International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 306-313

https://doi.org/10.70517/ijhsa46425

Research on multimodal image fusion of lung cancer based on deep image processing algorithm

Jianming Peng[1,2], Qian Liang[3], Huishen Yan[1] and Rui Liang[2,*]
[1] School of Medicine, Yangzhou Polytechnic College, Yangzhou, Jiangsu, 225009, China
[2] School of Basic Medical Science, Suzhou Vocational Health College, Suzhou, Jiangsu, 215009, China
[3] Scientific Research Department, Wuming Hospital of Guangxi Medical University, Nanning, Jiangsu, 530199, China
Corresponding authors: (e-mail: liangrui111222@126.com).

Abstract Fusion of multimodal images for the detection of lung cancer lesions can synthesize different modal image features and break through the limitation of single modality. In this paper, 500 lung cancer patients in a hospital were selected as research objects, and ROI segmentation and image filtering were successively performed on their lesion image group data to complete the preprocessing of the image group data. Then Pyradiomics technique is used to extract image features with clear contour details from tumor regions in magnetic resonance (MRI) type images based on five filters. Subsequently, a hierarchical multimodal feature and classifier fusion framework is proposed based on the MCF algorithm. The filtered image features are input into the framework, and each modal feature is selected individually, and the model training and information fusion are carried out in a hierarchical manner to build a prediction model based on multimodal features and classifiers. Compared with similar modeling algorithms, the correlation of most of the features extracted from CT and PET images by the prediction model in this paper reaches 0.2 or above, which shows excellent performance of multimodal image feature screening for lung cancer.

Index Terms lung cancer multimodal image, image features, classifier, filter, MCF algorithm

I. Introduction

Lung cancer is one of the most common cancers and the leading cause of cancer deaths worldwide [1]. According to official data, there will be about 19.3 million new cancer cases and nearly 10 million cancer deaths worldwide in 2023, of which about 1.8 million deaths will be due to lung cancer, accounting for about 18% of cancer deaths. Lung cancer is a malignant tumor that originates from malignant cells in lung tissue [2]. The mechanisms of lung cancer are complex and varied, and the most important causative factor is smoking, in which harmful substances such as nicotine, tar and carbon monoxide enter the lungs, triggering cellular gene mutations and leading to the gradual evolution of normal cells into malignant tumor cells [3], [4]. However, smoking is not the only risk factor, long-term exposure to second-hand smoke, occupational exposure to carcinogens, family genetic factors and air pollution are also closely related to the development of lung cancer, whose radioactive elements, polluting particles, and endogenous natural damages lead to gene mutations, thus developing lung cancer [5]-[8]. Symptoms of lung cancer are often not obvious in the early stages, but early diagnosis and early treatment are crucial to improve the cure rate and survival rate of lung cancer [9].

However, the probability of missed diagnosis for subtle nodules on chest radiographs is more than 35%, and the risk of false-positive probability of diagnosis on PET-CT scanning medical imaging is close to 30% [10], [11]. However, combining CT with sputum cytology with appropriate matching criteria can reduce lung cancer mortality [12]. In addition, because lung cancer is mostly detected at advanced stages, it is considered a highly metastatic type of cancer, with up to 65% of lung cancer patients eventually developing brain metastases in clinical practice [13]. Due to the better performance of magnetic resonance imaging (MRI) in examining brain metastases [14]. Conventional MRI images include T1-weighted imaging (TIWI), fluid-attenuated inversion recovery sequences (FLAIR), and diffusion-weighted imaging (DWI), which have different sensitivities to tumors. TIWI provides images of anatomical structures with better contrast, FLAIR can help detect abnormal changes around cerebrospinal fluid areas, and DWI is very helpful [15]-[17]. The combined application of these image imaging modalities helps physicians to better assess and localize brain metastases and select appropriate therapeutic measures for better management of patients with brain metastases [18], [19]. Therefore, the medical community and scientific and technological researchers have been exploring multimodal image fusion methods for lung cancer.

