PAPER

Multi-view learning for lymph node metastasis prediction using tumor and nodal radiomics in gastric cancer

IOP Publishing Phys. Med. Biol. 67 (2022) 055007 https://doi.org/10.1088/1361-6560/ac515b

Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

RECEIVED
18 August 2021

REVISED
22 January 2022

ACCEPTED FOR PUBLICATION
2 February 2022

PUBLISHED
24 February 2022

PAPER

Multi-view learning for lymph node metastasis prediction using tumor and nodal radiomics in gastric cancer

Jing Yang[1,2,3,6], **Li Wang**[4,5,6], **Jiale Qin**[3], **Jichen Du**[1], **Mingchao Ding**[1], **Tianye Niu**[1,2] and **Rencang Li**[4]

[1] Peking University Aerospace School of Clinical Medicine, Aerospace Center Hospital, Beijing 100049, People's Republic of China
[2] Shenzhen Bay Laboratory, Shenzhen 518118, People's Republic of China
[3] Women's Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, 310006, People's Republic of China
[4] Department of Mathematics, University of Texas at Arlington, Arlington, TX 76019-0408, United States of America
[5] Department of Computer Science and Engineering, University of Texas at Arlington, Arlington, TX 76019-0408, United States of America
[6] These authors contributed equally to this work, and should be considered as co-first authors.

E-mail: tyniu@qq.com

Keywords: radiomics, multi-view learning, lymph node metastasis, gastric cancer, preoperative prediction

Supplementary material for this article is available online

Abstract

Purpose. This study aims to develop and validate a multi-view learning method by the combination of primary tumor radiomics and lymph node (LN) radiomics for the preoperative prediction of LN status in gastric cancer (GC). *Methods.* A total of 170 contrast-enhanced abdominal CT images from GC patients were enrolled in this retrospective study. After data preprocessing, two-step feature selection approach including Pearson correlation analysis and supervised feature selection method based on test-time budget (FSBudget) was performed to remove redundance of tumor and LN radiomics features respectively. Two types of discriminative features were then learned by an unsupervised multi-view partial least squares (UMvPLS) for a latent common space on which a logistic regression classifier is trained. Five repeated random hold-out experiments were employed. *Results.* On 20-dimensional latent common space, area under receiver operating characteristic curve (AUC), precision, accuracy, recall and F1-score are 0.9531 ± 0.0183, 0.9260 ± 0.0184, 0.9136 ± 0.0174, 0.9468 ± 0.0106 and 0.9362 ± 0.0125 for the training cohort respectively, and 0.8984 ± 0.0536, 0.8671 ± 0.0489, 0.8500 ± 0.0599, 0.9118 ± 0.0550 and 0.8882 ± 0.0440 for the validation cohort respectively (reported as mean $\pm$ standard deviation). It shows a better discrimination capability than single-view methods, our previous method, and eight baseline methods. When the dimension was reduced to 2, the model not only has effective prediction performance, but also is convenient for data visualization. *Conclusions.* Our proposed method by integrating radiomics features of primary tumor and LN can be helpful in predicting lymph node metastasis in patients of GC. It shows multi-view learning has great potential for guiding the prognosis and treatment decision-making in GC.

1. Introduction

Gastric cancer (GC) is the fifth most common malignant tumor and the fourth leading cause of cancer-related death globally (Sung *et al* 2021). Lymph node (LN) metastasis (LNM) determines the extent of LN dissection and is one of the main independent prognostic factors in GC (Forman and Burley 2006, Cho *et al* 2007, Lum *et al* 2020). Patients with occult LNM not diagnosed by histological examination had a significantly poor prognosis compared to those without occult LNM (Huang *et al* 2013). Therefore, it is of great significance to evaluate the LN status for the improvement of prognosis in patients with GC.

In clinic, postoperative factors represented by tumor size, histological types and lymphovascular invasion have been identified as risk factors of LNM on GC patients (Shang-Guan *et al* 2018, Lin *et al* 2019). However, nearly half of patients who undergo surgical resection will experience relapse (Sada *et al* 2019). Another strategy

for evaluating the LN status can be relied on the comprehensive characteristics of primary tumors, but most of them need to be obtained from postoperative pathological examination (Wang *et al* 2020). A noninvasive medical strategy to avoid the risk and complications of operation is needed. Currently, computed tomography (CT) is the most widely used imaging modality for the evaluation of LN in GC. It mainly relies on morphological features, leading to unsatisfactory accuracy only at approximately 60% (Kim *et al* 2005a, 2005b, Lee *et al* 2010). Thus, accurate preoperative prediction for LNM of GC patients is still a great challenge.

Radiomics is a process that automatically extract high-dimensional quantitative features from imaging data, followed by further data analysis for clinical decision support (Lambin *et al* 2012, Gillies *et al* 2016, Larue *et al* 2017, Traverso *et al* 2018). Several works have shown that CT radiomics could facilitate the preoperative individualized prediction of LN status in GC (Feng *et al* 2019, Jiang *et al* 2019, Wang *et al* 2020, Gao *et al* 2020a, 2020b). For example, Gao *et al* (2020a) utilized multivariate logistic regression analysis to build a CT-based radiomics model for preoperative prediction of LNM in 463 early-stage GC patients, indicting good predictive power in the training and testing cohort with AUC values of 0.91 and 0.89, respectively. Wang *et al* (2020) constructed a radiomics nomogram to predict the LNM by using a random forest (RF) algorithm. Its AUC values also showed good discrimination in 247 consecutive GC patients (training cohort, 0.886 (0.808–0.941); testing cohort, 0.881 (0.759–0.956)). These studies mainly utilize tumor-only radiomics features or integrate with a small amount of clinicopathologic features (e.g. CT-reported LN status, serum biomarkers, T and N stage) to find non-invasive predictors of LNM. However, the occurrence of LN is accompanied by the development and progression of primary tumor. Relying solely on tumor radiomics cannot thoroughly reflect the heterogeneity underlying the complicated biological process of GC.

In addition to the above characteristics, our previous research found LN radiomics has the predictive power for discriminating GC patients with or without LNM (Yang *et al* 2020). And the LNM of patients can be better predicted by simply fusing primary tumors radiomics and LN radiomics. Wang *et al* (2021) also received favorable predictive accuracy in predicting No.3 LNM in T1-2 GC patients by integrating quantitative radiomics features of No.3 LN and primary tumors. These results suggest that the combination of tumor radiomics features and LN radiomics features may capture more information of heterogeneity and explain latent relationship between imaging characteristics and LNM risk in GC patients. Unfortunately, directly combining the two together may not comprehensively describe the information, thus limits the ability to accurately predict LN status.

2. Previous work and novel contributions

For LNM classification, tumor radiomics and LN radiomics could be considered as two views (data sources), which can be complementary and redundant to each other. Considering different views may have different number of features and are not directly comparable, how to effectively integrate information coming from different views becomes an urgent problem to be solved.

Multi-view learning has been attracting broad attention for leveraging the data collected from multiple views to overcome the limitations of single-view analysis (Sun 2013, Zhao *et al* 2017). In the field of medical image processing, multi-view learning has been increasingly used to deal with machine learning problems (such as segmentation (Liang *et al* 2020, Xia *et al* 2020), recognition (Wei *et al* 2019, Zhou *et al* 2021), and classification (Fratello *et al* 2017, Puyol-Anton *et al* 2019, Xie *et al* 2019, Thammasorn *et al* 2021)) of high-dimensional data represented by multiple distinct feature sets. There has been some related work with multi-view learning methods to establish statistically significant correlations between radiomics and clinical endpoints. Li *et al* (2020) presented a radiomics approach to design a multi-view network architecture for mammographic density classification. Lee *et al* (2020) proposed a multi-view data analysis approach using radiomics and dosiomics texture features in 388 patients with lung cancer radiotherapy to predict acute-phase weight loss. Nevertheless, the methodological research of multi-view CT radiomics aiming at GC patients is still unclear.

To the best of our knowledge, there is no research studying in practical applications of multi-view learning in GC at present. Herein, based on our previous study, we developed and validated a multi-view learning method using both primary tumor radiomics and LN radiomics for the LNM prediction of GC patients. The specific novel contributions of this work are as follows:

- Multi-view learning technology is introduced into the field of GC radiomics for the first time. A multi-view subspace learning method is utilized to exploit the complementarity of multi-view data at the decision level. It could find a latent subspace shared by each view and fuse them together in a shared view.
- We extract radiomics features of primary tumor and LN and embed them into a unified framework. Two-step feature selection method with an upper bound on the total cost used to access groups of features could remove redundant information and retain relevant information between multi-view radiomics features.

