

Retraction

Retracted: Modeling of Hidden Markov in Ultrasound Image-Assisted Diagnosis

Journal of Healthcare Engineering

Received 23 May 2023; Accepted 23 May 2023; Published 24 May 2023

Copyright © 2023 Journal of Healthcare Engineering. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1].

This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process. Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] L. Shao, Z. Zhou, H. Wu, J. Ni, and S. Li, "Modeling of Hidden Markov in Ultrasound Image-Assisted Diagnosis," *Journal of Healthcare Engineering*, vol. 2021, Article ID 5597591, 10 pages, 2021.

Research Article

Modeling of Hidden Markov in Ultrasound Image-Assisted Diagnosis

Liping Shao ¹, Zubang Zhou ¹, Hongmei Wu ², Jinrong Ni ³, and Shulan Li ¹

¹Department of Ultrasound, Gansu Provincial Hospital, Lanzhou, Gansu 730000, China

²Department of Ultrasound, Yinchuan Women and Maternal and Child Health Hospital, Ningxia 750000, China

³Department of Cardiac Surgery, The First Clinical Medical School of Lanzhou University, Lanzhou 730030, China

Correspondence should be addressed to Shulan Li; vivian138lan@163.com

Received 2 February 2021; Revised 25 February 2021; Accepted 3 April 2021; Published 12 April 2021

Academic Editor: Zhihan Lv

Copyright © 2021 Liping Shao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Different segmentation of lung nodules using the same segmentation algorithm can easily lead to excessive segmentation errors. Therefore, it is necessary to design an effective segmentation algorithm to improve image segmentation accuracy. Based on the hidden Markov model, this study processed the ultrasound images of pulmonary nodules to improve their diagnostic results. At the same time, this study was combined with the ultrasound image of lung nodules to process the ultrasound images. In addition, this study combines the convex hull algorithm for image processing, uses the improved vector method to repair, improves image recognizability, establishes a reliable feature extraction algorithm, and establishes a comprehensive diagnostic model. Finally, this study designed the test for performance analysis. Through experimental research, it can be seen that the model constructed in this study has certain clinical effects and can provide theoretical reference for subsequent related research.

1. Introduction

At present, there are more and more patients with clinical pulmonary nodules. The 1–30 mm lung nodules found by X-ray or chest CT account for 0.2% or 1%, respectively. Most of them are benign nodules, and malignant knots account for 20%. Meanwhile, as the size of nodules increases, the malignant rate fluctuates between 3% and 80%. When the diameter of nodules is between 4 and 7 mm, the incidence of malignant nodules is 1%; when the nodules are 8–20 mm in diameter, malignant nodules account for 18%; when the diameter of nodules is greater than 20 mm, the malignant rate is as high as 50%. At present, the nature of pulmonary nodules is determined by chest CT-guided percutaneous lung biopsy, bronchoscopy, or even open lung biopsy to confirm the pathology. These invasive procedures have the disadvantages of high risk and uneconomical and even bring unnecessary harm to patients with benign pulmonary nodules.

Since the 1990s, lung nodule detection based on CT images has gradually become an important research content

in the field of computer-aided diagnosis. Moreover, many universities and research institutions at home and abroad have carried out a lot of experimental work on the construction of lung nodule detection models.

In the exploration of research methods, many scholars took pulmonary nodules as the research object to conduct image segmentation. Wang et al. [1] first segmented the lung parenchyma based on the regional growth method and isolated the ROI from the lung parenchyma according to the Gaussian mixture model and the Hessian matrix. Then, they selected Tsallis entropy and Shannon entropy as descriptive features and used the support vector machine (SVM) to classify lung nodules and nonnodular regions. Meanwhile, José et al. [2] used clustering coarse segmentation using growing neural gas (GNG). Then, they separated the lung nodules from the tissues containing blood vessels and bronchus according to the three-dimensional distance transformation. Finally, the SVM was used to effectively identify the lung nodules based on the extracted shape features and texture features. Furthermore, Jacobs et al. [3] integrated fuzzy threshold, Gaussian matrix, mean

curvature, Hessian matrix, and other algorithms to extract the ROI. Then, they use local shape features and local divergence information as the feature expression of ROI and, finally, use weighted SVM to identify and detect lung nodules. Perandini et al. [4] performed segmentation of pulmonary nodules based on vascular and nodule enhancement filters proposed by Li et al. [5]. Then, they locate the cluster center of the lung nodules based on the divergence calculated based on the Gaussian template to achieve ROI extraction. Finally, the genetic algorithm-based classifier, artificial neural network (ANN), SVM, and three kinds of classifiers were used to compare and analyze the detection effect of lung nodules. Kim et al. [6] extracted the ROI by using the region growth algorithm and morphological operation and then reconstructed the spline surface based on the 3D spring model, which facilitated the extraction of relevant 3D gray features and shape features, and finally used ANN to detect the lung nodules. Guedes et al. [7] proposed a method of genetic algorithm template matching (GATM), which is first used for the rough detection of pulmonary nodule region, and then, the nodule misjudgment region is eliminated by combining the shape of the region and the gradient characteristic rule of the gray scale. Boididou et al. [8] first constructed a probabilistic statistical model based on lung nodule data and used the model to determine whether candidate nodules in CT images were connected to other lung tissues. They then form a fully automated 3D pixel segmentation of the pulmonary nodules, blood vessels, thoracic cavity, and lung parenchyma through a unified feature set and classifier and use a probabilistic classifier to determine the relationship between the candidate node and the lung tissue structure. Plapous et al. [9] first separated the lung parenchyma from the CT image, calculated the statistical parameters (such as mean, standard deviation, skewness, peak, fifth standard deviation, and sixth central moment) from the segmentation results, and finally obtained better classification results using BP neural network.

