Research Article

Construction of a Diagnostic Model for Lymph Node Metastasis of the Papillary Thyroid Carcinoma Using Preoperative Ultrasound Features and Imaging Omics

Chao Zhang [ID],[1] **Lihua Cheng** [ID],[1] **Weiwen Zhu** [ID],[1] **Jian Zhuang** [ID],[1] **Tong Zhao** [ID],[2] **Xiaoqin Li** [ID],[1] **and Wenfeng Wang** [ID][3]

[1]*Department of Ultrasound, The Affiliated Changzhou No. 2 People's Hospital with Nanjing Medical University, Changzhou 213004, China*
[2]*Graduate School, Dalian Medical University, Dalian 116000, China*
[3]*School of Science, Shanghai Institute of Technology, Shanghai 201418, China*

Correspondence should be addressed to Xiaoqin Li; xiaoqinl_cz@outlook.com

In this paper, we mainly adopted 337 patients who had undergone the surgery on lymph node metastasis of papillary thyroid carcinoma (PTC) as the sample population. In order to provide clinical reference for the intelligent decision-making in treatment plan and improvement of prognosis, we utilized ultrasound features and imaging features to construct five early diagnosis models for patients based on the ultrasound features, imaging features, and combined features. The model integrated with broad learning system (BLS) showed the best performance, with the area under the curve (AUC) of 0.857 (95% confidence interval (CI): 0.811–0.902)) and the accuracy of 0.805 (95% CI: 0.759–0.850). For demographic and clinical features, the prediction effect was also good, with the AUC more than 0.700.

1. Introduction

Papillary thyroid carcinoma (PTC) is one of the most common pathologic types of thyroid cancer [1]. The current clinical problem is to find regions where lymph node metastasis is prone to occur [2]. This problem is usually solved by utilizing the ultrasound technology, which is also the first choice for thyroid cancer examination. Ultrasound technology can determine whether the patient has cervical lymph node metastasis before surgery, which is of great significance for the selection of surgical methods, radiotherapy and chemotherapy, and the judgment of prognosis [3]. The major advantage of machine learning is that the learning model can improve treatment decisions for cancers and provide clinical references to improve the prognosis [4]. Deep learning models have been used in previous studies, but it takes a lot of time in training stage [5–7].

As an effective and efficient incremental learning system, broad learning system (BLS) can provide value for prediction model, which largely reduced the time cost of model training [5]. If combined with imaging omics, broad learning features can then be utilized in establishing the lymph node metastasis model [6–8]. Imaging omics is mainly based on the extraction and analysis of images features from CT, MRI, PET, and other medical images to quantitatively evaluate diseases such as thyroid papillary carcinoma and lymph nodes [9]. It can be used to diagnose diseases, predict prognosis, and analyze biological behavior of diseases [10]. Imaging omics was proved to be objective in image extraction of lymph node features in PTC and had important implications for prediction of clinical outcome [11–15]. Since imaging omics has been successfully applied to the diagnosis of thyroid cancer, lung cancer, liver cancer, breast cancer, and other diseases [16–22], it will also be employed in the present study.

TABLE 1: Sensitivity analysis before and after gap-filling.

Variables	Missing number	Before filling ($n = 428$)	After filling ($n = 428$)	Statistics	P
Carcinoembryonic antigen, $M(Q_1, Q_3)$	17 (3.97%)	1.42 (0.88, 2.10)	1.42 (0.89, 2.07)	$Z = -0.066$	0.948
Free triiodothyroxine, mean ± SD	8 (1.87%)	5.15 ± 1.31	5.16 ± 1.32	$t = -0.06$	0.949
Free thyroxine, mean ± SD	8 (1.87%)	18.29 ± 4.76	17.96 ± 3.70	$t = 1.12$	0.261
Thyroid stimulating hormone, $M(Q_1, Q_3)$	8 (1.87%)	1.93 (1.12, 3.17)	1.96 (1.15, 3.19)	$Z = 0.368$	0.713
Thyroid globulin antibody, $M(Q_1, Q_3)$	8 (1.87%)	19.04 (12.98, 66.89)	18.66 (12.97, 57.48)	$Z = -0.397$	0.691
Maximum diameter of nodule, $M(Q_1, Q_3)$	1 (0.23%)	0.80 (0.50, 1.20)	0.80 (0.50, 1.20)	$Z = -0.055$	0.956

To combine imaging omics with broad learning features, random forest is employed to develop the basic analytic models, which is a combination of decision trees [23]. Each decision tree is trained by randomly generating a new data set from the original data set. The result of random forest is the decision of most decision trees [24–28]. But a single model classification method is often prone to overfitting problem. Many scholars often improve the prediction accuracy through the combination of multiple single models, which is called classifier combination method. Random forest is an algorithm that proposed to solve the overfitting problem of a single decision tree model [29]. Random forest uses the bootstrap resampling method to extract multiple samples from the original samples and then conducts decision tree modeling for each bootstrap sample, and then synthesizes multiple decision trees for prediction, and obtains the final prediction result through voting [30, 31].

The organization of this article is as follows. We will use preoperative ultrasound features and image analyses to construct an early diagnosis model for lymph node metastasis in PTC in Section 2. These models will be performed, evaluated, and then integrated with BLS in Section 3.

2. Materials and Methods

2.1. Study Design and Population.
This study was a cross-sectional study which was approved by the Institutional Review Board of The Affiliated Changzhou No. 2 People's Hospital with Nanjing Medical University (approval number: [2021]KY021-01). The sample population was 337 patients who had undergone PTC surgery in Changzhou Second People's Hospital after inclusion and exclusion.

The inclusion criteria were as follows: (1) patients aged ≥18 years old; (2) PTC patients who received fine needle biopsy before operation and were confirmed; (3) patients without benign lesions or single malignant lesions; (4) patients who underwent extensive neck lymph node dissection; (5) patients with complete clinical data.

The exclusion criteria were as follows: (1) patients who received anticancer treatment such as radiotherapy and chemotherapy before operation; (2) patients without undergoing ultrasound examination before operation.

2.2. Missing Data Assessment.
There were 428 nodules in 337 patients. Noting that each nodule had two or more ultrasound images from different angles, there were a total of 973 ultrasound images for 428 nodules in 337 patients. Alternatively, a total of 428 data and 973 representative ultrasound images were collected. Missing values in the data were filled by random interpolation. Sensitivity analysis before and after gap-filling is shown in Table 1.