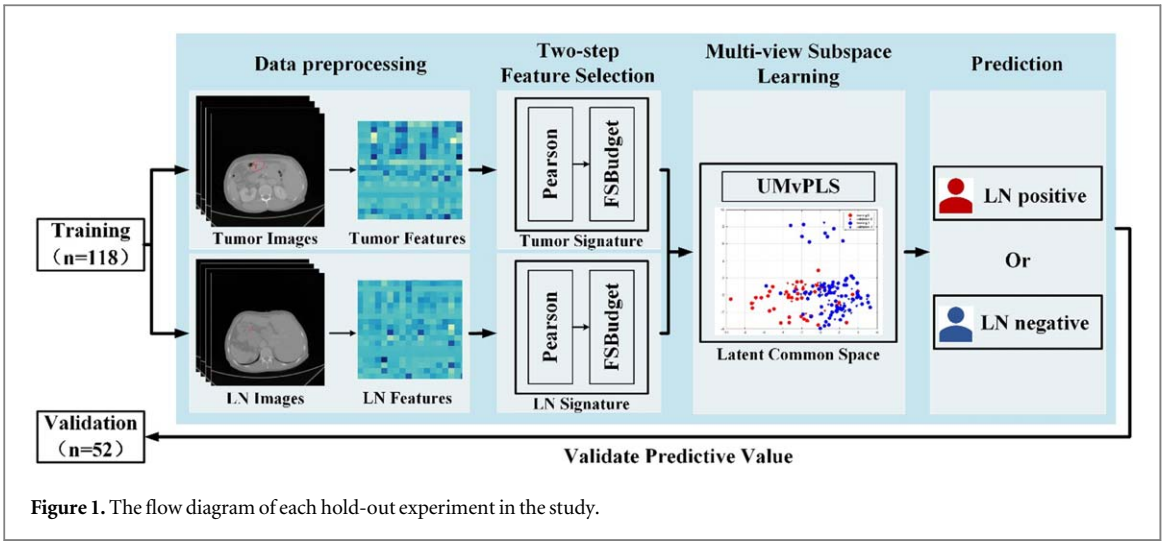

Figure 1. The flow diagram of each hold-out experiment in the study.

- For the preoperative prediction of LNM in GC, our multi-view learning model could provide advantages compared to processing the two branches separately. More useful information will be mined and the heterogeneity could be well represented. As a noninvasive prediction tool, it has the potential to be applied to other diseases to solve different clinical tasks.

3. Materials and methods

3.1. Patients

This study was approved by the Institutional Review Board of the Affiliated Hospital of Qingdao University (Qingdao, China). The requirement for informed consent was waived. One hundred and seventy GC patients who were treated surgically and confirmed pathologically at the hospital between May 2016 and April 2019 were enrolled in this retrospective study. There are 113 LN positive (GC patients with LNM) and 57 LN negative (GC patients without LNM) patients with contrast-enhanced abdominal CT images. The inclusion criteria are as follows: (1) the imaging examination performs <2 weeks before surgical resection. (2) GC diagnosis is histologically confirmed. (3) Lymphadenectomy is performed. (4) There are complete contrast-enhanced abdominal CT images, clinicopathologic characteristics, and tumor marker examination. (5) No combined malignant neoplasm, no distant metastasis, no preoperative neoadjuvant chemotherapy or radiotherapy are permitted. (6) The data quality is satisfactory for analysis. The exclusion criteria are as follows: (1) patients undergo preoperative treatments such as chemotherapy or radiotherapy. (2) Patients are diagnosed with advanced GC which has spread to other parts of the body (such as liver, lung, brain and bone). (3) No surgical resection or complete clinicopathologic characteristics. (4) Unclear CT images or imperfect preparation before CT. CT scanning protocol of all patients were derived. More details can be found in our previous study using this dataset for the first time (Yang *et al* 2020).

3.2. Study flow diagram

Five repeated random hold-out experiments were conducted in this study. In each hold-out experiment, patients are randomly drawn from the dataset at a close to 7:3 ratio to form training cohort and validation cohort. The flow diagram of the study for each hold-out experiment is shown in figure 1. The pipeline of this study includes four main steps: data preprocessing, two-step feature selection, multi-view subspace learning, and prediction. Data preprocessing mainly includes delineation of tumor volume of interest (VOI) and LN VOI, interpolation, and feature extraction. Senior radiologists utilized an open-source imaging platform (ITK-SNAP, version 3.6.0; www.itksnap.org) to delineate tumor VOI and LN VOI on CT scans. After interpolating into isotropic voxel spacing of $1 \times 1 \times 1\ \mathrm{mm}^3$, radiomics features of tumor VOI and LN VOI were extracted using PyRadiomics (van Griethuysen *et al* 2017) (more details shown in the previous work (Yang *et al* 2020)). Then two-step feature selection is performed in the training cohort. Redundant tumor radiomics features and LN radiomics features are removed using Pearson correlation analysis. Discriminative features are selected using a supervised feature selection method to obtain tumor signature and LN signature. Next, a multi-view subspace learning method takes patients with these two signatures and learns projection matrices for a latent common space on which logistic regression classifier is trained. Finally, the model is tested in the corresponding validation cohort.

3.3. Two-step feature selection

Before feature selection, all tumor and LN VOI features were normalized using the Z-score normalization so that all radiomics features have the same scale with zero mean and standard deviation via the formula

$$\hat{x}_{i,s} = \frac{x_{i,s} - \mu_s}{\sigma_s}, \tag{1}$$

where $x_{i,s}$ is the sth feature value of the ith patient, μ_s is the mean of the sth feature of all patients in the training cohort, σ_s is the deviation of the sth feature of all patients in the training cohort, and $\hat{x}_{i,s}$ is the normalized sth featrure value of the ith patient. The same formula is applied to the sth feature value of patients in the validation cohort with the fixed μ_s and σ_s obtained on the training cohort. In the following steps of the experiment, the normalized tumor and LN radiomics features were used unless specific statements have been provided.

In the training cohort, we applied the two-step feature selection approach to both tumor VOI features and LN VOI features to remove redundance for each type of inputs. Our goal is to select minimal number of features that are discriminative without degradation of classification performance. These selected features are considered as the radiomics signature for tumor and LN, respectively. To reach this goal, high correlated tumor VOI features were first filtered out using Pearson correlation analysis (Meng *et al* 2019, Wang *et al* 2019). Then a supervised feature selection method based on test-time budget (namely FSBudget) was designed to select tumor radiomics signature for the differentiation between LN positive and LN negative patients. The same process was conducted for LN VOI features and LN radiomics signature was also obtained.

Pearson correlation coefficient can be calculated for each pair of two radiomics features over either tumor or LN VOI features. The larger the coefficient is, the higher the positive correlation of the two features will be. Features with coefficients larger than a certain threshold are considered as redundant features and will be eliminated. Pearson correlation analysis is an unsupervised learning approach, so it cannot differentiate discriminative and non-discriminative features.

To select discriminative features, we took the FSBudget algorithm which learns a linear predictor by introducing binary indicator variables for selecting groups of features with an explicit budget constraint to upbound the total cost when the cost is available for each group. In our previous research (Wang *et al* 2018), a general and efficient algorithm had been proposed to solve the relaxation problem by leveraging the existing support vector machine (SVM) solvers with various loss functions. The FSBudg*et al*gorithm is a targeted modification on the basis of the previous algorithm to focus on medical radiomics research. Mathematically, FSBudget solves the following optimization problem

$$\min_{\theta,w,b,f} \frac{1}{2}\sum_{s=1}^{d}\frac{w_s^2}{\theta_s} + C\sum_{i=1}^{n} L(f_i, y_i): f_i = \sum_{s=1}^{d} w_s x_{i,s} + b, \sum_{s=1}^{d} c_s\theta_s \leqslant B, \theta_s \in \{0, 1\}, \forall\, s, i, \tag{2}$$

where $(x_i, y_i)_{i=1}^{n}$ is the training cohort with n patients and d features, x_i is the feature vector of the ith patient with the sth feature $x_{i,s}$, $y_i = 1$ for metastasis and 0 for non-metastasis, θ is an indicator vector with entry as either 0 or 1 for not selected or selected, respectively. In this study, each feature is treated as a group with uniform cost 1, that is $c_s = 1$. The total budget B becomes the expected number of features to be selected. As FSBudget is formulated with a general loss function L, it is easy to be adapted for different learning problems such as classification and regression. Our study aims to find discriminative features for binary classification, so the hinge loss is used. Given a training cohort, five-fold cross validation is conducted to select the optimal hyper-parameter C in the grid $\{0.01, 0.1, 1, 10, 100\}$ with respect to the classification accuracy.

