

PRESENTATION: E-POSTER

<https://doi.org/10.29289/25945394V36S11071>

1071 – Featuring extraction for predicting malignancy in images from breast ultrasound nodules

Daniel Guimarães Tiezzi, Isabela da Mata Tiezzi, Sarah Abud Recife, Ruth Morais Bonini, Licerio Miguel, Nilton Onari, Liliane Silvestre, Isabela Panzeri Carlotti Buzatto

¹Centro de Referência da Saúde da Mulher – Ribeirão Preto (SP), Brazil.

²Universidade de São Paulo, Faculdade de Medicina de Ribeirão Preto – Ribeirão Preto (SP), Brazil.

³Hospital das Clínicas de Ribeirão Preto – Ribeirão Preto (SP), Brazil.

⁴Hospital de Amor – Barretos (SP), Brazil.

Introduction: Ultrasonography (US) is the standard technology to analyse palpable or mammographically identified breast nodules. BI-RADS classification has been routinely used to select patients to core needle biopsy based on the risk of malignancy. However, due to the low predictive positive value (PPV) of this method, most patients undergo unnecessary procedures. Developing models for Computer-Aided Diagnosis (CAD) may optimize breast cancer diagnosis.

Method: An exploratory analysis evaluating the utility of featuring extraction as a tool to predict malignancy from breast images acquired by US examination. The BUSBRA dataset (1,064 unique images) was used in training / validation tasks and the final adjusted model was tested on a local dataset (229 images: 135 benign and 44 malignant). Feature extraction focused on the mask area and a total of 87 features was inferred using the radiomics library for Python. Logistic regression was used for feature selection (n=16). The train/validation dataset was split into a 7:3 ratio and cross validation (k=5) was used to verify the performance of the support vector machine (SVM), random forest (RF), logistic regression (LR), and gradient boosting decision tree (XGB) classifiers. The PPV and negative predictive values (PPV and NPV) were estimated and the final model was tuned to optimize the NPV, thereby minimizing missing cancer cases.

Results: The model mean accuracy was 0.87 ± 0.05 during the CV processes and the models performances were similar. After tuning the decision threshold to optimize the NPV, the adjusted model achieved 95% NPV with 67% PPV (accuracy=82%). The final prediction for the test set achieves 96% NPV and 51% PPV and 98% sensitivity with the RF classifier.

Conclusion: Featuring extraction from US images is an efficient method to automate tools that discriminate the etiology of breast nodules.

Ethics: Research Ethics Committee Approval (CAAE 41696820.4.1001.5440).

Keywords: breast neoplasm; diagnostic ultrasound; artificial intelligence.