

Article

A Multispectral Feature Selection Method Based on a Dual-Attention Network for the Accurate Estimation of Fractional Vegetation Cover in Winter Wheat

Runzhi Yang^{1,2}, Shanshan Li^{1,*}, Bing Zhang^{1,3} , Qianjun Jiao¹ , Dailiang Peng^{1,4}, Songlin Yang^{1,2} and Ruyi Yu⁵

¹ Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China; yangrunzhi22@mails.ucas.ac.cn (R.Y.); zhangbing@aircas.ac.cn (B.Z.); jiaoqj@radi.ac.cn (Q.J.); pengdl@aircas.ac.cn (D.P.); yangsonglin211@mails.ucas.ac.cn (S.Y.)

² School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China

³ College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 101408, China

⁴ International Research Center of Big Data for Sustainable Development Goals, Beijing 100094, China

⁵ WEIFU Intelligent Sensor (WU XI) Technology Company Limited, Wuxi 214000, China; ruyi.yu@weifu.com.cn

* Correspondence: lishanshan@aircas.ac.cn; Tel.: +86-137-2003-4690

Abstract: Spectral information plays a crucial role in fractional vegetation cover (FVC) estimation, and selecting the appropriate spectral information is essential for improving the accuracy of FVC estimation. Traditionally, spectral feature selection is primarily guided by physical mechanisms or empirical statistical models. This has led to the use of multispectral and hyperspectral images, which often result in missing or redundant information, thereby decreasing the efficiency and accuracy of FVC estimation. This study proposes a novel dual-attention network to select the feature bands of Sentinel-2 multispectral images for the accurate FVC estimation of winter wheat. In the first step, the importance of hyperspectral band reflectances was determined using simulated data from the PROSAIL model, by combining the dual-attention mechanism with the convolutional neural network (DAM-CNN). In the second step, the importance of Sentinel-2 multispectral bands was converted from the hyperspectral band importance identified in the previous stage, and subsequently ranked accordingly. Based on the feature ranking results, multispectral simulated data translated from hyperspectral simulated data were used for CNN training, and multispectral feature selection was conducted based on FVC accuracy. Finally, the selected features were assessed based on their performance in FVC estimation using a CNN model with real data. The experimental results indicate that during the key growth period of winter wheat, the combination of red, green, and red-edge bands significantly influences the FVC estimation accuracy. Band 3 (Green), band 4 (Red), band 5 (Red-edge 1), and band 6 (Red-edge 2) of Sentinel-2 satellite images contribute most significantly to winter wheat FVC estimation, achieving an accuracy comparable to that obtained using all bands, while reducing the training time by 19.1%, as confirmed by field survey data.

Keywords: fractional vegetation cover; dual-attention mechanism; convolutional neural network; feature selection; feature assessment



Citation: Yang, R.; Li, S.; Zhang, B.; Jiao, Q.; Peng, D.; Yang, S.; Yu, R. A Multispectral Feature Selection Method Based on a Dual-Attention Network for the Accurate Estimation of Fractional Vegetation Cover in Winter Wheat. *Remote Sens.* **2024**, *16*, 4441. <https://doi.org/10.3390/rs16234441>

Academic Editor: Farid Melgani

Received: 25 September 2024

Revised: 17 November 2024

Accepted: 23 November 2024

Published: 27 November 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Fractional vegetation cover (FVC) refers to the proportion of the ground area covered by the vertical projection of vegetation organs, including leaves, stems, and branches [1,2]. As a critical parameter for measuring land surface vegetation distribution [3], FVC plays a vital role in the atmosphere, pedosphere, hydrosphere, and biosphere [4,5]. Wheat is one of the primary grain crops in China and serves as a staple food in the northern regions. It is crucial for monitoring wheat growth, facilitating scientific management, and making informed agricultural decisions by accurately and promptly obtaining FVC information [6].

The estimation of FVC is primarily accomplished through two methods: field measurement and remote sensing estimation [7]. Due to its extensive coverage and strong capability for continuous observation, remote sensing has become a highly effective tool for regional and global FVC estimation [8,9]. The primary approaches in remote sensing estimation include empirical models, pixel unmixing models, physical model-based methods [9], and deep learning methods. Empirical models are straightforward and computationally efficient, offering high FVC estimation accuracy in localized areas. However, their results are limited by the selection of spectral bands and are heavily dependent on ground samples [10–12]. Pixel-unmixing models, while relatively simple and capable of providing accurate estimates, face challenges in endmember identification and have not fully considered the contribution of spectral bands to FVC estimation [13–15]. Physical model-based methods have clear physical significance and strong theoretical foundations, enabling a more straightforward interpretation of the relationships between parameters. However, these methods are often complex and typically require indirect estimation through machine learning techniques or lookup table (LUT) methods [16,17]. Among these approaches, deep learning has gained increasing attention due to its superior performance and high accuracy in large-scale FVC estimation [18–23]. Convolutional Neural Networks (CNNs) are typical models in deep learning that extract features from input data through convolution operations, capturing information at different levels, layer by layer. The convolutional layers are used to extract features from the data, while pooling layers reduce the computational load and prevent overfitting, allowing the model to progressively acquire more complex and abstract feature representations. In FVC estimation, CNN methods are widely applied and have shown significant advantages. Huang et al. employed a one-dimensional convolutional neural network (1D-CNN) and long short-term memory (LSTM) to pre-train their model, and subsequently fine-tuned it using data from HJ-2A/B remote sensing images, achieving high accuracy in FVC estimation for mountain regions [20]. Similarly, Wang et al. utilized a deep learning method based on a fully convolutional neural network to extract vegetation information under shadow conditions from high-dynamic range images, resulting in more accurate FVC estimates [22]. Despite the significant advantages of deep learning methods in FVC estimation, their limited interpretability hinders the precise identification of crucial input features, thereby limiting the reliability of model applications.

The canopy reflectance of vegetation is a key factor in FVC estimation, and is primarily influenced by leaf pigment content and the leaf area index (LAI). Chlorophyll content impacts reflectance in the visible spectrum, while LAI influences reflectance across the visible, near-infrared, and shortwave infrared bands [24–26]. As a result, reflectance across these spectral ranges significantly influences the FVC of winter wheat. With the development of hyperspectral remote sensing technology, hyperspectral remote sensing enhances spectral information across the visible, near-infrared, and shortwave infrared ranges, enabling the quantitative analysis of subtle spectral features of vegetation, thereby playing a critical role in improving the accuracy of FVC estimation [27–31].

While the increase in spectral information generally provides more features that are relevant to FVC estimation, it can also lead to data redundancy, potentially compromising model accuracy, stability, and computational efficiency [17,32]. For instance, in a study on post-fire plant community FVC estimation, a model built using five spectral bands did not show significant improvement in accuracy with the inclusion of additional bands, but instead experienced reduced stability and computational efficiency [33]. Thus, a critical challenge remains to improve FVC estimation accuracy while minimizing data redundancy.

Current spectral feature selection methods for FVC estimation primarily rely on physical mechanisms and expert knowledge [10,13,34]. For example, Wang et al. identified three high-score band reflectances by assessing the importance of reflectances in various spectral bands of a multispectral satellite [17], while Yu et al. employed a deep learning method for pre-training and subsequently fine-tuned the model using real data, focusing on combinations of 10 m resolution bands from a multispectral satellite [19]. However, most of these studies are based on multispectral data, neglecting the extensive information

available in hyperspectral data. Therefore, there is a need to identify an effective feature extraction method that can fully leverage the richer spectral information available in hyperspectral data.

In human vision, attention is naturally directed toward the most salient information when observing objects, a phenomenon known as the visual attention mechanism. In deep learning, the attention mechanism emulates this process by automatically focusing on key aspects of the input data while downplaying less critical information. This mechanism is extensively applied in various domains, such as image processing, natural language processing, and computer vision [35–38]. As a feature extraction method, the attention mechanism can extract both local and global features [39], and when employed as an independent module, it can significantly improve model accuracy [40,41]. These strengths provide a robust foundation for leveraging the attention mechanism in hyperspectral feature extraction for vegetation parameter estimation.

Self-attention is a specialized form of the attention mechanism that enables the model to learn the intrinsic relationships within the input data [42], and can be used as a plug-and-play feature extraction module within networks. The Squeeze-and-Excitation (SE) module [43] is also an attention mechanism that is particularly suited to enhance the performance of existing convolutional networks, especially in terms of lightweight and efficient design. For this study, self-attention can capture complex correlations between bands, enabling the network to more effectively identify and leverage the relationships between different bands.

This paper presents a novel method that combines a dual-attention mechanism with a CNN (DAM-CNN). The dual-attention mechanism is employed to determine the importance of hyperspectral band reflectances, while the CNN utilizes these band reflectances, weighted by their importance, for FVC estimation. The DAM-CNN framework identifies the importance of hyperspectral bands that yield the highest FVC estimation accuracy. This study began with hyperspectral bands derived from a physical model, aiming to extract spectral information that is relevant to FVC estimation, and combined deep learning with the dual-attention mechanism to minimize redundant data and improve efficiency without compromising accuracy. Lastly, the proposed method would be assessed using field survey data. The specific process of the method is illustrated in Figure 1.

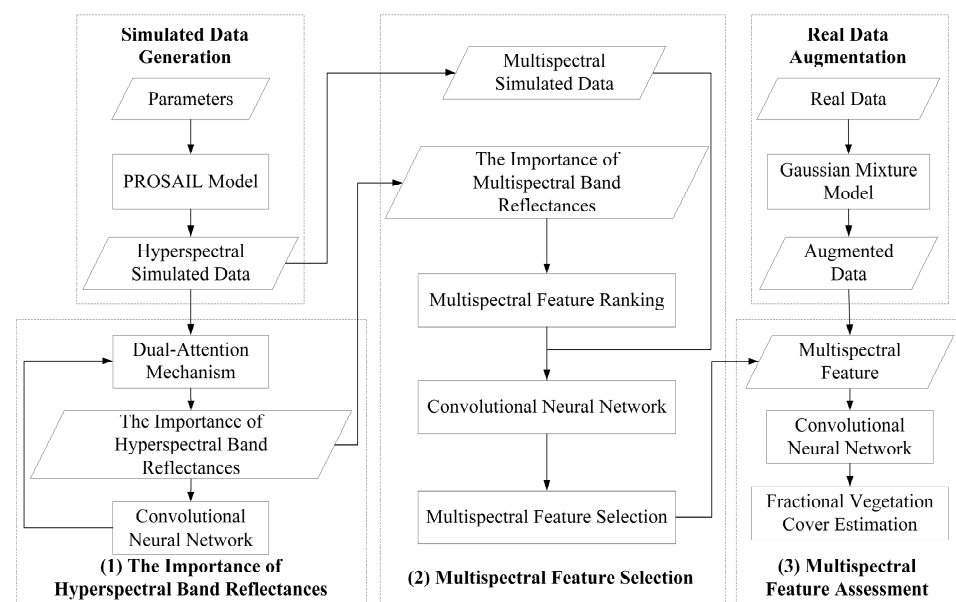


Figure 1. Flow chart of the proposed method.

(1) A large amount of hyperspectral simulated data was generated using the PROSAIL model and subsequently reduced to an appropriate sample size through Latin hypercube

sampling (LHS). The importance of hyperspectral bands was determined through the DAM-CNN;

(2) The importance of Sentinel-2 multispectral bands was converted from the importance of hyperspectral bands and subsequently ranked. The multispectral simulated data, translated from hyperspectral simulated data, were used for CNN training, with feature selection performed based on the feature ranking results;

(3) The limited real data was augmented using the Gaussian mixture model (GMM), creating an augmented dataset. Based on the selected multispectral features, FVC estimation was performed using this augmented dataset, and the final results were then assessed and validated by field survey FVC data.

2. Materials and Methods

2.1. Study Area and Real Data

This study focused on three areas (Figure 2). Study area (a) is located in the Xiaotangshan National Precision Agriculture Research Demonstration Base in Changping District, Beijing, and is referred to as Xiaotangshan. Study area (b) covers Dachang Hui Autonomous County and Sanhe City in Hebei Province, referred to as Sanhe. Study area (c) includes Xiong County, Rongcheng County, and Anxin County, all of which are part of the Xiong'an New Area in Hebei Province, referred to as Xiong'an in this paper. All three study areas are characterized by flat terrain and a warm temperate continental monsoon climate. The average annual precipitation ranges from 500 to 600 mm, with most rainfall occurring during the summer months. The primary crop in these areas is winter wheat.

