Performance Analysis of Low and High-Grade Breast Tumors Using DCE MR Images and LASSO Feature Selection

B. Priyadharshini[1], A. Mythili[1] and K. R. Anandh[2]

1. School of Electronics Engineering,Vellore Institute of Technology, Vellore-632014, Tamil Nadu, India
2. Department of Radiology, Cincinnati Children's Hospital Medical Centre, Cincinnati -45229, USA
mythili.asaithambi@vit.ac.in, priyadharshini.b2020a@vitstudent.ac.in, anandhmurali@gmail.com

***Abstract*— Breast cancer is a complex genetic disease with diverse morphological and biological characteristics. Generally, the grade of a breast tumor is a prognostic factor and representation of its potential aggressiveness. Presently, Dynamic Contrast-Enhanced MRI (DCE-MRI) has gained a predominant role in assessing tumor grades and vascular physiology. However, due to tumor heterogeneity, tumor-grade classification is still a daunting challenge for radiologists. Therefore, to unburden the tumor grading process, a study has been carried out with 638 patients taken from the Duke-Breast-Cancer-MRI database (431- low-grade & 207- high-grade). The clinicopathological characteristics such as ER receptors, PR receptors, HER2, Pathological Complete Response (PCR or non-PCR), Menopausal status, and Bi-lateral status have shown a high significance of p = <0.00001, <0.00001, 0.0023, <0.00001, 0.0262, and 0.0045 respectively.**

The LASSO (Least Absolute Shrinkage and Selection Operator) feature selection model has selected 8 optimal features out of 529 feature sets (from Duke-Breast-Cancer-MRI). The extracted features are involved in the classification of high-grade and low-grade tumors by using a collection of classifiers such as Linear Support Vector Machines (LSVM), Logistic regression (LR), Linear discriminant analysis (LDA), Gaussian Naïve Bayes (GNB), k-Nearest Neighbors (KNN), and Random Forest (RF). The outcome of the L-SVM and LR showed better performance metrics values among all classifiers. Hence, the acquired classification results disclose that histological grade prediction using radiomics would aid clinical management and prognosis.

Keywords*— *classifiers, tumor grade, classification, feature selection.

I. Introduction

Breast cancer is a highly prevalent global malignancy and the primary cause of cancer fatalities worldwide [1]. The American Cancer Society reports that there are sustained differences in breast cancer occurrences and mortalities [2]. Generally, the histologic tumor grade represents a morphological assessment of the tumor malignancy and aggressiveness, pivotal in clinical treatment planning and estimating patient prognosis [3][4]. Also, tumor-grade analysis is based on genomic alterations and the molecular profile of the tumor cells. Tumors are mostly classified either as Low grade (Grade I or II) or High grade (Grade III& Grade IV). Clinically, DCE MRI captures hyper-vascular lesions with high temporal resolution aiding accurate tumor grade differentiation by keeping reasonable spatial resolution [5]. Radiomics, an upcoming technology has proved its efficacy in quantitative image analysis benefits analyzing different types of breast tumors, classifying grades, predicting treatment outcomes, and determining the risk of recurrence [6]. Hence, DCE-MRI-based radiomics is a promising prognostic biomarker for clinical decision-making and preoperative survival prediction.

Recent studies have reported the significance of non-invasive diagnostic tools and identifying imaging alternatives for tumor characterization [7][8]. Rather than evaluating the size and stage of the tumors alone, it is imperative to analyze histological grades and subtypes for precise prediction and diagnosis [9]. Studies have also revealed that prompt neoadjuvant or adjuvant chemotherapy is essential for high-grade tumor patients, whereas low-grade patients may benefit from long-term follow-up [10]. However, 30–60% of low or intermediate-grade breast tumors are extremely variable in biology and morphology making cancer treatment challenging [11]. Therefore, there are significant difficulties in grouping tumors as low or intermediate-grade. Clinically, needle biopsy aids in revealing histological tumor grades preoperatively. However, morphological changes in the tumor may lead to erroneous tumor prediction, eventually resulting in an increased false positive rate and delayed diagnosis [12]. Hence, it is crucial to ascertain appropriate radiomics-based Machine Learning techniques to differentiate low- and high-grade breast tumors.