3.4. Multi-view subspace learning and prediction

Our proposed unsupervised multi-view partial least squares (UMvPLS) (Wang and Li 2020) was utilized to learn one function to model tumor signature and LN signature. Without loss of generality, let X_{LN} be the view LN with LN signature and X_{TU} be the view of tumor with tumor signature. Each column represents one patient in the training cohort. PLS aims to solve the following optimization problem

$$\max_{P_{LN},P_{TU}} \text{trace}(P_{LN}^T X_{LN} X_{TU}^T P_{TU}): \text{s.t.} P_{LN}^T P_{LN} = P_{TU}^T P_{TU} = I_k, \tag{3}$$

where P_{LN}, P_{TU} are the projection matrix from the number of features to the dimension k of the latent common space. Both inputs need to be centered so that the mean of n columns in X_{LN} and X_{TU} are zeros. To solve problem (3), the UMvPLS algorithm is used since it can guarantee the orthogonality constraints, and simultaneously maximize the covariance in the common space by relying on proven numerical linear algebra techniques, while existing methods often encounter numerical instabilities and offer no orthogonality guarantee on view-specific projection matrices. Orthogonal projections not only possess the nice property of metric preservation, but also provide a natural representation for data visualization similar to principal component analysis (PCA), but on two views.

In training phrase, we solve (3) using UMvPLS for the optimal P_{LN} and P_{TU}. For any given patient, the LN and tumor signatures (x_{LN} and x_{TU}, respectively) are projected to the common latent space via the following formulas:

$$z_{LN} = P_{LN}^T x_{LN}, \tag{4}$$

$$z_{TU} = P_{TU}^T x_{TU}. \tag{5}$$

Although two signatures are located in different feature spaces, both projected points reside in the same space. For classification performance, the concatenation of two projected points (i.e. $z = [z_{LN}; z_{TU}]$) often shows better performance, while the average value of two projected points (i.e. $z = (z_{LN} + z_{TU})/2$) can be used for data visualization. The logistic regression classifier is built on the common latent space with the concatenation of two views using the training cohort.

3.5. Performance evaluation

In the validation phrase, both tumor signature and LN signature of the validation cohort were first projected to the same latent common space via the learned projection matrices, and then the prediction was made by the learned classifier. Finally, the performances on the training and validation cohorts were evaluated. The discrimination performance was quantified by receiver operating characteristic (ROC) curve and area under the curve (AUC). Various performance metrics including precision, accuracy, recall and F1-score, were also used to measure the quality of the learned binary classifier. Here the evaluation indices are given by

$$\text{Precision} = \frac{TP}{TP + FP}, \tag{6}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \tag{7}$$

$$\text{Recall} = \frac{TP}{TP + FN}, \tag{8}$$

$$\text{F1} = \frac{2 \times TP}{2 \times TP + FP + FN}, \tag{9}$$

where TP is the number of the true positive, TN is the number of the true negative, FP is the number of the false positive and FN is the number of the false negative.

4. Results

4.1. Dataset

A total of 170 patients were enrolled in this retrospective study, including 112 men (mean age, 61.61 years; age range, 33–90 years) and 58 women (mean age, 58.76 years; age range, 37–80 years). In each experiment, there were 118 patients in the training cohort (consisting of 79 LN positive and 39 LN negative patients) and 52 patients in the validation cohort (consisting of 34 LN positive and 18 LN negative patients). After data preprocessing and Z-score normalization, 1561 tumor radiomics features and 833 LN radiomics features were obtained for each patient. These radiomics features mainly include three classes: shape, first order statistics and texture features. Texture features contain gray level cooccurence matrix (GLCM), gray level run length matrix (GLRLM), gray level size zone matrix (GLSZM), neigbouring gray tone difference matrix (NGTDM), and gray level dependence matrix (GLDM) features.

4.2. Two-step feature selection

After removing redundant radiomics features, more than 200 tumor radiomics features (mean number, 255.2; number range, 246–273) and more than 100 LN radiomics features (mean number, 190.2; number range, 187–194) were retained using Pearson correlation analysis with threshold 0.9. The numbers of these radiomics features are shown in tables S1 and S2 (available online at stacks.iop.org/PMB/67/055007/mmedia). FSBudget with budget 40 was applied to these tumor and LN radiomics features to obtain ~40 optimal tumor radiomics features and ~40 optimal LN radiomics features as tumor signature and LN signature individually. Details of the tumor signature and the LN signature are given in tables S3 and S4. Note that FSBudget does not return exactly 40 tumor radiomics features (mean number, 40.6; number range, 40–41) or LN radiomics features (mean number, 40.8; number range, 40–42), but very close to the budget cost. Further observation of each feature class indicates that there are great differences at five hold-out experiments. Tumor-only model or LN-only model with logistic regression classifier was constructed using the radiomics signature of tumor or LN. In the training cohort, the overall AUC values (reported as mean $\pm$ standard deviation), are 0.8522 ± 0.0242 and 0.9109 ± 0.0317 for the tumor-only and LN-only models respectively. The overall AUC values of the validation

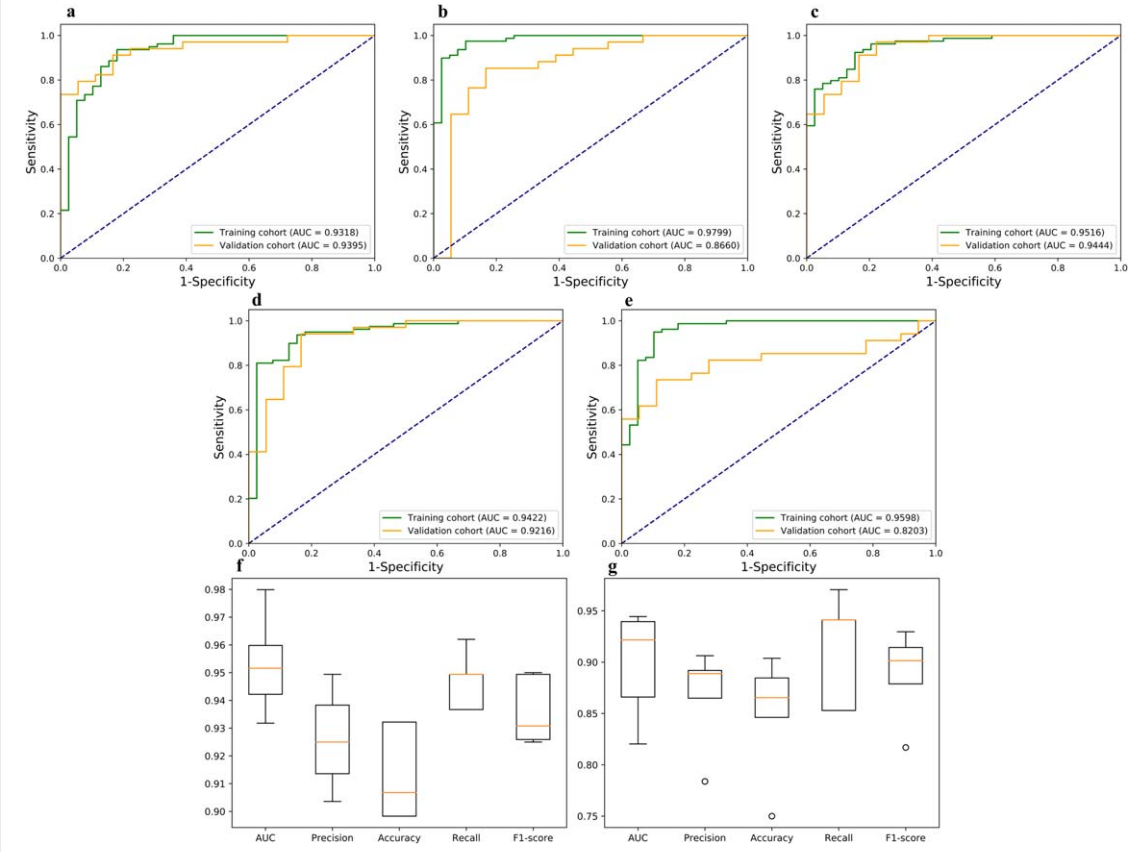

Figure 2. Performances of the UMvPLSR_20 model by five repeated random hold-out experiments. (a) to (e) are the ROC curves of Experiment 1 to Experiment 5. (f) and (g) are the boxplots of five performance metrics in the training and validation cohorts.

cohort in the tumor-only and LN-only models are 0.8127 ± 0.0941 and 0.8121 ± 0.0291, respectively. Detailed performance of the two single-view radiomics models is shown in tables S5 and S6.

