

Article

Application of Machine Learning Algorithm in Predicting Axillary Lymph Node Metastasis from Breast Cancer on Preoperative Chest CT

Soyoung Park ^{1,†} , Jong Hee Kim ^{2,†} , Yoon Ki Cha ^{2,*,‡}, Myung Jin Chung ^{2,*,‡}, Jung Han Woo ² and Subin Park ¹

¹ Department of Health Sciences and Technology, SAIHST, Sungkyunkwan University, Seoul 06351, Republic of Korea; thyoung94@gmail.com (S.P.); subinn.park@gmail.com (S.P.)

² Department of Radiology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul 06351, Republic of Korea; jonghk1101@naver.com (J.H.K.); jhwoo9984@gmail.com (J.H.W.)

* Correspondence: yoonki.cha@samsung.com (Y.K.C.); mj1.chung@samsung.com (M.J.C.)

† These authors contributed equally to this work.

‡ These authors contributed equally to this work.

Abstract: Axillary lymph node (ALN) status is one of the most critical prognostic factors in patients with breast cancer. However, ALN evaluation with contrast-enhanced CT (CECT) has been challenging. Machine learning (ML) is known to show excellent performance in image recognition tasks. The purpose of our study was to evaluate the performance of the ML algorithm for predicting ALN metastasis by combining preoperative CECT features of both ALN and primary tumor. This was a retrospective single-institutional study of a total of 266 patients with breast cancer who underwent preoperative chest CECT. Random forest (RF), extreme gradient boosting (XGBoost), and neural network (NN) algorithms were used. Statistical analysis and recursive feature elimination (RFE) were adopted as feature selection for ML. The best ML-based ALN prediction model for breast cancer was NN with RFE, which achieved an AUROC of 0.76 ± 0.11 and an accuracy of 0.74 ± 0.12 . By comparing NN with RFE model performance with and without ALN features from CECT, NN with RFE model with ALN features showed better performance at all performance evaluations, which indicated the effect of ALN features. Through our study, we were able to demonstrate that the ML algorithm could effectively predict the final diagnosis of ALN metastases from CECT images of the primary tumor and ALN. This suggests that ML has the potential to differentiate between benign and malignant ALNs.

Keywords: breast cancer; machine learning; axilla; lymph nodes; lymphatic metastasis; computed tomography



Citation: Park, S.; Kim, J.H.; Cha, Y.K.; Chung, M.J.; Woo, J.H.; Park, S. Application of Machine Learning Algorithm in Predicting Axillary Lymph Node Metastasis from Breast Cancer on Preoperative Chest CT. *Diagnostics* **2023**, *13*, 2953. <https://doi.org/10.3390/diagnostics13182953>

Academic Editor: Sung Chul Lim

Received: 10 August 2023

Revised: 5 September 2023

Accepted: 13 September 2023

Published: 14 September 2023



Copyright: © 2023 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/>).

1. Introduction

Being the most common cancer in women worldwide, breast cancer poses a great challenge to public health on a global scale [1]. Axillary lymph node (ALN) status is one of the most critical prognostic factors for patients with breast cancer [2]. In fact, identification of the presence of ALN metastasis plays a pivotal role in pathologic staging, prognosis, and guidance of treatment in patients with breast cancer [3].

Although widespread adaptation of MRI and PET/CT in breast cancer staging has improved sensitivity and specificity in the assessment of lymph node metastasis compared with CT, the long scan time may increase patients' discomfort [4]. The greater number of contraindications associated with MRI and the higher cost of a PET/CT scan are also obstacles to wide usage compared with other diagnostic modalities [4]. More significantly, chest contrast-enhanced CT (CECT) not only displays enhanced breast cancer tissue and ALN metastasis, but also excludes possible intrapulmonary and thoracic bone metastases [5]. As a result, chest CECT has been widely used for preoperative evaluation of breast cancer [5–7].

However, according to previous studies using CT and MRI to differentiate benign from malignant lymph nodes, there have been limitations in assessing accurate size or signal intensity threshold [8]. Despite its limitations in differentiating benign and malignant lymph nodes, CT is still a widely used modality. There is a great need for non-invasive and accurate methods for predicting ALN metastasis in breast cancer.

In the era of artificial intelligence, current developments in radiology focus on the improvement of decision support systems to maximize the potential role of non-invasive imaging modalities [9]. In the previous study by Zhou et al., they developed a deep learning algorithm to predict ALN metastasis in patients with clinically negative ALN using the US images of primary breast cancers [10]. The best-performing CNN achieved 85% sensitivity and 73% specificity compared to the 73% sensitivity and 63% specificity of the radiologists [10]. Recently, Chen et al. developed a CNN model to predict sentinel lymph node status based on dynamic contrast-enhanced MRI [11]. Their CNN model showed the best performance in tumor groups smaller than 0.2cm with an AUC of 0.081 at the internal validation set and 0.823 at the external test set 1 [11].

Furthermore, there have been a few studies predicting ALN metastasis based on preoperative CT images [12]. In the previous study by Yang et al., they developed a deep learning model with only features of ALN itself for the prediction of ALN metastasis [12]. However, according to the most widely used monograms and scoring system for predicting ALN metastasis in clinical practice [13–19], primary tumor characteristics, such as tumor size, are also considered, as well as features of ALN [20]. Therefore, the aim of our study was to investigate the potential of the machine learning (ML) model for predicting ALN metastasis with chest CECT, using not only the features of ALN itself but also the primary tumor.

2. Materials and Methods

2.1. Data Collection

The institutional review board (IRB) of Samsung Medical Center approved this retrospective study. The requirement for patient consent to use clinical data was waived by the IRB due to a retrospective study design (IRB file number: 2021-08-031). Between May 2019 and December 2019, 829 patients with breast cancer who underwent preoperative chest CECT with both pre- and post-contrast-enhancement protocols were included in this study. Inclusion criteria were (1) pathologically confirmed breast cancer patients who underwent preoperative chest CECT and (2) those who had no history of previous neoadjuvant chemotherapy. Exclusion criteria were as follows: (1) No visible primary tumor on chest CECT; (2) no visible suspicious ALN on chest CECT; (3) male patients; and (4) recurrent breast cancer after breast surgery. Finally, a total of 266 patients were selected for this study. The flow chart of the selection of study subjects is shown in Figure 1.

2.2. Feature Extraction

We collected age of patients as a demographic feature. Imaging features of primary tumor and ALN were collected by two thoracic radiologists with 16 and 30 years of experience.

The location of the primary tumor (T_site; left or right) and the maximal diameter of the primary tumor (T_size) were evaluated. Pattern and degree of enhancement of the primary tumor were also evaluated. Since there is no generally used standard for breast imaging or classification with chest CECT, we evaluated enhancement pattern of the primary tumor based on 5th BI-RADS lexicon for MRI [21]. Although enhancement characteristics of masses are divided into six types on BI-RADS lexicon, we reduced them into three categories. Since MRI provides greater soft tissue visualization than CT, detailed enhancement patterns, such as rim enhancement and enhancing internal septations, might only appear as heterogeneous enhancements on CT [22]. Therefore, enhancement patterns were divided into the following three categories: T_homogeneous, T_heterogeneous, and T_non-mass.

