









Article

Sentinel Lymph Node Metastasis on Clinically Negative Patients: Preliminary Results of a Machine Learning Model Based on Histopathological Features

Annarita Fanizzi ¹, Vito Lorusso ², Albino Biafara ³, Samantha Bove ⁴, Maria Colomba Comes ¹, Cristian Cristofaro ¹, Maria Digennaro ⁵, Vittorio Didonna ¹, Daniele La Forgia ^{6,*}, Annalisa Nardone ⁷, Domenico Pomarico ¹, Pasquale Tamborra ¹, Alfredo Zito ⁸, Angelo Virgilio Paradiso ⁵ and Raffaella Massafra ¹

- ¹ Struttura Semplice Dipartimentale di Fisica Sanitaria, I.R.C.C.S. Istituto Tumori “Giovanni Paolo II”, Viale Orazio Flacco 65, 70124 Bari, Italy; a.fanizzi@oncologico.bari.it (A.F.); mariac.comes@libero.it (M.C.C.); c.cristofaro@oncologico.bari.it (C.C.); v.didonna@oncologico.bari.it (V.D.); domenico.pomarico@hotmail.it (D.P.); p.tamborra@oncologico.bari.it (P.T.); r.massafra@oncologico.bari.it (R.M.)
- ² Unità Operativa Complessa di Oncologia Medica, I.R.C.C.S. Istituto Tumori “Giovanni Paolo II”, Viale Orazio Flacco 65, 70124 Bari, Italy; vitolorusso@me.com
- ³ Dipartimento di Economia e Finanza, Università degli Studi di Bari, 70124 Bari, Italy; a.biafara@studenti.uniba.it
- ⁴ Dipartimento di Matematica, Università degli Studi di Bari, 70124 Bari, Italy; s.bove9@studenti.uniba.it
- ⁵ Oncologia Medica Sperimentale, I.R.C.C.S. Istituto Tumori “Giovanni Paolo II”, Viale Orazio Flacco 65, 70124 Bari, Italy; madi1963@libero.it (M.D.); a.paradiso@oncologico.bari.it (A.V.P.)
- ⁶ Struttura Semplice Dipartimentale di Radiologia Senologica, I.R.C.C.S. Istituto Tumori “Giovanni Paolo II”, Viale Orazio Flacco 65, 70124 Bari, Italy
- ⁷ Unità Operativa Complessa di Radioterapia, IRCCS Istituto Tumori “Giovanni Paolo II”, Viale Orazio Flacco 65, 70124 Bari, Italy; a.nardone@oncologico.bari.it
- ⁸ Unità Operativa Complessa di Anatomia Patologica, I.R.C.C.S. Istituto Tumori “Giovanni Paolo II”, Viale Orazio Flacco 65, 70124 Bari, Italy; a.zito@oncologico.bari.it
- * Correspondence: d.laforgia@oncologico.bari.it; Tel.: +39-080-5555111



Citation: Fanizzi, A.; Lorusso, V.; Biafara, A.; Bove, S.; Comes, M.C.; Cristofaro, C.; Digennaro, M.; Didonna, V.; Forgia, D.L.; Nardone, A.; et al. Sentinel Lymph Node Metastasis on Clinically Negative Patients: Preliminary Results of a Machine Learning Model Based on Histopathological Features. *Appl. Sci.* **2021**, *11*, 10372. <https://doi.org/10.3390/app112110372>

Academic Editor:
Aleksander Mendyk

Received: 29 July 2021
Accepted: 28 October 2021
Published: 4 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: The reported incidence of node metastasis at sentinel lymph node biopsy is generally low, so that the majority of women underwent unnecessary invasive axilla surgery. Although the sentinel lymph node biopsy is time consuming and expensive, it is still the intra-operative exam with the highest performance, but sometimes surgery is achieved without a clear diagnosis and also with possible serious complications. In this work, we developed a machine learning model to predict the sentinel lymph nodes positivity in clinically negative patients. Breast cancer clinical and immunohistochemical features of 907 patients characterized by a clinically negative lymph node status were collected. We trained different machine learning algorithms on the retrospective collected data and selected an optimal subset of features through a sequential forward procedure. We found comparable performances for different classification algorithms: on a hold-out training set, the logistics regression classifier with seven features, i.e., tumor diameter, age, histologic type, grading, multiplicity, in situ component and Her2-neu status reached an AUC value of 71.5% and showed a better trade-off between sensitivity and specificity (69.4 and 66.9%, respectively) compared to other two classifiers. On the hold-out test set, the performance dropped by five percentage points in terms of accuracy. Overall, the histological characteristics alone did not allow us to develop a support tool suitable for actual clinical application, but it showed the maximum informative power contained in the same for the resolution of the clinical problem. The proposed study represents a starting point for future development of predictive models to obtain the probability for lymph node metastases by using histopathological features combined with other features of a different nature.

Keywords: sentinel lymph node biopsy; breast cancer; machine learning; histopathological features; clinically negative lymph node

1. Introduction

The prediction of lymph node involvement in breast cancer represents an important task which could reduce unnecessary surgery and improve the definition of oncological therapies [1–6]. The research of a trade-off strategy among time-consuming, expensive and invasive methodologies is the scientific goal of several research studies [7,8], especially with reference to breast cancer patients with clinically negative sentinel lymph nodes.

For patients with clinically negative lymph nodes at the clinical or radiological exam [9], the current guidelines provide the removal of sentinel lymph node biopsy (SLNB). These are the first axillary draining lymph nodes, known as “sentinel” lymph nodes [10–12]. However, the SLNB is time-consuming and expensive, and may lead to complications such as wound infection, seroma, paraesthesia, lymphedema and hematoma [7,13–17]. In patients with early stage breast cancer for whom there is a low incidence of axillary metastases (about 10–25%) it may be an unnecessary invasive procedure [7,12–14,18].

The aim of our work is to develop other less invasive and preferably cheaper diagnostic tools without compromising the diagnosis of patient care. In clinically negative patients, the availability of a new tool able to provide an accurate probabilistic estimation makes surgical actions unnecessary, and this would result in an immediate improvement in both the effectiveness and the quality of care [19].

In our previous work [20] we had generalized an open-source classification algorithm adding other prognostic factors with respect to those used in the original work [21–23]. In this study, we report the results of a multivariate analysis aimed at developing a sentinel lymph nodes status predictive model for patients with and characterized by a clinically negative lymph node status. Particularly, we evaluated the predictive power of different breast cancers’ clinical and immunohistochemical features. With respect to previous work, we introduced new prognostic factors such as Her2/neu, multiplicity and an in situ component and trained three different state-of-the-art machine learning algorithms on the retrospective collected data. Moreover, we implemented a sequential forward procedure to select an optimal subset of features, and we evaluated the results obtained in hold-out cross-validation.