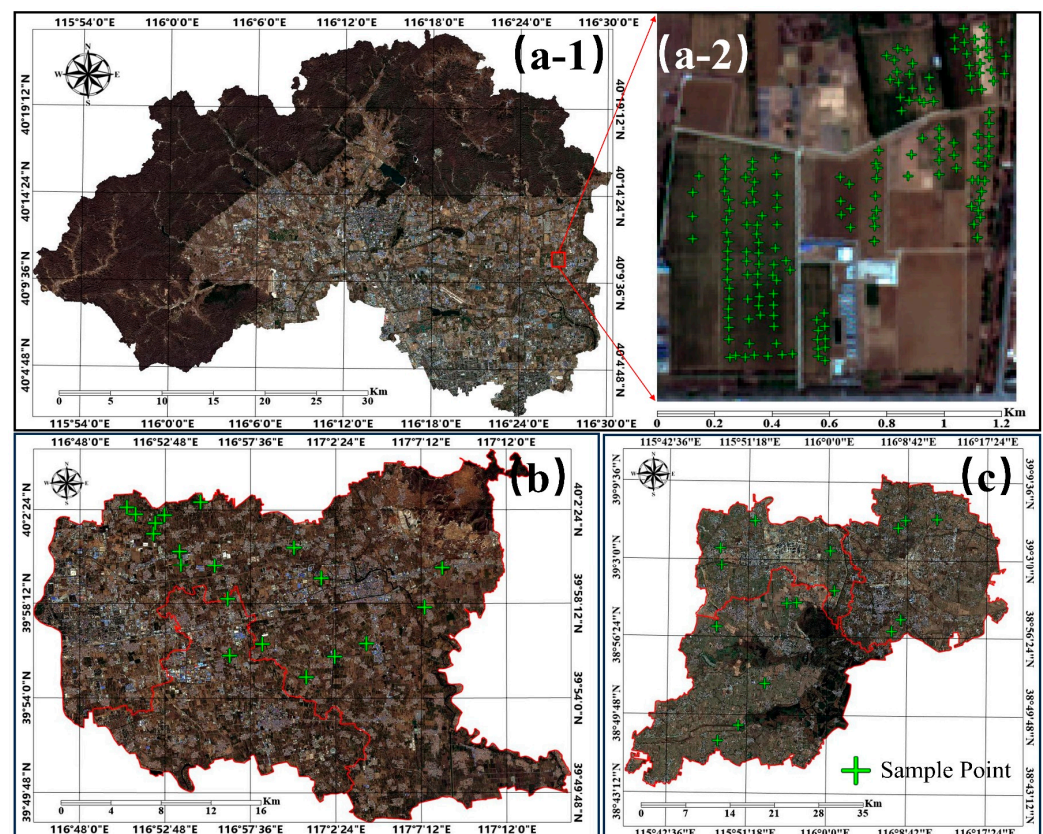


Figure 2. The research was conducted across three areas, including the Xiaotangshan National Precision Agriculture Research Demonstration Base (a-2) in the Changping District, Beijing (a-1), Dachang Hui Autonomous County and Sanhe City (b), and Xiong County, Rongcheng County, and Anxin County (c). The green plus signs represent the locations of the sample points in the study areas.

Ground truth images were captured using digital photography [44]. A digital camera was used to capture vertical images from several meters above the ground, thereby obtaining field-measured images. The FVC value was then calculated by determining the proportion of vegetation pixels to the total number of pixels in the image. The FVC extraction process involved converting the RGB image into the CIE $L^*a^*b^*$ ($L^*a^*b^*$) color space [45]. In this space, the channel a^* in $L^*a^*b^*$ specifically indicates the red–green content, effectively distinguishing green (vegetation) pixels from non-green (background) pixels. By applying an appropriate threshold to channel a^* , vegetation pixels were isolated, producing a binarized image in which vegetation and non-vegetation pixels were clearly separated.

Based on ground surveys and satellite imagery, ground sample plots were selected, with each plot containing five measurement points: one at the center and four at the corners. The field-measured FVC value for each plot was determined by calculating the average FVC value from the five measurement points [16]. This approach is known as the five-point sampling method, as illustrated in Figure 3.

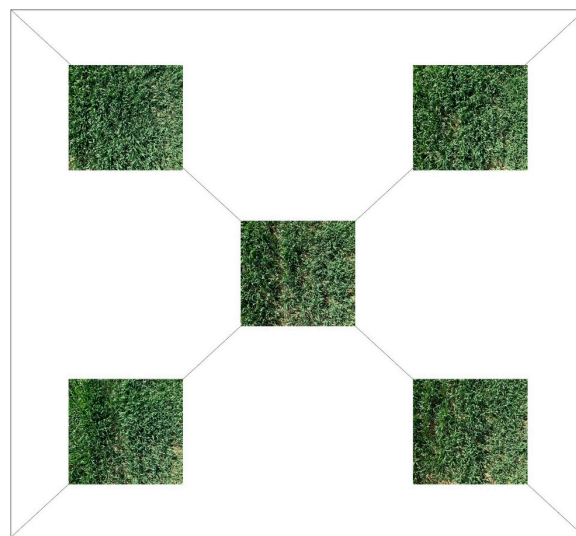


Figure 3. Schematic diagram of the five-point sampling method.

Field experiments were conducted in three study areas: Xiaotangshan, Sanhe, and Xiong'an. In all three areas, winter wheat is the dominant vegetation type. The experimental period extended from 2020 to 2023, covering the key growth period of winter wheat, including tillering, green-up, jointing, stem elongation, and heading. Throughout the experimental period, a total of 823 sample points were collected. The detailed data, including the sampling dates and the corresponding number of ground samples, are provided in Table 1.

Table 1. Sampling dates and sample sizes of ground samples in different study areas.

Study Area	Vegetation Type	Sampling Date	Sampling Size
Xiong'an	Winter Wheat	7 November 2020	121
		8 November 2020	
		1 April 2021	89
		2 April 2021	12
		14 April 2021	35
Xiaotangshan	Winter Wheat	28 April 2021	44
		4 March 2022	62
		8 April 2022	278
		29 April 2022	73
		21 March 2023	11
Sanhe	Winter Wheat	7 April 2023	32
		8 April 2023	66

The Sentinel-2 multispectral imaging satellite is equipped with a multispectral imager that provides data at three distinct spatial resolutions: 10 m, 20 m, and 60 m, covering 13 spectral bands from visible light to near-infrared and shortwave infrared (Table 2). Notably, in the red-edge region, Sentinel-2 includes three bands that are highly effective for monitoring vegetation.

Table 2. Overview of Sentinel-2 spectral bands [46].

Bands No.	Central Wavelength (nm)	Band Width (nm)	Spatial Resolution (m)
1	443	20	60
2	490	65	10
3	560	35	10
4	665	30	10
5	705	15	20
6	740	15	20
7	783	20	20
8	842	115	10
8a	865	20	20
9	945	20	60
10	1380	30	60
11	1610	90	20
12	2190	180	20

Through the Copernicus Data Space Ecosystem (<https://dataspace.copernicus.eu>, accessed on 10 April 2024), Bottom-of-Atmosphere (BOA) corrected reflectance data could be directly downloaded. To ensure the consistency and validity of both the ground samples and satellite data, it was essential to download satellite data corresponding to the time when the ground samples were collected.

Based on the latitude and longitude of the ground samples, the corresponding BOA reflectance data from the satellite were obtained and combined with the field-measured FVC to form the real data.

It is important to note that since band 10 is a cirrus band, the BOA reflectance data includes only 12 bands. In addition to band 10, both band 1 and band 9 are primarily used for atmospheric correction [17], and the blue band is susceptible to atmospheric influence, making its atmospheric correction challenging [17,47]. As a result, these bands were excluded from this study. Consequently, the nine bands primarily studied in this paper include band 3, band 4, band 5, band 6, band 7, band 8, band 8a, band 11, and band 12. To maintain consistent spatial resolution, band 5, band 6, band 7, band 8a, band 11, and band 12 were resampled to a 10 m spatial resolution.

2.2. The Importance of Hyperspectral Band Reflectances

2.2.1. Simulated Data Generation

The PROSAIL model generated simulated reflectance data in the 400–2500 nm spectral range, along with the corresponding FVC values. Given that the volume of the generated data could reach several million records, the dataset size was reduced using the LHS sampling approach to improve training efficiency.

1. Generation of Hyperspectral Simulated Data with the PROSAIL Model

The PROSAIL model generated hyperspectral simulated reflectance data, with the corresponding label data being FVC [16,17,19]. The input parameter ranges and step sizes for the model were determined based on expert knowledge [16,48–51] and field measurements.

The PROSAIL model is a combination of the PROSPECT model and the SAIL model. Its input parameters include structural parameter (N), chlorophyll content (C_{ab}), carotenoid content (C_{ar}), brown pigment content (C_{brown}), equivalent water thickness (C_w), dry matter content (C_m), anthocyanin content (C_{ant}), leaf area index (LAI), average leaf angle (ALA),

hotspot coefficient (Hot), solar zenith angle (SZ), viewing zenith angle (VZA), relative azimuth angle (RAZ), and soil reflectance. A detailed overview of the parameter settings for the PROSAIL model is provided in Table 3.

Table 3. Input parameters of the PROSAIL model.

Model	Parameter	Unit	Value Range	Step Size
PROSPECT	N	-	1–2	0.5
	C _{ab}	µg/cm ²	10–80	5
	C _{ar}	µg/cm ²	25% C _{ab}	-
	C _{brown}	-	0	-
	C _w	cm	0.005–0.015	0.005
	C _m	g/cm ²	0.003–0.006	0.001
	C _{ant}	µg/cm ²	2	-
SAIL	LAI	m ² /m ²	0.15–6.15	0.5
	ALA	°	30–55	5
	SZA	°	25–45	5
	Hot	-	0.1	-
	VZA	°	0	-
	RAZ	°	0	-

The soil reflectance data from China were obtained from the International Soil and Information Center (<http://www.isric.org>, accessed on 20 December 2023). Subsequently, similar soil reflectances were filtered using the spectral angle [52], and the average of these reflectances was calculated to reduce data redundancy. The method for calculating the spectral angle is presented below:

$$\alpha = \cos^{-1} \left(\frac{\sum_{i=1}^n x_i y_i}{\sqrt{\left(\sum_{i=1}^n x_i^2 \right) \left(\sum_{i=1}^n y_i^2 \right)}} \right), \quad (1)$$

where the spectral angle α represents the angular distance between two spectral vectors, $\mathbf{x} = (x_1, x_2, \dots, x_n)$ and $\mathbf{y} = (y_1, y_2, \dots, y_n)$. The range of α extends from 0 to $\pi/2$, with larger values indicating greater differences between the spectral vectors. In this study, soil reflectances with a spectral angle of less than 0.05 were considered sufficiently similar, and their average was calculated to generate a new representative soil reflectance. Consequently, 15 representative soil reflectances were selected for further analysis.

The labeled FVC data corresponding to the simulated reflectance were calculated using the gap fraction function of the ALA and LAI, as expressed in Equations (2) and (3) [17,53]:

$$P(\theta) = e^{-\lambda_0 (G(\theta, \theta_1) / \cos \theta) \times LAI}, \quad (2)$$

$$FVC = 1 - P(0), \quad (3)$$

where FVC is the fractional vegetation cover and LAI is the leaf area index. The term $P(\theta)$ represents the gap fraction function, where θ is the direction of the gap fraction. Since FVC is defined as being observed from the nadir direction, θ is typically set to 0. The function $G = (G(\theta, \theta_1))$ represents the orthogonal projection of unit leaf area in the direction θ , with θ_1 corresponding to the average leaf angle. The parameter λ_0 , which generally takes a value of 1, represents the dispersion or aggregation of leaves.

A total of 3,159,000 samples were generated using the PROSAIL model. However, training directly on this dataset would be computationally expensive and time-consuming due to its large volume. To address this issue, a sampling technique was applied to reduce the sample size and mitigate the computational burden.

