

Technical Note

Res2-Unet+, a Practical Oil Tank Detection Network for Large-Scale High Spatial Resolution Images

Bo Yu ¹, Fang Chen ^{1,2,3,4}, Yu Wang ^{5,*}, Ning Wang ^{1,2}, Xiaoyu Yang ⁶, Pengfei Ma ⁵, Chunyan Zhou ⁵ and Yuhuan Zhang ⁵

- ¹ Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China; yubo@radi.ac.cn (B.Y.); chenfang_group@163.com (F.C.); wangning171@mails.ucas.edu.cn (N.W.)
 - ² University of Chinese Academy of Sciences, Beijing 100049, China
 - ³ Hainan Key Laboratory of Earth Observation, Aerospace Information Research Institute, Chinese Academy of Sciences, Sanya 572029, China
 - ⁴ State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China
 - ⁵ Center for Satellite Application on Ecology and Environment, Ministry of Ecology and Environment, Beijing 100006, China; mapengfei@e-mail.secmep.cn (P.M.); zhouchunyan@e-mail.secmep.cn (C.Z.); zhangyuhuan@e-mail.secmep.cn (Y.Z.)
 - ⁶ Jinan Environmental Monitoring Center of Shandong Province, Jinan 250013, China; jnhbjyangxiaoyu@jn.shandong.cn
- * Correspondence: wangyu@e-mail.secmep.cn; Tel.: +86-18519661716



Citation: Yu, B.; Chen, F.; Wang, Y.; Wang, N.; Yang, X.; Ma, P.; Zhou, C.; Zhang, Y. Res2-Unet+, a Practical Oil Tank Detection Network for Large-Scale High Spatial Resolution Images. *Remote Sens.* **2021**, *13*, 4740. <https://doi.org/10.3390/rs13234740>

Academic Editor:
Melanie Vanderhoof

Received: 20 October 2021
Accepted: 18 November 2021
Published: 23 November 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: Oil tank inventory is significant for the economy and the military, as it can be used to estimate oil reserves. Traditional oil tank detection methods mainly focus on the geometrical characteristics and spectral features of remotely sensed images based on feature engineering. The methods have a limited application capability when the distribution pattern of ground objects in the image changes and the imaging condition varies largely. Therefore, we propose an end-to-end deep convolution network Res2-Unet+, to detect oil tanks in a large-scale area. The Res2-Unet+ method replaces the typical convolution block in the encoder of the original Unet method using hierarchical residual learning branches. A hierarchical branch is used to decompose the feature map into a few sub-channel features. To evaluate the generalization and transferability of the proposed model, we use high spatial resolution images from three different sensors in different areas to train the oil tank detection model. Images from yet another sensor in another area are used to evaluate the trained model. Three more widely used methods, Unet, Segnet, and PSPNet, are trained and evaluated for the same dataset. The experiments prove the effectiveness, strong generalization, and transferability of the proposed Res2-Unet+ method.

Keywords: oil tank detection; high spatial resolution image; oil reservation

1. Introduction

The oil tank, as a storage container of oil products, is a widely used piece of equipment in the petroleum, natural gas and petrochemical industries [1]. It is particularly important in storing and transferring oil and its related products, such as petroleum products, that are liquid at ambient temperature. The timely and accurate detection of oil tanks in a large-scale area is important in estimating oil reserves and provides data support to formulate policies related to oil production and reserves. Since oil tanks are usually located in residential areas or beside harbors, real-time oil tank detection is essential in assessing the threat of oil explosion and leakage. However, retrieving detailed records of oil tanks from the public domain is very difficult in China, necessitating alternative retrieval methods.

Recently, with the continuous development and maturity of remote sensing technology, increased amounts of high-spatial resolution satellite data, including from Quickbird,

WorldView, SPOT, Gaofen-1, Gaofen-2, Gaofen-6, and Ziyuan satellites, have been widely used in various object detection tasks, and have achieved fruitful results [2–4]. As an important energy storage device, the oil tank has become key in remote sensing exploration systems [5]. However, due to the variable illumination, viewing angles, and imaging quality of different remote sensed images, the edges of oil tanks in the remote sensed images are usually fuzzy, and their colors are not uniform. Moreover, oil tanks in different places vary significantly in their distribution pattern and in their complex of background objects, as shown in Figure 1. Oil tanks have different shapes, sizes, and roof materials. Omission error is easy to trigger when oil tanks are small, and commission error is likely to occur when background objects have similar spectral and geometrical characteristics. To deal with these issues, it is necessary to carry out research on the efficient detection technology of oil tanks from remote sensed images.



Figure 1. Oil tank demonstration: (a,b) two demonstration cases.

There has been much research on oil tank detection based on high spatial resolution images. Kushwaha et al. [6] proposed a knowledge-based strategy to detect bright oil tanks after morphological segmentation from high spatial resolution images. The Hough transform [7] and its related methods [8–10] have been widely used in oil tank detection [11], as they are efficient in finding circles. The shapes of shadow has been successfully used to supplement oil tank detection [12,13]. Owing to the development of image processing techniques, saliency enhancement has been employed to highlight oil tanks in images to reduce false detections [14]. Geometric characteristics, such as symmetry and contour shape, are used as criteria to eliminate background objects [15,16]. The synthesis of saliency enhancement and shapes promotes an unsupervised shape-guided model [Jing, 2018 #17], but it is sensitive to the boundary of adjacent objects in the image. Most of the above methods focus on the morphological characteristics of oil tanks using image segmentation techniques, which heavily rely on manually determined thresholds and criteria pre-defined by the researcher to detect oil tanks.