2.3. Image Preprocessing and Classification.
In the present study, Lasso regression filtering is used for image processing [32, 33]. The processes were to sample n original sample data with the sample size of N and each observation object had an equal probability of being selected, which was 1/N. The sample was regarded as the whole, and the subsamples sampled were regarded as samples from the sample. Such subsample was called the bootstrap sample. The sampling process can be formulated as follows. Let H(x) represent the random forest classification result, $h_i(x)$ represent the classification result of a single decision tree, Y represent the classification target, $I(\cdot)$ represent indicative function, and the random forest classification model adopt a simple voting strategy to complete the final classification.

(1) Each decision tree was generated by training sample X with sample size K and random vector θ_k

(2) Random vector sequence $\{\theta_k, \quad 1, \dots, K\}$ was independently and identically distributed

(3) Random forest was the set of all decision trees $\{h(X, \theta_k), \quad k = 1, 2, \dots, K\}$

Among these processes, each decision tree model $h(X, \theta_k)$ had one vote to select the classification result of input variable X: $H(x) = \max_Y \sum_{i=1}^{k} I(h_i(x) = Y)$.

The remaining variable of image feature was gray-level size zone matrix (GLSZM) entropy. The remaining three variables were gender, age, and carcinoembryonic antigen in the demographic information and clinical data, and the remaining four features were the maximum diameter of nodule in ultrasound features, aspect ratio, calcification, and relative capsule position. BLS was established for image classification through learning the variables in the model to obtain the output variables. In the process of image classification, broad learning mapped the input data, constructed the mapping features, and then activated the mapping features to enhance the features, and output the two parts together. We screened out the new features by using the loss function of the 1-norm in Lasso regression, and the new features were merged into the random forest as follows:

$$J(w, b) = \frac{1}{2m} \arg\min_{w,b} \sum_{i=1}^{m} (\hat{y}_i - y_i)^2 + \alpha \sum_{i=1}^{n} \|w_i\|. \quad (1)$$

2.4. Establishment of the Diagnostic Models.
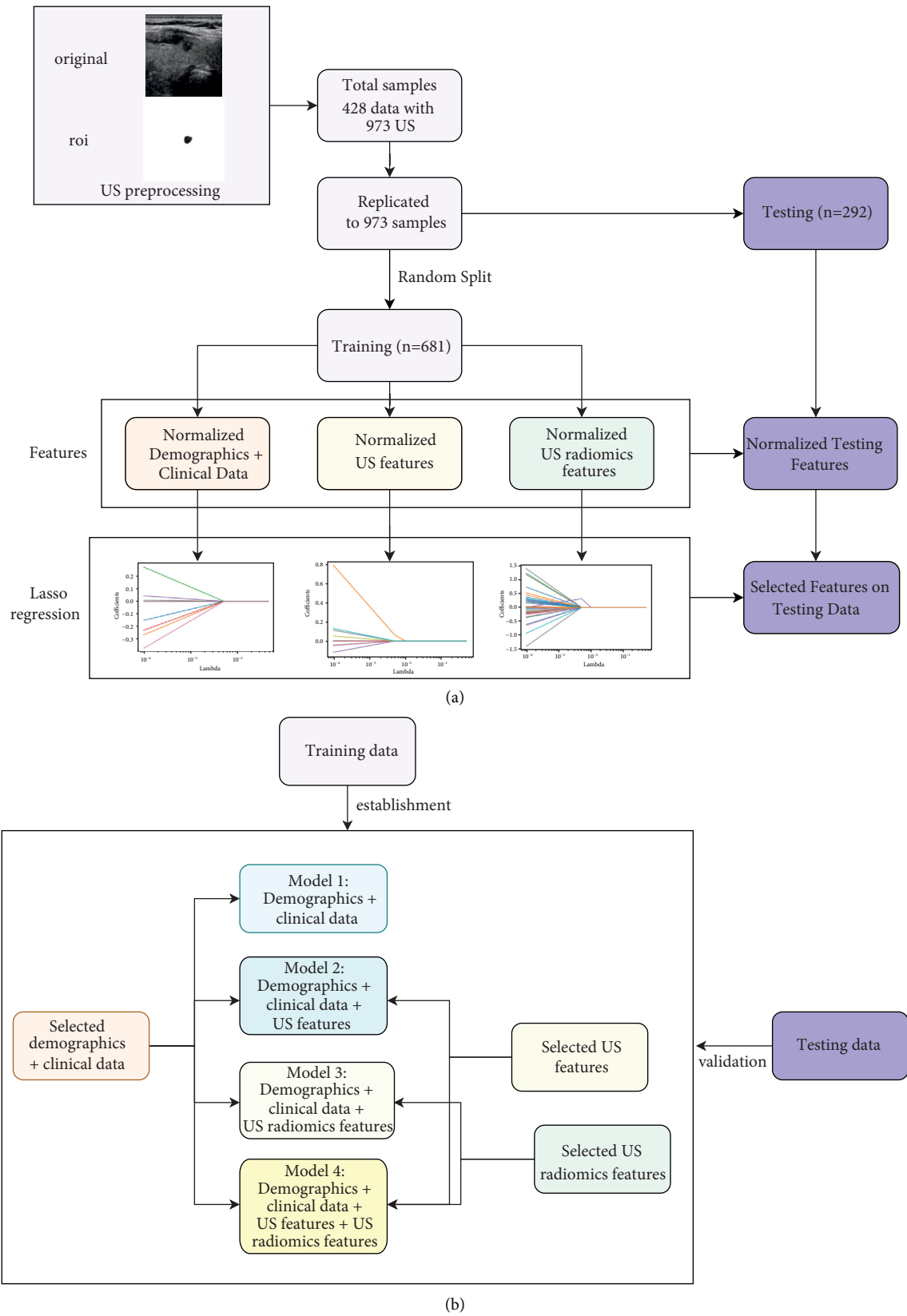For each nodule, the ROI was delineated according to the gray image

(a)

(b)

Figure 1: (a) Flow chart for the model development and validation. (b) Characteristics of the diagnostic models.

TABLE 2: Characteristics comparison for the training and testing sets.

