

Article

Deep Convolutional Neural Network for Large-Scale Date Palm Tree Mapping from UAV-Based Images

Mohamed Barakat A. Gibril ¹, Helmi Zulhaidi Mohd Shafri ^{1,*}, Abdallah Shanableh ^{2,3}, Rami Al-Ruzouq ^{2,3}, Aimrun Wayayok ⁴ and Shaiful Jahari Hashim ⁵

¹ Department of Civil Engineering and Geospatial Information Science Research Centre (GISRC), Faculty of Engineering, Universiti Putra Malaysia (UPM), Serdang 43400, Selangor, Malaysia; gs57386@student.upm.edu.my

² Department of Civil and Environmental Engineering, Faculty of Engineering, University of Sharjah, Sharjah 27272, United Arab Emirates; shanableh@sharjah.ac.ae (A.S.); ralruzouq@sharjah.ac.ae (R.A.-R.)

³ GIS and Remote Sensing Center, Research Institute of Sciences and Engineering, University of Sharjah, Sharjah 27272, United Arab Emirates

⁴ Department of Biological and Agricultural Engineering, Faculty of Engineering, Universiti Putra Malaysia (UPM), Serdang 43400, Selangor, Malaysia; aimrun@upm.edu.my

⁵ Department of Computer and Communication Systems Engineering, Faculty of Engineering, Universiti Putra Malaysia (UPM), Serdang 43400, Selangor, Malaysia; sjh@upm.edu.my

* Correspondence: helmi@upm.edu.my; Tel.: +60-3-97696459



Citation: Gibril, M.B.A.; Shafri, H.Z.M.; Shanableh, A.; Al-Ruzouq, R.; Wayayok, A.; Hashim, S.J. Deep Convolutional Neural Network for Large-Scale Date Palm Tree Mapping from UAV-Based Images. *Remote Sens.* **2021**, *13*, 2787. <https://doi.org/10.3390/rs13142787>

Academic Editors: Yubin Lan, Wenjiang Huang, Kyeong-Hwan Lee, Yali Zhang and Muhammad Naveed Tahir

Received: 2 June 2021

Accepted: 12 July 2021

Published: 15 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. 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/>).

Abstract: Large-scale mapping of date palm trees is vital for their consistent monitoring and sustainable management, considering their substantial commercial, environmental, and cultural value. This study presents an automatic approach for the large-scale mapping of date palm trees from very-high-spatial-resolution (VHSR) unmanned aerial vehicle (UAV) datasets, based on a deep learning approach. A U-Shape convolutional neural network (U-Net), based on a deep residual learning framework, was developed for the semantic segmentation of date palm trees. A comprehensive set of labeled data was established to enable the training and evaluation of the proposed segmentation model and increase its generalization capability. The performance of the proposed approach was compared with those of various state-of-the-art fully convolutional networks (FCNs) with different encoder architectures, including U-Net (based on VGG-16 backbone), pyramid scene parsing network, and two variants of DeepLab V3+. Experimental results showed that the proposed model outperformed other FCNs in the validation and testing datasets. The generalizability evaluation of the proposed approach on a comprehensive and complex testing dataset exhibited higher classification accuracy and showed that date palm trees could be automatically mapped from VHSR UAV images with an F-score, mean intersection over union, precision, and recall of 91%, 85%, 0.91, and 0.92, respectively. The proposed approach provides an efficient deep learning architecture for the automatic mapping of date palm trees from VHSR UAV-based images.

Keywords: date palm trees; tree species classification; semantic segmentation; fully convolutional neural networks; unmanned aerial vehicle (UAV)

1. Introduction

1.1. Background

The date palm tree (*Phoenix dactylifera* L.) is one of the oldest perennial fruit trees [1] and has been one of the most cultivated fruit trees since the Neolithic/Early Bronze Age [2]. The palm tree has unique and easily recognized characteristics, including a single trunk, palm leaves, and fronds. The crown of a date palm tree is densely covered with long pinnate leaves, which vary with the age of the tree and environmental conditions and can be as long as 4 m, on average [3]. The average height of date palm trees typically ranges from 15 m to 25 m [4]. Date palm trees can generally be grown in arid and semi-arid environments and are planted extensively on the Arabian Peninsula, in West Asia, and

in North Africa. These trees are resilient and capable of surviving in a very hot and dry climates and tolerating saline and alkaline soils [5]. Date palm trees may live for more than 100 years [6] if they are not attacked by pests (i.e., red palm weevil) or diseases. These trees play a considerable role in harsh arid and semi-arid environments by supporting and stabilizing desert ecosystems [6]. Date palm trees bear fruit at an average age of five years, with an average annual yield of 400–600 kg/tree/year and may continue to produce for up to 60 years [7]. According to the Food and Agriculture Organization [8], the world production of dates has increased from 1,852,592 tons in 1961 to 9,075,446 tons in 2019. The world's total harvested area increased six times from 1961 (240,972 ha) to 2019 (1,381,434 ha). The estimation of the population of palm trees and the harvest are derived on the basis of the total quantity of the produced dates, and accurate quantification of date palm trees is either limited or obsolete [9]. The precise information about the number, distribution, and health of date palm trees is crucial for sustainable management, disease and pest control, and yield estimation. Considering that palm trees are distributed over large agricultural and urban areas, the mapping and consistent monitoring of these trees using field-based surveys are impractical and can be laborious and time-intensive tasks.

Remote sensing technologies have substantially boosted the efficiency and accuracy of vegetation mapping, as they offer valuable and feasible tools for acquiring and observing large areas with comprehensive options for resolution [10–12]. A tremendous amount of satellite-based data is being collected and has been used extensively for the extraction of vegetation cover, forestry, and changes over the Earth's surface at regional and global scales [13–19]. However, satellite and piloted aircrafts are constrained by their ability to deliver adequate spatial and temporal resolutions, which are essential to several applications that require short revisit times; such applications include discriminating vegetation or crop types and monitoring their phenological stages and health [20,21]. The capabilities of unmanned aerial vehicles (UAVs) in acquiring images with flexible revisit scheduling at low altitudes with ultra-spatial and temporal resolutions have enabled the observation of small individual plants and the extraction of information at a fine scale that can support farmers in their decision making, improve agricultural production, and optimize the utilization of resources [22,23]. A plethora of studies have successfully employed UAV platforms to acquire red–green–blue (RGB), multispectral, hyperspectral, and thermal images for studying vegetation [24–31], invasive plants [32–34], plant diseases, pests and stresses [35–40], agriculture [41–46], and individual trees [47,48].

Given the formidable and increasing amount of remotely sensed data, a wide spectrum of machine learning (ML) techniques has been used and developed to extract meaningful information and harness the unprecedented sources of data for versatile earth-related applications. Deep learning (DL), as a subfield of machine learning and artificial intelligence, has received considerable attention in the field of remote sensing in the past few years and has increasingly been used in a wide range of applications. In the same manner as the function of the human brain, DL algorithms learn by establishing the natural relationships between input and output data through multilayered, interconnected deep neural network (DNN) architectures [49,50]. Different from the classical machine learning models, DNN is data-driven, which eliminates the need for the construction of manually hand-crafted features of hierarchical data representations; high-level deep features are automatically learned from an input of imagery datasets. DL outperforms classical ML algorithms by effectively tackling the curse of dimensionality and achieving a better and consistent level of classification accuracies from massive image datasets without a significant drop in accuracy [51]. Convolutional neural networks are one of the most widely used deep supervised learning models in a wide spectrum of remote sensing applications and have achieved extraordinary improvement in recent years in the classification of remotely sensed data [52,53]. The use of diverse CNNs in crops and plant phenology recognition [54–59], weed detection [60–62], agriculture [51,63], vegetation mapping [64–68], tree crown detection and mapping [69–72], and disease detection [73–77] has elicited considerable interest.

1.2. Related Work

The mapping and detection of individual tree crowns, tree/plant/vegetation species, crops, and wetlands from UAV-based images are achieved by diverse CNN architectures, which are used to perform different tasks, including path-based classification [78–87], object detection [88–97], and semantic segmentation [98–107]. Recently, semantic segmentation, a commonly used term in computer vision where each pixel within the input imagery is assigned to a particular class, has been a widely used technique in diverse earth-related applications [108]. Various architectures, such as fully convolutional networks (FCNs), SegNet [109], U-Net [110], and DeepLab V3+ [111], have been used successfully to delineate tree and vegetation species [70,98,100,101,103,105,106,112–124], crops [51,57,58,102,125,126], wetlands [107,127], and weeds [61,99] from various remotely sensed data. For instance, Freudenberg et al. [128] utilized U-Net architecture to detect oil and coconut palms from WorldView 2, 3 satellite images. Their approach, which achieved accuracies ranging from 89% to 92%, was proposed as a way to precisely monitor palm trees at large scales. To obtain oil palm plantation maps from high spatial-resolution satellite images, Dong et al. [129] proposed a U-Net structure with a residual channel attention unit and a conditional random field for post-processing. The study achieved an overall accuracy of 96.88% and a mean intersection-over-union of 90.58%. Morales et al. [105] semantically segmented *Mauritia flexuosa* palm trees from UAV images, which were acquired under different environments and light conditions on the basis of Google's DeepLab V3+ architecture. The presented DeepLab V3+ model outperformed four U-Net architectures and was able to distinguish young palms or palms partially covered by other types of vegetation. Torres et al. [100] evaluated five semantic segmentation architectures, including SegNet, U-Net, FC-DenseNet, and two DeepLab V3+ variants for segmenting single tree species. An intersection over union ranging from 77.1% to 92.5% was reported by the experimental analysis, which demonstrated the effectiveness of the evaluated architecture.

To the best of the authors' knowledge, the vast majority of date palm mapping studies focus on the utilization of the traditional machine learning algorithms, such as traditional maximum likelihood supervised classification [130], spectral indices and thresholding analysis [131,132], hybrid per-pixel classification approach [133], fuzzy logic [134], and DT rule-based object-based image analysis [135]. Only limited studies have been dedicated to using deep learning techniques to detect date palm trees [9,136]. The current study aims to (1) develop a deep semantic segmentation method based on U-Shape convolutional network (U-Net) architecture and a pre-trained deep residual network for large-scale mapping of date palm trees; (2) establish a comprehensive and versatile labeled dataset to support the development of the proposed semantic segmentation model for date palm trees from very-high spatial resolution unmanned aerial vehicle (UAV) images; (3) compare the performance of the proposed approach with those of different state-of-the-art semantic segmentation networks.

2. Study Area and Materials

2.1. Experimental Site

The study area is located in the eastern region of Ajman Emirate, United Arab Emirates (UAE). It is geographically located between latitude 25.36°N and 25.43°N and longitude 55.54°E and 55.63°E (World Geodetic System, 1984), as shown in Figure 1, and covers approximately 85 km². The climate of the UAE ranges from arid to hyper-arid, with a daily high temperature ranging between 24 and 42 °C, with mean temperatures of 18 °C–34 °C and extreme hot daytime temperatures occurring frequently, which reach above 40 °C in the summer season [137,138]. The majority of the UAE experiences sporadic and irregular rainfall in time and geographical distribution, whereas the average annual rainfall can be less than 6 mm in the interior of the southern desert and can reach almost 160 mm in the northern and eastern mountainous regions of the country [139].

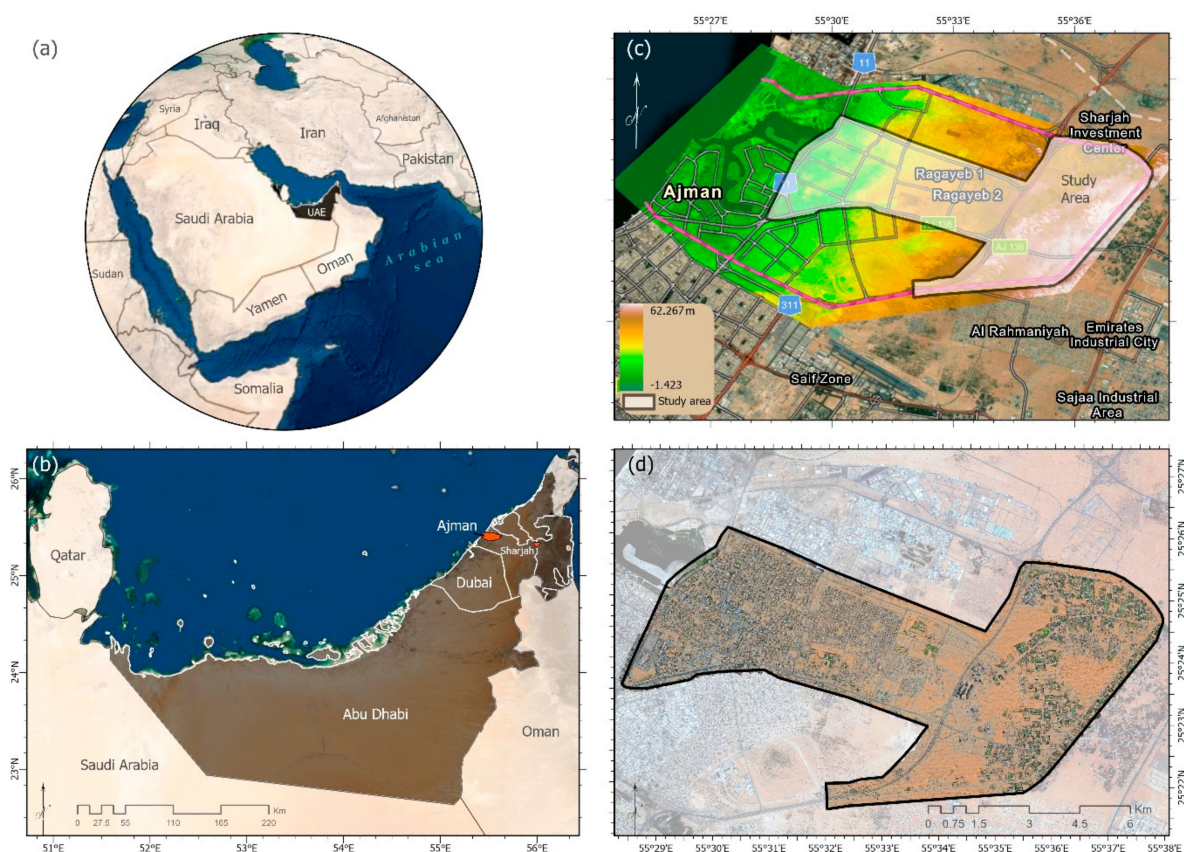


Figure 1. Geographical location of the study area: (a) the Middle East; (b) UAE; (c) location of the study area and the digital elevation model of Ajman Emirate; (d) UAV image of the study area.