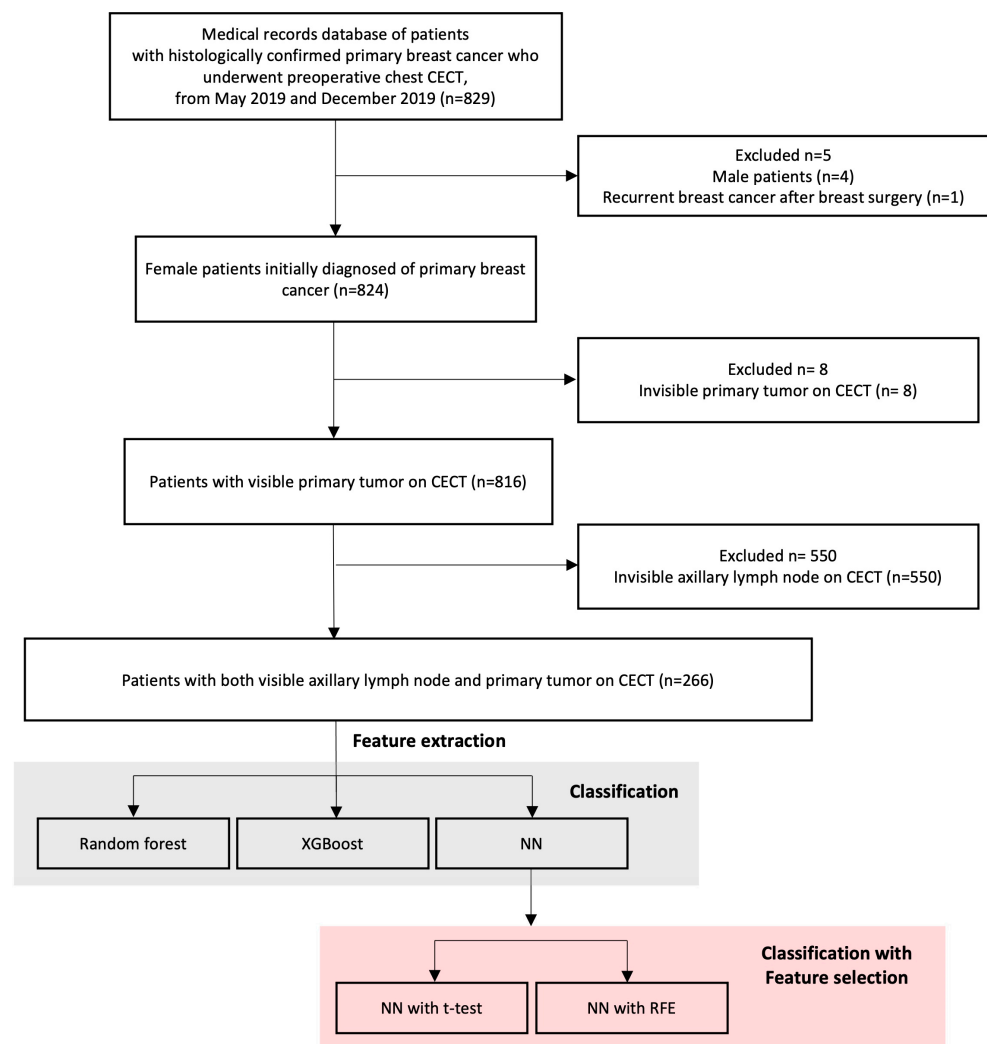


Figure 1. Flowchart of selecting study subjects for the development and evaluation of machine learning models for axillary lymph node metastasis prediction. NN: Neural network. RFE: recursive feature elimination.

T_homogeneous stands for uniform and confluent enhancement throughout the mass, which is same as BIRADS MRI lexicon [21]. T_heterogeneous is an integrated category of several BIRADS MRI lexicons, including heterogeneous enhancement, rim enhancement, and enhancing internal septations [21]. T_non-mass is defined as an enhancement that has neither a tri-dimensional mass nor typical mass characteristics [21].

Within T_homogeneous category, enhancement degree was also evaluated. It was categorized as follows: if the degree of contrast enhancement of primary tumor was similar to that of the most adjacent vessel, it was considered as a marked enhancement; if the enhancement degree of the tumor was higher than that of the ipsilateral latissimus dorsi or serratus anterior muscle, but lower than that of the vessel, it was considered a moderate degree; if the enhancement degree was similar to that of the ipsilateral latissimus dorsi or serratus anterior muscle, it was considered a mild enhancement; and if it was lower than that of the ipsilateral latissimus dorsi or serratus anterior muscle, it was considered a minimal enhancement [23]. To avoid subjective bias of reading features, two radiologists examined and marked a region of interest (ROI) to sign both primary tumor and ALN with the highest metastatic potential [23,24].

We extracted features by measuring Hounsfield unit (HU) value for the primary tumor, aorta, and ALN. T_average (average HU value of primary tumor) was measured as region of interest covering more than half of the tumor diameter. T_SD was the standard

deviation (SD) of HU values of the primary tumor. $A_average$ was the average HU value of attenuation of aorta. A_SD was SD of HU values of aorta. $N_average$ was the average HU value in the most suspected ALN metastasis in CECT. N_area was the approximate area of the most suspected ALN. Workflow of N_area measurement is described in Figure 2.

