IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 30, 2020, accepted May 7, 2020, date of publication May 22, 2020, date of current version June 11, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2996626

Computation of Atmospheric Optical Parameters Based on Deep Neural Network and PCA

CHEN CHUAN[1], CHENG ZHENG[1], LIU BO[2], AND LI LIGANG[2]

[1]Institute of Computer Application, China Academy of Engineering Physics, Mianyang 621900, China
[2]Key Laboratory of Electronics and Information Technology for Space System, National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China

Corresponding author: Li Ligang (liligang@nssc.ac.cn)

This work was supported in part by the China Academy of Engineering Physics Innovation Development Fund (Grant PY2019131) and in part by the Key Deployment Projects of the Chinese Academy of Science (Grant KGFZD -125-15-014).

ABSTRACT The effect of the atmosphere on the propagating energy during the remote sensing imaging simulation is one of the most critical factors affecting the quality of image. The classical methods retrieving atmospheric optical parameters (AOP) have shortcomings in GPU-based imaging simulation application. This paper proposed a method based on deep neural networks (DNN) and principal component analysis (PCA) to compute AOP. Firstly, MODTRAN is employed to obtain large numbers of AOP as original spectrum set in different weather and observation geometry conditions. Then the dimension of original spectrum is reduced by PCA. Next, a DNN is constructed and trained using compressed spectral signatures. Finally, estimated AOP are obtained through inverse PCA by decompressing the output of DNN. The results show that original AOP and estimated AOP have a high spectral similarity which relative error is less than 2%. Compared with the classical methods, DNN can be used to accurately and fast compute AOP with any kind of conditions in remote sensing imaging applications, without consuming large of graphic memory.

INDEX TERMS Atmospheric optical parameters, imaging simulation, deep neural networks, principal component analysis, atmospheric downward radiation.

I. INTRODUCTION

Remote sensing imaging simulation based on a 3D scene (RSIS-3D) uses computer simulation technology to simulate the imaging process of remote sensor, it can accurately analyze factors and mechanism that affecting the quality of image in the process of imaging. RSIS-3D owns significant application value in sensor design, remote sensing data processing, image quality evaluation and target characteristic research [1]–[5].

Compared with classical image synthesis methods [6], [7], RSIS-3D addresses complex 3D scene and requires a large amount of calculation. GPU based ray tracing method is employed to compute the at-sensor radiation of remote sensor parallel [8]. Several kinds of parameters of atmosphere will be involved in calculating of the at-sensor radiation, which are termed as atmospheric optical parameter (AOP), including transmittance, sky diffuse, path radiation and aerosol optical thickness. These parameters are one of the key factors affecting the quality of remote sensing imaging simulation. Besides, these parameters are also very important for many other remote sensing applications, e.g. atmospheric correction, satellite instrument calibration, earth surface investigation.

The best way to obtain accurate AOP is using observation instruments to measure the atmosphere synchronously. Hu measured solar radiation through atmosphere by CE318 sun-photometer, at the same time as MODIS imaging to investigate the atmospheric aerosol optical properties over Lake Taihu [9]. Yet this method is lack of flexibility, and it is very time-consuming and high cost. The most widely used method is to employ radiation transmission software, e.g. MODTRAN, to compute AOP according to weather condition and observation conditions. The Digital Imaging and Remote Sensing Image Generation (DIRSIG) used the MODTRAN radiation propagation model to calculate exo-atmospheric irradiance, emitted and scattered radiances and path transmission predictions, supporting the generation of real-time infrared scene [10]. This mothed works well in many other remote sensing applications [11]–[13], however it is unsuitable for RSIS-3D as the radiation transmission software is not compatible for GPU platform. The look-up table (LUT) method is another commonly used way to provide AOP for some remote sensing applications.

The associate editor coordinating the review of this manuscript and approving it for publication was Ah Hwee Tan.

Darvishzadeh *et al.* [14] generated several LUTs in different size to inverter the canopy's chlorophyll content and studied the effect of the LUT size on the retrieval accuracy by using analysis of variance. But LUT-based method is also restricted by the limitation of the sampling resolution of LUT and interpolation method [15], [16]. it will reduce computation efficiency of RSIS-3D as its large amount of graphics memory cost.

The above three typical methods have their own advantages and shortcomings for RSIS-3D on different views, as descripted in the table below.

