Combining 3D Radiative Transfer Model and Convolutional Neural Network to Accurately Estimate Forest Canopy Cover From Very High-Resolution Satellite Images

Decai Jin, Jianbo Qi, Huaguo Huang, and Linyuan Li

***Abstract*—Forest canopy cover (FCC) plays an important role in many ecological, hydrological and forestry applications. For large-scale applications, FCC is usually estimated from remotely sensed data by inverting radiative transfer models (RTMs) or using data-driven regressions. In this article, we proposed a hybrid model, which combines a 3-D RTM and transfer learning-based convolutional neural network (T-CNN), to estimate FCC from very high-resolution satellite images (e.g., Chinese GaoFen-2, 1 m resolution with 4 bands). Unlike common hybrid models that are purely trained with simulation data, T-CNN combines simulation data-based pre-training and actual data-based transfer learning, which is a widely used technique in artificial intelligence for fine-tuning models. The performance of T-CNN was compared with a random forest (RF) model and two general CNN models, including CNN trained with actual dataset only (data-CNN) and CNN trained with RTM simulation data only (RTM-CNN). Results on the independent validation dataset (not used in training stage) showed that T-CNN had higher accuracy (RMSE = 0.121, R^2 = 0.83), compared with RF (RMSE = 0.26, R^2 = 0.61), Data-CNN (RMSE = 0.142, R^2 = 0.81), and RTM-CNN (RMSE = 0.144, R^2 = 0.73), which indicates that T-CNN has a strong transferability. Tests on different training sizes showed that T-CNN (0.084 < RMSE < 0.108) provided constantly better performances than RF (0.116 < RMSE < 0.122) and data-CNN (0.103 < RMSE < 0.128), which demonstrates the potential of T-CNN as an alternative to RTM-based inversion and data-driven regressions to estimate FCC, especially when training data is imbalanced and inadequate.**

***Index Terms*—Three-dimensional (3-D) radiative transfer model, convolutional neural network (CNN), forest canopy cover (FCC), transfer learning.**

I. Introduction

FORESTS, covering nearly 30–40% of land surface across the globe, play an essential role in both the carbon cycle and water cycle. Hence, the healthy and sustainable management of forests is causing concerns, especially with climate anomalies and the increasing greenhouse effect [1], [2]. Forest canopy cover (FCC), which refers to the percentage of the ground covered by a vertical projection of the outermost perimeter of tree crowns [3], is widely used to assess the quality of forest and describe forest changes. It is also a relevant factor in climate changes, energy transfer, transpiration and photosynthesis [4]–[6]. According to the definition of forest given by the food and agriculture organization of the United Nations [7], FCC is the key indicator to define forest, which refers to an area with a tree canopy cover of more than 10% and with an extent of more than 0.5 hectare. This indicates that the accurate estimation of FCC in low coverage areas, where the distinguishment between understory and overstory should be considered, is substantially important [8]. However, traditional field measurements tend to be too laborious or inaccurate for large areas. Remote sensing, which provides the ability to observe the canopy from the near nadir view, currently, is the only technology that allows one to monitor land surfaces globally.

To retrieve FCC from remotely sensed data, parametric regression methods that establish an empirical relationship between field survey data and spectral data such as vegetation indices are usually used [9]–[11]. Alternatively, machine learning methods, such as random forest (RF), support vector machine (SVM), and artificial neural network (ANN), tend to outperform the commonly used parametric equations because of their nonlinear fitting ability [12], [13]. However, these methods are dependent on spectral data quality which is affected by the training sample size and the pre-processing approaches such as data normalization. Besides, some studies also attempt to estimate FCC using radiative transfer models (RTMs), such as PROSAIL, SCOPE on the basis of physical laws [14]. POLDER, MERIS and CYCLOPES products also combined different RTMs with machine learning methods to estimate fractional vegetation cover on large scales [15]–[17]. However, RTMs always simplify the 3-D structures and radiative transfer processes to a certain degree, which may reduce the estimation accuracy. Also, based on a physical assumption that pixel spectrum is a mixture of spectra from different components, spectral mixture analysis and pixel dichotomy model can be used [5], [8], [18], [72].

Manuscript received June 22, 2021; revised September 17, 2021; accepted October 15, 2021. Date of publication October 26, 2021; date of current version November 10, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 41571332, in part by the National Science Foundation of China under Grant 42001279, and in part by the National Key Research and Development Program of China under Grant 2017YFC0504003-4. *(Corresponding author: Jianbo Qi.)*

The authors are with the State Forestry and Grassland Administration Key Laboratory of Forest Resources and Environmental Management, Beijing Forestry University, Beijing 100083, China (e-mail: redbooks_jdc@gmail.com; jianboqi@bjfu.edu.cn; huaguo_huang@bjfu.edu.cn; lilinyuan@bjfu.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2021.3122509

However, the drawback is that it is difficult to obtain pure vegetation and background pixels that are required in these models. Thus, these methods are very limited in the area of dense understory vegetation due to difficulty in distinguishing between understory vegetation spectra and canopy spectra.

In the past decades, airborne light detection and ranging (LiDAR) sensors have been increasingly used in forestry surveys to estimate forest parameters [19], [20]. And the accuracy of estimating FCC using airborne LiDAR data is quite high [21]. However, regionally or globally mapping FCC by using LiDAR is still a challenge due to the limitation of flight time and distance, as well as the high costs. Therefore, many studies often treat LiDAR as reference data to validate other methods or combine them with other remotely sensed data to estimate FCC in a large area [1], [22].

Convolutional neural network (CNN) that is a special machine learning algorithm is designed to learn spatial features by introducing receptive field through convolutional kernels. Some studies have demonstrated that CNN outperforms traditional machine learning algorithms (RF, ANN, SVM), especially for tasks, such as image classification and regression [23], [24]. Another advantage of CNN is the ability to extract features automatically, thus avoiding the feature engineering part [25], which reduces the loss of information that may be induced by manual feature selection in traditional machine learning algorithms. As a result, CNN has gained great success in many remote sensing applications, including land cover classification and inversion of forest parameters [19], [26], [27]. For instance, combining unmanned aerial vehicle (UAV) data and CNN, i.e., data-CNN, Kattenborn *et al.* [28] estimated the FCC of different vegetation types and gained pretty good estimation accuracy. However, a drawback is that CNN usually requires a large number of training datasets, which is not always fulfilled in tasks that field measurements are not easy to perform or be accessible.

Typically, the dataset covering different canopy types and growth conditions is almost impossible to be obtained completely [14]. Hence, data-driven regression methods, such as RF and SVM are vulnerable, which are less likely to be extended to a larger area [29], [30]. Instead, transfer learning that can fine-tune pre-trained models (fitted with a big dataset) to achieve high prediction accuracy on a new dataset by introducing a small number of new data, is becoming a popular research in both artificial intelligence and computer vision. Besides, it can also effectively solve the problem of data imbalance and inadequacy [31], [32]. However, there are fewer studies using transfer learning in remote sensing, which is mainly limited by the fact that big remote sensing datasets with labels are rarely published and shared [27].

As an alternative to manually labeling training datasets, RTMs can efficiently provide a large amount of labeled training data that covers different conditions. It has been a faithful method to train CNN, i.e., RTM-CNN. For instance, Annala *et al.* [33] used simulation spectral data from RTM SLOP to train 1-D CNN, which demonstrates the great potential of RTM-driven CNN. However, commonly used 1-D RTMs abstract canopies into homogeneous horizontal layers, which can not explicitly represent the canopy spatial heterogeneity. More precisely, the overstory, which is the major indicator of forest, cannot be distinguished from the understory vegetation from a single spectrum, which may lead to estimation errors. These limitations indicate that 3-D RTMs have the potential to replace 1-D RTMs in inversion models, since 3D RTMs, such as DART [34], RAPID [35], and LESS [36], can simulate canopy reflectance spectra as well as images based on explicitly described structures. Compared to 1-D RTMs, images simulated by 3-D RTMs not only provide spectral information but also texture information. In fact, the possibility to use 3-D RTMs to estimate canopy parameters has already been explored by some researchers [37]–[40]. Therefore, 3-D RTMs may be an alternative for hybrid models, especially within the current scenario of increasing computation power and model acceleration [41], [42].