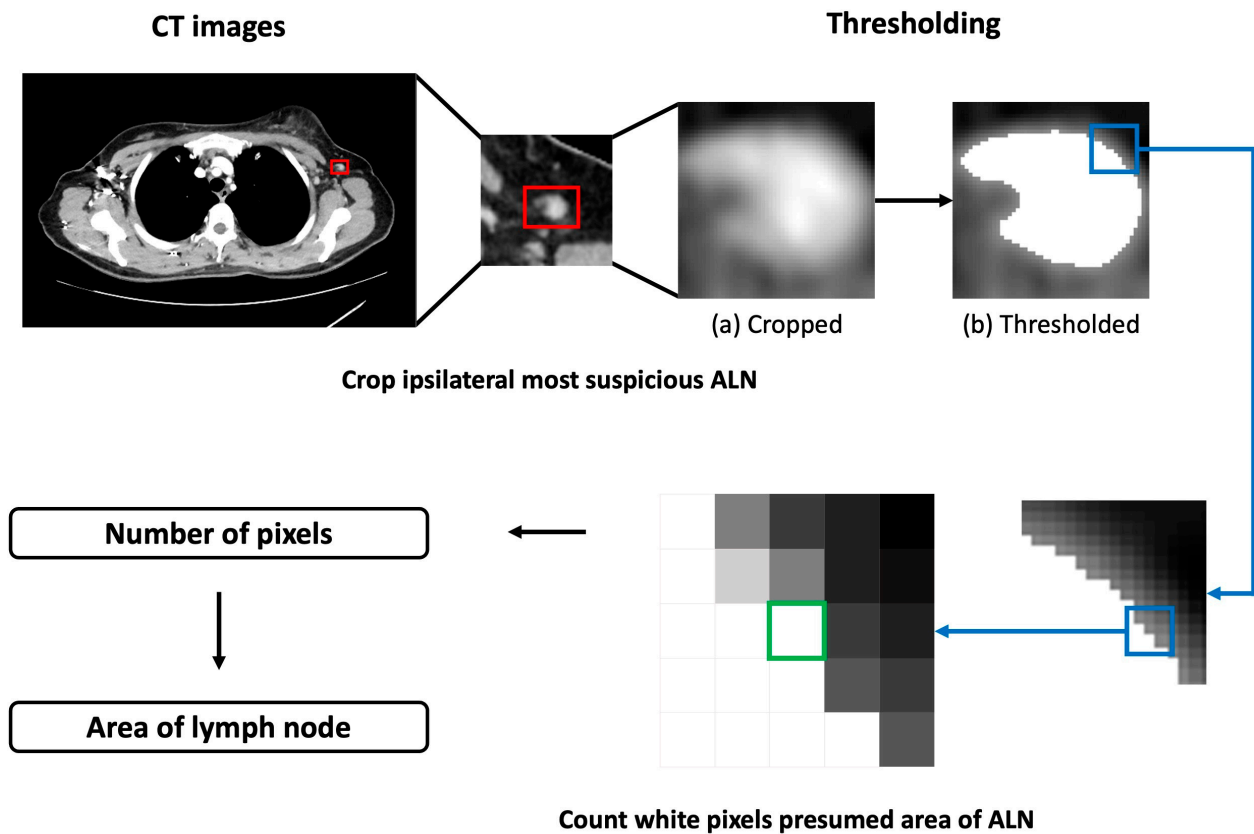


Figure 2. Flowchart of the measurement of the area of the lymph node suspected of metastasis. The red rectangle is the region of interest (ROI) of the most suspicious axillary lymph node (ALN). The green box is one of the white pixels.

We cropped the most suspicious ALN from isometric resampled CECT image. We then determined the threshold of the cropped image. Values above the *thresh* were set to be *maxValue*. Otherwise, values were unchanged, as shown in the following equation:

$$thresholded(x, y) = \begin{cases} maxValue & \text{if } cropped(x, y) > thresh \\ cropped(x, y) & \text{otherwise} \end{cases}$$

where $cropped(x, y)$ was the HU value for each pixel of cropped ALN image, $thresh$ was the threshold value, and $thresholded(x, y)$ was the thresholded image.

Next, the number of pixels in *maxValue* was counted. The value of *maxValue* used was white, which was (255, 255, 255) in RGB. Since we performed isometric resampling, the number of pixels could refer to the ALN area.

Lastly, since the degree of contrast enhancement might be affected by intrinsic factors, such as cardiac output and BMI, to eliminate the influence of intrinsic factors, we applied aorta-based corrected attenuation value [25]. Therefore, three variables, $\Delta(\text{Aorta} - \text{Tumor})$, $\Delta(\text{Aorta} - \text{Node})$, and $\Delta(\text{Tumor} - \text{Node})$ were generated by calculating each difference between three variables, $A_average$, $T_average$, and $N_average$.

Consequently, we collected a total of 15 variables (Age , T_site , T_size , $T_homogeneous$, $T_heterogeneous$, $T_non-mass$, $T_average$, T_SD , $N_average$, N_area , $A_average$, A_SD ,

$\Delta(\text{Aorta} - \text{Tumor})$, $\Delta(\text{Aorta} - \text{Node})$, and $\Delta(\text{Tumor} - \text{Node})$) for each patient. There was no missing value for any variable.

To analyze collected variables with ML, feature encoding methods were applied. For categorical variables in which there was no ordinal relationship, a one-hot encoding technique was used. For categorical features with ordinal relationships, they were encoded into zero and non-zero variables.

2.3. Feature Selection Methods

Feature selection is one of the most important processes that can improve the performance and efficiency of computation time. This is accomplished by selecting a subset of features that could derive an optimal performance of the model in the entire feature set while eliminating irrelevant, redundant, and noisy features that do not contribute to the model's performance [26,27]. If the number of data used for training is small, it might be difficult for the model to learn important relationships sufficiently [28]. Therefore, by using feature selection, we tried to facilitate the learning process with a limited amount of data while removing less important features [29].

We applied statistical methods and recursive feature elimination (RFE) to exclude the least important features and identify the meaningful feature subset used in the neural network (NN) model. Regarding statistical methods, a *t*-test was used for numerical variables and a chi-square test was used for categorical variables. These tests were performed using SPSS software for Windows ver. 27.0 (SPSS Inc., Chicago, IL, USA). For both statistical analyses, features with *p*-values less than 0.05 were regarded as statistically different between metastasis and non-metastasis.

We gathered only statistically significant variables and created one subset. The RFE could eliminate the least meaningful features iteratively until the area under the curve (AUC) reached the maximum value [26,30,31]. In the iterative process in RFE, one can gain the best feature subset that provides the highest AUC. RFE was performed using Python software (version 3.7.11).

2.4. Machine Learning Algorithms

Three ML algorithms were used to compare performances of models for classifying whether there was any breast cancer metastasis to ALN: random forest (RF) [32], extreme gradient boosting (XGBoost) [33], and NN. RF is an ensemble learning algorithm for classification, regression, and other tasks with a combination of decision trees that calculate the mode or mean/average of each tree. XGBoost is also an ensemble algorithm that builds multiple gradients boosted decision trees to maximize efficiency and performance [33]. NN algorithm mimics the operation of the human brain [34]. It is constructed with an input layer, an output layer, and singular or multiple hidden layers placed between two of them.

We tuned hyperparameters of NN via grid search at each RFE step to achieve better performance [35,36]; so, a learning rate of 0.001 for Adam optimizer [37], batch size of 10, and epochs of 300 were used. In the same way, hyperparameters of RF and XGBoost were also tuned. These ML-driven predictions of breast cancer metastasis to ALN were developed using TensorFlow (version 2.3.0), Keras (version 2.4.0), Scikit-learn (version 0.24.2) libraries, and Python software (version 3.7.11).

For more advanced experiments, we applied two feature selection algorithms, statistical analysis and RFE, to NN. The final workflow for predicting ALN metastasis is shown in Figure 3.

2.5. Evaluation of Machine Learning Models

To develop ML algorithms, the total dataset was split into a training set and a test set. The training set was used to teach breast cancer metastasis to ALN prediction algorithms, and the test set was used to evaluate trained algorithms. At the training phase, we used stratified 10-fold cross-validation. This randomly divided the dataset into 10 partitions (folds) while keeping the ratio of each label (e.g., positive and negative) distribution for each

fold. We then repeatedly validated and evaluated the model using 9 partitions as training data and the remaining 1 partition as validation data. These techniques assured a model's generalized performance and prevented overfitting [38]. Because outputs of ML-driven algorithms were probabilistic estimates of breast cancer metastasis in ALN, the performance for each model was evaluated using area under the receiver operating characteristic curve (AUROC) and confusion matrix, comparing them with histopathological examination results of ALN.

