Article

Hybrid Deep Learning Models with Sparse Enhancement Technique for Detection of Newly Grown Tree Leaves

Shih-Yu Chen [1,2], Chinsu Lin [3,*], Guan-Jie Li [1,2], Yu-Chun Hsu [1,2] and Keng-Hao Liu [4]

1 Department of Computer Science and Information Engineering, National Yunlin University of Science and Technology, Yunlin 64002, Taiwan; sychen@yuntech.edu.tw (S.-Y.C.); m10717003@gemail.yuntech.edu.tw (G.-J.L.); m10817021@yuntech.edu.tw (Y.-C.H.)
2 Artificial Intelligence Recognition Industry Service Research Center, National Yunlin University of Science and Technology, Yunlin 64002, Taiwan
3 Department of Forestry and Natural Resources, National Chiayi University, Chiayi 600355, Taiwan
4 Department of Mechanical and Electro-mechanical Engineering, National Sun Yat-sen University, Kaohsiung 80424, Taiwan; keng3@mail.nsysu.edu.tw
* Correspondence: chinsu@mail.ncyu.edu.tw

Abstract: The life cycle of leaves, from sprout to senescence, is the phenomenon of regular changes such as budding, branching, leaf spreading, flowering, fruiting, leaf fall, and dormancy due to seasonal climate changes. It is the effect of temperature and moisture in the life cycle on physiological changes, so the detection of newly grown leaves (NGL) is helpful for the estimation of tree growth and even climate change. This study focused on the detection of NGL based on deep learning convolutional neural network (CNN) models with sparse enhancement (SE). As the NGL areas found in forest images have similar sparse characteristics, we used a sparse image to enhance the signal of the NGL. The difference between the NGL and the background could be further improved. We then proposed hybrid CNN models that combined U-net and SegNet features to perform image segmentation. As the NGL in the image were relatively small and tiny targets, in terms of data characteristics, they also belonged to the problem of imbalanced data. Therefore, this paper further proposed 3-Layer SegNet, 3-Layer U-SegNet, 2-Layer U-SegNet, and 2-Layer Conv-U-SegNet architectures to reduce the pooling degree of traditional semantic segmentation models, and used a loss function to increase the weight of the NGL. According to the experimental results, our proposed algorithms were indeed helpful for the image segmentation of NGL and could achieve better kappa results by 0.743.

Keywords: deep learning; semantic segmentation; imbalanced data; robust principal component analysis; remote sensing image

1. Introduction

The key to the sustainable development of forest ecosystem resources is to protect the coverage of wild trees. Annual coverage is tracked continuously to relieve the influence of global warming or climate change. According to the forest resource assessment report of UNFAO (United Nations Food and Agricultural Organization), the change in the area of woods, forest biomass, carbon storage, and the improvement of healthy forests are periodic evaluation indexes used for sustainable global forest resources [1]. Therefore, using telemetry to monitor the health standards of forest ecosystems has an important effect on controlling global warming. Climatic change may influence climate events such as the budding and senescence of leaves [2]. Once trees perceive the signal to begin growing in early spring, treetop leaves begin to sprout, newly grown leaves (NGL) develop gradually, and growth of the crown diameter, height, diameter, and carbon storage is further increased [3]. Therefore, NGL can be regarded as the primary key to the response of trees to temperature change and providing key information for early detection of climate

change [4]. In forestry, telemetry has been used to investigate species classification [5,6], tree delimitation [7,8], and biomass productivity evaluation [9,10]. However, NGL may have a low probability of occurrence in detection, or they may be smaller than the size of the background. For example, NGL in the forest canopy or NGL in an injured crown in the forest canopy may not be detected effectively by traditional spatial domain image processing techniques [11–14]. In recent years, unmanned aerial vehicle (UAV) has become an important tool for crop inspection, environmental surveillance, and the detection of various earth surfaces. UAV helps in monitoring different environmental changes [15–17], including climate change, ecosystem change, urbanization, and habitat variation. Therefore, the goal of this paper is to investigate robust deep learning-based algorithms to diagnose the growth of trees, and even climatic change by detecting newly grown leaves (NGL) from UAV bitmap images.

Recent years have shown rapid development of deep learning techniques [18] in numerous perceptual tasks like object detection and image understanding, such as genes identification [19–23], cancer diagnosis [24,25], dental caries [26], biomedical/medical imaging [27–29], skin lesion [30,31], and crack detection [32,33]. In remote sensing applications, several deep learning frameworks have been presented. For example, in Reference [34], a deep convolutional encoder–decoder model for remote sensing image segmentation was proposed. The model applies the encoder network to extract high-level semantic features of ultra-high-resolution images and uses the decoder network to map the low-resolution encoder feature maps, and creates full input resolution feature maps for pixel-wise labeling. An end-to-end fully convolutional neural network (FCN) [35] was proposed in Reference [36] for the semantic segmentation of remote sensing imaging. An object-based Markov random field (MRF) model with auxiliary label fields was presented in Reference [37] to capture more macro and detailed information for the semantic segmentation of high spatial resolution remote sensing images. Furthermore, symmetrical dense-shortcut deep FCNs for the semantic segmentation of very-high-resolution remote sensing images were proposed in Reference [38]. On the other hand, semantic labeling of 3D urban models over complex scenarios was presented in Reference [39] based on convolutional neural network (CNN) and structure-from-motion techniques. The deep learning method used in this paper belongs to the field of semantic segmentation [35]. The method of semantic segmentation refers to classifying each pixel in an image [40]. The classification results presented by semantic segmentation have more spatial information than traditional classification models and they play an important role in many practical applications. For example, Hyeonwoo Noh et al. [41] discussed image understanding based on semantic segmentation, and Michael Treml et al. [42] investigated automatic driving based on semantic segmentation. In recent years, many related deep learning techniques have been intensively applied to research of leaves such as leaf recognition [43], plant classification [44,45], leaf disease classification, and detection [46–51]. Plant and leaf identification [52–54] using deep learning is a relatively new field. For example, Reference [43] used a basic and Pre-trained GoogleNet for leaf recognition. Reference [52] applied a CNN model to identify leaf veins. Reference [53] proposed a weed segmentation system by combing the segmentation algorithm and CNN models. Reference [54] proposed LeafNet, a CNN-based plant identification system. The LeafNet architecture is also similar to well-known AlexNet and CifarNet. According to the literature review in this section, there is no research that investigates the effect of the segmentation of hybrid deep learning manners for detection of a newly grown leaf (NGL) in detail. We believe our study first attempted to examine the detection of a newly grown leaf (NGL) using deep learning technique in high-resolution UAV images.

The NGL images used in this article are taken by the UAV drone with a pixel resolution of six centimeters; however, NGL can be smaller than six centimeters. From the image point of view, an NGL is smaller than one pixel (known as sub-pixels). If the popular deep learning models (SegNet [40,55–57], LF-SegNet [58], U-Net [59–64], U-SegNet [65], etc.) are used directly, some target information will be lost due to the excessive number of pooling

performed by the original methods. Therefore, we have added the skip-connection method to connect to the network layer in SegNet. Since the proportion of NGL in the overall image is small (which is an imbalanced data problem in machine learning), we have also adjusted the loss function to improve the detection rate of NGL. On the other hand, it would be helpful for target detection in an image to enhance possible target regions in advance using preprocessing techniques, such as robust principal component analysis (RPCA) or related matrix decomposition techniques based on low-rank and sparse properties [66]. RPCA or its related techniques have been successfully used in several applications of remote sensing or image processing. For example, in Reference [67], a joint reconstruction and anomaly detection framework was presented for compressive hyperspectral images based on Mahalanobis distance-regularized tensor RPCA. An edge-preserving rain removal technique was proposed in Reference [68] for light field image data based on RPCA. Moreover, the problem of background subtraction was formulated and solved by relying on a low-rank matrix completion framework in Reference [69]. Furthermore, an efficient rank-revealing decomposition framework based on randomization was presented in Reference [70] for reconstructions of low-rank and sparse matrices. More researches about RPCA or sparse representation can be found in References [71–75]. In this paper, we have proposed a sparse enhancement (SE) technique to strengthen possible targets in an image before applying the learned deep network for the detection of newly grown tree leaves since based on the characteristic of NGL, it can be considered as a sparse matrix in a remote sensing forest image.

In prior studies about NGL, Lin et al. [4] first investigated the hyperspectral target detection algorithms for the possibility of detecting NGL by using spectral information in active and passive manners. However, target detection algorithms were sensitive to the desired target in the final results. After that, Chen et al. [76] proposed the Optimal Signature Generation Process (OSGP) and adaptive window-based Constrained Energy Minimization (CEM) [77], which could provide steadier detection results and reduce the occurrence of false alarms. Another method is based on data preprocessing, in which the target of detection is enhanced by data preprocessing. Chen et al. [78] used preprocessing to enhance the signal of NGL and adopted the Weighted CEM to further reduce the misrecognition rate in the detection results. Differing from prior documents in terms of target detection point of view, this paper started with the segmentation of deep learning application which achieves pixel-level prediction by classifying each pixel according to its category and dividing the image into foreground and background for binary classification and target detection as previous studies. This paper combined hybrid models with U-NET and SegNet, reducing the number of pooling layers to keep the information of NGL and using skip connection to extract feature information from low-level information in the hopes of enhancing NGL features. Moreover, this paper used sparse enhancement preprocessing before the network architecture of deep learning. As the area of NGL to be detected in the forest image had sparse-like characteristics compared to the full image, the NGL in the original image was enhanced by this method. So, the goal of this paper is to investigate the possibility of hybrid deep learning architectures for the detection of NGL from UAV images and the performance of using sparse enhancement (SE) technique as preprocessing in our proposed models.

2. Materials and Methods

In this section, we describe the study site and the dataset in Section 2.1, followed by Sections 2.2 and 2.3, where we describe the current widely-used CNN models and loss function. Finally, in Sections 2.3–2.6, we present the proposed hybrid CNN architecture.

2.1. Description of the Study Site

In 2002, there were 17 species of trees planted in Taiwan and the area of hardwood forest was 188.59 hectares [4]. The 17 species included the tropical species "Swietenia macrophylla" in the south of Taiwan, which experiences defoliation for one to two weeks

during the middle to late March. New leaves grow rapidly within one to two days and are observable in the air. This study used an eBee Real-Time Kinematic (RTK) drone (developed by SenseFly, Switzerland) carrying a Canon PowerShot S110 camera flying at an altitude of 239.2 ft of the area on 12 July 2014 under very good weather condition. The weather is generally hot and humid. The mean annual temperature is 24.38 °C. Leaf development is a process of dynamic plant growth responding to plant physiology and environmental signals [79]. Leaves start to develop from the apical meristems of branches. It is unachievable to visualize a newly sprouted leaflet from the UAV at the initiation stage of the life cycle of a leaf. However, after a period of gradual development, the newly grown leaf is normally at a size of a centimeter and can be visualized from a distance. July is a good time to fly a UAV to detect newly grown leaves in the south of Taiwan. The spatial resolution of the ground sample distance of the images was 6.75 cm (known as a centimeter-level very high resolution (VHR) image). The research site was located in the Baihe District of Tainan City, Taiwan (23°20′ N, 120°27′ E) as shown in Figure 1. This image includes the red, green, and blue bands, with a dimension of 1000 × 1300 pixels and the actual area of the full image is 60 m × 78 m.
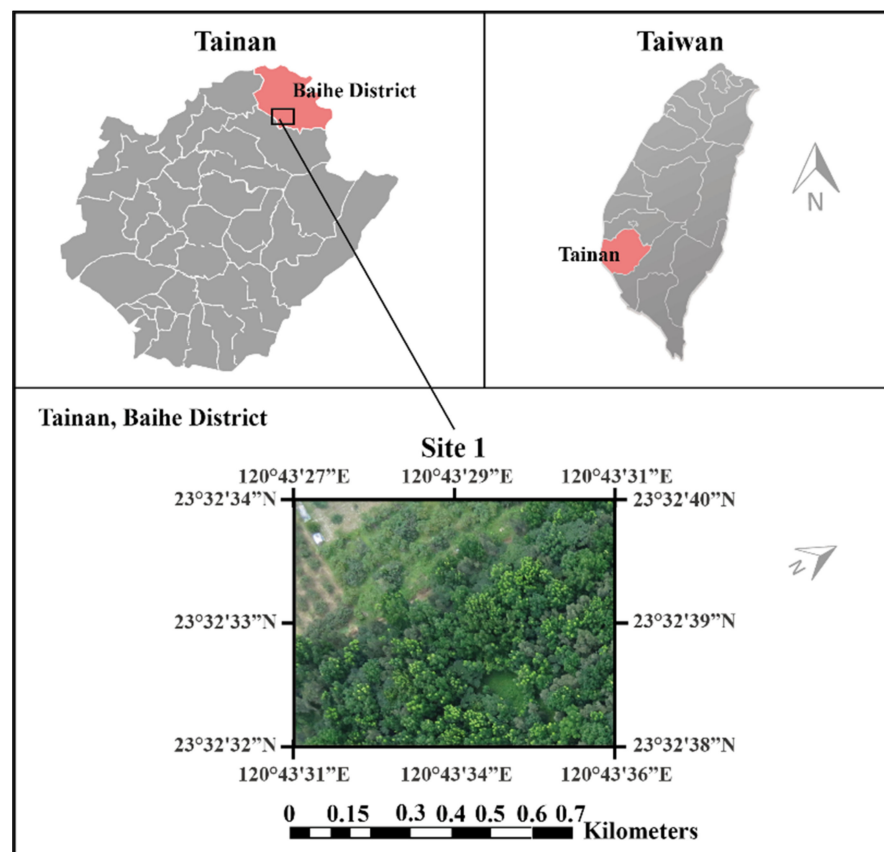

Figure 1. Study site.

Ground Truth

In 2008, a few permanent plots were deployed over the broadleaf forest for research of forest growth [80,81], where a series of ground inventory is annually conducted for stand dynamics [17]. The data have been successfully derived using the forest canopy height model [82] and have investigated the feasibility of global/local target detection algorithms for the detection of NGL by our prior studies [4,76,78]. As shown in Figure 2a,b, the NGLs can be visually inspected by experts due to their appearance. They are bright, light green, and amassed over the tree crowns. According to a row of several years of inventory, the ground truth of the NGL over the images was visually interpreted by two professors from

the Department of Forestry at National Chiayi University, Taiwan, and was also validated in situ. To quantify and compare the effects of different algorithms, there must be an NGL detection map as the ground truth of Area 1 and Area 2, as shown in Figure 2c,d. The NGL is only about 3–4% of the entire images.