Variables	Total ($n = 973$)	Training set ($n = 681$)	Testing set ($n = 292$)	Statistics	P
Gender, n (%)					
Male	207 (21.27)	143 (21.00)	64 (21.92)	$\chi^2 = 0.103$	0.748
Female	766 (78.73)	538 (79.00)	228 (78.08)		
Age, mean ± SD	44.74 ± 11.25	44.66 ± 11.05	44.91 ± 11.70	$t = -0.32$	0.752
BMI, mean ± SD	23.87 ± 3.39	23.90 ± 3.43	23.82 ± 3.29	$t = 0.30$	0.763
Carcinoembryonic antigen, $M(Q_1, Q_3)$	1.34 (0.86, 2.02)	1.34 (0.84, 2.02)	1.33 (0.89, 2.05)	$Z = 0.067$	0.947
The free triiodide, $M(Q_1, Q_3)$	5.10 (4.60, 5.50)	5.10 (4.60, 5.50)	5.00 (4.60, 5.50)	$Z = -0.201$	0.841
Free thyroxine, mean ± SD	18.61 ± 5.03	18.58 ± 4.80	18.67 ± 5.55	$t = -0.24$	0.812
T stimulating hormone, $M(Q_1, Q_3)$	1.91 (1.11, 3.19)	1.94 (1.15, 3.25)	1.78 (1.06, 2.99)	$Z = -1.374$	0.169
T globulin antibody, $M(Q_1, Q_3)$	20.92 (12.98, 78.98)	20.92 (13.04, 83.68)	19.11(12.27, 73.34)	$Z = -0.982$	0.326
Part, n (%)					
Ru	105 (10.79)	73 (10.72)	32 (10.96)		
Right middle school	266 (27.34)	199 (29.22)	67 (22.95)		
Lower right	148 (15.21)	101 (14.83)	47 (16.10)		
Left	73 (7.50)	47 (6.90)	26 (8.90)	$\chi^2 = 6.065$	0.416
Left middle school	237 (24.36)	167 (24.52)	70 (23.97)		
The lower left	100 (10.28)	64 (9.40)	36 (12.33)		
Isthmus	44 (4.52)	30 (4.41)	14 (4.79)		
Max diameter of nodule, $M(Q_1, Q_3)$	0.80 (0.54, 1.20)	0.80(0.53, 1.20)	0.80(0.57, 1.31)	$Z = 0.867$	0.386
Form, n (%)					
Rules	95 (9.76)	64 (9.40)	31 (10.62)		
Under-rule	318 (32.68)	217 (31.86)	101 (34.59)	$\chi^2 = 1.327$	0.515
Irregular	560 (57.55)	400 (58.74)	160 (54.79)		
Boundary, n (%)					
Clear	170 (17.47)	113 (16.59)	57 (19.52)		
Lack of clarity	368 (37.82)	261 (38.33)	107 (36.64)	$\chi^2 = 1.226$	0.542
Unclear or vague	435 (44.71)	307 (45.08)	128 (43.84)		
Aspect ratio, n (%)					
≤1	415 (42.65)	282 (41.41)	133 (45.55)		
>1	558 (57.35)	399 (58.59)	159 (54.45)	$\chi^2 = 1.431$	0.232
Composition, n (%)					
Cystic or almost totally cystic	1 (0.10)	1 (0.15)	0 (0.00)		
Capsule solidity	19 (1.95)	13 (1.91)	6 (2.05)	Fisher	1.000
Real or almost all real	953 (97.94)	667 (97.94)	286 (97.95)		
Echo, n (%)					
Isoechoic or hyperechoic	5 (0.51)	3 (0.44)	2 (0.68)		
Low echo	926 (95.17)	649 (95.30)	277 (94.86)	Fisher	0.898
Extremely low echo	21 (2.16)	14 (2.06)	7 (2.40)		
Mixed echo	21 (2.16)	15 (2.20)	6 (2.05)		
Calcification, n (%)					
No calcification	362 (37.20)	262 (38.47)	100 (34.25)		
Coarse calcification	62 (6.37)	47 (6.90)	15 (5.14)	Fisher	0.293
Eggshell calcification	6 (0.62)	5 (0.73)	1 (0.34)		
Microcalcification	543 (55.81)	367 (53.89)	176 (60.27)		
Relative coating position, n (%)					
Stay away from	398 (40.90)	286 (42.00)	112 (38.36)		
Cling	485 (49.85)	333 (48.90)	152 (52.05)	$\chi^2 = 1.123$	0.570
Breakthrough	90 (9.25)	62 (9.10)	28 (9.59)		
TI_DS classification, *n* (%)				$\chi^2 = 0.812$	0.937
LNM transfer, n (%)					
No	576 (59.20)	410 (60.21)	166 (56.85)	$\chi^2 = 0.953$	0.329
Yes	397 (40.80)	271 (39.79)	126 (43.15)		

T: thyroid; TI_DS: ultrasonic thyroid imaging and data system; LNM: lymph node metastases.

selected in the largest long axis cross section. The early diagnosis model of lymph node metastases (LNM) was constructed by combining the preoperative ultrasound features and ultrasound image features, as shown in Figures 1(a) and 1(b). The focus area of PTC was framed by the clinician, and then the imaging features of the focus area were extracted by the pyradiomics algorithm.

TABLE 3: The predictive performance of these models in the training and testing sets.

Models	Cut-off	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	AUC (95% CI)	Accuracy (95% CI)
Model 1[a]	0.348	0.849 (0.806–0.891)	0.851 (0.817–0.886)	0.790 (0.744–0.837)	0.895 (0.864–0.925)	0.913 (0.893–0.934)	0.850 (0.823–0.877)
Model 1[b]	0.348	0.738 (0.661–0.815)	0.771 (0.707–0.835)	0.710 (0.632–0.788)	0.795 (0.733–0.857)	0.813 (0.762–0.863)	0.757 (0.708–0.806)
Model 2[a]	0.437	0.808 (0.761–0.855)	0.868 (0.836–0.901)	0.802 (0.755–0.849)	0.873 (0.840–0.905)	0.913 (0.892–0.934)	0.844 (0.817–0.872)
Model 2[b]	0.437	0.730 (0.653–0.808)	0.789 (0.727–0.851)	0.724 (0.647–0.802)	0.794 (0.732–0.856)	0.818 (0.769–0.868)	0.764 (0.715–0.812)
Model 3[a]	0.360	0.886 (0.848–0.924)	0.827 (0.790–0.863)	0.772 (0.725–0.818)	0.916 (0.888–0.944)	0.941 (0.925–0.957)	0.850 (0.823–0.877)
Model 3[b]	0.360	0.762 (0.688–0.836)	0.723 (0.655–0.791)	0.676 (0.599–0.753)	0.800 (0.736–0.864)	0.821 (0.772–0.871)	0.740 (0.689–0.790)
Model 4[a]	0.500	0.823 (0.777–0.868)	0.966 (0.948–0.983)	0.941 (0.911–0.971)	0.892 (0.863–0.921)	0.984 (0.977–0.990)	0.909 (0.887–0.931)
Model 4[b]	0.500	0.667 (0.584–0.749)	0.910 (0.866–0.953)	0.848 (0.778–0.919)	0.782 (0.724–0.841)	0.857 (0.811–0.902)	0.805 (0.759–0.850)