2. Latin Hypercube Sampling

LHS is a refined Monte Carlo sampling technique introduced by McKay, Beckman, and Conover in 1979 [54]. Since then, LHS has been widely employed due to its ability to maintain sample comprehensiveness while reducing the sample size [55,56]. Compared to other sampling methods, LHS provides higher accuracy with a reduced sample size, offering significant advantages in terms of computational time and storage efficiency [57].

When sampling from one-dimensional FVC data, if 10 samples are required, the interval $(0, 1)$ is first divided into 10 equal sections. One data point is randomly selected from each section. These selected points are then mapped to the actual sample data through the inverse of the cumulative distribution function (CDF) of the FVC data. By applying this mapping, the FVC samples are obtained. This procedure is shown in Figure 4, where the blue line represents the CDF of FVC, and the red arrows indicate the mapping relationship for the 10 sampled FVC points.

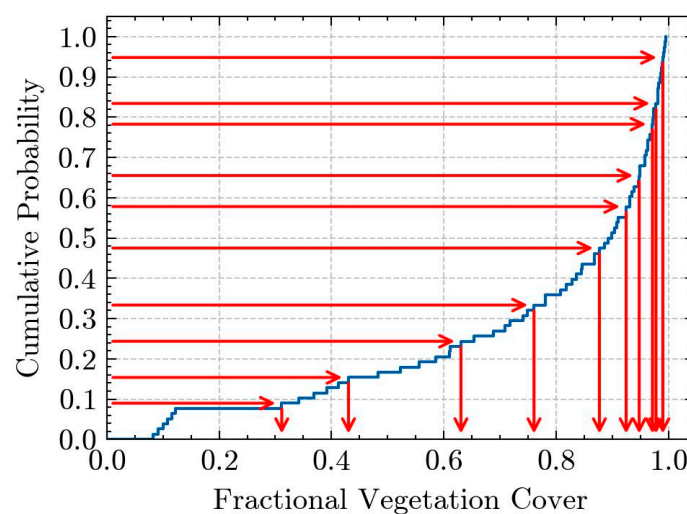


Figure 4. Mapping relationship of fractional vegetation cover (FVC) in Latin Hypercube Sampling (LHS).

2.2.2. DAM-CNN Model

This study proposes the DAM-CNN model by combining a CNN with a dual-attention mechanism. The dual-attention mechanism was designed to capture the relationships between spectral reflectances and FVC to determine the importance of hyperspectral band reflectances. The reflectances with varying importance levels served as the inputs to the CNN, while the outputs were the corresponding FVC label data. Through iterative training, the DAM-CNN model learned to identify the importance of hyperspectral bands that contributed to improving the accuracy of FVC estimation.

1. Dual-Attention Mechanism

The dual-attention mechanism proposed in this paper combines self-attention and output-attention mechanisms (Figure 5).

The self-attention mechanism, as a specialized feature extraction method, processes spectral information by considering the relationships between each spectrum in the sequence and all other spectra. For instance, in the transformation from x_1 to f_1 , the relationships between individual elements are considered. The spectral vector $\mathbf{x} = (x_1, x_2, \dots, x_n)$ is independently transformed through linear operations to produce three distinct vectors: the query vector $\mathbf{q} = (q_1, q_2, \dots, q_n)$, the key vector $\mathbf{k} = (k_1, k_2, \dots, k_n)$, and the value vector $\mathbf{v} = (v_1, v_2, \dots, v_n)$. The dot product of q_1 and $\mathbf{k} = (k_1, k_2, \dots, k_n)$ is computed, followed by a Softmax normalization, resulting in $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$, which represents the attention score of x_1 with respect to all other variables. $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ is then used to perform a weighted sum with $\mathbf{v} = (v_1, v_2, \dots, v_n)$, producing the output f_1 . The

calculation of f_1 takes into account the global context, as it incorporates the influence of all variables, while the attention mechanism ensures that more significant elements—those with higher attention scores—have a greater impact on the final output.

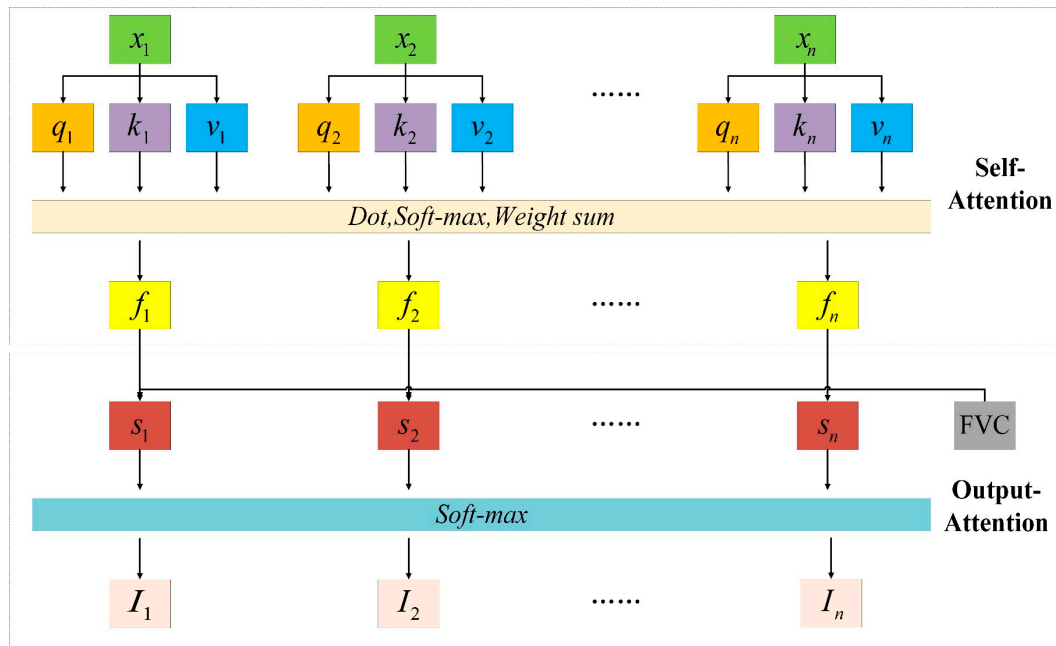


Figure 5. Schematic diagram of the dual-attention mechanism.

The output-attention mechanism focuses primarily on the relationship between the output and each input. Unlike traditional attention mechanisms, the output-attention mechanism explicitly incorporates the output as a key factor in determining the input importance. After the self-attention mechanism extracts spectral features, yielding $\mathbf{f} = (f_1, f_2, \dots, f_n)$, the output-attention mechanism calculates the attention score $\mathbf{s} = (s_1, s_2, \dots, s_n)$ by evaluating the relationship between the output FVC and the spectral features $\mathbf{f} = (f_1, f_2, \dots, f_n)$. $\mathbf{s} = (s_1, s_2, \dots, s_n)$ is then normalized through a Softmax operation to yield $\mathbf{I} = (I_1, I_2, \dots, I_n)$, which represents the importance of the input bands. The output-attention mechanism enables the model to emphasize spectral reflectances that have a greater influence on the FVC, allowing the importance of each band to more accurately reflect its relationship with the FVC.

2. Convolutional Neural Network

The CNN is a significant advancement over traditional neural networks and one of the most prominent algorithms in deep learning. The fundamental architecture of a CNN consists of several key components: an input layer, convolutional layers, pooling layers, fully connected layers, and an output layer.

The CNN utilized in this study consists of two convolutional layers and two pooling layers, which are employed to extract and refine more important spectral features from the input data. After the convolutional and pooling operations, the data are further processed through two fully connected layers. These fully connected layers retain critical feature information while simultaneously performing dimensionality reduction.

3. DAM-CNN

The combination of the dual-attention mechanism and CNN is illustrated in Figure 6. First, hyperspectral band reflectances and real FVC are used as inputs. After processing through the dual-attention mechanism, the importance of bands are obtained—essentially the weights of the hyperspectral band reflectances. These weighted reflectances are then input into the CNN to generate the estimated FVC. Finally, by calculating the Mean Squared Error (MSE) between the estimated and real FVC, the model continuously adjusts and

optimizes the network parameters using a backpropagation mechanism, allowing the dual-attention mechanism and CNN to focus more effectively on the information most relevant to FVC estimation.

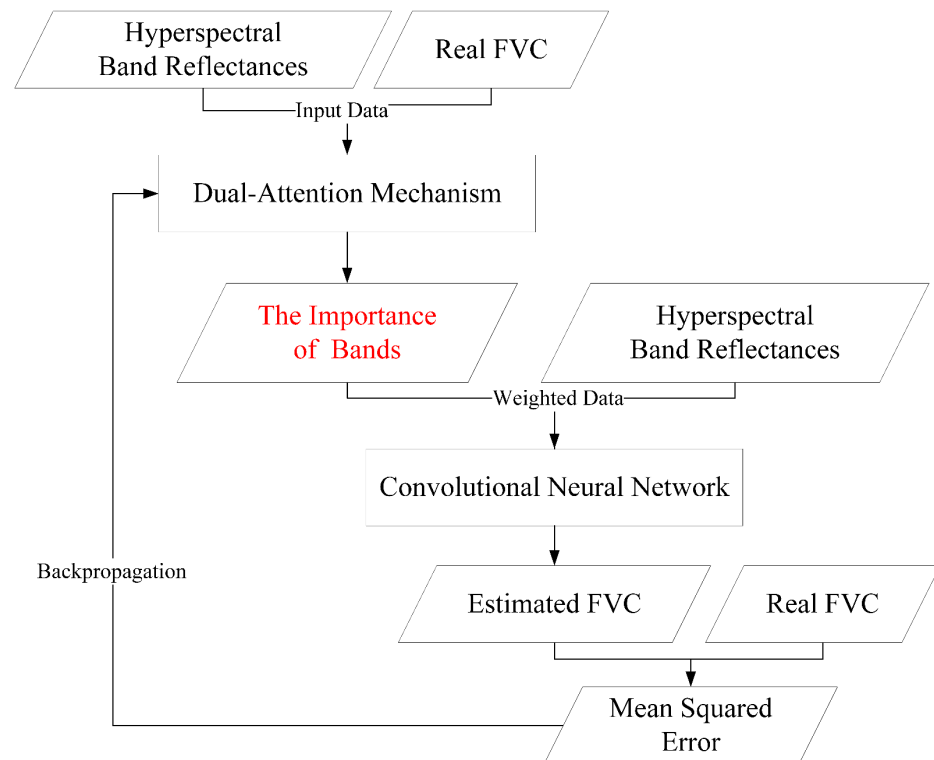


Figure 6. Schematic diagram of the DAM-CNN.

The final outputs of the entire network are the importance of bands. By analyzing the magnitude of these importance values, key band reflectances can be identified to use as model inputs. This approach improves model efficiency compared to using all band reflectances.

The training process of the DAM-CNN is divided into two stages.

Stage 1 involves the joint training of the dual-attention mechanism and the CNN. In the initial training stage, this study simultaneously optimizes the dual-attention mechanism and CNN to enable the two components to work in coordination, thereby providing the model with powerful feature extraction capabilities. Through joint training, the dual-attention mechanism and CNN are able to learn both global and local features of the data in the early stages. The goal of this phase is to enable the model's modules to work synergistically, gradually improving the model's initial performance.

Stage 2 involves freezing the CNN and training the dual-attention mechanism. In the second stage, the parameters of the CNN are frozen, and only the dual-attention mechanism is trained to further refine its performance on the existing feature representations. The purpose of this stage is to enable the dual-attention mechanism to capture the key features of the data more accurately. By optimizing the attention mechanism while keeping the CNN frozen, the model can further enhance its feature extraction ability based on the established feature representations.

The specific training parameters for the two stages are shown in Table 4.

Table 4. Specific training parameters for the two stages.

Training Stage	Learning Rate	Optimizer	Batch Size	Epoch	Loss Function
Stage 1	0.0001	Adam	128	2	MSE
Stage 2	0.0001	Adam	32	5	MSE

2.3. Multispectral Feature Selection

To obtain the feature band ranking of multispectral images, the importance of hyperspectral bands was converted into the multispectral band importance based on the Sentinel-2 satellite's spectral response function. The specific calculation for this conversion is provided in Equation (4) [58]:

$$W = \sum_{\lambda=1}^n W(\lambda) \times \beta(\lambda) / \sum_{\lambda=1}^n \beta(\lambda), \quad (4)$$

where W represents the multispectral band importance of Sentinel-2, and λ denotes the wavelength. The terms $W(\lambda)$ and $\beta(\lambda)$ refer to the hyperspectral band importance and Sentinel-2's spectral response function, respectively, at a given wavelength λ .