In this article, we proposed a hybrid model [transfer learning-based convolutional neural network (T-CNN)] to estimate FCC based on the combination of a 3-D RTM and transfer learning-based CNN. There are three main advantages as follows: a large number of simulation images that cover various FCC situations are generated by the 3-D RTM; the application of transfer learning effectively reduces the requirement for a large actual dataset, which is a must for most data-driven models; and the CNN architecture without complex feature engineering intelligently selects the optimal combination of features to reduce manual workload.

In Section II, study areas, field measurements, satellite and LiDAR data, as well as the processing of data will be introduced. Section III describes the simulation process of a 3-D RTM, the realization of the proposed T-CNN method and the inversion of FCC using other machine learning methods. In Section IV, the results of T-CNN are compared with other methods. We also assess the transferability of the proposed T-CNN and compare performances with different training sizes. Section V and VI are the discussion and conclusion remarks.

II. Materials

A. Study Area and Field Measurements

Two study areas, which respectively represent natural forest and planted forest, were selected in this article. The first area is located in the north region (50 ° 55 ‘ N, 121 ° 30 ’E) of Genhe Forest Reserve in Inner Mongolia, China [see Fig. 1(a)]. It has a cold continental monsoon climate with an average temperature of around −5.3°C, annual precipitation from 450 to 500 millimeters, and an average elevation of 1000 m. The average temperature of the growing season ranges from 8°C in May to 16°C in July. Precipitation is concentrated in August and September. The growing season is generally from early May to late September. The forest coverage rate is around 75%. Dominant species are conifers, including Dahurian Larch (*Larix gmelinii*), White Birch (*Betula platyphylla* Suk). The understory vegetation of Dahurian Larch is mainly Rhododendron (*Rhododendron simsii* Planch.) or Ledum (*Ledum palustre* L). The height of Rhododendron is around 1.5 m, while Ledum is less than 0.3 m. The other study area is located in Chengde (42 ° 23 ’ N, 117 ° 19 ’ E), China [see Fig. 1(a)]. The average elevation is around 1500 m and the forest coverage rate is close to 80%.

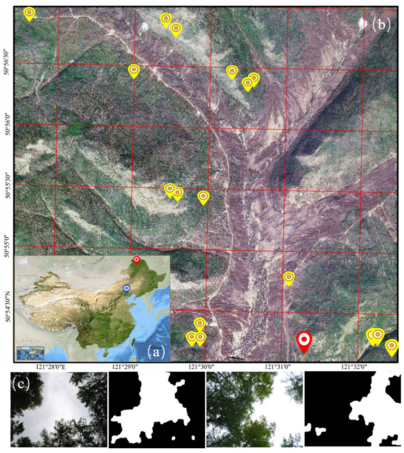

Fig. 1. Overview of the study area. (a) The Geographical location of Genhe (red point) and Chengde (blue point) in China. (b) Satellite image of Genhe area, including the distribution of field plots. (c) Original digital photographs (left) and the corresponding FCC images (right) after morphology operation based on the threshold segmentation.

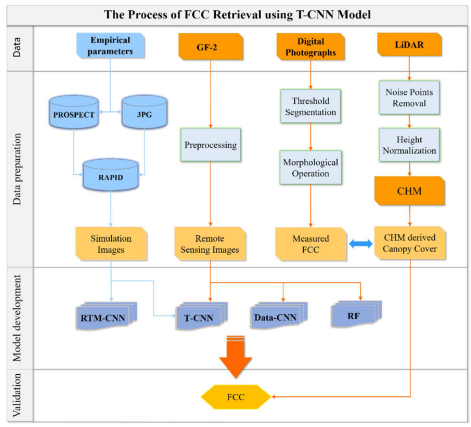

Fig. 2. Workflow of estimating FCC on the basis of T-CNN model.

The tree species mainly consist of Larix principis-rupprechtii (*Larix principis-rupprechtii* Mayr) and spruce [*Picea abies* (L.) H.Karst]. The understory vegetation mainly includes honeysuckle (*Lonicera japonica* Thunb), Wild Rose (*R. multiflora*).

FCCs were measured in 18 plots [45 m × 45 m, Fig. 1(b)] in the Genhe area by using upward photography [see Fig. 1(c)]. Upward photography provided a near-vertical direction observation, similar to the definition of FCC, which is also a conventional measurement of canopy cover. Approximately 50 upward photographs were taken using a Nikon D3000 along two diagonals across the plot in each plot. The size of the photograph is 2592 × 3872 pixels and the field of view is about 66.4°× 47.9°. The algorithm of Korhonen and Heikkinen [43] was employed to separate the crown pixels from the sky pixels, which uses a threshold to classify photographs into binary images and then use the morphology method to remove within-crown gaps. Finally, the FCC of each photograph is the fraction of crown pixels.

B. GaoFen-2 Data

The GaoFen-2 (GF-2) is a multispectral (514, 546, 656, and 822 nm), very-high-resolution (panchromatic band: 1 m; multispectral band: 4 m) and wide-swath (about 45 km) imaging mission, which supports the monitoring of vegetation, water bodies and soil covers. The GF-2 satellite was launched in 2014, which enables a revisit time of five days. However, GF-2 is only available from 2014, thus the image used in Genhe was obtained in June 2015. In addition, due to the weather with many clouds in Chengde, the other GF-2 imagery was obtained in June 2017. The preprocessing of GF-2 image was conducted in ENVI. First, the radiometric calibration coefficient of GaoFen-2 was used to calibrate digital numbers into radiance value. Second, the Gram-Schmidt Pan Sharpening method was used to fuse panchromatic imagery with 1 m resolution and multispectral imagery with 4 m resolution. And then, the quick atmosphere correction (QUAC) was used to correct the atmosphere effect to obtain a reflectance image. In addition, previous studies showed that there is a significant relationship between normalized difference vegetable index (NDVI) and canopy cover [8], [14], [18], therefore, NDVI was computed and added as the fifth channel of GF-2 images.

C. Airborne LiDAR Data

The airborne LiDAR with a small footprint was obtained from August 16 to September 25, 2012, in Genhe area. The airborne platform, which is equipped with a Leica ALS60, flies at an altitude of nearly 1800 m above ground. The scan angle is less than 35° and the point cloud density is around 8 points/m^2. In Chengde, the LiDAR data was obtained with the same instrument of Genhe. The flight height is also around 1800 m above ground, which produces a point cloud density around 8 points/m^2.

To obtain the canopy height model (CHM), the noise points were first removed manually, and cloth simulation filtering algorithm [44] was then used to separate ground points from nonground points. Finally, the CHM was generated using a pit-free algorithm embed in lidR package with a resolution of 0.5 m [45], [46].

III. Methods

The retrieving process mainly consists of three parts: data preparation; model development; and model validation, as shown in Fig. 2. All the methods including hybrid inversion model (T-CNN), CNN fitted with actual dataset only (data-CNN), CNN fitted with RTM simulated dataset only (RTM-CNN), and RF were implemented with Python. In addition, for testing the overall performance of these methods, the actual remote sensing data (Genhe) was split into training dataset (TR-dataset) and testing dataset (TE-dataset). Further, from one site to another site is challenging [47], [73], [74], thus the developed models were further validated on the independent dataset (IN-dataset). The IN-dataset was collected in a geographically different site, Chengde and never used to train the model.

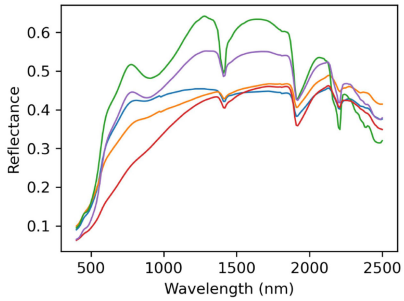

Fig. 3. Reflectance of the five soils used to represent the possible range of soil spectral.

A. Data Preparation

1) Actual Dataset: TR-dataset and TE-dataset are from GF-2 and the corresponding FCC that is derived from CHM in Genhe. GF-2 images were first georeferenced with CHM through manually selected control points. In addition, since the higher height of understory vegetation is around 1.5 m to 2 m, CHM derived FCC was computed as the percentage of pixels with an appropriate height value that larger than 2.5 m within a certain extent

$$\text{Canopy Cover} = \frac{\text{CHM}_{\text{value}>2.5}}{\text{CHM}_{\text{total}}} \tag{1}$$

Where $\text{CHM}_{\text{value}>2.5}$ means the pixels that CHM is higher than 2.5 belong to the canopy. $\text{CHM}_{\text{total}}$ represents all pixels of CHM.