In this work, an attempt has been made to differentiate tumor grades non-invasively using supervised learning algorithms along with dimensionality reduction methods for better performance outcomes. To accomplish the aforementioned process, a total of 638 clinically proven breast cancer patients and 529 radiomics features (from Duke-Breast-Cancer-MRI) were involved in this study. In addition, statistical analysis of clinicopathologic characteristics was tabulated to reveal a high association of these features in precise tumor classification. Subsequently, the LASSO (Least Absolute Shrinkage and Selection Operator) feature selection model has selected 8 optimal radiomics features from whole feature sets. The resulting features have been examined for the

multicollinearity test using the Pearson Correlation coefficient heat map. Lastly, a collection of classifiers is involved in classifying low- and high-grade breast cancer patients.

II. Materials and Method

A. MRI protocol

The DCE MR image data sets used in this analysis are taken from The Cancer Imaging Archive (TCIA) [13]. A 1.5T or 3T scanner capturing axial breast MRIs involves a fat-saturated gradient echo T1-weighted pre-contrast sequence, non-fat saturated T1-weighted sequence, and four post-contrast T1-weighted sequences taken after the IV administration of contrast agent (weight-based protocol of 0.2mL/kg). The median time of 131s was used between two post-contrast sequences. One of eight fellowship-trained breast imagers performed annotation. A graphical User Interface was generated for each patient that displayed the following MRI sequences: pre-contrast, first post-contrast, and pre-contrast subtracted from first post-contrast. Further, the tumors are delineated by 3D boxes supplied by the reader.

B. Feature Selection

The segmentation mask was extracted with the aid of the reader's annotation (3D box) followed by fuzzy C-means automatic segmentation. Computerized in-house software was used in extracting 529 features named tumor enhancement, shape, enhancement of tissues surrounding, texture, and shape [14]. Also, LASSO regression analysis techniques are frequently employed in feature selection and binary classification purposes [15][16]. As the Lasso regression produces sparse models, it is implemented in the feature selection process and expressed mathematically as

$$L_{lasso}(\hat{\beta}) = \sum_{i=1}^{n} \left(y_i - x_i'\hat{\beta}\right)^2 + \alpha \sum_{j=1}^{m} \left|\hat{\beta}_j\right| \quad (1)$$

Equation (1) represents the cost function with the L1 penalty term and α as a hyperparameter that tunes the penalty term. The whole analysis was performed using Python (V3.6) and the model has selected 8 optimal features out of 529 with high significant difference (p-value 0.05 or AUC > 0.5) for breast tumor classification. Further, a pairwise Pearson Correlation Coefficient Matrix (PCCM) identified high-correlated feature pairs for better performance outcomes.

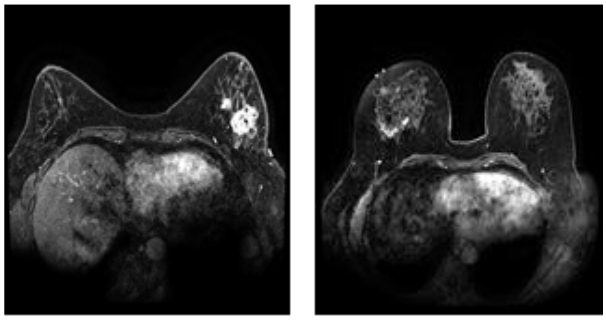
C. Classifiers

A collection of classifiers such as Logistic regression (LR), k-nearest Neighbors (KNN), Linear discriminant analysis (LDA), Gaussian Naïve Bayes (GNB), Linear Support Vector Machines (LSVM), and Random Forest (RF) was implemented for the classification. The patient sample was trained and tested by dividing the datasets into training samples (80%) and testing samples (20%). Furthermore, 5-fold cross-validation has been conducted on the training and testing datasets. The performance of different classification models was analyzed by using evaluation matrices such as Accuracy, Sensitivity, Area Under the receiver operating characteristic Curve (AUC), specificity, F1-score, Precision, Positive Predictive Value (PPV), and Negative Predictive Value (NPV).