Similarly, the hyperspectral simulated data obtained through LHS sampling were translated into Sentinel-2 multispectral simulated data. The FVC estimation was subsequently conducted by using different numbers of multispectral simulated band reflectances as inputs to the CNN, aiming to determine the number of feature bands.

In addition to the multispectral features obtained through DAM-CNN, this study compared the multispectral feature selection methods used in references [17,19]. The random forest regression used in reference [17] is an ensemble learning method that can be used to evaluate feature importance. By calculating the contribution of each feature's split during model training, random forest quantifies the impact of each feature on the model's predictions. Reference [19] involves manually selecting feature combinations, primarily comparing the impact of Sentinel-2's 10 m resolution band combinations on the accuracy of FVC estimation.

2.4. Multispectral Feature Assessment

2.4.1. Real Data Augmentation

The band reflectances from the Sentinel-2 satellite, along with the corresponding field-measured FVC, constituted the real data. Compared to the hundreds of thousands or even millions of simulated data points, the relatively small amount of real data could result in underfitting and unstable parameter updates, leading to suboptimal performance in practical applications. To address this limitation, data augmentation techniques are commonly utilized. These methods enhance the diversity of the training dataset by generating new data, improving the model's robustness and its ability to generalize effectively.

The GMM is a classic method for generating new data, which is widely utilized in data augmentation due to its ability to capture the complexity of data distributions. In remote sensing research, existing data augmentation methods, such as the image-level label-based self-supervised method [59], image composition technique [60], and the conversion of simulated samples into real images [61], provide significant advantages in image data augmentation. However, these methods are primarily suited for image data rather than numerical data. As a statistical probability model, the GMM can accurately model the interrelationships between multiple band reflectance data, generating samples that are similar in distribution to the original data. This ensures that the augmented data maintain consistency with the statistical properties of the original data.

The process begins by selecting an optimal number of Gaussian components, typically using the Bayesian information criterion (BIC) or Akaike information criterion (AIC). Afterward, the GMM model generates new band reflectances by learning the characteristics of the real band reflectances. Finally, a random forest (RF) model is used to predict new FVC values for the new band reflectances by learning the relationship between the real band reflectances and the real FVC.

In this study, a total of 823 real samples were divided into 658 training samples and 165 testing samples, with an 8:2 ratio. The GMM was used to generate an augmented sample set of 65,800 samples from the 658 training samples for training the deep learning model, and 165 real samples were used for accuracy assessment.

2.4.2. Assessment of Selected Feature Bands

The augmented dataset and 165 real samples were employed to assess the results of feature selection. This assessment was divided into two components: (1) comparing the FVC estimation accuracy and efficiency between all bands and the proposed feature bands; and (2) comparing the FVC estimation accuracy and efficiency between feature bands selected by methods from other studies and those chosen by the method proposed in this study.

2.5. Evaluation Metrics

This study selected two evaluation metrics to assess the accuracy of FVC estimation: root mean square error (RMSE) and the coefficient of determination (R^2). RMSE quantifies the overall error between the model's predicted values and the actual observed values, while R^2 indicates the explanatory power of the independent variables for the dependent variable in the regression model. The formulas for RMSE and R^2 are shown in Equations (5) and (6):

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (5)$$

$$RMSE = \sqrt{\sum_{i=1}^n (\hat{y}_i - y_i)^2 / n}, \quad (6)$$

where n is the total number of samples, y_i is the real FVC value of the i -th sample, $\hat{y}_i$ is the estimated FVC value of the i -th sample, and $\bar{y}$ is the average of the real FVC values.

3. Results

The results are divided into three sections:

(1) The importance analysis of hyperspectral band reflectances, including the sample size for Latin Hypercube Sampling and the solution for the importance of hyperspectral band reflectances. The sample size for LHS was determined, and the DAM-CNN model utilized the sampled data as input to solve for hyperspectral band importance.

(2) Multispectral feature selection, including a solution for the importance of multispectral band reflectances, outcomes of the multispectral feature selection, and other multispectral feature selection methods. After converting the hyperspectral band importance into multispectral band importance, feature ranking was conducted. To determine the number of selected feature bands, different feature quantities were selected for training based on the feature ranking results, using multispectral simulated data translated from hyperspectral simulated data. Finally, experiments with other feature selection methods were conducted.

(3) Outcomes of the multispectral feature assessment. Comparative experiments on different feature bands were conducted based on the augmented data, including accuracy and performance comparison experiments.

3.1. The Importance Analysis of Hyperspectral Band Reflectances

3.1.1. Sample Size for Latin Hypercube Sampling

Li et al. studied the relationship between the similarity and error change between sampled data of the LHS and the original data under varying sample sizes, and found that the trends of both were generally consistent [62]. The study indicated that an optimal sample size exists, which not only ensures model accuracy but also enhances model efficiency. This optimal sample size can be identified by analyzing changes in similarity.

Since FVC and band reflectances exhibit a one-to-one correspondence, this study exclusively sampled FVC, with the corresponding band reflectances matched accordingly.

In this study, FVC samples of 3,000,000 (3M), 2,800,000 (2.8M), 2,600,000 (2.6M), 2,400,000 (2.4M), 2,200,000 (2.2M), 2,000,000 (2M), 1,800,000 (1.8M), 1,600,000 (1.6M), 1,400,000 (1.4M), 1,200,000 (1.2M), 1,000,000 (1M), 800,000 (0.8M), 600,000 (0.6M), 400,000 (0.4M), and 200,000 (0.2M) were taken (M is defined as one million). To ensure the reliability and stability of the

experiments, each FVC sample dataset was resampled 10 times. The similarity between the sampled FVC data and the original FVC data was calculated using the Euclidean distance, and the average of these calculations was used as the final measure of similarity.

As shown in Figure 7, the line chart illustrates the change in Euclidean distance as the FVC sample size decreases, while the bar chart shows the change in the rate of Euclidean distance with decreasing FVC sample sizes.

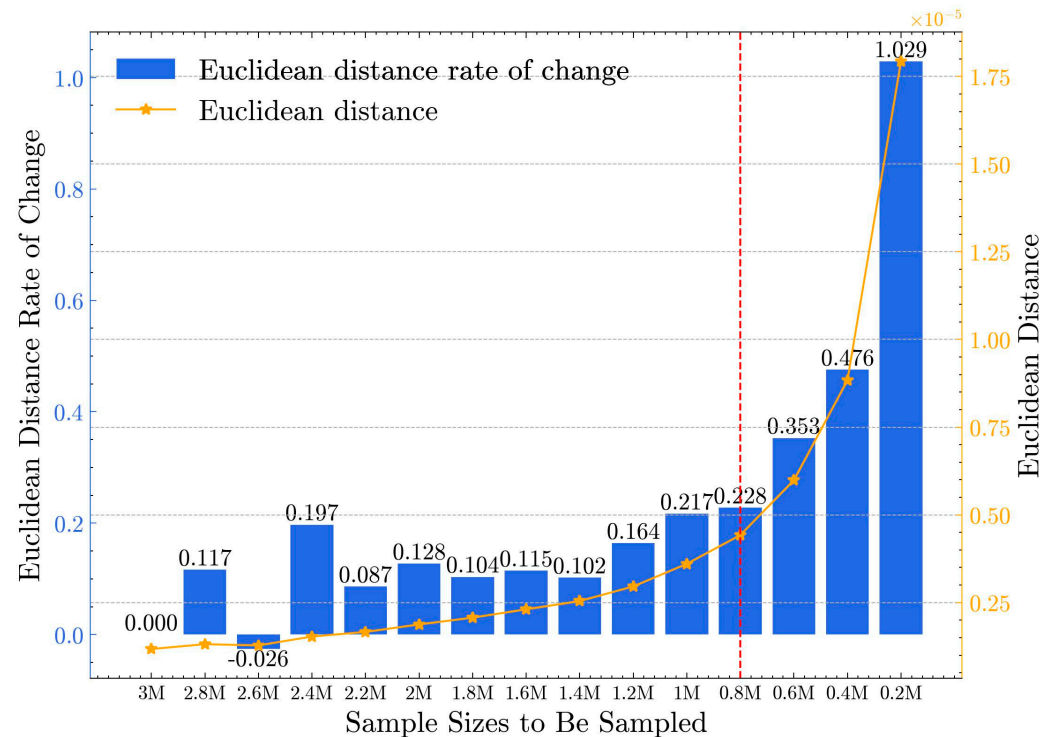


Figure 7. Euclidean distance and its rate of change across various FVC sample sizes.

The formula for calculating the rate of change of the Euclidean distance when the FVC sample size is 2,800,000 is shown in Equation (7):

$$r_{2.8M} = (d_{2.8M} - d_{3M}) / d_{3M}, \quad (7)$$

where $r_{2.8M}$ represents the rate of change of the Euclidean distance as the FVC sample size decreases from 3,000,000 to 2,800,000, $d_{2.8M}$ denotes the Euclidean distance at the FVC sample size of 2,800,000, and d_{3M} denotes the Euclidean distance at the FVC sample size of 3,000,000. The method for calculating the rate of change of the Euclidean distance remains consistent across various FVC sample sizes.

From the line and bar charts, it is evident that there is a substantial change in the Euclidean distance between the FVC sample sizes of 800,000 and 600,000. According to Li et al.'s research, the accuracy also experiences a significant change in this range [62].

Therefore, by using LHS, 800,000 FVC samples were obtained, and the corresponding band reflectances were matched accordingly. This process resulted in a training set comprising $800,000 \times 2102$, where the 800,000 rows represent 800,000 samples, and each row contains 2102 data points, corresponding to reflectances from 400 to 2500 nm and the associated FVC values.

3.1.2. Solution for the Importance of Hyperspectral Band Reflectances

Figure 8 shows the importance of hyperspectral bands within the 400–2500 nm wavelength range for FVC estimation. Based on the importance distribution, some bands show greater significance in FVC estimation than others, suggesting that these specific bands have a more substantial impact on the model's accuracy.

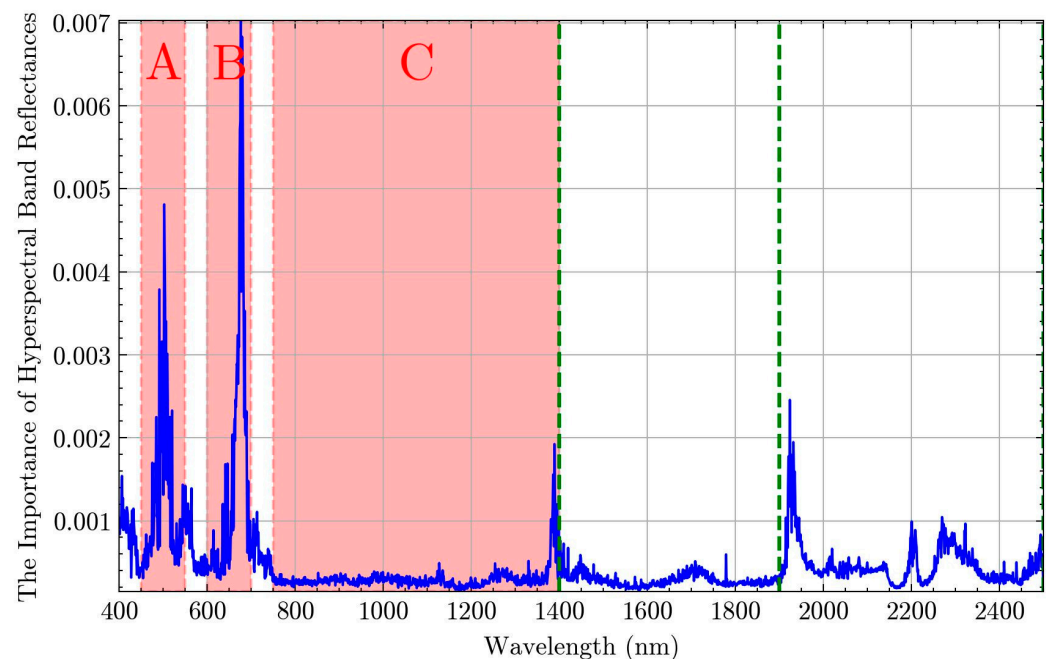


Figure 8. Distribution of the importance of hyperspectral band reflectances in FVC estimation covering the 400–2500 nm wavelength range.

