

Received 13 March 2024, accepted 8 May 2024, date of publication 13 May 2024, date of current version 20 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3400026

RESEARCH ARTICLE

Convolutional Neural Network-Based Multiscale Feature Selection and Evaluation in Image Segmentation

CAO DI¹, CAO JIAN-NONG^{2,3}, DENG LIANG⁴, AND LOU LI-PING⁵

¹School of Earth Science and Resources, Chang'an University, Xi'an 710054, China

²Key Laboratory of Degraded and Unused Land Consolidation Engineering, Ministry of Land and Resources, Xi'an 710054, China

³School of Geology Engineering and Geomatics, Chang'an University, Xi'an 710054, China

⁴China Coal Technology and Engineering Group Xi'an Research Institute Company Ltd., Xi'an 710054, China

⁵Yellow River Engineering Consulting Company Ltd., Zhengzhou 450000, China

Corresponding author: Cao Di (RenesmeeSwan@126.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 41571346; in part by the Key Laboratory of Degraded and Unused Land Consolidation Engineering of the Ministry of Land and Resources under Grant SXDJ2017-10 and Grant 2016KCT-23; in part by the Natural Science Basic Research Plan in Shaanxi Province of China under Grant 2018JQ4027; and in part by the Fundamental Research Funds for the Central Universities under Grant 310826171014, Grant 310826160320, and Grant 300102268647.

ABSTRACT Multiscale image segmentation based on artificial neural networks is a hot topic in research on remote sensing image processing. However, the establishment and evaluation of pooling models and selection of feature operators lack clear standards. Based on the biological visual multiscale perception mechanism, this study combines classical wavelet theory with convolutional neural network theory to establish 10 sets of geometric operators and construct the corresponding multiscale image feature pyramids. Statistical analysis shows that the 10 sets of operators exhibit two types of information transmission characteristics, that is, balanced and growth. The obtained image features become more fragmented as operator complexity increases. After excluding the two operator groups with high complexities, the remaining eight groups were applied to the convolutional neural network image-segmentation algorithm. Eight pooling models were established to obtain the corresponding multiscale image features, perform convolution operations, and generate multiscale segmentation results for remote sensing images. The evaluation results reveal that the high complexity of the feature operators is unfavorable for feature transmission and preservation, and compared with operators having the information transmission characteristics of growth, those with balanced information transmission characteristics show better performance in convolutional neural network image segmentation. The segmentation accuracy was improved by 1.5%–2%. The conformity of the segmentation results was improved by 1%–1.5%. Finally, the degree of interclass chaos is reduced by 4.1%–10%.

INDEX TERMS Image segmentation, remote sensing, multiscale feature pyramid, pooling model, convolutional neural networks.

I. INTRODUCTION

Under the driving force of the multiscale perception mechanism [1], human vision can acquire visual information of different scales when facing a complete visual object, and capture the corresponding information features at each scale [2]. Simulating this multiscale perception mechanism of

human vision is a hot topic in image-understanding research. The main multiscale image processing method was initially based on dimensionality reduction analysis. That is, two-dimensional (2D) image information is compressed to one dimension before being deconstructed. The most widely used methods are Fourier transform, Gabor transform, short-time Fourier transform, wavelet transform [3], multiscale geometric analysis [4], ridgelet, curvelet [5], contourlet [6], bandelet, wedgelet, beamlet, and shearlet transform [7]. These methods


The associate editor coordinating the review of this manuscript and approving it for publication was Yi Zhang .

TABLE 1. Artificial neural network-based segmentation algorithm.

Segmentation Algorithm	Architecture
AlexNet	CNN
ZFNet	AlexNet
VGGNet-16	VGGNet
GoogLeNet-Inception v1	Inception module
ResNet	Residual module
Fully Convolutional Network [11]	VGG-16(FCN)
U-net [12]	FCN
Dilated Convolutions [13]	FCN
RefineNet [14]	U-Net
PSPNet [15]	ResNet
Large Kernel Matters [16]	ResNet
SegNet [17]	VGG-16 + Decoder
Bayesian SegNet [18]	SegNet
DeepLab [19]	VGG-16/ResNet-101
MINC-CNN [20]	GoogleNet(FCN)
CRFasRNN [21]	FCN-8s
Dilation [22]	VGG-16
ENet [23]	ENet bottleneck
Multiscale-CNN-Raj	VGG-16(FCN)
Multiscale-CNN-Eigen	Custom
Multiscale-CNN-Roy	Multiscale-CNN-Eigen
Multiscale-CNN-Bian	FCN
ParseNet [24]	VGG-16
ReSeg	VGG-16 + ReNet
LSTM-CF [25]	Fast R-CNN + DeepMask
2D-LSTM [26]	MDRNN
RCNN	MDRNN
DAG-RNN [27]	Elman network
SDS	R-CNN + BoxCNN
Deep Mask [28]	VGG-A
Sharp Mask [29]	DeepMask
MultiPathNet [30]	Fast R-CNN + DeepMask
Huang-3DCNN	3DCNN [31]
PointNet [32]	Own MLP-based
Clockwork Convnet [33]	FCN
BiSeNet	CNN
Feature Pyramid Network	CNN
High-Resolution Network	ResNet/DenseNet

are dimensionality reduction analysis algorithms for images. Although the dimensionality reduction means more thorough information deconstruction, it also leads to fragmentation and loss of high-dimensional information. Therefore, nondimensional reduction analysis has become one of the most important directions in image-understanding development. A. Cohen, I. Daubechies, R.Q. Jia, and R. Long initially extended the classical wavelet analysis theory to high dimensions. However, related academic research progress was slow, and practical applications were limited to the replacement of true 2D images with 2D wavelet tensors [8], [9], [10]. The introduction of 2D wavelet tensors partially compensates for the loss caused by the dimensionality reduction and attempts to reduce the redundancy, which is an attempt to analyze two-dimensional images in pseudo two-dimensional space. However, this improvement cannot overcome the limit imposed by the dimensionality reduction itself. With the emergence of deep learning and its integration with theories (e.g., artificial neural networks) in 2006, image-segmentation algorithms based on convolutional neural networks have

rapidly developed. Consequently, the nondimensional reduction analysis of images based on the multiscale perception mechanism of biological vision was realized. This means that we no longer analyze image information by one dimension or pseudo two-dimension method, but by constructing a multi-scale space to analyze two-dimensional image information in two-dimensional spaces. The relative segmentation method differentiates hundred of model structures (Table 1).

The establishment of multiscale feature maps involves image feature selection and information filtration in the context of convolutional neural network (CNN)-based image-segmentation algorithms. Techniques such as Lp and spectral pooling are commonly employed to accomplish this.

$$A_k^l(i, j) = \left[\sum_{x=1}^f \sum_{y=1}^f A_k^l(s_0i + x, s_0j + y)^p \right]^{\frac{1}{p}} \quad (1)$$

Lp pooling is a pooling model inspired by the hierarchical structure found in the visual cortex of biological vision. It computes the Lp norm of the features within each pooling region, thereby leading to a more robust and stable feature representation. The parameter “p” can assume any positive value, typically 1, 2, and ∞ . Lp pooling is reduced to average pooling for $p = 1$, where the mean value of the features within the pooling region is used as the output. For $p = 2$, Lp pooling corresponds to Euclidean distance pooling, using the L2 norm of the features within the pooling region as the output. Lp pooling becomes max pooling as p approaches ∞ , where the maximum feature value within the pooling region is used as the output.

Random selection of the p value is called random pooling. Lp random pooling randomly selects a value within the region according to a specific probability distribution to ensure that most of the non-maximal activation signals can proceed to the next level. In mixed pooling, p is a linear combination of the mean and max pooling.

$$A_k^l = \lambda L_1(A_k^l) + L_\infty(A_k^l), \lambda \in [0, 1] \quad (2)$$

Incorporating techniques (e.g., Lp and spectral pooling) for multiscale feature selection and information filtration enables CNN-based image-segmentation algorithms to generate more accurate and robust feature representations, thereby enhancing the overall performance of image-segmentation tasks.

However, clear standards for the establishment of pooling models and selection of feature operators during multiscale image feature selection and information filtering are currently lacking. The influence of different models on the CNN-based image-segmentation results remain unclear.