The advent of machine learning methods has made it possible to learn the criteria and the thresholds from training samples. This has improved oil tank detection by enlarging the study area with more diverse oil tanks. Support vector machines (SVMs) have been used to construct models to separate oil tanks from other background objects [17,18]. Contrario Clustering has been proposed to reduce false alarms in oil tank detection [19]. Moreover, with the popularity and the outstanding performance of deep convolutional networks in computer vision [20], convolutional networks have been employed to detect oil tanks [1,18]. However, these networks are mostly used for extracting features after image pre-segmentation based on the histograms of oriented gradients (HOGs) [21] and speeded up robust features (SURFs) [22]. The synthesis with other manually engineered features can enhance oil tanks, but can also remove other oil tanks as background objects. Multi-scale feature learning of oil tanks with different background objects is lacking in the models above. In this paper, we propose an end-to-end deep neural network Res2-Unet+ to detect oil tanks with various shapes, sizes, and illumination conditions in a large-scale area

based on high spatial resolution images. The network structure of Res2-Unet+ is modified from the typical network structure of Unet [23], which has been widely used for detecting various objects of interest based on remotely sensed images [24]. The main contributions of our paper are listed below:

- (1) An end-to-end deep neural network to detect oil tanks.
- (2) An enhanced learning capability of multi-scale oil tank features by adopting a hierarchical residual feature learning module to learn features at a more granular level, and to broaden the range of the receptive field of each network layer.
- (3) An increase in the potential for applications by training and evaluating the proposed oil tank detection model in different large-scale areas.

2. Related Works

We propose an end-to-end semantic segmentation framework, Res2-Unet+, to detect oil tanks in high spatial resolution remotely sensed images. This section introduces research concerning semantic segmentation deep learning models for object detection using remotely sensed images.

Semantic segmentation is a widely used technique in land cover classification and object detection, and assigns a semantic label to each pixel in an image [25]. A fully convolutional network (FCN) [26] is a significant deep neural network framework proposed for semantic segmentation. An FCN consists of three parts: a backbone convolution layer composed of typical network structures, such as ResNet [27] and VGG [28]; an up-sampling layer, which rescales the feature map to the size of the input image; and a skip layer, which optimizes the output result image. However, FCN has many unavoidable drawbacks, such as low precision, insensitivity to fine detail, and a lack of spatial consistency. In addition to FCN, SegNet [29] and atrous convolution [30] have been proposed to improve the segmentation resolution of the object of interest and capture multi-scale context information. Based on atrous convolution, a series of DeepLab models have been developed to achieve better segmentation by adopting multiple sampling rates [31]. However, a conditional random field (CRF) is required to finetune the segmentation results of the models above for detecting objects in remotely sensed images [32]. Research has been conducted on employing PSPNet [4] and CPAN [33] to detect buildings using corresponding multi-scale feature learning modules, but such research requires considerable computation time. Unet is a widely used semantic segmentation framework for various object detection tasks concerning remotely sensed images, as it is easy and highly efficient to implement [34]. Many networks have been proposed by improvement of Unet structure for object detection, such as Unet++ [35]. To the best of our knowledge, there has been little research on end-to-end semantic segmentation frameworks to detect oil tanks for large-scale applications. Therefore, this paper proposes an efficient and reliable end-to-end deep neural network based on Unet to detect various oil tanks a large-scale area.

3. Methods

A detailed network structure of our proposed Res2-Unet+ framework is presented in Figure 2. The input image, with a size of $3 \times 512 \times 512$ pixels, is firstly convoluted using an initial convolution module, and encoded to a feature map with 2048 channels using continuous convolution steps. The map is decoded to an output result image with a size of $2 \times 512 \times 512$ pixels. The initial convolution module is a set of three convolution operations with a kernel of a 3×3 element and a stride of 2, 1, and 1. The concatenation between features in the encoding and decoding stages at different scales integrates the characteristics of oil tanks at different levels, which increases accuracy. Because the objective is oil tank detection, the output result image is a 0–1 binary image, wherein 0 intensity indicates background objects and 1 represents oil tanks.

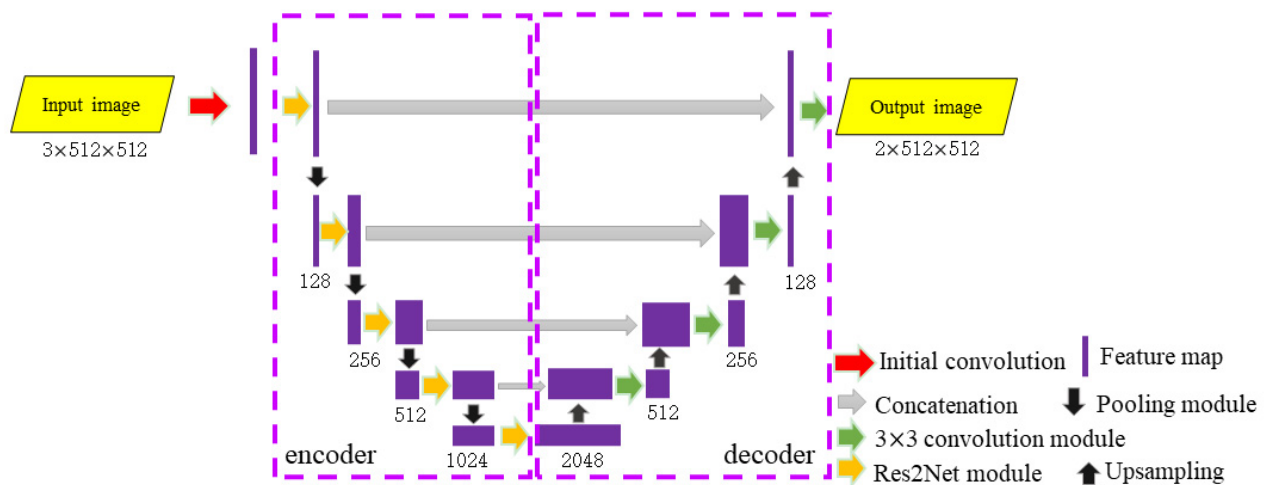


Figure 2. Flow chart of our proposed Res2-Unet+.

