International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 3062-3075

https://doi.org/10.70517/ijhsa464253

Research on Image Segmentation Algorithm Combining Multi-angle Image Analysis and Feature Selection Techniques

Yanpin Mei[1,*]
[1] Yangzhou Polytechnic College, Yangzhou, Jiangsu, 225009, China
Corresponding authors: (e-mail: wind845@163.com).

Abstract In this paper, perimeter, roundness, boundary roughness, rectangularity, and center of mass displacement are selected as candidate features, the electric coal image dataset is collected, the set of feature vectors is selected based on the importance, relevance, and distinguishability of the features, and the KPCA algorithm is used to reduce the unnecessary image feature data and reduce the vector dimensions to obtain the optimal subset of the features, and the color features, such as color moments of the image, and the grayscale covariance are extracted matrix, Tamura texture features and filter these features. Then an image segmentation network CLUNS based on convolutionalized recurrent neural network is proposed. The classification and recognition results show that compared with other segmentation algorithms, the segmentation accuracy of the segmentation network proposed in this paper is 97.82%, and compared with the original CLUNS network algorithm the proposed algorithm improves the segmentation accuracy by 4.30 percentage points, and significantly reduces the loss rate in the validation set, respectively, by 10.16 The proposed algorithm has a significant advantage in running time, good generalization ability and stability, and provides a reference for subsequent quantitative image detection.

Index Terms image processing, neural network, feature extraction, KPCA algorithm

I. Introduction

As an important auxiliary means of clinical diagnosis, medical images provide doctors with important pathology cut-away diagrams without harming patients, providing an important basis for clinical diagnosis [1]. However, manual diagnosis is not flawless. For example, clinicians may make less accurate diagnostic and therapeutic judgments because of their shallow experience or fatigue accumulated from long hours of work [2]. In addition, due to the influence of human subjective factors, the diagnostic results of different experts for the same patient may sometimes be different, which greatly increases the time for patients to confirm the diagnosis [3]. However, with the continuous development of technology, classical image segmentation algorithm has become a powerful tool for medical image analysis [4]. It is able to delineate the structure and organization in medical images, providing a deeper understanding of the image content. These algorithms are able to effectively extract key features in the image, providing doctors with more comprehensive information and helping to improve the accuracy of diagnosis [5]-[7].

Image segmentation is the process of dividing an image into different regions based on the attributes of the image (color, brightness, texture, etc.), which is an intermediate step in image processing and an important part of pattern recognition and image understanding [8]. Traditional image segmentation algorithms, commonly based mainly on clustering algorithm and fuzzy entropy algorithm. Literature [9] proposed a new multilevel image segmentation clustering algorithm which performs well in standard benchmark tests and has the advantages of simplicity, generality and computational efficiency. Literature [10] constructed a new hybrid image segmentation method using a combination of clustering algorithm and black hole algorithm, where the black hole algorithm is able to define the fitness of each segment and find the most matching segments for merging, overcoming the limitations of using clustering algorithm alone. Literature [11] combines k-means clustering with an improved subtractive clustering algorithm in order to determine more accurate initial cluster centers of mass, this new hybrid image segmentation method overcomes the limitations of the standard k-means algorithm by determining more accurate initial cluster centers of mass. Literature [12] proposed an improved fuzzy entropy clustering (IFEC) algorithm for Magnetic Resonance Imaging (MRI) brain image segmentation, which combines local spatial and grayscale information to make it insensitive to noise and preserves image details without the need for parameter selection during the clustering process. Literature [13] introduces a new algorithm for cancer medical image segmentation based on level set fuzzy entropy, which is practically tested on three types of medical images and outperforms other segmentation algorithms. Literature [14] constructed a new image segmentation algorithm based on fuzzy theory, improved discrete gray wolf optimizer and local information aggregation technique, which outperforms other

optimization-based and fuzzy entropy-based models. Literature [15] incorporates spatial information one into the image thresholding process through a novel gray-level local fuzzy entropy (GLLFE) histogram, which is able to improve the quality of the image segmentation results compared with the traditional thresholding methods. In general, classical image segmentation algorithms have inherent defects such as rigid processing, contrast sensitivity, and inability to deal with complex structures, which limits their application in complex medical image analysis [16]. In addition, the above image segmentation algorithms usually require manual parameter tuning and lack a comprehensive understanding of image contextual information [17].

With the rise of deep learning, image segmentation algorithms based on models such as convolutional neural networks (CNNs), long short-term memory networks (LSTMs), recurrent neural networks (RNNs), and generative adversarial networks (GANs) are gradually becoming an effective means to overcome these limitations [18]-[20]. Literature [21] proposed an unsupervised image segmentation algorithm based on CNN algorithm which utilizes squeezing and excitation networks to capture the interdependencies between features and combines them with image enhancement techniques to produce better segmentation results. Literature [22] proposes an image segmentation algorithm based on a bi-convolutional LSTM network, including an encoder network that captures multimodal feature interactions, and a decoder network that refines multimodal features with spatial attention to produce an accurate segmentation mask. Literature [23] constructed an image segmentation algorithm for optical coherence tomography (OCT) based on RNN networks and found it to be more competitive with CNN methods. In practice, segmented images are often disturbed by complex environments, and the above segmentation methods fail to take into account the influence of spatial information of neighboring pixels on clustering, which leads to the influence of the sensitivity of noise, resulting in unsatisfactory image segmentation results, misclassification, and unclear edges.