According to the research status of the pulmonary nodule detection algorithm, although the construction method of the lung nodule CAD model is different, it basically includes the following three important steps: (1) ROI segmentation of pulmonary nodules, (2) feature extraction and selection, and (3) classification recognition of candidate regions. Among them, ROI segmentation is the premise of feature level fusion, which provides an important research object for further research of feature level fusion. However, there are still some shortcomings in the ROI segmentation algorithm of lung nodules. The main reason is that most segmentation algorithms based on two-dimensional images will lose the spatial structure information of ROI, which is not conducive to the extraction of three-dimensional features. Moreover, the segmentation algorithm is designed with less consideration for the adjacent relationship between the lung nodules and other tissues, and the same segmentation algorithm for different types of lung nodules is likely to cause excessive segmentation errors. Therefore, it is necessary to design an effective segmentation algorithm to improve the segmentation accuracy of ROI. Feature extraction and selection and classification recognition of

targets can be attributed to the main content of image feature level fusion method. In order to further analyze the problems in the detection algorithm of pulmonary nodules, based on the hidden Markov model, this study studied the ultrasound images of pulmonary nodules and strived to improve the application of this model in the diagnosis of pulmonary nodules.

2. Research Method

2.1. Segmentation Process of Lung Nodule Images. The process of segmentation of lung parenchyma is mainly divided into three steps: initial contour segmentation of the lung, lung contour repair, and lung parenchymal extraction. The overall process of the algorithm is shown in Figure 1 [10].

- (1) Binary image: in the CT image of the lung, it includes not only high-density areas such as bones, soft tissues, and blood vessels around the chest but also low-density areas such as lung parenchyma and trachea, as well as background interference areas such as scan beds and clothes outside the chest. Objects or tissues of different densities have different CT values. Therefore, if a uniform threshold is used for segmentation, the effect of lung image obtained is poor and has a greater influence on subsequent processing. Therefore, the original lung image is binarized using the OTSU threshold method.
- (2) Remove trachea and other backgrounds: first, the binary image is subjected to a region filling hole operation, and the background region (lung parenchyma and trachea, etc.) inside the white portion is filled with white. Next, the filled image is subtracted from the original binary image to obtain a binary image including lung parenchyma and trachea and other interference items, as shown in Figure 2 [11].

For a point set S on a plane, the convex hull is the smallest convex polygon that can contain all points in the point set S . The most widely used is the Graham scanning method, and the algorithm is shown in Figure 3 [12].

The 2D convex hull algorithm mainly consists of 3 steps:

Step 1: start selection and point set sorting. For a set of points with n points on the plane, we first select the point with the smallest Y coordinate. If there are multiple points, we select the point with the smallest X coordinate as the starting point, i.e., P_0 point. After that, the algorithm scans all other points counterclockwise from the starting point for sorting, sorting from the positive angle between the point P_i ($i = 1, 2, 3, \dots, n-1$) to the point P_0 and the X -axis from small to large. If there is the same angle, the point with a small Y coordinate is preferred. After sorting, the point set is [13]

$$S = \{P_0, P_1, \dots, P_{n-1}\}. \quad (1)$$

Step 2: according to the sort order, the points are taken Graham scan. It is known that P_0 is a bump, and it is

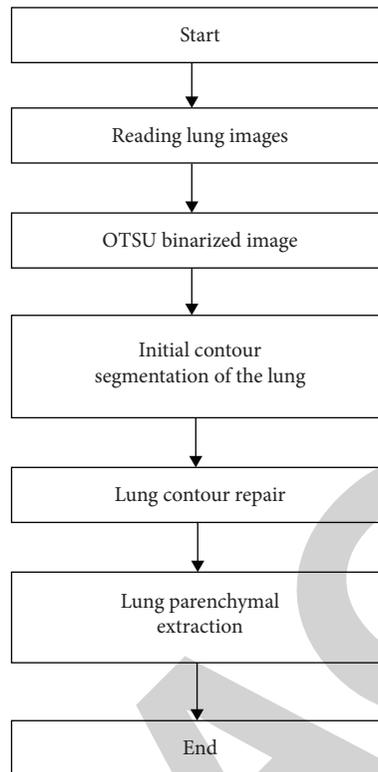


FIGURE 1: Flowchart of segmentation of lung parenchyma.

