Received: 19 July 2021 | Revised: 13 September 2021 | Accepted: 30 September 2021

DOI: 10.1002/mp.15298

RESEARCH ARTICLE

MEDICAL PHYSICS

3D gray density coding feature for benign-malignant pulmonary nodule classification on chest CT

BingBing Zheng[1,#] | **Dawei Yang**[2,3,#] | **Yu Zhu**[1] | **Yatong Liu**[1] | **Jie Hu**[2] | **Chunxue Bai**[2,3]

[1] School of Information Science and Engineering, East China University of Science and Technology, Shanghai, China

[2] Department of Pulmonary Medicine, Shanghai Respiratory Research Institute, Zhongshan Hospital, Fudan University, Shanghai, China

[3] Shanghai Engineering Research Center of Internet of Things for Respiratory Medicine, Shanghai, China

Correspondence
Yu Zhu, School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China.
Email: zhuyu@ecust.edu.cn
Chunxue Bai, Department of Pulmonary Medicine, Shanghai Respiratory Research Institute, Zhongshan Hospital, Fudan University, Shanghai 200032, China.
Email: bai.chunxue@zs-hospital.sh.cn

[#]BingBing Zheng and Dawei Yang are co-first authors.

Funding information
Shanghai Pujiang Program, Grant/Award Number: 20PJ1402400; Science and Technology Commission of Shanghai Municipality, Grant/Award Number: 20DZ2261200; Shanghai Engineer & Technology Research Center of Internet of Things for Respiratory Medicine, Grant/Award Number: 20DZ2254400

Abstract

Purpose: Early detection is significant to reduce lung cancer-related death. Computer-aided detection system (CADs) can help radiologists to make an early diagnosis. In this paper, we propose a novel 3D gray density coding feature (3D GDC) and fuse it with extracted geometric features. The fusion feature and random forest are used for benign–malignant pulmonary nodule classification on Chest CT.

Methods: First, a dictionary model is created to acquire codebook. It is used to obtain feature descriptors and includes 3D block database (BD) and distance matrix clustering centers. 3D BD is balanced and randomly selecting from benign and malignant pulmonary nodules of training data. Clustering centers is got by clustering the distance matrix, which is the distance between every two blocks in 3D BD. Then, feature descriptor is obtained by coding the pulmonary nodule with codebook, and 3D GDC feature is the result of histogram statistics on feature descriptor. Second, geometric features are extracted for fusion feature. Finally, random forest is performed for benign–malignant pulmonary nodule classification with fusion feature of the 3D gray density coding feature and the geometric features.

Results: We verify the effectiveness of our method on the public LIDC-IDRI dataset and the private ZSHD dataset. For LIDC-IDRI dataset, compared with other state-of-the-art methods, we achieve more satisfactory results with 93.17 $\pm$ 1.94% for accuracy and 97.53 $\pm$ 1.62% for AUC. As for private ZSHD dataset, it contains a total of 238 lung nodules from 203 patients. The accuracy and AUC achieved by our method are 90.0% and 93.15%.

Conclusions: The results show that our method can provide doctors with more accurate results of benign–malignant pulmonary nodule classification for auxiliary diagnosis, and our method is more interpretable than 3D CNN methods, which can provide doctors with more auxiliary information.

KEYWORDS
3d gray density coding feature, benign–malignant classification, geometric features, pulmonary nodules, random forest

1 | INTRODUCTION

Lung cancer is a common cancer and the leading cause of cancer-related deaths in both man and woman worldwide.[1] The recent 5-year survival rate for all stages combined is 18.6% but decreases further to 5% for advanced-stage disease.[2,3] Early detection is significant to reduce cancer-related death.[4] Computed tomography (CT) scanning technique is one of the most important techniques in medical imaging. Compared with x-ray, CT technology provides doctors with more information through higher sensitivity and resolution. According to the study of National Lung Screening Trial, the use of low-dose CT scans in lung cancer screening

TABLE 1 Related works about benign–malignant lung nodule classification on 3D

Methods	Dataset	Methods	Class	Result
Al-Shabi et al.[20]	LIDC-IDRI[19]	Proposed 3D axial-attention	Benign–malignant	AUC:96.17 Acc:92.81 Sen:92.36
Afshar et al.[21]	LIDC-IDRI	3D Multi-scale capsule network	Benign–malignant	Acc:93.12 AUC:96.4
Xie et al.[22]	LIDC-IDRI	Proposed fusion 3D feature and Adaboost-BPNN	Benign–malignant	AUC:96.65 (malignancy rate 3 were discarded)
Yan et al.[23]	LIDC-IDRI	A novel 3D CNN	Benign–malignant	Acc:87.4 Sen:89.4 Spe85.2
Shen et al.[24]	LIDC-IDRI	Multi-crop Convolutional Neural Network	Low malignancyHigh malignancy	Acc:87.14 AUC:93 Spe:93 Sen:77
Xie et al.[25]	LIDC-IDRI	Multi-view knowledge-based collaborative (MV-KBC) deep model	Benign–malignant	Acc:91.60 AUC:95.70
Ren et al.[27]	LIDC-IDRI	Proposed MRC-DNN	Benign–malignant	Acc:0.90 Sen:0.81 Spe:0.95
Tong et al[28]	LIDC-IDRI	3D-CNN and SVM with MKL algorithms	Benign–malignant	Acc:90.65 Sen:87.50 Spe:94.12

reduced the possibility of smokers' lung cancer related death about 20%.[5]

Radiologists can make a correct diagnosis with these medical imaging techniques, but too many images may greatly increase workload. Farther, with the rapidly increasing demand in health care and physical examination, a large number of experienced radiologists are required for screening. Mistakes caused by various reasons during manual screening can lead to unnecessary personnel injury. To alleviate the problem, computer aided diagnosis system (CADs) is a good choose. CADs can automatically locate and analyze pulmonary nodules to provide doctors with auxiliary diagnosis and even possible treatment options.[6–9] It can greatly reduce the workload of radiologists. This paper focuses on the pulmonary nodules' benign and malignant classification method with 3D features extraction based on PET/CT images. The main innovations are as follows:

1. This paper presents a novel 3D gray density coding (3D GDC) feature. First, 3D pulmonary nodules are coded by establishing a dictionary model. Then, histogram statistics is performed on coding results to get 3D GDC feature. 3D GDC feature can characterize the distribution of different gray density levels in pulmonary nodules, which represents significant discrimination between benign and malignant pulmonary nodules on CT images. The experiments prove the effectiveness of 3D GDC for benign-malignant pulmonary nodule classification.
2. Geometric features (GFs) of 3D pulmonary nodules are extracted and fused with extract 3D GDC feature as the final feature, including depth, width and height of external matrix, longest diameter, and shortest diameter. It can provide additional auxiliary information to obtain more satisfactory diagnostic results.
3. We evaluate the proposed methods on two datasets. For public LIDC-IDRI dataset, the proposed fusion feature can get state-of-the-art results with 97.53 ± 1.62% for AUC and 93.17 ± 1.94% for accuracy. For private ZSHD dataset, significant results are obtained with 93.15% for AUC and 90.0% for accuracy.