(**a**) Area 1   (**b**) Area 2

(**c**) Area 1 ground-truth   (**d**) Area 2 ground-truth
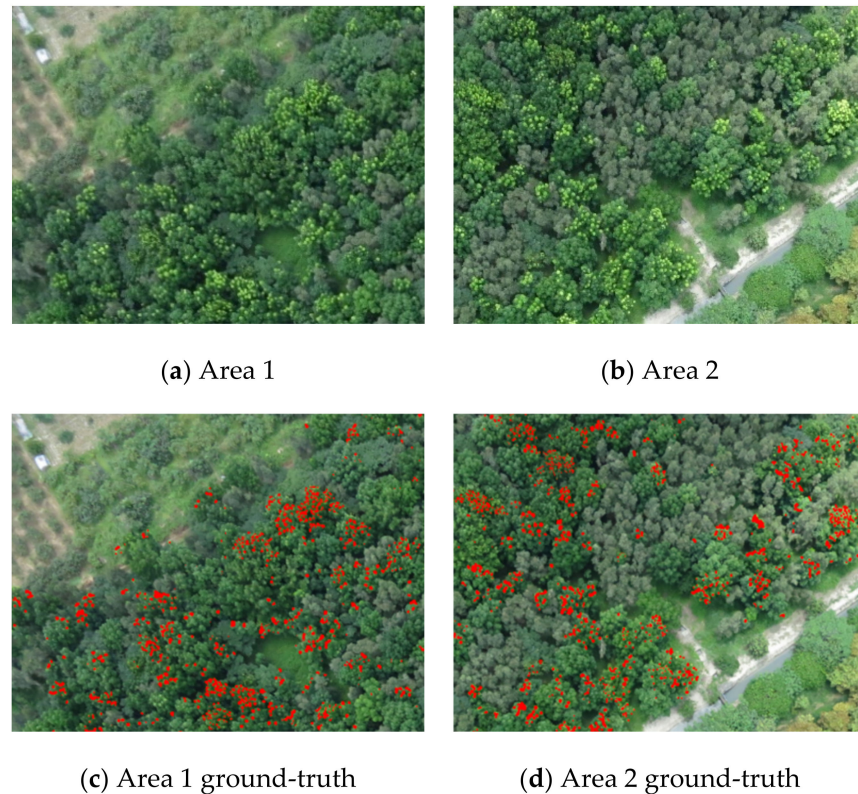
Figure 2. The original images and the corresponding Ground-truths of Area 1 and Area 2.

2.2. Deep Learning Models

2.2.1. SegNet

The SegNet network architecture was proposed by the research team of Vijay and Alex at the University of Cambridge [55]. The architecture is divided into a convolutional encoding layer, convolutional layers corresponding to the max-pooling layer, and an up-sampling layer to replace the original max-pooling layer. The reduced feature map is restored to the size of an input image. Finally, the softmax classification layer is applied. The architecture is shown in Figure 3 (there are 26 convolutional layers, five max-pooling layers, and five up-sampling layers). For an image imported into SegNet (if the image size is W × H), the feature map is obtained through the convolutional layers and standardized operation, and the feature map is reduced by the max-pooling layer to a certain scale for the upsampling operation. The feature map is deconvolved to the size of the original input image, therefore, numerous feature maps that are as large as the original image can be obtained. The result is imported into the softmax classifier to analyze the class of each pixel. This paper performed modification using SegNet as the fundamental model, and the pooling frequency of SegNet was reduced to maintain the information of tiny targets.
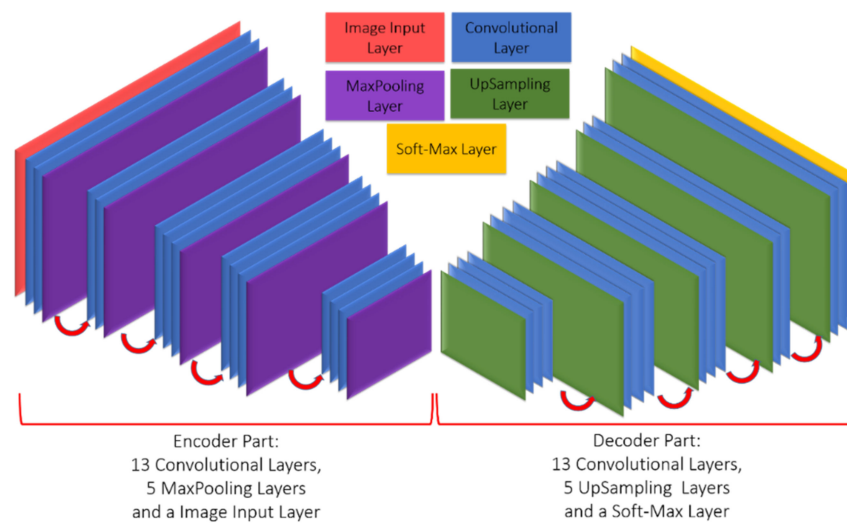
Figure 3. SegNet architecture.

2.2.2. U-Net

U-Net [60] is a segmentation network proposed by Olaf Ronneberger et al. at the ISBI(International Symposium on Biomedical Imaging) Challenge in 2015 (they won the championship). In medical images, Zongwei Zhou et al. [83] proposed the deep supervision model extended from U-Net for training. Moreover, Ozan Oktay et al. [84] used a new AG model in U-Net, in which the sensitivity and precision of prediction were enhanced. The U-shaped network structure is shown in Figure 4. The overall network structure is similar to SegNet as it uses the skip connection technique to compensate for the spatial information lost after up-sampling of the bottom layer.
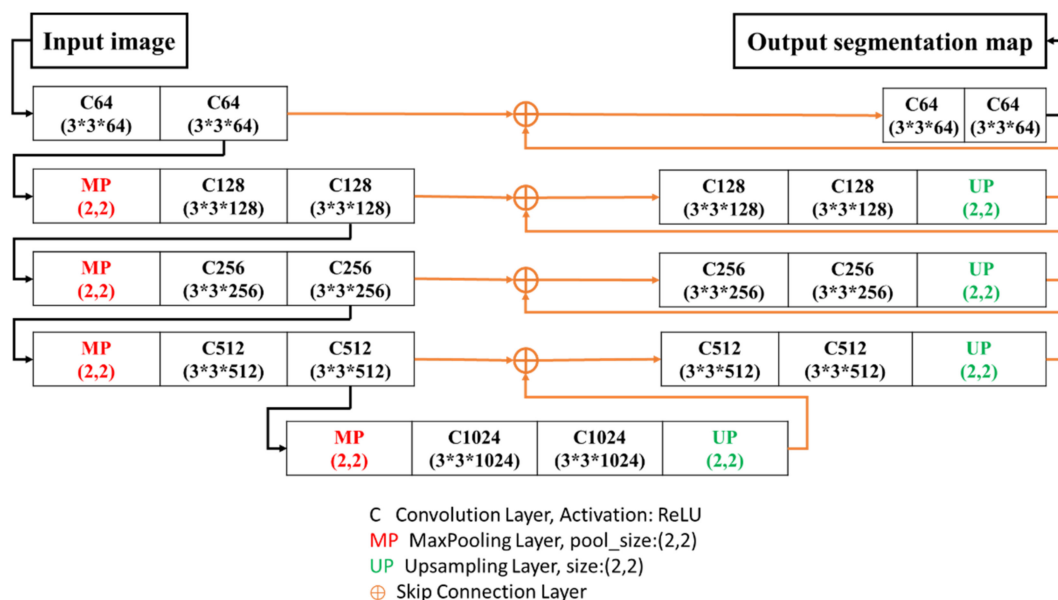
Figure 4. U-Net architecture.

2.3. Loss Function of the Convolutional Neural Network

In machine learning, the loss function is used to estimate the extent of the error in a model. It is calculated according to the prediction value and the true value. A smaller error indicates the model has a certain recognition capability. The NGL detection classification method in this paper could be regarded as a dichotomy, as the loss function is mostly

classified by Binary cross-entropy [85]. However, the NGL accounts for only a small part of the overall data, meaning it is an imbalanced data problem in machine learning. To enable the model to detect as many NGL as possible, we have used the following loss function method with adjustable data weights.

2.3.1. Binary Cross-Entropy (Binary CE or CE)

In deep learning, the common loss function for a binary problem is CE [85], which is expressed as follows:

$$\text{Binary CE}(p, \hat{p}) = -(p\log(\hat{p}) + (1-p)\log(1-\hat{p})) \tag{1}$$

where p is the ground-truth and $\hat{p}$ is the result predicted for the NGL category. The log loss of all samples represents the average of each sample loss, while the loss for a perfect classification model is 0. However, for the two categories of equal importance, if the sample quantity is extremely unbalanced, it is likely to be ignored when calculating the loss for a small number of samples.

2.3.2. Weighted Cross-Entropy (WCE)

For the problem of an uneven sample number, Aurelio et al. [86] proposed a new method, known as weighted cross-entropy (WCE). WCE imports hyperparameter β to increase the weight of the NGL sample, and the range of β is (1,2). The equation is expressed as follows:

$$\text{WCE}(p, \hat{p}) = -(\beta\, p\log(\hat{p}) + (1-p)\log(1-\hat{p})) \tag{2}$$

2.3.3. Balanced Cross-Entropy (BCE)

The loss function used by Shiwen Pan et al. [87] and us is the BCE method. It is different from WCE because the weight of the NGL sample in calculating the loss is increased and the weight of non-NGL samples is suppressed, as expressed in Equation (3):

$$\text{BCE}(p, \hat{p}) = -(\beta p\log(\hat{p}) + (1-\beta)(1-p)\log(1-\hat{p})) \tag{3}$$

2.4. Hybrid Convolutional Neural Networks

The classical SegNet model was used as the main skeleton of the model. The original SegNet model has more pooling times. The low-level features of tiny targets will blur or even disappear. Therefore, the number of pooling layers was reduced at first. To increase the detection rate, the model could be downsized on the other hand and the loss of tiny targets could be reduced while maintaining the accuracy. When the skip connection (SC) of U-Net was used, the spatial information of the same level was connected upward in the up-sampling process on the bottom layer, and then convolution of the up-sampling was performed. Finally, the batch normalization was added to each convolutional layer [88] to guarantee data stability.

This paper proposed four models based on SegNet and the Skip connection method. Table 1 describes the codes used in the network architecture diagram, wherein C (the convolution layer) uses the unified parameter, Kernel size: (3,3) the 3 by 3 kernel map is used for convolution, padding: same a layer of 0 is added to the outer ring of diagram before convolution, Activation: ReLU let the data be larger than 0, C64 and C128 represent 64 and 128 masks, the kernel size of the outer layer is (1,1). Moreover, the activation used the sigmoid function, in which the data were set as 0~1 for handling binary problems. The complete architecture is introduced in the next section.

Table 1. Model code description.

Code	Code Description	Parameter Information
C	Convolution Layer	Kernel Size: (3, 3) Padding: same Activation: ReLU Filter Size -> C64, C128, C256
BN	BatchNormalization Layer	
MP	MaxPooling Layer	
UP	UpSampling Layer	
Out	Convolution Layer	Kernel Size: (1, 1) Padding: same Activation: sigmoid Filter Size: 1
⊕	Skip Connection Layer	

2.4.1. 3-Layer SegNet (3L-SN)

Firstly, to maintain the information of tiny targets, the pooling depth in the first extended model which is based on the SegNet model, is reduced from five layers to three layers. Thus, the information of tiny targets would not be lost for pooling, and the number of weights to be trained for the whole network could be reduced. Figure 5 shows the 3-Layer SegNet architecture after simplifying SegNet, in which only three layers of convolution are maintained. Batch normalization was performed after each convolution to ensure the data would not encounter the gradient problem and to prevent the next layer of activation function from being out of action or saturated.
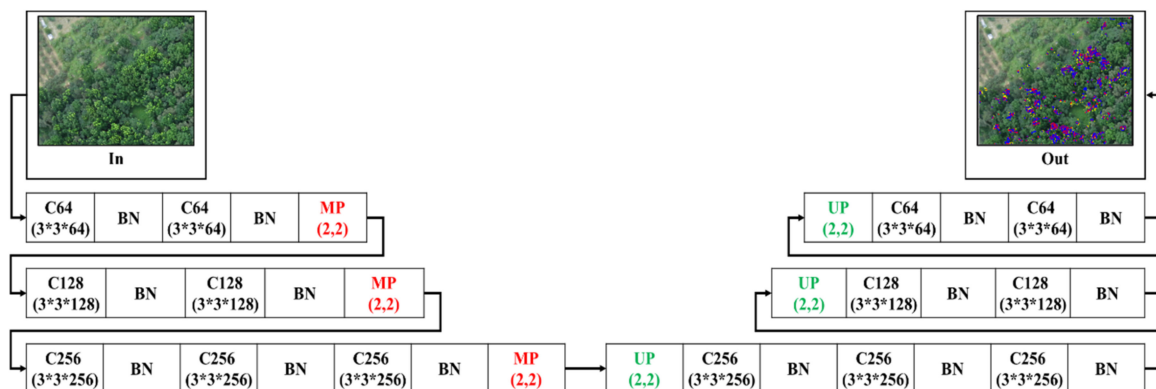
Figure 5. 3-Layer SegNet architecture.

2.4.2. 3-Layer U-SegNet (3L-USN)

Afterward, based on the 3-Layer SegNet network model, the skip connection (SC) of U-Net was imported. This concept was proposed by P. Kumar et al. [65] and Daimary D. et al. [89]. This paper differs in that the skip connection was applied after SegNet pooling depth was reduced by two layers. The purpose of this architecture was to enhance the spatial information of the sample in the course of up-sampling, to enhance the detection capability. Figure 6 shows the 3-Layer skip-connection SegNet network architecture.
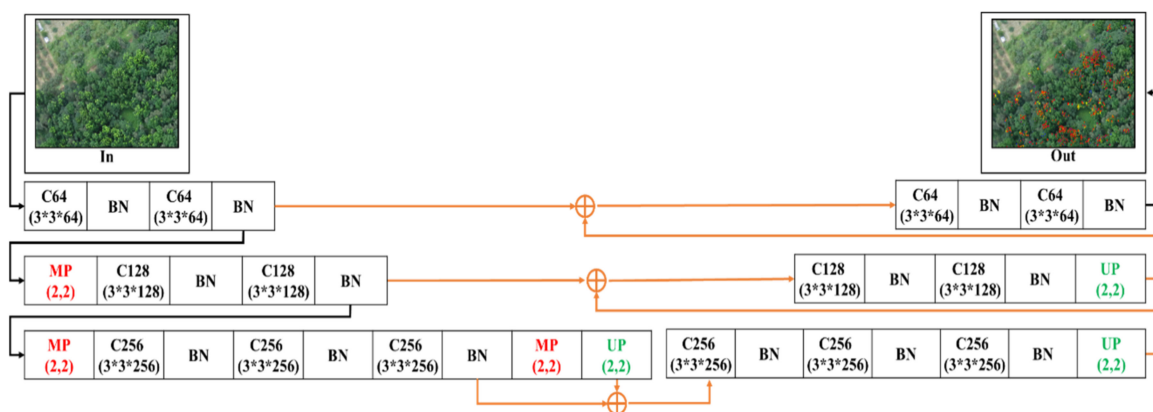
Figure 6. 3-Layer skip connection (SC) SegNet architecture.

2.4.3. 2-Layer U-SegNet (2L-U-SN)

Figure 7 shows that one more pooling layer was reduced while maintaining the skip connection following the network model in the previous section, and the number of convolutions of the convolutional layer was regulated. The purpose was to validate whether there was still a certain detection capability for NGL under a two-layer pooling degree.