The study was conducted to extract features from three aspects: color, texture, and shape, and different feature selection methods such as decision tree, linear discriminant analysis, and kernel principal component analysis were used to select features, to select the optimal subset, to reduce the number of unnecessary features, and to reduce the dimensionality of the data by using the KPCA algorithm for feature selection. The electric coal image is used as a research dataset to extract the color features such as color moments and other color features of the image and gray scale covariance matrix and Tamura texture features. And a three-dimensional segmentation network combining convolutional neural network and recurrent neural network is proposed, while the cross-entropy loss and loss function based on image segmentation index are flexibly used for the category imbalance problem of semantic segmentation, and the difference between this method and the existing segmentation methods in terms of segmentation accuracy and segmentation speed is comparatively analyzed.

II. Image segmentation method based on multi-angle image feature analysis and selection

II. A.Image feature extraction method

II. A. 1) Color characteristics

The color features reflect the overall characteristics of the image, and the first-order moments, second-order moments and third-order moments of the color are usually used to express the color features of the image.

Let $P(j,i)$ be the value of the i th color component of the j th pixel of the image, then the first-order moment is:

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N} P_{ji} \tag{1}$$

It represents the mean value of the color of the area to be measured.

The second order distance is:

$$\sigma_i = \left(\frac{1}{N}\sum_{j=1}^{N}(P_{ij}-\mu_i)^2\right)^{1/2} \tag{2}$$

It represents the color variance, i.e. inhomogeneity, of the area to be measured.

The third order distance is:

$$S_i = \left(\frac{1}{N}\sum_{j=1}^{N}(P_{ij}-\mu_i)^3\right)^{1/3} \tag{3}$$

The third-order moments reflect the asymmetry of the colors. If the image color is perfectly symmetrical, its value should be zero.

II. A. 2) Texture characteristics

(1) Gray scale co-occurrence matrix

Calculating the grayscale co-occurrence matrix of an image is the most classical method to characterize texture. The gray level co-occurrence matrix can better reflect the correlation in the gray level space by calculating the joint distribution of the frequency of occurrence of two gray level pixel pairs that are separated from each other in the image.

The gray level co-occurrence matrix is specified as follows: assume that the operator O denotes some relative positional relationship between two pixels and the image $f(x,y)$ has L gray levels. Let G be a matrix whose elements g_{ij} are the number of pixel pairs in the image f that satisfy the positional relationship specified by operator O, where gray levels x_i and x_j occur, where $i \le L$ and $j \le L$. The matrix G thus obtained, consisting of the elements g_{ij}, is the grayscale co-occurrence matrix. In the grayscale co-occurrence matrix, it is known that the total number of pixel pairs n satisfying the position operator O is equal to the sum of the elements of the matrix G. The co-occurrence matrix can be normalized using equation (4) as follows:

$$P_{ij} = \frac{g_{ij}}{n} \tag{4}$$

As a result, the gray scale co-occurrence matrix G is obtained by normalizing the normalized co-occurrence matrix G_n, which is calculated as shown in Equation (5):

$$G_n = \frac{G}{n} \tag{5}$$

(2) Extraction of texture features

Based on the above analysis of texture features, six texture features of the catalyst particle segmentation region are calculated in this section, which are: maximum probability, contrast, correlation, homogeneity, energy and entropy based on the gray level co-occurrence matrix (GLCM). The strongest response in the grayscale co-occurrence matrix, i.e., the maximum value of the element P_{ij} is the maximum probability, which is calculated as shown in (6):

$$\text{Maximum probability} = \max_{i,j}(P_{ij}) \tag{6}$$

Contrast ratio is a measure of the contrast between a pixel in an image and the grayscale of its neighboring pixels, which is zero for a constant image, and is calculated by the following formula (7):

$$\text{Contrast} = \sum_{i=1}^{K}\sum_{j=1}^{K}(i-j)^2 P_{ij} \tag{7}$$

Correlation is a measure of how a pixel in an image is related to its neighboring pixels. The correlation of a perfectly positively correlated image is 1, the correlation of a negatively correlated image is -1, and the correlation of a constant image is calculated as NaN. the calculation of correlation is shown in equations (8) and (9):

$$\text{Relevance} = \sum_{i=1}^{K}\sum_{j=1}^{K}\frac{(i-m_u)(j-m_c)P_{ij}}{\sigma_u \sigma_c} \tag{8}$$

Among them:

$$\begin{cases} m_u = \sum_{i=1}^{K} iP(i) \\ m_c = \sum_{j=1}^{K} jP(j) \\ \sigma_u^2 = \sum_{i=1}^{K} (i-m_u)^2 P(i) \\ \sigma_c^2 = \sum_{j=1}^{K} (j-m_c)^2 P(j) \\ P(i) = \sum_{j=1}^{K} P_{ij} \\ P(j) = \sum_{i=1}^{K} P_{ij} \end{cases} \tag{9}$$

Homogeneity is a measure of the closeness of the elements in the gray level co-occurrence matrix of an image to its diagonal elements, which is calculated by the following formula (10):

$$\text{Homogeneity} = \sum_{i=1}^{K}\sum_{j=1}^{K} \frac{P_{ij}}{1+|i-j|} \tag{10}$$

Calculation of energy and entropy based on the grayscale co-occurrence matrix is also used to measure the stochasticity of the data as shown in (11) and (12):

$$\text{Energy} = \sum_{i=1}^{K}\sum_{j=1}^{K} P_{ij}^2 \tag{11}$$

$$\text{Entropy} = -\sum_{i=1}^{K}\sum_{j=1}^{K} P_{ij} \log_2 P_{ij} \tag{12}$$

As a result, this section extracts six texture features from the segmented catalyst particle region based on the gray scale co-occurrence matrix.