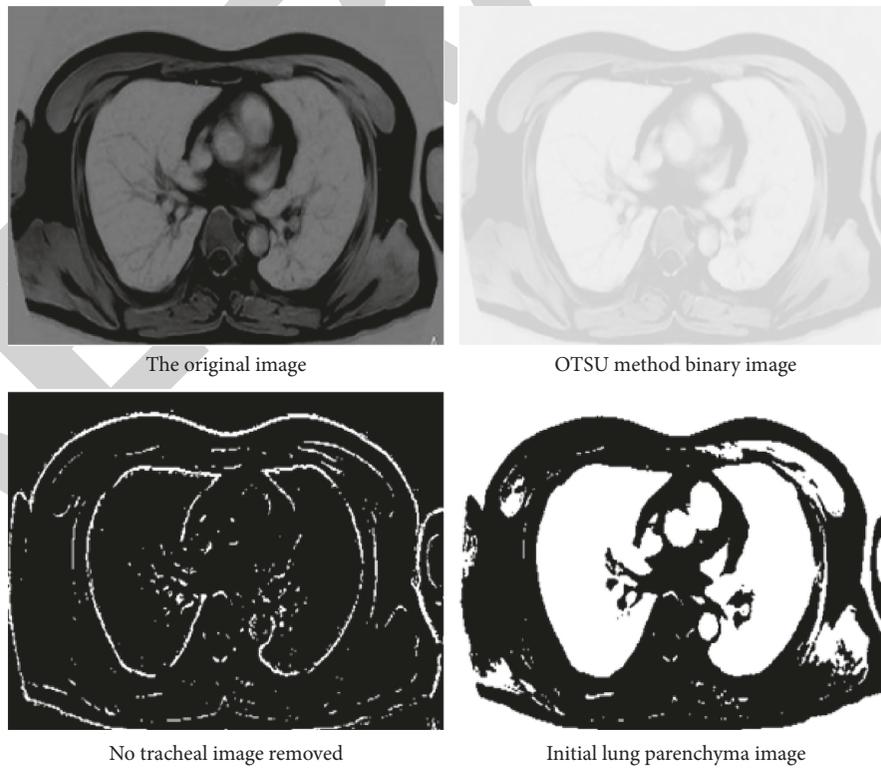


FIGURE 2: Initial contour segmentation of the lungs.

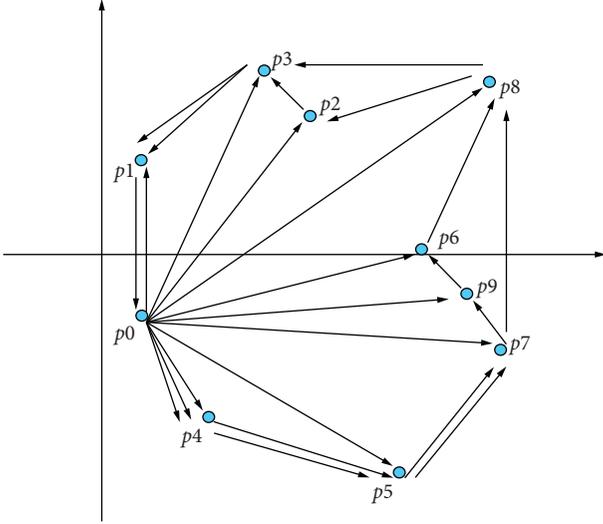


FIGURE 3: Schematic diagram of convex hull algorithm.

first determined whether P_1 is a bump. The cross-multiplication result of vector P_1, P_0 and P_2, P_1 formed by P_2, P_1 , and P_0 are calculated to determine whether P_1 is a bump, and the condition is [14]

$$P_1P_0 \times P_2P_1 = X_1Y_2 - X_2Y_1 \geq 0. \quad (2)$$

In the formula, $P_1P_0 = (X_1, Y_1), P_2P_1 = (X_2, Y_2)$. If P_1 is a bump, the algorithm then determines if the next point P_2 is a bump, i.e., $P_3P_2 \times P_2P_1 \geq 0$ and so on.

Step 3 (backtracking operation): if the current point P_k is not a bump, that is, $P_{k+1}P_k \times P_kP_{k-1} < 0$ ($k = 1, 2, 3, \dots, n-2$), the algorithm performs a backtracking operation: The vector cross-multiplication $P_{k+1}P_{k-1} \times P_{k-1}P_{k-2}$ ($k = 2, 3, \dots, n-2$) of the previous point P_{k-1} of the current point is calculated. If P_{k-1} is not a bump, then the vector cross-multiplication $P_{k+1}P_{k-2} \times P_{k-2}P_{k-3}$ of the previous point P_{k-2} of P_{k-1} is determined and so on. Until the calculation point is a bump, the backtracking ends. After that, the concavity of the next point P_{k+1} is recalculated until all points are considered. All the bumps obtained are two-dimensional convex hull point sets, and the sequential connection point sets obtain two-dimensional convex hulls [15].

In this paper, the segments to be repaired are divided into two categories according to the number N of points between the two convex hull points of the segment to be repaired. The number of points is less than the threshold N_0 which is the first type and is mainly the outer contour depression and can be directly repaired by directly connecting the two convex points. The number of points is larger than the threshold N_0 which is the second type and is mainly the inner contour depression and cannot be directly repaired. In this paper, the improved vector bump method is used to repair the internal cavity. Among them, the empirical value is $N_0 = 120$ [16].

2.2. Pulmonary Parenchymal Repair of Improved Vector Method. Aiming at the problem that the convex hull algorithm is used to repair the crossover effect of the internal contour of the lung parenchyma, this paper uses the improved vector method to repair. For simple polygons, it is a simpler method to use the cross-multiplication of vectors to determine the concavity and convexity of polygon vertices. The vertex concavity is judged by calculating the crossover of the two vectors a and b formed by the two points to be processed and the two adjacent points [17].

The bumps obtained by the original vector product are processed.

Step 1: when the number of points between two adjacent bumps (points with r greater than 0) is $n \geq 0$, the sum of the vector crosses of all the points between the two points is calculated. If $\text{sum} \geq 0$, the bump is retained, otherwise it is not retained.