Figure 7. 2-Layer SC SegNet architecture.
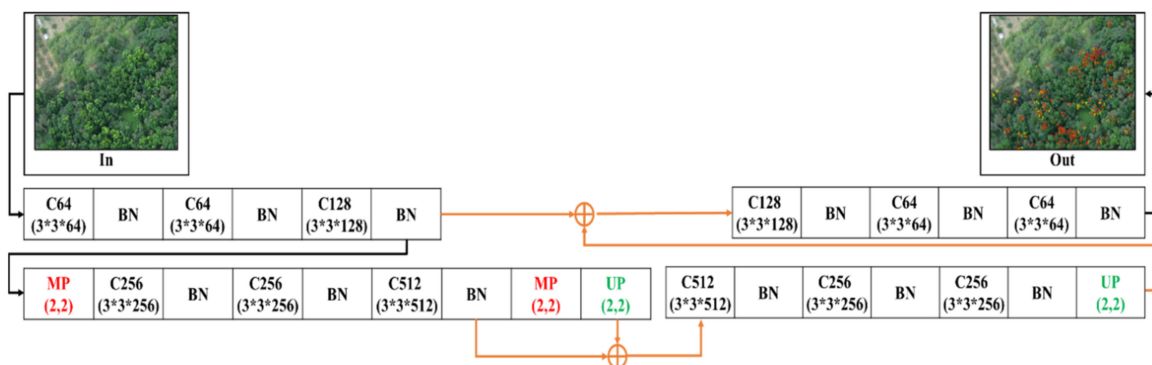
2.4.4. 2-Layer Conv-U-SegNet (2L-Conv-USN)

Figure 8 shows the bottom layer combined with a set of a convolutional layer and BN layer in the encoding stage based on the 2L-U-SN network model, which was used to extract feature information from low-level information in the hopes of enhancing the detection efficiency.
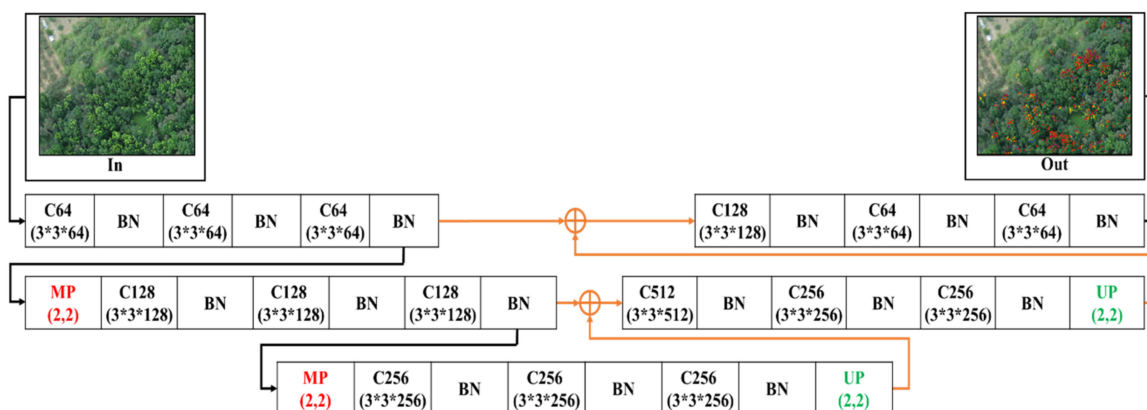

Figure 8. 2-Layer Conv-U-SegNet architecture.

2.5. Robust Principal Component Analysis (RPCA) and Sparse Enhancement (SE)

To enhance the efficiency of the deep learning network for NGL detection, this study proposed a processing method based on Robust Principal Component Analysis (RPCA) [67–69,90–92], and to perform image enhancement for the probable NGL areas in the forest image in advance. The SegNet deep learning network was trained and tested by an enhanced image. RPCA could separate the sparse matrix and low-rank matrix from the input image (regarded as a matrix). RPCA is usually used in image recognition and dimension reduction based on high dimensional data, and it performs linear conversion of the original data by calculating the maximum feature vector. The general data matrix contains the main structural information and probable slight noise information. RPCA decomposes the input matrix into two matrices which are added up; one is a low-rank matrix (including the main structural information) in which the columns or rows are linearly correlated. The other is a sparse matrix (including noise information), expressed as Equation (4):

$$\min_{L,S} \text{rank}(L) + \gamma ||S||_0, \text{ subject to } P = L + S, \tag{4}$$

where P is the original data (original input image), L is the low-rank matrix, rank (L) is used for calculating the rank of matrix L, S is the sparse matrix, $||S||_0$ is the l_0-norm of S (number of nonzero coefficients in S), and γ is the regulation parameter. As the operation of l_0-norm is relatively complex, Equation (4) has been proved convertible to Equation (5):

$$\min_{L,S} \text{rank}(L) + \lambda ||S||_1, \text{ subject to } P = L + S, \tag{5}$$

wherein λ is the regularization parameter and $||S||_1$ is the l_1-norm of S (the sum of the absolute values of the coefficients in S). The input image P can be decomposed into low-rank matrix L and sparse matrix S by optimizing Equation (2), wherein L is the principal component of the image and S is the image noise with sparsity. In this study as the NGL area to be detected in the forest image had sparsity-like characteristics compared with the full image, preprocessing of the RPCA decomposition was performed for the forest image (P) to be processed, in which P = L + S. The position of the nonzero element (or one with a relatively larger absolute value) in S (sparse matrix) is extremely likely and corresponds to the location of the NGL in the original forest image. To enhance the NGL area in the original forest image, we proposed adding sparse matrix S to the original image P to obtain the enhanced image E = P + S. This step is called sparse enhancement (SE), as shown in Figure 9. This method was proposed in Reference [78] and was combined with the network architecture of deep learning for the first time in this paper. The original image was enhanced by sparse enhancement. The fundamental purpose of SE was to enhance the probable NGL area in the original forest image P to obtain the enhanced image E = P + S. The deep learning network was trained based on the enhanced image (Output P + S in Figure 9) and the corresponding ground truth.

According to the results of Reference [78], the enhancement effect of linear combination maintains the original spatial information, and it is enough to enhance the portion of NGL. Therefore, this paper uses the P+S image of the original image plus the sparse image for detection. As RPCA has different statistical methods, and according to the characteristics of different robust statistics, the calculated effects will also be different. We have used RPCA with different kernels such as the GA kernel (Grassmann Averages) [93], OPRMF kernel (Online Probabilistic Approach to Robust Matrix) [94], flip-SPCP-max-QN kernel (stable principal component pursuit) [95], and GoDec (Go Decomposition) [96], to find kernels with characteristics matching the NGL characteristics to enhance the accuracy of NGL detection.
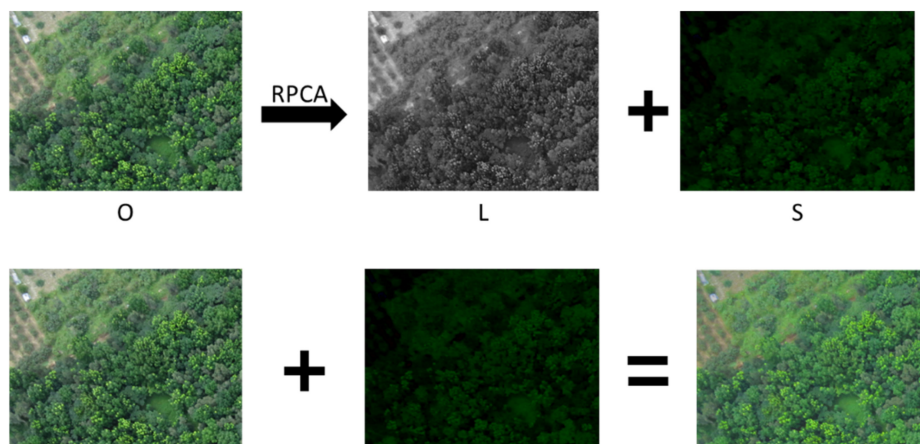
Figure 9. Sparse enhancement flowchart.

2.6. Hybrid Convolutional Neural Network with Sparse Enhancement

The NGL image was processed by SE, and the sparse image in the original image was enhanced to highlight the NGL part in the image. Afterward, the enhanced NGL image was split into training and test datasets. As this paper had two images (Area 1 and Area 2), when Area 1 was used as training data, Area 2 would be used as test data. When Area 2 was used as training data, Area 1 would be used as test data.

2.6.1. Training Data Set

The training images were generated by sliding windows, and the window size was 260×200. To increase the training data volume, the original image was segmented and slid rightwards for only 65 pixels (red window moving rightwards to the cyan window in Figure 10). When it slid to the edge, the window would move down 50 pixels and then slide to the left edge. These actions were repeated until the complete image is captured. There were 289 images split from the image as training data, and each image was 260×200. Figure 11 shows the training data generation of Area 1 and Area 2.

Figure 10. Training data process.
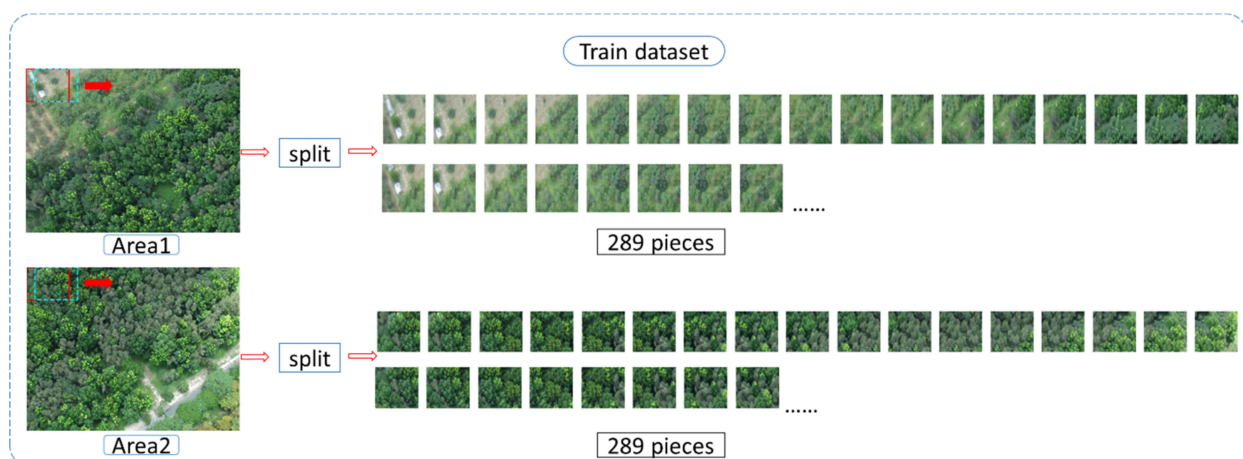
2.6.2. Testing Data Set

The sliding window split method is also used for the testing data, and the window size was 200×260. Differing from the training data which slid rightwards only for 65 pixels, the test data slid rightwards 260 pixels at a time to avoid image overlap during the test (the red window moving to the cyan window in Figure 11). When the window slid to the edge,

it would move down 200 pixels and then return to the left edge. The aforesaid action was repeated until the complete image was captured. There were 25 images split from the test data, and each image was 260×200. Figure 11 shows the test data splitting flowchart of Area 1 and Area 2.

Figure 11. Data testing process.

After the training data and test data were split, model training was performed, as shown in Figure 12. First of all, the training data and ground-truth were trained using the model. The trained model predicted the test data and the performance was evaluated with the ground-truth of the testing data.
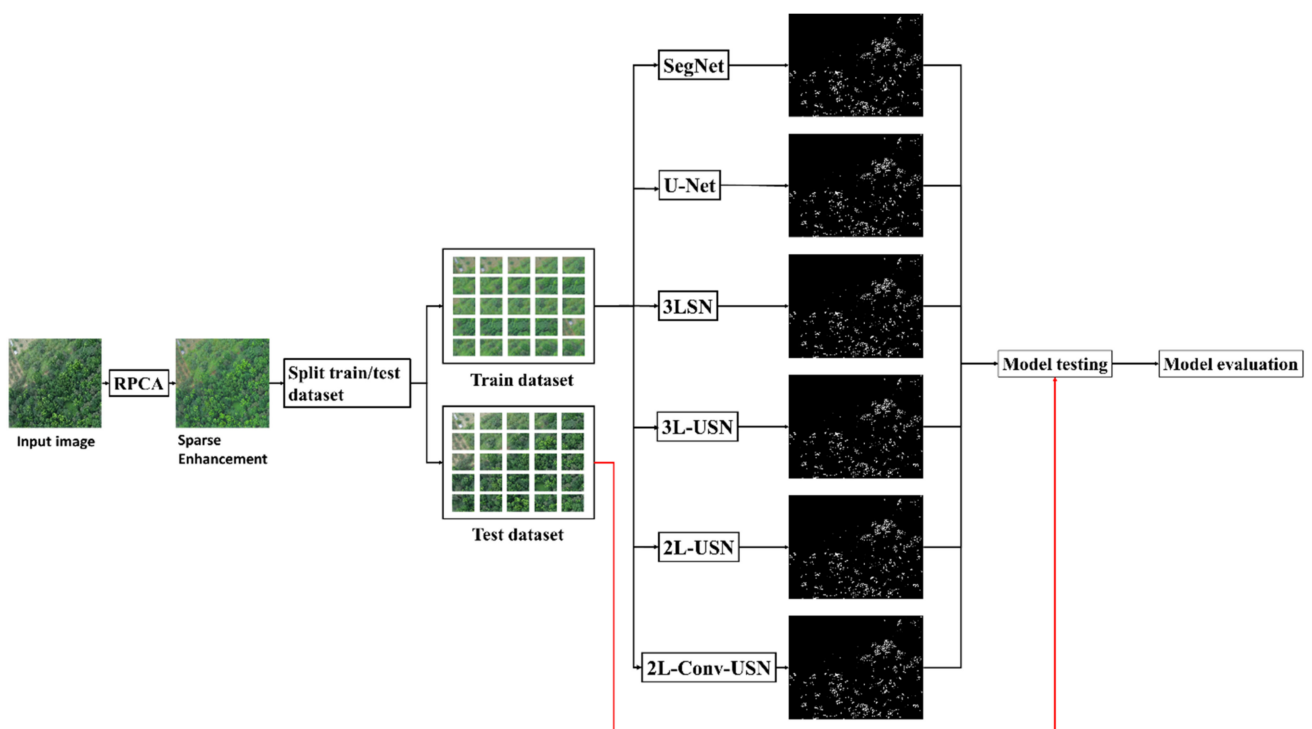
Figure 12. Flowchart of the hybrid convolutional neural network with the sparse enhancement technique.

2.7. Evaluation of Detection Results

Two methods for evaluating accuracy were used in this paper. The first one was receiver operating characteristic (ROC) [97] analysis, which has recently received considerable interest in a wide range of applications in signal processing and medical diagnosis. To evaluate the detection performance of a ROC analysis, which makes use of a curve plotted as a function of detection probability, PD or True Positive, TP versus false alarm probability, PF or False Positive, FP are commonly used to assess the effectiveness of a detector. As an alternative to the use of ROC curves, the area under the curve (AUC), which has been widely used in medical diagnosis, was also calculated by the area under

the ROC curve. Another method Cohen's kappa coefficient (Kappa) [98] can measure the consistency between two classes. In image processing, the ROC curve is used to measure the effectiveness of the detector. Cohen's kappa is an algorithm that evaluates the results of binarization and calculates consistency.

2.7.1. ROC Curve

Receiver operating characteristic (ROC) analysis has been widely used in signal processing and communications to assess the efficiency of a detector. Figure 13 shows schematic diagram of ROC curve. The main concept of ROC analysis is a binary classification model, which has two classes of output results, e.g., correct/incorrect, match/mismatch, target/non-target, etc. ROC analysis which makes use of a curve to plot detection power (PD), or true positive (TP) versus false alarm probability (PF), or false positive (PF) is a commonly used evaluation tool to assess the effectiveness of a detector. As an alternative to the use of ROC curves, the area under the curve (AUC) is another index for evaluating the performance of detectors.
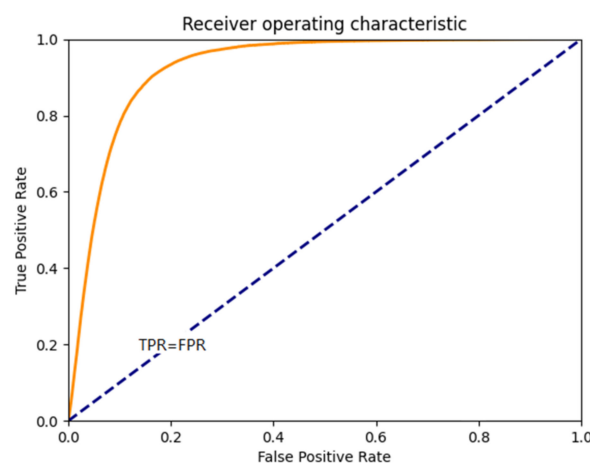
Figure 13. Receiver operating characteristic (ROC) curve.

