Journal of
Applied Remote Sensing

RemoteSensing.SPIEDigitalLibrary.org

Multiple geometry atmospheric correction for image spectroscopy using deep learning

Fangcao Xu
Guido Cervone
Gabriele Franch
Mark Salvador

SPIE.

Fangcao Xu, Guido Cervone, Gabriele Franch, Mark Salvador, "Multiple geometry atmospheric correction for image spectroscopy using deep learning," *J. Appl. Remote Sens.* **14**(2), 024518 (2020), doi: 10.1117/1.JRS.14.024518

Multiple geometry atmospheric correction for image spectroscopy using deep learning

Fangcao Xu,[a,*] Guido Cervone,[a] Gabriele Franch,[b] and Mark Salvador[c]

[a]Pennsylvania State University, Institute for Computational and Data Sciences, Department of Geography, University Park, Pennsylvania, United States
[b]University of Trento, FBK Research Centre, Italy
[c]Zi INC, Washington D.C., United States

Abstract. The goal of this research is to develop a general deep learning solution for atmospheric correction and target detection using multiple hyperspectral scenes. It is assumed that the scenes differ only in range and viewing angles, that they are acquired in rapid sequence using an airborne sensor orbiting a target, and that the target and the atmosphere remain invariant within the time scale of the collection. Several hundred thousand hyperspectral simulations were performed using the MODTRAN model and were used to train the deep learning solution, as well as to validate the proposed method. The input to the deep learning solution is a matrix of the simulated radiances at the sensor as function of wavelength and elevation angles. The output is atmospheric upwelling, downwelling, and transmission. This solution is repeated for all or a subset of pixels in the scene. We focus on emissive properties of targets, and simulations are performed in the longwave infrared between 7.5 and 12 μm. Results show that the proposed method is computationally efficient and it can characterize the atmosphere and retrieve the target spectral emissivity within one order of magnitude errors or less when compared with the original MODTRAN simulations. © *2020 Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JRS.14.024518]

Keywords: hyperspectral imaging analysis; deep learning; convolutional neural network; atmospheric correction; surface reflectivity retrieval; target detection.

Paper 200162 received Feb. 28, 2020; accepted for publication May 22, 2020; published online Jun. 1, 2020.

1 Introduction

The increasing spectral–spatial resolution of hyperspectral imaging (HSI) enables more accurate analysis for target detection and classification. The general goal of image spectroscopy, or hyperspectral remote sensing, is to measure emitted or reflected radiation at hundreds of narrow and continuous spectral wavelengths collected remotely from air-borne or space-borne platforms.[1]

Hyperspectral image analysis is used to identify with confidence surface solid materials, liquids, or atmospheric gases based on their spectral signatures.[2–5] Except when temperature reaches absolute zero, all objects emit electromagnetic radiation, and most objects also reflect electromagnetic radiation emitted by other sources.[6] One of the most important steps in remote sensing image analysis consists of taking into account the atmospheric radiative transfer process, which causes attenuation of the signal as measured at the sensor.[7] A careful atmospheric characterization and correction is paramount to quantify the atmospheric effects and to derive accurate surface target properties.

The current state-of-the-art approach to accounting for atmospheric influence on hyperspectral scenes is based on algorithms and tools developed in the early 1990s and continues to pervade in operational, research, and development activities.[8–11] These methods have demonstrated operational success and are ensconced in existing algorithms and software. However, many simplifying assumptions or expedient processing steps are made in this process.

*Address all correspondence to Fangcao Xu, E-mail: xfangcao@psu.edu

1931-3195/2020/$28.00 © 2020 SPIE

At the foundation of these approaches is the fundamental radiative transfer equation (RTE), which defines the reflective and emissive spectral radiance components of the image scene and its three-dimensional surroundings, as well as detailed atmosphere profiles (e.g., gas and aerosols) as a function of altitude and pressure. Because of the combined effects of variations in a material's physical properties, sensor noise, as well as uncertainty in atmospheric parameters and the difficulty associated with temperature-emissivity separation, leads to inherent variation observed in remote sensing data and makes the parameter estimation of the traditional RTE very challenging.[12]

The most significant error occurs in atmospheric correction in which a single geometric solution for all elements of the RTE is applied to every pixel of a spectral image.[13–15] We know this solution to be expedient, but it is also error inducing. Another significant shortcoming consists of analyzing a single hyperspectral image at a time. In other words, every image is an island. This individual image analysis is driven primarily by the limits of sensors and collection platforms with optimal scene collection consisting of a nadir looking scan of the target area. These existing solutions lag behind the revolution in computational power, artificial intelligence, and agile sensors and platforms.

These solutions have been shown to fail under varied environmental conditions, obscuration due to clouds, and target material conditions. It can be said that today's analysis is effective in performing material identification in optimal collection conditions. It can also be said that, in nonoptimal conditions such as cloudy scenes, shadows, intimate mixtures of materials, liquid spills and residues, or any combination of the above, material identification of solids, liquids, or gases is ineffective, unrepeatable, or subject to unknown levels of uncertainty. We hypothesize that this gap in effectiveness stems from the incomplete solution of the RTE that a single hyperspectral image and existing analysis methods afford.

It is theoretically possible to achieve higher accuracy by analyzing spectra relative to the same target pixels if they are acquired in multiple individual scenes collected in rapid sequence and at different angles. Collecting these scenes requires gimballed instruments with attitudes that can be dynamically adjusted during collection, which overcomes the fixed nadir looking geometries of past instruments.[16] Using advanced algorithms that can ingest and analyze multiple scenes with varying atmospheric contributions is also required. There are currently no established solutions to the RTE that take into account spectra acquired simultaneously from multiple angle geometries.

In addition to imaging a particular scene location, the agile sensor can also image the background and the atmosphere itself to better estimate additional components of the RTE. While the previous generation of sensors has a scene revisit time of minutes to days, the current generation, and even more so the next, has a scene revisit time of as little as a few seconds. The ability to collect rapid sequences of hyperspectral scans under different angles provides an unprecedented dataset. This combination of multiple scans and geometric diversity forces us to re-evaluate the fundamental RTE for its full spatial dimensionality and temporal component.

There are three objectives in our research:

1. Expand the solution of the current RTE for full geometric diversity to exploit multiscan hyperspectral images simultaneously;
2. Develop a deep learning approach to estimate different components of the expanded RTE under different atmospheric conditions;
3. Apply the deep learning solution to the target detection.

The motivation of this research is to provide a well-defined deep learning solution to compensate for atmospheric influence on the total radiance received at the sensor and to retrieve the target spectral properties that enable the identification of target materials in scenarios and environments, of which no current or planned hyperspectral image exploitation system is capable. Compared with traditional state-of-the-art approaches using the RTE, this approach makes use of the computer ability of artificial intelligence, which can better learn the nonlinear relationship and complex interactions between atmosphere and different radiative components passing through it. This can enhance the current state-of-the-art in hyperspectral remote sensing research and drive future hyperspectral sensor performance requirements and concepts of atmospheric characterization and target detection operations.

To properly test the proposed approach, two steps must occur.

1. Develop a solution using simulated data because it is necessary to test the proposed approach in an idealized environment in which all variables can be controlled and an estimation of the sources of error can be precisely quantified. Furthermore, deep learning-based solutions require hundreds of thousands of examples and counter-examples to learn. It is probably unfeasible and not cost effective to train the system on real-world collected data; further, simulated data should contain the general underlying relationships that are also present in the real-world data and that the system has to learn. Finally, an initial simulated solution is necessary for the collection of real-world data; basically it drives what data to collect in terms of target and ground truth and what aspects of the method should be particularly tested in a real-world setting.
2. Collect hyperspectral remote sensing scenes that can help validate the proposed approach in controlled but realistic real-world settings. Data collection is generally time consuming in terms of preparing for the collection and the collection itself, especially when that ground truth measurements are to be collected at the same time as the aerial collection. To test the methodology, it is necessary to collect atmospheric characteristics along with target characteristics, which greatly increases the complexity of the collection. Furthermore, data collections can be very expensive, especially in the longwave part of the spectrum.

This research describes the first of these two necessary steps, namely providing results of the methodology as applied to simulated data. The results presented in this article were used to plan a data collection campaign, called Nittany Radiance 2019, that occurred on the premises of the Pennsylvania State University campus in April 2019.[17] The analysis of Nittany Radiance 2019 is still ongoing and will be reported at a later time. Several of the assumptions made with regards to spectral range, resolution, and collection geometry are chosen to meet the characteristics of Blue Heron. This instrument, built by the Harris corporation, has two focal plane arrays with 258 detectors each, 256 spectral channels, and a spectral range of nominally 7.5 to 13.5 μm.[18]

This article is organized as follows: Sec. 2 discusses the methods, including current state-of-the-art solutions to RTE. This section describes the overall solution and introduces the assumptions made, as well as the expected errors that we expect to achieve. Section 3 presents the results obtained and the overall accuracy achieved. Finally, Sec. 4 summarizes the research in terms of the results obtained and limitations to the proposed solution.

2 Methods

Different surface targets absorb or reflect solar and other objects' electromagnetic radiation in different ways. The spectral reflective or emissive properties of a surface target depend on the material type, its physical and chemical conditions, and the solar and sensor geometries. The different spectral signatures of surface materials make it possible to classify them from remote measurements.

2.1 *Radiative Transfer Equation*

A purely physics based state-of-the-art RTE describes the flow of radiative energy through different media as a mathematical process of absorption, reflection, and scattering from a target to a sensor. For the case of a downlooking view at a target, the total observed spectral radiance is denoted as

$$L_\lambda = \left\{ E_{s\lambda} \cos\sigma\tau_1(\lambda)\frac{r(\lambda)}{\pi} + \varepsilon(\lambda)L_{T\lambda} + [F(L_{ds\lambda} + L_{de\lambda}) + (1-F)(L_{bs\lambda} + L_{be\lambda})]r(\lambda) \right\}\tau_2(\lambda) + L_{us\lambda} + L_{ue\lambda}, \quad (1)$$

where s represents the solar component, ε is the self-emitted thermal component, b is the background involved component, and d and u are the downwelling and upwelling components,

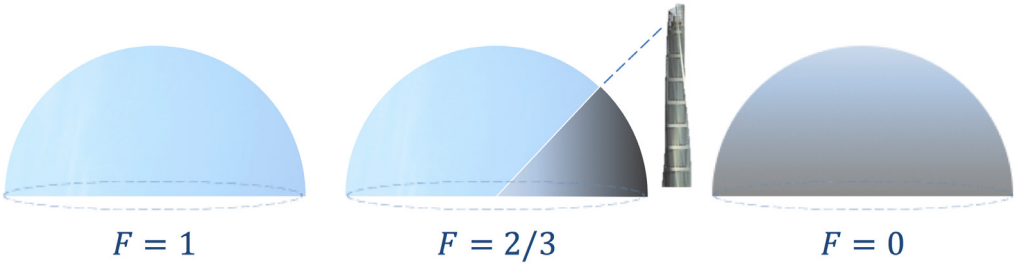

Fig. 1 Shape factor F.

Table 1 Parameters in RTE.

Parameters	Values
$E_{s\lambda}$	Solar spectral irradiance
σ	Solar incident angle
$r(\lambda)$	Target spectral reflectivity
$\varepsilon(\lambda)$	Target spectral emissivity
F	Shape factor of visible atmosphere to the target
τ_1	Atmospheric transmission from the Sun to the target
τ_2	Atmospheric transmission from the surface to the sensor
$L_{T\lambda}$	Spectral radiance of a blackbody at temperature T
$L_{ds\lambda} + L_{d\varepsilon\lambda}$	Solar and atmospheric downwelling spectral radiance
$L_{bs\lambda} + L_{b\varepsilon\lambda}$	Background reflected and self-emitted spectral radiance
$L_{us\lambda} + L_{u\varepsilon\lambda}$	Solar and atmospheric upwelling spectral radiance

respectively.[10] The shape factor F represents the fraction that the hemisphere is obscured by background objects, shown in Fig. 1. All other elements in Eq. (1) are listed in Table 1.

The RTE requires the calculation of the direct or scattered solar and atmospheric downwelling and upwelling radiance, the atmospheric transmission, and the target reflectivity. Numerous computational codes are available to solve the RTE. Their accuracy and speed of calculations depend on the implemented parameters and models. For example, Pstar uses the discrete ordinate method (DOM), which is applied to a single homogeneous layer to obtain the discrete ordinate solution and then the atmospheric molecular and aerosol scattering are calculated with a small number of discrete ordinate directions. It is efficient for an optically thin atmosphere profile, such as clear sky and background aerosol.[19]

Another solution to the RTE is found by MYSTIC, which uses a Monte–Carlo optimization method (MC).[20,21] Compared with the DOM, the MC method allows users to define complex clouds and aerosols and inhomogeneous atmosphere profiles with the guarantee of a higher accuracy. However, it is more computational expensive and has its own inherent statistical noise.

SOSVRT uses the successive-order-scattering method, which models the atmospheric scattering effect by the integration of all orders of scattering along the line-of-sight (LOS) path.[22] It is accurate in quantifying the atmospheric scattering radiance, but it is inefficient at achieving the computational convergence for optical thick medium or oblique observation angles.

Finally, RAY uses the doubling–adding technique to approximate the vertical inhomogeneous atmospheric layers by a large number of small parallel homogeneous layers.[23] It provides a trade-off between the time and accuracy, which can save additional running time because of its symmetrical and reciprocal relationships for calculating.

It was found that all of these codes give very similar results, which is intuitive because they try to solve numerically the same physical phenomenon.[24] Herein lies the vulnerability of a traditional mathematical-based solution that relies on a single viewpoint. Because the geometry of the problem is vastly simplified, the single solution is applied to every pixel of the scene regardless of its geometry or local conditions to arrive at the target.

There is a potential transformative solution to improve the current state-of-the-art solutions that a single hyperspectral image affords by leveraging recent advances in computational and data science. Artificial intelligence and machine learning (ML) in general, and deep learning techniques more specifically, give a new perspective on how to solve the RTE using the availability of large amounts of multiple hyperspectral image scans. We hypothesize that this new multiscan geometry can provide a more accurate atmospheric characterization and can target detection by analyzing spectra of same target pixels acquired in multiple individual scenes collected in rapid sequence and at different angles.

This research proposes using a deep learning approach to learn the relationship between the total radiance received at the sensor L_λ and different solar and atmospheric components as a function of geometric angles that can be used for atmospheric correction in real time and with limited computational resources, so it can be efficiently implemented into a real-time hardware solution.

2.2 *Data Simulation and Collection*

The proposed deep learning solution consists of two main steps. The first is the training of the network, which is generally a time consuming and computationally expensive task, and the instantiation of the solution, which is a very quick process. Deep learning networks require large amounts of data to internally build a concept representation of the relationships between the input and output, and thus it is more practical to use simulated versus observed data to generate the required large dataset.

Synthetic data are generated using the MODTRAN software to simulate the total radiance, upwelling, downwelling, and atmospheric transmission for different geometric angles, atmospheric models, time of the day and year, and specific targets with varying reflectivities.[8,25] It is assumed that data are collected at different angles using an airborne sensor, shown in Fig. 2(a). Several simplifications and assumptions are made here:

1. The airborne sensor is assumed to have a spectral resolution of 17.5 nm between 7.5 and 12 μm, which reflects the characteristics of the Blue Heron sensor.
2. The target is located at $(40.7934, -77.86)$, which lies within the Pennsylvania State University campus.
3. The sensor is assumed to be orbiting around the target with a fixed range of 5 km.
4. The target is located in an open area. Therefore, the shape factor $F = 1$ and the background components $L_{bs\lambda}, L_{b\varepsilon\lambda}$ can be omitted from Eq. (1):

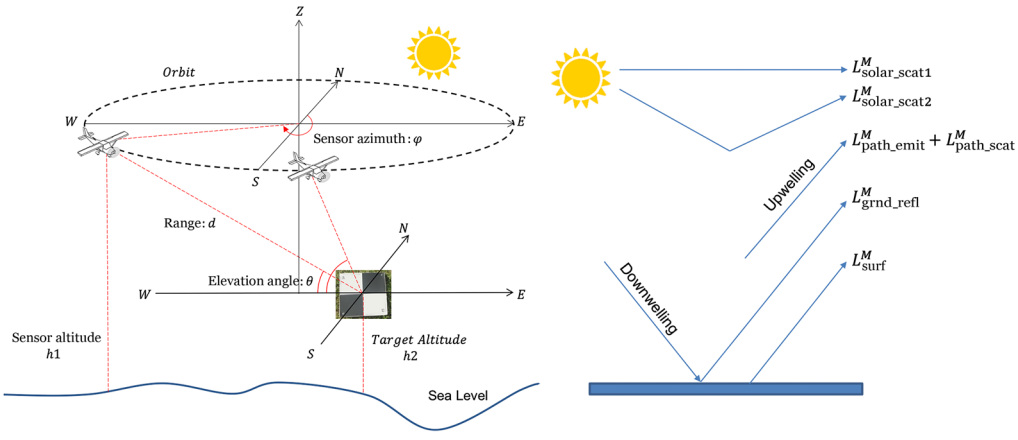

Fig. 2 Geometric setting and MODTRAN simulated components.

$$L_\lambda = \left[E_{s\lambda} \cos \tau_1(\lambda) \frac{r(\lambda)}{\pi} + \varepsilon(\lambda) L_{T\lambda} + (L_{ds\lambda} + L_{d\varepsilon\lambda}) r(\lambda)\right] \tau_2(\lambda) + L_{us\lambda} + L_{u\varepsilon\lambda}. \quad (2)$$

5. The simulated target used for training is an opaque Lambertian gray body with a constant emissivity across the spectrum, although this constant value varies across the simulations.
6. The multiple simulated targets used for the validation have varying emissivity that correspond to real-world objects, such as polyethylene.

The current work provides a framework to learn the relationship between the at-sensor radiance and different atmospheric and solar components under a given geometric setting and can be easily adjusted to exploit different spectral and physical range parameters. All parameters used for the MODTRAN simulation are listed in Table 2. The different radiative components that MODTRAN simulated are shown in Fig. 2(b).

The at-sensor total radiance L_λ simulated via MODTRAN consists of the ground reflected radiance, surface self-emitted radiance, two atmospheric upwelling radiances, and two solar upwelling radiances, represented as

$$L_\lambda = L^M_{\text{grnd\_refl}} + L^M_{\text{surf}} + L^M_{\text{path\_emit}} + L^M_{\text{path\_scat}} + L^M_{\text{solar\_scat1}} + L^M_{\text{solar\_scat2}}, \quad (3)$$

where $L^M_{\text{path\_emit}}$ is the atmospheric LOS thermal radiance and $L^M_{\text{path\_scat}}$ is the atmospheric thermal radiance scattered to the LOS path. $L^M_{\text{solar\_scat1}}$ is the solar single scattering radiance and $L^M_{\text{solar\_scat2}}$ is the solar multiple scattering radiance. The corresponding relationship between MODTRAN simulated radiative components and those in Eq. (2) is listed in Table 3.

As shown in Table 3, the simulated at-sensor ground reflected radiance $L^M_{\text{grnd\_refl}}$ involves an unknown target spectra variable [i.e., $r(\lambda)$]. Therefore, the downwelling radiance must be separated from the MODTRAN simulated ground reflected radiance $L^M_{\text{grnd\_refl}}$:

Table 2 Parameters in MODTRAN simulation.

Parameters	Values
Wavelength λ	[7.5, 12] by every 0.0175 (μm)
Day of the year	[1 to 365]
Time of the day	[2, 6, 10, 14, 18, 22] (:00)
Reflectivity $r(\lambda)$	[5, 10, 15, 30, 50, 80, 100] (%)
Elevation angle θ	[30 to 90] by every 5 (deg)
Azimuth ϕ	[0 to 360] by every 5 (deg)
Target temperature T	320 (K)
Shape factor F	1

Table 3 Different decomposition of L_λ.

MODTRAN	RTE
$L^M_{\text{grnd\_refl}}$	$\left[\frac{E_{s\lambda} \cos \sigma\tau_1(\lambda)}{\pi} + L_{ds\lambda} + L_{d\varepsilon\lambda}\right] r(\lambda)\tau_2(\lambda)$
L^M_{surf}	$\varepsilon(\lambda) L_{T\lambda} \tau_2(\lambda)$
$L^M_{\text{solar\_scat1}} + L^M_{\text{solar\_scat2}}$	$L_{us\lambda}$
$L^M_{\text{path\_emit}} + L^M_{\text{path\_scat}}$	$L_{u\varepsilon\lambda}$

$$L^{M}_{\text{down}} = \frac{L^{M}_{\text{grnd\_refl}}}{\tau_2(\lambda) \times r(\lambda)} = \frac{E_{s\lambda} \cos \sigma \tau_1(\lambda)}{\pi} + L_{ds\lambda} + L_{d\varepsilon\lambda}. \tag{4}$$

As mentioned above, the goal of using deep learning is to autonomously learn the attribute–value relationships between specific targets and the parameterization of the RTE to characterize the atmosphere under different geometric angles and time. For MODTRAN simulations, the learning task is to estimate the atmospheric transmission $\tau_2(\lambda)$, downwelling radiance L^{M}_{down}, solar and atmospheric upwelling radiance $L^{M}_{\text{solar\_scat1}} + L^{M}_{\text{solar\_scat2}}$ and $L^{M}_{\text{path\_emit}} + L^{M}_{\text{path\_scat}}$, given as input the at-sensor total radiance L_λ from millions of simulated data.

2.3 *Deep Neural Network*

The deep learning model used in this research is a convolutional neural network (CNN), which can automatically and hierarchically extract and learn spectral features from HSI data itself with little or no human supervision. It has been successfully applied to many hyperspectral image analyses, such as spectral unmixing, segmentation, target detection, and classification with ML methods in recent years.[26–30]

Many CNN architectures have been developed in recent years to address different spectral, spatial, or spectral–spatial feature analysis. Charmisha et al.[31] took the pixel vectors as input and presented a vectorized CNN architecture to perform the classification of HSI data in the spectral domain. More complex CNNs will utilize the spectral–spatial information together for a better feature extraction. Wei et al.[32] proposed a hierarchical deep framework called spectral–spatial response to jointly learn spectral and spatial features through the template matching for land cover analysis. Shi and Pun[33] combined the spectral–spatial feature extracted by the CNN model with a multiscale hierarchical recurrent neural network to capture the spatial relations of local features at different scales. Mou et al.[34] adopted an end-to-end recurrent convolutional neural network with several convolutional layers at the beginning to extract spectral–spatial features for change detection in earth observation.

However, as Signoroni et al. suggested in their review paper, most recent HSI-DL studies exploiting CNNs include the classification as a final goal, which is highly related to the pixel labeling or feature extraction as a classifier.[35] The DL-driven RTE solution to characterize the atmosphere and accurately quantify its radiative attenuation effects on the at-sensor received signal is still needed.

Therefore, we propose a fundamentally different research question. While related research focuses on direct pixel classification (or target detection), the proposed approach focuses on characterizing the atmosphere, which is used to extract the pixel reflectivity (or emissivity) from the total sensor radiance. The main advantage of the proposed approach is that, in theory, it can work with any target pixel, whereas methods that use a classifier for a direct pixel classification are generally limited to signatures that were used for training.

To fully exploit the rich spectral–spatial information obtained from multiscan hyperspectral images and learn the complex nonlinear relationship between the observed at-sensor signal and different radiative components that contribute to it, a low-dimensional representation of radiative features needs to be extracted and learned via the deep neural network.

The proposed CNN has two parts: the encoder and the decoder. The encoder is a network that compresses the input through convolutional layers into a latent space vector. It has three convolution blocks (conv_block) and one convolution layer (conv_layer) in which each conv_block contains a convolution layer + a batch normalization layer + an activation layer (leaky rectified linear unit). The function of the last convolution layer in the encoder is to convert the feature maps into a latent space vector. The decoder network takes this latent space vector and performs upsampling with transposed convolution operations to reconstruct the output, and it also has three convolution blocks and one transposed convolution layer (conv_trans) in which each conv_block contains a transposed convolution layer + a batch normalization layer + an activation layer (rectified linear unit). The function of the last transposed convolution layer in the decoder is to convert the feature maps into four 256×8 matrices. The network structure and its parameters are illustrated in Fig. 3 and Table 4.

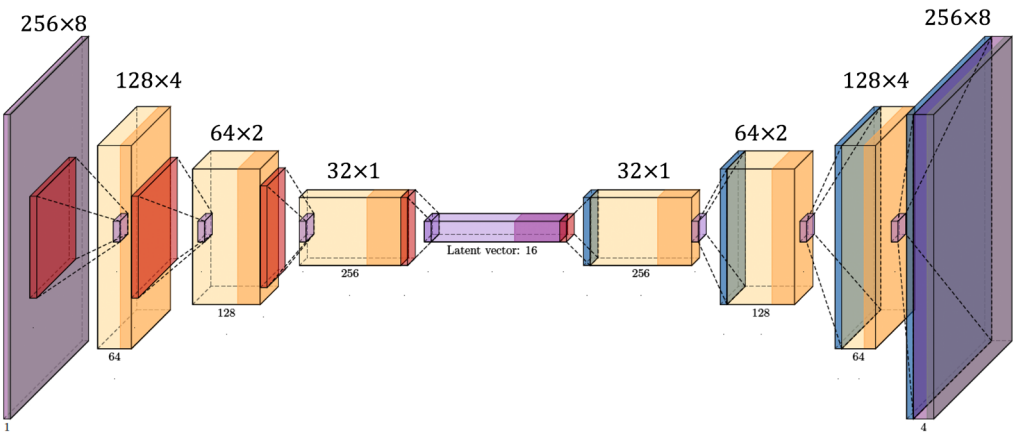

Fig. 3 Autoencoder CNN structure.

Table 4 Parameters of the network structure.

Operation layer		Number of kernels	Kernel size	Stride	Padding	Number of parameters	Data size ($nc \times h \times w$)
Input data		—	—	—	—	0	$1 \times 256 \times 8$
Encoder	conv_block	64	$4 \times 4 \times 1$	2×2	1×1	1088	$64 \times 128 \times 4$
	conv_block	128	$4 \times 4 \times 64$	2×2	1×1	2176	$128 \times 64 \times 2$
	conv_block	256	$4 \times 4 \times 128$	2×2	1×1	4352	$256 \times 32 \times 1$
	conv_layer	16	$32 \times 1 \times 256$	1×1	0×0	528	$16 \times 1 \times 1$
Decoder	conv_block	256	$32 \times 1 \times 16$	1×1	0×0	8192	$256 \times 32 \times 1$
	conv_block	128	$4 \times 4 \times 256$	2×2	1×1	2176	$128 \times 64 \times 2$
	conv_block	64	$4 \times 4 \times 128$	2×2	1×1	1088	$64 \times 128 \times 4$
	conv_trans	4	$4 \times 4 \times 64$	2×2	1×1	68	$4 \times 256 \times 8$

Each input in this proposed CNN is the at-sensor total radiance, observed at different viewing angles from multiscan hyperspectral images for a specific day of the year, time of the day, and specific target, as listed in Table 2. Instead of using a vector to represent the spectral radiance at different wavelength bands, this input is a 256×8 matrix as a function of both angle and wavelength, where 256 rows represent the 256 wavelength bands between 7.5 and 12 μm and eight columns represent for eight geometric angles.

Each input matrix has four corresponding outputs, and each output is also a 256×8 matrix having the same wavelengths and geometric angles as rows and columns. The values in these four outputs are the atmospheric transmission rate, downwelling radiance, and solar and atmospheric upwelling radiance that contribute to the at-sensor total radiance, respectively.

The CNN network is trained on the NVIDIA Quadro P5000 GPU with four workers, using PyTorch[36] to learn the parameters that can map each input matrix (i.e., at-sensor total radiance) to its four corresponding output matrices ($4 \times 256 \times 8$) by minimizing the mean square error (MSE) as a loss function for all 299,040 input and output matrices simulated for different atmospheric models, time of the day, day of the year, and reflectivities. The optimization function used in this network is the adaptive moment estimation (Adam) with a learning rate 10^{-4} and a batch size of 128 for 10,000 epochs. The total training time on 299,040 matrices is around 60 min while the inference time to predict four output matrices given any at-sensor total radiance matrix as input is 2 to 3 s.

The reason to choose the encoder–decoder architecture is because this model is able to project the original input into a new space and generate compressed, extended, or even equally

dimensional outputs rather than carry out classification tasks, and it perfectly fits our research purpose.[37] The comparison between deep learning estimations and the results from the MODTRAN computations is given in Sec. 3. This leads to a measure of both accuracy and precision using multiple scans for any given image scan prior to performing the target detection.

2.4 *Target Spectra Retrieval*

After the DL model is well-trained, the spectra of surface or near-surface targets can be retrieved with predicted $\tau_2(\lambda)$, L^M_{down}, $L_{us\lambda}$, and $L_{u\varepsilon\lambda}$, together with the observed at-sensor total radiance L_λ. Recalling Eq. (2) and Table 3, the remaining unknown variables are just $\varepsilon(\lambda)$, $r(\lambda)$, and $L_{T\lambda}$. $L_{T\lambda}$ is also referred to as the Planck equation, which describes the radiation emitted by a blackbody

$$L_{T\lambda} = 2\ hc^2\lambda^{-5}(e^{\frac{hc}{\lambda kT}} - 1)^{-1}, \tag{5}$$

where T is the target temperature in Kelvin, h is Planck's constant, c is the speed of light, and k is the Boltzmann gas constant. $L_{T\lambda}$ can be calculated, assuming T is known. In addition, according to the energy equilibrium and Kirchoff's law, all of the incident flux can only be transmitted, absorbed, or reflected. The absorptivity of a surface is numerically equal to its emissivity. Thus, the reflectivity of an opaque surface has the following relationship with its emissivity:

$$\tau(\lambda) + \alpha(\lambda) + r(\lambda) = 1 \quad \tau(\lambda) \approx 0 \quad \alpha(\lambda) = \varepsilon(\lambda) \quad \varepsilon(\lambda) + r(\lambda) = 1, \tag{6}$$

where $\tau(\lambda)$ is the transmissivity, $\alpha(\lambda)$ is the absorptivity, $r(\lambda)$ is the reflectivity, and $\varepsilon(\lambda)$ is the emissivity.

Substituting Eqs. (4) and (6) into Eq. (2), $r(\lambda)$ is solved as

$$L^M_{\text{down}} r(\lambda) + [1 - r(\lambda)]L_{T\lambda} = \frac{L_\lambda - L_{us\lambda} - L_{u\varepsilon\lambda}}{\tau_2(\lambda)} \quad r(\lambda) = \frac{(L_\lambda - L_{us\lambda} - L_{u\varepsilon\lambda})/\tau_2(\lambda) - L_{T\lambda}}{L^M_{\text{down}} - L_{T\lambda}}. \tag{7}$$

It is intuitive to see that the target spectral reflectivity $r(\lambda)$ now is a function of the observed total radiance L_λ, predicted solar and atmospheric components [i.e., L^M_{down}, $\tau_2(\lambda)$, $L_{us\lambda}$, $L_{u\varepsilon\lambda}$], and calculable blackbody thermal radiation $L_{T\lambda}$, assuming a known T. Comparing this retrieved $r(\lambda)$ with the spectral features in the database, such as the NIST or USGS spectral library, the target material or similar materials can be identified. This DL solution only takes the observed total radiance as input from multiscan hyperspectral images to characterize the atmosphere and retrieve the target spectra, which can be implemented into real-time target detection unlike that of traditional state-of-the-art approaches.

3 Results

This section first conducts an experiment to understand which geometries of multiple scans and spectral bands are necessary for an accurate atmospheric and background characterization. After identifying the optimal combinations of angles, azimuth, and spectra for the data simulation and collection, millions of data are simulated via the MODTRAN. Then, the proposed autoencoder neural network is trained to estimate solar and atmospheric components of upwelling, downwelling, and transmission as a function of angles. Finally, the target spectral properties are retrieved, and an error estimate is provided.

3.1 *Model Parameter Configuration*

The geometries to best characterize the solar and atmospheric influence are determined in two steps. First, azimuths from 0 deg to 360 deg by every 5 deg are simulated for the LWIR spectrum between 7.5 and 12 μm with a spectral resolution of 17.5 nm, with all other parameters listed in Table 2 fixed, to test how MODTRAN simulated radiative components change as a function of

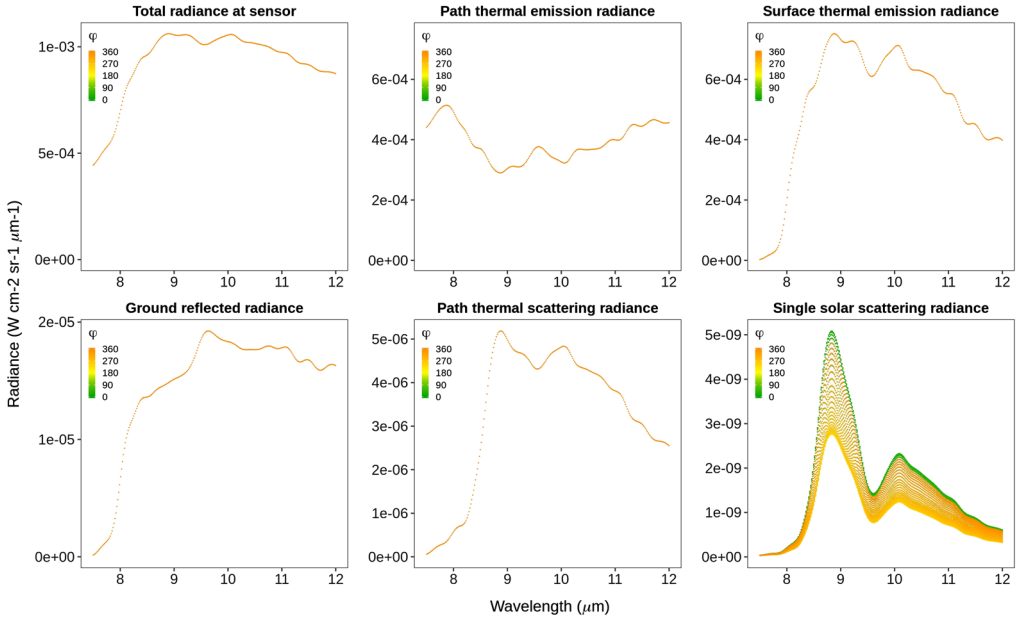

Fig. 4 Simulated radiance components at different azimuths ($\theta = 30$ deg, 2 *PM*).

azimuth and wavelength. Next, the azimuth is fixed and the elevation angles vary from 30 deg to 90 deg to research the influence of the elevation angles.

Figure 4 shows MODTRAN simulated radiance for different azimuths with a fixed elevation angle at 30 deg at the local time of 2 pm on April 18. The target reflectivity is set as 0.1. It was found that the azimuth only has obvious influence on the solar scattering radiance, but the atmospheric and surface self-emitted thermal radiance; in addition, the ground-reflected radiance does not vary much, assuming the target is a Lambertian surface.

Then, the solar scattering radiance at different azimuths is projected as a pie chart at the peak of the spectrum (i.e., 8.845 μm). At 2 pm, the Sun's azimuth is 203 deg from the north (0 deg) and the highest value of the solar scattering radiance that the sensor receives is at 23 deg. In other words, when the sensor is facing the Sun, the received solar scattering radiance is highest, as shown in Fig. 5.

Another finding is that solar radiance is at least three orders of magnitudes smaller than self-emitted thermal components for the LWIR spectrum, which means the solar contributions are

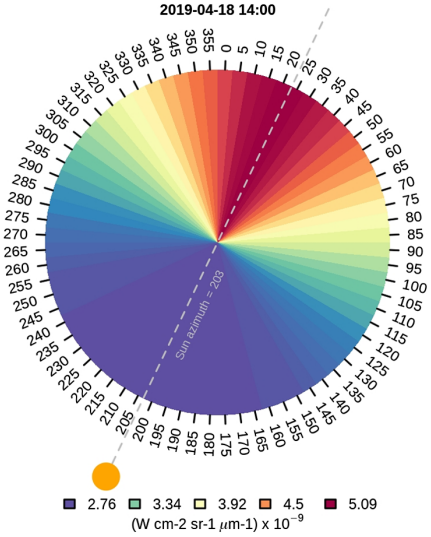

Fig. 5 Solar radiance as a function of azimuth.

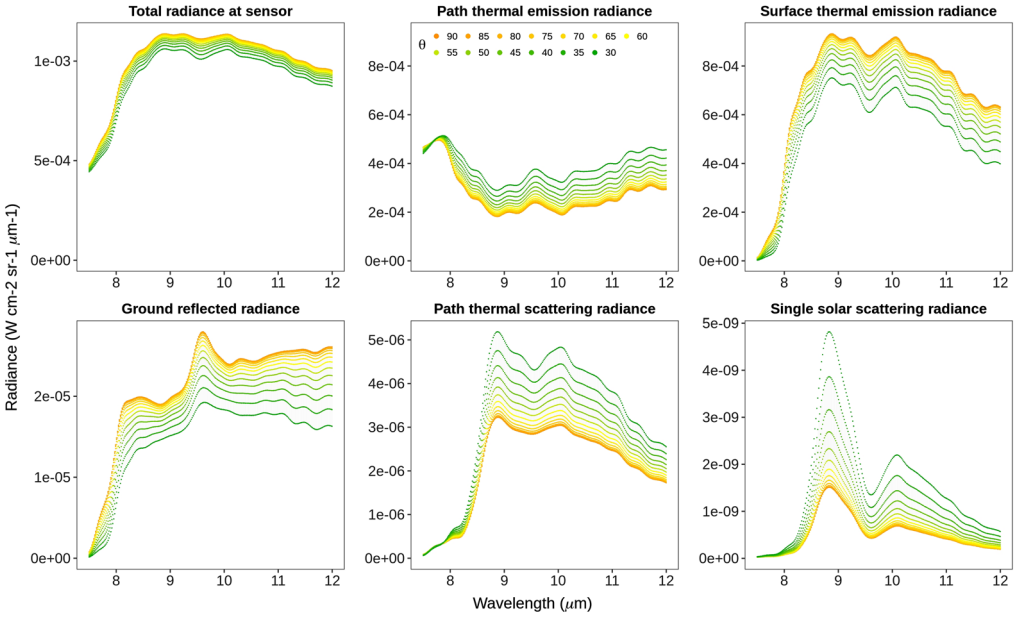

Fig. 6 Simulated radiance components at different elevation angles ($\phi = 0$ deg, 2 *PM*).

negligible. These analyses lead to a further simplification of RTE, and the reflectitivity calculation by removing the solar components from Eqs. (2), (3), (4), and (7) as

$$L^M_{\text{down}} = L_{d\varepsilon\lambda}; L_\lambda = [\varepsilon(\lambda)L_{T\lambda} + L_{d\varepsilon\lambda}r(\lambda)]\tau_2(\lambda) + L_{u\varepsilon\lambda}$$
$$r(\lambda) = \frac{(L_\lambda - L^M_{\text{path\_emit}} - L^M_{\text{path\_scat}})/\tau_2(\lambda) - L_{T\lambda}}{L^M_{\text{down}} - L_{T\lambda}}. \tag{8}$$

Continuing the previous experiment, Fig. 6 shows MODTRAN-simulated radiative components with different elevation angles but a fixed azimuth at the same location and time for the same target. It is intuitive to see the relationship between each component, and the elevation angle is nonlinear. The upwelling thermal radiances $L^M_{\text{path\_emit}}$ and $L^M_{\text{path\_scat}}$ decrease when the elevation angle increases from 30 deg to 90 deg (vertically looking up), opposite to the surface thermal emission L^M_s and ground reflected radiance $L^M_{\text{grnd\_refl}}$. This is because the LOS path will go through warmer and denser air when the elevation angle is lower, resulting in a higher atmospheric thermal emission and scattering. However, the transmission decreases due to a higher atmospheric absorption caused by the warmer and denser air near the ground; therefore, the at-sensor L^M_{surf} and $L^M_{\text{grnd\_refl}}$ are smaller when the elevation angle is lower.

To draw a conclusion, the azimuth and solar components do not have significant impacts on the at-sensor total radiance for the LWIR spectrum, but the elevation angle does. It is necessary to integrate the elevation angle into the proposed DL solution as a starting point to better characterize atmosphere and retrieve the target spectra.

3.2 *Atmospheric Component Learning*

The proposed neural network is trained to characterize the atmosphere by estimating four atmospheric radiative components at different elevation angles. Each MODTRAN simulated component is a vector across 256 wavelength bands for a specific elevation angle, time of the day, day of the year, and target, plotted as a single colored line in Fig. 6. For the neural network training purpose, the downwelling radiance L^M_{down} is first separated from the ground reflected radiance $L^M_{\text{grnd\_refl}}$, as Eq. (4) suggests. Then, vectors across 256 wavelength bands at eight selected elevation angles for each component are converted to a matrix, as shown in Fig. 7.

The total number of the training matrix transformed from millions of simulated MODTRAN data is 299,040 (i.e., $365 \times 6 \times 7 \times 4 \times 5$), which represents 365 days of the year, six times of the day, seven targets of different reflectivities under six seasonal or yearly atmospheric models

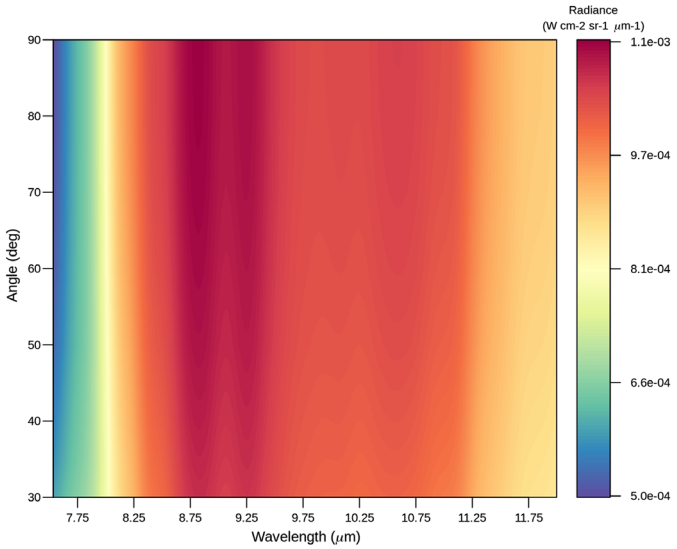

Fig. 7 Total radiance received at the sensor.

(numerically equal to four yearly models) for five radiative components (i.e., one input and four corresponding outputs) (i.e., L_λ, τ_2, L^M_{down}, $L^M_{\text{path\_emit}}$, and $L^M_{\text{path\_scat}}$).

The proposed DL solution is validated using the trained neural network to identify the spectral characteristics of targets that were not part of the training. Polyethylene, retrieved from the

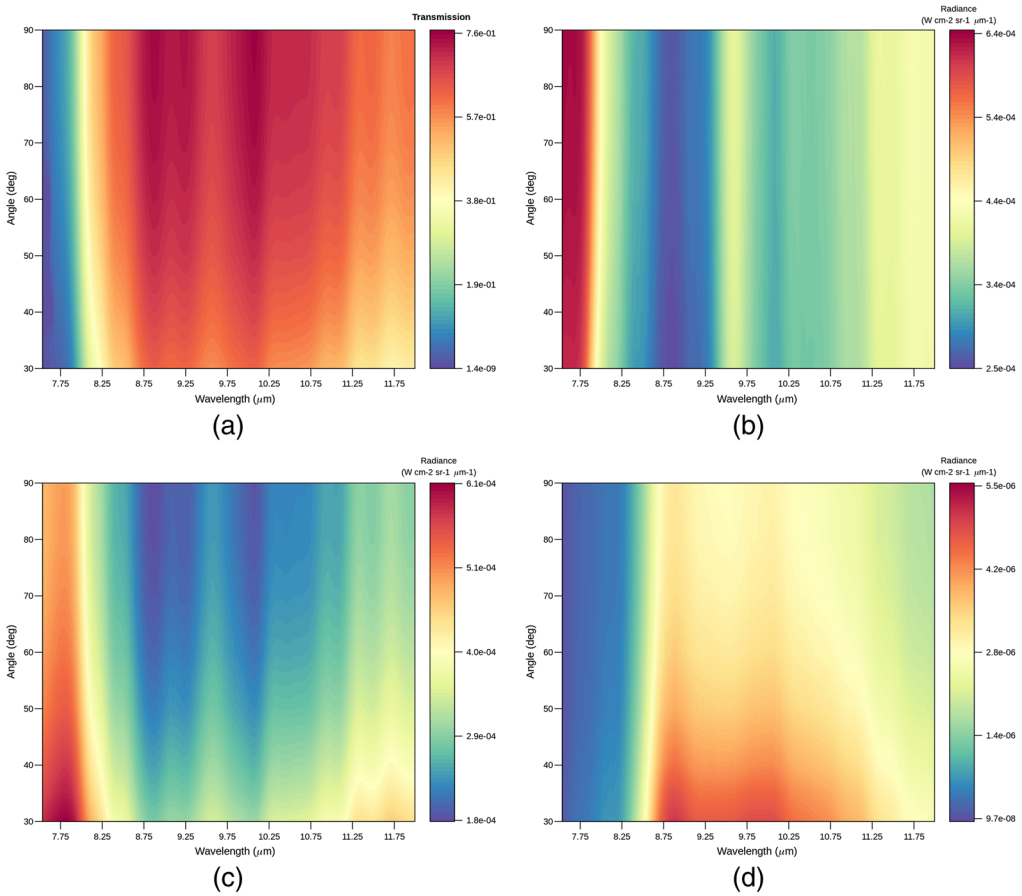

Fig. 8 Predicted atmospheric radiative components. (a) Predicted transmission, (b) predicted downwelling radiance, (c) predicted path thermal emission, and (d) predicted path thermal scattering.

NIST Chemistry WebBook as well as 21 real-world ground targets measured by a contact reflectometer, is used for the validation. Its spectral reflectivities are not constant across the spectrum, which is different from any of the targets simulated in the training dataset with constant reflectivities. However, the trained DL model should be able to characterize the atmospheric parameters (transmission, upwelling, and downwelling) independent of the target reflectivity.

Figure 7 shows the polyethylene's total radiance received at the sensor as input. Figures 8 and 9 are the predicted τ_2, L^M_{down}, $L^M_{\text{path\_emit}}$, and $L^M_{\text{path\_scat}}$ as well as their associated residuals via the trained neural network, given the at-sensor total radiance of polyethylene at different angles as input.

The downwelling radiance has a little bit larger error compared with other components because it is separated from the ground reflected radiance before the training process. Referring to Eq. (4), it makes sense that L^M_{down} is more sensitive to the changes of the target spectra, and the error of the transmission τ_2 will also be directly propagated to it. However, it is clear to see all predicted errors are one or more orders of magnitude smaller than the original signal, which means that the trained neural network is able to characterize the atmosphere at different angles, only given the at-sensor total radiance as input.

3.3 *Target Spectra Retrieval*

The target spectral reflectivity now can be solved with the observed at-sensor total radiance and predicted atmospheric radiative components via Eq. (8). Figure 10 shows the retrieved spectral reflectivities at different elevation angles for the polyethylene from NIST Chemistry WebBook and the large black plastic board from 21 ground measured targets. The associated root

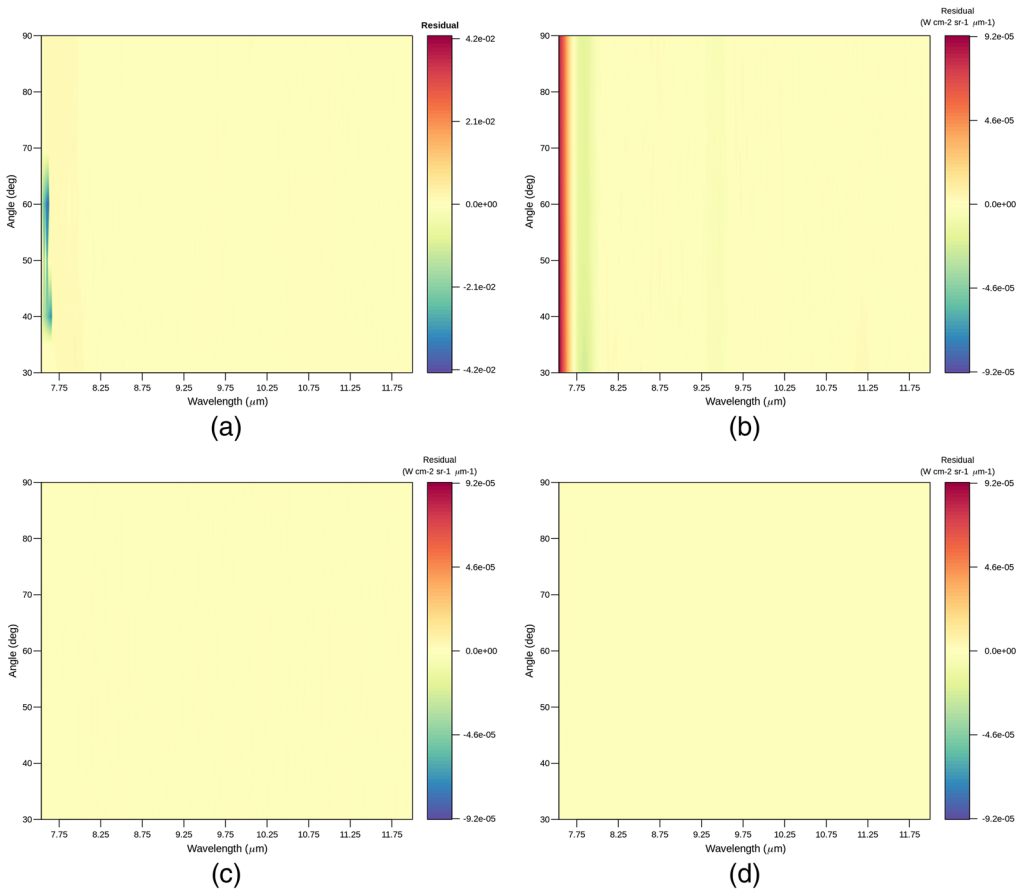

Fig. 9 Residuals of predicted atmospheric radiative components. (a) Residual of transmission, (b) residual of downwelling radiance, (c) residual of path thermal emission, and (d) residual of path thermal scattering.

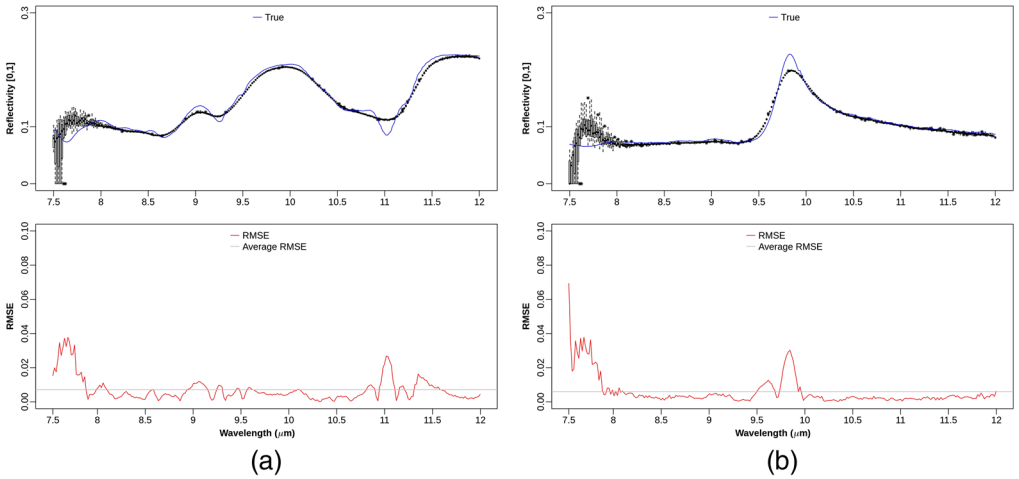

Fig. 10 Retrieved reflectivity at different elevation angles for (a) polyethylene and (b) black plastic board.

mean square errors (RMSEs) between retrieved reflectivities and their true values are also given below.

Several observations are made from comparing the retrieved spectral reflectivities with their true values:

1. The retrieved spectral reflectivities between 7.5 and 7.75 μm have a huge variance at different elevation angles, but they are almost the same for the remaining spectrum. This is due to a larger error of predicted transmission at several angles between 7.5 and 7.75 μm, as shown in Fig. 8
2. The error of retrieved reflectivities becomes larger when the original target reflectivity has a sharp change. For example, the deviation of retrieved reflectivity of the polyethylene from its true values is much larger at around 11 μm; the same occurs for the black plastic board at around 9.7 μm

Overall, the average RMSE of retrieved reflectivities for two real-world materials are both below 0.02 even counting the largest error at the beginning of the LWIR spectrum. This proposed neural network trained on the Lambertian targets with constant reflectivities across the spectrum is thus able to accurately characterize the atmosphere and retrieve the target spectral properties without a detailed atmospheric profile. It is also much more computationally efficient than using a numerical solution to do atmospheric correction and target detection for each pixel in a scene. This additional computational efficiency means that the proposed solution is particularly effective for real-time hardware implementations.

4 Conclusion

This research presents an artificial intelligence/deep learning-based solution to characterize the atmosphere at different vantage points and to retrieve target spectral properties. The results show that this proposed approach can (1) estimate atmospheric components of upwelling, downwelling, and transmission as a function of angles, given as input the at-sensor total radiance, within one order of magnitude errors or less when compared with MODTRAN solutions for the longwave infrared spectrum; (2) retrieve the target spectral properties; and (3) give an estimate of the error for the retrieved target spectra.

In a realistic operational use, the complexities can be summarized as follows: (1) the proposed deep learning solution is theoretically extendable to the full spectrum. However, the LWIR spectrum solutions shown are simpler because the solar components and the influence of varying azimuths are negligible compared with self-emitted thermal components and elevation angles; (2) it might be not possible to collect all of the geometries simulated, and the effect of missing data or interpolated data on the solution is unclear.

Future work will address the two identified shortcomings of extending the current research to full spectrum and developing a computationally tractable solution to deal with missing and incomplete data. Furthermore, a more rigorous testing of the methodology, including field collections of upwelling, downwelling, and transmission to compare the computational solution to a real-world scenario, is planned.

Acknowledgments

This research was supported by the Defense Advanced Research Projects Agency (DARPA) award FA-8650-19-1-7905.

References

1. M. Teke et al., "A short survey of hyperspectral remote sensing applications in agriculture," in *6th Int. Conf. Recent Adv. Space Technol.*, IEEE, pp. 171–176 (2013).
2. D. Landgrebe, "Hyperspectral image data analysis," *IEEE Signal Process. Mag.* **19**(1), 17–28 (2002).
3. D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Process. Mag.* **19**(1), 29–43 (2002).
4. F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.* **42**(8), 1778–1790 (2004).
5. G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.* **43**(6), 1351–1362 (2005).
6. J. B. Campbell and R. H. Wynne, *Introduction to Remote Sensing*, Guilford Press, New York (2011).
7. A. M. Carleton, *Satellite Remote Sensing in Climatology*, Belhaven Press, London (1991).
8. S. M. Adler-Golden et al., "Atmospheric correction for shortwave spectral imagery based on MODTRAN4," *Proc. SPIE* **3753**, 61–70 (1999).
9. H. Mayer, "Automatic object extraction from aerial imagery: a survey focusing on buildings," *Comput. Vision Image Understanding* **74**(2), 138–149 (1999).
10. J. R. Schott, *Remote Sensing: The Image Chain Approach*, Prentice Hall, London (2007).
11. P. Ghamisi et al., "Advances in hyperspectral image and signal processing: a comprehensive overview of the state of the art," *IEEE Geosci. Remote Sens. Mag.* **5**(4), 37–78 (2017).
12. D. Manolakis et al., "Longwave infrared hyperspectral imaging: principles, progress, and challenges," *IEEE Geosci. Remote Sens. Mag.* **7**(2), 72–100 (2019).
13. A. French, J. Norman, and M. Anderson, "A simple and fast atmospheric correction for spaceborne remote sensing of surface temperature," *Remote Sens. Environ.* **87**(2–3), 326–333 (2003).
14. D. Manolakis, D. Marden, and G. A. Shaw, "Hyperspectral image processing for automatic target detection applications," *Lincoln Lab. J.* **14**(1), 79–116 (2003).
15. O. Hagolle et al., "A multi-temporal and multi-spectral method to estimate aerosol optical thickness over land, for the atmospheric correction of FormoSat-2, LandSat, VENμS and Sentinel-2 images," *Remote Sens.* **7**(3), 2668–2691 (2015).
16. M. Z. Salvador, "Expanding the dimensions of hyperspectral imagery to improve target detection," *Proc. SPIE* **9988**, 998807 (2016).
17. G. Cervone and M. Z. Salvador, "Nittany radiance 2019 longwave hyperspectral experiment," in *AGU Fall Meeting*, AGU (2019).
18. B. M. Rankin, J. B. Broadwater, and M. Smith, "Anomalous pixel replacement and spectral quality algorithm for longwave infrared hyperspectral imagery," in *IEEE Int. Geosci. Remote Sensing Symp.*, pp. 4316–4319 (2018).
19. C. Siewert, "A discrete-ordinates solution for radiative-transfer models that include polarization effects," *J. Quant. Spectrosc. Radiat. Transfer* **64**(3), 227–254 (2000).
20. B. Mayer, "Radiative transfer in the cloudy atmosphere," in *EPJ Web Conf.*, EDP Sciences, Vol. 1, pp. 75–99 (2009).
21. C. Emde et al., "The impact of aerosols on polarized sky radiance: model development, validation, and applications," *Atmos. Chem. Phys.* **10**(2), 383–396 (2010).

22. M. Duan, Q. Min, and D. Lü, "A polarized radiative transfer model based on successive order of scattering," *Adv. Atmos. Sci.* **27**(4), 891–900 (2010).
23. E. Zege and L. Chaikovskaya, "New approach to the polarized radiative transfer problem," *J. Quant. Spectrosc. Radiat. Transfer* **55**(1), 19–31 (1996).
24. A. A. Kokhanovsky et al., "Benchmark results in vector atmospheric radiative transfer," *J. Quant. Spectrosc. Radiat. Transfer* **111**(12–13), 1931–1946 (2010).
25. A. Berk et al., "MODTRAN® 6: a major upgrade of the MODTRAN® radiative transfer code," in *6th Workshop Hyperspectral Image and Signal Process.: Evolut. Remote Sens.*, IEEE, pp. 1–4 (2014).
26. M. Castelluccio et al., "Land use classification in remote sensing images by convolutional neural networks," arXiv:1508.00092 (2015).
27. M. Noor et al., "Hyperspectral image enhancement and mixture deep-learning classification of corneal epithelium injuries," *Sensors* **17**(11), 2644 (2017).
28. S. Yu, S. Jia, and C. Xu, "Convolutional neural networks for hyperspectral image classification," *Neurocomputing* **219**, 88–98 (2017).
29. Y. Koga, H. Miyazaki, and R. Shibasaki, "A CNN-based method of vehicle detection from aerial images using hard example mining," *Remote Sens.* **10**(1), 124 (2018).
30. J. Nalepa et al., "Unsupervised segmentation of hyperspectral images using 3-D convolutional autoencoders," *IEEE Geosci. Remote Sens. Lett.* (2020).
31. K. Charmisha, V. Sowmya, and K. Soman, "Dimensionally reduced features for hyperspectral image classification using deep learning," *Lect. Notes Electr. Eng.* **500**, 171–179 (2018).
32. Y. Wei, Y. Zhou, and H. Li, "Spectral–spatial response for hyperspectral image classification," *Remote Sens.* **9**(3), 203 (2017).
33. C. Shi and C.-M. Pun, "Multi-scale hierarchical recurrent neural networks for hyperspectral image classification," *Neurocomputing* **294**, 82–93 (2018).
34. L. Mou, L. Bruzzone, and X. X. Zhu, "Learning spectral-spatial-temporal features via a recurrent convolutional neural network for change detection in multispectral imagery," *IEEE Trans. Geosci. Remote Sens.* **57**(2), 924–935 (2019).
35. A. Signoroni et al., "Deep learning meets hyperspectral image analysis: a multidisciplinary review," *J. Imaging* **5**(5), 52 (2019).
36. A. Paszke et al., "Pytorch: an imperative style, high-performance deep learning library," in *Adv. Neural Inf. Process. Syst.*, pp. 8024–8035 (2019).
37. M. Paoletti et al., "Deep learning classifiers for hyperspectral imaging: a review," *ISPRS J. Photogramm. Remote Sens.* **158**, 279–317 (2019).

Fangcao Xu is a PhD student at Pennsylvania State University. Her research focuses on the development and application of artificial intelligence/deep learning algorithms for hyperspectral imaging analysis in remote sensing. She received her first MS degree in urban spatial analytics from the University of Pennsylvania and her second MS degree in geography from the Pennsylvania State University in 2015 and 2019, respectively.

Guido Cervone is a professor of geography, meteorology, and atmospheric science and is associate director of the Institute for Computational and Data Sciences. His fields of expertise are geoinformatics, machine learning, and remote sensing. His research focuses on the development and application of computational algorithms for the analysis of spatio–temporal remote sensing, numerical modeling, and social media "Big Data."

Gabriele Franch is completing his PhD in computer science between FBK Research Center and the University of Trento, Italy. He has been a visiting scholar at PSU and has worked in FBK and several startups in the last 10 years. Now, he is a researcher at FBK, applying deep learning to weather nowcasting, to help civil protections in Italy to prevent floods.

Mark Salvador is a president and principal scientist at Zi INC, a company focused on applied remote sensing research. His fields of expertise include computational science, remote sensing, and field experimentation. His primary research focus is hyperspectral remote sensing for civil, commercial, and defense operations.