

Dual-Concentrated Network With Morphological Features for Tree Species Classification Using Hyperspectral Image

Zhengqi Guo, Mengmeng Zhang [✉], Wen Jia [✉], Jinxin Zhang, and Wei Li [✉], *Senior Member, IEEE*

Abstract—At present, deep learning is a hot topic in the field of the classification of hyperspectral image (HSI), and it has aroused wide attention. However, in fine-grained classification tasks, such as tree species classification, the uncertain spectrum remains the major factor restraining the classification performance. In order to solve the dilemma of forest tree species classification, a dual-concentrated network with morphological features (DNMF) is proposed. First, mathematical morphology is used to extract the morphological features of HSI. Then, coarse-grained information is extracted from the original hyperspectral data, and fine-grained information is extracted from morphological features. After that, both morphological representations and spectral inputs are fed into DNMF, and the overall evaluation index and visual image are obtained. The advantage of DNMF is that it decouples the spatial and spectral information, and a multisource information fusion process is then simulated. Accordingly, DNMF obtains high tree species classification accuracy. In order to verify the superiority of DNMF, we choose Gaofeng State-owned Forest Farm in Guangxi Province and the Belgium dataset, which was collected near the western part of Belgium as the research area. Related experiments demonstrate that the DNMF model achieves clearly better classification performance over other competitive baselines.

Index Terms—Dual-concentrated network with morphological features (DNMF), hyperspectral image (HSI), mathematical morphology, tree species.

I. INTRODUCTION

HYPERSPECTRAL remote sensing technology uses hyperspectral sensors to image the target area in dozens to hundreds of continuous bands. At the same time, it contains rich spatial information and spectral information. With the improvement of sensor technology, high-resolution image has attracted more and more attention [1], [2]. In recent years, airborne hyperspectral image (HSI) technology has gradually matured, and the resolution of HSI has been significantly improved. In

Manuscript received 9 May 2022; revised 7 July 2022; accepted 6 August 2022. Date of publication 17 August 2022; date of current version 2 September 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 62001023 and Grant 42101403, in part by the China Postdoctoral Science Foundation under Grant 2021M700441, and in part by the National Key Research and Development Program of China under Grant 2019YFE0126700. (*Corresponding authors: Mengmeng Zhang; Wen Jia.*)

Zhengqi Guo, Mengmeng Zhang, Jinxin Zhang, and Wei Li are with the School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China (e-mail: 3120210840@bit.edu.cn; 7520200002@bit.edu.cn; zjx2016@foxmail.com; liwei089@ieee.org).

Wen Jia is with the Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, Beijing 100091, China (e-mail: jiawen@ifrit.ac.cn). Digital Object Identifier 10.1109/JSTARS.2022.3199618

2016, Pang et al. [3] designed the multisensor airborne system LiCHy (light detection and ranging, CCD, and hyperspectral). LiCHy can obtain HSI with a spatial resolution of 0.2 m. With high-resolution data blocks, HSI enables more accurate land cover analysis. Ren et al. [4] sliced the HSI data into data blocks for efficient computation of the covariance matrix and obtained excellent classification performance in 2014. Azar et al. [5] used spectral blocks with spatial groups to effectively exploit spectral and spatial information. In 2017, Zhong et al. [6] used the residual network to train and classify urban hyperspectral data with a spatial resolution of 1.3 m and achieved excellent results. Also, Zhong et al. [6] concluded that hyperspectral remote sensing earth observation technology is widely used in different fields, such as environmental monitoring, agricultural survey, mining survey, atmospheric research, and other fields. Although hyperspectral datasets are widely used, there are great differences between commonly used hyperspectral datasets and forestry hyperspectral datasets. To be specific, the spectral curves of different ground types in the commonly used public datasets are quite different. For example, Kumar et al. [7] used the open hyperspectral data (Indian pines data) that contains 16 ground object species, and different classes in this dataset represent clearly different spectral curves. However, the spectral reflectance in the same genus of tree species is very similar, and it is necessary to propose a new classification algorithm to solve the problem of “different objects with a similar spectrum” in tree species classification. Based on this classification challenge, the determination of forest tree species has become a new hotspot in forest hyperspectral data applications. In forest application, it can obtain spatial information on tree species composition and distribution pattern in forest, and establish a foundation for tree species classification, forest map drawing, and forest diversity protection. For example, Korpela et al. [8] proposed to use airborne data to classify tree species in 2010. Marcinkowska et al. [9] aimed to discover the potential of hyperspectral remote sensing data for mapping forest vegetation ecosystems in 2014. On this basis, Marcinkowska et al. [10] identified tree species in a large area of forest vegetation, which was conducive to the drawing of forest maps and the protection of forest diversity. Now many research try to introduce light detection and ranging (LiDAR) data into spectrum-driven tree species recognition task to further improve the classification performance [11], [12], [13]. Dalponte et al. [11] analyzed the use of airborne hyperspectral and LiDAR data for tree species classification in the Southern

Alps, and the diagnosis spectral resolution can be integrated with the rich geometrical information together. Matthew et al. [12] combined the tree crown information (such as the maximum tree height) in LiDAR data with the spectral reflectance knowledge of HSI, and final classification process was implemented by support vector machine. Similar with tree species classification, many fine-grained recognition tasks rely heavily on the full use of multivariate data, and some researchers have proposed some derivative algorithms based on multivariate data fusion. In 2013, Li et al. [14] proposed convolutional recurrent neural network (CRNN), which used two ways of the classification rule set generation for the CRNN evaluation. In 2018, Li et al. [15] proposed a coupled sparse tensor factorization (CSTF)-based approach for fusing hyperspectral and multispectral images. In 2018, Xu et al. [16] developed a two-tunnel convolutional neural network (CNN) framework to extract spectral-spatial features from HSI. Dian et al. [17] regulated hyperspectral and multispectral image fusion by CNN denoiser. In 2020, Hang et al. [18] proposed an efficient and effective framework to fuse hyperspectral and LiDAR data using two coupled CNN.

Returning to the problem of tree species identification, sometimes it is difficult for us to obtain the same LiDAR/multispectral image as the HSI shooting location. But we often have unprocessed external information that is underutilized. If the unprocessed external information can be used, the structural relationship of the original spatial information can be maintained. In 2021, Hang et al. [19] proposed a multitask generative adversarial network (MTGAN) to take advantage of the rich information from unlabeled samples. Another research found that prior information helps to improve CNN's learning ability. According to this, Hang et al. [20] further proposed an attention-aided CNN model for spectral-spatial classification of HSIs. To weaken the influence of "different objects with a similar spectrum" in the new algorithm under small training samples, we plan to use the fine-grained information (such as morphological features) to compensate for the ignored characteristics reflecting the classes differences in the coarse-grained details. Fine-grained image analysis is a long-standing problem in computer vision and pattern recognition, and it has been widely used in the real world [21]. For example, Yin et al. [22] proposed to use fine-grained image information (two local dynamic pose features) to solve the problem that people with similar clothes are difficult to distinguish. Furthermore, there are more fine-grained information in images to be mined. In recent years, mathematical morphology has been gradually applied to the mining of fine-grained spatial information of images and applied in HSI classification. Many morphological algorithms have proposed some feature attributes to enhance local geometric information accordingly. For example, Li et al. [23] proposed to capture the compactness attribute of circular shape and the elongation attribute of long strip shape. So, it is feasible to extract fine-grained information from HSI by the morphological algorithm.

Before mining fine-grained information in HSIs, it is necessary for us to understand the development of mathematical morphology algorithm. In the development of mathematical morphology, based on previous research, Plaza et al. [24] proposed morphological profiles (MP), which is an adaptive

mathematical morphology algorithm introducing the concept of MP into HSI processing. MP can effectively extract the spatial information of the image and achieve the effective classification of fuzzy objects. Since then, a series of mathematical morphology algorithms have been proposed. For example, Mura et al. [25] proposed attribute profiles (AP) based on attribute filtering (AF). In 2016, Ghamisi et al. [26] proposed the extinction profiles (EP) component tree (maximum or minimum tree) based on extinction filtering, which sheers the tree nodes based on extinction value to extract features. In 2019, Li et al. [23] proposed the local contain profiles (LCP), which is designed based on the topology tree. Compared with the component tree, the topological tree significantly enhances the stability of the classification process based on regional contexts in the image. In 2021, Cao et al. [27] creatively proposed a threshold-based local contain profile (TLCP) on the basis of EP and LCP. Based on the topology tree, TLCP flexibly set thresholds according to different physical properties of objects and retains more spatial information than LCP. In 2021, Hou et al. [28] proposed multiple morphological profiles (MMPs) to enhance the utilization of HSIs. Aforementioned morphological features enhance spatial information usage in HSI and improve the performance of HSI classification.

