ARL-TR-9475 ● JUNE 2022

DEVCOM ARMY RESEARCH LABORATORY

In-situ Atmospheric Intelligence for Hybrid Power Grids: Volume 6 (Convolutional Neural Networks for Whole Sky Imager Data Analysis)

by Michael S Lee, Michael S D'Arcy, and Gail T Vaucher

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

ARL-TR-9475 ● JUNE 2022

In-situ Atmospheric Intelligence for Hybrid Power Grids: Volume 6 (Convolutional Neural Networks for Whole Sky Imager Data Analysis)

Michael S Lee, Michael S D'Arcy, and Gail T Vaucher
DEVCOM Army Research Laboratory

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED (From - To)
June 2022	Technical Report	November 2020–March 2022

4. TITLE AND SUBTITLE
In-situ Atmospheric Intelligence for Hybrid Power Grids: Volume 6 (Convolutional Neural Networks for Whole Sky Imager Data Analysis)

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Michael S Lee, Michael S D'Arcy, and Gail T Vaucher

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
DEVCOM Army Research Laboratory
ATTN: FCDD-RLC-ED
Aberdeen Proving Ground, MD 21005

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-TR-9475

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES
ORCID ID: Michael S Lee, 0000-0002-0419-6069

14. ABSTRACT

Automated analysis of whole sky imager (WSI) images can provide atmospheric intelligence for near-future solar power generation. Atmospheric intelligence in the form of input to solar radiation models can subsequently inform optimization algorithms of an energy management system, enabling efficient and effective power distribution. In this work, we developed three separate convolutional neural network models to assess WSI images including image segmentation by cloud type, estimation of thin/thick cloud percentage, and inference of current solar radiation. A cloud segmentation model was trained with limited cloud example data and was moderately successful, except when identifying a bright clear sky. A cloud quantification model was trained with hand-labeled percentages and was accurate enough for use in energy management system simulations. Finally, a solar radiation model was trained on a year-round data set from the National Renewable Energy Laboratory. This model was able to infer the total solar radiation within 50 W/m^2 using a camera sensor limited to only visible wavelengths. Future studies will include model accuracy improvements with more training data and new prediction models of future sky conditions based on past WSI image sequences.

15. SUBJECT TERMS
machine learning, convolutional neural networks, deep learning, whole sky imager, WSI images, cloud detection, solar radiation, percent cloud cover, Military Information Sciences

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Michael S Lee
Unclassified	Unclassified	Unclassified	UU	24	19b. TELEPHONE NUMBER (Include area code) (410) 278-5888

Contents

List of Figures

List of Tables

Acknowledgments

We would like to thank Dr R Jane, Mr M Berman, Dr J Hyatt, and Mr M Ziemann for helpful discussions, and Professor Gordon Parker and Mr Thomas Price for their energy management system simulation applications.

This work was supported in part by high-performance computer time and resources from the DOD High Performance Computing Modernization Program.

1. Introduction

Automated computational evaluation of cloud cover using machine learning (ML) is an important postprocessing tool for the analysis of satellite weather imagery[1] and ground-based whole sky imager (WSI) imaging.[2] One benefit of WSI imagery over satellite imagery is it can ascertain local sky conditions in remote locations. This extracted knowledge can be leveraged by hybrid energy optimization algorithms, where solar and wind power are part of the power resources.

The original ML algorithms developed for satellite imagery combined manually designed feature/texture detectors with low-parameter classification schemes such as support vector machines and logistical regression.[1,3-7] Recent advances in ML, specifically, convolutional neural networks (CNNs), popularized by the success of AlexNet,[8] allow for automated learning of relevant textural features to help maximize classification accuracy, albeit at the expense of higher computational cost, more trainable parameters, and subsequently more required training data.

Different types of ML-based analysis can be applied to WSI images,[2,9–10] starting with a simple image classification of an entire image. For a detailed classification of features in an image at a pixel-by-pixel level, a semantic segmentation (SS) model can be used.[11] SS has become practical thanks to the development of CNNs, and the technique was simplified using skip connections that link encoder and decoder layers of equal size, as established in U-net autoencoders (UAEs). Note: UAEs were first invented to segment biomedical images.[12] SS is also proving to be useful in the field of driver-assistance and self-driving vehicles, as it can identify the key features of the visuals ahead (e.g., road, road lines, cars, pedestrians, trees).[13]

For this study, the investigation began by applying the SS technique to WSI images, as described in Section 2. In addition, a simpler ML problem of evaluating percentage cloud cover in a WSI Image was pursued. The goal was to determine what percentage of the sky was thick clouds, thin clouds, and clear sky without detailed segmentation. The softmax activation function provided a natural final layer for compositional estimation as its output, which by definition, added up to unity. However, for training the model, instead of using a categorical cross-entropy loss term, a continuous output was sought using a mean-squared error metric.

Finally, a direct estimation of solar radiation (SR) from standard camera images[14–16] that were generally sensitive only to the visual spectrum (0.3–0.7 μm wavelength) was explored. The pyranometers used to accurately determine total SR sample from a wider spectrum, between 0.3 and 3 μm; therefore, from a narrow visual spectrum image, the ML algorithm needed to infer the radiative energy

beyond the visual spectrum. An elaboration of these techniques and their application is provided in the next section.

2. Methods

This work investigated three different WSI analysis tasks and applications: semantic segmentation of cloud types, the evaluation of cloud percentage, and estimation of SR. Underlying all three model applications was the CNN, which was used to learn and identify features of WSIs at multiple length scales, which then yielded activation values that were mathematically combined to form a desired output. In comparison, early ML methods required researchers to handcraft texture classifiers.[1,3] Despite the CNN advancements, consulting these original studies serves to expand the types of textural classifiers that could be adopted into a neural network topology.

2.1 Semantic Segmentation

The first model explored was very ambitious given the lack of hand-labeled data. The WSI image was segmented, block by block, into seven sky categories: cumulus, stratocumulus, cirrus, cirrocumulus, stratocirrus, stratus, and clear sky. Since manually labeling images was labor intensive, we searched the Internet for example pictures of each class. The approximately 20 reference images were carefully cropped or masked, removing irrelevant regions. Each masked reference image was sliced into 32 × 32-pixel images and used as input to train the neural network classifier, as seen in Fig. 1. A dropout layer with a fraction of 0.05 was placed between each convolutional layer to improve generalization accuracy. During inference, we generated sliding square windows of the input to achieve higher-resolution segmentation.

By far, the most challenging aspects of this model were the unique nature of the camera specification, camera settings and ambient lighting conditions between the reference training images, and the actual WSI images being analyzed. Therefore, even with some simple brightness and contrast adjustments, the model (trained on Internet cloud examples) was challenged by the test images.

In addition to the neural network model, a non-ML sun locator was introduced to moderately improve the segmentations, especially blanking out the sun. The locator function began by converting pixel values to the range [0, 1], applying the strong contrast operator, $f(x) = x^{12}$, and then computing the image's center of mass, which was the first candidate center point of a circle enclosing the sun. The center of mass was then recalculated within a 100-pixel radius of the first computed center. The radius of the sun circle was computed as the average of the x- and y-axis

standard deviations of the image at the second center. The resulting circle was filled to create a sun mask and applied to the final ML output image to remove the pixels associated with the sun. This algorithm did not account for lens glare and other artifacts caused by direct sunlight impinging on the lens and camera sensor.

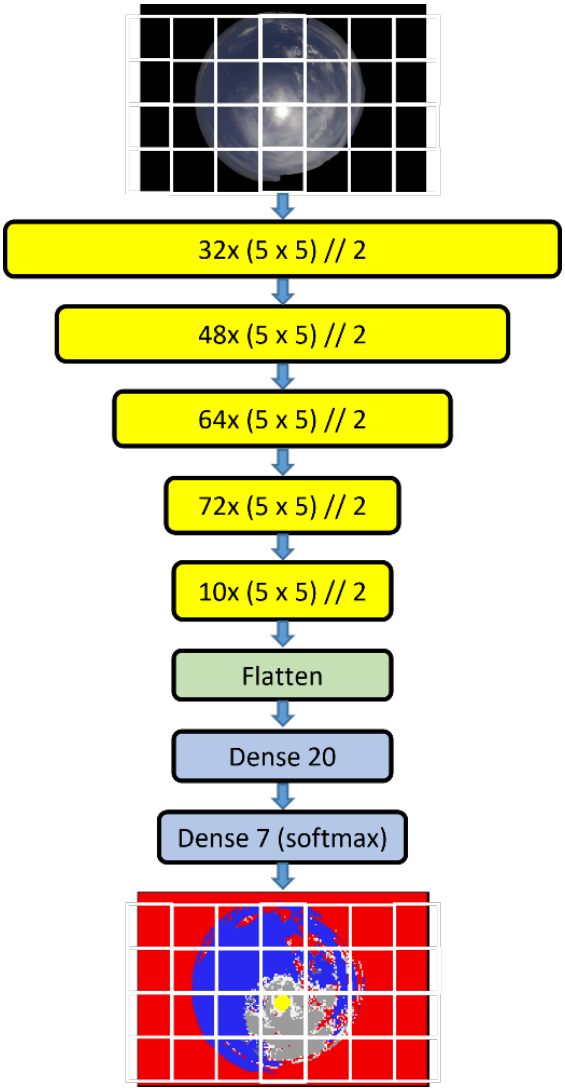


Fig. 1 Neural network topology for the cloud segmentation model used in this work. The "//" symbol means strided downsampling. The preceding *N*x refers to the number of kernels, *N*, in the convolutional layer.

2.2 Percentage Sky Composition

Given that the semantic segmentation model lacked consistent and labeled training data, a model was developed with more modest goals and manually labeled data. First, the problem was reduced to simply estimating the percentage of three classes: thick clouds, thin clouds, and clear sky. Second, a subset of images ($N = 200$) was hand-labeled. These images used the same camera, protocol, and location of the

images to which the model would later be applied. For this study, the data set was from the Atmospheric Intelligence for Hybrid Power Grid (AIHPG) WSI (detailed in Section 2.4). The neural network used is shown in Fig. 2. The model was very compact for a CNN, as it contained only 73,803 trainable weights. The model was trained on 160 images, validated on 20 images, and tested on 20 images. Preprocessing strategies of these particular images were explored and a simple brightening/gamma correction formula was devised:

$$y_{red} = x_{red}^{0.7},$$
$$y_{green} = x_{green}^{0.63},$$
$$y_{blue} = x_{blue}^{0.6125}.$$

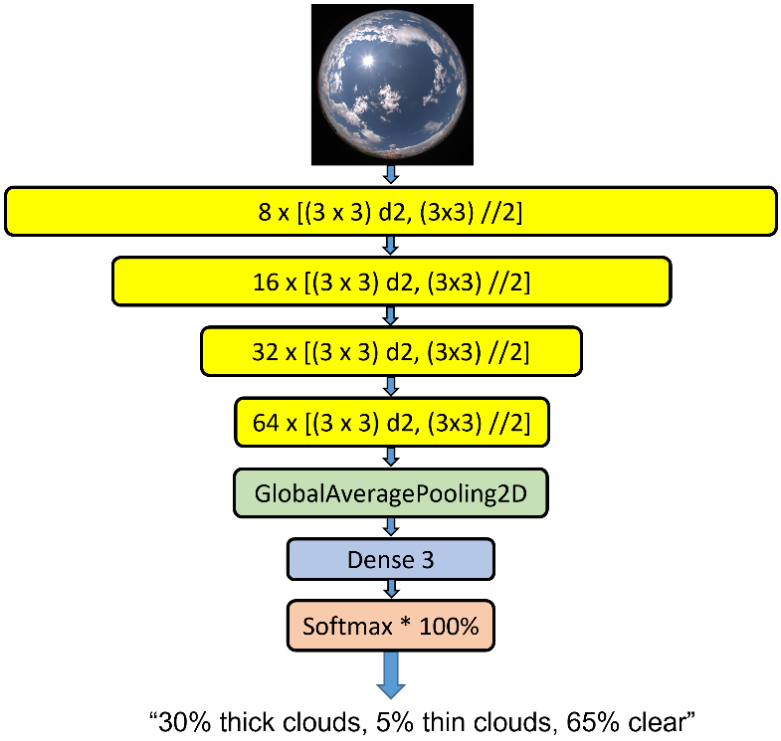


Fig. 2 Neural network topology for percent cloud composition model. "(3 × 3) d2, (3 × 3)//2" means a 3 × 3 kernel layer with a dilation rate = 2 followed by a 3 × 3 kernel layer with stride = 2 downsampling.

2.3 Estimation of Solar Radiation

Finally, using data from the National Renewable Energy Laboratory (NREL),[17] WSI images were paired to simultaneous SR outputs from their CM22 pyranometer to develop a model that estimates the SR from the WSI image. Multiple research groups have already looked at this problem for both WSI[15–16] and satellite images.[14] The largest issue beyond achieving small error bars was generalizability, that is, training a model that could perform well on images collected over a different test

site. We are unaware of any studies where solar output was regressed against WSI images for more than one site. Indeed, to build a truly generalized model would require uniform data collection over many locations throughout the globe, reflecting variations in locations, skies, and climates.

Within the field of neural-network-based computer vision, supervised regression problems (e.g., comparing images against continuous variables) are not as common as classification. Deep regression is used for applications such as counting objects,[18] human pose estimation, depth estimation, and age estimation.[19] Like previous efforts, we assumed that we could take a general CNN classifier architecture and simply remove the softmax layer, as seen in Fig. 3, which is normally used to perform multiclass logistical regression. In addition, the output activations were scaled to match the expected output values, such that the pre-output values ranged from [–0.5, 0.5]. In this way, the learned weights remained centered around a mean of 0 and the variance was close to 1. This scaling was not entirely necessary but helped with model initialization prior to training.

Many advanced topologies have been developed with CNNs since the original formulation and breakthrough model AlexNet;[8] however, for simplicity, a similar topology consisting of a sequence of 5 × 5 kernel layers was used.

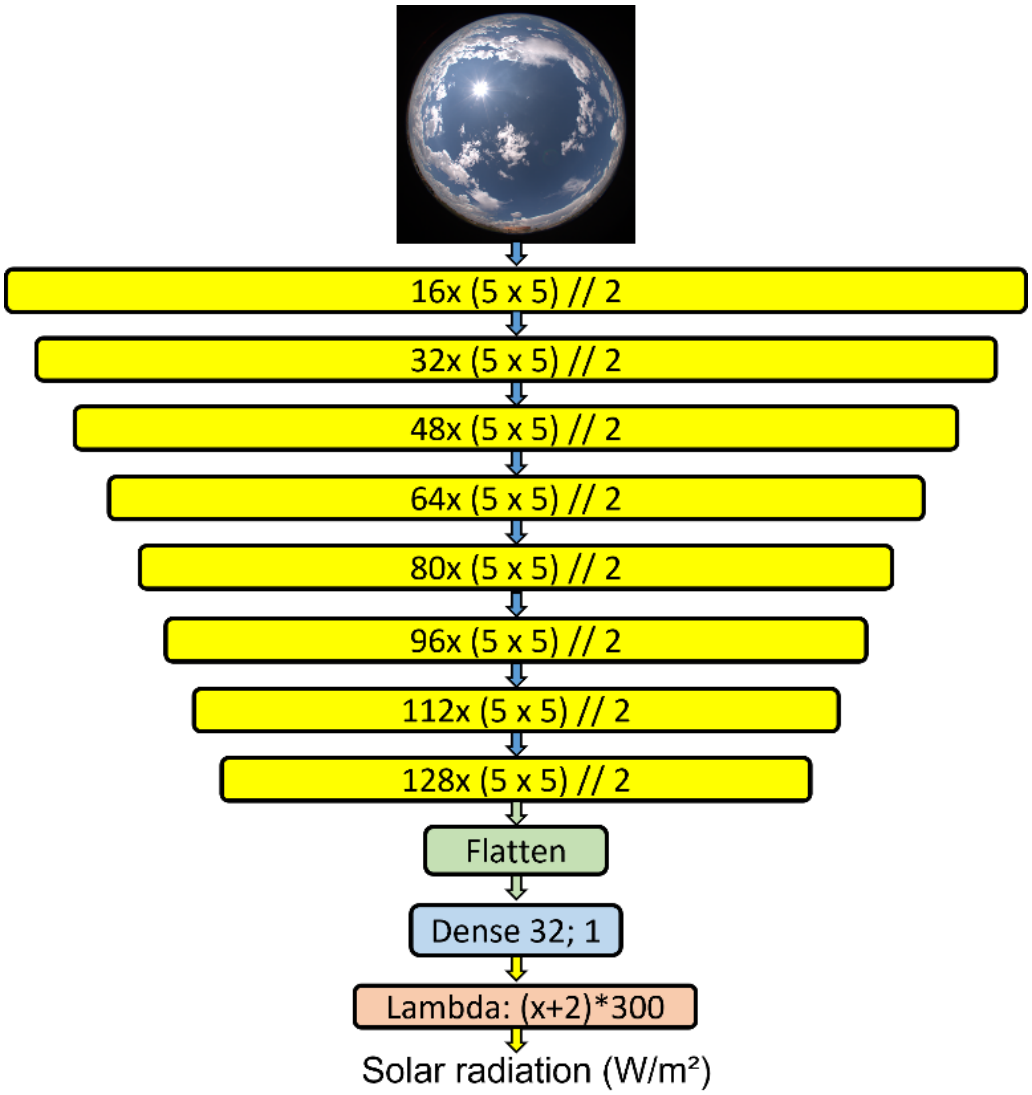


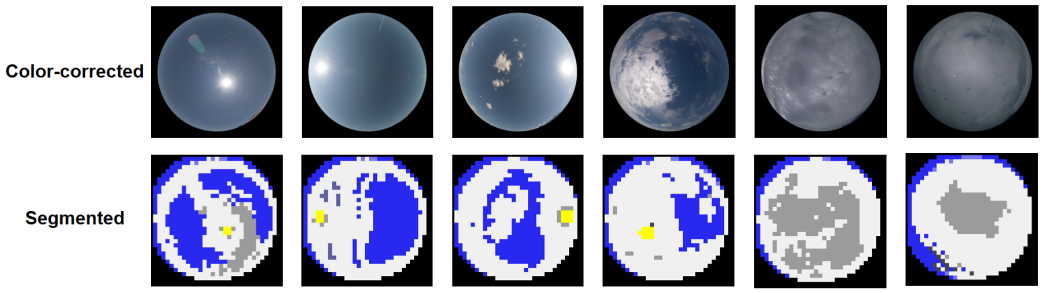
Fig. 3 SR estimation model used in this work

2.4 AIHPG Whole Sky Imager

The AIHPG WSI data were acquired using a Raspberry Pi 4 Model B computer connected to a Raspberry Pi HQ camera (Raspberry Pi Foundation, UK), which contained a Sony IMX477 sensor. The camera had a CS-mount lens with an f1.5 aperture and 1.55-mm focal length. The camera was enclosed in a Tadashi dome with a 0.9 neutral density filter. The exposure time varied between 500 µs to 100 ms such that each recorded image had just above an average pixel value of 75 (out of a maximum value of 255). Snapshots were recorded every 3 min during the day over a period of several months.

3. Results

The first model attempted to segment cloud type, pixel by pixel, in the WSI images. Random exemplary results are presented in Fig. 4. The most prominent issue as seen in the first two sets of images from the left is the over-identification of clouds under a clear sky. In the first image, the reflection of the sun in the lens is interpreted as cloud cover. Furthermore, the white horizons, which reflect lens and natural geometric/atmospheric distortions, are mostly characterized as clouds. The sun locator model performed satisfactorily, as it picked up the correct location and masked out some of the segmentation for the clear and, in examples 3 and 4, partly cloudy cases.



Cumulus	48%	58%	66%	74%	50%	67%
Cirrus	1%	0.5%	1%	1%	-	1%
Stratus	10%	1.5%	0.5%	1%	42%	19.5%
Sky	41%	40%	32.5%	24%	8%	12.5%

Fig. 4 Two clear, two partly cloudy, and two overcast sky cases evaluated by the segmentation model (left to right). Legend for segmented outputs: sky (blue), cumulus (white), stratus (gray), sun (yellow), other (black).

The second model was also supervised. For this model, about 200 images were manually labeled in terms of percentages of thick, thin, and clear sky. The model converged in about 400 epochs, as seen in Fig. 5. While it was less descriptive than the first model, it appeared to be more accurate and useful for application purposes. In Table 1, many of the results are qualitatively correct; however, some issues are seen in the fourth and fifth images. In the fourth image, the model mildly overpredicts clear sky before the sun completely comes into view in the morning. In the fifth image, the model also misses the faded cirrus background and incorrectly predicts a mostly clear sky. Since this image already has the sun over the horizon for the evening, this particular result would not be problematic for SR prediction purposes.

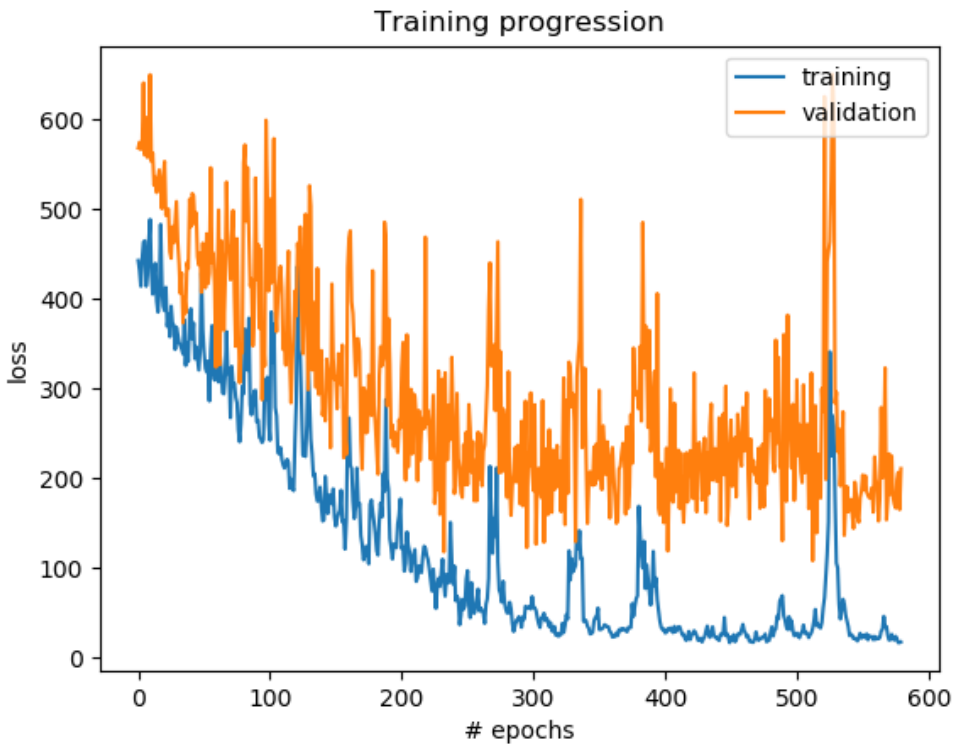
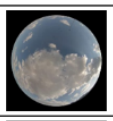
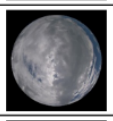
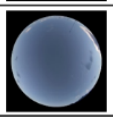
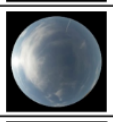
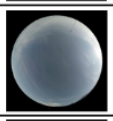
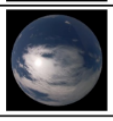


Fig. 5 **Training and validation loss during training of the cloud percent composition model**

Table 1 **Prediction vs. manual labeling of a representative test set of WSI images**

Whole Sky Image	Ground Truth			Prediction		
	Thick %	Thin %	Sky %	Thick %	Thin %	Sky %
	60	0	40	66	3	30
	95	0	5	97	0	2
	3	0	97	3	3	93
	15	75	10	1	57	41
	10	90	0	0	6	92
	30	20	50	56	1	42

The third model computed SR directly from visual images. In Fig. 6, the validation loss converged after about 100 epochs, even as the training error continued to slope downward. Figure 7 compares estimated and actual SR for four distinct days in 2021. The top-right graph demonstrates that the SR model can infer the time of day from the sun position in the sky. The general agreement in all four graphs was satisfactory, with only a few minor issues. For example, in the top-left graph, the lower SR values after midday were not estimated, indicating perhaps that the SR was overpredicted when clouds were covering the sun. In the bottom-left graph, the model may shoot for an average SR value rather than pinpoint precise estimates during an overcast sky. Figure 8 compares estimated and actual values on a scatter plot. Here, larger actual SR values are more likely to have significant overestimation or consistent but smaller underestimation.

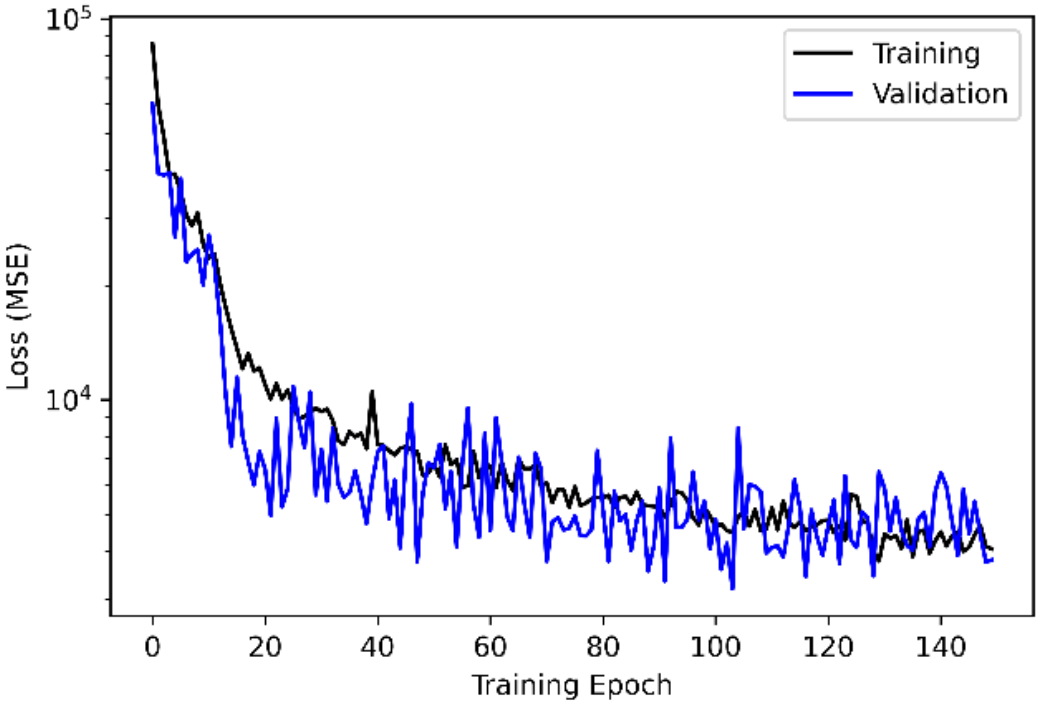


Fig. 6 Training loss progress for SR model

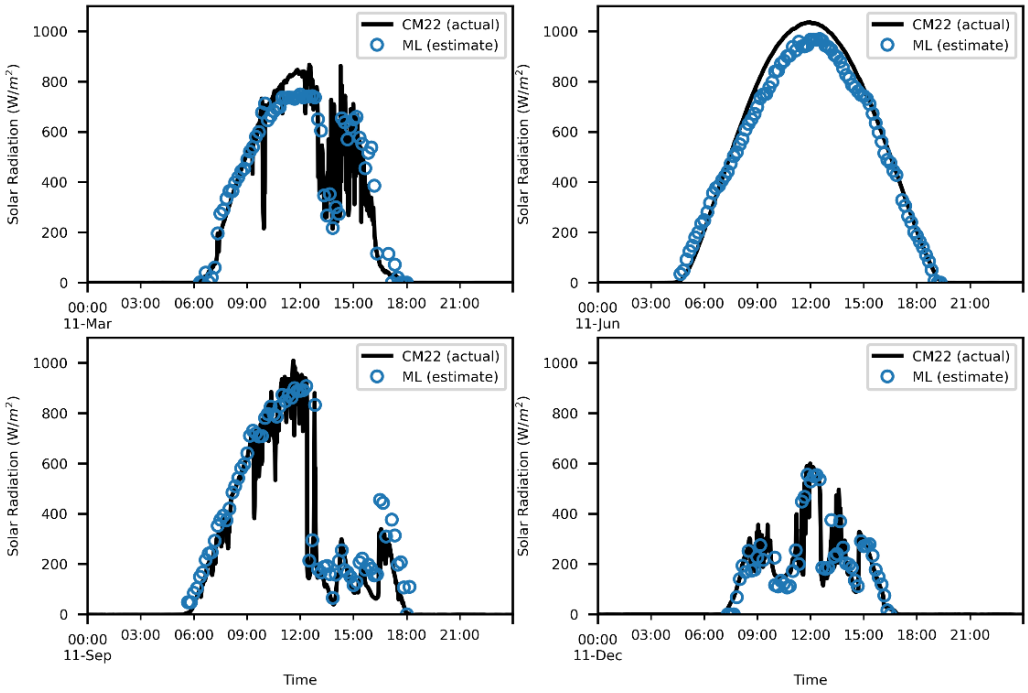


Fig. 7 ML-estimated vs. actual SR over four different days in 2021

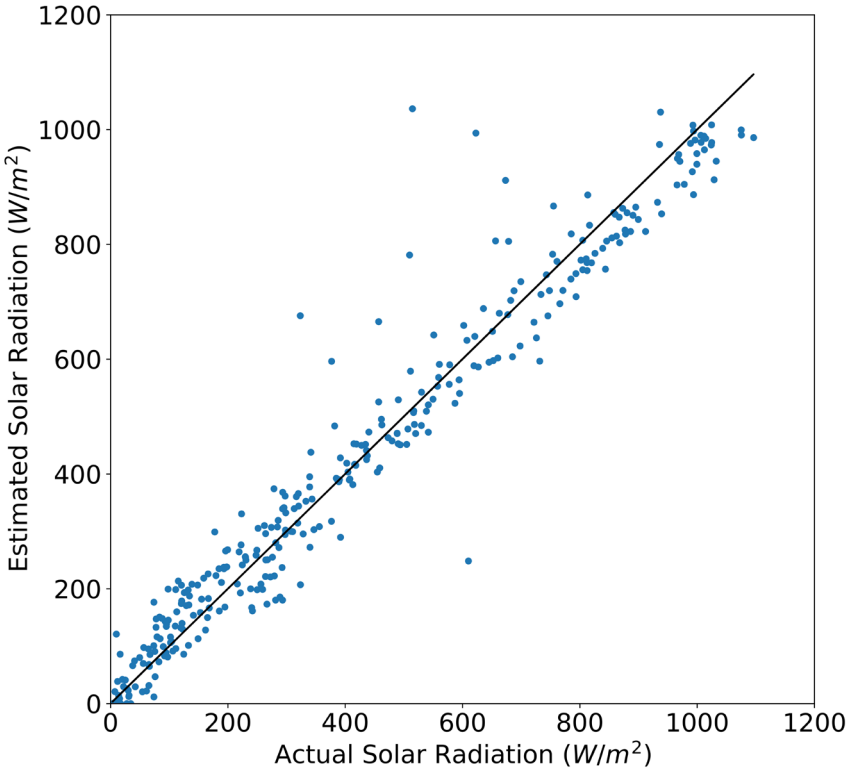


Fig. 8 Estimated vs. actual SR for test set using SR model in this work

4. Discussion

This work explored three distinct supervised learning tasks for analyzing WSI images. The semantic segmentation method fell short most likely due to the lack of manually labeled segmentations that are normally used for these applications. For example, in autonomous-car driving and unmanned aerial systems applications, hand-labeled segmentation masks are an integral component of the research effort. One automated solution is to synthetically generate large amounts of 3-D scenes with realistic graphic engines where the viewing perspective is set to the expected field of view of the attached cameras. In our work, the manual segmentation was replaced with a data set of sizeable real examples of each cloud type. This proved to be inadequate because of discrepancies among the camera properties, perspective, and lighting conditions compared to the actual data tested (WSI images). One possible method to try in the future would be to augment the training data with random variations of brightness, contrast, and color to increase the sensitivity of the model to out-of-distribution data.

The second model diagnosed thick and thin cloud composition proved more reliable as it was trained on hand-labeled WSI images. Furthermore, the training images were obtained from the same camera and exposure settings as those used in testing. In addition, the complexity of the required analysis was significantly reduced, in that only the total fraction of two cloud types and clear sky was expected as output. The results of this model can be used as input into a solar radiation prediction component of a broader energy management system.[20–21]

Finally, a CNN model was developed to estimate SR from WSI images. Results for the extreme conditions of low light and maximum light were accurate; however, medium bright conditions had a small percentage of overpredicted "outliers" and a consistent but small underprediction.

5. Conclusion

This work began with the optimistic goal of classifying multiple cloud types at a pixel-by-pixel level. While this objective could still be possible, it would require a significant amount of hand-labeled photos to ensure accurate and consistent ground truths are utilized. With only a limited amount of ground-truth data taken from different cameras, lenses, exposure speed, and aperture size, the semantic segmentation problem was significantly more challenging. General success will likely only come with an accumulation of hand-labeled data from multiple different systems imaging multiple different climates at all times of year.

A subset problem of estimating the total composition of thick/thin clouds and clear sky for an image required only a small hand-labeled data set, presuming the imaging system and environment remained constant. The ML percent cloud cover model was moderately successful and is now being incorporated into a comprehensive energy optimization system.

A direct estimation of SR from WSI images was demonstrated. To do this required collocated data collected over a single year. Correspondence with ground-truth SR measurements was good but not perfect. However, the model's implicit ability to estimate solar position/time of day was encouraging.

6. References

1. Azimi-Sadjadi MR, Zekavat SA. Cloud classification using support vector machines. IGARSS 2000. IEEE 2000 International Geoscience and Remote Sensing Symposium. Taking the pulse of the planet: the role of remote sensing in managing the environment. Proceedings 2; 2000.

2. Allmen MC, Kegelmeyer Jr, WP. The computation of cloud-base height from paired whole-sky imaging cameras. Journal of Atmospheric and Oceanic Technology. 1996;13(1):97–113.

3. Tian B, Azimi-Sadjadi MR, Vonder Haar TH, Reinke D. Neural network-based cloud classification on satellite imagery using textural features. Proceedings of International Conference on Image Processing. 1997;3:209–212.

4. Wohlfarth K, Schröer C, Klaß M, Hanekes S, Venhaus M, Kauffman S, Wilhelm T, Wohler C. Dense cloud classification on multispectral satellite imagery. 2018 10th IAPR Workshop on Pattern Recognition in Remote Sensing (PRRS). 2018;1–6.

5. Hong Y, Hsu K-L, Sorooshian S, Gao X. Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. Journal of Applied Meteorology. 2004;43(12):1834–1853.

6. Escrig H, Batlles FJ, Alonso J, Baena FM, Bosch JL, Salbidegoitia IB, Burgaleta JI. Cloud detection, classification and motion estimation using geostationary satellite imagery for cloud cover forecast. Energy. 2013;55:853–859.

7. Lee J, Weger RC, Sengupta SK, Welch RM. A neural network approach to cloud classification. IEEE Transactions on Geoscience and Remote Sensing. 1990;28(5):846–855.

8. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems. 2012;25:1097–1105.

9. Heinle A, Macke A, Srivastav A. Automatic cloud classification of whole sky images. Atmospheric Measurement Techniques. 2010;3(3):557–567.

10. Ghonima MS, Urquhart B. Chow CW, Shields JE, Cazorla A, Kleissl J. A method for cloud detection and opacity classification based on ground-based sky imagery. Atmospheric Measurement Techniques. 2012;5(11):2881–2892.

11. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015;3431–3440.

12. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical Image Computing and Computer-Assisted Intervention. 2015;234–241.

13. Treml M, Aronja-Medina JA, Unterthiner T, Durgesh R, Friedmann F, Schuberth P, Mayr A, Heusel M, Hofmarcher M, Widrich M, Nessler V, Hochreiter S. Speeding up semantic segmentation for autonomous driving. Semantic Scholar; 2016 Oct 15. [accessed 2022 May 9]. https://www.semanticscholar.org/paper/Speeding-up-Semantic-Segmentation-for-Autonomous-Treml-Arjona-Medina/d1c091bf9402f1caf13892a3fae39326507401be?msclkid=10fc69fecfae11ec87caffbc8ab9b4a4.

14. Polo J, Zarzalejo LF, Ramírez L. Solar radiation derived from satellite images. In: Modeling solar radiation at the Earth's surface. Springer, Berlin & Heidelberg; 2008. p 449–462.

15. Alonso-Montesinos J, Batlles FJ. The use of a sky camera for solar radiation estimation based on digital image processing. Energy. 2015;90:377–386.

16. Al-lahham A, Theeb O, Elalem K, Alshawi T. Alshebeili S. Sky imager-based forecast of solar irradiance using machine learning. Electronics. 2020;9(10):1700.

17. ASI-16 Sky Imager Gallery. NREL Solar Radiation Research Laboratory (SRRL); 2017 [accessed 2022 Feb 1]. https://midcdmz.nrel.gov/apps/imagergallery.pl?SRRLASI.

18. Onoro-Rubio D, López-Sastre RJ. Towards perspective-free object counting with deep learning. European Conference on Computer Vision. 2016;615–629.

19. Lathuilière S, Mesejo P, Alameda-Pineda X, Horaud R. A comprehensive analysis of deep regression. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2019;42(9):2065–2081.

20. Vaucher G, Welch A. Atmospheric renewable energy research, volume 6 (atmospheric renewable energy field study no. 3 and preparation for tactical power in-situ atmospheric intelligence). Army Research Laboratory (US); 2018 Nov. Report No.: ARL-TR-8582.

21. Vaucher G, Berman M, Parker G, Lee M, D'Arcy S, Jane R, Price T. In-situ atmospheric intelligence for hybrid power grids: Volume 2 (Automated Data Flow Tests). CCDC Army Research Laboratory (US); 2020 Sep. Report No.: ARL-TR-9060.

List of Symbols, Abbreviations, and Acronyms

AIHPG	atmospheric intelligence for hybrid power grid
ARL	Army Research Laboratory
CNN	convolutional neural network
DEVCOM	US Army Combat Capabilities Development Command
DOD	Department of Defense
ML	machine learning
NREL	National Renewable Energy Laboratory
PRRS	Pattern Recognition in Remote Sensing
SR	solar radiation
SS	semantic segmentation
UAE	U-net autoencoder
WSI	whole sky imager

1 (PDF)	DEFENSE TECHNICAL INFORMATION CTR DTIC OCA
1 (PDF)	DEVCOM ARL FCDD RLD DCI TECH LIB
11 (PDF) 5 (HC)	DEVCOM ARL FCDD RLC E B MACCALL T JAMESON FCDD RLC ED M S DARCY M LEE J HYATT M ZIEMANN G VAUCHER (5 HC + PDF) R RANDALL FCDD RLS RP M BERMAN B GEIL R JANE