2. Materials and Methods

2.1. Experimental Data

The dataset used in our analysis is composed by the histological outcomes of 907 patients, registered in the period 2015–2018 and referred to Istituto Tumori “Giovanni Paolo II” in Bari (Italy), which resulted negative at both clinical and instrumental examination and had undergone the one-step nucleic acid amplification (OSNA) procedure. This procedure is time consuming and expensive, but it is still the intra-operative exam with the highest performance (it currently has a sensitivity of 87.5–100% and a specificity of 90.5–100%) [12–14].

We considered the patients with clinically negative lymph nodes who did not have suspicious signs in axillary ultrasound, which is a routine examination during the presurgical staging phase of the armpit, or patients that resulted negative after a fine needle aspiration biopsy following the identification of axillary changes on instrumental examination.

For each patient, we collected several prognostic factors characterizing the tumor evaluated on postoperative specimen pathology. The retrospective observational study was approved by the Scientific Board of the Istituto Tumori “Giovanni Paolo II” and carried out according the Helsinki Statement. Based on our regulation on retrospective studies, all patients who gave consent to use the data for scientific purposes were recruited.

2.2. Histological Evaluation Procedure

We collected the immunohistochemical data carried out by our Pathological Anatomy department, such as tumor size (T, measured in mm), histological subtype (ductal, lobular, other special types), estrogen receptor expression (ER, Pos/Neg), progesterone receptor expression (PgR, Pos/Neg), histological grade (G, Elston–Ellis scale: 1, 2, 3), cellular

marker for proliferation (Ki67, Pos/Neg with cut-off 20%), human epidermal growth factor receptor-2 (HER2/neu: 0, 1+, 2+, 3+), tumor multiplicity (Pos/Neg), presence of carcinoma in situ associated with invasive component (Pos/Neg), and also the sentinel lymph nodes status (N, Pos/Neg) required in the associated supervised approach.

The tumor grade G was defined by the Elston–Ellis modification of the Scarff–Bloom–Richardson grading system on a three-grade scale, i.e., grade G1 (low grade), G2 (intermediate grade) or G3 (high grade), where a lower grade indicates a better prognosis [24]. The histological exam was performed through multiple biopsy sampling with 14–16 G core under ultrasound guidance. Immunohistochemical expression is categorized according to the following molecular subtypes on the basis of the St. Gallen convention [25], using a threshold equal to 20% for Ki67 [26]: luminal A (ER+ and/or PR+, HER2- and low Ki67), luminal B (ER+ and/or PR+, and HER2+ or HER2- with high value of Ki67), HER2 positive (ER/PR- and HER2+) and triple negative (ER-, PR- and HER2-).

2.3. Statistical Analysis

In order to evaluate the association between each clinical feature and the sentinel lymph node status, we used the Mann–Whitney test for the age feature measured on an interval scale, whereas we used the Chi-square or Spearman test for all the other features that were measured on a nominal or ordinal scale, respectively. A result was considered statistically significant when the *p*-value was less than 0.05.

Partitioning Around Medoids (PAM) algorithm was employed to identify non-supervised clustering in *k*-groups. Specifically, the PAM algorithm tries to minimize the mean square error, trying to reduce the distance between the points of a cluster and the point that, among the observed data, is located more centrally, called the Medoid. For the optimal estimate of the *k*-medoids (and therefore of the *k* clusters), we used the analysis of the silhouette that allows us to graphically visualize the quality of the clustering. The silhouette index is generally used to identify the optimal number of groups in a hierarchical cluster and as a synthetic indicator to evaluate the overall quality of clustering [27]. Its advantage is the low computational complexity and the simple rules of interpretation.

In order to show the results of the cluster analysis in a bivariate space, we applied a Multiple Correspondence Analysis for dimensionality reducing.

2.4. Classification Models

Three different classification models were trained to predict the sentinel lymph node positivity. We used well-known machine learning methods, which were Random Forest (RF), logistic regression, and Naïve Bayesian.

Random Forests is a well-known ensemble machine learning classifier, which generally provides good performance with low over-fitting [28]. RF provides an embedded method for feature selection: it takes advantage of its own feature selection process and performs classification at the same time. There are two measures of importance for each feature: the first one measures how much the accuracy decreases when a feature is excluded to the forest, the second one measures the decrease in Gini impurity when a feature is chosen to split a node of a tree. In our work, we used this second method. A standard configuration of RF was adopted with 100 trees and 20 features (as described in Breiman (24)) randomly selected at each split because more complicated architectures did not give any significant classification improvement. Moreover, in order to control the over-fitting risk, we have fixed a small number of observations per tree leaf, such as five.

A logistic regression prediction model measures the underlying relationships between features and patient outcomes existing within the data [29]. The accuracy of a logistic regression model is mainly judged by considering discrimination and calibration. Discrimination is the model capability to correctly assign a higher risk of an outcome to the patients who are truly at higher risk, whereas calibration is the model capability to assign the correct average absolute level of risk, i.e., to accurately estimate the probability of the patient outcome.

The Naive Bayes classifier is a probabilistic machine learning model that is used to fulfill classification tasks based on the Bayes theorem [30]. It requires a small amount of training data to estimate the necessary parameters; despite their apparently over-simplified assumptions, Naive Bayes classifiers worked well in many real-world situations.

Default parameters were used for logistic regression and Naive Bayes classifiers.

The considered classifiers are based on a different approach to solve the classification task. A RF classifier is an ensemble technique that combines several decision trees to calculate the predicted class, so that the forecasts made by decision trees, which may be individually inaccurate, aim to improve performance and reduce over-fitting when combined together. The Naïve Bayesian classifier is based on a probabilistic approach to solve a classification problem, whereas the logistic regression model exploits a mathematical approach to the problem through the estimation of the classification score performance by means of a linear combination of features.

Feature importance techniques and classification models were performed using the MATLAB R2018a (Mathworks, Inc., Natick, MA, USA) software.

2.5. Performance Evaluation

We performed the hold-out cross-validation procedure in order to evaluate the classification performance of each learning model to lymph node status. Specifically, we carried out a feature selection analysis and developed the learning classification models on 70% of the samples (training set) randomly selected. Then, we evaluated the obtained results on the remaining 30% of 907 clinically negative patients.

On the training hold-out training set, in order to identify a subset of features with higher diagnostic power, we developed a forward stepwise feature selection. The sequential forward selection algorithm identified a subset of features that best predicted the expected result by sequentially adding at the each step the features that improve the classification performances on 100 ten-fold cross-validation rounds. Specifically, the selection of features according to their importance is driven by the AUC index: in each series of cross-validation, we appended to the prognostic factors sequence the one endowed with the highest median value of AUC distribution. Then, we proceeded with the evaluation of the new distribution for each sequence obtained by adding a feature in the remaining set and comparing the associated median until the remaining prognostic factors were empty.