Atmosphere acts different optical features under different meteorological conditions and observational geometries. The transmission of light in the atmosphere can be seen as a complex black box system and can be expressed as a composite function which mapping meteorological conditions and observational geometries to AOP spectral signature. DNN is able to approximate any continuous composite function mapping one finite dimensional space to another finite dimensional space with arbitrary precision if the DNN satisfies the following two conditions: 1) the output layer of the DNN adopts a linear activation function, 2) having at least one hidden layer and the hidden layer has a ''compressed'' property. This is so called the universal approximation theorem of DNN, which indicates that DNN can be used to represent any composite function [17], [18]. Therefore, DNN can be used to imitate the complex black box system of atmosphere and precisely calculate AOP.

In this study we propose a method that using DNN to calculate these AOP. Depending on remote sensing imaging model, a large number of AOP spectrum are computed by MODTRAN under different weather conditions and observation geometries. Consideration the high dimensions of the spectral signature, PCA is employed to reduction the dimension of spectrum and the compressed spectrum are treated as training and test sample of constructed DNN, and the nonlinear relationships between weather conditions, observation geometries and AOP are described by the well-trained DNN.

II. METHODOLOGY

A. ATMOSPHERIC OPTICAL PARAMETERS USED IN RSIS-3D

A part of radiation emitted from solar is scattered back to space and the other goes through the atmosphere and reach at the earth, and reflected by earth surface and goes through the atmosphere again, which carrying on the characteristic information of atmosphere and surface, then been captured by satellite sensor [8]. The received energy of remote sensor operating in visible and near infrared bands $L(\lambda)$ is:

$$\begin{aligned} L(\lambda) = PSF_\lambda((E_s^0(\lambda) * \tau_{down} * BRDF(\theta_s, \theta_v, \varphi, \lambda) \\ + \int_{\varphi=0}^{2\pi} \int_{\theta=0}^{\pi/2} L_{down}(\theta, \varphi, \lambda) \frac{BRDF(\theta, \theta_v, \varphi, \lambda)}{\pi} d\theta d\varphi) * \tau_{up} \\ + L_{path}(\theta_v, \varphi_v, \lambda)) \end{aligned} \quad (1)$$

where $PSF_\lambda(x)$ is point spread function which are used to characterize the spatial modulation effect of atmosphere, λ means there variables are wavelength dependent, $BRDF(\theta_i, \theta_o, \varphi)$ is bidirectional reflectance distribution function of earth surface and the other variables are called atmospheric optical parameters, they are:

1. $E_s^0(\lambda)$: the spectral irradiance of solar radiation outside the atmosphere, is mainly affected by the distance between the solar and the earth which depends on the date of observation.
2. $\tau_{down}(\lambda)$: The spectral transmittance of the atmosphere along the incident direction of sun light from top of atmosphere to earth surface, is mainly affected by atmospheric mode, aerosol mode, visibility, target altitude and solar zenith angle.
3. $L_{down}(\theta, \varphi, \lambda)$: The downward spectral radiance from a sampling direction with zenith θ and azimuth φ in the 2π space above the target surface, also called sky light, is mainly affected by observation date, atmospheric mode, aerosol mode, visibility, target altitude, solar zenith angle, incident zenith θ and relative azimuth φ between solar and sampling direction.
4. $\tau_{up}(\theta_V, \varphi_V, \lambda)$ and $L_{path}(\theta_V, \varphi_V, \lambda)$: The upward spectral transmittance and path radiation from earth surface to satellite sensor, are mainly affected by atmospheric mode, aerosol mode, visibility, target altitude, and satellite observation zenith angle. In addition, path radiation also affected by observation date, the solar zenith and relative azimuth between the sunlight and line of sight of remote sensor.
5. $\tau_{mol}(\lambda)$ and $\tau_{aer}(\lambda)$: The optical thickness of atmospheric molecules of very layering of atmosphere in a stratified atmospheric model, are mainly affected by atmospheric mode, aerosol mode and visibility, which are used to characterize the spatial modulation effect of atmosphere in $PSF_\lambda(x)$

In RSIS-3D, these AOP are the most critical factors during the calculation of at-sensor radiation as the complex components and instabilities of atmosphere. To accurately and fast acquire atmospheric radiation transmittance feature is the key to ensure the accuracy of at-sensor radiation.

Among these five class of AOP, atmospheric downward radiation, $L_{down}(\theta, \varphi, \lambda)$ is the most instable as it has 8 influencing factors and the most complicated mechanism. Therefore, this paper takes downward radiation as an example to demonstrate how to construct and train a DNN to calculate AOP.

