The 2023 IEEE Signal Processing in Medicine and Biology Symposium

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

Priyadharshini. B*, Mythili A*, Anandh K R**
*School of Electronics Engineering, Vellore Institute of Technology, Vellore-632014, Tamil Nadu, India.
**Department of Radiology, Cincinnati Children's Hospital Medical Center, Cincinnati -45229, United States

Introduction

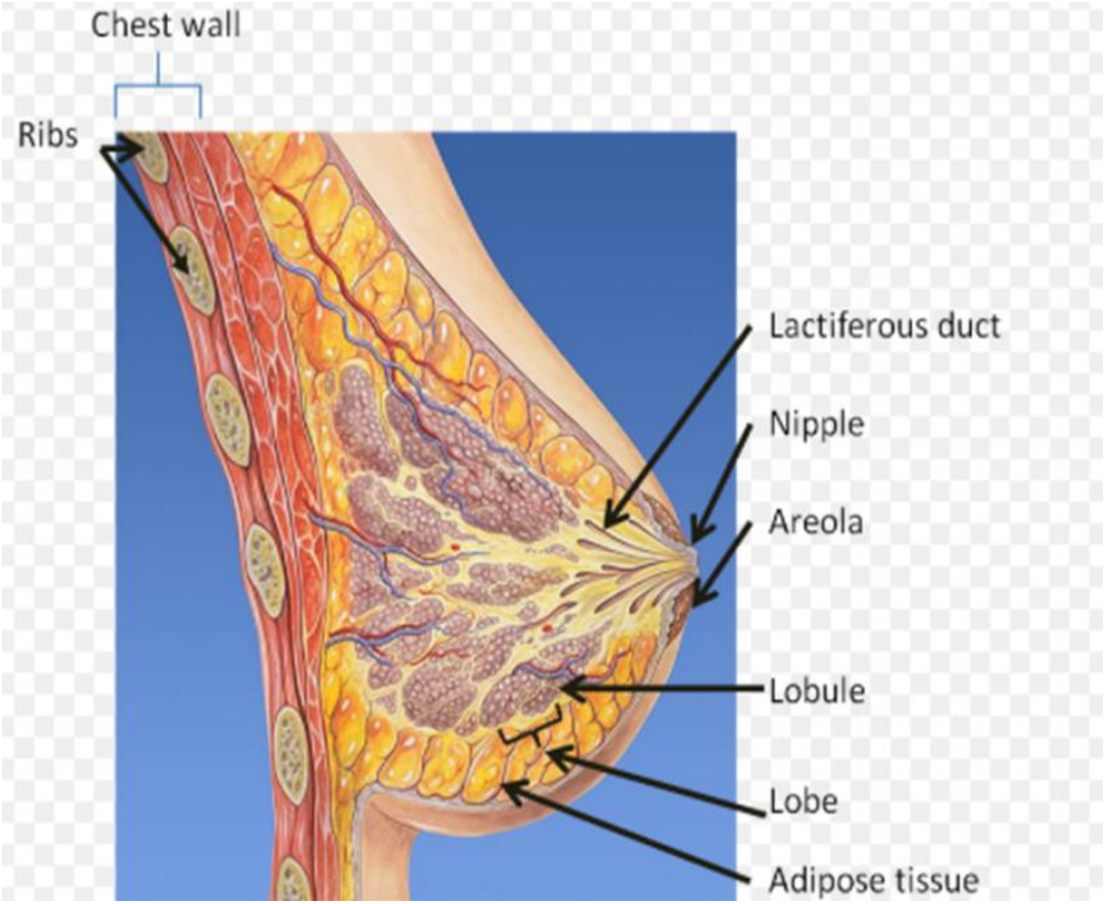
Breast Cancer Facts - In 2020, there were 2.3 million women diagnosed with breast cancer. *(WHO 2020) .*

Causes of Breast cancer-Genetic, Environment, Early menstruation, and Late menopause.

Breast Anatomy- Parts-Ducts, Lobes, Lobules, Lymph node.