Performances of each classification model on both training and test hold-out sets are assessed in terms of Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) curve and also F1 score, accuracy, sensitivity and specificity calculated by choosing the optimal threshold through the Youden's index [31].

The performance measures at each feature selection step on the hold-out training set were estimated on 100 ten-fold cross-validation rounds and summarized in terms of the median, first and third quartile.

3. Results

3.1. Statistical Analysis Results

The samples' characteristics are summarized in Table 1. Nine hundred and seven patients aged between 23 and 92 years (with a median, first, and third quartile of 51, 61 and 70 years, respectively) were involved in the study. In total, 296 patients were positive on histological examination of the sentinel lymph nodes, whereas 611 patients resulted negative.

The association test between the sentinel lymph node status and each prognostic factor is essential to understand whether there is a relationship between the two variables. Our experimental results have shown that there was a relationship between the sentinel lymph node status and tumor diameter, ki67, grading, histological type, multiplicity and also age (p -value < 0.05). All the other variables were not significantly associated with sentinel lymph node status.

Table 1. Observed number of patients with positive sentinel lymph nodes according to the considered prognostic factors.

	<i>n</i> Patients	<i>n</i> Positive		<i>n</i> Patients	<i>n</i> Positive
Overall	907	296	ER		
Age §			Negative	94	29
21–40	33	14	Positive	813	267
41–50	128	46	PgR		
51–60	151	50	Negative	202	65
61–70	164	54	Positive	705	231
71–80	114	33	Ki67 *		
>80	45	17	Negative	518	148
Diameter *			Positive	389	148
T1a	46	6	HER2/neu		
T1b	191	30	0	676	227
T1c	420	132	1 +	123	33
T2	250	128	2 +	59	24
Grading *			3 +	49	12
G1	274	62	Multiplicity §		
G2	418	150	Negative	710	205
G3	215	84	Positive	197	91
Histologic type §			In situ component		
Ductal	704	246	Absent	397	133
Lobular	147	41	Present	510	163
Special type	56	9			

* *p*-value Spearman Test < 0.05; § *p*-value Chi-Square Test < 0.05; § *p*-value Mann–Whitney Test < 0.05.

3.2. Classification Results on Hold-Out Training Set

The training set of hold-out validation sampling was made up of 635 clinically negative patients with 421 negative sentinel lymph node cases (control cases) and 214 positive cases at postoperative histological survey.

Figure 1 shows the performance of the three classifiers considered in terms of AUC value (median, first and third quartiles' values) at increasing the number of features selected on the hold-out training set (Table 2).

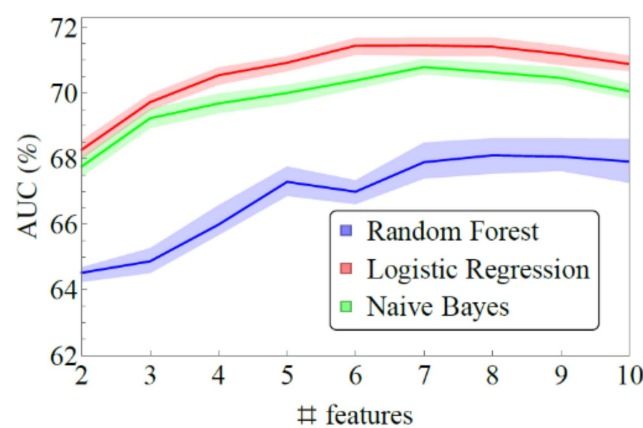


Figure 1. AUC distribution of the three classification models trained on a hold-out training set by increasing the number of features according to our feature selection order as indicated in the table. The solid line stands for the median yielded by 100 ten-fold cross validation, while the shaded region corresponds to the interquartile range.

Table 2. Features selected by each classifier in the different steps of the stepwise procedure implemented.

Number of Features	RF	Logistic Regression	Naive Bayesian
2	Diameter + Multiplicity	Diameter + Age	Diameter + Histologic Type
3	+Histologic Type	+Histologic Type	+Her2/Neu
4	+Grading	+Grading	+Age
5	+In Situ	+Multiplicity	+Grading
6	+Ki67	+In Situ	+Multiplicity
7	+Age	+Her2/Neu	+In Situ
8	+ER	+PGR	+PGR
9	+Her2/Neu	+Ki67	+ER
10	+PgR	+ER	+Ki67

The logistic regression classifier seems to always achieve slightly higher performances than the others, although overcoming the Naïve Bayesian classifier of a few percentage points. The best performance was reached with seven features for the Naïve Bayesian and logistic regression classifiers, whereas eight features were needed for the RF classifier.

In Table 3, we report the performances obtained for each classifier by using the subset of features that maximized performance. Although the performances were comparable in terms of accuracy, the logistics regression classifier reached an accuracy of 71.5% and showed a better trade-off between sensitivity and specificity than the other two classifiers (69.4 and 66.9%, respectively) when seven features, i.e., tumor diameter, age, histologic type, grading, multiplicity, in situ component and Her2-neu status, were considered.

Table 3. Median performance of the best models calculated on the hold-out training set by means of the sequential forward feature selection evaluated on 100 ten-fold cross validation rounds for each of the used classification algorithms. The prediction performances are summarized in terms of median, 1st and 3rd quartile.

Classifier	# Features	AUC (%)	F1 Score	Accuracy (%)	Sensitivity (%)	Specificity (%)
RF	8	68.1 (65.9–68.6)	54.5 (48.9–59.5)	67.7 (65.9–68.8)	57.9 (52.2–61.7)	72.3 (68.9–76.8)
Logistic Regression	7	71.5 (71.1–71.7)	59.2 (55.60–63.7)	67.9 (66.0–68.7)	69.4 (67.3–74.3)	66.9 (61.9–70.0)
Naive Bayesian	7	70.8 (70.6–71.0)	58.4 (54.7–61.8)	66.3 (65.2–67.6)	70.6 (66.4–73.0)	63.9 (61.2–68.0)

3.3. Cluster Analysis on Hold-Out Training Set

The classification performance does not exceed 68% accuracy with any of the used classifiers. In order to highlight any profile characterizing the cases not correctly identified by the classifiers, we implemented unsupervised hierarchical analysis of the groups. On the whole dataset, on the basis of the relative Silhouette index, we identified the optimal number of groups in which to divide the samples (Figure 2). In particular, in Figure 3, we represent the three groups with respect to the first two main components. For the sake of simplicity, we separated the control case clusters from those made up of patients who tested positive for the sentinel lymph node and in each of them, we reported the cases (24 false negative and 25 false positive) that were never correctly classified in all the validation rounds by each of the three implemented classifiers. Evidently, such cases are not characterized by a particular pattern of the used features.