B. SPECTRUM REDUCTION DIMENSION WITH PCA

Principal component analysis (PCA), a widely used method for feature extraction and data analyses, can be employed to reduce the dimension of spectral signature without losing too much information. The goal of PCA is to condense the original spectrum set into a feature space via linear transformation so that the new compressed spectrum signature has no

correction between each other. The compressed spectrum sets are in a feature space which is defined by the eigenvectors of covariance matrix C of the original spectrum sets. The covariance matrix C, $C = E\left[(x-\mu)(x-\mu)^T\right]$, with x as the spectral signature and $\mu = E(x)$ as the mathematical expectation of the spectrum sets, is practice calculations replaced by an estimated $\hat{C}$:

$$\hat{C} = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \hat{\mu}\right)\left(x_i - \hat{\mu}\right)^T \tag{2}$$

where x is a n-dimension spectral signature and $\hat{\mu}$ is the estimated mean vector over n sample spectral signatures.

The eigenvectors $u_1, u_2 \ldots u_n$ and corresponding eigenvalues $\lambda_1, \lambda_2 \ldots \lambda_n$, ordered in decreasing order, are calculated from covariance matrix $\hat{C}$. The eigenvalues describe the amount of information recovered by the corresponding eigenvector and give the components in order of significance. In fact, it turns out that the eigenvector with the highest eigenvalue is the principle component of the data set [20]. The first principal component u_1 contains maximum information as its corresponding eigenvalue λ_1 is the maximum, the second principal component u_2 contains about the following maximum information as its corresponding eigenvalue λ_2 is the 2nd-maximum, etc. Information loss decreases by a step to the next.

The first k eigenvectors are chosen as a new spectral bases if their contained information are great than a threshold p defined by the user, such as 99%, and components with lesser significance are ignored.

$$p = \frac{\sum_{i=1}^{k}\lambda_i}{\sum_{i=1}^{n}\lambda_i} \tag{3}$$

The number of chosen eigenvector k is clearly smaller than original spectrum dimension n. Even a small value of $k = 4 \ldots 10$ is enough for high quality reconstruction of original spectral sets [20]. And the lossy information defined as reconstruction error J_e is:

$$J_e = 1 - p = \frac{\sum_{i=k+1}^{n}\lambda_i}{\sum_{i=1}^{n}\lambda_i} \tag{4}$$

For compression purposes, eigenvectors $u_i, i = 1 \ldots k$, and matrix $x^T u_i$ are needed. To decompress, the estimation of the original spectral signature is received from:

$$x^* = \sum_{i=1}^{p}\left(x^T u_i\right)u_i \tag{5}$$

C. DEEP NEURAL NETWORK

Deep learning methods can be described as a common pattern: specific training and test data sets, cost functions, optimization processes and model structures. Deep neural network is a high-efficiency and classical model of deep learning. This paper employees DNN as learning model to estimate AOP.

Key points of constructing DNN as deep learning model include preparing and preprocessing training data, determining network structure and training the network, while network structure determined by depth and width of the network and weight and biases of each neural, and the network training process includes determining cost function and training algorithm.

1) TRAINING DATA AND PREPROCESSING

It is unrealistic to measure the actual downward radiation as training samples of DNN under different visibilities and solar zenith angles. Therefore, MODTRAN is employed to calculate the downward radiation under different conditions, which are training and testing samples of DNN.

In atmospheric radiation transfer model, there are 8 main factors affecting the downward radiation of the atmosphere above the target hemisphere space, including atmospheric mode, aerosol mode, observation date, visibility, target altitude, solar zenith angle, sampling zenith angle θ and sampling relative azimuth φ. Among these parameters, atmospheric mode and aerosol mode are discrete enumeration type and the others are continuous.

More input variables in DNN, more detail characteristics can be learned, and more difficulty to train it as its nonlinear increasing complexity. In this research, four most important factors, visibility, solar zenith angle, sampling zenith angle and sampling relative azimuth, are selected as input variables. Target is assumed at sea level which means target altitude is 0. For different atmospheric mode, aerosol mode and date, the same method can be employed to train networks for different combinations.

Weather condition is mainly characterized by visibility, which is very good when visibility is 23KM while it is poor when visibility is 5KM. Therefore, visibility ranges from 5Km to 23Km with an interval of 1km, for a total of 19 different visibility are selected as input variable.

For a visible and near infrared remote sensor, it is not suitable for earth imaging when solar zenith is greater than 60°. Therefore, solar zenith ranges from 0° to 60° with an interval of 5°, for a total of 13 different solar zeniths are selected as input variable.

Incident radiation caused by skylight can be ignored when sampling zenith angle is close to 90°, and MODTRAN would crash when encountering a situation of $\tan 90°$. Therefore, sampling zenith angle ranges from 0° to 85° with an interval of 5°, for a total of 18 different sampling zenith.

Atmospheric down-radiation is symmetrically distributed with main plane of solar. Therefore, samples relative azimuth ranges from 0° to 180° with an interval of 10°, for a total of 19 different relative azimuth.

Atmospheric down-radiation is wavelength-depended. For visible and near infrared, wavelength ranges from 400nm to 1000nm with an interval of 5nm. Therefore, a radiation

spectral signature with 121 bands will obtained as output of DNN.

As above, the four input variables can compose a total of $19 \times 13 \times 18 \times 19 = 84{,}474$ combinations, forming 84,474 different observation conditions. So an 84474*4 matrix and 84474*121 matrix are constructed to present input independent variable and dependent spectral radiation, respectively.

To avoid numerical calculation errors while training DNN preventing the saturation of neuron output caused by excessive net input, and to improve training speed, the input data and output spectrum value are normalized to [0,1], respectively.

2) STRUCTURE OF DNN

It takes two steps to construct a DNN to calculate AOP. Firstly, definition input layer and out layer of DNN. To calculate down-radiation by DNN, the input of DNN is a 4D vector ($X = (x_1, x_2, x_3, x_4)$)is a vector consists of visibility, solar zenith, sampling zenith and samples relative azimuth) and the desired output is a 121-D vector. It is must be a very complex DNN with such a high-dimension output and a huge challenge for training algorithm. Wang *et al.* [19] converted this complex DNN with high-dimension output into many simple DNNs with 1-dimension output and trained these simple DNN respectively, but some questions were introduced during the process of training and application. In this research, we use PCA to tackle this issue. The samples spectral signatures are compressed by PCA to satisfy the ability of DNN and the top k principal components are selected as basis of compressed samples. So, the output of DNN is a k-D vector. Then inverse PCA operation (iPCA) is executed and estimated spectral signature(121D) is obtained. The mean square error (MSE) between estimated and original spectral signature is employed as lost function.

Secondly, definition hidden layer depth and width and their activation function. As of today, there is no theoretical research to guide how to set up number of hidden layer and number of neurons in each hidden layer, depending on the scale and internal relation of the training data, even depending on adopted training algorithm. Empirical method is employed to determine these hidden layer parameters, starting from ''shallow and thin'' to ''deep and fat'' network, training DNN and measuring its performance. If the loss function cannot converge or gap between output and target does not meet expected threshold, increasing the number of neurons and the number of layers and retraining the new network. According to the universal approximation theorem, we employed sigmoid function as activation function of hidden layer.

The constructed DNN can be represented by:

$$\mathbf{z}^l = \sigma\left(\mathbf{W}^l * \mathbf{z}^{l-1} + \mathbf{b}^l\right) \tag{6}$$

where $\mathbf{W}^l$ and $\mathbf{b}^l$ are weights and biases of l^{th} layer, respectively, $\sigma\ (u)$ is sigmoid function serving as activation function and $\mathbf{z}^l$ is output of l^{th} layer, the input of first layer $\mathbf{z}^0$ is input variable x, and output of last layer $\mathbf{z}^L$ is output radiation y. The corresponding DNN can be show as follows:

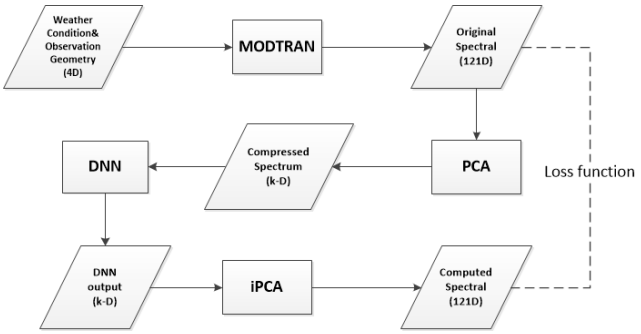


FIGURE 1. Dataflow of DNN-based method to compute AOP.

