OPTIMIZING RADIOMICS TECHNIQUE AS POTENTIAL RADIOGRAPH BIOMARKER VIA REPRODUCIBILITY STUDY

Report by: Dr. Muhammad Khalis Abdul Karim,
Department of Physics, Faculty of Science, Universiti Putra Malaysia.
e-mail: mkhalis@upm.edu.my

Since its introduction, radiomics features has been acknowledge for its fundamental methods for machine learning development in the medical imaging field. Basically, radiomics technique is a high throughput analysis that applies advanced computational approaches to convert image data from the selected region into high dimensional feature data, assuming the data provide information that could be used as a potential predictive biomarker. The quantitative features from the technique have been utilized extensively in cancer research and incorporate with machine learning to improve breast cancer prognosis. These features which originate from various sources of diagnostic information also overcomes the limitation of observer. However, the most critical parts in radiomics technique are reproducibility. Reproducibility describes the performances of radiomics measurements using different techniques or observers, or even from different diagnostic centers. To obtain accurate and precise results, the extracted features should be optimized to estimate patient survival analysis and boost treatment selection and monitoring for each patient.

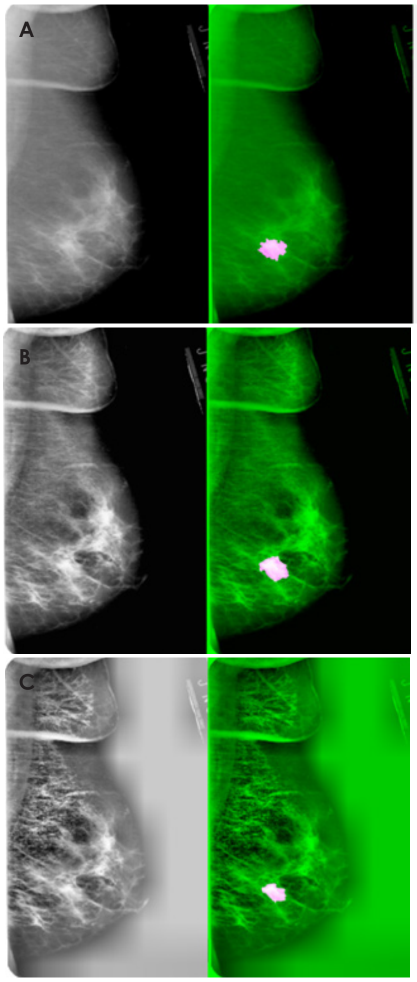

Figure 1: Mammogram radiograph with A) no image enhancement B) AHE enhancement C) CLAHE enhancement and segmented tumor for each dataset.

Hence, this work focuses the reproducibility of tumor segmentation during quantitative image extraction. Although physicians commonly use manual segmentation, this method is a time-consuming process that has more substantial inter-observer variability. To ensure higher accuracy in semiautomatic segmentation, pre-processing image enhancement is vital (Fig.1). Therefore, we assessed the robustness reproducibility of radiomics features of breast cancer through 2D mammograms. As illustrate in Fig 2, 37 radiomics imaging features were included and classified into three main features (6 tumor intensity histogram-based features, 22 textural features and 9 shape-based features). We found three quantitative imaging features were robust and had higher reproducibility when semiautomatic tumor segmentation with enhancer, Contrast Limited Adaptive Histogram Equalization (CLAHE) and Adaptive Histogram Equalization (AHE) techniques ($p<0.05$) were applied. The CLAHE set had higher reproducibility for most GLCM-based features and shape-based features compared to AHE techniques. The results showed that most of these features achieved high reproducibility scores when contrast enhancement and semi-automatic segmentation were applied to the image dataset. This work shows the existence in variation for the radiomics features extracted from tumor region and differed significantly with the image enhancement techniques. Semiautomatic segmentation with image enhancement using CLAHE algorithm gave the best result and was a better alternative than manual delineation as the first two techniques yielded reproducible descriptors.

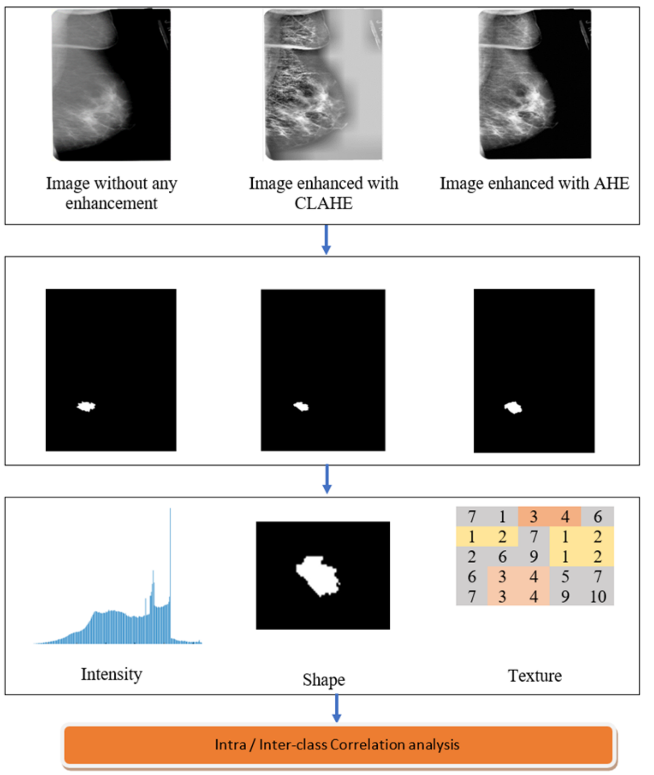

Figure 2: Reproducibility analysis of radiomics features in the 2D mammograms. (a) Three datasets were selected in our study with different image enhancement. (b) Three datasets were segmented using semi-automatic segmentation and manual delineation. Intensity, shape and textural transformed features were extracted from every dataset. The reproducibility of the radiomics features was measured by different indicators.