in test objects indicated an average relationship of $relative\_contrast_{raw} = -relative\_contrast_{processed}$ where the minus sign refers to the change of relative intensities, i.e. regions where x-ray intensity is high are dark in processed images. Because of this relationship, it was possible to adapt the templates such that they could be used for insertion in processed images. As the template would be multiplied with the original image, it was important that the default value of the template was kept at 1.0 such that image pixels within the template region that did not contain a microcalcification were unchanged when the template was applied. When mirroring the template over 1 with the relationship $processed\_value = 2 - raw\_value$, the relative contrast was correctly modified whereas the background values did not change. This modification of the template ensured a correct simulation of a lesion into processed images of the CEM dataset.

2.4 Choice of Location of Insertion in the Breast

Another important aspect was how to determine the location of insertion of the lesion into the breast image. Since microcalcifications could occur almost anywhere in the breast tissue, it could be argued a completely random choice of location would suffice. A preliminary radiomics analysis of typical locations indicated by radiologists showed relatively high skewness values. To obtain a more realistic result, this parameter was therefore used to guide the choice of insertion location. Next to this, regions with clips that had very high intensity and often created artefacts, were filtered out by eliminating regions with a very large range in pixel values.

To ensure both locations of the inserted cluster in the CC and MLO view of the breast corresponded, the position relative to the nipple was used as an extra parameter.[13] By first determining a location in CC view and then only considering locations in MLO view that had a similar distance and angle between the lines connecting the nipple with the lesion and the nipple with the pectoralis center, the correspondence was realized and subject to further investigation as this is just one choice out of many possibilities. Finally a rotation of the 3-dimensional cluster accounted for the correct projections in CC and MLO view.

2.5 Validation

To evaluate the simulation procedure, radiologists were asked to check the classification of the newly created clusters into benign/malignant categories next to their realism. In a validation study setup with the ViewDEX software package[14] 40 real as well as 40 synthetic cases were shown to radiologists experienced in mammography. For both sets half of the cases contained benign clusters whereas the others were malignant. Radiologist were not informed about this distribution to ensure avery case was evaluated without any prior knowledge. Only the breast in which the lesion was present was considered in both the CC and the MLO view. For each case the study examined three aspects. (1) 'How confident are you this cluster is realistic?' Answers could be given on a 6-level confidence scale with 1 denoting the respondent was extremely confident the cluster was synthetic whereas 6 represented extreme confidence the cluster at hand was real. (2) 'How confident are you this cluster is malignant?' The answers were again given on a 6-level confidence scale ranging from 1 for most probably benign to 6 for most probably malignant. (3) 'Do the locations of the cluster in CC and MLO view correspond?' For this question the possible answers consisted of 'yes', 'undecided' and 'no'. To check the method, all cases were shown to the readers in a random order ensuring no preliminary knowledge on the real or synthetic CEM cases could impact the results.

In order to investigate the actual impact of the correspondence of chosen locations, a second validation study was set up. The framework and distribution of images was identical to the first study, only now the dataset was split in half. For the first half only one of the CC and MLO views was chosen whereas the second half still included both views of all investigated breasts. By doing so the impact of the correspondence of location was eliminated from the first half allowing the comparison of results with and without this influence. As the study itself already emphasized the focus on the location, the third question 'Do the locations of the cluster in CC and MLO view correspond?' was dropped.

3. RESULTS

3.1 Synthetic Data of Simulated Microcalcification Clusters

The methods in Section 2.2 resulted in a dataset of 188 individual malignant microcalcifications. A dataset of 200 individual benign microcalcifications was created as well. Some differences in properties between both sets existed. Figure 2 shows the distribution of the largest diameter of each of the microcalcifications, which is similar to the distribution discussed by Warren et al.[9] As expected the benign microcalcifications were smaller (all within a range of 0.03 mm to 0.6 mm) than the malignant microcalcifications where the largest diameters ranged from 0.195 mm to 1.29 mm. Another difference was the shape of the microcalcifications. As expected the benign microcalcifications had a high level of sphericity, between 0.7 and 0.9, with some outliers up to 0.5 for the smaller microcalcifications. For the malignant microcalcifications on the other hand, the sphericity ranged from 0.2 to a maximum value of 1.0. So some structures appeared elongated or spiculated whereas others were close to perfect spheres.