4.3. Multi-view subspace learning

In this section, UMvPLS was used to explore the complementary information between LN and tumor signatures by projecting them into a latent common space. The UMvPLS radiomics model with the dimension of the latent common space set to 20 is code-named UMvPLSR_20. The space dimension was determined by our previous study (Yang *et al* 2020). ROC curves and five performance metrics of the UMvPLSR_20 model are described in figure 2 (more details are shown in table S7). The overall AUC values of the training cohorts and the validation cohorts are 0.9531 ± 0.0183 and 0.8984 ± 0.0536 respectively. The other four performance metrics also show the good predictive power for discriminating patients with and without LNM in patients of GC (training cohort: precision, 0.9260 ± 0.0184; accuracy, 0.9136 ± 0.0174; recall, 0.9468 ± 0.0106; F1-score, 0.9362 ± 0.0125; validation cohort: precision, 0.8671 ± 0.0489; accuracy, 0.8500 ± 0.0599; recall, 0.9118 ± 0.0550; F1-score, 0.8882 ± 0.0440).

We further reduced the dimension of the latent common space from 20 to 2, in order to demonstrate the capability of our proposed method for data visualization. The proposed method for an intuitive 2-dimensional latent common space is code-named UMvPLSR_2. We show the projected points of UMvPLSR_2 with both training and validation cohorts in figure 3. In five hold-out experiments, training-0 and training-1 mean LN negative and LN positive patients in training cohort, respectively. Similarly, validation-0, and validation-1 mean LN negative and LN positive patients in validation cohort, respectively. UMvPLSR_2's ROC curves by five repeated random hold-out experiments are shown in figures 4(a)–(e), and compared using Delong test in table S10. There is no significant difference in ROC curves between any two experiments (all P values > 0.05). Five performance metrics of the UMvPLSR_2 model in the training and validation cohorts are described in figures 4(f) and (g). The overall AUC, precision, accuracy, recall and F1-score are 0.8593 ± 0.0226, 0.8583 ± 0.0320, 0.8220 ± 0.0247, 0.8810 ± 0.0212 and 0.8691 ± 0.0163 for the training cohort respectively. As for the validation cohorts, AUC of 0.8660 ± 0.0573, precision of 0.8615 ± 0.0605, accuracy of 0.8308 ± 0.0534, recall of 0.8882 ± 0.0483 and F1-score of 0.8732 ± 0.0376 are obtained in five experiments.

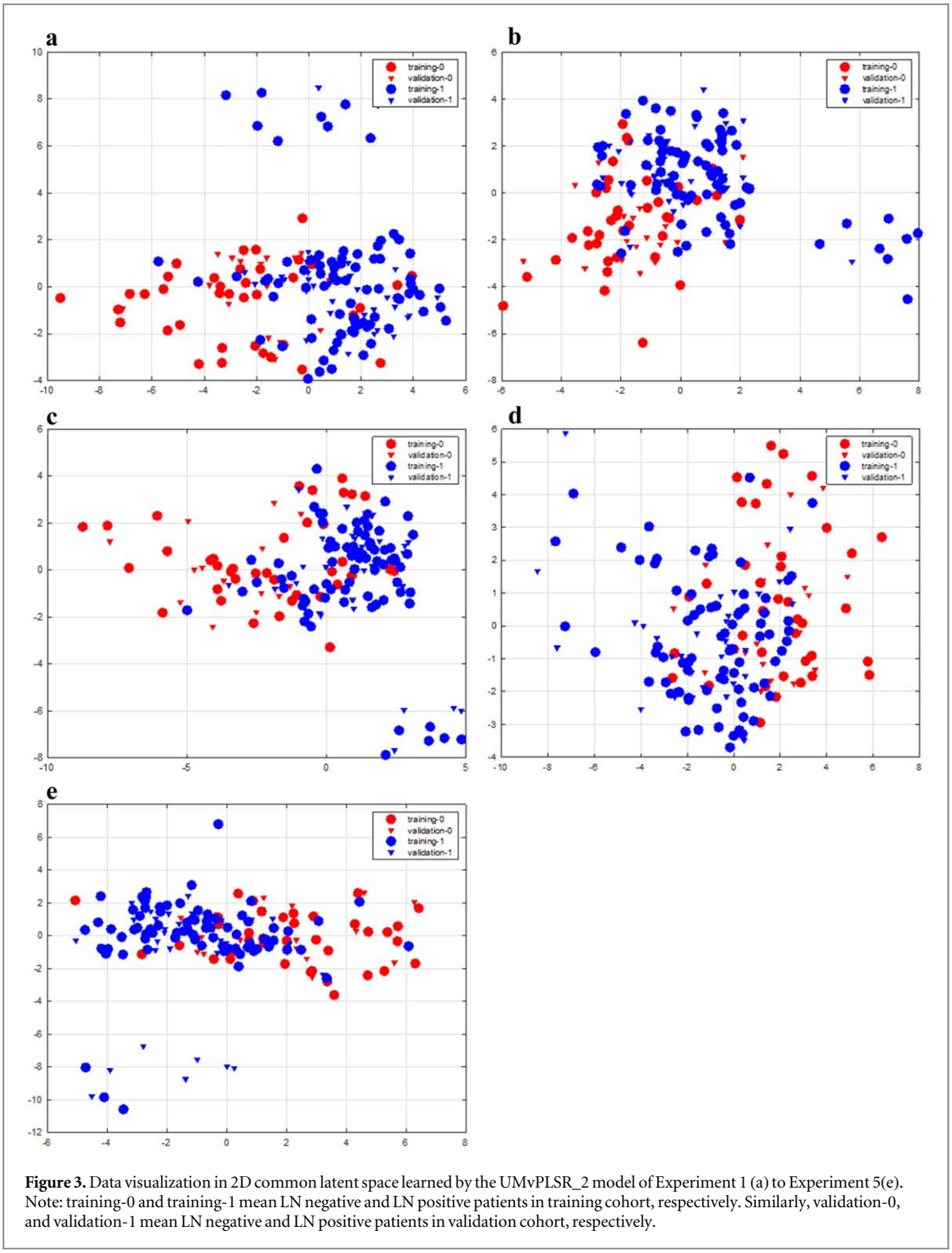

Figure 3. Data visualization in 2D common latent space learned by the UMvPLSR_2 model of Experiment 1 (a) to Experiment 5(e). Note: training-0 and training-1 mean LN negative and LN positive patients in training cohort, respectively. Similarly, validation-0, and validation-1 mean LN negative and LN positive patients in validation cohort, respectively.

More details are provided in table S8. Results demonstrate our multi-view radiomics model, whether in 20-dimensional space or 2-dimensional space, has good predictive power for LNM on GC patients.