Although the traditional morphological algorithm has the advantages of crucial textures extraction and high-efficiency calculation, the classification performance of these methods is still limited and strongly subject to manual parameter selection as investigated by Dalponte et al. [29]. Therefore, a more efficient and autonomous classification framework is urgently needed. Deep learning methods, which act more dynamically to provide automation features, have been extensively employed for remote sensing image feature extraction and classification. In particular, Hinton [30] found that CNN is one kind of deep network that involves fewer parameters than a fully connected network does. CNN can directly process two-dimensional (2-D) images and reduce the setting burden of manual parameters. It mainly consists of three parts: convolution layer, pooling layer, and fully connection layer. In 2017, Raczko and Zagajewski [31] applied CNN to tree species classification for the first time, forming a new HSI classification framework. Based on the CNN framework, scholars effectively improve the accuracy of HSI classification and reduce the complexity of the manual intervention. Chen et al. [32] and Zhao et al. [33] jointly used dimension reduction and deep learning techniques for spectral and spatial feature extraction so as to achieve better classification performance. Wang et al. [34], [35] proposed two kinds of generic models to solve the classification difficulties caused by intraclass differences and interclass similarities. Lee et al. [36] proposed the contextual deep CNN, which can optimally explore local contextual interactions by jointly exploiting local spatio-spectral relationships of neighboring pixel. However, as investigated by Xue et al. [37], the problem of overfitting and information loss remain the greatest challenges in deep networks.

In view of the characteristics of forest hyperspectral remote sensing images, a DNMF is proposed, which solves the shortcoming of traditional morphology methods that have low classification accuracy and are susceptible to the Hughs phenomenon.

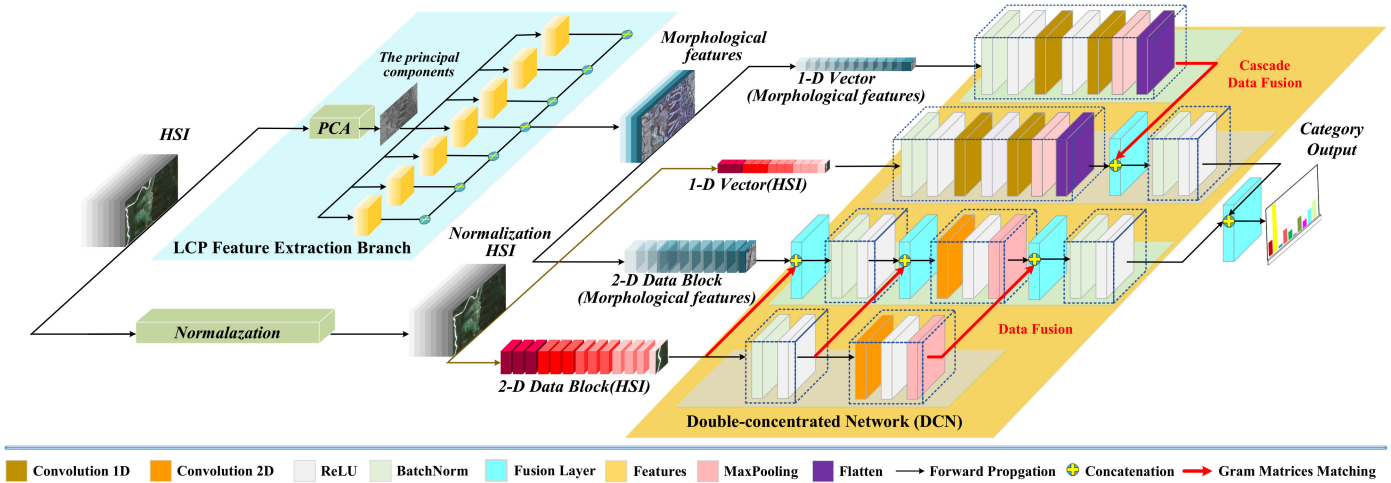


Fig. 1. Basic structure of the DNMF, including mathematical morphology for LCP feature extraction, the architecture of the DCN, and data fusion.

The proposed method realizes the decoupling of spatial-spectral information from a new perspective and couples it in an appropriate way, which significantly improves the performance of HSI classification. As a morphological feature extraction algorithm, LCP is applied to utilize the fine-grained spatial information of the input image. To be specific, in the first step, mathematical morphological features (LCP) are firstly extracted from the HSI. The second step is to extract the 1-D vector and 2-D data block from the original hyperspectral data and the morphological features. The third step is to acquire the final classification result based on a cross-converged dual-concentrated network.

The main contributions of this article are as follows.

- 1) In this article, we have creatively proposed a DNMF for tree-species classification. The spatial information and spectral information are first decoupled by morphological feature extraction, and the morphological features and the original data are recoupled in the collaborative utilization of spectral dimension [one-dimensional (1-D) vector] and spatial dimension (2-D data block). It imitates the process of multiple information fusion.
- 2) Due to the characteristics of forest vegetation itself, it is difficult to obtain a large number of samples, which sets obstacles for obtaining the classification results of tree species. This method can obtain high classification accuracy of tree species in the case of small samples, which is suitable for the present status of forest hyperspectral data. Also, it has solved the problem of “different objects with a similar spectrum” in tree species classification.

II. DUAL-CONCENTRATED NETWORK WITH MORPHOLOGICAL FEATURES

We propose a new HSI classification framework, which is called DNMF. The proposed framework is mainly composed of three parts: morphological feature (LCP) extraction of the HSI, HSI standardization, and multigranularity features cross fusion based on the dual centralized network. The overall architecture of the proposed classification framework is shown in Fig. 1.

Part A introduces the LCP features extraction. Part B explains the detailed settings of dual-concentrated network (DCN). Part C describes the feature fusion and HSI normalization.

A. Mathematical Morphology for LCP Feature Extraction

Before deep learning algorithms were applied to HSI, morphological algorithms were relatively common in HSI classification, but it does not mean that the traditional morphological algorithm is backward. In this article, the morphological feature extraction algorithm (LCP) is used to decouple HSI information, which contributes to more significant texture information acquisition, and effectively improving the utilization of spatial information.

Among the traditional morphological methods, EP and LCP are proposed in recent years. Both EP and LCP can effectively extract spatial information in HSI and have auxiliary functions for enhancing spatial information in woodland HSI. Specifically, EP is built on component tree (Max Tree/Min tree). Component tree is built on the basis of the pixel values between the connected fields in pixels, so the component tree is susceptible to external factors. For example, the presence of clouds and shadows affects the pixel values, the components tree will change accordingly. Hence, the feature extraction process is affected inevitably. Compared with EP, LCP is extracted based on the construction of topological tree, where the topological tree is built based on the inclusion relationship between connected areas. It is noteworthy that the inclusion relationship is less sensitive to the external factors, which significantly improves the stability compared with the component tree based on pixel values.