2 | RELATED WORKS

Many researchers have proposed methods for CADs related to lung cancer, such as the detection of lung nodules,[10,11] reduction of false positives of lung nodules,[12,13] and the segmentation of lung nodules.[14,15] For the benign–malignant pulmonary nodule classification, there have been many excellent works. Some of them are based on 2D models[16–18] and others are based on 3D models. Table 1 summarizes some related literature about 3D methods.

Shen et al.[22] presented a 3D multi-crop convolutional neural network (MC-CNN) to automatically extract nodule salient information by employing a novel multi-crop pooling strategy which crops different regions from convolutional feature maps and then applies max-pooling

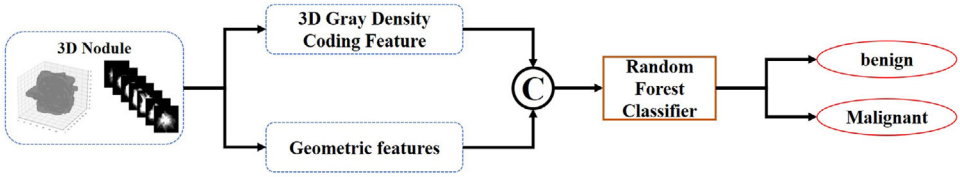

FIGURE 1 The flow chart of the proposed method

different times. MC-CNN was verified on the LIDC-IDRI dataset which includes 880 low malignancy-suspicious nodules (LMN) and 495 high malignancy-suspicious nodules (HMN). LMN means that the average malignancy rating in the annotation is less than 3 and HMN is greater than 3. The best accuracy, AUC, and sensitivity were 87.14%, 93%, and 77%, respectively. Wang et al.[26] proposed a 3D joint deep learning model where the segmentation can better facilitate the classification of pulmonary GGNs. It was a cascade architecture with both segmentation and classification networks. As a trainable preprocessing module, the segmentation model provides the classification-guided attention weight map for the original CT data to achieve better diagnostic performance.

Ren et al.[27] proposed a novel CNN named regularized classification deep neural network (MRC-DNN) to benign–malignant pulmonary nodule classification on 3D CT. On verification on LIDC-IDRI dataset, the proposed method obtained 0.90 for accuracy with 0.81 for sensitivity and 0.95 for specificity. Tong et al.[28] proposed CADs for lung nodule diagnosis based on 3D-CNN and SVM with multiple kernel learning (MKL) algorithms. The system not only explores the computed tomography (CT) scans, but also the clinical information of patients like age, smoking history, and cancer history. To extract deeper image features, a 34-layer 3D residual network (3D-ResNet) is employed. Heterogeneous features including the extracted image features and the clinical data are learned with MKL.

3 | METHOD

The proposed method mainly includes two parts. The first part is to extract 3D GDC feature and the geometric features of pulmonary nodules on CT images. Then combining the two features, random forest is applied for benign–malignant pulmonary nodule classification, as shown in Figure 1. We will introduce the proposed method in detail as following:

3.1 | 3D gray density coding feature

The Hounsfield unit (HU) value of pulmonary nodules on CT images is highly correlated with the classification of benign and malignant pulmonary nodules. Therefore, the gray density is one of the important features of pulmonary nodules. In this paper, 3D GDC feature is proposed for benign–malignant pulmonary nodule classification on chest CT.

The extraction of 3D GDC feature is divided into three steps. First, the dictionary model is created to obtain a codebook, as shown in Figure 2. The codebook involves 3D block database (BD) and clustering centers C_{BD}. Second, pulmonary nodules are coded to obtain the feature descriptors with the obtained codebook, as shown in Figure 5. Finally, the 3D GDC feature is obtained by using histogram statistics. Following, the method will be described in detail.

3.1.1 | Dictionary model

Dictionary model is created to obtain codebook. Codebook includes 3D BD and clustering centers and is used to code pulmonary nodules for 3D feature descriptors.

3D block database (BD) is the important part of extracting 3D GDC feature. It contains 3D blocks that are randomly extracted from 3D pulmonary nodules of training data. The number and the size of blocks in 3D BD will determine the effect of feature expression. For the number of blocks, in order to ensure the diversity and category balance of blocks in BD, a same number of 3D blocks are randomly selected from different categories of pulmonary nodules. For the size of blocks, because the sizes of pulmonary nodules are between 3 and 30 mm, 3D blocks of appropriate size should be selected. Large sizes are suitable for big pulmonary nodules, while small size for small nodules.

Suppose the size of 3D block is $p_D \times p_W \times p_H$ (D, W, and H mean the depth, length, and width of block), a sliding window with the same size is used to slide on the 3D pulmonary nodules of training data to obtain 3D blocks. Only the 3D blocks corresponding to the foreground region are saved. Then, N blocks are randomly selected from the saved blocks to form 3D BD, as shown in Figure 2. We extract 800 3D blocks from each category (benign and malignant), a total of 1600 for 3D BD.

Then, the distances between each two blocks in 3D BD are calculated to form the distance matrix H_{BD}, which is a symmetric matrix of size $N \times N$. Each row (or

