Data synthesis for a convolutional neural network to discriminate solid masses and cysts in dual-energy digital mammography: Simulation study

Marian Qian[1], Andrey Makeev[2], Bahaa Ghammraoui[2], Andreu Badal[2], and Stephen J. Glick[2]

[1]Thomas Jefferson High School for Science and Technology, Alexandria, VA

[2]DIDSR/OSEL/CDRH, U.S. Food and Drug Administration, Silver Spring, MD

Introduction

Suspicious lesions found in screening mammography are sometime biopsied for more in-depth analysis of breast tissue, but only about 10-30% of breast biopsies turn out to be cancer. This causes undue patient anxiety and increased healthcare costs resulting in many women putting off future screenings [4]. Training a CNN to eventually differentiate solid masses from benign cysts can prevent the need for further screenings.

The generation of data for training a covolutional neural network (CNN) to differentiate between solid masses and cysts in mammograms will be what is primarily described as CNNS require thousands of images to train from scratch and this classification is based off of very slight differences in mass and cyst linear attenuation coefficients (μ_{cyst} & μ_{mass}) at high energy levels [3]. The hypothesis of this study is that images acquired at these differing kVp settings can be combined into a single RGB image, resulting in data CNNs can use.

Methods

(Fig 1.2) Each breast phantom consisted of adipose and glandular tissue and contained 47 lesions.

Generation of 3-D breast masses [1]

(Fig 1.1) Diameters, incremented by 1 mm from 4 mm to 8 mm, represented masses and cysts.

Total 100 phantoms were used, with 20 phantoms for each diameter and for masses and cysts.

Placed into anthropomorphic digital breast phantoms [2] developed by Christian Graff

Placed small spheres in same regions as lesions

MC-GPU simulation on phantom

(Eqn 1.3) μ_{cyst} & μ_{mass} from basis functions of aluminum (μ_{Al}) and poly(methyl methacrylate) (μ_{PMMA}) and measured values of $\mu_{cyst\ \&\ mass}$ from 10 mm thick lesion samples to find coefficients α and β in order to generate a complete table of μ constants [3].

MC-GPU (made at FDA) simulations with μ_{cyst} & μ_{mass} at 26 and 50 kVp

OpenCV Hough Circles to find centers of masses

Use determined centers to extract ROIs

(Fig 1.4) Each lesion was extracted using IDL and Python programs.

Fill RGB red and green channel with 26 and 50 kVp images

(Fig 1.5) Fill last blue channel was filled with zeroes.

Average glandular dose received by the breast was 1 mGy and 2 mGy with 26 kVp and 50 kVp spectra respectively.

Ending with 9,400 images (masses and cysts total)

Results and Discussion

- Erhard, et al. conducted a clinical study for this type of classification using spectral mammography [4] and showed good performance.
- Our study used dual-energy mammograms instead, with images simulated from Monte Carlo-GPU (https://github.com/DIDSR/VICTRE_MCGPU/) because of the the limited data available from Erhard, et al. and the difficulty to acquire such images.
- Our finished dataset of masses and cysts contains a total of 9,400 images (200 phantoms x 47 lesions), with each image having a region of interest of 150 by 150 voxels.
- (Fig 2.1) Because μ_{cyst} decreases slightly at higher energies than μ_{mass} (solid masses and cysts have differing μ as a function of energy) and the pixel values for the images taken at 26 kVp and 50 kVp are still maintained in an RGB image, the difference in attenuation at these energies might provide a way for the CNN to classify the lesions.

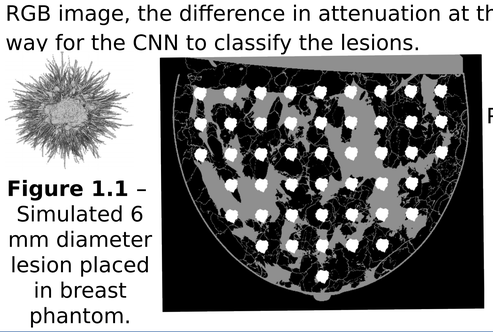


Figure 1.1 – Simulated 6 mm diameter lesion placed in breast phantom.

Figure 1.2 – Placement of 47 masses with diameter of 6 mm in 3-D digital breast.

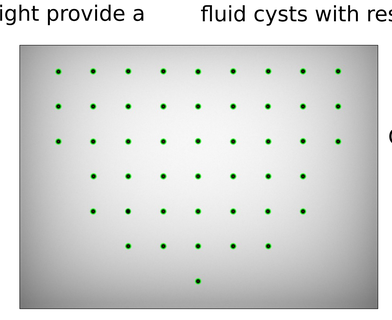


Figure 1.4 – After applying OpenCV Hough Circles to identify centers of masses.