In this paper, we firstly briefly summarize the preprocessing process of lung cancer pathology slice data based on ROI segmentation and image filtering in imaging histology. Secondly, it elaborates the definition of original

image, and the basic concepts and mathematical formulas of the five pre-filters. Then we propose a hierarchical multimodal feature and classifier fusion framework, analyze the module structure and training steps of the framework, and build a prediction model based on multimodal features and classifiers. Finally, we compare the performance of different types of image feature screening between the prediction model in this paper and similar model algorithms, and evaluate the image fusion and classification effect of the prediction model in this paper to test the feasibility of the designed model.

II. Prediction model based on multimodal features and classifiers

II. A.Preprocessing of image set data

In this paper, 500 lung cancer patients in a hospital are taken as research objects, and 500 lung cancer patients, including 250 lymph node-positive patients and 250 lymph node-negative patients, are divided into a training set (including a validation set) and a test set according to the ratio of 3:2 by using a randomized division without destroying the balance of the samples. Among them, the training set was used for model training and the test set was used for model evaluation. The following are the with preprocessing methods for the imaging set data.

(1) ROI segmentation: ITK-SNAP software is an interactive medical image segmentation tool, and ROIs can be easily segmented using ITK-SNAP software. In this study, the tumor region was segmented by three experienced radiologists using ITK-SNAP software. For each patient's CT image, one radiologist read and outlined the tumor region, and then another radiologist reviewed the results to confirm whether the tumor region was accurately outlined, and if there was a controversy, the outlining was redone after a discussion to reach a consensus. The final CT images and outlined Mask images were saved in Nifti format.

(2) Image filtering: this study not only extracts features from the original CT images, but also from the filtered images, the filtering methods used in this study include Gaussian Laplace (Log) filtering and wavelet filtering.Log filtering mainly consists of two steps, the first step is to perform Gaussian filtering on the image, and the second step is to perform Laplace filtering on the image. Among them, Gaussian filtering reduces the noise and acts as a smoothing effect on the image, and Laplace filtering highlights the regions of the image where the intensity changes rapidly.Wavelet filtering mainly consists of three steps. The first step is to wavelet transform the image using wavelets (e.g., haar wavelets), the second step is to process the wavelet coefficients based on some rule (e.g., threshold filtering) to remove noise from the image, and the third step is to reconstruct the signal using wavelet inverse transform to obtain an approximation of the original image.

II. B.Extraction of Imaging Histological Features

In order to extract high-throughput features from the tumor region in MRI images, features are extracted from MRI images using Pyradiomics, an open-source Python package that is often used to extract features from images in medical formats.

The pre-processed MRI image needs to be processed by 5 pre-filters before feature extraction, different filters can help to extract richer feature information from the MRI image. These five pre-filters are: original image, Laplace Gaussian filter (LoG), square filter, square root filter, logarithmic filter, and exponential filter. The following is a detailed description of the above filters.

Original image: It is the original image that is not processed in any way and is used directly for the next step of feature extraction. The features extracted from the original image belong to the most direct features extracted from the image, which are often able to represent the tumor in the most direct way.

LoG filter: is an edge enhancement filter. It is used to emphasize areas in an image that vary in gray level, which has the effect of making the edges and contours of objects in the image clearer.LoG filter consists of two types of filters, Gaussian filter and Laplace filter. The Gaussian filter is used to make the image smoother and then the Laplace filter is used which will make the contours of the image clearer.LoG filter can be defined as equation (1):

$$\nabla^2 G(x,y,z,\sigma) = \frac{1}{(\sigma\sqrt{2\pi})^3} e^{-\frac{x^2+y^2+z^2}{2\sigma^2}} \tag{1}$$

where $G(x,y,z,\sigma)$ is defined as a Gaussian filter and ∇^2 is defined as a Laplace filter.