False negatives are mostly related to group 2, characterized by patients with mildly aggressive tumors, i.e., small T1b and T1c, low grading (G1 or G2), negative ki67 and an absent in situ component. On the contrary, the false positives are afferent to group 1 or 3, characterized by T1c and T2 tumors, ki67 positive, medium-high grading G3 or G2 and components present in situ (Table 4).

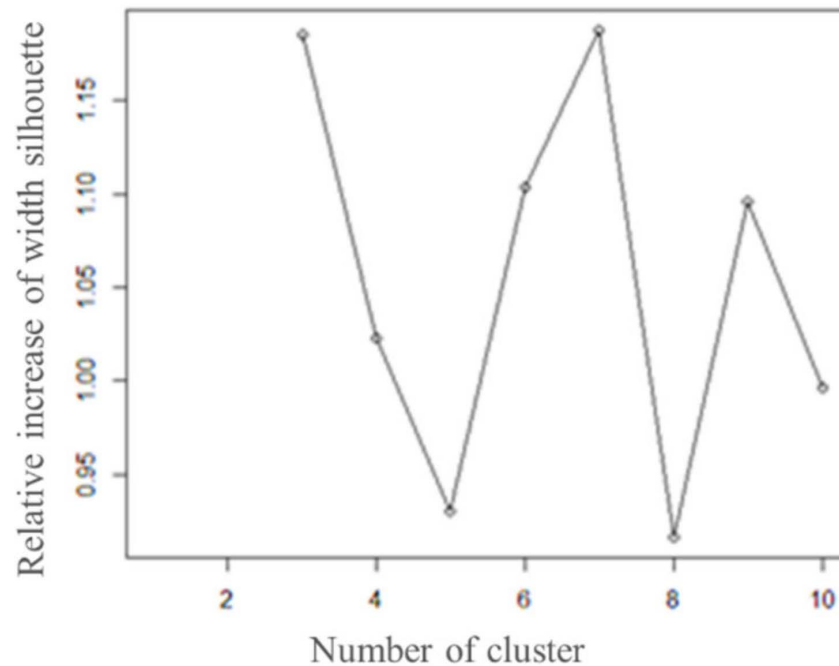


Figure 2. Relative increases in width silhouette related to PAM clustering implemented on hold-out training set.

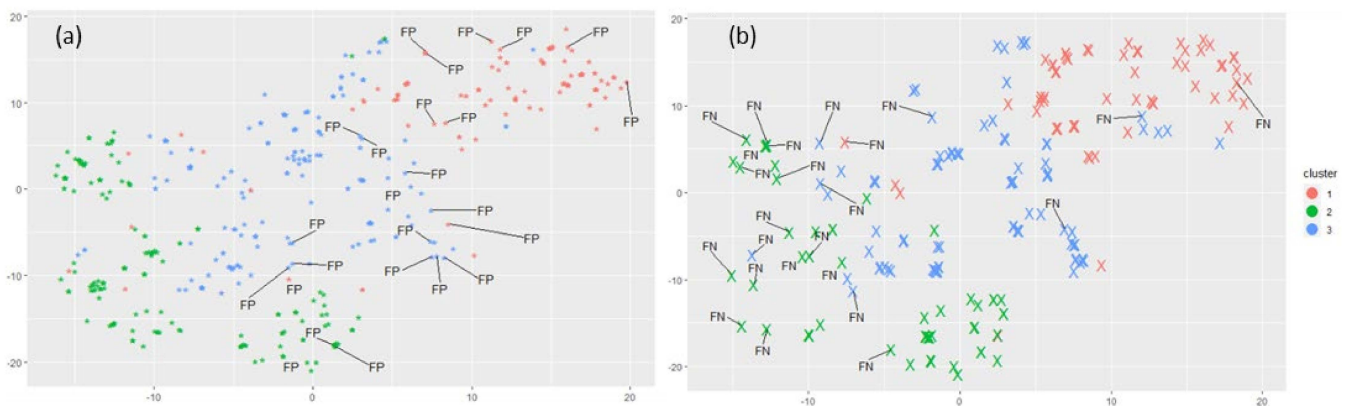


Figure 3. Clustering of positive (a) and negative (b) cases in three groups identified by PAM algorithm on overall hold-out training set. In the clustering plot of positive cases, we have indicated the FN, whereas in the clustering plot of negative cases we have indicated the FP.

3.4. Classification Results on Hold-Out Test Set

In order to evaluate the robustness of obtained results, we calculated the performance classification of the three best models on the hold-out test set by using hold-out validation sampling, which was made up of 272, composed of 190 control cases and 82 cases with negative and positive sentinel lymph nodes, respectively.

The performance of each classifier loses accuracy by about five percentage points (Table 5). However, the logistic regression model still shows a sensitivity higher than 68% and a specificity of about 60%, thus losing about six percentage points compared to the results observed on the hold-out training set.

Table 4. Frequency distribution (percentage) of each group identified by PAM algorithm on the overall hold-out training set.

Features		Cluster	1	2	3
		n (Pos/Neg)	164 (66/98)	207 (70/137)	264 (96/168)
Histologic Type	Ductal		152 (92.7%)	146 (70.5%)	214 (81.1%)
	Lobular		5 (3.1%)	33 (15.9%)	29 (11.0%)
	Special		7 (4.3%)	28 (13.5%)	21 (8.0%)
Diameter	T1a		5 (3.1%)	12 (5.8%)	14 (5.3%)
	T1b		21 (12.8%)	62 (30.0%)	42 (15.9%)
	T1c		40 (24.4%)	94 (45.4%)	147 (55.7%)
	T2		98 (59.8%)	39 (18.8%)	61 (23.1%)
Age	Median		57	61	61
ER	Neg		61 (37.2%)	0.0	3 (1.1%)
	Pos		103 (62.9%)	207 (100%)	261 (98.8%)
PgR	Neg		112 (68.3%)	20 (9.7%)	15 (5.7%)
	Pos		52 (31.7%)	187 (90.3%)	249 (94.3%)
Ki67	Neg		13 (7.9%)	154 (74.4%)	65 (24.6%)
	Pos		151 (92.1%)	53 (25.6%)	199 (75.4%)
Grading	G1		10 (6.1%)	130 (62.8%)	35 (13.3%)
	G2		15 (9.1%)	71 (34.3%)	201 (76.1%)
	G3		139 (84.8%)	6 (2.9%)	28 (10.6%)
Her2/Neu	0		93 (56.7%)	177 (85.5%)	202 (76.5%)
	1+		19 (11.6%)	26 (12.6%)	33 (12.5%)
	2+		20 (12.2%)	3 (1.4%)	23 (8.7%)
Multiplicity	3+		32 (19.5%)	1 (0.5%)	6 (2.3%)
	Neg		123 (75.0%)	177 (85.5%)	192 (72.7%)
	Pos		41 (25.0%)	30 (14.5%)	72 (27.3%)
In Situ	Neg		54 (32.9%)	153 (73.9%)	59 (22.3%)
	Pos		110 (67.1%)	54 (26.1%)	205 (77.7%)

Table 5. Classification performance calculated on the hold-out test set by using the subset of features that maximized performance for each classifier.

