



Advanced Machine Learning and Deep Learning Approaches for Remote Sensing

Gwanggil Jeon ^{1,2}

¹ Department of Embedded Systems Engineering, Incheon National University, Incheon 22012, Republic of Korea; gjeon@inu.ac.kr

² Energy Excellence & Smart City Lab., Incheon National University, Incheon 22012, Republic of Korea

1. Introduction

Unlike field observation or field sensing, remote sensing is the process of obtaining information about an object or phenomenon without making physical contact. According to recent research results, technologies such as artificial intelligence-based deep learning show the potential to overcome the problems of image and video signal processing faced in remote sensing. These technologies generally require the help of high-speed image processing devices such as GPUs, and high computing performance is essential. Through the development of these devices, remote sensing technology, and aerial sensor technology, the scientific community can now monitor Earth with high-resolution images and secure huge quantities of earth observation data. These capacities stem from the fast, accurate and highly reliable technology based on artificial intelligence. The papers published in this Special Issue describe recent advances in big data processing and artificial intelligence-based technologies for remote sensing technologies. A total of 17 papers were published in this Special Issue.

2. Overview of Contributions

The most significant obstacle to optical remote sensing imaging is clouds. In the contribution by Ma et al., entitled “Cloud Removal from Satellite Images Using a Deep Learning Model with the Cloud-Matting Method”, the authors introduce a technique for the removal of clouds from satellite images by paying attention to image overlap and using a method that considers ground surface reflection and cloud top reflection as a linear mixture of image elements [1]. To this end, a two-step convolutional neural network is used to extract cloud transparency information and then generate ground surface information for thin cloud regions. The authors test the proposed model on simulated and ALCD data sets. The model successfully recovers the surface information of the thin cloud region when thick and thin clouds coexist and does so without significantly damaging the information of the original image.

The use of semantic segmentation technology, being a core component of computer vision in remote sensing images, is currently widely applied. The majority of the remote sensing image semantic segmentation methods are based on CNN, but recently transformer-based technology is also widely applied. In the contribution by Li et al. “RCCT-ASPPNet: Dual-Encoder Remote Image Segmentation Based on Transformer and ASPP”, the authors propose RCCT-ASPPNet, which includes a dual encoder structure of RCCT and ASPP [2]. The RCCT uses transformers to fuse global multiscale semantic information, and residual structures are used to connect inputs and outputs. ASPP, performed based on CNN, can extract contextual information about high-level semantics and spatial and channel information through the application of CBAM.

The SAR-ATR method uses unlabeled measured data and labeled simulated data to improve performance. This is due to the problem that there is not a significant quantity of labeled measurement data, and as such this method is currently widely used. In the



Citation: Jeon, G. Advanced Machine Learning and Deep Learning Approaches for Remote Sensing. *Remote Sens.* **2023**, *15*, 2876. <https://doi.org/10.3390/rs15112876>

Received: 29 May 2023

Accepted: 30 May 2023

Published: 1 June 2023



Copyright: © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

contribution by Zhang et al., entitled “Azimuth-Aware Discriminative Representation Learning for Semi-Supervised Few-Shot SAR Vehicle Recognition”, the authors propose a method for designing two AADR loss functions that suppress the intra-class variation of samples with large azimuth differences [3]. Through cosine similarity, they simultaneously magnify the difference between classes of samples with the same azimuthal angle in the feature embedding space. The unlabeled measurement data of the MSTAR dataset are assigned labels of a more similar category among the Sarsim and Sample datasets.

Big data and parameter tuning are essential in the twin process of training and using convolutional neural networks. This process consumes an extensive temporal and computing resources. To improve this, this paper proposes a new lightweight model called FlexibleNet. The contribution by M. Awad, entitled “FlexibleNet: A New Lightweight Convolutional Neural Network Model for Estimating Carbon Sequestration Qualitatively Using Remote Sensing”, proposed a scaling-based model of “width, depth and resolution” [4]. Unlike conventional methods that arbitrarily scale the “width, depth, and resolution” factors, FlexibleNet scales the network width, depth, and resolution uniformly using a fixed set of scaling factors. Experiments have shown that the FlexibleNet model exhibits higher robustness and lower parameter tuning requirements on smaller datasets compared to conventional models.

In the contribution by Ravishankar et al., published under the name of “Capacity Estimation of Solar Farms Using Deep Learning on High-Resolution Satellite Imagery”, the authors propose a deep learning framework for detecting solar power plants via the application of semantic segmentation convolutional neural networks to satellite images [5]. They also propose a model that predicts the energy generation capacity of the detected solar power plant facility. According to their research results, the proposed deep learning model achieved high performance indicators by showing an accuracy of 96.87% and a Jaccard index value of 95.5%. In addition, the average error of the energy generation capacity prediction model was 4.5%. In this study, 23,000 images of 256×256 size were used.

In recent ocean studies, ocean wave parameters, such as SWH, are being actively predicted. Remote sensing has dramatically increased the available quantity of marine data, and artificial intelligence technologies have demonstrated the ability to process big data and derive meaningful insights from them. In the contribution by Atteia et al., entitled “Deep-Learning-Based Feature Extraction Approach for Significant Wave Height Prediction in SAR Mode Altimeter Data”, the authors propose a deep learning-based hybrid approach for SWH prediction using satellite SAR data [6]. Several hybrid feature sets are created using the proposed approach and SWH is modeled using GPR and NNR. SAR mode altimeter data from Sentinel-3A missions, calibrated with field buoy data, were used to train and evaluate the SWH model.

There has been substantial progress in the segmentation of remote sensing images based on deep learning in recent years. However, existing remote sensing image segmentation techniques have two limitations: (1) object detection performance in various scales is poor in complex scene segmentation; (2) feature reconstruction for accurate segmentation is difficult. In order to improve this problem, the contribution by Ma et al., entitled “Deep-Separation Guided Progressive Reconstruction Network for Semantic Segmentation of Remote Sensing Images”, proposed the use of a deep separation-induced progressive reconstruction network [7]. This study made two major contributions. First, the authors design a decoder composed of progressive reconstruction blocks that capture detailed features at various resolutions by utilizing multi-scale qualities obtained from different receptive fields. Second, they use deep features to detect objects of different scales by proposing a deep separation module that classifies various classes based on semantic features. On the basis of testing on two optical remote sensing image datasets, the proposed network shows the best performance among the comparison targets.

In the contribution by Yang et al., entitled “A Multi-Dimensional Deep-Learning-Based Evaporation Duct Height Prediction Model Derived from MAGIC Data”, an EDH prediction network using MLP is proposed [8]. A multidimensional EDH prediction

model is constructed using spatial and temporal “additional data” derived from meteorological measurements. The experimental results reveal the following. (1) Compared with the NPS model, the root mean square error of the weather-MLP-EDH model is 54%. (2) RMSE can be reduced through the contribution of spatial and temporal parameters. (3) The meteorological parameters can be appended to the multilayer-MLP-EDH model so that measurements can fit well at both large and small scales, and the error is improved by 77.51% compared to the NPS model. The proposed model can greatly improve the prediction accuracy of EDH.

Despite many advances in remote sensing imaging technology, remote sensing imaging struggles to meet application requirements due to its low resolution. In order to obtain high-resolution remote sensing images, the authors apply super-resolution techniques to restore and reconstruct remote sensing images. Super-resolution technology solves the quality degradation problem of remote sensing image acquisition systems and efficiently restores images. In the contribution by Wang et al., entitled “A Review of Image Super-Resolution Approaches Based on Deep Learning and Applications in Remote Sensing”, a study on a super-resolution method in deep learning-based remote sensing images is conducted [9]. To this end, the research background of image super-resolution technology is explained, and details such as training and test data sets, image quality and model performance evaluation methods, and model design principles are explained.

The contribution by Wang et al., published under the title “Real-Time Vehicle Sound Detection System Based on Depthwise Separable Convolution Neural Network and Spectrogram Augmentation”, proposes a lightweight model for intelligent sensor system and vehicle detection [10]. Vehicle detection is a binary problem that classifies vehicles or non-vehicles. Deep neural networks have shown high performance in many signal processing applications. However, the performance of deep neural networks depends on big data. Data about issues such as vehicle tracking are limited, making the application of data augmentation technology essential. The proposed algorithm applies mel spectrogram broadening before extracting MFCC features in order to improve the robustness of the system. As the results of the experiment, the final frame-level accuracy achieved was 94.64%, and 34% of the parameters were reduced after compression.

An image whose image quality is degraded due to atmospheric turbulence is additionally affected by noise. The added noise defeats basic signal processing techniques. Since conventional widely used optimization methods are performed under the assumption that there is no noise, noise removal and deblurring must be independently performed in advance in order to use these techniques. The contribution by Shu et al., entitled “Blind Restoration of Atmospheric Turbulence-Degraded Images Based on Curriculum Learning”, proposes the use of an NSRN (noise suppression-based restoration network) for image degradation due to turbulence [11]. The noise suppression module is designed to learn low-order subspaces from turbulence-degraded images, the asymmetric U-NET module is used for blurry image deconvolution, and the fine deep back-projection (FDBP) module is used to reconstruct sharp images. It is used for multi-level functional fusion. They also propose an improved learning strategy to incrementally train the network with the purpose of achieving a good performance through a local-to-global, easy-to-difficult learning method. According to the experimental results, the method based on NSRN showed excellent performance with PSNR 30.1dB and SSIM 0.9.

Sea surface temperature (SST) joins the widely used physical parameters in oceanography and meteorology. In addition to direct measurement and remote sensing, models for SST data have been developed to obtain SST. Since the ocean is a comprehensive and complex dynamic system, the distribution and variability of SST are affected by a variety of factors. In the contribution by Guo et al., entitled “Prediction of Sea Surface Temperature by Combining Interdimensional and Self-Attention with Neural Networks,” a multivariate long short-term memory (LSTM) model is proposed that uses wind speed and air pressure at sea level as inputs along with SST in order to overcome this problem and improve prediction accuracy [12]. In addition, for model optimization, a position encoding matrix

and multi-dimensional input are studied. In addition, a self-attention strategy is adopted to smooth the data during the training process. According to the experimental results, the proposed model is superior to the LSTM alone model and the model with only SST as input.

In the contribution of Qu et al., submitted under the title “Mode Recognition of Orbital Angular Momentum Based on Attention Pyramid Convolutional Neural Network”, the authors propose an OAM mode detection technique based on AP-CNN in order to solve the problem of lack of accuracy in existing OAM detection systems for vortex optical communication [13]. They introduce segmented image classification to exploit the low-level detailed features of the vortex beam superposition and the similar light intensity distribution of plane wave interferograms. ResNet18 is used as the backbone of AP-CNN, and a technique for the detection of subtle differences in light intensity in images is developed by adopting a dual path structure. According to the experimental results, AP-CNN improved accuracy by up to 7% and reduced false mode identification by 3% in the confusion matrix of superimposed vortex modes compared to ResNet18.

Improving the quality of low-light images is a key factor in the interpretation of the surface state of remote sensing images. In the contribution by Rasheed et al., entitled “An Empirical Study on Retinex Methods for Low-Light Image Enhancement”, the authors aim to produce images with higher contrast, noise suppression, and better quality in their low-light versions [14]. Recently, an image enhancement method based on the Retinex theory has received a lot of attention. Therefore, the authors conduct a study to compare the Retinex-based low-light enhancement method with other state-of-the-art low-light enhancement methods and to determine the generalization ability and computational cost. They use experimental results to compare the robustness of Retinex-based methods with other low-light enhancement techniques using different test data sets. Various evaluation criteria are used to compare the results, and an average ranking system is proposed to rank quality enhancement methods.

Weather factors, such as bad weather, can occur when performing land classification through remote sensing, which is a major cause of poor sensing performance. This limitation can be reduced by several factors, such as low-quality aerial imagery and inefficient fusion of multimodal representations. Therefore, it is essential to build a reliable framework capable of robustly coding remote sensing images. In the contribution by Shi et al. on the multimodal convergence and attention mechanism, entitled “Towards Robust Semantic Segmentation of Land Covers in Foggy Conditions”, the authors use HRNet techniques to extract basic features and then use the spectral and spatial representation learning module to extract spectral–spatial representations [15]. In addition, in order to bridge the gap between heterogeneous devices, the authors propose the use of a multimodal Representation fusion module.

Remote sensing images with high temporal and spatial resolution are important for monitoring land surface changes, vegetation changes, and natural disaster surveillance. However, it is difficult to directly obtain high-resolution remote sensing images, and thus the deployment of space–time convergence technology to obtain remote sensing images is receiving a lot of attention. In the contribution by Li et al., entitled “Enhanced Multi-Stream Remote Sensing Spatiotemporal Fusion Network Based on Transformer and Dilated Convolution”, the authors propose a deep learning model with high accuracy and robustness to better extract spatiotemporal information from remote sensing images [16]. The proposed model is EMSNet, which extends the existing MSNet. Dilated convolution is used to extract temporal information and reduce parameters. The authors further adapt the improved transformer encoder to image fusion techniques and enhance it again to effectively extract spatiotemporal information. Experimental results show that EMSNet improved SSIM by 15.3% in the CIA dataset, ERGAS by 92.1% in the LGC dataset, and RMSE by 92.9% in the AHB dataset when compared to MSNet.

LFMC is an important indicator used to assess wildfire risk and fire spread rate. In the contribution by Xie et al. “Retrieval of Live Fuel Moisture Content Based on Multi-

Source Remote Sensing Data and Ensemble Deep Learning Model”, the authors propose two ensemble models that combine deep learning models in order to further improve the inspection accuracy of LFMC [17]. One is a layered ensemble model based on LSTM, TCN and LSTM-TCN models, and the other is an Adaboost ensemble model based on an LSTM-TCN model. Measured LFMC data, MODIS, Landsat-8, and Sentinel-1 remote sensing data and auxiliary data, such as canopy height and land cover in wildfire-prone areas in the western United States, are selected as study subjects. As a result of the search, remote sensing data of different groups are compared. The experimental results suggest that the LFMC search accuracy is higher than that of single-source remote-sensing data since the use of multi-source data can incorporate the advantages of different types of remote-sensing data. The proposed ensemble model can better extract the non-linear relationship between LFMC and remote sensing data.

3. Conclusions

This Special Issue introduces 17 research findings on advanced machine learning and deep learning approaches for remote sensing. Based on the research results introduced here, it is expected that further development and research in the field of artificial intelligence-based remote sensing will yield results in the future.

Acknowledgments: I thank the authors who published their research results in this Special Issue and the reviewers who reviewed their papers. I also thank the editors for their hard work and perseverance in making this Special Issue a success.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Ma, D.; Wu, R.; Xiao, D.; Sui, B. Cloud Removal from Satellite Images Using a Deep Learning Model with the Cloud-Matting Method. *Remote Sens.* **2023**, *15*, 904. [\[CrossRef\]](#)
2. Li, Y.; Cheng, Z.; Wang, C.; Zhao, J.; Huang, L. RCCT-ASPPNet: Dual-Encoder Remote Image Segmentation Based on Transformer and ASPP. *Remote Sens.* **2023**, *15*, 379. [\[CrossRef\]](#)
3. Zhang, L.; Leng, X.; Feng, S.; Ma, X.; Ji, K.; Kuang, G.; Liu, L. Azimuth-Aware Discriminative Representation Learning for Semi-Supervised Few-Shot SAR Vehicle Recognition. *Remote Sens.* **2023**, *15*, 331. [\[CrossRef\]](#)
4. Awad, M.M. FlexibleNet: A New Lightweight Convolutional Neural Network Model for Estimating Carbon Sequestration Qualitatively Using Remote Sensing. *Remote Sens.* **2023**, *15*, 272. [\[CrossRef\]](#)
5. Ravishankar, R.; AlMahmoud, E.; Habib, A.; de Weck, O.L. Capacity Estimation of Solar Farms Using Deep Learning on High-Resolution Satellite Imagery. *Remote Sens.* **2023**, *15*, 210. [\[CrossRef\]](#)
6. Atteia, G.; Collins, M.J.; Algarni, A.D.; Samee, N.A. Deep-Learning-Based Feature Extraction Approach for Significant Wave Height Prediction in SAR Mode Altimeter Data. *Remote Sens.* **2022**, *14*, 5569. [\[CrossRef\]](#)
7. Ma, J.; Zhou, W.; Qian, X.; Yu, L. Deep-Separation Guided Progressive Reconstruction Network for Semantic Segmentation of Remote Sensing Images. *Remote Sens.* **2022**, *14*, 5510. [\[CrossRef\]](#)
8. Yang, C.; Wang, J.; Shi, Y. A Multi-Dimensional Deep-Learning-Based Evaporation Duct Height Prediction Model Derived from MAGIC Data. *Remote Sens.* **2022**, *14*, 5484. [\[CrossRef\]](#)
9. Wang, X.; Yi, J.; Guo, J.; Song, Y.; Lyu, J.; Xu, J.; Yan, W.; Zhao, J.; Cai, Q.; Min, H. A Review of Image Super-Resolution Approaches Based on Deep Learning and Applications in Remote Sensing. *Remote Sens.* **2022**, *14*, 5423. [\[CrossRef\]](#)
10. Wang, C.; Song, Y.; Liu, H.; Liu, H.; Liu, J.; Li, B.; Yuan, X. Real-Time Vehicle Sound Detection System Based on Depthwise Separable Convolution Neural Network and Spectrogram Augmentation. *Remote Sens.* **2022**, *14*, 4848. [\[CrossRef\]](#)
11. Shu, J.; Xie, C.; Gao, Z. Blind Restoration of Atmospheric Turbulence-Degraded Images Based on Curriculum Learning. *Remote Sens.* **2022**, *14*, 4797. [\[CrossRef\]](#)
12. Guo, X.; He, J.; Wang, B.; Wu, J. Prediction of Sea Surface Temperature by Combining Interdimensional and Self-Attention with Neural Networks. *Remote Sens.* **2022**, *14*, 4737. [\[CrossRef\]](#)
13. Qu, T.; Zhao, Z.; Zhang, Y.; Wu, J.; Wu, Z. Mode Recognition of Orbital Angular Momentum Based on Attention Pyramid Convolutional Neural Network. *Remote Sens.* **2022**, *14*, 4618. [\[CrossRef\]](#)
14. Rasheed, M.T.; Guo, G.; Shi, D.; Khan, H.; Cheng, X. An Empirical Study on Retinex Methods for Low-Light Image Enhancement. *Remote Sens.* **2022**, *14*, 4608. [\[CrossRef\]](#)
15. Shi, W.; Qin, W.; Chen, A. Towards Robust Semantic Segmentation of Land Covers in Foggy Conditions. *Remote Sens.* **2022**, *14*, 4551. [\[CrossRef\]](#)

16. Li, W.; Cao, D.; Xiang, M. Enhanced Multi-Stream Remote Sensing Spatiotemporal Fusion Network Based on Transformer and Dilated Convolution. *Remote Sens.* **2022**, *14*, 4544. [[CrossRef](#)]
17. Xie, J.; Qi, T.; Hu, W.; Huang, H.; Chen, B.; Zhang, J. Retrieval of Live Fuel Moisture Content Based on Multi-Source Remote Sensing Data and Ensemble Deep Learning Model. *Remote Sens.* **2022**, *14*, 4378. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.