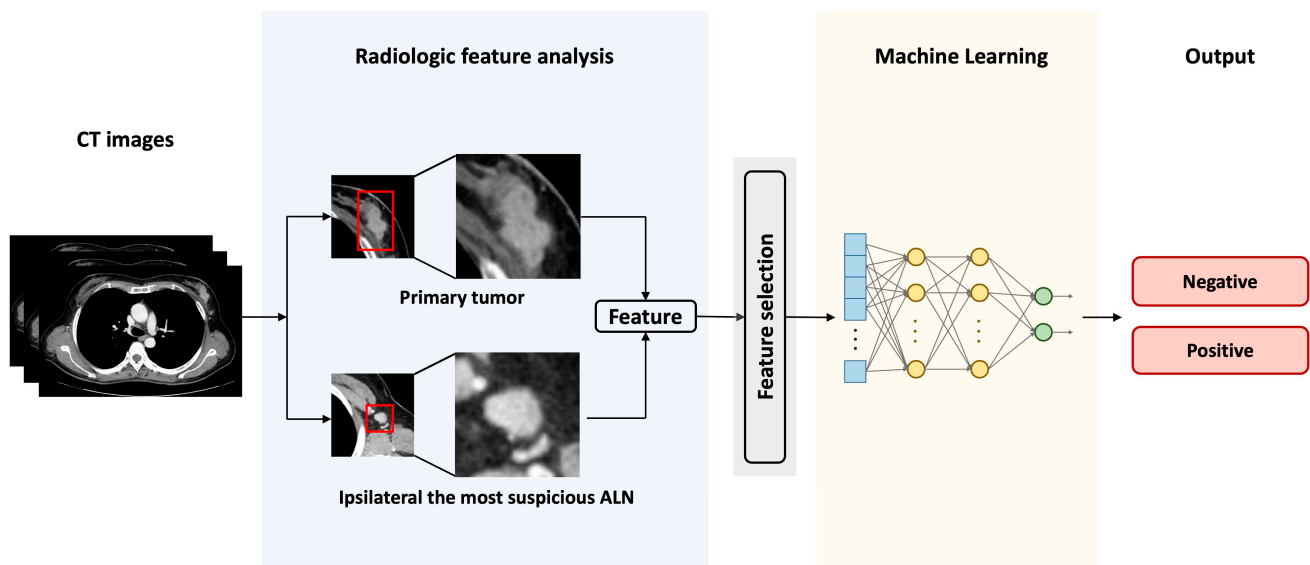


Figure 3. Schematic illustration of the machine learning based prediction framework for axillary lymph node (ALN) metastasis of breast cancer. The red rectangle is the region of interest (ROI) of the primary tumor and the most suspicious axillary lymph node (ALN).

3. Results

3.1. Data Collection and Feature Engineering

All 266 patients were females with a median age of 50 years (range, 23–81 years). Among the 266 patients with breast cancer and most suspicious ALN metastasis, 186 patients had ALN metastasis while 80 patients did not have ALN metastasis. Table 1 shows demographics and characteristics of the collected dataset.

T_{site} did not show significant differences between metastatic and benign ALN groups. ($p = 0.593$). T_{size} and N_{area} were significantly associated with ALN metastasis ($p < 0.05$). Regarding the enhancement pattern of the primary tumor, 234 patients had T_{homogeneous} features (87.97%), 8 patients had T_{heterogeneous} features (3.01%), and 24 patients had T_{non-mass} features (9.02%). In the T_{homogeneous} group, 3 patients had a minimal enhancement, 33 patients had a mild enhancement, 115 patients had a moderate enhancement, and 83 patients had a marked enhancement. Average and SD of HU values were 85.96 ± 22.02 (T_{average}) and 11.71 ± 5.84 (T_{SD}) for the primary tumor and 174.39 ± 21.46 (A_{average}) and 8.65 ± 2.32 (A_{SD}) for the aorta of all patients. None of these four features was associated with ALN metastasis. On the other hand, among the two features extracted from ALN, N_{area} (884.01 ± 282.24 , $p < 0.05$) was statistically significant. Three features that presented a difference in enhancement degree did not show meaningful characteristics in statistical analysis.

Table 1. Demographics and characteristics of the collected dataset.

	All Patients	ALN Metastasis (–)	ALN Metastasis (+)	p-Value
Age (years)	50 (23–81)	49.5 (23–77)	50 (32–81)	–
Primary tumor				
T_site (left/right)	133:133	38:42	95:91	–
T_size (mm)	18.72 ± 10.03	21.27 ± 13.54	17.62 ± 7.85	<0.05
Enhancement pattern				
T_homogeneous ^a (E ₁ :E ₂ :E ₃ :E ₄)	234 (87.97%) (3:33:115:83)	65 (81.25%) (1:13:30:21)	169 (90.86%) (2:20:85:62)	<0.05
T_heterogeneous	8 (3.01%)	3 (3.75%)	5 (2.69%)	–
T_non-mass	24 (9.02%)	12 (15.00%)	12 (6.45%)	<0.05
Enhancement degree (HU)				
T_average	85.96 ± 22.02	88.78 ± 24.64	84.75 ± 20.74	–
T_SD	11.71 ± 5.84	11.26 ± 5.79	11.90 ± 5.87	–
Lymph node				
N_average	81.44 ± 21.31	83.68 ± 21.46	80.48 ± 21.23	–
N_area	884.01 ± 282.24	940.35 ± 243.31	859.78 ± 294.73	<0.05
Aorta				
A_average	174.39 ± 21.46	173.17 ± 19.24	174.92 ± 22.38	–
A_SD	8.65 ± 2.32	8.33 ± 2.10	8.78 ± 2.40	–
Calibrated enhancement degree^b				
Δ(Aorta – Tumor)	88.43 ± 26.34	84.39 ± 27.11	90.17 ± 25.89	–
Δ(Aorta – Node)	92.95 ± 26.13	89.49 ± 22.09	94.44 ± 27.60	–
Δ(Tumor – Node)	4.52 ± 27.16	5.11 ± 28.87	4.27 ± 26.48	–

Values are presented as number of patients, median (range), number of patients per category separated by colons, or mean value ± standard deviation. ALN: axillary lymph node; HU: Hounsfield Unit. ^a Categorized by the degree of enhancement of primary tumor. (E₁) Minimal enhancement. (E₂) Mild enhancement. (E₃) Moderate enhancement. (E₄) Marked enhancement. ^b Enhancement degree corrected through the elimination of intrinsic factors, such as BMI and cardiac output.

3.2. Performance Comparisons of ML Models for Detecting ALN Metastasis

Among three classification models which trained all features, NN recorded higher AUROC (0.71 ± 0.13) than RF (0.63 ± 0.11) and XGBoost (0.62 ± 0.08). NN with statistical analysis showed a lower AUROC (0.63 ± 0.12) performance than NN but similar to RF and XGBoost. The NN with RFE reached the highest AUROC of 0.76 ± 0.11 and accuracy of 0.74 ± 0.12 when compared to three classification models without feature selection and to another NN model with feature selection (NN with statistical analysis). Table 2 summarized results of five experiments for classifying ALN metastasis.

Table 2. Performance comparison of machine learning models for predicting ALN metastasis.