This study addresses this issue by leveraging the multiscale perception mechanism observed in biological vision in combination with classical wavelet and multiscale geometric analysis theories. We devised 10 sets of feature operators and constructed the corresponding multiscale feature pyramids to conduct a preliminary analysis and categorization

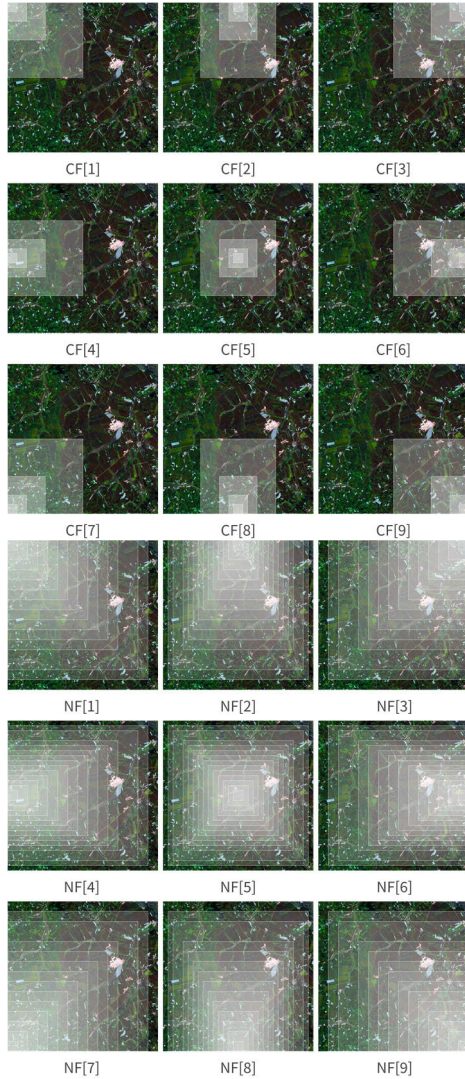


FIGURE 1. Traditional and new multiscale pyramid models.

of these operators. Subsequently, we incorporated different feature operators into the pooling model within the framework of the CNN-based image-segmentation algorithm. This enables us to extract multiscale image features through convolutional operations, which results in the generation of segmented outputs at various scales for remote sensing images. Finally, a comprehensive analysis and evaluation of the image-segmentation results are performed. The experimental results show that the experimental process can effectively screen out high-quality operators, remarkably and generally increasing the precision of remote sensing image segmentation by 1.5%–2%, improving intraclass conformity by 1%–1.5%, and reducing intraclass chaos by 4.1%–10.1%. Overall, the experimental results reveal that the proposed screen operator process provides explicit guidelines for establishing a multiscale feature pyramid model during multiscale image feature selection and information filtering, ultimately enhancing the accuracy of the image-segmentation results.

The remainder of this paper is organized as follows: Section II defines the multiscale and builds the multiscale feature pyramid with different geometric moments for remote sensing images. Section III presents experiments in multi-spectral remote sensing images to confirm the performance of different geometric moments, and Section IV concludes the paper and discusses future research directions.

II. METHODOLOGY

This section defines the multiscale feature pyramid and geometric moments of the images. We also conducted a multidimensional statistical analysis of the constructed multiscale feature pyramid model.

A. DEFINITION OF THE MULTISCALE FEATURE PYRAMID

Image multiscale analysis refers to the spatial series $\{V_j\}_{j \in \mathbb{Z}}$ in scale-space $L^2(\mathbb{R})$. The series meets certain conditions, including monotonicity, approximability, scalability, translation invariance, and the Riesz base [34]. Guan Yu Jing provided a generalized multiscale analysis theory that departs from the traditional theory [35]. In this new theory, different scales are built based on different scale functions in the scale-space.

In a generalized multiscale analysis theory and in scale-space $L^2(\mathbb{R})$, $\{V_j\}_{j \in \mathbb{Z}}$ is a spatial series presented as $V_j \subset L^2(\mathbb{R})$, $V_j \subset V_{j+1}$, $j \in \mathbb{Z}$. For each $\{V_j\}_{j \in \mathbb{Z}}$, there exists $\varphi_j(x)$, $\{\varphi_j(x - k/2^j)\}_{k \in \mathbb{Z}}$, which is the orthonormal basis of V_j .

The traditional model is advantageous in terms of computation, decay rates, and globality of information perception. However, the interval distance between two adjacent layers increases exponentially with the establishment of the multiscale model, leading to missing information layers. The new multiscale pyramid model was supported by good continuity in the multiscale analysis framework. It can solve the missing layer problem and is closer to the multiscale perception mechanism in human vision, which is the current direction for multiscale image analysis development.

B. DEFINITION OF THE GEOMETRIC MOMENTS OF IMAGES

Moment is used to describe the characteristics of a random variable in mathematical statistics. It was applied to demonstrate the geometrical characteristics of the image analysis. Stable geometrical characteristics are not affected by light, noise, or geometric deformation. Therefore, we introduce geometric moments for the subsequent multiscale analysis based on a generalized multiscale analysis. The gray distribution of the image target area D is set as follows:

$$f(x, y) \quad (x, y) \in D \quad (3)$$

The origin moment of order $p + q$ for D is

$$m_{pq} = \iint_D x^p y^q f(x, y) dx dy \quad (p, q = 0, 1, 2 \dots) \quad (4)$$

TABLE 2. Moment operators.

Model	Moment Operator
M_a	$\phi_{01} = \eta_{01} + \eta_{10}$
M_b	$\phi_{02} = \eta_{01} + \eta_{10}$ (noise reduction)
M_c	$\phi_{03} = \eta_{11}$
M_d	$\phi_{H1} = \eta_{20} + \eta_{02}$
M_e	$\phi_{H2} = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$
M_f	$\phi_{H3} = (\eta_{30} - 3\eta_{12})^2 + (\eta_{03} + 3\eta_{21})^2$
M_g	$\phi_{H4} = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$
M_h	$\phi_{H5} = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})^2 - 3(\eta_{21} + \eta_{03})^3 + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$
M_i	$\phi_{H6} = (\eta_{30} - \eta_{03})(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^3 + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$
M_j	$\phi_{H7} = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$

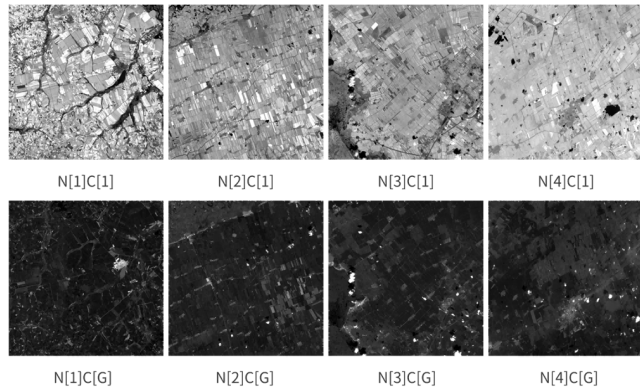


FIGURE 2. Multiscale pyramid models. (Four image slices, N [1]-N [4]. Two channels, C [1] and C[G]).

The central moment of order $p + q$ for D is

$$\mu_{pq} = \iint_D (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy \quad (p, q = 0, 1, 2 \dots) \quad (5)$$

The relationship between the origin and central moments is presented as follows:

$$\mu_{pq} = \sum_{k=0}^p \sum_{l=0}^q \binom{p}{k} \binom{q}{l} (-1)^{k+l} m_{p-k, q-l} m_{10}^k m_{01}^l m_{00}^{-(k+l)} \quad (6)$$

The normalized central moment is obtained as

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^r} \quad (r = \frac{p+q+2}{2}, p+q = 2, 3 \dots) \quad (7)$$

For the moment of an image, the lower its order, the more sensitive it becomes to the extreme value and the lower its computational complexity becomes: $\phi_{01}, \phi_{02}, \phi_{03}$.

Hu. M. K proposed the invariant moment concept and constructed seven invariant moments (i.e., $\phi_{H1}, \phi_{H2}, \phi_{H3}, \phi_{H4}, \phi_{H5}, \phi_{H6}, \phi_{H7}$) using the second- and third-order normalized central moments. The invariant moments include rotation, scaling, and translation invariance. However, in keeping the image characteristics, ϕ_{H1}, ϕ_{H2} , and ϕ_{H3} showed higher stabilities than $\phi_{H4}, \phi_{H5}, \phi_{H6}$, and ϕ_{H7} . ϕ_{H2} and ϕ_{H3} performed well in keeping boundaries in multiscale image segmentation.

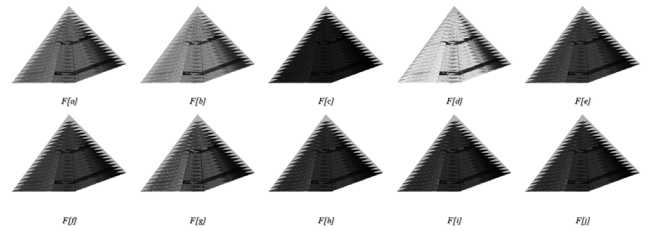


FIGURE 3. Multiscale pyramid models. (We built 10 sets of multiscale pyramid models for 2 channels of 4 images with 10 features, Feature[a]-Feature[j]. In this figure, a pyramid represents the set of all scale slices of a channel in an image slice under that feature. Taking the 10 pyramids shown in the figure as a group, we have established a total of $2 \times 4 = 8$ groups.)