Squared Filter: is calculated by extracting the squares of the gray values of pixels in the image and then scaling them to the range of variation of pixel gray values in the original image using normalization. The function of this filter is mainly to enhance the contrast of the image and the details in the image will be clearer after using this filter. The formula is given in equation (2):

$$f(x)=\left(\frac{1}{\sqrt{\max(|x|)}}x\right)^2 \tag{2}$$

Square Root Filter: The calculation method is similar to that of the square filter, except that the square root filter calculates the square root of the gray value of an image pixel. The role of this filter is opposite to that of the square filter, and its role is to reduce the contrast of the image. Its formula is defined in equation (3):

$$f(x)=\begin{cases}\sqrt{\max(|x|)x}, & if\ x\geq 0\\ -\sqrt{-\max(|x|)x}, & if\ x<0\end{cases} \tag{3}$$

Logarithmic filter: It is calculated by calculating the logarithm of the absolute value +1 of the gray value of the original image pixel and scaling it to the range of variation of the original image gray value after the calculation. Its formula is defined as equation (4):

$$f(x)=\begin{cases}c\log(x+1), & if\ x\geq 0\\ -c\log(-x+1), & if\ x<0\end{cases}, where\ c=\frac{\max(|x|)}{\log(\max(|x|)+1)} \tag{4}$$

Exponential filter: is to extract the exponent of the gray value of the pixel in the image and scaled to the range of variation of the gray value of the original image after calculation as in equation (5).

$$f(x)=e^{\frac{\log(\max(|x|)+1)}{\max(|x|)}x} \tag{5}$$

II. C.Multimodal Features and Multi-Classifier Hierarchical Fusion Framework

In the hierarchical fusion framework based on the MCF algorithm, the features of each modality are individually feature-selected and then fed into multiple base classifiers for model training and information fusion in a hierarchical manner, with the classifiers and modalities fused through the MCF of the Level-1 and Level-2 modules, respectively.

During the training process, the main function of the Level-1 module is to generate weights for multiple base classifiers in each modality. In the training process, the K-fold cross-validation is used for the training set of each modality to obtain the prediction probability of each base classifier, for each fold, and the prediction probability of the k-fold of the i-classifier can be expressed as equation (6):

$$\hat{P}_i^k\left(i=1,2,\ldots,N,k=1,2,\ldots,K\right) \tag{6}$$

At the same time, the prediction performance evaluation index corresponding to each fold can be obtained, such as accuracy, AUC, sensitivity, specificity, etc., then the evaluation matrix $\hat{D}_{N,M}^k$ can be constructed, and according to the MCF method then the weights of the base classifiers are calculated as $\hat{\omega}_i^m, i=1,2,\ldots,N$, and $i=1,2,\ldots,N$ is obtained from the weight matrix $\tilde{W}$ by taking the mean by column. At the same time, the Level-1 module also outputs the predicted probability $\tilde{P}^m$ of the current modal m training set as a whole, which is calculated by doing a weighted sum of the validation set of the k th-fold and the corresponding row of weights in the weighting matrix $\tilde{W}$ as in Equation (7):

$$\tilde{P}_m^k=\sum_{i=1}^{N}\hat{P}_i^k\cdot\hat{\omega}_i^k \tag{7}$$

The predicted probabilities from this training set are mainly used for training the Level-2 module.

The main intention of training the Level-2 module is to assign weights to each modality, i.e., the process of multimodal fusion. Using the prediction probability $\tilde{P}^m$ of the whole training set outputted by Level-1, the prediction performance evaluation index under modality m is calculated, and then a new evaluation matrix $\tilde{D}_{m,M}$ is constructed, so that the weights of modalities $\tilde{W}^i$, $i=1,2,\ldots,m$ can be calculated by the formula in the MCF method.