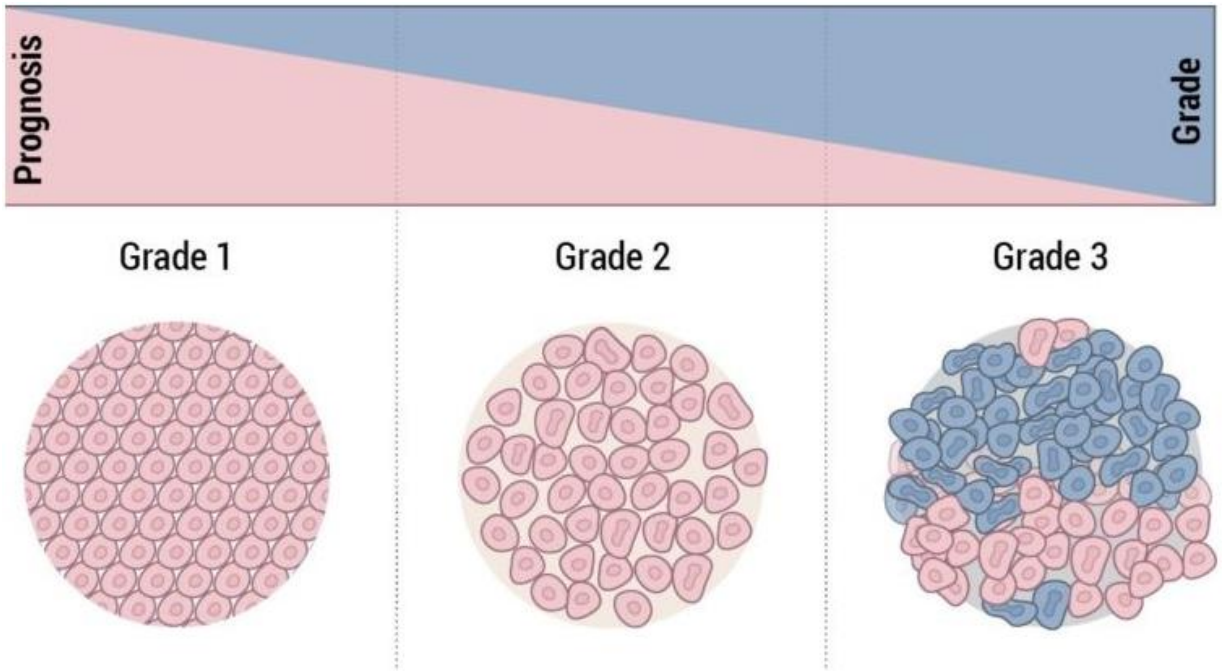
DCE MRI for breast imaging produces high-resolution images for women at high risk of breast cancer and is also effective for evaluating dense female breasts

- Detect microscopic lesions in a (potentially) large volume of tissue. High temporal resolution while preserving high spatial resolution.



Introduction- Grade Facts

- Grade –Prognostic factor and aggressive potential
- **The 3 factors for one of the scoring systems are (the** Nottingham Histologic Score system)
    - amount of gland formation
    - nuclear features
    - mitotic activity



Motivation

- Low grades (Grade I and Grade II) are less aggressive and show an avascular nature with less proliferation of tumors.
- High Grade is a more aggressive, highly intense, highly vascularized, and heterogenous large mass where necrotic, and apoptotic processes take place in the tumor.
- Needle biopsy may be a misinterpretation of the actual grade due to tumor heterogeneity.
- It is essential to ascertain suitable machine learning methods for differentiating low and high-grade breast tumors.

Aim & Objectives

Aim: To analyze Radiomics-based low and high-grade DCE-MR breast tumor classification with a collection of classifiers using LASSO feature selection

Objectives:

- Analysis of clinicopathological characteristics
- Feature selection by LASSO model
- classification of high-grade and low-grade tumors by using a collection of classifiers

Materials & Methods

Dataset description

- A total of 638 patients included in our study where 431(67.55%) were low-grade and 207 (32.44%).

- A total of 529 features named tumor enhancement, shape, enhancement of tissues surrounding, texture, and shape were extracted from the segmented tumor

Feature Selection

- LASSO regression analysis techniques are frequently employed in feature selection and binary classification.