Classifier	# Features	Accuracy (%)	Sensitivity (%)	Specificity (%)
RF	8	63.6	52.4	68.4
Logistic Regression	7	62.1	68.3	59.7
Naïve Bayesian	7	61.8	52.4	65.8

4. Discussion

In patients with clinically node-negative breast cancer, current international guidelines require sentinel lymph node biopsy (SLNB) [11–14]. Although it is currently the most-performed exam, it is a time consuming and expensive procedure and is also extremely invasive, and could result in a number of side effects [7,18,32]. However, in patients with early stage breast cancer the incidence of axillary metastases is low, therefore that procedure may be an unnecessary invasive procedure.

The aim of our work was to develop a sentinel lymph nodes status predictive model for clinically negative patients that could replace the SLNB. Starting from a set of 907 patients, for each of them we collected several prognostic features, such as age, tumor size, histological subtype, estrogen receptor expression, progesterone receptor expression, histological grade, cellular marker for proliferation, human epidermal growth factor receptor-2, multiplicity, in situ component and also the sentinel lymph nodes status. Subsequently, we trained three different classifiers combined with a sequential forward feature selection algorithm on the hold-out training set to select the feature subset which reached the highest

value of AUC. Then, we trained the three previously selected models on the hold-out test set and we compared their performances in terms of accuracy, sensitivity and specificity.

On the hold-out training set, the best classification model was the Logistic Regression algorithm on a subset of seven features, i.e., tumor size, age, histologic type, grading, multiplicity, in situ component and Her2-neu status. This model reached a median AUC equals to 71.5%, a median accuracy of 67.9%, sensitivity equal to 69.4% and specificity of 66.9%.

Furthermore, on the same training set we performed a cluster analysis, in order to identify the positive or negative contribution of different feature subsets for classification purposes. The cluster analysis highlighted the correlation among some feature subsets and the realization of either false-positive or false-negative cases. Indeed, we observed that false negatives were afferent to features such as small T1b and T1c tumors, low grading (G1 or G2), negative ki67 and absent in situ component; instead, false positives were related to features as T1c and T2 tumors, ki67 positive, medium-high grading G3 or G2 and present during the in situ component.

Finally, we observed that the performance of the best Logistic Regression model, as well as the performances of the other ones, were lower on the hold-out test set. Even though the best model in terms of accuracy and specificity resulted the RF classifier on a subset of eight features, the only model with a sensitivity still greater than 68% was the Logistic Regression algorithm, which had the same result as the best model of the hold-out training set. The state-of-the-art model is characterized by many works, which propose non-sentinel lymph nodes status predictive models based on features of different nature [33–37]. On the contrary, there are a low number of studies whose aim was the development of a sentinel lymph nodes status predictive model through the analysis of histological features [20,38–41]. Thus far, the nomogram developed by the researchers of Memorial Sloan-Kettering Cancer Center (MSKCC) (a, b, c) is the most widely used model to predict the likelihood of SLN metastasis. The baseline model (a) reached an AUC value of 75.4%. In subsequent validation studies, comparable performances were achieved (b, c). However, these models were evaluated on cohorts of non-clinically negative patients and therefore different from the sample object of our study.

The model developed in [41] for prediction of sentinel lymph nodes metastasis reached an AUC value equal to 88.3% by considering some histological features, such as tumor size, and lymph vascular invasion in ER-positive and HER2-negative (ER+/HER2-), but also genetic ones.

In our previous work [20], we evaluated the usefulness of the CancerMath tool to predict the sentinel lymph node status for clinically negative patients referred to our institute. In this validation study, by using tumor size, age, histologic type, grading, expression of estrogen receptor, and progesterone receptor ki67 and HER2 on the independent test set, the model showed an accuracy of 53.8%, which is much lower than the accuracy value achieved thanks to the analysis exposed here.

The current work differs from the other mentioned research studies because of the use of a machine learning approach for the analysis. Although the model can be further optimized by evaluating other feature selection techniques or classification algorithms on a very extensive experimental database, the only histological features did not allow us to develop a support instrument suitable for actual clinical application. What emerges, and which is important to underline to a public of biomedical data scientists or clinician with a transactional interest field, is that the maximum informative power contained in the only characteristics considered for the prediction of the sentinel lymph node status does not exceed 62–63% on the independent test. For this reason, encouraged by recent studies with improved results in the prediction of the lymph node metastasis probability thanks to the joint use of histopathologic and radiomic features [6,40–54], in our future works we will also involve radiomic features extracted from first-level radiological examinations, such as ultrasound and mammography.

5. Conclusions

In patients with clinically negative lymph nodes, the incidence of lymph nodes is generally low. Although for this type of patient the sentinel lymph node biopsy (SLNB) is the intra-operative exam with the highest performance, it often represents a useless invasive procedure with possible serious complications. Therefore, an important clinical task is to develop a procedure that could surrogate SLNB without compromising the quality of care. In this work, we presented a preliminary model to predict the sentinel lymph node status. We trained different machine learning algorithms on tumor histopathology features but the performances evaluated on a hold-out test set reached an accuracy classification of about 64%. Therefore, the model trained on histopathologic features is yet not suitable for clinical use for the prediction of metastatic lymph nodes in clinically negative patients. However, in future works we will explore other information sources, such as radiomics and genetics, in order to use them conjointly with clinical data considered in this study. By reaching high levels of accuracy, the use of such a support system would have a high clinical impact, either avoiding the sentinel lymph node procedure or reducing the time and cost of surgical interventions, unnecessary axillary dissections with related comorbidities.

Author Contributions: Conceptualization, A.F., D.L.F., D.P. and R.M.; data curation, A.F., M.D., D.P., A.Z., A.V.P. and R.M.; formal analysis, A.F., A.V.P. and R.M.; methodology, A.F.; resources, V.L. and A.Z.; software, A.F., A.B. and D.P.; supervision, R.M.; writing—original draft, A.F., D.P. and R.M.; writing—review and editing, A.F., V.L., A.B., S.B., M.C.C., C.C., M.D., V.D., D.L.F., A.N., D.P., P.T., A.Z., A.V.P. and R.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Scientific Board of Istituto Tumori ‘Giovanni Paolo II’—Bari, Italy.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. *Cancers* 2021; 13, 352 10 of 12.