	Classification			Classification with Feature Selection	
	RF	XGBoost	NN	NN with Stats ^a	NN with RFE [*]
AUROC	0.63 ± 0.11 (0.47–0.84)	0.62 ± 0.08 (0.51–0.74)	0.71 ± 0.13 (0.49–0.84)	0.63 ± 0.12 (0.49–0.87)	0.76 ± 0.11 (0.61–0.97)
Accuracy	0.70 ± 0.07 (0.56–0.81)	0.66 ± 0.06 (0.56–0.74)	0.72 ± 0.12 (0.44–0.85)	0.61 ± 0.13 (0.44–0.85)	0.74 ± 0.12 (0.56–0.92)
PPV	0.72 ± 0.04	0.73 ± 0.04	0.90 ± 0.09	0.92 ± 0.07	0.93 ± 0.06
NPV	0.48 ± 0.35	0.45 ± 0.23	0.61 ± 0.20	0.45 ± 0.11	0.59 ± 0.19
Sensitivity	0.91 ± 0.08	0.82 ± 0.08	0.71 ± 0.24	0.49 ± 0.21	0.69 ± 0.29
Specificity	0.18 ± 0.13	0.28 ± 0.13	0.75 ± 0.24	0.88 ± 0.12	0.88 ± 0.12

Values are mean ± standard deviation of 10-fold cross-validation results and values in the bracket are minimum–maximum values of 10-fold cross-validation results. RF: random forest; NN: neural network; RFE: recursive feature elimination; AUROC: area under the receiver operating characteristic curve; PPV: positive predictive value; NPV: negative predictive value. ^a NN with statistical analysis. ^{*} RFE with lymph features (same model as used in Table 3).

Table 3. Classification performances with or without lymph node features for predicting ALN metastasis.

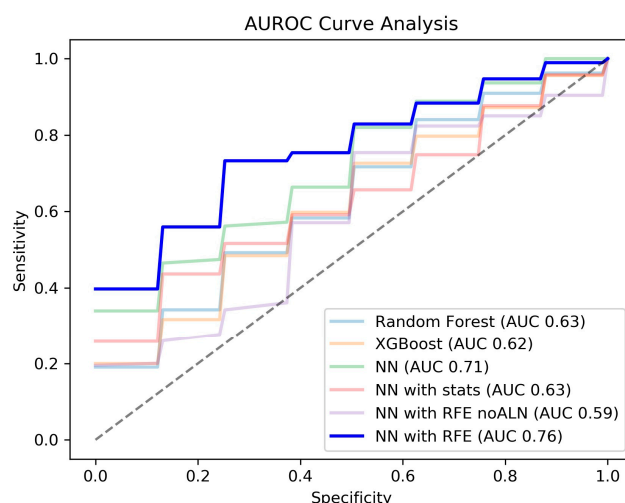
	RFE without Lymph Features	RFE with Lymph Features *
AUROC	0.59 ± 0.17 (0.34–0.96)	0.76 ± 0.11 (0.61–0.97)
Accuracy	0.66 ± 0.18 (0.37–0.88)	0.74 ± 0.12 (0.56–0.92)
PPV	0.88 ± 0.11	0.93 ± 0.06
NPV	0.53 ± 0.17	0.59 ± 0.19
Sensitivity	0.63 ± 0.32	0.69 ± 0.29
Specificity	0.71 ± 0.25	0.88 ± 0.12

Values are mean $\pm$ standard deviation of 10-fold cross-validation results and values in the bracket are minimum–maximum values of 10-fold cross-validation results. AUROC: area under the receiver operating characteristic curve; PPV: positive predictive value; NPV: negative predictive value. * NN with RFE (same model as used in Table 2).

All 15 features were inserted into RF, XGBoost, and NN. For classification with feature selection, four features (T_size, T_homogeneous, T_non-mass, and N_area) selected via statistical feature selection were used for NN with stats, and eight features (T_size, T_average, N_average, N_area, A_average, Δ (Aorta – Tumor), Δ (Aorta – Node), and Δ (Tumor – Node)) selected via RFE were used for NN with RFE.

To validate the importance of ALN-related features, we also tried the NN with RFE model without ALN-related features. The results of the performance comparison of NN with ALN features and NN without ALN features are shown in Table 3. Since there were four ALN-related features (N_average, N_area, Δ (Aorta – Node), and Δ (Tumor – Node)), the number of input variables for RFE without lymph features was 11.

During RFE without lymph features experiment, the best performance was shown when there were four input features (T_size, T_average, A_average, and Δ (Aorta – Tumor)). Compared with RFE with lymph features, all evaluation variables of RFE with lymph features showed higher performances. Especially, AUROC and Specificity increased the most from 0.59 ± 0.17 to 0.76 ± 0.11 and 0.71 ± 0.25 to 0.88 ± 0.12 . Figure 4 shows the AUROC curves of all six tested ML algorithms.

**Figure 4.** Receiver operating characteristic (AUROC) curve of the six tested algorithms.

4. Discussion

Despite its clinical importance, the prediction of lymph node metastasis is still challenging in the field of imaging. As shown in our study, radiologic features between benign and malignant ALN evaluated by two radiologists showed ambiguous results. There was no significant difference in the average enhancement degree between benign and malignant

lymph nodes. Furthermore, the sizes of primary tumor and lymph nodes were even larger in benign ALN groups. This finding might be partly due to a selection bias caused by our radiologic assessing method, which evaluates the largest suspicious lymph nodes, even for benign ALN cases.

Our ambiguous result regarding the size and enhancement degree between benign and malignant ALN is well correlated with the previous studies on limitations of CT and MRI in the evaluation of metastatic lymph nodes [8]. Although previous studies which differentiated metastatic lymph nodes based on size using CT had controversial results, they generally showed low accuracy [8,39,40]. Furthermore, several studies have investigated the utility of dynamic contrast enhancement for differentiating normal from metastatic lymph nodes [8,41–43]. Although they observed a significantly longer time to peak, lower peak enhancement, and slower washout slope for malignant lymph nodes, they also indicated the difficulty in standardizing acquisition parameters to obtain reproducible data [8]. This might be due to the fact that conventional human vision evaluation is dependent on the knowledge and familiarity of the radiologists.

Various ML approaches have been proposed for classifying or detecting breast cancer metastasis in ALN [44–46]. Song et al. [45] proposed a ML-based radiomics model for predicting ALN metastasis in invasive ductal breast cancer (IDC). The volume of interests (VOIs) were drawn in the primary tumor on the PET scan, and texture features were extracted in 100 consecutive IDC patients who underwent surgical resection of the primary tumor and/or ALN dissection. They used XGBoost to select features and evaluate the model. Zheng et al. [44] developed a deep learning radiomics (DLR) model that classified the ALN extent in early stage breast cancer. They used the US, shear wave elastography (SWE) images, and clinicopathological data of 584 patients. Their combined DLR model offers a non-invasive imaging biomarker which can predict the status of ALN for patients with early stage breast cancer. Zhang et al. [46] developed a multiparametric MRI model combined with ensemble learning to predict ALN metastasis in invasive breast cancer. They fused axial T2WI, DWI, and DCE-MRI models and achieved an AUC of 0.913.

In this study, we trained ML models to predict the ALN metastasis of breast cancer patients using clinical information and imaging features of both the primary tumor and ALN on preoperative chest CECT images. Additionally, in order to avoid overfitting, when we trained models directly from images using a small amount of data, we extracted handcrafted knowledge-based features and adapted them for training.