The left side of Fig. 1 is a brief illustration of the extraction process of LCP, while the detailed settings of LCP are divided into three steps, as shown in Fig. 2. First, the original hyperspectral data blocks are processed through the principal component analysis (PCA) algorithm, and corresponding principal components are selected according to our demand for information. Second, topological trees are constructed from the principal components. Third, we set the same extinction filter parameters for the topological trees constructed by different

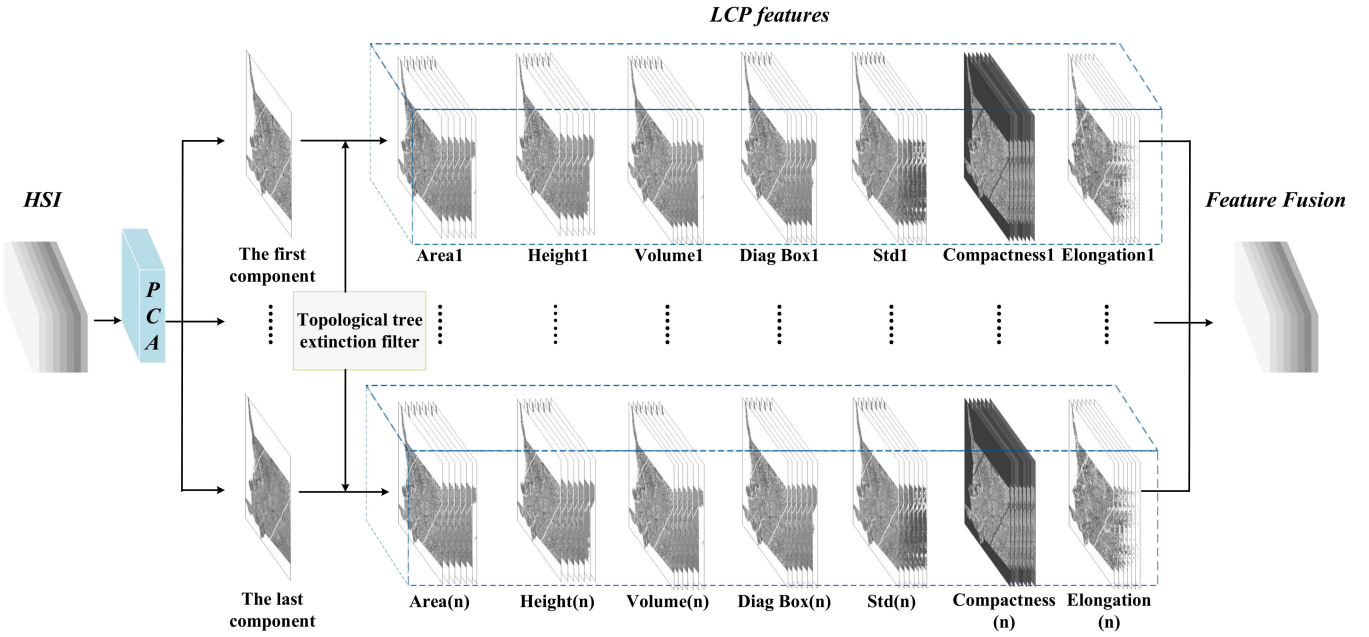


Fig. 2. Flow chart of feature extraction and integration based on LCP.

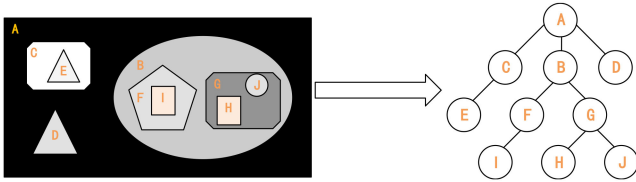


Fig. 3. Basic structure of the topology tree.

components. The parameters contain seven extinction values corresponding to seven pruned topological trees. Fourth, we reconstruct the seven pruned topological trees corresponding to each attribute (area, height, volume, etc.) and obtain seven reconstructed images. Finally, we stack seven images of each attribute to obtain a 49-dimensional morphological feature block. In order to better understand the feature extraction and integration of LCP, we introduce in detail the construction process of the topological trees and the principle of topological tree extinction filtering.

The construction of the topological tree can be explained by a simple image inclusion relationship. In Fig. 3, A represents the largest connected area that contains all pixels in the whole image. A is the root node, B, C, and D represent the second largest connected region, and B, C, and D are also child nodes of A. B contains regions F and G, so F and G are the child nodes of B. Following these principles, the topology tree is constructed.

Extinction filtering is the most critical part of the whole process of feature extraction and integration based on LCP after the construction of the topology tree. The extinction value in the topology tree is defined as follows: assuming that M is the local minimum connected region in image \mathbf{X} , and $\psi = (\varphi_\lambda)_\lambda$ is a series of decreasing connected inverse spread transforms. $\varepsilon_\varphi(M)$ represents the corresponding extinction value associated

with Ψ . If it is the global maximum λ value after extinction filtering, M is still the minimum connected region of $\varphi_\varphi(X)$. The definition of extinction value can be given by the following formula:

$$\varphi_\mu(M) = \sup\{\lambda \geq 0 | \forall \leq \lambda, M \subset \text{Min}(\varphi_\mu(x))\} \quad (1)$$

where $\text{Min}(\varphi_\mu(x))$ is a set containing all the minimum connected regions of $\varphi_\mu(x)$.

EF is a connected filter whose principle is to delete or retain the connected area corresponding to the leaf node and its branch nodes. Let $\text{Max}(X) = \{M_1, M_2, \dots, M_N\}$ be the set of minimum connected regions in the image \mathbf{X} , and each M_i ($i = 1, 2, \dots, N$) has an extinction value [defined by (1)]. First, M_i is sorted in decreasing order according to ω_i . Then, according to threshold n , the first n M_i are selected, and the corresponding leaf nodes are marked; finally, the branches with marked leaf nodes are retained according to the filtering strategy, while the branches with unmarked leaf nodes are cut off. This filtering process can be defined as follows:

$$EF^n = A_g^\nu \quad (2)$$

where EF obtains A_g^ν through reconstruction operation. g is a function for selecting markers. g can be expressed as follows:

$$g = \text{Max}_{i=1}^n(M_i^*) \quad (3)$$

where Max is the operation of selecting the minimum connected area with the largest extinction value, and M_i^* is the smallest connected area corresponding to the i highest extinction value.

After determining corresponding extinction value and minimum connection area of each node, we need to set an appropriate extinction value for EF to extract features and eliminate noise areas. The threshold is set as a^m , ($a = 1, 2, 3, \dots, m = 0, 2, 3, \dots, s - 1$), where a is the basic parameter and s is the number of thresholds.

B. Architecture and Details of the DCN

CNN-based classification is a hot topic in hyperspectral area in the recent years. However, inherent nonlinearity between materials and the corresponding spectral profiles still severely restrict the performance of existing CNN-based methods. To change this phenomenon, we have proposed a HSI processing architecture based on a DNMF.

The right side of the Fig. 1 is a schematic diagram of DCN. In particular, the DCN can be divided into two parts: fine-grained information processing branch and coarse-grained information processing branch. The standardization section will be introduced in Section C.

In the architecture of DCN, we design a new strategy to recouple the spectral-spatial information. Morphological features are extracted by LCP for collecting fine-grained information, while the normalized HSI retains the coarse-grained hyperspectral input, thus constructing a coarse-to-fine information set. In order to make the recoupling process in a more stable manner, the fine-grained information (morphological features) and the coarse-grained information are separately analyzed from the spatial and spectral dimension, respectively. Specifically, both the normalization HSI and the LCP features are divided into 2-D data blocks and 1-D vectors. The processing tunnels of fine-grained and coarse-grained inputs are consistent with each other. Further, taking the processing of normalization HSI as an example, the information feed forward settings are introduced as follow. The size of the normalized HSI is $H \times W \times D$, where 2-D input blocks and 1-D vector are first cropped from the input. We select based on the target pixel to construct the 1-D vector in the normalization HSI. In order to maintain the stability and registration of recoupling, the position of the target pixel of morphological features is the same as that of the normalization HSI. The dimension of the 1-D vector is $1 \times B$. Here, B is the number of the feature panels in the normalization HSI input, which is also the same as the bands number in original HSI. Particularly, Zhang et al. [38] demonstrated that an appropriate increase in the size of the input data helps to improve classification performance. So, the appropriate window size of $S \times S$ is set according to the practical demands. To keep the architecture robust, the 2-D data block is extracted from the same target pixel in the 1-D vector, which also contains the same bands number with original HSI. Taking the target pixel as the center, a cube with size of $S \times S \times B$ is selected as the 2-D block.

Finally, we classify according to the dimension of the input quantity and then fuse related features according to the class. In 1-D fusion channel, the network contains two 1-D convolutional layers, one batch normalization layer, two ReLU layers, a max-pooling layer, and the flatten layer. Uezato et al. [39] explained that the spectral curve is continuous. According to the characteristics of spectral curves provided by spectral dimensions of HSIs, 1-D vectors are fused in a cascade manner in Fig. 1. To ensure the consistency of spatial and spectral information of HSI, the structure of the 2-D fusion channel is also similar to that of the 1-D fusion channel. The difference between the two channels is that the 2-D fusion channel contains two fusions. Both the 1-D fusion channel and 2-D fusion channel pass through a batch normalization layer and a ReLU layer. Classification results were

obtained based on the fused features of 1-D and 2-D fusion channels.

