Semantic Segmentation Networks for Forest Applications

Subjects: Meteorology & Atmospheric Sciences

Contributor: See Ven Lim , Mohd Asyraf Zulkifley , Azlan Saleh , Adhi Harmoko Saputro , Siti Raihanah Abdani

Deforestation remains one of the key concerning activities around the world due to commodity-driven extraction, agricultural land expansion, and urbanization. The effective and efficient monitoring of national forests using remote sensing technology is important for the early detection and mitigation of deforestation activities. Deep learning techniques have been vastly researched and applied to various remote sensing tasks, whereby fully convolutional neural networks have been commonly studied with various input band combinations for satellite imagery applications, but very little research has focused on deep networks with high-resolution representations, such as HRNet.

forest   remote sensing   deep learning   CNN   U-Net

1. Introduction

Forests provide significant ecological importance, including but not limited to biodiversity maintenance, habitats for various flora and fauna, the mitigation of climate change effects, watershed protection [1], erosion protection, carbon sequestration [2], and precipitation level maintenance. Forests also provide economic benefits, with raw materials such as timber [3], food, and medicine, which can drive commercial activities, contribute to people's livelihoods, and lead the development of the national economy. Malaysia's tropical rainforests constitute one of the twelve mega-diverse ecosystems in the world, housing approximately 152,000 fauna species and 15,000 flora species [4]. However, forests nowadays are subjected to uncontrolled deforestation due to various reasons, such as commodity extraction, agricultural expansion, and urbanization, and deforestation remains one of the most concerning issues around the world. Based on the public data from Global Forest Watch created by a collaboration between the University of Maryland, Google USGS, and NASA, for the years 2000 to 2020, Malaysia experienced a net change of −1.12 Mha in tree cover. From 2016 to 2021, the total tree cover loss in Malaysia was 2.43 Mha, where 91.4% of the total loss (2.22 Mha) was caused by commodity-driven deforestation [5].

The Ministry of Energy and Natural Resources (KeTSA) [6] has been monitoring the forest status for many decades through a national forest monitoring program. The Ministry has deployed remote sensing technology to speed up the survey process while reducing the human labor needed for forest monitoring efforts. Initially, panchromatic aerial photographs were used as a remote sensing imaging source, before being replaced by satellite imagery starting from 1991 due to advancements in satellite technology and its effectiveness in remote sensing applications. Previously, the main approach for remote-sensing-based monitoring systems relied on a combination

of elementary spectral bands—for example, the normalized difference vegetation index (NDVI) to distinguish green vegetation's spectral features [6].

There has been growing interest in applying automated techniques to remote sensing to rapidly identify the forest cover area, especially by using conventional machine learning techniques. Several studies have implemented conventional machine learning algorithms in the field of forestry-related remote sensing tasks, such as decision trees, random forest classification, and support vector machines. However, conventional machine learning algorithms are dependent on the type of information or features set by the algorithm's designer and have a limited capability to extract complex and deep features by themselves. The performance of conventional machine learning algorithms can also be very case-specific, which limits the scalability of such machine learning models to other applications. Therefore, many researchers prefer to employ deep learning methods instead of conventional machine learning algorithms, even though they are usually only effective for large datasets [7].

Deep learning methods, which have been applied in several forest-related applications, are a subtype of machine learning that enables feature learning from a set of large data. It is a type of representation learning method that can learn features directly from raw data for accurate detection and classification results. Usually, its architecture consists of composite multilevel representations, using non-linear functions to transform the representations from one-level to higher-level and more abstract representations, enabling the model to achieve complex feature learning [8]. There are various types of deep learning models available, such as convolutional neural networks and recurrent neural networks. Convolutional neural networks (CNNs) are built specifically to handle images in the form of multiple arrays. In terms of accuracy, flexibility, and rapid processing, CNNs perform better than conventional methods [9]. The classic CNN architecture uses multiple convolution and pooling operations to extract useful features from images before being passed to the fully connected layers. Furthermore, the fully connected layers can be replaced with upsampling operations for image segmentation purposes, creating fully convolutional neural networks. Further research has produced various improvised state-of-the-art architectures based on CNNs, such as the Siamese neural network, DeepLab, and U-Net. According to Elizar et al. [10], with the development of fully convolutional neural networks, the segmentation performance based on deep learning methods has improved dramatically in the past few years, especially when compared to the conventional machine learning approach.