To match the resolution of estimated FCC with field-measured FCC, the GF-2 images and CHM were clipped into small tiles, each of which has the same size of 48 m × 48 m, which produced 5700 tiles in Genhe area. As mentioned in the introduction, LiDAR is a relatively expensive measurement, thus reducing data volume would effectively decrease costs. Therefore, we focus more on whether the proposed model could generate a well accuracy with a small dataset. In order to test the model performance over a small dataset, we set aside 10% of all the tiles for TR-dataset and 90% for TE-dataset, which is different from common machine learning where the training dataset is usually much larger than the testing dataset.

IN-dataset was mainly set to test the transferability of a machine learning model and was not used in the training process. In this article, it has the same processing protocol as TR-dataset and TE-dataset, the only difference is that the GF-2 and CHM data used in IN-dataset are from a different location, i.e., Chengde area, where canopy spatial arrangement has apparent difference with Genhe area. IN-dataset has 11 020 tiles.

2) Simulation Dataset: In this article, a Radiosity Applicable to Porous IndiviDual Objects for directional reflectance over complex vegetated scenes (RAPID) model is used to simulate GF-2 imagery. RAPID is a 3-D RTM by using the concept of porous individual thin objects, which was proposed in 2013 by Huang *et al.* [35]. With consisting development, the RAPID model can simulate optical, LiDAR, thermal, and microwave signals with a unified 3D scene and input parameters [48]–[50]. RAPID reasonably simplifies the 3-D structure of vegetation, thus improving the efficiency of radiative transfer simulation in large scenes. In other words, it is characterized by batch simulating images efficiently, which is more suitable for training CNN that requires a larger number of images. Thus, the newest version of RAPID (downloaded from[1]) was used to simulate multispectral images.

TABLE I
INPUT PARAMETERS OF RAPID SIMULATIONS

Parameter	Notation (Unit)	Range
Diameter at breast height	DBH [cm]	5.2-30.2
Tree height	H [m]	6-16
Crown width	Cw [m]	1.03-3.32
Crown length	Cl [m]	1.93-9.41
Leaf area index	LAI [m^2/m^2]	0.53-11.4
Tree location	(X, Y) [m]	0-48
Bands	B [nm]	514, 546, 656, 822
Observed view	OV [°]	0
Sun location	(Azimuth, Zenith) [°]	125-250, 0-60
N	/	1-3 (step:0.5)
Chlorophyll content	Cab [ug/cm^2]	10-90 (step:10)
Brown pigment	Cbrown	0.001
Dry matter content	Cm[g/cm^2]	0.02
Equivalent water thickness	Caw[cm]	0.005

A prerequisite to simulate images using 3-D RTM is to build virtual scenes. In order to make the virtual forest scene more reasonable, we used physiological processes predicting growth model (3-PG) to simulate the input parameters of RAPID. 3-PG is a forest growth model based on simple plant physiological principles [51], which has been used to estimate forest productivity in many countries and regions. As a result, with the use of the 3-PG, the number of parameter combinations for RAPID was effectively reduced. For example, the input parameters of RAPID, such as mean diameter at breast height, leaf area index, stem density was generated by 3-PG. The locations of trunks were randomly distributed within the plot. The leaf spectra were obtained from the PROSPECT model by inputting different combinations of leaf parameters to cover various land surface conditions. In fact, in the case of just using red, green, blue, and near-infrared band, the variation of chlorophyll content and structure parameter *N* has covered possible distribution range of common leaf spectra. After testing, fixing brown pigment, dry matter content and equivalent water thickness does not lose accuracy, thus we only modified chlorophyll content and *N* for the sake of simplicity, as given in Table I. For illumination and observation geometries, nadir view was always guaranteed, but the solar zenith and azimuth ranged from 0° to 60° and from 125° to 250°, respectively. The varying solar angles can also emulate the local solar angle change induced by rugged terrains, since previous studies showed that solar position and terrain both have a certain degree of influence on the inversion of forest parameters [52].

The soil spectra data used in this article was downloaded from the International Soil Reference and Information Centre.[2] These soil data contained a wide range of soil types with different soil brightness in the world [53], including 253 Chinese soil profiles. To remove redundant soil data and decrease calculation volume,

[1][Online]. Available: http://www.3dforest.cn/
[2][Online]. Available: http://www.isric.org

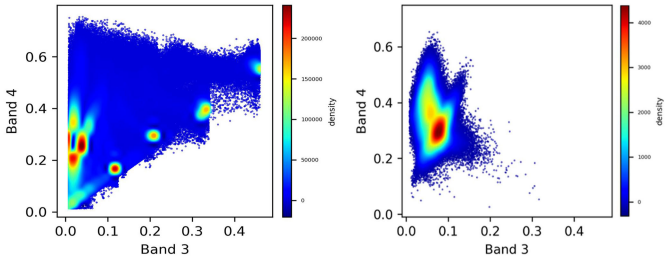

Fig. 4. Density scatter plots of near infrared (band 4) and red band (band 3) reflectance. On the left, the scatter distribution for simulation dataset. On the right, the distribution of real dataset from GF-2 in study area.

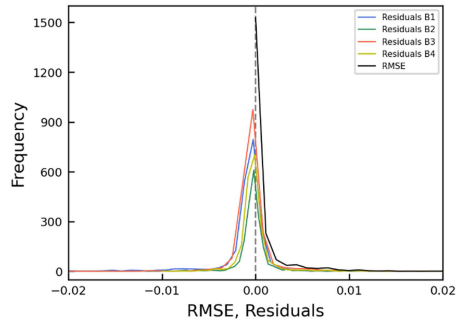

Fig. 5. Distribution of the mismatch values (RMSE) as computed by (2) over a database of 1000 actual spectra.

we employed mean values of similar profiles. The spectral angle mapper [14] was used to filter similar soil profiles with a spectral angle smaller than 0.1. In the end, five representative soil reflectance (see Fig. 3) were determined from original soil data. Based on the parameter combinations in Table I, a total of 65318 images were finally simulated.

First evaluation of the simulations was performed by simply comparing the red (band 3) and near infrared (band 4) reflectance values (see Fig. 4). As expected, the soil points are more concentrated in the bottom and right part of the scatter plot. Further evaluating was executed by comparing real dataset, which was composed of 150 000 random sampling pixels from GF-2 data in study area. Results shows that the differences exist mainly on the right side of the scatter plot, which is due to the real soil pixels have small percentage in study area (the average FCC was both larger than 0.6). The distribution of simulation pixels has completely covered the real dataset distribution.

However, more important is the realism with which simulated spectra follow the actual pixels. The mismatch, φ, between each randomly selected pixels (Ω) and the nearest simulated spectra ($\hat{\gamma}$) was computed according to (2) over the 150 000 actual spectra (γ) for the four bands (b) considered here

$$\varphi = \min_{\Omega} \left(\sqrt{\frac{\sum_{b=1}^{4} (\gamma_b - \hat{\gamma}_b)^2}{4}} \right). \tag{2}$$

Fig. 5 shows that most of RMSEs calculated by (2) and residuals of each band are concentrated around zero, which indicates that simulated reflectance is very close to an actual case.

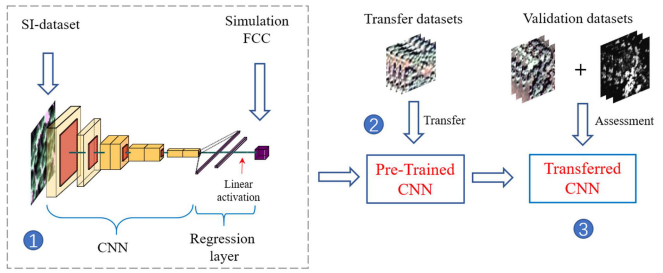

Fig. 6. Workflow of the proposed method. First, the simulation datasets from RAPID are learned by CNN. Then, the learned model is fine-tuned with remotely sensed data, achieving a transferred inversion of FCC. Finally, the field measurements and CHM-derived FCC are used to validate the model.