C. Data Fusion and HSI Normalization

In the architecture of method, morphological features contain a large amount of fine-grained texture information, which greatly enhances spatial information. Spectral information in forest tree species data shares relatively weak discriminative power, while the enhanced spatial information strongly improve the whole model's performance on tree species classification.

The recoupling of diversified information is embedded in the DCN branch of the proposed model. First, morphological data and standardized HSI are preprocessed into 2-D data blocks as input, and then a 1-D vector is obtained based on 2-D data blocks. First, morphological features and standardized HSI are preprocessed into 2-D data blocks as input, and then 1-D vectors were obtained based on 2-D data blocks. The fusion method of morphological features and standard HSI 1-D vector is cascade fusion, which is designed for better utilize the spectral information. The fusion method of morphological features and standard HSI 2-D data blocks includes three times fusions, which effectively enhances the utilization of spatial information.

In the pretreatment step, HSI has been processed by a standardized process. Due to the influence of weather and light on HSIs, the spectral intensity will change. To reduce the differences in spectral curves between the same tree species and facilitate final classification, we normalize each band of the input HSI. In practical application, each band is normalized and the normalization HSI is obtained by feature concatenation operation. This filtering process can be defined as follows:

$$\text{Normalization}_i = \frac{\text{HSI}_i - \text{Min}(\text{HSI}_i)}{\text{Max}(\text{HSI}_i) - \text{Min}(\text{HSI}_i)} \quad (4)$$

where *Normalization* is obtained by subtracting the difference between the minimum value of this band and the spectral intensity of each pixel in the i band, and then it is divided by the difference between the maximum value and the minimum value. Normalization_i is the i th normalization, HSI_i is the i th band of the HSI. $\text{Max}(\text{HSI}_i)$ represents the pixel with the largest spectral intensity among all pixels in the i th band of HSI, and $\text{Min}(\text{HSI}_i)$ represents the pixel with the smallest spectral intensity among all pixels in the i th band of HSI.

III. EXPERIMENTS AND ANALYSIS

For the proposed DNMF framework, all the programs are implemented using Python language, and the network is constructed using Keras¹ and TensorFlow² deep learning framework. TensorFlow is an open-source software library for numerical computation using data flow graphs, and Keras can be seen as a simplified interface to Tensorflow.

A. Experimental Data

The performance of the proposed DNMF framework is evaluated on two forestry datasets, the GSFF dataset, and the Belgium

¹<https://github.com/fchollet/keras>

²<http://tensorflow.org/>

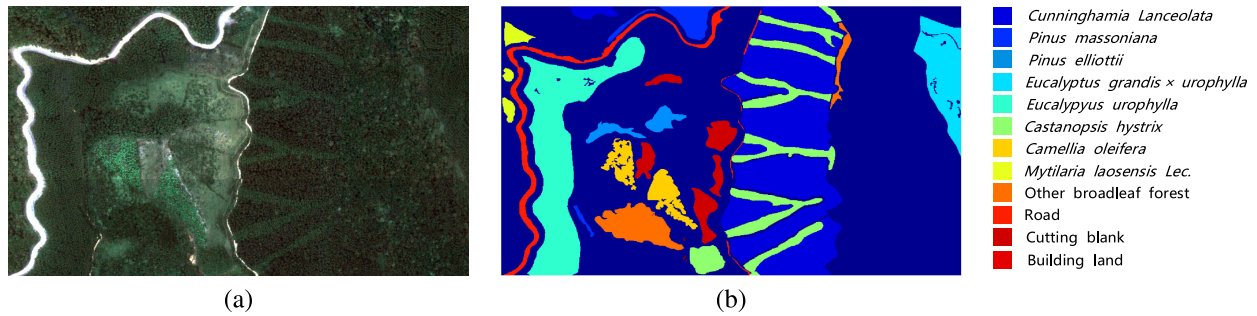


Fig. 4. (a) False-color image of the GSFF data, (b) Ground truth of the GSFF data.

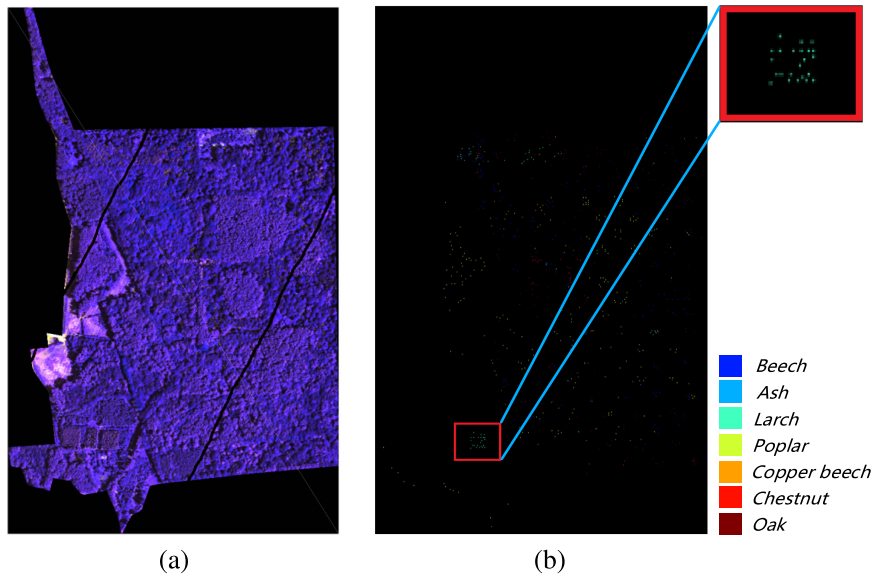


Fig. 5. (a) False-color image of the Belgium data, (b) Ground truth of the Belgium data.

dataset. The GSFF dataset is illustrated in Fig. 4, and Belgium dataset is illustrated in Fig. 5. There are few sample points in the ground truth of Belgium data, and the overall display effect is unsatisfactory, so some sample points are enlarged for display.

1) *GSFF Dataset*: In 2019, Zhang et al. [40] proposed GSFF data for the first time in the study of tree species classification using 3-D-CNN model to process HSI. Considering that it is often difficult to obtain enough samples for model training in actual remote sensing image classification tasks, in order to make the model usable in practical applications, randomly selecting 2.5% labeled pixels per class for training and all the other pixels in the ground-truth map for testing. Zhang et al. [40] proposed the GSFF dataset, which consists of 906×572 pixels, and it was gathered by the AISA Eagle II sensor in GSFF in Guangxi province in South China. There are 125 spectral channels covering the range from 400 to 990 nm with a spatial resolution of 1 m [41]. The GSFF dataset originally has 12 different land-cover classes, containing 9 forest vegetation categories.

2) *Belgium Dataset*: Considering that it is often difficult to obtain enough samples for model training in actual remote

sensing image classification tasks, in order to make the model usable in practical applications, randomly selecting 19.8% labeled pixels per class for training and all the other pixels in the ground-truth map for testing. Liao et al. [42] proposed the Belgium dataset, which consists of 649×1079 pixels, and it was gathered by the airborne prism experiment (APEX) near the western part of Belgium. There are 286 spectral channels covering the range from 400 to 1000 nm with a spatial resolution of 1.5 m. The Belgium dataset originally has 7 forest vegetation categories, containing 1450 trees.

B. Parameters Tuning

The numbers of training and testing samples of the GSFF dataset are listed in Table I, and the numbers of training and testing samples of the Belgium dataset are listed in Table II. Parameter setting can greatly affect the classification performance of the proposed framework, so the performance with different parameter setting is first evaluated.