TABLE 3. Moment operators.

Pattern	Model
Balanced mode:	$F[a], F[b], F[d]$
Growth mode:	$F[c], F[e], F[f], F[g], F[h], F[i], F[j]$

The 10 sets of feature operators in this work were $M_a, M_b, M_c, M_d, M_e, M_f, M_g, M_h, M_i$, and M_j (Table 2). Before M_b construction, the image was first denoised to reduce the extreme value interference among these operators to a low-order moment.

The image multiscale feature pyramid level used in the experiment was 50. The sampling mode was denoted by NF (1), which mean that the sampling window was established from the top-left corner of the image with consistent hierarchical intervals.

$$W_N = \{(x_n, y_n) = Feature_{N \times N}, n = 1, 2, \dots, N\} \quad (N : 1 \cdot 50)$$

(Feature: $F[a], F[b], F[c], F[d], F[e], F[f], F[g], F[h], F[i], F[j]$.)

W : Sampling window.
 N : Level.

The experimental data comprised multispectral images from the Gaofen-6 satellite of the China High-resolution Earth Observation System (CHEOS) program. As shown in Figure 2, four randomly selected image slices measuring 1500×1500 px were chosen. Each slice contained two channels.

C. MULTISCALE FEATURE PYRAMID ANALYSIS

In this study, the multi-scale feature pyramid is conceptualized as a three-dimensional feature space. Within the multi-scale feature pyramid, slices at all scales are perceived as a continuous whole. We conduct longitudinal profile analysis, cross section analysis of this three-dimensional feature space across different levels, to extract the characteristics of information transmission in various dimensions within this space. Additionally, by comprehensively comparing information transmission characteristics across different dimensions,

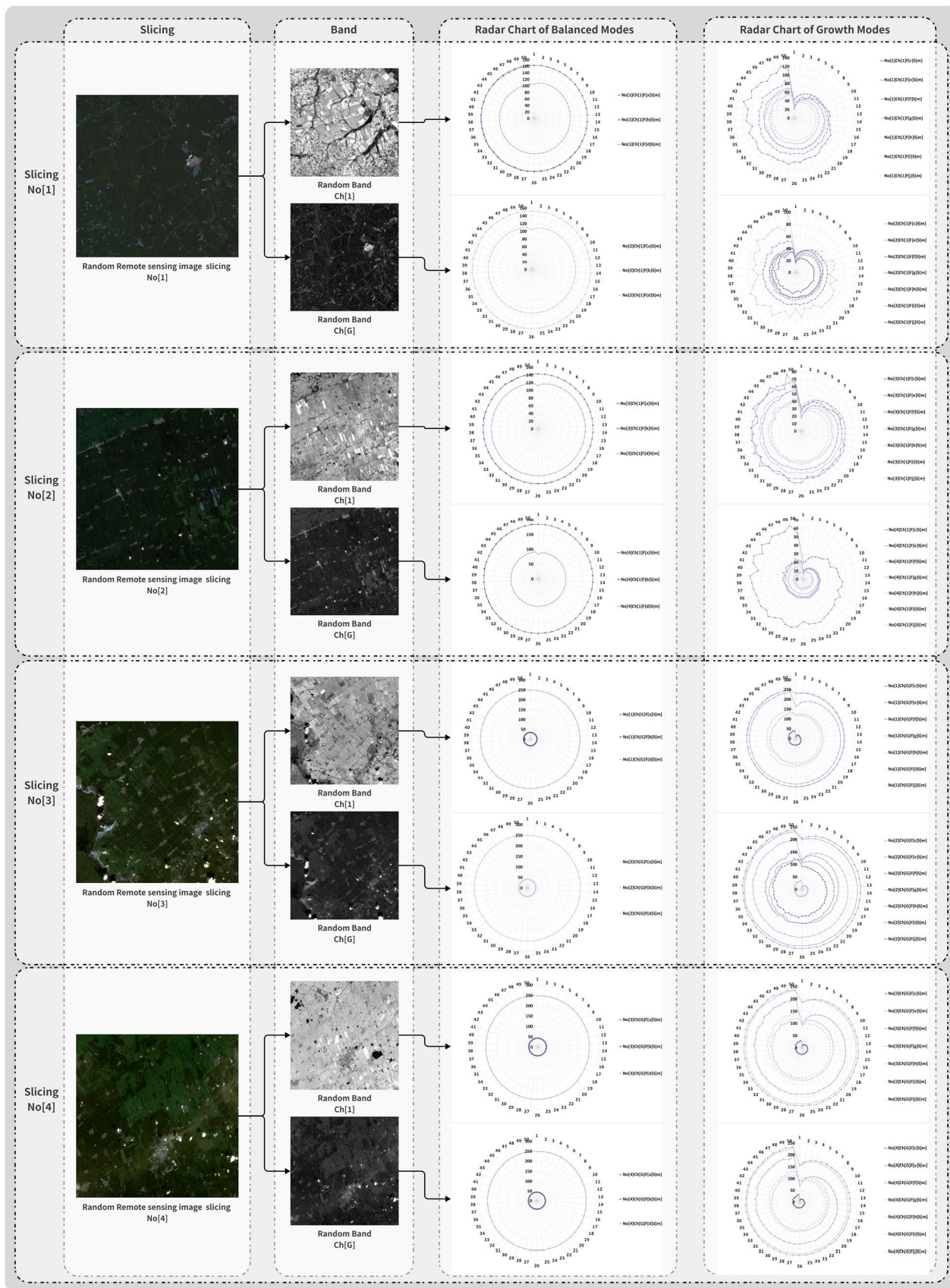


FIGURE 4. Longitudinal profile statistical analysis.

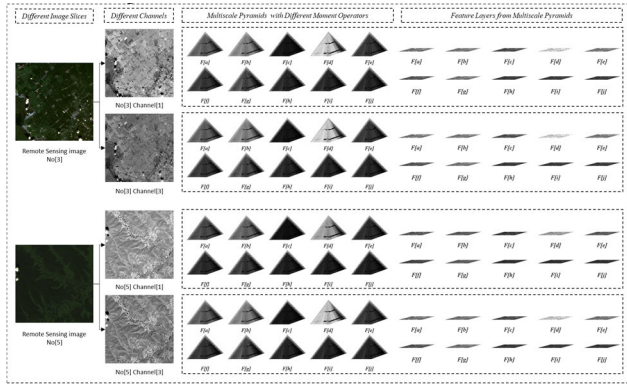


FIGURE 5. Cross section of the multiscale pyramid models (1).

we categorize and select features of the feature space and the corresponding feature operators.

1) LONGITUDINAL PROFILE ANALYSIS

In longitudinal profile analysis, we conducted experiments on a set of 8 images, each composed of 4 random image slices with 2 channels, using feature operators, F[a]-F[j] to establish the multiscale pyramid, shown in Figure 4. To obtain comprehensive experimental results, we set the scale levels relatively high at 50. We conducted statistical analysis on the total brightness values of the images at each scale in the multiscale pyramid. The statistical results show that the brightness values exhibit two distinct patterns of change across different scales: balance mode and growth mode. In the balance mode, the brightness values exhibit minimal variations across different scales, resulting in a nearly circular pattern in the radar chart. In the growth mode, the brightness values exhibit a trend of change with the scales, leading to a spiral pattern in the radar chart. Based on these statistical results, we tentatively categorized the feature operators into two groups: the balance mode group and the growth mode group. (Table 3).

2) CROSS SECTION ANALYSIS

In cross section analysis, we conducted a comparison on slices of the same scale but belonging to different feature spaces, as shown in figure 5.