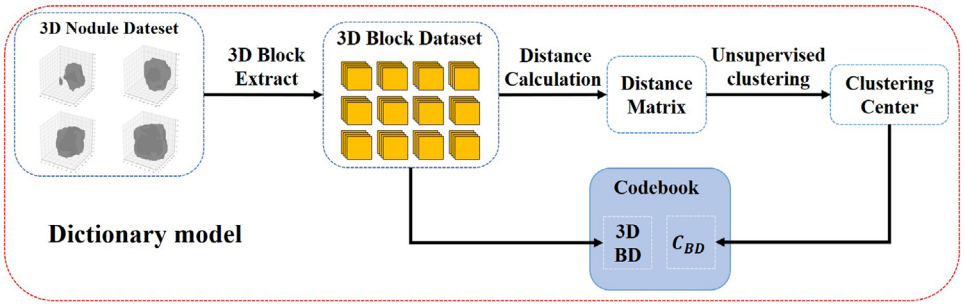

FIGURE 2 Dictionary model structure diagram

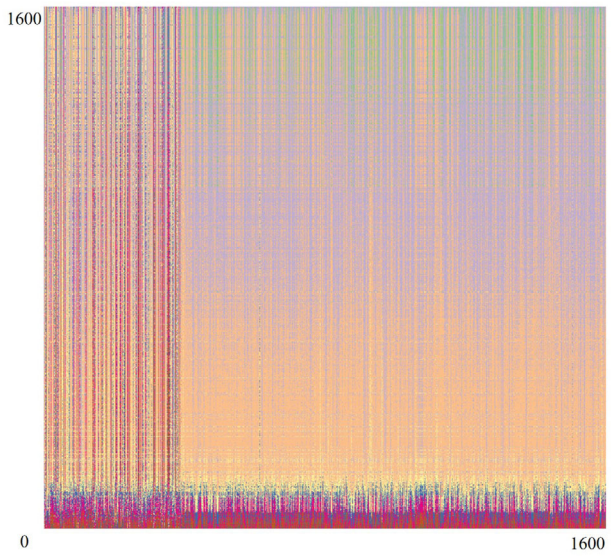

FIGURE 3 Visualization of distance matrix

column) in H_{BD} is the distance vector between a block with all blocks in the 3D BD. The commonly used vector distance calculation methods include Euclidean (EU), Canberra (CA), Cosine (COS), and City Block (CB). The experimental results show that using CA distance can get the best results. The distance formulas for any two equal length vectors p and q are as follows:

$$d_{CA}\ (p, q) = \sum_{i=1}^{n} \frac{|p_i - q_i|}{|p_i| + |q_i|}, \tag{1}$$

where n is the length of the vector; p *and* q are the vectors to be calculated.

Distance matrix can be seen as N distance vectors. We sorted them by their average values, and the visualization is shown in Figure 3. Then, K-means[29] unsupervised clustering method is used to distance matrix clustering because the algorithm is simple, has fast convergence and good effect. As for the number of clusters N_C, the obvious aggregation can be seen from Figure 3, and the range of cluster number can be set between 6 and 14. We perform related ablation experiment for the number of clusters and the best result can be obtained when N_C is 10, as shown in Figure 4. Cluster centers C_{BD} are obtained and each center corresponds to a 3D gray density coding level.

So far, the dictionary model is created and the codebook is obtained, which includes 3D BD and clustering centers C_{BD}. Next, codebook is used to code all 3D pulmonary nodules for 3D feature descriptors.

3.1.2 | Coding process and feature extraction

For a 3D pulmonary nodule of size $D \times W \times H$, a sliding window with a size of $(p_D \times p_W \times p_H)$ is used to slide on the nodule. The sliding window is the same size as the blocks in 3D BD. Stride is set as 1 and zero padding is used for the edge points. For each sliding window, we can get a 3D block b_p (p means the position of the slide window in the 3D pulmonary nodule). Finally, a total of $D \times W \times H$ blocks are obtained, as the 3D block in Figure 5.

For the obtained 3D blocks, the distance vector d_p is obtained by calculating the distances between the 3D block with blocks in 3D BD of codebook. Then, obtained distance vector is compared with the C_{BD} in codebook. The 3D gray density coding level L_{b_p} of the 3D block is the label of the clustering center in C_{BD} corresponding to the minimum distance with d_p. In particular, if the center point (position of the block) of the 3D block is the background area, the distance matrix is not required. The 3D gray density coding level L_{b_p} is set −1. The gray density coding level of each position of the 3D pulmonary nodule is obtained to form a 3D pulmonary nodule feature descriptor, as shown in Figure 6b. The above process is shown in Equations (2) and (3):

$$d_p = \left[d_{CA}\left(b_p, b_m\right) \ \textit{for}\ b_m\ \textit{in}\ BD \right] \tag{2}$$

$$L_{b_p} = Label\left(\underset{n\,=\,1\sim10}{\operatorname{argmin}}\ d_{CA}\left(d_p, C_{BD_n}\right) \right) \tag{3}$$

where p means the position of the block in the 3D pulmonary nodule; n means the number of clusters.