The majority of the studies conducted in remote-sensing-based forestry applications use a deep learning approach, specifically convolutional neural networks, to effectively learn features for the accurate classification of forest images. The most used architecture is U-Net, due to its success in performing various segmentation tasks relating to semantic alienation. In general, the CNN architecture for segmentation tasks follows the encoder–decoder topology, whereby the encoder extracts information from low-resolution representations, while the decoder reconstructs high-resolution representations from the extracted low-resolution feature maps. However, very few studies have used high-resolution and multi-resolution fusion networks for remote sensing tasks, which is expected to yield semantically rich and spatially precise segmentation performance [11]. The minimum number of input bands required for effective forest classification also varies across studies, resulting in varying performance. Certain satellite spectral bands may not always be publicly available, and the creation of vegetation indices also requires additional data pre-processing. However, CNNs may be able to achieve sufficient classification quality without

needing a large number of input bands. They can extract the necessary information for forest mapping not only based on the pixel color but also based on the pixel context [12].

2. Detecting Method for Forest Cover and Deforestation

There are two main types of image classification used in remote sensing applications, which are pixel-based classification and object-based classification. Pixel-based classification assigns a class to every pixel in an image. It can be further divided into unsupervised and supervised methods. On the other hand, object-based classification groups pixels into objects that have representative vector shapes in terms of size and geometry.

A CNN is a popular deep learning model used for image processing and object detection, and it has multiple hidden layers [10][13]. Activation functions make the CNN more nonlinear and improve its expression ability. Pooling layers in CNNs perform downsampling and extract important features while reducing the dimensions of hidden layers. The fully connected layer connects the last few layers of the CNN and acts as a classifier to determine the probability of a pixel belonging to a certain class. The output layer consists of one neuron per class category, and all neurons are connected to the fully connected neurons [14].

Dong et al. [15] developed a fusion model of a CNN and random forest (RF) to classify subtropical areas in Taihuyuan, China, using satellite imagery. The model replaces the fully connected layer of the CNN with the RF classifier, resulting in improved performance. However, the model is computationally expensive due to its size and the number of inputs. Khan et al. [16] used a CNN to detect forest changes in Melbourne, Australia, using satellite imagery. They used bounding boxes to label the changes and found that the deep CNN model had higher accuracy and mean IoU compared to other methods. However, accurately producing bounding boxes for forest regions can be difficult, and the method cannot predict forest cover areas.

Fully convolutional networks (FCN) are used to obtain high-level representations from low-level representations by substituting the fully connected layer of the CNN with locally connected layers [17]. This replaced section forms the decoder part of the FCN, creating an encoder–decoder topology. Some improved variants of FCN-based networks are Siamese neural networks, DeepLab, and U-Net.

Siamese neural networks (SNN) consist of two identical CNNs that share weights during encoding. An SNN can also identify similarities and differences between inputs by computing distance metrics. Therefore, SNNs have been applied in visual and change detection tasks, including video target tracking [18] and landscape change detection [19]. Guo et al. [20] utilized a fully convolutional SNN to identify forest changes in Nanning and Fuzhou, China, using Landsat-8 satellite imagery. They introduced a modified version of Caye Daudt et al.'s model [21], called Siamese, which employs concatenation weight-sharing and subtraction weight-sharing methods. This modification prioritizes change information in different layers while preserving detailed image information. The results showed that the method achieved accurate deforestation and afforestation detection and good IoU scores.

Chen et al. [22] introduced the DeepLab series of advanced deep learning segmentation models in 2016. The latest version, DeepLabv3+, has a similar design to previous models but with some differences. Andrade et al. [23] utilized DeepLabv3+ to detect deforestation and found that this model outperformed other models. DeepLabv3+ also performs well with limited dataset sizes, demonstrating its superior generalization capacity. However, they found that there was a potential bias in the trained models due to the large imbalance between deforested and non-deforested areas. Ferreira et al. [24] applied deep learning to map Brazil nut trees in the Amazonian rainforest using WorldView-3 satellite imagery by adopting the DeepLabv3+ architecture with three different encoder backbones: ResNet-18, ResNet-50, and MobileNetV2. In their study, DeepLabv3+ with all three different backbones achieved almost similar accuracy in mapping Brazil nut trees. The researchers noted that the shadows of Brazil nut trees were important features for proper mapping.

U-Net is a popular deep learning architecture designed for semantic segmentation tasks. It was designed by Ronneberger et al. [25] originally for biomedical image segmentation. Since then, U-Net has shown tremendous success in various image segmentation tasks for various applications, including forestry-related tasks. One modification in the U-Net architecture is the use of a large number of feature channels in the upsampling part, allowing context information propagation to higher-resolution layers. Abdani et al. [26] enhanced the multi-scale capability of U-Net by incorporating an SPP module. Bragagnolo et al. [27] used U-Net to map forest cover changes in the Amazon rainforest and compared its performance with that of other deep learning architectures. The results showed that U-Net and ResNet50-SegNet had high accuracy and F1 scores for forest cover mapping segmentation. U-Net also had the lowest training time among the benchmarked architectures. One advantage of the researchers' approach is the model's ability to tolerate some misclassification without significantly affecting the performance.

A study by Wagner et al. [12] examined Amazon forest cover and deforestation from 2015 to 2022 using U-Net with the pixel-based approach. Images were obtained from the Planet API with the PlanetNICFI R package v1.0.4. Forest cover masks were created with the K-textures algorithm. U-Net achieved high accuracy and F1-scores in forest cover segmentation, validated using airborne LiDAR data. However, the model faced difficulties in detecting deforestation, particularly in identifying burnt areas.

References

1. Worm, B.; Barbier, E.B.; Beaumont, N.; Duffy, J.E.; Folke, C.; Halpern, B.S.; Jackson, J.B.C.; Lotze, H.K.; Micheli, F.; Palumbi, S.R.; et al. Impacts of Biodiversity Loss on Ocean Ecosystem Services. Science 2006, 314, 787–790.

2. Bonan, G.B. Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests. Science 2008, 320, 1444–1449.

3. Schulze, K.; Malek, Ž.; Verburg, P.H. Towards Better Mapping of Forest Management Patterns: A Global Allocation Approach. For. Ecol. Manag. 2019, 432, 776–785.

4. Ministry of Energy Natural Resource. Sixth National Report of Malaysia to the Convention on Biological Diversity; Ministry of Energy Natural Resource: Putrajaya, Malaysia, 2019; Available online: https://www.cbd.int/doc/nr/nr-06/my-nr-06-en.pdf (accessed on 15 January 2023).

5. Potapov, P.; Hansen, M.C.; Pickens, A.; Hernandez-Serna, A.; Tyukavina, A.; Turubanova, S.; Zalles, V.; Li, X.; Khan, A.; Stolle, F.; et al. The Global 2000–2020 Land Cover and Land Use Change Dataset Derived From the Landsat Archive: First Results. Front. Remote Sens. 2022, 3, 856903.

6. Kementerian Sumber Asli Alam Sekitar Iklim dan Perubahan National Forest Monitoring System—REDD PLUS. Available online: https://redd.ketsa.gov.my/mrvframework/national-forest-monitoringsystem/ (accessed on 15 January 2023).

7. Nathiratul Athriyah, A.M.; Muhammad Amir, A.K.; Zaki, H.F.M.; Zulkifli, Z.A.; Hasbullah, A.R. Incremental Learning of Deep Neural Network for Robust Vehicle Classification. J. Kejuruter. 2022, 34, 843–850.

8. LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. Nature 2015, 521, 436–444.

9. Nafea, M.M.; Tan, S.Y.; Jubair, M.A.; Abd, M.T. A Review of Lightweight Object Detection Algorithms for Mobile Augmented Reality. Int. J. Adv. Comput. Sci. Appl. 2022, 13, 536–546.

10. Elizar, E.; Zulkifley, M.A.; Muharar, R.; Hairi, M.; Zaman, M. A Review on Multiscale-Deep-Learning Applications. Sensors 2022, 22, 7384.

11. Wang, J.; Sun, K.; Cheng, T.; Jiang, B.; Deng, C.; Zhao, Y.; Liu, D.; Mu, Y.; Tan, M.; Wang, X.; et al. Deep High-Resolution Representation Learning for Visual Recognition. IEEE Trans. Pattern Anal. Mach. Intell. 2019, 43, 3349–3364.