II. A. 3) Shape characteristics

(1) Perimeter

Perimeter calculates the length of the boundary of a connected region. Boundary adjacent pixels exist in two cases, one is adjacent pixels side by side that is, adjacent pixels are located in the same column or the same row, in this case the distance between adjacent pixels is 1; the other case is the boundary of the two adjacent pixels are located on the diagonal, in this case the distance between adjacent pixels is $\sqrt{2}$. The boundary perimeter calculation formula, as shown in equation (13).

$$L = \sum_{i=0}^{N} \sqrt{(x_{i+1}-x_i)^2 + (y_{i+1}-y_i)^2} \tag{13}$$

where, (x_i, y_i), (x_{i+1}, y_{i+1}) are the coordinates of neighboring pixel points on the boundary of the connected region.

(2) Roundness

Roundness is used to describe the complexity of the boundaries of things. Compared with things with relatively regular shapes such as flashlights, fluorescent lights, and metal reflections, images have irregular shapes, and the interference of regular shiny objects is excluded by circularity. The maximum circularity value is 1, and the smaller the value, the more irregular the object is. The formula for calculating the circularity C is shown in equation (14).

$$C = \frac{4\pi S}{L^2} \tag{14}$$

where, L is the perimeter of the connected area; S is the area of the connected area.

(3) Boundary roughness

Boundary roughness can represent the complexity of the image boundary. Boundary roughness is described using the ratio of the perimeter of the convex envelope of the connected region to the perimeter of the image, the formula for boundary roughness, as shown in equation (15).

$$B = \frac{L_c}{L} \tag{15}$$

where, L_c is the perimeter of the convex packet of the connected region; L is the perimeter of the connected region.

(4) Rectangularity

The rectangularity degree indicates the degree of similarity between the connected region and the smallest external rectangle. The formula of rectangular degree R is defined as shown in equation (16).

$$R = \frac{S}{S_R} \tag{16}$$

where, S_R is the smallest rectangular area that contains the region where the object is located; S is the area of the suspected region.

(5) Center of mass displacement

The center of mass is defined as the average of the pixel coordinates in the region, the center of mass itself is not used as a criterion, it is mainly used to solve the displacement, the coordinates of each image pixel in a region are (x_j, y_j), where $j = 0,1,2,\cdots,n-1$, then the coordinates of the center of mass of the i th frame (x_i, y_i) can be obtained by Eq. (17).

$$(x_i, y_i) = \left(\frac{1}{n}\sum_{j=0}^{n-1} x_j, \frac{1}{n}\sum_{j=0}^{n-1} y_j \right) \tag{17}$$

Center of mass displacement d : if the center of gravity of the current frame is (x_i, y_i) and the center of gravity of the previous frame is (x_{i-1}, y_{i-1}), then the current frame displacement is as shown in equation (18).

$$d = \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} \tag{18}$$

II. B.Feature selection

II. B. 1) General process of feature selection

A target object can have many features at the same time, some features are beneficial to target recognition, some are not, and the relationship between various features is relatively unintuitive, some features will cancel each other when selected simultaneously, and some features are more correlated with each other, which may result in redundancy when selected at the same time. In the case of a certain number of samples, the more data makes the training model consume more time, which will reduce the classification accuracy. Feature selection is the process of selecting a subset of features from all feature vectors from the most to the least, removing similarities and reserving differences, to achieve a better classification result when classifying and recognizing images. The process of feature selection is to filter the features with high degree of importance and reduce the number of unnecessary features, so as to effectively reduce the number of dimensions, reduce the amount of computation, so as to make the model more capable of generalization and reduce the shortcomings caused by overfitting.

II. B. 2) Feature selection methods

(1) Manual selection

One of the most commonly used and effective methods for selecting features is manual selection, the advantage is that the experimenter can select the features subjectively by himself, and select the necessary features artificially by virtue of his own experience, the disadvantage is that the subset recognition effect of manual selection is related to the experimenter's experience, and it is easy to lead to errors due to the cumbersome data.

(2) Principal Component Analysis

Principal Component Analysis (PCA) is a multivariate statistical algorithm used to discard the features that carry a small amount of data information, retaining the important features. PCA is often used for linear dimensionality

reduction of high-dimensional data, and is now used in many fields, when the problem involves multiple variables or data of high dimensionality, and the correlation between the variables is obvious, borrowing this method is easier to simplify the problem, and to capture the main aspects of the matter [24].

(3) Linear Discriminant Analysis

Linear discriminant analysis (LDA) is based on the algorithmic basic principles of maximizing the spacing between sample classes and minimizing the spacing within sample classes to achieve feature selection. Like PCA, LDA is also a linear dimensionality reduction method, the difference is that the LDA algorithm requires the input of labeled feature data, which is a supervised dimensionality reduction algorithm.

(4) Decision Tree Method

Decision Tree (DTM) is a feature selection method for forward selection. The feature set selection can be realized by pruning operation on the fully grown decision tree.The DTM method uses information gain as an evaluation function index. Information gain indicates the degree of information complexity reduction under the same conditions. The greater the information entropy, the less the feature affects the recognition result, the more important the feature is.

Suppose feature A has n values and the information entropy of the data set D about A is denoted:

$$H_A(D) = -\sum_i^n \frac{|D_i|}{|D|} \log \frac{|D_i|}{|D|} \tag{19}$$

The information gain $IG_R(D,A)$ of the categorical variables can be obtained from the information entropy after selecting the feature A, which is calculated as:

$$IG_R(D,A) = \frac{IG(D,A)}{H_A(D)} \tag{20}$$

II. B. 3) KPCA nonlinear feature selection

Kernel Principal Component Analysis (KPCA) is a nonlinear upgrade of PCA algorithm, which can eliminate redundancy and spatial correlation between sample data and extract nonlinear principal element features containing the main data information. KPCA algorithm, for the input matrix X, adopts the principle of nonlinear mapping to transform all the samples in X to a high-dimensional or even infinite-dimensional feature space, and then performs PCA downscaling of it to reduce the dimensionality by switching dimension space to reduce the computational difficulty and avoid complex operations [25].