Specifically, within the visible spectrum, Window B (600–700 nm) exhibits the highest importance, highlighting the critical role of red band reflectance in FVC estimation. Window A (450–550 nm) corresponds to the transition region where chlorophyll absorption and reflection occur, making the reflected light from this wavelength range rich in physiological information about vegetation.

In the near-infrared spectrum (Window C: 750 nm–1400 nm), vegetation generally exhibits high reflectance. However, the vegetation reflectance curve in this range typically forms a plateau, indicating minimal changes in reflectance and a flat trend. This phenomenon complicates the distinction of features in the near-infrared spectrum, making it difficult to extract clear feature differences during data processing.

In the shortwave infrared spectrum region (1400 nm–2500 nm), due to water absorption, the reflectances exhibit significant valleys near 1400 nm, 1900 nm, and 2500 nm (green dashed line), and thus these bands are commonly used for monitoring vegetation water content. In this study on FVC estimation, the importance of the bands at 1400 nm, 1900 nm, and 2500 nm is higher, indicating their critical role in the estimation process. This further confirms the sensitivity of these bands in monitoring vegetation water content.

3.2. Multispectral Feature Selection

3.2.1. Solution for the Importance of Multispectral Band Reflectances

Based on the spectral response function of the Sentinel-2 satellite, the importance of hyperspectral bands within the 400–2500 nm range was converted to the importance of multispectral bands. The distribution of multispectral band importance values across the nine bands of Sentinel-2 is shown in Figure 9.

The analysis results clearly indicate that the reflectance of band 4 (Red) in Sentinel-2 multispectral data is the most important, with its importance value being significantly higher than that of the other bands. This highlights the irreplaceable role of red band reflectance in the FVC estimation process. The reflectance of the red band is closely linked to chlorophyll content in plants, and since chlorophyll is a critical factor influencing FVC, variations in the red band reflectance directly impact the accuracy of FVC estimation. The prominent importance of band 4 reflectance indicates that variations in the red band reflectance are often a key factor in FVC estimation.

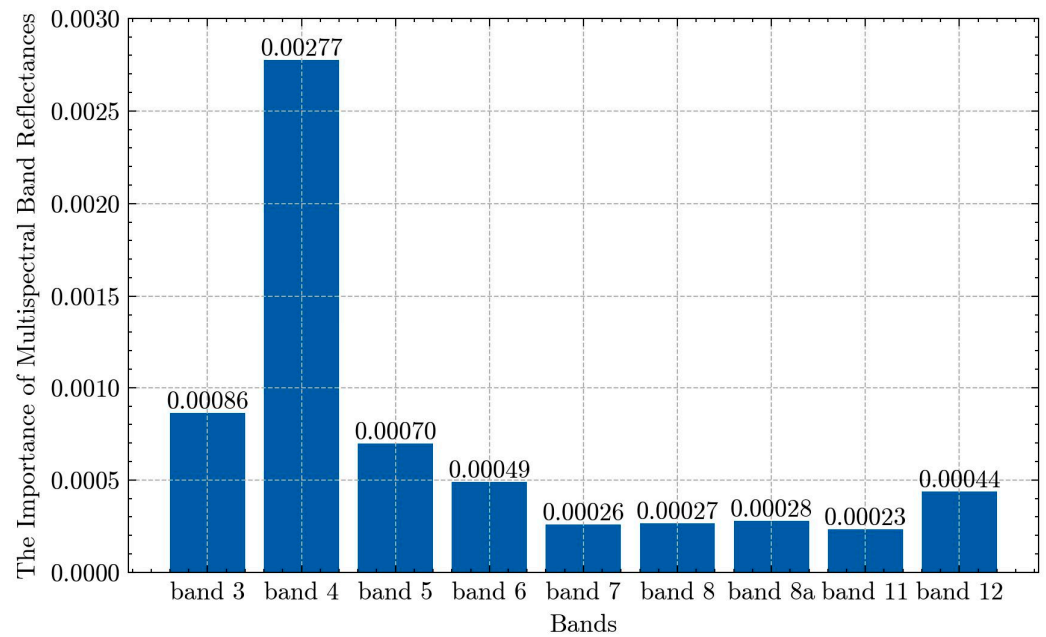


Figure 9. The importance values of multispectral band reflectances in Sentinel-2 bands (band 3, band 4, band 5, band 6, band 7, band 8, band 8a, band 11, and band 12).

The reflectance of band 3 (Green) also exhibits relatively high importance. This band typically represents the color of vegetation perceived by the human eye and is indicative of vegetation health. Thus, band 3 reflectance holds particular significance in FVC analysis.

Additionally, the red-edge band reflectances (band 5 and band 6) are distinctive spectral features of green vegetation that are closely associated with chlorophyll content, and they play a crucial role in FVC estimation. The variations in the red-edge band reflectances can signal alterations in plant health and physiological activity, providing a unique advantage in the FVC estimation process.

3.2.2. Outcomes of the Multispectral Feature Selection

The nine multispectral bands were ranked based on their importance: band 4, band 3, band 5, band 6, band 12, band 8a, band 8, band 7, and band 11. The results of the feature ranking were known, and the next experiment was to determine how many features to select.

The experiment began with the most important band reflectance (band 4) and progressively incorporated the reflectances of other bands until all nine band reflectances were included in the model. The training process for each combination was repeated 10 times, and the average values of RMSE and R^2 were calculated. The results are shown in Figure 10.

As the CNN model's input increases from one band reflectance (band 4) to four band reflectances (band 4, band 3, band 5, and band 6), the RMSE significantly decreases from 0.1035 to 0.0200, while the R^2 value rises from 0.8476 to 0.9943. This represents an 80.7% reduction in the RMSE and a 17.3% increase in R^2 , indicating that the addition of band 3, band 5, and band 6 substantially improves the model accuracy. However, when the input increases from four band reflectances (band 4, band 3, band 5, and band 6) to nine band reflectances (band 4, band 3, band 5, band 6, band 12, band 8a, band 8, band 7, and band 11), the RMSE only further decreases from 0.0200 to 0.0181, and the R^2 slightly increases from 0.9943 to 0.9953. This translates into a 9.5% reduction in the RMSE and only a 0.1% increase in R^2 , indicating that the addition of band 12, band 8a, band 8, band 7, and band 11 does not significantly improve the model's accuracy. These findings suggest that while adding more important bands significantly boosts the accuracy, further increasing the number of bands beyond a certain point leads to accuracy stabilization, with no substantial improvements.

Therefore, this study selected four feature bands (band 3, band 4, band 5, and band 6) for FVC estimation.

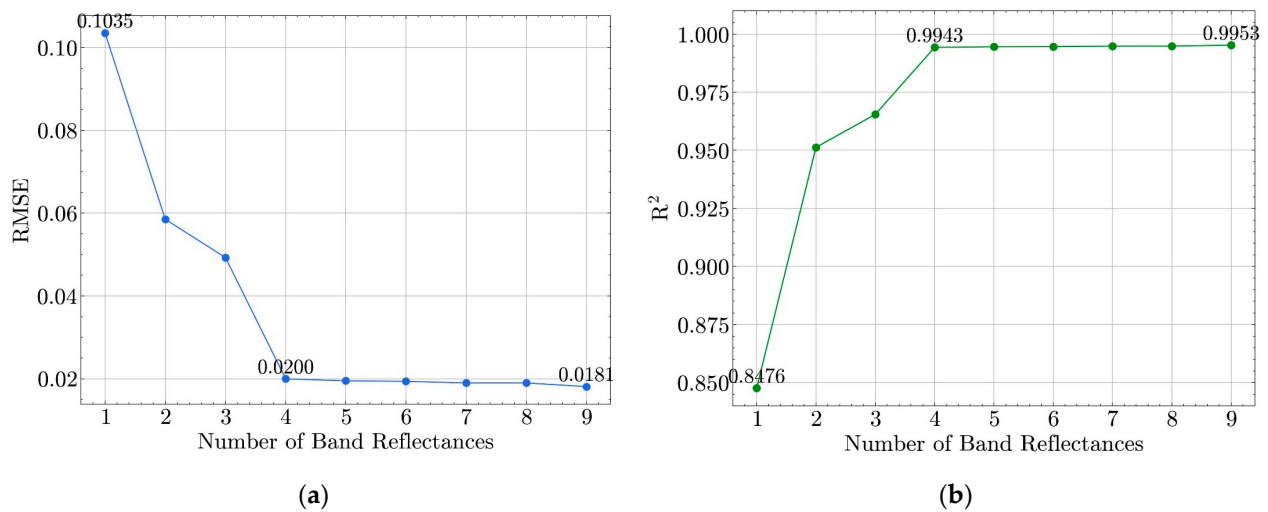


Figure 10. RMSE (a) and R^2 (b) average for CNN model using reflectances of different bands as inputs.

To further validate the importance of band 3, band 4, band 5, and band 6, additional experiments were conducted. In these experiments, four bands were randomly selected for training, and verification was conducted using simulated test data. The RMSE results from 100 experiments are shown in Figure 11.

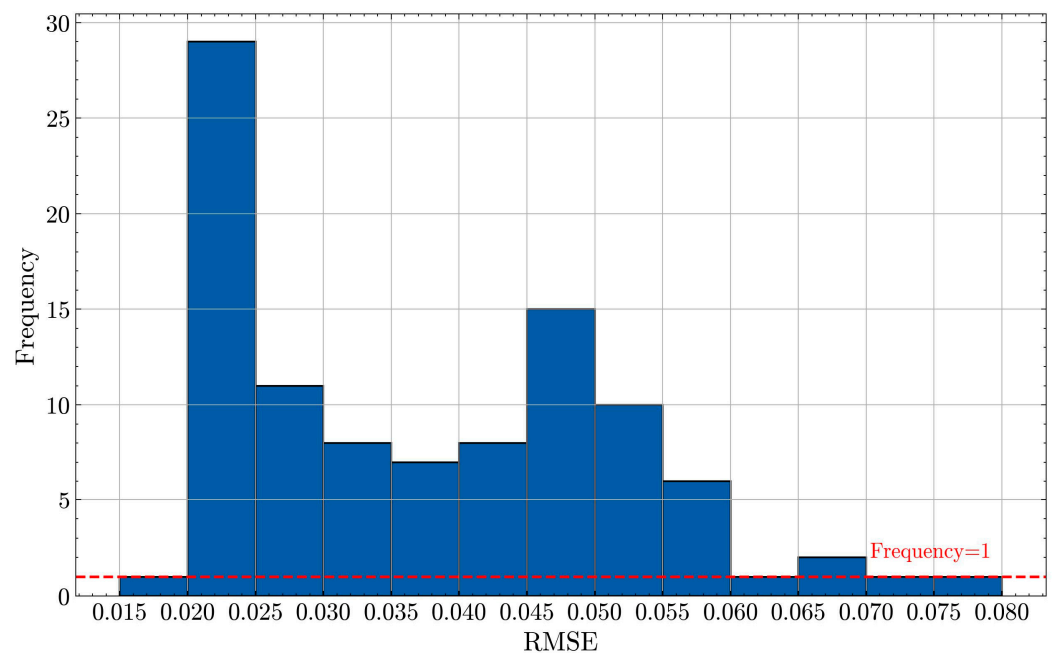


Figure 11. RMSE distribution in 100 experiments with 4 randomly selected bands for training.

It can be observed that the vast majority of RMSE values are above 0.020, while the accuracy for band 3, band 4, band 5, and band 6 is 0.02. The red line marks the position where the frequency is 1, indicating that out of the 100 randomly selected combinations of 4 bands, only one combination yields an RMSE below 0.020, and that combination is band 3, band 4, band 5, and band 6. This result further confirms the importance of these four bands in FVC estimation.

3.2.3. Other Multispectral Feature Selection Methods

1. Random Forest Regression

The band importance scores from the random forest regression are shown in Figure 12a. Based on the magnitude of the importance scores, the bands are ranked as follows: band 4, band 8a, band 6, band 7, band 5, band 12, band 8, band 11, and band 3.