In line with the work by [36], a Res2Net module is adopted in the encoder section to gradually learn the multi-scale features of oil tanks. A comparison of the structure of a commonly used 3×3 convolution module and the detailed network structure of the Res2Net module is demonstrated in Figure 3. In a typical 3×3 convolution module, the input feature map is firstly convoluted by a 1×1 element kernel, generating Feature X, which is further convoluted by a 3×3 element kernel, generating Feature Y. Feature Y is concatenated with the input feature after convolution by a 1×1 element kernel to create an output feature. Instead of a direct 3×3 convolution on Feature X, the Res2Net module decomposes the n -channel feature map into four sub-channel feature maps, x_1 , x_2 , x_3 , and x_4 . Each decomposed feature map has a channel size of $\frac{n}{4}$, and experiences 3×3 convolution in a different way. The decomposed feature map x_1 is directly assigned as the output feature map y_1 . The decomposed feature map x_2 is convoluted by a 3×3 element kernel with stride of 1 to generate feature map m_2 , which is, in turn, used for generating output feature map y_2 . As demonstrated in Figure 3, feature map m_3 is generated by a 3×3 convolution with stride of 1 on the summary of feature map m_2 and x_3 . Feature map m_3 is assigned as output feature map y_3 and added with decomposed feature map x_4 . Similarly, with the network branch of decomposed feature maps x_2 and x_3 , output feature map y_4 is calculated and concatenated with output feature maps y_1 , y_2 , and y_3 to create Feature Y. Such multi-branch learning can enhance the multi-scale feature learning ability by learning at a more granular level and enlarging the receptive field. This is in contrast with the single-branch learning in commonly used 3×3 convolution modules.

In addition to the original Res2Net module from [36] in our proposed network Res2-Unet+, the order of activation function and batch normalization after each convolution operation in the convolution module has been transposed that activation function comes before batch normalization, as demonstrated in Figure 2. Such an operation is conducted according to the conclusion by [37]; that is, the non-negative response of the activation function ReLU (as illustrated in Equation (1)) may update the weight of the model in an unideal way.

$$y_{relu} = \max(0, x) \quad (1)$$

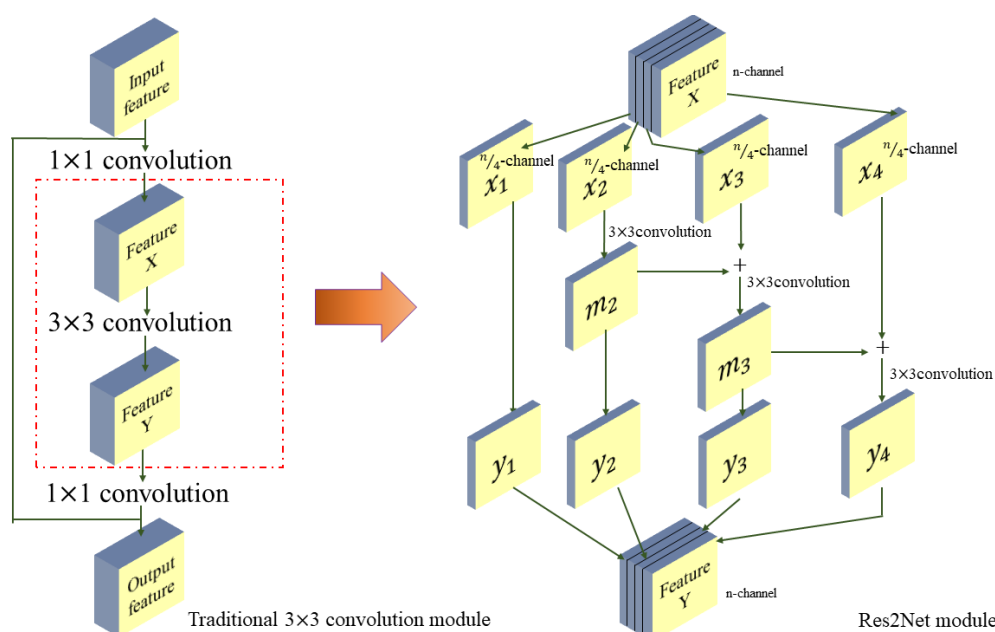


Figure 3. Comparison of the network structures of a traditional 3×3 convolution module and the original Res2Net module from [36].

4. Experiments and Results

In order to evaluate the robustness and transferability of the proposed framework for oil tank detection, the proposed Res2-Unet+ is trained and evaluated using remotely sensed images from multiple sensors in different places, obtained from the Center for Satellite Application on Ecology and Environment, Ministry of Ecology and Environment. We used images covering part of Dongying, Shandong Province, from the Ziyuan satellite (shown in Figure 4); Cangzhou, Hebei Province, from the Gaofen-1 satellite (shown in Figure 5); and Tangshan, Hebei Province, from the Gaofen-6 satellite (shown in Figure 6) to train the model. Such a strategy can enhance model generalization for images with different imaging conditions—different ground objects and distribution patterns. Images covering part of Yantai, Shandong Province, from the Gaofen-2 satellite, (shown in Figure 7) are used for evaluating the trained model. Detailed imaging information from the images used in our study is listed in Table 1. The evaluation image has a different spatial resolution from the training images, with different imaging sensors in different places. Such a strategy can be used to evaluate the transferability of the trained model. The ground truth oil tanks are visually interpreted by three experienced interpreters and validated by another experienced interpreter.

Table 1. General information of images used in this study.

Sensor	City	Time	Spatial Resolution	Train/Validate
Gaofen-1	Cangzhou	14 March 2020	2 m	Train
Gaofen-2	Yantai	20 July 2020	1 m	Validate
Gaofen-6	Tangshan	2 June 2020	2 m	Train
Ziyuan	Dongying	29 April 2020	2 m	Train

As shown in Figures 4–7, the oil tanks are located in both rural and urban areas. Background objects form a complex with large illumination variations and textural differences. Figure 4 shows that the oil tanks are mostly located in fragmented areas within factories, and neighboring background objects comprise small buildings with similar, bright spectral characteristics. The case in Figure 5 is comparatively more easily to detect. Farmland is the main background object, and the oil tanks are mainly located in open squares. However, the oil tanks in Figure 6 are scattered, and the background objects include densely distributed buildings, open squares, and other objects with complex spectral and textural

characteristics. In contrast with from the cases in Figures 4–6, the oil tanks in Figure 7 vary largely in different sizes and the background objects are mostly bright buildings, which are easy to be confused with each other. The oil tanks and background object distribution patterns differ significantly among different study areas. Such complexity provides abundant samples to enhance the generalization ability while training the model, and to evaluate the transferability of the trained model.

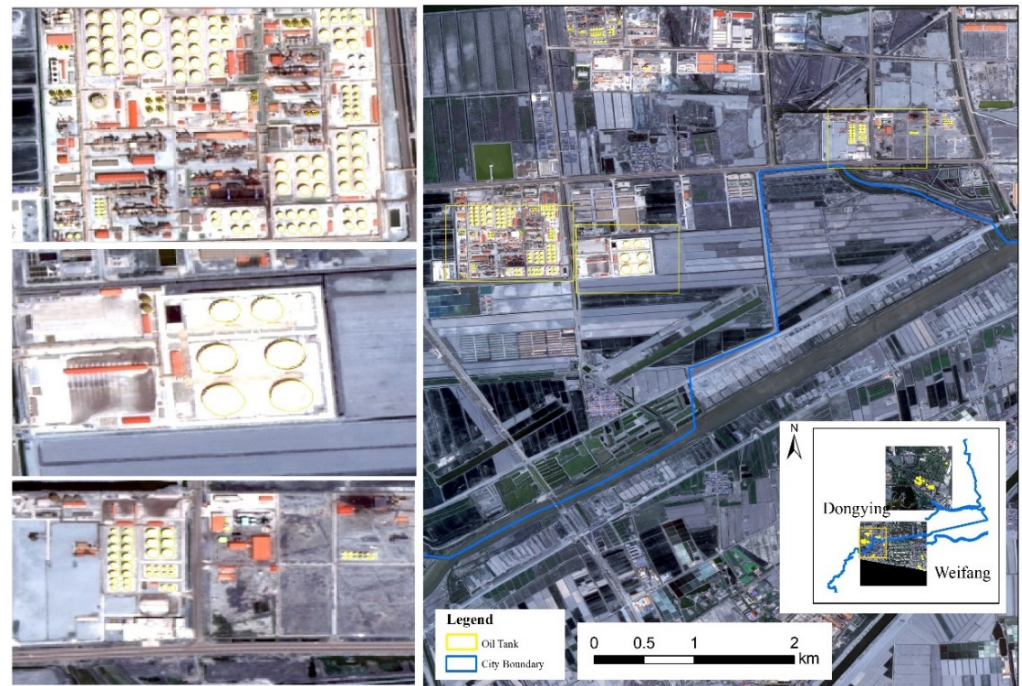


Figure 4. Study area of Dongying, Shandong Province, from a false color composition image from the sensor of Ziyuan.

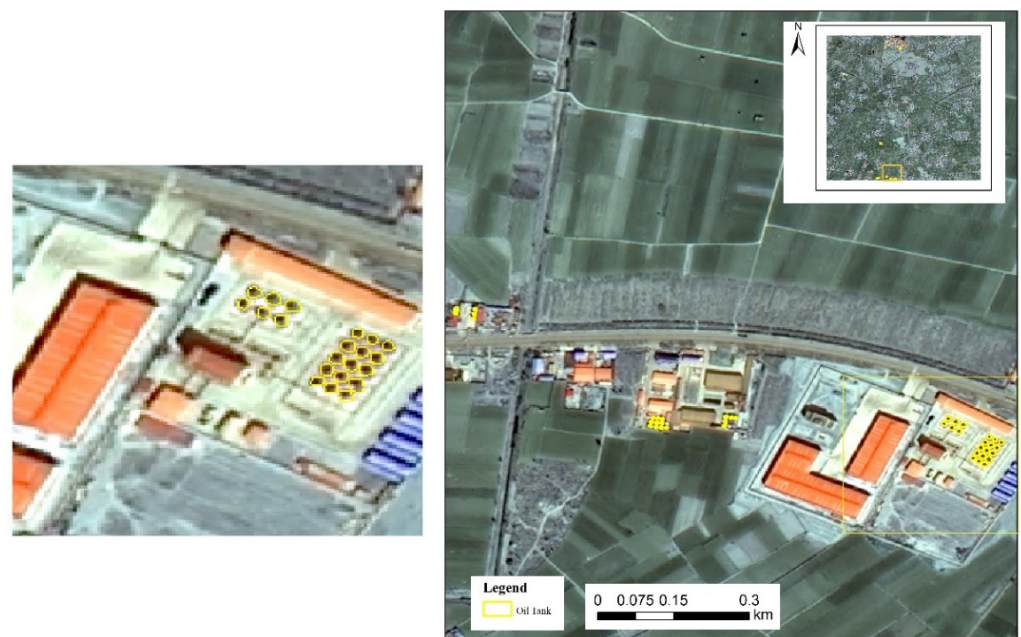


Figure 5. Study area of Cangzhou, Hebei Province, from a false color composition image from the Gaofen-1 sensor.

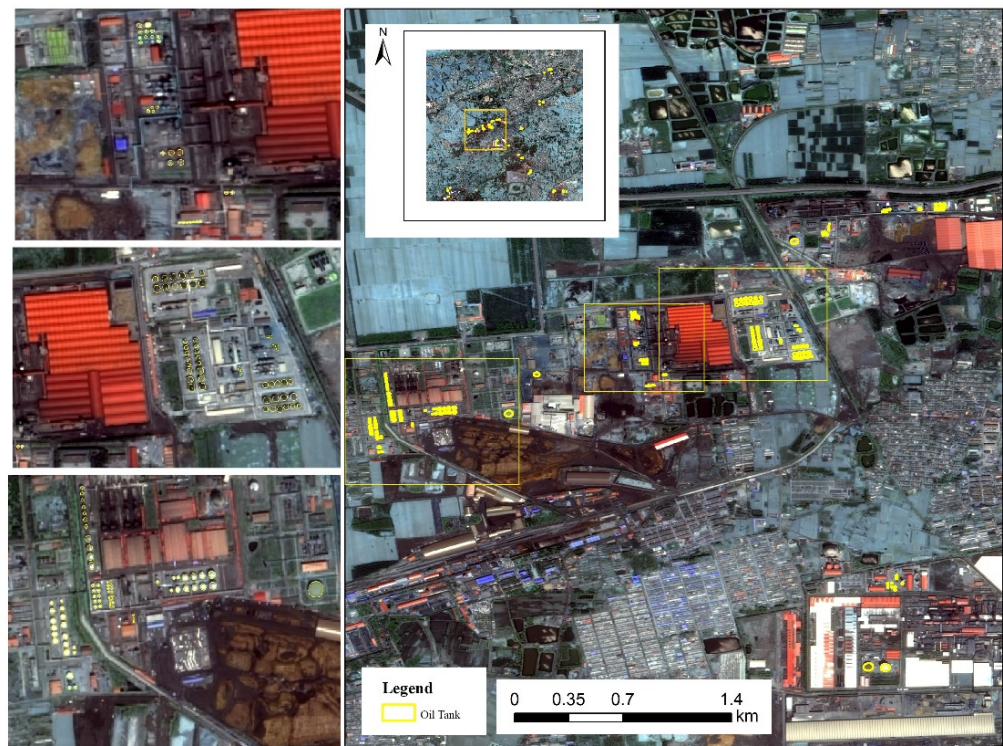


Figure 6. Study area of Tangshan, Hebei Province, from a false color composition image from the Gaofen-6 sensor.

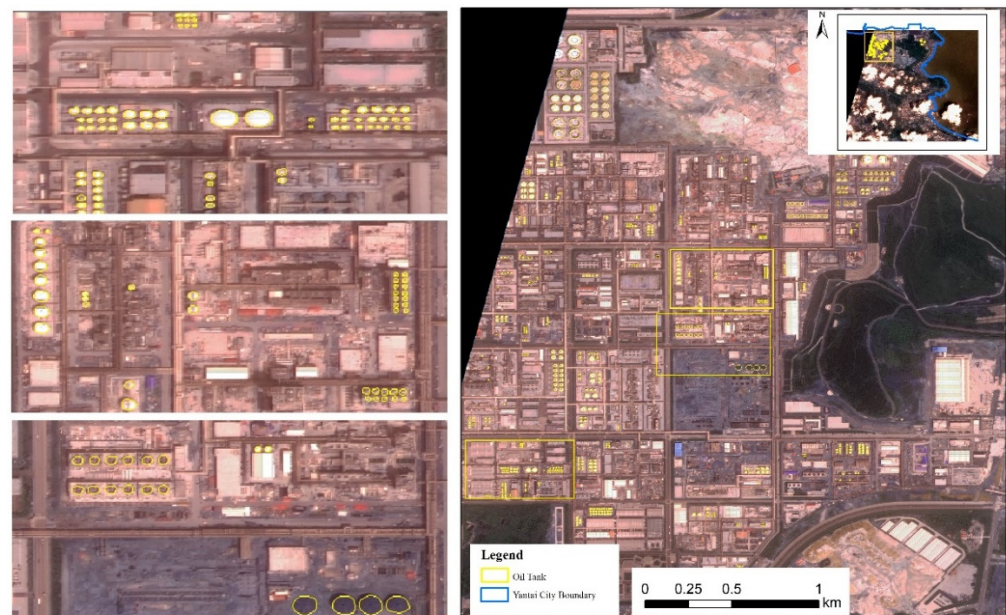


Figure 7. Study area of Yantai, Shandong Province, from a false color composition image from the Gaofen-2 sensor.

The images used in our study comprise three channels: the near infra-red channel, green channel, and blue channel, which is set through multiple trials. A combination of the channels can endow the background objects with distinctive spectral characteristics. Our proposed framework is implemented by ourselves in the Pytorch environment, with the operating system Ubuntu 16.04. The model is trained using three GPUs of TITAN X—each has a storage memory of 2 GB. Due to limited of GPU memory storage, it is difficult to load the original training and evaluating images directly. They are cropped into patches

with a size of 512×512 pixels. Moreover, since training and evaluating images are from different datasets, all the image patches in training and evaluating datasets have been used for model construction and evaluation. Stochastic Gradient Descent (SGD) optimizing [38] is adopted to optimize our model, and the initial learning rate is set at 0.01. However, an unwanted local minimum or saddle point is easily obtained when converging the model. A momentum of 0.9, therefore, is added to restrain SGD oscillation. Moreover, in order to avoid overfitting, a weight decay of 0.0001 is added before parameter regularization. Equation (2) encapsulates how our framework adopts *BCEloss* to train the oil tank detection model. The model converges to a minimum loss at the 156th epoch in our study. For comparative purposes, the original Unet, Segnet, and PSPNet frameworks are trained with the same experimental settings to evaluate whether our proposed Res2-Unet+ improves the detection performance of oil tanks compared with using other typical networks.

$$BCEloss(x_i, y_i) = -[y_i \log x_i + (1 - y_i) \log(1 - x_i)] \quad (2)$$

In order to conduct a visual comparison of the detected oil tanks using our proposed Res2-Unet+ method and other widely used methods, we randomly selected four patches with oil tanks from the evaluation images. The corresponding original images, ground truth images, and detected results using Res2-Unet+, Unet, Segnet, and PSPNet are displayed in Figure 8. Clearly, the proposed Res2-Unet+ framework can detect most oil tanks accurately. In the first case in Figure 8(a.1), Unet, Segnet, and PSPNet all falsely detect bright background buildings as oil tanks, while Res2-Unet+ successfully distinguish the bright buildings as background objects (shown in green squares). In Figure 8(c.1,d.1), the oil tanks detected by PSPNet are easily confused with each other, and those with relatively small apparent spectral differences with respect to the spectra in neighboring objects are easily omitted by PSPNet, as shown in Figure 8(b.5). However, Unet, Segnet, and Res2-Unet+ can detect most oil tanks without many false alarms in the cases of Figure 8(b.1,c.1,d.1). This indicates that the spatial pooling module adopted in PSPNet is not quite suitable for oil tank detection, especially when oil tanks are small and among background objects comprising small bright buildings.

To form an objective and comprehensive evaluation of the detection performance of oil tanks, we calculate several evaluation statistics based on the detected result images, referring to the ground truth oil tanks. In accordance with the evaluation strategy adopted in recent published work in object detection [24,32–34], the precision, recall, F1-measure, and Intersection-Over-Union (IOU) are calculated to evaluate the detection performance of each method. As indicated in Equations (3)–(6), *Precision* is the percentage of detected oil tank pixels that are ground truth oil tank, *Recall* is the percentage of ground truth oil tank pixels that are detected, *F1-measure* is calculated as a general performance indicator to balance *precision* and *recall*, *IOU* evaluates the performance concerning the overlap ratio between detected oil tank pixels and ground truth pixels. The averages of the calculated evaluation statistics for all the evaluation images from the Gaofen-2 satellite using each method are displayed in Table 2. Moreover, to evaluate the implementation efficiency, FPS (Frames per second) of each method is collected and listed in Table 2 as well.