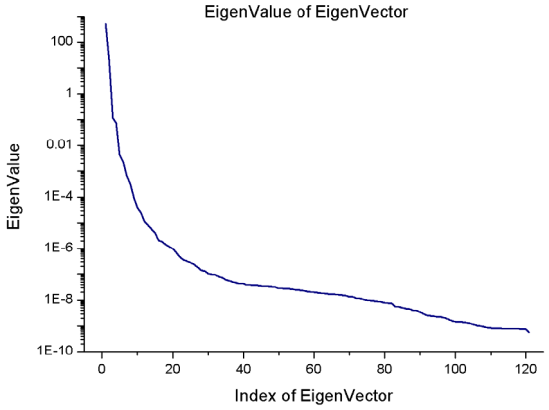


FIGURE 2. The eigenvalues of covariance matrix in decreasing order.

TABLE 1. Advantages and limitations of typical AOP obtaining method.

	Advantages	Limitations
Field Measurement	accurate	inflexible, costly
MODTRAN-based	accurate	Incompatible for GPU platform
LUT-based	offline computed; can be used on GPU parallel	low accuracy as interpolation; graphics memory costly

Besides, a linear output function is employed as the activation function of output layer. And 80% of 84,474 samples are treated as training set and the other 20% are test set.

III. RESULT AND DISCUSSION

A. SPECTRAL COMPRESSION RESULTS WITH PCA

The dimension of spectral signatures calculated by MODTRAN is up to 121 bands. It is difficult to directly train DNN with such high-dimension output data. PCA is employed to reduce the dimension of the original spectral signatures without too much information lossy.

Fig. 2 shows the eigenvalues of covariance matrix of original spectral signature sets in decreasing order and Table. 1 shows the ratio of contribution of several components.

TABLE 2. Ratio of contribution (RoC) of top four components.

	PC1	PC2	PC3	PC4
RoC	96.04%	3.93%	0.022%	0.014%
Accumulative RoC	96.04%	99.97%	99.98%	99.99%

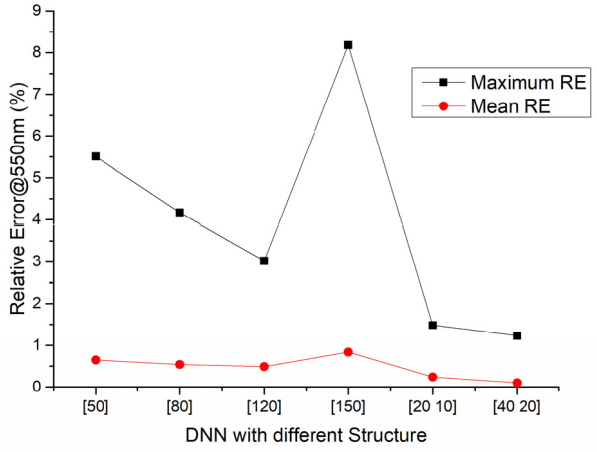


FIGURE 3. Performance of different structural neural network.

The first principal component describes the greater percentage (96.04%), while the second component illustrate about 3.93 of total variance. It can be found that the accumulative variance of the top four components is up to 99.99%. So, it could be concluded that the top four components could almost express the total spectral information.

B. DNN TRAINING RESULTS

To train the constructed DNN, gradient descent (GD) combined algorithm with backpropagation (BP) is employed as training algorithm. The weights and biases of DNN are updated according to the gradient of every mini-batch training samples.

We set the average relative error between DNN output (estimated spectrum) and target (original spectrum) is not greater than 0.1% and the maximum relative error is not greater than 2%. According to this standard, we start from a single hidden layer with 50 neurons, and gradually increases the number of neurons and the number of layers until the network performance reaches the above preset target. Fig. 3 shows relative error of spectral signature at 550nm of DNN with different structure.

The figure shows the relative error of DNN with a single or two hidden layer(s) or different number of neurons for atmospheric down-radiation at 550nm, respectively. It shows that network performance with two hidden layers is significantly better than network with only one hidden layer. The performance with one hidden layer increases with the number of neurons increases from 50 to 120, but it decreases when the hidden layer has 150 neurons. It shows that the more complexity of network, the higher requirements for training algorithm, and adopted gradient descent algorithm cannot treat such a large network. Average error is less than 0.5% while maximum relative error is still greater than 3% when the single hidden layer of DNN has 120 neurons. Comparing with a single hidden layer DNN, performance of two hidden layers DNN improves significantly. Performance of two hidden layers DNN with [20 10] neurons is much better than one hidden layer DNN with 120 neural. Performance of DNN satisfies the preset requirements when neurons of the double hidden layer reach at [40 20]. The result DNN is shown as:

C. COMPUTING AOP UNDER ARBITRARY CONDITIONS

Trained DNN in previous section learn the inherent relationship between input quads (visibility, solar zenith angle, sampling zenith angle and relative azimuth) and atmospheric down-radiation from a large number of spectral samples, and can be used to calculate atmospheric down-radiation instead of MODTRAN. Other 1000 different conditions are randomly generated within the range of 4 input variables and their corresponding downward radiation are also calculated through the trained DNN and compared with downward radiation calculated by MODTRAN. Fig. 5 shows the downward radiation spectrum curves calculated by two methods when visibility is 20km, the solar zenith angle is 32°, the sampling zenith angle is 18° and the relative azimuth is 125°. It shows that the two spectrums, calculated by MODTRAN and DNN respectively, are almost identical.

Spectral similarity scale is employed to measure the difference between two spectrums. Comparing with distance-based or angle-based spectral similarity measurement method, the spectral similarity scale method can simultaneously measure amplitudes and shapes of two spectrums. Spectral similarity value (SSV) representing spectral similarity scale is defined as:

$$SSV = \sqrt{d^2 + \left(1 - r^2\right)^2} \tag{7}$$

where d is the spectrally generalized 2-norm geometric distance, used to measure the difference in amplitude between two spectrums; r is correlation coefficient, used to measure the difference in shape between two spectrums, defined as:

$$\begin{aligned} d &= \sqrt{\sum \frac{(x_i - y_i)^2}{\max(x_i, y_i)\, N}} \\ r &= \frac{\sum (x_i - \mu_x)\left(y_i - \mu_y\right)}{(N-1)\,\sigma_x \sigma_y} \end{aligned} \tag{8}$$

where N represents the dimension of the spectral vector, μ represents average of two spectrums, σ represents standard deviation of two spectrums, x and y are the reference spectrum and the test spectrum identification. To suppress influence of magnitude of spectrum on similarity, the max value of spectrums is divided to constraint d between 0-1.

SSV ranges from 0 to $\sqrt{2}$. The smaller the value, the higher the similarity of the two spectrums; otherwise, the lower the similarity. SSV has a minimum value of 0 when two spectrums are completely coincident, and SSV has a

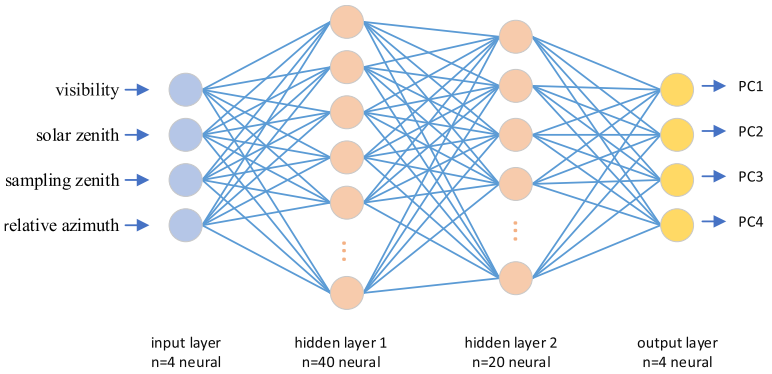


FIGURE 4. Structure of trained DNN.

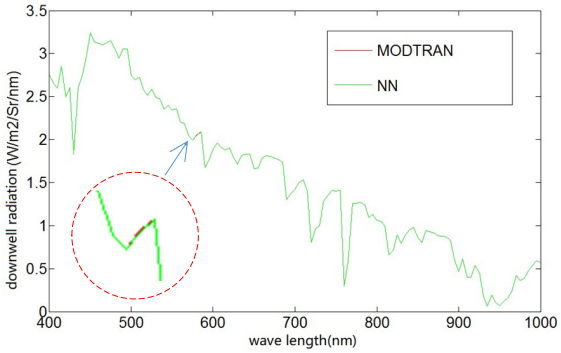


FIGURE 5. Generalization ability of DNN.

TABLE 3. SSVs between original and estimated spectral signatures.

	Maximum	Minimum	Average
DNN-based method	0.330	0.004	0.007
LUT-based method	0.175	0.000	0.079

maximum value $\sqrt{2}$ when two spectrums are perpendicular or opposite to each other.

1000 SSVs, between original (MODTRAN's output) and estimated spectral signature (DNN calculated), are obtained for different conditions. To compare DNN-based method and LUT-based method, a lookup table is also constructed by original spectral signatures and other 1000 SSVs, between original and LUT-based method estimated spectrum, are obtained. Table 3 shows the statistical value of these SSVs.