Consistent with the previous experimental results, the balanced group(F[a], F[b], F[d]) and the growth group(F[c], F[e], F[f], F[g], F[h], F[i], F[j]) exhibit significant differences. The feature space belonging to the balance mode shows a higher degree of detail preservation and is more suitable for capturing detailed information, while the feature space belonging to the growth mode is more sensitive to edge information and is more suitable for edge extraction

Meanwhile, through the analysis of Feature[i] and Feature[j] in Figure 6, we observed that the utilization of high-order geometric feature operators resulted in a significant increase in computational load and fragmented

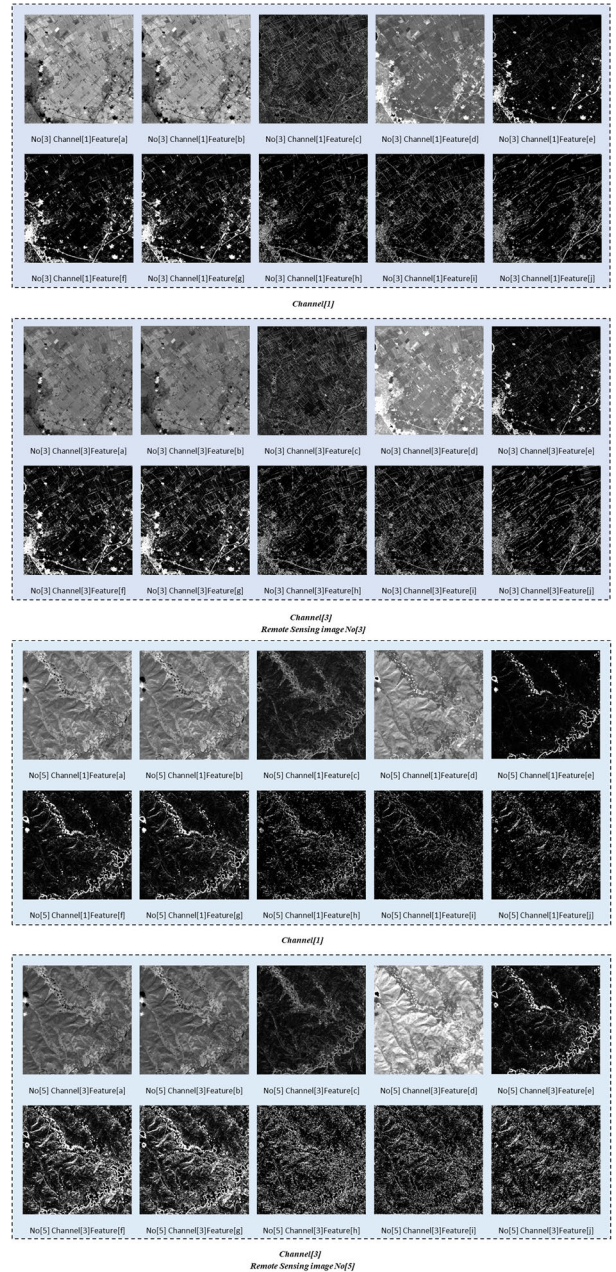


FIGURE 6. Cross section of the multiscale pyramid models (2).

information, thereby reducing the efficiency of information transmission. In the subsequent experimental procedure, excessively complex operators were filtered out to alleviate the computational burden and enhance information transmission efficiency. This ensured a more optimal and efficient feature extraction process.

III. IMAGE SEGMENTATION AND EVALUATION

The following is an overview of the proposed model: (1) remote sensing image preprocessing, (2) convolutional calculation of the multiscale feature pyramid, (3) spectral and shape feature extraction, (4) upsampling and multiscale convolution, and (5) evaluation.

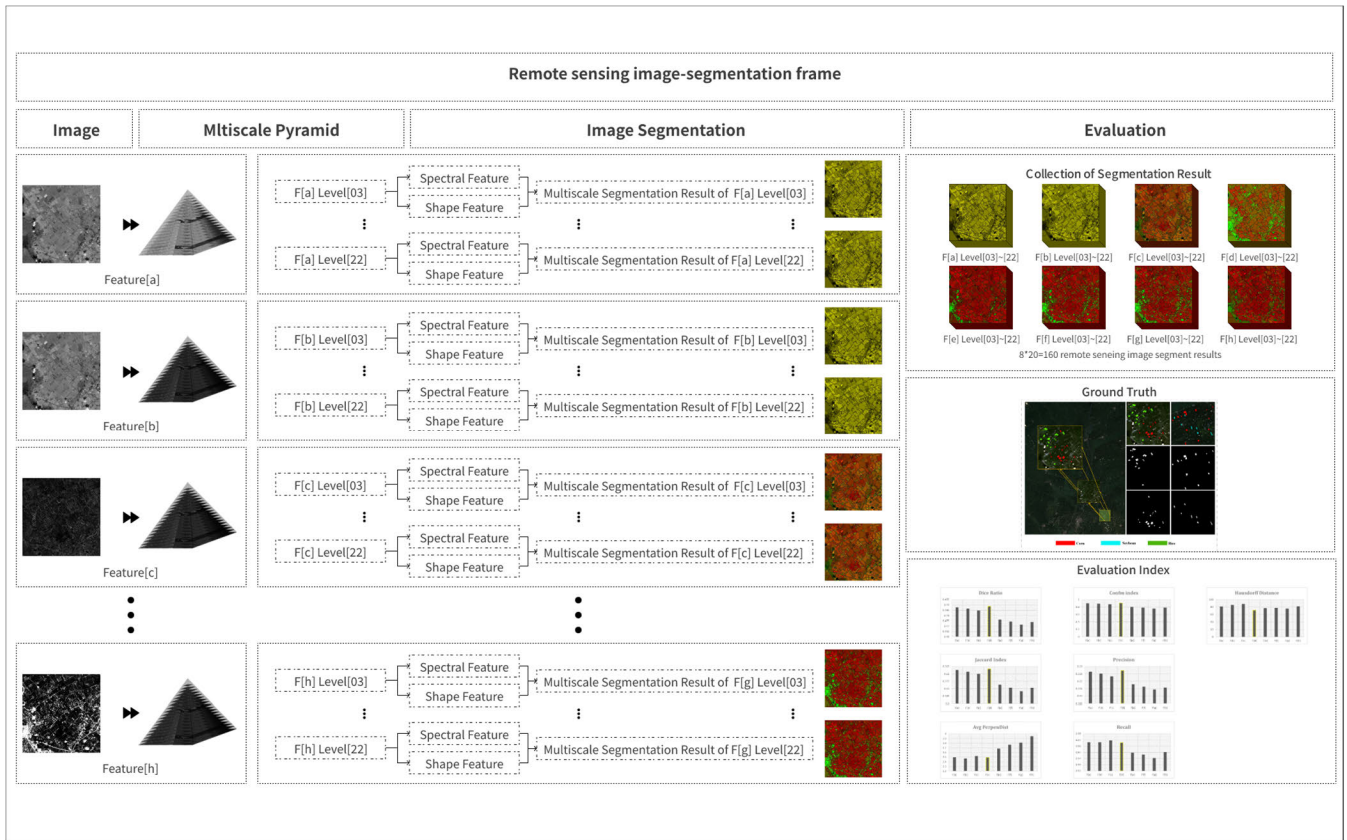


FIGURE 7. Remote sensing image-segmentation frame.

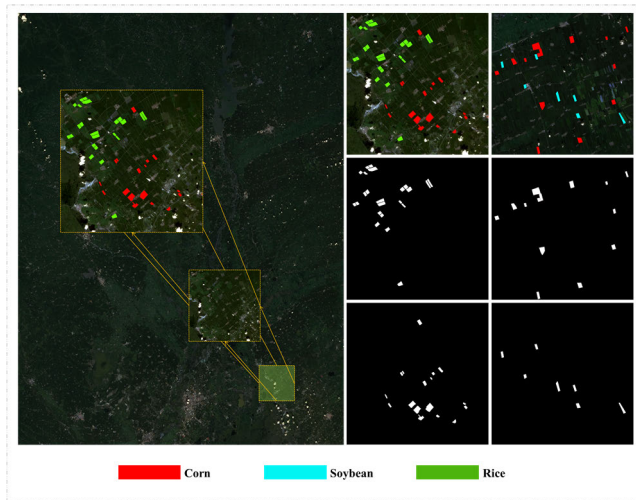


FIGURE 8. Ground-truth data.

A. REMOTE SENSING IMAGE SEGMENTATION FRAME

The spectral and shape feature-extraction algorithm is presented below [36]. This algorithm uses multi-scale classification network evolution for segmentation, and integrates spectral and shape features. The weighted value of the two is used as the heterogeneity evaluation criterion for minimum

heterogeneity merging, and finally obtains the segmentation result.

Spectral heterogeneity:

$$h_{spectral} = \sum_c w_c (n_1 (\delta_{mc} - \delta_{1c}) + n_2 (\delta_{mc} - \delta_{2c})) \quad (8)$$

Shape heterogeneity:

$$h_{compact} = n_1 \left(\frac{l_m}{\sqrt{n_m}} - \frac{l_1}{\sqrt{n_1}} \right) + n_2 \left(\frac{l_m}{\sqrt{n_m}} - \frac{l_2}{\sqrt{n_2}} \right) \quad (9)$$

$$h_{smooth} = n_1 \left(\frac{l_m}{b_m} - \frac{l_1}{b_1} \right) + n_2 \left(\frac{l_m}{b_m} - \frac{l_2}{b_2} \right) \quad (10)$$

$$h_{shape} = w_{smooth} h_{smooth} + (1 - w_{smooth}) h_{compact} \quad (11)$$

Spectral and Shape heterogeneity:

$$h = w_{spectral} h_{spectral} + (1 - w_{spectral}) h_{shape}$$

h : heterogeneity.

w : The weight.

n : The area of the region.

