Improving Tumor Detection Accuracy in Mammograms of Dense Tissue Breast Cancer Patients with Machine Learning Methods and Image Processing Techniques

Paimagam, Akshra (School: North Carolina School of Science and Mathematics)

Breast cancer is the second leading cause of cancer mortality in women, contributing to roughly 40,000 deaths annually. While breast cancer can pose a threat to any woman, those with dense breast tissue have an increased risk of developing aggressive breast cancer as opposed to women with fatty breast tissue. This can be attributed to surplus glandular tissue, which obscures visualization in mammograms and maximizes difficulties in correctly spotting tumors. Due to potential confusion between white glandular tissue and tumors, misdiagnosis is common with dense breasts. For this reason, breast cancer diagnosis is 80-98% correct for women with fatty breasts but only 50% accurate for those with dense breasts. For radiologists, this disparity pairs with the onset of visual fatigue from examining overwhelming numbers of mammograms, which results in diminished diagnostic accuracy. Given the psychological ramifications that false diagnoses have on patients, the issue prompted this study on developing and optimizing the accuracy of computational algorithms used for the diagnosis of breast cancer. A graphical user interface (GUI) was devised to analyze pixel values, detect tumor edges through contour detection, and classify tumors as malignant or benign through a machine learning model. The highest-performing algorithm for both types of breast tissue was random forest classifier, producing a diagnostic accuracy of 87.65% for fatty and 84.75% for dense, and a sensitivity of 82% and specificity of 96.77% for fatty tissue and a sensitivity of 82.58% and specificity of 87.33% for dense tissue. These results demonstrate an improvement in diagnostic accuracy with the GUI from traditional mammography.

Awards Won: