Winter Wheat Yield Estimation from Multitemporal Remote Sensing Images based on Convolutional Neural Networks

4 authors:

Haowei Mu
Lanzhou Jiaotong University
9 PUBLICATIONS   **15** CITATIONS

SEE PROFILE

Xuewei Dang
Lanzhou Jiaotong University
8 PUBLICATIONS   **47** CITATIONS

SEE PROFILE

Liang Zhou
Nanjing univeristiy
40 PUBLICATIONS   **288** CITATIONS

SEE PROFILE

Bo Yuan
Lanzhou Jiaotong University
3 PUBLICATIONS   **6** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Urban planning View project

enviromental geography View project

Winter Wheat Yield Estimation from Multi-temporal Remote Sensing Images Based on Convolutional Neural Networks

Haowei Mu, Liang Zhou*, Xuewei Dang, Bo Yuan
Faculty of Geomatics,Lanzhou Jiaotong University
National-Local Joint Engineering Research Center of Technologies and Applications for National Geographic State Monitoring
Gansu Provincial Engineering Laboratory for National Geographic State Monitoring
Lanzhou, China
zhougeo@126.com

Abstract—The development of deep learning and big data technology has introduced information and intelligent techniques to agricultural remote sensing estimation. The deep learning methods represented by Convolutional Neural Network (CNN) have abilities to extract the depth-dependent features of crop growth. In the field of crop yield estimation, the core challenge is to utilize CNN to extract the relevant information from remote sensing image. In this paper, we apply histogram dimensionality reduction and time series fusion to generate the input layer of CNN. In view of the data characteristics, the CNN network structure was designed to extract the features of winter wheat growth from multi-temporal MODIS images for yield estimation in North China. The results showed that the estimated yield of winter wheat based on time-series remote sensing images is highly correlated with statistical data, with Pearson's r of 0.82, RMSE of 724.72 kg·hm^{-2}. In the case of sufficient statistical data, the provincial model performs better. CNN is able to mine more relevant information and has higher robustness. It also provides a technical reference for estimating large-scale crop yield.

Keywords—Winter Wheat, Yield Estimation, CNN, North China

I. INTRODUCTION

Food is the basis of human survival and development. In recent years, tremendous efforts are needed to end hunger, achieve food security and improved nutrition and promote sustainable agriculture[1]. The key to these efforts is to timely and accurately estimate crop yields to regulate food market safety.

In terms of agricultural remote sensing, yield estimation is always a core problem. New advances in satellite data acquisition and processing offer promise for monitoring agricultural lands globally[2]. Machine learning (ML) techniques have been utilized for crop monitoring and yield estimation/prediction using remotely sensed data[3]. With the development of artificial intelligence and big data technology, agricultural remote sensing applications will be intelligent and informationized upgrades. Nowadays, Deep-learning (DL) algorithms, which learn the representative and discriminative features hierarchically from the data, have recently become a hotspot in the machine-learning area and have been introduced into the geoscience and remote sensing community for remote sensing big data analysis[4]. Compared with traditional methods, the neural network (NN) is more efficient and accurate than Multiple Linear Regression (MLR)[5] and support vector regression (SVR)[6]. It is capable of extracting important features for crop yields from the input data. Many scholars have proposed feasible and effective NN structures to achieve excellent performance in large-scale crop yield estimation. Such as Multiple Restricted Boltzmann Machine (MRBM)[6], Spiking Neural Network (SNN)[7], Deep Gaussian Process (DGP)[8], Deep Transfer Learning (DTL)[9], etc. In especial, You et al.[8] adds a DGP to Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM) network, and outperforming all the competing approaches. At the same time, a novel dimension reduction technique is introduced, which enables training of NNs and automatically learns useful features even when labeled training data are scarce. In the demand of remote sensing applications, it is found that deep learning has great application potential in remote sensing data analysis. In the field of DL, CNN has become one of the state-of-the-art algorithms. To timely and efficiently estimate crop yields, CNN is potential to mine insight into the underlying mechanisms that drive real-world phenomena such as the growth of different crops[10]. It is very promising since it can extract high-level semantic information from the basic features of the original pixels based on remote sensing data.