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

$$F1 - measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (5)$$

$$IoU = \frac{TP}{TP + TN + FP} \quad (6)$$

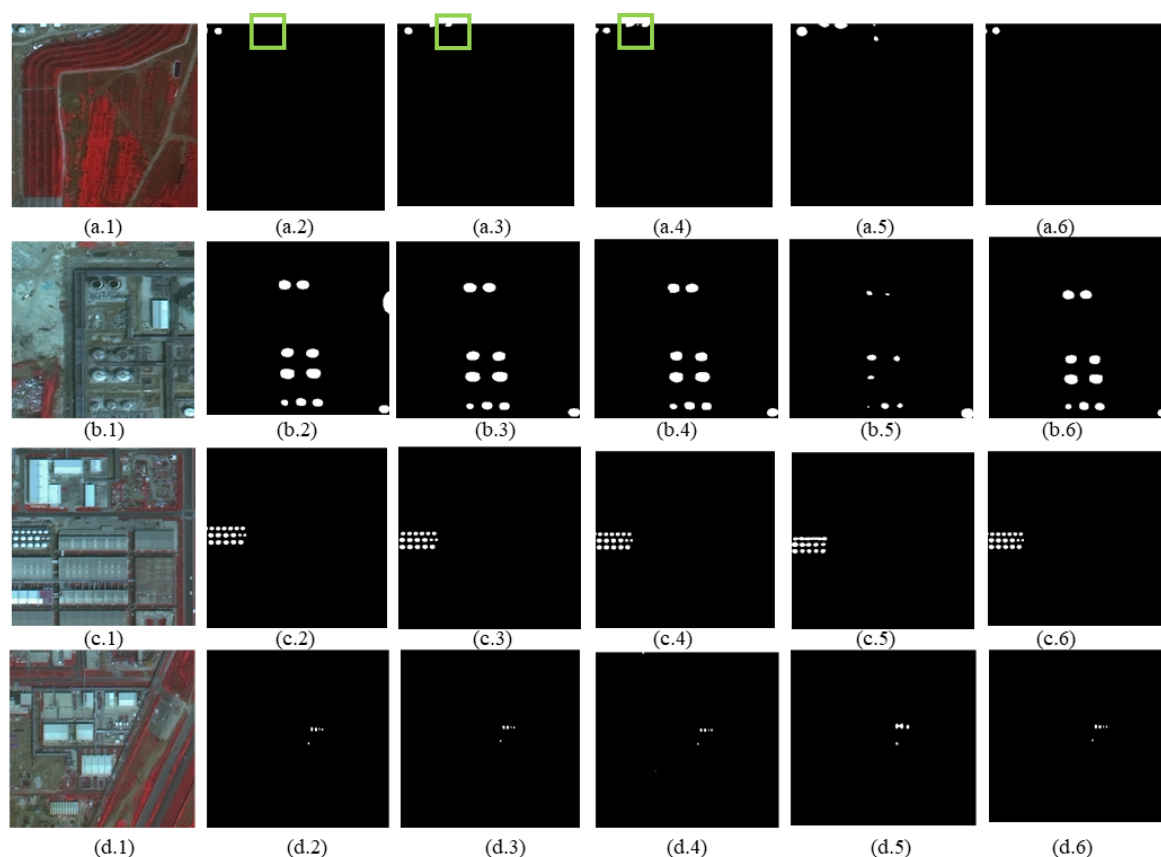


Figure 8. Examples of the detection of oil tanks using four methods: (*.1) original images; (*.2) ground truth images; (*.3) results using Unet; (*.4) results using Segnet; (*.5) results using PSPNet; and (*.6) results using Res2-Unet+.

Table 2. Evaluation statistics of oil tank detection from the evaluation images of the Gaofen-2 satellite (%).

Method	Precision	Recall	F1-Measure	IOU	FPS
Res2-Unet+	94.03	89.78	91.57	85.14	103.28
Unet	90.96	89.11	89.69	84.22	105.32
Segnet	89.12	89.38	88.77	82.89	90.23
PSPNet	74.47	60.87	60.98	47.45	40.87

Clearly, Res2-Unet+ achieves the best performance with the highest *precision*, *recall*, *F1-measure*, and *IOU* among the four comparing methods. The precision of Res2-Unet+ is 3% higher than the second-highest precision, found in Unet. This indicates that the Res2net module is effective in capturing the multi-scale features of oil tanks at a granular level. By enlarging the receptive fields during convolution operation, our proposed Res2-Unet+ can increase the precision of oil tank detection and decrease the omission rate of oil tanks that are difficult to distinguish from other complex background objects. Notably, PSPNet does not perform satisfactorily in detecting oil tanks in the evaluation images. One possible reason is that the spatial pyramid pooling module can filter out the multi-scale features of oil tanks, especially when they are small, constituting a few pixels only. Moreover, as shown in Table 1, since evaluation images have different spatial resolutions from different imaging sensors of different study areas with training images, the proposed Res2-Unet+ has stronger transferability in oil detection. In terms of implementation efficiency, our proposed Res2-Unet+ is similar high efficient with Unet due to both FPS higher than 100, while PSPNet is low efficient with FPS of lower than 50. Therefore, with a high efficiency and transferability, our proposed Res2-Unet+ can be recognized as with stronger applicability for practical applications.

5. Discussion

The above analysis shows that the performance of our method relies heavily on the size of an oil tank. Therefore, we analyzed the detection performance of our proposed Res2-Unet+ as a function of oil tank area. The area of all the ground truth oil tanks in the evaluation images ranges from 1 m² to 19,326.5 m². We grouped the oil tanks into five categories according to the area of each oil tank with manually set criteria, 1–400 m², 400–800 m², 800–1200 m², 1200–1600 m², and greater than 1600 m². The average precision, recall, F1-measure, and IOU of all the oil tanks in each category are calculated and summarized in Figure 9. The evaluation statistics increase immediately when the area of oil tanks is larger than 400 m². Our proposed method performs best with regard to oil tank detection for the area 400–800 m². The other three evaluation statistics decrease gradually with increasing area, especially IOU. Precision stays high and smooth with the continuously increasing size of oil tanks after the area getting greater than 400 m². This indicates that oil tank detection using the proposed method is accurate for areas greater than 400 m². However, the proposed method can easily omit oil tank pixels when they have a large area. For oil tanks with an area larger than 800 m², a major proportion of the oil tank pixels can be easily detected by the proposed method; but the number of pixels that are confused with neighboring bright background objects increases greatly with increasing oil tank area. To deal with this problem, enlarging the variability of the training samples of large oil tanks may be one possible solution in future studies.