From these sets of microcalcifications unique clusters were generated for each synthetic case. To apply the method for the creation of a large set of microcalcification clusters, 250 simulated benign and 250 simulated malignant clusters were constructed. The distribution of the number of microcalcifications per cluster in Figure 3 visualizes how the malignant clusters generally contained more microcalcifications. The smallest benign cluster had a size of 2.25 mm with 8 individual microcalcifications and the most microcalcifications (23) could be found in a cluster of 5.88 mm. For the malignant set the number of microcalcifications ranged from 18 to 32 with a mean value of 22.5 with the size of the smallest cluster being 2.25 mm and the largest being 5.99 mm. These values are close to the characteristics in[10] where the mean number of microcalcifications in a benign cluster is 16.9 and in a malignant cluster 24.4.

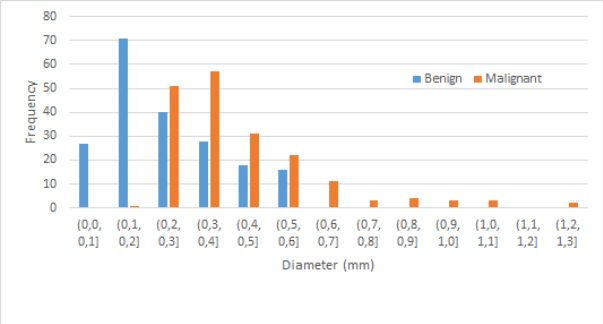


Figure 2: Frequency distribution of largest diameter of microcalcifications

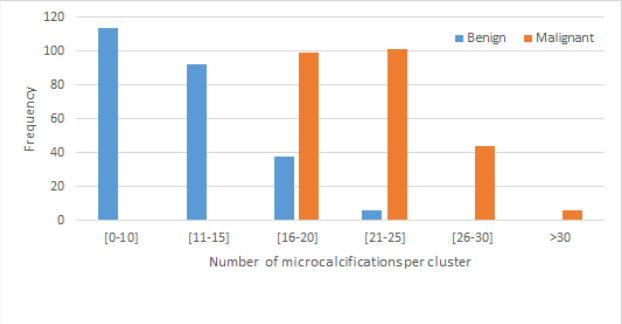
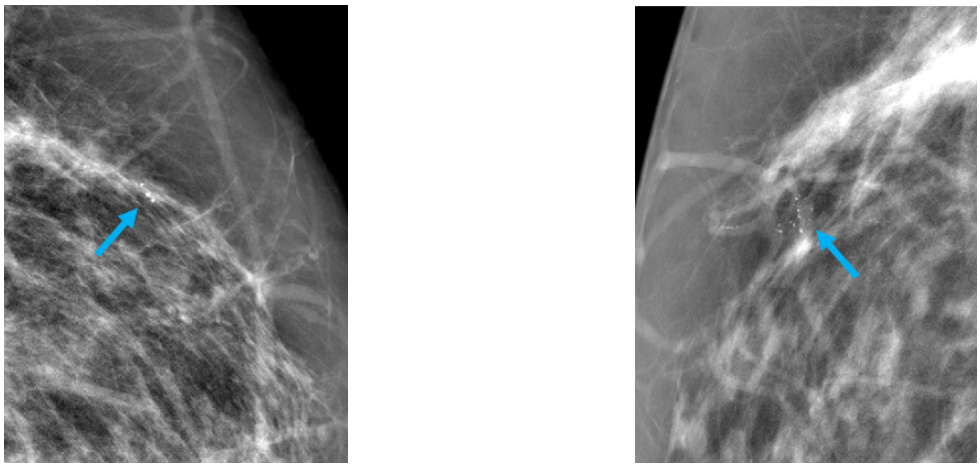