In this paper, Fig 1 shows North China (110.22~122.71°E, 31.39~42.62°N) is selected as the research area. Mainly including Henan, Hebei, Shanxi and Shandong provinces. Most areas of Shanxi and northern Hebei are removed by cleaning the statistical data. The main reason is that the altitude causes the winter wheat to be planted less, and secondly some areas with abnormal statistics are deleted.
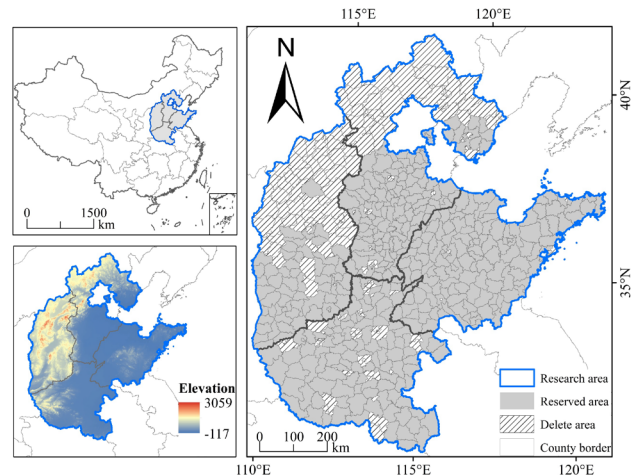

Fig. 1.   Study area description

II. Data Sets

A. MODIS Images

MODIS data has high temporal resolution and can acquire multi-temporal images that are sufficient to monitor crop yields. In this paper, 19 estimation indexes related to winter wheat growth were selected from 6 different MODIS products. The selected indexes include:

1) surface reflectance data of 7 different wavelengths in MOD09A1.

2) the surface temperature of day and night in MYD11A2.

3) NDVI and EVI included in MOD13A1 and MYD13A1.

4) Fraction of Photosynthetically Active Radiation (FPAR) and Leaf Area Index (LAI) in MOD15A2H.

5) Total Evapotranspiration (ET), Average Latent Heat Flux (LE), Total Potential Evapotranspiration (PET) and Average Potential Latent Heat Flux (PLE) in MOD16A2.

6) Gross Primary Productivity (GPP) and Net Photosynthesis (PsnNet) in MOD17A2H.

B. Data Pre-processing

According to the growing season of winter wheat, the data time range is from the 273rd day of the year to the 185th day of the next year. The image data ranging from 2006 to 2016 are selected. Therefore, The total experimental MODIS images amount to 6376, and are processed by Shell and GDAL. The MOD13A1 and MYD13A1 images were fused in time series to ensure the integrity of the time. The MOD15A2H image was missing on the 49th day of 2016, so it was replaced by averaging the images on days 41 and 57. At the same time, the grain classification data in MOD12Q1 is used to mask non-cereal pixels. All MODIS data is converted into a UTM projection system based on WGS-84 ellipsoid. Multiple types of images are resampled, spliced and cropped, and finally fused into 163,944 images that contain 19 bands.

TABLE I.   DIMENSION REDUCTION RANGE FOR EACH BAND

Band	Time (d)	Resolution (m)	Scale	Dimension-reduce range
sur_refl1 (620-670 nm)				
sur_refl2 (841-876 nm)				
sur_refl3 (459-479 nm)				
sur_refl4 (545-565 nm)	8	500	0.0001	1,5000
sur_refl5 (1230-1250 nm)				
sur_refl6 (1628-1652 nm)				
sur_refl7 (2105-2155 nm)				
LST_Day	8	1000	0.02	13000,16000
LST_Night				13000,15000
ET	8	500	0.1	1,300
PET				1,800
LE			10000	1,800
PLE				1,2000
GPP	8	500	0.0001	1,200
PsnNet				1,350
NDVI	16	500	0.0001	1,9000
EVI				1,9000
FPAR	8	500	0.01	1,75
LAI			0.1	1,30

III. Method

A. Histogram Extraction

In this paper, we assume that the winter wheat yield is not related to the positions of the image pixels. Based on this assumption, the histogram is utilized as an image feature since it contains no spatial information. It should be noted that the growth of winter wheat is only closely reflected by the partial value range of each selected index. Therefore, the histograms - which contain 36 bins - are generated and normalized within the partial ranges of the selected indexes as shown in Table 1.