In the testing process, the test set is input into the trained base classifier after feature selection, and the predicted probability of the test set is obtained as in equation (8):

$$P_i^m(i=1,2,\ldots,N) \tag{8}$$

The meaning is the probability of prediction of classifier i under modality m. The weights of the base classifiers output by using Eq. and Level-1 are fused at the classifier level to obtain the predicted probability of the test set under modality m after fusion as in Eq. (9):

$$P^m = \sum_{i=1}^{N} P_i^m \cdot \hat{\omega}_i^m \tag{9}$$

$\hat{\omega}_i^m$ is the weight of each base classifier obtained from the training set under modality m. For fusion at the modal level, the weights of each modality output from Level-2 are used to obtain the final prediction probability as in Eq. (10):

$$P = \sum_m P^m \cdot \hat{w}^m \tag{10}$$

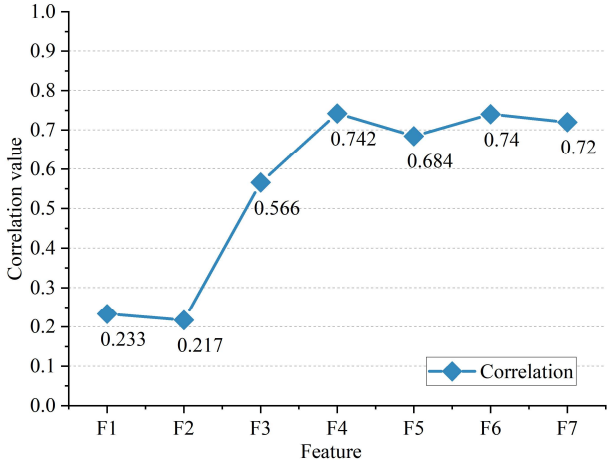
III. Performance analysis of the prediction model

III. A. Comparison of image feature screening results

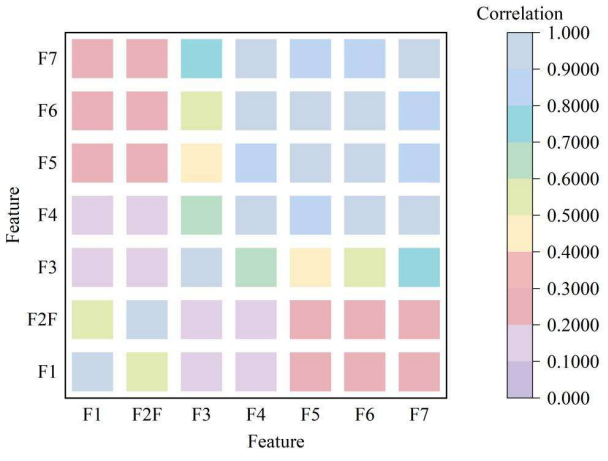
In the training cohort (n=300), 216 patients carried EGFR-sensitive mutations (EGRF+) and 5 patients had wild-type EGFR (EGRF-). In the test cohort (n=200), 97 patients carried EGFR-sensitive mutations. This section unfolds a detailed analysis of the commonly used VGG_TRAD prediction model with the features selected by the model in this paper.

The correlation between features and features, and EGFR-sensitive mutations was visualized, and Figure 1 shows the correlation coefficients between CT and PET image features and targets in this paper's model. Where (F1)LLH_GLRLM_LGLRE, (F2)LLh_GLCM_MP, (F3)CT_F153, (F4)CT_F482, (F5)CT_F494, (F6)CT_F844, (F7)CT_F1024, (F8)OFM, (F9)OF1OP, (F10)LLL_FM, (F11)LLL F10P, (F12)pet_F469 are the features filtered by VGG_TRAD model.