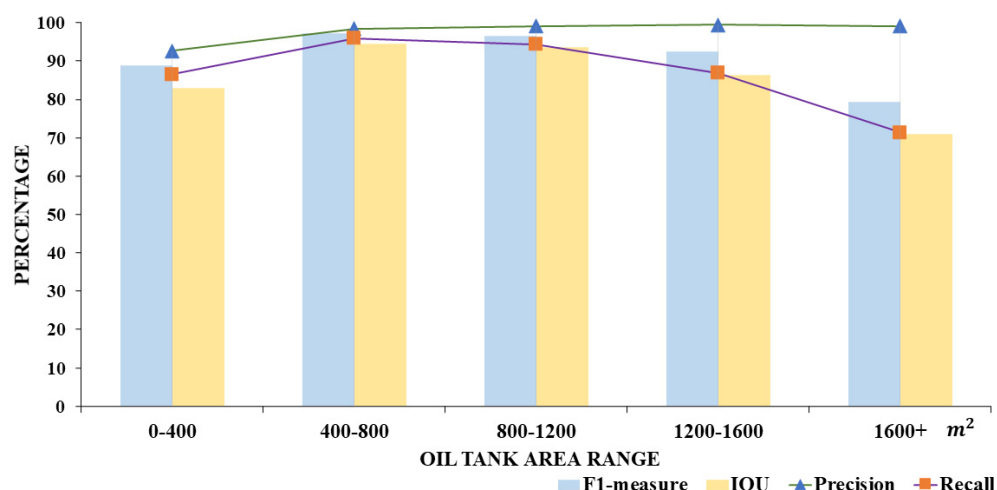


Figure 9. Evaluation statistics as a function of oil tank area using Res2-Unet+.

6. Conclusions

In this paper, we propose an end-to-end deep learning framework to detect oil tanks in high spatial resolution images. The proposed Res2-Unet+ framework adopts a Res2Net module in the encoder to learn the multi-scale features of oil tanks in a hierarchical manner. Such granular level learning enlarges the receptive field in the residual convolution. This improves the accuracy in detecting oil tanks with multiple sizes and colors. Compared with three other widely used network structures, Unet, Segnet, and PSPNet, Res2-Unet+ achieves the best performance by detecting most oil tanks accurately with the smallest number of false alarm pixels. The spatial pyramid pooling module employed in PSPNet proves to be unsuitable in detecting small oil tanks from background objects with similar spectral appearances. One possible reason is due to the small size of oil tanks, which is easily to be omitted during the continuous pooling. Moreover, our proposed Res2-Unet+ framework performs best when oil tanks have an area of between 400 m² and 800 m². The accuracy decreases gradually with increasing oil tank area from 800 m². This may be overcome by adding more training samples of large oil tanks in future studies. Generally, our proposed Res2-Unet+ method has strong transferability and generalization when faced

with images from different sensors in different study areas and with different background object distribution patterns.

Author Contributions: Conceptualization, B.Y. and Y.W.; methodology, B.Y. and F.C.; validation, P.M., C.Z., X.Y. and Y.Z.; formal analysis, B.Y.; investigation, B.Y.; resources, Y.W. and N.W.; writing—original draft preparation, B.Y.; writing—review and editing, B.Y.; funding acquisition, F.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 41871345, the Strategic Priority Research Program of the Chinese Academy of Sciences, grant number XDA19030101, and Guangxi Innovation-driven Development Special Project, grant number GuiKe-AA20302022.

Data Availability Statement: The implementation codes are available from [yubozuzu123/res2_unet](https://github.com/yubozuzu123/res2_unet) (github.com), accessed on 21 October 2021.