III. Results and Discussion

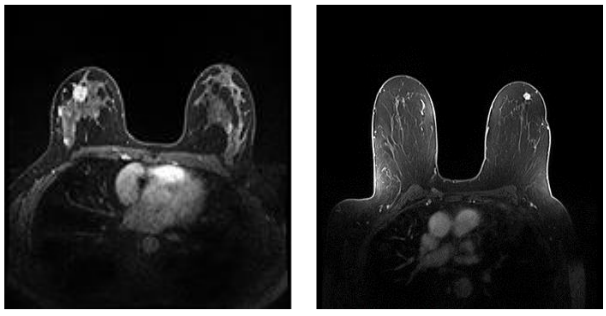
A. Representative Set of Breast DCE MR Images

A total of 638 patients were segregated and included in our study where 431(67.55%) are low-grade and 207 (32.44%) are high-grade with an average age of 52.29±11.25. Recent studies reveal that cancer is an age-related disease as its incidence and fatalities increase with age [17][18]. Additionally, the overall occurrence is higher in aged persons whereas younger cancer-positive patients show more aggressiveness resulting in poorer outcomes. Consequently, Figure 1(a & b) is the



(a) (b)

Figure 1. Representative set of breast DCE MR Images of two different High-grade patients acquired in the axial plane (a) one can appreciate high-intensity tumor and (b) one cannot appreciate high-intensity tumor.



(a) (b)

Figure 2. Representative set of breast DCE MR Images of two different Low-grade patients acquired in the axial plane (a) one can appreciate moderate-intensity tumor and (b) one cannot appreciate moderate-intensity tumor.

representative breast DCE MR images of two different high-grade 56- and 46-year-old patients, respectively. Figure 1 (a) shows a high-intensity, vascularized heterogeneous tumor mass with necrosis. Similarly, a low-intensity, low-vascularized small tumor mass is observed in Figure 1 (b). Figure 2 (a & b) is the representative breast DCE MR images of two different low-grade 36- and 42-year-old patients, respectively. Figure 2 (a) shows a small circular-shaped mass with an un-notable necrotic component. In Figure 2 (b) an undetectable tumor shape and size where the tumor presence is invisible from the original image and eventually poses challenges for the radiologist to locate the tumor precisely. Nevertheless, Figure 1 (a & b) and Figure 2 (a & b) demonstrate different sets of DCE MR breast images which are of high and low grade respectively, due to variations in the appearance of tumor volume it is still challenging for radiologists to identify tumors in these image sets resulting in preoperative erroneous predictions. To overcome the aforementioned challenges in distinguishing tumor grades, Radiomics-based machine-learning techniques have been employed in this study.

B. Clinicopathologic Characteristics

Clinicopathological features such as ER receptors, PR receptors, HER2, Pathological Complete Response (PCR or non-PCR), Menopausal status, and Bi-lateral status showed a high significance of p = <0.00001, <0.00001,0.0023, <0.00001, 0.0262, and 0.0045 respectively depicted in Table 1. Since the statistical analysis showed highly significant differences between the characteristics and tumor grades, it aided in accurate tumor classification with better classifier performance.

C. LASSO Analysis

Eight optimal radiomics features were selected by the LASSO model for classifying low and high-grade breast cancer patients. Figure 3 shows (a) LASSO coefficients as a function of alpha and (b) Mean square error on each fold as a function of alpha. In Figure 3 (a) moving from left to right, the model contains many features with a high magnitude coefficient, and as alpha increases the coefficient estimates become approximately zero by neglecting redundant features. Figure 3 (b) shows the Mean square error value for different cross-validation. A 5-fold cross-validation was involved in evaluating the optimal value of alpha as the alpha value differs for each fold. The necessity of optimizing the hyperparameter reveals the model's computational performance. Consequently, the calculation of mean squared- error across each model demonstrated the model's performance. Moreover, the cross-validation approach is employed as the selected features are large, and adequate samples are at each fold. The black thick line in Figure 3 (b) shows the average value of the error across folds. Furthermore, the Pearson Correlation Coefficient between each pair of these selected features was then computed by generating heatmaps for multicollinearity analysis rendering the interpretability of the model.