B. Transfer Learning-Based Convolutional Neural Network

CNN is a multilayer stacked network, which mainly includes input layer, hidden layer (convolutional, pooling, fully connected), and output layer. Compared with general CNN training, the proposed T-CNN mainly has two stages (see Fig. 6), including model pretraining and transfer learning. In model pretraining, we used the simulation dataset (SI-dataset) to pre-train the CNN. The purpose of this step is to initialize the neural network weights. Usually, CNN is composed of several blocks and fully connected layers at the end. The block generally consists of one or two convolution layers and a special pooling layer. The convolution layer is responsible for extracting data feature maps and the pooling layer focuses on reducing the dimensionality of feature maps and presenting redundant information. The fully connected layer is used to implement regression or classification tasks. A linear activation function was applied before the output of the fully connected layer. Where the linear activation function stands for regression function with the mathematical formulation

$$y_i = \sum_i (w_i x_i) + b_i \tag{3}$$

where y_i is the output of the model, w_i is the weight between ith feature and output, x_i is the ith feature value and b_i is the bias. In addition, the mean absolute percentage error function was used as the loss function

$$L\ (Y, f\,(x)) = \frac{1}{m} \sum_{i=1}^{m} \left| \frac{Y - f\,(x)}{Y} \right| \tag{4}$$

where Y and $f(x)$ represent the reference and prediction FCC, respectively, and m is the number of samples.

The study employed a gradient descent algorithm called Adam optimizer, which can dynamically adjust the gradient descent process and accelerate model convergence. In this article, we used a dropout layer, which randomly samples the connections between neurons of different fully connected layers in a certain ratio, and early stopped the training process to avoid overfitting.

At the transfer learning stage, only TR-dataset was used to fine-tune the model. In particular, the first three layers of CNN were frozen and the rest layers were retrained with the TR-dataset. Other settings, such as activation function, were the same as those in the previous stage.

TABLE II
DATASETS USED FOR DIFFERENT INVERSION MODELS

Methods	Training	Transfer learning	Validation
T-CNN			
- Pre-training	SI-dataset	\	TE-dataset, IN-dataset
- Transfer	\	TR-dataset	
Data-CNN	TR-dataset	\	TE-dataset, IN-dataset
RTM-CNN	SI-dataset	\	TE-dataset, IN-dataset
RF	TR-dataset	\	TE-dataset, IN-dataset

C. Comparison With Other Machine Learning Algorithms

RF introduced by Breiman [54] is a machine learning algorithm commonly used for FCC regression [22], [55]. RF, which is composed of many decision trees, obtains classification or regression results by voting or calculating the average value. Therefore, RF is more robust than general decision trees. Besides, RF uses the bootstrap method to carry out random sampling of the training samples, which prevents the overfitting of the model. However, it would lead to a certain percentage (about 36.8%) of samples (out-of-bag error data) not being sampled. Hence, out-of-bag error data can be used to verify robustness with cross-validation. One of the cores of RF is feature selection. In this article, we chose all bands (B1–B4), vegetable indexes (NDVI, Ratio Vegetation Index, Difference Vegetation Index), and texture (mean, variance, contrast, dissimilarity, homogeneity, energy, correlation, angular second moment), to build feature dataset. Besides, to remove the collinearities among features, recursive feature elimination was used to filter features. In addition, we also tested the number of estimators and depth of RF by a large number of trials.

Data-CNN employed the same CNN architecture as the one used in the pre-training stage of T-CNN. The only difference between them was the input data. Specifically, Data-CNN used the TR-dataset, while T-CNN, in the pre-training stage, used the SI-dataset.

RTM-CNN also has exactly the same architecture as T-CNN and Data-CNN. Compared with T-CNN, it only has the pre-training stage. This model is mainly for assessing the inversion performance based purely on SI-dataset, which is a very widely used approach in remote sensing inversion studies. The application details of dataset are given in Table II.

D. Accuracy Assessment

To evaluate the accuracy of the model, we used the coefficient of determination (R^2), root-mean-squared error (RMSE) and bias as assessment indicators

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - y_i)^2}{n}} \tag{5}$$

$$\mathrm{Bias} = \frac{\sum_{i=1}^{n}(xi - yi)}{n} \tag{6}$$

where x_i and y_i are the cover and the predicted cover of the forest. n means the quantity of validation data.

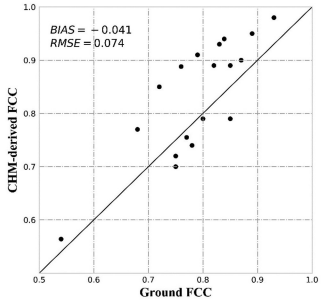

Fig. 7. Validation of the CHM-derived FCC using field photography.

IV. RESULTS

A. Reliability of CHM-Derived FCC

Fig. 7 shows the comparison between CHM-derived FCC and photograph-derived FCC. The result shows that there is a correlation with an RMSE of 0.074 and a Bias of −0.041, which indicated that CHM-derived FCC employed by the study can be considered as an accurate FCC estimation.

B. Performance of Data-Driven Models

For RF, Data-CNN and T-CNN, TR-dataset, TE-dataset and IN-dataset were used to validate the estimation accuracy and the results were given in Table III. Considering the CHM-derived FCC (including TR-dataset and TE-dataset) at Genhe, the T-CNN model ($R^2 = 0.80$, RMSE = 0.109) outperformed RF ($R^2 = 0.75$, RMSE = 0.122) and Data-CNN ($R^2 = 0.74$, RMSE = 0.126) models on the TE-dataset. But the performance of Data-CNN was better than RF at TR-dataset.

A similar goodness-of-fit validation was implemented on the IN-dataset. Fig. 8 shows that the relationship between CHM-derived FCC and the corresponding prediction of different models. In terms of RMSE and R^2, results showed that accuracy of T-CNN ($R^2 = 0.83$, RMSE = 0.121) was higher than RF ($R^2 = 0.61$, RMSE = 0.26) and Data-CNN ($R^2 = 0.81$, RMSE = 0.142). In addition, noticed that the transfer performance of RF was quite poor. For example, FCC was overestimated at low canopy cover and was underestimated at high canopy cover.

The validation results (see Table III) from the ground FCC at Genhe showed that T-CNN ($R^2 = 0.06$) did not always perform well based on R^2, which may be due to the small validation sample size and imbalanced ground FCCs that are generally greater than 0.6 in Genhe. Even so, the accuracy of T-CNN (RMSE = 0.105 and bias = -0.045) has already been closed to that of RF (RMSE = 0.079 and Bias = 0.019).

C. RTM-CNN Performance

FCC can also be inverted by RTM-CNN trained with SI-dataset. Unlike data-driven models, RTM-CNN did not use any actual measurements, i.e., GF-2 images and the corresponding CHM-derived FCC. We compared its estimation accuracy with IN-dataset and the full of the dataset from Genhe, i.e., the total of TR-dataset and TE-dataset. As shown in Fig. 9, RTM-CNN has better performance on the IN-dataset than on the Genhe dataset

TABLE III
RANDOM FOREST, DATA-DRIVEN DEEP LEARNING (DATA-CNN) AND DEEP TRANSFER LEARNING (T-CNN) ACCURACY STATISTICS ON TRANSFER DATASET, TEST DATASET, GROUND MEASUREMENTS (GROUND), AND INDEPENDENT DATASET

	RTM-CNN				RF				Data-CNN				T-CNN			
	TR	TE	Ground	IN	TR	TE	Ground	IN	TR	TE	Ground	IN	TR	TE	Ground	IN
Sample size	570	5130	18	8950	570	5130	18	8950	570	5130	18	8950	570	5130	18	8950
R^2	0.65	0.61	0.10	0.73	0.88	0.75	0.10	0.61	0.98	0.74	0.05	0.81	0.99	0.80	0.06	0.83
RMSE	0.15	0.152	0.108	0.144	0.085	0.122	0.079	0.260	0.033	0.126	0.133	0.142	0.024	0.109	0.105	0.121
Bias	-0.028	-0.031	-0.047	0.038	-0.0006	-0.005	0.019	0.202	-0.008	-0.019	-0.056	0.082	-0.0004	-0.011	-0.045	0.054