12. Wagner, F.H.; Dalagnol, R.; Silva-Junior, C.H.; Carter, G.; Ritz, A.L.; Hirye, M.C.; Ometto, J.P.H.B.; Saatchi, S. Mapping Tropical Forest Cover and Deforestation with Planet NICFI Satellite Images and Deep Learning in Mato Grosso State (Brazil) from 2015 to 2021. Remote Sens. 2022, 15, 521.

13. LeCun, Y.; Kavukcuoglu, K.; Farabet, C. Convolutional Networks and Applications in Vision. In Proceedings of the ISCAS 2010–2010 IEEE International Symposium on Circuits and Systems: Nano-Bio Circuit Fabrics and Systems, Paris, France, 30 May–2 June 2010; pp. 253–256.

14. Phung, V.H.; Rhee, E.J. A High-Accuracy Model Average Ensemble of Convolutional Neural Networks for Classification of Cloud Image Patches on Small Datasets. Appl. Sci. 2019, 9, 4500.

15. Dong, L.; Du, H.; Mao, F.; Han, N.; Li, X.; Zhou, G.; Zhu, D.; Zheng, J.; Zhang, M.; Xing, L.; et al. Very High Resolution Remote Sensing Imagery Classification Using a Fusion of Random Forest and Deep Learning Technique—Subtropical Area for Example. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2020, 13, 113–128.

16. Khan, S.H.; He, X.; Porikli, F.; Bennamoun, M. Forest Change Detection in Incomplete Satellite Images with Deep Neural Networks. IEEE Trans. Geosci. Remote Sens. 2017, 55, 5407–5423.

17. Long, J.; Shelhamer, E.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 2014, 39, 640–651.

18. Tao, R.; Gavves, E.; Smeulders, A.W.M. Siamese Instance Search for Tracking. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 1420–1429.

19. Chen, H.; Wu, C.; Du, B.; Zhang, L.; Wang, L. Change Detection in Multisource VHR Images via Deep Siamese Convolutional Multiple-Layers Recurrent Neural Network. IEEE Trans. Geosci. Remote Sens. 2020, 58, 2848–2864.

20. Guo, Y.; Long, T.; Jiao, W.; Zhang, X.; He, G.; Wang, W.; Peng, Y.; Xiao, H. Siamese Detail Difference and Self-Inverse Network for Forest Cover Change Extraction Based on Landsat 8 OLI Satellite Images. Remote Sens. 2022, 14, 627.

21. Caye Daudt, R.; Le Saux, B.; Boulch, A. Fully Convolutional Siamese Networks for Change Detection. In Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; pp. 4063–4067.

22. Chen, L.C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. arXiv 2018, arXiv:1802.02611.

23. Andrade, R.B.; Costa, G.A.O.P.; Mota, G.L.A.; Ortega, M.X.; Feitosa, R.Q.; Soto, P.J.; Heipke, C. Evaluation of Semantic Segmentation Methods for Deforestation Detection in the Amazon. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 2020, XLIII-B3, 1497–1505.

24. Ferreira, M.P.; Lotte, R.G.; D'Elia, F.V.; Stamatopoulos, C.; Kim, D.-H.; Benjamin, A.R. Accurate Mapping of Brazil Nut Trees (Bertholletia excelsa) in Amazonian Forests Using WorldView-3 Satellite Images and Convolutional Neural Networks. Ecol. Inform. 2021, 63, 101302.

25. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; Part III 18. Springer International Publishing: Cham, Switzerland, 2015.

26. Abdani, S.R.; Zulkifley, M.A.; Mamat, M. U-Net with Spatial Pyramid Pooling Module for Segmenting Oil Palm Plantations. In Proceedings of the IEEE International Conference on Artificial Intelligence in Engineering and Technology, IICAIET 2020, Kota Kinabalu, Malaysia, 26–27 September 2020; pp. 1–5.

27. Bragagnolo, L.; da Silva, R.V.; Grzybowski, J.M.V. Amazon Forest Cover Change Mapping Based on Semantic Segmentation by U-Nets. Ecol. Inform. 2021, 62, 101279.