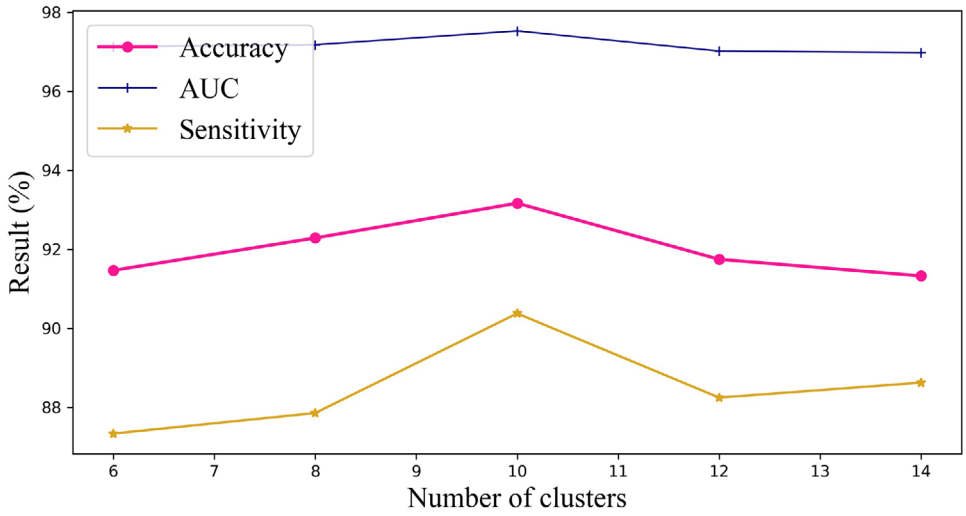

FIGURE 4 Visualization of ablation experiments about cluster number

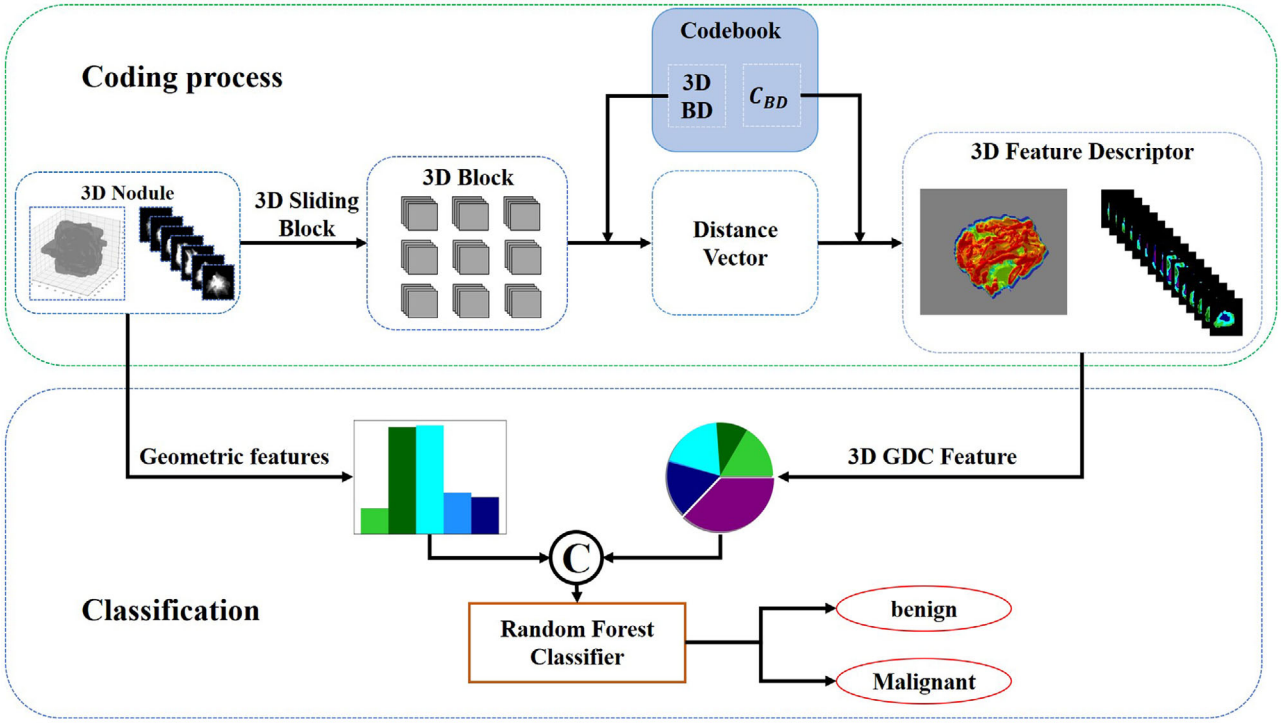

FIGURE 5 Coding, feature extraction, and classification schematics of the proposed method

At last, we statistics the number of different gray density coding levels to get the final 3D GDC feature. The results are shown as Figure 6.

3.2 | Geometric features

3D GDC feature can represent the distribution of different gray density levels of pulmonary nodules, but it cannot reflect the geometric information related to 3D pulmonary nodules (such as size, volume, longest diameter, and shortest diameter), which have a strong correlation with the benign and malignant of pulmonary nodules. So, some geometric features of 3D pulmonary nodules are extracted, which are depth, width, height, longest diameter, and shortest diameter. These features will be concatenated with 3D GDC feature for benign–malignant pulmonary nodule classification. Figure 7 shows the visualization results of geometric features of some typical pulmonary nodules on LIDC-IDRI dataset.

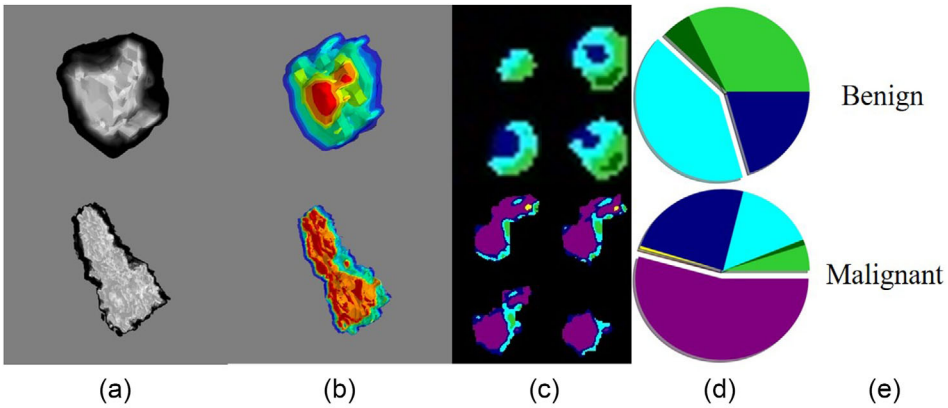

FIGURE 6 Visualization of the intermediate results of the 3D GDC feature extraction process of typical pulmonary nodules on LIDC-IDRI. (a) The 3D pulmonary nodule; (b) the 3D feature descriptor of nodule; (c) the 3D gray density coding level of the middle layers in the 3D pulmonary nodule; (d) the 3D gray density coding feature (3D GDC); (e) the benign and malignant label of 3D pulmonary nodule

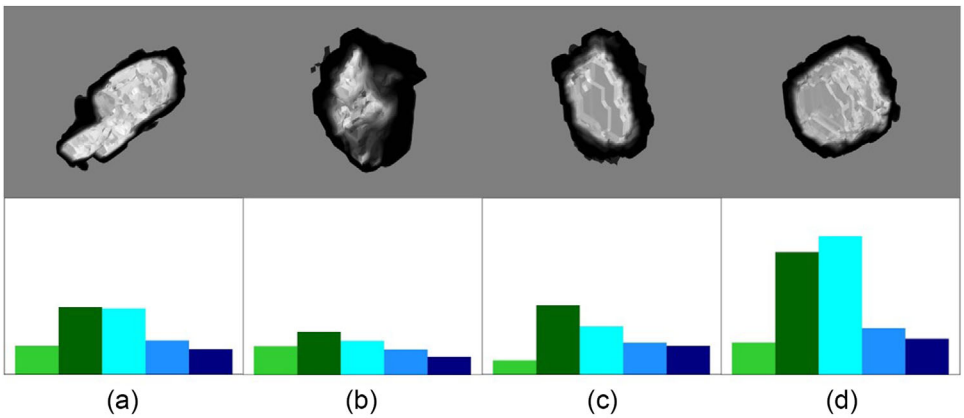

FIGURE 7 Visualization of geometric features of typical pulmonary nodules. The first row is the 3D pulmonary nodule. The second row is the visualization of geometric features. (a,b) Benign pulmonary nodules; (c,d) malignant pulmonary nodules

3D GDC feature and geometric features are concatenated for classification of benign and malignant pulmonary nodules on LIDC-IDRI dataset.