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

References

1. Zalpour, M.; Akbarizadeh, G.; Alaei-Sheini, N. A new approach for oil tank detection using deep learning features with control false alarm rate in high-resolution satellite imagery. *Int. J. Remote Sens.* **2020**, *41*, 2239–2262. [\[CrossRef\]](#)
2. Chen, F.; Yu, B.; Xu, C.; Li, B. Landslide detection using probability regression, a case study of wenchuan, northwest of chengdu. *Appl. Geogr.* **2017**, *89*, 32–40. [\[CrossRef\]](#)
3. Izadi, M.; Mohammadzadeh, A.; Haghighattalab, A. A new neuro-fuzzy approach for post-earthquake road damage assessment using ga and svm classification from quickbird satellite images. *J. Indian Soc. Remote Sens.* **2017**, *45*, 965–977. [\[CrossRef\]](#)
4. Yu, B.; Yang, L.; Chen, F. Semantic segmentation for high spatial resolution remote sensing images based on convolution neural network and pyramid pooling module. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 3252–3261. [\[CrossRef\]](#)
5. Liu, Z.; Zhao, D.; Shi, Z.; Jiang, Z. Unsupervised saliency model with color markov chain for oil tank detection. *Remote Sens.* **2019**, *11*, 1089. [\[CrossRef\]](#)
6. Kushwaha, N.K.; Chaudhuri, D.; Singh, M.P. Automatic bright circular type oil tank detection using remote sensing images. *Def. Sci. J.* **2013**, *63*, 298–304. [\[CrossRef\]](#)
7. Yuen, H.; Princen, J.; Illingworth, J.; Kittler, J. Comparative study of hough transform methods for circle finding. *Image Vis. Comput.* **1990**, *8*, 71–77. [\[CrossRef\]](#)
8. Atherton, T.J.; Kerbyson, D.J. Size invariant circle detection. *Image Vis. Comput.* **1999**, *17*, 795–803. [\[CrossRef\]](#)
9. Li, B.; Yin, D.; Yuan, X.; Li, G.-Q. Oilcan recognition method based on improved hough transform. *Opto-Electron. Eng.* **2008**, *35*, 30–34.
10. Weisheng, Z.; Hong, Z.; Chao, W.; Tao, W. Automatic oil tank detection algorithm based on remote sensing image fusion. In Proceedings of the 25th Anniversary IGARSS 2005 IEEE International Geoscience and Remote Sensing Symposium, Seoul, Korea, 9–29 July 2005; pp. 3956–3958.
11. Cai, X.; Sui, H.; Lv, R.; Song, Z. Automatic circular oil tank detection in high-resolution optical image based on visual saliency and hough transform. In Proceedings of the 2014 IEEE Workshop on Electronics, Computer and Applications, Ottawa, ON, Canada, 8–9 May 2014; pp. 408–411.
12. Ok, A.O. A new approach for the extraction of aboveground circular structures from near-nadir vhr satellite imagery. *IEEE Trans. Geosci. Remote Sens.* **2013**, *52*, 3125–3140. [\[CrossRef\]](#)
13. Xu, H.; Chen, W.; Sun, B.; Chen, Y.; Li, C. Oil tank detection in synthetic aperture radar images based on quasi-circular shadow and highlighting arcs. *J. Appl. Remote Sens.* **2014**, *8*, 083689. [\[CrossRef\]](#)
14. Yao, Y.; Jiang, Z.; Zhang, H. Oil tank detection based on salient region and geometric features. In Proceedings of the Optoelectronic Imaging and Multimedia Technology III, Beijing, China, 9–11 October 2014; p. 92731G.
15. Jing, M.; Zhao, D.; Zhou, M.; Gao, Y.; Jiang, Z.; Shi, Z. Unsupervised oil tank detection by shape-guide saliency model. *IEEE Geosci. Remote Sens. Lett.* **2018**, *16*, 477–481. [\[CrossRef\]](#)
16. Ok, A.O.; Başeski, E. Circular oil tank detection from panchromatic satellite images: A new automated approach. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 1347–1351. [\[CrossRef\]](#)
17. Xia, X.; Liang, H.; RongFeng, Y.; Kun, Y. Oil tank extraction in high-resolution remote sensing images based on deep learning. In Proceedings of the 26th International Conference on Geoinformatics, Kunming, China, 28–30 June 2018; pp. 1–6.
18. Zhang, L.; Shi, Z.; Wu, J. A hierarchical oil tank detector with deep surrounding features for high-resolution optical satellite imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 4895–4909. [\[CrossRef\]](#)
19. Tadros, A.; Drouyer, S.; Gioi, R.G.v.; Carvalho, L. Oil tank detection in satellite images via a contrario clustering. In Proceedings of the IGARSS 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020; pp. 2233–2236.

20. Chen, Y.; Li, L.; Whiting, M.; Chen, F.; Sun, Z.; Song, K.; Wang, Q. Convolutional neural network model for soil moisture prediction and its transferability analysis based on laboratory vis-nir spectral data. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *104*, 102550. [[CrossRef](#)]
21. Dalal, N.; Triggs, B. Histograms of oriented gradients for human detection. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–25 June 2005; pp. 886–893.
22. Bay, H.; Ess, A.; Tuytelaars, T.; Van Gool, L. Speeded-up robust features (surf). *Comput. Vis. Image Underst.* **2008**, *110*, 346–359. [[CrossRef](#)]
23. 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 Intervention, Munich, Germany, 5–9 October 2015; pp. 234–241.
24. Ivanovsky, L.; Khryashchev, V.; Pavlov, V.; Ostrovskaya, A. Building detection on aerial images using u-net neural networks. In Proceedings of the 24th Conference of Open Innovations Association (FRUCT), Moscow, Russia, 8–12 April 2019; pp. 116–122.
25. Wang, P.; Chen, P.; Yuan, Y.; Liu, D.; Huang, Z.; Hou, X.; Cottrell, G. Understanding convolution for semantic segmentation. In Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Tahoe, NV, USA, 12–15 March 2018; pp. 1451–1460.
26. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
27. He, K.; Zhang, X.; Ren, S.; Sun, J. Identity mappings in deep residual networks. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 11–14 October 2016; pp. 630–645.
28. Sengupta, A.; Ye, Y.; Wang, R.; Liu, C.; Roy, K. Going deeper in spiking neural networks: Vgg and residual architectures. *Front. Neurosci.* **2019**, *13*, 95. [[CrossRef](#)] [[PubMed](#)]
29. 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](#)]
30. Yu, F.; Koltun, V. Multi-scale context aggregation by dilated convolutions. *arXiv* **2015**, arXiv:1511.07122.
31. 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.* **2017**, *40*, 834–848. [[CrossRef](#)] [[PubMed](#)]
32. Liu, P.; Liu, X.; Liu, M.; Shi, Q.; Yang, J.; Xu, X.; Zhang, Y. Building footprint extraction from high-resolution images via spatial residual inception convolutional neural network. *Remote Sens.* **2019**, *11*, 830. [[CrossRef](#)]
33. Sebastian, C.; Imbriaco, R.; Bondarev, E.; de With, P.H. Contextual pyramid attention network for building segmentation in aerial imagery. *arXiv* **2020**, arXiv:2004.07018.
34. He, N.; Fang, L.; Plaza, A. Hybrid first and second order attention unet for building segmentation in remote sensing images. *Sci. China Inf. Sci.* **2020**, *63*, 1–12. [[CrossRef](#)]
35. Zhou, Z.; Siddiquee, M.; Tajbakhsh, N.; Liang, J. Unet++: A nested u-net architecture for medical image segmentation. In Proceedings of the 4th Deep Learning in Medical Image Analysis (DLMIA) Workshop, Granada, Spain, 20 September 2018.
36. Gao, S.; Cheng, M.-M.; Zhao, K.; Zhang, X.-Y.; Yang, M.-H.; Torr, P.H. Res2net: A new multi-scale backbone architecture. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 652–662. [[CrossRef](#)] [[PubMed](#)]
37. Chen, G.; Chen, P.; Shi, Y.; Hsieh, C.-Y.; Liao, B.; Zhang, S. Rethinking the usage of batch normalization and dropout in the training of deep neural networks. *arXiv* **2019**, arXiv:1905.05928.
38. Bottou, L. Large-scale machine learning with stochastic gradient descent. In Proceedings of the COMPSTAT'2010, Paris, France, 22–27 August 2010; pp. 177–186.