1) *The Patch Size*: The performance of different sizes of the image patch is tested. Each experiment is evaluated by

TABLE I
NUMBERS OF TRAINING AND TESTING SAMPLES FOR THE GSFF DATASET

No.	Class	Training	Test
1	Cunninghamia Lanceolata	2019	78723
2	Pinus massoniana	134	5206
3	Pinus elliottii	104	3803
4	Eucalyptus grandis × urophylla	373	14558
5	Eucalyptus urophylla	1007	39267
6	Castanopsis hystrix	637	24824
7	Camellia oleifera	75	2925
8	Mytilaria laosensis Lec.	208	8115
9	Other broadleaf forest	269	10472
10	Road	262	10210
11	Cutting blank	254	9916
12	Building land	2	66
-	Total	5342	208350

TABLE II
NUMBERS OF TRAINING AND TESTING SAMPLES FOR THE BELGIUM DATASET

No.	Class	Training	Test
1	Beech	80	321
2	Ash	13	54
3	Larch	23	93
4	Poplar	83	333
5	Copper beech	16	64
6	Chestnut	13	54
7	Oak	60	243
-	Total	288	1162

TABLE III
CLASSIFICATION PERFORMANCE (%) OF THE PROPOSED DNMF FRAMEWORK WITH DIFFERENT WINDOW SIZES USING THE GSFF DATASET AND THE BELGIUM DATASET

Dataset	Metrics	OA	AA	Kappa
GSFF	11 × 11	97.39	97.60	96.17
	13 × 13	97.86	97.48	97.30
	15 × 15	97.95	97.22	97.40
Belgium	5 × 5	83.81	83.67	79.27
	7 × 7	86.48	86.36	82.86
	9 × 9	75.11	74.17	68.11

the overall accuracy (OA), the average accuracy (AA), and the Kappa coefficient. Experimental results demonstrate that the size of the image block has a certain impact on the classification performance of different datasets.

Table III lists that the window size of 15 × 15 offers better performance than others on the GSFF data. However, the Kappa coefficient of 15 × 15 is almost the same as the Kappa coefficient of 13 × 13, indicating that increasing the size of the input window could not continue to improve the consistency of classification. At the same time for improving computing efficiency and save computing resources. To sum up, we choose the window size of 15 × 15 for the following experiment with the GSFF dataset.

Table III concludes that the window size of 7 × 7 offers better performance than others on the Belgium dataset. To sum up, we choose the window size of 7 × 7 for the following experiment with the Belgium dataset.

2) *The Learning Rate*: The learning rate controls the convergence of the model by determining how far the weights move in the gradient direction in a mini-batch. Because the initial random weights are far away from the optimal value, a

TABLE IV
CLASSIFICATION PERFORMANCE (OA%) OF THE PROPOSED DNMF WITH VARYING LEARNING RATES

Dataset	Stage	Learning rate	OA
GSFF	1	0.01	85.46
		0.001	91.55
		0.0001	88.99
	2	0.001	95.42
		0.0001	97.31
		0.00001	92.59
Belgium	1	0.01	61.43
		0.001	77.86
		0.0001	57.14
	2	0.001	63.57
		0.0001	82.14
		0.00001	80.00

dual-concentrate network of HSI is trained with a large learning rate during the first stage. When the training of the HSI network is completed, the weights of the HSI branches are fixed. Mathematical morphology features are transmitted phase by phase, and the network is fine-tuned with a small learning rate. Adam optimizer is selected to derive the optimal learning rate combination. Specifically, Table IV further lists the optimal learning rate of two training stages with the GSFF dataset and the Belgium dataset.

3) *Analysis of the Number of Training Samples*: Labeled samples are very crucial in machine learning. However, data labeling is time-consuming and laborious in remote sensing data. In particular, it is difficult for forest surveys to mark the samples of HSIs, including remote locations, large forest areas, and other issues. Therefore, it is very important to use a small amount of labeled data to construct tree species classification model.

Here, we randomly select [10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%] samples from training set. The experiments are conducted ten times; then, the average results are reported. As shown in Fig. 6, the OA values tend to decrease as the number of training samples decreases. Especially, significant effect deterioration occurs in the result of 3DCNN in Fig. 6. This is because fewer training samples tend to overfitting. In particular, the total sample of the Belgium dataset is only 1450, which is far less than the total sample of GSFF. DNMF framework has the best stability on the Belgium dataset and the GSFF dataset. In the case of small samples, DNMF is the best for tree species identification. On the GSFF dataset, DNMF achieves 98.66% accuracy in the case of 10% training samples, which is significantly better than other deep learning algorithms. Besides, on the Belgium dataset, DNMF achieves 70.63% accuracy with 10% training samples, which is superior to other deep learning algorithms. To be concluded, DNMF obtains more representative morphological features by decoupling spatial-spectral information and recouples them during the training process in DCN, so as to maintain the stability of the algorithm and the accuracy of tree species recognition in the case of small samples.

4) *The Normalization Module*: In the DNMF framework, HSIs are normalized according to bands to obtain normalization

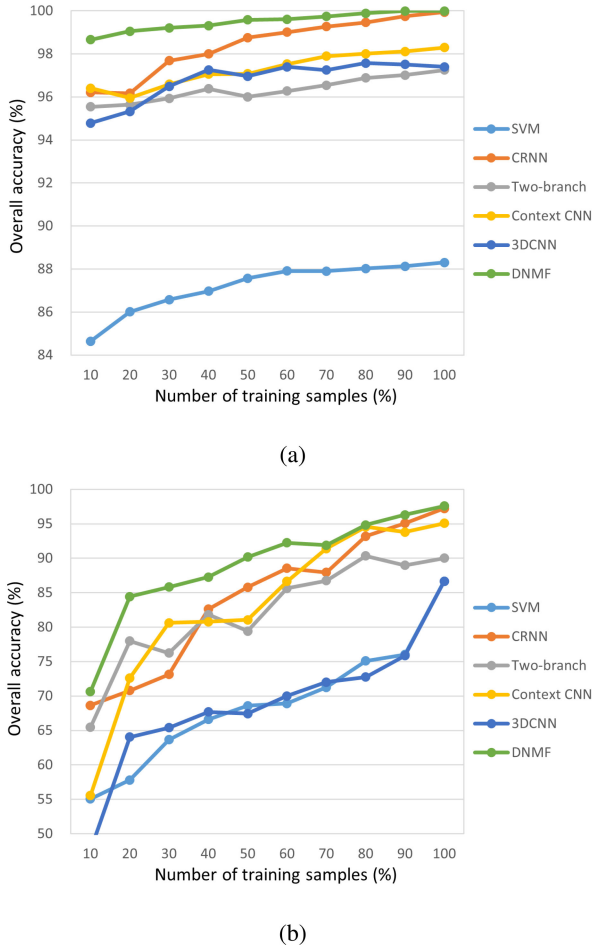


Fig. 6. Relationship between OA and the number of training samples on different dataset, (a) GSFF dataset, (b) Belgium dataset.

TABLE V
CLASSIFICATION PERFORMANCE (%) OF THE PROPOSED DNMF FRAMEWORK WITH OR WITHOUT NORMALIZATION OPERATOR USING THE GSFF DATASET AND THE BELGIUM DATASET

Dataset	Normalization	OA	AA	Kappa
GSFF	✓	97.86	97.48	97.30
	×	97.34	97.33	96.64
Belgium	✓	86.48	86.36	82.86
	×	82.52	82.15	77.71

HSI, which plays a vital role in the final fusion classification. However, it is difficult to know to what extent the normalization HSI plays a role in the final classification results. Therefore, we test the experimental results of the DNMF framework on the GSFF dataset and the Belgium dataset with or without normalization operation. Table V concludes that the DNMF framework with the normalization operator offers better performance than the DNMF framework without the normalization operator on the GSFF dataset and the Belgium dataset.

On the GSFF dataset, the overall accuracy (OA), the average accuracy (AA), and the Kappa coefficient are significantly improved by 0.15%–0.76% in Table V. This result shows that the spectral curves of the same ground object in the GSFF

dataset have little difference, and the normalization operation has little impact on the DNMF framework.

On the Belgium dataset, the overall accuracy (OA), the average accuracy (AA), and the Kappa coefficient are significantly improved by 3.96%–5.15% in Table V. This result shows that the spectral curves of the same ground object in the Belgium dataset are quite different. The normalization operation significantly improves the classification results of the DNMF framework and has a great impact on the DNMF framework.

C. Classification Performance

To demonstrate the performance of the proposed DNMF framework preprocessed remote sensing data, some traditional and state-of-the-art methods are compared, such as SVM, extreme deep learning, CRNN [14], Two-Branch CNN [16], Context CNN [36] and 3DCNN [40].

1) *Comparison of Classification Performance*: For a fair comparison, all the training and testing samples are set the same as in Tables I, III, and IV. OA, AA, Kappa coefficients, and classification accuracy are used to evaluate each model before and after LCP feature extraction branches is incorporated.