Figure 3: Frequency distribution of number of microcalcifications per cluster



(a) Benign cluster

(b) Malignant cluster

Figure 4: Examples of simulated clusters inserted into processed low-energy backgrounds of the GE CEM dataset

The images in Figure 4 illustrate a benign and a malignant simulated cluster inserted into a processed low-energy image of the CEM dataset. In each case the cluster was inserted in a location with high skewness. The malignant case contained more individual microcalcifications which were more spiculated. As a preliminary result these images confirmed the synthetic cases with inserted clusters have a realistic appearance.

3.2 Validation Results

In the validation study it was found that the receiver operating characteristic (ROC) curves (visualized in Figure 5) of the malignancy classification of the synthetic cases had a higher area under the curve (AUC)-score than the curves of the real cases. The latter represented AUC values of 0.75 (confidence interval $[0.62-0.88]$) and of 0.66 (confidence interval $[0.51-0.80]$) whereas the synthetic cases resulted in AUC values of 0.94 (confidence interval $[0.89-0.99]$) and 0.97 (confidence interval $[0.94-0.99]$) for radiologist A and radiologist B respectively. A higher AUC value signified a higher level of confidence in classifying the cases at hand into 'benign' and 'malignant' categories with 1.0 denoting a perfect score.

Next to the malignancy classification, radiologists could often distinguish the synthetic from the real cases. The ROC-curves in Figure 6 correspond to the confidence of the realism classification. Radiologist A had more experience in reading GE images and obtained high AUC values of 0.84 (confidence interval $[0.73-0.93]$) for the benign and of 0.93 (confidence interval $[0.85-0.99]$) for the malignant cases. Radiologist B obtained lower AUC values of 0.70 (confidence interval $[0.57-0.82]$) and of 0.71 (confidence interval $[0.57-0.83]$) for the benign and malignant cases respectively. As the curves show there is no significant difference in the classification of benign or malignant cases denoting both types had a similar level of realism.

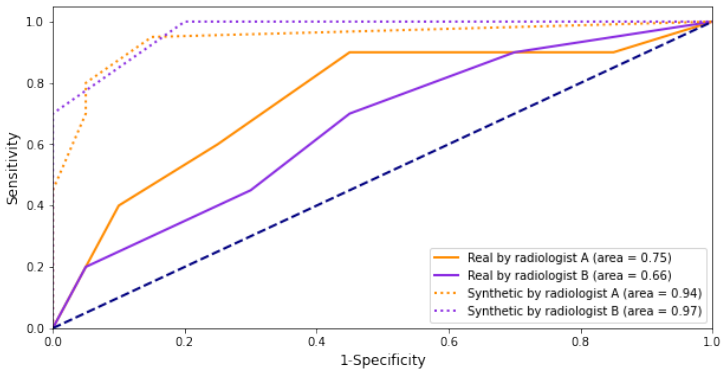


Figure 5: ROC-curves for the classification task of the cluster into 'benign' and 'malignant' categories for each radiologist. The full lines represent the classification task for the real cases whereas the dotted lines represent the synthetic cases.

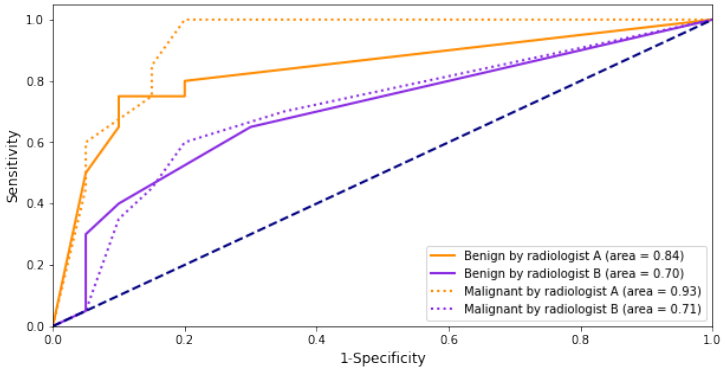


Figure 6: ROC-curves for the classification task of the cluster into 'real' and 'synthetic' categories for each radiologist. The full lines represent the classification task for the benign cases whereas the dotted lines represent the malignant cases.