True positive (TP) and false positive (FPR) are used as the evaluation values of target detection accuracy. The accuracy is defined as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{P} + \text{N}} \tag{6}$$

2.7.2. Cohen's Kappa

Cohen's kappa coefficient is a statistical evaluation method for measuring consistency between two classes. Cohen's kappa evaluates the results of binarization and calculates consistency. It is calculated by an error matrix identical to ROC curve, as shown in Table 2.

Table 2. Error matrix of Cohen's kappa.

Error Matrix		Ground Truth		Total
		NGL (p)	Non-NGL (n)	
Detection	NGL (p′)	True Positive P_a	False Positive P_b	$P_a + P_b$
	Non-NGL (n′)	False Negative P_c	True Negative P_d	$P_c + P_d$
Total		$P_a + P_c.$	$P_b + P_d$	Total Pixels N

Where P_a is the true value of NGL and what is detected is NGL, P_b is the true value of NGL but what is detected is not NGL, P_c is not the true value of NGL but what is detected is NGL, P_d is not the true value of NGL and what is detected is not NGL. Cohen's kappa can be defined in the following:

$$P_o = \frac{P_a + P_d}{P_a + P_b + P_c + P_d} = \frac{P_a + P_d}{N} \tag{7}$$

$$P_e = P_{Yes} + P_{No} = \frac{P_a + P_b}{N} \cdot \frac{P_a + P_c}{N} + \frac{P_c + P_d}{N} \cdot \frac{P_b + P_d}{N} \tag{8}$$

$$K = \frac{P_o - P_e}{1 - P_e} = 1 - \frac{1 - P_o}{1 - P_e} \tag{9}$$

Coefficient K, as calculated by Cohen's kappa is $-1\sim1$; if the detection results are completely correct, the K value is equal to 1; if K is equal to 0, the detection result is all NGLs or all non-NGL.

3. Results

This section is divided into three stages to compare the related experimental results. Section 3.1 performs a comparative analysis of the proposed parameters of BCE after the parameters are selected. The SE generated by RPCA with different kernels is tested in Section 3.2. Finally, Section 3.3 performs a comparative analysis of all the proposed CNNs.

3.1. BCE Parameter

This section discusses the data difference in the extended model proposed in this paper when the hyperparameter β of Loss function BCE is different, and the performance of each model is calculated and evaluated. This paper used ROC, ACC, and Kappa as the criteria of the evaluation score, because the effects of hyperparameter β on AUC and Kappa were compared. Figures 14 and 15 only show the data of the flip-SPCP-max-QN kernel corresponding to different hyperparameter of β. According to the data drawing list of Area 1, the models had better performance when β = 0.999, and various models had smaller differences when β = 0.995, but the Kappa is lower. The performance of β = 0.99 in AUC was only a few percentage points different from β = 0.999, but the Kappa had a good performance in the two areas; therefore, β = 0.99 was selected as the parameter in subsequent experiments.
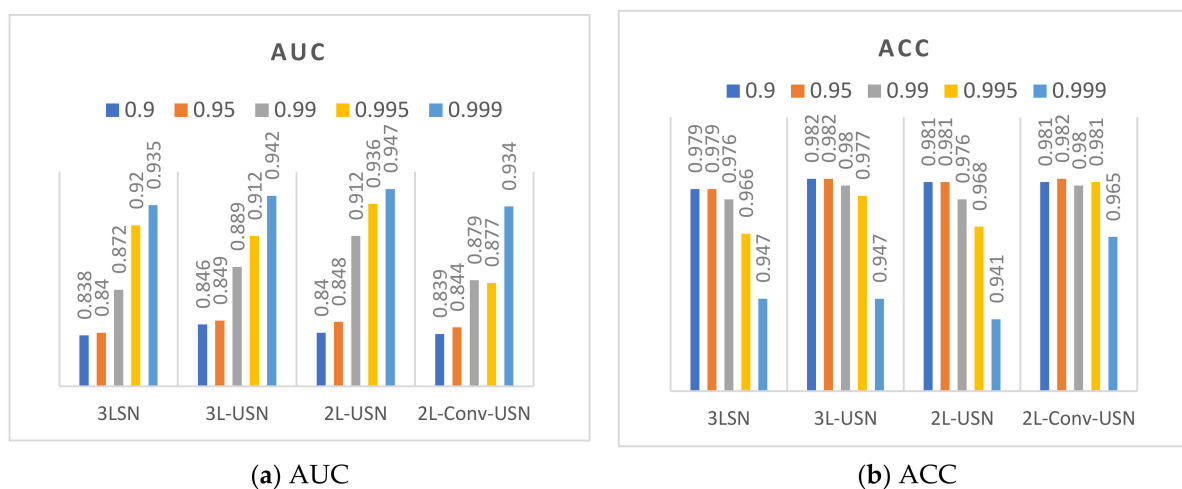
(**a**) AUC

(**b**) ACC

Figure 14. *Cont.*
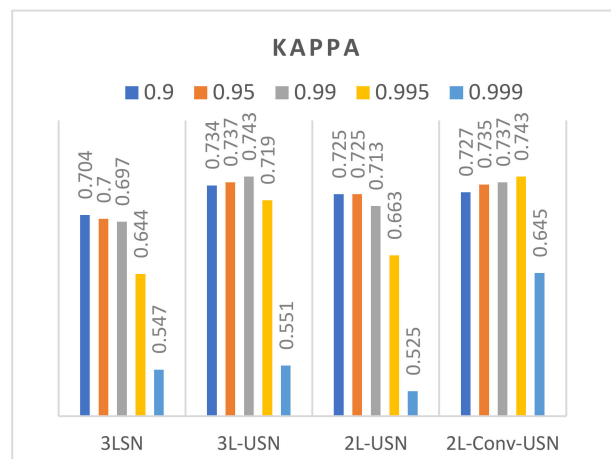
(**c**) Kappa

Figure 14. Data bar chart of Area 1 using five hyperparameters of β and four models.
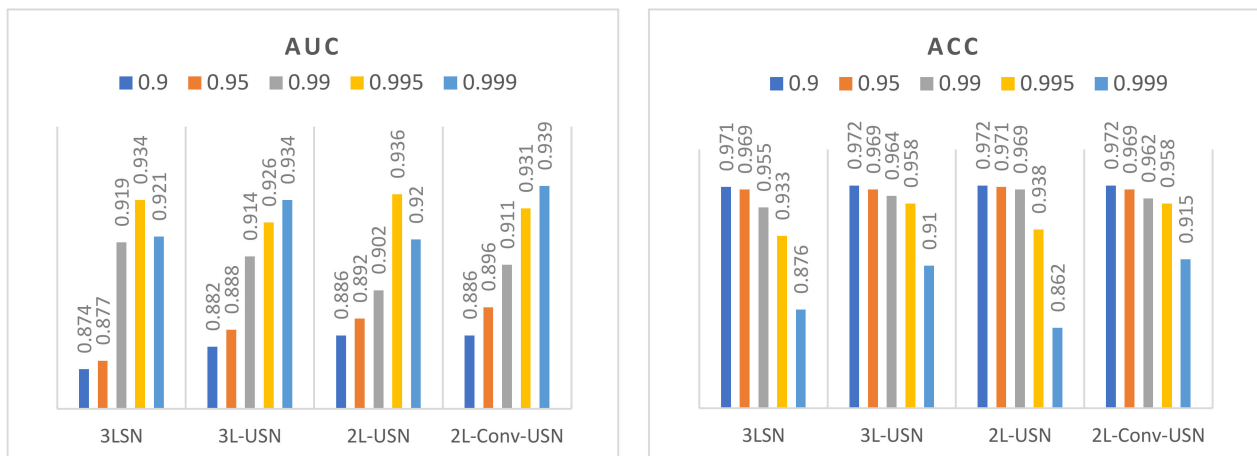
(**a**) AUC

(**b**) ACC
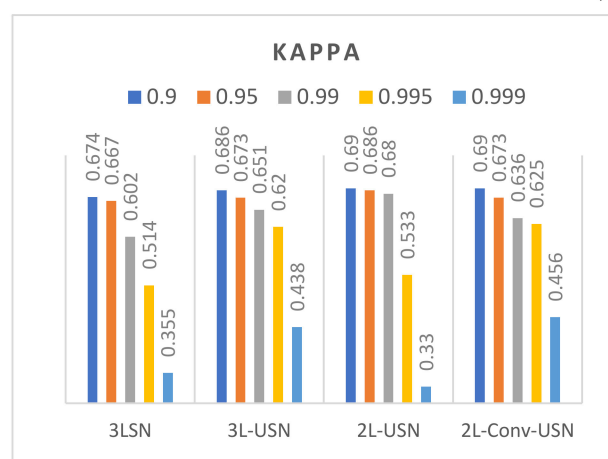
(**c**) Kappa

Figure 15. Data bar chart of Area 2 using five hyperparameters of β and four models.

3.2. Results of Different Kernels in SE

This section combines the SE preprocessed by RPCA with different kernels with the SegNet and U-Net of the BCE lost function and the proposed models used in the previous sections using Area 1 and Area 2. To review the differences among different kernels after

preprocessing, the data in the previous section were sequenced according to AUC, true positive rate (TPR), FPR, and Kappa to compile a line chart, as shown in Figures 16 and 17. According to the data drawing list of Area 1, the flip-SPCP-max-QN had better performance in detection (TPR) but had a larger increase in the amplitude of the misrecognition rate (FPR); therefore, its Kappa value was influenced greatly. With the GoDec method, the false alarm rate is relatively low, but the detection rate was also relatively low. To sum up, OPRMF had a relatively average performance in detection. In terms of Area 2, the GoDec kernel could not inhibit the background well, and the other three kernels had similar detection effects.
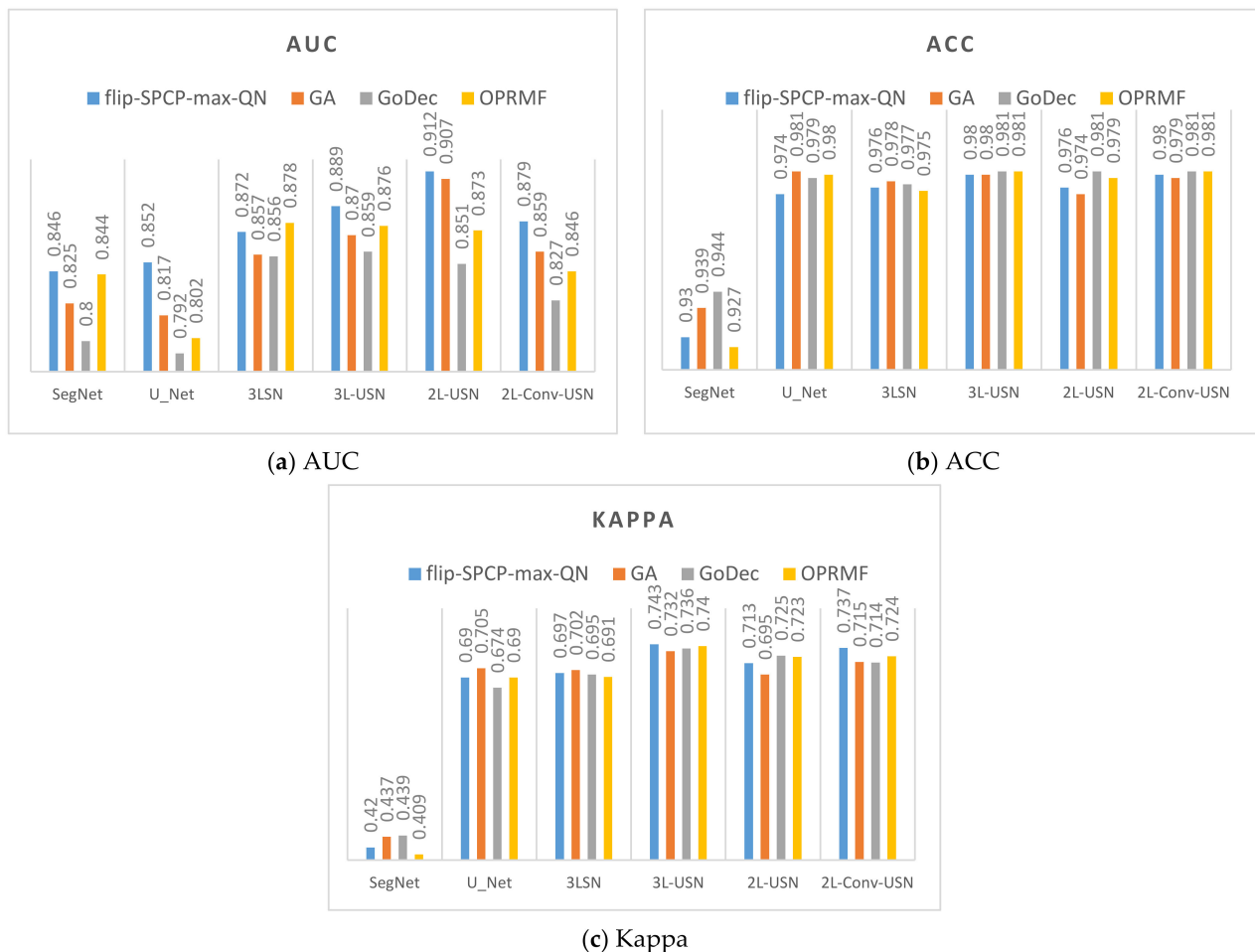

(**a**) AUC

(**b**) ACC

(**c**) Kappa

Figure 16. Data bar chart of the area under the curve (AUC) and Kappa of Area 1 using four kernels + four models.

3.3. Results of Hybrid CNN Models

This section shows the comprehensive comparative results of various deep learning models using Area 1 and Area 2. This paper used the ROC, ACC, and Kappa as the criteria of the evaluation score. Table 3 shows the detection results of the original SegNet, U-Net, and our proposed model of Area 1 without SE. AUC, TPR, FPR, ACC, and Kappa were used for various evaluations. To compare with prior studies, in References [4,76,78], the global and local target detection results including Adaptive Coherence Estimator (ACE), Target Constrained Interference Minimized Filter (TCIMF), Constrained Energy Minimization (CEM), Subset CEM, Sliding Window-CEM (SW CEM), Adaptive Sliding Window-CEM (ASW CEM), and Weighted Background Suppression (WBS) version of above detectors were used for comparison. The results of target detection methods had better performance in AUC and TPR, but our proposed deep learning models especially in 3L-USN had the lowest FPR, and the highest ACC and Kappa were used up to 0.981 and 0.741, respectively.