The best performance was achieved by NN with RFE, and it reached an AUROC of 0.72 ± 0.09 and an accuracy of 0.64 ± 0.15 . Interestingly, our experiments showed that the RFE provides a better feature subset to improve the performance of NN than statistical analysis with our dataset. Statistical methods, such as t-test and chi-square, could be used to compare means and analyze if there is any probable difference between two groups. Since features are analyzed independently, feature dependencies could be ignored, leading to a poor feature subset [47]. On the other hand, since RFE uses sequential elimination, it interacts with classifiers with the advantage of capturing feature dependencies [48].

To understand the importance of ALN-related features for predicting ALN metastasis, we conducted a comparative experiment. As a result of comparing the performance of the NN with the RFE model along with or without ALN-related features from chest CECT, all performance measurements increased, showing that analyzing both primary tumor and ALN has a positive effect on predicting ALN metastasis. Particularly, specificity increased from 0.71 ± 0.25 to 0.88 ± 0.12 . Due to the increase in specificity via the addition of CECT features of ALNs and primary tumors, unnecessary treatment processes should be avoidable, and procedure-related side effects may also be prevented [49,50].

Our work has notable strengths compared with other previous studies predicting breast cancer metastasis using ML models. To the best of our knowledge, most of the data types or methods used in previous studies could be divided into two types: (1) multiple image modalities, such as ultrasonography, X-ray, CT, or MRI, for extracting features in a single area (either primary tumor, lymph node, or other organs); and (2) multiple

image modalities, such as ultrasonography, X-ray, CT, or MRI, for extracting features in multiple areas (one image modality for each area) [12,44,51,52]. However, unlike previous studies, we attempted to predict metastasis by considering both the primary tumor and ALN in chest CECT via the extraction of features from two areas using a single image modality [12]. Compared to multimodal image, single modal image research has fewer training parameters, so it benefits in computational cost and is less limited in building layers and channels [53].

This study has several limitations. Firstly, we used the experimental dataset from a single institution to train and evaluate the ML model. It is often difficult to gather a large-scale medical image dataset due to issues such as annotation cost or privacy [54,55]. However, since chest CECT images may differ depending on the scanner, performance may deteriorate if external validation is performed. Moreover, two radiologists went through the additional process of labeling the pattern and the degree of enhancement of the primary tumor while conducting this study, which means it required a lot of annotation cost to use unstructured readings for training. Furthermore, pathologic and imaging correlations were not performed for ALNs one by one. In cases of multiple ALNs on the ipsilateral side of the primary tumor, we only evaluated the most suspicious ALN. Therefore, collected imaging features of lymph node might show ambiguous results regarding lymph node size and enhancement. Finally, several cases with non-visible primary tumor or lymph nodes on CT were excluded. Although primary tumor and lymph nodes are not delineated on CT, such cases might also have pathologically confirmed lymph node metastasis. Characteristic findings of those cases were not evaluated in this study.

In further studies with expanded data collections, security and privacy issues should be considered for a more widespread application in the clinical setting [56]. Also, to increase the utilization of additionally collected data, it is necessary to go through the process of generating structured readings [57] by finding important keywords that are used repeatedly in the readings of various radiologists [57,58].

5. Conclusions

In conclusion, we adopted ML methods to classify ALN metastasis of breast cancer patients using clinical information and chest CECT image features. Although we used a limited amount of data, we were able to demonstrate that our ML models could classify ALN metastasis of breast cancer by using appropriate feature selection methods considering both primary tumor and ALN features. Imaging features of both ALNs and primary tumors from chest CECT play an important role in ML models for the ALN metastasis of breast cancer. With further validation in a larger population and model calibration, our ML technique might serve as an important decision support tool in clinical applications when predicting ALN metastasis in breast cancer.

Author Contributions: S.P. (Soyoung Park); Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Visualization, Writing—original draft, Writing—review and editing, J.H.K.; Formal analysis, Methodology, Validation, Visualization, Writing—original draft, Writing—review and editing, Y.K.C.; Project administration, Investigation, Supervision, Validation, Writing—review and editing, M.J.C.; Project administration, Resources, Supervision, Validation, Writing—review and editing, J.H.W.; Writing—review and editing, S.P. (Subin Park); Resources, Software. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health Welfare, Republic of Korea (grant number: HR21C0885), Future Medicine 20*30 Project of the Samsung Medical Center (SMC1210791).

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board (or Ethics Committee) of the Samsung Medical Center, Sungkyunkwan University School of Medicine (IRB No. SMC 2021-08-031).

Informed Consent Statement: Patient consent was waived due to the retrospective nature of this study and the anonymization of patient data.

Data Availability Statement: The datasets generated during this study are not publicly available due to privacy and ethical considerations; however, anonymized data can be provided by the corresponding author upon reasonable request and with the approval of the ethics committee. Researchers interested in accessing the data should contact the corresponding author for further information.

Conflicts of Interest: The authors declare no conflict of interest, financial or otherwise, in relation to this study or its publication.