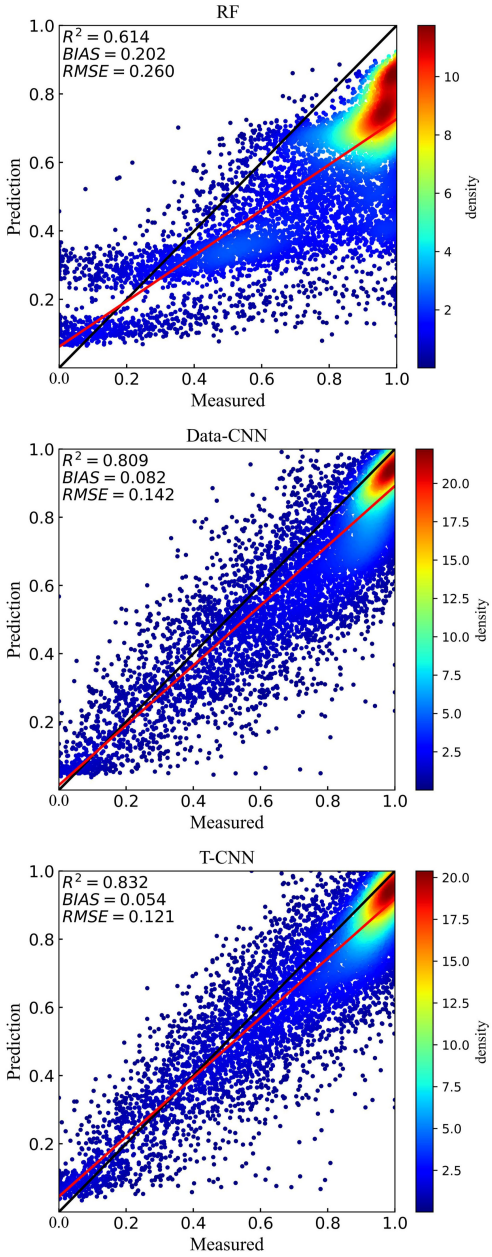

Fig. 8. Performance of different models on the IN-dataset.

in terms of RMSE or R^2, but we also found more discrete points on the IN-dataset.

D. Influences of Training Dataset Size

Fig. 10 showed the accuracies (RMSE) of RF, Data-CNN and T-CNN under different training size. Please noted that the training dataset and validation dataset here refer to GF-2 images and the corresponding CHM-derived FCC from Genhe. Considering the performance of T-CNN model, higher prediction accuracy was obtained for any training sample size and RMSE decreased significantly from 0.108 to 0.084 with increasing sample size. For RF model, RMSE was insensitive to training size, but had a larger error (between 0.116 and 0.122) than T-CNN. For Data-CNN model, RMSE decreased from 0.128 to 0.103 with the increase of training size. In addition, the performance of Data-CNN model was worse than RF when the sample size was small. Subsequently, it surpassed the accuracy of RF and approached T-CNN, as the number of training data size increased.

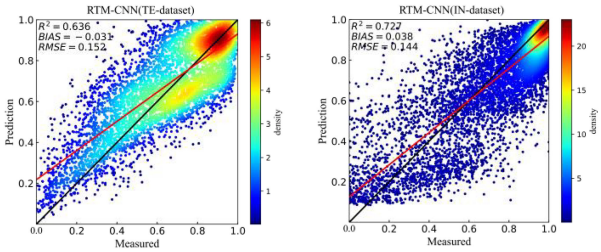

Fig. 9. Performance of RTM-CNN on the different datasets.

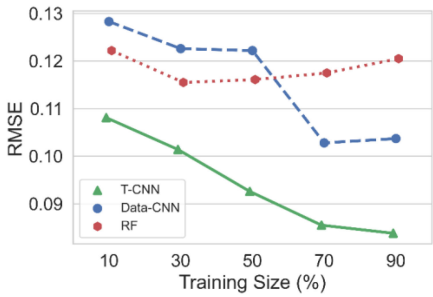

Fig. 10. Prediction performance of RF, Data-CNN and T-CNN models with respective samples size.

V. DISCUSSION

Accurate estimate of FCC is of great importance for policy makers and practical applications. The transferable prediction of the FCC is at the core of this article. The fundamental assumption underlying our method is that deep transfer learning can be trained with simulation data generated by RTMs, and be transferred to specific regions through a small number of real samples.

A. T-CNN Model Performance on FCC Estimation

FCC prediction accuracy was higher when CNN was employed than other machine learning methods (see Table III). This finding was also consistent with Kattenborn *et al.* [28] who used CNN to regress the canopy cover based on images from UAV. Similarly, Pullanagari *et al.* [47] confirmed that 1-D-CNN has a higher accuracy of nitrogen content estimation based on a large number of multiposition and multitime hyperspectral data. But these studies often require large amounts of data to support model development, which is a challenge in forestry due to the complexity and difficulty of forest surveys. The application of hybrid models can help to reduce the need for actual datasets [27]. For instance, Annala *et al.* [33] presented a promising test to successfully invert Cab using a hyperspectral simulation dataset from RTM SLOP instead of hyperspectral measurement data from UAV. Previous studies have focused more on hyperspectral data and 1-D-CNN, and rarely ventured into the application of multi-spectral images. As a further succession and extension of previous studies, this research successfully inverts FCC based on simulation dataset from RAPID combined with CNN (here refers to RTM-CNN) as shown in Fig. 9. The prediction accuracy was further improved by using transfer learning which is hot topic in the computer vision. It was worth noting that we only used a small amount of data (TR-dataset) for transfer but obtained better accuracy than Data-CNN.

More complex models, e.g., CNN, are more likely to overfitting that model performs well in training dataset while performs badly in the testing dataset. Compared to RF, T-CNN and Data-CNN had significant overfitting problems (see Table III), where R^2 on TR-dataset was 0.99, while $R2$ on TE-dataset was much lower. Thus, the performance of models on the independent dataset was quite important for evaluating the robustness. As shown in Fig. 8, T-CNN model has the best performance with an RMSE of 12.1% on IN-dataset, which is also the highest accuracy within the previously published studies (RMSE 12-18%) on optically-based FCC estimation in the boreal forest [8], [56], [57] or conifer forest sites [12], [58]. In addition, we obviously observed RF has an overestimation of low cover and underestimation of high cover, which was consistent with Sexton *et al.* [59], Hadi *et al.* [8] and DiMiceli *et al.* [60]. The overestimation may be caused by the understory, which usually has 20–45% contribution in the canopy reflectance [61], especially in low and middle areas where the understory is incorrectly recognized as overstory since only reflectance information, instead of height, is used. But this phenomenon was not observed in T-CNN and Data-CNN, which may be attributed to simulation dataset providing priors and realistic constraints in model training [62] and the more effective feature extraction of CNN than manual feature selection of RF [23], [63].

B. Evaluating the Predictive Performance of RTM-CNN

Although, RTMs have been successfully inverted for many forest parameters, e.g., leaf area index, canopy cover, nitrogen content, etc., what has been a source of criticism is its ill-posed inversion. This research evaluated the performance of RTM-CNN on two different datasets and found that its estimation accuracy was lower than that of the data-driven models (see Table III). In particular, there were significantly more discrete predictions on IN-dataset (see Fig. 9), which may be due to the differences between simulation data and real remote sensing data. And the reason for these differences is often because of the complexity of canopy and uncertainties related to measurements, which may lead to ill-posed inversion [64]. Just as RTM's simplification of the canopy shape might be not consistent with the real canopy and the distribution of trees did not correspond to the distribution under real nature conditions in this research. Koetz *et al.* [65] used LiDAR to reconstruct the 3-D structure of the single tree and inverted LAI, Cab, canopy cover through RTM. They found that canopy cover estimates based on the real 3-D structure of a single tree resulted in a 22.2% lower RMSE and a 17% increase in terms of R2. In summary, inversion of RTM does not use any measurements, but its ill-posed problem still affects the estimation accuracy of FCC.

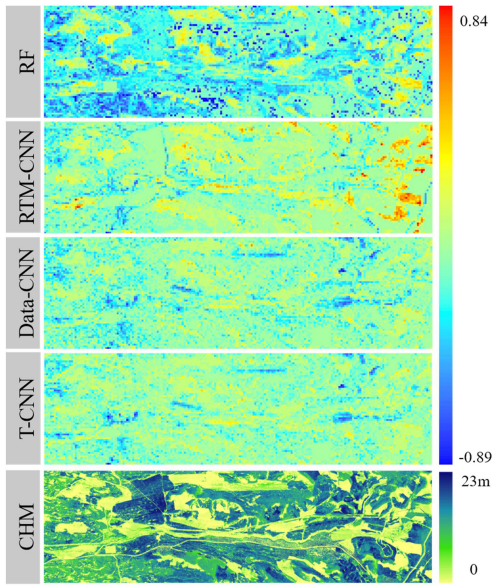

Fig. 11. Distribution of estimation error in IN-dataset using RF, RTM-CNN, Data-CNN, and T-CNN model. Negative values mean that the cover is underestimated.

