1. Introduction

The American Cancer Society estimates that there will be over 270,000 new cases of breast cancer in 2019 [1]. If these diagnoses are made soon after the cancer first develops through mammography screening, the chance of survival is over 90% [2]. Radiologist have had assistance from automatic computer interpretation or Computer Aided Detection (CAD) systems that support radiologists in making decisions [3-6]. These systems have been useful by the side of radiologists, but work has yet to be done on utilizing radiologist input within the CAD system itself.

Our approach utilizes radiologist gist readings alongside a CAD system that uses a deep CNN and transfer learning to detect subtle abnormalities in full-field mammography. Radiologist gist data was taken by exposing radiologists to a unilateral mammogram of a breast with no abnormalities, a breast with an abnormality, or a breast contralateral to a breast with an abnormality for 250 milliseconds [7]. They were asked to rate the mammogram from 100, definitely normal, to 0, definitely call back. We see improvements to the CAD system when implementing a voting system that takes radiologist input into account within the model.

1. Methods and Materials

2.1 The dataset

The mammography dataset contained 220 unilateral full-field digital mammograms obtained from 110 unique patients at Brigham and Women’s Hospital. Three classes were in these images – no malignancy (110 images), malignancy (66 images), or contralateral to the breast with a malignancy (44 images).

2.2 Radiologist gist data

For the radiologist gist data, 10 radiologists were shown 120 images of the 220 where 40 were completely normal, 40 had a confirmed malignancy, and 40 were normal but contralateral to a breast with a malignancy. They were shown the image for 250 milliseconds and asked to report on a scale from 0, recommending the patient return for further examination, to 100, the scan is normal. The readings from the 10 radiologists were averaged for each image, giving the final radiologist response.

The final classification for the radiologist is set to 0 (normal) if the average response is greater than 50 and set to 1 (malignant) otherwise. A “confidence score” for each image is calculated based on the average response’s difference from 50.

2.3 Transfer learning

For the CAD system, a transfer learning model based on VGG19 was used. This is a deep CNN trained to classify over 1000 classes from the ImageNet corpus. The last hidden layer logits that come from the pre-trained network are saved for each of the 220 images. Those logits, referred to as the “codes”, are used as inputs into a linear support vector machine for binary classification as normal or abnormal.

From the dataset, we use only the normal and malignant images for training. Images were downscaled to 244x244 pixels to fit the pre-trained model properly. A 70/30 split is used for training and testing. Confidence values are obtained for each image, defined as the output’s distance from the hyperplane separating the classes.

2.4 Voting system

A simple voting system is used which will take the response that has a higher confidence between the radiologist and transfer learning model.

1. Results

3.1 Radiologist results

3.2 Transfer learning results

3.3 Voting system results

1. Discussion
2. Conclusion

[1] <https://cancerstatisticscenter.cancer.org/#!/cancer-site/Breast>

[2] M. Kalager, M. Zelen, F. Langmark, and H. O. Adami. Effect of screening mammography on breast cancer mortality in norway. New England Journal of Medicine, 363(13):1203–1210, 2010.

[3] B. Q. Huynh, H. Li, and M. L. Giger. Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. Journal of Medical Imaging, 3(3):034501–034501, 2016.

[4] Giger M. L., Boone J., Chan H., “History and status of CAD and quantitative image analysis,” Med. Phys. 35(12), 5799–5820 (2008).

[5] Hadjiiski L., Sahiner B., Chan H.-P., “Advances in CAD for diagnosis of breast cancer,” Curr. Opin. Obstet. Gynecol. 18(1), 64–70 (2006).10.1097/01.gco.0000192965.29449.da

[6] Giger M. L., Karssemeijer N., Schnabel J. A., “Breast image analysis for risk assessment, detection, diagnosis, and treatment of cancer,” Annu. Rev. Biomed. Eng. 15, 327–357 (2013).10.1146/annurev-bioeng-071812-152416

[7] K. K. Evans, T. M. Haygood, J. Cooper, A.-M. Culpan, and J. M. Wolfe. A half second glimpse often lets radiologists identify breast cancer cases even when viewing the mammogram of the opposite breast. Proceedings of the National Academy of Sciences, 113(37):10292–10297, 2016.