4.4. Performance comparison

4.4.1. Multi-view method versus single-view method

Compared to the tumor-only model and the LN-only model, the UMvPLSR_20 model' AUC is better in each experiment except for the validation cohort of Experiment 3 (table 1). In five validation cohorts, the overall precision, accuracy, recall and F1-score increases by 4%, 13%, 17%, and 10% for the tumor-only model respectively, and 10%, 13%, 8%, and 9% for the LN-only model respectively (tables S5–S7). All performance metrics of the UMvPLSR_20 model are globally better than those of single-view methods, which designating the

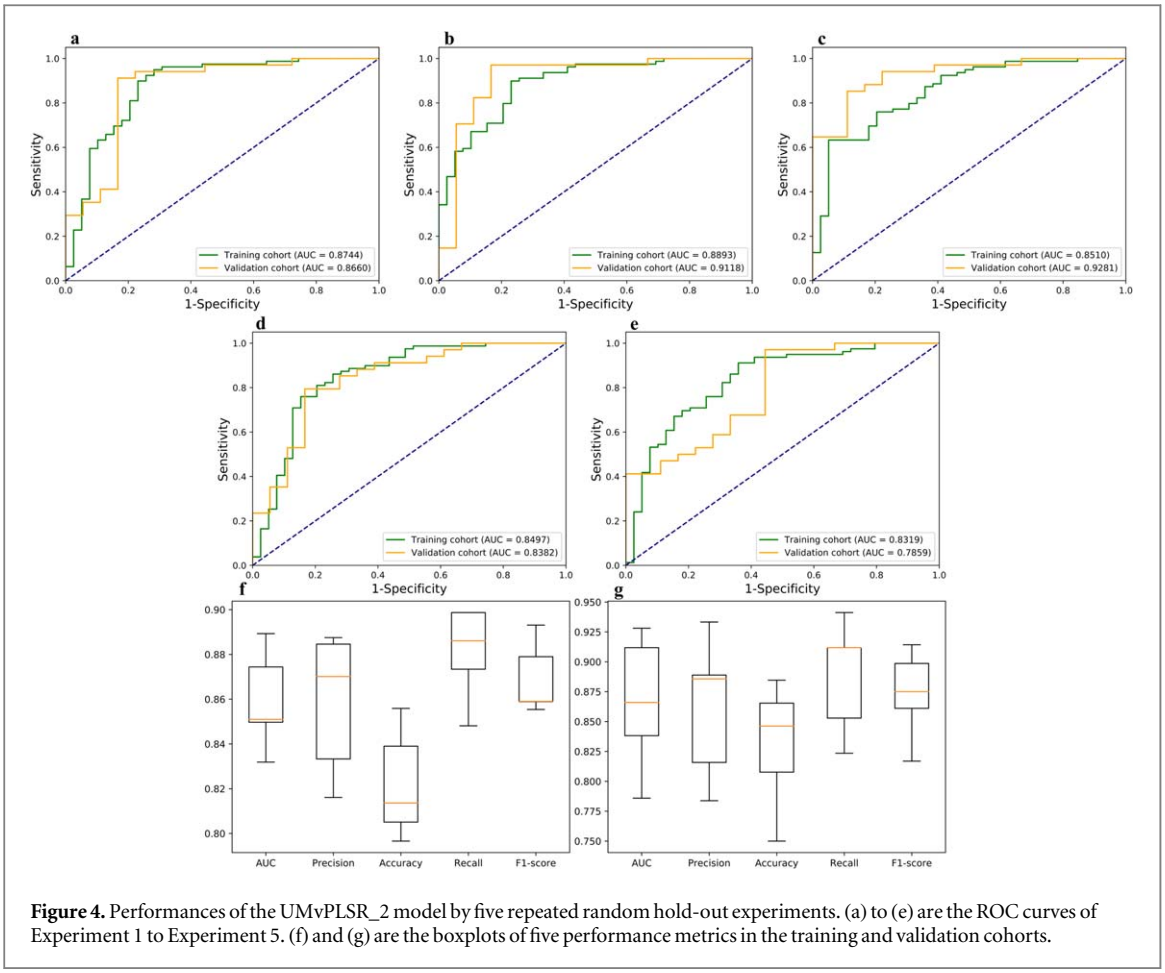

Figure 4. Performances of the UMvPLSR_2 model by five repeated random hold-out experiments. (a) to (e) are the ROC curves of Experiment 1 to Experiment 5. (f) and (g) are the boxplots of five performance metrics in the training and validation cohorts.

Table 1. The AUC values of the tumor-only model, the LN-only model, the UMvPLSR_20 model, and the UMvPLSR_2 model by five repeated random hold-out experiments.

Model	1	2	3	4	5	Mean	Std
			Training cohort				
Tumor-only	0.8588	0.8711	0.8121	0.8488	0.8702	0.8522	0.0242
LN-only	0.8952	0.9588	0.9228	0.8750	0.9030	0.9109	0.0317
UMvPLSR_20	**0.9318**	**0.9799**	**0.9516**	**0.9422**	**0.9598**	**0.9531**	0.0183
UMvPLSR_2	0.8744	0.8893	0.8510	0.8497	0.8319	0.8593	0.0226
			Validation cohort				
Tumor-only	0.7582	0.8284	**0.9542**	0.8203	0.7026	0.8127	0.0941
LN-only	0.8431	0.8105	0.7859	0.8399	0.7810	0.8121	0.0291
UMvPLSR_20	**0.9395**	0.8660	0.9444	**0.9216**	**0.8203**	**0.8984**	0.0536
UMvPLSR_2	0.8660	**0.9118**	0.9281	0.8382	0.7859	0.8660	0.0573

Std is the abbreviation of standard deviation.

The highest AUC value of each experiment in the training or validation cohort is in a bold.

performance on concatenated features outperforms that of individual view. It implies that LN and tumor could contain complementary information for the LNM prediction in GC.

For UMvPLSR_2, it is clear that 2D space loses certain information in comparison with the UMvPLSR_20 model since UMvPLSR_2's overall performance decreases by 3% at the worst. However, the average value of each performance metric in UMvPLSR_2 is still higher than that of the tumor-only model and the LN-only model at five validation cohorts (tables S5, S6, S8). Even if the dimension of the latent common space is 2, the classification performance does not degrade too much, and the visualization results demonstrate that two classes can be separated by a linear classifier with decent performance. In short, the proposed multi-view radiomics strategy provide advantages compared to single-view radiomics strategy in our validation cohorts.

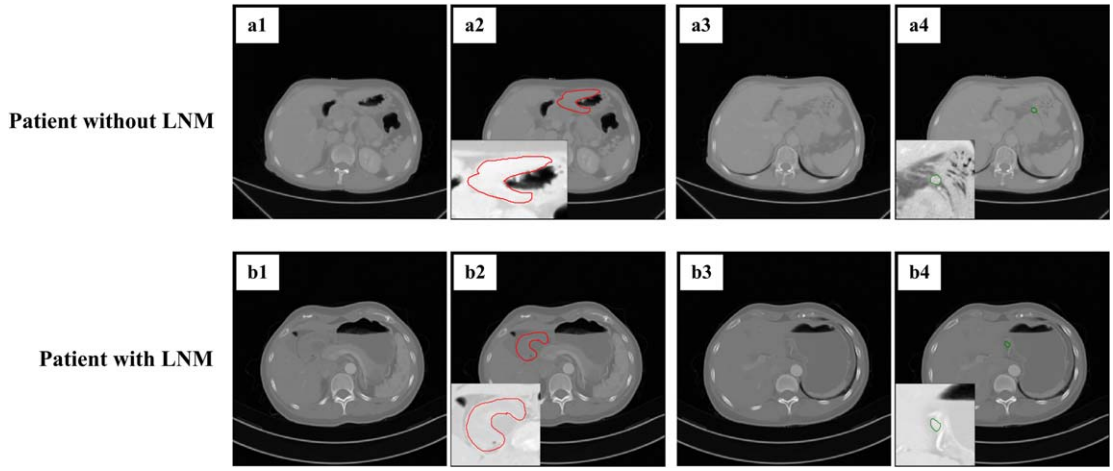

Figure 5. Examples of clinical images. (a1)–(a4) are from a 58 year old female patient without LNM. The TNM staging of this patient is T3N0M0. (b1)–(b4) are from a 47 year old female patient with LNM. Her TNM staging is T3N1M0. Figures with the number 1 and 2 are the middle slice of the tumor on transverse section (red line is the contour of tumor). Figures with the number 3 and 4 are the middle slice of the largest LN visible on transverse section (green line is the contour of LN).

Table 2. The AUC values of the UMvPLSR_20 model, the UMvPLSR_2 model and the previously proposed radiomics signature model by five repeated random hold-out experiments.

Model	1	2	3	4	5	Mean	Std
			Training cohort				
UMvPLSR_20	**0.9318**	**0.9799**	**0.9516**	**0.9422**	**0.9598**	**0.9531**	0.0183
UMvPLSR_2	0.8744	0.8893	0.8510	0.8497	0.8319	0.8593	0.0226
Radiomics signature [(Yang *et al* 2020)]	0.9296	0.9523	0.9400	0.9195	0.9179	0.9319	0.0129
			Validation cohort				
UMvPLSR_20	**0.9395**	0.8660	**0.9444**	**0.9216**	0.8203	**0.8984**	0.0536
UMvPLSR_2	0.8660	**0.9118**	0.9281	0.8382	0.7859	0.8660	0.0573
Radiomics signature [(Yang *et al* 2020)]	0.8742	0.8627	0.8088	0.8824	**0.8448**	0.8546	0.0261

Std is the abbreviation of standard deviation.