The principle of KPCA dimensionality reduction is to assume that there are N samples $x_1, x_2, \cdots, x_n \in R^m$, each of which is m-dimensional, and the nonlinear mapping function is Φ, and the covariance matrix of the mapping to the high-dimensional space is denoted as:

$$C = \frac{1}{N} \sum_{i,j=1}^{N} \Phi(x_i)\Phi(x_j)^T \tag{21}$$

Let the matrix eigenvalues be λ and the eigenvectors be V, then the corresponding eigenequation for the covariance matrix C is:

$$CV = \lambda V \tag{22}$$

Make an inner product of the mapping function $\Phi(x_k)$ with the above equation, i.e:

$$\langle \Phi(x_k), CV \rangle = \lambda \langle \Phi(x_k), V \rangle, k = 1, 2, \cdots, N \tag{23}$$

Then the eigenvector V of the covariance matrix C can be calculated as:

$$V = \sum_{i=1}^{N} \alpha_i \Phi(x)\Phi(x_i) \tag{24}$$

where α_i is the coefficient in the expression of the eigenvector V, which is obtained by associating the above four equations:

$$\lambda\sum_{i=1}^{N}\alpha_i\langle\Phi(x_i),\Phi(x_k)\rangle = \frac{1}{N}\sum_{i=1}^{N}\alpha_i\langle\Phi(x_i),\sum_{j=1}^{N}\Phi(x_j)\rangle\langle\Phi(x_i),\Phi(x_k)\rangle \tag{25}$$

Define the kernel matrix K_{ij} for $N*N$:

$$K_{ij} = K(x_i,x_j) = \langle\Phi(x_i),\Phi(x_j)\rangle \; i,j = 1,2,\cdots,N \tag{26}$$

where $K(x_i,x_j)$ is the inner product of the kernel function. Equation (25) can be simplified from equation (26):

$$K\alpha = N\lambda\alpha \text{ Which } \alpha = (\alpha_1,\alpha_2,\cdots,\alpha_N)^T \tag{27}$$

α is the eigenvector corresponding to the kernel function matrix K .

II. C.Image segmentation network based on CLUNS network structure

II. C. 1) U-Net-S

The underlying network structure for 2D image segmentation in CLUNS is a modified U-Net. The text names this structure U-Net-S, where S stands for Symmetrical. The downsampling and upsampling parts of the original U-Net are not strictly symmetrical, so the feature maps in the downsampling process need to be cropped in the jump-connected part. This cropping will bring some information loss, and the output size of the asymmetric network structure is inconsistent with the output size. The modified U-Net-S, on the other hand, has a strictly symmetric structure. The implementation method is to add padding in the convolution operation, which can be realized by setting the padding parameter of the convolution layer to “same” in the deep learning framework Keras2.

II. C. 2) Spatial context learning

CLUNS uses convolutional long and short-term memory for spatial context learning. CLUNS uses long-short-term memory, abbreviated as LSTM, to learn spatial context by replacing matrix multiplication with convolution, which preserves the spatial information of the LSTM input data.

The α is the feature vector corresponding to the kernel function matrix K .

LSTM is a kind of recurrent neural network (RNN).LSTM is based on RNN and adds the structure of gate. First of all, RNN has hidden state for memorizing the information of the previous time step, a simple RNN unit can be represented as:

$$\begin{aligned} h_t &= \phi(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \\ o_t &= \phi(W_{ho}h_t) \end{aligned} \tag{28}$$

That is, the hidden state h_t of the t time step is calculated based on the input x_t and the hidden state h_{t-1} of the previous time step and activated by the activation function, and the output O_t is calculated based on the hidden state h_t of the current time step. The basic LSTM unit is obtained by adding input gates i , forgetting gates f and output gates o to the RNN. Without loss of generality, the gates in a typical LSTM cell (note: there are many variants of LSTM) can be represented as:

$$\begin{aligned} i_t &= \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \\ f_t &= \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \\ o_t &= \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \end{aligned} \tag{29}$$

The activation function uses sigmoid. The hidden state of the LSTM needs to be computed with the help of the memory cell C . The candidate memory cell $\tilde{C}$ needs to be computed before computing C , which is computed as follows:

$$\tilde{C}_t = \tanh\left(W_{xc}x_t + W_{hc}h_{t-1} + b_c\right) \tag{30}$$

where the activation function $\tanh$. It is important to note that input gates, forgetting gates and output gates are called gates because the values in them are controlled in the $[0,1]$ interval by activation during computation, which is similar to a switch in an electrical circuit. Memory cells can then decide to memorize or forget certain information

by opening and closing the switch, computationally obtained using element-wise multiplication of matrices (denoted as ⊙):

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{31}$$

Hidden states are computed with the help of output gates and memory cells:

$$h_t = o_t \odot \tanh(C_t) \tag{32}$$

The hidden state is also used as the output of this unit.

II. D.Loss Functions in Image Segmentation

The choice of loss function is critical for semantic segmentation tasks, especially for medical images. This is because the target to be segmented in medical images often occupies only a small region in the whole image, which manifests itself as a category imbalance problem in the semantic segmentation of images. In this case, if equal attention is given to the positive and negative categories, then the negative category, which occupies a larger proportion, will dominate in the loss function, thus causing the problem of not classifying the positive category accurately enough. For the segmentation problem of binary classification, the commonly used loss function is the binary cross-entropy loss, which can be expressed as follows (y denotes the true category label, and $\hat{y}$ denotes the predicted value):

$$L(y,\hat{y}) = -y\log\hat{y} - (1-y)\log(1-\hat{y}) \tag{33}$$

The most direct way to solve the category imbalance is to add a proportionality factor to the cross-entropy loss function, which is used to control the attention to the loss of different categories. Introducing the proportionality factor α denotes the attention to the positive category, then the corresponding attention to the negative category is $1-\alpha$, and so the cross-entropy loss function for category balancing is obtained:

$$L(y,\hat{y}) = -\alpha y\log\hat{y} - (1-\alpha)(1-y)\log(1-\hat{y}) \tag{34}$$