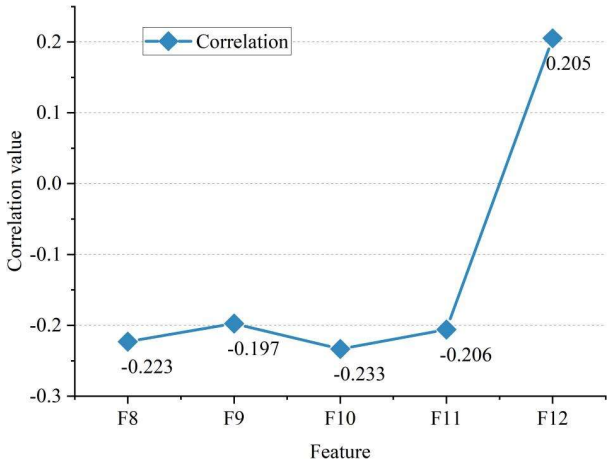
Figure 1(a) shows the strength of correlation between different CT image features and EGFR (epidermal growth factor receptor) sensitive mutations. (F1)LLH_GLRLM_LGLRE and (F2)LLh_GLCM_MP, which are superficial tumor features, have shorter bars, while deep tumor features starting with “CT_F” have relatively longer bars. This indicates that deep features have higher correlation with EGFR-sensitive mutations and are more predictive than shallow features. Figure 1(b) demonstrates the correlation between different CT image features. The results show high correlations between tumor depth features, especially between (F4)CT_F482, (F5)CT_F494, (F6)CT_F844 and (F7)CT_F1024, which show a correlation close to 1, suggesting that they can reflect similar tumor characteristics. Combining the results of Fig. 1(a)-(b), it can be learned that the similar tumor characteristics responded by the CT depth features in the VGG_TRAD model may have a great potential for identifying the characteristics of EGFR-sensitive mutations. Similarly, Fig. 1(c) and (d) reflect the correlation between PET image features and EGFR-sensitive mutations, and the correlation between PET image features themselves, respectively. The results show that both shallow and deep tumor features are less sensitive to mutated tumors. This suggests that the VGG_TRAD model may not identify the tumor metabolic information in PET images well.



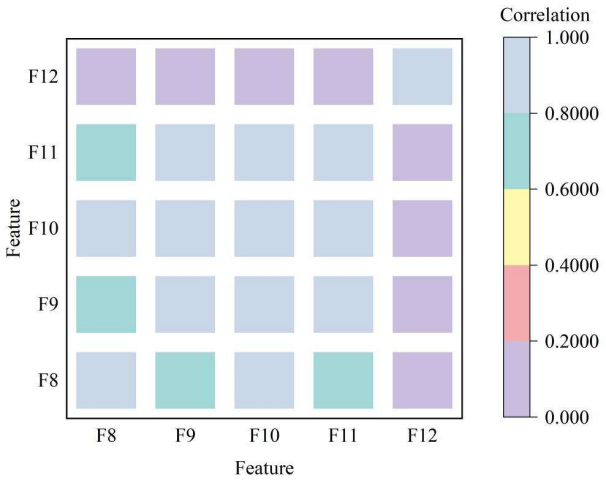
(a) The characteristic correlation values of F1-F7



(b) The characteristic correlation values of F1-F7 are strong or weak



(c) The characteristic correlation values of F8-F12



(d) The characteristic correlation values of F8-F12 are strong or weak

Figure 1: Image features and correlation coefficients

The performance of the correlation coefficients of CT-type and PET-type image features and targets in the prediction model of this paper is shown in Figure 2 for LLH_GLRLM_HGLRE, LHL_GLDM_LDHGLE, CT_F186, CT_F190, CT_F228, CT_F245, CT_F246, CT_F254, CT_F268, CT_ F302, CT_F304, CT_F311, CT_F503, CT_F638, CT_F659, CT_F943, CT_F986, PET_F429, PET_F601, PET_F675, PET_F708, PET_F988, and PET_F999, a total of 23 features, using numbers 1-23 for numbering, where 3-16 are CT type images and 17-23 are PET type images. Unlike the features screened in the VGG_TRAD model, among the features screened in the prediction model in this paper, most of the image correlations were 0.2 and above for both CT-type and PET-type image features, which had a strong correlation with EGFR-sensitive mutations.