Table 1. Statistical Analysis of the Clinical and Histopathological Features

	Low grade	**High grade**	**p-value**
No of subjects	431(67.5%)	207(32.4 %)	
Age(Mean±SD)	54.69±10.86	49.9±11.6	0.6921
Estrogen receptors status			<.00001
Positive	376(87.2%)	105(50.7%)	
Negative	55(12.8%)	102(49.2%)	
Progesteron receptor status			<.00001
Positive	338(78%)	75(36.23%)	
Negative	93(21.57%)	132(63.8%)	
HER2 status			0.00239
Positive	62(13.3%)	50(24.1%)	
Negative	369(85.6%)	157 (75.8%)	
Response status			<.00001
PCR	9(2.08%)	37 (17.87%)	
Non-PCR	83(19.25%)	59 (28.50%)	
Not Available	332(77.0%)	103 (49.7%)	
Others	7(1.62%)	9 (4.34%)	
Menopausal Status			0.02629
Premenopausal	179(41.5%)	109(52.6%)	
Postmenopausal	241(55.9%)	95(45.8%)	
Not Available	11(2.55%)	3(1.44%)	
Bilateral status			0.00451
Bilateral	25(5.80%)	2(0.96%)	
Non-Bilateral	406(94.6%)	205(99%)	

Generally, a strong correlation between the features defines multicollinearity, and addressing this statistical fact depicts the validity and robustness of the model [19]. As a result of multicollinearity analysis, there were no feature sets with a correlation coefficient >0.8. The 8 final features selected are tabulated in Table 2. Figure 4 shows a Heatmap of the Pearson correlation coefficient of selected features and the obtained maximum correlation coefficient was only 0.56. Therefore, all eight optimal features were retained for further analysis.

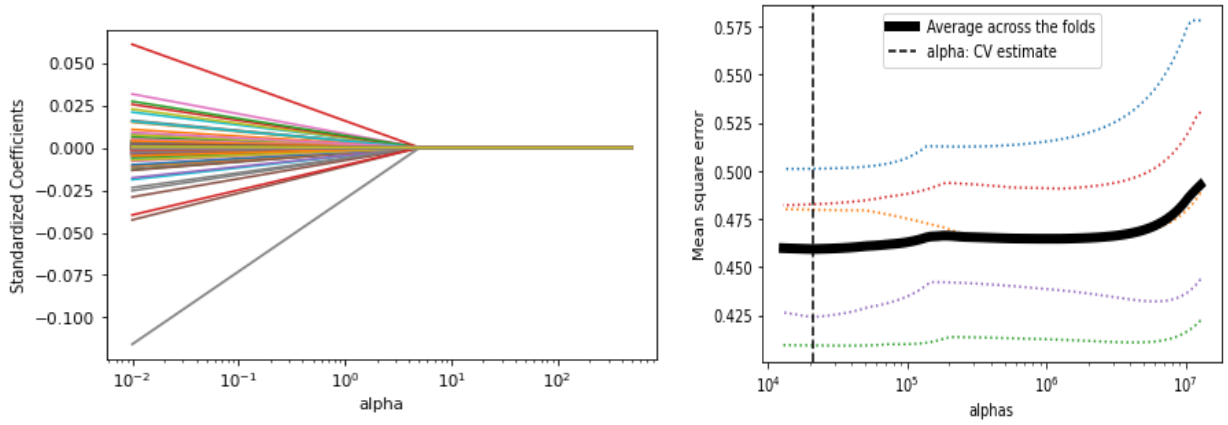


Figure 3. LASSO (a) LASSO coefficients as a function of alpha and (b) Mean Square Error as a function of alpha.