3.3 | Random forest

Random forest[30] is a classifier that uses multiple trees to train and test samples. The advantages of the algorithm are fast learning speed, good robustness, simple structure, and good effect. In this paper, the extracted fusion feature is used as input to train random forest for benign and malignant classification of pulmonary nodules.

4 | EXPERIMENTS AND RESULTS

4.1 | Dataset

We verified the effectiveness of the proposed method on public LIDC-IDRI dataset and private ZSHD dataset, respectively. Details of the two datasets are described as follows:

LIDC-IDRI was created by seven academic centers and eight medical imaging companies, which contains 1018 cases. Each subject includes images from a clinical thoracic CT scan and an associated XML file that records the results of a two-phase image annotation process performed by four experienced thoracic radiologists.[31]

Referring to the lung nodule list of Reeves et al.,[32] combined with XML annotation files, this paper extracted the pulmonary nodules, which were labeled by at least three radiologists. The size of these pulmonary nodules ranged from 3 to 30 mm and all of them were marked with characteristics. Pulmonary nodules were classified into five groups according to risk classification. We set pulmonary nodules with risk of 1, 2 as benign, and those with risk of 4, 5 as malignant. The numbers of benign and malignant pulmonary nodules are 388 and 391. We randomly divided the data into 5-equipartition for five-fold cross validation. The ratio of the training set to the

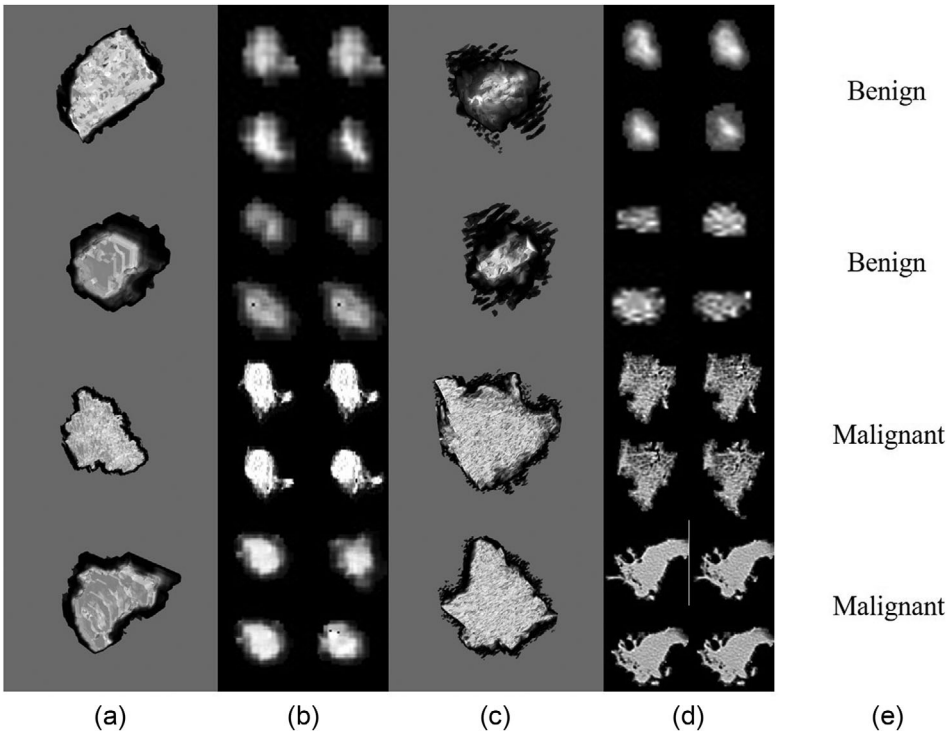

FIGURE 8 Display of 3D pulmonary nodules. (a,c) Visualization of 3D pulmonary nodules from LIDC-IDRI and ZSHD datasets, respectively; (b,d) the middle layers in the corresponding 3D pulmonary nodules; (e) the benign and malignant label of 3D pulmonary nodule

TABLE 2 Summary of the number of extracted pulmonary nodules in LIDC-IDRI. (1, 2 benign; 4, 5 malignant)

	Benign	Malignancy	Total
Train	308	312	620
Test	80	79	159
Total	388	391	779

TABLE 3 Summary of the number of different categories of pulmonary nodules in ZSHD

	Benign	Malignancy	Total
Train	70	118	188
Test	19	31	50
Total	89	149	238

test set is 4:1. Table 2 summarizes the number of different categories of extracted pulmonary nodules in LIDC-IDRI.

The private ZSHD dataset is provided by the Zhongshan hospital in Shanghai with pathologic reports. It contains CT scans from 203 patients which consist of 238 3D pulmonary nodules with 89 benign nodules and 149 malignant nodules. The slice thickness of CT images varies from 0.625 to 5 mm and most of them are 1 mm. The diameters of the nodules are 3–30 mm. We randomly divided the dataset into training set and test set at a 4:1 ratio. Table 3 summarizes the number of different categories of pulmonary nodules in ZSHD.

We rescaled the pulmonary nodules to 1 mm × 0.5 mm × 0.5 mm for LIDC-IDRI and ZSHD. Figure 8 shows the display of 3D pulmonary nodules from LIDC-IDRI and ZSHD datasets.

4.2 | Evaluation

The metrics employed to quantitatively evaluate classification are accuracy, AUC, and sensitivity. AUC (area under roc curve) is a standard indicator to measure the classification model. Accuracy represents the classification results, as shown in Equation (4):

$$Accuracy = \frac{TN+TP}{N+P} \tag{4}$$

The sensitivity measures the classifier's ability to identify positive samples:

$$\text{Sensitivity} = \frac{TP}{FN+TP} \tag{5}$$

where TN represents true negative, TP represents true positive, FN represents false negative, N represents negative, and P represents positive.

TABLE 4 Ablation experimental results on the number and size of 3D blocks in 3D BD based on ZSHD

Scheme	Acc (%)	AUC (%)	Sen (%)
$(3\times5\times5)_{1600}$	**90.0**	**93.15**	**93.54**
$(3\times7\times7)_{1600}$	90.0	93.68	93.54
$(3\times9\times9)_{1600}$	84.0	91.34	80.64
$1200_{3\times5\times5}$	84.0	90.1	93.87
$2000_{3\times5\times5}$	86.0	91.68	80.64

$(3\times5\times5)_{1600}$ represents the size of 3D blocks is $3\times5\times5$ and the number of blocks is 1600.
Abbreviations: Acc, accuracy; AUC, area under curve; Sen, sensitivity.