When questioning the radiologists on the aspects they had found relevant to distinguish synthetic from real cases the choice of location was often stated. The figures in Table 1 denote that mainly the synthetic cases had locations that either did not or only marginally corresponded. While the chosen location in one view was mostly considered plausible, it was the combination of both projections of CC and MLO view that led the radiologists to believe a case at hand contained a simulated cluster. Apart from the location, the intensity of the simulated microcalcifications also impacted their realism. The radiologists considered the synthetic microcalcifications of higher contrast and with sharper boundary. Upon a more detailed investigation, a mistake was found in the

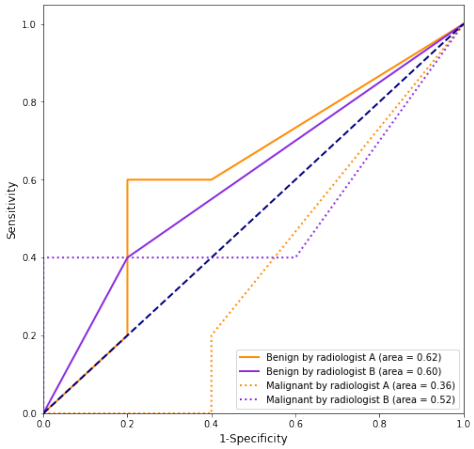
Table 1: Results for question (3) 'Do the locations of the cluster in CC and MLO view correspond?' in the ViewDEX validation study

		Yes	Maybe	No
Radiologist A	Real	36	2	2
	Synthetic	27	3	10
Radiologist B	Real	33	5	2
	Synthetic	29	7	4

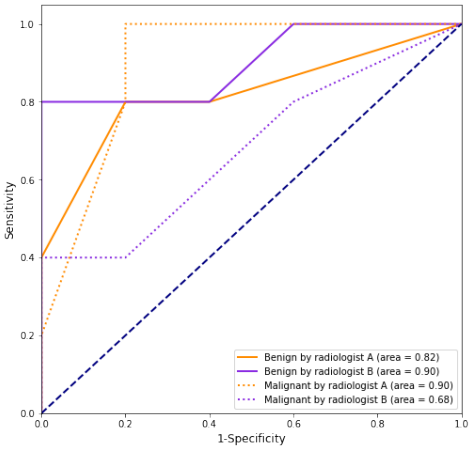
parameters that were input in the simulation tool. This was adjusted.

The second study was performed after the first investigation and used two datasets, one with both views and one with one view only. When only one view at the time is shown, the AUC values of the ROC-curves in Figure 7a range from 0.36 for the classification of malignant cases for radiologist B to 0.62 for the classification of benign cases for radiologist A. As these values were close to 0.5, it is denoted radiologist can no longer confidently distinguish between real and synthetic cases. Figure 7b on the other hand, visualizes ROC-curves with AUC values ranging from 0.68 for the classification of malignant cases by radiologist B to 0.90 for the the classification of benign cases by radiologist B. These high values indicate the (unfortunate) improvement in the performance to recognize synthetic cases when both views are shown.

In the second study, also the classification into benign versus malignant lesions was examined and these results remained similar to the earlier results shown in Figure 5. Both Figure 8a and Figure 8b confirm the high AUC values for synthetic cases and the lower AUC values for real cases.