[a]Using the training set; [b]using the testing set. PPV: positive predictive value; NPV: predictive value; AUC: area under the curve; CI: confidence internal.
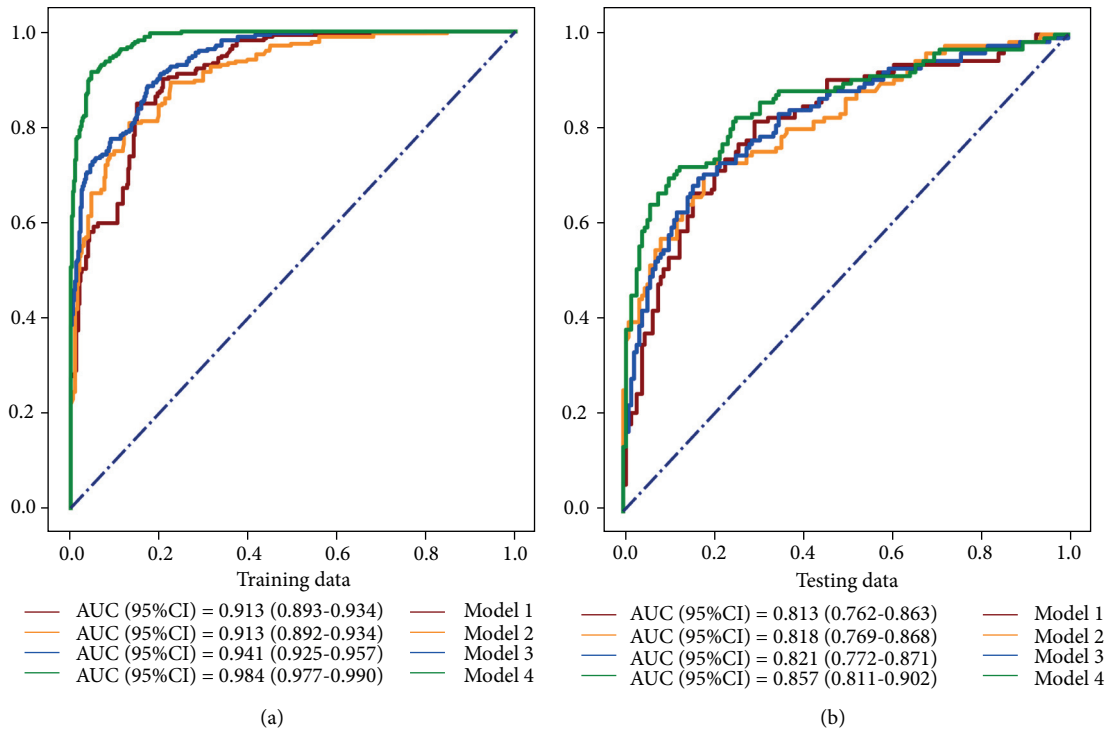

FIGURE 2: ROC curves of the four models in training and testing sets.

TABLE 4: The predictive performance of lymph node metastasis by BLS feature learning.

Models	Cut-off	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	AUC (95% CI)	Accuracy (95% CI)
Model 4[a]	0.500	0.823 (0.777–0.868)	0.966 (0.948–0.983)	0.941 (0.911–0.971)	0.892 (0.863–0.921)	0.984 (0.977–0.990)	0.909 (0.887–0.931)
Model 4[b]	0.500	0.667 (0.584–0.749)	0.910 (0.866–0.953)	0.848 (0.778–0.919)	0.782 (0.724–0.841)	0.857 (0.811–0.902)	0.805 (0.759–0.850)
Model 5[a]	0.465	0.959 (0.936–0.983)	0.973 (0.958–0.989)	0.959 (0.936–0.983)	0.973 (0.958–0.989)	0.995 (0.992–0.998)	0.968 (0.954–0.981)
Model 5[b]	0.465	0.778 (0.705–0.850)	0.843 (0.788–0.899)	0.790 (0.719–0.862)	0.833 (0.777–0.890)	0.853 (0.806–0.901)	0.815 (0.771–0.860)

[a]Using the training set; [b]using the testing set. PPV: positive predictive value; NPV: predictive value; AUC: area under the curve; CI: confidence internal.

Combined training
AUC (95%CI) = 0.984 (0.977-0.990)

Combined testing
AUC (95%CI) = 0.857 (0.811-0.902)

(a)

BLS features training
AUC (95%CI) = 0.995 (0.992-0.998)

BLS features testing
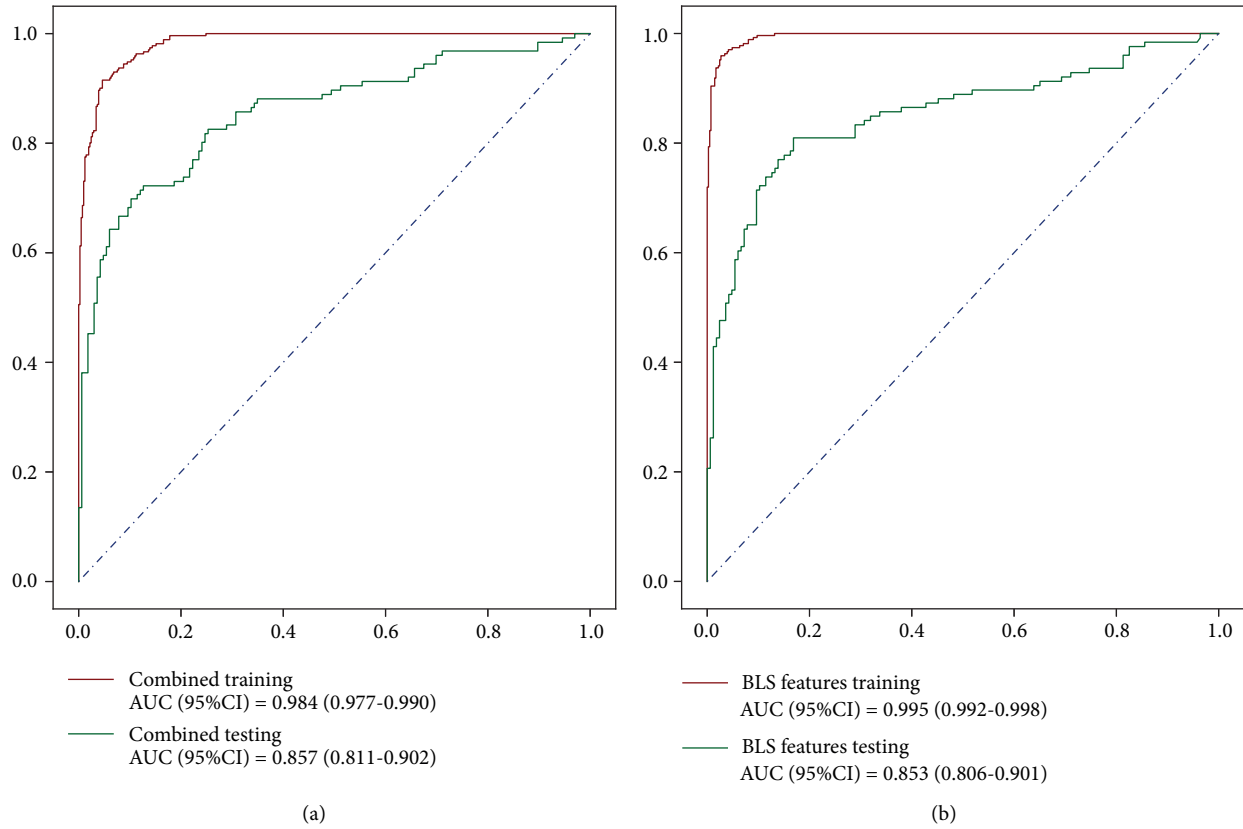AUC (95%CI) = 0.853 (0.806-0.901)