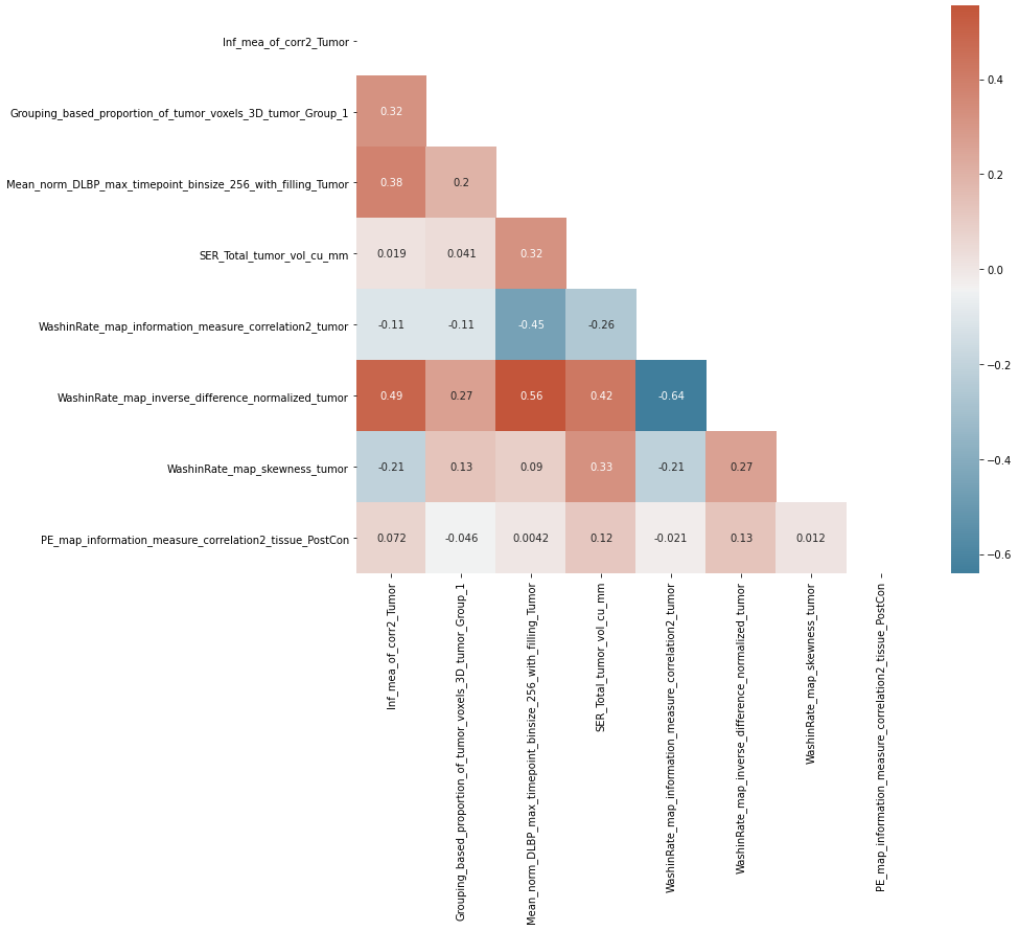


Figure 4. Pearson Correlation Coefficient Heatmap of the features selected for predicting tumor grades.

D. Performance Analysis of Classifiers

The performance of six classifier models was analyzed by testing datasets involving LASSO-selected features for the classification of tumors. Table 3 shows the performance analysis of different classifiers using the evaluation matrices with 67.55%, and 32.44% low and high-grade patients respectively. Generally, the model performance was evaluated by its F1 score, Accuracy, and AUC as the F1 score gives the combined effect of recall and precision and accuracy is model's integrity.