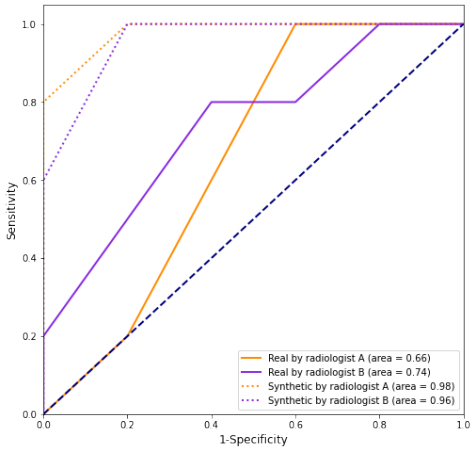


(a) ROC-curves for study with only one view shown

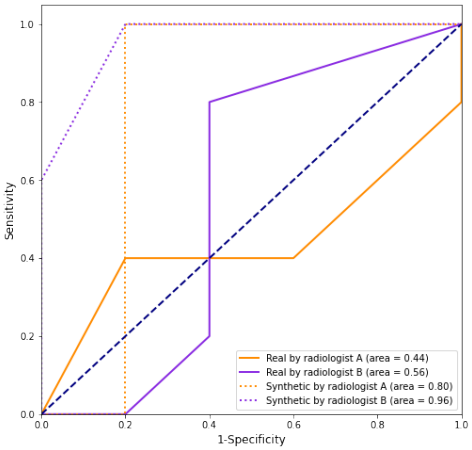


(b) ROC-curves for study with both views shown

Figure 7: ROC-curves for the classification task of the clusters into 'real' and 'synthetic' categories for each radiologist. The full lines represent the classification task for the benign cases whereas the dotted lines represent the malignant cases.



(a) ROC-curves for study with only one view shown



(b) ROC-curves for study with both views shown

Figure 8: ROC-curves for the classification task of the clusters into 'benign' and 'malignant' categories for each radiologist. The full lines represent the classification task for the real cases whereas the dotted lines represent the synthetic cases.

4. DISCUSSION

These results illustrated the capability of the method discussed in Section 2.2 to create a large set of realistic synthetic microcalcification clusters. The distribution of the size and shape could be related to figures discussed previously in literature.[8–10] Radiologists also confirmed the realism of the structure of the clusters and the shape of the microcalcifications. As they could classify the synthetic cases with a high AUC value into either the group of benign or malignant cases, the parameters used to construct the different types of clusters were also verified. This high AUC value as compared to the real cases could be explained by the clearly defined characteristics of the clusters and a lower level of complexity.

For the simulation into mammography images on the other hand, it was found the realism was affected by the difficulty in determining a corresponding location in CC and MLO view. In a study with only one view shown, radiologist could not confidently distinguish between real and synthetic cases, whereas studies with both views shown simultaneously resulted in higher AUC values. Improvement could potentially be made by considering a larger number of radiomics features describing the background region and relating the structures in the location of insertion in the two views.

5. CONCLUSION

The method described in this work is capable of creating a large set of benign as well as malignant microcalcification clusters with a realistic appearance as is validated by radiologists. When these are inserted into a low-energy (contrast-enhanced) mammography background image the appearance is confirmed by radiologist to be realistic and proves we can generate synthetic data. Even though the chosen characteristics of the simulated clusters allow for easier classification between benign and malignant clusters for synthetic data than for real clusters, improvements can still be made in terms of choice of location. Such a dataset is expected to be useful for tuning computer-aided diagnosis tools, for data augmentation or to investigate the emerging CEM technique when limited clinical data is available.

ACKNOWLEDGEMENTS

This work is part of the global partnership between KU Leuven and Maastricht University. We are grateful for the collaboration between the Department of imaging and pathology at KU Leuven and GROW - School for oncology and developmental biology at Maastricht University as well as for the assistance from the departments of radiology at UZ Leuven and Maastricht UMC+.