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

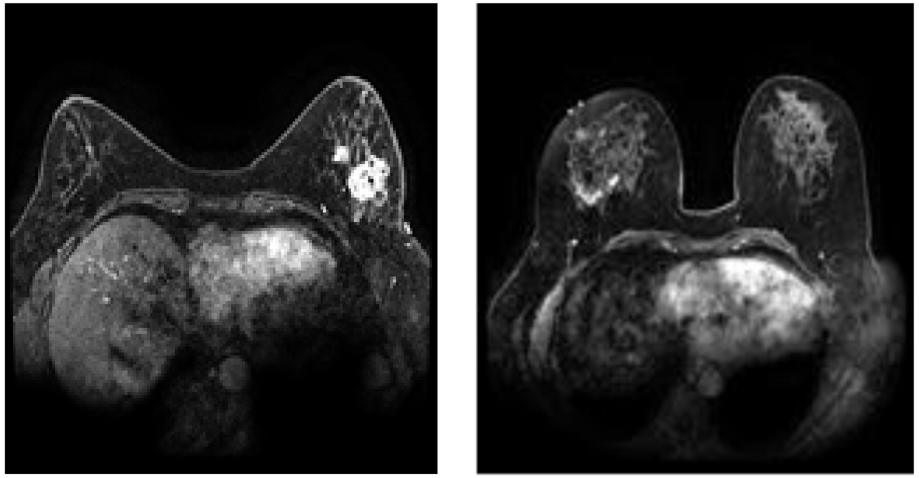
- Pairwise Pearson Correlation Coefficient Matrix (PCCM) identified high-correlated feature pairs

Materials & Methods

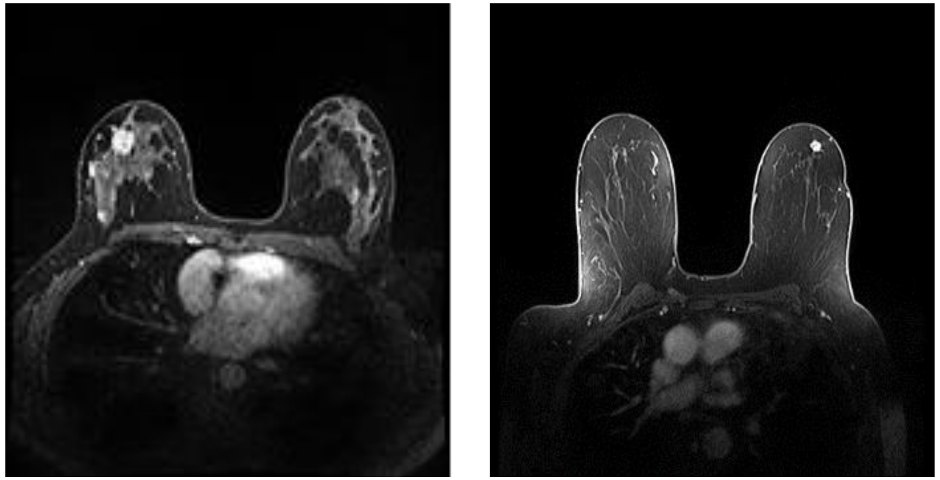
Classifiers

- Logistic regression (LR), k-nearest Neighbors (KNN), Linear discriminant analysis (LDA), Gaussian Naïve Bayes (GNB), Linear Support Vector Machines (LSVM), and Random Forest (RF) were implemented for the classification of Low and High grade
- 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).