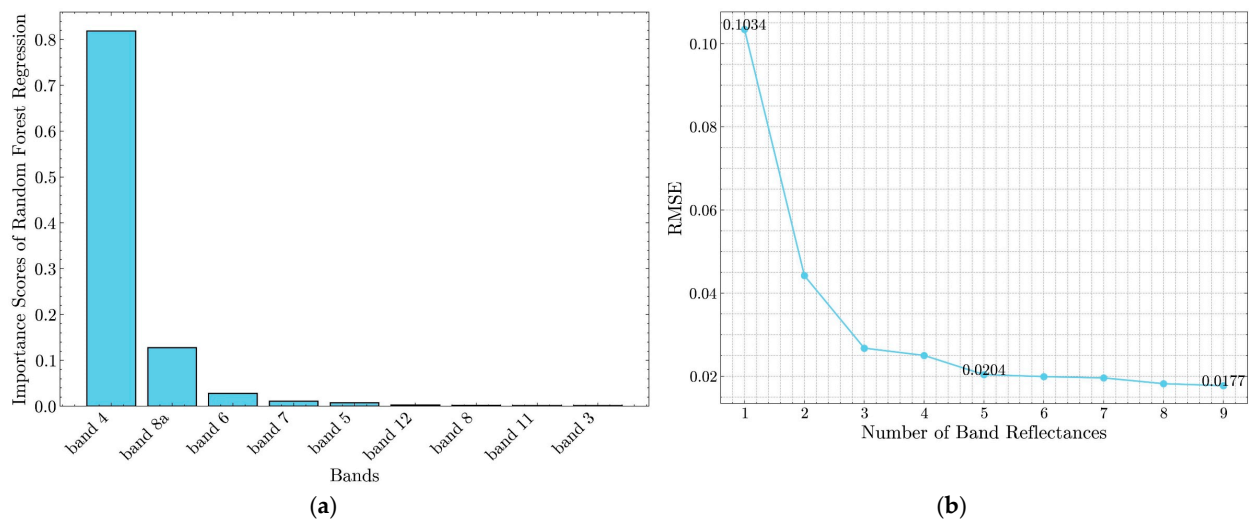


Figure 12. Band importance scores of random forest regression (a) and RMSE for different numbers of bands (b).

The RMSE obtained with different numbers of bands as CNN inputs, where bands were added sequentially based on their importance ranking, is shown in Figure 12b. As the number of bands increased from one (band 4) to five (band 4, band 5, band 6, band 7, and band 8a), the RMSE decreased from 0.1034 to 0.0204, representing a total reduction of 80.3%. However, when the number of bands was further increased from five to nine, the RMSE decreased by only 0.0027. Therefore, the key features identified by random forest regression are band 4, band 5, band 6, band 7, and band 8a.

2. Manual Feature Combination Selection

Manually selected feature combinations compared the following two sets: band 4 and band 8; and band 3, band 4, and band 8. The RMSE obtained when each of these combinations was used as CNN input is shown in Figure 13.

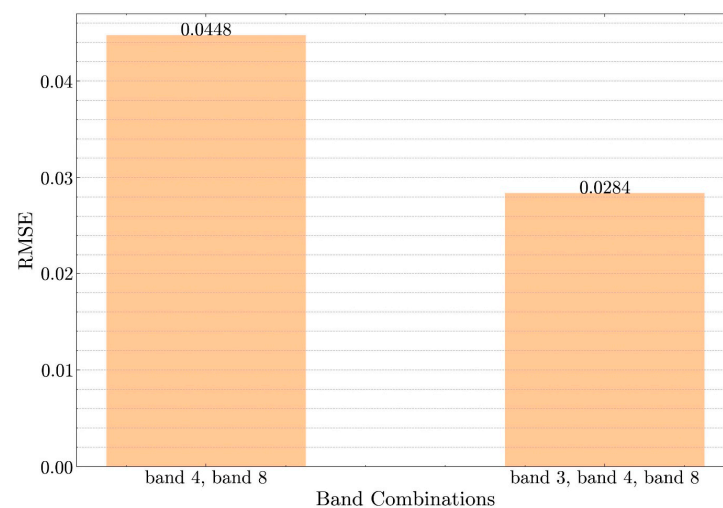


Figure 13. The RMSE obtained from the two different band combinations.

As shown in the figure, the combination of band 3, band 4, and band 8 achieved the highest accuracy. Therefore, the manually selected feature combination identifies band 3, band 4, and band 8 as the optimal features.

3.3. Outcomes of the Multispectral Feature Assessment

To illustrate the outcomes of the multispectral feature assessment, two comparative experiments were conducted: one focusing on accuracy comparison and the other on efficiency comparison. The features compared include the full spectral bands, the bands selected in this study (band 3, band 4, band 5, and band 6), the bands selected using the method from reference [17] (band 4, band 5, band 6, band 7, and band 8a), and the bands selected using the method from reference [19] (band 3, band 4, and band 8). The training data for the experiment consist of 65,800 augmented samples, and the testing is conducted on 165 real data samples.

3.3.1. FVC Estimation Accuracy Assessment

Figure 14a shows the test results of the model built with the nine band reflectances (band 3, band 4, band 5, band 6, band 7, band 8, band 8a, band 11, and band 12) along with FVC on real data. The results show that the RMSE is 0.090 and the R^2 value is 0.869, indicating high accuracy. This demonstrates that the GMM method is highly applicable to FVC estimation, and the generated data effectively meet the requirements of deep learning models, resulting in more stable model performance. However, there is still some discrepancy between these results and those obtained from the simulated data. In fact, simulated data represent spectral reflectance and FVC under ideal conditions, whereas real data are unavoidably influenced by environmental factors, leading to differences in accuracy.

Figure 14b shows the results of the selected band 3, band 4, band 5, and band 6 (RMSE = 0.091, R^2 = 0.865). When compared to the results of all nine bands (RMSE = 0.090, R^2 = 0.869), the RMSE and R^2 values are nearly identical, with the RMSE differing by only 0.001 and R^2 differing by only 0.004. This indicates that the selection of band 3, band 4, band 5, and band 6 not only significantly reduces the volume of data required for model training but also preserves a high level of accuracy.

Figure 14c,d show the FVC estimation results based on band 3, band 4, and band 8, and band 4, band 5, band 6, band 7, and band 8a, respectively.

Compared to the FVC estimation results based on band 3, band 4, and band 8 (RMSE = 0.096, R^2 = 0.853) and band 4, band 5, band 6, band 7, and band 8a (RMSE = 0.094, R^2 = 0.858), the results from the selected band 3, band 4, band 5, and band 6 (RMSE = 0.091, R^2 = 0.865) in this study show a slight improvement in accuracy. These findings suggest that, during the key growth period of winter wheat, the combination of red, green, and red-edge bands has a greater influence on FVC estimation than the combinations of red, green, and near-infrared bands or red, near-infrared, and red-edge bands.

In conclusion, when estimating FVC for winter wheat using Sentinel-2 data, selecting band 3, band 4, band 5, and band 6 for modeling can effectively reduce the data volume while maintaining a high level of accuracy.

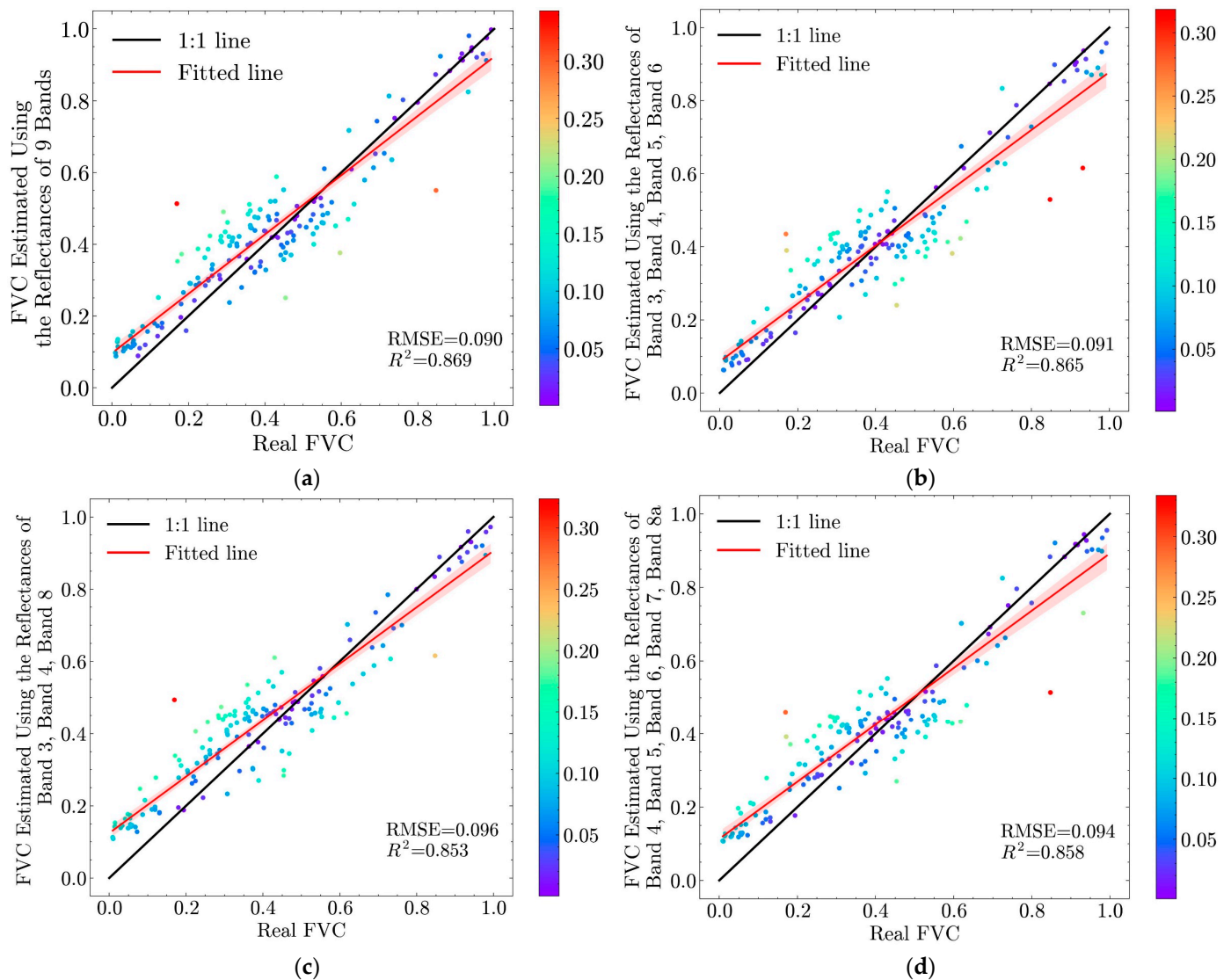


Figure 14. The FVC estimation results based on 9 bands (a), band 3, band 4, band 5, and band 6 (b), band 3, band 4, and band 8 (c), and band 4, band 5, band 6, band 7, and band 8a (d).

3.3.2. FVC Estimation Efficiency Assessment

The efficiency comparison experiment was divided into two parts: first, the performance improvement of the LHS during the calculation process of the importance of hyperspectral bands; second, the performance enhancement of band 3, band 4, band 5, and band 6 obtained after feature selection.

The experimental setup in this study employed the deep learning framework PyTorch for model construction, training, and testing, with Python as the programming language. The experimental environment was based on the Linux operating system and utilized dual Intel(R) Xeon(R) Silver 4316 processors, providing a total of 80 logical processors. The base frequency of the processors was 2.30 GHz, with a maximum turbo frequency of 3.40 GHz. The system was equipped with a total of 251 GiB of memory.

The training process of the DAM-CNN is divided into two stages: Stage 1 involves the joint training of the dual-attention mechanism and the CNN. Stage 2 involves freezing the CNN and training the dual-attention mechanism.

LHS reduced the number of samples from 3,159,000 to 800,000. As shown in Table 5, the training time for 800,000 samples was reduced by 74.2% in Stage 1 and 75.0% in Stage 2 compared to using 3,159,000 samples. Overall, it took 328,937.392 s to calculate the importance of hyperspectral bands using 3,159,000 samples and 82,877.583 s using

800,000 samples. Therefore, during the importance calculation process, LHS reduced the time by 74.8%.