2.2. UAV Image Acquisition and Preprocessing

A commercial-grade off-the-shelf fixed-wing UAV (eBee-plus, senseFly[®], Cheseaux-sur-Lausanne, Switzerland) was used to acquire the VHSR images used in this research. The UAV system was equipped with a senseFly S.O.D.A (sensor optimized for drone applications) RGB camera (20 MP digital compact camera with a focal length of 28 mm that acquires VHR visible-color images: red (660 nm), green (520 nm), blue (450 nm)) onboard an inertial measurement unit and Global Navigation Satellite System (GNSS) with real-time kinematic/postprocessed kinematic (RTK/PPK) modes to enable high horizontal accuracy. Flight missions were planned and undertaken using senseFly's eMotion flight controller and data management software. Following the preflight planning and manual launch of the eBee-plus, flight sessions were managed independently by the onboard autopilot. Flight missions were undertaken at an average flying height of 100 m above elevation data (AED), with 80% longitudinal and lateral overlaps. The utilized elevation data, provided in senseFly's eMotion software, were based on a 90 m resolution digital elevation model derived from the Shuttle Radar Topography Mission (SRTM) combined with other data sources (i.e., ASTER GDEM, SRTM30, cartographic data) [140]. Flight lines were oriented perpendicular to the direction of the prevailing wind on the day of the survey. During the flights, a ground-based Trimble R10 GNSS receiver was used in static mode as a base reference station. The preprocessing of the acquired image data was initiated by rectifying the drone locations where the images were captured during the flight, using the PPK mode. Specifically, the ground GNSS RINEX (receiver independent exchange format) data and drone-based GNSS data (drone flight log file) were processed using eMotion software. Then, pix4Dmapper software (v.4.4.10; Prilly, Switzerland) was used to import the geotagged overlapping images and develop the orthomosaic of the study area. The

final product was one orthomosaic RGB image with an average ground sampling distance of 5 cm.

2.3. Labeled Data

Having high-quality and sufficient training data is critical for machine learning algorithms. For the sake of labeling remotely sensed data for semantic segmentation, which requires each pixel in an image to be assigned to a category number (class), a corresponding binary mask was manually prepared for the UAV data. In preparing the binary mask, date palm tree pixels were encircled using ArcGIS Desktop software (v.10.7) to indicate the presence of date palm trees in the study area (Figure 2). The corresponding ground truth data served as a benchmark for the training and evaluation of the implemented models. In this study, the labeling process was comprehensive enough to cover the entire dataset and thereby incorporate as many versatile contexts as possible (i.e., palm trees in farms with vegetation and soil backgrounds and palm trees in urban environments). Given the fine details in the VHSR UAV data, the processing and analysis of large UAV images are demanding and may consume much time and memory. Moreover, resampling these data results in the loss of spatial resolution. As convolutional layers involve extensive computations, the VHSR orthomosaic UAV data and the corresponding mask were divided into equal-sized image tiles (512×512) to cope with the GPU memory limitations. An image-label pair was produced for each image tile in the study area and its corresponding mask (Figure 2). The generated image tiles were divided into three distinct sets: 65% of the data was used for training, 15% was used for validation, and 20% was used for testing purposes. Overall, 11,754 image tiles were selected for training and were artificially enlarged thrice through data augmentation by rotating the image-label pair by 90° , 180° , and 270° using the Sk-Image and Scipy libraries in Python. A total of 2300 image tiles were selected from the generated image tiles for validation purposes. Meanwhile, 3900 image tiles were kept for testing the generalizability of the model. The total number of image tiles used in the current study was enough to develop an efficient DL model for date palm tree mapping, as it is greater than the number of image tiles used in several successful studies [58,103,114,115,122,123,127,128,141].

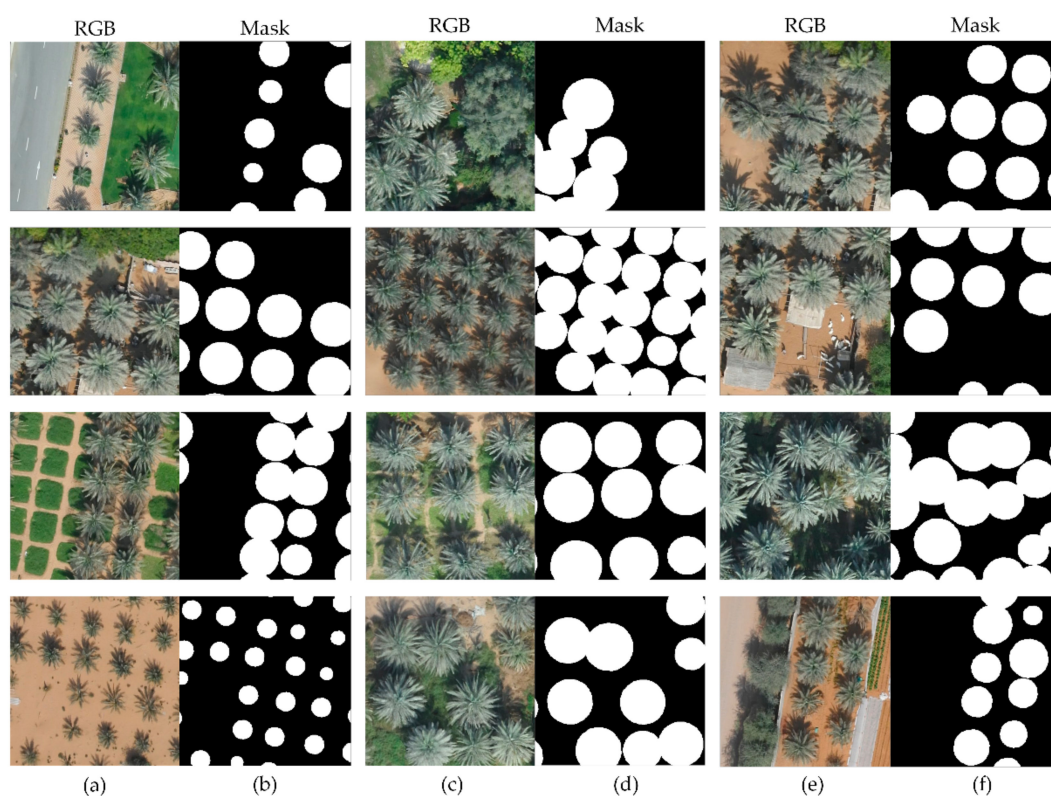


Figure 2. Samples of different image-label pairs: (a,c,e) image tiles and (b,d,f) their corresponding masks.

3. Methodology

Different from patch-based classification and object detection techniques that are based on CNNs, fully convolutional neural networks (FCNs) could be used to delineate the boundary and position of individual date palm trees by performing pixel-level semantic segmentation. This section is dedicated to providing a brief description of the proposed U-Net architecture, the compared FCN networks (e.g., DeepLab V3+ and PSPNet), the utilized segmentation evaluation metrics, and the experimental setup.

3.1. U-Net

U-Net, a U-shaped architecture originally proposed for biomedical image semantic segmentation, is one of the commonly used FCN architectures in studies to classify remotely sensed data for multiple applications. It is a symmetric CNN architecture that comprises the encoder (capturing the context in the input image), the bottleneck, and the decoder (mapping and restoring the contextual information back to the original resolution). The encoder part in a U-Net architecture, a contracting path comprising a set of convolutional and max-pooling layers, receives the input image patches and produces an increased number of down-sampled feature maps on the basis of the depth of the network. The decoder part, an expanding path comprising a set of convolutional, concatenation, and upsampling layers, seeks to retrieve the precise locations and fine characteristics of the features that have been learned by the encoder to semantically segment images. Such retrieval is usually achieved by continuously upsampling feature maps and concatenating them with learned high-resolution features obtained from the corresponding blocks from the encoder.

In this study, a deep residual learning network (ResNet) [142] was considered the encoder backbone of the U-Net network for extracting features from input datasets. ResNet, an network architecture that was motivated by the design of Visual Geometry Group network (VGG) [143], was designed to solve the problem of deep gradient explosion and gradient vanishing when the number of layers in the network increases. ResNet architecture encompasses several sets of blocks (i.e., a sequence of convolution, batch normalization, and ReLu) that implement a specific type of connection method, which is referred to as shortcut connection or skip connection. The output feature maps of a particular layer (x) are forwarded and added to a deeper layer ($y = F(x) + x$). The depth of the ResNet may vary according to the basis of the number of the designed residual layers. ResNet-18, ResNet-50, and ResNet-101 are some common examples of ResNet variations. In this paper, a pertained ResNet-50 based on ImageNet was used as a backbone to increase classification performance and the generalizability of the proposed model. Additional implementation details of the proposed approach are shown in Figure 3 and Table 1.

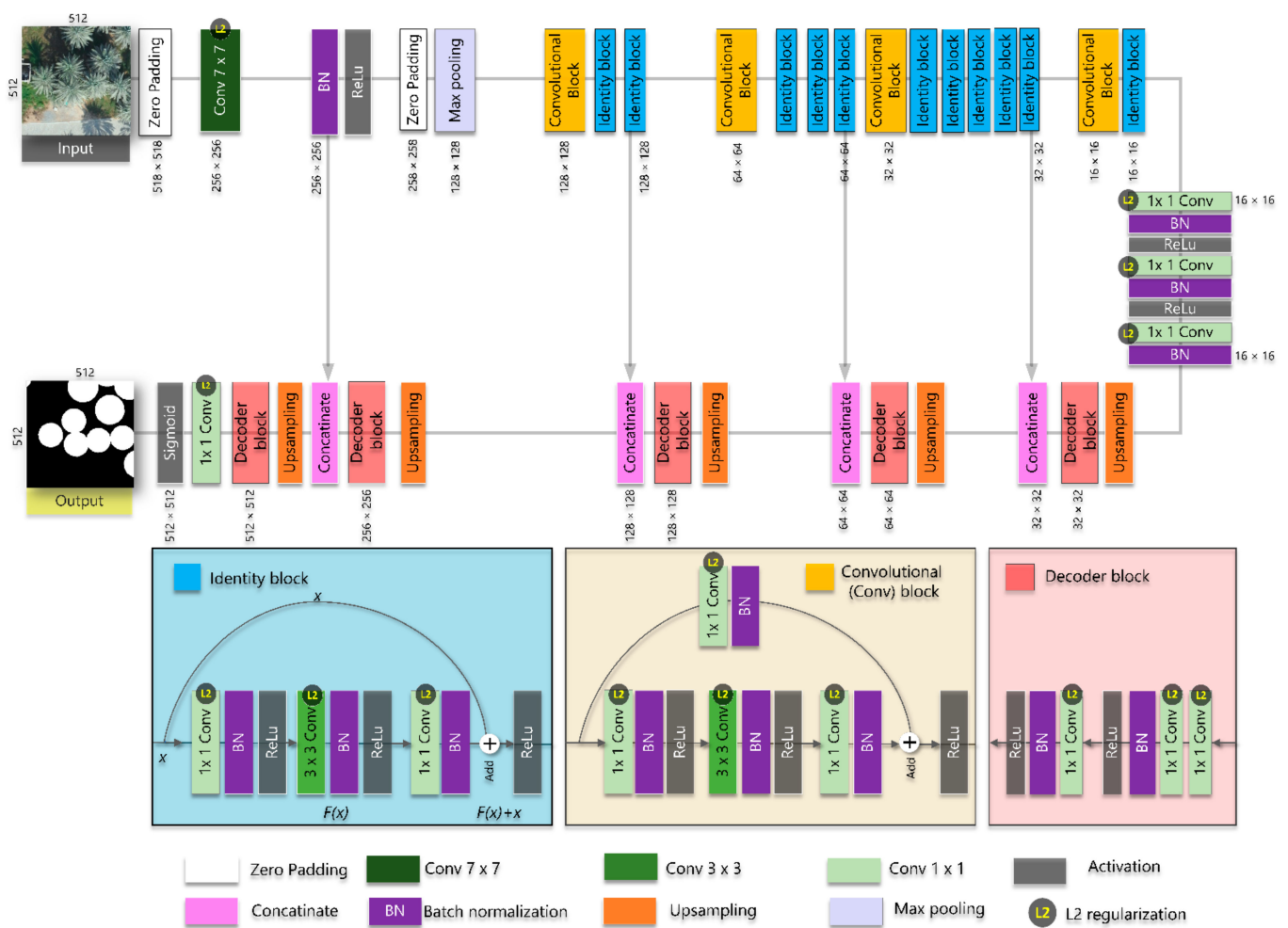


Figure 3. U-Net architecture based on ResNet-50 adopted in this study.

Table 1. The architecture of the proposed U-Net network.

Path	Unit	Kernel Size (k), Feature Map (fm)	Output Size (Width × Height × Channels)
Input			512 × 512 × 3
Encoder	ZeroPadding2D		518 × 518 × 3
	Conv2D	k = (7 × 7), fm = 64	
	Batch normalization + Relu	k = (3 × 3), fm = 64	256 × 256 × 64
	ZeroPadding2D	Fm = 64	258 × 258 × 64
	MaxPooling2D	k = (3 × 3), fm = 64	128 × 128 × 64
	Convltional block 2 × Identity block	$\left(\begin{array}{l} k = (1 \times 1), fm = 64 \\ k = (3 \times 3), fm = 64 \end{array} \right) \times 3$	128 × 128 × 256
	Convltional block 3 × Identity block	$\left(\begin{array}{l} k = (1 \times 1), fm = 256 \\ k = (3 \times 3), fm = 128 \end{array} \right) \times 4$	64 × 64 × 512
	Convltional block 5 × Identity block	$\left(\begin{array}{l} k = (1 \times 1), fm = 256 \\ k = (3 \times 3), fm = 256 \end{array} \right) \times 6$	32 × 32 × 1024
	Convltional block Identity block	$\left(\begin{array}{l} k = (1 \times 1), fm = 512 \\ k = (3 \times 3), fm = 512 \end{array} \right) \times 2$	16 × 16 × 2048
		$k = (1 \times 1), fm = 2048$	

Table 1. Cont.