Equation 1.3

$$\mu_{mass}(E) = \alpha_1 \mu_{PMMA}(E) + \beta_1 \mu_{Al}(E)$$
$$\mu_{cyst}(E) = \alpha_2 \mu_{PMMA}(E) + \beta_2 \mu_{Al}(E)$$

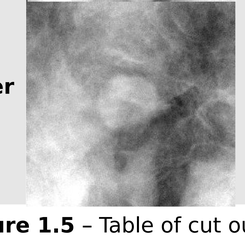
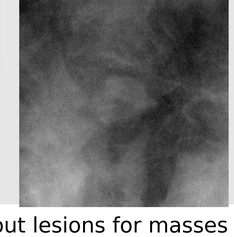
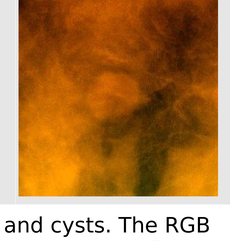
Energy Spectrum	26 kVp	50 kVp	26 kVp (red) & 50 kVp (green)
Solid Mass 6 mm diameter			
Fluid Cysts 6 mm diameter			

Figure 1.5 – Table of cut out lesions for masses and cysts. The RGB image consists of the red channel for 26 kVp, green channel for 50 kVp, and blue channel filled with zeroes.

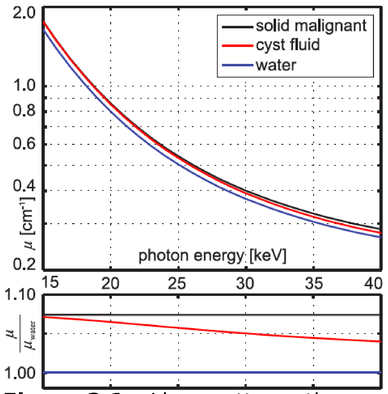


Figure 2.1 – Linear attenuation coefficients of solid masses and fluid cysts with respect to water [3].

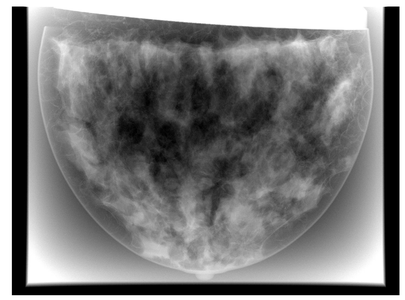


Figure 1.6 – Resulting mammogram for 6 mm masses at 50 kVp after simulation time 39 minutes with Nvidia GeForce GTX 1080 Ti.

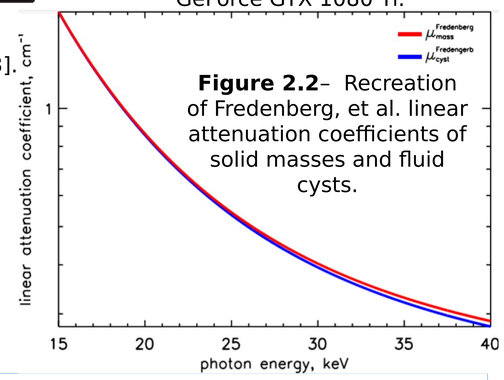


Figure 2.2– Recreation of Fredenberg, et al. linear attenuation coefficients of solid masses and fluid cysts.

Conclusion and Future Work

The next step for this project is to train the CNN on these 94,000 images using a ResNet 34 or 50 architecture [5]. Although the images look similar to the human eye, we hypothesize that the CNN will be able to differentiate the two to a certain extent because of the differing relationship between high- and low-energy linear attenuation coefficients for cysts and masses. We plan on using the Python Keras library to implement the model and will continue to fine tune the network's hyper-parameters as well. Because we currently do not have any way to validate how realistic our generated images are to sample lesions, we could also train a GANs model with the discriminator using images from Erhand, et al. who used real tissue samples from patients.

Acknowledgements

Thanks to ORISE for providing the funding for this research project, as well as Christian Graff for supplying the 3-D breast phantoms.

References

[1] de Sisternes L, Brankov GJ, Zysk MA. A computational model to generate simulated three-dimensional breast masses. *Med Phys*. 2015;42:94-1117.
[2] Graff CG. A new open-source multi-modality digital breast phantom. *Proc. SPIE.* 2016:978309-978310.
[3]Fredenberg E, Kilburn-Toppin F, Willsher P, et al. Measurement of breast-tissue x-ray attenuation by spectral mammography: solid lesions. *Phys Med Biol.* 2016;51:2595-2612.
[4] Erhard K, Killburn-Toppin F, Wills P, et al. Characterization of cystic lesions by spectral mammography: Results of a Clinical Pilot Study. *Invest Radiol.* 2016;51:34-347.
[5] Kaiming He, et al. Identity Mappings in Deep Residual Networks. *Microsoft Research.* In: CoRR abs/1603.05027. 2016. URL : http://arxiv.org/abs/1603.05027