Results and Discussions



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

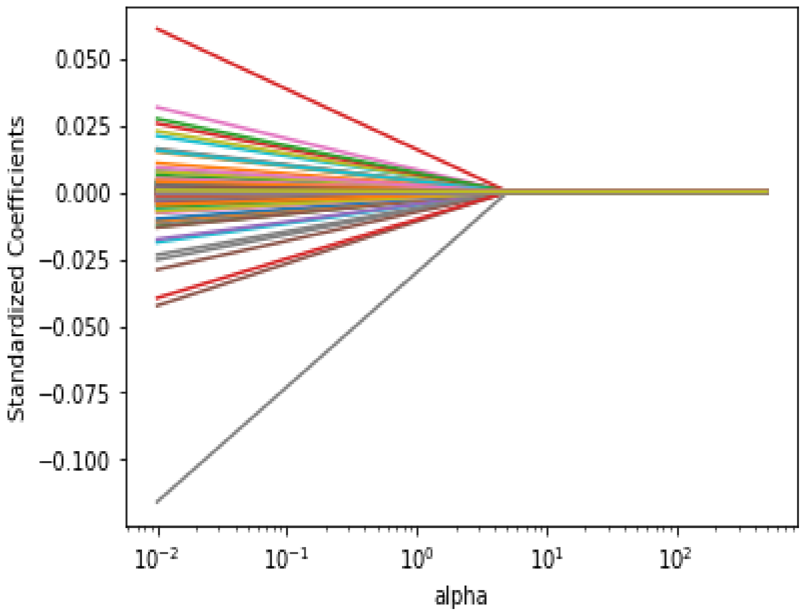


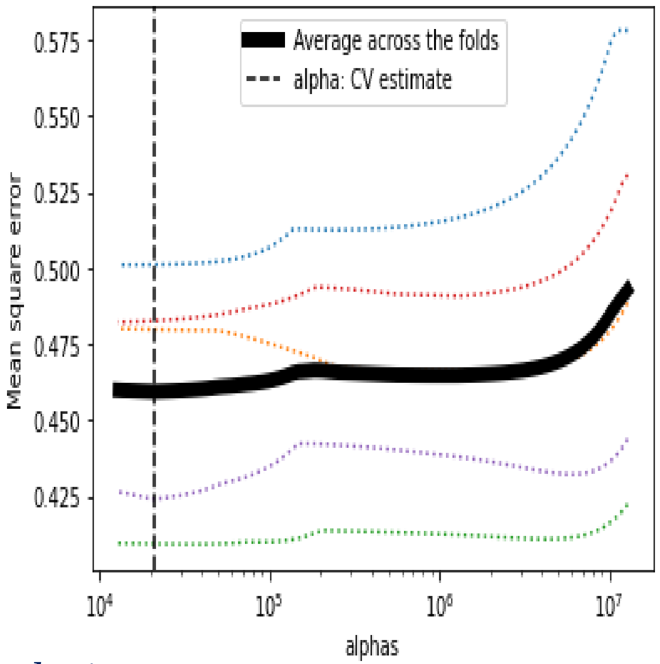
Representative set of breast DCE MR Images of two different Low-grade patients acquired in the axial plane (a) one can appreciate moderate-intensity tumor and (b) one cannot appreciate moderate-intensity tumor.

Results and Discussions

Clinicopathologic Characteristics

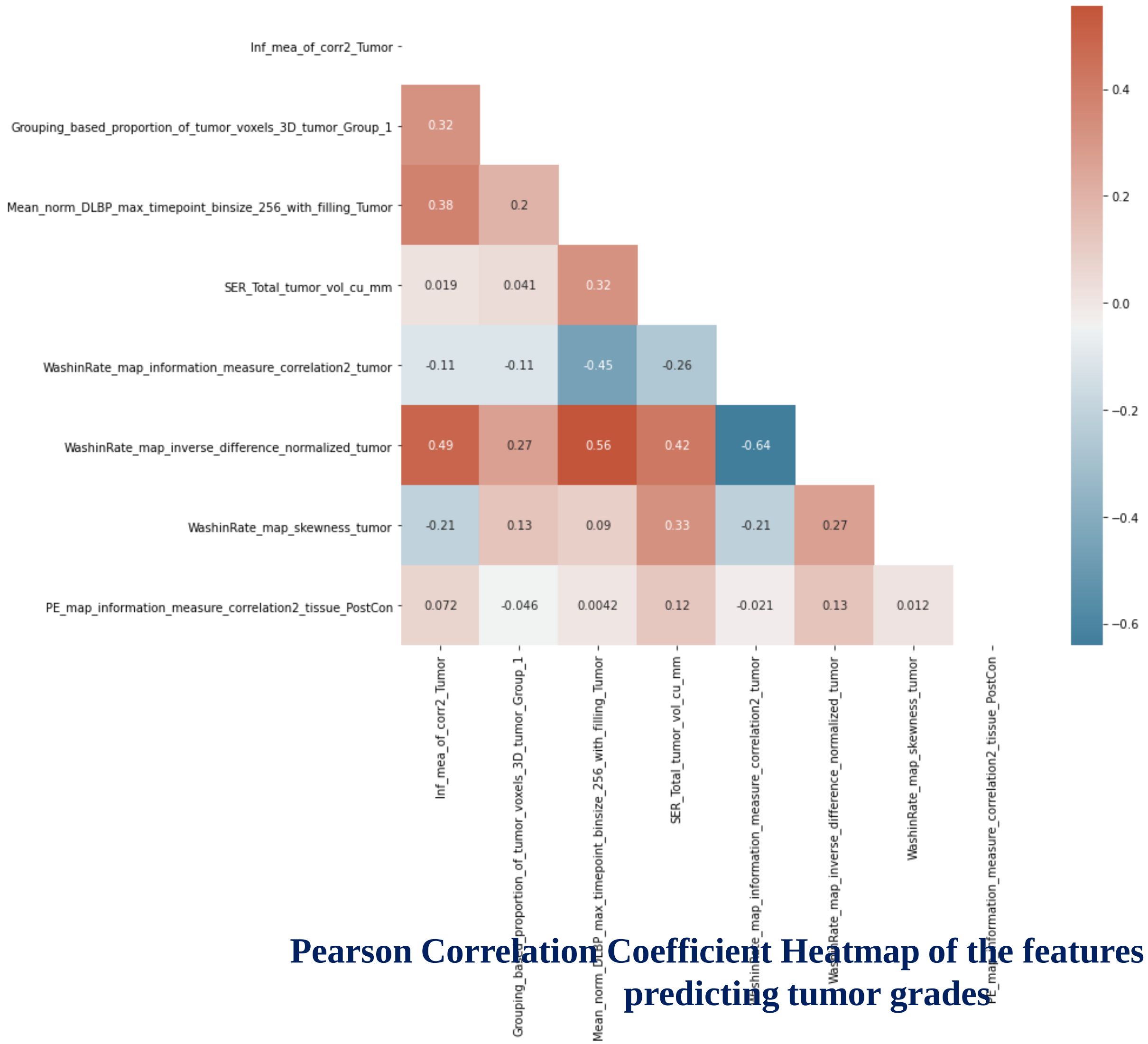
	Low grade	High grade	p-value
No of subjects	431(67.55%)	207(32.44 %)	
Age(Mean±SD)	54.69±10.86	49.90 ± 11.61	0.6921
Estrogen receptors status			<.00001
Positive	376(87.23%)	105(50.72%)	
Negative	55(12.76%)	102(49.27%)	
Progesteron receptor status			<.00001
Positive	338(78%)	75(36.23%)	
Negative	93(21.57%)	132(63.76%)	
HER2 status			0.00239
Positive	62(13.38%)	50(24.15%)	
Negative	369(85.61%)	157 (75.84%)	
Response status			<.00001
PCR	9(2.08%)	37 (17.87%)	
Non-PCR	83(19.25%)	59 (28.50%)	
Not Available	332(77.03%)	103 (49.75%)	
Others	7(1.62%)	9 (4.34%)	
Menopausal Status			0.02629
Premenopausal	179(41.53%)	109(52.65%)	
Postmenopausal	241(55.91%)	95(45.89%)	
Not Available	11(2.55%)	3(1.44%)	
Bilateral status			0.00451
Bilateral	25(5.80%)	2(0.96%)	
Non -Bilateral	406(94.66%)	205(99.03%)	