Although the cross-entropy loss for category balancing can reduce the undesirable effects of category imbalance to some extent, it introduces additional hyperparameters. A better solution is to use a loss function related to the evaluation metrics of image segmentation. The evaluation metrics commonly used in image segmentation are Dice similarity coefficient and IoU (called Jaccard similarity coefficient), DSC can be expressed as:

$$DSC(Y,\hat{Y}) = \frac{2(|Y|\cap|\hat{Y}|)}{|Y|+|\hat{Y}|} \tag{35}$$

From the expression of DSC, we can see that DSC measures the proportion of overlap between the predicted value and the target value. The higher the proportion of overlap between the two, the closer the DSC is to 1. Based on the DSC the Dice loss can be obtained, i.e.:

$$L_{dice}(Y,\hat{Y}) = 1 - DSC(Y,\hat{Y}) \tag{36}$$

The expression for IoU is:

$$J(Y,\hat{Y}) = \frac{|Y\cap\hat{Y}|}{|Y\cup\hat{Y}|} = \frac{|Y\cap\hat{Y}|}{|Y|+|\hat{Y}|-|Y\cap\hat{Y}|} \tag{37}$$

The loss function corresponding to the IoU is called the Jaccard distance and is denoted as:

$$D_{Jaccard}(Y,\hat{Y}) = 1 - J(Y,\hat{Y}) \tag{38}$$

Both Dice loss and Jaccard distance are effective in reducing the adverse effects due to category imbalance during semantic segmentation network training, but also suffer from training instability. Alternatively, these loss functions can be combined and stacked as needed.

In this paper, different loss functions are used for the training of 2D U-Net-S and 3D CLUNS. U-Net-S uses cross-entropy superimposed on Dice's loss as the loss function, while CLUNS uses Jaccard's distance as the loss function. This is because U-Net-S is responsible for segmenting the general outline of the thyroid gland from the original CT image, a process that requires a balanced consideration of segmentation accuracy and evaluation metrics based on overlap rate.

III. Experimental results and analysis

III. A. Experimental preparation

III. A. 1) Experimental environment

Hardware environment: intel(r)core(tm)i5760@2.8 ghz.

Software environment: microsoft windows 7 service pack 1, matlab r2014a, oracle database 10g.

III. A. 2) Experimental image library

The image database comes from the background database of “Real-time Fuel Management and Intelligent Analysis System”, and the coal has been pre-processed with a total of 3500 images, with the image size of 250×250. The acquisition of coal images is obtained by photographing the wagons from a long distance during the sampling of coal.

III. B. Image feature extraction and selection analysis

The purpose of feature selection is to select the features that best represent the electric coal image from a large number of features. The features of power coal image can be selected from color features and texture features. The color features of the coal image are selected from color moments and color histograms, and the texture features of the coal image are selected from grayscale covariance matrix and Tamura texture features.

(1) Color feature selection: the HSV color space is selected in color feature extraction, and the color moments include the first, second, and third order moments of each HSV component, and the gray scale of the electric coal image is divided into 9 intervals, with the gray scale ranging from 5 to 50, and the spacing distance is 5. Fig. 1 shows the trend of the first, second, and third order moments of the color hue H with the change of the gray scale, in which the horizontal axis represents the gray scale in the coal quality indexes, and the vertical axis represents the eigenvalues. From the observation in the figure, it can be seen that the values of the second-order and third-order moments of hue H and the trend of change with ash are basically the same, and one of them can be selected when the color moment feature is chosen.

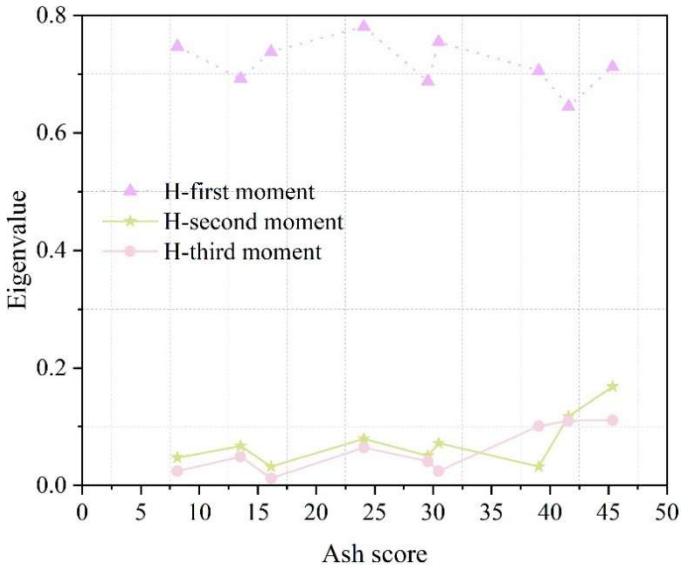

Figure 1: The trend of the first, second and third order moments of hue h with ash content

Figure 2 shows the trend of the first, second and third order moments of saturation S with the gray level. From the observation in the figure, it can be seen that the trend of the values of the second-order and third-order moments of saturation S with the gray level is basically the same, and one of them can be selected in the feature selection.