References

1. Siegel, R.L.; Miller, K.D.; Fuchs, H.E.; Jemal, A. Cancer statistics, 2022. *CA A Cancer J. Clin.* **2022**, *72*, 7–33. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Quiet, C.A.; Ferguson, D.J.; Weichselbaum, R.R.; Hellman, S. Natural history of node-negative breast cancer: A study of 826 patients with long-term follow-up. *J. Clin. Oncol.* **1995**, *13*, 1144–1151. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Tafreshi, N.K.; Kumar, V.; Morse, D.L.; Gatenby, R.A. Molecular and Functional Imaging of Breast Cancer. *Cancer Control* **2010**, *17*, 143–155. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Yang, C.; Dong, J.; Liu, Z.; Guo, Q.; Nie, Y.; Huang, D.; Qin, N.; Shu, J. Prediction of Metastasis in the Axillary Lymph Nodes of Patients With Breast Cancer: A Radiomics Method Based on Contrast-Enhanced Computed Tomography. *Front. Oncol.* **2021**, *11*, 726240. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Ashiba, H.; Nakayama, R. Computerized evaluation scheme to detect metastasis in sentinel lymph nodes using contrast-enhanced computed tomography before breast cancer surgery. *Radiol. Phys. Technol.* **2019**, *12*, 55–60. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Urata, M.; Kijima, Y.; Hirata, M.; Shinden, Y.; Arima, H.; Nakajo, A.; Koriyama, C.; Arigami, T.; Uenosono, Y.; Okumura, H.; et al. Computed tomography Hounsfield units can predict breast cancer metastasis to axillary lymph nodes. *BMC Cancer* **2014**, *14*, 730. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Li, J.; Zhang, Y.; Zhang, W.; Gao, Y.; Jia, S.; Guo, J. Contrast enhanced computed tomography is indicative for angiogenesis pattern and display prognostic significance in breast cancer. *BMC Cancer* **2014**, *14*, 672. [\[CrossRef\]](#)
8. Torabi, M.; Aquino, S.L.; Harisinghani, M.G. Current concepts in lymph node imaging. *J. Nucl. Med.* **2004**, *45*, 1509–1518.
9. Samiei, S.; Granzier, R.W.Y.; Ibrahim, A.; Primakov, S.; Lobbes, M.B.L.; Beets-Tan, R.G.H.; van Nijnatten, T.J.A.; Engelen, S.M.E.; Woodruff, H.C.; Smidt, M.L. Dedicated Axillary MRI-Based Radiomics Analysis for the Prediction of Axillary Lymph Node Metastasis in Breast Cancer. *Cancers* **2021**, *13*, 757. [\[CrossRef\]](#)
10. Zhou, L.-Q.; Wu, X.-L.; Huang, S.-Y.; Wu, G.-G.; Ye, H.-R.; Wei, Q.; Bao, L.-Y.; Deng, Y.-B.; Li, X.-R.; Cui, X.-W.; et al. Lymph Node Metastasis Prediction from Primary Breast Cancer US Images Using Deep Learning. *Radiology* **2019**, *294*, 19–28. [\[CrossRef\]](#)
11. Chen, M.; Kong, C.; Lin, G.; Chen, W.; Guo, X.; Chen, Y.; Cheng, X.; Chen, M.; Shi, C.; Xu, M.; et al. Development and validation of convolutional neural network-based model to predict the risk of sentinel or non-sentinel lymph node metastasis in patients with breast cancer: A machine learning study. *eClinicalMedicine* **2023**, *63*, 102176. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Yang, X.; Wu, L.; Ye, W.; Zhao, K.; Wang, Y.; Liu, W.; Li, J.; Li, H.; Liu, Z.; Liang, C. Deep Learning Signature Based on Staging CT for Preoperative Prediction of Sentinel Lymph Node Metastasis in Breast Cancer. *Acad. Radiol.* **2020**, *27*, 1226–1233. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Barranger, E.; Coutant, C.; Flahault, A.; Delpech, Y.; Darai, E.; Uzan, S. An axilla scoring system to predict non-sentinel lymph node status in breast cancer patients with sentinel lymph node involvement. *Breast Cancer Res. Treat.* **2005**, *91*, 113–119. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Degnim, A.C.; Reynolds, C.; Pantvaidya, G.; Zakaria, S.; Hoskin, T.; Barnes, S.; Roberts, M.V.; Lucas, P.C.; Oh, K.; Koker, M.; et al. Nonsentinel node metastasis in breast cancer patients: Assessment of an existing and a new predictive nomogram. *Am. J. Surg.* **2005**, *190*, 543–550. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Hwang, R.F.; Krishnamurthy, S.; Hunt, K.K.; Mirza, N.; Ames, F.C.; Feig, B.; Kuerer, H.M.; Singletary, S.E.; Babiera, G.; Meric, F.; et al. Clinicopathologic factors predicting involvement of nonsentinel axillary nodes in women with breast cancer. *Ann. Surg. Oncol.* **2003**, *10*, 248–254. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Kohrt, H.E.; Olshen, R.A.; Bermas, H.R.; Goodson, W.H.; Wood, D.J.; Henry, S.; Rouse, R.V.; Bailey, L.; Philben, V.J.; Dirbas, F.M.; et al. New models and online calculator for predicting non-sentinel lymph node status in sentinel lymph node positive breast cancer patients. *BMC Cancer* **2008**, *8*, 66. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Pal, A.; Provenzano, E.; Duffy, S.W.; Pinder, S.E.; Purushotham, A.D. A model for predicting non-sentinel lymph node metastatic disease when the sentinel lymph node is positive. *Br. J. Surg.* **2008**, *95*, 302–309. [\[CrossRef\]](#)
18. Saidi, R.F.; Dudrick, P.S.; Remine, S.G.; Mittal, V.K. Nonsentinel lymph node status after positive sentinel lymph node biopsy in early breast cancer. *Am. Surg.* **2004**, *70*, 101–105; discussion 105, discussion 105. [\[CrossRef\]](#)
19. Van Zee, K.J.; Manasseh, D.M.; Bevilacqua, J.L.; Boolbol, S.K.; Fey, J.V.; Tan, L.K.; Borgen, P.I.; Cody, H.S., 3rd; Kattan, M.W. A nomogram for predicting the likelihood of additional nodal metastases in breast cancer patients with a positive sentinel node biopsy. *Ann. Surg. Oncol.* **2003**, *10*, 1140–1151. [\[CrossRef\]](#)