(b)

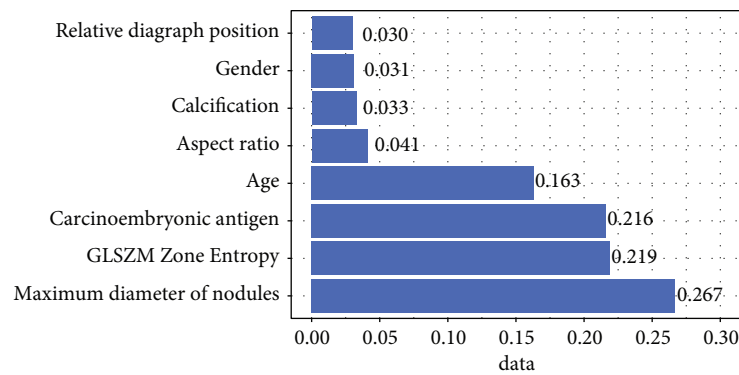Figure 3: The ROC curve of Model 4 and Model 5 in training and testing sets.

Figure 4: Features importance map.

The strategy in construction of the five diagnostic models was different. In Model 1, only demographic information and clinical data were used. In Model 2, we combined demographic information, clinical data, and ultrasound features. Model 3 combined demographic information, clinical data, and imaging features. Model 4 combined the demographic information, clinical data, ultrasound features, and imaging features. Broad learning was used to learn the variables in Model 4, and new variables were obtained, which were incorporated into the random forest model to obtain Model 5.

The data set was randomly divided into 7 : 3 training set and testing set, which were then normalized, respectively.

Lasso regression was used to filter features in the training set, and then the prediction model was constructed.

The area under the curve (AUC), accuracy, sensitivity, and specificity were used to evaluate the model. Then, the importance of features was expressed by using the feature importance map.

3. Results and Discussion

3.1. Diagnostic Performance of the Five Models. As shown in Table 2, because the data were randomly divided into training sets and testing sets, the ultrasonic features were compared in balance. Because their P values were all >0.05,
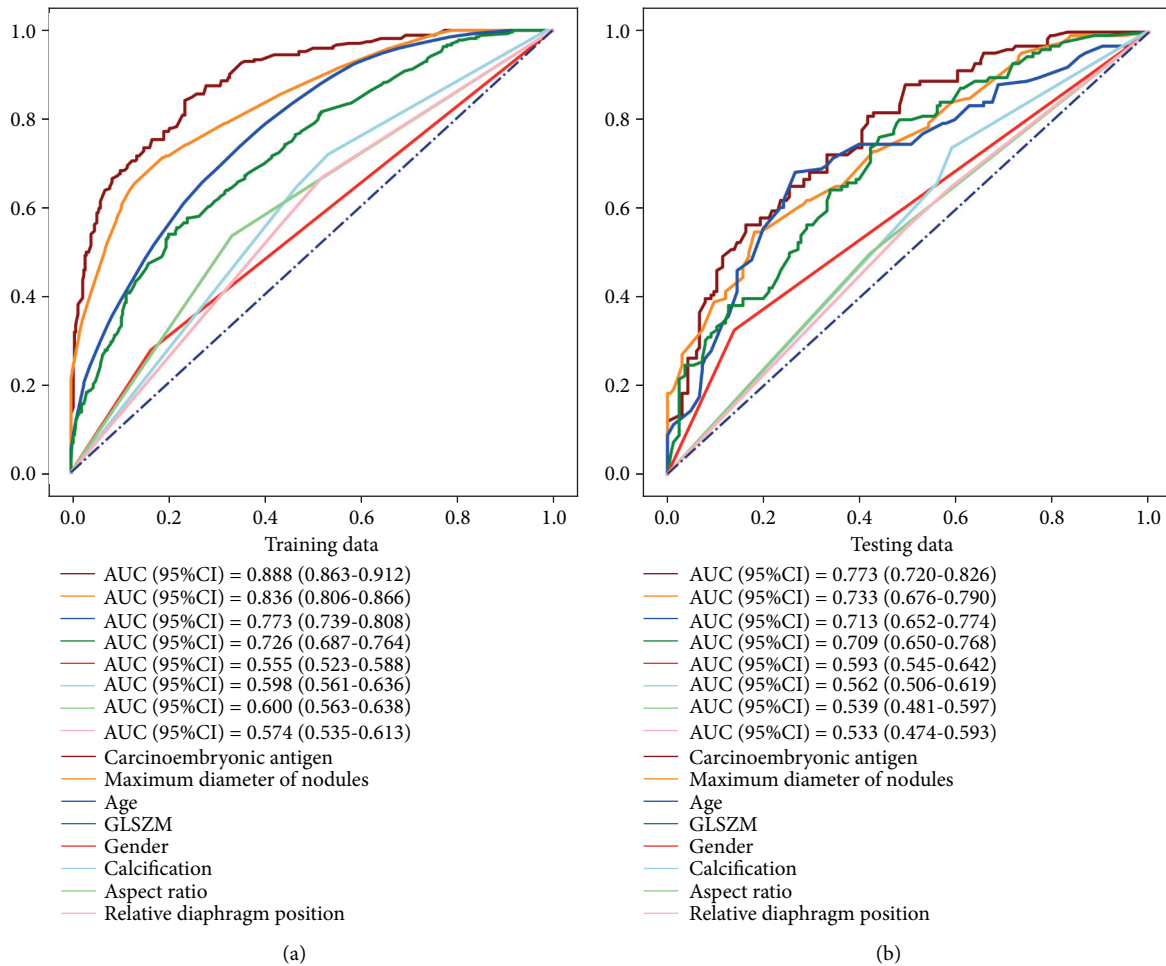
FIGURE 5: Ultrasonic features AUC: testing set on the left and training set on the right (carcinoembryonic antigen, maximum diameter of nodule, age, GLSZM zone entropy, sex, calcification, aspect ratio, and relative capsule position in turn, which are modified according to translation.

the difference between training sets and testing sets was not statistically significant. This confirmed that the performances between training sets and testing sets were comparable.