δ : The standard deviation

l : The edge length of this area.

b : The perimeter of the minimum bounding rectangle enclosing this area. (12)

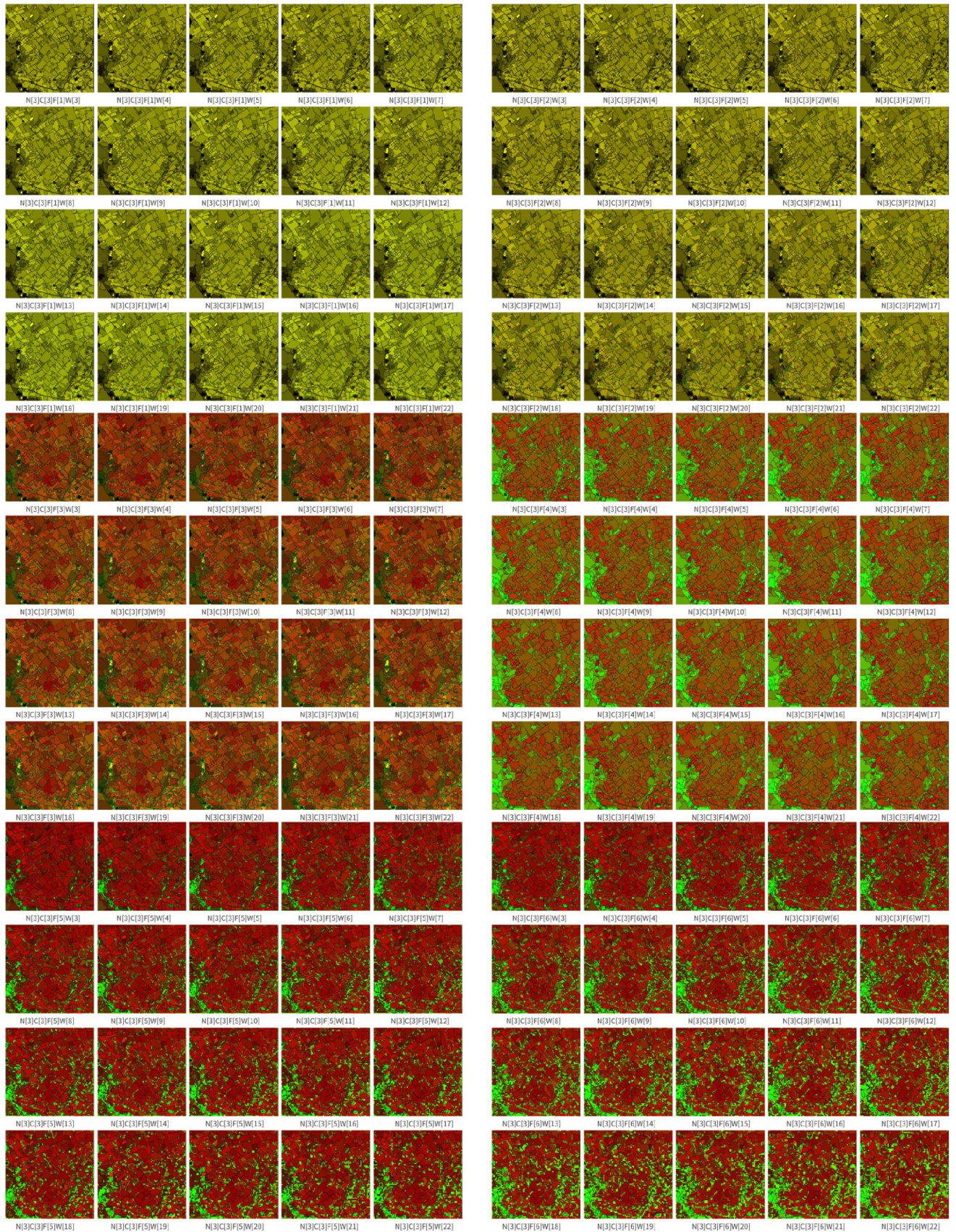


FIGURE 9. Remote sensing image-segmentation result.

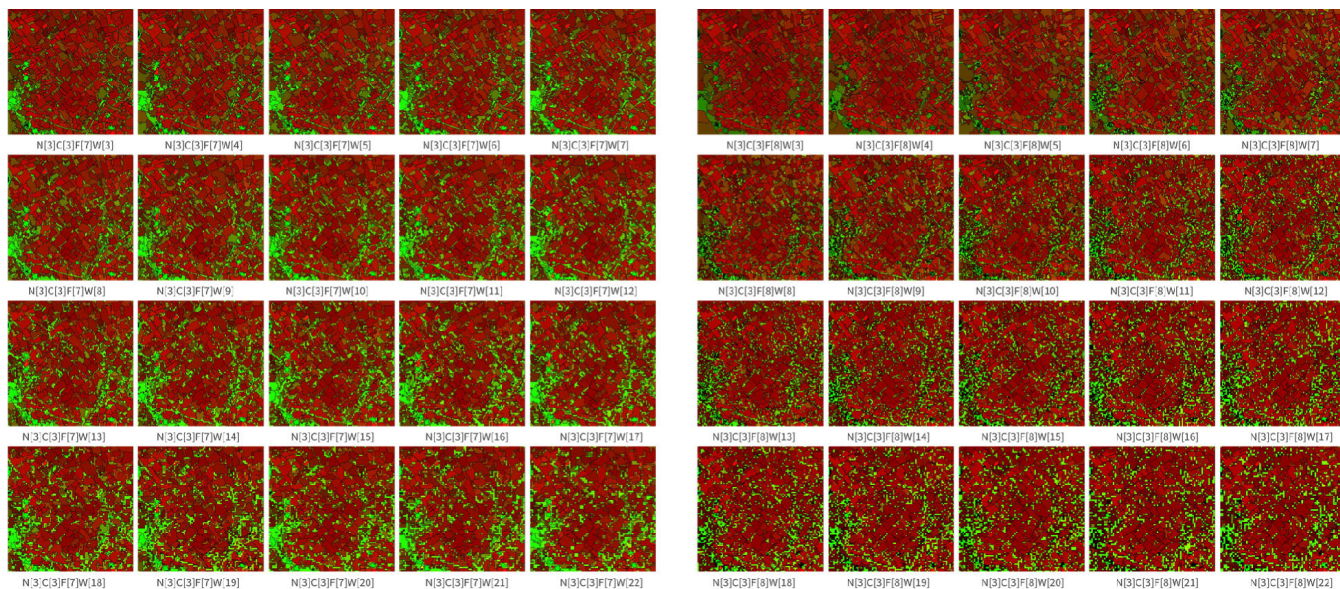


FIGURE 9. (Continued.) Remote sensing image-segmentation result.

The experimental data consisted of multispectral images from the Gaofen-6 satellite in the CHEOS program. Randomly selected image slices measuring 1500×1500 px were chosen. Each slice contained eight channels (i.e., band alpha 1, band alpha 2, band alpha 3, band alpha 4, band alpha 5, band R, band B, and band G). The CNNs extracted feature maps for each spectrum after the multispectral remote sensing image was preprocessed. The convolution kernel included Ma, Mb, Mc, Md, Me, Mf, Mg, and Mh.

B. SEGMENTATION RESULTS

Fig. 9 presents the multispectral remote sensing image-segmentation results, including 20×8 sub-images.

The multispectral remote sensing image-segmentation results contained many broken patches; thus, we could not visually observe the differences between the experimental and control groups. Herein, we introduced the ground-truth data and evaluation indices.

In the ground-truth data, red, blue, and green represent corn, soybeans, and rice, respectively. We converted the ground-truth data into binary images for the subsequent evaluation.

C. EVALUATION

Seven commonly used and mutually independent evaluation metrics were employed to evaluate the experimental results. The evaluation was performed on the segmentation results of 160 remote sensing images, as shown in Table 4. We divided the segmentation results of the 160 images into eight groups based on the variations in their multiscale feature pyramid operators: F[a], F[b], F[c], F[d], F[e], F[f], F[g], and F[h]. We calculated the means of the evaluation results for the eight groups and presented them as seven histograms (Figure. 10).

(1) The Dice coefficient derived from the kappa statistic as a reliability measure was used to calculate the balance between the intersection area and the average sum of the individual areas. With this evaluation result, the balanced mode group means (F[a], F[b], F[d]) showed a higher similarity with the ground-truth data than the growth mode group (F[c], F[e], F[f], F[g], F[h]). The F[d] indicates depicted the highest similarity (0.489).