B. Sample Generation

Taking the county of Henan Province as an example to illustrate the sample generation process, each county produces a sample in a growing season. Each index shown in Figure 2 is synthesized into a 36×36 matrix in the time dimension. One of the 36 indicates the characteristics of the histogram extraction, and the other 36 indicates the growth season of 36 days (with an interval of 8 days). Finally, the 36×36×19 three-dimensional matrix generated by the 19 bands serves as the input layer of CNN. The output layer employs the corresponding winter wheat yield statistics (in kilograms per hectare) to generate a CNN sample.
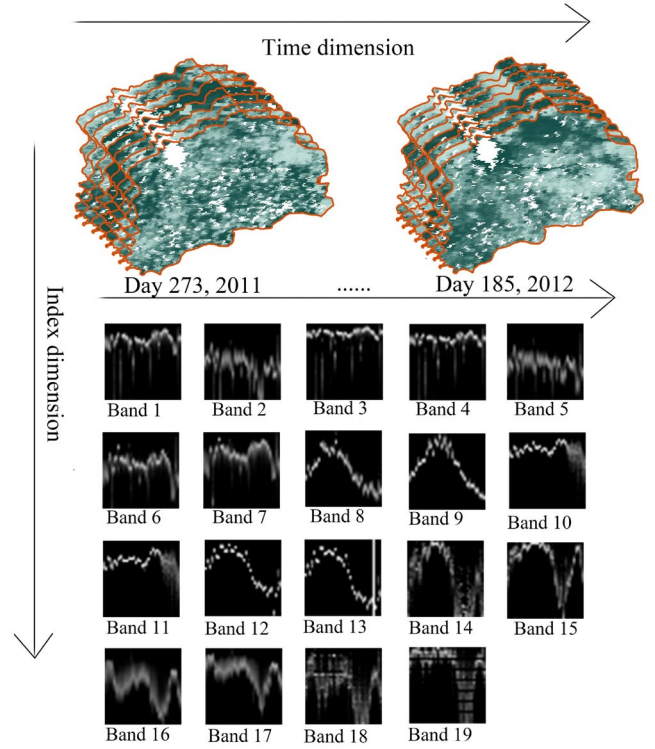

Fig. 2.   Sample generation process

C. CNN-Based Regression

The applied convolutional neural network is designed according to the characteristics of the generated samples, and the small convolution kernel is used. The multi-layer stacking of the convolution kernel reduces the number of parameters, increases the network depth and enhances the network capacity and complexity, and thus fits the complex process of crop growth.

The neural network is composed of the input layer, 7 convolution layers, 7 activation layers, 7 batch normalization layers, 3 dropout layers, 2 full connection layers and output layer as shown in Fig.3. Considering the facts that the batch

normalization algorithm is able to improve the generalization ability of the model, and too many dropout layers will reduce the prediction ability of the model, the dropout layer is used after all the batch normalization layers[11]. Moreover, in order to address the over-fitting problem of the model, the L2 parameter regularization is added to improve the generalization ability of the model.
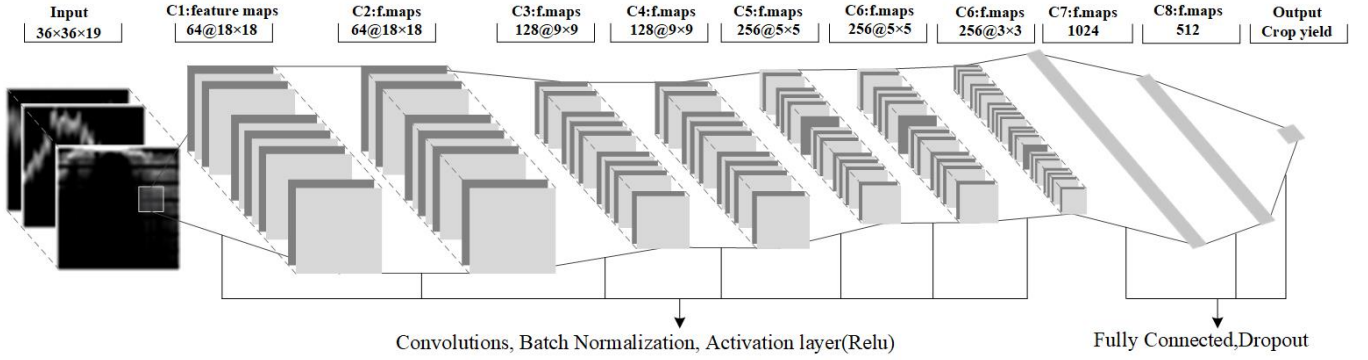

Fig. 3. Regression-based convolutional neural network structure

IV. EXPERIMENTAL RESULT

The algorithm in this paper is implemented by TensorFlow which is an end-to-end open-source platform for machine learning. Model accelerated training through GPU. The experiments are generated on a PC equipped with an Intel Core i7 with 3.7 GHz and Nvidia GeForce GTX 1080Ti graphics card.