The highest AUC value of each experiment in the training or validation cohort is in bold.

4.4.2. Our method versus previously proposed method

The performance of the UMvPLSR_20 model was then compared with that of the previously proposed radiomics model (Yang *et al* 2020). The previous radiomics signature was obtained by a four-step feature selection method on the same dataset as ours. The radiomics signature model was constructed by the radiomics signature based on logistic regression classifier (Yang *et al* 2020). Table 2 manifests that the average performance of the UMvPLSR_20 model (training cohort, 0.9531; validation cohort, 0.8984) is superior to that of our previous model (training cohort, 0.9319; validation cohort, 0.8546). For a comprehensive comparison, we then focused on the performance of the training and validation cohorts in each experiment. In addition to the validation cohort of Experiment 5, the AUC value of the UMvPLSR_20 model outperforms that of the previous radiomics model in all training and validation cohorts. As for the UMvPLSR_2 model, the average AUC value of validation cohorts slightly surpasses that of the previous model (UMvPLSR_2 model, 0.8660; previous radiomics model, 0.8546). Taken together, our proposed method shows better predictive power than the previously proposed method with a simple fusion strategy for discriminating patients with or without LNM in GC.

4.4.3. Our method versus other baseline methods

In this section, we compared our proposed UMvPLSR_20 method with other baseline methods using eight classifiers including Decision Tree (Breiman *et al* 2017), RF (Breiman 2001), Adaboost (Freund and Schapire 1997), Gradient Boosting (Friedman 2001), Naïve Bayes (Zhang 2004), SVM-Recursive Feature Elimination (SVM-RFE) (Guyon *et al* 2002), L1-Logistic regression (Fan *et al* 2008) and Multi-layer Perceptron (MLP) (Hinton 1990) in terms of three different settings: tumor-only, LN-only and the combination of tumor and LN radiomics features. As eight baseline methods do not take two set of features as input, so the

concatenation of two sets of features is used for comparisons. The validation results are reported in terms of five metrics (table 3). Results show that our method generally outperform others except MLP works the best for precision (5% higher than ours). The performance of the UMvPLSR_20 model and eight baseline methods on tumor-only or LN-only radiomics signatures are shown in tables S11 and S12. Tables S11 and S12 are results with our single-view data, so their performance will not be that important even though either does not show the best results, but our UMvPLSR_20 model still outperform their concatenation of two sets of features as well as the single-view baselines.

5. Discussion

Radiomics features of primary tumors and LN show positive role in predicting LNM in GC, yet only a simple combination of the two has been used in previous research. In this study, we developed and validated a multi-view learning radiomics method to discriminate between LN positive and LN negative on GC patients. To the best of our knowledge, this is the first time that a multi-view learning technology to study patients with GC.

The scope and number of LNM determine the extent of LN dissection in surgery, and directly affect the prognosis in patients with GC (Forman and Burley 2006, Cho *et al* 2007, Lum *et al* 2020). LNM occurs with the progression of primary tumor, which is a complex and continuous process (figure 5). Only relying on primary tumor radiomics without considering LN radiomics may not thoroughly discriminate patients with or without LNM. It urges that radiomics features extracted from LN is coming into our view. We found that both primary tumor radiomics and LN radiomics have the predictive power for differentiating GC patients with or without LNM, which is consistent with previous studies (Yang *et al* 2020, Wang *et al* 2021) (table 1). It is worth noting that the AUC value of the LN-only model is greater than that of the tumor-only model in three out of five validation cohorts. Our findings point to the potential benefit of imaging characteristics of LN, showing the traditional strategy solely based on imaging characteristics of primary tumors is not comprehensive for the LNM prediction in GC.

Since multiple views of radiomics features are jointly considered when performing classification prediction in medical tasks, the multi-view learning strategy has been involved in recent researches. Puyol-Anton *et al* proposed two multi-view machine learning algorithms, using a multimodal cardiac motion atlas from 3-dimensional magnetic resonance and 3-dimensional ultrasound data, for the classification of dilated cardiomyopathy patients. The multi-view Laplacian support vector machines algorithm achieved the best performance (global approach: 92.71%; regional approach: 94.32%), illustrating the automated diagnostic pipeline is an important aid in the quantification of the contractility and function of the left ventricular myocardium (Puyol-Anton *et al* 2019). Thammasorn *et al* developed a novel nearest-neighbor validation strategy based on small-sample medical imaging data to optimize multi-view triplet network for classification. The strategy was superior to other common deep representation learning baselines in dealing with several medical tasks, such as radiation therapy delivery error prediction and sarcoma survival prediction (Thammasorn *et al* 2021). These studies show great potentials of multi-view learning technologies in classification of clinical tasks. In this study, we explored the application of a multi-view learning technique to the prediction of LN status on GC patients.

The major contribution of our work is to exploit the complementarity of different views and seek the consensus among them to fuse multi-view data on a feature level. After data preprocessing, all tumor and LN radiomics features were selected by two-step feature selection methods including Pearson correlation analysis and FSBudg*et al*gorithm, respectively. Comparing to sequential forward floating selection (SFFS) algorithm (Pudil *et al* 1994) used in our previous study (Yang *et al* 2020), FSBudget has the following advantages: (i) FSBudget is a supervised feature selection approach with a global objective, but SFFS is a greedy approach; (ii) FSBudget can select varying number of features with a proper budget B, while SFFS often gets stuck for selecting moderate number of features due to its inclusive and exclusive operations with no convergence guarantee; (iii) FSBudget is scalable for high-dimensional data, while SFFS can be extremely slow on a large number of features. Then UMvPLS, one of multi-view subspace learning methods, was utilized to fuse the information of tumor radiomics and LN radiomics to disambiguate the result made by one single view. Among various multi-view learning methods, multi-view subspace learning is a widely studied approach that learns a latent common space such that the projected data points of each view preserve certain properties. For example, canonical correlation analysis (CCA) (Hardoon *et al* 2004) aims to maximize the correlation of two views and partial least squares (PLS) (Wold *et al* 1984) aims to find the maximum covariance with orthogonal projection matrices. In this study, we have two considerations: (i) the cross-view covariance/correlation is important to explore the complementary information provided by each other; (ii) data visualization is important for exploratory analysis in medical domain. With the two considerations, our recently proposed UMvPLS algorithm based on PLS was chosen for our subspace learning since it maximizes the cross-covariance and learns orthogonal projection

Table 3. The performance of the UMvPLSR_20 model and eight baseline methods on the combination of tumor and LN radiomics by five repeated random hold-out experiments.