Step 2: when the number of adjacent two bumps is $n = 0$, one of the bumps is reserved. According to experimental experience, the algorithm retains the next bump, as shown in Figure 4 [18].

Step 3: on the basis of Steps 1 and 2, one of the two bumps in the remaining convex points is rounded off. This step is used to remove the situation where two bumps may be closer in the steps of Steps 1 and 2. For example, the vector product of consecutive 3 points is greater than zero. According to the experimental experience, it is judged whether or not a bump is discarded by the number of points between two adjacent bumps, and the empirical value is $M_0 = 5$. If it is less than the threshold M_0 , the previous bump is discarded, and the next bump is retained, otherwise no operation is performed. After the above three steps, a new bump is obtained.

2.3. Feature Extraction of This Study. This paper is mainly for the automatic detection of isolated solid nodules and adhesion pleural nodules. Therefore, in the feature extraction, only the two types of nodules are extracted. The extracted features in this paper are mainly divided into gray features, geometric features, shape features, texture features and so on, as shown in Table 1 [19].

- (1) Grayscale features (gray mean M and variance σ^2): assuming that $I(i, j)$ is the gray value of an arbitrary point in the image, and N is the length and width of the image, the corresponding image gray mean and gray variance calculation formulas are as follows [20]:

$$M = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N I(i, j), \quad (3)$$

$$\sigma^2 = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N I(i, j). \quad (4)$$

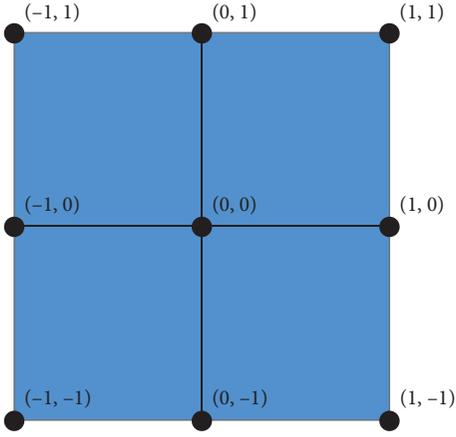


FIGURE 4: 3 * 3 coordinate point map.

TABLE 1: Feature extraction of this study.

Feature category	Specific characteristics
Grayscale feature	Gray mean, gray variance, and gray entropy
Geometric features	Area, diameter, and perimeter
Shape feature	Long axis, short axis, linearity, eccentricity, convexity, and squareness
Texture feature	Energy, contrast, and inverse moment

Gray level entropy H : the gray level entropy is represented by H . Image entropy is a statistical form of feature that reflects the amount of average information in an image. Moreover, the one-dimensional entropy of an image represents the amount of information contained in the aggregated features of the grayscale distribution in the image. Assuming that P_i is set to represent the proportion of pixels whose gray value is i in the image, the formula for calculating the unary gray entropy of the gray image is as follows:

$$H = - \sum_{i=0}^{255} P_i \log P_i. \quad (5)$$

Among them, P_i is a gray value corresponding to the probability of occurrence in the image, which is obtained by gray histogram statistics.

- (2) Geometric features: the diameter of the region is indicated by D , which is defined herein as the diameter of a circle having the same area as the region of interest. $D = \sqrt{4 * \text{Area}/P_i}$ can be obtained by deriving the area A .

Area perimeter L : currently, there are many ways to calculate the perimeter of a region. Among the methods of chain code, Freeman chain code is the most widely used. In this paper, by extracting the contour boundary of the region of interest, the number of boundary points is equivalent to the perimeter of the equivalent replacement region, which achieves the purpose of fast calculation.

- (3) Shape feature (long axis h and short axis o): the main axis extracted in this paper is the long axis length of the candidate region based on the regionprops method in MATLAB, which is defined as the long axis length of the ellipse with the same standard second-order central moment as the region (in the sense of pixels). The minor axis is defined as the short axis length of the ellipse having the same standard second-order center moment as the region (in the sense of pixels). The long and short axes are represented in the area as shown in Figure 5.

When calculating the long and short axes, we first define the $p + q$ order center moment of the area image as μ_{pq} , and the calculation formula is

$$\mu_{pq} = \sum_i \sum_j (i - i_0)^p (j - j_0)^q I(i, j). \quad (6)$$

Among them, i_0 and j_0 are the centroid abscissa and ordinate of the region, respectively. Then, the major and minor axes of the regional equivalent ellipse are

$$h = 2 \left[\frac{2 \left(\mu_{20} + \mu_{02} + \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2} \right)}{\mu_{00}} \right]^{1/2}, \quad (7)$$

$$o = 2 \left[\frac{2 \left(\mu_{20} + \mu_{02} - \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2} \right)}{\mu_{00}} \right]^{1/2}. \quad (8)$$

The regional linearity x is also known as the aspect ratio or the fine length. The length and width are, respectively, the length and width of the smallest rectangle MER of the candidate region, and the corresponding calculation formula is as follows:

$$X = \frac{W_{\text{MER}}}{H_{\text{MER}}}. \quad (9)$$

This feature was chosen primarily for the differentiation of tubular vessels and round nodules. The eccentricity feature is defined as the eccentricity of an ellipse with the same standard second-order central moment as the region, that is, the ratio of the major axis to the minor axis of the ellipse. Among them, the length of the major axis is a and the length of the minor axis is b :

$$\text{Ec} = \frac{a}{b}. \quad (10)$$

The theory shows that if a and b are closer, that is, the ellipse having the same standard second-order central moment as the region is closer to a circle, the Ec value is closer to 1. However, it does not necessarily mean that the candidate area is close to a circle. The calculation formula is

$$\text{So} = \frac{\text{Area}}{\text{ConvexArea}}. \quad (11)$$

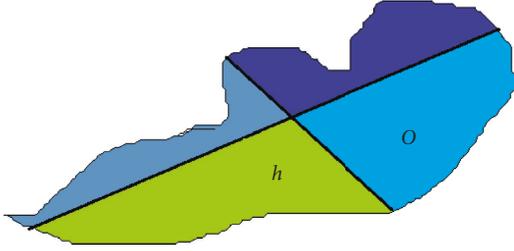


FIGURE 5: The long and short axes of the area graphic.

