RETRIEVAL OF LEAF AREA INDEX AND LEAF CHLOROPHYLL CONTENT FROM HYPERSPECTRAL DATA USING DEEP LEARNING NETWORKS

B. Hu* [1], W. M. Jung[1], J. Liu [2], J. Shang[2]
[1] Dept of Earth and Space Science and Engineering, York University, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada- baoxin@yorku.ca, wjung1008@gmail.com
[2] Agriculture and Agri-Food Canada, 960 Carling Avenue, Ottawa, Ontario, K1A 0C6, Canada- (jiangui.liu, Jiali.Shang) @agr.gc.ca

Commission III, WG III/4

KEY WORDS: Deep learning, Convolutional neural network, Autoencoder, Leaf area index, Leaf chlorophyll content, Hyperspectral

ABSTRACT:

This study aimed to exploit the use of deep learning networks in the retrieval of the biophysical and biochemical parameters of vegetation canopies. Convolutional Neural Network (CNN), network with only fully connected layers, referred as dense network (DNN), and Autoencoder (AE) were investigated to retrieve leaf area index (LAI) and leaf chlorophyll content. Hyperspectral data simulated by the coupled PROSPECT and SAIL model were used for training and validation. The real CASI hyperspectral data in 50 spectral channels ranging from 522.4 nm to 894.2 nm collected over three agricultural crop fields during the growing season of 2001 were used, together with the in-situ measured LAI and leaf chlorophyll content, as independent test set. Occlusion analysis was also employed to determine the important spectral bands at which reflectance made more contributions to the retrieval with a CNN and interpret the latent variables of the AE. Satisfactory results from these deep learning networks were obtained, compared with ground truth. The DNN with the input of the vegetation indices sensitive to LAI and leaf chlorophyll content ($MTVI2$ and $TCARI/OSAVI$) generated the best results with $R^2$of 0.86 for LAI and 0.55 for leaf chlorophyll content.

1. INTRODUCTION

Biophysical and biochemical parameters of vegetation canopies, such as leaf area index (LAI) and leaf chlorophyll content, play crucial roles in precision agricultural management, forest ecology monitoring, and global climate change studies. Accurate and robust retrieval of these parameters from remotely sensed data remains a challenge. The commonly used methods are based on either statistical regression or physical model inversion (Ali et al., 2020). Feature selection is critical and challenging in the existing methods. Statistical regression methods are based on the empirical relationship between the variable of interest, such as LAI, and the selected features. In physical model inversion, the selected features are used in the cost function. Deep learning, in contrast, involves automatic learning from examples, allowing features to be extracted directly from data. The potential of deep learning is therefore attracting a lot of attention in the field of remote sensing (Zhang et al., 2016). Considerable efforts in the application of deep learning are put into qualitative interpretation of remotely sensed data. Yet, deep learning in quantitative remote sensing is largely unexplored (Tan et al., 2019 and Apolo-Apolo et al., 2020). This objective of this study was to exploit the use of deep learning networks in the retrieval of LAI and leaf chlorophyll content of vegetation canopies. Convolutional Neural Network (CNN), network with only fully connected layers, referred as dense network (DNN), and Autoencoder (AE) (LeCun et al 2015) were investigated. We aimed to (1) establish strategies, workflows, and configurations of CNN, DNN and AE for the retrieval of LAI and leaf chlorophyll content of agricultural crops, (2) provide interpretation to the extracted features and retrieval results obtained from CNN, DNN and AE models, and (3) evaluate CNN, DNN and AE, in the comparison with the commonly used statistical regression.

2. DATA USED

For this study, both real and simulated hyperspectral remotely sensed data were used. The real data were acquired by a Compact Airborne Spectrographic Imager (CASI) over the former Greenbelt Farm of Agriculture and Agri-Food Canada, Ottawa (45˚18' N, 75˚45' W) in 2001 during three intensive field campaign (IFC) on June 13 (IFC1), June 26 (IFC2), and July 19 (IFC 3). The study area included three crops, corn, wheat, and soybean. The data covered spectral regions from 500 nm to 1044.3 nm in 72 bands and had a spatial resolution of 2 m by 2m. Due to presence of noise in the first and last several bands, only 50 spectral bands covering from 522.4 nm to 894.2 nm were used. CASI data were processed to surface reflectance through calibration, atmospheric correction, and geo-referencing (Haboudane et al., 2004). Simultaneously with CASI data acquisition, ground truth was collected at fourteen ground locations that are marked on the false color composite of CASI images shown in Figure 1. A LI-COR LAI-2000 Plant Canopy Analyzer set was used to quantify non-destructive total LAI and a LI_COR LI-3000 area meter for determining separately destructive LAI of green and dead leaves. The destructive leaf area measurements were used to calculate the proportion of green LAI. At these locations, leaf tissues were also collected for laboratory determination of leaf chlorophyll concentration. Details of the field measurements were presented in Pattey et al. (2001) and Strachan et al (2002).