20. Chang, J.M.; Leung, J.W.T.; Moy, L.; Ha, S.M.; Moon, W.K. Axillary Nodal Evaluation in Breast Cancer: State of the Art. *Radiology* **2020**, *295*, 500–515. [\[CrossRef\]](#)
21. Sickles, E.A.; D'Orsi, C.J.; Bassett, L.W.; Appleton, C.M.; Berg, W.A.; Burnside, E.S. *ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System*; American College of Radiology: Reston, VA, USA, 2013; pp. 39–48.
22. Muller, N.L. Computed tomography and magnetic resonance imaging: Past, present and future. *Eur. Respir. J.* **2002**, *19*, 3S–12S. [\[CrossRef\]](#)
23. Choi, Y.J.; Kim, T.H.; Cha, Y.J.; Son, E.J.; Gweon, H.M.; Park, C.H. Incidental Breast Lesions on Chest CT: Clinical Significance and Differential Features Requiring Referral. *J. Korean Soc. Radiol.* **2018**, *79*, 303–310. [\[CrossRef\]](#)
24. Park, J.; Kim, C.; Cha, Y.K.; Chung, M.J. Missed and Detected Incidental Breast Cancers on Contrast Enhanced Chest CT: Detection Rates and CT Features. *Diagnostics* **2023**, *13*, 1522. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Ruppert-Kohlmayer, A.J.; Uggowitz, M.; Meissnitzer, T.; Ruppert, G. Differentiation of Renal Clear Cell Carcinoma and Renal Papillary Carcinoma Using Quantitative CT Enhancement Parameters. *Am. J. Roentgenol.* **2004**, *183*, 1387–1391. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Cho, B.H.; Yu, H.; Kim, K.-W.; Kim, T.H.; Kim, I.Y.; Kim, S.I. Application of irregular and unbalanced data to predict diabetic nephropathy using visualization and feature selection methods. *Artif. Intell. Med.* **2008**, *42*, 37–53. [\[CrossRef\]](#)
27. Vasantha, M.S.; Bharathi, D.V.S.; Dhamodharan, R.K. Medical Image Feature, Extraction, Selection and Classification. *Int. J. Eng. Sci. Technol.* **2010**, *2*, 2071–2076.
28. Zhang, Y.; Ling, C. A strategy to apply machine learning to small datasets in materials science. *Npj Comput. Mater.* **2018**, *4*, 25. [\[CrossRef\]](#)
29. Zhang, X.; Yu, L.; Yin, H.; Lai, K.K. Integrating data augmentation and hybrid feature selection for small sample credit risk assessment with high dimensionality. *Comput. Oper. Res.* **2022**, *146*, 105937. [\[CrossRef\]](#)
30. Cho, B.H.; Yu, H.; Lee, J.; Chee, Y.J.; Kim, I.Y.; Kim, S.I. Nonlinear support vector machine visualization for risk factor analysis using nomograms and localized radial basis function kernels. *IEEE Trans. Inf. Technol. Biomed.* **2008**, *12*, 247–256. [\[CrossRef\]](#)
31. Cho, B.H.; Lee, D.Y.; Park, K.A.; Oh, S.Y.; Moon, J.H.; Lee, G.I.; Noh, H.; Chung, J.K.; Kang, M.C.; Chung, M.J. Computer-aided recognition of myopic tilted optic disc using deep learning algorithms in fundus photography. *BMC Ophthalmol.* **2020**, *20*, 407. [\[CrossRef\]](#)
32. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
33. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
34. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Larochelle, H.; Erhan, D.; Courville, A.; Bergstra, J.; Bengio, Y. An empirical evaluation of deep architectures on problems with many factors of variation. In Proceedings of the 24th International Conference on Machine Learning—ICML'07, Corvallis, OR, USA, 20–24 June 2007.
36. Yang, L.; Shami, A. On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing* **2020**, *415*, 295–316. [\[CrossRef\]](#)
37. Diederik, Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2017**, arXiv:1412.6980.
38. Santos, M.S.; Soares, J.P.; Abreu, P.H.; Araujo, H.; Santos, J. Cross-Validation for Imbalanced Datasets: Avoiding Overoptimistic and Overfitting Approaches [Research Frontier]. *IEEE Comput. Intell. Mag.* **2018**, *13*, 59–76. [\[CrossRef\]](#)
39. Kau, R.J.; Alexiou, C.; Laubenbacher, C.; Werner, M.; Schwaiger, M.; Arnold, W. Lymph node detection of head and neck squamous cell carcinomas by positron emission tomography with fluorodeoxyglucose F 18 in a routine clinical setting. *Arch. Otorhinolaryngol.-Head Neck Surg.* **1999**, *125*, 1322–1328. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Dwamena, B.A.; Sonnad, S.S.; Angobaldo, J.O.; Wahl, R.L. Metastases from non-small cell lung cancer: Mediastinal staging in the 1990s—Meta-analytic comparison of PET and CT. *Radiology* **1999**, *213*, 530–536. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Fischbein, N.J.; Noworolski, S.M.; Henry, R.G.; Kaplan, M.J.; Dillon, W.P.; Nelson, S.J. Assessment of metastatic cervical adenopathy using dynamic contrast-enhanced MR imaging. *AJNR Am. J. Neuroradiol.* **2003**, *24*, 301–311.
42. Laissy, J.P.; Gay-Depassier, P.; Soyer, P.; Dombret, M.C.; Murciano, G.; Sautet, A.; Aubier, M.; Menu, Y. Enlarged mediastinal lymph nodes in bronchogenic carcinoma: Assessment with dynamic contrast-enhanced MR imaging. Work in progress. *Radiology* **1994**, *191*, 263–267. [\[CrossRef\]](#)
43. Szabó, B.K.; Aspelin, P.; Kristoffersen Wiberg, M.; Tot, T.; Boné, B. Invasive breast cancer: Correlation of dynamic MR features with prognostic factors. *Eur. Radiol.* **2003**, *13*, 2425–2435. [\[CrossRef\]](#)
44. Zheng, X.; Yao, Z.; Huang, Y.; Yu, Y.; Wang, Y.; Liu, Y.; Mao, R.; Li, F.; Xiao, Y.; Wang, Y.; et al. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. *Nat. Commun.* **2020**, *11*, 1236. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Song, B.-I. A machine learning-based radiomics model for the prediction of axillary lymph-node metastasis in breast cancer. *Breast Cancer* **2021**, *28*, 664–671. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Zhang, X.; Liu, M.; Ren, W.; Sun, J.; Wang, K.; Xi, X.; Zhang, G. Predicting of axillary lymph node metastasis in invasive breast cancer using multiparametric MRI dataset based on CNN model. *Front. Oncol.* **2022**, *12*, 1069733. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Piao, Y.; Piao, M.; Park, K.; Ryu, K.H. An ensemble correlation-based gene selection algorithm for cancer classification with gene expression data. *Bioinformatics* **2012**, *28*, 3306–3315. [\[CrossRef\]](#) [\[PubMed\]](#)

48. Bolón-Canedo, V.; Sánchez-Marño, N.; Alonso-Betanzos, A. A review of feature selection methods on synthetic data. *Knowl. Inf. Syst.* **2013**, *34*, 483–519. [[CrossRef](#)]
49. O'Toole, J.; Miller, C.L.; Specht, M.C.; Skolny, M.N.; Jammallo, L.S.; Horick, N.; Elliott, K.; Niemierko, A.; Taghian, A.G. Cording following treatment for breast cancer. *Breast Cancer Res. Treat.* **2013**, *140*, 105–111. [[CrossRef](#)] [[PubMed](#)]
50. Nieves Maldonado, S.M.; Pubul Núñez, V.; Argibay Vázquez, S.; Macías Cortiñas, M.; Ruibal Morell, Á. Axillary web syndrome following sentinel node biopsy for breast cancer. *Rev. Esp. Med. Nucl. Imagen Mol. (Engl. Ed.)* **2016**, *35*, 325–328. [[CrossRef](#)]
51. Kim, E.J.; Kim, S.H.; Kang, B.J.; Choi, B.G.; Song, B.J.; Choi, J.J. Diagnostic value of breast MRI for predicting metastatic axillary lymph nodes in breast cancer patients: Diffusion-weighted MRI and conventional MRI. *Magn. Reson. Imaging* **2014**, *32*, 1230–1236. [[CrossRef](#)]
52. Joo, S.; Ko, E.S.; Kwon, S.; Jeon, E.; Jung, H.; Kim, J.-Y.; Chung, M.J.; Im, Y.-H. Multimodal deep learning models for the prediction of pathologic response to neoadjuvant chemotherapy in breast cancer. *Sci. Rep.* **2021**, *11*, 18800. [[CrossRef](#)]
53. Lai, Z.; Deng, H. Medical Image Classification Based on Deep Features Extracted by Deep Model and Statistic Feature Fusion with Multilayer Perceptron. *Comput. Intell. Neurosci.* **2018**, *2018*, 2061516. [[CrossRef](#)]
54. Wang, S.; Li, C.; Wang, R.; Liu, Z.; Wang, M.; Tan, H.; Wu, Y.; Liu, X.; Sun, H.; Yang, R.; et al. Annotation-efficient deep learning for automatic medical image segmentation. *Nat. Commun.* **2021**, *12*, 5915. [[CrossRef](#)]
55. Khan, B.; Fatima, H.; Qureshi, A.; Kumar, S.; Hanan, A.; Hussain, J.; Abdullah, S. Drawbacks of Artificial Intelligence and Their Potential Solutions in the Healthcare Sector. *Biomed. Mater. Devices* **2023**, 1–8. [[CrossRef](#)]
56. Ansari, A.A.; Mishra, B.; Gera, P.; Khan, M.K.; Chakraborty, C.; Mishra, D. Privacy-Enabling Framework for Cloud-Assisted Digital Healthcare Industry. *IEEE Trans. Ind. Inform.* **2022**, *18*, 8316–8325. [[CrossRef](#)]
57. Cao, Y.; Zhou, Z.; Chakraborty, C.; Wang, M.; Wu, Q.M.J.; Sun, X.; Yu, K. Generative Steganography Based on Long Readable Text Generation. *IEEE Trans. Comput. Soc. Syst.* **2022**, 1–11. [[CrossRef](#)]
58. Tang, M.; Gandhi, P.; Md; Zou, C.; Blakey, J.; Luo, X. Progress Notes Classification and Keyword Extraction using Attention-based Deep Learning Models with BERT. *arXiv* **2019**, arXiv:1910.05786. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.