ConvexArea is the smallest convex polygon area, and Area is the region area. At the same time, the pixel ratio Ex in the region and its minimum boundary rectangle, also known as the rectangle, is calculated as

$$Ex = \frac{\text{Area}}{\text{BoxArea}}. \quad (12)$$

Among them, Area is the region area, BoxArea is the minimum bounding rectangle area, and Ex actually reflects the extent of the area's extended range. Roundness S: roundness is one of the commonly used methods for measuring circularity, reflecting the proximity of the region to the circle, which is defined as

$$S = \frac{4\pi\text{Area}}{L^2}. \quad (13)$$

When $S = 1$, the shape is a circle, and the larger the S, the more irregular the shape. The Fourier descriptor has deformation independence such as translation, rotation, and scaling and has better robustness when representing object features. In this paper, based on MATLAB's countercontour extraction algorithm, the region contour is obtained and stored in a two-dimensional matrix. After that, the coordinates of the contour boundary points are represented by complex numbers and then converted and stored in the one-dimensional array T to facilitate the late fast Fourier transform. The coordinate complex number is expressed as follows:

$$t_i = x_i + jy_i. \quad (14)$$

Among them, j is a complex symbol and i is a corresponding boundary point number. The Fourier transform is performed on the T array (in this case, the fast Fourier transform is used), and the transform is a one-dimensional Fourier transform, that is, the time domain signal is converted into a frequency domain signal. The corresponding transformation formula is

$$a(u) = \frac{1}{N} \sum_{i=0}^{N-1} t_i e^{-j2\pi ui/N}. \quad (15)$$

In the formula, N is the number of boundary points, $u = 0, 1, \dots, N-1, \dots, 1$. Since the data obtained by fft

is not symmetrical with the frequency, the fftshift operation is performed to make the images of the positive semiaxis portion and the negative half-axis portion symmetrical about the respective centers. After that, the DC component of fft is moved to the center of the spectrum. That is, assuming that the spectral image is divided into four parts by one horizontal line and one vertical line, the fftshift operation exchanges the divided four-part diagonal and the diagonal matrix.

- (4) Texture features: the methods for extracting image texture features can be roughly divided into four categories: statistical-based methods, model-based methods, signal processing-based methods, and structural analysis-based methods. The first three categories are more commonly used.

There are more than a dozen image features extracted based on GLCM, but there are many features of correlation. According to the research by Ulaby and Baraldi et al., several of the features with the strongest discriminative ability are extracted, which are contrast, energy, and correlation.

- (1) The energy is represented by ASM, that is, the sum of the squares of the probability values of all the elements in the gray level co-occurrence matrix, which is a reflection of the uniformity in the gray value distribution of the image and the thickness of the texture. If all values of the co-occurrence matrix are equal, the energy value is small. If some values are large and other values are small, the corresponding energy value is larger. When the element distribution is concentrated, the corresponding energy value is larger, and the corresponding calculation formula is as follows:

$$\text{ASM} = \sum_i \sum_j P(i, j)^2. \quad (16)$$

- (2) Contrast, also known as contrast, is represented by CON, which is a reflection of whether the image is clear and the depth of the groove. Among them, if the groove of the texture is deeper, the corresponding contrast will be larger, and the corresponding visual effect will be clearer. Conversely, if the grooves are shallow, the contrast will be smaller, and the effect will be more blurred. The corresponding calculation formula is as shown in (10):

$$\text{CON} = \sum_i \sum_j (i - j)^2 P(i, j). \quad (17)$$

- (3) Correlation, also known as the inverse moment, is represented by IDM, which reflects the homogeneity of the image texture and measures how much the image texture changes locally. A large value indicates a lack of variation between different regions of the image texture, and the locality is very uniform. The corresponding calculation formula is as shown in (11):

$$\text{IDM} = \sum_i \sum_j \frac{P(i, j)}{1 + (i - j)^2} \quad (18)$$

In order to describe the texture information of the candidate nodule part obtained by segmentation more accurately, the candidate region is segmented again by the method of minimum enclosing rectangle, which performs texture feature extraction for the smallest rectangular region instead of texture analysis for the whole image.

3. Results

In order to reduce the computational complexity of the convex hull algorithm, the obtained initial lung region is subjected to boundary tracking, and the boundary is implemented by a two-dimensional convex hull algorithm. The obtained convex hull point set is as shown in Figure 6(a). The red circle in the figure indicates the convex point set and the blue is the initial lung outline. In order to determine the recessed area of the lung contour to be repaired, the ratio of the Euclidean distance between the adjacent two convex hull points and the arc length between the two points is calculated for the initial lung contour convex blunt point set. The experience threshold $K_0 = 9.67$ is selected to obtain the contour depression between the two convex hull points that need to be repaired, which is defined as the segment to be repaired, as shown in Figure 6(b). The section to be repaired has a total of 4 segments, which are determined by 8 convex hull points, and is represented by red circles in Figure 6.