REFERENCES

[1] Lobbes, M., Smidt, M., Houwers, J., Tjan-Heijnen, V., and Wildberger, J., "Contrast enhanced mammography: Techniques, current results, and potential indications," *Clinical Radiology* **68**(9), 935 – 944 (2013).

[2] James, J. and Tennant, S., "Contrast-enhanced spectral mammography (CESM)," *Clinical Radiology* **73**(8), 715 – 723 (2018).

[3] Li, L., Roth, R., Germaine, P., Ren, S., Lee, M., Hunter, K., Tinney, E., and Liao, L., "Contrast-enhanced spectral mammography (CESM) versus breast magnetic resonance imaging (MRI): A retrospective comparison in 66 breast lesions," *Diagnostic and Interventional Imaging* **98**(2), 113–123 (2017).

[4] Lewin, J. M. and Yaffe, M. J., [*A history of contrast-enhanced mammography*], 1–21, Springer International Publishing, Cham (2019).

[5] Lalji, U. C., Jeukens, C. R., Houben, I., Nelemans, P. J., van Engen, R. E., van Wylick, E., Beets-Tan, R. G., Wildberger, J. E., Paulis, L. E., and Lobbes, M. B., "Evaluation of low-energy contrast-enhanced spectral mammography images by comparing them to full-field digital mammography using EUREF image quality criteria," *European Radiology* **25**(10), 2813–2820 (2015).

[6] Shaheen, E., Van Ongeval, C., Zanca, F., Cockmartin, L., Marshall, N., Jacobs, J., Young, K. C., R. Dance, D., and Bosmans, H., "The simulation of 3D microcalcification clusters in 2D digital mammography and breast tomosynthesis," *Medical Physics* **38**(12), 6659–6671 (2011).

[7] Le Gal, M., Chavanne, G., and Pellier, D., "Diagnostic value of clustered microcalcifications discovered by mammography (apropos of 227 cases with histological verification and without a palpable breast tumor)," *Bulletin du cancer* **71**(1), 57–64 (1984).

[8] Burnside, E. S., Ochsner, J. E., Fowler, K. J., Fine, J. P., Salkowski, L. R., Rubin, D. L., and Sisney, G. A., "Use of microcalcification descriptors in BI-RADS 4th edition to stratify risk of malignancy," *Radiology* **242**(2), 388–395 (2007).

[9] Warren, L. M., Mackenzie, A., Dance, D. R., and Young, K. C., "Comparison of the x-ray attenuation properties of breast calcifications, aluminium, hydroxyapatite and calcium oxalate," *Physics in Medicine and Biology* **58**(7) (2013).

[10] Fondrinier, E., Lorimier, G., Guerin-Boblet, V., Bertrand, A. F., Mayras, C., and Dauver, N., "Breast microcalcifications: Multivariate analysis of radiologic and clinical factors for carcinoma," *World Journal of Surgery* **26**(3), 290–296 (2002).

[11] Shaheen, E., De Keyzer, F., Bosmans, H., Dance, D. R., Young, K. C., and Van Ongeval, C., "The simulation of 3D mass models in 2D digital mammography and breast tomosynthesis," *Medical Physics* **41**(8), 10–12 (2014).

[12] Vancoillie, L., Marshall, N., Cockmartin, L., Vignero, J., Zhang, G., and Bosmans, H., "Verification of the accuracy of a hybrid breast imaging simulation framework for virtual clinical trial applications," *Journal of Medical Imaging* **7**(04), 1 (2020).

[13] Van Engeland, S., Timp, S., and Karssemeijer, N., "Finding corresponding regions of interest in mediolateral oblique and craniocaudal mammographic views," *Medical Physics* **33**(9), 3203–3212 (2006).

[14] Svalkvist, A., Svensson, S., Hagberg, T., and Båth, M., "Viewdex 3.0 - Recent development of a software application facilitating assessment of image quality and observer performance," *Radiation Protection Dosimetry* **195**(3-4), 372–377 (2021).