A. Sample Description

The availability of statistical data is crucial for the selection of the experimental region. Therefore, due to the unavailability of the statistical data of Shandong in 2008 and after 2010, they are not used to construct experimental samples. Through data organization and cleaning, Henan,

Hebei, Shanxi and Shandong provinces retained 105, 116, 60 and 132 counties respectively. Therefore, totally 3,619 samples were constructed.

B. Result of Model

Taking the year 2016 as an example(Shandong Province is 2010), the spatial distribution of the model error in the training set is shown in Fig.4.(a). The error of the whole training set is less than 1000 kg·hm^{-2}, mainly distributed in Shanxi and Hebei provinces. The spatial distribution of the model error in the validation set is shown in Fig.4.(b). Each province produced outliers above 1000 kg·hm^{-2}, but most of the errors were less than 1000 kg·hm^{-2}, The error is randomly distributed, Shandong and Shanxi have more outliers, less in Hebei.

(a)Training Set Error Spatial Distribution

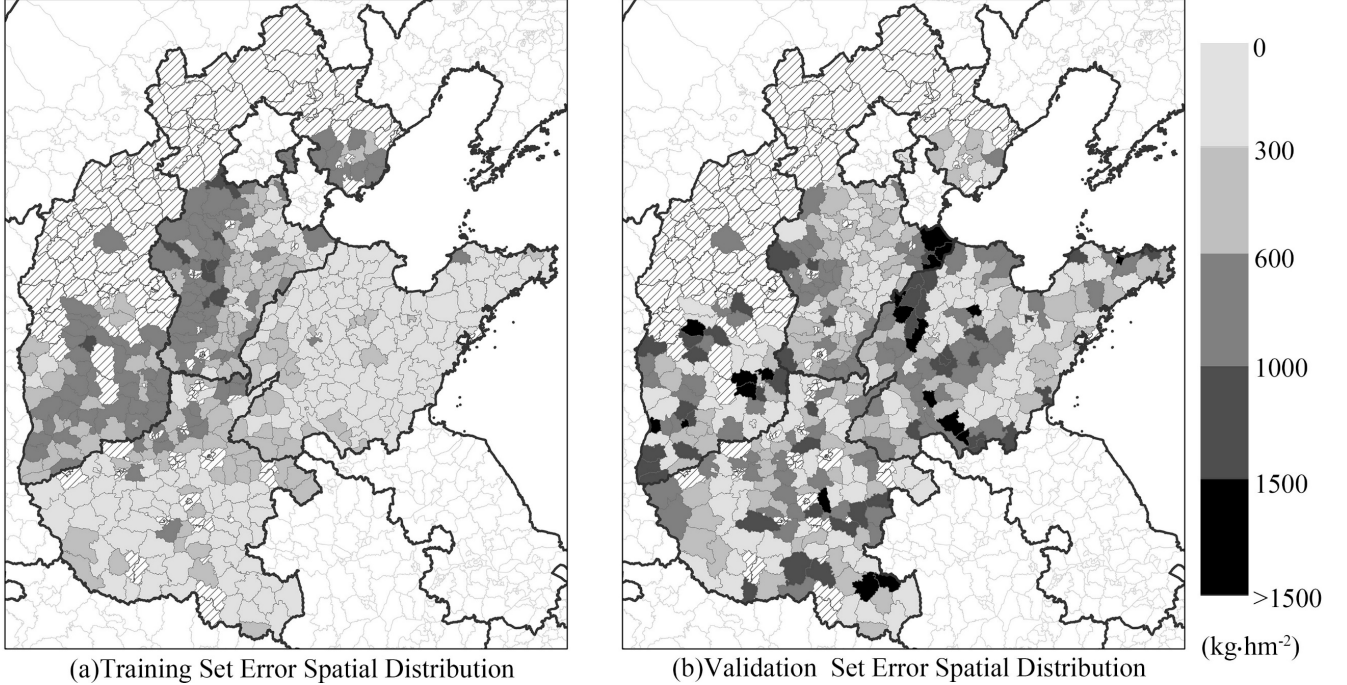(b)Validation Set Error Spatial Distribution

Fig. 4. Model error distribution

The results in the training set are shown in Table 2. It achieved good performances in different areas. The model performance of the validation set is shown in Table 3. Among them, the model achieved the best performance in Henan and Hebei, since there are sufficient samples in the two provinces. Comparatively, the estimation results of Shandong and Shanxi are unsatisfactory, but there are better performances in the training set than in the other two provinces, indicating that the models in the two regions have been over-fitting. The reason may be that Shanxi and Shandong have fewer samples, and CNN is more dependent on the number of samples. As a result, the experimental results prove that the samples generated by

histogram extraction and time fusion can be processed by CNN.