In this paper, the vector method is improved. First, the extracted bumps are divided into two categories. One type is a bump that has a significant effect on repair and is defined as a necessary bump, mainly concentrated on the contour point, as shown in Figure 7(a), the letters A, B, etc. The other type of bumps that have no significant effect on repair is defined as unnecessary bumps, mainly concentrated on jagged boundaries, as indicated by the white arrow in Figure 7(b). In order to reduce the amount of calculation, we need to remove unnecessary bumps as much as possible while retaining the necessary bumps.

The patched boundary is shown in Figure 8.

Since the GLCM is a pixel-level calculation, it takes too much time for the entire image to be calculated, and the calculation task is large. The main factors affecting the calculation amount are three parameters: image gray level G , two-pixel point distance D , and direction θ .

In this paper, the parameters are set from these three aspects to achieve the purpose of reducing the amount of calculation. First, the CT grayscale image of the lung is compressed to 8 levels, and the distance D is selected to be 1. Since the direction is symmetrical, this article is to ensure that the main direction is considered. To reduce the amount of calculation, the horizontal direction (i.e., 0°) and the two directions are selected. The corresponding vector parameters are $[0 \ D]$ and $[-D \ D]$ as shown in Figure 9.

The data used in the feature extraction phase of this paper is based on the results of candidate nodule extraction. After some candidate nodules were removed by layering

method, 20 features including gradation features, morphological features, geometric features, and texture features of 710 candidate nodule samples were extracted. Among them, the number of samples containing true nodules (i.e., positive samples) is 95, and the number of pseudonodule samples (i.e., negative samples) is 615. Some candidate nodule features are shown in Table 2.

4. Discussion and Analysis

For the lung contour depression caused by nodules, based on the traditional two-dimensional convex hull algorithm, this paper uses the improved vector bump method to focus on repairing the internal contour depression of the lung and compares it with the traditional convex hull algorithm and the widely used corner detection method. Corner detection algorithm patching: the corner point is the point where the gray level changes sharply or the curvature value of the edge curve in the two-dimensional image is the important object feature of the image, which plays a key role in pattern recognition and analysis. In this paper, the corner point detection of the lung contour is extracted based on the global and local curvature values. The algorithm can accurately extract good and rough corners, and the calculation amount is small, which is widely used in corner detection.

In order to extract the relevant features of the candidate region, the calculation of the regional gray mean and variance is performed for each candidate nodule region obtained after the candidate nodule segmentation and is based on the gray histogram statistics in the region. It should be noted that when calculating the gray value or variance of a certain region, since the gray value of the background region is zero, all the points with the gray value of zero need to be removed when performing gray scale statistics on the entire image, thereby obtaining the histogram statistics of the corresponding candidate nodule regions.

Figure 7(a) shows the bumps of the inner contour of the initial lung found using the original vector bump method. The red circle in the figure is the obtained bump, and the area marked by the white tip indicates that the bump is too concentrated, which is often caused by noise and the like. The resulting initial pulmonary contour is shown in Figure 7(b). The white arrow is partially jagged and processed by the improved vector method, which successfully removes these unnecessary bumps while preserving the necessary bumps for repair. In this test data, the number of contour points on the inner side of the initial lung is 331, and the original vector bump method finds 88 bumps. However, the improved vector bump method removes unnecessary bumps that have no significant effect on the repair and obtains 16 bump numbers, which greatly reduces the amount of subsequent processing calculations.

Through the above analysis, three laws can be drawn. Rule one: the vector consisting of two adjacent coordinates of the lung parenchyma contour in the two-dimensional plane has only eight cases of $(-1, 1)$, $(0, 1)$, $(1, 1)$, $(1, 0)$, $(1, -1)$, $(0, -1)$, $(-1, -1)$, and $(-1, 0)$, that is, within the range of t points centered on the point, as shown in Figure 6. Among them, $(0, 0)$ is the point to be processed and the rest

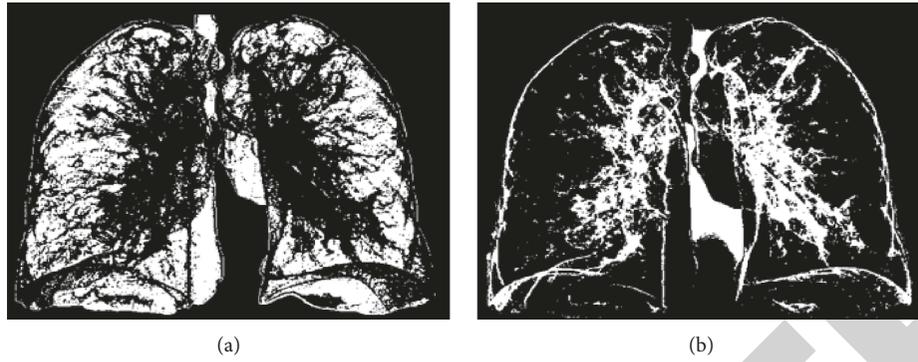


FIGURE 6: Results of the convex hull algorithm. (a) Initial lung wide convex bulge. (b) Convex segment to be repaired.



FIGURE 7: Improved vector bump method. (a) Original vector result. (b) Improved vector method results.