As shown in Table 3, it can be found from the prediction results that Model 4 performed best in the testing set, which combined ultrasound features and imaging features in our data set, with the AUC of 0.857 (95% confidence internal (CI): 0.811–0.902) and the accuracy of 0.805 (95%CI: 0.759–0.850). The receiver operator characteristic (ROC) curves of the four models are shown in Figure 2.

3.2. Prediction Results after Integrated with Broad Learning. Eight features in Model 4 were included in BLS model to get 106 features, and then 5 features were screened out by Lasso using $\alpha = 0.004$. Five features put into the stochastic forest prediction model to predict whether lymph node metastasis occurred are shown in Table 4. Figure 3 shows the ROC curve of Model 4 and Model 5 in training and testing sets.

3.3. Discussion on the Importance of Model Features. Since the prediction results of Model 4 and Model 5 were relatively close, and the difference was not statistically

significant, we finally chose Model 4 because of its high interpretability. From the map of feature importance (Figure 4), it can be found that the most important variable was the maximum diameter of nodules, followed by GLSZM zone entropy in imaging features, and the third was carcinoembryonic antigen.

Overall, carcinoembryonic antigen and age were the best predictors of demographic and clinical features. Among ultrasonic features, the maximum diameter of the nodule was the best predictor. Imaging features also predict well, as seen in Figure 5.

As the most common thyroid malignancy, the papillary thyroid cancer is associated with cervical lymph node metastases in 30% to 90% of patients [34]. The lymph node dissection (LND) is the mainstay treatment for clinically evident cervical lymph node metastases [35]. So far, surgical treatment options in the literature include the traditional radical LND, the modified radical LND, the selective LND, and a "berry picking" resection in which only the grossly abnormal lymph nodes are excised [36–42]. The selective LND represents a compartment-based resection based on documented lymph node metastases [43, 44]. This study constructed diagnostic models through an integration of the

random forest and BLS, which was demonstrated to be a successful attempt to break the related bottlenecks in the future.

Before constructing the diagnostic models for lymph node metastasis, we considered using the ultrasound features of lymph nodes as input. But cervical lymph nodes are widely distributed (mainly in 6 regions) and there are some limitations in the feature recognition of cervical lymph nodes by ultrasound, especially the lymph nodes in the central area, as well as the special anatomical structures such as posterior trachea, posterior esophagus, retropharyngeal area, and mediastinum, which cannot be displayed well by ultrasound [45]. Meanwhile, researches in modeling, diagnosis, and treatment have confirmed that some ultrasonic features of primary lesions are related to lymph node metastasis [46, 47]. Our experiment also demonstrated that we can better identify lymph node metastasis in different regions through imaging features of primary lesions. This was the reason why no lymph node features were used as input in the present study.

Unlike the analysis of normal cancer [48, 49], lymph node metastasis is detected through postoperative pathology (gold standard) [50]. The inclusion criteria of this study were those who underwent extensive neck lymph node dissection to ensure the accurate diagnosis of LNM. The potential risk of lymph node metastasis has led to many PTC patients receiving total thyroidectomy, lymph node dissection, and other treatments, resulting in widespread overtreatment. Therefore, we hope to build a diagnostic model of preoperative LNM to help realize accurate identification of high-risk patients with LNM in this population to reduce overtreatment.

We constructed five early diagnosis models of LNM by combining the preoperative ultrasound features and ultrasound image features. These models were chosen for the convenience to extract the imaging features of the focus area of PTC. The strategy in construction of the five diagnostic models was different, and finally, we founded that the top three parameters are more important than the others. These results present further evidence for a systematic review and meta-analysis in previous studies, which indicated that patient gender is a factor associated with lymph node metastasis in T1 colorectal cancer [51]. The clinical significance lies in helping the clinicians in early diagnosis, which not only reduces the workload of clinicians but also cut off the suffering of patients [52–54]. The previous studies utilized deep learning algorithms for detection of lymph node metastases, while broad learning algorithms were rarely utilized [54]. Deep learning models spend too much time in the training stage and BLS can greatly reduce the time cost in training the model [53, 54]. This was also the major innovation of our study.

4. Conclusion

In this paper, five early diagnostic models were developed from random forest and integrated with the BLS to obtain experimental results with population informatics, clinical data, and ultrasonic characteristics. There was no significant difference between the combining of BLS and random forest and random forest was chosen to make predictions for the model. The most important feature map statistics show the maximum diameter of the nodule. It was the most important variable, followed by the GLSZM zone entropy and hence should be employed in subsequent studies [55–60].

Data Availability

All the data to support the experiments and findings in this study are available from the corresponding authors upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] N. Tohnosu, S. Onoda, and K. Isono, "Ultrasonographic evaluation of cervical lymph node metastases in esophageal cancer with special reference to the relationship between the short to long axis ratio (S/L) and the cancer content," *Journal of Clinical Ultrasound*, vol. 17, no. 2, pp. 101–106, 2010.

[2] S. Natsugoe, H. Yoshinaka, T. Morinaga et al., "Ultrasonographic detection of lymph-node metastases in superficial carcinoma of the esophagus," *Endoscopy*, vol. 28, no. 8, pp. 674–679, 1996.

[3] Y. Liu, J. W. Zhang, and Q. M. Wang, "Diagnostic efficacy of ultrasonographic characteristics of thyroid carcinomain predicting cervical lymph node metastasis," *Ultrasound in Medicine and Biology*, vol. 42, no. 1, pp. 68–74, 2016.

[4] J. M. Fitzpatrick, C. N. Sternberg, F. Saad et al., "Treatment decisions for advanced genitourinary cancers: from symptoms to risk assessment," *European Urology Supplements*, vol. 8, no. 9, pp. 738–746, 2009.

[5] C. L. Philip Chen and Z. L. Liu, "Broad learning system: an effective and efficient incremental learning system without the need for deep architecture," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 99, pp. 10–24, 2018.

[6] P. Oehr, "Omics"-based imaging in cancer detection and therapy," *Personalized Medicine*, vol. 3, no. 1, pp. 19–32, 2006.

[7] S. M. Han, H. J. Lee, and J. Y. Choi, "Computer-aided prostate cancer detection using texture features and clinical features in ultrasound image," *Journal of Digital Imaging*, vol. 21, no. S1, pp. 121–133, 2008.

[8] J. E. Belizário, B. A. Sangiuliano, M. Perez-Sosa, B. V Santos, and G Machado-Santelli, "Advances in the integration of optical and mass spectrometry molecular imaging technologies: from omics data to molecular signature discovery," *Discovery Medicine*, vol. 20, no. 112, pp. 393–401, 2015.