TABLE II. MODEL ERROR IN TRAINING SET

	Henan	Hebei	Shandong	Shanxi	North China
Pearson's r	0.98	0.93	0.94	0.95	0.93
RMSE	242.39	425.21	300.19	534.99	386.46

TABLE III. MODEL ERROR IN VALIDATION SET

	Henan	Hebei	Shandong	Shanxi	North China
Pearson's r	0.87	0.78	0.56	0.69	0.82
RMSE	657.93	586.80	758.14	963.05	724.72

C. Robustness Verification

Training different models based on the neural network structure proposed in this paper without artificially setting hyperparameters. The important hyperparameters are as follows: learning rate is 0.001, weight decay is 0.005, and dorp out is 0.25. Verify the robustness of the model on conditions that guarantee the objectivity of the experiment. The training model is selected separately for five years. The Pearson's r of the validation set is higher and the RMSE is around 1000 kg·hm^{-2}. The method of this paper has high robustness and can be used crop yield estimation in the future.

TABLE IV. MODEL ROBUSTNESS VERIFICATION

	2006	2008	2010	2012	2014
Pearson's r	0.55	0.71	0.67	0.76	0.76
RMSE	1081.78	992.38	1074.49	938.52	945.89

V. CONCLUSION AND DISCUSS

The convolutional neural network model can effectively extract the features related to winter wheat yield from remote sensing images. It deals with the constraints of the traditional statistical model on the fitting of complex relationships and does not require field sample collection. This end-to-end model has the advantage of synergism. After sufficient and proper training, the model is able to efficiently predict future winter wheat yield in a real-time manner.

In this paper, the method of histogram extraction is used for dimensionality reduction of remote sensing images, which has a large loss of information. In the process of sample generation, outliers are produced in the time dimension. Therefore, it is possible to consider a more advanced algorithm for dimensionality reduction and time series smoothing in future research.

ACKNOWLEDGMENT

REFERENCES

[1] United Nations. "Transforming our world: The 2030 agenda for sustainable development," General Assembley 70 session, 2015.

[2] DB. Lobell, D. Thau, C. Seifert, E. Engle, and B. Little, "A scalable satellite-based crop yield mapper," Remote Sensing of Environment, 164, 2015, pp. 324-333.

[3] H. Aghighi, M. Azadbakht, D. Ashourloo, HS. Shahrabi, and S. Radiom, "Machine Learning Regression Techniques for the Silage Maize Yield Prediction Using Time-Series Images of Landsat 8 OLI," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11(12), 2018, pp. 4563-4577.

[4] L. Zhang, L. Zhang, and B. Du. "Deep learning for remote sensing data: A technical tutorial on the state of the art," IEEE Geoscience and Remote Sensing Magazine 4.2, 2016, pp. 22-40.

[5] B. Ji, Y. Sun, S. Yang, and J. Wan. "Artificial neural networks for rice yield prediction in mountainous regions," The Journal of Agricultural Science, 145(3), 2007, pp. 249−261.

[6] K. Kuwata, R. Shibasaki. "Estimating crop yields with deep learning and remotely sensed data," Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International. IEEE, 2015, pp. 858−861.

[7] P. Bose, N.K. Kasabov, L. Bruzzone, R.N. Hartono. "Spiking neural networks for crop yield estimation based on spatiotemporal analysis of image time series," IEEE Transactions on Geoscience and Remote Sensing, 54.11, 2016, pp. 6563-6573.

[8] J. You, X. Li, M. Low, D. Lobell, S. Ermon. "Deep Gaussian Process for Crop Yield Prediction Based on Remote Sensing Data," AAAI, 2017, pp. 4559−4566.

[9] A.X. Wang, C. Tran, N. Desai, D. Lobell, and S. Ermon, "Deep Transfer Learning for Crop Yield Prediction with Remote Sensing Data," Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies. ACM, 2018, pp. 50.

[10] M. Sabini, G. Rusak, B. Ross, "Understanding Satellite-Imagery-Based Crop Yield Predictions," Stanford, 2017.

[11] S Ioffe, C Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," arXiv preprint arXiv:1502.03167, 2015.