(2) The Hausdorff distance is a measure of the maximum discrepancy between two true subsets in the space [37]. This demonstrates the inhomogeneity between the segmented blocks in image segmentation. A larger Hausdorff distance value results in a higher inhomogeneity between the segmented blocks. The differences among the means of the eight groups of experimental results were insignificant, but F[d] showed the shortest Hausdorff distance value of 71.540.

(3) The Jaccard index measures the ratio of the intersection area of the two sets divided by the area of their union. In this work, F[d] had the highest Jaccard index value of 0.323, denoting the highest similarity with ground-truth data.

(4) The average perpendicular index measures the average vertical distance between the experimental results and the ground-truth data. Compared to the growth mode group means (F[c], F[e], F[f], F[g], F[h]), the balanced mode group means (F[a], F[b], F[d]) exhibited shorter distances of 3.494, 3.472, and 3.493.

(5) Conformity coefficient is a global similarity coefficient used to measure the ratio of correctly segmented pixels to the number of incorrectly segmented pixels. F[d] obtained a high conformity coefficient of 2.323.

(6) Precision is the ratio of the number of correctly segmented pixels to the sum of the correctly and incorrectly

TABLE 4. Evaluation.

Evaluation		Scales																			
Result	Evaluation index	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
F[a]E[d]	Dice Ratio	0.49	0.49	0.49	0.48	0.49	0.49	0.49	0.49	0.49	0.48	0.48	0.49	0.49	0.48	0.49	0.48	0.49	0.49	0.49	0.49
F[b]E[h]	Hausdorff Distance	56.89	104.31	105.82	57.72	59.23	106.23	56.29	66.37	101.49	97.42	64.35	110.49	66.37	104.71	61.81	83.52	105.59	65.00	76.90	82.68
F[a]E[j]	Jaccard Index	0.33	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32
F[a]E[a]	Average Perpendicular	3.29	3.59	3.60	3.68	3.53	3.54	3.44	3.41	3.52	3.44	3.43	3.52	3.53	3.64	3.43	3.51	3.37	3.47	3.46	3.49
F[a]E[c]	Conformity	0.96	0.88	0.90	0.86	0.91	0.92	0.91	0.89	0.91	0.90	0.92	0.92	0.90	0.85	0.90	0.89	0.92	0.88	0.90	0.89
F[a]E[p]	Precision Index	0.33	0.32	0.33	0.32	0.33	0.33	0.33	0.32	0.33	0.33	0.33	0.33	0.33	0.32	0.33	0.32	0.33	0.32	0.33	0.33
F[a]E[r]	Recall Index	0.97	0.97	0.97	0.97	0.97	0.96	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.96	0.97	0.97	0.96	0.97	0.96	0.97
Evaluation		Scales																			
Result	Evaluation index	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
F[b]E[d]	Dice Ratio	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.48	0.48	0.49	0.49	0.48	0.49	0.48	0.49	0.49	0.48	0.48
F[b]E[h]	Hausdorff Distance	108.75	68.66	57.28	108.08	110.60	91.59	108.69	101.60	109.17	87.37	87.13	68.25	72.84	72.37	64.40	101.60	69.32	72.24	78.49	75.59
F[b]E[j]	Jaccard Index	0.32	0.32	0.32	0.32	0.32	0.33	0.32	0.32	0.32	0.32	0.32	0.32	0.33	0.32	0.32	0.32	0.32	0.32	0.32	0.32
F[b]E[a]	Average Perpendicular	3.42	3.38	3.48	3.43	3.54	3.49	3.41	3.55	3.50	3.60	3.52	3.51	3.43	3.38	3.67	3.32	3.37	3.45	3.43	3.57
F[b]E[c]	Conformity	0.91	0.92	0.88	0.89	0.92	0.95	0.90	0.89	0.89	0.87	0.85	0.89	0.94	0.87	0.89	0.87	0.88	0.89	0.85	0.84
F[b]E[p]	Precision Index	0.33	0.33	0.32	0.33	0.33	0.33	0.33	0.33	0.33	0.32	0.32	0.33	0.33	0.32	0.33	0.32	0.33	0.32	0.33	0.32
F[b]E[r]	Recall Index	0.97	0.97	0.97	0.97	0.97	0.96	0.97	0.96	0.97	0.97	0.97	0.97	0.97	0.96	0.96	0.97	0.97	0.96	0.96	0.96
Evaluation		Scales																			
Result	Evaluation index	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
F[c]E[d]	Dice Ratio	0.49	0.49	0.49	0.48	0.49	0.49	0.48	0.49	0.49	0.48	0.48	0.49	0.49	0.48	0.49	0.48	0.48	0.48	0.48	0.48
F[c]E[h]	Hausdorff Distance	104.66	99.57	96.75	111.89	96.25	93.48	54.13	70.61	102.11	104.66	108.26	108.75	72.01	88.53	63.95	54.82	67.23	105.10	63.95	105.60
F[c]E[j]	Jaccard Index	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32
F[c]E[a]	Average Perpendicular	3.51	3.33	3.60	3.66	3.32	3.51	3.50	3.46	3.46	3.59	3.73	3.50	3.47	3.65	3.55	3.41	3.70	3.52	3.59	3.48
F[c]E[c]	Conformity	0.90	0.91	0.92	0.87	0.89	0.88	0.85	0.88	0.90	0.83	0.83	0.89	0.88	0.86	0.89	0.87	0.87	0.86	0.84	0.86
F[c]E[p]	Precision Index	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33
F[c]E[r]	Recall Index	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97
Evaluation		Scales																			
Result	Evaluation index	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
F[d]E[d]	Dice Ratio	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.48	0.49	0.49	0.49	0.49	0.48	0.49	0.48	0.49	0.49
F[d]E[h]	Hausdorff Distance	102.47	56.75	74.71	72.45	57.07	61.85	57.07	110.07	92.31	83.36	64.20	56.75	56.52	78.85	62.63	80.08	81.63	57.07	55.32	69.64
F[d]E[j]	Jaccard Index	0.32	0.33	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.33	0.32	0.33	0.32	0.33	0.32	0.32	0.33
F[d]E[a]	Average Perpendicular	3.66	3.51	3.61	3.56	3.43	3.58	3.46	3.50	3.47	3.56	3.53	3.59	3.59	3.51	3.31	3.48	3.38	3.52	3.39	3.22
F[d]E[c]	Conformity	0.92	0.95	0.89	0.89	0.92	0.92	0.89	0.90	0.90	0.91	0.87	0.90	0.93	0.89	0.95	0.87	0.93	0.87	0.92	0.94
F[d]E[p]	Precision Index	0.33	0.33	0.32	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.32	0.33	0.33	0.33	0.33	0.32	0.33	0.32	0.33	0.33
F[d]E[r]	Recall Index	0.97	0.97	0.97	0.97	0.97	0.96	0.97	0.97	0.97	0.97	0.97	0.96	0.97	0.96	0.96	0.96	0.96	0.97	0.96	0.96
Evaluation		Scales																			
Result	Evaluation index	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
F[e]E[d]	Dice Ratio	0.49	0.49	0.49	0.48	0.48	0.47	0.48	0.47	0.48	0.48	0.47	0.47	0.47	0.47	0.47	0.47	0.47	0.47	0.47	0.47
F[e]E[h]	Hausdorff Distance	98.62	106.98	59.03	68.45	50.22	73.55	101.39	74.55	57.07	71.39	73.06	57.71	70.83	82.98	73.50	81.30	101.60	69.35	112.01	57.57
F[e]E[j]	Jaccard Index	0.32	0.32	0.32	0.32	0.32	0.31	0.32	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31
F[e]E[a]	Average Perpendicular	3.54	3.56	3.59	3.53	3.56	3.70	3.72	3.74	3.70	3.58	3.79	3.79	3.59	3.70	3.96	3.79	3.81	3.81	3.64	3.58
F[e]E[c]	Conformity	0.89	0.88	0.88	0.85	0.83	0.77	0.85	0.78	0.81	0.79	0.76	0.77	0.78	0.77	0.76	0.77	0.79	0.76	0.76	0.78
F[e]E[p]	Precision Index	0.33	0.32	0.33	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.31	0.31	0.31	0.31	0.31	0.31	0.32
F[e]E[r]	Recall Index	0.96	0.96	0.95	0.95	0.95	0.95	0.95	0.94	0.95	0.95	0.94	0.95	0.95	0.94	0.95	0.94	0.95	0.95	0.95	0.95
Evaluation		Scales																			
Result	Evaluation index	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
F[f]E[d]	Dice Ratio	0.49	0.48	0.48	0.48	0.47	0.47	0.47	0.48	0.47	0.47	0.47	0.47	0.47	0.47	0.47	0.47	0.47	0.47	0.47	0.47
F[f]E[h]	Hausdorff Distance	99.92	70.49	100.65	100.88	47.51	59.67	99.36	106.23	50.77	63.29	80.02	50.60	63.79	101.24	77.00	67.74	68.01	90.51	66.37	91.05
F[f]E[j]	Jaccard Index	0.32	0.32	0.32	0.32	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31
F[f]E[a]	Average Perpendicular	3.58	3.49	3.51	3.55	3.63	3.73	3.70	3.57	3.75	3.76	3.89	3.83	3.84	3.97	3.89	3.76	3.94	3.91	4.06	3.97
F[f]E[c]	Conformity	0.88	0.85	0.87	0.84	0.78	0.78	0.76	0.82	0.77	0.77	0.76	0.74	0.77	0.76	0.75	0.76	0.76	0.71	0.75	0.73
F[f]E[p]	Precision Index	0.33	0.32	0.33	0.32	0.32	0.32	0.31	0.32	0.32	0.32	0.31	0.31	0.32	0.31	0.31	0.31	0.31	0.31	0.31	0.31
F[f]E[r]	Recall Index	0.96	0.95	0.95	0.95	0.95	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.95	0.94	0.95	0.95
Evaluation		Scales																			
Result	Evaluation index	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
F[g]E[d]	Dice Ratio	0.49	0.48	0.48	0.47	0.47	0.47	0.47	0.47	0.47	0.47	0.47	0.46	0.47	0.47	0.46	0.47	0.47	0.46	0.47	0.47
F[g]E[h]	Hausdorff Distance	107.34	73.93	49.34	107.34	65.79	65.73	56.75	63.03	68.01	64.54	80.22	74.00	74.24	105.85	62.37	67.03	53.76	88.46	87.86	97.00
F[g]E[j]	Jaccard Index	0.32	0.32	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.30	0.31	0.30	0.31	0.30	0.31	0.30	0.31	0.31
F[g]E[a]	Average Perpendicular	3.55	3.43	3.64	3.89	3.76	3.91	3.96	3.73	3.82	3.71	3.74	3.93	3.80	3.66	4.12	3.97	3.88	3.93	3.87	3.96
F[g]E[c]	Conformity	0.92	0.88	0.82	0.76	0.74	0.74	0.73	0.77	0.74	0.74	0.77	0.69	0.75	0.77	0.70	0.72	0.75	0.69	0.73	0.72
F[g]E[p]	Precision Index	0.33	0.33	0.32	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.32	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31
F[g]E[r]	Recall Index	0.95	0.94	0.95	0.95	0.94	0.93	0.93	0.94	0.93	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.95	0.94	0.94
Evaluation		Scales																			
Result	Evaluation index	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
F[h]E[d]	Dice Ratio	0.49	0.48	0.48	0.48	0.47	0.47	0.47	0.47	0.48	0.47	0.47	0.48	0.47	0.48	0.47	0.47	0.47	0.46	0.47	0.48
F[h]E[h]	Hausdorff Distance	102.96	78.03	101.24	63.95	94.43	108.75	105.59	56.32	48.01	57.72	61.01	59.91	109.88	57.49	89.81	76.00	104.00	96.32		