Table VI lists classification results of GSFF data. Compared with other methods before the fusion of morphological features branch in all methods, the proposed framework DNMF has a significant improvement in overall results, the overall accuracy is 97.68%. The kappa coefficient is also significantly improved, indicating that the consistency of the proposed DNMF framework in the GSFF dataset is improved without the morphological feature extraction branches. After the fusion of the morphological feature extraction branch in all methods, each method has improved the classification results of each category. DNMF framework still has the highest classification accuracy among the nine categories of tree species, indicating that the proposed network is relatively suitable for forest vegetation classification. Within each model, the OA, AA, and Kappa coefficient have increased significantly after proper fusion of the mathematical morphological feature extraction branch. As a traditional feature extraction method, mathematical morphology has improved the classification accuracy of SVM the most, reaching 14.49%. In the deep learning network, the lifting accuracy of each model is 3.18% for CRNN, 1.63% for Two-Branch CNN, 0.43% for 3DCNN, 2.76% for Context CNN, and 0.91% for Context CNN. Morphological features are useful for species classification.

Besides, in each classification algorithm, we separately use the morphological features for training, and the results are shown in Table VI. The results show that the DNMF framework has the best classification results. However, the classification results of different deep learning algorithms show that the simple use of morphological features will affect the tree species recognition results. The joint utilization of original HSI and morphological features can obtain the best tree species classification performance.

Table VII lists classification results of Belgium data. Unlike the GSFF dataset, the seven types of features in the Belgium

TABLE VI
COMPARISON OF THE CLASSIFICATION ACCURACY (%) AMONG THE PROPOSED METHOD AND THE BASELINES USING THE GSFF DATA

No.	Class(train/test)	Performance																	
		SVM			CRNN [14]			Two-branch CNN [16]			Context CNN [36]			3D-CNN [40]			Proposed "DNMF"		
		HSI	MF	HSI+MF	HSI	MF	HSI+MF	HSI	MF	HSI+MF	HSI	MF	HSI+MF	HSI	MF	HSI+MF	HSI	MF	HSI+MF
1	Cunninghamia Lanceolata(2019/78723)	70.18	71.93	75.51	94.46	90.50	96.37	95.78	89.72	96.59	94.54	94.50	95.00	94.34	91.92	94.27	97.06	6.10	97.41
2	Pinus massoniana(134/5206)	38.22	41.69	23.97	73.32	65.46	92.02	89.57	58.48	92.51	89.64	61.39	91.28	77.69	27.80	80.38	93.94	81.02	93.69
3	Pinus elliotii(104/4070)	31.06	42.22	51.49	77.95	64.39	90.37	90.39	57.59	92.35	96.67	60.89	85.50	81.47	46.03	82.16	97.60	78.68	97.90
4	Eucalyptus grandis×urophylla(373/14558)	63.86	98.55	98.46	98.06	99.90	99.94	97.16	99.87	99.98	98.04	99.94	98.49	90.92	95.40	92.06	99.87	100	100
5	Eucalyptus urophylla(1007/39267)	54.48	76.84	83.49	97.87	99.35	98.90	98.00	98.80	99.51	96.19	98.25	98.91	96.19	96.92	96.42	99.69	98.88	99.86
6	Castanopsis hystrix(637/24824)	42.76	79.58	77.77	81.96	91.42	89.80	86.62	85.85	90.40	79.66	91.33	87.77	85.04	90.26	85.62	90.78	92.14	93.06
7	Mytilaria laosensis Lec.(75/2925)	45.17	72.63	83.07	94.51	87.41	89.96	93.71	93.50	94.58	88.08	89.76	97.72	90.15	86.84	92.54	99.34	95.37	99.16
8	Camellia oleifera(208/8115)	67.63	85.87	77.60	97.36	45.37	98.57	96.97	96.56	98.80	92.85	96.76	97.89	94.87	93.71	96.62	99.56	98.17	99.55
9	Other broadleaf forest(269/10372)	56.49	84.83	85.16	85.57	81.93	94.74	91.35	94.33	95.21	78.74	94.73	93.68	88.44	92.95	88.75	94.22	96.47	97.77
10	Road(262/10246)	98.94	70.69	96.55	98.99	95.79	98.86	99.08	91.59	99.07	98.31	86.73	98.82	97.92	79.59	98.17	98.89	93.97	99.02
11	Cutting blank(254/9916)	90.04	47.18	72.06	99.48	75.78	99.62	98.90	79.48	99.60	99.39	93.83	99.79	98.25	88.60	99.37	99.84	98.63	99.83
12	Building land(2/68)	25.00	0	0	51.16	0	6.06	52.27	0	17.14	82.26	2.90	99.21	44.71	0	26.32	30.00	43.68	18.67
	OA	65.64	75.70	80.13	93.04	87.05	96.22	95.06	90.86	96.69	92.52	93.55	95.28	91.81	89.50	92.24	97.15	95.83	97.75
	AA	56.99	64.33	79.97	87.56	74.78	87.93	90.82	78.81	89.65	91.20	80.92	95.34	86.67	74.17	86.06	91.73	89.43	91.33
	Kappa	54.62	67.59	73.87	91.29	83.68	95.24	93.76	88.40	95.83	90.69	91.80	94.10	89.71	86.67	90.24	96.39	94.73	97.16

TABLE VII
COMPARISON OF THE CLASSIFICATION ACCURACY (%) AMONG THE PROPOSED METHOD AND THE BASELINES USING THE BELGIUM DATA

No.	Class(train/test)	Performance																	
		SVM			CRNN [14]			Two-branch CNN [16]			Context CNN [36]			3D-CNN [40]			Proposed "DNMF"		
		HSI	MF	HSI+MF	HSI	MF	HSI+MF	HSI	MF	HSI+MF	HSI	MF	HSI+MF	HSI	MF	HSI+MF	HSI	MF	HSI+MF
1	Beech(80/321)	41.48	40.35	49.89	66.77	53.05	68.58	74.80	56.11	74.07	59.47	50.81	72.55	59.08	51.64	57.78	80.72	56.58	83.25
2	Ash(13/54)	0	34.38	61.54	56.07	0	55.10	72.73	0	75.27	48.65	25.58	60.55	28.57	0	43.68	77.67	42.22	87.13
3	Larch(23/93)	54.93	62.16	72.09	62.94	1.94	69.14	84.26	0	85.71	81.14	61.08	84.85	30.67	18.69	60.67	89.12	78.45	88.65
4	Poplar(83/333)	47.97	48.20	70.44	85.12	48.31	84.18	87.81	56.32	89.46	88.19	55.43	85.23	79.53	51.63	86.38	92.80	70.94	91.30
5	Copper beech(16/64)	0	0	81.40	80.50	0	81.67	99.22	9.20	99.22	88.89	18.18	96.92	54.17	10.67	59.50	93.43	32.76	98.46
6	Chestnut(13/54)	0	0	25.00	48.19	0	43.68	61.36	0	61.54	25.81	9.09	58.06	18.39	0	43.01	77.42	25.81	80.00
7	Oak(61/243)	28.97	29.38	42.35	56.97	3.92	59.88	67.67	34.22	67.46	62.14	32.88	45.03	39.24	28.05	51.36	77.17	46.81	80.87
	OA	43.15	42.55	57.80	69.94	39.53	70.80	78.98	44.53	79.15	70.80	45.65	72.61	57.41	41.90	64.05	84.75	58.05	86.48
	AA	24.76	17.87	57.53	65.22	15.32	70.22	78.43	22.26	78.90	69.55	36.15	71.55	44.24	22.95	57.48	84.52	50.51	86.36
	Kappa	23.23	22.50	44.38	61.33	16.69	62.22	72.94	25.20	73.21	62.30	28.86	65.07	44.06	21.00	53.67	80.54	45.71	82.69

dataset are completely tree species. Compared with other methods before the fusion of morphological features branch in all methods, the proposed framework DNMF has a significant improvement in overall results, the overall accuracy is 86.48%. The kappa coefficient is also significantly improved, indicating that the consistency of the proposed DNMF framework in the Belgium dataset is improved without the morphological feature extraction branch. After the fusion of the morphological feature extraction branches in all methods, each method has improved the classification results of each category. DNMF framework still has the highest classification accuracy among the seven categories of tree species. Within each model, the OA, AA, and Kappa coefficient have increased significantly after proper fusion of the mathematical morphological feature extraction branch. As a traditional feature extraction method, mathematical morphology has improved the classification accuracy of SVM the most, reaching 14.65%. In the deep learning network, the lifting accuracy of each model is 0.86% for CRNN, 0.17% for Two-Branch CNN, 6.63% for 3DCNN, 1.89% for Context CNN, and 1.73% for Context CNN. Morphological features are useful for species classification.