Path	Unit	Kernel Size (k), Feature Map (fm)	Output Size (Width × Height × Channels)
Bottleneck	Conv2D	$k = (1 \times 1), fm = 512$	$16 \times 16 \times 512$
	Batch normalization + Relu	$k = (1 \times 1), fm = 512$	$16 \times 16 \times 512$
	Conv2D	$k = (3 \times 3), fm = 512$	$16 \times 16 \times 512$
	Batch normalization + Relu	$k = (3 \times 3), fm = 512$	$16 \times 16 \times 512$
	Conv2D	$k = (1 \times 1), fm = 2048$	$16 \times 16 \times 2048$
	Batch normalization	$k = (1 \times 1), fm = 2048$	$16 \times 16 \times 2048$
Decoder	Upsampling2D	$fm = 2048$	$32 \times 32 \times 2048$
	Decoder block	$\begin{pmatrix} k = (1 \times 1), fm = 2048 \\ k = (1 \times 1), fm = 2048 \\ k = (1 \times 1), fm = 2048 \end{pmatrix}$	$32 \times 32 \times 2048$
	Concatenate_1	$fm = 3072$	32, 32, 3072
	Upsampling2D	$fm = 3072$	64, 64, 3072
	Decoder block	$\begin{pmatrix} k = (1 \times 1), fm = 1024 \\ k = (1 \times 1), fm = 1024 \\ k = (1 \times 1), fm = 1024 \end{pmatrix}$	64, 64, 1024
	Concatenate005F2	$fm = 1536$	64, 64, 1536
	Upsampling2D	$fm = 1536$	128, 128, 1536
	Decoder block	$\begin{pmatrix} k = (1 \times 1), f = 512 \\ k = (1 \times 1), f = 512 \\ k = (1 \times 1), f = 512 \end{pmatrix}$	128, 128, 512
	Concatenate_3	$f = 768$	128, 128, 768
	Upsampling2D	$f = 768$	256, 256, 768
	Decoder block	$\begin{pmatrix} k = (1 \times 1), fm = 256 \\ k = (1 \times 1), fm = 256 \\ k = (1 \times 1), fm = 256 \end{pmatrix}$	256, 256, 256
	Concatenate_4	$fm = 320$	256, 256, 320
	Upsampling2D	$fm = 320$	512, 512, 320
	Decoder block	$\begin{pmatrix} k = (1 \times 1), fm = 64 \\ k = (1 \times 1), fm = 64 \\ k = (1 \times 1), fm = 64 \end{pmatrix}$	512, 512, 64
Output	Conv2D + sigmoid	$k = (1 \times 1), fm = 1$	512, 512, 1

3.2. DeepLabV3+

The family of DeepLab architectures proposed by the Google research team adopts multiscale atrous (i.e., holes) convolutions to solve the problems of multiscale objects in image segmentation. Different from the traditional convolution operation, the atrous convolution maintains the same resolution of features without increasing the number of parameters [144]. Four DeepLab architectures have been proposed over the past few years. First, the first version of DeepLab architectures, DeepLab V1, incorporates deep convolutional neural networks and probabilistic graph models (i.e., conditional random field). Second, the DeepLab V2 introduces Atrous Spatial Pyramid Pooling (ASPP) mechanism for extracting multiscale contextual information by using multiple parallel dilated convolutions with different dilation rates [145,146]. While the DeepLab V3 utilizes an improved ASPP module [111], the latest DeepLab architecture, that is DeepLab V3+ [147], improves previous DeepLab versions by introducing a decoder to refine segmentation results and produce more distinctive boundaries. Overall, DeepLab V3+ architecture encompasses an encoder, ASPP module, and a decoder (Figure 4). The adopted encoder network serves as a feature extractor, which reduces feature maps and captures the rich semantic information. The design of the encoder varies depending on the adopted backbone network. Multilevel features of the input image are captured through the ASPP mechanism to solve the multiscale problem of image segmentation of objects. Eventually, the decoder gradually retrieves the spatial information to produce more refined and sharp segmentation results [147]. In

this study, the performance of the proposed approach was compared with those of two variants of DeepLab V3+, based on ResNet-50 [142] and Xception [148] backbone networks.

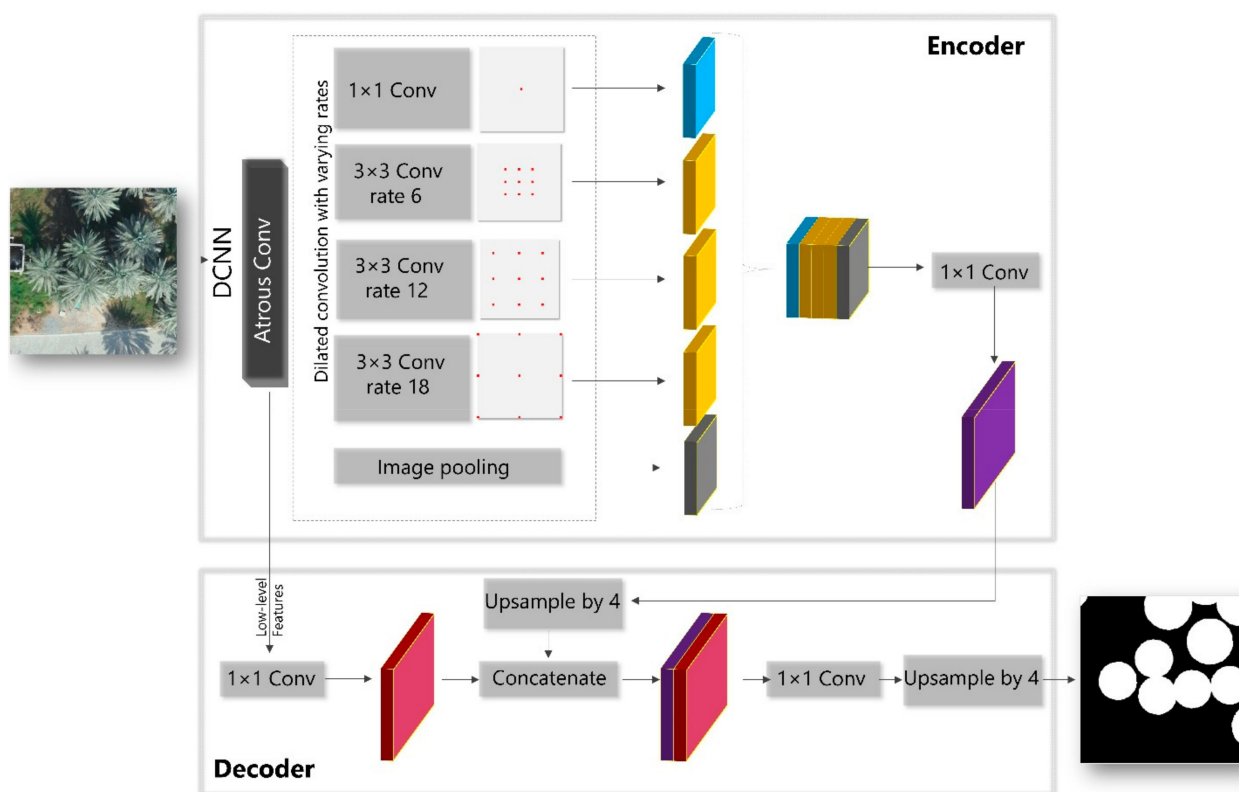


Figure 4. DeepLab V3+ architecture.

3.3. Pyramid Scene Parsing Network

Similar to DeepLab V3+, the pyramid scene parsing network (PSPNet) [149] utilizes a spatial pyramid pooling module between the encoder and decoder structure to capture global contextual information [150] and integrate multiscale features by controlling the size of the receptive field [151]. As shown in Figure 5, feature maps are extracted by the encoder (adopted CNN architecture), and then a series of parallel poolings is used with different grid scales for aggregating contextual information from various regions in extracted features and obtaining a broad spectrum of information. Convolved low-dimension feature maps are then upsampled through bilinear interpolation, concatenated, and ultimately fed to a convolution layer with a proper activation function to extract a probability map(s). The backbone network of the adopted PSPNet was based on ResNet-50 [142] in this study.

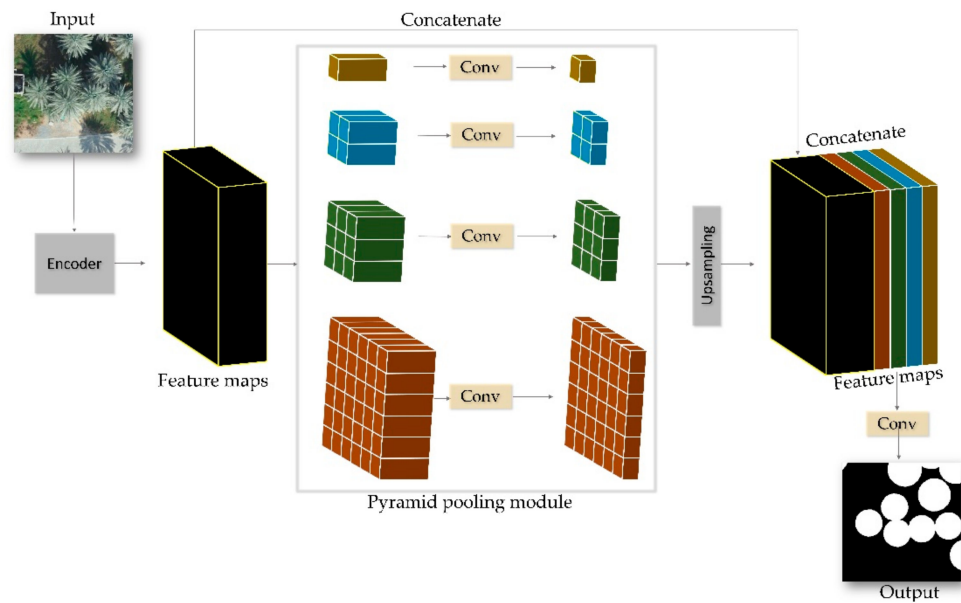


Figure 5. PSPNet architecture.

3.4. Evaluation Metrics

To quantitatively evaluate and analyze the performance of various semantic segmentation architectures for detecting date palm trees, various pixel-by-pixel accuracy measures were utilized. Dice similarity coefficient (DSC) (also known as the F-score) and Mean Intersection-Over-Union (Mean-IOU) (also known as the Jaccard Index) metrics were utilized to evaluate the performance of the different trained models on independent testing datasets. These measures are generally used to compute the amount of agreement between the semantically segmented pixels (CNN output) and the hand-annotated masks. These measures can mathematically be expressed in accordance with Equations (1)–(4). Their computed values range from 0 to 1, wherein the value of 1 indicates the utmost similarity between the predicted and labeled mask (high segmentation accuracy), and 0 shows no similarity between them.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2)$$

$$\text{DSC}(m, p) = 2 \times \frac{|m \cap p|}{|m| + |p|} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

$$\text{Mean-IOU}(m, p) = \frac{|m \cap p|}{|m| + |p| - |m \cap p|} \quad (4)$$

where m denotes the binary ground truth mask, and p represents the predicted semantic segmentation. TP , TN , FP , and FN symbolize the numbers of true positive, true negative, false positive, and false negative, respectively.

3.5. Loss Function

The Dice loss [152], which is a regionally based loss that optimizes the network by using the dice coefficient, was used in this study on the basis of the empirical evaluation. This loss can mitigate the class imbalance between the foreground class (i.e., palm trees) and the background class in binary segmentation tasks [153,154]. Equation (4) expresses the formulation of Dice loss (L_{Dice}).

$$L_{Dice} = 1 - \frac{2 \sum_i^N p_i m_i + \zeta}{\sum_i^N p_i^2 + \sum_i^N m_i^2 + \zeta}, \quad (5)$$

where $p_i \in \{0, 1\}$ is the predicted probability (sigmoid output) of the i th pixel in the image, m_i is the labeled mask of the i th pixel in the image ($0 \leq m_i \leq 1$), and ζ is a minimal constant value to avoid the division by zero problems in the denominator.

3.6. Experimental Setup

Segmentation models were built using the TensorFlow deep learning framework and executed using multiple graphics processing unit (GPU) hardware. The data parallelism approach using the TensorFlow MirroredStrategy, which enables synchronous distributed training on multiple GPUs in one server, was applied to fit the semantic segmentation models on a Linux cluster, with the following specifications: Intel[®], Xeon[®], 2.3 GHz, 512 GB of RAM, eight NVIDIA Tesla K80 (GK210GL) GPUs, and 100 TB of storage. Figure 6 depicts the distributed training strategy used in this study. Eight replicas were created, and each variable in the segmentation model was mirrored across all replicas. The global batch size was set to 32, and the global batch was divided into small minibatches over the eight GPUs. Each GPU independently performed forward and backward parallel passes, and the gradients for the different batches of datasets were computed separately. An independent validation set (2300 image-mask pairs) was used after each training epoch to compute the loss and accuracy of the trained model, thereby avoiding overfitting, and then evaluate its generalizability. Dice loss was computed between outputs of the replicated models and corresponding masks of input image sets as an objective function. The dice coefficient was used to evaluate the segmentation outputs. The gradients of the objective function were gathered and averaged, and an identical update was applied to each independent network.

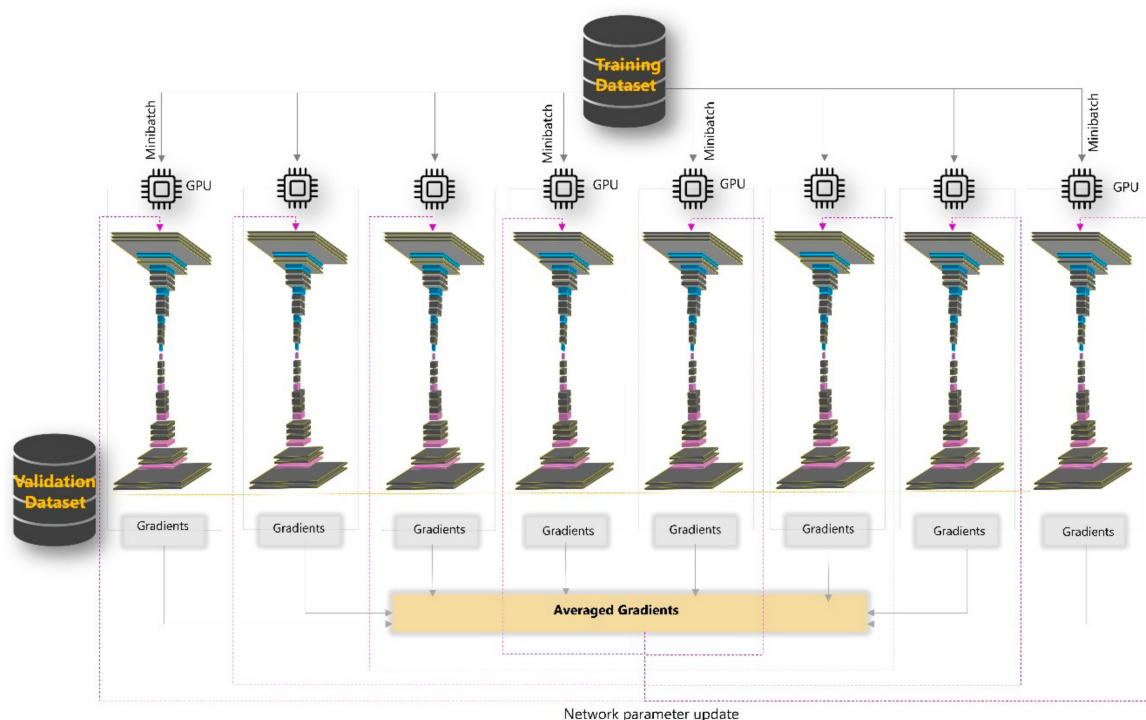


Figure 6. Mirrored distributed training strategy.