Data Availability Statement: The data presented in this study are available on request to the corresponding author. The data are not publicly available because they are the property of Istituto Tumori ‘Giovanni Paolo II’—Bari, Italy.

Acknowledgments: This work was supported by funding from the Italian Ministry of Health “Ricerca Finalizzata 2018”.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Farnsworth, R.H.; Achen, M.G.; Stacker, S.A. The evolving role of lymphatics in cancer metastasis. *Curr. Opin. Immunol.* **2018**, *53*, 64–73. [[CrossRef](#)]
2. Nagahashi, M.; Ramachandran, S.; Rashid, O.M.; Takabe, K. Lymphangiogenesis: A new player in cancer progression. *World J. Gastroenterol.* **2010**, *16*, 4003–4012. [[CrossRef](#)] [[PubMed](#)]
3. Argentiero, A.; de Summa, S.; di Fonte, R.; Iacobazzi, R.M.; Porcelli, L.; da Vià, M.; Brunetti, O.; Azzariti, A.; Silvestris, N.; Solimando, A.G. Gene expression comparison between the lymph node-positive and -negative reveals a peculiar immune microenvironment signature and a therapeutic role for WNT targeting in pancreatic ductal adenocarcinoma: A pilot study. *Cancers* **2019**, *11*, 942. [[CrossRef](#)] [[PubMed](#)]
4. Egger, M.E.; Stevenson, M.; Bhutiani, N.; Jordan, A.C.; Scoggins, C.R.; Philips, P.; Martin, R.C.; McMasters, K.M. Age and Lymphovascular Invasion Accurately Predict Sentinel Lymph Node Metastasis in T2 melanoma Patients. *Ann. Surg. Oncol.* **2019**, *26*, 3955–3961. [[CrossRef](#)] [[PubMed](#)]
5. Falck, A.K.; Fernö, M.; Bendahl, P.-O.; Rydén, L. St Gallen molecular subtypes in primary breast cancer and matched lymph node metastases—aspects on distribution and prognosis for patients with luminal tumours: Results from a prospective randomised trial. *BMC Cancer* **2013**, *13*, 558. [[CrossRef](#)] [[PubMed](#)]
6. Qiu, S.-Q.; Zeng, H.-C.; Zhang, F.; Chen, C.; Huang, W.-H.; Pleijhuis, R.G.; Wu, J.-D.; van Dam, G.M.; Zhang, G.-J. A nomogram to predict the probability of axillary lymph node metastasis in early breast cancer patients with positive axillary ultrasound. *Sci. Rep.* **2016**, *6*, 21196. [[CrossRef](#)] [[PubMed](#)]

7. Land, S.R.; Kopec, J.A.; Julian, T.B.; Brown, A.M.; Anderson, S.J.; Krag, D.N.; Christian, N.J.; Costantino, J.P.; Wolmark, N.; Ganz, P.A. Patient-reported outcomes in sentinel node-negative adjuvant breast cancer patients receiving sentinel-node biopsy or axillary dissection: National Surgical Adjuvant Breast and Bowel Project phase III protocol B-32. *J. Clin. Oncol.* **2010**, *28*, 3929–3936. [[CrossRef](#)] [[PubMed](#)]
8. Diotaiuti, S.; de Summa, S.; Altieri, R.; Dantona, C.; Tommasi, S.; di Gennaro, M.; Rubini, G.; Pastena, M.I.; Argentiero, A.; Zito, F.A.; et al. Biomarker phenotyping drives clinical management in axillary sentinel node: A retrospective study on women with primary breast cancer in 2002. *Oncol. Lett.* **2020**, *20*, 2469–2476. [[CrossRef](#)] [[PubMed](#)]
9. Gruppo di lavoro AIRO per la Patologia Mammaria. *Consensus Per L'irradiazione delle Stazioni Linfonodali Mammarie*; Associazione Italiana Radioterapia Oncologica: Milano, Italy, 2016.
10. Early and locally advanced breast cancer: Diagnosis and management. In *NICE Guideline NG101, Evidence Reviews*; National Institute for Health and Care Excellence: London, UK, 2018.
11. Tsujimoto, M.; Nakabayashi, K.; Yoshidome, K.; Kaneko, T.; Iwase, T.; Akiyama, F.; Kato, Y.; Tsuda, H.; Ueda, S.; Sato, K.; et al. One-step Nucleic Acid Amplification for Intraoperative Detection of Lymph Node Metastasis in Breast Cancer Patients. *Clin. Cancer Res.* **2007**, *13*, 4807–4816. [[CrossRef](#)] [[PubMed](#)]
12. Galimberti, V.; Fontana, S.R.; Maisonneuve, P.; Steccanella, F.; Vento, A.R.; Intra, M.; Luini, A. Sentinel node biopsy after neoadjuvant treatment in breast cancer: Five-year follow-up of patients with clinically node-negative or node-positive disease before treatment. *Eur. J. Surg. Oncol.* **2016**, *42*, 361–368. [[CrossRef](#)] [[PubMed](#)]