Table 5. Time spent on different training processes with sample sizes of 3,159,000 and 800,000.

Sample Size	Time Spent on Stage 1 (s)	Time Spent on Stage 2 (s)
3,159,000	67,075.519	261,861.873
800,000	17,284.960	65,592.623

Under the same model and training parameters, the training and testing times were compared for models constructed using the following band combinations: band 3, band 4, band 5, band 6, band 7, band 8, band 8a, band 11, and band 12 of Sentinel-2; the bands selected in this study (band 3, band 4, band 5, and band 6); the bands selected using the method from reference [19] (band 3, band 4, and band 8); and the bands selected using the method from reference [17] (band 4, band 5, band 6, band 7, and band 8a).

As shown in Table 6, the testing time for all four experiments was under 1 s, so the comparison primarily focuses on training time. The training time for the four bands selected in this study was reduced by 19.1% compared to using the nine bands. Compared to the FVC estimation results using band 3, band 4, and band 8, the combination of the four bands selected in this study slightly improved the accuracy while increasing the training time by only 6.5%. Additionally, compared to band 4, band 5, band 6, band 7, and band 8a, the bands used in this study not only improved the accuracy but also enhanced the efficiency.

Table 6. Comparison of training and testing times in FVC estimation for different bands.

Input Bands (B Denotes Band)	Training Time (s)	Testing Time (s)
B3, B4, B5, B6, B7, B8, B8a, B11, B12	94.542	0.058
B3, B4, B5, B6	76.461	0.054
B3, B4, B8	71.789	0.053
B4, B5, B6, B7, B8a	76.473	0.055

The results indicate that the four selected bands significantly improve the experimental efficiency.

4. Discussion

This paper proposes a dual-attention mechanism network to select the feature bands from Sentinel-2 multispectral images for the FVC estimation of winter wheat. The method combines a dual-attention mechanism with a CNN, forming the DAM-CNN framework. The dual-attention mechanism considers the impact of FVC when determining the importance of hyperspectral bands, resulting in importance values that better reflect bands with a higher correlation with FVC. During the iterative process, the CNN updates the importance of the bands, ultimately leading to a higher accuracy in FVC estimation.

In existing studies on FVC estimation, many spectral feature selection methods have not fully considered the information in hyperspectral bands that could aid in FVC estimation, potentially leading to the omission of important spectral information or the introduction of redundant features. For example, Wang et al. only selected usable features from multispectral satellite data, without utilizing the rich spectral details provided by hyperspectral data [17]. In Yu et al.'s study, the impact of manually selected Sentinel-2 10 m spatial resolution band combinations on FVC estimation was compared, which could lead to the omission of important features or the introduction of redundant ones [19].

The DAM-CNN method is a feature selection approach based on hyperspectral information. It automatically identifies features that are highly relevant to FVC while effectively reducing the influence of redundant information by making full use of the rich spectral data. DAM-CNN uses simulated hyperspectral data to ensure that the feature selection process incorporates sufficient sources of spectral information. Additionally, through

multiple convolutional layers and the dual-attention mechanism, DAM-CNN effectively identifies features with low relevance to FVC, thereby reducing information redundancy and improving the model's robustness and reliability.

This method not only overcomes the issue of insufficient spectral coverage commonly found in traditional multispectral feature selection approaches but also enhances the model's efficiency in utilizing spectral data through effective feature filtering, thereby avoiding noise interference caused by redundant features. Compared to traditional methods, DAM-CNN demonstrates potential advantages in handling complex spectral information and FVC estimation, making it an effective tool for feature selection in FVC estimation.

The PROSAIL model used in this study demonstrates high generalizability and applicability. This model is a widely recognized radiative transfer model that has been successfully applied to various plant types and environmental conditions. Its solid physical foundation enables it to adapt well to different crops and regions. Additionally, the proposed DAM-CNN method effectively selects features through multiple convolutional layers and a dual-attention mechanism. This method is also applicable to other crops with rich spectral features, such as rice, maize, and others. Moreover, due to the high flexibility of the DAM-CNN method in processing different spectral information, it is suitable for cross-regional applications. Future research could experiment with other regions and crop types to further assess the generalization ability of DAM-CNN across regions and for different crop types.

The distribution of the importance of hyperspectral bands indicates that the blue-green range, red range, and the strong absorption range of water in the shortwave infrared region play key roles in FVC estimation. During the multispectral feature assessment stage, this paper primarily verifies the importance of band 3, band 4, band 5, and band 6 of Sentinel-2 satellite (green, red, and red-edge bands) for winter wheat FVC estimation. Hyperspectral satellites, such as EOS AM-1, HJ-1, and Gaofen-5, provide data with higher spectral resolution than traditional multispectral satellites. The next step in this study is to validate the hyperspectral features using hyperspectral data.

5. Conclusions

This paper proposes a new method that combines a dual-attention mechanism with a CNN to accurately estimate the FVC of winter wheat. By determining the importance of hyperspectral bands using simulated data and converting them into the importance of multispectral bands, feature ranking and selection can be performed, allowing multispectral feature assessment on the field survey dataset.

The multispectral feature selection results indicate that the reflectances of band 3, band 4, band 5, and band 6 of Sentinel-2 contribute the most to the key growth period of winter wheat FVC estimation. Further validation on real data compared these four band reflectances not only with the results of all nine band reflectances but also with the band reflectances selected based on methods from other studies. The findings demonstrate that the four selected band reflectances maintain high accuracy while improving training efficiency.

Author Contributions: Conceptualization, S.L., B.Z., Q.J. and D.P.; methodology, R.Y. (Runzhi Yang) and S.L.; software, R.Y. (Runzhi Yang); validation, R.Y. (Runzhi Yang); formal analysis, R.Y. (Runzhi Yang); investigation, R.Y. (Ruyi Yu), and S.Y.; resources, S.L.; writing—original draft preparation, R.Y. (Runzhi Yang); writing—review and editing, R.Y. (Runzhi Yang) and S.L.; visualization, R.Y. (Runzhi Yang). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China (No. 2021YFB3900503).

Data Availability Statement: The soil reflectance used in this study can be obtained from the International Soil and Information Center (<http://www.isric.org>, accessed on 20 December 2023). The satellite data can be obtained from the Copernicus Data Space Ecosystem (<https://dataspace.copernicus.eu>, accessed on 10 April 2024).

Acknowledgments: We would like to thank the editors and reviewers of the journal for their valuable comments and suggestions. We also appreciate Python for helping us build deep learning models and generate plots.

Conflicts of Interest: Author Ruyi Yu was employed by the company WEIFU Intelligent Sensor (WU XI) Technology Company Limited. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Ding, X.; Wang, Q.; Tong, X. Integrating 250 m MODIS Data in Spectral Unmixing for 500 m Fractional Vegetation Cover Estimation. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *111*, 102860. [\[CrossRef\]](#)
- Lambin, E.F.; Strahler, A.H. Indicators of Land-Cover Change for Change-Vector Analysis in Multitemporal Space at Coarse Spatial Scales. *Int. J. Remote Sens.* **1994**, *15*, 2099–2119. [\[CrossRef\]](#)
- Liu, D.; Yang, F.; Liu, S. Estimating Wheat Fractional Vegetation Cover Using a Density Peak K-Means Algorithm Based on Hyperspectral Image Data. *J. Integr. Agric.* **2021**, *20*, 2880–2891. [\[CrossRef\]](#)
- Wu, D.; Wu, H.; Zhao, X.; Zhou, T.; Tang, B.; Zhao, W.; Jia, K. Evaluation of Spatiotemporal Variations of Global Fractional Vegetation Cover Based on GIMMS NDVI Data from 1982 to 2011. *Remote Sens.* **2014**, *6*, 4217–4239. [\[CrossRef\]](#)
- Qin, W.; Zhu, Q.; Zhang, X.; Li, W.; Fang, B. Review of vegetation covering and its measuring and calculating method. *J. Northwest A&F Univ. (Nat. Sci. Ed.)* **2006**, *34*, 163–170. (In Chinese)
- Weiss, M.; Jacob, F.; Duveiller, G. Remote Sensing for Agricultural Applications: A Meta-Review. *Remote Sens. Environ.* **2020**, *236*, 111402. [\[CrossRef\]](#)
- Andreatta, D.; Gianelle, D.; Scotton, M.; Dalponte, M. Estimating Grassland Vegetation Cover with Remote Sensing: A Comparison Between Landsat-8, Sentinel-2, and PlanetScope Imagery. *Ecol. Indic.* **2022**, *141*, 109102. [\[CrossRef\]](#)
- Leprieur, C.; Kerr, Y.H.; Mastorchio, S.; Meunier, J.C. Monitoring vegetation cover across semi-arid regions: Comparison of remote observations from various scales. *Int. J. Remote Sens.* **2000**, *21*, 281–300. [\[CrossRef\]](#)
- Yan, K.; Gao, S.; Chi, H.; Qi, J.; Song, W.; Tong, Y.; Mu, X.; Yan, G. Evaluation of the Vegetation-Index-Based Dimidiate Pixel Model for Fractional Vegetation Cover Estimation. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–14. [\[CrossRef\]](#)
- Dymond, J.R.; Stephens, P.R.; Newsome, P.F.; Wilde, R.H. Percentage Vegetation Cover of a Degrading Rangeland from SPOT. *Int. J. Remote Sens.* **1992**, *13*, 1999–2007. [\[CrossRef\]](#)
- North, P.R.J. Estimation of fAPAR, LAI, and Vegetation Fractional Cover from ATSR-2 Imagery. *Remote Sens. Environ.* **2002**, *80*, 114–121. [\[CrossRef\]](#)
- Hill, M.J.; Zhou, Q.; Sun, Q.; Schaaf, C.B.; Palace, M. Relationships Between Vegetation Indices, Fractional Cover Retrievals, and the Structure and Composition of Brazilian Cerrado Natural Vegetation. *Int. J. Remote Sens.* **2017**, *38*, 874–905. [\[CrossRef\]](#)
- Jing, X.; Yao, W.; Wang, J.; Song, X. A Study on the Relationship Between Dynamic Change of Vegetation Coverage and Precipitation in Beijing's Mountainous Areas During the Last 20 Years. *Math. Comput. Model.* **2011**, *54*, 1079–1085. [\[CrossRef\]](#)
- Zhang, X.; Liao, C.; Li, J.; Sun, Q. Fractional Vegetation Cover Estimation in Arid and Semi-Arid Environments Using HJ-1 Satellite Hyperspectral Data. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *21*, 506–512. [\[CrossRef\]](#)
- Zhan, Y.; Meng, Q.; Wang, C.; Li, J.; Zhou, K.; Li, D. Fractional Vegetation Cover Estimation Over Large Regions Using GF-1 Satellite Data. In Proceedings of the Land Surface Remote Sensing II, Beijing, China, 8 November 2014; p. 92604B.
- Jia, K.; Liang, S.; Gu, X.; Baret, F.; Wei, X.; Wang, X.; Yao, Y.; Yang, L.; Li, Y. Fractional Vegetation Cover Estimation Algorithm for Chinese GF-1 Wide Field View Data. *Remote Sens. Environ.* **2016**, *177*, 184–191. [\[CrossRef\]](#)
- Wang, B.; Jia, K.; Liang, S.; Xie, X.; Wei, X.; Zhao, X.; Yao, Y.; Zhang, X. Assessment of Sentinel-2 MSI Spectral Band Reflectances for Estimating Fractional Vegetation Cover. *Remote Sens.* **2018**, *10*, 1927. [\[CrossRef\]](#)
- Li, W.; Buitenwerf, R.; Munk, M.; Bøcher, P.K.; Svenning, J.-C. Deep-Learning Based High-Resolution Mapping Shows Woody Vegetation Densification in Greater Maasai Mara Ecosystem. *Remote Sens. Environ.* **2020**, *247*, 111953. [\[CrossRef\]](#)
- Yu, R.; Li, S.; Zhang, B.; Zhang, H. A Deep Transfer Learning Method for Estimating Fractional Vegetation Cover of Sentinel-2 Multispectral Images. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [\[CrossRef\]](#)
- Huang, Y.; Zhou, X.; Lv, T.; Tao, Z.; Zhang, H.; Li, R.; Zhai, M.; Liang, H. The Retrieval of Forest and Grass Fractional Vegetation Coverage in Mountain Regions Based on Spatio-Temporal Transfer Learning. *Remote Sens.* **2023**, *15*, 4857. [\[CrossRef\]](#)
- Wang, Z.; Song, D.; He, T.; Lu, J.; Wang, C.; Zhong, D. Developing Spatial and Temporal Continuous Fractional Vegetation Cover Based on Landsat and Sentinel-2 Data with a Deep Learning Approach. *Remote Sens.* **2023**, *15*, 2948. [\[CrossRef\]](#)
- Wang, Z.; Chen, W.; Xing, J.; Zhang, X.; Tian, H.; Tang, H.; Bi, P.; Li, G.; Zhang, F. Extracting Vegetation Information from High Dynamic Range Images with Shadows: A Comparison Between Deep Learning and Threshold Methods. *Comput. Electron. Agric.* **2023**, *208*, 107805. [\[CrossRef\]](#)
- He, J.; Lyu, D.; He, L.; Zhang, Y.; Xu, X.; Yi, H.; Tian, Q.; Liu, B.; Zhang, X. Combining Object-Oriented and Deep Learning Methods to Estimate Photosynthetic and Non-Photosynthetic Vegetation Cover in the Desert from Unmanned Aerial Vehicle Images with Consideration of Shadows. *Remote Sens.* **2023**, *15*, 105. [\[CrossRef\]](#)