Since limited ground truth data were available, labelled data for the training of deep learning networks were generated by the widely used PROSPECT and SAIL models and hereafter they are referred as the simulated data. The PROSPECT model (Jacquemoud and Baret, 1990) assumes a uniform distribution of water and pigments throughout the leaves and a constant leaf surface roughness. Its model outputs, leaf reflectance and

* Corresponding author

transmittance are used as the inputs for the SAIL model. The SAIL model (Verhoef, 1984) assumes vegetation canopies as uniform infinitely extended turbid media with small Lambertian scatters randomly distributed within the canopies. In addition to leaf reflectance and transmittance, its input parameters also include the LAI, leaf angle distribution function, soil reflectance, and imaging geometry (solar zenith angle view zenith angle and relative azimuth angle). The coupled PROSPECT and SAIL model was employed to generate a series of canopy reflectance values at the wavelengths corresponding to those in CASI data. Several simulation scenarios were exploited in this study and the one used is list in Table 1. The imaging geometry simulated was consistent with CASI data acquisition. Soil reflectance was measured on the ground using a field spectrometer. In the current study, only LAI and leaf chlorophyll content were kept as variables and retrieved to be consistent with the CASI data covering only the visible and near-infrared spectral ranges where canopy reflectance is mainly sensitive to LAI and leaf chlorophyll content.

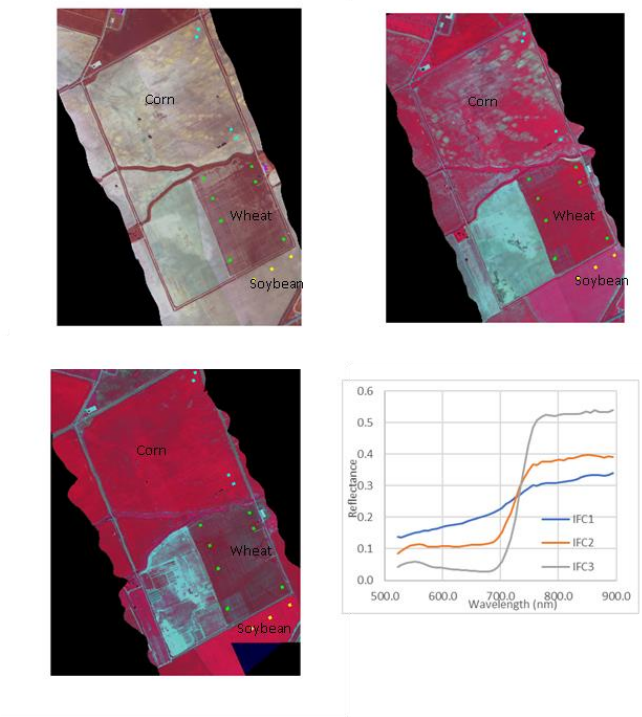

Figure 1. The false color composite images of IFC1 (June 13), IFC2 (June 26), and IFC 3 (July 19) in 2001 (clockwise) together with the locations where ground truth data were collected. The reflectance spectra are the mean values from the uppermost ground location in the corn field.

3. METHODOLOGY

3.1 CNN for the retrieval of LAI and leaf chlorophyll content

For the CNN, the input was the reflectance at 50 spectral channels. We attempted many configurations and all of them tended to have the problem of overfitting. For this paper, the network with the fewest parameters was used. There were three convolution layers with 16, 32, and 64 filters, respectively and each filter is 3 by 1. Each convolution layer was followed by a max-pooling layer. There were four dense layers with the dimensions of 128, 64, 32, 16, respectively. The output had two nodes representing LAI and leaf chlorophyll content. The summary of the model is provided in Table 2.

The training of the CNN model was done with the previously described simulated data. There were in total 9768 samples and 90% was used for training and the rest for validation.

Parameters	Values
Leaf chlorophyll content	5 to 70 at step of 5 μg/cm^2
Internal structural parameter	1.55
Dry matter content	0.0035 μg/cm^2
Leaf water content	0.01 cm
LAI	0.05 to 0.1 at step of 0.01 and 0.1 to 0.7 at step of 0.1
Leaf angle distribution function	Spherical
Soil reflectance	Field measured
Solar angle	23˚
View angle	Nadir (0˚)
Relative azimuth angle	0˚

Table 1. Simulation input

Layer	Output shape	Parameter number
Input	(50,1)	0
Convolution 1	(50, 16)	64
Max Pooling	(25, 16)	0
Convolution 2	(25, 32)	1568
Max Pooling	(13, 32)	0
Convolution 3	(13, 64)	6208
Max Pooling	(7, 64)	0
Flatten	448	0
Dense 1	128	57472
Dense 2	64	8256
Dense 3	32	2080
Dense 4	16	528
Dense 5 (to output)	2	34

Table 2. Summary of the CNN model used

3.2 DNN for the retrieval of LAI and leaf chlorophyll content

In this study, we also attempted to use the fully connected layers only. For this type of networks, we did three experiments: one with the same input as that for CNN (i.e the reflectance at 50 spectral channels) and the other two with vegetation indices as input. For the first one, the network often did not converge for all of the configurations attempted. As a result, we omit it in this paper. In the second experiment, two commonly used vegetation indices developed to estimate leaf chlorophyll content and LAI were employed as input. They are the $MTVI2$ (Equation (1) that was proposed by Haboudane et al. (2004) to retrieve LAI and $\frac{TCARI}{OSAVI}$ (Equation 2) by Haboudane et al. (2002) for leaf chlorophyll content.

$$MTVI2 = \frac{1.5[1.2\,(R_{800}-R_{550})-2.5(R_{670}-R_{550})]}{\sqrt{(2R_{800}+1)^2-(6R_{800}-5\sqrt{R_{670}})-0.5}}, \quad (1)$$

where R=reflectance with the subscript as wavelength in nm

$$TCARI = 3\left[(R_{700}-R_{670})-0.2(R_{700}-R_{550})\left(\frac{R_{700}}{R_{670}}\right)\right];$$

$$OSAVI = \frac{(1+0.16)(R_{800}-R_{670})}{(R_{800}+R_{670}+0.16)}, \quad (2)$$

where R=reflectance with the subscript as wavelength in nm

For the network, there were four fully connected layers with the numbers of nodes of 64, 32,16, and 8, respectively. The input layer had two nodes taking the values of these two indices. In the third experiment, in additional to $MTVI2$ and $\frac{TCARI}{OSAVI}$, there indices in the form of normalized difference vegetation index were used. They were $\frac{R_{750}-R_{705}}{R_{750}+R_{705}}$, $\frac{R_{850}-R_{730}}{R_{850}+R_{730}}$, $\frac{R_{850}-R_{570}}{R_{850}+R_{570}}$. The network configuration was same as that in the second experiment except for the input.

3.3 AE for the retrieval of LAI and leaf chlorophyll content

For the AE, both the input and output were the reflectance at 50 spectral channels. The extracted features to reconstruct the reflectance spectra were represented by the values in the latent nodes. This process was unsupervised. Table 3 is the summary of the encoder and decoder.

The training of the AE was attempted using two sets of data. The first one was the previously described simulated data (9768 samples) and the second derived from real CASI imagery (115728 samples). Similar to the training of the CNN, 90% of the samples was used for training and the rest for validation. Once trained, the latent variables were used to predict leaf chlorophyll content and LAI. This was a supervised process. Two linear models were established to predict LAI and leaf chlorophyll content, respectively. The predictors were non-zero latent variables. For the investigation of using real CASI data for AE training, 12 of the samples with ground truth were used to determine the coefficients of the linear models and the rest for testing. The selection of these 12 samples had an impact on the results depending on its dynamic ranges in LAI and leaf chlorophyll content. For this study, they were randomly selected.

Encoder		
Layer	Output shape	Parameter number
Input	(50,1)	0
Convolution 1	(50, 16)	64
Max Pooling	(25, 16)	0
Convolution 2	(25, 32)	1568
Flatten	800	0
Dense	8	6408
Decoder		
Input	8	0
Dense	800	7200
Re-shape	(25,32)	0
Convolution 1	(25, 32)	3104
Up sampling	(50, 32)	0
Convolution 2	(50, 16)	1552
Convolution 3	(50,1)	49

Table 3. Summary of the AE used

3.4 Evaluation of the deep learning networks

As mentioned earlier, samples derived from CASI imagery were used for independent testing and there were 14 plots where ground truth was collected. For three field campaigns, there were 42 samples for LAI and 32 for leaf chlorophyll content due to missing measurements. Scatter plots of the estimated values vs. field measurements were made and R^2 values were calculated.

In addition, the performance of the deep learning networks was also compared with the method using empirical relationship between LAI and leaf chlorophyll content and vegetation indices. In this study, LAI and leaf chlorophyll content were calculated using indices in Equations 3 and 4, respectively (Haboudane et al., 2002 and 2004).

$$LAI = 0.2227e^{3.6566*MTVI2} \quad (3)$$

$$Chl_{ab} = -30.194ln\left(\frac{TCARI}{OSAVI}\right) - 18.363 \quad (4)$$

where R=reflectance with the subscript as wavelength in nm

3.5 Occlusion sensitivity analysis of the trained CNN

To gain understanding of the learned CNN network, the occlusion sensitivity technique (Zeiler and Fergus, 2014) was used. This technique is commonly used for image classification and implemented by masking out a spatial region one by one to reveal whether the masked region affect the classification result. In this study, occlusion was applied to individual spectral bands to show the influence of the reflectance in each band on the retrieval of LAI and leaf chlorophyll content. 97 simulated samples were used for this investigation. It is worth mentioning that these samples were not used for the training. With the trained CNN, the LAI and leaf chlorophyll content were predicted for these 97 simulated samples and used as baseline for comparison. In the occlusion sensitivity analysis, for a given band to be occulated, a random number between 0 and 1 was generated for each of the 97 samples. The values for LAI and leaf chlorophyll content for these samples were then predicted by the trained CNN. The difference in the form of root mean square error (RMSE) between the baseline LAI and leaf chlorophyll content and the predicted ones was calculated. A large RMSE meant the reflectance in this particular band had a large impact on the retrieval of LAI and leaf chlorophyll content. In this investigation, the RMSE values were calculated for the LAI and leaf chlorophyll content separately. In addition to comparing with the baseline data, the retrieved values in the occlusion analysis were also compared with the true LAI and leaf chlorophyll content that were used for the simulation. Similar analysis was also applied to the samples derived from CASI hyperspectral imagery (the independent test set).

3.6 Interpretation of AE latent variables

To understand the meaning of the AE latent variables, we plot the reconstructed reflectance spectrum for one pixel (as an example) by setting individual non-zero latent variables zero and by keeping individual non-zero latent variables (set other variables zero). We also calculated the RMSE between the reflectance of individual spectral bands reconstructed by removing or keep individual latent variables and the reference obtained using all latent variables.

4. RESULTS AND ANALYSIS

4.1 The retrieved LAI and leaf chlorophyll content

This analysis was done for the AE trained using the samples extracted from real CASI imagery. The scatter plots between the estimated LAI and leaf chlorophyll content and those measured

on the ground and laboratory are shown in Figures 2-6 for CNN, DNN with two indices, DNN with five indices, and AE-linear models, respectively. The comparisons between the estimated LAI and leaf chlorophyll content obtained based on the empirical relationships with vegetation indices of $MTVI2$ and $\frac{TCARI}{OSAVI}$ are shown in Figure 7. From these figures, the following observations could be made.

(1) The R^2 values were higher for LAI than for leaf chlorophyll content for all methods, compared with the ground truth.

(2) Compared with the traditional method (based on the empirical relationship with vegetation indices), all supervised methods (CNN, DNN) performed much better in the estimation of leaf chlorophyll content with the R^2values of around 0.5 vs 0.27. However, the R^2values remained low. For LAI, CNN obtained a comparable R^2 value (0.72 vs 0.73), but the DNNs had better results, especially for DNN with 2 indices with the R^2 value as high as 0.82.

(3) The DNN with two vegetations indices ($MTVI2$ and $TCARI/OSAVI$) generated the best result with R^2 of 0.86 for LAI and 0.55 for leaf chlorophyll content. Increasing the number of vegetation indices did not produce improved results in the DNN.

(4) With the same vegetation indices, $MTVI2$ and $TCARI/OSAVI$, a dense layer network had a much higher R^2 in the prediction of LAI (0.86) and leaf chlorophyll content (0.55), compared with that (0.73 and 0.28 for LAI and leaf chlorophyll content, respectively) obtained based on the empirical relationships reported in the literature (Haboudane et al., 2002 and 2004). It is worth mentioning that the vegetation indices and their empirical relationships with LAI and leaf chlorophyll content reported in Haboudane et al. (2002 and 2004) were developed for CASI imagery over the same area.

(5) Among the deep learning networks, AE-linear models performed poorly for the retrieval of LAI but had a comparable result for the retrieval of leaf chlorophyll content. In term of the comparison between the two training strategies investigated, the performance was better for leaf chlorophyll content, and similar for LAI by using the real CASI data, compared with that by using the simulated data. By using real CASI data, the AE was trained by a large number of samples to better represent the real data. However, the linear estimation of LAI and leaf chlorophyll content and was done with a limited number of training samples (10). It is worth mentioning that the samples used for generating Figure 6 is different from those for other figures (1-5 and 7).

It is also worth mentioning that for all three deep learning networks investigated in this study, a very strong 1:1 correlation in the LAI and leaf chlorophyll content was observed based on training and validation data (both from simulation). This indicated a poor generalization for these models.

4.2 The important spectral bands in the trained CNN

The results of the occlusion sensitivity analysis are shown in Figure 8 for the simulated data and in Figure 9 for CASI hyperspectral data. The top two panels were the results obtained by comparing the estimated LAI and leaf chlorophyll content with truth (Figure 8) and ground truth (Figure 9), while the bottom two were based on the comparison with the baseline obtained from undisturbed data.

It is revealed by Figure 8 that based on the simulated data, the similar results were obtained no matter whether the truth or the baseline as the reference for comparison in the occlusion analysis. This is because the results from the undisturbed data were very close to the truth. This observation was consistent with that could made from this figure, which was any disturbance to the reflectance of any spectral bands would lead to a higher RMSE in the retrieval of LAI and leaf chlorophyll content using the trained CNN. In contrast, by comparing the top and bottom two panels in Figure 9, one can notice different phenomena. When the estimated LAI and leaf chlorophyll content were compared with the ground measurements, it was observed that disturbance in some spectral bands increased the RMSE compared with un-disturbed data (horizontal lines), which for some spectral bands, the RMSE was reduced. In addition, the important wavelengths revealed by the top two panels and those by the bottom two were different. The former used ground truth as the reference and the latter employed the LAI and leaf chlorophyll content obtained from undisturbed data. The above observations might further demonstrate that the simulated data could not fully represent the real CASI observations.

A closer examination of Figure 8 revealed that based on the simulated data, the reflectance at the following wavelengths had stronger influence on the estimation of leaf chlorophyll content (from strong to weak) 672.4 nm, 755.7 nm, 763.3 nm, 732.9 nm, and 770.9 nm. These wavelengths are in the red and red-edge region of the spectrum. For the estimation of LAI, the reflectance in the red region from 627.2 nm to 679.9 nm had the strongest impact, and the reflectance in the visible region had a stronger influence than that in the Near infrared. The same observations could be made from Figure 9 by using the results from the undisturbed data as reference. Reflectance in these wavelengths is known to be sensitive to leaf chlorophyll content and/or LAI. The disagreement with the accumulated knowledge in the literature is that reflectance in the near infrared had a less impact on the retrieval of LAI than others. It might be caused by some bias exiting in the simulated data.

Using the ground truth as the reference (top two panels in Figure 9), the important wavelengths for the retrieval of LAI were in the near infrared region, while they were around 552, and 710nm for leaf chlorophyll content.

4.3 The meaning of latent variables of the trained AE

For a random selected test pixel in the CASI imagery, the original reflectance and that reconstructed by setting individual non-latent variables zeros are shown in Figure 10. The reflectance reconstructed using all latent variables was very similar to the original reflectance. For the trained AE, nodes 1, 4, and 6 were non-zeros. For the same pixel, the reflectance spectra reconstructed by keeping individual latent variables (others set to 0) are shown in Figure 11. The RMSE between the reflectance of each spectral band reconstructed using all latent variables (reference) and that using individual latent variables over the test samples is shown in Figure 12. The RMSE between the reference and the reflectance reconstructed by setting individual latent variables zeros over the test samples is shown in Figure 13.

As shown in Figures 10 and 11, the higher reflectance of green leaves in the near infrared band was reflected by latent variable (node) 4; latent variable 1 captured the absorption of leaf chlorophyll content of the red radiation of green leaves; and the average reflectance over the spectral bands was reflected by node 6. These observations made based on one test sample were

consistent with the RMSE calculated over all test samples (Figures 12 and 13). For example, as shown in Figure 12, by setting latent variable 1 zero, the largest RMSE occurs around 680 nm, the absorption bands of leaf chlorophyll.

5. CONCLUSIONS

In this study, three deep learning networks, CNN, DNN and AE, were exploited to retrieve LAI and leaf chlorophyll content. Learning from simulated data, deep learning networks investigated in this study, CNN, DNN, and AE showed promising results in the retrieval of LAI and leaf chlorophyll content using CASI imagery. The deep dense network (with fully connected layers only) with the input of the vegetation indices sensitive to LAI and leaf chlorophyll content ($MTVI2$ and $TCARI/OSAVI$) generated the best results. Compared with the ground truth, the R^2 was 0.86 for LAI and 0.55 for leaf chlorophyll content. This demonstrated the importance of domain knowledge in the applications of deep learning in the retrieval of parameter retrieval.

Occlusion analysis showed that the wavelengths at which reflectance had the most contribution toward the estimation of LAI and leaf chlorophyll content using a CNN were consistent with those known in the literature (sensitive to either parameter). On-going investigations are focused on adding more simulated data in the training process, seeking additional real hyperspectral data with ground truth for training and independent testing, and developing methodology to incorporation of domain knowledge in the network architecture. The results obtained by the analysis of the AE implied that reflectance spectra could be reconstructed by the latent variables. Through linear modelling, these latent variables could be used to estimate LAI and leaf chlorophyll content with reasonable accuracies. However advanced methods are needed to improve the accuracies.

In this study, limited ground truth was used to validate the estimated LAI and leaf chlorophyll content. Further validation is needed to reach more reliable conclusions.

ACKNOWLEDGEMENTS

We would like to thanks The Natural Sciences and Engineering Research Council (NSERC) of Canada for financial support.

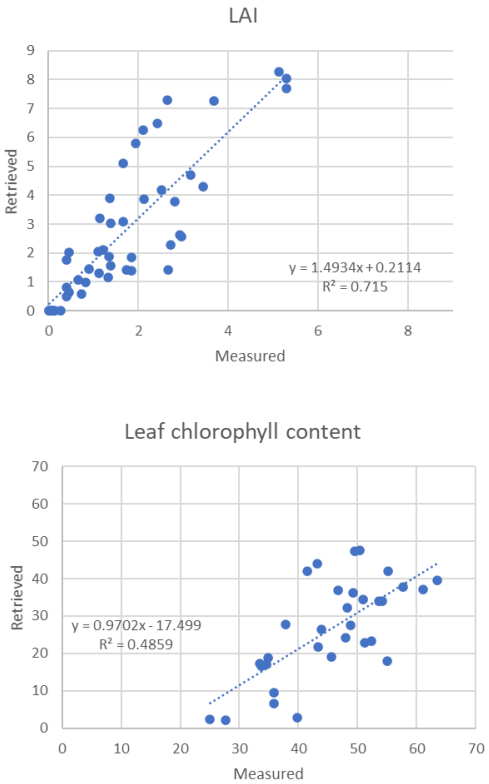

Figure 2. The scatter plots of the estimated LAI (top) and leaf chlorophyll content (bottom) using the CNN versus the ground measured values.

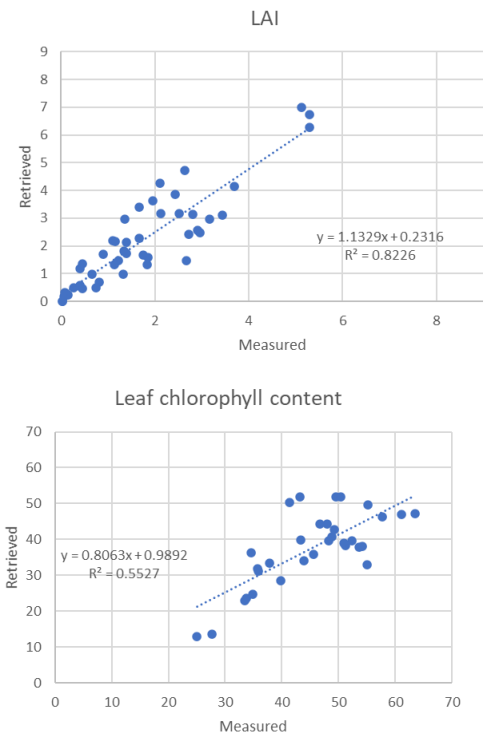

Figure 3. The scatter plots of the estimated LAI (top) and leaf chlorophyll content (bottom) using the DNN (2 indices) versus the ground measured values.

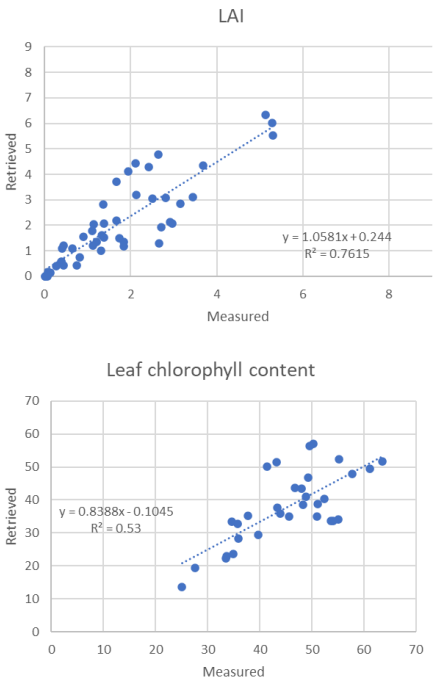

Figure 4. The scatter plots of the estimated LAI (top) and leaf chlorophyll content (bottom) using the DNN (five indices) versus the ground measured values.

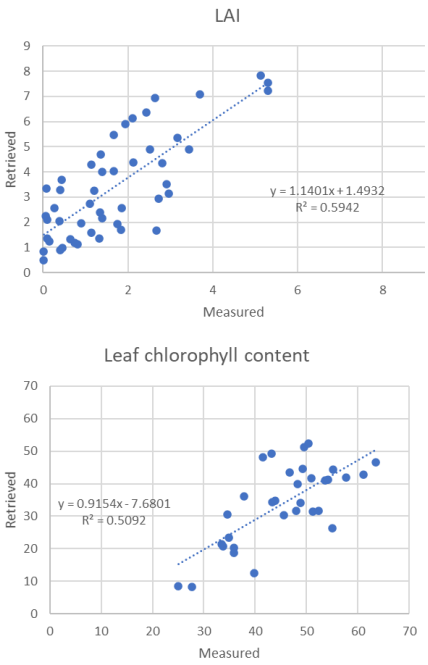

Figure 5. The scatter plots of the estimated LAI (top) and leaf chlorophyll content (bottom) using the AE and linear models versus the ground measured values. The training of AE was carried out using simulated data.

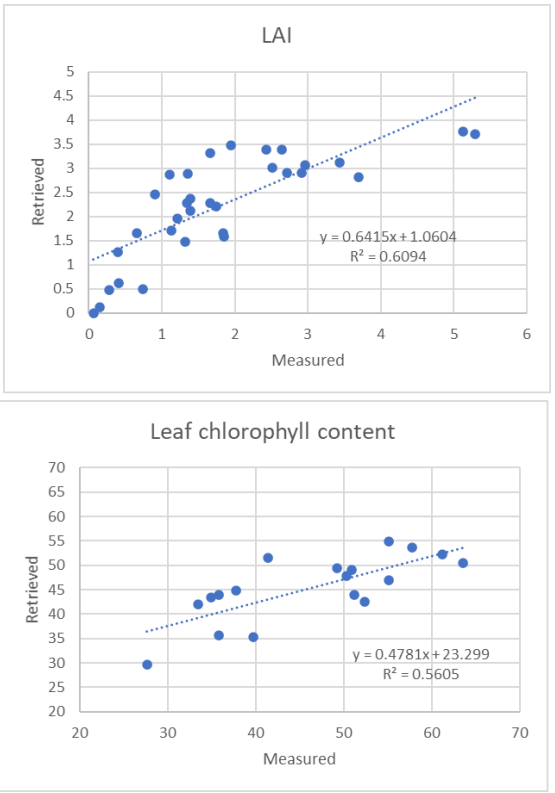

Figure 6. The scatter plots of the estimated LAI (top) and leaf chlorophyll content (bottom) using the AE and linear models versus the ground measured values. The training of AE was carried out using real CASI data.

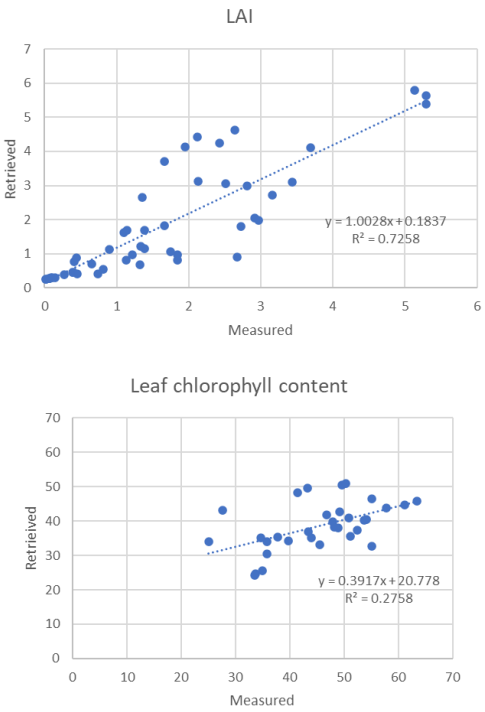

Figure 7. The scatter plots of the estimated LAI (top) and leaf chlorophyll content (bottom) using the empirical relationships with $MTVI2$ and $\frac{TCARI}{OSAVI}$, respectively.

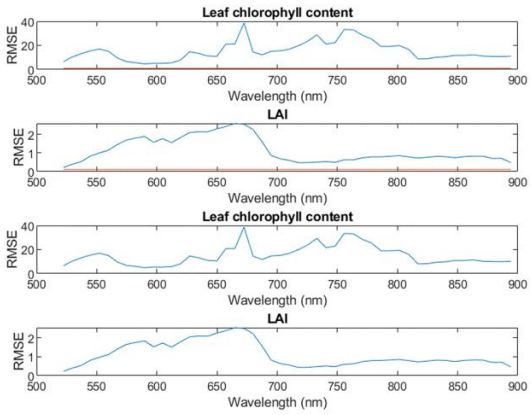

Figure 8. The effect of reflectance in individual spectral bands on the retrieval of LAI and leaf chlorophyll content based on the simulated data. Top two panels: the Root Mean Square Error (RMSE) between the retrieved leaf chlorophyll content and LAI and the truth. The horizontal lines represent the RMSE between the values estimated from the original simulated data (baseline, without any occlusion) and the truth. The bottom two panels: the RMSE in the retrieved values between the baseline and using data with random values for individual spectral bands.

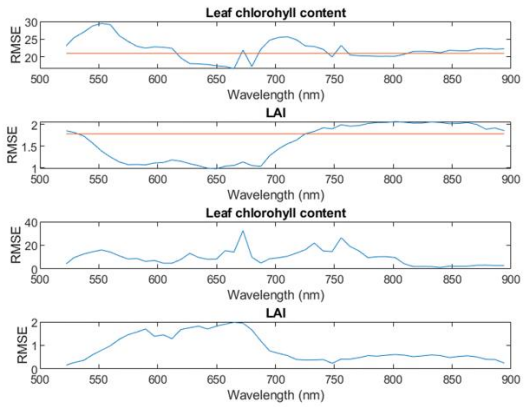

Figure 9. The effect of reflectance in individual spectral bands on the retrieval of LAI and leaf chlorophyll content based on hyperspectral CASI data. Top two panels: the RMSE between the retrieved leaf chlorophyll content and LAI and the ground measurements. The horizontal lines represent the RMSE between the values estimated from the original CASI data (baseline, without any occlusion) and the truth. The bottom two panels: the RMSE in the retrieved values between the baseline and using data with random values for individual spectral bands.

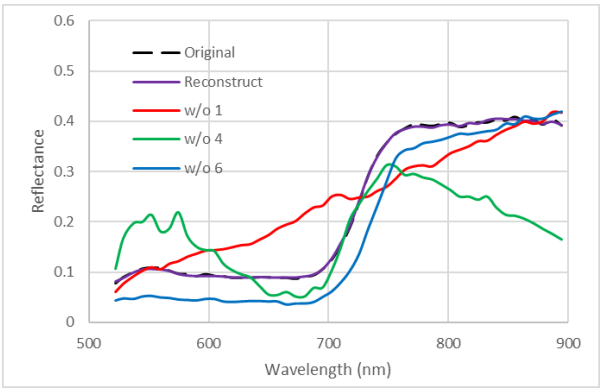

Figure 10. The original reflectance of a randomly selected test pixel, the reflectance reconstructed by using all latent nodes, and that by setting individual latent variables as zero. There were three non-zero latent notes (1, 4, and 6). The notation of w/o 1, as an example, means the reflectance was reconstructed by setting latent node 1 as zero.

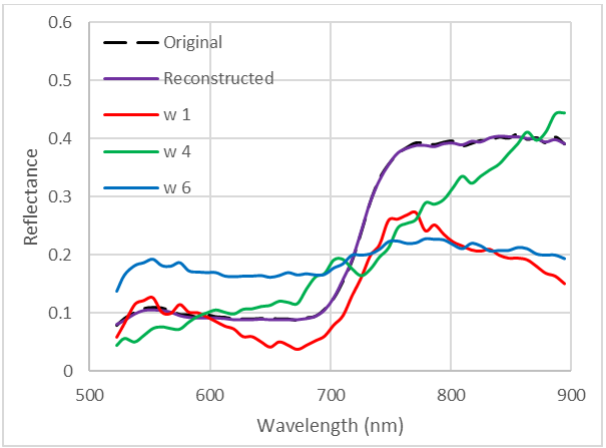

Figure 11. The original reflectance of a randomly selected test pixel, the reflectance reconstructed by using all latent nodes, and that by keeping individual latent variables (setting others as zero). There were three non-zero latent notes (1, 4, and 6). The notation of w 1, as an example, means the reflectance was reconstructed by only using latent node 1.

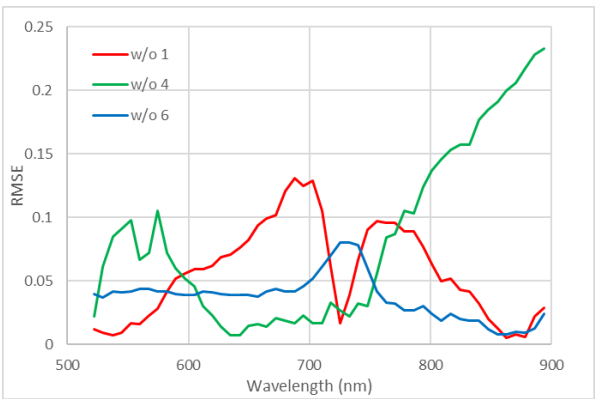

Figure 12. The effects of the three non-zero latent variables exhibited by the RMSE calculated using all test samples in the reflectance of each spectral band between the reflectance reconstructed using all latent variables (reference) and that by setting individual variables zero.

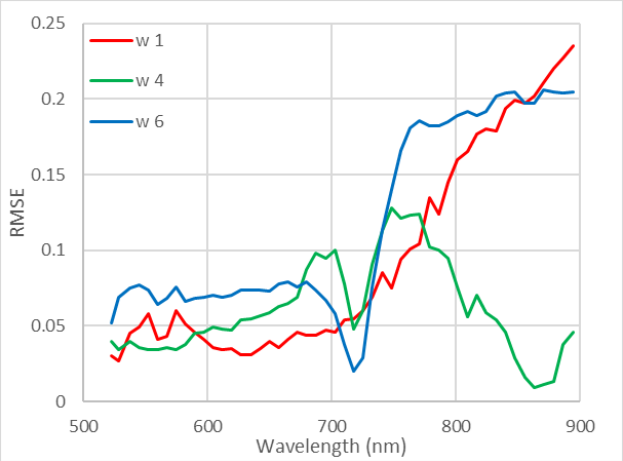

Figure 13. The effects of the four non-zero latent variables exhibited by the RMSE calculated using all test samples in the reflectance of each spectral band between the reflectance reconstructed using all latent variables (reference) and that using individual variables (setting others zero). The notation of w 1, as an example, means the reflectance was reconstructed by only using latent node 1.

REFERENCES

Ali, A. M., Darvishzadeh, R. et al, 2020. Machine learning methods' performance in radiative transfer model inversion to retrieve plant traits from Sentinel-2 data of a mixed mountain forest. *International Journal of Digital Earth*, DOI: 10.1080/17538947.2020.1794064.

Apolo-Apolo, O.E., Pérez-Ruiz, M. et al., 2020. A mixed data-based deep neural network to estimate leaf area index in wheat breeding trials. *Agronomy* 10, 175.

Jacquemoud, S., and Baret, F., 1990. Prospect: A model for leaf optical properties spectra. *Remote Sens. Environ*., 34, 75–91.

LeCun, Y., Bengio, Y. and Hinton, G., 2015. Deep learning. *Nature*, 521(7553), 436-444

Pattey, E., Strachan, I. B., Boisvert, J. B., Desjardins, R. L., and McLaughlin, N., 2001. Effects of nitrogen application rate and weather on corn using micrometeorological and hyperspectral reflectance measurements. Agric. For. *Meteorol*., 108, 85– 99.

Strachan, I. B., Pattey, E., and Boisvert, J. B., 2002. Impact of nitrogen and environmental conditions on corn as detected by hyperspectral reflectance. *Remote Sens. Environ*., 80, 213– 224.

Tan, J., NourEldeen, N., et al., 2019. Deep learning convolutional neural network for the retrieval of land surface temperature from AMSR2 data in China. *Sensors*, 19, 2987.

Verhoef, W., 1984. Light scattering by leaf layers with application to canopy reflectance modeling: The SAIL model. *Remote Sens. Environ*., 16, 125– 141.

Zeiler M.D., and Fergus R., 2014. Visualizing and Understanding Convolutional Networks. In: Fleet D., Pajdla T., Schiele B., Tuytelaars T. (eds) Computer Vision – ECCV 2014. ECCV 2014. Lecture Notes in Computer Science, vol 8689. Springer, Cham.