The above table shows that original and DNN-based method estimated spectrums have a high similarity, with average SSV is 0.007, a maximum SSV is 0.330, and a median SSV is 0.004. The minimum SSV of LUT-based method is 0.000 because the randomly generated condition is exact matched in LUT and not necessary to interpolate. The result shows that DNN can be used to quickly calculate the atmospheric down-radiation under any conditions and has a good generalization ability.

D. USAGE OF DNN IN RSIS-3D

A well-trained DNN are represented by a composed function and its parameters (weights and biases of neural) can rewritten in form of matrix. To calculate the at-sensor radiation in RSIS-3D, this matrix is loaded into GPU and operated with input quads and the outputs are handle by iPCA to obtain the desired AOP during the ray-tracing process.

Compared with the conventional LUT-based method that storing the whole AOP samples in different combination, DNN-based method only needs to store a trained DNN which only 60 neural and a 4*121 matrix, which only cost less than 2.4KB memory while LUT-based method will consume more than 39MB to store these 84474 original atmospheric downward spectral signature. DNN-based method can save a lot of precious graphics memory when using GPU to calculate at-sensor radiance.

IV. CONCLUSION

Atmosphere is the most important transmission medium in energy propagation of remote sensing imaging. AOP are critical factors for remote sensing imaging simulation and determine the accuracy of the simulation results. This paper proposed a method using DNN to calculate these parameters and demonstrated how to construct and train the DNN. For characteristics of optical parameters are sensitive to wavelength, multiple simple DNNs parallel methods are adopted, which greatly reduces the network training. The training results and generalization ability of DNN show that the relative error between atmospheric down-radiation calculated by DNN and MODTRAN is less than 2%.

REFERENCES

[1] R. A. Schowengerdt, *Remote Sensing: Models and Methods for Image Processing*. San Diego, CA, USA: Academic, 2010.

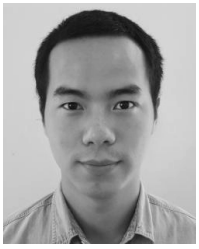
[2] X. Zhu, F. Wu, T. Wu, and C. Zhao, "Remote sensing imaging simulation and cloud removal," *Int. J. Modern Phys. B*, vol. 31, nos. 16–19, Jul. 2017, Art. no. 1744079.

[3] K. E. Joyce, S. E. Belliss, S. V. Samsonov, S. J. McNeill, and P. J. Glassey, "A review of the status of satellite remote sensing and image processing techniques for mapping natural hazards and disasters," *Prog. Phys. Geography, Earth Environ.*, vol. 33, no. 2, pp. 183–207, Jun. 2009.

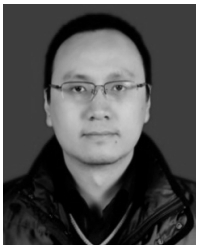
[4] R. O. Green, M. L. Eastwood, C. M. Sarture, T. G. Chrien, M. Aronsson, B. J. Chippendale, J. A. Faust, B. E. Pavri, C. J. Chovit, M. Solis, M. R. Olah, and O. Williams, "Imaging spectroscopy and the airborne Visible/Infrared imaging spectrometer (AVIRIS)," *Remote Sens. Environ.*, vol. 65, no. 3, pp. 227–248, Sep. 1998.

[5] Y. I. Wei-Ning, "Introduction of optical remote sensing imaging simulation system," *J. Atmos. Environ. Optics.*, vol. 2, no. 6, pp. 459–463, Jun. 2007.