(**a**) AUC

(**b**) ACC
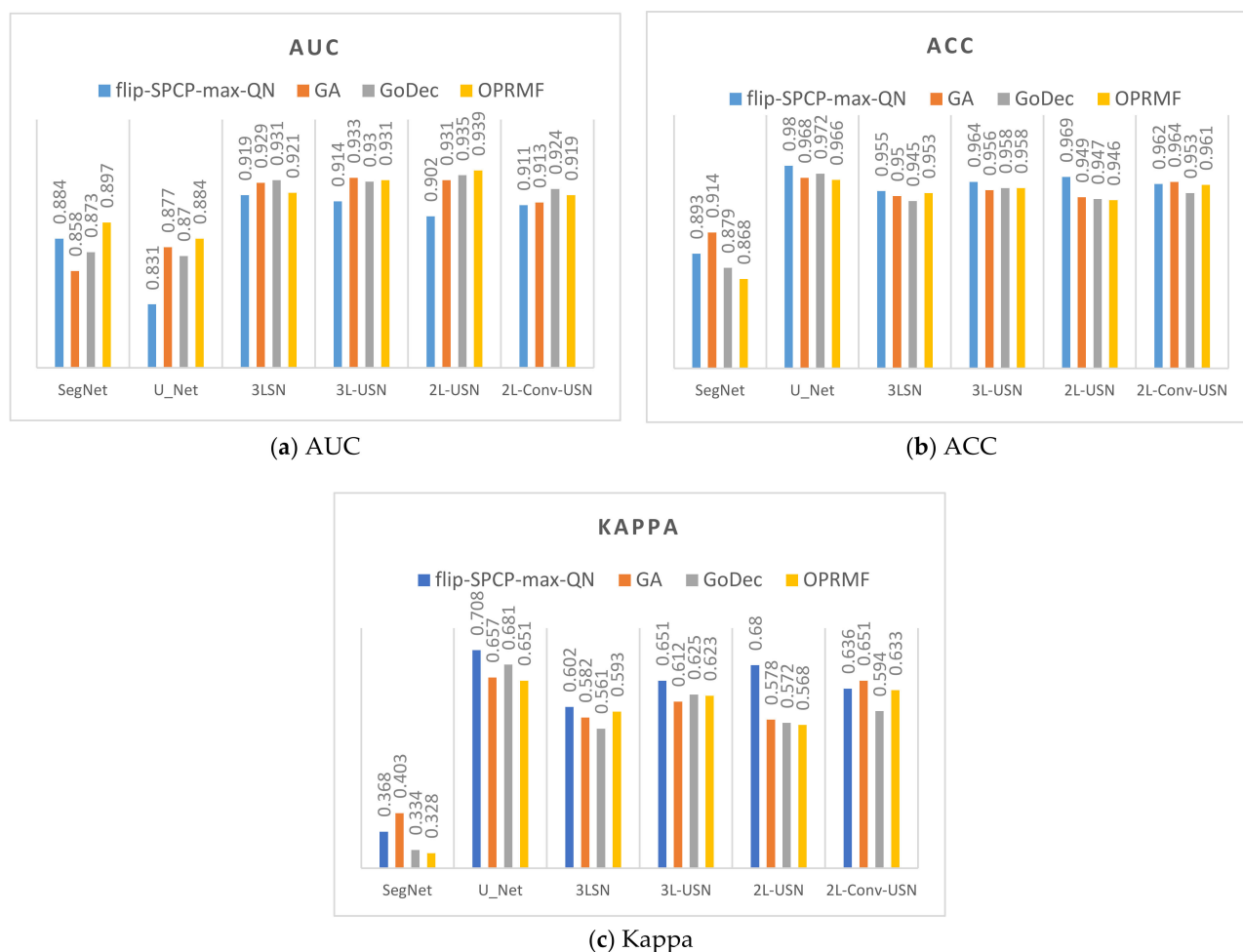
(**c**) Kappa

Figure 17. Data bar chart of the AUC and Kappa of Area 2 using four kernels + four models.

Table 3. Detection comparison among global, local target detection, SegNet, U-Net, and our proposed deep learning models of Area 1 without sparse enhancement (SE).

	AUC	PD/TPR	PF/FPR	ACC	Kappa
Global Target Detection					
Traditional ACE	0.9417	0.8827	0.1099	0.8897	0.3389
WBS-ACE (ED)	0.9375	0.8697	0.1113	0.888	0.3313
Traditional TCIMF	0.95737	0.8903	0.1076	0.8924	0.3472
WBS-TCIMF(ED)	0.9583	0.8906	0.1082	0.8917	0.3458
Traditional CEM	0.9606	0.8915	0.1022	0.8976	0.3604
WBS-CEM(ED)	0.9713	0.9029	0.0738	0.9254	0.4483
Sparse WCEM	0.971	0.9067	0.08	0.9194	0.4288
Local Target Detection					
	AUC	PD/TPR	PF/FPR	ACC	Kappa
Subset CEM	0.9685	0.9099	0.0874	0.9125	0.4074
Subset WBS-CEM	0.9725	0.9177	0.0827	0.9173	0.4248
SW CEM	0.9694	0.9151	0.0811	0.9188	0.4289
SW WBS-CEM	0.9749	0.9204	0.0729	0.9269	0.4588
ASW CEM	0.9704	0.9257	0.0753	0.9248	0.4526
ASW WBS-CEM	0.9781	0.9325	0.0673	0.9327	0.4847

Table 3. *Cont.*

Deep Learning Based (Proposed)					
Model	**AUC**	**TPR**	**FPR**	**ACC**	**Kappa**
SegNet	0.79	0.621	0.042	0.946	0.438
U-Net	0.886	0.799	0.026	0.967	0.653
3LSN	0.859	0.73	0.012	0.978	0.706
3L-USN	0.773	0.757	0.01	0.981	0.741
2L-USN	0.856	0.821	0.009	0.98	0.727
2L-Conv-USN	0.853	0.715	0.009	0.981	0.73

The red/green values indicate the best/worst performance in each evaluation metric.

Table 4 shows the detection results of all methods with SE. The SE shown in Table 4 used the flip-SPCP-max-QN kernel with its higher accuracy, but the loss function only used the BCE method ($\beta = 0.99$). Again, the results of target detection methods had better performance in AUC and TPR, but our 3L-USN had the lowest FPR, and the highest ACC and Kappa were used up to 0.98 and 0.743, respectively.
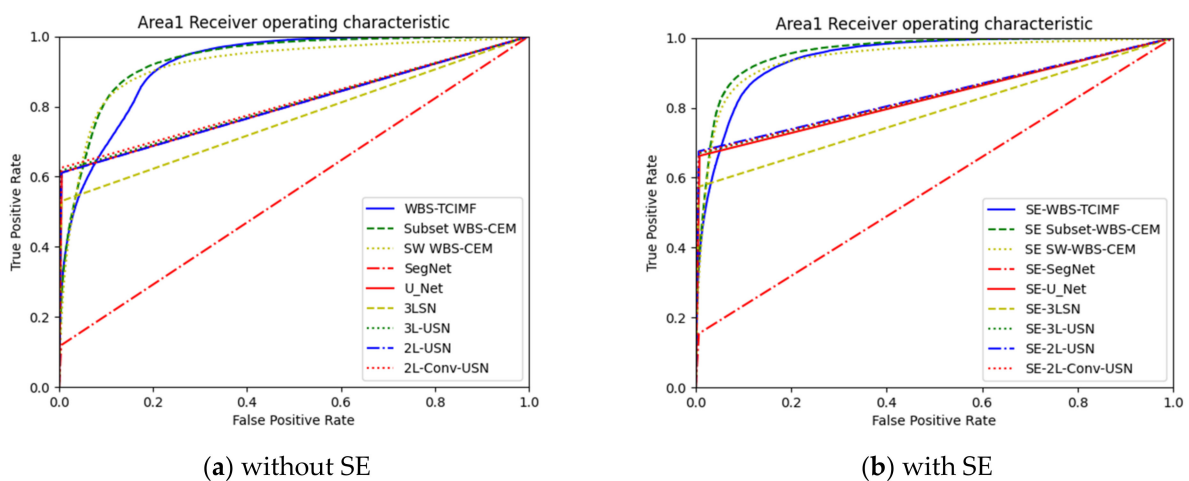
Table 4. Detection comparison among the global and local target detection results, SegNet, U-Net, and our proposed deep learning models of Area 1 with SE.

Global Target Detection					
	AUC	**PD/TPR**	**PF/FPR**	**ACC**	**Kappa**
SE-WBS-ACE	0.9658	0.9154	0.0887	0.9115	0.4058
SE-WBS-TCIMF	0.9691	0.8963	0.0749	0.9241	0.4417
SE-WBS-CEM	0.9713	0.9029	0.0738	0.9254	0.4483
Sparse WCEM	0.971	0.9067	0.08	0.9194	0.4288
Local Target Detection					
	AUC	**PD/TPR**	**PF/FPR**	**ACC**	**Kappa**
SE Subset-WBS-CEM	0.9749	0.9159	0.0681	0.9313	0.4744
SE SW-WBS-CEM	0.9766	0.9241	0.0681	0.9316	0.4778
SE ASW-WBS-CEM	0.9796	0.9324	0.0616	0.9382	0.5076
Deep Learning Based (Proposed)					
Model	**AUC**	**TPR**	**FPR**	**ACC**	**Kappa**
SE-SegNet	0.846	0.751	0.063	0.93	0.42
SE-U-Net	0.852	0.719	0.014	0.974	0.69
SE-3L SegNet	0.872	0.76	0.015	0.976	0.697
SE-3L-U-SegNet	0.889	0.789	0.012	0.98	0.743
SE-2L-U-SegNet	0.912	0.844	0.019	0.976	0.713
SE-2L-Conv-U-SegNet	0.879	0.77	0.011	0.98	0.737

The red/green values indicate the best/worst performance in each evaluation metric.

Figure 18 plots the ROC curves on the results in Tables 3 and 4. Since target detection algorithms have very similar results, we only plot WBS-TCIMF, Subset WBS-CEM, and SW WBS-CEM. In addition, deep learning models provide hard decisions. In this case, their ROC curves are rectangles.

Table 5 shows the detection results of global and local target detection, SegNet, U-Net, and the proposed models of Area 2. In Area 2, the Kappa of U-Net was 0.692, which was the best among various models but provides the worst AUC, and TPR. Local target detection of Subset CEM has the best AUC but Kappa is not acceptable. However, our proposed 3L-SN and 3L-USN provide second-highest ACC and Kappa, slightly lower AUC but overall performance is better than others.

(**a**) without SE   (**b**) with SE

Figure 18. ROC curve of the global and local target detection results, SegNet, U-Net, and our proposed deep learning mod-els in Area 1.

Table 5. Detection comparison among the global and local target detection results, SegNet, U-Net, and our proposed deep learning models of Area 2 without SE.

	AUC	TPR	FPR	ACC	Kappa
Global Target Detection					
Traditional ACE	0.943	0.899	0.131	0.871	0.325
WBS-ACE	0.948	0.910	0.140	0.862	0.312
Traditional TCIMF	0.952	0.907	0.138	0.864	0.315
WBS-TCIMF	0.888	0.925	0.230	0.777	0.202
Traditional CEM	0.950	0.903	0.132	0.870	0.324
WBS-CEM	0.893	0.925	0.220	0.787	0.212
Local Target Detection					
Subset CEM	0.952	0.893	0.114	0.887	0.359
Subset WBS-CEM	0.917	0.923	0.180	0.824	0.255
SW CEM	0.921	0.849	0.110	0.889	0.351
SW WBS-CEM	0.923	0.903	0.146	0.856	0.299
ASW CEM	0.942	0.875	0.097	0.901	0.390
ASW WBS-CEM	0.932	0.918	0.141	0.862	0.313
Deep Learning Based (Proposed)					
SegNet	0.867	0.819	0.086	0.91	0.397
U-Net	0.806	0.617	0.006	0.98	0.692
3L SegNet	0.916	0.87	0.038	0.958	0.618
3L-USN	0.931	0.906	0.044	0.954	0.605
2L-USN	0.934	0.928	0.059	0.94	0.541
2L-Conv-USN	0.932	0.912	0.048	0.951	0.587

The red/green values indicate the best/worst performance in each evaluation metric.

Table 6 shows the detection results of all methods in Area 2 with Sparse enhancement. In Area 2, again, the Kappa of U-Net was 0.708, which was the best among various models

but provide the worst AUC and TPR. 3L-USN provides second-highest ACC and Kappa, slightly lower AUC but overall performance is better than others.
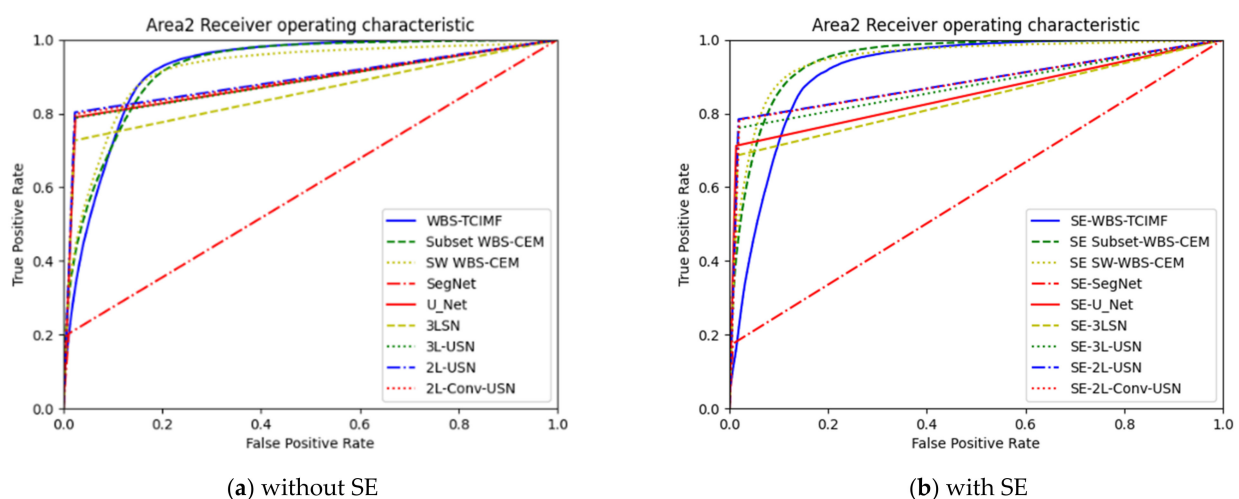
Table 6. Detection comparison among the global and local target detection results, SegNet, U-Net, and our proposed deep learning models of Area 2 with SE.

|	AUC	TPR	FPR	ACC	Kappa
Global Target Detection					
SE-WBS-ACE	0.950	0.911	0.126	0.876	0.339
SE-WBS-TCIMF	0.917	0.909	0.162	0.841	0.276
SE-WBS-CEM	0.916	0.916	0.172	0.832	0.265
Local Target Detection					
	AUC	TPR	FPR	ACC	Kappa
Subset SE-WBS-CEM	0.939	0.921	0.140	0.863	0.316
SW SE-WBS-CEM	0.938	0.898	0.120	0.881	0.347
ASW SE-WBS-CEM	0.949	0.924	0.117	0.885	0.363
Deep Learning Based (Proposed)					
Model	AUC	TPR	FPR	ACC	Kappa
SegNet	0.884	0.874	0.106	0.893	0.368
U-Net	0.831	0.669	0.008	0.98	0.708
3LSN	0.919	0.881	0.042	0.955	0.602
3L-USN	0.914	0.86	0.032	0.964	0.651
2L-USN	0.902	0.828	0.024	0.969	0.68
2L-Conv-USN	0.911	0.855	0.033	0.962	0.636

The red/green values indicate the best/worst performance in each evaluation metric.