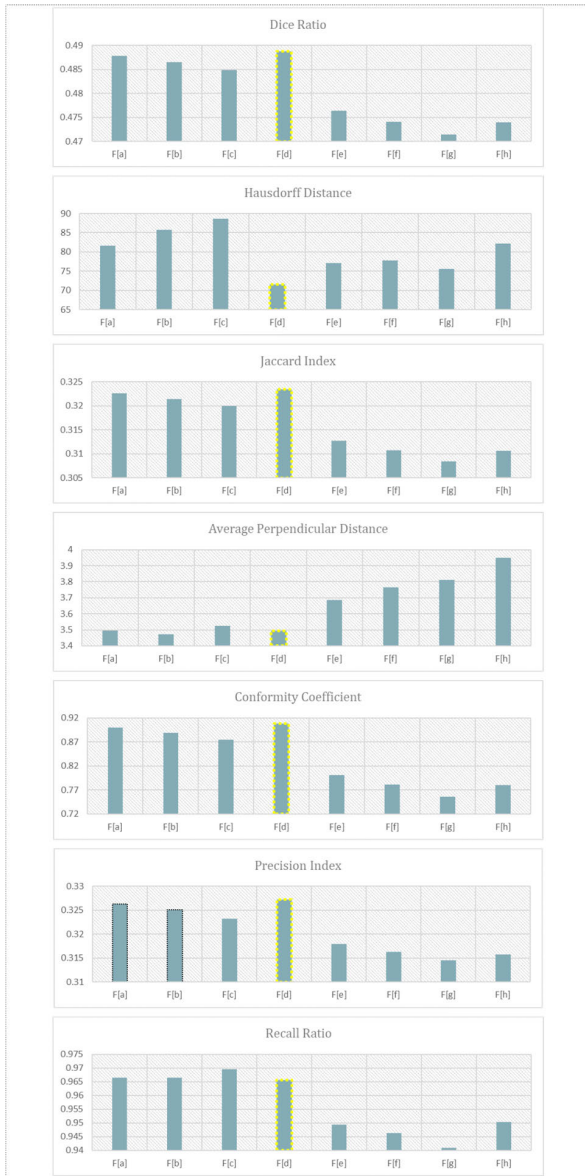


FIGURE 10. Multiscale image-segmentation results.

IV. CONCLUSION

In this study, 10 sets of geometric operators were established and the corresponding multiscale image feature pyramids were constructed. Statistical analysis showed that operators have two types of information transmission characteristics: balanced and growth. The obtained image features became more fragmented as operator complexity increased. After excluding the two operator groups with excessively high complexities, the remaining segmentation algorithm was applied to the remaining eight groups. The segmentation results obtained on authentic multispectral remote sensing images verified that the multiscale feature pyramid built with balanced mode operators had a better performance than the growth mode, especially in F[d]. An experimental segmentation evaluation demonstrated that compared to the precision of the growth mode group, the balanced mode group was generally improved by 1.5%–2%, the intraclass conformity

was improved by 1%–1.5%, and the intraclass chaos was reduced by 4.1%–10.1%. Overall, the experimental results indicate that during the construction of the pooling model, this experimental process can provide a relatively clear classification, screening, and evaluation method for the selection of feature operators. The construction of pooling models and the selection of feature operators are the foundation for constructing multi-scale spaces and various artificial neural network algorithms, as well as the key foundation for remote sensing image segmentation.

A potential limitation of this study is that although we have analyzed, classified, and evaluated feature operators, our classification of feature operators is still not detailed enough. For example, within the balanced group, F [d] performed significantly better than F [a] and F [b]. Although the statistical curves of F [d] and F [a], F [b] are very close in trend, there is a certain difference in numerical values. Therefore, we consider combining the previously proposed “deep-learning-based multiscale pyramid sieve and analysis module” theory in subsequent experiments to expand the scope of the experiment, further discuss whether the balanced mode group can be further differentiated, clarify its differentiation mechanism, comprehensively establish the optimal operator screening mechanism, and form a relatively complete technical system [38].

ACKNOWLEDGMENT

The authors would like to extend their heartfelt appreciation to Prof. Jian-Nong Cao, Liang Deng, and Li-Ping Lou for their guidance and assistance in the preparation of this article.

REFERENCES

- [1] L. I. Yan-Jun and Z. Ke, *Vision Bionica Image Guidance Technique and Application*. Beijing, China: National Defense Industry Press, 2006.
- [2] W. Nan-Nan and L. Gui-Ling, “Video quality evaluation models based on human visual properties,” *J. Image Graph.*, vol. 6, no. 6, pp. 523–527, 2001.
- [3] C. A. O. Jian-Nong, W. Ping-Lu, and D. Yu-Wei, “Automatic extraction technique of residential areas in high resolution remote sensing image,” *J. Wuhan Univ. Inf. Sci. Ed.*, vol. 39, no. 7, pp. 831–837, 2014.
- [4] C. En-Qing, W. Jian-Bo, and Q. I. Lin, “Novel multiscale edge tracking algorithm for edge detection,” *Appl. Res. Comput.*, vol. 32, no. 12, pp. 3801–3804, 2015.
- [5] E. J. Candes and D. L. Donoho, “Curvelets—A surprisingly effective nonadaptive representation for objects with edges,” *Curve and Surface Fitting: Saint-Malo*. Nashville, TN, USA: Vanderbilt Univ. Press, 1999, pp. 105–120.
- [6] M. N. Do and M. Vetterli, “The contourlet transform: An efficient directional multiresolution image representation,” *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2091–2106, Dec. 2005, doi: [10.1109/TIP.2005.859376](https://doi.org/10.1109/TIP.2005.859376).
- [7] W.-Q. Lim, “The discrete shearlet transform: A new directional transform and compactly supported shearlet frames,” *IEEE Trans. Image Process.*, vol. 19, no. 5, pp. 1166–1180, May 2010, doi: [10.1109/TIP.2010.2041410](https://doi.org/10.1109/TIP.2010.2041410).
- [8] A. Cohen and I. Daubechies, “Non-separable bidimensional wavelet bases,” *Revista Matemática Iberoamericana*, vol. 9, no. 1, pp. 51–137, Apr. 1993.
- [9] R. Q. Jia, “Multiresolution of l_p spaces,” *J. Math. Anal. Appl.*, vol. 184, no. 3, pp. 620–639, Jun. 1994, doi: [10.1006/jmaa.1994.1226](https://doi.org/10.1006/jmaa.1994.1226).
- [10] L. Rui-Lin, *High-Dimensional Wavelet Analysis*. New York, NY, USA: World Publishing Corporation, 1995.
- [11] E. Shelhamer, J. Long, and T. Darrell, “Fully convolutional networks for semantic segmentation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640–651, Apr. 2017.