24. Broge, N.H.; Leblanc, E. Comparing Prediction Power and Stability of Broadband and Hyperspectral Vegetation Indices for Estimation of Green Leaf Area Index and Canopy Chlorophyll Density. *Remote Sens. Environ.* **2001**, *76*, 156–172. [\[CrossRef\]](#)
25. Sun, J.; Shi, S.; Yang, J.; Chen, B.; Gong, W.; Du, L.; Mao, F.; Song, S. Estimating Leaf Chlorophyll Status Using Hyperspectral Lidar Measurements by PROSPECT Model Inversion. *Remote Sens. Environ.* **2018**, *212*, 1–7. [\[CrossRef\]](#)
26. Yue, J.; Guo, W.; Yang, G.; Zhou, C.; Feng, H.; Qiao, H. Method for Accurate Multi-Growth-Stage Estimation of Fractional Vegetation Cover Using Unmanned Aerial Vehicle Remote Sensing. *Plant Methods* **2021**, *17*, 51. [\[CrossRef\]](#)
27. Gaffney, R.; Augustine, D.J.; Kearney, S.P.; Porensky, L.M. Using hyperspectral imagery to characterize rangeland vegetation composition at process-relevant scales. *Remote Sens.* **2021**, *13*, 4603. [\[CrossRef\]](#)
28. Rigge, M.; Bunde, B.; Postma, K.; Oliver, S.; Mueller, N. Application of Normalized Radar Backscatter and Hyperspectral Data to Augment Rangeland Vegetation Fractional Classification. *Remote Sens.* **2024**, *16*, 2315. [\[CrossRef\]](#)
29. Robichaud, P.R.; Lewis, S.A.; Laes, D.Y.M.; Hudak, A.T.; Kokaly, R.F.; Zamudio, J.A. Postfire soil burn severity mapping with hyperspectral image unmixing. *Remote Sens. Environ.* **2007**, *108*, 467–480. [\[CrossRef\]](#)
30. Yue, Y.; Wang, K.; Zhang, B.; Jiao, Q.; Liu, B.; Zhang, M.Y. Remote sensing of fractional cover of vegetation and exposed bedrock for karst rocky desertification assessment. *Procedia Environ. Sci.* **2012**, *13*, 847–853. [\[CrossRef\]](#)
31. Guerschman, J.P.; Hill, M.J.; Renzullo, L.J.; Barrett, D.J.; Marks, A.S.; Botha, E.J. Estimating fractional cover of photosynthetic vegetation, non-photosynthetic vegetation and bare soil in the Australian tropical savanna region upscaling the EO-1 Hyperion and MODIS sensors. *Remote Sens. Environ.* **2009**, *113*, 928–945. [\[CrossRef\]](#)
32. Mutanga, O.; Adam, E.; Cho, M.A. High-Density Biomass Estimation for Wetland Vegetation Using WorldView-2 Imagery and Random Forest Regression Algorithm. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *18*, 399–406. [\[CrossRef\]](#)
33. Fernández-Guisuraga, J.M.; Verrelst, J.; Calvo, L.; Suárez-Seoane, S. Hybrid Inversion of Radiative Transfer Models Based on High Spatial Resolution Satellite Reflectance Data Improves Fractional Vegetation Cover Retrieval in Heterogeneous Ecological Systems After Fire. *Remote Sens. Environ.* **2021**, *255*, 112304. [\[CrossRef\]](#)
34. Graetz, R.D.; Pech, R.P.; Davis, A.W. The Assessment and Monitoring of Sparsely Vegetated Rangelands Using Calibrated Landsat Data. *Int. J. Remote Sens.* **1988**, *9*, 1201–1222. [\[CrossRef\]](#)
35. Jiang, Y.; Li, Y.; Zou, S.; Zhang, H.; Bai, Y. Hyperspectral Image Classification with Spatial Consistence Using Fully Convolutional Spatial Propagation Network. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 10425–10437. [\[CrossRef\]](#)
36. Pan, H.; Zhao, X.; Ge, H.; Liu, M.; Shi, C. Hyperspectral Image Classification Based on Multiscale Hybrid Networks and Attention Mechanisms. *Remote Sens.* **2023**, *15*, 2720. [\[CrossRef\]](#)
37. Choi, H.; Cho, K.; Bengio, Y. Fine-grained attention mechanism for neural machine translation. *Neurocomputing* **2018**, *284*, 171–176. [\[CrossRef\]](#)
38. Wang, D.; Xiang, S.; Zhou, Y.; Mu, J.; Zhou, H.; Irampaye, R. Multiple-Attention Mechanism Network for Semantic Segmentation. *Sensors* **2022**, *22*, 4477. [\[CrossRef\]](#)
39. Liu, J.; Chen, L.; Shi, R.; Xu, J.; Liu, A. Distant Supervised Relation Extraction with Hierarchical Attention Mechanism. In Proceedings of the 7th International Conference on Communication and Information Processing, Beijing, China, 7 March 2022; pp. 44–50.
40. Liu, J.; Yang, J.; Liu, K.; Xu, L. Ocean Current Prediction Using the Weighted Pure Attention Mechanism. *J. Mar. Sci. Eng.* **2022**, *10*, 592. [\[CrossRef\]](#)
41. Chen, C.; Wu, X.; Chen, A. A Semantic Segmentation Algorithm Based on Improved Attention Mechanism. In Proceedings of the 2020 International Symposium on Autonomous Systems (ISAS), Guangzhou, China, 6–8 December 2020; pp. 244–248.
42. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention is all you need. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 6000–6010.
43. Hu, J.; Shen, L.; Sun, G. Squeeze-and-Excitation Networks. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.
44. Zhou, Q.; Robson, M. Automated Rangeland Vegetation Cover and Density Estimation Using Ground Digital Images and a Spectral-Contextual Classifier. *Int. J. Remote Sens.* **2001**, *22*, 3457–3470. [\[CrossRef\]](#)
45. Liu, Y.; Mu, X.; Wang, H.; Yan, G. A Novel Method for Extracting Green Fractional Vegetation Cover from Digital Images. *J. Veg. Sci.* **2012**, *23*, 406–418. [\[CrossRef\]](#)
46. Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; et al. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sens. Environ.* **2012**, *120*, 25–36. [\[CrossRef\]](#)
47. Jiang, Z.; Huete, A.R.; Didan, K.; Miura, T. Development of a Two-Band Enhanced Vegetation Index Without a Blue Band. *Remote Sens. Environ.* **2008**, *112*, 3833–3845. [\[CrossRef\]](#)
48. Yu, Q.; Yang, G.; Wang, C. Chlorophyll inversion of winter wheat based on ground hyperspectral data and PROSAIL model. *Sci. Surv. Mapp.* **2019**, *44*, 96–102+136. (In Chinese)
49. Sun, Q.; Jiao, Q.; Chen, X.; Xing, H.; Huang, W.; Zhang, B. Machine Learning Algorithms for the Retrieval of Canopy Chlorophyll Content and Leaf Area Index of Crops Using the PROSAIL-D Model with the Adjusted Average Leaf Angle. *Remote Sens.* **2023**, *15*, 2264. [\[CrossRef\]](#)

50. Qu, Y.; Zhang, Y.; Wang, J. A Dynamic Bayesian Network Data Fusion Algorithm for Estimating Leaf Area Index Using Time-Series Data from In Situ Measurement to Remote Sensing Observations. *Int. J. Remote Sens.* **2012**, *33*, 1106–1125. [[CrossRef](#)]
51. Li, Z.; Jin, X.; Yang, G.; Drummond, J.; Yang, H.; Clark, B.; Li, Z.; Zhao, C. Remote Sensing of Leaf and Canopy Nitrogen Status in Winter Wheat (*Triticum aestivum* L.) Based on N-PROSAIL Model. *Remote Sens.* **2018**, *10*, 1463. [[CrossRef](#)]
52. Kruse, F.A.; Lefkoff, A.; Boardman, J.W.; Heidebrecht, K.; Shapiro, A.; Barloon, P.; Goetz, A. The Spectral Image Processing System (SIPS)—Interactive Visualization and Analysis of Imaging Spectrometer Data. *Remote Sens. Environ.* **1993**, *44*, 145–163. [[CrossRef](#)]
53. Nilson, T. A Theoretical Analysis of the Frequency of Gaps in Plant Stands. *Agric. Meteorol.* **1971**, *8*, 25–38. [[CrossRef](#)]
54. McKay, M.D.; Beckman, R.J.; Conover, W.J. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics* **2000**, *42*, 55–61. [[CrossRef](#)]
55. Yin, J.; Lu, W.; Xin, X.; Zhang, L. Application of Monte Carlo sampling and Latin Hypercube sampling methods in pumping schedule design during establishing surrogate model. In Proceedings of the 2011 International Symposium on Water Resource and Environmental Protection, Xi'an, China, 20–22 May 2011; pp. 212–215.
56. Bang, B.j.; Kwon, H.j.; Kim, Y.H. Effective Sampling Method for Statistical Leakage Analysis: Hybrid Sampling Method. In Proceedings of the 2017 International SoC Design Conference (ISOCC), Seoul, Republic of Korea, 5–8 November 2017; pp. 302–303.
57. Iordanis, I.; Koukouvinos, C.; Silou, I. Classification accuracy improvement using conditioned Latin Hypercube Sampling in Supervised Machine Learning. In Proceedings of the 2022 12th International Conference on Dependable Systems, Services and Technologies (DESSERT), Athens, Greece, 9–11 December 2022; pp. 1–5.
58. Yang, L.; Jia, K.; Liang, S.; Wei, X.; Yao, Y.; Zhang, X. A Robust Algorithm for Estimating Surface Fractional Vegetation Cover from Landsat Data. *Remote Sens.* **2017**, *9*, 857. [[CrossRef](#)]
59. Zhao, M.; Hu, X.; Zhang, L.; Meng, Q.; Chen, Y.; Bruzzone, L. Beyond Pixel-Level Annotation: Exploring Self-Supervised Learning for Change Detection with Image-Level Supervision. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 1–16. [[CrossRef](#)]
60. Davydzienka, T.; Tahmasebi, P.; Carroll, M. Improving remote sensing classification: A deep-learning-assisted model. *Comput. Geosci.* **2022**, *164*, 105123. [[CrossRef](#)]
61. Xiao, Q.; Liu, B.; Li, Z.; Ni, W.; Yang, Z.; Li, L. Progressive Data Augmentation Method for Remote Sensing Ship Image Classification Based on Imaging Simulation System and Neural Style Transfer. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2021**, *14*, 9176–9186. [[CrossRef](#)]
62. Li, S.; Yang, R.; Chen, B.; Xu, S.; Chen, J.; Wu, Y.; Feng, X. Method, Device, and Equipment for Determining Sample Size for Vegetation Coverage Estimation. CN Patent 202310822103.5, 19 September 2023. (In Chinese).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.