The hyperparameters of all segmentation models were empirically tuned by conducting a set of experiments. The encoder part of all models was initialized using ImageNet pretrained weights, and these weights were fine-tuned with further training. Among the stochastic gradient-based optimization algorithms, the adaptive momentum estimation (ADAM) optimizer [155] was chosen in this study for all FCN networks because of its efficiency in improving convergence and dealing with vanishing learning rates [156]. All segmentation models were trained for 100 epochs by using the ADAM optimizer with an initial learning rate of 0.001 and momentum hyperparameters β_1 and β_2 of 0.9 and

0.999, respectively. The training process continued until the model converged or until a maximum number of epochs was reached. Various techniques, including early stopping, were used to avoid overfitting. Early stopping was applied to control the training when the performance in the validation dataset degraded over a particular number of consecutive epochs. L2 regularization was also added to all convolutional layers to avoid overfitting.

4. Results

4.1. Evaluation of Segmentation Performance

The performance of the proposed segmentation model (U-Net based on a pretrained ResNet-50) for mapping date palm trees was compared with different state-of-the-art fully convolutional networks, including U-Net (VGG-16 backbone network), DeepLab V3+ (ResNet-50 backbone network), DeepLab V3+ (Xception backbone network), and PSPNet (ResNet-50 backbone network). Figure 7 displays the evolution of the loss and dice similarity coefficient curves of the proposed approach over the epochs. Several accuracy measures, including precision, recall, F-score, and Mean-IOU metrics, were applied to the validation dataset for all models to evaluate the performance of the proposed architecture against other segmentation architectures quantitatively, as shown in Figure 8. Among them, the proposed model outperformed other segmentation architectures, followed by PSPNet, DeepLab V3+ (Xception backbone network), U-Net (VGG-16 backbone network), and DeepLab V3+ (ResNet-50 backbone network). The proposed approach achieved an F-score, Mean-IOU, precision, and recall of 92%, 85%, 0.92, and 0.91, respectively. A precision metric of 0.92 indicates positive detections relative to the labeled data (consistency of 92% between 2300 labeled and predicted images).

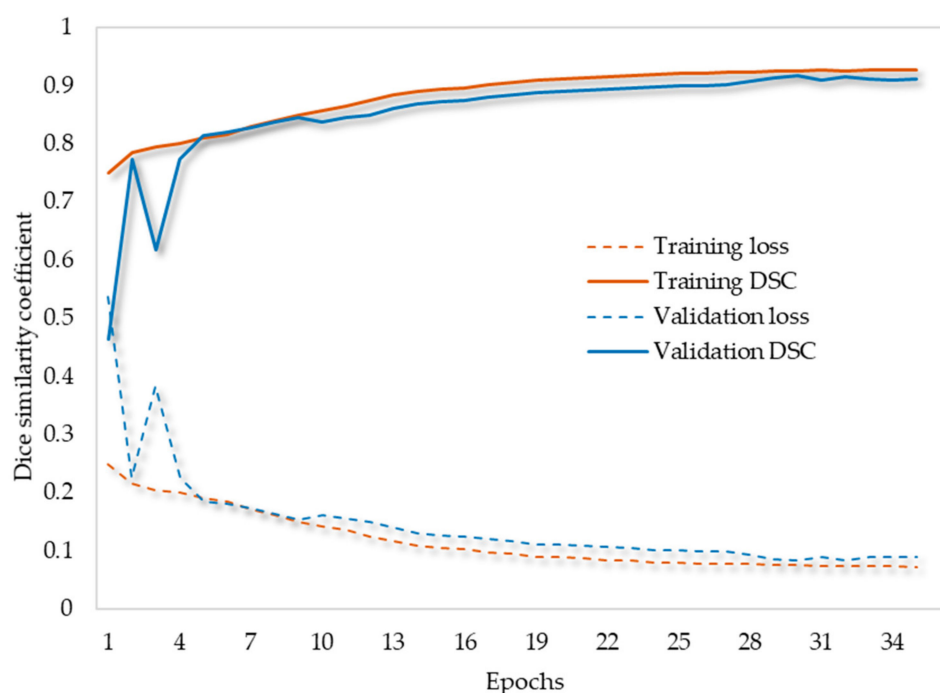


Figure 7. Metric evolution over epochs of the proposed approach.

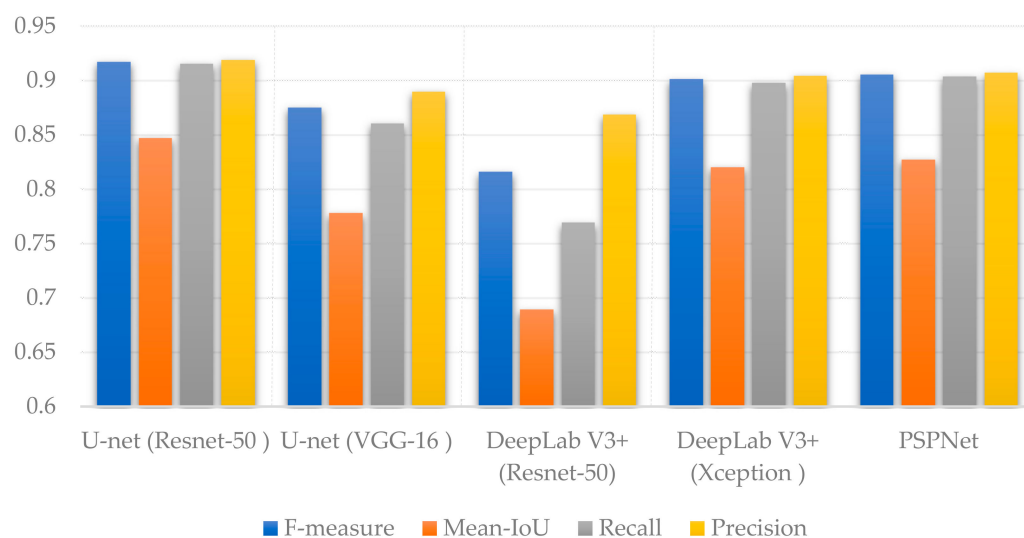


Figure 8. Evaluation metrics for all models obtained from the validation dataset.

The output of semantic segmentation models is a probability map ranging from 0 to 1, thereby indicating the probability of the presence or absence of date palm trees in an image. Here, a threshold value greater than 0.5 was applied to the probability map to derive the segmentation results. Figure 9 shows six randomly selected images from the validation dataset and their corresponding masks and experimental results of the five segmentation models. The original image and ground truth are shown on the left side of the image, and the result of the proposed approach is illustrated on the right side of the image. All segmentation models provided satisfactory segmentation results, with an F-score ranging from 81% to 92% and Mean-IOU ranging from 78% to 85%. However, a disparity can be observed between the results of the five segmentation models in terms of the size and boundary of the detected palm trees. Quantitative and visual analyses showed that the proposed approach presents significant potential in mapping date palm trees from UAV images, because it provides a satisfactory delineation of date palm trees.

4.2. Generalizability Evaluation

As described in Section 2.3 a total of 3900 images, extracted from the VHRS orthomosaic UAV product, were selected as the testing dataset to evaluate the generalization capability of the proposed network. Figure 10 illustrates the segmentation quality metric for semantic segmentation of the testing dataset achieved by the trained deep learning models. Figure 11 displays nine randomly selected images from the testing dataset and their corresponding masks and segmentation outputs of the trained models. The proposed segmentation model demonstrates excellent generalization capacity because it achieved an F-score of 91% and Mean-IOU of 85%. The compared segmentation models also maintained a similar range of accuracies. The proposed model demonstrates an efficient model for date palm tree mapping from UAV images via the comparative evaluation of segmentation results.

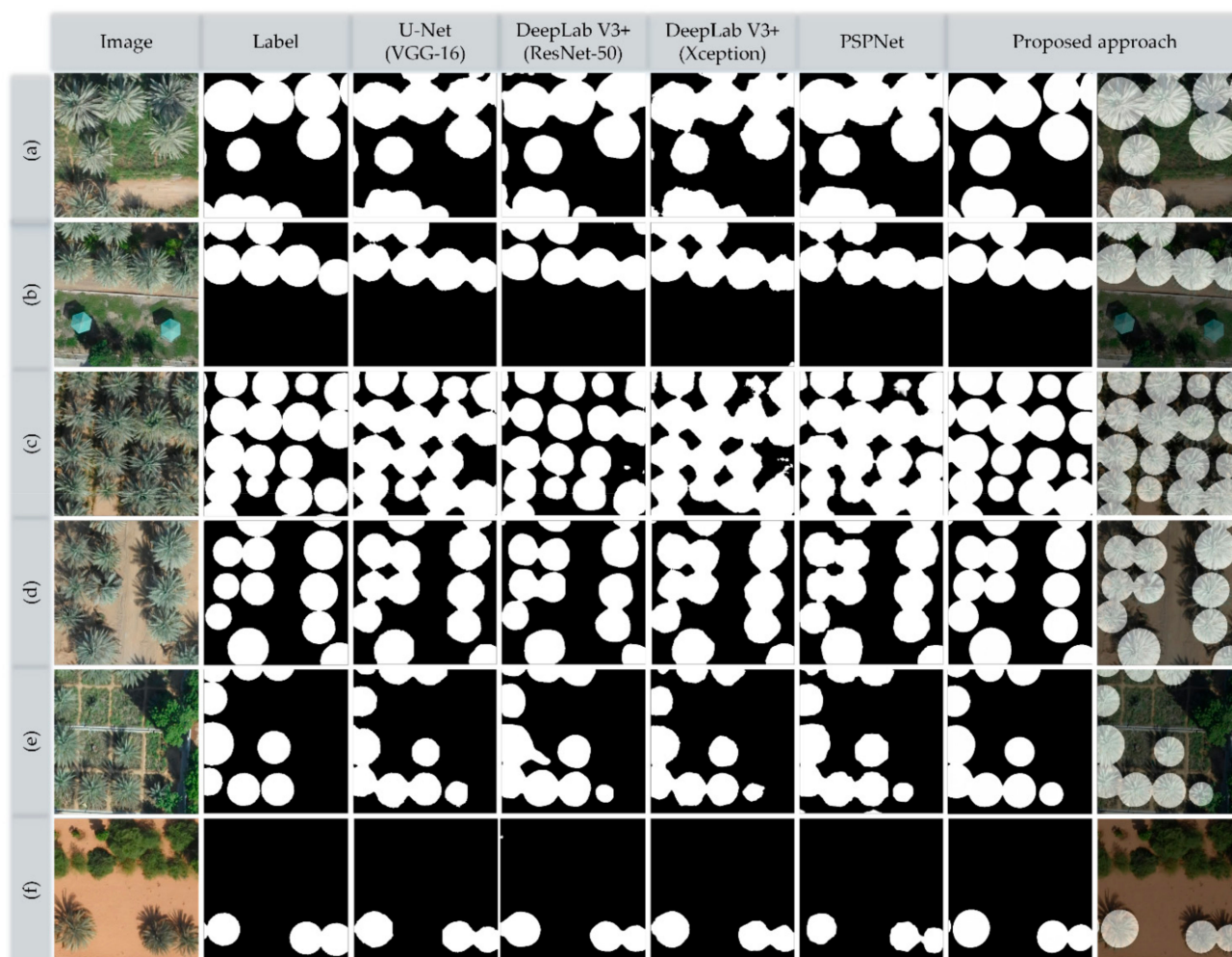


Figure 9. Comparison of six randomly selected images (a–f) from the validation dataset and their corresponding segmentation results. The left column shows the selected RGB images, followed by the corresponding ground truth label, and segmentation results of the evaluated models including U-Net (based on VGG-16 network), DeepLab V3+ (based on ResNet-50 network), DeepLab V3+ (based on Xception network), PSPNet and the proposed approach.

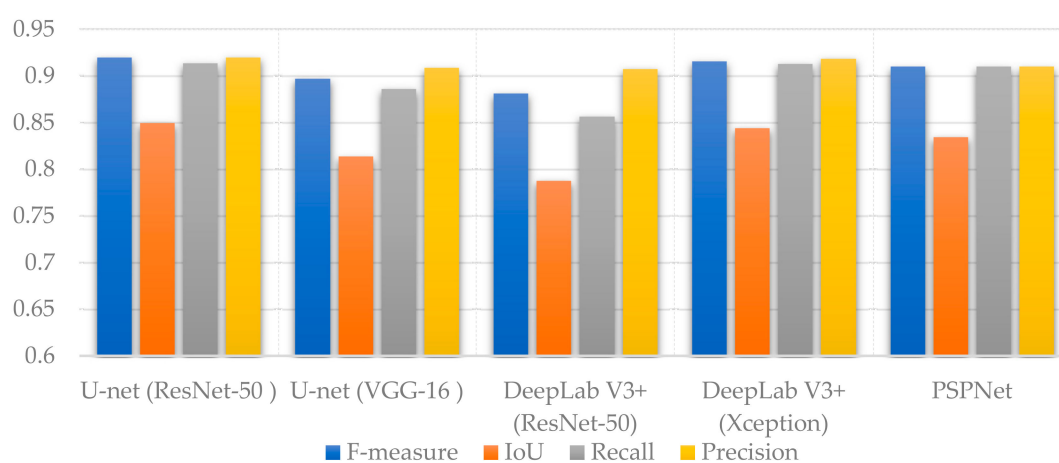


Figure 10. Summary of evaluation metrics of segmentation models obtained from the testing dataset.

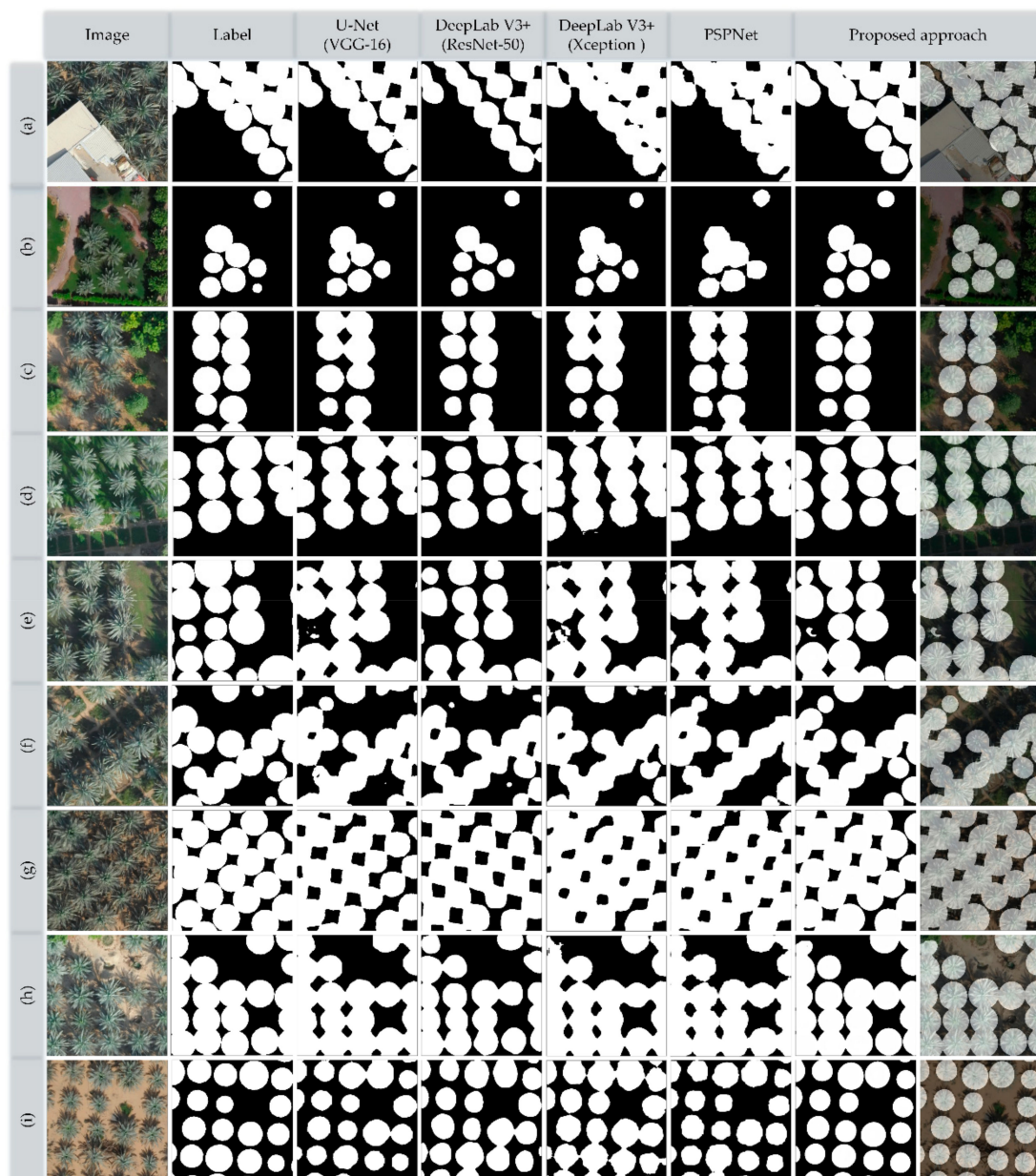


Figure 11. Comparison of nine randomly selected images (a–i) from the testing dataset and their corresponding segmentation results. The left side of the figure shows the selected RGB images followed by the corresponding ground truth label and classification results of the evaluated models.

5. Discussion

Large-scale mapping of date palm trees is essential for their consistent monitoring and sustainable management, given the substantial commercial, environmental, and landscaping values of date palm trees. Mapping and monitoring date palm trees using ground surveys is challenging because these trees are planted in different agricultural and urban environments. The increasing availability and continuous development of commercial UAV systems have amplified the popularity and utilization of UAV-based images in a wide range of earth-related studies. Different from satellite-based images, large-scale UAV images are acquired in different seasons, flight heights, spatial resolutions, weather conditions, sunlight angles, and image illuminations. Developing an accurate transferable approach for large-scale mapping of date palm trees from UAV images can be challenging because feature values may vary significantly based on the source of data, image object

segmentation level, and intraclass variability among classes, given the dependence of traditional machine learning techniques on the selection of shallow handcrafted features (i.e., band ratio, color invariants, and geometrical features). Thus, misclassification is expected when traditional machine learning is applied to different imageries [108].

In the current study, a deep semantic segmentation model based on U-Net architecture and a deep residual network was proposed for the large-scale mapping of date palm trees. A pretrained ResNet-50 based on ImageNet was adopted in the encoder module of U-Net. A comprehensive labeled dataset was developed to support the development of the proposed semantic segmentation for date palm trees from very-high spatial resolution (VHSR) UAV images. The labeled dataset was compiled from different agricultural and urban environments with a substantial variance in tree crown sizes, shapes, ages, health status, and backgrounds. The model was trained on eight GPUs (NVIDIA™ Tesla K80) through synchronous distributed training. The developed model was evaluated with independent validation and testing datasets. The performance of the proposed model was also compared with different advanced segmentation networks with various encoder backbones, including two variants of the DeepLab V3+ [147] (based on pretrained ResNet-50 and Xception backbones), PSPNet [149] (based on pretrained ResNet-50), and U-Net (based on a pretrained VGG-16) [143]. All segmentation models were tested on an NVIDIA™ Titan RTX graphics card with 24 Gb RAM. Table 2 compares the number of parameters, training time per epoch, and the testing time of the evaluated segmentation models.

Table 2. Training and testing details for each model.

Model	Backbone	No. of Trainable Parameters (M)	Training Time/Epoch (m)	Test Time (s)/Image
U-Net	ResNet-50	~157.280	~75	~0.17
U-Net	VGG-16	~25.858	~43	~0.1
DeepLab V3+	ResNet-50	~17.795	~25	~0.07
DeepLab V3+	Xception	~21.558	~29	~0.09
PSPNet	ResNet-50	~46.631	~65	~0.14

The proposed approach maintained high accuracy in the validation and testing datasets and indicated that date palm trees can be mapped with an average F-score (>91%) and Mean-IOU (>85%). With a F-score that ranges from 88% to 92% and a Mean-IOU that ranges from 78% to 85%, all of the evaluated segmentation models provided satisfactory segmentation results on the testing dataset. U-net architecture based on ResNet-50 architecture outperformed other segmentation models, with a F-score ranging from 1.2%–10% and Mean-IOU ranging from 2%–16%. This was followed by PSPNet (ResNet-50 Backbone), DeepLab V3+ (Xception Backbone), U-net (VGG-16 Backbone), and DeepLab V3+ (ResNet-50 Backbone). Numerous studies that have used and compared different deep semantic segmentation architectures for tree, crop, and vegetation mapping have reported similar ranges of segmentation metrics [57,70,100,102,117,125]. For instance, five semantic segmentation architectures, including SegNet, U-Net, FC-DenseNet, and DeepLab V3+ (based on Xception and MobileNetV2 backbones), were evaluated in Reference [100] for segmenting threatened single tree species from UAV-based images. An intersection-over-union that ranges from 77.1% to 92.5% and F1-score between 87.0% and 96.1% were reported by their experimental analysis. FC-DenseNet and U-Net models were superior to DeepLab V3+ (MobileNetV2), SegNet, and DeepLab V3+ (Xception). Cao and Zhang [112] utilized an improved U-Net model by replacing the convolutional layer in the U-Net network with a residual unit of ResNet for classifying different tree species from high-spatial-resolution airborne images. The developed approach was then followed by post-classification processing using conditional random fields to obtain smoother tree boundaries. An overall classification accuracy of 87% was achieved by the improved U-Net network. Ferreira et al. [70] achieved high accuracy by incorporating ResNet-18 in the DeepLab V3+ architecture to detect and classify Amazonian palm species from UAV images.

The proposed model in this study shows an efficient approach for date palm tree mapping from UAV images. It can segment date palm trees in relatively complex agricultural and urban environments and where palm trees are partially obscured by higher trees and shadow. Figures 9 and 11 depict the segmentation outputs of randomly selected images (512×512 pixels) from the validation and testing dataset. Although the differences in evaluation scores between ResUnet-50 and some of the evaluated architectures may not appear significant, it demonstrates better delineation of date palm trees. Considering that it is computationally intensive to train and test large UAV imagery in a deep semantic segmentation model, the UAV data were split into smaller tiles (512×512 pixels) and fed to the trained network and predicted the presence of date palm trees. In the splitting process of a large image, an overlap between the tiles may be considered to ensure better delineation of the palm trees around the edges of the generated image tiles. The final prediction is reconstructed by merging the segmentation outputs of the classified tiles. For instance, Figure 12 shows the segmentation output of the proposed model for different image tiles with larger sizes (5120×5120 pixels) without performing any post-processing operations. However, a minor misclassification might be encountered in the reconstructed product. For instance, when a palm tree is divided into two separate images, the shape/size of these predicted palm trees might vary slightly. In addition, some minor vertical lines can be observed, as shown in Figure 12d–f. These errors can be refined by post-processing computer-vision operations [70].

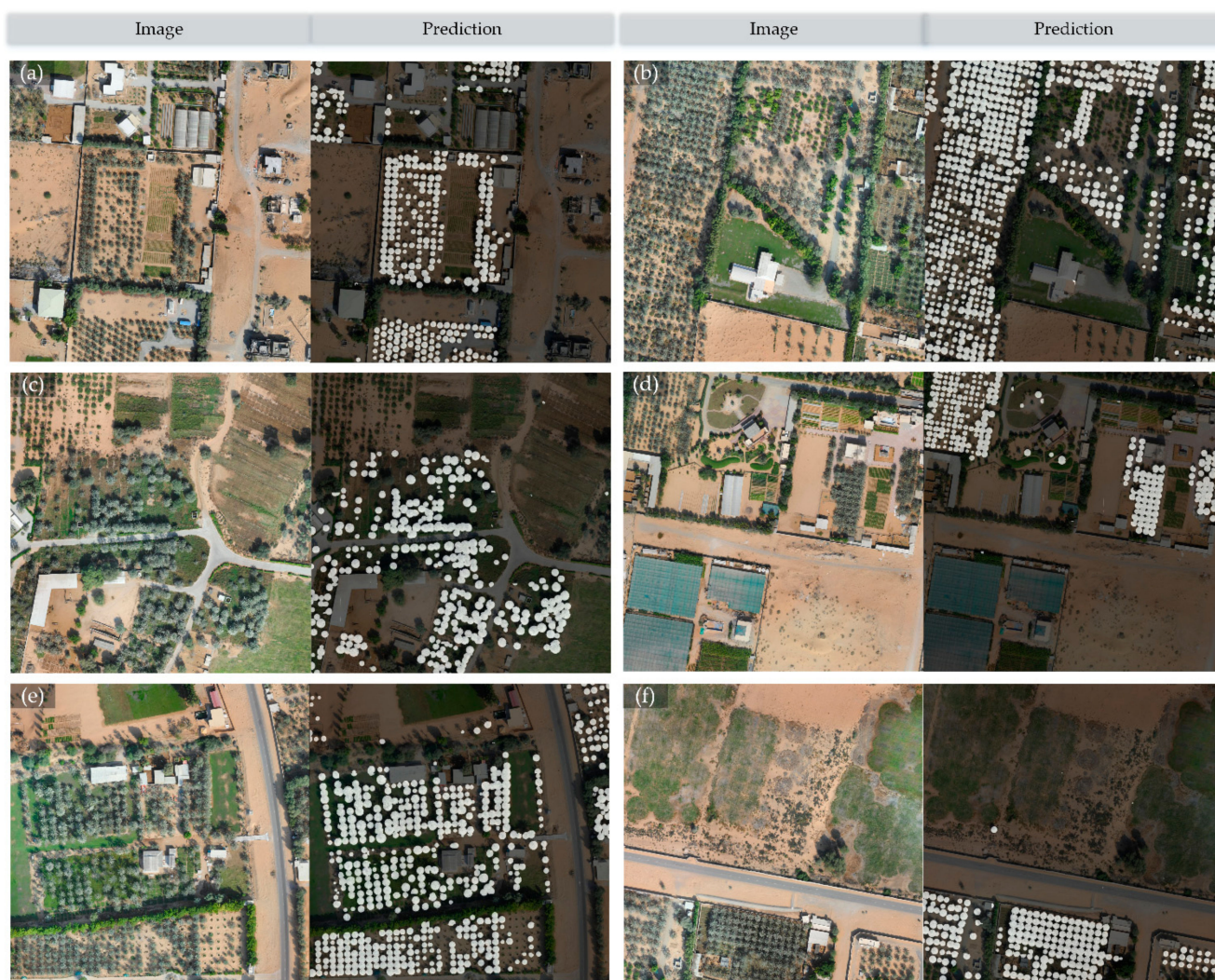


Figure 12. (a–f) Segmentation output of the proposed model for different randomly selected image tiles with larger sizes.

6. Conclusions

This study presented an automatic approach for the large-scale mapping of date palm trees from VHRS UAV images based on a deep semantic segmentation model. A pre-trained deep residual learning framework (ResNet-50) was used as the backbone of the encoder module of a U-Net. A large and diverse labeled dataset was created to aid in the development of the proposed semantic segmentation model. A distributed training strategy was used to train the model on multiple GPUs. The proposed segmentation model was evaluated with different state-of-the-art fully convolutional networks, including U-Net (VGG-16), PSPNet (based on ResNet-50), and two variations of DeepLab V3+ (ResNet-50 and Xception backbones). Experimental results showed that the proposed approach was superior to other semantic segmentation models in validation and testing datasets, achieving an F-score of 91% and Mean-IOU of 85%. The proposed deep fully convolutional network is an efficient tool for the accurate mapping and delineation of date palm tree VHRS UAV images, thereby building and updating geospatial databases and enabling consistent monitoring of date palm trees.

Author Contributions: Conceptualization, M.B.A.G. and H.Z.M.S.; methodology, M.B.A.G. and H.Z.M.S.; formal analysis, M.B.A.G.; writing—original draft preparation, M.B.A.G.; writing—review and editing, M.B.A.G., H.Z.M.S., A.S. and R.A.-R.; visualization, M.B.A.G.; Resources, H.Z.M.S., A.S. and R.A.-R.; supervision, H.Z.M.S., A.S., A.W. and S.J.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing is not applicable to this article.

Acknowledgments: The authors would like to acknowledge Universiti Putra Malaysia for the financial support, the municipality of Ajman for providing remotely sensed data of the study area, and the University of Sharjah for providing the high performance computing cluster used in this research.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Riad, M. The date palm sector in Egypt. *CIHEAM Options Mediterr.* **1996**, *53*, 45–53.
2. Tengberg, M. Beginnings and early history of date palm garden cultivation in the Middle East. *J. Arid Environ.* **2012**, *86*, 139–147. [CrossRef]
3. Zaid, A.; Wet, P.F. *Chapter I: Botanical and Systematic Description of the Date Palm*; FAO: Rome, Italy, 2002. Available online: [http://www.fao.org/docrep/006.Y4360E/y4360e05.htm](http://www.fao.org/docrep/006/Y4360E/y4360e05.htm) (accessed on 31 March 2018).
4. Spennemann, D.H.R. Review of the vertebrate-mediated dispersal of the date palm, *Phoenix dactylifera*. *Zool. Middle East.* **2018**, *64*, 283–296. [CrossRef]
5. Chao, C.C.T.; Krueger, R.R. The date palm (*Phoenix dactylifera* L.): Overview of biology, uses, and cultivation. *HortScience* **2007**, *42*, 1077–1082. [CrossRef]
6. Kurup, S.S.; Hedar, Y.S.; Al Dhaheri, M.A.; El-heawiety, A.Y.; Aly, M.A.M.; Alhadrami, G. Morpho-physiological evaluation and RAPD markers-assisted characterization of date palm (*Phoenix dactylifera* L.) varieties for salinity tolerance Morpho-physiological evaluation and RAPD markers-assisted characterization of date palm (*Phoenix dactylifera* L.). *J. Food Agric. Environ.* **2009**, *7*, 503–507.
7. Al-Alawi, R.; Al-Mashiqri, J.H.; Al-Nadabi, J.S.M.; Al-Shihi, B.I.; Baqi, Y. Date palm tree (*Phoenix dactylifera* L.): Natural products and therapeutic options. *Front. Plant Sci.* **2017**, *8*, 845. [CrossRef] [PubMed]
8. FAOSTAT. Available online: <http://www.fao.org/faostat/en/#data/QC> (accessed on 9 March 2021).
9. Culman, M.; Delalieux, S.; Van Tricht, K. Individual palm tree detection using deep learning on RGB imagery to support tree inventory. *Remote Sens.* **2020**, *12*, 3476. [CrossRef]
10. Pei, F.; Wu, C.; Liu, X.; Li, X.; Yang, K.; Zhou, Y.; Wang, K.; Xu, L.; Xia, G. Monitoring the vegetation activity in China using vegetation health indices. *Agric. For. Meteorol.* **2018**, *248*, 215–227. [CrossRef]
11. Xie, Y.; Sha, Z.; Yu, M. Remote sensing imagery in vegetation mapping: A review. *J. Plant Ecol.* **2008**, *1*, 9–23. [CrossRef]
12. Malatesta, L.; Scholte, P.T.; Vitale, M. Vegetation mapping from high-resolution satellite images in the heterogeneous arid environments of Socotra Island (Yemen). *J. Appl. Remote Sens.* **2019**. [CrossRef]

13. Zhao, A.; Zhang, A.; Liu, J.; Feng, L.; Zhao, Y. Assessing the effects of drought and “Grain for Green” Program on vegetation dynamics in China’s Loess Plateau from 2000 to 2014. *CATENA* **2019**, *175*, 446–455. [\[CrossRef\]](#)
14. Marston, C.; Aplin, P.; Wilkinson, D.; Field, R.; O’Regan, H. Scrubbing Up: Multi-scale investigation of woody encroachment in a Southern African savannah. *Remote Sens.* **2017**, *9*, 419. [\[CrossRef\]](#)
15. Spiekermann, R.; Brandt, M.; Samimi, C. Woody vegetation and land cover changes in the Sahel of Mali (1967–2011). *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *34*, 113–121. [\[CrossRef\]](#)
16. Hilker, T.; Lyapustin, A.I.; Hall, F.G.; Myneni, R.; Knyazikhin, Y.; Wang, Y.; Tucker, C.J.; Sellers, P.J. On the measurability of change in Amazon vegetation from MODIS. *Remote Sens. Environ.* **2015**, *166*, 233–242. [\[CrossRef\]](#)
17. Gärtner, P.; Förster, M.; Kurban, A.; Kleinschmit, B. Object based change detection of Central Asian Tugai vegetation with very high spatial resolution satellite imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *31*, 110–121. [\[CrossRef\]](#)
18. Kumagai, K. Verification of the analysis method for extracting the spatial continuity of the vegetation distribution on a regional scale. *Comput. Environ. Urban. Syst.* **2011**, *35*, 399–407. [\[CrossRef\]](#)
19. Disney, M. Remote sensing of vegetation: Potentials, limitations, developments and applications. In *Canopy Photosynthesis: From Basics to Applications*; Springer: Dordrecht, The Netherlands, 2016; pp. 289–331.
20. Senthilnath, J.; Kandukuri, M.; Dokania, A.; Ramesh, K.N. Application of UAV imaging platform for vegetation analysis based on spectral-spatial methods. *Comput. Electron. Agric.* **2017**, *140*, 8–24. [\[CrossRef\]](#)
21. Nebiker, S.; Annen, A.; Scherrer, M.; Oesch, D. A light-weight multispectral sensor for micro UAV Opportunities for very high resolution airborne remote sensing. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2008**, *37*, 1193–1199.
22. Candiago, S.; Remondino, F.; De Giglio, M.; Dubbini, M.; Gattelli, M. Evaluating multispectral images and vegetation indices for precision farming applications from UAV images. *Remote Sens.* **2015**, *7*, 4026–4047. [\[CrossRef\]](#)
23. Xiang, H.; Tian, L. Development of a low-cost agricultural remote sensing system based on an autonomous unmanned aerial vehicle (UAV). *Biosyst. Eng.* **2011**, *108*, 174–190. [\[CrossRef\]](#)
24. Komárek, J.; Klouček, T.; Prošek, J. The potential of unmanned aerial systems: A tool towards precision classification of hard-to-distinguish vegetation types? *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *71*, 9–19. [\[CrossRef\]](#)
25. Weil, G.; Lensky, I.; Resheff, Y.; Levin, N. Optimizing the timing of unmanned aerial vehicle image acquisition for applied mapping of woody vegetation species using feature selection. *Remote Sens.* **2017**, *9*, 1130. [\[CrossRef\]](#)
26. Husson, E.; Reese, H.; Ecke, F. Combining spectral data and a DSM from UAS-images for improved classification of non-submerged aquatic vegetation. *Remote Sens.* **2017**, *9*, 247. [\[CrossRef\]](#)
27. Michez, A.; Piégay, H.; Lisein, J.; Claessens, H.; Lejeune, P. Classification of riparian forest species and health condition using multi-temporal and hyperspatial imagery from unmanned aerial system. *Environ. Monit. Assess.* **2016**, *188*, 146. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Lisein, J.; Michez, A.; Claessens, H.; Lejeune, P. Discrimination of deciduous tree species from time series of unmanned aerial system imagery. *PLoS ONE* **2015**, *10*, e0141006. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Prošek, J.; Šimová, P. UAV for mapping shrubland vegetation: Does fusion of spectral and vertical information derived from a single sensor increase the classification accuracy? *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *75*, 151–162. [\[CrossRef\]](#)
30. Mishra, N.; Mainali, K.; Shrestha, B.; Radenz, J.; Karki, D. Species-level vegetation mapping in a Himalayan treeline ecotone using unmanned aerial system (UAS) imagery. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 445. [\[CrossRef\]](#)
31. Ishida, T.; Kurihara, J.; Viray, F.A.; Namuco, S.B.; Paringit, E.C.; Perez, G.J.; Takahashi, Y.; Marciano, J.J. A novel approach for vegetation classification using UAV-based hyperspectral imaging. *Comput. Electron. Agric.* **2018**, *144*, 80–85. [\[CrossRef\]](#)
32. Müllerová, J.; Brůna, J.; Bartaloš, T.; Dvořák, P.; Vítková, M.; Pyšek, P. Timing is important: Unmanned aircraft vs. satellite imagery in plant invasion monitoring. *Front. Plant Sci.* **2017**, *8*. [\[CrossRef\]](#)
33. Abeyasinghe, T.; Simic Milas, A.; Arend, K.; Hohman, B.; Reil, P.; Gregory, A.; Vázquez-Ortega, A. Mapping invasive *Phragmites australis* in the old woman creek estuary using UAV remote sensing and machine learning classifiers. *Remote Sens.* **2019**, *11*, 1380. [\[CrossRef\]](#)
34. Gaston, K.J.; Gonzalez, F.; Mengersen, K.; Gaston, K.J. UAVs and machine learning revolutionising invasive grass and vegetation surveys in remote arid lands. *Sensors* **2018**, *18*, 605. [\[CrossRef\]](#)
35. Calderón, R.; Navas-Cortés, J.A.; Lucena, C.; Zarco-Tejada, P.J. High-resolution airborne hyperspectral and thermal imagery for early detection of Verticillium wilt of olive using fluorescence, temperature and narrow-band spectral indices. *Remote Sens. Environ.* **2013**, *139*, 231–245. [\[CrossRef\]](#)
36. Nishar, A.; Richards, S.; Breen, D.; Robertson, J.; Breen, B. Thermal infrared imaging of geothermal environments and by an unmanned aerial vehicle (UAV): A case study of the Wairakei—Tauhara geothermal field, Taupo, New Zealand. *Renew. Energy* **2016**, *86*, 1256–1264. [\[CrossRef\]](#)
37. Pérez-Ortiz, M.; Peña, J.M.; Gutiérrez, P.A.; Torres-Sánchez, J.; Hervás-Martínez, C.; López-Granados, F. A semi-supervised system for weed mapping in sunflower crops using unmanned aerial vehicles and a crop row detection method. *Appl. Soft Comput.* **2015**, *37*, 533–544. [\[CrossRef\]](#)
38. Wu, Z.; Ni, M.; Hu, Z.; Wang, J.; Li, Q.; Wu, G. Mapping invasive plant with UAV-derived 3D mesh model in mountain area—A case study in Shenzhen Coast, China. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *77*, 129–139. [\[CrossRef\]](#)
39. Zhang, X.; Han, L.; Dong, Y.; Shi, Y.; Huang, W.; Han, L.; González-Moreno, P.; Ma, H.; Ye, H.; Sobeih, T. A deep learning-based approach for automated yellow rust disease detection from high-resolution hyperspectral UAV images. *Remote Sens.* **2019**, *11*, 1554. [\[CrossRef\]](#)

40. Vanegas, F.; Bratanov, D.; Powell, K.; Weiss, J.; Gonzalez, F. A novel methodology for improving plant pest surveillance in vineyards and crops using UAV-based hyperspectral and spatial data. *Sensors* **2018**, *18*, 260. [\[CrossRef\]](#)
41. Cummings, A.R.; Karale, Y.; Cummings, G.R.; Hamer, E.; Moses, P.; Norman, Z.; Captain, V. UAV-derived data for mapping change on a swidden agriculture plot: Preliminary results from a pilot study. *Int. J. Remote Sens.* **2017**, *38*, 2066–2082. [\[CrossRef\]](#)
42. Liu, H.; Zhu, H.; Wang, P. Quantitative modelling for leaf nitrogen content of winter wheat using UAV-based hyperspectral data. *Int. J. Remote Sens.* **2017**, *38*, 2117–2134. [\[CrossRef\]](#)
43. Mesas-Carrascosa, F.J.; Clavero Rumbao, I.; Torres-Sánchez, J.; García-Ferrer, A.; Peña, J.M.; López Granados, F. Accurate ortho-mosaicked six-band multispectral UAV images as affected by mission planning for precision agriculture proposes. *Int. J. Remote Sens.* **2017**, *38*, 2161–2176. [\[CrossRef\]](#)
44. Yue, J.; Feng, H.; Jin, X.; Yuan, H.; Li, Z.; Zhou, C.; Yang, G.; Tian, Q. A comparison of crop parameters estimation using images from UAV-mounted snapshot hyperspectral sensor and high-definition digital camera. *Remote Sens.* **2018**, *10*, 1138. [\[CrossRef\]](#)
45. Kanning, M.; Kühling, I.; Trautz, D.; Jarmer, T. High-resolution UAV-based hyperspectral imagery for LAI and chlorophyll estimations from wheat for yield prediction. *Remote Sens.* **2018**, *10*, 2000. [\[CrossRef\]](#)
46. Wei, L.; Yu, M.; Zhong, Y.; Zhao, J.; Liang, Y.; Hu, X. Spatial-spectral fusion based on conditional random fields for the fine classification of crops in UAV-borne hyperspectral remote sensing imagery. *Remote Sens.* **2019**, *11*, 780. [\[CrossRef\]](#)
47. Nevalainen, O.; Honkavaara, E.; Tuominen, S.; Viljanen, N.; Hakala, T.; Yu, X.; Hyypä, J.; Saari, H.; Pölönen, I.; Imai, N.N.; et al. Individual tree detection and classification with UAV-Based photogrammetric point clouds and hyperspectral imaging. *Remote Sens.* **2017**, *9*, 185. [\[CrossRef\]](#)
48. Dos Santos, A.A.; Marcato Junior, J.; Araújo, M.S.; Di Martini, D.R.; Tetila, E.C.; Siqueira, H.L.; Aoki, C.; Eltner, A.; Matsubara, E.T.; Pistori, H.; et al. Assessment of CNN-based methods for individual tree detection on images captured by RGB cameras attached to UAVs. *Sensors* **2019**, *19*, 3595. [\[CrossRef\]](#)
49. Bakambekova, A.; James, A.P. Deep learning theory simplified. In *Deep Learning Classifiers with Memristive Networks; Modeling and Optimization in Science and Technologies*; Springer Nature Switzerland AG: Cham, Switzerland, 2020; Volume 14, pp. 41–55. [\[CrossRef\]](#)
50. Zhang, L.; Zhang, L.; Du, B. Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geosci. Remote Sens. Mag.* **2016**, *4*, 22–40. [\[CrossRef\]](#)
51. Malambo, L.; Popescu, S.; Ku, N.; Rooney, W.; Zhou, T.; Moore, S. A deep learning semantic segmentation-based approach for field-level sorghum panicle counting. *Remote Sens.* **2019**, *11*, 2939. [\[CrossRef\]](#)
52. Ampatzidis, Y.; Partel, V. UAV-based high throughput phenotyping in citrus utilizing multispectral imaging and artificial intelligence. *Remote Sens.* **2019**, *11*, 410. [\[CrossRef\]](#)
53. Zhou, K.; Ming, D.; Lv, X.; Fang, J.; Wang, M. CNN-based land cover classification combining stratified segmentation and fusion of point cloud and very high-spatial resolution remote sensing image data. *Remote Sens.* **2019**, *11*, 2065. [\[CrossRef\]](#)
54. Nevavuori, P.; Narra, N.; Lipping, T. Crop yield prediction with deep convolutional neural networks. *Comput. Electron. Agric.* **2019**, *163*, 104859. [\[CrossRef\]](#)
55. Yalcin, H. Plant phenology recognition using deep learning: Deep-Pheno. In Proceedings of the 2017 6th International Conference on Agro-Geoinformatics, Fairfax, VA, USA, 7–10 August 2017; pp. 1–5.
56. Ma, L.; Liu, Y.; Zhang, X.; Ye, Y.; Yin, G.; Johnson, B.A. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS J. Photogramm. Remote Sens.* **2019**, *152*, 166–177. [\[CrossRef\]](#)
57. Ji, S.; Zhang, Z.; Zhang, C.; Wei, S.; Lu, M.; Duan, Y. Learning discriminative spatiotemporal features for precise crop classification from multi-temporal satellite images. *Int. J. Remote Sens.* **2020**, *41*, 3162–3174. [\[CrossRef\]](#)
58. Bosilj, P.; Aptoula, E.; Duckett, T.; Cielniak, G. Transfer learning between crop types for semantic segmentation of crops versus weeds in precision agriculture. *J. Field Robot.* **2020**, *37*, 7–19. [\[CrossRef\]](#)
59. Lv, Y.; Zhang, C.; Yun, W.; Gao, L.; Wang, H.; Ma, J.; Li, H.; Zhu, D. The delineation and grading of actual crop production units in modern smallholder areas using RS Data and Mask R-CNN. *Remote Sens.* **2020**, *12*, 1074. [\[CrossRef\]](#)
60. Bah, M.D.; Dericquebourg, E.; Hafiane, A.; Canals, R. Deep learning based classification system for identifying weeds using high-resolution UAV imagery. In *Proceedings of the Science and Information Conference*; Springer: Cham, Switzerland, 2019; pp. 176–187.
61. Huang, H.; Deng, J.; Lan, Y.; Yang, A.; Deng, X.; Zhang, L. A fully convolutional network for weed mapping of unmanned aerial vehicle (UAV) imagery. *PLoS ONE* **2018**, *13*, e0196302. [\[CrossRef\]](#)
62. Hasan, A.S.M.M.; Sohel, F.; Diepeveen, D.; Laga, H.; Jones, M.G.K. A survey of deep learning techniques for weed detection from images. *Comput. Electron. Agric.* **2021**, *184*, 106067. [\[CrossRef\]](#)
63. Mazzia, V.; Comba, L.; Khaliq, A.; Chiaberge, M.; Gay, P. UAV and machine learning based refinement of a satellite-driven vegetation index for precision agriculture. *Sensors* **2020**, *20*, 2530. [\[CrossRef\]](#) [\[PubMed\]](#)
64. Sharpe, S.M.; Schumann, A.W.; Yu, J.; Boyd, N.S. Vegetation detection and discrimination within vegetable plasticulture row-middles using a convolutional neural network. *Precis. Agric.* **2019**, 1–14. [\[CrossRef\]](#)
65. Bayr, U.; Puschmann, O. Automatic detection of woody vegetation in repeat landscape photographs using a convolutional neural network. *Ecol. Inform.* **2019**, *50*, 220–233. [\[CrossRef\]](#)
66. Ganchenko, V.; Doudkin, A. Agricultural vegetation monitoring based on aerial data using convolutional neural networks. *Opt. Mem. Neural Netw.* **2019**, *28*, 129–134. [\[CrossRef\]](#)

67. Neupane, B.; Horanont, T.; Hung, N.D. Deep learning based banana plant detection and counting using high-resolution red-green-blue (RGB) images collected from unmanned aerial vehicle (UAV). *PLoS ONE* **2019**, *14*, e0223906. [\[CrossRef\]](#)
68. Kattenborn, T.; Leitloff, J.; Schiefer, F.; Hinz, S. Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2021**, *173*, 24–49. [\[CrossRef\]](#)
69. Braga, J.R.G.; Peripato, V.; Dalagnol, R.; Ferreira, M.P.; Tarabalka, Y.; Aragão, L.E.O.C.; de Campos Velho, H.F.; Shiguemori, E.H.; Wagner, F.H. Tree crown delineation algorithm based on a convolutional neural network. *Remote Sens.* **2020**, *12*, 1288. [\[CrossRef\]](#)
70. Ferreira, M.P.; de Almeida, D.R.A.; de Almeida Papa, D.; Minervino, J.B.S.; Veras, H.F.P.; Formighieri, A.; Santos, C.A.N.; Ferreira, M.A.D.; Figueiredo, E.O.; Ferreira, E.J.L. Individual tree detection and species classification of Amazonian palms using UAV images and deep learning. *For. Ecol. Manag.* **2020**, *475*, 118397. [\[CrossRef\]](#)
71. Roslan, Z.; Long, Z.A.; Husen, M.N.; Ismail, R.; Hamzah, R. Deep learning for tree crown detection in tropical forest. In Proceedings of the 2020 14th International Conference on Ubiquitous Information Management and Communication (IMCOM), Taichung, Taiwan, 3–5 January 2020. [\[CrossRef\]](#)
72. Weinstein, B.G.; Marconi, S.; Bohlman, S.; Zare, A.; White, E. Individual tree-crown detection in rgb imagery using semi-supervised deep learning neural networks. *Remote Sens.* **2019**, *11*, 1309. [\[CrossRef\]](#)
73. Kerkech, M.; Hafiane, A.; Canals, R. Deep leaning approach with colorimetric spaces and vegetation indices for vine diseases detection in UAV images. *Comput. Electron. Agric.* **2018**, *155*, 237–243. [\[CrossRef\]](#)
74. Dang, L.M.; Ibrahim Hassan, S.; Suhyeon, I.; Sangaiah, A.K.; Mehmood, I.; Rho, S.; Seo, S.; Moon, H. UAV based wilt detection system via convolutional neural networks. *Sustain. Comput. Inform. Syst.* **2018**. [\[CrossRef\]](#)
75. Hasan, M.; Tanawala, B.; Patel, K.J. Deep learning precision farming: Tomato leaf disease detection by transfer learning. *SSRN Electron. J.* **2019**. [\[CrossRef\]](#)
76. Castela Tetila, E.; Brandoli Machado, B.; Menezes, G.K.; da Silva Oliveira, A.; Alvarez, M.; Amorim, W.P.; de Souza Belete, N.A.; da Silva, G.G.; Pistori, H. Automatic recognition of soybean leaf diseases using UAV images and deep convolutional neural networks. *IEEE Geosci. Remote Sens. Lett.* **2019**, 1–5. [\[CrossRef\]](#)
77. Bajpai, G.; Gupta, A.; Chauhan, N. Real time implementation of convolutional neural network to detect plant diseases using internet of things. In *International Symposium on VLSI Design and Test*; Springer: Singapore, 2019; pp. 510–522.
78. Kattenborn, T.; Eichel, J.; Wiser, S.; Burrows, L.; Fassnacht, F.E.; Schmidtlein, S. Convolutional neural networks accurately predict cover fractions of plant species and communities in unmanned aerial vehicle imagery. *Remote Sens. Ecol. Conserv.* **2020**, 1–15. [\[CrossRef\]](#)
79. Hamylton, S.M.; Morris, R.H.; Carvalho, R.C.; Roder, N.; Barlow, P.; Mills, K.; Wang, L. Evaluating techniques for mapping island vegetation from unmanned aerial vehicle (UAV) images: Pixel classification, visual interpretation and machine learning approaches. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *89*, 102085. [\[CrossRef\]](#)
80. Fan, Z.; Lu, J.; Gong, M.; Xie, H.; Goodman, E.D. Automatic tobacco plant detection in UAV images via deep neural networks. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 876–887. [\[CrossRef\]](#)
81. Yang, Q.; Shi, L.; Han, J.; Yu, J.; Huang, K. A near real-time deep learning approach for detecting rice phenology based on UAV images. *Agric. For. Meteorol.* **2020**, *287*, 107938. [\[CrossRef\]](#)
82. Nezami, S.; Khoramshahi, E.; Nevalainen, O.; Pölönen, I.; Honkavaara, E. Tree species classification of drone hyperspectral and RGB imagery with deep learning convolutional neural networks. *Remote Sens.* **2020**, *12*, 1070. [\[CrossRef\]](#)
83. Qian, W.; Huang, Y.; Liu, Q.; Fan, W.; Sun, Z.; Dong, H.; Wan, F.; Qiao, X. UAV and a deep convolutional neural network for monitoring invasive alien plants in the wild. *Comput. Electron. Agric.* **2020**, *174*, 105519. [\[CrossRef\]](#)
84. Bonet, I.; Caraffini, F.; Pena, A.; Puerta, A.; Gongora, M. Oil palm detection via deep transfer learning. In Proceedings of the 2020 IEEE Congress on Evolutionary Computation (CEC), Glasgow, UK, 19–24 July 2020. [\[CrossRef\]](#)
85. Tao, H.; Li, C.; Zhao, D.; Deng, S.; Hu, H.; Xu, X.; Jing, W. Deep learning-based dead pine tree detection from unmanned aerial vehicle images. *Int. J. Remote Sens.* **2020**, *41*, 8238–8255. [\[CrossRef\]](#)
86. Zhang, C.; Xia, K.; Feng, H.; Yang, Y.; Du, X. Tree species classification using deep learning and RGB optical images obtained by an unmanned aerial vehicle. *J. For. Res.* **2020**. [\[CrossRef\]](#)
87. Nguyen, H.T.; Caceres, M.L.L.; Moritake, K.; Kentsch, S.; Shu, H.; Diez, Y. Individual sick fir tree (*Abies mariesii*) identification in insect infested forests by means of UAV images and deep learning. *Remote Sens.* **2021**, *13*, 260. [\[CrossRef\]](#)
88. Safonova, A.; Guirado, E.; Maglinets, Y.; Alcaraz-Segura, D.; Tabik, S. Olive tree biovolume from uav multi-resolution image segmentation with mask r-cnn. *Sensors* **2021**, *21*, 1617. [\[CrossRef\]](#) [\[PubMed\]](#)
89. Fromm, M.; Schubert, M.; Castilla, G.; Linke, J.; McDermid, G. Automated detection of conifer seedlings in drone imagery using convolutional neural networks. *Remote Sens.* **2019**, *11*, 2585. [\[CrossRef\]](#)
90. Ocer, N.E.; Kaplan, G.; Erdem, F.; Kucuk Matci, D.; Avdan, U. Tree extraction from multi-scale UAV images using Mask R-CNN with FPN. *Remote Sens. Lett.* **2020**, *11*, 847–856. [\[CrossRef\]](#)
91. Pulido, D.; Salas, J.; Rös, M.; Puettmann, K.; Karaman, S. Assessment of tree detection methods in multispectral aerial images. *Remote Sens.* **2020**, *12*, 2379. [\[CrossRef\]](#)
92. Liu, X.; Ghazali, K.H.; Han, F.; Mohamed, I.I. Automatic detection of oil palm tree from UAV images based on the deep learning method. *Appl. Artif. Intell.* **2021**, *35*, 13–24. [\[CrossRef\]](#)
93. Zheng, J.; Fu, H.; Li, W.; Wu, W.; Yu, L.; Yuan, S.; Tao, W.Y.W.; Pang, T.K.; Kanniah, K.D. Growing status observation for oil palm trees using Unmanned Aerial Vehicle (UAV) images. *ISPRS J. Photogramm. Remote Sens.* **2021**, *173*, 95–121. [\[CrossRef\]](#)