Figure 19 plots the ROC curves on the results in Tables 4 and 5. To further highlight the detection results, Figures 20 and 21 use three colors to mark false information (blue), hits (red), and target misses (yellow).

As the training sample, Area 1 had a smaller forest coverage, and only the right half had coverage. The detection result of Area 2 was worse than that of Area 1.

(**a**) without SE　　　　　　　　　　　　　　　　　　　　(**b**) with SE

Figure 19. ROC curve of the global and local target detection results, SegNet, U-Net, and our proposed deep learning models in Area 2.
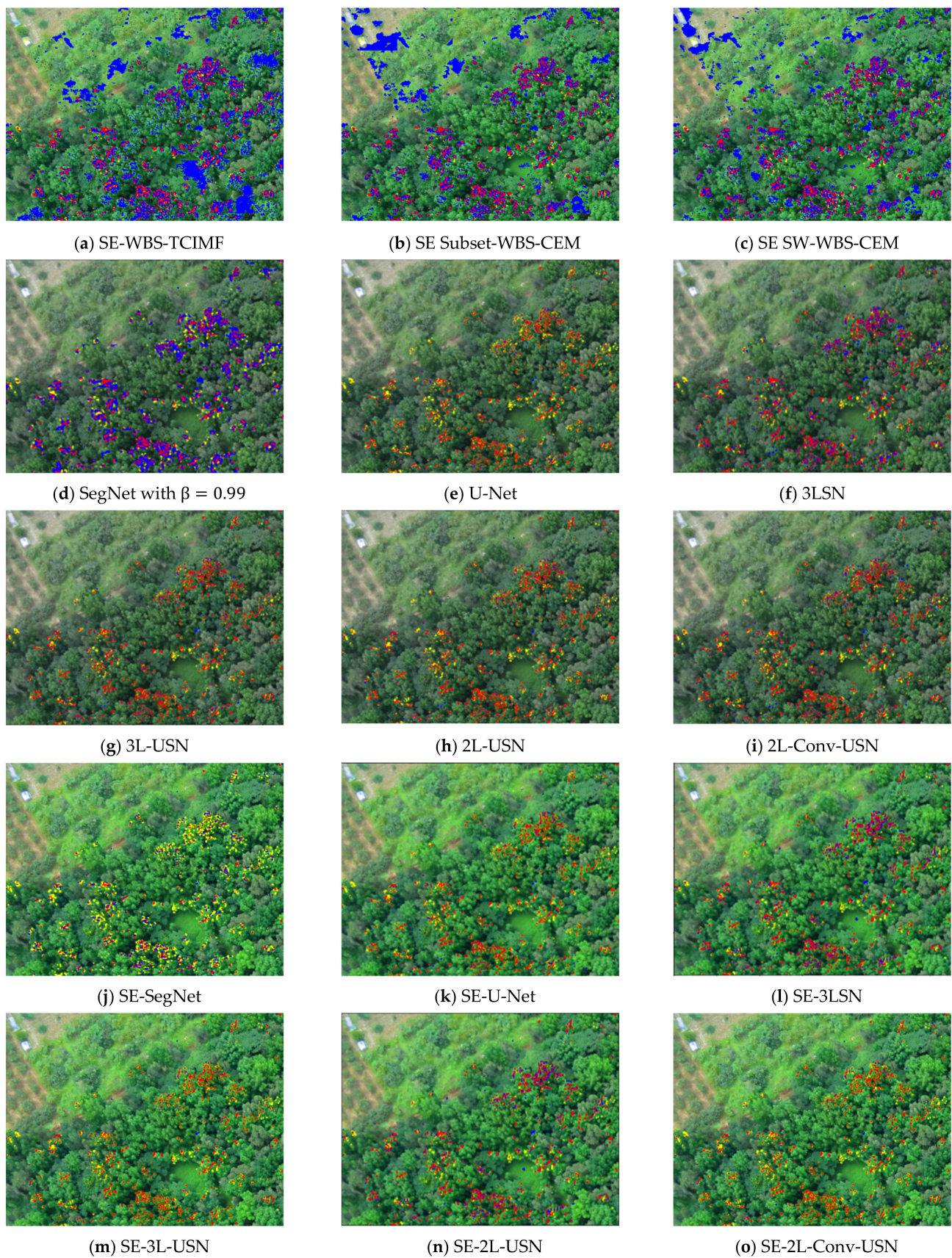
Figure 20. Resulting images of the global and local target detection results, SegNet, U-Net, and our proposed deep learning models in Area 1.
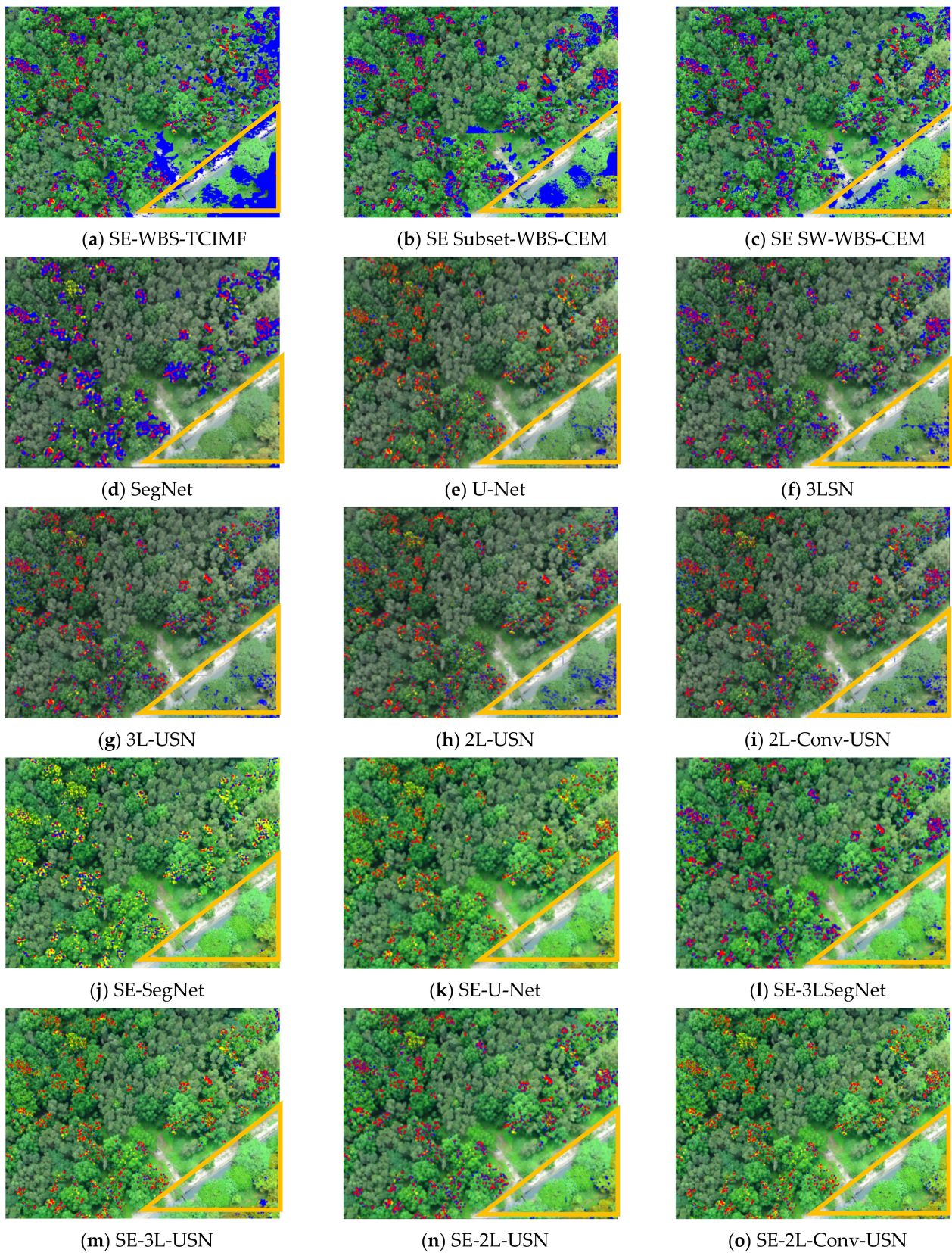
Figure 21. Resulting images of the global and local target detection results, SegNet, U-Net, and our proposed deep learning models in Area 2.

It is obvious our proposed deep learning models in Figure 21j–o have very few blue points (false alarm), especially in the right lower corner. According to the above resulting

images and data, the original SegNet had excessive layers of pooling and the feature information of tiny targets had disappeared, which influenced the detection rate. As our hybrid CNN architecture reduced the pooling degree, the NGL detection rate was increased greatly. Figures 22 and 23 show the data comparative histograms of SegNet, U-Net, and the proposed models of Area 1 and Area 2.
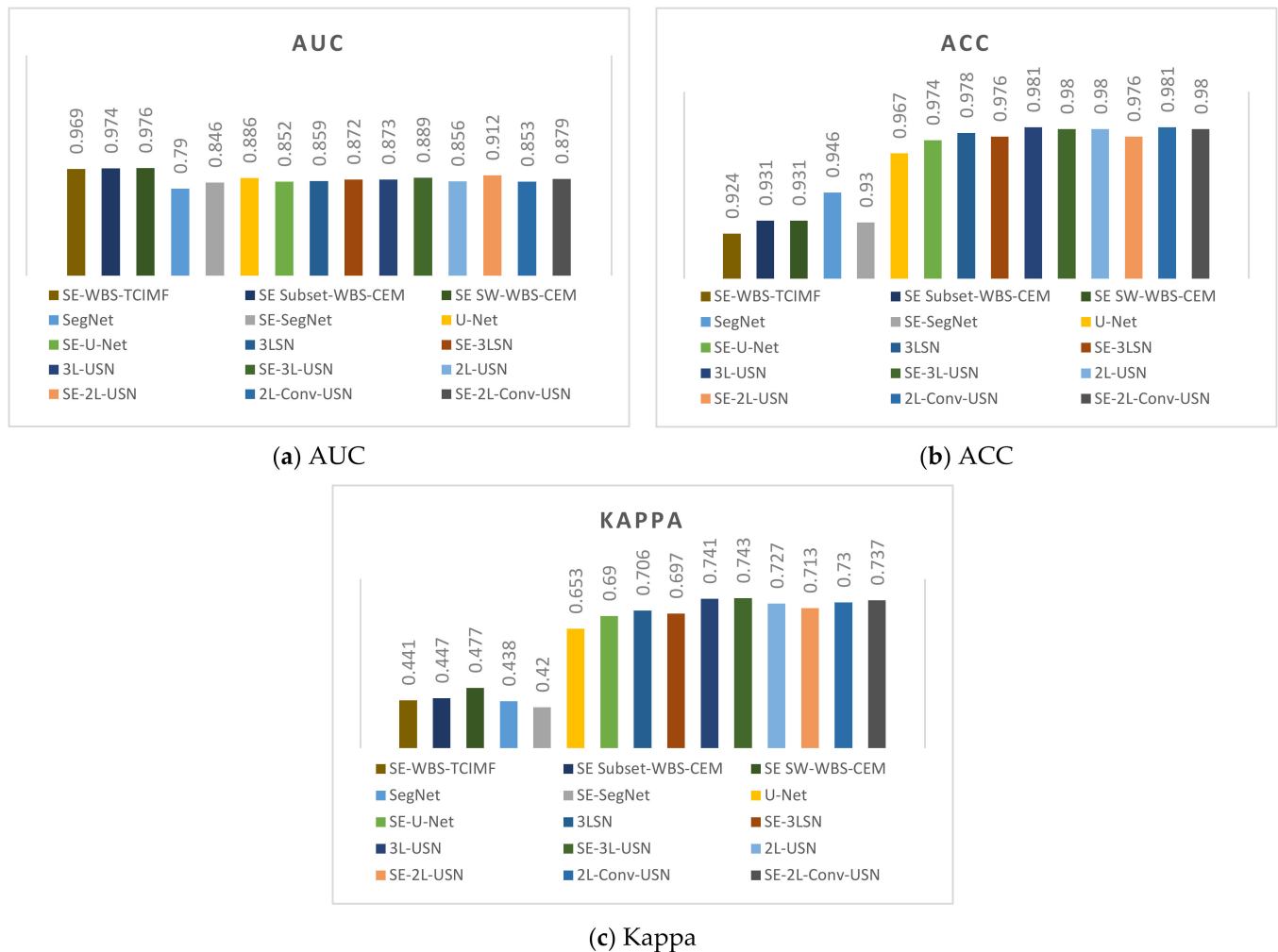

(**a**) AUC

(**b**) ACC

(**c**) Kappa

Figure 22. Model data comparative histogram of Area 1.
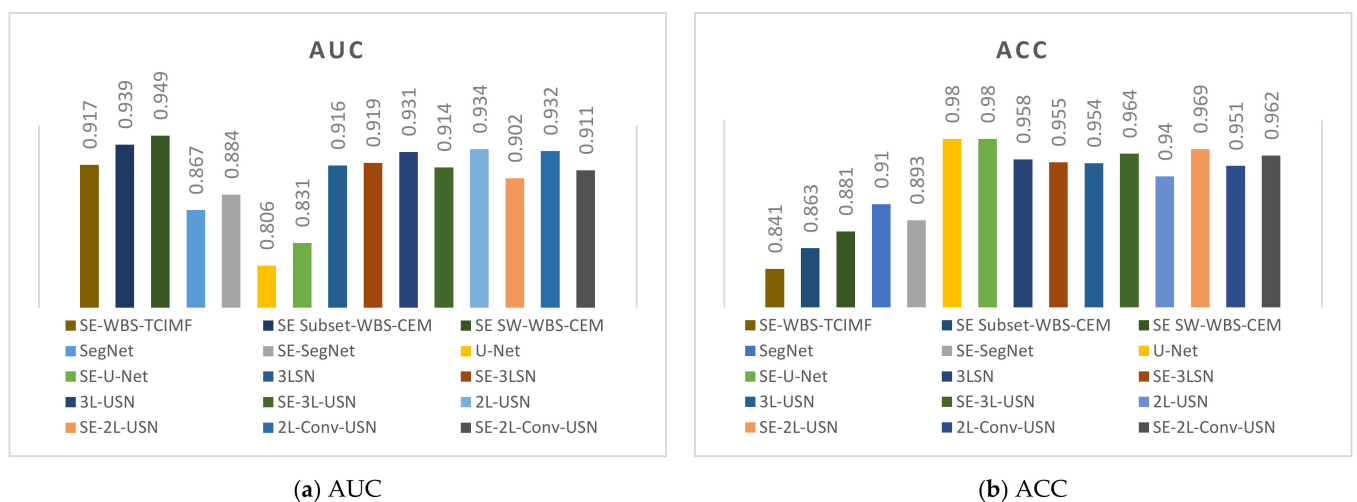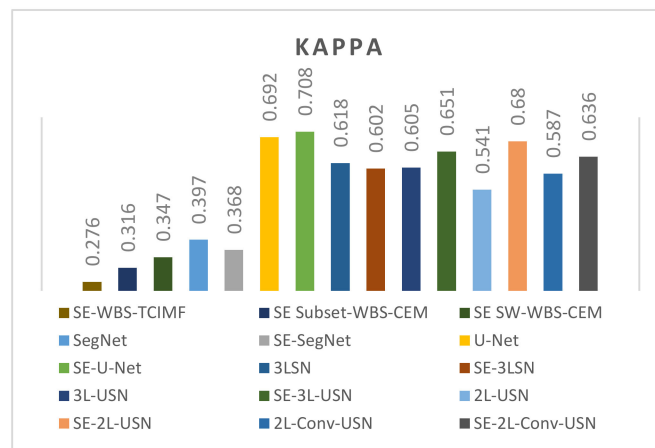

(**a**) AUC

(**b**) ACC

Figure 23. *Cont*.