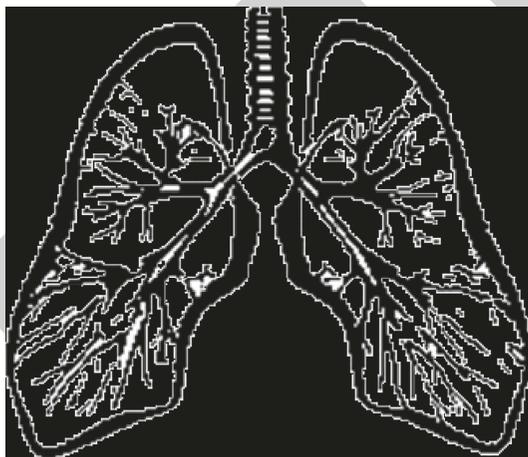


FIGURE 8: Repair results of the improved vector method.

are the adjacent points of the point. The result r of the cross-multiplication between the two points of the pending point and the two adjacent points is only $-2, -1, 0, 1, 2$. Among them, r is greater than zero, which is a bump, r is less than zero, which is a pit, and r is equal to zero, which is a three-point collinear. Rule 2: unnecessary bumps mainly appear on jagged boundaries. The distribution of the corresponding

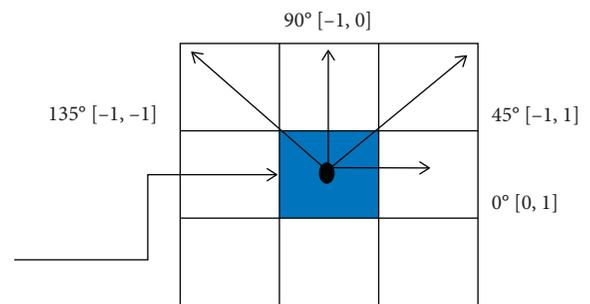


FIGURE 9: Parameter diagram of the direction vector.

vector product result r satisfies $r > 0$, $r < 0$, and $r = 0$. That is, when the vector product of the points between adjacent two bumps is less than zero, a depression is formed. Among them, $r > 0$ indicates that the point is a bump, $r < 0$ indicates that the point is a pit, and $r = 0$ indicates that the point is collinear with two adjacent points. Rule 3: the necessary bumps generally appear in pairs, and the number of points between the two necessary bumps is zero or greater than zero. When the number is greater than zero, the vector product r value of the corresponding point is zero, that is, the points between the two necessary bumps only appear collinear, and there is no recess.

TABLE 2: Partial candidate nodule features.

	Eccentricity	Convexity	Circularity	Compactness	Entropy	Boundary delineator
ROI	0.527	0.945	0.775	8.762	0.002	0.072
ROI	0.897	0.432	0.233	86.165	0.053	0.004
ROI	0.912	0.940	0.760	13.386	0.002	0.052
ROI	0.953	0.780	0.396	25.127	0.004	0.239
ROI	0.825	0.950	0.874	7.752	0.001	0.093
ROI	0.857	0.964	0.708	11.998	0.002	0.058
ROI	0.921	0.777	0.563	17.953	0.002	0.042

In this paper, the lung parenchymal contour repair process has also been improved and is mainly divided into four steps:

Step 1: according to the bump extracted by the improved vector method, the Euclidean distance from one bump to the other bump is calculated from the first bump. If the Euclidean distance is less than the set threshold D , the algorithm proceeds to the second step. If the bump to be processed reaches the nearest bump or the distance to the other bump is greater than the threshold D , the next bump of the to-be-processed bump is selected to perform Step 1 operation again.

Step 2: the ratio t of the arc length to the distance between the bump to be processed and the closest bump is calculated. If $t > T$, the two bumps need to be repaired, and the points on the original boundary between the two bumps are stored in an intermediate variable m , otherwise no repair is needed. Among them, T is the empirical threshold and takes the value $T = 1.3$.

Step 3: the algorithm continues to calculate the ratio t between the arc length and the distance of the bump to be processed and all other bumps satisfying the condition and performs Step 2 operation. Until the distance between the to-be-processed bump and the i th bump is greater than the set threshold D , the algorithm stops and moves to Step 1 to reselect the next pending point.

Step 4: the points in the intermediate variable m are deleted on the boundary of the initial lung contour, and the obtained new coordinate points are connected by straight lines to form a patched edge.

5. Conclusion

Based on the hidden Markov model, this study studied the ultrasound images of pulmonary nodules and sought to improve the application of this model in the diagnosis of pulmonary nodules. The process of segmentation of lung parenchyma is mainly divided into three steps: pulmonary initial contour segmentation, lung contour repair, and lung parenchymal extraction. Aiming at the problem that the convex hull algorithm is used to repair the crossover effect of the internal contour of the lung parenchyma, this paper uses the improved vector method to repair. For simple polygons, it is a simpler method to use the cross-multiplication of vectors to determine the concavity and convexity of polygon

vertices. In addition, this paper is mainly for the automatic detection of isolated solid nodules and adhesion pleural nodules, so in the feature extraction, only the two types of nodules are extracted. Moreover, the extracted features in this paper are mainly divided into grayscale features, geometric features, shape features, texture features, and so on. Meanwhile, in order to describe the texture information of the candidate nodule part obtained by segmentation more accurately, the candidate area is segmented again by the method of minimum enclosing rectangle, and texture feature extraction is performed for the smallest rectangular area instead of texture analysis for the whole image. The research results show that the algorithm proposed in this study has certain effects on the ultrasound image recognition of pulmonary nodules, which can be gradually applied to clinical practice.