Table 2. LASSO Selected Features

Selected Features
Inf_mea_of_corr2_Tumor'
Grouping_based_proportion_of_tumor_voxels_3D_tumor_Group_1
Mean_norm_DLBP_max_timepoint_binsize_256_with_filling_Tumor
SER_Total_tumor_vol_cu_mm
WashinRate_map_information_measure_correlation2_tumor
WashinRate_map_inverse_difference_normalized_tumor
WashinRate_map_skewness_tumor
PE_map_information_measure_correlation2_tissue_PostCon

Table 3. Performance Analysis of Different Classifiers for Categorizing Low and High-Grade.

Classifiers	Accuracy (%)	AUC	Sensitivity (%)	F1-score	Specificity (%)	Precision	NPV
LD	74.6	0.78	91.53	0.82	39.68	0.75	0.69
LR	75.6	0.76	92.30	0.83	41.26	0.77	0.72
GNB	73.6	0.74	90.76	0.82	38.09	0.75	0.67
L-SVM	77.9	0.79	96.15	0.86	39.52	0.82	0.58
C-KNN	73.6	0.70	91.53	0.82	36.50	0.74	0.67
RF	74.4	0.71	91.36	0.84	30.18	0.76	0.57

Linear Discriminant- (LD), Logistic Regression- (LR), Gaussian Naïve Bayes –(GNB), Linear Support Vector Machine- (L-SVM), Cosine k- Nearest Neighbor- (C-KNN), Random Forest (RF)

Similarly, for binary classification, the threshold and scale invariance properties of AUC aid in evaluating the model's performance. From Table 3, L-SVM shows an Accuracy of 77.9%, an F1 score of 0.86, and an AUC of 0.79 predicting high true positives thereby rendering interrelationship among different features for two class classifications. LR has shown an NPV of 0.72 which is high among all classifiers with a high true negative to the total negative predictions. Moreover, better performance of the model reveals the higher AUC values which can distinguish positive and negative classes which are visualized in Figure 5. Further, the study has been extended by involving 50% low and 50% high-grade patients for balanced data set analysis to avoid biases in the models. The outcome of the additional experiments with balanced datasets showed an increased specificity of 85.36% in LD and LR whereas L-SVM showed an AUC of 0.79, Accuracy of 79%, and F1 score of 0.88 among other classifiers. Also, the study will focus on a comparative analysis of different traditional approaches as an extension of this work. In conclusion, the study employing a Radiomics-based Machine Learning algorithm to classify tumor grades shows that prognostic discrimination is possible without guidance or prior knowledge of breast cancer biology for algorithm training.

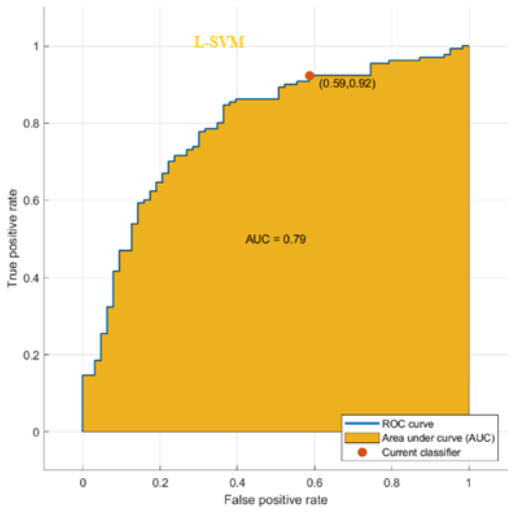


Figure 5. AUC for L-SVM Classifier.

IV. Summary

This study experimented to classify breast tumor grades using different classifiers. LASSO feature selection method with optimal hyperparameter selection has selected 8 optimal features for the evaluation process. The clinical and histopathological characteristics tabulation revealed highly significant differences

between the clinical parameters and tumor grades. A Pearson Correlation Heat Map has been generated for the feature's multi-collinearity identification. Lastly, the collection of classifiers was involved in tumor grade classification.

REFERENCES

[1] L. Wilkinson and T. Gathani, "Understanding breast cancer as a global health concern.," *Br. J. Radiol.,* vol. 95, no. 1130, p. 20211033, Feb. 2022.

[2] V. F. Grabinski and O. W. Brawley, "Disparities in Breast Cancer.," *Obstet. Gynecol. Clin. North Am*., vol. 49, no. 1, pp. 149–165, Mar. 2022.