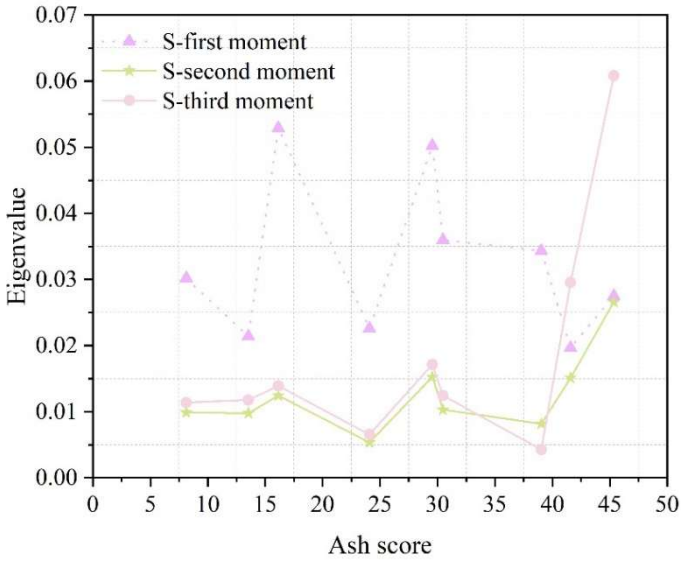

Figure 2: The first, second and third moments of saturation s with respect to ash content

Fig. 3 shows the trend of the first, second and third order moments of luminance V with gray scale. The second-order moments and third-order moments are taken as one of the two for feature selection.

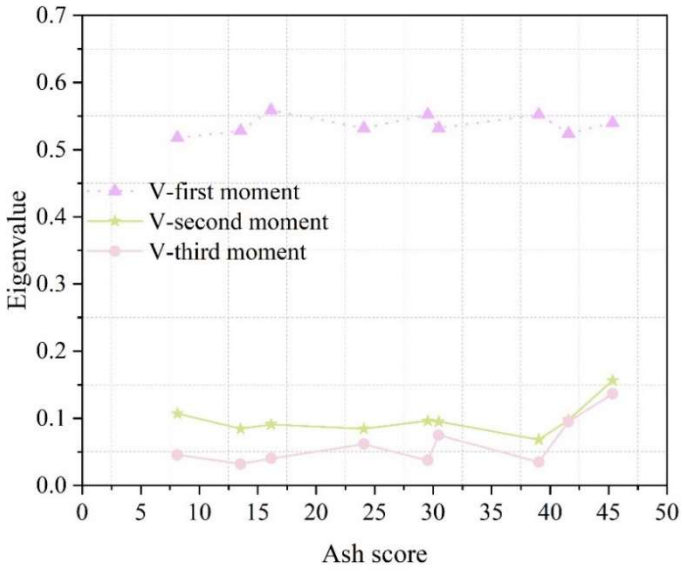

Figure 3: The trend of the first, second and third moments of brightness v with ash content

From the above analysis of color moments, it can be seen that hue first order moment, hue second order moment, saturation first order moment, saturation second order moment, luminance first order moment and luminance second order moment are selected from the color moments as the color features for image recognition in order to reduce the amount of computation and to speed up the speed of image processing.

The color histogram has the invariance of translation and rotation, but due to improper quantization, the pixels of some color units are very sparse, and the cumulative histogram is used in order to eliminate this zero-value error. In the selection of color features, the first and second order moments of each component of the HSV of the electric coal image and the cumulative histogram after non-equal spacing quantization are selected.

(2) Texture feature selection: the features of the gray scale covariance matrix mainly include energy, contrast, entropy, and correlation, and the trend of the mean and standard deviation of contrast with the gray scale is shown in Fig. 4. From the observation in the figure, it can be seen that the electric coal images with different gray scores have different mean values of contrast, but the standard deviation is basically the same, and the standard deviation

of contrast can be disregarded in the feature selection. In the selection of the gray scale covariance matrix texture features, the mean and variance of energy, entropy, correlation and the mean value of contrast can be selected.

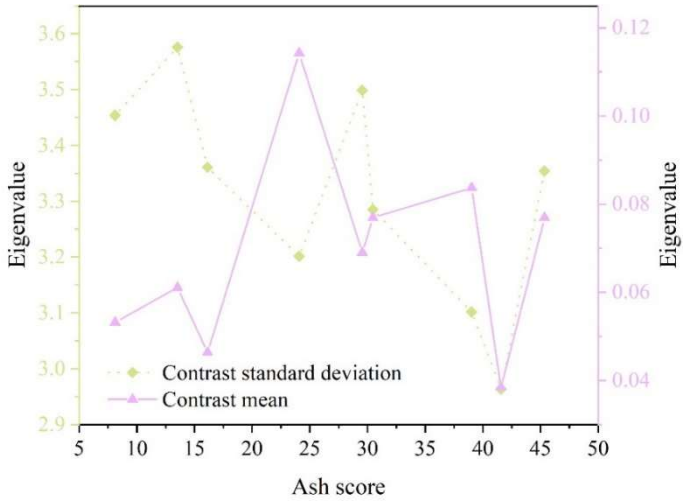

Figure 4: Trend of mean contrast and standard deviation with ash content

The texture features obtained from Tamura texture features give more image subjectively than the texture obtained from the gray scale covariance matrix and better reflect the texture features, Tamura texture features mainly include roughness, contrast and directionality, etc., Fig. 5 shows the trend of the change of Tamura texture features with the gray scale.

From the observation in the figure, it can be seen that the value of roughness and contrast, as well as the trend of change with gray level are different for the electric coal images with different gray levels, but the directionality is almost equal and close to 0. For the electric coal images, the directionality can be eliminated from the consideration of the Tamura texture features.