Data Availability

The data used to support the findings of the study cannot be shared as the authors were not given permission.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Liping Shao and Zubang Zhou have contributed equally to this work.

References

- [1] B. Wang, X. Tian, Q. Wang et al., "Pulmonary nodule detection in CT images based on shape constraint CV model," *Medical Physics*, vol. 42, no. 3, pp. 1241–1254, 2015.
- [2] R. José, M. C. Oliveira, and P. M. D. Azevedo-Marques, "Cloud-based NoSQL open database of pulmonary nodules for computer-aided lung cancer diagnosis and reproducible research," *Journal of Digital Imaging*, vol. 29, no. 6, p. 716, 2016.
- [3] C. Jacobs, E. M. Van Rikxoort, K. Murphy, M. Prokop, C. M. Schaefer-Prokop, and B. van Ginneken, "Computer-aided detection of pulmonary nodules: a comparative study using the public LIDC/IDRI database," *European Radiology*, vol. 26, no. 7, pp. 2139–2147, 2016.
- [4] S. Perandini, G. A. Soardi, M. Motton et al., "Enhanced characterization of solid solitary pulmonary nodules with Bayesian analysis-based computer-aided diagnosis," *World Journal of Radiology*, vol. 8, no. 8, pp. 729–734, 2016.

- [5] A. K. Dhara, S. Mukhopadhyay, A. Dutta et al., "Content-based image retrieval system for pulmonary nodules: assisting radiologists in self-learning and diagnosis of lung cancer," *Journal of Digital Imaging*, 2017.
- [6] J. Kim, D.-H. Jeong, D. Lee, and H. Jung, "User-centered innovative technology analysis and prediction application in mobile environment," *Multimedia Tools and Applications*, vol. 74, no. 20, pp. 8761–8779, 2015.
- [7] Á. L. V. Guedes, R. G. d. A. Azevedo, and S. D. J. Barbosa, "Extending multimedia languages to support multimodal user interactions," *Multimedia Tools and Applications*, vol. 76, no. 4, pp. 5691–5720, 2017.
- [8] C. Boididou, S. E. Middleton, Z. Jin et al., "Verifying information with multimedia content on twitter: a comparative study of automated approaches," *Multimedia Tools & Applications*, no. 6, pp. 1–27, 2017.
- [9] C. Plapous, S. A. Berrani, B. Besset et al., "A low-complexity audio fingerprinting technique for embedded applications," *Multimedia Tools and Applications*, vol. 77, no. 2, 2017.
- [10] A. Álvarez, C. Mendes, M. Raffaelli et al., "Automating live and batch subtitling of multimedia contents for several European languages," *Multimedia Tools and Applications*, vol. 75, no. 18, pp. 10823–10853, 2016.
- [11] A. K. Sangaiah, P. R. Subramaniam, and X. Zheng, "A combined fuzzy DEMATEL and fuzzy TOPSIS approach for evaluating GSD project outcome factors," *Neural Computing and Applications*, vol. 26, no. 5, pp. 1025–1040, 2015.
- [12] U. Ayegül, Y. Demir, Güzeli, and Cüneyt, "A new facial expression recognition based on curvelet transform and online sequential extreme learning machine initialized with spherical clustering," *Neural Computing and Applications*, vol. 27, no. 1, pp. 131–142, 2016.
- [13] A. K. Sangaiah, J. Gopal, A. Basu et al., "An integrated fuzzy DEMATEL, TOPSIS, and ELECTRE approach for evaluating knowledge transfer effectiveness with reference to GSD project outcome," *Neural Computing & Applications*, vol. 1, pp. 1–13, 2015.
- [14] "Multispectral palmprint recognition using multiclass projection extreme learning machine and digital shearlet transform," *Neural Computing and Applications*, vol. 27, no. 1, pp. 143–153, 2016.
- [15] P. Liu and L. Shi, "The generalized hybrid weighted average operator based on interval neutrosophic hesitant set and its application to multiple attribute decision making," *Neural Computing and Applications*, vol. 26, no. 2, pp. 457–471, 2015.
- [16] S. M. Mahbubur Rahman, S. P. Lata, and T. Howlader, "Bayesian face recognition using 2D Gaussian-Hermite moments," *EURASIP Journal on Image and Video Processing*, vol. 2015, no. 1, p. 35, 2015.
- [17] P. Y. Lin and Y. H. Chen, "High payload secret hiding technology for QR codes," *EURASIP Journal on Image and Video Processing*, vol. 2017, no. 1, p. 14, 2017.
- [18] T. M. Khan, M. A. Khan, Y. Kong et al., "Stopping criterion for linear anisotropic image diffusion: a fingerprint image enhancement case," *Eurasip Journal on Image & Video Processing*, vol. 2016, no. 1, p. 6, 2016.
- [19] S. Al-Maadeed, F. Ferjani, S. Elloumi et al., "A novel approach for handedness detection from off-line handwriting using fuzzy conceptual reduction," *EURASIP Journal on Image and Video Processing*, vol. 2016, no. 1, p. 1, 2016.
- [20] "A comprehensive survey of handwritten document benchmarks: structure, usage and evaluation," *EURASIP Journal on Image and Video Processing*, vol. 2015, no. 1, p. 46, 2015.