[3] A. I. Baba and C. Câtoi. Comparative Oncology. Bucharest (RO): *The Publishing House of the Romanian Academy*; 2007. Chapter 3, TUMOR CELL MORPHOLOGY. Available from: https://www.ncbi.nlm.nih.gov/books/NBK9553/

[4] S. M. Telloni, "Tumor Staging and Grading: A Primer.*," Methods Mol. Biol.*, vol. 1606, pp. 1–17, 2017.

[5] H. Chen, Y. Min, K. Xiang, J. Chen, and G. Yin, "DCE-MRI Performance in Triple Negative Breast Cancers: Comparison with Non-Triple Negative Breast Cancers.," *Curr. Med. imaging*, vol. 18, no. 9, pp. 970–976, 2022.

[6] Y. Zhou, G. Zhou, J. Zhang, C. Xu, F. Zhu, and P. Xu, "DCE-MRI based radiomics nomogram for preoperatively differentiating combined hepatocellular-cholangiocarcinoma from mass-forming intrahepatic cholangiocarcinoma.," *Eur. Radiol*., vol. 32, no. 7, pp. 5004–5015, Jul. 2022.

[7] R. A. Gatenby, O. Grove, and R. J. Gillies, "Quantitative imaging in cancer evolution and ecology.," *Radiology*, vol. 269, no. 1, pp. 8–15, Oct. 2013.

[8] J. Wu et al., "Radiological tumor classification across imaging modality and histology.," *Nat. Mach. Intell*., vol. 3, pp. 787–798, Sep. 2021.

[9] F. Cardoso et al., "Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up†.," *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol*., vol. 30, no. 8, pp. 1194–1220, Aug. 2019.

[10] E. A. Rakha et al., "Breast cancer prognostic classification in the molecular era: the role of histological grade.," *Breast Cancer Res*., vol. 12, no. 4, p. 207, 2010.

[11] C. Sotiriou et al., "Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis.," *J. Natl. Cancer Inst*., vol. 98, no. 4, pp. 262–272, Feb. 2006.

[12] E. A. Martin et al., "Current status of biopsy markers for the breast in clinical settings.," *Expert Rev. Med. Devices*, vol. 19, no. 12, pp. 965–975, Dec. 2022.

[13] K. Clark et al., "The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository," *J. Digit. Imaging*, vol. 26, no. 6, pp. 1045–1057, 2013.

[14] A. Saha et al., "A machine learning approach to radiogenomics of breast cancer: a study of 922 subjects and 529 DCE-MRI features.," *Br. J. Cancer*, vol. 119, no. 4, pp. 508–516, Aug. 2018.

[15] Y. Huang et al., "Multi-Parametric MRI-Based Radiomics Models for Predicting Molecular Subtype and Androgen Receptor Expression in Breast Cancer.," *Front. Oncol*., vol. 11, p. 706733, 2021.

[16] H. Lu and J. Yin, "Texture Analysis of Breast DCE-MRI Based on Intratumoral Subregions for Predicting HER2 2+ Status.," *Front. Oncol*., vol. 10, p. 543, 2020,.

[17] A. I. Rozhok and J. DeGregori, "The evolution of lifespan and age-dependent cancer risk.," *Trends in cancer*, vol. 2, no. 10, pp. 552–560, Oct. 2016.

[18] M. C. White, D. M. Holman, J. E. Boehm, L. A. Peipins, M. Grossman, and S. J. Henley, "Age and cancer risk: a potentially modifiable relationship.," *Am. J. Prev. Med*., vol. 46, no. 3 Suppl 1, pp. S7-15, Mar. 2014.

[19] K. I. Sundus, B. H. Hammo, M. B. Al-Zoubi, and A. Al-Omari, "Solving the multicollinearity problem to improve the stability of machine learning algorithms applied to a fully annotated breast cancer dataset," *Informatics Med. Unlocked*, vol. 33, p. 101088, 2022.