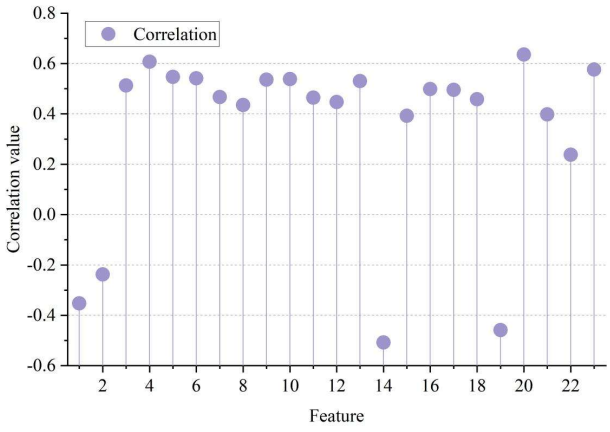


Figure 2: The correlation coefficient of image features in the RES TRAD model

III. B. Fusion of Predictive Models and Assessment of Classification Effectiveness

III. B. 1) Feature Screening Based on Principal Component Analysis

Principal Component Analysis (PCA), as one of the commonly used feature screening techniques, can reduce the dimensionality of the dataset while retaining the original information of the data as much as possible. In this section, the dimensionality reduction effect analysis of different modal (m, n) image features extracted from the model of this paper is carried out, where m is 210-dimensional high-dimensional features and n is 150 air-frequency features.

The 300-dimensional high-dimensional fusion features obtained from the fusion of different modal features are screened using PCA dimensionality reduction, the screening effect is shown in Table 1, and the final dimension of the selected high-dimensional fusion features is 210 dimensions. Setting the final dimension of the air-frequency

features as 150 dimensions, after fusion with the high-dimensional fusion features again using PCA to obtain a global feature dimension of 270, the dimensionality reduction effect is shown in Table 1.

Table 1: The effect of dimensionality reduction by PCA

Dimension	High-dimensional fusion features (%)	Global feature (%)
30	94.67	98.49
60	94.35	97.79
90	96.51	98.87
120	95.43	98.49
150	95.3	98.55
180	94.54	98.94
210	96.83	99.06
240	94.54	97.98
270	95.56	99.51
300	94.41	98.3

III. B. 2) Effect of high-dimensional feature fusion

The optimal features obtained by dimensionality reduction of different modal features extracted with the modeling method in this paper are fused into 300-dimensional deep learning features by tandem operation, and 210-dimensional high-level features are obtained after dimensionality reduction by PCA again, and the optimal features obtained by dimensionality reduction of different modal features extracted with the modeling method in this paper are fused into 300-dimensional deep learning features by tandem operation, and 210-dimensional high-level features obtained by dimensionality reduction of different modal features extracted with PCA again, and 210-dimensional high-level features are obtained after dimensionality reduction by PCA again, and 210-dimensional high-level features are obtained after dimensionality reduction of different modal features extracted with PCA again. Bagging, (HF8) SVM, (HF9) LGBM, and (HF10) XGB, a total of 10 types of classifiers using 5-fold cross-validation to compare the accuracy scores before and after the fusion are shown in Fig. 3. It can be seen that the accuracy of the fused features is higher than that of the extracted features of the model approach of this paper, and the highest in the (HF10) XGB classifier is up to 95.82%.

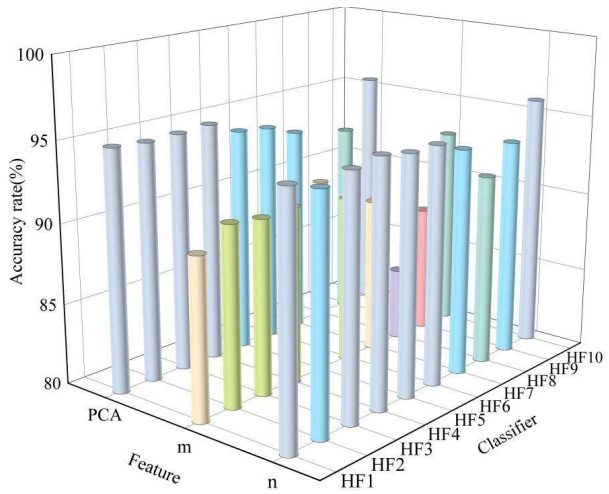


Figure 3: m,n features and high-dimensional fusion features ACC

III. B. 3) Validation of global feature effects

The experimentally obtained 210-dimensional high-dimensional features and 150 air-frequency features (low-dimensional fusion features) are fused, and the 270-dimensional global features are obtained by PCA dimensionality reduction, and the F1 scores before and after the fusion are compared by using 5-fold cross-validation in Fig. 4, in which the F1 scores of low-dimensional fusion features and memory high-dimensional fusion features after fusion are greatly improved compared with those before fusion, and the F1 scores of all the classifiers are 95.00% and above, indicating that low-dimensional fusion features and high-dimensional fusion features can effectively improve the classification effect. The F1 scores of all 10 classifiers are 95.00% and above,