Method	1	2	3	4	5	Mean	Std
			AUC				
RandomForest	0.9281	**0.9069**	0.8611	0.8922	0.8562	0.8889	0.0305
DecisionTree	0.6250	0.7941	0.7606	0.7230	0.7459	0.7297	0.0640
Adaboost	0.9134	0.7288	0.8627	0.7794	0.8219	0.8212	0.0716
GradientBoosting	0.8987	0.9036	0.8660	0.8578	**0.8848**	0.8822	0.0199
NaiveBayes	0.7876	0.6993	0.7647	0.7467	0.6716	0.7340	0.0476
SVM-RFE	0.8922	0.7345	0.9101	0.7516	0.6569	0.7891	0.1086
L1-LogisticRegression	0.9346	0.7533	0.8333	0.8660	0.8203	0.8415	0.0663
MLP	0.4779	0.4975	0.5613	0.5123	0.4575	0.5013	0.0393
UMvPLSR_20	**0.9395**	0.8660	**0.9444**	**0.9216**	0.8203	**0.8984**	0.0536
			Precision				
RandomForest	0.7857	0.8378	0.8378	0.8250	0.7778	0.8128	0.0290
DecisionTree	0.7097	0.8611	0.8485	0.8182	0.7879	0.8051	0.0604
Adaboost	0.8421	0.7143	0.7632	0.8000	0.7442	0.7727	0.0497
GradientBoosting	0.8378	0.8378	0.8485	0.8158	0.8205	0.8321	0.0136
NaiveBayes	0.8462	0.8000	0.8333	0.7895	0.7368	0.8012	0.0428
SVM-RFE	0.8611	0.8182	0.8378	0.7879	0.7241	0.8058	0.0530
L1-LogisticRegression	0.8611	0.8276	0.8438	0.8000	**0.8286**	0.8322	0.0226
MLP	0.6458	0.6538	0.6400	0.6400	0.6538	0.6467	0.0069
UMvPLSR_20	**0.8649**	**0.9063**	**0.8919**	**0.8889**	0.7838	**0.8671**	0.0489
			Accuracy				
RandomForest	0.8077	0.8269	0.8269	0.8462	0.7308	0.8077	0.0451
DecisionTree	0.5962	**0.8462**	0.7885	0.7500	0.7115	0.7385	0.0938
Adaboost	0.8462	0.6346	0.7308	0.8077	0.7500	0.7538	0.0809
GradientBoosting	0.8269	0.8269	0.7885	0.8077	**0.8269**	0.8154	0.0172
NaiveBayes	0.5192	0.4615	0.5000	0.5577	0.5192	0.5115	0.0349
SVM-RFE	0.8462	0.7500	0.8269	0.7115	0.5962	0.7462	0.1003
L1-LogisticRegression	0.8462	0.7115	0.7692	0.7500	0.7885	0.7731	0.0498
MLP	0.6154	0.6538	0.6154	0.6154	0.6538	0.6308	0.0211
UMvPLSR_20	**0.8654**	**0.8462**	**0.9038**	**0.8846**	0.7500	**0.8500**	0.0599
			Recall				
RandomForest	**0.9706**	0.9118	0.9118	**0.9706**	0.8235	0.9176	0.0603
DecisionTree	0.6471	0.9118	0.8235	0.7941	0.7647	0.7882	0.0962
Adaboost	0.9412	0.7353	0.8529	0.9412	0.9412	0.8824	0.0907
GradientBoosting	0.9118	0.9118	0.8235	0.9118	0.9412	0.9000	0.0446
NaiveBayes	0.3235	0.2353	0.2941	0.4412	0.4118	0.3412	0.0847
SVM-RFE	0.9118	0.7941	0.9118	0.7647	0.6176	0.8000	0.1220
L1-LogisticRegression	0.9118	0.7059	0.7941	0.8235	0.8529	0.8176	0.0761
MLP	0.9118	**1.0000**	0.9412	0.9412	**1.0000**	**0.9588**	0.0395
UMvPLSR_20	0.9412	0.8529	**0.9706**	0.9412	0.8529	0.9118	0.0550
			F1-score				
RandomForest	0.8684	0.8732	0.8732	0.8919	0.8000	0.8614	0.0355
DecisionTree	0.6769	0.8857	0.8358	0.8060	0.7761	0.7961	0.0780
Adaboost	0.8889	0.7246	0.8056	0.8649	0.8312	0.8230	0.0635
GradientBoosting	0.8732	0.8732	0.8358	0.8611	**0.8767**	0.8640	0.0168
NaiveBayes	0.4681	0.3636	0.4348	0.5660	0.5283	0.4722	0.0793
SVM-RFE	0.8857	0.8060	0.8732	0.7761	0.6667	0.8015	0.0882
L1-LogisticRegression	0.8857	0.7619	0.8182	0.8116	0.8406	0.8236	0.0451
MLP	0.7561	0.7907	0.7619	0.7619	0.7907	0.7723	0.0170
UMvPLSR_20	**0.9014**	**0.8788**	**0.9296**	**0.9143**	0.8169	**0.8882**	0.0440

Std is the abbreviation of standard deviation.

The highest AUC value of each experiment in the training or validation cohort is in a bold.

matrices for better data visualization, simultaneously. Finally, the UMvPLSR_20 model was built by the logistic regression classifier on the 20-dimensional latent common space.

In the training and validation cohorts, each performance index of the UMvPLSR_20 model exceeds 0.91 and 0.85 respectively (table S7). The overall performance of the UMvPLSR_20 model (table S7) is higher than that of the tumor-only model (table S5) or the LN-only model (table S6). These results show that compared with single-view radiomics model, multi-view inputs have obviously benefits in increasing performance, demonstrating that the advanced features of the UMvPLSR_20 model could provide robust expression of the input images. Meanwhile, table 2 shows the overall performance of UMvPLSR_20 is better in contrast to the previously proposed radiomics method (Yang *et al* 2020). It is obvious that the features selected by our proposed method are more informative, which confirms that the effectiveness of our method to predict the LN status on GC patients. This is probably attributed to its ability to implicitly account for 3D information by fusing tumor view and LN view into a common latent common space rather than simple fusion. We then reduced the dimension of the latent common space to 2 for generating the UMvPLSR_2 model. For UMvPLSR_2, tumor view and LN view are directly comparable and can easily be merged together. It forms a 2-dimensional common space that is supposed to preserve the information from each view, facilitating the visualization of the classification results (figure 3). The overall performance of UMvPLSR_2 has slightly worse than UMvPLSR_20, but it is still greater than two single-view radiomics models and is conducive to doctors' judgment (figure 4 and table S8). In addition, our UMvPLSR_20 method also outperforms other methods with eight commonly used machine learning classifiers (table 3). In general, our proposed multi-view radiomics analysis strategy could significantly improve performance for predicting LNM in GC patients, compared to the traditional radiomics analysis strategy.

Future work will focus on incorporating non-imaging information (such as genomics, metabolic and semantic features) and other modality data (when the LN images are not available), in order to form a more comprehensive multi-view learning strategy. Furthermore, we are collecting more gastric cancer data from other clinical centers. A multi-center gastric cancer database is being developed to provide more clinical decision support. Other abdominal tumor data will also be considered to test the same predictive pipeline for more clinical tasks in the future.

6. Conclusions

In summary, this study proposed a multi-view learning method to learn radiomics features of primary tumor and LN. Compared with single-view radiomics model, the previous radiomics method, and other baseline methods, our predictive pipeline can be more effective for the LNM prediction in GC. It provides medical professionals with an effective noninvasive strategy to pick a choice for GC patients' therapy, and has potential to be applied to other clinical tasks.

Acknowledgments

This work was supported by Supported by Beijing Natural Science Foundation (Z210008), Natural Science Foundation of China (NSFC Grant No. 81871351, 81827804, 81972849). We thank the staff in the Affiliated Hospitals of Qingdao University for their technical assistance.

Conflict of interest

The authors declare no conflicts of interest.

ORCID iDs

Tianye Niu https://orcid.org/0000-0003-4181-3641

References

Breiman L 2001 Random forests *Mach. Learn.* **45** 5–32

Breiman L, Friedman J H, Olshen R A and Stone C J 2017 *Classification and Regression Trees* (New York: Routledge)

Cho B C, Jeung H C, Choi H J, Rha S Y, Hyung W J, Cheong J H, Noh S H and Chung H C 2007 Prognostic impact of resection margin involvement after extended (D2/D3) gastrectomy for advanced gastric cancer: a 15-year expereince at a single institute *J. Surg. Oncol.* **95** 461–8