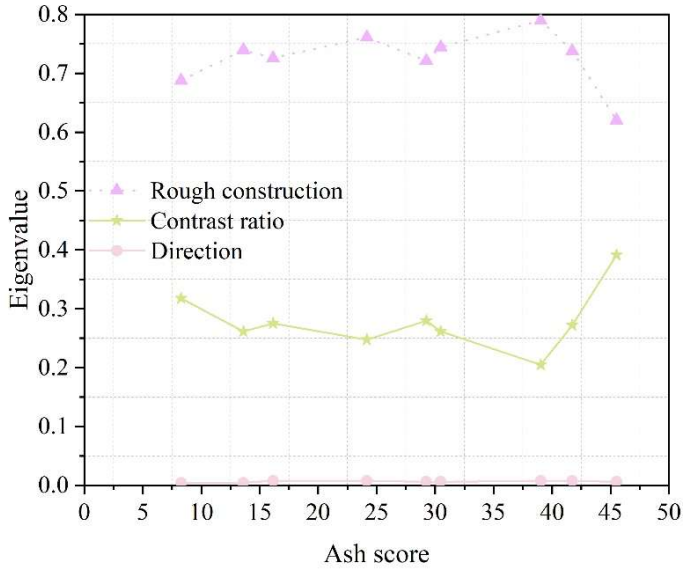

Figure 5: Tamura trend of texture characteristics with ash content

With the above analysis, for the electric coal image, the image features selected are the first and second order moments and cumulative histograms of H, S, and V. The texture features selected are the mean and variance of energy, entropy, and correlation in the gray level covariance matrix, the mean of contrast, and the roughness and contrast in the Tamura texture feature.

III. C. Image segmentation results and analysis

III. C. 1) Segmentation results

The proposed 3D segmentation network CLUNS based on convolutionalized recurrent neural network can segment the image clearly and get the target image segmentation accuracy as shown in Table 1. From the table, it can be seen that the segmentation algorithm in this paper has the highest segmentation accuracy on the training and validation sets, the lowest loss rate on the training and validation sets, and the shortest time consumption. In particular, the accuracy is improved by 4.38, 4.07 and 4.30 percentage points and the loss rate is reduced by 7.27, 10.33 and 10.16 percentage points on the validation set compared to PLVQNN, FPN and U-Net-S networks, respectively.

Table 1: Segmentation results

Algorithm	Training set Accuracy/%	Verification set Val_accuracy/%	Training set loss/%	Verification set val_loss/%	Processing time Time/step/s
PLVQNN	97.49	93.44	18.12	19.52	4
FPN	93.86	93.75	29.63	22.58	17
U-Net-S	93.37	93.52	19.27	22.41	2
Improved CLUNS	98.53	97.82	6.22	12.25	0.452

Fig. 6-Fig. 7 show the changes in the accuracy and loss rate curves before and after the improvement of the CLUNS network. From the figures, it can be seen that the improvement method proposed in this paper improves the accuracy ACCURACY significantly during the iteration process, and the loss rate LOSS decreases faster. Compared with the original CLUNS network model, the average processing time per STEP is the shortest, 0.439 s. The improved CLUNS network can select the final features from only one segmentation branch, and this selection brings speed gain. Therefore, the segmentation algorithm of CLUNS, a 3D segmentation network based on convolutionalized recurrent neural network proposed in this paper, solves the problem of insufficient segmentation results due to insufficient feature extraction, improves the segmentation accuracy, and also brings about speed gain, which is suitable for solving the problem of segmentation of complex background images.

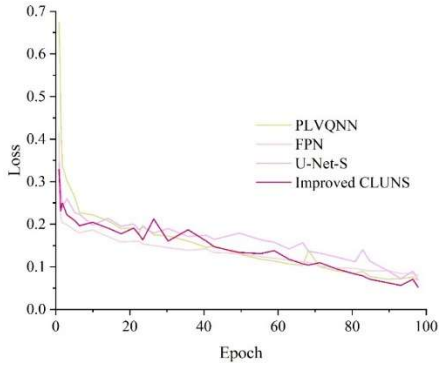

Figure 6: Improved the loss curve before and after

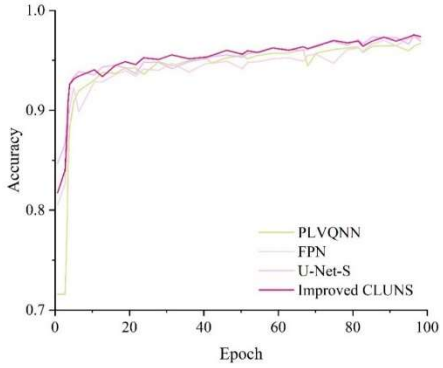

Figure 7: Improved accuracy curve before and after improvement

III. C. 2) Comparative analysis

The segmentation accuracies obtained in this paper are compared with the segmentation results of electric coal images studied by previous researchers, and the results are shown in Table 2. In terms of electric coal crown image segmentation, the segmentation results obtained by different segmentation methods are not the same, in which the size of the input image also has an impact on the segmentation algorithm. Compared with the existing conventional mainstream electric coal canopy image segmentation algorithms, the segmentation accuracy of the algorithm proposed in this paper is 95.73%, which is an improvement over the segmentation accuracy of other methods.

Table 2: Compares segmentation result with previous studies

Algorithm	Image size	Segmentation accuracy/%
Improved CLUNS	512×512	95.73
EarSegNet	400×400	75.31
Beeses learning algorithm	5280×3956	71.45
Improve Otsu	320×237	83.52

IV. Conclusion

In this paper, CLUNS network is proposed by combining convolutional neural network and recurrent neural network used for sequence data modeling, and KPCA algorithm is used to process the feature vectors, extract the color features such as color moments of the image and the grayscale covariance matrices, Tamura texture features, and screen these features, and conduct experimental analysis of the extracted image features. The experimental results show that on the same dataset, the segmentation accuracy of the method proposed in this paper is higher than that of PLVQNN (93.44%), FPN (93.75%) and U-Net-S (93.52%), and the segmentation accuracy has been significantly improved with the shortest time consumption.