94. Barmpoutis, P.; Kamperidou, V.; Stathaki, T. Estimation of extent of trees and biomass infestation of the suburban forest of Thessaloniki (Seich Sou) using UAV imagery and combining R-CNNs and multichannel texture analysis. In Proceedings of the Twelfth International Conference on Machine Vision (ICMV 2019), Amsterdam, The Netherlands, 16–18 November 2019. [\[CrossRef\]](#)
95. Weinstein, B.G.; Marconi, S.; Bohlman, S.A.; Zare, A.; White, E.P. Cross-site learning in deep learning RGB tree crown detection. *Ecol. Inform.* **2020**, *56*, 101061. [\[CrossRef\]](#)
96. Liu, Y.; Cen, C.; Che, Y.; Ke, R.; Ma, Y.; Ma, Y. Detection of maize tassels from UAV RGB imagery with faster R-CNN. *Remote Sens.* **2020**, *12*, 338. [\[CrossRef\]](#)
97. Osco, L.P.; dos Santos de Arruda, M.; Marcato Junior, J.; da Silva, N.B.; Ramos, A.P.M.; Moryia, É.A.S.; Imai, N.N.; Pereira, D.R.; Creste, J.E.; Matsubara, E.T.; et al. A convolutional neural network approach for counting and geolocating citrus-trees in UAV multispectral imagery. *ISPRS J. Photogramm. Remote Sens.* **2020**, *160*, 97–106. [\[CrossRef\]](#)
98. Yang, M.-D.; Tseng, H.H.; Hsu, Y.C.; Tsai, H.P. Semantic segmentation using deep learning with vegetation indices for rice lodging identification in multi-date UAV visible images. *Remote Sens.* **2020**, *12*, 633. [\[CrossRef\]](#)
99. Huang, H.; Lan, Y.; Yang, A.; Zhang, Y.; Wen, S.; Deng, J. Deep learning versus object-based image analysis (OBIA) in weed mapping of UAV imagery. *Int. J. Remote Sens.* **2020**, *41*, 3446–3479. [\[CrossRef\]](#)
100. Torres, D.L.; Feitosa, R.Q.; Happ, P.N.; La Rosa, L.E.C.; Junior, J.M.; Martins, J.; Bressan, P.O.; Gonçalves, W.N.; Liesenberg, V. Applying fully convolutional architectures for semantic segmentation of a single tree species in urban environment on high resolution UAV optical imagery. *Sensors* **2020**, *20*, 563. [\[CrossRef\]](#)
101. Kattenborn, T.; Eichel, J.; Fassnacht, F.E. Convolutional neural networks enable efficient, accurate and fine-grained segmentation of plant species and communities from high-resolution UAV imagery. *Sci. Rep.* **2019**, *9*, 17656. [\[CrossRef\]](#) [\[PubMed\]](#)
102. Liu, C.; Li, H.; Su, A.; Chen, S.; Li, W. Identification and grading of maize drought on RGB images of UAV based on improved U-net. *IEEE Geosci. Remote Sens. Lett.* **2021**, *18*, 198–202. [\[CrossRef\]](#)
103. Wu, J.; Yang, G.; Yang, H.; Zhu, Y.; Li, Z.; Lei, L.; Zhao, C. Extracting apple tree crown information from remote imagery using deep learning. *Comput. Electron. Agric.* **2020**, *174*, 105504. [\[CrossRef\]](#)
104. Liu, T.; Abd-Elrahman, A.; Morton, J.; Wilhelm, V.L. Comparing fully convolutional networks, random forest, support vector machine, and patch-based deep convolutional neural networks for object-based wetland mapping using images from small unmanned aircraft system. *GISci. Remote Sens.* **2018**, *55*, 243–264. [\[CrossRef\]](#)
105. Morales, G.; Kemper, G.; Sevillano, G.; Arteaga, D.; Ortega, I.; Telles, J. Automatic segmentation of *Mauritia flexuosa* in unmanned aerial vehicle (UAV) imagery using deep learning. *Forests* **2018**, *9*, 736. [\[CrossRef\]](#)
106. Kentsch, S.; Caceres, M.L.L.; Serrano, D.; Roure, F.; Diez, Y. Computer vision and deep learning techniques for the analysis of drone-acquired forest images, a transfer learning study. *Remote Sens.* **2020**, *12*, 1287. [\[CrossRef\]](#)
107. Tang, T.Y.; Fu, B.L.; Lou, P.Q.; Bi, L. Segnet-based extraction of wetland vegetation information from UAV images. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *42*, 375–380. [\[CrossRef\]](#)
108. Al-Ruzouq, R.; Gibril, M.B.A.; Shanableh, A.; Kais, A.; Hamed, O.; Al-Mansoori, S.; Khalil, M.A. Sensors, features, and machine learning for oil spill detection and monitoring: A review. *Remote Sens.* **2020**, *12*, 3338. [\[CrossRef\]](#)
109. Badrinarayanan, V.; Handa, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling. *arXiv* **2015**, arXiv:1505.07293.
110. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2015; Volume 9351, pp. 234–241.
111. Chen, L.-C.; Papandreou, G.; Schroff, F.; Adam, H. Rethinking atrous convolution for semantic image segmentation. *arXiv* **2017**, arXiv:1706.05587.
112. Cao, K.; Zhang, X. An improved Res-UNet model for tree species classification using airborne high-resolution images. *Remote Sens.* **2020**, *12*, 1128. [\[CrossRef\]](#)
113. Flood, N.; Watson, F.; Collett, L. Using a U-net convolutional neural network to map woody vegetation extent from high resolution satellite imagery across Queensland, Australia. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *82*, 101897. [\[CrossRef\]](#)
114. Xiao, C.; Qin, R.; Huang, X. Treetop detection using convolutional neural networks trained through automatically generated pseudo labels. *Int. J. Remote Sens.* **2020**, *41*, 3010–3030. [\[CrossRef\]](#)
115. Wagner, F.H.; Sanchez, A.; Tarabalka, Y.; Lotte, R.G.; Ferreira, M.P.; Aidar, M.P.M.; Gloor, E.; Phillips, O.L.; Aragão, L.E.O.C. Using the U-net convolutional network to map forest types and disturbance in the Atlantic rainforest with very high resolution images. *Remote Sens. Ecol. Conserv.* **2019**, *5*, 360–375. [\[CrossRef\]](#)
116. Nogueira, K.; Santos, J.A.; Cancian, L.; Borges, B.D.; Silva, T.S.F.; Morellato, L.P.; Torres, R.S. Semantic segmentation of vegetation images acquired by unmanned aerial vehicles using an ensemble of ConvNets. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; Volume 2, pp. 3787–3790.
117. Bhatnagar, S.; Gill, L.; Ghosh, B. Drone image segmentation using machine and deep learning for mapping raised bog vegetation communities. *Remote Sens.* **2020**, *12*, 2602. [\[CrossRef\]](#)
118. Wagner, F.H.; Sanchez, A.; Aidar, M.P.M.; Rochelle, A.L.C.; Tarabalka, Y.; Fonseca, M.G.; Phillips, O.L.; Gloor, E.; Aragão, L.E.O.C. Mapping Atlantic rainforest degradation and regeneration history with indicator species using convolutional network. *PLoS ONE* **2020**, *15*, e0229448. [\[CrossRef\]](#)

119. Liu, J.; Wang, X.; Wang, T. Classification of tree species and stock volume estimation in ground forest images using deep learning. *Comput. Electron. Agric.* **2019**, *166*, 105012. [\[CrossRef\]](#)
120. Kentsch, S.; Karatsiolis, S.; Kamilaris, A.; Tomhave, L.; Lopez Caceres, M.L. Identification of tree species in Japanese forests based on aerial photography and deep learning. In *Advances and New Trends in Environmental Informatics*; Springer: Cham, Switzerland, 2020.
121. Korznikov, K.A.; Kislov, D.E.; Altman, J.; Doležal, J.; Vozmishcheva, A.S.; Krestov, P.V. Using u-net-like deep convolutional neural networks for precise tree recognition in very high resolution rgb (Red, green, blue) satellite images. *Forests* **2021**, *12*, 66. [\[CrossRef\]](#)
122. Ayhan, B.; Kwan, C. Tree, shrub, and grass classification using only RGB images. *Remote Sens.* **2020**, *12*, 1333. [\[CrossRef\]](#)
123. Ayhan, B.; Kwan, C.; Budavari, B.; Kwan, L.; Lu, Y.; Perez, D.; Li, J.; Skarlatos, D.; Vlachos, M. Vegetation detection using deep learning and conventional methods. *Remote Sens.* **2020**, *12*, 2502. [\[CrossRef\]](#)
124. Ayhan, B.; Kwan, C.; Larkin, J.; Kwan, L.M.; Skarlatos, D.P.; Vlachos, M. Deep learning models for accurate vegetation classification using RGB image only. In Proceedings of the SPIE Defense + Commercial Sensing, Online Only, 21 April 2020. [\[CrossRef\]](#)
125. Wang, S.; Xu, Z.; Zhang, C.; Zhang, J.; Mu, Z.; Zhao, T.; Wang, Y.; Gao, S.; Yin, H.; Zhang, Z. Improved winter wheat spatial distribution extraction using a convolutional neural network and partly connected conditional random field. *Remote Sens.* **2020**, *12*, 821. [\[CrossRef\]](#)
126. Lin, Z.; Guo, W. Sorghum panicle detection and counting using unmanned aerial system images and deep learning. *Front. Plant Sci.* **2020**, *11*, 1–13. [\[CrossRef\]](#) [\[PubMed\]](#)
127. Du, L.; McCarty, G.W.; Zhang, X.; Lang, M.W.; Vanderhoof, M.K.; Li, X.; Huang, C.; Lee, S.; Zou, Z. Mapping forested wetland inundation in the delmarva peninsula, USA using deep convolutional neural networks. *Remote Sens.* **2020**, *12*, 644. [\[CrossRef\]](#)
128. Freudenberg, M.; Nölke, N.; Agostini, A.; Urban, K.; Wörgötter, F.; Kleinn, C. Large scale palm tree detection in high resolution satellite images using U-Net. *Remote Sens.* **2019**, *11*, 312. [\[CrossRef\]](#)
129. Dong, R.; Li, W.; Fu, H.; Gan, L.; Yu, L.; Zheng, J.; Xia, M. Oil palm plantation mapping from high-resolution remote sensing images using deep learning. *Int. J. Remote Sens.* **2020**, *41*, 2022–2046. [\[CrossRef\]](#)
130. Mihi, A.; Nacer, T.; Chenchouni, H. *Monitoring Dynamics of Date Palm Plantations from 1984 to 2013 Using Landsat Time-Series in Sahara Desert Oases of Algeria*; Springer: Berlin/Heidelberg, Germany, 2019; ISBN 9783030014407.
131. Mulley, M.; Kooistra, L.; Bierens, L. High-resolution multisensor remote sensing to support date palm farm high-resolution multisensor remote sensing to support date palm farm management. *Agriculture* **2019**, *9*, 26. [\[CrossRef\]](#)
132. Shareef, M.A. Estimation and mapping of dates palm using landsat-8 images: A case study in Baghdad City. In Proceedings of the 2018 International Conference on Advanced Science and Engineering (ICOASE), Duhok, Iraq, 9–11 October 2018; pp. 425–430.
133. Issa, S.; Dahy, B.; Saleous, N. Mapping and assessing above ground biomass (AGB) of date palm plantations using remote sensing and GIS: A case study from Abu Dhabi, United Arab Emirates. In Proceedings of the Remote Sensing for Agriculture, Ecosystems, and Hydrology XXI, Strasbourg, France, 9–11 September 2019.
134. Mazlounzadeh, S.M.; Shamsi, M.; Nezamabadi-pour, H. Fuzzy logic to classify date palm trees based on some physical properties related to precision agriculture. *Precis. Agric.* **2010**, 258–273. [\[CrossRef\]](#)
135. Al-Ruzouq, R.; Shanableh, A.; Barakat, A. Gibril, M.; AL-Mansoori, S.; Al-Ruzouq, R.; Shanableh, A.; Barakat, A. Gibril, M.; AL-Mansoori, S. Image segmentation parameter selection and ant colony optimization for date palm tree detection and mapping from very-high-spatial-resolution aerial imagery. *Remote Sens.* **2018**, *10*, 1413. [\[CrossRef\]](#)
136. Culman, M.; Delalieux, S.; Van Tricht, K. Palm tree inventory from aerial images using retinanet. In Proceedings of the 2020 Mediterranean and Middle-East Geoscience and Remote Sensing Symposium (M2GARSS), Tunis, Tunisia, 9–11 March 2020; pp. 314–317.
137. Tadesse, W.; Halila, H.; Jamal, M.; Assefa, S.; Oweis, T.; Baum, M. Role of sustainable wheat production to ensure food security in the CWANA region. *J. Exp. Biol. Agric. Sci.* **2017**, *5*. [\[CrossRef\]](#)
138. Yilmaz, A.G.; Shanableh, A.; Merabtene, T.; Atabay, S.; Kayemah, N. Rainfall trends and intensity-frequency-duration relationships in Sharjah City, UAE. *Int. J. Hydrol. Sci. Technol.* **2020**, *10*, 487–503. [\[CrossRef\]](#)
139. Murad, A.A.; Nuaimi, H.; Hammadi, M. Comprehensive assessment of water resources in the United Arab Emirates (UAE). *Water Resour. Manag.* **2007**, *21*, 1449–1463. [\[CrossRef\]](#)
140. *senseFly eMotion 3 User Manual*; Revision 1.9; senseFly Parrot Group: Cheseaux-sur-Lausanne, Switzerland, 2018.
141. Martins, G.B.; La Rosa, L.E.C.; Happ, P.N.; Filho, L.C.T.C.; Santos, C.J.F.; Feitosa, R.Q.; Ferreira, M.P. Deep learning-based tree species mapping in a highly diverse tropical urban setting. *Urban For. Urban Green.* **2021**, *64*, 127241. [\[CrossRef\]](#)
142. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
143. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. In Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015—Conference Track Proceedings, San Diego, CA, USA, 7–9 May 2015.
144. Jiang, Y.; Liu, W.; Wu, C.; Yao, H. Multi-scale and multi-branch convolutional neural network for retinal image segmentation. *Symmetry* **2021**, *13*, 365. [\[CrossRef\]](#)
145. Chen, L.-C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs. *arXiv* **2014**, arXiv:1412.7062.

146. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 834–848. [\[CrossRef\]](#)
147. Chen, L.C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2018; Volume 11211 LNCS, pp. 833–851.
148. Chollet, F. Xception: Deep Learning with Depthwise Separable Convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017.
149. Zhao, H.; Shi, J.; Qi, X.; Wang, X.; Jia, J. Pyramid Scene Parsing Network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 6230–6239. [\[CrossRef\]](#)
150. Krestenitis, M.; Orfanidis, G.; Ioannidis, K.; Avgerinakis, K.; Vrochidis, S.; Kompatsiaris, I. Oil spill identification from satellite images using deep neural networks. *Remote Sens.* **2019**, *11*, 1762. [\[CrossRef\]](#)
151. Cui, B.; Fei, D.; Shao, G.; Lu, Y.; Chu, J. Extracting raft aquaculture areas from remote sensing images via an improved U-net with a PSE structure. *Remote Sens.* **2019**, *11*, 2053. [\[CrossRef\]](#)
152. Milletari, F.; Navab, N.; Ahmadi, S.A. V-Net: Fully convolutional neural networks for volumetric medical image segmentation. In Proceedings of the 2016 4th International Conference on 3D Vision, Stanford, CA, USA, 25–28 October 2016; pp. 565–571. [\[CrossRef\]](#)
153. Ma, J. Segmentation Loss Odyssey. *arXiv* **2020**, arXiv:2005.13449.
154. Lan, M.; Zhang, Y.; Zhang, L.; Du, B. Global context based automatic road segmentation via dilated convolutional neural network. *Inf. Sci.* **2020**, *535*, 156–171. [\[CrossRef\]](#)
155. Kingma, D.P.; Ba, J.L. Adam: A method for stochastic optimization. In Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015—Conference Track Proceedings, San Diego, CA, USA, 7–9 May 2015; pp. 1–15.
156. Paoletti, M.E.; Haut, J.M.; Plaza, J.; Plaza, A. Deep learning classifiers for hyperspectral imaging: A review. *ISPRS J. Photogramm. Remote Sens.* **2019**, *158*, 279–317. [\[CrossRef\]](#)