Besides, in each classification algorithm, we separately use the morphological features after feature extraction for training, and the results are shown in Table VII. The results show that the DNMF framework has the best classification results. However, the classification results of different deep learning algorithms show that the simple use of morphological features will affect the tree species recognition results. The joint utilization of original HSI and morphological features can obtain the best tree species classification performance.

2) *Visual Comparison:* For visual evaluation of the proposed DNMF framework, the visual map of each aforementioned method for the GSFF data is illustrated in Fig. 7. In Fig. 7, it can be seen that in different models, the visual images obtained by the fusion of the mathematical morphology feature extraction branch are closer to the ground truth map than before the fusion with less noise. Among all the visual images, the proposed DNMF framework is closest to the true value and has the least noise.

IV. CONCLUSION

In this article, a framework based on DNMF processing channel has been proposed. Compared with the existing processing methods, DNMF framework creatively integrates the morphological feature extraction branch into the HSI processing process, decouples the spatial information from HSI and obtains richer and refined spatial information. In different HSI processing frameworks, the integration of LCP morphological feature extraction branches improved the experimental results to varying degrees, which fully demonstrated that the morphological features strongly enhance the spatial information utilization in the process of forest HSI processing. DNMF framework performed best among all baselines because both spectral, fine-grained spatial, and coarse-grained spatial information have been fully utilized in the proposed method. Thus, DNMF achieved superior classification performance even with small training samples of forest HSIs.

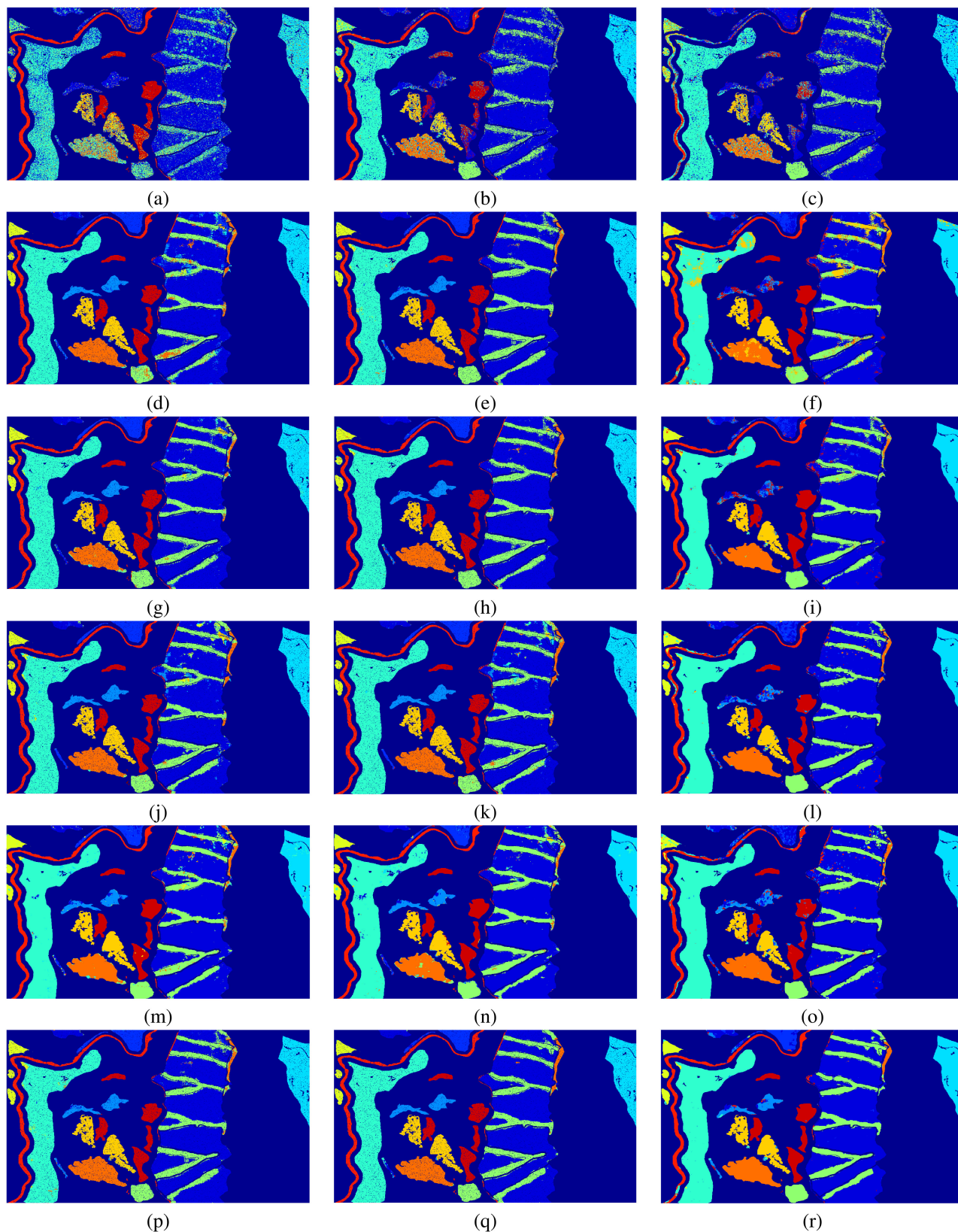


Fig. 7. Classification maps from the proposed DNMF and the baselines on the GSFF Data, (a) SVM: 65.64%, (b) SVM(HSI+MF): 80.13%, (c) SVM(MF): 75.70%, (d) CRNN: 93.04%, (e) CRNN(HSI+MF): 96.22%, (f) CRNN(MF): 87.05%, (g) Two-branch CNN: 95.06%, (h) Two-branch CNN(HSI+MF): 96.69%, (i) Two-branch CNN(MF): 90.86%, (j) Context CNN: 92.52%, (k) Context CNN(HSI+MF): 95.28%, (l) Context CNN(MF): 93.55%, (m) 3DCNN: 91.81%, (n) 3DCNN(HSI+MF): 92.24%, (o) 3DCNN(MF): 89.50%, (p) Proposed “DNMF”: 97.15%, (q) Proposed “DNMF”(HSI+MF): 97.75%, (r) Proposed “DNMF”(MF): 95.83%.