(**c**) Kappa

Figure 23. Model data comparative histogram of Area 2.

4. Discussion

Many intriguing findings can be observed from Tables 3 and 4. Firstly, prior studies [4,76,78] used hyperspectral target detection algorithms. An advantage of hyperspectral target detection algorithms is they only require one spectral signature (target of interest) as an input parameter, while other prior knowledge such as multiple targets of interest or background is not required. However, the biggest problem is that these algorithms only provide soft decisions. As thresholding is required for the final decision, selecting a good threshold is a big challenge. This paper used deep learning models, which represent hard decisions without selecting thresholds. Moreover, global and local target detection methods are filter-based algorithms that apply a finite impulse response (FIR) filter to pass through the target of interest while minimizing and suppressing the background using a specific constraint [99]. It can detect many similar targets but also generates a very high false alarm rate. If the detection power or true positive rate (TPR) is the priority without considering a false alarm, using target detection algorithms is the appropriate manner. However, our proposed deep learning models provide a little lower AUC and TPR but achieve the highest overall accuracy and Kappa with the lowest false positive rate. Moreover, target detection methods are very sensitive since it only needs one input signature. Deep learning methods have more stable results but require a certain amount of training samples. Therefore, target detection and deep learning methods start from different points of view which can be applied in different circumstances. If you have limited information and concern computing time, matched filter-based target detection can meet demands. On the other hand, if you already have a certain amount of information and computing time is not your major concern, deep learning methods can fulfill your needs.

Secondly, our proposed models had much better results than the original SegNet and U-Net, meaning the information of tiny target NGL could be maintained by reducing the pooling degree, and U-SegNet with increased skip connection could enhance the spatial information of the sample. The performance in AUC, ACC, and Kappa were much better than the original SegNet and U-Net, especially in Kappa, which was up to 0.74.

Thirdly, NGL can be characterized by four unique features. First, they have unexpected presence. Secondly, they have a very low probability of occurrence, Thirdly, they have a relatively small sample population. Last but most important, their signatures are distinct from their surrounding pixels. As the features of NGL are coincident with the characteristics of sparsity, the NGL signal can be enhanced by our proposed sparse enhancement (SE) technique even areas may differ in the appearance of NGL. Tables 7 and 8 list different models after SE. The enhanced models were enhanced in AUC and TPR. According to the SegNet model result in Area 1, the AUC of the enhanced image was greatly increased by 0.56 and the TPR was increased by 0.13, but there were some false alarms; therefore

the Kappa decreased slightly. To sum up the results, the overall NGL enhancement was effective and the AUC and TPR of each model of Area 1 were increased significantly. Last but not the least, according to the data comparison table of Area 2 shown in Table 8, the AUC and TPR of various models were reduced but the ACC and Kappa were increased significantly. It could be observed in Figure 18 that the blue false alarm area in the right lower corner after SE was reduced greatly; therefore, the performance in ACC and Kappa was upgraded as the FPR decreased greatly. The AUC and TPR were reduced slightly but remained acceptable; therefore, the overall performance of SE in Area 1 indicated that the detection rate was increased, and the false alarm was reduced effectively in Area 2. Generally, the Kappa was improved significantly. As the NGL accounted for only 3% of the full image, the Kappa was relatively objective for evaluation. Another reason was that Area 1 is used as a training sample for testing Area 2. Area 1 has less NGL coverage indicates fewer training samples and Area 2 includes more species of trees in the right lower corner; therefore, the overall performance in Area 2 was worse than that in Area 1. To sum up the above data, the SE preprocessing method had a better detection effect on various models.

Table 7. Area 1 experiment data comparison table.

Area	Model	AUC	TPR	FPR	ACC	Kappa
1	**SegNet**	+0.560▲	+0.130▲	+0.021▲	−0.016▼	−0.018▼
	U-Net	−0.034▼	−0.080▼	−0.080▼	+0.008▲	+0.037▲
	3LSN	+0.013▲	+0.030▲	+0.003▲	−0.002▼	−0.009▼
	3L-USN	+0.114▲	+0.032▲	+0.032▲	−0.001▼	+0.002▲
	2L-USN	+0.056▲	+0.013▲	+0.010▲	−0.004▼	−0.014▼
	2L-Conv-USN	+0.026▲	+0.055▲	+0.002▲	−0.001▼	+0.007▲

Table 8. Area 2 experiment data comparison table.

Area	Model	AUC	TPR	FPR	ACC	Kappa
2	**SegNet**	+0.017▲	+0.055▲	+0.020▲	−0.017▼	−0.029▼
	U-Net	+0.025▲	+0.052▲	+0.002▲	0.000▲	+0.016▲
	3LSN	+0.003▲	+0.011▲	+0.004▲	−0.003▼	−0.016▼
	3L-USN	−0.017▼	−0.046▼	−0.012▼	+0.010▲	+0.001▲
	2L-USN	−0.032▼	−0.100▼	−0.035▼	+0.029▲	+0.141▲
	2L-Conv-USN	−0.021▼	−0.057▼	−0.015▼	+0.011▲	+0.049▲

The red/green values indicate the best/worst performance in each evaluation metric.

5. Conclusions

The use of telemetry to monitor the health of forest ecosystems has an important influence on controlling global warming. Prior studies have investigated and proposed target detection algorithms for NGL detection. It can achieve high detection power but cause a very high false alarm rate as well and only provide a soft decision, not a hard decision. To address the above issues, this paper makes several contributions. First and foremost, this paper applied deep learning in the SegNet network architecture and proposed four models, 3LSN, 3L-USN, 2L-USN, and 2L-Conv-USN by decreasing the number of pooling and adding skip connection (SC) technique. The original SegNet model has a high pooling frequency that can cause the low-level features of tiny targets to blur or even disappear; therefore, we reduced the number of layers of pooling. On the one hand, the model could be downsized; on the other hand, the loss of tiny targets could be reduced while maintaining information to increase the detection performance. When the skip connection (SC) of U-Net was used, the spatial information of the same level was connected up through up-sampling on the bottom layer. In comparison to SegNet and U-Net, the four models had better performance in NGL detection of tiny targets and training time. Secondly, this paper used the weighted BCE loss function since NGL in the

image belongs to imbalanced data; the weight of NGL can be increased, and the weight of the background was suppressed at the same time. Third, the SE technique was used as preprocessing to enhance NGL signal before entering deep learning models. According to the experimental results, after Area 1 was preprocessed by SE, the NGL detection rate was increased successfully, and the number of false alarms was reduced greatly in Area 2. Last but not the least, this paper conducts a complete comparative analysis, analyzing the pros and cons of global/local target detection algorithms and deep learning methods. To sum up, this paper is believed to be the first work that highlighted the use of hybrid deep learning models for the detection of NGL. The experimental results demonstrate the overall performance of our proposed models outperforms local/global target detection algorithms and original state of art deep learning models.

Author Contributions: Conceptualization, S.-Y.C.; data curation, C.L.; formal analysis, G.-J.L. and Y.-C.H.; funding acquisition, S.-Y.C.; investigation, S.-Y.C.; methodology, S.-Y.C.; project administration, S.-Y.C.; resources, C.L.; software, S.-Y.C., G.-J.L., and Y.-C.H.; supervision, S.-Y.C.; validation, S.-Y.C., G.-J.L., Y.-C.H. and K.-H.L.; visualization, S.-Y.C.; writing—original draft, S.-Y.C. and Y.-C.H.; writing—review and editing, C.L. and K.-H.L. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data and source codes are available from the authors upon reasonable request.

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

References

1. Food and Agriculture Organization of the United Nations. *Global Forest Resources Assessment 2015: How are the world's Forests Changing?* 2nd ed.; Food & Agriculture Org.: Rome, Italy, 2018.
2. Lin, C.; Dugarsuren, N. Deriving the Spatiotemporal NPP Pattern in Terrestrial Ecosystems of Mongolia Using MODIS Imagery. *Photogramm. Eng. Remote Sens.* **2015**, *81*, 587–598. [CrossRef]
3. Lin, C.; Thomson, G.; Popescu, S.C. An IPCC-compliant technique for forest carbon stock assessment using airborne LiDAR-derived tree metrics and competition index. *Remote Sens.* **2016**, *8*, 528. [CrossRef]
4. Lin, C.; Chen, S.-Y.; Chen, C.-C.; Tai, C.-H. Detecting newly grown tree leaves from unmanned-aerial-vehicle images using hyperspectral target detection techniques. *ISPRS J. Photogramm. Remote Sens.* **2018**, *142*, 174–189. [CrossRef]
5. Götze, C.; Gerstmann, H.; Gläßer, C.; Jung, A. An approach for the classification of pioneer vegetation based on species-specific phenological patterns using laboratory spectrometric measurements. *Phys. Geogr.* **2017**, *38*, 524–540. [CrossRef]
6. Burai, P.; Deák, B.; Valkó, O.; Tomor, T. Classification of Herbaceous Vegetation Using Airborne Hyperspectral Imagery. *Remote Sens.* **2015**, *7*, 2046–2066. [CrossRef]
7. Mohan, M.; Silva, C.A.; Klauberg, C.; Jat, P.; Catts, G.; Cardil, A.; Hudak, A.T.; Dia, M. Individual Tree Detection from Unmanned Aerial Vehicle (UAV) Derived Canopy Height Model in an Open Canopy Mixed Conifer Forest. *Forests* **2017**, *8*, 340. [CrossRef]
8. Wallace, L.; Lucieer, A.; Watson, C.S. Evaluating Tree Detection and Segmentation Routines on Very High Resolution UAV LiDAR Data. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 7619–7628. [CrossRef]
9. Dugarsuren, N.; Lin, C. Temporal variations in phenological events of forests, grasslands and desert steppe ecosystems in Mongolia: A remote sensing approach. *Ann. For. Res.* **2016**, *59*, 175–190. [CrossRef]
10. Popescu, S.C.; Zhao, K.; Neuenschwander, A.; Lin, C. Satellite lidar vs. small footprint airborne lidar: Comparing the accuracy of aboveground biomass estimates and forest structure metrics at footprint level. *Remote Sens. Environ.* **2011**, *115*, 2786–2797. [CrossRef]
11. Zeng, M.; Li, J.; Peng, Z. The design of Top-Hat morphological filter and application to infrared target detection. *Infrared Phys. Technol.* **2006**, *48*, 67–76. [CrossRef]

12. Gao, C.; Meng, D.; Yang, Y.; Wang, Y.; Zhou, X.; Hauptmann, A.G. Infrared Patch-Image Model for Small Target Detection in a Single Image. *IEEE Trans. Image Process.* **2013**, *22*, 4996–5009. [CrossRef]
13. Debes, C.; Zoubir, A.M.; Amin, M.G. Enhanced Detection Using Target Polarization Signatures in Through-the-Wall Radar Imaging. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 1968–1979. [CrossRef]
14. Qi, S.; Ma, J.; Tao, C.; Yang, C.; Tian, J. A Robust Directional Saliency-Based Method for Infrared Small-Target Detection Under Various Complex Backgrounds. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 495–499. [CrossRef]
15. Lo, C.-S.; Lin, C. Growth-Competition-Based Stem Diameter and Volume Modeling for Tree-Level Forest Inventory Using Airborne LiDAR Data. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 2216–2226. [CrossRef]
16. Lin, C.; Popescu, S.C.; Huang, S.C.; Chang, P.T.; Wen, H.L. A novel reflectance-based model for evaluating chlorophyll concentrations of fresh and water-stressed leaves. *Biogeosciences* **2015**, *12*, 49–66. [CrossRef]
17. Lin, C.; Tsogt, K.; Zandraabal, T. A decompositional stand structure analysis for exploring stand dynamics of multiple attributes of a mixed-species forest. *For. Ecol. Manag.* **2016**, *378*, 111–121. [CrossRef]
18. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
19. Hasan, M.A.; Lonardi, S. DeeplyEssential: A deep neural network for predicting essential genes in microbes. *BMC Bioinform.* **2020**, *21*, 367. [CrossRef]
20. Le, N.Q.K.; Do, D.T.; Hung, T.N.K.; Lam, L.H.T.; Huynh, T.-T.; Nguyen, N.T.K. A Computational Framework Based on Ensemble Deep Neural Networks for Essential Genes Identification. *Int. J. Mol. Sci.* **2020**, *21*, 9070. [CrossRef] [PubMed]
21. Le, N.Q.K.; Huynh, T.-T. Identifying SNAREs by Incorporating Deep Learning Architecture and Amino Acid Embedding Representation. *Front. Physiol.* **2019**, *10*, 1501. [CrossRef] [PubMed]
22. Song, J.; Liu, G.; Jiang, J.; Zhang, P.; Liang, Y. Prediction of Protein–ATP Binding Residues Based on Ensemble of Deep Convolutional Neural Networks and LightGBM Algorithm. *Int. J. Mol. Sci.* **2021**, *22*, 939. [CrossRef]
23. Le, N.Q.K. Fertility-GRU: Identifying Fertility-Related Proteins by Incorporating Deep-Gated Recurrent Units and Original Position-Specific Scoring Matrix Profiles. *J. Proteome Res.* **2019**, *18*, 3503–3511. [CrossRef]
24. Munir, K.; Elahi, H.; Ayub, A.; Frezza, F.; Rizzi, A. Cancer Diagnosis Using Deep Learning: A Bibliographic Review. *Cancers* **2019**, *11*, 1235. [CrossRef]
25. Chougrad, H.; Zouaki, H.; Alheyane, O. Deep Convolutional Neural Networks for breast cancer screening. *Comput. Methods Programs Biomed.* **2018**, *157*, 19–30. [CrossRef]
26. Lee, J.-H.; Kim, D.-H.; Jeong, S.-N.; Choi, S.-H. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *J. Dent.* **2018**, *77*, 106–111. [CrossRef]
27. Yu, B.; Wang, Y.; Wang, L.; Shen, D.; Zhou, L. Medical Image Synthesis via Deep Learning. *Adv. Exp. Med. Biol.* **2020**, *1213*, 23–44. [CrossRef] [PubMed]
28. Hu, W.; Zhang, Y.; Li, L. Study of the Application of Deep Convolutional Neural Networks (CNNs) in Processing Sensor Data and Biomedical Images. *Sensors* **2019**, *19*, 3584. [CrossRef] [PubMed]
29. Suzuki, K. Overview of deep learning in medical imaging. *Radiol. Phys. Technol.* **2017**, *10*, 257–273. [CrossRef] [PubMed]
30. Harangi, B. Skin lesion classification with ensembles of deep convolutional neural networks. *J. Biomed. Inform.* **2018**, *86*, 25–32. [CrossRef] [PubMed]
31. El-Khatib, H.; Popescu, D.; Ichim, L. Deep Learning–Based Methods for Automatic Diagnosis of Skin Lesions. *Sensors* **2020**, *20*, 1753. [CrossRef] [PubMed]
32. Cha, Y.J.; Choi, W.; Büyüköztürk, O. Deep learning-based crack damage detection using convolutional neural networks. *Comput.-Aided Civ. Infrastruct. Eng.* **2017**, *32*, 361–378. [CrossRef]
33. Dung, C.V.; Anh, L.D. Autonomous concrete crack detection using deep fully convolutional neural network. *Autom. Constr.* **2019**, *99*, 52–58. [CrossRef]
34. Wei, X.; Guo, Y.; Gao, X.; Yan, M.; Sun, X. A new semantic segmentation model for remote sensing images. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 1776–1779. [CrossRef]
35. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015.
36. Sun, S.; Yang, L.; Liu, W.; Li, R. Feature Fusion Through Multitask CNN for Large-scale Remote Sensing Image Segmentation. *arXiv* **2018**, arXiv:1807.09072v1.
37. Zheng, C.; Zhang, Y.; Wang, L. Semantic Segmentation of Remote Sensing Imagery Using an Object-Based Markov Random Field Model with Auxiliary Label Fields. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 3015–3028. [CrossRef]
38. Chen, G.; Zhang, X.; Wang, Q.; Dai, F.; Gong, Y.; Zhu, K. Symmetrical Dense-Shortcut Deep Fully Convolutional Networks for Semantic Segmentation of Very-High-Resolution Remote Sensing Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 1633–1644. [CrossRef]
39. Lotte, R.G.; Haala, N.; Karpina, M.; Aragão, L.E.O.E.C.D.; Shimabukuro, Y.E. 3D Façade Labeling over Complex Scenarios: A Case Study Using Convolutional Neural Network and Structure-From-Motion. *Remote Sens.* **2018**, *10*, 1435. [CrossRef]
40. Badrinarayanan, V.; Handa, A.; Cipolla, R. SegNet: A deep convolutional encoder-decoder architecture for robust semanticpixel-wise labelling. *arXiv* **2015**, arXiv:1505.07293.