TABLE 5 Ablation experimental results on different features on LIDC-IDRI

Characteristics	Acc (%)	AUC (%)	Sen (%)
3D GDC	88.28 ± 2.13	94.52 ± 2.35	86.10 ± 3.22
3D GDC + GFs	93.17 ± 1.94	97.53 ± 1.62	90.38 ± 2.18

Abbreviations: Acc, accuracy; AUC, area under curve; Sen, sensitivity.

4.3 | Ablation experiment

3D block database (BD) is the most important part of extracting 3D GDC feature. The number of 3D BD and size of 3D blocks will determine the effect of feature expression. We performed ablation experiments, as shown in Table 4.

For the size of block, Table 4 indicates that $3\times5\times5$ and $3\times7\times7$ can achieve almost the same effect, but larger block size $(3\times9\times9)_{1600}$ tends to get lower effect. In order to reduce the computing time, we choose $3\times5\times5$ as the block size in this paper. As for the number of blocks, the best classification result can be obtained when the number is set as 1600.

Ablation experiment was also performed to verify the effectiveness of 3D GDC feature and geometrical features (GFs) on LIDC-IDRI dataset. The results are shown in Table 5.

Table 5 shows that average accuracy can reach to 88.28% by using only 3D GDC feature, while AUC, sensitivity of 94.52% and 86.10% are reasonable, respectively. By fusing 3D GDC feature and GFs ($p < 0.05$), the average accuracy, AUC, and sensitivity of 93.17%, 97.53%, and 90.38% are achieved. The best results of the three indicators showed the effectiveness of the fusion feature extracted.

4.4 | Results on LIDC-IDRI

For LIDC-IDRI dataset, we compared with some excellent methods. References [23,26] are based on the simple 3D CNN, and the authors in [24] use 3D CNN to extract multi-crop fusion feature for classification. The authors in [28] fuse the feature extracted by 3D CNN and other clinical information features, and finally classify the pulmonary nodules by SVM. The authors in [27] employ 3D CNN to combine manifold learning regularization. Reference [20] is based on 3D axial-attention (self attention). Table 6 summarizes the experimental results.

TABLE 6 Results of benign and malignant classification on LIDC-IDRI

Methods	Acc (%)	AUC (%)	Sen (%)
Wang et al.[26]	86.79	93.78	84.81
Yan et al.[23]	87.4	94.7	89.4
Shen et al.[24]	87.18	92.45	90.10
Ren et al.[27]	90.0	–	81.0
Tong et al.[28]	90.65	–	87.50
Al-Shabi et al.[20]	92.81	96.37	**92.36**
Ours	93.17 ± 1.94	97.53 ± 1.62	90.38 ± 2.18

Abbreviation: Acc, accuracy; AUC, area under curve; Sen, sensitivity.
Best in bold.

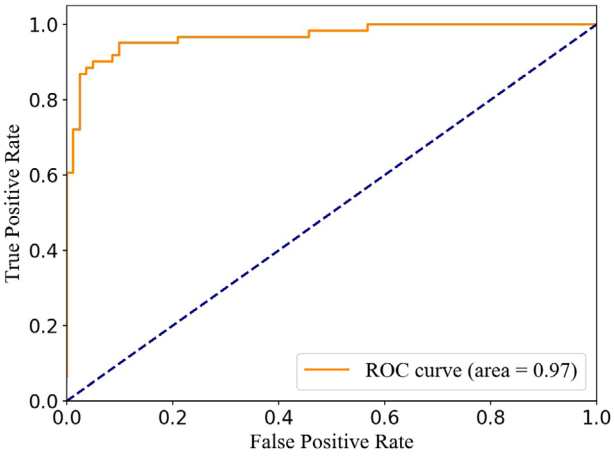

FIGURE 9 ROC of experiment results on LIDC-IDRI testing data

Table 6 shows that the proposed method has a better effect than other state-of-the-art 3D CNN methods. We can achieve average 93.17% for accuracy, 97.53% for AUC, and 90.38% for sensitivity. Although sensitivity is 1.98% lower than in reference [20], the proposed method can obtain better accuracy and AUC. Figure 9 shows the ROC of classification results on LIDC-IDRI dataset.

4.5 | Results on ZSHD

For ZSHD dataset, we compared with three excellent methods, including Zheng et al.,[16] Shen et al.,[24] and Wang et al.[26] Among them, we converted the 2D network STM-Net proposed by Zheng et al.[16] into 3D for experiment. The experimental results of Shen et al.[24] and Wang et al.[26] were reproduced by us. Table 7 summarized the experimental results.

Table 7 indicates that the proposed method has a better effect than other methods. Compared with the

TABLE 7 Results of benign and malignant classification on ZSHD

Methods	Acc (%)	AUC (%)	Sen (%)	p-value
Zheng et al.[16]	84.0	89.31	87.10	1.7×10^{-6}
Shen et al.[24]	82.0	86.39	80.65	4.3×10^{-5}
Wang et al.[26]	86.0	88.79	87.10	3.7×10^{-8}
Ours	**90.0**	**93.15**	**93.54**	

Best in bold.
Abbreviation: Acc, accuracy; AUC, area under curve; Sen, sensitivity.

other three methods, our method has the most satisfactory accuracy, AUC, and sensitivity. Experimental results proved that the method we proposed can better assist the doctor in the classification of benign and malignant pulmonary nodules, and the *p*-values are smaller than 0.01, which means that the proposed method is significantly different from other methods.

5 | DISCUSSION

This paper proposes a novel 3D GDC feature for benign–malignant pulmonary nodule classification on chest CT. 3D GDC feature is fused with geometric features (GFs) of pulmonary nodules for fusion feature. Then the random forest is used for classification.

The proposed 3D GDC feature can characterize the distribution of different gray density levels in pulmonary nodules, which is meaningful to classification of benign and malignant. First, in order to extract 3D GDC feature, a dictionary model is created to obtain codebook including 3D BD and clustering centers C_{BD}. 3D BD is acquired by randomly selected same number of certain size 3D blocks from benign and malignant pulmonary nodules of training data. Then, the distance matrix is calculated between every two blocks, and it is clustered by K-means to get C_{BD}. The clustering centers represent different gray density coding levels. Next, the gray density descriptor is obtained by encoding each location of the 3D pulmonary nodule. The 3D GDC feature can be obtained by applying histogram statistics. Figure 10 shows some visual results in 3D GDC extraction. In benign pulmonary nodules, the proportion of low-grade gray density (light color) in 3D GDC features of pulmonary nodules is higher. While in malignant pulmonary nodules, the proportion of high-grade gray density (dark color) is higher. It illustrates that the 3D GDC features have significant distinction property for benign and malignant classification and significant interpretability. For the LIDC-IDRI dataset, using only 3D GDC features, the average accuracy of the proposed method in fivefold cross validation reaches to 88.28%, AUC is 94.52%, and sensitivity is 86.10%. This fully reflects the effectiveness of 3D GDC feature.