- [12] J. Li, "A multi-scale U-Net for semantic segmentation of histological images from radical prostatectomies," in *Proc. AMIA Annu. Symp.*, Apr. 2017, pp. 1140–1148.
- [13] Q. Zhang, Q. Yuan, J. Li, Z. Yang, and X. Ma, "Learning a dilated residual network for SAR image despeckling," *Remote Sens.*, vol. 10, no. 2, p. 196, Jan. 2018, doi: 10.3390/rs10020196.
- [14] R. N. Rajaram, E. Ohn-Bar, and M. M. Trivedi, "RefineNet: Refining object detectors for autonomous driving," *IEEE Trans. Intell. Vehicles*, vol. 1, no. 4, pp. 358–368, Dec. 2016, doi: 10.1109/TIV.2017.2695896.
- [15] H. Fang and F. Lafarge, "Pyramid scene parsing network in 3D: Improving semantic segmentation of point clouds with multi-scale contextual information," *ISPRS J. Photogramm. Remote Sens.*, vol. 154, pp. 246–258, Aug. 2019.
- [16] C. Peng, X. Zhang, G. Yu, G. Luo, and J. Sun, "Large kernel matters—Improve semantic segmentation by global convolutional network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1743–1751.
- [17] V. Badrinarayanan, A. Kendall, and C. R. SegNet, "A deep convolutional encoder–decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.
- [18] A. Kendall, V. Badrinarayanan, and R. Cipolla, "Bayesian SegNet: Model uncertainty in deep convolutional encoder–decoder architectures for scene understanding," 2015, *arXiv:1511.02680*.
- [19] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, Apr. 2018, doi: 10.1109/TPAMI.2017.2699184.
- [20] H. Zhang, J. Xue, and K. Dana, "Deep TEN: Texture encoding network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 708–717.
- [21] Z. Yan, H. Zhang, Y. Jia, T. Breuel, and Y. Yu, "Combining the best of convolutional layers and recurrent layers: A hybrid network for semantic segmentation," 2016, *arXiv:1603.04871*.
- [22] Y. Wei, H. Xiao, H. Shi, Z. Jie, J. Feng, and T. S. Huang, "Revisiting dilated convolution: A simple approach for weakly- and semi-supervised semantic segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7268–7277.
- [23] A. Paszke, A. Chaurasia, S. Kim, and E. Culurciello, "ENet: A deep neural network architecture for real-time semantic segmentation," 2016, *arXiv:1606.02147*.
- [24] W. Liu, A. Rabinovich, and A. C. Berg, "Parse net: Looking wider to see better," 2015, *arXiv:1506.04579*.
- [25] Z. Li, "LSTM-CF: Unifying context modelling and fusion with LSTMs for RGB-D scene labeling," in *Proc. European Conf. Comput. Vis.*, Amsterdam, The Netherlands. New York, NY, USA: Springer, Oct. 2016, pp. 541–557.
- [26] W. Byeon and T. M. Breuel, "Supervised texture segmentation using 2D LSTM networks," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 4373–4377.
- [27] P. Baldi, and G. Pollastri, "The principled design of large-scale recursive neural network architectures-DAG-RNNs and the protein structure prediction problem," *J. Mach. Learn. Res.*, vol. 4, pp. 575–602, Dec. 2003.
- [28] P. O. Pinheiro, R. Collobert, and P. Dollár, "Learning to segment object candidates," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 1–7.
- [29] P. O. Pinheiro, T. Y. Lin, R. Collobert, and P. Dollár, "Learning to refine object segments," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Cham, Switzerland: Springer, 2016, pp. 75–91.
- [30] S. Zagoruyko, A. Lerer, T.-Y. Lin, P. Pinheiro, S. Gross, S. Chintala, and P. Dollár, "A MultiPath network for object detection," in *Proc. Brit. Mach. Vis. Conf.*, 2016, doi: 10.5244/c.30.15.
- [31] S. Ji, W. Xu, M. Yang, and K. Yu, "3D convolutional neural networks for human action recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 221–231, Jan. 2013, doi: 10.1109/TPAMI.2012.59.
- [32] R. Q. Charles, H. Su, M. Kaichun, and L. J. Guibas, "PointNet: Deep learning on point sets for 3D classification and segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 77–85.
- [33] H. Zhao, X. Qi, X. Shen, J. Shi, and J. Jia, "ICNet for real-time semantic segmentation on high-resolution images," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Sep. 2018, pp. 405–420.
- [34] D. Y. S. Yun, *Research on Evaluation Model of Human Vision in Stereo Video Based on Multi-Scale Analysis and Similarity China*. Jilin, China: Ji Lin Univ., 2011.
- [35] G. Y. Jing, "General multiscale analysis and Isocratic multiscale analysis and application," Doctoral dissertation, Jilin Univ., Jilin, China, 2000.
- [36] Z. J. Ping, "Analysis of image segmentation based on multi-scale method," *Coal Technol.*, vol. 32, no. 1, pp. 218–219, 2013.
- [37] H.-H. Chang, A. H. Zhuang, D. J. Valentino, and W.-C. Chu, "Performance measure characterization for evaluating NeuroImage segmentation algorithms," *NeuroImage*, vol. 47, no. 1, pp. 122–135, Aug. 2009.
- [38] D. Cao, J.-N. Cao, Q. Zhu, L.-J. Lou, and N.-Z. Xiao, "Deep learning-based multiscale pyramid sieve and analysis module in image segmentation," *IEEE Access*, vol. 11, pp. 22307–22319, 2023, doi: 10.1109/ACCESS.2023.3252400.



CAO DI was born in Henan, China, in 1988. She received the B.S. degree in geographic information systems and the M.S. degree in geological engineering from Chang'an University, Xi'an, China, in 2012 and 2014, respectively, where she is currently pursuing the Ph.D. degree in geoscience information systems.

Her research interests include image understanding, remote sensing image analysis, and geographic information systems.

Miss. Cao's awards and honors include the Hibiscus Student Fellowship (School of Earth Science and Resources, Chang'an University).



CAO JIAN-NONG was born in Xi'an, China, in 1963. He received the B.S. degree from Wuhan University, Wuhan, China, in 1987, the M.S. degree from Northwest University, Xi'an, in 2000, and the Ph.D. degree from Wuhan University, in 2005.

Since 2005, he has been with Chang'an University. He participated in national-level project and provincial-level topics, including "Research on feature structured multiscale analysis method

for information extraction from high-resolution remote sensing images" and "Hyperspectral image segmentation based on high-dimensional Markov network structure statistics." His research interests include remote sensing technology and application (including image understanding and image pattern recognition), high-resolution remote sensing information extraction technology, Lidar point cloud information extraction, photogrammetry and remote sensing technology (including photogrammetry and oblique photography), geographic information systems and application, 3S basic theory, and technology application. He is a member of the Jiu San Community.



DENG LIANG was born in Jiangxi, in 1985. He received the B.S. degree from East China University of Technology, Nanchang, China, in 2009, and the M.S. degree in software engineering from Chang'an University, Xi'an, China.

In 2011, he was with Esri (China) Information Technology Company Ltd., mainly responsible for development and implementation. Since November 2021, he has been an Architect with China Coal Technology and Engineering Group Xi'an Research Institute Company Ltd., mainly responsible for core product planning, architecture design, and core code development.



LOU LI-PING was born in Zhengzhou, in 1979. She received the B.S. degree in surveying and mapping engineering from Wuhan University, Wuhan, China, in 2007.

After graduation, she has been with Yellow River Engineering Consulting Company Ltd. Her professional title is a Senior Engineer. Her research interests include remote sensing technology, photogrammetry, and cartography.

...