LASSO Analysis

Results and Discussions



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

Results and Discussions

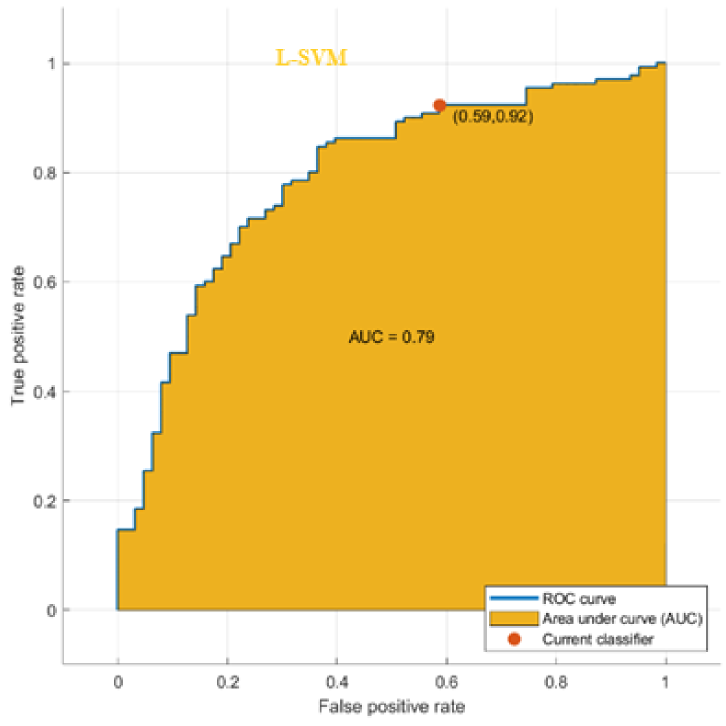
Performance Analysis of Different Classifiers for Categorizing Low and High-Grade

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

Results and Discussions

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



AUC for L-SVM Classifier

Conclusions

- An experiment was conducted 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.
- For the feature’s multi-collinearity identification, a Pearson Correlation Heat Map has been generated.
- 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.

References

[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.