[9] K. Zetzmann, I. Ludolph, R. E. Horch, and A. M. Boss, "Imaging for treatment planning in lipo-and lymphedema," *Phlebologie*, vol. 49, no. 2, pp. 72–78, 2020.

[10] A. Holzinger, B. Haibe-Kains, and I. Jurisica, "Why imaging data alone is not enough: AI-based integration of imaging,

omics, and clinical data," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 46, no. 9, pp. 2722–2730, 2019.

[11] D. De Groot, F. Kuper, M. Radonjic et al., "Imaging and "omics" in rats demonstrate impact of juvenile versus developmental exposure to TBTO," *Reproductive Toxicology*, vol. 32, no. 2, pp. 165-166, 2011.

[12] J.-K. Hériché, S. Alexander, and J. Ellenberg, "Integrating imaging and omics: computational methods and challenges," *Annual Review of Biomedical Data Science*, vol. 2, no. 1, pp. 175–197, 2019.

[13] D. De Groot, D. De Groot, F. Kuper et al., "Imaging and omics in developing and juvenile rats after exposure to TBTO," *Toxicology Letters*, vol. 211, no. supp-S, p. S155, 2012.

[14] A. J. Benjamin, M. M. Buschmann, A. Schneider et al., "Can comprehensive imaging analysis with analytic morphomics and geriatric assessment predict serious complications in patients undergoing pancreatic surgery?" *Journal of Gastrointestinal Surgery*, vol. 21, no. 6, pp. 1009–1016, 2017.

[15] J. A. Disselhorst, M. A. Krueger, S. M. M. Ud-Dean et al., "Linking imaging to omics utilizing image-guided tissue extraction," *Proceedings of the National Academy of Sciences*, vol. 115, no. 13, pp. E2980–E2987, 2018.

[16] D. Viola, G. Materazzi, L. Valerio et al., "Prophylactic central compartment lymph node dissection in papillary thyroid carcinoma: clinical implications derived from the first prospective randomized controlled single institution study," *Journal of Clinical Endocrinology & Metabolism*, vol. 100, no. 4, pp. 1316–1324, 2015.

[17] H. Y. Chai, L. K. Wee, T. T. Swee, ShH Salleh, and L. Y Chea, "An artifacts removal post-processing for epiphyseal region-of-interest (EROI) localization in automated bone age assessment (BAA)," *BioMedical Engineering Online*, vol. 10, no. 1, pp. 87–22, 2011.

[18] A. Faisal, S.-C. Ng, S.-L. Goh, and K. W. Lai, "Knee cartilage segmentation and thickness computation from ultrasound images," *Medical, & Biological Engineering & Computing*, vol. 56, no. 4, pp. 657–669, 2018.

[19] S. M. Lewis, M. L. Asselin-Labat, N. Quan et al., "Spatial omics and multiplexed imaging to explore cancer biology," *Nature Methods*, pp. 1–16, 2021.

[20] T. Smets, T. De Keyser, T. Tousseyn, E. Waelkens, and B. De Moor, "Correspondence-aware manifold learning for microscopic and spatial omics imaging: a novel data fusion method bringing mass spectrometry imaging to a cellular resolution," *Analytical Chemistry*, vol. 93, no. 7, pp. 3452–3460, 2021.

[21] V. Kumar, Y. Gu, S. Basu et al., "Radiomics: the process and the challenges," *Magnetic Resonance Imaging*, vol. 30, no. 9, pp. 1234–1248, 2012.

[22] R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley, "Deep learning for healthcare: review, opportunities and challenges," *Briefings in Bioinformatics*, vol. 19, no. 6, pp. 1236–1246, 2017.

[23] P. Jiang, H. Wu, W. Wang, W. Ma, X. Sun, and Z. Lu, "MiPred: classification of real and pseudo microRNA precursors using random forest prediction model with combined features," *Nucleic Acids Research*, vol. 35, pp. W339–W344, 2007.

[24] V. F. Rodriguez-Galiano, B. Ghimire, J. Rogan, M. Chica-Olmo, and J. P. Rigol-Sanchez, "An assessment of the effectiveness of a random forest classifier for land-cover classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 67, pp. 93–104, 2012.

[25] R. L. Lawrence, S. D. Wood, and R. L. Sheley, "Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (randomForest)," *Remote Sensing of Environment*, vol. 100, no. 3, pp. 356–362, 2006.

[26] H. Bart, B. Stefan, V. R. Frans, and D. B. Pieter, "A flexible integrative approach based on random forest improves prediction of transcription factor binding sites," *Nucleic Acids Research*, vol. 40, no. 14, p. e106, 2012.

[27] Y. Qi, J. Klein-Seetharaman, and Z. Bar-Joseph, "Random forest similarity for protein-protein interaction prediction from multiple sources," *Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing*, vol. 10, pp. 531–542, 2005.

[28] W. G. Touw, J. R. Bayjanov, L. Overmars et al., "Data mining in the Life Sciences with Random Forest: a walk in the park or lost in the jungle?" *Briefings in Bioinformatics*, vol. 14, no. 3, pp. 315–326, 2013.

[29] T. Shi, D. Seligson, A. S. Belldegrun, A. Palotie, and S. Horvath, "Tumor classification by tissue microarray profiling: random forest clustering applied to renal cell carcinoma," *Modern Pathology*, vol. 18, no. 4, pp. 547–557, 2005.

[30] J. Butaye, H. Jacquemyn, and M. Hermy, "Differential colonization causing non-random forest plant community structure in a fragmented agricultural landscape," *Ecography*, vol. 24, no. 4, pp. 369–380, 2001.

[31] J. C.-W. Chan and D. Paelinckx, "Evaluation of random forest and adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery," *Remote Sensing of Environment*, vol. 112, no. 6, pp. 2999–3011, 2008.

[32] D. W. Hosmer, T. Hosmer, C. S. Le, and S. Lemesshow, "A comparison of goodness-of-fit tests for the logistic regression model," *Statistics in Medicine*, vol. 16, no. 9, pp. 965–980, 2015.

[33] B. Foa and B. Ga, "Fitting logistic regression models to assess vitamin D deficiency with clinical parameters in chronic hepatitis B patients," *Infectious Disease Modelling*, vol. 6, pp. 612–617, 2021.

[34] N. R. Caron and O. H. Clark, "Papillary thyroid cancer: surgical management of lymph node metastases," *Current Treatment Options in Oncology*, vol. 6, no. 4, pp. 311–322, 2005.

[35] T. P. J. Hennessy, "Lymph node dissection," *World Journal of Surgery*, vol. 18, no. 3, pp. 367–372, 1994.