Funding

This research was supported by the 2023 University-level key research project: 2023 "Qinglan Project" of Jiangsu Universities.

References

[1] Huang, X., & Tsechpenakis, G. (2009). Medical image segmentation. In Information discovery on electronic health records (pp. 269-308). Chapman and Hall/CRC.
[2] Lai, M. (2015). Deep learning for medical image segmentation. arXiv preprint arXiv:1505.02000.
[3] Chen, L., Bentley, P., Mori, K., Misawa, K., Fujiwara, M., & Rueckert, D. (2018). DRINet for medical image segmentation. IEEE transactions on medical imaging, 37(11), 2453-2462.
[4] Sharma, N., & Aggarwal, L. M. (2010). Automated medical image segmentation techniques. Journal of medical physics, 35(1), 3-14.
[5] Norouzi, A., Rahim, M. S. M., Altameem, A., Saba, T., Rad, A. E., Rehman, A., & Uddin, M. (2014). Medical image segmentation methods, algorithms, and applications. IETE Technical Review, 31(3), 199-213.
[6] Ramesh, K. K. D., Kumar, G. K., Swapna, K., Datta, D., & Rajest, S. S. (2021). A review of medical image segmentation algorithms. EAI Endorsed Transactions on Pervasive Health & Technology, 7(27).
[7] Guo, Z., Li, X., Huang, H., Guo, N., & Li, Q. (2019). Deep learning-based image segmentation on multimodal medical imaging. IEEE Transactions on Radiation and Plasma Medical Sciences, 3(2), 162-169.
[8] Qi, J., Yang, H., & Kong, Z. (2022, October). A review of traditional image segmentation methods. In 5th International Conference on Computer Information Science and Application Technology (CISAT 2022) (Vol. 12451, pp. 400-406). SPIE.
[9] Olugbara, O. O., Adetiba, E., & Oyewole, S. A. (2015). Pixel intensity clustering algorithm for multilevel image segmentation. Mathematical Problems in Engineering, 2015(1), 649802.
[10] Dhanachandra, N., Chanu, Y. J., & Singh, K. M. (2023). A new hybrid image segmentation approach using clustering and black hole algorithm. Computational Intelligence, 39(2), 194-213.
[11] Tongbram, S., Shimray, B. A., & Singh, L. S. (2021). Segmentation of image based on k-means and modified subtractive clustering. Indonesian Journal of Electrical Engineering and Computer Science, 22(3), 1396-1403.
[12] Verma, H., Agrawal, R. K., & Kumar, N. (2014). Improved fuzzy entropy clustering algorithm for MRI brain image segmentation. International journal of imaging systems and technology, 24(4), 277-283.
[13] Maolood, I. Y., Al-Salhi, Y. E. A., & Lu, S. (2018). Thresholding for medical image segmentation for cancer using fuzzy entropy with level set algorithm. Open Medicine, 13(1), 374-383.
[14] Li, L., Sun, L., Kang, W., Guo, J., Han, C., & Li, S. (2016). Fuzzy multilevel image thresholding based on modified discrete grey wolf optimizer and local information aggregation. Ieee Access, 4, 6438-6450.
[15] Zheng, X., Tang, Y., & Hu, W. (2018). Image thresholding based on gray level-fuzzy local entropy histogram. IEEJ Transactions on Electrical and Electronic Engineering, 13(4), 627-631.
[16] Lei, T., Liu, P., Jia, X., Zhang, X., Meng, H., & Nandi, A. K. (2019). Automatic fuzzy clustering framework for image segmentation. IEEE Transactions on Fuzzy Systems, 28(9), 2078-2092.
[17] Prabu, C., Bavithiraja, S. V. M. G., & Narayanamoorthy, S. (2016). A novel brain image segmentation using intuitionistic fuzzy C means algorithm. International Journal of Imaging Systems and Technology, 26(1), 24-28.

[18] Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., & Terzopoulos, D. (2021). Image segmentation using deep learning: A survey. IEEE transactions on pattern analysis and machine intelligence, 44(7), 3523-3542.
[19] Ghosh, S., Das, N., Das, I., & Maulik, U. (2019). Understanding deep learning techniques for image segmentation. ACM computing surveys (CSUR), 52(4), 1-35.
[20] Liu, X., Song, L., Liu, S., & Zhang, Y. (2021). A review of deep-learning-based medical image segmentation methods. Sustainability, 13(3), 1224.
[21] Ilyas, T., Khan, A., Umraiz, M., & Kim, H. (2020). Seek: A framework of superpixel learning with cnn features for unsupervised segmentation. Electronics, 9(3), 383.
[22] Ye, L., Liu, Z., & Wang, Y. (2020). Dual convolutional LSTM network for referring image segmentation. IEEE Transactions on Multimedia, 22(12), 3224-3235.
[23] Kugelman, J., Alonso-Caneiro, D., Read, S. A., Vincent, S. J., & Collins, M. J. (2018). Automatic segmentation of OCT retinal boundaries using recurrent neural networks and graph search. Biomedical optics express, 9(11), 5759-5777.
[24] Jean Claude Uwamahoro,John Bosco Habarulema,Dalia Buresova,Nigussie Mezgebe Giday,Valence Habyarimana,Kateryna Aksonova... & Ange Cynthia Umuhire. (2025). Simultaneous evaluation of solar activity proxies during geomagnetic storms using principal component analysis: Case study of the African low and mid-latitude regions. Journal of Atmospheric and Solar-Terrestrial Physics,270,106477-106477.
[25] Eider Barba,David Casamichana,Pedro Figueiredo,Fábio Yuzo Nakamura & Julen Castellano. (2024). The Use of Principal Component Analysis for Reduction in Sleep Quality and Quantity Data in Female Professional Soccer. Sensors,25(1),148-148.