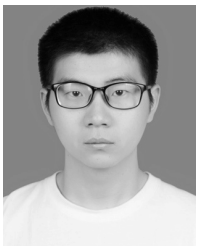
[6] Y. Liu, A. Wong, and P. Fieguth, "Remote sensing image synthesis," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2010, pp. 2467–2470.
[7] S. D. Kuo, J. R. Schott, and C. Y. Chang, "Synthetic image generation of chemical plumes for hyperspectral applications," *Opt. Eng.*, vol. 34, no. 4, pp. 1047–1056, Apr. 2000.
[8] C. Chen, "Key techniques studying on high-resolution remote sensing imaging simulation," Ph.D. dissertation, Science Island Branch Graduate School, Univ. Sci. Technol. China, Hefei, China, 2018.
[9] F.-C. Hu, B. Zhang, Z.-C. Chen, X. Zhang, J.-S. Li, Q. Shen, and Q.-X. Tong, "Investigation of atmospheric aerosol optical properties over lake taihu using MODIS and CE318 data," in *Proc. Int. Symp. Photoelectron. Detection Imag., Rel. Technol. Appl.*, Sep. 2008, Art. no. 66251F.
[10] J. S. Sanders and S. D. Brown, "Utilization of DIRSIG in support of real-time infrared scene generation," *Proc. SPIE*, vol. 4029, pp. 278–285, Jul. 2000.
[11] Q. Gu, X. Chen, and D. Yang, "Verification technology of remote sensing camera satellite imaging simulation based on ray tracing," *Proc. SPIE*, vol. 10339, Aug. 2017, Art. no. 103391M.
[12] A. Kylling, "Fast simulation tool for ultraviolet radiation at the Earth's surface," *Opt. Eng.*, vol. 44, no. 4, Apr. 2005, Art. no. 041012.
[13] J.-P. Gastellu-Etchegorry, T. Yin, N. Lauret, E. Grau, J. Rubio, B. D. Cook, D. C. Morton, and G. Sun, "Simulation of satellite, airborne and terrestrial LiDAR with DART (I): Waveform simulation with quasi-Monte Carlo ray tracing," *Remote Sens. Environ.*, vol. 184, pp. 418–435, Oct. 2016.
[14] R. Darvishzadeh, A. A. Matkan, and A. D. Ahangar, "Inversion of a radiative transfer model for estimation of rice canopy chlorophyll content using a lookup-table approach," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 4, pp. 1222–1230, Aug. 2012.
[15] P. R. Colarco, S. Gassó, C. Ahn, V. Buchard, A. M. da Silva, and O. Torres, "Simulation of the ozone monitoring instrument aerosol index using the NASA goddard Earth observing system aerosol reanalysis products," *Atmos. Meas. Techn.*, vol. 10, no. 11, pp. 4121–4134, Nov. 2017.
[16] Y. Peng, G. He, Z. Zhang, T. Long, M. Wang, and S. Ling, "Study on atmospheric correction approach of Landsat-8 imageries based on 6S model and look-up table," *J. Appl. Remote Sens.*, vol. 10, no. 4, Oct. 2016, Art. no. 045006.
[17] S. Sonoda and N. Murata, "Neural network with unbounded activation functions is universal approximator," *Appl. Comput. Harmon. Anal.*, vol. 43, no. 2, pp. 233–268, Sep. 2017.
[18] B. Hanin, "Universal function approximation by deep neural nets with bounded width and ReLU activations," *Mathematics*, vol. 7, no. 10, p. 992, Oct. 2019, doi: 10.3390/math7100992.
[19] H. D. Wang, X. S. Ma, and Z. Yang, "Computing the atmospheric point spread function by artificial neural Networks," *Laser Optoelectron. Prog.*, vol. 53, no. 10, pp. 250–255, Jan. 2016.
[20] A. Elrharras, R. Saadane, M. Wahbi, and A. Hamdoun, "Neural networks and PCA for spectrum sensing in the context of cognitive radio," in *Proc. Medit. Conf. Inf. Commun. Technol.*, Berlin, Germany: Springer, 2016, pp. 173–181.
[21] X. Zhu D. Tuia, L. Mou, G.-S. Xia, L. Zhang, F. Xu, and F. Fraundorfer, "Deep learning in remote sensing: A review," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 8–36, Dec. 2017.



CHEN CHUAN received the Ph.D. degree in computer application technology from the University of Science and Technology of China, Hefei, China, in 2018. He is currently an Assistant Researcher with the Institute of Computer Application, China Academy of Engineering Physics. His research interests include remote sensing information process, machine learning, and digital mock up.



CHENG ZHENG received the master's degree in mechanical manufacturing and automation from Northwestern Polytechnical University, Xi'an, China, in 2007. He is currently pursuing the Ph.D. degree in electronics and information with the University of Electronic Science and Technology of China. He is also working at the Institute of Computer Application, China Academy of Engineering Physics. His research interests include image process, machine learning, and virtual reality.



LIU BO received the B.E. degree from the North University of China, Taiyuan, China, in 2018. He is currently pursuing the master's degree in computer applications with the University of Chinese Academy of Sciences. He is also working at the Chinese Academy of Sciences. His current research interest includes object detection based on simulated image sets.

LI LIGANG received the Ph.D. degree in optical engineering from the Chinese Academy of Sciences, in 2006. He is currently a Professor with the Chinese Academy of Sciences. His main research interests include optical imaging simulation system technology, spatial information processing, and other aspects of research.