[36] R. Groningen, "Extended lymph node dissection for gastric cancer: who may benefit? Final results of the randomized Dutch gastric cancer group trial," *Journal of Clinical Oncology*, vol. 23, no. 23, pp. 5404-5405, 2005.

[37] G. J. Jager, R. A. Heesakkers, and J. A. Witjes, "The importance of the extent of pelvic-lymph-node dissection in the diagnosis of lymph-node metastases in prostate cancer-Reply," *The Lancet Oncology*, vol. 9, no. 10, pp. 915–917, 2008.

[38] N. Martini, "Mediastinal lymph node dissection for lung cancer. The Memorial experience," *Chest Surgery Clinics of North America*, vol. 5, no. 2, pp. 189–203, 1995.

[39] M. Noguchi, "Benefits of D_2 lymph node dissection for patients with gastric cancer and pN0 and pN1 lymph node metastases," *British Journal of Surgery*, vol. 83, no. 7, pp. 1144–1147, 2010.

[40] P. Shen, J. M. Guenther, L. A. Wanek, and D. L. Morton, "Can elective lymph node dissection decrease the frequency and mortality rate of late melanoma recurrences?" *Annals of Surgical Oncology*, vol. 7, no. 2, pp. 114–119, 2000.

[41] S. M. Keller, S. Adak, H. Wagner, and D. H. Johnson, "Mediastinal lymph node dissection improves survival in patients with stages II and IIIa non-small cell lung cancer," *The Annals of Thoracic Surgery*, vol. 70, no. 2, pp. 358–365, 2000.

[42] B. H. Bochner, H. W. Herr, and V. E. Reuter, "Impact of separate versus en bloc pelvic lymph node dissection on the number of lymph nodes retrieved in cystectomy specimens," *The Journal of Urology*, vol. 166, no. 6, pp. 2295-2296, 2001.

[43] D. E. Elder, D. Guerry, M. Vanhorn et al., "The role of lymph node dissection for clinical stage I malignant melanoma of intermediate thickness (1.51-3.99 mm)," *Cancer*, vol. 56, no. 2, pp. 413–418, 2015.

[44] M. E. Allaf, G. S. Palapattu, B. J. Trock, H. B. Carter, and P. C. Walsh, "Anatomical extent of lymph node dissection: impact on men with clinically localized prostate cancer," *The Journal of Urology*, vol. 172, no. 5, pp. 1840–1844, 2004.

[45] H. K. Su, L. L. Dos Reis, M. A. Lupo et al., "Striving toward standardization of reporting of ultrasound features of thyroid nodules and lymph nodes: a multidisciplinary consensus statement," *Thyroid*, vol. 24, no. 9, pp. 1341–1349, 2014.

[46] M. Yoshimasa, Y. Masahiro, H. Shotaro, M. Takata, and W. Nishio, "Chemosensitivity of lung cancer: differences between the primary lesion and lymph node metastasis," *Oncology Letters*, vol. 1, no. 2, p. 345, 2010.

[47] H. A. Weinberger and D. Stetten, "Extensive secondary axillary lymph node carcinoma without clinical evidence of primary breast lesion," *Surgery*, vol. 29, no. 2, pp. 217–222, 1951.

[48] Z. Jahanzad, Y. M. Liew, M. Bilgen et al., "Regional assessment of LV wall in infarcted heart using tagged MRI and cardiac modelling," *Physics in Medicine and Biology*, vol. 60, no. 10, pp. 4015–4031, 2015.

[49] K. Azira, S. C. Ng, M. L. Yih, and K. W. Lai, "An overview on image registration techniques for cardiac diagnosis and treatment," *Cardiology Research and Practice*, vol. 2018, Article ID 1437125, 15 pages, 2018.

[50] R. Narinder, "Epidural technique for postoperative pain: gold standard no more?" *Regional Anesthesia and Pain Medicine*, vol. 37, no. 3, pp. 310–317, 2012.

[51] K. Ichimasa, S.-E. Kudo, H. Miyachi et al., "Patient gender as a factor associated with lymph node metastasis in T1 colorectal cancer: a systematic review and meta-analysis," *Molecular and Clinical Oncology*, vol. 6, no. 4, pp. 517–524, 2017.

[52] J. A. Golden, "Deep learning algorithms for detection of lymph node metastases from breast cancer: helping artificial intelligence Be seen," *JAMA*, vol. 318, no. 22, pp. 2184–2186, 2017.

[53] K. G. Andersen, E. K. Aasvang, N. Kroman, and N. Kehlet, "Intercostobrachial nerve handling and pain after axillary lymph node dissection for breast cancer," *Acta Anaesthesiologica Scandinavica*, vol. 58, no. 10, pp. 1240–1248, 2015.

[54] H. H. N. Pham, M. Futakuchi, A. Bychkov, T. Furukawa, K. Kuroda, and J. Fukuoka, "Detection of lung cancer lymph node metastases from whole-slide histopathologic images using a two-step deep learning approach," *American Journal Of Pathology*, vol. 189, no. 12, pp. 2428–2439, 2019.

[55] A. Briganti, M. L. Blute, J. H. Eastham et al., "Pelvic lymph node dissection in prostate cancer," *European Urology*, vol. 55, no. 6, pp. 1251–1265, 2009.

[56] R. Essner, A. Conforti, M. C. Kelley et al., "Efficacy of lymphatic mapping, sentinel lymphadenectomy, and selective complete lymph node dissection as a therapeutic procedure for early-stage melanoma," *Annals of Surgical Oncology*, vol. 6, no. 5, pp. 442–449, 1999.

[57] H. Shimada, I. Endo, S. Togo, A. Nakano, T. Izumi, and G. Nakagawara, "The role of lymph node dissection in the treatment of gallbladder carcinoma," *Cancer*, vol. 79, no. 5, pp. 892–899, 2015.

[58] J. H. M. Blom, H. van Poppel, J. M. Maréchal et al., "Radical nephrectomy with and without lymph-node dissection: final results of European organization for research and treatment of cancer (EORTC) randomized phase 3 trial 30881," *European Urology*, vol. 55, no. 1, pp. 28–34, 2009.

[59] N. Altorki, M. Kent, C. Ferrara, and J. Port, "Three-field lymph node dissection for squamous cell and adenocarcinoma of the esophagus," *Annals of Surgery*, vol. 236, no. 2, pp. 177–183, 2002.

[60] H. Hayashi, T. Ochiai, H. Shimada, and Y. Gunji, "Prospective randomized study of open versus laparoscopy-assisted distal gastrectomy with extraperigastric lymph node dissection for early gastric cancer," *Surgical Endoscopy*, vol. 19, no. 9, pp. 1172–1176, 2005.