indicating that the low-dimensional fusion features and high-dimensional fusion features can effectively improve the classification effect. Global features can utilize the advantages of both low-dimensional and high-dimensional features to fully express the information contained in the image, so as to achieve better classification results.

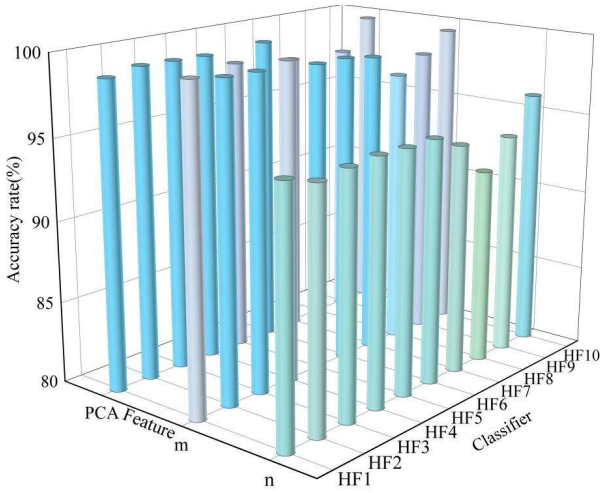


Figure 4: Comparison of F1 scores of the three groups of fused features

IV. Conclusion

In this paper, a prediction model based on a hierarchical fusion framework of multimodal features and classifiers is proposed, which combines the interactions between multimodal features and relies on multiple classifiers to increase the flexibility and diversity, and is able to accurately and efficiently extract the features of lung cancer foci in different sequences of MRI images. With the support of this framework, the designed prediction model not only has good multimodal lung cancer image feature screening ability and prediction performance, but also has superior lung cancer image classification accuracy and efficiency. Most of the extracted image features have correlations of 0.2 and above. After the dimensionality reduction and fusion of the extracted multimodal features, the accuracy can reach up to 95.82% on the XGB classifier, and the F1 scores of the 10 classifiers can reach 95.00% and above.

Contribution

# Jianming Peng and Qian Liang have contributed equally to this work.

Funding

1. The Excellent Teaching Team Fund of the Cyan-Blue Project at Yangzhou Polytechnic College;
2. Suzhou Applied Basic Research Program (No. SYWD2024101);
3. Key Scientific Research Project of Suzhou Vocational Health College (No. SZWZY202402);
4. the Youth Fund of the Cyan-Blue Project at Jiangsu Province;
5. the Academic Leader Fund of the Cyan-Blue Project at Jiangsu Province.

References

[1] Tao, M. H. (2019). Epidemiology of lung cancer. Lung Cancer and Imaging, 4-1.
[2] Sainz de Aja, J., Dost, A. F., & Kim, C. F. (2021). Alveolar progenitor cells and the origin of lung cancer. Journal of internal medicine, 289(5), 629-635.
[3] Aisner, D. L., Sholl, L. M., Berry, L. D., Rossi, M. R., Chen, H., Fujimoto, J., ... & Kwiatkowski, D. J. (2018). The impact of smoking and TP53 mutations in lung adenocarcinoma patients with targetable mutations—the Lung Cancer Mutation Consortium (LCMC2). Clinical Cancer Research, 24(5), 1038-1047.
[4] Jeon, J., Holford, T. R., Levy, D. T., Feuer, E. J., Cao, P., Tam, J., ... & Meza, R. (2018). Smoking and lung cancer mortality in the United States from 2015 to 2065: a comparative modeling approach. Annals of internal medicine, 169(10), 684-693.
[5] Akhtar, N., & Bansal, J. G. (2017). Risk factors of Lung Cancer in nonsmoker. Current problems in cancer, 41(5), 328-339.
[6] Brey, C., Gouveia, F. T., Silva, B. S., Sarquis, L. M. M., Miranda, F. M. D. A., & Consonni, D. (2020). Lung cancer related to occupational exposure: an integrative review. Revista gaucha de enfermagem, 41, e20190378.
[7] Benusiglio, P. R., Fallet, V., Sanchis-Borja, M., Coulet, F., & Cadranel, J. (2021). Lung cancer is also a hereditary disease. European Respiratory Review, 30(162).
[8] Xue, Y., Wang, L., Zhang, Y., Zhao, Y., & Liu, Y. (2022). Air pollution: A culprit of lung cancer. Journal of hazardous materials, 434, 128937.
[9] Kim, J., Lee, H., & Huang, B. W. (2022). Lung cancer: diagnosis, treatment principles, and screening. American family physician, 105(5), 487-494.