13. Banerjee, S.M.; Michalopoulos, N.V.; Williams, N.R.; Davidson, T.; el Sheikh, S.; McDermott, N.; Tran-Dang, M.-A.; Davison, S.; Keshtgar, M.R. Detailed evaluation of one step nucleic acid (OSNA) molecular assay for intra-operative diagnosis of sentinel lymph node metastasis and prediction of non-sentinel nodal involvement: Experience from a London Teaching Hospital. *Breast* **2014**, *23*, 378–384. [[CrossRef](#)]
14. Szychta, P.; Westfal, B.; Maciejczyk, R.; Smolarz, B.; Romanowicz, H.; Krawczyk, T.; Zadrożny, M. Intraoperative diagnosis of sentinel lymph node metastases in breast cancer treatment with onestep nucleic acid amplification assay (OSNA). *Arch. Med. Sci.* **2016**, *6*, 1239–1246. [[CrossRef](#)] [[PubMed](#)]
15. Fanfani, F.; Monterossi, G.; Ghizzoni, V.; Rossi, E.D.; Dinoi, G.; Inzani, F.; Fagotti, A.; Alletti, S.G.; Scarpellini, F.; Nero, C.; et al. One-Step Nucleic Acid Amplification (OSNA): A fast molecular test based on CK19 mRNA concentration for assessment of lymph-nodes metastases in early stage endometrial cancer. *PLoS ONE* **2018**, *13*, e0195877. [[CrossRef](#)]
16. Yan, M.; Abdi, M.A.; Falkson, C. Axillary Management in Breast Cancer Patients: A comprehensive review of the key trials. *Clin. Breast Cancer* **2018**, *18*, e1251–e1259. [[CrossRef](#)]
17. DiSipio, T.; Rye, S.; Newman, B.; Hayes, S. Incidence of unilateral lymphedema after breast cancer: A systematic review and meta-analysis. *Lancet Oncol.* **2013**, *14*, 500–515. [[CrossRef](#)]
18. Giuliano, A.E.; Hunt, K.K.; Ballman, K.V.; Beitsch, P.D.; Whitworth, P.W.; Blumencranz, P.W.; Leitch, A.M.; Saha, S.; McCall, L.M.; Morrow, M. Axillary dissection vs no axillary dissection in women with invasive breast cancer and sentinel node metastasis: A randomized clinical trial. *JAMA* **2011**, *305*, 569–575. [[CrossRef](#)] [[PubMed](#)]
19. Mazo, C.; Kearns, C.; Mooney, C.; Gallagher, W.M. Clinical Decision Support Systems in Breast Cancer: A Systematic Review. *Cancers* **2020**, *12*, 369. [[CrossRef](#)]
20. Fanizzi, A.; Pomarico, D.; Paradiso, A.; Bove, S.; Didonna, V.; Diotaiuti, S.; Giotta, F.; La Forgia, D.; Latorre, A.; Pastena, M.I.; et al. Predicting of Sentinel Lymph Node Status in Breast Cancer Patients with Clinically Negative Nodes: A Validation Study. *Cancers* **2021**, *13*, 352. [[CrossRef](#)] [[PubMed](#)]
21. Available online: <http://www.lifemath.net/cancer/> (accessed on 15 March 2020).
22. Chen, L.L.; Nolan, M.E.; Silverstein, M.J.; Mihm, M.C.; Sober, A.J.; Tanabe, K.K.; Smith, B.L.; Younger, J.; Michaelson, J.S. The impact of primary tumor size, lymph node status, and other prognostic factors on the risk of cancer death. *Cancer* **2009**, *115*, 5071–5083. [[CrossRef](#)] [[PubMed](#)]
23. Michaelson, J.S.; Chen, L.L.; Silverstein, M.J.; Mihm, M.C.; Sober, A.J.; Tanabe, K.K.; Smith, B.L.; Younger, J. How cancer at the primary site and in the lymph nodes contributes to the risk of cancer death. *Cancer* **2009**, *115*, 5095–5107. [[CrossRef](#)] [[PubMed](#)]
24. Egner, J.R. AJCC Cancer Staging Manual. *J. Am. Med. Assoc.* **2010**, *304*, 1726. [[CrossRef](#)]
25. Coates, A.S.; Winer, E.P.; Goldhirsch, A.; Gelber, R.D.; Gnant, M.; Piccart-Gebhart, M.; Thürlimann, B.; Senn, H.-J.; Andre, F.; Baselga, J.; et al. Tailoring therapies-improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer. *Ann. Oncol.* **2015**, *26*, 1533e46. [[CrossRef](#)] [[PubMed](#)]
26. Tay, T.K.Y.; Thike, A.A.; Pathmanathan, N.; Jara-Lazaro, A.R.; Iqbal, J.; Sng, A.S.H.; Ye, H.S.; Lim, J.C.T.; Koh, V.C.Y.; Tan, J.S.Y.; et al. Using computer assisted image analysis to determine the optimal Ki67 threshold for predicting outcome of invasive breast cancer. *Oncotarget* **2018**, *9*, 11619–11630. [[CrossRef](#)]
27. Dudek, A. Silhouette Index as Clustering Evaluation Tool. In *Conference of the Section on Classification and Data Analysis of the Polish Statistical Association*; Springer: Cham, Switzerland, 2019; pp. 19–33.
28. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
29. Meurer, W.J.; Tolles, J. Logistic regression diagnostics: Understanding how well a model predicts outcomes. *JAMA* **2017**, *317*, 1068–1069. [[CrossRef](#)]
30. Zhang, H. The optimality of naive Bayes. *AA* **2004**, *1*, 3.
31. Youden, W.J. Index for rating diagnostic tests. *Cancer* **1950**, *3*, 32–35. [[CrossRef](#)]