C. Spatial Distribution of Estimation Errors

The prediction of FCC is often disturbed by many factors (understory, terrain, illumination, and zenith angles). Hence, accurately predicting FCC is still a very difficult task. As shown in Figs. 8 and 9, some scatters that are away from the 1:1 line have significant errors. To further illustrate the error distribution, the error map of different models on IN-dataset was shown in Fig. 11. The result shows that the error maps of different inversion models have similar spatial distribution. Compared with CHM, we found overestimation often occurs in bare ground with grass and underestimation is more likely to occur at the edge of the forest, which is clearly reflected in the error maps of all models. Apparently, the grass has similar reflectance with tree crown, thus grassland may be wrongly recognized as tree crown. For the underestimation around the border between forest and grassland, it is may be caused by the mixture pixel that contains both grassland and tree crowns, which dramatically changes the

spectra and texture information. Nevertheless, compared with RT, T-CNN and Data-CNN still reduce the impact of these factors.

D. Influence of Training Size for Model Development

Machine learning algorithms are particularly appealing in the remote sensing field due to the powerful nonlinear fitting ability [66], [67]. However, the imbalance and scarce training data is common in large area natural resource applications using remote sensing, such as forest classification [29], which affects the prediction accuracy. At present, most studies take more attention to how the machine learning algorithm itself overcomes this problem. But the algorithm itself is still troubled by the amount of input information. This article provides new thinking from the perspective of data enhancement and data simulation. As shown in Fig. 10, T-CNN has the best performance against the different training sizes. Please noted that the accuracy of Data-CNN may be close to T-CNN in the case of sufficient training samples. However, labeling a large amount of training data is extremely expensive for Data-CNN. T-CNN is thus also a good alternative to Data-CNN. Similarly, Ma *et al.* [68] have used simulation data to expand the training dataset for Data-CNN implementing image segmentation, and achieved the aim of reducing the real labeled samples at the same level of accuracy.

E. Model Transferability

Typically, many regression methods are accustomed to splitting the dataset into training and test dataset. The fact that the test dataset is randomly sampled and validates the accuracy of the model from the same data source is not a good representation of the robustness of the model. This approach may cause high accuracy at the test dataset while it is less likely to extend on an independent dataset. Thus, some studies used the datasets from different locations or times to validate model transferability. Thus, in the article, an IN-dataset from Chengde, which was not included in the TR-dataset and TE-dataset, was used to validate model transferability.

We observed that both Data-CNN and T-CNN with CNN architecture have better transferability than RF (see Fig. 8), which may be attributed to the powerful feature extraction capability of CNN that helps to identify and extract the key features. For instance, Pullanagari *et al.* [47] transferred 1-D-CNN model to other sites and gained highest estimation accuracy for nitrogen content compared to other machine learning methods. In addition, T-CNN performed better than Data-CNN, which is due to the richness of simulation dataset achieved by samples from different parameter combinations. Just like Siegmann and Jarmer [69] improved prediction accuracy for LAI through incorporating a large range of LAI data. Even though independent site was geographically and temporally different, the performance of T-CNN still was best, which indicates powerful transferability and reliability.

F. Effects of Different Resolutions for T-CNN

In recent years, there has been an increasing number of high-resolution FCC mapping studies [70], which is beneficial for our quantitative management of forests. High-resolution FCC mapping as the input parameters for some physical and statistical models helps to improve model accuracy. For instance, Gastón *et al.* [71] evaluated the influences of the different vegetation resolutions for multiscale habitat selection models and found that high-resolution FCC helps to estimate species' habitat. Although this article mainly studied the FCC at a coarser resolution (48 m), the FCC of different resolutions can also be estimated by modifying the input size of T-CNN. As shown in Fig. 12, the estimation results of different resolutions (10, 20, 30 m) are compared under 60 m resolution. It can be seen that FCC predictions under different resolutions are generally consistent with average differences of all pixels less than 5%. This indicates that the proposed T-CNN can be used to estimate FCC under different resolutions without considering scale effect.

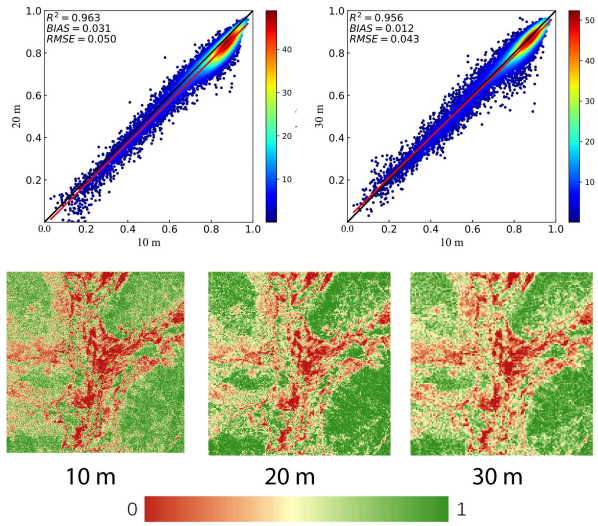

Fig. 12. Relationships between 10 m and 20 m, 30 m of FCC for T-CNN are shown (upper). The FCCs of different resolutions are aggregated to 60 m resolution for comparison. The FCC mappings with different resolutions are shown (bottom).

VI. Conclusion

This article proposed T-CNN method to estimate FCC, which combines 3-D RTM RAPID and transfer learning based on CNN. The significance of this method is threefold. First, transfer learning provides an effective approach to fill the gap between simulated dataset and actual dataset, which makes the inversion algorithm more adaptive to different areas. Second, T-CNN achieved relatively high accuracy by using only a small amount (10%) of the actual dataset. It also mitigates the data imbalance problem, since the simulation dataset used to train the model has already covered a wide range of conditions. Finally, overestimation of low cover and underestimation of high cover, which is common in many empirical models, have been optimized. Compared with RF, Data-CNN and RTM-CNN, T-CNN achieved the highest prediction accuracy.

Despite its effectiveness, T-CNN still needs actual measurements, which may be not always fulfilled, especially when mapping forest parameters on large scales. Future work may

focus on how to narrow the gap between remote sensing images and simulated images by using GAN network, such as cycle-GAN, which may completely remove the requirement for actual measurements.

REFERENCES

[1] C. Mahoney, R. Hall, C. Hopkinson, M. Filiatrault, A. Beaudoin, and Q. Chen, "A forest attribute mapping framework: A pilot study in a northern Boreal forest, northwest territories, Canada," *Remote Sens.*, vol. 10, no. 9, p. 1338, 2018.

[2] T. F. Stocker *et al.*, Climate Change 2013: The physical science basis. contribution of working group I to the fifth assessment report of ipcc the intergovernmental panel on climate change, 2014.

[3] T. Gschwantner *et al.*, "Common tree definitions for National Forest Inventories in Europe," *Silva Fennica*, vol. 43, no. 2, 2009.

[4] G. J. Williamson, L. D. Prior, M. R. Grose, R. M. B. Harris, and D. M. J. S. Bowman, "Projecting canopy cover change in Tasmanian Eucalypt forests using dynamically downscaled regional climate models," *Regional Environ. Change*, vol. 14, no. 4, pp. 1373–1386, 2014.

[5] W. Song, X. Mu, G. Ruan, Z. Gao, L. Li, and G. Yan, "Estimating fractional vegetation cover and the vegetation index of bare soil and highly dense vegetation with a physically based method," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 58, pp. 168–176, 2017.

[6] P. J. Sellers, "Modeling the exchanges of energy, water, and carbon between continents and the atmosphere," *Science*, vol. 275, no. 5299, pp. 502–509, 1997.

[7] "Global forest resources assessment updata 2005," Food Agriculture Org. United Nations, Rome, Italy, 2004.

[8] L. K. Hadi, A. Hovi, P. Rönnholm, and M. Rautiainen, "The accuracy of large-area forest canopy cover estimation using landsat in Boreal region," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 53, pp. 118–127, 2016.

[9] N. K. Patel, R. K. Saxena, and A. Shiwalkar, "Study of fractional vegetation cover using high spectral resolution data," *J. Indian Soc. Remote Sens.*, vol. 35, no. 1, pp. 73–79, 2007.