[10] Bradley, S. H., Abraham, S., Callister, M. E., Grice, A., Hamilton, W. T., Lopez, R. R., ... & Neal, R. D. (2019). Sensitivity of chest X-ray for detecting lung cancer in people presenting with symptoms: a systematic review. British Journal of General Practice, 69(689), e827-e835.
[11] Feng, M., Yang, X., Ma, Q., & He, Y. (2017). Retrospective analysis for the false positive diagnosis of PET-CT scan in lung cancer patients. Medicine, 96(42), e7415.
[12] Spiro, S. G., Shah, P. L., Rintoul, R. C., George, J., Janes, S., Callister, M., ... & Hackshaw, A. (2019). Sequential screening for lung cancer in a high-risk group: randomised controlled trial: LungSEARCH: a randomised controlled trial of Surveillance using sputum and imaging for the EARly detection of lung Cancer in a High-risk group. European Respiratory Journal, 54(4).
[13] Wang, G., Xu, J., Qi, Y., Xiu, J., Li, R., & Han, M. (2019). Distribution of brain metastasis from lung cancer. Cancer Management and Research, 9331-9338.
[14] Zhu, D., Shao, Y., Yang, Z., Cheng, A., Xi, Q., Liang, X., & Chu, S. (2023). Magnetic resonance imaging characteristics of brain metastases in small cell lung cancer. Cancer Medicine, 12(14), 15199-15206.
[15] van der Zwaag, W., Buur, P. F., Fracasso, A., van Doesum, T., Uludağ, K., Versluis, M. J., & Marques, J. P. (2018). Distortion-matched T1 maps and unbiased T1-weighted images as anatomical reference for high-resolution fMRI. Neuroimage, 176, 41-55.
[16] Azad, R., Tayal, M., Azad, S., Sharma, G., & Srivastava, R. K. (2017). Qualitative and quantitative comparison of contrast-enhanced fluid-attenuated inversion recovery, magnetization transfer spin echo, and fat-saturation T1-weighted sequences in infectious meningitis. Korean Journal of Radiology, 18(6), 973-982.
[17] Usuda, K., Funazaki, A., Maeda, R., Sekimura, A., Motono, N., Matoba, M., & Uramoto, H. (2017). Economic benefits and diagnostic quality of diffusion-weighted magnetic resonance imaging for primary lung cancer. Annals of Thoracic and Cardiovascular Surgery, 23(6), 275-280.
[18] Zoccarato, M., Ferrati, C., & D'Errico, I. (2023). Superficial Brainstem FLAIR Hyperintensity and Restricted Diffusion Leading to the Detection of Non–Small-Cell Lung Cancer. Neurology, 101(2), 88-89.
[19] Sun, L., Luan, S., Ye, X., Chen, J., Shi, J., Zhu, H., ... & Yu, H. (2024). Combination of Contrast-enhanced FlAIR and Contrast-enhanced T1WI: A Quick and Efficient Method in Detecting Brain Metastases of Lung Cancers. Current Medical Imaging, 20(1), e15734056288838.