REFERENCES

- [1] B. Hou et al., "Object detection in high-resolution panchromatic images using deep models and spatial template matching," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 2, pp. 956–970, Feb. 2020.
- [2] X. Sun et al., "Automated high-resolution earth observation image interpretation: Outcome of the 2020 Gaofen challenge," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 8922–8940, 2021.
- [3] Y. Pang et al., "Lichy: The Caf's LiDAR, CCD and hyperspectral integrated airborne observation system," *Remote Sens.*, vol. 8, no. 5, 2016, Art. no. 398.
- [4] J. Ren et al., "Effective feature extraction and data reduction in remote sensing using hyperspectral imaging [applications corner]," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 149–154, Jul. 2014.
- [5] S. Ghanbari Azar et al., "Hyperspectral image classification based on sparse modeling of spectral blocks," *Neurocomputing*, vol. 407, pp. 12–23, 2020.
- [6] Z. Zhong et al., "Deep residual networks for hyperspectral image classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2017, pp. 1824–1827.
- [7] A. S. Kumar et al., "Hyperspectral image classification by a variable interval spectral average and spectral curve matching combined algorithm," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 12, no. 4, pp. 261–269, 2010.
- [8] I. Korpela, H. O. Ørka, M. Maltamo, T. Tokola, and J. Hyypää, "Tree species classification using airborne LiDAR - effects of stand and tree parameters, downsizing of training set, intensity normalization, and sensor type," *Silva Fennica*, vol. 44, no. 2, pp. 319–339, 2010.
- [9] A. Marcinkowska et al., "Mapping vegetation communities of the karkonosze national park using apex hyperspectral data and support vector machines," *Miscellaneous Geographica*, vol. 18, no. 2, pp. 23–29, 2014.
- [10] A. Modzelewska, F. E. Fassnacht, and K. Stereńczak, "Tree species identification within an extensive forest area with diverse management regimes using airborne hyperspectral data," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 84, 2020, Art. no. 101960.
- [11] M. Dalponte, L. Bruzzone, and D. Gianelle, "Tree species classification in the southern alps with very high geometrical resolution multispectral and hyperspectral data," in *Proc. IEEE 3rd Workshop Hyperspectral Image Signal Process., Evol. Remote Sens.*, 2011, pp. 1–4.
- [12] C. Matthew et al., "Mapping savanna tree species at ecosystem scales using support vector machine classification and BRDF correction on airborne hyperspectral and LiDAR data," *Remote Sens.*, vol. 4, no. 11, pp. 3462–3480, 2012.
- [13] Y. Zhu et al., "Joint feature extraction for multi-source data using similar double-concentrated network," *Neurocomputing*, vol. 450, pp. 70–79, 2021.
- [14] W. Li, L. Cao, D. Zhao, X. Cui, and J. Yang, "Crnn: Integrating classification rules into neural network," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, 2013, pp. 1–8.
- [15] S. Li, R. Dian, L. Fang, and J. M. Bioucas-Dias, "Fusing hyperspectral and multispectral images via coupled sparse tensor factorization," *IEEE Trans. Image Process.*, vol. 27, no. 8, pp. 4118–4130, Aug. 2018.
- [16] X. Xu et al., "Multisource remote sensing data classification based on convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 937–949, Feb. 2018.
- [17] R. Dian, S. Li, and X. Kang, "Regularizing hyperspectral and multispectral image fusion by CNN denoiser," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 3, pp. 1124–1135, Mar. 2021.
- [18] R. Hang et al., "Classification of hyperspectral and LiDAR data using coupled CNNs," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 7, pp. 4939–4950, Jul. 2020.
- [19] R. Hang et al., "Classification of hyperspectral images via multitask generative adversarial networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 2, pp. 1424–1436, Feb. 2021.
- [20] R. Hang et al., "Hyperspectral image classification with attention-aided CNNs," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 3, pp. 2281–2293, Mar. 2021.
- [21] X. Sun et al., "Fair1m: A benchmark dataset for fine-grained object recognition in high-resolution remote sensing imagery," *ISPRS J. Photogrammetry Remote Sens.*, vol. 184, pp. 116–130, 2022.
- [22] J. Yin, A. Wu, and W. S. Zheng, "Fine-grained person re-identification," *Int. J. Comput. Vis.*, vol. 128, no. 12, pp. 1654–1672, 2020.
- [23] W. Li et al., "Feature extraction for hyperspectral images using local contain profile," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 12, pp. 5035–5046, Dec. 2019.
- [24] A. Plaza et al., "A new approach to mixed pixel classification of hyperspectral imagery based on extended morphological profiles," *Pattern Recognit.*, vol. 37, no. 6, pp. 1097–1116, 2004.
- [25] M. D. Mura et al., "Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 3, pp. 542–546, May 2011.
- [26] N. He et al., "Hyperspectral images classification by fusing extinction profiles feature," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2017, pp. 2267–2270.
- [27] D. Cao et al., "Hyperspectral and infrared image collaborative classification based on morphology feature extraction," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 4405–4416, 2021.
- [28] Z. Hou et al., "Hyperspectral change detection based on multiple morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–12, 2022.
- [29] M. Dalponte et al., "The role of spectral resolution and classifier complexity in the analysis of hyperspectral images of forest areas," *Remote Sens. Environ.*, vol. 113, no. 11, pp. 2345–2355, 2009.
- [30] G. E. Hinton et al., "Improving neural networks by preventing co-adaptation of feature detectors," *Comput. Sci.*, vol. 3, no. 4, pp. 212–223, 2012.
- [31] E. Raczko and B. Zagajewski, "Comparison of support vector machine, random forest and neural network classifiers for tree species classification on airborne hyperspectral apex images," *Eur. J. Remote Sens.*, vol. 50, no. 1, pp. 144–154, 2017.
- [32] Y. Chen et al., "Deep feature extraction and classification of hyperspectral images based on convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6232–6251, Oct. 2016.
- [33] W. Zhao and S. Du, "Spectral-spatial feature extraction for hyperspectral image classification: A dimension reduction and deep learning approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4544–4554, Aug. 2016.
- [34] X. Wang et al., "Caps-TripleGAN: GAN-assisted Capsnet for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 7232–7245, Sep. 2019.
- [35] X. Wang et al., "Cva²e: A conditional variational autoencoder with an adversarial training process for hyperspectral imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 8, pp. 5676–5692, Aug. 2020.
- [36] H. Lee and H. Kwon, "Going deeper with contextual CNN for hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 26, no. 10, pp. 4843–4855, Oct. 2017.
- [37] Z. Xue et al., "Attention-based second-order pooling network for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 11, pp. 9600–9615, Nov. 2021.
- [38] M. Zhang, W. Li, and Q. Du, "Diverse region-based CNN for hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2623–2634, Jun. 2018.
- [39] T. Uezato et al., "A novel endmember bundle extraction and clustering approach for capturing spectral variability within endmember classes," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 11, pp. 6712–6731, Nov. 2016.
- [40] B. Zhang, L. Zhao, and X. Zhang, "Three-dimensional convolutional neural network model for tree species classification using airborne hyperspectral images," *Remote Sens. Environ.*, vol. 247, 2020, Art. no. 111938.
- [41] W. Jia et al., "A kernel-driven BRDF approach to correct airborne hyperspectral imagery over forested areas with rugged topography," *Remote Sens.*, vol. 12, no. 3, 2020, Art. no. 432.
- [42] W. Liao et al., "Deep learning for fusion of apex hyperspectral and full-waveform Lidar remote sensing data for tree species mapping," *IEEE Access*, vol. 6, pp. 68716–68729, 2018.



Zhengqi Guo received the B.S. degree in electronic information engineering from Beijing University of Chemical Technology, Beijing, China, in 2021. He is currently working toward the M.S. degree in information and communication engineering with the School of Information and Electronics, Beijing Institute of Technology, Beijing, China.

His research interests include image processing, multisource data fusion, and pattern recognition.



Mengmeng Zhang received the B.S. degree in computer science and technology from the Qingdao University of Science and Technology, Qingdao, China, in 2014, and the Ph.D. degree in control science and engineering from Beijing University of Chemical Technology, Beijing, China, in 2019.

She is currently an Assistant Professor with the School of Information and Electronics, Beijing Institute of Technology, Beijing, China. Her research interests include remote sensing image processing and pattern recognition.



Jinxin Zhang received the B.S. degree in mathematics and applied mathematics and the M.S. degree in mathematics from the College of Science from China Agricultural University (CAU), Beijing, China, in 2014 and 2016, respectively, and the Ph.D. degree in computer science and technology from the Department of College of Information and Electrical Engineering, CAU, in 2021.

He is currently a Post-Doctoral Researcher with the Beijing Institute of Technology, Beijing, China. His current research interests include dimensionality reduction, image classification, and semisupervised learning.



Wen Jia received the B.E. degree in geographic information system from the Foshan University, Foshan, China, in 2012, the M.A. degree in forest management from Southwest Forestry University, Kunming, China, in 2015, and the Ph.D. degree in forest management from Chinese Academy of Forestry, Beijing, China, in 2019.

He was a Visiting Student with the University of British Columbia, Vancouver, Canada, from 2016 to 2017. He is currently an Assistant Professor with the Department of Forest Application of Remote Sensing,

Institute of Forest Resource Information Techniques, Chinese Academy of Forestry. His main research focuses on the radiometric correction of hyperspectral remote sensing image, tree species mapping, and forest monitoring.



Wei Li (Senior Member, IEEE) received the B.E. degree in telecommunications engineering from Xi-dian University, Xi'an, China, in 2007, the M.S. degree in information science and technology from Sun Yat-Sen University, Guangzhou, China, in 2009, and the Ph.D. degree in electrical and computer engineering from Mississippi State University, Starkville, MS, USA, in 2012. Subsequently, he spent one year as a Postdoctoral Researcher with the University of California, Davis, CA, USA. He is currently a Professor with the School of Information and Electronics,

Beijing Institute of Technology, Beijing, China. His research interests include hyperspectral image analysis, pattern recognition, and data compression.

Dr. Li is an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS (SPL) and the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (JSTARS).