[10] J. M. B. Carreiras, J. M. C. Pereira, and J. S. Pereira, "Estimation of tree canopy cover in evergreen Oak Woodlands using remote sensing," *Forest Ecol. Manage.*, vol. 223, no. 1-3, pp. 45–53, 2006.

[11] P. R. J. North, "Estimation of fAPAR, LAI, and vegetation fractional cover from ATSR-2 imagery," *Remote Sens. Environ.*, vol. 80, no. 1, pp. 114–121, 2002.

[12] D. Stojanova, P. Panov, V. Gjorgjioski, A. Kobler, and S. Džeroski, "Estimating vegetation height and canopy cover from remotely sensed data with machine learning," *Ecol. Inform.*, vol. 5, no. 4, pp. 256–266, 2010.

[13] M. Karlson, M. Ostwald, H. Reese, J. Sanou, B. Tankoano, and E. Mattsson, "Mapping tree canopy cover and aboveground biomass in Sudano-Sahelian woodlands using landsat 8 and random forest," *Remote Sens.*, vol. 7, no. 8, pp. 10017–10041, 2015.

[14] K. Jia *et al.*, "Fractional vegetation cover estimation algorithm for chinese GF-1 wide field view data," *Remote Sens. Environ.*, vol. 177, pp. 184–191, 2016.

[15] J.-L. Roujean and L. Roselyne, "Global mapping of vegetation parameters from POLDER multiangular measurements for studies of surface-atmosphere interactions: A pragmatic method and its validation," *J. Geophys. Res.*, vol. 107, no. D12, 2002.

[16] F. Baret *et al.*, "LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION," *Remote Sens. Environ.*, vol. 110, no. 3, pp. 275–286, 2007.

[17] F. J. García-Har, F. Camacho-de Coca, and J. Meliá Miralles, "Intercomparison of SEVIRI/MSG and MERIS/ENVISAT biophysical products over Europe and Africa," in *Proc. 2nd MERIS/(A)ATSR User Workshop, Frascati*, 2008.

[18] G. Jiapaer, X. Chen, and A. Bao, "A comparison of methods for estimating fractional vegetation cover in arid regions," *Agricultural Forest Meteorol.*, vol. 151, no. 12, pp. 1698–1710, 2011, doi: 10.1016/j.agrformet.2011.07.004.

[19] N. Lang, K. Schindler, and J. D. Wegner, "Country-wide high-resolution vegetation height mapping with Sentinel-2," *Remote Sens. Environ.*, vol. 233, p. 111347, 2019.

[20] Q. Ma, Y. Su, and Q. Guo, "Comparison of canopy cover estimations from airborne LiDAR, aerial imagery, and satellite imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 9, pp. 4225–4236, Sep. 2017, doi: 10.1109/jstars.2017.2711482.

[21] L. Korhonen, I. Korpela, J. Heiskanen, and M. Maltamo, "Airborne discrete-return LiDAR data in the estimation of vertical canopy cover, angular canopy closure and leaf area index," *Remote Sens. Environ.*, vol. 115, no. 4, pp. 1065–1080, 2011.

[22] O. S. Ahmed, S. E. Franklin, M. A. Wulder, and J. C. White, "Characterizing stand-level forest canopy cover and height using landsat time series, samples of airborne LiDAR, and the random forest algorithm," *ISPRS J. Photogramm. Remote Sens.*, vol. 101, pp. 89–101, 2015.

[23] H. Gjoreski, J. Bizjak, M. Gjoreski, and M. Gams, "Comparing deep and classical machine learning methods for human activity recognition using wrist accelerometer," in *Proc. the IJCAI-16 Workshop Deep Learn. Artif. Intell.*, 2016.

[24] C. Zhang *et al.*, "Joint deep learning for land cover and land use classification," *Remote Sens. Environ.*, vol. 221, pp. 173–187, Feb. 2019.

[25] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Comput. Electron. Agriculture*, vol. 147, pp. 70–90, 2018.

[26] Q. Yuan *et al.*, "Deep learning in environmental remote sensing: Achievements and challenges," *Remote Sens. Environ.*, vol. 241, p. 111716, 2020.

[27] T. Kattenborn, J. Leitloff, F. Schiefer, and S. Hinz, "Review on convolutional neural networks (CNN) in vegetation remote sensing," *ISPRS J. Photogramm. Remote Sens.*, vol. 173, pp. 24–49, 2021.

[28] T. Kattenborn *et al.*, "Convolutional neural networks accurately predict cover fractions of plant species and communities in unmanned aerial vehicle imagery," *Remote Sens. Ecol. Conserv.*, vol. 6, no. 4, pp. 472–486, 2020.

[29] A. Mellor, S. Boukir, A. Haywood, and S. Jones, "Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the ensemble margin," *ISPRS J. Photogramm. Remote Sens.*, vol. 105, pp. 155–168, May 2015.

[30] C. Huang, L. S. Davis, and J. R. G. Townshend, "An assessment of support vector machines for land cover classification," *Int. J. Remote Sens.*, vol. 23, no. 4, pp. 725–749, 2010.

[31] A. X. Wang, C. Tran, N. Desai, D. Lobell, and S. Ermon, "Deep transfer learning for crop yield prediction with remote sensing data," in *Proc. 1st ACM SIGCAS Conf. Comput. Sustain. Soc.*, 2018.

[32] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010, doi: 10.1109/TKDE.2009.191.

[33] L. Annala, E. Honkavaara, S. Tuominen, and I. Pölönen, "Chlorophyll concentration retrieval by training convolutional neural network for stochastic model of leaf optical properties (SLOP) inversion," *Remote Sens.*, vol. 12, no. 2, 2020.

[34] J. P. Gastellu-Etchegorry, V. Demarez, V. Pinel, and F. Zagolski, "Modeling radiative transfer in heterogeneous 3-D vegetation canopies," *Remote Sens. Environ.*, vol. 58, no. 2, pp. 131–156, Nov. 1996.

[35] H. Huang, W. Qin, and Q. Liu, "RAPID: A radiosity applicable to porous individual objects for directional reflectance over complex vegetated scenes," *Remote Sens. Environ.*, vol. 132, pp. 221–237, 2013.

[36] J. Qi *et al.*, "LESS: Large-scale remote sensing data and image simulation framework over heterogeneous 3D scenes," *Remote Sens. Environ.*, vol. 221, pp. 695–706, 2019.

[37] D. Kimes, J. Gastellu-Etchegorry, and P. Estève, "Recovery of forest canopy characteristics through inversion of a complex 3D model," *Remote Sens. Environ.*, vol. 79, no. 2, pp. 320–328, Feb. 2002.

[38] A. Banskota *et al.*, "Investigating the utility of wavelet transforms for inverting a 3-D radiative transfer model using hyperspectral data to retrieve forest LAI," *Remote Sens.*, vol. 5, no. 6, pp. 2639–2659, 2013.

[39] A. Banskota *et al.*, "An LUT-based inversion of DART model to estimate forest LAI from hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 3147–3160, Jun. 2015.

[40] T. Miraglio, K. Adeline, M. Huesca, S. Ustin, and X. Briottet, "Joint use of PROSAIL and DART for fast LUT building: Application to gap fraction and leaf biochemistry estimations over sparse oak stands," *Remote Sens.*, vol. 12, no. 18, p. 2925, 2020.

[41] J. Qi, T. Yin, D. Xie, and J.-P. Gastellu-Etchegorry, "Hybrid scene structuring for accelerating 3D radiative transfer simulations," *Remote Sens.*, vol. 11, no. 22, p. 2637, 2019.

[42] H. Huang *et al.*, "A realistic structure model for large-scale surface leaving radiance simulation of forest canopy and accuracy assessment," *Int. J. Remote Sens.*, vol. 30, no. 20, pp. 5421–5439, 2009.

[43] L. Korhonen and J. Heikkinen, "Automated analysis of in situ canopy images for the estimation of forest canopy cover," *Forest Sci.*, vol. 55, pp. 323−334, 2009.