32. Cabioglu, N.; Yazici, M.S.; Arun, B.; Broglio, K.R.; Hortobagyi, G.N.; Price, J.E.; Sahin, A. CCR7 and CXCR4 as Novel Biomarkers Predicting Axillary Lymph Node Metastasis in T1 Breast Cancer. *Clin. Cancer Res.* **2005**, *11*, 5686–5693. [[CrossRef](#)] [[PubMed](#)]
33. Houvenaeghel, G.; Lambaudie, E.; Classe, J.M.; Mazouni, C.; Giard, S.; Cohen, M.; Faure, C.; Charitansky, H.; Rouzier, R.; Daraï, E.; et al. Lymph node positivity in different early breast carcinoma phenotypes: A predictive model. *BMC Cancer* **2019**, *19*, 45. [[CrossRef](#)]
34. Chen, K.; Liu, J.; Li, S.; Jacobs, L. Development of nomograms to predict axillary lymph node status in breast cancer patients. *BMC Cancer* **2017**, *17*, 561. [[CrossRef](#)] [[PubMed](#)]
35. Ahn, S.K.; Kim, M.K.; Kim, J.; Lee, E.; Yoo, T.-K.; Lee, H.-B.; Kang, Y.J.; Kim, J.; Moon, H.-G.; Chang, J.M.; et al. Can we skip intraoperative evaluation of sentinel lymph nodes? Nomogram predicting involvement of three or more axillary lymph nodes before breast cancer surgery. *Cancer Res. Treat. Off. J. Korean Cancer Assoc.* **2017**, *49*, 1088. [[CrossRef](#)]
36. Bonsang-Kitzis, H.; Mouttet-Boizat, D.; Guillot, E.; Feron, J.G.; Fourchette, V.; Alran, S.; Pierga, J.-Y.; Cottu, P.; Lerebours, F.; Stevens, D.; et al. Medico-economic impact of MSKCC non-sentinel node prediction nomogram for ER-positive HER2-negative breast cancers. *PLoS ONE* **2017**, *12*, e0169962. [[CrossRef](#)]
37. Balachandran, V.P.; Gonen, M.; Smith, J.J.; DeMatteo, R.P. Nomograms in oncology: More than meets the eye. *Lancet Oncol.* **2015**, *16*, e173–e180. [[CrossRef](#)]
38. Bevilacqua, J.L.B.; Kattan, M.W.; Fey, J.V.; Cody, H.S., III; Borgen, P.I.; Van Zee, K.J. Doctor, what are my chances of having a positive sentinel node? A validated nomogram for risk estimation. *J. Clin. Oncol.* **2007**, *25*, 3670–3679. [[CrossRef](#)] [[PubMed](#)]
39. Chen, J.Y.; Chen, J.J.; Yang, B.L.; Liu, Z.B.; Huang, X.Y.; Liu, G.Y.; Wu, J. Predicting sentinel lymph node metastasis in a Chinese breast cancer population: Assessment of an existing nomogram and a new predictive nomogram. *Breast Cancer Res. Treat.* **2012**, *135*, 839–848. [[CrossRef](#)] [[PubMed](#)]
40. Klar, M.; Foeldi, M.; Markert, S.; Gitsch, G.; Stickeler, E.; Watermann, D. Good prediction of the likelihood for sentinel lymph node metastasis by using the MSKCC nomogram in a German breast cancer population. *Ann. Surg. Oncol.* **2009**, *16*, 1136–1142. [[CrossRef](#)] [[PubMed](#)]
41. Okuno, J.; Miyake, T.; Sota, Y.; Tanei, T.; Kagara, N.; Naoi, Y.; Shimoda, M.; Shimazu, K.; Kim, S.J.; Noguchi, S. Development of prediction model including microRNA expression for sentinel lymph node metastasis in ER-positive and HER2-negative breast cancer. *Ann. Surg. Oncol.* **2021**, *28*, 310–319. [[CrossRef](#)] [[PubMed](#)]
42. Dong, Y.; Feng, Q.; Yang, W.; Lu, Z.; Deng, C.; Zhang, L.; Lian, Z.; Liu, J.; Luo, X.; Pei, S.; et al. Preoperative prediction of sentinel lymph node metastasis in breast cancer based on radiomics of T2-weighted fat-suppression and diffusion-weighted MRI. *Eur. Radiol.* **2018**, *28*, 582–591. [[CrossRef](#)]
43. Luo, J.; Ning, Z.; Zhang, S.; Feng, Q.; Zhang, Y. Bag of deep features for preoperative prediction of sentinel lymph node metastasis in breast cancer. *Phys. Med. Biol.* **2018**, *63*, 245014. [[CrossRef](#)]
44. Liu, M.; Mao, N.; Ma, H.; Dong, J.; Zhang, K.; Che, K.; Duan, S.; Zhang, X.; Shi, Y.; Xie, H. Pharmacokinetic parameters and radiomics model based on dynamic contrast enhanced MRI for the preoperative prediction of sentinel lymph node metastasis in breast cancer. *Cancer Imaging* **2020**, *20*, 65. [[CrossRef](#)]
45. Liu, J.; Sun, D.; Chen, L.; Fang, Z.; Song, W.; Guo, D.; Ni, T.; Liu, C.; Feng, L.; Xia, Y.; et al. Radiomics analysis of dynamic contrast-enhanced magnetic resonance imaging for the prediction of sentinel lymph node metastasis in breast cancer. *Front. Oncol.* **2019**, *9*, 980. [[CrossRef](#)]
46. Fanizzi, A.; Basile, T.M.A.; Losurdo, L.; Bellotti, R.; Bottigli, U.; Dentamaro, R.; Didonna, V.; Fausto, A.; Massafra, R.; Moschetta, M.; et al. A Machine Learning Approach on Multiscale Texture Analysis for Breast Microcalcification Diagnosis. *BMC Bioinform.* **2020**, *21*, 91. [[CrossRef](#)] [[PubMed](#)]
47. Fanizzi, A.; Basile, T.M.A.; Losurdo, L.; Bellotti, R.; Bottigli, U.; Campobasso, F.; Didonna, V.; Fausto, A.; Massafra, R.; Tagliafico, A.; et al. Ensemble DiscreteWavelet Transform and Gray-Level Co-Occurrence Matrix for Microcalcification Cluster Classification in Digital Mammography. *Appl. Sci.* **2019**, *9*, 5388. [[CrossRef](#)]
48. Losurdo, L.; Fanizzi, A.; Basile, T.M.A.; Bellotti, R.; Bottigli, U.; Dentamaro, R.; Didonna, V.; Lorusso, V.; Massafra, R.; Tamborra, P.; et al. Radiomics Analysis on Contrast-Enhanced Spectral Mammography Images for Breast Cancer Diagnosis: A Pilot Study. *Entropy* **2019**, *21*, 1110. [[CrossRef](#)]
49. Fanizzi, A.; Losurdo, L.; Basile, T.M.A.; Bellotti, R.; Bottigli, U.; Delogu, P.; Diacono, D.; Didonna, V.; Fausto, A.; Lombardi, A.; et al. Fully Automated Support System for Diagnosis of Breast Cancer in Contrast-Enhanced Spectral Mammography Images. *J. Clin. Med.* **2019**, *8*, 891. [[CrossRef](#)] [[PubMed](#)]
50. Basile, T.M.A.; Fanizzi, A.; Losurdo, L.; Bellotti, R.; Bottigli, U.; Dentamaro, R.; Didonna, V.; Fausto, A.; Massafra, R.; Moschetta, M.; et al. Microcalcification Detection in Full-Field Digital Mammograms: A Fully Automated Computer-Aided System. *Phys. Med.* **2019**, *64*, 1–9. [[CrossRef](#)]
51. Fanizzi, A.; Basile, T.M.A.; Losurdo, L.; Bellotti, R.; Tangaro, S.; La Forgia, D.; Didonna, V.; Massafra, R.; Tamborra, P.; Moschetta, M.; et al. Hough transform for microcalcification detection in digital mammograms. *Appl. Digit. Image Process. XL* **2017**, *10396*, 41. [[CrossRef](#)]
52. Losurdo, L.; Fanizzi, A.; Basile, T.M.A.; Bellotti, R.; Bottigli, U.; Dentamaro, R.; Didonna, V.; Fausto, A.; Massafra, R.; Monaco, A.; et al. A Combined Approach of Multiscale Texture Analysis and Interest Point/Corner Detectors for Microcalcifications Diagnosis. In *Bioinformatics and Biomedical Engineering; IWBBIO 2018; Lecture Notes in Computer Science*; Springer: Cham, Switzerland, 2018; pp. 302–313. [[CrossRef](#)]

-
53. La Forgia, D.; Fanizzi, A.; Campobasso, F.; Bellotti, R.; Didonna, V.; Lorusso, V.; Moschetta, M.; Massafra, R.; Tamborra, P.; Tangaro, S.; et al. Radiomic Analysis in Contrast-Enhanced Spectral Mammography for Predicting Breast Cancer Histological Outcome. *Diagnostics* **2020**, *10*, 708. [[CrossRef](#)] [[PubMed](#)]
 54. Yang, J.; Wang, T.; Yang, L.; Wang, Y.; Li, H.; Zhou, X.; Zhao, W.; Ren, J.; Li, X.; Tian, J.; et al. Preoperative prediction of axillary lymph node metastasis in breast cancer using mammography-based radiomics method. *Sci. Rep.* **2019**, *9*, 4429. [[CrossRef](#)]