41. Noh, H.; Hong, S.; Han, B. Learning deconvolution network for semantic segmentation. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1520–1528.

42. Treml, M.; Arjona-Medina, J.; Unterthiner, T.; Durgesh, R.; Friedmann, F.; Schuberth, P.; Mayr, A.; Heusel, M.; Hofmarcher, M.; Widrich, M.; et al. Speeding up semantic segmentation for autonomous driving. In Proceedings of the Conference on Neural Information Processing Systems (NIPS 2016), Barcelona, Spain, 5–10 December 2016.

43. Jeon, W.-S.; Rhee, S.-Y. Plant Leaf Recognition Using a Convolution Neural Network. *Int. J. Fuzzy Lg. Intell. Syst.* **2017**, *17*, 26–34. [CrossRef]

44. Kaya, A.; Keceli, A.S.; Catal, C.; Yalic, H.Y.; Temucin, H.; Tekinerdogan, B. Analysis of transfer learning for deep neural network based plant classification models. *Comput. Electron. Agric.* **2019**, *158*, 20–29. [CrossRef]

45. Nkemelu, D.K.; Omeiza, D.; Lubalo, N. Deep convolutional neural network for plant seedlings classification. *arXiv* **2018**, arXiv:1811.08404.

46. Lv, M.; Zhou, G.; He, M.; Chen, A.; Zhang, W.; Hu, Y. Maize Leaf Disease Identification Based on Feature Enhancement and DMS-Robust Alexnet. *IEEE Access* **2020**, *8*, 57952–57966. [CrossRef]

47. Dalal, T.; Singh, M. Review Paper on Leaf Diseases Detection and Classification Using Various CNN Techniques. In *Mobile Radio Communications and 5G Networks*; Springer International Publishing: Singapore, 2021; pp. 153–162.

48. Tm, P.; Pranathi, A.; SaiAshritha, K.; Chittaragi, N.B.; Koolagudi, S.G. Tomato Leaf Disease Detection Using Convolutional Neural Networks. In Proceedings of the 2018 Eleventh International Conference on Contemporary Computing (IC3), Noidia, India, 2–4 August 2018; pp. 1–5.

49. Gandhi, R.; Nimbalkar, S.; Yelamanchili, N.; Ponkshe, S. Plant disease detection using CNNs and GANs as an augmentative approach. In Proceedings of the 2018 IEEE International Conference on Innovative Research and Development (ICIRD), IEEE, Bangkok, Thailand, 11–12 May 2018; pp. 1–5.

50. Mohanty, S.P.; Hughes, D.P.; Salathé, M. Using Deep Learning for Image-Based Plant Disease Detection. *Front. Plant Sci.* **2016**, *7*, 1419. [CrossRef]

51. Naik, S.; Shah, H. Classification of Leaves Using Convolutional Neural Network and Logistic Regression. In *ICT Systems and Sustainability*; Springer: Singapore, 2021; pp. 63–75.

52. Grinblat, G.L.; Uzal, L.C.; Larese, M.G.; Granitto, P.M. Deep learning for plant identification using vein morphological patterns. *Comput. Electron. Agric.* **2016**, *127*, 418–424. [CrossRef]

53. Dos Santos Ferreira, A.; Freitas, D.M.; da Silva, G.G.; Pistori, H.; Folhes, M.T. Weed detection in soybean crops using ConvNets. *Comput. Electron. Agric.* **2017**, *143*, 314–324. [CrossRef]

54. Barré, P.; Stöver, B.C.; Müller, K.F.; Steinhage, V. LeafNet: A computer vision system for automatic plant species identification. *Ecol. Inform.* **2017**, *40*, 50–56. [CrossRef]

55. Badrinarayanan, V.; Kendall, A.; Cipolla, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef]

56. Pradhan, P.; Meyer, T.; Vieth, M.; Stallmach, A.; Waldner, M.; Schmitt, M.; Popp, J.; Bocklitz, T. Semantic Segmentation of Non-linear Multimodal Images for Disease Grading of Inflammatory Bowel Disease: A SegNet-based Application. In Proceedings of the 8th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2019), Prague, Czech Republic, 19–21 February 2019; pp. 396–405.

57. Khagi, B.; Kwon, G.-R. Pixel-Label-Based Segmentation of Cross-Sectional Brain MRI Using Simplified SegNet Architecture-Based CNN. *J. Healthc. Eng.* **2018**, *2018*, 3640705. [CrossRef] [PubMed]

58. Mittal, A.; Hooda, R.; Sofat, S. LF-SegNet: A fully convolutional encoder–decoder network for segmenting lung fields from chest radiographs. *Wirel. Pers. Commun.* **2018**, *101*, 511–529. [CrossRef]

59. Falk, T.; Mai, D.; Bensch, R.; Çiçek, Ö.; Abdulkadir, A.; Marrakchi, Y.; Böhm, A.; Deubner, J.; Jäckel, Z.; Seiwald, K.; et al. U-Net: Deep learning for cell counting, detection, and morphometry. *Nat. Methods* **2019**, *16*, 67–70. [CrossRef]

60. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted, Munich, Germany, 5–9 October 2015; pp. 234–241.

61. Esser, P.; Sutter, E. A Variational U-Net for Conditional Appearance and Shape Generation. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 8857–8866.

62. Von Eicken, T.; Basu, A.; Buch, V.; Vogels, W. U-Net: A user-level network interface for parallel and distributed computing. *ACM SIGOPS Oper. Syst. Rev.* **1995**, *29*, 40–53. [CrossRef]

63. Dong, H.; Yang, G.; Liu, F.; Mo, Y.; Guo, Y. Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks. In Proceedings of the Annual Conference on Medical Image Understanding and Analysis, Edinburgh, UK, 11–13 July 2017; Springer: Cham, Switzerland, 2017; pp. 506–517.

64. Jansson, A.; Humphrey, E.; Montecchio, N.; Bittner, R.; Kumar, A.; Weyde, T. Singing voice separation with deep u-net convolutional networks. In Proceedings of the International Society for Music Information Retrieval Conference, Suzhou, China, 23–27 October 2017.

65. Kumar, P.; Nagar, P.; Arora, C.; Gupta, A. U-SegNet: Fully convolutional neural network based automated brain tissue segmentation tool. *arXiv* **2018**, arXiv:1806.04429.

66. Candès, E.J.; Li, X.; Ma, Y.; Wright, J. Robust principal component analysis? *J. ACM* **2011**, *58*, 1–37. [CrossRef]

67. Xu, Y.; Wu, Z.; Chanussot, J.; Wei, Z. Joint Reconstruction and Anomaly Detection from Compressive Hyperspectral Images Using Mahalanobis Distance-Regularized Tensor RPCA. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 2919–2930. [CrossRef]
68. Tan, C.H.; Chen, J.; Chau, L.P. Edge-preserving rain removal for light field images based on RPCA. In Proceedings of the 22nd International Conference on Digital Signal Processing (DSP), London, UK, 23–25 August 2017; pp. 1–5.
69. Rezaei, B.; Ostadabbas, S. Background Subtraction via Fast Robust Matrix Completion. In Proceedings of the 2017 IEEE International Conference on Computer Vision Workshops (ICCVW), Venice, Italy, 22–29 October 2017; pp. 1871–1879.
70. Kaloorazi, M.F.; De Lamare, R.C. Low-rank and sparse matrix recovery based on a randomized rank-revealing decomposition. In Proceedings of the 2017 22nd International Conference on Digital Signal Processing (DSP), London, UK, 23–25 August 2017; pp. 1–5.
71. Dao, M.; Kwan, C.; Ayhan, B.; Tran, T.D. Burn scar detection using cloudy MODIS images via low-rank and sparsity-based models. In Proceedings of the 2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP), Washington, DC, USA, 7–9 December 2016; pp. 177–181.
72. Lee, P.H.; Chan, C.C.; Huang, S.L.; Chen, A.; Chen, H.H. Blood vessel extraction from OCT data by short-time RPCA. In Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016; pp. 394–398.
73. Chai, Y.; Xu, S.; Yin, H. An Improved ADM algorithm for RPCA optimization problem. In Proceedings of the 32nd Chinese Control Conference, Xi'an, China, 26–28 July 2013; pp. 4476–4480.
74. Wen, F.; Zhang, Y.; Gao, Z.; Ling, X. Two-Pass Robust Component Analysis for Cloud Removal in Satellite Image Sequence. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 1090–1094. [CrossRef]
75. Lin, Z.; Chen, M.; Ma, Y. The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-Rank Matrices. *arXiv* **2010**, arXiv:1009.5055.
76. Chen, S.Y.; Lin, C.; Tai, C.H.; Chuang, S.J. Adaptive Window-Based Constrained Energy Minimization forDetection of Newly Grown Tree Leaves. *Remote Sens.* **2018**, *10*, 96. [CrossRef]
77. Chang, C.-I. *Hyperspectral Data Processing: Algorithm Design and Analysis*; Wiley: Hoboken, NJ, USA, 2013.
78. Chen, S.-Y.; Lin, C.; Chuang, S.-J.; Kao, Z.-Y. Weighted Background Suppression Target Detection Using Sparse Image Enhancement Technique for Newly Grown Tree Leaves. *Remote Sens.* **2019**, *11*, 1081. [CrossRef]
79. Bar, M.; Ori, N. Leaf development and morphogenesis. *Development* **2014**, *141*, 4219–4230. [CrossRef]
80. Lin, C.; Lin, C.-H. Comparison of carbon sequestration potential in agricultural and afforestation farming systems. *Sci. Agricola* **2013**, *70*, 93–101. [CrossRef]
81. Lin, C.; Wang, J.-J. The effect of trees spacing on the growth of trees in afforested broadleaf stands on cultivated farmland. *Q. J. Chin. For.* **2013**, *46*, 311–326.
82. Lin, C. Improved derivation of forest stand canopy height structure using harmonized metrics of full-waveform data. *Remote Sens. Environ.* **2019**, *235*, 111436. [CrossRef]
83. Zhou, Z.; Siddiquee, M.M.R.; Liang, N.T.J. Unet++: A nested u-net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*; Springer: Cham, Switzerland, 2018; pp. 3–11.
84. Oktay, O.; Schlemper, J.; Folgoc, L.L.; Lee, M.; Heinrich, M.; Misawa, K.; Mori, K.; McDonagh, S.; Hammerla, N.Y.; Kainz, B.; et al. Attention U-Net: Learning where to Look for the Pancreas. *arXiv* **2018**, arXiv:1804.03999.
85. De Boer, P.-T.; Kroese, D.P.; Mannor, S.; Rubinstein, R.Y. A Tutorial on the Cross-Entropy Method. *Ann. Oper. Res.* **2005**, *134*, 19–67. [CrossRef]
86. Aurelio, Y.S.; de Almeida, G.M.; de Castro, C.L.; Braga, A.P. Learning from imbalanced data sets with weighted cross-entropy function. *Neural Process. Lett.* **2019**, *50*, 1937–1949. [CrossRef]
87. Pan, S.; Zhang, W.; Zhang, W.; Xu, L.; Fan, G.; Gong, J.; Zhang, B.; Gu, H. Diagnostic Model of Coronary Microvascular Disease Combined with Full Convolution Deep Network with Balanced Cross-Entropy Cost Function. *IEEE Access* **2019**, *7*, 177997–178006. [CrossRef]
88. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the 32nd International Conference on Machine Learning, Lille, France, 6–11 July 2015.
89. Daimary, D.; Bora, M.B.; Amitab, K.; Kandar, D. Brain TumorSegmentation from MRI Images using Hybrid Convolutional NeuralNetworks. *Procedia Comput. Sci.* **2020**, *167*, 2419–2428. [CrossRef]
90. Wright, J.; Peng, Y.; Ma, Y.; Ganesh, A.; Rao, S. Robust Principal Component Analysis: Exact Recovery of Corrupted Low-Rank Matrices by Convex Optimization. In Proceedings of the Neural Information Processing Systems, NIPS, Vancouver, BC, Canada, 7–10 December 2009.
91. Bouwmans, T.; Zahza, E. Robust PCA via Principal Component Pursuit: A Review for a Comparative Evaluation in Video Surveillance. *Comput. Vis. Image Underst.* **2014**, *122*, 22–34. [CrossRef]
92. Vaswani, N.; Bouwmans, T.; Javed, S.; Narayanamurthy, P. Robust PCA and Robust Subspace Tracking. *IEEE Signal Process. Mag.* **2017**, *35*, 32–55. [CrossRef]
93. Hauberg, S.; Feragen, A.; Black, M.J. Grassmann Averages for Scalable Robust PCA. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–27 June 2014; pp. 3810–3817.
94. Wang, N.; Yao, T.; Wang, J.; Yeung, D.Y. A Probabilistic Approach to Robust Matrix Factorization. In Proceedings of the European Conference on Computer Vision, Firenze, Italy, 7–13 October 2012; pp. 126–139.

95. Aravkin, A.; Becker, S.; Cevher, V.; Olsen, P. A variational approach to stable principal component pursuit. *arXiv* **2014**, arXiv:1406.1089.
96. Zhou, T.; Tao, D. GoDec: Randomized low-rank & sparsity matrix decompositionin noisy case. In Proceedings of the 28th International Conference on Machine Learning, Bellevue, WA, USA, 28 June–2 July 2011.
97. Albertella, A.; Sacerdote, F. Spectral analysis of block averaged data in geopotential global model determination. *J. Geod.* **1995**, *70*, 166–175. [CrossRef]
98. Poor, H.V. *An Introduction to Detection and Estimation Theory*, 2nd ed.; Springer: New York, NY, USA, 1994.
99. Chen, S.-Y.; Chang, C.-Y.; Ou, C.-S.; Lien, C.-T. Detection of Insect Damage in Green Coffee Beans Using VIS-NIR Hyperspectral Imaging. *Remote Sens.* **2020**, *12*, 2348. [CrossRef]