Fan R-E, Chang K-W, Hsieh C-J, Wang X-R and Lin C-J 2008 LIBLINEAR: a library for large linear classification *J. Mach. Learn. Res.* **9** 1871–4
Feng Q-X, Liu C, Qi L, Sun S-W, Song Y, Yang G, Zhang Y-D and Liu X-S 2019 An Intelligent clinical decision support system for preoperative prediction of lymph node metastasis in gastric cancer *J. Am. Coll. Radiol.* **16** 952–60
Forman D and Burley V J 2006 Gastric cancer: global pattern of the disease and an overview of environmental risk factors *Best Pract. Res. Clin. Gastroenterol.* **20** 633–49
Fratello M, Caiazzo G, Trojsi F, Russo A, Tedeschi G, Tagliaferri R and Esposito F 2017 Multi-view ensemble classification of brain connectivity images for neurodegeneration type discrimination *Neuroinformatics* **15** 199–213
Freund Y and Schapire R E 1997 A decision-theoretic generalization of on-line learning and an application to boosting *J. Comput. Syst. Sci.* **55** 119–39
Friedman J H 2001 Greedy function approximation: a gradient boosting machine *Ann. Stat.* **29** 1189–232
Gao X, Ma T, Cui J, Zhang Y, Wang L, Li H and Ye Z 2020a A CT-based radiomics model for prediction of lymph node metastasis in early stage gastric cancer *Acad. Radiol.* **28** E155-E164
Gao X, Ma T, Cui J, Zhang Y, Wang L, Li H and Ye Z 2020b A radiomics-based model for prediction of lymph node metastasis in gastric cancer *Eur. J. Radiol.* **129** 109069
Gillies R J, Kinahan P E and Hricak H 2016 Radiomics: images are more than pictures, they are data *Radiology* **278** 563–77
van Griethuysen J J M, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, Beets-Tan R G H, Fillion-Robin J-C, Pieper S and Aerts H J W L 2017 Computational radiomics system to decode the radiographic phenotype *Cancer Res.* **77** E104–7
Guyon I, Weston J, Barnhill S and Vapnik V 2002 Gene selection for cancer classification using support vector machines *Mach. Learn.* **46** 389–422
Hardoon D R, Szedmak S and Shawe-Taylor J 2004 Canonical correlation analysis: an overview with application to learning methods *Neural Comput.* **16** 2639–64
Hinton G E 1990 *Machine Learning* (Amsterdam: Elsevier) pp 555–610
Huang J-Y, Xu Y-Y, Li M, Sun Z, Zhu Z, Song Y-X, Miao Z-F, Wu J-H and Xu H-M 2013 The prognostic impact of occult lymph node metastasis in node-negative gastric cancer: a systematic review and meta-analysis *Ann. Surg. Oncol.* **20** 3927–34
Jiang Y *et al* 2019 Radiomics signature on computed tomography imaging: association with lymph node metastasis in patients with gastric cancer *Frontiers Oncol.* **9** 340
Kim A Y, Kim H J and Ha H K 2005a Gastric cancer by multidetector row CT: preoperative staging *Abdominal Imaging* **30** 465–72
Kim H J, Kim A Y, Oh S T, Kim J S, Kim K W, Kim P N, Lee M G and Ha H K 2005b Gastric cancer staging at multi-detector row CT gastrography: comparison of transverse and volumetric CT scanning *Radiology* **236** 879–85
Lambin P *et al* 2012 Radiomics: extracting more information from medical images using advanced feature analysis *Eur. J. Cancer* **48** 441–6
Larue R T H M, Defraene G, De Ruysscher D, Lambin P and Van Elmpt W 2017 Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures *Br. J. Radiol.* **90** 1070
Lee I J, Lee J M, Kim S H, Shin C-I, Lee J Y, Kim S H, Han J K and Choi B I 2010 Diagnostic performance of 64-channel multidetector CT in the evaluation of gastric cancer: differentiation of mucosal cancer (T1a) from submucosal involvement (T1b and T2) *Radiology* **255** 805–14
Lee S H, Han P, Hales R K, Voong K R, Noro K, Sugiyama S, Haller J W, McNutt T R and Lee J 2020 Multi-view radiomics and dosiomics analysis with machine learning for predicting acute-phase weight loss in lung cancer patients treated with radiotherapy *Phys. Med. Biol.* **65** 19
Li C, Xu J, Liu Q, Zhou Y, Mou L, Pu Z, Xia Y, Zheng H and Wang S 2020 Multi-view mammographic density classification by dilated and attention-guided residual learning *IEEE/ACM Trans. Comput. Biol. Bioinf.* **18** 1003–13
Liang S, Kim-Han T, Nie D, Zhang Y and Shen D 2020 Multi-view spatial aggregation framework for joint localization and segmentation of organs at risk in head and neck CT images *IEEE Trans. Med. Imaging* **39** 2794–805
Lin J-X *et al* 2019 Risk factors of lymph node metastasis or lymphovascular invasion for early gastric cancer: a practical and effective predictive model based on international multicenter data *Bmc Cancer* **19** 1
Lum C Y, Huang K-H, Chen M-H, Fang W-L, Chao Y, Lo S-S, Li A F-Y, Wu C-W and Shyr Y-M 2020 The clinicopathological characteristics and prognosis of patients with node-positive gastric cancer after curative surgery *J. Chin. Med. Assoc.* **83** 751–5
Meng X *et al* 2019 Preoperative radiomic signature based on multiparametric magnetic resonance imaging for noninvasive evaluation of biological characteristics in rectal cancer *Eur. Radiol.* **29** 3200–9
Pudil P, Novovicova J and Kittler J 1994 Floating search methods in feature-selection *Pattern Recognit. Lett.* **15** 1119–25
Puyol-Anton E, Ruijsink B, Gerber B, Amzulescu M S, Langet H, De Craene M, Schnabel J A, Piro P and King A P 2019 Regional multi-view learning for cardiac motion analysis: application to identification of dilated cardiomyopathy patients *IEEE Trans. Biomed. Eng.* **66** 956–66
Sada Y H, Smaglo B G, Tan J C, Cao H S T, Musher B L and Massarweh N N 2019 Prognostic value of nodal response after preoperative treatment of gastric adenocarcinoma *J. Natl Comprehensive Cancer Netw.* **17** 161–8
Shang-Guan X-C *et al* 2018 Preoperative lymph node size is helpful to predict the prognosis of patients with stage III gastric cancer after radical resection *Surg. Oncol.* **27** 54–60
Sun S 2013 A survey of multi-view machine learning *Neural Comput. Appl.* **23** 2031–8
Sung H, Ferlay J, Siegel R L, Laversanne M, Soerjomataram I, Jemal A and Bray F 2021 Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries *CA—Cancer J. Clin.* **71** 209–49
Thammasorn P, Chaovalitwongse W A, Hippe D S, Wootton L S, Ford E C, Spraker M B, Combs S E, Peeken J C and Nyflot M J 2021 Nearest neighbor-based strategy to optimize multi-view triplet network for classification of small-sample medical imaging data *IEEE Trans. Neural Netw. Learn. Syst.* **0** 1–15
Traverso A, Wee L, Dekker A and Gillies R 2018 Repeatability and reproducibility of radiomic features: a systematic review *Int. J. Radiat. Oncol. Biol. Phys.* **102** 1143–58
Wang L and Li R-C 2020 A scalable algorithm for large-scale unsupervised multi-view partial least squares *IEEE Trans. Big Data* **0** 1-11
Wang L, Zhu D and Chi Y 2018 Efficient test-time predictor learning with group-based budget *32nd AAAI Conf. on Artificial Intelligence*
Wang X, Li C, Fang M, Zhang L, Zhong L, Dong D, Tian J and Shan X 2021 Integrating No.3 lymph nodes and primary tumor radiomics to predict lymph node metastasis in T1-2 gastric cancer *BMC Med. Imaging* **21** 58
Wang X *et al* 2019 Can peritumoral radiomics increase the efficiency of the prediction for lymph node metastasis in clinical stage T1 lung adenocarcinoma on CT? *Eur. Radiol.* **29** 6049–58
Wang Y, Liu W, Yu Y, Liu J-J, Xue H-D, Qi Y-F, Lei J, Yu J-C and Jin Z-Y 2020 CT radiomics nomogram for the preoperative prediction of lymph node metastasis in gastric cancer *Eur. Radiol.* **30** 976–86

Wei W, Dai Q, Wong Y, Hu Y, Kankanhalli M and Geng W 2019 Surface-electromyography-based gesture recognition by multi-view deep learning *IEEE Trans. Biomed. Eng.* **66** 2964–73
Wold S, Albano C, Dunn W, Edlund U, Esbensen K, Geladi P, Hellberg S, Johansson E, Lindberg W and Sjöström M 1984 *Chemometrics* (Berlin: Springer) pp 17–95
Xia Y, Yang D, Yu Z, Liu F, Cai J, Yu L, Zhu Z, Xu D, Yuille A and Roth H 2020 Uncertainty-aware multi-view co-training for semi-supervise d me dical image segmentation and domain adaptation *Med. Image Anal.* **65** 101766
Xie Y, Zhang J and Xia Y 2019 Semi-supervised adversarial model for benign-malignant lung nodule classification on chest CT *Med. Image Anal.* **57** 237–48
Yang J *et al* 2020 Integrating tumor and nodal radiomics to predict lymph node metastasis in gastric cancer *Radiother. Oncol.* **150** 89–96
Zhang H 2004 The optimality of naive Bayes *AA* **1** 3
Zhao J, Xie X, Xu X and Sun S 2017 Multi-view learning overview: Recent progress and new challenges *Inf. Fusion* **38** 43–54
Zhou J, Zhang Q and Zhang B 2021 An automatic multi-view disease detection system via collective deep region-based feature representation *Future Gener. Comput. Syst.—Int. J. Esci.* **115** 59–75