[44] W. Zhang *et al.*, "An Easy-to-use airborne LiDAR data filtering method based on cloth simulation," *Remote Sens.*, vol. 8, no. 6, p. 501, 2016.
[45] J.-R. Roussel *et al.*, "LidR: An R package for analysis of airborne laser scanning (ALS) data," *Remote Sens. Environ.*, vol. 251, p. 112061 2020.
[46] A. Khosravipour, A. K. Skidmore, M. Isenburg, T. Wang, and Y. A. Hussin, "Generating Pit-free canopy height models from airborne LiDAR," *Photogramm. Eng. Remote Sens.*, vol. 80, no. 9, pp. 863–872, 2014.
[47] R. R. Pullanagari, M. Dehghan-Shoar, I. J. Yule, and N. Bhatia, "Field spectroscopy of canopy nitrogen concentration in temperate grasslands using a convolutional neural network," *Remote Sens. Environ.*, vol. 257, 2021.
[48] K. Du *et al.*, "Simulation of ku-band profile radar waveform by extending radiosity applicable to porous individual objects (RAPID2) model," *Remote Sens.*, vol. 12, no. 4, 2020.
[49] H. Huang *et al.*, "Extending RAPID model to simulate forest microwave backscattering," *Remote Sens. Environ.*, vol. 217, pp. 272–291, 2018.
[50] H. Huang and R. H. Wynne, "Simulation of LiDAR waveforms with a time-dependent radiosity algorithm," *Can. J. Remote Sens.*, vol. 39, pp. S126–S138, 2014.
[51] J. J. Landsberg and R. H. Warning, "A generalised model of forest productivity using simplified concepts of radiation-use efficiency, carbon balance and partitioning," *Forest Ecol. Manage.*, vol. 95, pp. 209–228, 1997.
[52] T. Hwang, C. Song, P. V. Bolstad, and L. E. Band, "Downscaling real-time vegetation dynamics by fusing multi-temporal MODIS and landsat NDVI in topographically complex terrain," *Remote Sens. Environ.*, vol. 115, no. 10, pp. 2499–2512, Oct. 2011.
[53] K. Shepherd, C. Palm, C. Gachengo, and B. Vanlauwe, "Rapid characterization of organic resource quality for soil and livestock management in tropical agroecosystems using near-infrared spectroscopy," *Agronomy J.*, vol. 95, pp. 1314–1322, 2003.
[54] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, pp. 5–32, 2001.
[55] Q. Zhao, F. Wang, J. Zhao, J. Zhou, S. Yu, and Z. Zhao, "Estimating forest canopy cover in black locust (Robinia pseudoacacia L.) Plantations on the loess plateau using random forest," *Forests*, vol. 9, no. 10, 2018.
[56] H. S. J. Zald *et al.*, "Integrating landsat pixel composites and change metrics with LiDAR plots to predictively map forest structure and aboveground biomass in Saskatchewan, Canada," *Remote Sens. Environ.*, vol. 176, pp. 188–201, Apr. 2016.
[57] L. Korhonen, J. Heiskanen, and I. Korpela, "Modelling LiDAR -derived boreal forest canopy cover with SPOT 4 HRVIR data," *Int. J. Remote Sens.*, vol. 34, no. 22, pp. 8172–8181, 2013.
[58] A. M. S. Smith, M. J. Falkowski, A. T. Hudak, J. S. Evans, A. P. Robinson, and C. M. Steele, "A cross-comparison of field, spectral, and LiDAR estimates of forest canopy cover," *Can. J. Remote Sens.*, vol. 35, no. 5, pp. 447–459, 2014.
[59] J. O. Sexton *et al.*, "Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with LiDAR -based estimates of error," *Int. J. Digit. Earth*, vol. 6, no. 5, pp. 427–448, 2013.
[60] C. M. DiMiceli, M. L. Carroll, R. A. Sohlberg, C. Huang, M. C. Hansen, and J. R. G. Townshend, *Annual Global Automated MODIS Vegetation Continuous Fields (MOD44B) at 250m Spatial Resolution For Data Years Beginning Day 65,2000–2010 Collection 5 Percent Tree Cover*, College Park, MD, USA: Univ. Maryland, 2011.
[61] M. Rautiainen and P. Lukeš, "Spectral contribution of understory to forest reflectance in a boreal site: An analysis of EO-1 hyperion data," *Remote Sens. Environ.*, vol. 171, pp. 98–104, 2015.
[62] M. Reichstein *et al.*, "Deep learning and process understanding for data-driven earth system science," *Nature*, vol. 566, no. 7743, pp. 195–204, Feb 2019.
[63] C. Yoo, D. Han, J. Im, and B. Bechtel, "Comparison between convolutional neural networks and random forest for local climate zone classification in mega urban areas using landsat images," *ISPRS J. Photogramm. Remote Sens.*, vol. 157, pp. 155–170, Nov. 2019.
[64] B. Combal *et al.*, "Retrieval of canopy biophysical variables from bidirectional reflectance using prior information to solve the ill-posed inverse problem," *Remote Sens. Environ.*, vol. 84, pp. 1–15, 2002.
[65] B. Koetz *et al.*, "Fusion of imaging spectrometer and LIDAR data over combined radiative transfer models for forest canopy characterization," *Remote Sens. Environ.*, vol. 106, no. 4, pp. 449–459, 2007.
[66] G. Mountrakis, J. Im, and C. Ogole, "Support vector machines in remote sensing: A review," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 3, pp. 247–259, May 2011.
[67] D. J. Lary, A. H. Alavi, A. H. Gandomi, and A. L. Walker, "Machine learning in geosciences and remote sensing," *Geosci. Front.*, vol. 7, no. 1, pp. 3–10, Jan. 2016.
[68] B. Ma *et al.*, "Data augmentation in microscopic images for material data mining," *NPJ Comput. Mater.*, vol. 6, no. 1, 2020.
[69] B. Siegmann and T. Jarmer, "Comparison of different regression models and validation techniques for the assessment of wheat leaf area index from hyperspectral data," *Int. J. Remote Sens.*, vol. 36, no. 18, pp. 4519–4534, 2015.
[70] L. Gao *et al.*, "Remote sensing algorithms for estimation of fractional vegetation cover using pure vegetation index values: A review," *ISPRS J. Photogramm. Remote Sens.*, vol. 159, pp. 364–377, 2020.
[71] A. Gastón *et al.*, "Species' habitat use inferred from environmental variables at multiple scales: How much we gain from high-resolution vegetation data?," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 55, pp. 1–8, 2017.
[72] K. Yan *et al.*, "Evaluation of the vegetation-index-based dimidiate pixel model for fractional vegetation cover estimation," *IEEE Trans. Geosci. Remote Sens.*, to be published, doi: 10.1109/tgrs.2020.3048493.
[73] K. Yan *et al.*, "Evaluation of MODIS LAI/FPAR product collection 6. Part 2: Validation and intercomparison," *Remote Sens.*, vol. 8, no. 6, pp. 1–14, 2016.
[74] K. Yan *et al.*, "Performance stability of the MODIS and VIIRS LAI algorithms inferred from analysis of long time series of products," *Remote Sens. Environ.*, vol. 260, p. 112438, 2021.

Decai Jin received the B.S. degree in forestry from the Centre South University Forestry and Technology, Changsha, China, in 2018. He is currently working toward the M.S. degree in forest management from Beijing Forestry University, Beijing, China.

His research interests include the extraction of vegetation biophysical parameters, radiative transfer modeling, deep learning.

Jianbo Qi received the Ph.D. degree in cartography and geographical information system from Beijing Normal University, Beijing, China, in 2019.

He was a joint-Ph.D. student with Paul Sabatier University, Toulouse, France, from 2016 to 2018. He is currently an Assistant Professor with Beijing Forestry University. His research interests include 3-D radiative transfer modeling, realistic forest scene simulation, and vegetation parameter retrieval.

Huaguo Huang received the B.Sc. degree in forestry and the M.Sc. degree in forest management from Beijing Forestry University, Beijing, China, in 2001 and 2004, respectively, and the Ph.D. degree in quantitative remote sensing from the Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing, China, in 2007.

Since 2007, he has been with Forestry College, Beijing Forestry University, where he is currently a Full Professor. His research interests include 3-D modeling of reflectance, thermal emission, LiDAR and microwave scattering of land surfaces (vegetation in particular), and remote sensing applications in agriculture and forest environment.

Linyuan Li received the B.S. degree in remote sensing science and technology from the Chang'an University, Xi'an, China, in 2014, and the Ph.D. degree in cartography and geographical information system from Beijing Normal University, Beijing, China.

He was a Visiting Scholar with the EMMAH, National Institute of Agronomic Research, Avignon, France. He is currently an Assistant Professor with Beijing Forestry University. His research interests include the extraction of vegetation biophysical parameters (e.g., fraction of vegetation cover and leaf area index) and physiological parameters (e.g., crop water content) from UAV data.