Although 3D GDC feature can well characterize the gray density distribution of pulmonary nodules, it cannot characterize the geometric features such as the size and volume, which are correlated with the benign and malignant. In this paper, geometric features of pulmonary nodules are extracted, including the length, width, height, longest diameter, and shortest diameter. Then, the 3D GDC feature and GFs are fused for benign and malignant classification. Compared with using 3D GDC alone, the average accuracy, AUC and sensitivity are increased ($p < 0.05$) by 4.89%, 3.01%, and 4.28%, respectively. It shows that the extracted GFs features of pulmonary nodules are effective.

As shown in Table 6, compared with other state-of-the-art 3D CNN methods, the proposed method achieves more satisfactory results. Compared to simple 3D CNN,[23,26] the accuracy, AUC, and sensitivity of our method are increased more than 5.77%, 2.83%, and 0.98%, respectively. Compared to multi-crop fusion 3D CNN feature,[24] significant improvements are obtained. Farther, compared to fusion features of 3D CNN and Clinical Information,[28] the improvements of 2.52% for accuracy and 2.88% for sensitivity are achieved. As for 3D axial-attention,[22] although sensitivity was 1.98% lower, the proposed method can get remarkable higher accuracy and AUC. Moreover, the proposed 3D GDC features and geometric features are more interpretable than the 3D CNN method, which can provide more meaningful auxiliary diagnostic information for doctors.

6 | CONCLUSION

This paper proposes a novel 3D GDC feature, which is fused with extracted geometric features for classification of benign and malignant pulmonary nodules. Random forest classifier is used. 3D GDC feature can characterize the distribution of different gray density levels in pulmonary nodules. The proportion of low-grade gray density of benign pulmonary nodules is higher, while for malignant pulmonary, the proportion of high-grade gray density is higher. It represents that 3D GDC feature has satisfactory discrimination and strong interpretability for classification of benign and malignant pulmonary nodules. Geometric features of 3D pulmonary nodules are extracted to fuse with 3D GDC, which are meaningful and explainable for the benign and malignant of pulmonary nodules. The experimental results verify the proposed method. In public dataset LIDC-IDRI, 93.17% for accuracy, 97.53% for AUC, and 90.38% for sensitivity are obtained. Compared with the state-of-the-art 3D CNN methods, the proposed method not only achieves more satisfactory results, but also has better interpretability. For ZSHD, combined with 3D GDC feature and GFs, the proposed method achieved 90% for accuracy, 93.15% for AUC, and 93.54% for sensitivity, which are higher

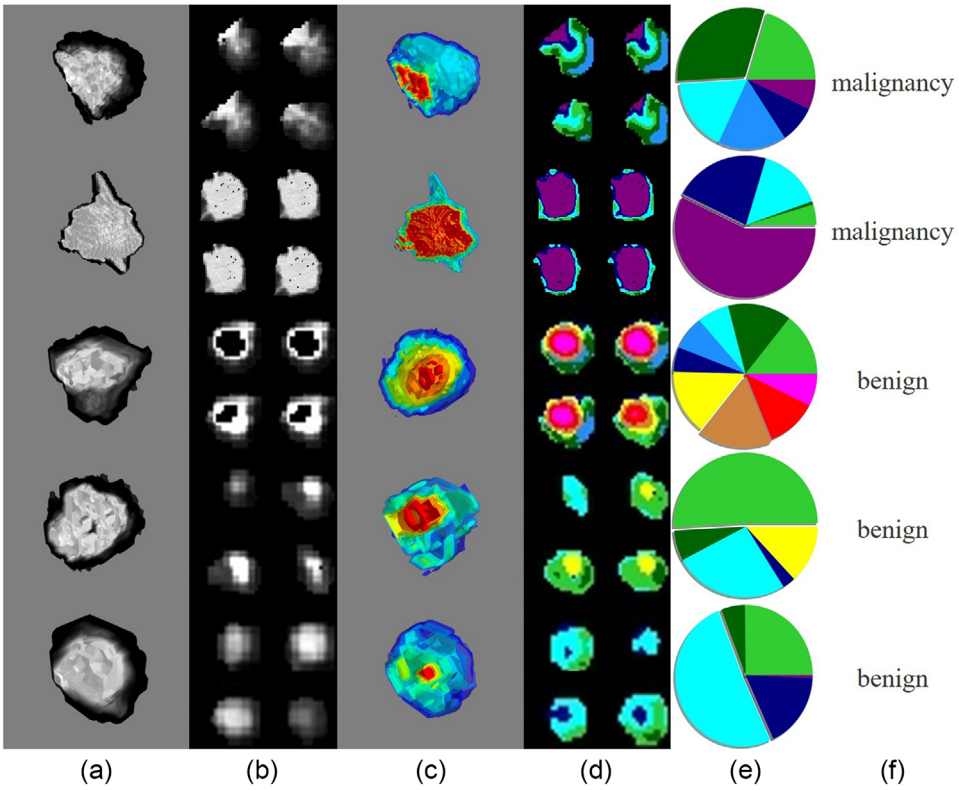

FIGURE 10 Visualization of the intermediate results of the 3D GDC feature extraction process of LIDC-IDRI dataset. (a) The 3D pulmonary nodules; (b) the middle layers in the corresponding 3D pulmonary nodules; (c) the gray density feature descriptors of nodules; (d) the 3D gray density coding level of the middle layers corresponding to (b); (e) the of 3D GDC features; (f) the benign and malignant label of 3D pulmonary nodules

than other state-of-the-art methods. The experimental results show that our method can provide doctors auxiliary analysis in the diagnosis of benign and malignant pulmonary nodules.

ACKNOWLEDGMENTS

The authors greatly appreciate the supports of the Shanghai Pujiang Program (20PJ1402400), the Science and Technology Commission of Shanghai Municipality (20DZ2261200), and the Shanghai Engineer & Technology Research Center of Internet of Things for Respiratory Medicine (20DZ2254400).

CONFLICT OF INTEREST

The author(s) declared no conflicts of interest with respect to the research, authorship, and publication of this paper.

DATA AVAILABILITY STATEMENT

Research data are not shared.

REFERENCES

1. Siegel RL, Miller KD, Goding Sauer A, et al. Colorectal cancer statistics, 2020. *CA Cancer J Clin*. 2020;70(3):145-164.
2. Huang P, Park S, Yan R, et al. Added value of computer-aided CT image features for early lung cancer diagnosis with small pulmonary nodules: a matched case-control study. *Radiology*. 2018;286(1):286-295.
3. Calheiros JLL, de Amorim LBV, de Lima LL, et al. The effects of perinodular features on solid lung nodule classification. *J Digit Imaging*. 2021;34:798-810.
4. Blandin Knight S, Crosbie PA, Balata H, et al. Progress and prospects of early detection in lung cancer. *Open Biol*. 2017;7(9):170070.
5. Inoue M, Nakatsuka S, Jinzaki M. Cryoablation of early-stage primary lung cancer. *Biomed Res Int*. 2014;2014:521691.
6. Cai L, Long T, Dai Y, et al. Mask R-CNN-based detection and segmentation for pulmonary nodule 3D visualization diagnosis. *IEEE Access*. 2020;8:44400-44409.
7. Jiang H, Shen F, Gao F, et al. Learning efficient, explainable and discriminative representations for pulmonary nodules classification. *Pattern Recognit*. 2021;113:107825.
8. Massion PP, Antic S, Ather S, et al. Assessing the accuracy of a deep learning method to risk stratify indeterminate pulmonary nodules. *Am J Respir Crit Care Med*. 2020;202(2):241-249.
9. Zhu H, Han G, Lin C, et al. Two-way MR-forest based growing path classification for malignancy estimation of pulmonary nodules. *IEEE J Biomed Health Inform*. 2021;25(10):3752-3762.
10. Suzuki K, Otsuka Y, Nomura Y, et al. Development and validation of a modified three-dimensional U-Net deep-learning model for automated detection of lung nodules on chest CT images from the lung image database consortium and Japanese datasets. *Acad Radiol*. 2020. https://doi.org/10.1016/j.acra.2020.07.030
11. Tan M, Wu F, Yang B, et al. Pulmonary nodule detection using hybrid two-stage 3D CNNs. *Med Phys*. 2020;47(8):3376-3388.

12. Wu Z, Ge R, Shi G, et al. MD-NDNet: a multi-dimensional convolutional neural network for false-positive reduction in pulmonary nodule detection. *Phys Med Biol*. 2020;65(23):235053.
13. Farhangi MM, Petrick N, Sahiner B, et al. Recurrent attention network for false positive reduction in the detection of pulmonary nodules in thoracic CT scans. *Med Phys*. 2020;47(5):2150-2160.
14. Qin Y, Zheng H, Huang X, et al. Pulmonary nodule segmentation with CT sample synthesis using adversarial networks. *Med Phys*. 2019;46(3): 218-1229.
15. Roy R, Chakraborti T, Chowdhury AS. A deep learning-shape driven level set synergism for pulmonary nodule segmentation. *Pattern Recognit Lett*. 2019;123:31-38.
16. Zheng J, Yang D, Zhu Y, et al. Pulmonary nodule risk classification in adenocarcinoma from CT images using deep CNN with scale transfer module. *IET Image Proc*. 2020;14(8):1481-1489.
17. Yeh YC, Kadota K, Nitadori J, et al. International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society classification predicts occult lymph node metastasis in clinically mediastinal node-negative lung adenocarcinoma. *Eur J Cardiothorac Surg*. 2016;49(1):e9-e15.
18. Le V, Yang D, Zhu Y, et al. Quantitative CT analysis of pulmonary nodules for lung adenocarcinoma risk classification based on an exponential weighted grey scale angular density distribution feature. *Comput Methods Programs Biomed*. 2018;160:141-151.
19. Armato SG III, McLennan G, Bidaut L, et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. *Med Phys*. 2011;38(2):915-931.
20. Al-Shabi M, Shak K, Tan M. 3D axial-attention for lung nodule classification. *Int J Comput Assist Radiol Surg*. 2021;16:1319–1324.
21. Afshar P, Oikonomou A, Naderkhani F, et al. 3D-MCN: a 3D multi-scale capsule network for lung nodule malignancy prediction. *Sci Rep*. 2020;10(1):1-11.
22. Xie Y, Zhang J, Xia Y, et al. Fusing texture, shape and deep model-learned information at decision level for automated classification of lung nodules on chest CT. *Inform Fusion*. 2018;42:102-110.
23. Yan X, Pang J, Qi H, et al. Classification of lung nodule malignancy risk on computed tomography images using convolutional neural network: a comparison between 2d and 3d strategies. In: Asian Conference on Computer Vision. Springer, Cham, 2016:91-101.
24. Shen W, Zhou M, Yang F, et al. Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification. *Pattern Recognit*. 2017;61:663-673.
25. Xie Y, Xia Y, Zhang J, et al. Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT. *IEEE Trans Med Imaging*. 2018;38(4):991-1004.
26. Wang D, Zhang T, Li M, et al. 3D deep learning based classification of pulmonary ground glass opacity nodules with automatic segmentation. *Comput Med Imaging Graph*. 2021;88:101814.
27. Ren Y, Tsai MY, Chen L, et al. A manifold learning regularization approach to enhance 3D CT image-based lung nodule classification. *Int J Comput Assist Radiol Surg*. 2020;15(2):287-295.
28. Tong C, Liang B, Su Q, et al. Pulmonary nodule classification based on heterogeneous features learning. *IEEE J Sel Areas Commun*. 2020;39(2):574-581.
29. Hartigan JA, Wong MA. Algorithm AS 136: a k-means clustering algorithm. *J R Stat Soc Ser C Appl Stat*. 1979;28(1):100-108.
30. Liaw A, Wiener M. Classification and regression by random forest. *R News*. 2002; 2(3):18–22.
31. Clark K, Vendt B, Smith K, et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. *J Digit Imaging*. 2013;26(6):1045-1057. https://doi.org/10.1007/s10278-013-9622-7.
32. Reeves AP, Biancardi AM, The Lung Image Database Consortium (LIDC) Nodule Size Report. Release: 2011-10-27-2. At: /lidc/, October 27, 2011

How to cite this article: Zheng BB, Yang D, Zhu Y, Liu Y, Hu J, Bai C. 3D gray density coding feature for benign-malignant pulmonary nodule classification on chest CT. *Med Phys*. 2021;48:7826–7836.
https://doi.org/10.1002/mp.15298